República Federativa do Brasil
Ministério do Desenvolvimento, Indústria e do Comércio Exterior
Instituto Nacional da Propriedade Industrial.

(21) **PI0709598-8 A2**

BRPI0709598A2

(22) Data de Depósito: 19/03/2007
(43) Data da Publicação: 19/07/2011
(RPI 2115)

(51) Int.Cl.:
C07K 16/00 2006.01
C12N 15/13 2006.01
C12N 5/10 2006.01
A61K 39/395 2006.01
G06F 19/00 2006.01

(54) Título: **COMPOSIÇÕES DE POLIPTÍDEOS ESTABILIZADOS**

(30) Prioridade Unionista: 17/03/2006 US 60/783,622, 09/06/2006 US 60/812,688, 08/12/2006 US 60/873,802, 08/12/2006 US 60/873,996, 17/03/2006 US 60/783,622, 09/06/2006 US 60/812,688, 09/06/2006 US 60/812,688, 08/12/2006 US 60/873,802, 17/03/2006 US 60/783,622, 08/12/2006 US 60/873,802, 08/12/2006 US 60/873,996

(73) Titular(es): Biogen Idec MA INC.

(72) Inventor(es): Alexey Alexandrovich Lugovskoy, Brian Robert Miller, Lisa J. Croner, Norman Wang, Scott Glaser, Stephen Demarest, William B. Snyder, Xiufeng Wu

(74) Procurador(es): Dannemann ,Siemsen, Bigler & Ipanema Moreira

(86) Pedido Internacional: PCT US2007006883 de 19/03/2007

(87) Publicação Internacional: WO 2007/109254de 27/09/2007

(57) Resumo: COMPOSIÇÕES DE POLIPEPTÍDEOS ESTABILIZADOS. A presente invenção refere-se, no mínimo em parte, ao desenvolvimento de moléculas de ligação estabilizadas que consistem em ou compreendem uma scFv estabilizada e métodos para preparar similares moléculas estabilizadas.

PI0709598-8

Ref nº 020080143/112/RJ

(17/11/08)

Relatório Descritivo da Patente de Invenção para **"COMPOSIÇÕES DE POLIPEPTÍDEOS ESTABILIZADOS".**

Pedido de Patente Relacionados

O pedido de patente reivindica o benefício de prioridade para os seguintes Pedidos de patente provisória dos Estados Unidos, cada um dos quais é por este incorporado por meio de referência em sua totalidade para todos os fins:

1) Pedido de patente provisória dos Estados Unidos Nº 60/783,622, depositado em 17 de março de 2006, intitulado "STABILIZED POLYPEPTIDES AND METHODS FOR EVALUATING AND INCREASING THE STABILITY OF SAME"

2) Patente de patente provisória dos Estados Unidos Nº 60/812.688, depositado em 9 de junho de 2006, intitulado "STABILIZED POLYPEPTIDES AND METHODS FOR EVALUATING AND INCREASING THE STABILITY OF SAME"

3) Pedido de patente provisória dos Estados Unidos Nº 60/873.502, depositado em 8 de dezembro de 2006, intitulado "METHODS FOR EVALUATING AND ENHANCING BIOPHYSICAL PROPERTIES OF POLYPEPTIDES"

4) Pedido de patente provisória dos Estados Unidos Nº 60/873.996, depositado em 8 de dezembro de 2006, intitulado "STABILIZED POLYPEPTIDE COMPOSITIONS".

O conteúdo de quaisquer patentes, Pedidos de patente, e referências citadas do início ao fim deste relatório descritivo são por este incorporadas por meio de referência em suas totalidades.

Antecedentes da Invenção

A estabilidade protéica é atualmente reconhecida como um problema central para o desenvolvimento e aumento em escala de proteínas, por exemplo, moléculas de anticorpos. As seqüências de domínio variável de cadeia pesada e leve de anticorpos estão sob pressão seletiva e podem variar significativamente em seqüência de um anticorpo para outro (Wu e Kabat, 1970). Há um grande número de genes de cadeia pesada variável de linha germinal

funcional (~40; Tomlinson et al., 1992; Tomlinson et al., 1994; Matsuda e Honjo, 1996), variável kappa (~40; Meindl et al., 1989; Cox et al., 1994; Barbie e Lefranc, 1998) e variável lâmbda (~30; Williams et al., 1996; Kawasaki, et al., 1997) embutidos dentro do genoma humano. A capacidade do sistema imune para criar muitas combinações de linhas germinais de domínio variável de cadeia pesada e de cadeia leve é um mecanismo para criar diversidade de anticorpos (Edelman, 1959; Frañek, 1961). Diversidade adicional é derivada pela inserção de uma região de peptídeo de ligação variável, o peptídeo de ligação J (mais peptídeo de ligação D para cadeias pesadas), entre os domínios variáveis e os domínios constantes (Leder, 1982). Subseqüente à seleção de anticorpos de linha germinal contra um antígeno específico, células B expressando um único anticorpo com linhas germinais de domínio variável selecionado e peptídeos (D-)J-conectivos sofrem o processo de mutação hipersomática dentro dos domínios variáveis de cadeia pesada e leve para aumentar a afinidade de anticorpo para o antígeno (Leder, 1982; Tomlinson et al., 1996). Inserções ou eliminações hipersomáticas dentro dos domínios variáveis também são comumente observadas, embora em uma freqüência muito menor do que a mutação hipersomática (de Wildt et al., 1999). Portanto, um anticorpo maduro selecionado contra um antígeno em particular terá seqüências de domínio variável altamente únicas as quais não são necessariamente otimizadas para estabilidade.

Estabilidade pobre pode afetar a capacidade de um anticorpo ou domínio de anticorpo dobrar adequadamente quando expressada em vários sistemas celulares, por exemplo, em sistemas de expressão bacteriana ou mamífera. Dobramento incorreto ou estabilidade pobre também podem resultar em populações fracionadas de material não-funcional (Martsev et al., 1998) e/ou anticorpos com a tendência para formar grandes agregados solúveis os quais são potencialmente perigosos ou imunogênicos quando usados terapeuticamente. Outros problemas de estabilidade incluem redobramento impedido, degradação, avidez prejudicada, agregação, ou perda de atividade depois de armazenamento.

Problemas de estabilidade não estão limitados a anticorpos que

ocorrem naturalmente, mas podem ocorrer em moléculas de ligação antigênica produzidas por engenharia genética também, tais como bibliotecas de anticorpos recombinantes (Hoogenboom, 2005), fragmentos de anticorpos (Holt et al., 2003; Todorovska et al., 2001; Worn and Plückthun, 2001; Reiter e Pastan, 1996) e anticorpos terapêuticos engenheirados genética ou humanizados para fins de fabricação e clínicos (Ewert et al., 2004; Carter and Merchant, 1997). Além disso, formas multivalentes de anticorpos, tais como moléculas biespecífica podem ter problemas de estabilidade. Apesar de anticorpos biespecíficos serem desejáveis devido a sua utilidade terapêutica em seres humanos, sua produção é imprevisível. Por exemplo, anticorpos biespecíficos podem resultar em substancial redução na estabilidade térmica, produção de produto, ou na formação de agregados de baixo peso molecular.

Alterações não naturais para os domínios VH e V_L também podem surgir ao humanizar anticorpos de roedores. Geralmente a humanização é realizada enxertando regiões determinantes de complementaridade (CDRs) de roedores sobre o domínio variável de linha germinal humana mais similar (Hurle e Gross, 1994). As linhas germinais escolhidas para humanização podem não ser usadas com grande freqüência pelo sistema imune e freqüentemente são necessárias alterações dentro das estruturas humanas para obter adequada ligação antigênica.

Proteínas instáveis sofrem de muitos problemas, incluindo um ou mais de: inadequabilidade para aumentar em escala a produção em biorreatores (por exemplo, devido ao baixo rendimento, níveis significativos de subprodutos indesejados tais como produto não reunido, e/ou material agregado), dificuldades em purificação de proteína, e inadequabilidade para preparação farmacêutica e uso (por exemplo, devido a níveis significativos de produtos de decomposição, qualidade pobre do produto, e/ou propriedades farmacocinéticas desfavoráveis). A instabilidade dentro dos domínios variáveis de anticorpos, Fabs, Fvs, ou Fvs de cadeia única (scFvs) pode resultar em muitos destes problemas. Construtos de scFv em particular têm demonstrado problemas com estabilidade, solubilidade, expressão, agregação, pro-

dutos de decomposição, e manufaturabilidade global tanto em sistemas de expressão bacterianos quanto em mamíferos (vide Worn e Pluckthun, 2001). Adicionalmente, a incorporação de moléculas scFv em produtos de anticorpos recombinantes estáveis de modo diverso freqüentemente confere estes traços geralmente indesejáveis ao novo desenho recombinante.

Por conseguinte, existe a necessidade de métodos aprimorados para prever a estabilidade de proteínas, tais como anticorpos, e para moléculas de ligação de antígeno estáveis aprimoradas que são adequadas para produção escalável. Também existe a necessidade de métodos aprimorados que possibilitem a produção escalável de uma população de moléculas de ligação de antígeno estáveis que são adequadas para aplicações farmacêuticas.

Sumário da Invenção

A invenção se baseia, no mínimo em parte, no desenvolvimento de composições de polipeptídeos aprimoradas, por exemplo, moléculas de ligação aprimoradas, compreendendo ou consistindo em moléculas scFv com aprimorada estabilidade e métodos para preparar as mesmas.

Em um aspecto, a invenção pertence a uma população de moléculas scFv estabilizadas, em que as moléculas scFv estabilizadas compreendem

i) um ligante de scFv interposto entre um domínio V_H e um domínio V_L, em que os domínios VH e V_L são encadeados por uma ligação dissulfeto entre um aminoácido no domínio V_H e um aminoácido no domínio V_L, e em que os domínios VH e V_L das moléculas scFv têm valores de Tm de mais de 55°C; ou

ii) um ligante de scFv interposto entre um domínio V_H e um domínio V_L, em que os domínios VH e V_L são encadeados por uma ligação dissulfeto entre um aminoácido no domínio V_H e um aminoácido no domínio V_L, e em que moléculas scFv têm uma T_{50} de mais de 49°C.

Em outro aspecto, a invenção pertence a molécula scFv estabilizada em que a molécula scFv estabilizada compreende

a. um ligante de scFv tendo a seqüência de aminoácidos (Gly_4

Ser)$_n$ interposta entre um domínio V_H e um domínio V_L, em que os domínios VH e V_L são encadeados por uma ligação dissulfeto entre o aminoácido 44 de V_H e o aminoácido 100 de V_L;

ii) um domínio V_H e um domínio V_L em que a molécula tem no mínimo uma substituição selecionada entre um grupo consistindo em:

a) substituição de um aminoácido na posição de Kabat 13 de VH;

b) substituição de um aminoácido na posição de Kabat 16 de VH;

c) substituição de um aminoácido na posição de Kabat 46 de VL;

d) substituição de um aminoácido na posição de Kabat 49 de VL;

e) substituição de um aminoácido na posição de Kabat 50 de VL;

f) substituição de aminoácidos nas posições de Kabat 49 e 50 de VL;

g) substituição de aminoácido na posição de Kabat 101 de VH;

h) substituição de aminoácido na posição de Kabat 20 de VH;

i) substituição de aminoácido na posição de Kabat 48 de VH;

j) substituição de aminoácido na posição de Kabat 3 de VL;

k) substituição de aminoácido na posição de Kabat 55 de VH;

l) substituição de aminoácido na posição de Kabat 67 de VH;

m) substituição de aminoácido na posição de Kabat 6 de VH ;

n) substituição de aminoácido na posição de Kabat 32 de VH;

o) substituição de aminoácido na posição de Kabat 49 de VH;

p) substituição de aminoácido na posição de Kabat 43 de VH;

q) substituição de aminoácido na posição de Kabat 72 de VH ;

r) substituição de aminoácido na posição de Kabat 79 de VH;

s) substituição de aminoácido na posição de Kabat 50 de VL;

t) substituição de aminoácido na posição de Kabat 75 de VL;

u) substituição de aminoácido na posição de Kabat 80 de VH;

v) substituição de aminoácido na posição de Kabat 83 de VH;

ou

iii) um domínio V_H e um domínio V_L com uma interface estabilizada entre VH e VL da scFv em que a molécula tem no mínimo uma substituição em uma posição de aminoácido diretamente dentro da interface

e/ou em uma posição de aminoácido que *estruturas* a interação entre VH e VL, em que a no mínimo uma posição de aminoácido é selecionada entre um grupo consistindo em: 47H, 37H, 45H, 46H, 51H, 59H, 109H, 14H, 26H, 27H, 40H, 69H, 103H, 29H, 38H, 44H, 49H, 55H, 63H, 54H, 68H, 72H, 74H, 79H, 82bH, 8H, 32H, 39H, 41H, 62H, 85H, 91H, 24H, 60H, 37L, 36L, 44L, 59L, 57L, 64L, 46L, 48L, 63L, 67L, 68L, 39L, 45L, 47L, 54L, 56L, 79L, 85L, 98L, 58L, 74L, 77L, 83L, 89L, e 104L, de acordo com a numeração de Kabat, e em que a molécula scFv estabilizada tem uma T_{50} maior do que a de uma molécula scFv carecendo da substituição.

Em outro aspecto, a invenção pertence a uma molécula scFv estabilizada compreendendo um ligante de scFv $(Gly_4\ Ser)_4$ interposto entre um domínio V_H e um domínio V_L, em que os domínios VH e V_L são encadeados por uma ligação dissulfeto entre um aminoácido no domínio V_H e um aminoácido no domínio V_L.

Em outra modalidade, a invenção pertence a uma molécula scFv estabilizada compreendendo um domínio V_H e um domínio V_L em que a molécula tem no mínimo uma série de substituições selecionadas entre o grupo consistindo em:

a) uma substituição de um aminoácido na posição de Kabat 16 de VH, por exemplo, com um glutamato ou glutamina e uma substituição de um aminoácido na posição de Kabat 46 de VL, por exemplo, com uma lisina;

b) uma substituição de um aminoácido na posição de Kabat 16 de VH, por exemplo, com um glutamato ou glutamina; uma substituição de um aminoácido na posição de Kabat 46 de VL, por exemplo, com uma lisina; e uma substituição de um aminoácido na posição de Kabat 55 de VH, por exemplo, com uma glicina; e

c) uma substituição de um aminoácido na posição de Kabat 16 de VH, por exemplo, com um glutamato ou glutamina; uma substituição de um aminoácido na posição de Kabat 46 de VL, por exemplo, com uma lisina; uma substituição de um aminoácido na posição de Kabat 55 de VH, por exemplo, com uma glicina; e substituição de um aminoácido na posição de Kabat 101 de VH, por exemplo, com um ácido aspártico.

Em outra modalidade, a molécula scFv estabilizada tem uma estabilidade equivalente a de um fragmento Fab convencional sob condições de estímulo térmico.

Em uma modalidade, a estabilidade é avaliada medindo a capacidade para ligar a uma molécula alvo.

Em uma modalidade, a molécula scFv compreende uma seqüência de aminoácidos codificados por uma seqüência de nucleotídeos selecionada entre o grupo consistindo em SEQ ID NOs:9, 14, 16, e 18.

Em uma modalidade, a molécula scFv compreende uma seqüência de aminoácidos selecionada entre o grupo consistindo em SEQ ID NOs:10, 15, 17, e 19.

Em outra modalidade, a invenção pertence a uma proteína de fusão compreendendo uma molécula scFv estabilizada.

Em uma modalidade, a proteína de fusão compreende no mínimo dois sítios de ligação antigênica.

Em outro aspecto, a invenção pertence a uma população de moléculas de ligação de antígenos multivalentes estabilizadas, em que as moléculas de ligação estabilizadas compreendem no mínimo uma molécula scFv estabilizada compreendendo um ligante de scFv interposto entre um domínio V_H e um domínio V_L, em que os domínios VH e V_L são encadeados por uma ligação dissulfeto entre um aminoácido no domínio V_H e um aminoácido no domínio V_L, e em que a população de moléculas de ligação estabilizadas compreendem proteínas solúveis monoméricas das quais não mais de 10% está presente em forma agregada.

Em uma modalidade, as moléculas de ligação de antígenos multivalentes estabilizadas são expressadas em células CHO ou NS0.

Em outro aspecto, a invenção pertence a uma população de moléculas de ligação de antígenos multivalentes estabilizadas, em que cada molécula de ligação de antígeno multivalente compreende

i) no mínimo uma molécula scFv compreendendo um ligante de scFv interposto entre um domínio V_H e um domínio V_L, em que os domínios VH e V_L são encadeados por uma ligação dissulfeto entre um aminoácido no

domínio V_H e um aminoácido no domínio V_L, e em que os domínios VH e V_L da no mínimo uma molécula scFv têm uma Tm de mais de 55°C; ou

ii) no mínimo uma molécula scFv compreendendo um ligante de scFv interposto entre um domínio V_H e um domínio V_L, em que os domínios VH e V_L são encadeados por uma ligação dissulfeto entre um aminoácido no domínio V_H e um aminoácido no domínio V_L, e em que a no mínimo uma molécula scFv da molécula de ligação de antígeno multivalente tem uma T_{50} de mais de 49°C.

Em outra modalidade, a invenção pertence a uma molécula de ligação de antígenos multivalentes estabilizada compreendendo

i) no mínimo uma molécula scFv estabilizada, em que a molécula scFv estabilizada compreende um ligante de scFv $(Gly_4\ Ser)_n$ interposto entre um domínio V_H e um domínio V_L e em que os domínios VH e V_L são encadeados por uma ligação dissulfeto;

ii) no mínimo uma molécula scFv estabilizada compreendendo um ligante de scFv tendo a seqüência de aminoácidos $(Gly_4\ Ser)_n$ interposto entre um domínio V_H e um domínio V_L, em que os domínios VH e V_L são encadeados por uma ligação dissulfeto entre o aminoácido 44 de V_H e o aminoácido 100 de V_L;

iii) no mínimo uma molécula scFv estabilizada, em que a molécula scFv estabilizada compreende no mínimo uma substituição selecionada entre um grupo consistindo em:

a) substituição de um aminoácido na posição de Kabat 13 de VH;
b) substituição de um aminoácido na posição de Kabat 16 de VH;
c) substituição de um aminoácido na posição de Kabat 46 de VL;
d) substituição de um aminoácido na posição de Kabat 49 de VL;
e) substituição de um aminoácido na posição de Kabat 50 de VL;
f) substituições de aminoácidos nas posições de Kabat 49 e 50 de VL;
g) substituição de aminoácido na posição de Kabat 101 de VH;
h) substituição de aminoácido na posição de Kabat 20 de VH;
i) substituição de aminoácido na posição de Kabat 48 de VH;

j) substituição de aminoácido na posição de Kabat 3 de VL;

k) substituição de aminoácido na posição de Kabat 55 de VH;

l) substituição de aminoácido na posição de Kabat 67 de VH;

m) substituição de aminoácido na posição de Kabat 6 de VH;

n) substituição de aminoácido na posição de Kabat 32 de VH;

o) substituição de aminoácido na posição de Kabat 49 de VH;

p) substituição de aminoácido na posição de Kabat 43 de VH;

q) substituição de aminoácido na posição de Kabat 72 de VH;

r) substituição de aminoácido na posição de Kabat 79 de VH;

s) substituição de aminoácido na posição de Kabat 50 de VL;

t) substituição de aminoácido na posição de Kabat 75 de VL;

u) substituição de aminoácido na posição de Kabat 80 de VH;

v) substituição de aminoácido na posição de Kabat 83 de VH; ou

iv) um domínio V_H e um domínio V_L com uma interface estabilizada entre VH e VL da scFv em que a molécula tem no mínimo uma substituição em uma posição de aminoácido diretamente dentro da interface e/ou em uma posição de aminoácido que *estruturas* a interação entre VH e VL, em que a no mínimo uma posição de aminoácido é selecionada entre um grupo consistindo em: 47H, 37H, 45H, 46H, 51H, 59H, 109H, 14H, 26H, 27H, 40H, 69H, 103H, 29H, 38H, 44H, 49H, 55H, 63H, 54H, 68H, 72H, 74H, 79H, 82bH, 8H, 32H, 39H, 41H, 62H, 85H, 91H, 24H, 60H, 37L, 36L, 44L, 59L, 57L, 64L, 46L, 48L, 63L, 67L, 68L, 39L, 45L, 47L, 54L, 56L, 79L, 85L, 98L, 58L, 74L, 77L, 83L, 89L, e 104L, de acordo com a numeração de Kabat,

Em outro aspecto, a invenção pertence a uma molécula de ligação de antígenos multivalentes estabilizada compreendendo no mínimo uma molécula scFv estabilizada compreendendo um ligante de scFv $(Gly_4\ Ser)_4$ interposto entre um domínio V_H e um domínio V_L, em que os domínios VH e V_L são encadeados por uma ligação dissulfeto entre um aminoácido no domínio V_H e um aminoácido no domínio V_L.

ii) no mínimo uma molécula scFv estabilizada compreendendo um ligante de scFv tendo a seqüência de aminoácidos $(Gly_4\ Ser)_n$ interposto entre um domínio V_H e um domínio V_L, em que os domínios VH e V_L são en-

cadeados por uma ligação dissulfeto entre o aminoácido 44 de V_H e o aminoácido 100 de V_L; ou

iii) no mínimo uma molécula scFv estabilizada, em que a molécula scFv estabilizada compreende no mínimo uma substituição selecionada entre um grupo consistindo em:

a) substituição de um aminoácido (por exemplo, glutamina) na posição de Kabat 3 de VL, por exemplo, com uma alanina, uma serina, uma valina, um ácido aspártico, ou uma glicina;

b) substituição de um aminoácido (por exemplo, serina) na posição de Kabat 46 de VL, por exemplo, com leucina;

c) substituição de um aminoácido (por exemplo, serina) na posição de Kabat 49 de VL, por exemplo, com tirosina ou serina;

d) substituição de um aminoácido (por exemplo, serina ou valina) na posição de Kabat 50 de VL, por exemplo, com serina, treonina, e arginina, ácido aspártico, glicina, ou lisina;

e) substituições de aminoácidos (por exemplo, serina) na posição de Kabat 49 e (por exemplo, serina) na posição de Kabat 50 de VL, respectivamente com tirosina e serina; tirosina e treonina; tirosina e arginina; tirosina e glicina; serina e arginina; ou serina e lisina;

f) substituição de um aminoácido (por exemplo, valina) na posição de Kabat 75 de VL, por exemplo, com isoleucina;

g) substituição de um aminoácido (por exemplo, prolina) na posição de Kabat 80 de VL, por exemplo, com serina ou glicina;

h) substituição de um aminoácido (por exemplo, fenilalanina) na posição de Kabat 83 de VL, por exemplo, com serina, alanina, glicina, ou treonina;

i) substituição de um aminoácido (por exemplo, ácido glutâmico) na posição de Kabat 6 de VH, por exemplo, com glutamina;

j) substituição de um aminoácido (por exemplo, lisina) na posição de Kabat 13 de VH, por exemplo, com glutamato;

k) substituição de um aminoácido (por exemplo, serina) na posição de Kabat 16 de VH, por exemplo, com glutamato ou glutamina;

l) substituição de um aminoácido (por exemplo, valina) na posição de Kabat 20 de VH, por exemplo, com uma isoleucina;

m) substituição de um aminoácido (por exemplo, asparagina) na posição de Kabat 32 de VH, por exemplo, com serina;

n) substituição de um aminoácido (por exemplo, glutamina) na posição de Kabat 43 de VH, por exemplo, com lisina ou arginina;

o) substituição de um aminoácido (por exemplo, metionina) na posição de Kabat 48 de VH, por exemplo, com uma isoleucina ou uma glicina;

p) substituição de um aminoácido (por exemplo, serina) na posição de Kabat 49 de VH, por exemplo, com glicina ou alanina;

q) substituição de um aminoácido (por exemplo, valina) na posição de Kabat 55 de VH, por exemplo, com uma glicina;

r) substituição de um aminoácido (por exemplo, valina) na posição de Kabat 67 de VH, por exemplo, com uma isoleucina ou uma leucina;

s) substituição de um aminoácido (por exemplo, ácido glutâmico) na posição de Kabat 72 de VH, por exemplo, com aspartato ou asparagina;

t) substituição de um aminoácido (por exemplo, fenilalanina) na posição de Kabat 79 de VH, por exemplo, com serina, valina, ou tirosina; e

u) substituição de um aminoácido (por exemplo, prolina) na posição de Kabat 101 de VH, por exemplo, com um ácido aspártico.

Em outro aspecto, a invenção pertence a uma molécula de ligação de antígeno multivalente compreendendo no mínimo uma molécula scFv estabilizada, em que a molécula scFv estabilizada compreende no mínimo uma série de substituições selecionadas entre o grupo consistindo em:

a) uma substituição de um aminoácido na posição de Kabat 16 de VH, por exemplo, com um glutamato ou glutamina e uma substituição de um aminoácido na posição de Kabat 46 de VL, por exemplo, com uma lisina;

b) uma substituição de um aminoácido na posição de Kabat 16 de VH, por exemplo, com um glutamato ou glutamina; uma substituição de um aminoácido na posição de Kabat 46 de VL, por exemplo, com uma lisina;

e uma substituição de um aminoácido na posição de Kabat 55 de VH, por exemplo, com uma glicina; e

c) uma substituição de um aminoácido na posição de Kabat 16 de VH, por exemplo, com um glutamato ou glutamina; uma substituição de um aminoácido na posição de Kabat 46 de VL, por exemplo, com uma lisina; uma substituição de um aminoácido na posição de Kabat 55 de VH, por exemplo, com uma glicina; e substituição de um aminoácido na posição de Kabat 101 de VH, por exemplo, com um ácido aspártico.

Em uma modalidade, no mínimo uma molécula scFv estabilizada é geneticamente fundida a um anticorpo. Em outra modalidade, duas moléculas scFv estabilizadas são geneticamente fundidas a um anticorpo.

Em uma modalidade, no mínimo uma molécula scFv estabilizada é geneticamente fundida ao término carbóxi da cadeia leve ou pesada do anticorpo.

Em uma modalidade, no mínimo uma molécula scFv estabilizada é geneticamente fundida ao término amino de uma cadeia leve ou cadeia pesada do anticorpo.

Em uma modalidade, uma molécula de ligação da invenção compreende no mínimo um sítio de ligação que liga a uma molécula preferencialmente expressada sobre células de câncer.

Em uma modalidade, o domínio V_H de uma molécula scFv da invenção é derivado do anticorpo BHA10.

Em uma modalidade, o domínio V_L de uma molécula scFv da invenção é derivado do anticorpo BHA10.

Em uma modalidade uma molécula de ligação da invenção compreende no mínimo um sítio de ligação que é específico para uma molécula selecionada entre o grupo consistindo em HER1, HER3, CD80, CD86, PD-1, CTLA4, B7-H4, RON, CD200, CD4, BAF R, EGFR, IGFR, VEGFR, um membro da família de receptores TNF, um receptor Tie, MET, IGF1, IGF2, TNF, um ligante de TNF, IL-6, TWEAK, Fn14, CD20, CD23, CRIPTO, HGF, $\alpha4\beta1$ integrina, $\alpha5\beta1$ integrina, $\alpha6\beta4$ integrina, e $\alpha V\beta6$ integrina.

Em uma modalidade, uma molécula de ligação da invenção

compreende no mínimo um sítio de ligação que liga a uma molécula envolvida em modulação de reações imunes. Em uma modalidade, a molécula é um receptor de Fc.

Em uma modalidade, uma molécula de ligação da invenção compreende no mínimo um sítio de ligação que liga a uma molécula envolvida em modulação de angiogênese. Em uma modalidade, uma molécula de ligação da invenção liga a VEGF ou angiopoitina.

Em uma modalidade, uma molécula de ligação da invenção compreende no mínimo um sítio de ligação que liga a um alvo neurológico.

Em uma modalidade, uma molécula de ligação da invenção é multiespecífica. Em uma modalidade, uma molécula de ligação da invenção é biespecífica.

Em outra modalidade, uma molécula de ligação da invenção liga a dois membros da família de receptores de TNF. Em uma modalidade, uma molécula de ligação da invenção liga a LTβR e Trail R2. Em uma modalidade, uma molécula de ligação da invenção compreende uma molécula scFv BHA10 geneticamente fundida a um anticorpo 14A2.

Em uma modalidade, uma molécula de ligação da invenção compreende uma cadeia pesada tendo uma seqüência selecionada entre o grupo consistindo em SEQ ID NO: 33, SEQ ID NO: 35, SEQ ID NO: 37, SEQ ID NO: 47, SEQ ID NO: 49, SEQ ID NO: 51, SEQ ID NO: 53, e SEQ ID NO: 55.

Em uma modalidade, uma molécula de ligação da invenção compreende uma seqüência de nucleotídeos codificando um polipeptídeo da molécula de qualquer uma das reivindicações 4 a 8, 11, 12, e 18 a 22.

Em uma modalidade, a invenção pertence a uma molécula de ácido nucléico compreendendo uma seqüência de nucleotídeos a qual codifica a molécula scFv estabilizada ou molécula de ligação compreendendo uma molécula scFv estabilizada da invenção.

Em uma modalidade, a molécula de ácido nucléico é em um vetor. Em outra modalidade, a invenção pertence a uma célula hospedeira compreendendo um vetor similar.

Em uma modalidade, a célula hospedeira é uma célula hospedeira mamífera, por exemplo, uma célula CHO ou uma célula NS0.

Em uma modalidade, a invenção pertence a uma população de moléculas de ligação compreendendo moléculas scFv, a população tendo no mínimo 10% menos agregados relativos a uma célula hospedeira a qual expressa uma população de moléculas de ligação compreendendo moléculas scFv convencionais.

Em um aspecto, a invenção pertence a um método para produzir uma molécula de ligação estabilizada, compreendendo cultivar uma célula hospedeira sob condições tais que a molécula de ligação estabilizada seja produzida.

Em uma modalidade, no mínimo 10 mg da molécula de ligação estabilizada seja produzida para cada litro do meio de cultura de células hospedeiras e em que não mais de 10% da molécula de ligação está presente em forma de agregado.

Em um aspecto, a invenção pertence a método para tratar um sujeito que se beneficiaria de tratamento com uma molécula de ligação da invenção compreendendo administrar a molécula de ligação ao sujeito de tal modo que ocorre tratamento.

Em uma modalidade, o sujeito está sofrendo de uma doença ou um distúrbio selecionados entre o grupo consistindo em câncer, uma doença ou distúrbio auto-imune, e uma doença ou distúrbio neurológico.

Em um aspecto, a invenção pertence a um método parra estabilizar uma molécula scFv compreendendo fundir geneticamente um domínio V_H e um domínio V_L usando um ligante de scFv $(Gly_4\ Ser)_n$ e produzir o domínio V_H para compreender uma cisteína no aminoácido 44 e produzir o domínio V_L para compreender uma cisteína no aminoácido 100 para formar uma molécula scFv estabilizada.

Em um aspecto, a invenção pertence a um método para produzir um anticorpo multivalente estabilizado compreendendo uma molécula scFv estabilizada, o método compreendendo fundir geneticamente uma molécula scFv estabilizada ao término amino ou ao término carbóxi de uma cadeia

leve ou pesada de uma molécula de anticorpo.

Em uma modalidade, a invenção pertence a um método para produzir uma molécula scFv estabilizada compreendendo substituir no mínimo um aminoácido na interface VH/VL da molécula scFv e/ou no mínimo um aminoácido que *estruturas* a uma ou mais substituições da interface VH/VL da molécula scFv simultaneamente melhoram a estabilidade térmica de ambos os domínios VH e V_L da molécula scFv comparada com uma molécula scFv convencional.

Em ainda outra modalidade, a invenção pertence a um método para produzir uma molécula scFv estabilizada compreendendo substituir no mínimo um aminoácido no domínio VH ou domínio VL que covaria com dois ou mais aminoácidos na interface entre os domínios VH e VL.

Em ainda outra modalidade, a invenção pertence a um método para produzir uma proteína de fusão estabilizada compreendendo uma molécula scFv compreendendo substituir no mínimo um aminoácido na interface VH/VL da molécula scFv e/ou no mínimo um aminoácido no *estrutura* VH/VL da molécula scFv e fundir geneticamente a molécula scFv a um polipeptídeo para deste modo produzir uma proteína de fusão estabilizada.

Em ainda outra modalidade, a invenção pertence a um método para melhorar a estabilidade de uma molécula multivalente compreendendo no mínimo uma molécula scFv, o método compreendendo introduzir no mínimo uma mutação estabilizante em a no mínimo uma molécula scFv para deste modo melhorar a estabilidade da molécula multivalente.

Em ainda outra modalidade, a invenção pertence a um método para melhorar a estabilidade de uma molécula scFv compreendendo introduzir no mínimo uma mutação estabilizante em a molécula scFv para deste modo melhorar a estabilidade de uma molécula scFv.

Em ainda outro aspecto, a invenção pertence a um método para fabricação em larga scala de uma proteína de fusão estabilizada, o método compreendendo:

(a) conduzir um ensaio de estabilidade térmica de uma molécula scFv para determinar a estabilidade térmica de uma molécula scFv candidata;

(b) comparar a estabilidade térmica da molécula scFv candidata com um controle adequado para identificar moléculas scFv estabilizadas tendo aumentada estabilidade térmica em relação ao controle,

(c) selecionar as moléculas scFv estabilizadas identificado na etapa (b);

(d) fundir geneticamente no mínimo uma molécula scFv estabilizada a proteína para formar uma proteína de fusão estabilizada;

(e) transfectar uma célula hospedeira mamífera com uma molécula de ácido nucléico codificando a proteína de fusão estabilizada,

(g) cultivar a célula hospedeira da etapa (f) sob condições tais que a proteína de fusão estabilizada é expressada;

em que a proteína de fusão estabilizada é expressada com não mais de 10% de agregação de proteína quando cultivada em 10 L ou mais de meio de cultura.

Em ainda outro aspecto, a invenção pertence a um método para determinar a estabilidade térmica de uma dobra de proteína a partir de um polipeptídeo da superfamília de imunoglobulinas (Ig) compreendendo uma seqüência candidata de aminoácidos, o método compreendendo as etapas de:

a) proporcionar um alinhamento de uma série de referência curada de seqüências correspondentes a uma dobra de Ig do polipeptídeo;

b) calcular a covariação entre resíduos aminoácidos das seqüências do alinhamento para gerar dados de covariação;

c) determinar um escore de covariação de seqüência posição-específico para posições de aminoácido dentro da seqüência candidata a partir de covariações dentro dos dados de covariação; e

d) armazenar ou fornecer resultados do escore de covariação de seqüência posição de aminoácido-específico como uma medida da estabilidade do polipeptídeo.

Em uma modalidade, o método compreende repetir no mínimo as etapas c) e d) para uma pluralidade de polipeptídeos da superfamília de imunoglobulinas (Ig). Em ainda outra modalidade, o método compreende

adicionalmente a etapa de selecionar entre a pluralidade de polipeptídeos da superfamília de imunoglobulinas (Ig) um polipeptídeo da superfamília das Ig para produção, com base nos escores de covariação posição de seqüência específica das seqüências candidatas.

Em outra modalidade, o método compreende adicionalmente a etapa de produzir o polipeptídeo selecionado entre a superfamília das Ig e preparar este para uso terapêutico.

Em outra modalidade, o polipeptídeo da superfamília das Ig derivado de uma proteína selecionada entre o grupo consistindo em uma imunoglobulina ou fragmento de ligação de antígeno da mesma, um receptor de imunoglobulina, uma proteína de adesão celular, uma integrina, um alergeno, um receptor de células T, e um complexo de histocompatibilidade maior (MHC).

Em uma modalidade, o polipeptídeo da superfamília das Ig é selecionada entre o grupo consistindo em uma região variável de cadeia pesada (V_H), uma região variável de cadeia leve (V_L), e um anticorpo de cadeia única (scFv).

Em outra modalidade, a dobra de Ig é selecionada entre o grupo consistindo em uma dobra classe V, uma dobra classe I, uma dobra classe C1, e uma dobra classe C2.

Em uma modalidade, a série de referência curada ou o alinhamento tem uma diversidade de no mínimo 50%.

Em uma modalidade, cada seqüência do alinhamento tem menos de 90% de identidade com cada outra seqüência do alinhamento.

Em outra modalidade, no mínimo uma seqüência é de uma espécie mamífera e no mínimo uma seqüência é de uma espécie não-mamífera.

Em uma modalidade, a resíduos da seqüência candidata são atribuídos escores de covariação posição específicos positivos para satisfazer covariações positivas encontradas em posições correspondentes no alinhamento.

Em uma modalidade, aos resíduos da seqüência candidata são

atribuídos escores de covariação posição específicos negativos para satisfazer covariações negativas covariações encontradas em posições correspondentes no alinhamento.

Em uma modalidade, a covariação positiva tem um coeficiente de associação phi (Φ) de cerca de +0,25 a cerca de +1,0.

Em uma modalidade, a covariação negativa tem um coeficiente de associação phi (Φ) de cerca de -0,25 a cerca de -1,0.

Em uma modalidade, o alinhamento é selecionado entre o grupo consistindo em um alinhamento de seqüência à base de estrutura, um alinhamento de seqüência à base de seqüência, e um alinhamento estrutural à base de estrutura.

Em uma modalidade, o escore de covariação posição específico é determinado para todos os pares de resíduos possíveis em cada posição de resíduo do alinhamento.

Em uma modalidade, o alinhamento é obtido (i) gerando um modelo HMM (Hidden Markov Model) a partir de um alinhamento inicial; (ii) adquirindo seqüências adicionais com o HMM; e (iii) alinhando as seqüências adicionais com o alinhamento inicial.

Em um aspecto, a invenção pertence a um programa de computador ou produto de programa de computador ou meio legível no computador tendo uma série de instruções executáveis por um processador as quais quando executadas fazem com que o processador realize um método da invenção.

Em ainda outra modalidade, a invenção pertence a um programa de computador ou produto de programa de computador ou meio legível no computador adequado para uso em um dispositivo eletrônico compreendendo dados correspondentes ao alinhamento do método da invenção.

Em ainda outra modalidade, a invenção pertence a um programa de computador ou produto de programa de computador ou meio legível no computador adequado para uso em um dispositivo eletrônico compreendendo os dados de covariação de um método da invenção.

Em uma modalidade, a invenção pertence a um dispositivo ar-

ranjado para realizar o método da invenção.

Em um aspecto, a invenção pertence a um sistema em um dispositivo de computação, compreendendo:

(i) um processo de compilação de seqüência inicial;

(ii) uma localização de armazenamento capaz de conterr dados de seqüência do alinhamento de qualquer uma das reivindicações 58 a 74;

(iii) uma facilidade de análise, a qual analisa programaticamente covariação entre pares de resíduos dentro do alinhamento para obter dados de covariação;

(iv) um dispositivo de saída com interface com o referido dispositivo eletrônico, o referido dispositivo de saída dando saída aos dados de covariação referidos.

Em uma modalidade, a invenção pertence a um meio em um dispositivo de computação contendo etapas executáveis para realizar um método compreendendo quaisquer das etapas do método da invenção.

Em ainda outra modalidade, o display de dados de freqüência de uso de resíduos compreende uma matriz compreendendo (i) posições de resíduos contíguos correspondentes à seqüência de polipeptídeo; e (ii) freqüências de uso de resíduos para um tipo de resíduo em cada posição de resíduo.

Em ainda outra modalidade, a freqüência de uso de resíduos é representada simbolicamente.

Em uma modalidade, todos os 20 aminoácidos naturais, intervalos, e resíduos ambíguos são representados. Em uma modalidade, freqüência de uso de resíduos se correlaciona com tamanho do símbolo.

Em uma modalidade, freqüência de uso de resíduos é correlacionada positivamente com tamanho do símbolo.

Em uma modalidade, (i) as posições de resíduos são apresentadas em colunas; e (ii) as freqüências de uso de resíduos são apresentadas em linhas.

Em outra modalidade, covariação entre os resíduos covariantes é representada por um revestimento de rede.

Em um aspecto, a invenção pertence a uma interface gráfica do usuário compreendendo um display gráfico de:

(i) um layout de grade compreendendo uma coleção de dados de freqüência de uso de resíduos para uma série de referência de seqüências de polipeptídeos; e

(ii) um revestimento de co-variação.

Em uma modalidade, a interface compreende adicionalmente: (iii) um display de uma seqüência de interesse.

Em um aspecto, a invenção pertence a um método para produzir um polipeptídeo da superfamília de imunoglobulinas (Ig) com uma propriedade biofísica aprimorada, o referido método compreendendo as etapas de:

a) proporcionar um alinhamento de uma série de referência curada de seqüências correspondentes a uma dobra de Ig do polipeptídeo;

b) calcular a covariação entre resíduos das seqüências do alinhamento para identificar resíduos covariáveis;

c) substituir um resíduo do polipeptídeo que falha em satisfazer uma covariação com o resíduo covariável encontrado em uma posição correspondente no alinhamento.

Em outra modalidade, a propriedade biofísica é selecionada entre o grupo consistindo em estabilidade térmica, perfil de desdobramento com pH, remoção estável de glicosilação, solubilidade, função bioquímica, e combinações das mesmas.

Em uma modalidade, a função bioquímica é selecionada entre o grupo consistindo em reconhecimento de um alvo protéico ou uma porção química, cuja porção química é selecionada entre o grupo consistindo em fosfato, metil, acetil, lipídeo, detergente, íon metálico, halogênio, RNA, DNA, oligonucleotídeo, e oligonucleosídeo.

Em uma modalidade, o polipeptídeo da superfamília das Ig é derivado de uma proteína selecionada entre o grupo consistindo em um anticorpo ou fragmento de ligação de antígeno do mesmo, um receptor de imunoglobulina, uma proteína de adesão celular, uma integrina, um alergeno, um receptor de células T, e uma molécula de complexo de histocompatibili-

dade maior (MHC).

Em uma modalidade, o polipeptídeo da superfamília das Ig é selecionada entre o grupo consistindo em uma região variável de cadeia pesada (V_H), uma região variável de cadeia leve (V_L), e um anticorpo de cadeia única (sFv).

Em uma modalidade, a dobra de Ig é selecionada entre o grupo consistindo em uma dobra classe V, uma dobra classe I, uma dobra classe C1, e uma dobra classe C2.

Em outra modalidade, o polipeptídeo da superfamília das Ig é um anticorpo modificado selecionado entre o grupo consistindo em um anticorpo de domínio, um anticorpo humanizado, um anticorpo humano, um anticorpo monoclonal não-humano, um anticorpo quimérico, um anticorpo biespecífico, um anticorpo contendo scFv, e um anticorpo com domínio eliminado.

Em uma modalidade, os resíduos covariáveis são parte de uma característica estrutural selecionada entre o grupo consistindo em uma ligação dissulfeto, uma ponte de sal, uma porção de uma superfície ou bolsa de ligação de ligante, e uma rede de van Der Waals, ligação hidrogênio, e/ou interações carga-carga.

Em outra modalidade, a invenção pertence a um método para produzir um anticorpo ou anticorpo modificado com uma aprimorada estabilidade, o referido método compreendendo as etapas de:

a) proporcionar um modelo estrutural de alta resolução de um anticorpo template que é experimentalmente validado como estável;

b) usando o modelo estrutural para identificar uma interface VH/VL e/ou resíduo de *estruturação* de interface no anticorpo modelo;

c) usando um modelo de homologia para identificar um resíduo da interface VH/VL correspondente do anticorpo ou anticorpo modificado que é importante para estabilizar a interface; e

d) substituir o resíduo da interface VH/VL correspondente no anticorpo ou anticorpo modificado com o resíduo da interface da proteína modelo.

Em uma modalidade, o resíduo da interface da proteína modelo encobre no mínimo 10 $Å^2$ de área superficial na interface.

Em uma modalidade, o anticorpo modificado é selecionado entre o grupo consistindo em um anticorpo de domínio, um anticorpo humanizado, um anticorpo humano, um anticorpo monoclonal não-humano, um anticorpo quimérico, um anticorpo biespecífico, um anticorpo contendo scFv, e um anticorpo com domínio eliminado.

Em outra modalidade, a invenção pertence a um polipeptídeo produzido pelo método da invenção.

Em uma modalidade, o polipeptídeo apresenta um aprimoramento em uma propriedade biofísica selecionada entre o grupo consistindo em estabilidade térmica, perfil de desdobramento com pH, remoção estável de glicosilação, solubilidade, função bioquímica, e combinações das mesmas.

Em uma modalidade, a função bioquímica é afinidade ou especificidade de ligação de ligante.

Em outro aspecto, a invenção pertence a um método para produzir uma biblioteca compreendendo moléculas scFv estabilizadas, o método compreendendo:

(a) proporcionar uma série de referência de seqüências correspondentes a uma seqüência de domínio variável de uma molécula scFv;

(b) identificar um aminoácido dentro do domínio variável o qual está ausente ou é encontrado em baixas freqüências em uma posição correspondente na série de referência; e

(c) combinar esta informação com dados de covariação sugerindo que modificar o aminoácido identificado com baixa freqüência na série de referência pode satisfazer limites de covariação existentes dentro do domínio variável; e

(d) substituir o aminoácido da etapa (b) com um aminoácido estabilizante candidato para deste modo produzir uma biblioteca compreendendo moléculas scFv estabilizadas.

Em uma modalidade, o aminoácido da etapa (b) tem uma classi-

ficação de consenso de menos de 0,5. Em outra modalidade, o aminoácido da etapa (b) está presente dentro de menos de 10% das seqüências na série de referência. Em outra modalidade, o aminoácido da etapa (b) é um resíduo não-consenso.

Em uma modalidade, o aminoácido estabilizante candidato é encontrado em uma posição correspondente dentro da série de referência.

Em ainda outra modalidade, o aminoácido estabilizante candidato é um aminoácido de consenso.

Em uma modalidade, o aminoácido estabilizante candidato é identificado por uma análise de uma estrutura 3-D da seqüência de região variável. Em outra modalidade, o aminoácido estabilizante é identificado por um cálculo de recondensação de cadeia lateral.

Em ainda outro aspecto, a invenção pertence a um método para produzir uma molécula scFv estabilizada, o método compreendendo:

(a) proporcionar uma biblioteca de scFv designada de acordo com o método da reivindicação 56,

(b) triar a biblioteca de scFv com uma prova de ensaio térmico para identificar moléculas scFv candidatas,

(c) comparar a estabilidade térmica de uma molécula scFv candidata da biblioteca de scFv com um controle adequado,

por meio do qual um aumento na estabilidade térmica de uma molécula scFv candidata em relação ao controle identifica a molécula scFv candidata como uma molécula scFv estabilizada.

Em outra modalidade, a invenção pertence a uma molécula scFv estabilizada identificado usando um método da invenção ou uma proteína de fusão compreendendo uma molécula scFv estabilizada da invenção.

Em outro aspecto, a invenção pertence a um método para prever a estabilidade de uma proteína candidata compreendendo uma seqüência de aminoácidos tendo seqüências de domínios candidatos, o método compreendendo:

(a) proporcionar uma série de referência de seqüências de aminoácidos correspondentes a uma seqüência de domínio candidato da proteí-

na candidata;

(b) determinar freqüências de resíduos em posições de aminoácidos individuais dentro da seqüência de domínio de ensaio para obter uma classificação de consenso; e

(c) usar a classificação de consenso para prever a estabilidade da proteína candidata, em que a classificação de consenso se correlaciona com a estabilidade da proteína candidata.

Em outro aspecto, a invenção pertence a um método para prever a estabilidade de uma proteína candidata, o método compreendendo:

(a) proporcionar uma série de referência de seqüências correspondentes a uma seqüência de domínio de ensaio da proteína candidata;

(b) determinar freqüências de resíduos em posições de aminoácidos individuais dentro da seqüência de domínio de ensaio para obter uma classificação de consenso; e

(c) usar a classificação de consenso para prever a estabilidade da proteína candidata, em que a classificação de consenso se correlaciona com a estabilidade da proteína candidata.

Em outra modalidade, a invenção pertence a um método para prever a estabilidade de uma proteína candidata, o método compreendendo:

(a) proporcionar uma série de referência de seqüências correspondentes a uma seqüência de domínio de ensaio da proteína candidata;

(b) determinar freqüências de resíduos em posições de aminoácidos dentro da seqüência de domínio de ensaio para obter uma classificação de consenso;

(c) determinar freqüências de resíduos em posições de aminoácidos correspondentes dentro das seqüências da série de referência para determinar uma classificação de consenso média;

(d) comparar a classificação de consenso com a classificação de consenso média para determinar um escore de seqüência; e

(e) usar o escore de seqüência para prever a estabilidade da proteína candidata, em que a classificação de consenso está correlacionada diretamente com a estabilidade da proteína candidata.

Em ainda outra modalidade, a invenção pertence a um método para prever a estabilidade do anticorpo candidato ou anticorpo modificado candidato, o método compreendendo:

(a) proporcionar uma série de referência de seqüências de domínios VH correspondentes a uma seqüência de domínio VH de ensaio do anticorpo ou anticorpo modificado candidato;

(b) determinar freqüências de resíduos em posições de aminoácido dentro da seqüência de domínio VH de ensaio para obter uma classificação de consenso;

(c) determinar freqüências de resíduos em posições de aminoácidos correspondentes dentro das seqüências da série de referência para determinar uma classificação de consenso média;

(d) comparar a classificação de consenso com a classificação de consenso média para determinar um escore de seqüência; e

(e) usar o escore de seqüência para prever a estabilidade do anticorpo ou anticorpo modificado candidato, em que a classificação de consenso está diretamente correlacionada com a estabilidade do anticorpo ou anticorpo modificado candidato.

Em uma modalidade, a série de referência compreende seqüências de proteína de proteínas com a mesma classe de dobra de proteína que a proteína candidata.

Em outra modalidade, a série de referência compreende seqüências ortólogas.

Em uma modalidade, a série de referência compreende seqüências humanas, por exemplo, seqüências de VH humanas, por exemplo, da mesma classe de Kabat.

Em uma modalidade, a série de referência compreende seqüências de linha germinal de VH humano.

Em uma modalidade, a seqüência de domínio de ensaio é uma porção da seqüência de domínio da proteína candidata.

Em uma modalidade, a classificação de consenso é determinada para cada posição da seqüência de ensaio para obter uma classificação

de consenso total, em que a classificação de consenso total se correlaciona com a estabilidade da proteína.

Em outra modalidade, a classificação de consenso total é calculada usando a fórmula:

$$\text{escore} = \sum_i \frac{h_i(r)}{c_i(r)}$$

em que:

$c_i(r)$ equivale à freqüência do resíduo de consenso em uma posição de aminoácido da seqüência de consenso,

$h_i(r)$ equivale à freqüência do resíduo de ensaio de uma posição de aminoácido da seqüência de ensaio, e

i equivale ao número de posições de aminoácido dentro da seqüência de ensaio.

Em uma modalidade, a estabilidade da proteína é determinada comparando a classificação de consenso da proteína candidata com uma classificação de consenso de um controle adequado.

Em outra modalidade, a estabilidade da proteína é determinada comparando a classificação de consenso da proteína candidata com a classificação de consenso da proteína candidata perfeita.

Em ainda outra modalidade, a estabilidade é determinada comparando o escore da seqüência da proteína candidata com um controle adequado.

Em outra modalidade, a estabilidade é determinada comparando o escore da seqüência do anticorpo ou anticorpo modificado candidato com um controle adequado.

Em uma modalidade, a proteína é uma proteína multidomínio, e em que a seqüência de domínio alvo é derivada do domínio menos estável da proteína multidomínio.

Em ainda outra modalidade, a proteína é um anticorpo modificado selecionado entre o grupo consistindo em um anticorpo camelídeo, anticorpo humanizado, um anticorpo humano, um anticorpo monoclonal não-humano, um anticorpo quimérico, um anticorpo biespecífico, um anticorpo

contendo scFv, e um anticorpo com domínio eliminado.

Em outra modalidade, a invenção pertence a um método para selecionar uma proteína candidata para expressão, compreendendo determinar a estabilidade da proteína candidata usando o método da invenção, em que a proteína candidata é selecionada para expressão se sua classificação de consenso ou seu escore de seqüência for preditivo de alta estabilidade.

Em outra modalidade, a invenção pertence a um método para selecionar uma proteína candidata para expressão, compreendendo determinar a estabilidade da proteína candidata usando o método da invenção, em que a proteína candidata é selecionada para expressão se seu escore de covariação específico da posição da seqüência for preditivo de alta estabilidade.

Em outra modalidade, a invenção pertence a um método para selecionar uma seqüência de ensaio de região variável de imunoglobulina aceitadora para uso na humanização de um anticorpo doador, o método compreendendo determinar a estabilidade da seqüência candidata usando o método da invenção, em que a seqüência candidata é selecionada se seu escore de seqüência for preditivo de alta estabilidade.

Em ainda outra modalidade, a invenção pertence a um método para selecionar uma seqüência de ensaio de região variável de imunoglobulina aceitadora para uso na humanização de um anticorpo do doador, o método compreendendo determinar a estabilidade da seqüência candidata usando o método da invenção, em que a seqüência candidata é selecionada de seu escore de covariação posição específica é preditiva de alta estabilidade.

Em outra modalidade, a seqüência de ensaio é uma seqüência de linha germinal humana.

Breve Descrição dos Desenhos

A Figura 1A mostra a seqüência de DNA (SEQ ID NO: 3) codificando um construto BHA10 scFv convencional compreendendo um ligante (Gly4Ser)3 (indicado em tipo em negrito). A Figura 1B mostra a seqüência

de aminoácidos (SEQ ID NO: 4) do construto BHA10 scFv convencional.

A Figura 2 representa os resultados de medições de calorimetria de varredura diferencial (DSC) com fragmentos Fab ou scFv convencionais purificados de BHA10. A Figura 2A representa os resultados de um estudo por DSC comparando o desdobramento de fragmentos Fab e scFv convencionais de BHA10. A Figura 2B representa os resultados de varreduras por DSC do scFv de BHA10 purificado em taxas de varredura de 1°C/min e 2°C/min.

A Figura 3A representa a seqüência de DNA (SEQ ID NO: 9) codificando um construto BHA10 scFv estabilizado com VH44/VL100 dissulfeto estabilizado compreendendo um ligante (Gly4Ser)3 (indicado em tipo em negrito). A Figura 3B mostra a seqüência de aminoácidos (SEQ ID NO: 10) do construto BHA10 scFv estabilizado com VH44/VL100 dissulfeto. Os resíduos cisteína formando a ligação dissulfeto VH44/VL100 são indicado em tipo em negrito e em itálico.

A Figura 4A representa a seqüência de DNA (SEQ ID NO: 14) codificando um BHA10 scFv estabilizado compreendendo um ligante $(Gly_4Ser)_4$ (indicado em tipo em negrito). A Figura 4B mostra a seqüência de aminoácidos (SEQ ID NO: 15) do construto $(Gly_4Ser)_4$ BHA10 scFv. A anotação é a mesma que na Figura 4A.

A Figura 5A representa a seqüência de DNA (SEQ ID NO: 16) codificando um BHA10 scFv compreendendo um ligante $(Gly_4Ser)_5$ (indicado em tipo em negrito). A Figura 5B representa a seqüência de aminoácidos (SEQ ID NO: 17) do construto $(Gly_4Ser)_5$ BHA10 scFv. A anotação é a mesma que na Figura 5A.

A Figura 6 representa os resultados de uma análise Western Blot para comparar os níveis de expressão de um BHA10 scFv convencional (alameda 1) e as moléculas BHA10 scFv establilizadas da invenção (alamedas 2-4). Cada amostra foi eletroforesada sob tanto condições redutoras (painel esquerdo) quanto não-redutoras (painel direito) .

A Figura 7A representa a seqüência de DNA (SEQ ID NO: 18) codificando um BHA10 scFv estabilizado com VH44/VL100 dissulfeto com-

preendendo um ligante $(Gly_4Ser)_4$ (indicado em tipo em negrito). A Figura 7B representa a seqüência de aminoácidos (SEQ ID NO: 19) do VH44/VL100 dissulfeto + BHA10 scFv estabilizado com $(Gly_4Ser)_4$. A anotação é a mesma que na Figura 7A. Os resíduos cisteína formando a ligação dissulfeto VH44/VL100 são indicados em tipo em negrito e em itálico.

A Figura 8 representa os resultados de uma prova de ensaio térmico na qual as estabilidades térmicas das moléculas scFv BHA10 estabilizadas da invenção foram comparadas com a de uma molécula scFv BHA10 convencional. A temperatura na qual 50% das moléculas scFv conservam sua atividade de ligação (T50) é indicada na legenda da figura.

A Figura 9 representa os resultados de análises por Calorimetria de Varredura Diferencial (DSC) realizadas com um BHA10 scFv convencional e um BHA10 $V_H44:V_L100/(Gly_4Ser)_4$ scFv.

A Figura 10 mostra resultados de ligação do corante fluorescente hidrofóbico sulfonato de 1-anilino-8-naftalina (ANS) a BHA10 scFv convencional, BHA10 $(Gly_4Ser)_4$ scFv, e BHA10 $V_H44:V_L100/(Gly_4Ser)_4$ scFv.

A Figura 11 representa os resultados de análise de freqüência de resíduos dos domínios VH e VL de BHA10. A Figura 11A lista posições de biblioteca de VH de BHA10 para triagem com base em análise de freqüência de resíduos de seqüências de domínio variável de IgG. A Figura 11B lista posições de biblioteca de VL de BHA10 para triagem com base em análise de freqüência de resíduos de seqüências de domínio variável de IgG.

A Figura 12 representa os resultados de uma prova de ensaio térmico na qual foram avaliados os efeitos de estabilizar mutações de BHA10 scFv sobre estabilidade térmica e a aditividade de mutações estabilizantes. (GS3 deonota $(G_4S)3$ e GS4 denota $(G_4S)4$; SS denota uma ligação dissulfeto).

A Figura 13 representa anticorpos biespecíficos LTβR/TRAIL-R2 estabilizados exemplares ("Hercules" anticorpos) da invenção. Os anticorpos Hercules são formados pela fusão de uma molécula scFv de BHA10 estabilizada da invenção a um anticorpo de IgG 14A2. A molécula scFv pode ser fundida ao C-término ou N-término da cadeia pesada (C-Hercules ou N_H-

Hercules) ou ao N-término da cadeia leve (N_L-Hercules).

A Figura 14 é um diagrama esquemático representando as reações de PCR seqüenciais usadas para fundir BHA10 scFvs convencionais e produzidos ao término amino (Figura 14A) ou término carboxila (Figura 14B) de uma cadeia pesada 14A2.

A Figura 15A representa a seqüência de DNA (SEQ ID NO: 28) de uma cadeia leve 14A2 quimérica compreendendo peptídeo de sinal (sublinhado). A Figura 15B mostra a seqüência de aminoácidos (SEQ ID NO: 29) de cadeia leve 14A2 quimérica.

A Figura 16 representa a seqüência de DNA (SEQ ID NO: 30) da cadeia pesada de um BHA10 scFv N_H-Hercules convencional.

A Figura 17 representa a seqüência de aminoácidos (SEQ ID NO: 31) da cadeia pesada de um BHA10 scFv N_H-Hercules convencional.

A Figura 18 representa a seqüência de DNA (SEQ ID NO: 32) da cadeia pesada de um BHA10 scFv $(Gly_4Ser)_4$ N_H- Hercules.

A Figura 19 representa a seqüência de aminoácidos (SEQ ID NO: 33) da cadeia pesada de um BHA10 scFv $(Gly_4Ser)_4$ N_H-Hercules.

A Figura 20 representa a seqüência de DNA (SEQ ID NO: 34) da cadeia pesada de um BHA10 scFv V_H44:V_L100 N-Hercules.

A Figura 21 representa a seqüência de aminoácidos (SEQ ID NO: 35) da cadeia pesada de um BHA10 scFv V_H44:V_L100 N_H-Hercules.

A Figura 22 representa a seqüência de DNA (SEQ ID NO: 36) da cadeia pesada de um BHA10 scFv V_H44:V_L100/$(Gly_4Ser)_4$ N_H-Hercules.

A Figura 23 representa a seqüência de aminoácidos (SEQ ID NO: 37) da cadeia pesada de um BHA10 scFv V_H44:V_L100/$(Gly_4Ser)_4$ N_H-Hercules.

A Figura 24 representa a seqüência de DNA (SEQ ID NO: 44) da cadeia pesada de um BHA10 scFv C-Hercules convencional.

A Figura 25 representa a seqüência de aminoácidos (SEQ ID NO: 45) da cadeia pesada de um BHA10 scFv C-Hercules convencional.

A Figura 26 representa a seqüência de DNA (SEQ ID NO: 46) da cadeia pesada de um BHA10 scFv $(Gly_4Ser)_4$ C- Hercules.

A Figura 27 representa a seqüência de aminoácidos (SEQ ID NO: 47) da cadeia pesada de um BHA10 scFv $(Gly_4Ser)_4$ C-Hercules.

A Figura 28 representa a seqüência de DNA (SEQ ID NO: 48) da cadeia pesada de um BHA10 scFv V_H44:V_L100 C-Hercules.

A Figura 29 representa a seqüência de aminoácidos (SEQ ID NO: 49) da cadeia pesada de um BHA10 scFv V_H44:V_L100 C-Hercules.

A Figura 30 representa a seqüência de DNA (SEQ ID NO: 50) da cadeia pesada de um BHA10 scFv V_H44:V_L100/$(Gly_4Ser)_4$ C-Hercules.

A Figura 31 representa a seqüência de aminoácidos (SEQ ID NO: 51) da cadeia pesada de um BHA10 scFv V_H44:V_L100/$(Gly_4Ser)_4$ C-Hercules.

A Figura 32 mostra resultados de análises Western Blot de anticorpos biespecíficos C- e N_H- Hercules expressados transitoriamente em células CHO. O painel esquerdo é analisado sob condições não-redutoras e o painel direito sob condições redutoras.

A Figura 33 representa os resultados de um ELISA para avaliar a atividade de ligação dos anticorpos Hercules estabilizados da invenção a R2 receptores TRAIL. A Figura 33A representa os resultados com anticorpos C-Hercules estabilizados. A Figura 33B representa os resultados com anticorpos N_H Hercules estabilizados. "wt" é anticorpo N_H Hercules convencional. "GS4" é um anticorpo N_H Hercules establilizado compreendendo um (Gly4Ser)4 scFv estabilizado. "ds" é um anticorpo N_H Hercules estabilizado compreendendo um scFv estabilizado compreendendo o ligante de dissulfeto VH44/VL100.

A Figura 34 representa os resultados de um ELISA para avaliar a atividade de ligação dos anticorpos Hercules estabilizados da invenção a LTβR receptores. A Figura 34A representa os resultados com anticorpos Hercules de C-terminal estabilizados. A Figura 34B representa os resultados com anticorpos N_H Hercules estabilizados. "wt" é anticorpo N_H Hercules convencional. "GS4" é um anticorpo N_H Hercules estabilizado compreendendo um $(Gly_4Ser)_4$ scFv estabilizado. "ds" é um anticorpo N_H Hercules estabilizado compreendendo um scFv estabilizadp compreendendo o ligante de dis-

sulfeto VH44/VL100.

A Figura 35 mostra os resultados de análise SEC de anticorpos biespecíficos C-Hercules estabilizados depois de cromatografia de Proteína A.

A Figura 36 representa os resultados de análise SDS-PAGE de anticorpos biespecíficos Hercules estabilizados purificados. A Figura 36A representa resultados com anticorpos biespecíficos NH-Hercules. A Figura 36B representa resultados com anticorpos biespecíficos C-Hercules. 5 ug de amostra foram carregados em cada alameda.

A Figura 37 mostra os resultados de análise SEC analítica de anticorpos biespecíficos N-Hercules estabilizados purificados. O Painel A mostra o perfil de N-Hercules com V_H44:V_L100 BHA10 scFv depois de cromatografia de Proteína A e o Painel B depois de SEC do preparativo. O Painel C mostra o perfil de N-Hercules com V_H44:$V_L100/(G_4S)_4$ BHA10 scFv depois de cromatografia de Proteína A e o Painel D depois de SEC do preparativo.

A Figura 38 mostra os resultados de análise SEC analítica de anticorpos biespecíficos C-Hercules estabilizados purificados. O Painel A mostra o perfil de C-Hercules com V_H44:V_L100 BHA10 scFv depois de cromatografia de Proteína A e o Painel B depois de SEC do preparativo. O Painel C mostra o perfil de C-Hercules com V_H44:$V_L100/(G_4S)_4$ BHA10 scFv depois de cromatografia de Proteína A e o Painel D depois de SEC do preparativo.

A Figura 39 mostra resultados de análises por Calorimetria de Varredura Diferencial (DSC) realizadas com C-Hercules com V_H44:$V_L100/(G_4S)_4$ BHA10 scFv e C-Hercules com BHA10 scFv convencional.

A Figura 40 representa os resultados de um ELISA para avaliar a atividade de ligação biespecífica dos anticorpos Hercules estabilizados da invenção para TRAIL-R2 e LTβR Ig receptores. "N-"Hercules" Cys e C-"Hercules" Cys são anticorpos Hercules compreendendo scFvs de N_H- ou C-terminal tendo a ligação dissulfeto VH44/VL100. "N-"Hercules"Cys e N-"Hercules" DM" são anticorpos Hercules compreendendo o scFv de N_H- ou C-terminal compreendendo o dissulfeto VH44/VL100 e o ligante (Gly4Ser)4.

O pentâmero CBE11 é um anticorpo controle com atividade de ligação de LTβR monoespecífica.

A Figura 41 representa os efeitos de anticorpos biespecíficos e monoespecíficos N- e C-Hercules estabilizados sobre o crescimento de células tumorais. As Figuras 41A a 41D representam efeitos sobre o crescimento de células tumorais WiDr, células tumorais Me180, células tumorais MDA231, e células HUVEC, respectivamente.

A Figura 42 representa um diagrama esquemático de um "minibody" dimérico de duas cadeias estabilizado compreendendo scFvs estabilizadas. O "minibody" exemplar contém uma primeira porção de cadeia compreendendo um scFv estabilizado com especificidade de ligação para antígeno TRAIL R2 e uma segunda porção de cadeia compreendendo um scFv estabilizado com especificidade de ligação para o antígeno LTβR. A orientação dos domínios VH e VL no scFv pode ser alterada e as especificidades de ligação respectivas podem ser alteradas.

A Figura 43 representa um diagrama esquemático de um "minibody" tetravalente dimérico de duas cadeias biespecífico estabilizado exemplar ("minibody" tetravalente N-scFv estabilizado biespecífico) compreendendo fragmentos de scFv estabilizado da invenção anexados aos términos amino. O "minibody" tetravalente bivalente estabilizado exemplar contém um primeira porção de cadeia compreendendo 2 scFvs estabilizados com especificidade de ligação para antígeno TRAIL R2 e uma segunda porção de cadeia compreendendo 2 scFvs estabilizados com especificidade de ligação para o antígeno LTβR. Outras configurações também são possíveis, por exemplo, o "minibody" tetravalente biespecífico também pode ser construído de tal modo que cada porção de cadeia contém 2 fragmentos de scFv estabilizados com diferentes especificidades. Em outra modalidade, a orientação dos domínios VH e VL no scFv pode ser alterada. Em outra modalidade, menos do que todos os scFvs são estabilizados.

A Figura 44 representa um diagrama esquemático de um "minibody" tetravalente dimérico de duas cadeias biespecífico estabilizado exemplar ("minibody" tetravalente C-scFv biespecífico estabilizado) compreen-

dendo fragmentos de scFv estabilizado anexados a ambos os términos carboxila de um "minibody" bivalente. O "minibody" tetravalente bivalente estabilizado exemplar contém uma primeira porção de cadeia compreendendo 2 scFvs estabilizados com especificidade de ligação para antígeno TRAIL R2 e uma segunda porção de cadeia compreendendo 2 scFvs estabilizados com especificidade de ligação para o antígeno LTβR. Outras configurações também são possíveis, por exemplo, os minibodies tetravalentes diméricos de duas cadeias biespecíficos também podem ser construídos de tal modo que cada cadeia contém 2 fragmentos de scFv estabilizada com diferentes especificidades. Em outra modalidade, a orientação dos domínios domínios VH e VL no scFv pode ser alterada. Em outra modalidade, menos do que todos os scFvs são estabilizados.

A Figura 45 mostra um diagrama esquemático de um diabody dimérico de quatro cadeias biespecífico estabilizado compreendendo scFvs estabilizados da invenção. Outras configurações também são possíveis, por exemplo, o "minibody" tetravalente dimérico de duas cadeias biespecífico estabilizado também pode ser construído de tal modo que cada braço contém fragmentos de scFv estabilizados com diferentes especificidades. A orientação dos domínios VH e VL pode ser alterada. Em outra modalidade, menos do que todos os scFvs são estabilizados.

A Figura 46 mostra um diagrama esquemático de um anticorpo scFv tetravalente dimérico de quatro cadeias biespecífico estabilizado (anticorpo tetravalente C-scFv estabilizado) compreendendo um scFv estabilizado anexado ao término carboxila de CH3 e um peptídeo de ligação de articulação. A orientação dos domínios VH e VL no scFv estabilizado pode ser alterada. Alternativamente, os fragmentos de scFv estabilizados podem ser anexados aos términos amino ou das porções de cadeia pesada ou leve para formar anticorpos tetravalentes N_H-scFv ou anticorpos tetravalentes N_L-scFv, respectivamente.

A Figura 47 mostra um diagrama esquemático de um anticorpo biespecífico eliminado o domínio CH2 scFv tetravalente de quatro cadeias establizado (anticorpo biespecífico eliminado o domínio CH2 tetravalente C-

scFv estabilizado) compreendendo um scFv estabilizado anexado ao término carboxila de CH3 e um peptídeo de ligação de articulação. Cada cadeia pesada porção do anticorpo biespecífico contém uma região Fv com especificidade de ligação para o antígeno TRAIL R2 e uma região estabilizada scFv com especificidade de ligação para o antígeno LTβR. A orientação dos domínios VH e VL no scFv estabilizado pode ser alterada e as especificidades de ligação antigênica respectivas podem ser alteradas. Em outra modalidade, menos do que todos os scFvs são estabilizados.

A Figura 48 mostra um diagrama esquemático de um anticorpo biespecífico eliminado o domínio CH2 scFv tetravalente de quatro cadeias estabilizado (anticorpo biespecífico eliminado o domínio CH2 tetravalente N_H-scFv estabilizado) compreendendo um scFv estabilizado anexado ao término amino de VH e compreendendo um peptídeo de ligação de articulação. Cada cadeia pesada porção do anticorpo biespecífico contém uma região Fv com especificidade de ligação para o antígeno TRAIL R2 e uma região scFv estabilizada com especificidade de ligação para o antígeno LTβR. A orientação dos domínios VH e VL no scFv estabilizado pode ser alterada e as especificidades de ligação antigênica respectivas podem ser alteradas. Em outra modalidade, menos do que todos os scFvs são estabilizados.

A Figura 49 mostra um diagrama esquemático de um anticorpo biespecífico eliminado o domínio CH2 scFv tetravalente de quatro cadeias estabilizado (anticorpo biespecífico eliminado o domínio CH2 tetravalente N_L-scFv estabilizado) compreendendo um scFv estabilizado anexado ao término amino de VL e compreendendo um peptídeo de ligação de articulação. Cada cadeia pesada porção do anticorpo biespecífico contém uma região Fv com especificidade de ligação para o antígeno TRAIL R2 e uma região scFv estabilizada com especificidade de ligação para o antígeno LTβR. A orientação dos domínios VH e VL na scFv estabilizada pode ser alterada e as especificidades de ligação antigênica respectivas podem ser alteradas. Em outra modalidade, menos do que todos os scFvs são estabilizados.

A Figura 50 representa curvas de T50 de BHA10 scFv selvagem (círculos fechados) versus BHA10 scFVs contendo o VL_S46L (Figura 15A,

círculos abertos), VH_V55G (Figura 15B, círculos abertos), e VH P101D (círculos abertos) (Figura 15C) mutações estabilizantes.

A Figura 51 representa curvas de DSC do VH-(GS)4-VL-6His BHA10 scFv selvagem (linha cinza) versus BH10 scFvs contendo as mutações estabilizantes VL S46L (linha preta) (Figura 51A) e VH V55G (linha preta) (Figura 51B).

A Figura 52 representa curvas de DSC do VH-(GS)4-VL-6His BHA10 scFv selvagem (linha cinza) versus a BH10 scFv contendo contendo a mutação estabilizante VH_P101D (linha preta).

A Figura 53 representa uma análise SDS-Page dos scFvs selvagens ("WT (G4S)X3") e mutantes estabilizados. Todos os scFvs mutantes contêm um peptídeo de ligação gly-ser da fórmula $(Gly_4\ Ser)_4$ entre o VH de N-terminal e o VL de C-terminal ("(G4S)X4") . A identidade da scFv em cada alameda é representada. As alamedas 12-14 contêm BHA10 scFv tendo uma ligação dissulfeto entre VH44 e VL110. A (G4S)X4/dissulfeto scFv estabilizada são denotadas 'DM' para 'duplo mutante'. As duas alamedas finais incluem mutações designadas que também foram estabilizantes para scFvs encadeado a dissulfeto.

A Figura 54A representa curvas de DSC do ("VH-(GS)4-VL-6His") BHA10 scFv selvagem (linha tracejada), scFv mutante VH S16E (linha cinza delgada), scFv mutante VL S46L (linha preta delgada), scFv mutante VH V55G (linha cinza espessa), e scFv mutante VH P101D (linha preta espessa). A Figura 54B representa a fluorescência dependente da temperatura de 10 mM de ANS em PBS (círculos abertos) ou na presença de BHA10 scFv selvagem ("VH-(GS)4-VL-6His") (círculos cinzas fechados), scFv mutante VH S16E (losangos), scFv mutante VL S46L (quadrados), scFv mutante VH V55G (triângulos invertidos), e scFv mutante VH P101D (triângulos). As linhas através das curvas são os ajustes para um modelo de desdobramento de dois estados.

A Figura 55 representa fluorescência ANS em 15 °C na presença de scFvs selavgens e mutantes. A fluorescência ANS induzida por cada scFv é plotada contra a TM do domínio VH do mesmo scFv conforme deter-

minado por DSC (Figura 55A) e o aumento dependente de temperatura em fluorescência ANS devida à transição de desdobramento de cada scFv (Figura 55B).

A Figura 56 representa cromatografia de exclusão de tamanho (SEC) de altas concentrações de N-Hercules estabilizado (XWU028; BHA10 scFv V_H44:V_L100/$(Gly_4Ser)_4$ N- Hercules; Figura 56A), C-Hercules estabilizado (XWU036; (BHA10 scFv V_H44:V_L100/$(Gly_4Ser)_4$ C-Hercules; Figura 56B) e o anticorpo BHA10 selvagem (Figura 56C) em seis pontos de tempo (T = 0, T = 1 semana, T = 2 semanas, T = 1 mo, T = 2 mo, e T = 3 mo) e baixas temperaturas (2 a 8 °C). O eixo x é tempo (minutos) e o eixo y é mAU (unidades de absorvência) em 280 nm. Não foram observadas alterações significativas no perfil de eluição com o curso de tempo.

A Figura 57 representa análise SDS-PAGE de XWU028 e XWU036 antes e depois de 3 meses de armazenamento em 2 a 8 ºC. Painel A. As amostras foram não reduzidas (designadas como "NR"). O MW esperado, ~200 kDa, é observado para as amostras biespecíficas. BHA10 roda em seu MW esperado de ~150 kDa. As alamedas 1 a 4 são amostras XWU028. As alamedas 5 a 8 são amostras XWU036. As alamedas 9 a 10 são amostras BHA10 de IgG. Painel B. As amostras foram reduzidas com DTT (designadas como "Red" para reduzidas). Os MWs esperados da cadeia pesada, ~75 kDa, e da cadeia leve, ~25 kDa, são observados para as amostras biespecíficas sob condições redutoras. As cadeias pesadas e leves de BHA10 rodaram em seus MWs esperados de 50 kDa e 25 kDa, respectivamente, sob condições redutoras. As alamedas 1, 2, 6, e 7 são amostras XWU028. As alamedas 3, 4, 8, e 9 são amostras XWU036. As alamedas 5 e 10 são amostras BHA10 de IgG.

A Figura 58 representa análises de massa intacta de XWU028 (Painel A) e XWU036 (Painel B) em T = 0, T = 1 mo, e T = 3 mo. Múltiplos picos são observados para cada proteína devido a vários níveis de sialilação dos carboidratos N-encadeados no CH2. A distribuição de carboidratos dos biespecíficos é típica de proteínas IgG1 de rotina.

A Figura 59 representa os resultados de um ensaio ELISA san-

duíche medindo ligação biespecífica de amostras de soro contendo Hercules de N- terminal (XWU028; Figura 56A) ou Hercules de C- terminal (XWU036; Figura 56B) a TRAIL-R2 e LTβR receptores.

A Figura 60 representa os resultados de um experimento avaliando a atividade *in vivo* relativa de anticorpos biespecíficos estabilizados (XWU028 (losangos) e XWU036 (quadrados abertos)), anticorpos biespecíficos não estabilizados (hCBE11 (triângulo fechado)) e anticorpos monoespecíficos administrados sozinhos (hBHA10 (triângulo invertido) e ch14A2 (losango)) e em combinação (triângulo aberto) contra um modelo de camundongo de xenoenxerto tumoral. A dosagem começou no dia 13. hCBE11 v. VC: $P < 0,001$; hBHA10 v. VC: $P < 0,001$; ch14A2 v. VC: $P < 0,01$; Hercules-II XWU028 v. VC. : $P < 0,001$; Hercules-II XWU028 v.hBHA10: $P < 0,05$; Hercules-II XWU028 v. ch14A2: $P < 0,005$; Hercules-II XWU036 v. VC: $P < 0,001$; Hercules-II XWU036 v. hBHA10: $P < 0,01$; Hercules-II XWU036 v. ch14A2: $P < 0,05$; combo v. V.C. : $P < 0,001$

A Figura 61 representa a seqüência de DNA de cadeia pesada (SEQ ID NO: 52) para anticorpo biespecífico C-Hercules BHA10 scFv V_H S16E + V_L S46L.

A Figura 62 representa a seqüência de aminoácidos de cadeia pesada (SEQ ID NO: 53) para anticorpo biespecífico C-Hercules BHA10 scFv V_H S16E + V_L S46L.

A Figura 63 representa a seqüência de DNA de cadeia pesada (SEQ ID NO: 54) para anticorpo biespecífico C-Hercules BHA10 scFv V_H44-V_L100/V_H S16E + V_L S46L.

A Figura 64 representa a seqüência de aminoácidos de cadeia pesada (SEQ ID NO: 55) para anticorpo biespecífico C-Hercules BHA10 scFv V_H44-V_L100/V_H S16E + V_L S46L.

A Figura 65 representa os resultados de cromatografia de exclusão de tamanho analítica (SEC) de elutriados de Proteína A de sobrenadantes contendo C-Hercules com o V_H S16E estabilizado + V_L S46L BHA10 scFv (XWU054) e C- Hercules com ("WT") BHA10 scFv convencional.

A Figura 66 representa géis SDS-PAGE de Hercules de C-

terminal purificado com o V_H S16E + V_L S46L BHA10 scFv (Figura 66A) e Hercules de C-terminal purificado com o V_H44:V_L100/V_H S16E + V_L S46L BHA10 scFv (Figura 66B).

A Figura 67 mostra os perfis de eluição SEC analítica de Hercules de C-terminal com o V_H S16E + V_L S46L BHA10 scFv (Figura 67A) e Hercules de C-terminal com o V_H44:V_L100/V_H S16E + V_L S46L BHA10 scFv (Figura 67B) subseqüente à etapa de purificação de proteína A inicial.

A Figura 68 mostra os resultados de provas de proliferação de células tumorais *in vitro* da atividade de Hercules de C-terminal com o V_H S16E + V_L S46L BHA10 scFv (XWU054) e Hercules de C-terminal com o V_H44:V_L100/V_H S16E + V_L S46L BHA10 scFv (XWU055). Tanto a células MDA231 (Figura 68A) quanto WiDr (Figura 68B) foram administradas anticorpo. A atividade para ambos os anticorpos Hercules foi comparada com anticorpo biespecífico XWU036 estabilizado e anticorpos monoespecíficos correspondentes (hBHA10 e ch14A2) administrados sozinhos ou em combinação.

A Figura 69A representa uma curva de desdobramento de DSC de BIIB7 IgG1 tomada sob condições idênticas às usadas para os 18 anticorpos humanos(humanizados) BIIB remanescentes (linha sólida). O Fab individual, transições C_H2 e C_H3 aparentes dentro da curva de DSC são marcadas e esboçadas com linhas pontilhadas. A Figura 69B representa as desdobramento curvas de desdobramento de DSC do anticorpo BIIB7 em ambos os formatos IgG1 e IgG4.

A Figura 70 representa curvas de desdobramento térmico de quatro anticorpos IgG1 humanos(humanizados), BIIB16, BIIB6, BIIB4 e BIIB1. As transições de desdobramento dos domínios C_H2 e C_H3 para todos os quatro construtos de IgG1 são idênticas, enquanto as transições de desdobramento de Fab são altamente variáveis.

A Figura 71 representa resultados de classificação de consenso para 182 seqüências de referência de domínio variável de anticorpo humano e as seqüências de consenso de subclasses individuais. A. Escores de seqüência subclasse V_H derivadas por análise de freqüência de resíduos con-

tra um banco de dados de V_H mamífero. B. Distribuição de escores de seqüência de V_H de referência selecionada entre NCBI. Escores de seqüência de V_H são agrupados de acordo com a subclasse. C. Escores de seqüência de consenso subclasse V_k derivados por análise de freqüência de resíduos contra um banco de dados de V_k mamífero. D. Distribuição de escores de seqüência de V_k de referência selecionada entre NCBI. Escores de seqüência de V_k são agrupados de acordo com a subclasse.

A Figura 72 representa gráficos bidimensionais comparando escores de seqüência para 18 anticorpos BIIB com os Fab T_Ms correspondentes conforme medido por DSC. A Figura 72A representa escores de BIIB1-18 V_H plotados contra as T_Ms de Fab. A identidade de cada ponto pode ser encontrada na Tabela 20. Os V_Hs de BIIB15 e BIIB16 contêm inserções / eliminações incomuns em suas CDRs. Os resultados para BIIB15 e BIIB16 são apresentados como quadrados. A T_M de BIIB18 foi etiquetada artificialmente como 57,2 °C (para combinar a menor T_M medida de BIIB17) uma vez que sua própria T_M foi imensurável devido à falta de expressão. A T_M de BIIB18 é provavelmente muito menor com base em seu escore de seqüência extremamente baixo e sua incapacidade para expressar. Os resultados de Fab para BIIB18 são apresentados como um triângulo. Os resultados de V_H de anticorpos BIIB restantes são apresentados como círculos. A Figura 72B representa escores de V_k de BIIB1-18 plotados contra as T_Ms de Fab. Os símbolos são os mesmos que descritos para o gráfico (A), exceto que os escores são para V_k.

A Figura 73 representa estruturas representativas para a classe V, classe C, e classe I de dobras de Ig.

A Figura 74 representa um histograma das extensões de seqüência de Dobras de Ig classe V, I, C1, e C2 obtidas da SCOP.

A Figura 75 representa um histograma de coeficientes de correlação de valor f calculado para cada combinação de resíduos dentro da série de dados classe C.

A Figura 76 mostra o resultado da ferramenta visual NAPMAP (NAPMAP Visual Tool) a qual inclui uma porção da plataforma de alinhamen-

to soba a qual a análise de covariação é revestida. A plataforma apresenta colunas contendo todos os aminoácidos incluindo intervalos (tipo de resíduo) para cada posição dentro da região V (nº da posição do resíduo).

A Figura 77 representa um revestimento de estatística de covariação de Dobras de Ig sobre o modelo NAPMAP. A seqüência de interesse é mostrada como uma linha curvada através dos aminoácidos relevantes. Covariações são mostradas como barras retas entre os resíduos aminoácidos covariáveis relevantes.

A Figura 78 mostra um detalhe da análise de covariação de BHA10 V_L salientando a covariação entre Tyr na posição de Kabat 36 (posição de resíduo nº 48) e a mutação Ser→Leu na posição de Kabat 46 (posição de resíduo nº 67) conforme descrito no Exemplo 7.

A Figura 79 mostra um detalhe da análise de covariação de BHA10 V_H salientando a covariação conflitante entre Val na posição de Kabat 67 (posição de resíduo nº 100), Thr na posição de Kabat 70 (posição de resíduo nº 104), e a mutação Met→Leu na posição de Kabat 80 (posição de resíduo nº 116).

A Figura 80 (SEQ ID NO: 63) mostra a seqüência de DNA de filamento único de anticorpo p5E8 PRIMATIZED® tetravalente de C-terminal de cadeia pesada compreendendo um scFv convencional.

A Figura 81 (SEQ ID NO: 64) mostra a seqüência de aminoácidos de anticorpo p5E8 PRIMATIZED® tetravalente de C-terminal de cadeia pesada compreendendo um scFv convencional.

A Figura 82A (SEQ ID NO: 65) mostra a seqüência de DNA de filamento único de cadeia leve p5E8 PRIMATIZED®. Seqüência de peptídeo de sinal é mostrada como sublinhada. A Figura 82B (SEQ ID NO: 66) mostra a seqüência de aminoácidos de cadeia leve p5E8 PRIMATIZED®.

A Figura 83A (SEQ ID NO: 67) mostra seqüência de DNA de filamento único de p5E8 (VL/VH) scFv PRIMATIZED® convencional. Seqüência de peptídeo de sinal é mostrada como sublinhada. A Figura 83B (SEQ ID NO: 68) mostra seqüência de aminoácidos de p5E8 (VL/VH) scFv PRIMATIZED® convencional.

A Figura 84A (SEQ ID NO: 69) mostra seqüência de DNA de filamento único de p5E8 (VH/VL) scFv PRIMATIZED® convencional. A Figura 84B (SEQ ID NO: 70) mostra seqüência de aminoácidos de p5E8 (VH/VL) scFv PRIMATIZED® convencional.

A Figura 85 representa curvas T50 de p5E8 (VH/VL) scFv convencional (círculos preenchidos) versus p5E8 (VL/VH) scFv convencional (círculos abertos).

A Figura 86 representa os resultados de um experimento avaliando a atividade *in vivo* relativa de anticorpos "Hercules-II" biespecíficos estabilizados (XWU036 (círculo fechado)) e anticorpos monoespecíficos administrados sozinhos (hBHA10 (triângulo fechado invertido) e ch14A2 (losango fechado)) e em combinação (triângulo aberto invertido) contra um modelo de camundongo de xenoenxerto tumoral de câncer de mama humano (MDA-MB-231).

A Figura 87 representa resíduos VH envolvidos em redes de Covariação que são importantes para manutenção de interface (Figura 87A) e resíduos da interface envolvidos em fazer contatos diretos com a interface VH/VL e que são mapeados diretamente para a superfície de domínio VH (Figura 87B). Setas indicam posições de resíduos fora da interface real que análise de covariação informa como sendo importante para estabilizar a interação VH/VL e para estabilidade de VH/VL. A localização de resíduos CDR3 hipervariáveis que fazem contatos diretos na interface entre VH e VL são encerrados dentro de uma caixa.

A Figura 88 representa resíduos VL envolvidos em redes de Covariação que são importantes para manutenção de interface (vide a Figura 88A) e resíduos da interface envolvidos em fazer contatos diretos com a interface VH/VL e que são mapeados diretamente com a superfície de VL (Figura 88B). Setas indicam posições de resíduos fora da interface real que análise de covariação informa como sendo importante para estabilizar a interação VH/VL e para estabilidade de VH/VL. A localização de resíduos CDR3 hipervariáveis que fazem contatos diretos na interface entre VH e VL são encerrados dentro de uma caixa.

A Figura 89 representa um ambiente exemplar adequado para praticar uma modalidade da invenção.

A Figura 90 é um fluxograma de uma seqüência exemplar de etapas que podem ser seguidas por uma modalidade da invenção de modo a determinar um escore de covariação de seqüência que pode ser usado como uma medida da estabilidade de um polipeptídeo.

A Figura 91 é um fluxograma de uma seqüência exemplar de etapas que podem ser seguidas por uma modalidade da invenção de modo a determinar e usar uma classificação de consenso para prever a estabilidade de uma proteína candidata.

A Figura 92 é um fluxograma de uma seqüência exemplar de etapas que podem ser seguidas por uma modalidade da invenção de modo a determinar e usar uma classificação de consenso média como uma medida da estabilidade de uma proteína candidata.

A Figura 93 é um fluxograma de uma seqüência exemplar de etapas que podem ser seguidas por uma modalidade da invenção de modo a usar um resíduo substituto de um polipeptídeo para determinar uma seqüência do polipeptídeo aprimorado.

As Figuras 94A e 94B mostram as seqüências de aminoácidos do BHA10 scFVs estabilizados contendo a mutação estabilizante S46L(VL) (SEQ ID NO: 137), e a mutação estabilizante V55G(VH) (SEQ ID NO: 138). A mutação estabilizante é indicada pelo resíduo envolvido por caixa. A seqüência lider, peptídeo de ligação gly/ser, e domínio CH1 são indicados pelos resíduos sublinhados, em negrito, e em itálico, respectivamente.

Descrição Detalhada da Invenção

A invenção se baseia, no mínimo em parte, no desenvolvimento de moléculas de ligação alvo estabilizadas que consistem em ou compreendem uma molécula scFv estabilizada e métodos para preparar similares moléculas de ligação estabilizadas. Além disso, a presente invenção proporciona métodos para designar proteínas estáveis, tais como moléculas de anticorpos.

As moléculas scFv estabilizadas da presente invenção são es-

pecialmente úteis para produzir moléculas estáveis multiespecíficas, por exemplo, moléculas biespecíficas. As moléculas de ligação estabilizadas da invenção, por exemplo, moléculas de ligação multiespecíficas podem ser expressada estavelmente em cultura, são adequadas para produção em larga escala, e são estáveis in vivo.

A invenção também se baseou, no mínimo em parte, no desenvolvimento de moléculas de ligação estabilizadas que consistem em ou compreendem uma molécula scFv estabilizada e métodos para preparar similares moléculas de ligação estabilizadas. Além disso, a presente invenção proporciona métodos para designar proteínas estáveis e para prever a estabilidade de proteínas, tais como moléculas de anticorpos.

Antes de descrição adicional da invenção, por conveniência, alguns termos são descritos abaixo:

I. Definições

Conforme usado aqui, neste Pedido de patente, o termo "molécula scFv" inclui moléculas de ligação as quais consistem em um domínio variável de cadeia leve (VL) ou porção do mesmo, e um domínio variável de cadeia pesada (VH) ou porção do mesmo, em que cada domínio variável (ou porção do mesmo) é derivado dos mesmos ou de diferentes anticorpos. Moléculas scFv preferencialmente compreendem um ligante de scFv interposto entre o domínio VH e o domínio VL. Moléculas scFv são conhecidas na técnica e são descritas, por exemplo, na patente dos Estados Unidos Nº 5.892.019, Ho et al. 1989. Gene 77:51; Bird et al. 1988 *Science* 242:423; Pantoliano et al. 1991. *Biochemistry* 30:10117; Milenic et al. 1991. *Cancer Research* 51:6363; Takkinen et al. 1991. *Protein Engineering* 4:837.

Um "ligante de scFv" conforme usado aqui, neste Pedido de patente, se refere a uma porção interposta entre os domínios VL e VH do scFv. Ligantes de scFv preferencialmente mantêm a molécula scFv em uma conformação de ligação antigênica. Em uma modalidade, um ligante de scFv compreende ou consiste de um peptídeo ligante de scFv. Em algumas modalidades, um peptídeo ligante de scFv compreende ou consiste de um peptídeo de ligação gly-ser. Em outras modalidades, um ligante de scFv com-

preende uma ligação dissulfeto.

Conforme usado aqui, neste Pedido de patente, o termo "peptídeo de ligação gly-ser" se refere a um peptídeo que consiste de resíduos glicina e serina. Um peptídeo de ligação gly-ser exemplar compreende a seqüência de aminoácidos $(Gly_4\ Ser)_n$. Em uma modalidade, n = 1. Em uma modalidade, n = 2. Em outra modalidade, n = 3. Em uma modalidade preferencial, n = 4, isto é, $(Gly_4\ Ser)_4$. Em outra modalidade, n = 5. Em ainda outra modalidade, n = 6.

Outro peptídeo de ligação gly-ser exemplar compreende a seqüência de aminoácidos $Ser(Gly_4Ser)_n$. Em uma modalidade, n = 1. Em uma modalidade, n = 2. Em uma modalidade preferencial, n = 3. Em outra modalidade, n = 4. Em outra modalidade, n = 5. Em ainda outra modalidade, n = 6.

Conforme usado aqui, neste Pedido de patente, o termo "ligação dissulfeto" se refere à ligação covalente formada entre dois átomos de enxofre. O aminoácido cisteína compreende um grupamento tiol que pode formar uma ligação dissulfeto ou ponte com um segundo grupamento tiol.

Conforme usado aqui, neste Pedido de patente, o termo "molécula scFv convencional" se refere a uma molécula scFv a qual não é uma molécula scFv estabilizada. Por exemplo, uma típica convencional carece de mutações estabilizantes e compreende um domínio VH e um domínio VL encadeado por um ligante $(G_4S)3$.

Uma "molécula scFv estabilizada" da invenção é uma molécula scFv compreendendo no mínimo uma mudança ou alteração comparada com uma molécula scFv convencional a qual resulta em estabilização da molécula scFv. Conforme usado aqui, neste Pedido de patente, o termo "mutação estabilizante" inclui uma mutação a qual confere reforço da estabilidade de proteína (por exemplo, estabilidade térmica) à molécula scFv e/ou a uma proteína maior compreendendo a referida molécula scFv. Em uma modalidade, a mutação estabilizante compreende a substituição de um aminoácido desestabilizante com uma substituição de aminoácido que confere reforço da estabilidade de proteína (aqui, neste Pedido de patente, um "aminoácido estabilizante"). Em uma modalidade, a mutação estabilizante é uma

na qual a extensão de um ligante de scFv foi otimizada. Em uma modalidade, uma molécula scFv estabilizada da invenção compreende uma ou mais substituições de aminoácidos. Por exemplo, em uma modalidade, uma mutação estabilizante compreende uma substituição de no mínimo um resíduo aminoácido cuja substituição resulta em um aumento na estabilidade da interface de VH e VL de uma molécula scFv. Em uma modalidade, o aminoácido está dentro da interface. Em outra modalidade, o aminoácido é um o qual *estruturas* a interação entre VH e VL. Em outra modalidade, uma mutação estabilizante compreende substituir no mínimo um aminoácido no domínio VH ou domínio VL que covaria com dois ou mais aminoácidos na interface entre os domínios VH e VL. Em outra modalidade, a mutação estabilizante é uma na qual no mínimo um resíduo cisteína é introduzido (isto é, é produzida em um ou mais dos domínios VH ou VL) de tal modo que os domínios VH e VL são encadeados por no mínimo uma ligação dissulfeto entre um aminoácido no domínio VH e um aminoácido no domínio VL. Em algumas modalidades preferenciais, uma molécula scFv estabilizada da invenção é uma molécula na qual tanto a extensão do ligante de scFv é otimizada quanto no mínimo um resíduo aminoácido é substituído e/ou os domínios VH e VL são encadeados por uma ligação dissulfeto entre um aminoácido no domínio VH e um aminoácido no domínio VL. Em uma modalidade, mais de uma das mutações estabilizantes descritas aqui, neste Pedido de patente, podem ser feitas em uma molécula scFv.

Em uma modalidade, uma ou mais mutações estabilizantes feitas a uma molécula scFv simultaneamente melhora a estabilidade térmica de ambos os domínios VH e VL da molécula scFv comparada com uma molécula scFv convencional.

Preferencialmente, uma população de uma ou mais das moléculas scFv estabilizadas da invenção é expressada como uma população de proteínas solúveis monoméricas. Em uma modalidade, não mais de 10% está presente em forma agregada. Em uma modalidade, as moléculas scFv estabilizadas da população podem compreender a mesma mutação estabilizante ou uma combinação de mutações estabilizantes. Em outras modalida-

des, as moléculas scFv estabilizadas individuais da população compreendem diferentes mutações estabilizantes.

As moléculas scFv estabilizadas em questão podem ser usadas sozinhas para ligar a uma molécula alvo ou podem ser encadeadas a outro polipeptídeo para formar moléculas de ligação estabilizadas as quais compreendem uma molécula scFv estabilizada. Por exemplo, uma molécula de ligação da invenção pode compreender uma molécula scFv encadeada a uma segunda molécula scFv ou uma molécula não -scFv, por exemplo, que confere especificidade de ligação alvo, tal como um anticorpo.

Conforme usado aqui, neste Pedido de patente, o termo "estabilidade de proteína" se refere a uma medida reconhecida na técnica da manutenção de uma ou mais propriedades físicas de uma proteína em resposta a uma condição ambiental (por exemplo, uma temperatura elevada ou reduzida). Em uma modalidade, a propriedade física é a manutenção da estrutura covalente da proteína (por exemplo, a ausência de clivagem proteolítica, oxidação indesejada ou deamidação). Em outra modalidade, a propriedade física é a presença da proteína em um estado adequadamente dobrado (por exemplo, a ausência de agregados ou precipitados solúveis ou insolúveis). Em uma modalidade, a estabilidade de uma proteína é medida avaliando uma propriedade biofísica da proteína, por exemplo, estabilidade térmica, perfil de desdobramento com pH, remoção estável de glicosilação, solubilidade, função bioquímica (por exemplo, capacidade para ligar a uma proteína (por exemplo, um ligante, um receptor, um antígeno, e etc.) ou porção química, e etc.), e/ou combinações das mesmas. Em outra modalidade, a função bioquímica é demonstrada pela afinidade de ligação de uma interação. Em uma modalidade, uma medida da estabilidade da proteína é estabilidade térmica, isto é, resistência a estímulo térmico. A estabilidade pode ser medida usando métodos conhecidos na técnica e/ou descritos aqui, neste Pedido de patente.

Os domínios VL e VH de uma molécula scFv são derivados de uma ou mais moléculas de anticorpos. Também será entendido por uma pessoa versada na técnica que as regiões variáveis das moléculas scFv da

invenção podem ser modificadas de tal modo que variam em seqüência de aminoácidos da molécula de anticorpo da qual foram derivadas. Por exemplo, em uma modalidade, podem ser feitas substituições de nucleotídeo ou aminoácido levando a substituições conservativas ou alterações em resíduos aminoácidos (por exemplo, em CDR e/ou resíduos de framework). Alternativamente ou além disso, podem ser feitas mutações a resíduos aminoácidos de CDR para otimizar a ligação de antígeno usando técnicas reconhecidas na técnica. As moléculas de ligação da invenção mantêm a capacidade de ligar a antígeno.

Conforme usado aqui, neste Pedido de patente, o termo "derivado de" uma proteína designada se refere à origem do polipeptídeo. Em uma modalidade, a seqüência de polipeptídeo ou aminoácidos que a qual é derivada de um polipeptídeo de partida em particular é uma seqüência de região variável (por exemplo, uma VH ou VL) ou seqüência relacionada com a mesma (por exemplo, uma região de framework ou CDR). Em uma modalidade, a seqüência de aminoácidos a qual é derivada de um polipeptídeo de partida em particular não é contígua. Por exemplo, em uma modalidade, uma, duas, três, quatro, cinco, ou seis CDRs são derivadas de um anticorpo de partida. Em uma modalidade, a seqüência de polipeptídeo ou aminoácidos que é derivada de um polipeptídeo de partida em particular ou seqüência de aminoácidos tem uma seqüência de aminoácidos que é essencialmente idêntica àquela da seqüência de partida ou uma porção da mesma, em que a porção consiste de no mínimo 3 a 5 aminoácidos, 5 a 10 aminoácidos, no mínimo 10 a 20 aminoácidos, no mínimo 20 a 30 aminoácidos, ou no mínimo 30 a 50 aminoácidos, ou a qual é identificável de modo diverso para uma pessoa versada na técnica como tendo sua origem na seqüência de partida.

Uma molécula de ácido nucléico isolada codificando uma molécula scFv estabilizada ou uma porção da mesma pode ser criada introduzindo uma ou mais substituições, adições ou eliminações de nucleotídeos na seqüência de nucleotídeos de uma molécula scFv convencional ou uma imunoglobulina da qual é derivada de tal modo que uma ou mais substitui-

ções, adições ou eliminações de aminoácidos são introduzidas na proteína codificada. Mutações podem ser introduzidas por técnicas de rotina, tais como mutagênese sítio-dirigida e mutagênese mediada por PCR. Em uma modalidade, substituições de aminoácido conservativas são feitas em um ou mais resíduos aminoácidos não-essenciais. Uma "substituição de aminoácido conservativa" é uma na qual o resíduo aminoácido é substituído com um resíduo aminoácido tendo uma cadeia lateral similar. Famílias de resíduos aminoácidos tendo cadeias laterais similares foram definidas na técnica, incluindo cadeias laterais básicas (por exemplo, lisina, arginina, histidina), cadeias laterais acidíferas (por exemplo, ácido aspártico, ácido glutâmico), cadeias laterais polares não carregadas (por exemplo, glicina, asparagina, glutamina, serina, treonina, tirosina, cisteína), cadeias laterais não-polares (por exemplo, alanina, valina, leucina, isoleucina, prolina, fenilalanina, metionina, triptofano), cadeias laterais β-ramificadas (por exemplo, treonina, valina, isoleucina) e cadeias laterais aromáticas (por exemplo, tirosina, fenilalanina, triptofano, histidina). Portanto, um resíduo aminoácido é um polipeptídeo imunoglobulina pode ser substtuído com outro resíduo aminoácido da mesma família de cadeia lateral. Em outra modalidade, uma cadeia de aminoácidos pode ser substituída com uma cadeia estruturalmente similar que difere em ordem e/ou composição de membros da família de cadeia lateral. Em outra modalidade, uma mutação é introduzida de modo a introduzir no mínimo uma molécula cisteína no domínio VH e no domínio VL e, deste modo, introduzir uma ligação dissulfeto na molécula scFv. Em outra modalidade, um aminoácido de uma molécula scFv convencional pode ser substituído com um aminoácido tendo propriedades físicas (por exemplo, espaciais) ou funcionais similares. Preferencialmente, aminoácidos substituídos em moléculas scFv convencionais são compatíveis com a integridade da interface V_L/V_H, conformações de CDR, e dobramentos V_H e/ou V_L.

Alternativamente, em outra modalidade, mutações podem ser introduzidas aleatoriamente ao longo de toda ou parte da seqüência codificando imunoglobulina.

As moléculas scFv estabilizadas da invenção ou polipeptídeos

compreendendo as moléculas scFv estabilizadas são moléculas de ligação, isto é, ligam a uma molécula alvo de interesse, por exemplo, um antígeno. Quando uma molécula scFv estabilizada da invenção é fundida a uma segunda molécula, a segunda molécula também pode conferir uma especificidade de ligação à proteína de fusão.

As moléculas de ligação da invenção consistem em moléculas scFv (por exemplo, um domínio VH e um domínio VL ligados por um ligante de scFv) ou compreendem uma molécula scFv estabilizada da invenção.

Em uma modalidade, as moléculas de ligação da invenção são monovalentes, isto é, compreendem um sítio de ligação alvo (por exemplo, como no caso de uma molécula scFv). Em uma modalidade, as moléculas de ligação da invenção são multivalentes, isto é, compreendem maisd eum sítio de ligação alvo. Em outra modalidade, as moléculas de ligação compreendem no mínimo dois sítios de ligação. Em uma modalidade, as moléculas de ligação compreendem dois sítios de ligação. Em uma modalidade, as moléculas de ligação compreendem três sítios de ligação. Em outra modalidade, as moléculas de ligação compreendem quatro sítios de ligação. Em outra modalidade, as moléculas de ligação compreendem mais de quatro sítios de ligação.

Em uma modalidade, as moléculas de ligação da invenção são monômeros. Em outra modalidade, as moléculas de ligação da invenção são multímeros. Por exemplo, em uma modalidade, as moléculas de ligação da invenção são dímeros. Em uma modalidade, os dímeros da invenção são homodímeros, compreendendo duas subunidades monoméricas idênticas. Em outra modalidade, os dímeros da invenção são heterodímeros, compreendendo duas subunidades monoméricas não-idênticas. As subunidades do dímero podem compreender uma ou mais cadeias de polipeptídeo. Por exemplo, em uma modalidade, os dímeros compreendem no mínimo duas cadeias de polipeptídeo. Em uma modalidade, os dímeros compreendem duas cadeias de polipeptídeo. Em outra modalidade, os dímeros compreendem quatro cadeias de polipeptídeo (por exemplo, como no caso de moléculas de anticorpos).

Conforme usado aqui, neste Pedido de patente, o termo "valência" se refere ao número de sítios de ligação alvo potenciais em um polipeptídeo. Cada sítio de ligação alvo liga especificamente uma molécula alvo ou sítio específico sobre uma molécula alvo. Quando um polipeptídeo compreende mais de um sítio de ligação alvo, cada sítio de ligação alvo pode ligar especificamente as mesmas moléculas ou diferentes (por exemplo, pode ligar a diferentes moléculas, por exemplo, diferentes antígenos, ou diferentes epítopes sobre a mesma molécula).

O termo "especificidade" se refere à capacidade para ligar especificamente (por exemplo, imunorreagir com) um dado alvo. Um polipeptídeo pode ser monoespecífico e conter um ou mais sítios de ligação os quais ligam especificamente um alvo ou um polipeptídeo pode ser multiespecífico (por exemplo, biespecífico ou triespecífico) e conter dois ou mais sítios de ligação os quais ligam especificamente os mesmos alvos ou diferentes. Ligação específica pode ser conferida por uma molécula scFv estabilizada da invenção e/ou uma porção não scFv à qual uma molécula scFv estabilizada da invenção é ligada.

Em uma modalidade, uma molécula de ligação da invenção é multiespecífica. Por exemplo, em uma modalidade, uma molécula de ligação multiespecífica da invenção é uma molécula biespecífica (por exemplo, anticorpo, "minibody", anticorpo com domínio eliminado, ou proteína de fusão, anticorpos de domínio único (por exemplo, camelídeos, de tubarão, seres humanos) compreendendo no mínimo uma molécula scFv estabilizada) tendo especificidade de ligação para no mínimo dois alvos, por exemplo, mais de uma molécula alvo ou mais de um epítope sohre a mesma molécula alvo. Em uma modalidade, uma molécula multiespecífica tem no mínimo um sítio de ligação específico para uma molécula orientada para redução ou eliminação e uma molécula alvo sobre uma célula. Em outra modalidade, uma molécula multiespecífica tem no mínimo um sítio de ligação alvo específico para uma molécula orientada para redução ou eliminação e no mínimo um sítio de ligação específico para um fármaco. Em ainda outra modalidade, uma molécula multiespecífica tem no mínimo um sítio de ligação específico para uma

molécula orientada para redução ou eliminação e no mínimo um sítio de ligação específico para um pró-fármaco.

Em uma modalidade, uma molécula multiespecífica compreende uma especificidade para uma molécula solúvel e uma especifidade para uma molécula da superfície celular. Em outra modalidade, uma molécula multiespecífica tem duas especificidades de ligação para dois alvos presentes sobre uma ou mais moléculas solúveis. Em outra modalidade, uma molécula multiespecífica tem duas especificidades de ligação para dois alvos presentes sobre uma ou mais moléculas da superfície celular (as quais podem estar presentes sobre uma ou mais células).

Em uma modalidade, as moléculas de ligação têm no mínimo um sítio de ligação alvo específico para uma molécula o qual media um efeito biológico (por exemplo, o qual modula ativação celular (por exemplo, por ligação a um receptor da superfície celular e resultando em transmissão ou inibição de um sinal de ativação ou inibitório), o qual resulta em morte da célula (por exemplo, por um caminho induzido por sinal celular, por fixação do complemento ou exposição a uma carga útil presente sobre a molécula de ligação), ou o qual modula uma doença ou um distúrbio em um sujeito (por exemplo, mediando ou promovendo morte celular, promovendo lise de um coágulo de fibrina ou promovendo formação de coágulo, ou modulando a quantidade de uma substância a qual está biodisponível (por exemplo, reforçando ou reduzindo a quantidade de um ligante tal como TNF□ no sujeito)).

Em outra modalidade, as moléculas de ligação da invenção ligam no mínimo um alvo que transduz um sinal para uma célula, por exemplo, ligando a um receptor da superfície celular, tal como um receptor da família de TNF. Por "transduz um sinal" significa que por ligação à célula, a molécula de ligação converte a influência extracelular sobre o receptor da superfície celular em uma resposta celular, por exemplo, modulando um caminho de transdução de sinal.

Em uma modalidade, as moléculas de ligação ligam no mínimo um sítio de ligação alvo específico para uma molécula orientada para redução ou eliminação, por exemplo, uma antígeno da superfície celular ou um

antígeno solúvel. Em uma modalidade, a ligação da molécula de ligação ao alvo resulta em redução ou eliminação do alvo, por exemplo, de um tecido ou da circulação. Em outra modalidade, as moléculas de ligação têm no mínimo um sítio de ligação específico para uma molécula que pode ser usado para detectar a presença de uma molécula alvo (por exemplo, para detectar um contaminante ou diagnosticar uma condição ou distúrbio). Em ainda outra modalidade, uma molécula de ligação da invenção compreende no mínimo um sítio de ligação que orienta a molécula de ligação para um sítio específico em um sujeito (por exemplo, para uma célula tumoral ou coágulo sangüíneo).

Em uma modalidade preferencial, uma molécula multiespecífica é um anticorpo tetravalente que em quatro sítios de ligação. Uma molécula tetravalente pode ser biespecífica e bivalente para cada especificidade. Descrição adicional de moléculas biespecíficas exemplares é proporcionada abaixo.

Moléculas de ligação preferenciais da invenção compreendem seqüências de aminoácidos de região de framework e constante derivadas de uma seqüência de aminoácidos humana. No entanto, polipeptídeos de ligação podem compreender seqüências de região de framework e/ou constante derivadas de outra espécie mamífera. Por exemplo, moléculas de ligação compreendendo seqüências murinas podem ser apropriadsa para algumas aplicações. Em uma modalidade, uma região de framework de primata (por exemplo, primata não-humano), porção de cadeia pesada, e/ou porção de articulação pode ser incluída nas moléculas de ligação em questão. Em uma modalidade, um ou mais aminoácidos murinos podem estar presentes na região de framework de um polipeptídeo de ligação, por exemplo, uma seqüência de aminoácidos de framework de primata humano ou não - humano pode compreender uma ou mais retro mutações de aminoácidos nas quais o resíduo aminoácido murino correspondente está presente e/ou pode compreender uma ou mutações para um resíduo aminoácido diferente não encontrado no anticorpo murino de partida. Moléculas de ligação preferenciais da invenção são menos imunogênicas do que anticorpos murinos.

Uma proteína de "fusão" ou quimérica compreende uma primeira seqüência de aminoácidos encadeada a uma segunda seqüência de aminoácidos com a qual não é naturalmente encadeado na natureza. As seqüências de aminoácidos podem existir normalmente em proteínas separadsa que são trazidas juntas no polipeptídeo de fusão ou podem existir normalmente na mesma proteína mas são colocadas em uma nova disposição no polipeptídeo de fusão. Uma proteína de fusão pode ser criada, por exemplo, por síntese química, ou criando e traduzindo um polinucleotídeo no qual as regiões do peptídeo são codificadas na relação desejada.

O termo "heterólogo" conforme aplicado a um polinucleotídeo ou um polipeptídeo, significa que o polinucleotídeo ou polipeptídeo é derivado de uma entidade genotipicamente distinta daquela da entidade à qual está sendo comparada. Por exemplo, um polinucleotídeo heterólogo ou antígeno pode ser derivado de uma espécie diferente, tipo celular diferente de um indivíduo, ou o mesmo tipo de célula ou diferente de indivíduos distintos.

O termo "domínio de ligação de ligante" ou "porção de ligação de ligante" conforme usado aqui, neste Pedido de patente, se refere a qualquer receptor nativo (por exemplo, receptor da superfície celular) ou qualquer região ou derivado da mesma retendo no mínimo uma capacidade de ligação de ligante qualitativa, e preferencialmente a atividade biológica de um receptor nativo correspondente.

O termo "domínio de ligação de receptor" ou "porção de ligação de receptor" conforme usado aqui, neste Pedido de patente, se refere a qualquer ligante nativo ou qualquer região ou derivado da mesma retendo no mínimo uma capacidade de ligação de receptor qualitativa, e preferencialmente a atividade biológica de um ligante nativo correspondente.

Em uma modalidade, as moléculas de ligação da invenção são moléculas de "anticorpo" ou "imunoglobulina" estabilizadas, por exemplo, moléculas de anticorpo ou imunoglobulina que ocorrem naturalmente (ou um fragmento de ligação de antígeno das mesmas) ou moléculas de anticorpos produzidas geneticamente que ligam antígeno de uma maneira similar a moléculas de anticorpos e que compreendem uma molécula scFv da invenção.

Conforme usado aqui, neste Pedido de patente, o termo "imunoglobulina" inclui um polipeptídeo tendo uma combinação de duas cadeias pesadas e duas leves quer ou não possua qualquer imunorreatividade específica relevante. "Anticorpos" se refere a similares montagens as quais têm atividade imunorreativa específica conhecida significativa para um antígeno de interesse (por exemplo, um antígeno associado a tumor). Anticorpos e imunoglobulinas compreendem cadeias leves e pesadas, com ou sem uma ligação covalente intercadeia entre as mesmas. Estruturas de imunoglobulina básicas em sistemas de vertebrados são relativamente bem-entendidas.

Conforme será discutido em mais detalhes abaixo, o termo genérico "imunoglobulina" compreende cinco classes distintas desta que podem ser distinguidas bioquimicamente. Todas as cinco classes de anticorpos estão dentro do âmbito da presente invenção, a seguinte discussão geralmente será dirigida à classe de moléculas de imunoglobulina IgG. Com respeito a IgG, imunoglobulinas compreendem duas cadeias de polipeptídeo leves idênticas de peso molecular de aproximadamente 23.000 Dáltons, e duas cadeias pesadas idênticas de peso molecular de 53.000 a 70.000. As quatro cadeias são ligadas por ligações dissulfeto em uma configuração em "Y" em que as cadeias leves agrupam por meio de chave as cadeias pesadas iniciando na boca do "Y" e continuando através da região variável.

Tanto as cadeias leves quanto pesadas são divididas em regiões de homologia estrutural e funcional. Os termos "constante" e "variável" são usados funcionalmente. A este respeito, será reconhecido que os domínios variáveis tanto das porções de cadeia leve (VL) quanto pesada (VH) determinam reconhecimento e especificidade antigênicos. Ao invés, os domínios constantes de cadeia leve (CL) e de cadeia pesada (CH1, CH2 ou CH3) conferem importantes propriedades biológicas tais como secreção, mobilidade transplacentária, ligação de receptor de Fc, ligação de complemento, e similares. Por convenção a numeração dos domínios de região constante aumenta à medida que se tornam mais distais do sítio de ligação de antígeno ou término amino do anticorpo. O N-término é uma região variável e no C-término é uma região constante; os domínios CH3 e CL na verda-

de compreendem o carbóxi-término da cadeia pesada e leve, respectivamente.

Mutações estabilizantes para moléculas scFv podem ser feitas a aminoácidos na CDR e/ou nas regiões de framework de uma cadeia pesada variável e/ou cadeia leve variável scFv. Conforme usado aqui, neste Pedido de patente, o termo "resíduos aminoácidos de CDR de região variável" inclui aminoácidos em uma CDR ou região determinante de complementaridade conforme identificado usando métodos à base de seqüência ou estrutura. Conforme usado aqui, neste Pedido de patente, o termo "CDR" ou "região determinante de complementaridade" significa os sítios combinando antígenos não contíguos encontrados dentro da região variável tanto de polipeptídeos de cadeia pesada quanto leve. Estas regiões particulares foram descritas por Kabat et al., J. Biol. Chem. 252, 6609-6616 (1977) e Kabat et al., *Sequences of protein of immunological interest.* (1991), e por Chothia et al., J. Mol. Biol. 196:901-917 (1987) e por MacCallum et al., J. Mol. Biol. 262:732-745 (1996) onde as definições incluem sobreposição ou subgrupos de resíduos aminoácidos quando comparados um com o outro. Os resíduos aminoácidos os quais englobam as CDRs conforme definido por cada uma das referências citadas acima são determinados para comparação. Preferencialmente, o termo "CDR" é uma CDR conforme definido por Kabat com base em comparações de seqüências.

	Definições de CDR		
	Kabat[1]	Chothia[2]	MacCallum[3]
V_H CDR1	31-35	26-32	30-35
V_H CDR2	50-65	53-55	47-58
V_H CDR3	95-102	96-101	93-101
V_L CDR1	24-34	26-32	30-36
V_L CDR2	50-56	50-52	46-55
V_L CDR3	89-97	91-96	89-96

[1]A numeração de resíduos segue a nomenclatura de Kabat et al., supra

[2]A numeração de resíduos segue a nomenclatura de Chothia et al., supra

[3]A numeração de resíduos segue a nomenclatura de MacCallum et al., supra

Conforme usado aqui, neste Pedido de patente, o termo "resíduos aminoácidos de framework (FR) de região variável" se refere aos aminoácidos na região de framework de uma cadeia de Ig. O termo "região de framework" ou "região FR" conforme usado aqui, neste Pedido de patente, inclui o resíduos aminoácidos que são parte da região variável, mas não são parte das CDRs (por exemplo, usando a definição de CDRs de Kabat). Portanto, um framework de região variável tem entre cerca de 100 a 120 aminoácidos de extensão mas inclui somente os aminoácidos fora das CDRs. Para o exemplo específico de uma região variável de cadeia pesada e para as CDRs conforme definido por Kabat *et al.*, região de framework 1 corresponde ao domínio da região variável englobando os aminoácidos 1 a 30; a região de framework 2 corresponde ao domínio da região variável englobando os aminoácidos 36 a 49; a região de framework 3 corresponde ao domínio da região variável englobando os aminoácidos 66 a 94, e a região de framework 4 corresponde ao domínio da região variável dos aminoácidos 103 à extremidade da região variável. As regiões de framework para a cadeia leve são separadas de modo similar por cada uma das CDRs da região variável de cadeia leve. De modo similar, usando a definição de CDRs de Chothia *et al.* ou McCallum *et al.* os limites da região de framework são separados pelos términos de CDR respectivos conforme descrito acima. Em modalidades preferenciais, as CDRs são conforme definido por Kabat.

Em anticorpos que ocorrem naturalmente, as seis CDRs presentes sobre cada anticorpo monomérico são seqüências de aminoácidos curtas e não-contíguas que são posicionadsa especificamente para formar o sítio de ligação antigênica enquanto o anticorpo assume sua configuração tridimensional em um ambiente aquoso. O restante dos domínios variáveis de cadeia pesada e leve mostram menos variabilidade intermolecular na seqüência de aminoácidos e são denominados as regiões de framework. As regiões de framework largamente adotam uma conformação de β-folha e as CDRs formam alças as quais conectam, e em alguns casos formam parte, da estrutura de β-folha. Portanto, estas regiões de framework agem formando um *estrutura* que proporciona posicionamento das seis CDRs em orienta-

ção correta por interações intercadeia, não-covalentes. O sítio de ligação antigênica formado pelas CDRs posicionadas define uma superfície complementar ao epítope sobre o antígeno imunorreativo. Esta superfície complementar promove a ligação não-covalente do anticorpo ao epítope de antígeno imunorreativo. A posição de CDRs pode ser prontamente identificada por uma pessoa com conhecimento regular da técnia.

Conforme previamente indicado, as estruturas de subunidades e configuração tridimensional das regiões constantes das várias classes de imunoglobulina são de conhecimento geral. Conforme usado aqui, neste Pedido de patente, o termo "domínio VH" inclui o domínio variável de amino terminal de uma cadeia pesada de imunoglobulina. Conforme usado aqui, neste Pedido de patente, o termo "domínio VL" inclui o domínio variável de amino terminal de uma cadeia leve de imunoglobulina.

O termo "fragmento" se refere a uma parte ou porção de um polipeptídeo (por exemplo, um anticorpo ou uma cadeia de anticorpo) compreendendo menos resíduos aminoácidos do que um polipeptídeo intacto ou completo. O termo "fragmento de ligação de antígeno" se refere a um fragmento de polipeptídeo de uma imunoglobulina ou anticorpo que liga antígeno ou compete com anticorpo intacto (isto é, com o anticorpo intacto do qual foram derivados) para ligação de antígeno (isto é, ligação específica). Conforme usado aqui, neste Pedido de patente, o termo "fragmento" de uma molécula de anticorpo inclui fragmentos de ligação de antígeno de anticorpos, por exemplo, uma cadeia leve de anticorpo (VL), uma cadeia pesada de anticorpo (VH), um anticorpo de cadeia única (scFv), um fragmento F(ab')2, um fragmento Fab, um fragmento Fd, um fragmento Fv, e um fragmento de anticorpo de domínio único (DAb). Fragmentos podem ser obtidos, por exemplo, através de tratamento químico ou enzimático de um anticorpo ou cadeia de anticorpo intacto ou completo ou por meios recombinantes.

Conforme usado aqui, neste Pedido de patente, o termo "sítio de ligação" compreende uma região de um polipeptídeo a qual é responsável por ligar seletivamente a uma molécula alvo de interesse (por exemplo, um antígeno, ligante, receptor, substrato ou inibidor). Domínios de ligação com-

preendem no mínimo um sítio de ligação alvo. Domínios de ligação típicos incluem um anticorpo domínio variável, um domínio de ligação de receptor de um ligante, um domínio de ligação de ligante de um receptor ou um domínio enzimático.

Moléculas de ligação da invenção podem ser preparadas usando técnicas que são conhecidas na técnica. Em uma modalidade, os polipeptídeos da invenção são "produzidos de modo recombinante," isto é, são produzidos usando tecnologia de DNA recombinante. Técnicas típicas para preparar similares moléculas são discutidas em mais detalhes abaixo.

Em uma modalidade, uma molécula de ligação da invenção é um anticorpo que ocorre naturalmente ao qual uma molécula scFv estabilizada foi fundida. Em uma modalidade, uma molécula de ligação da invenção é um anticorpo modificado ao qual uma molécula scFv estabilizada foi fundida. Conforme usado aqui, neste Pedido de patente, o termo "anticorpo modificado" inclui formas sintéticas de anticorpos as quais são alteraas de tal modo que não ocorrem naturalmente. Em outra modalidade, uma molécula de ligação da invenção é uma proteína de fusão compreendendo no mínimo uma molécula scFv.

Em modalidades preferenciais, um polipeptídeo da invenção não provocará um reação imune nociva em um ser humano.

Em uma modalidade, uma molécula de ligação da invenção compreende uma região constante, por exemplo, uma região constante de cadeia pesada. Em uma modalidade, uma região constante similar é modificada comparada com uma região constante selvagem. Isto é, os polipeptídeos da invenção revelados aqui, neste Pedido de patente, podem compreender alterações ou modificações para um ou mais dos três domínios constantes de cadeia pesada (CH1, CH2 ou CH3) e/ou para o domínio de região constante de cadeia leve (CL). Modificações exemplares incluem adições, eliminações ou substituições de um ou mais aminoácidos em um ou mais domínios. As alterações referidas podem ser incluídas para otimizar a função efetora, a meia-vida, e etc.

Conforme usado aqui, neste Pedido de patente, o termo "malig-

nidade" se refere a um tumor não-benigno ou um câncer. Conforme usado aqui, neste Pedido de patente, o termo "câncer" inclui uma malignidade caracterizada por crescimento celular desregulado ou descontrolado. Cânceres típicos incluem: carcinomas, sarcomas, leucemias, e linfomas. O termo "câncer" inclui tumores malignos primários (por exemplo, aqueles cujas células não migraram para sítios no corpo do sujeito diferentes do sítio do tumor original) e tumores malignos secundários (por exemplo, aqueles que surgem de metástase, a migração de células tumorias para sítios secundários que são diferentes do sítio do tumor original).

Conforme usado aqui, neste Pedido de patente, o termo "engenheirado" inclui manipulação de moléculas de polipeptídeo ou ácido nucléico por meios sintéticos (por exemplo, por técnicas recombinantes, síntese de peptídeo in vitro, por ligação enzimática ou química de peptídeos ou alguma combinação destas téncnicas). Preferencialmente, as moléculas de ligação da invenção são preparadas usando os métodos referidos.

Conforme usado aqui, neste Pedido de patente, os termos "encadeado," "fundido" ou "fusão" são usados de modo intercambiável. Estes termos se referem a unir juntos os dois ou mais elementos ou componentes, por quaisquer meios inclusive conjugação química ou meios recombinantes. Preferencialmente os polipeptídeos os quais são fundidos são fundidos geneticamente, isto é, são fundidos usando tecnologia de DNA recombinante. Uma "fusão em estrutura" se refere à união de dois ou mais frames de leitura aberta (ORFs) para formar um ORF contínuo mais longo, em uma maneira que mantém o frame de leitura correto dos ORFs originais. Portanto, a proteína de fusão resultante é uma única proteína contendo dois ou mais segmentos que correspondem aos polipeptídeos codificados pelos ORFs originais (cujos segmentos não são normalmente ligados deste modo na natureza.) Embora o frame de leitura seja portanto preparado contínuo do início ao fim dos segmentos fundidos, os segmentos podem ser fisicamente ou espacialmente separados, por exemplo, por seqüência encadeadora de scFv em estrutura.

No contexto de polipeptídeos, uma "seqüência linear" ou uma

"seqüência" é uma ordem de aminoácidos em um polipeptídeo em uma direção de terminal amino para carboxila na qual resíduos que são adjacentes uns aos outros na seqüência são contíguos na estrutura primária do polipeptídeo.

Conforme usado aqui, neste Pedido de patente, a expressão "sujeito que se beneficiaria de administração de uma molécula de ligação" inclui sujeitos, tais como sujeitos mamíferos, que se beneficiaria de administração de uma molécula de ligação usada, por exemplo, para detecção de um antígeno reconhecido por uma molécula de ligação (por exemplo, para um procedimento diagnóstico) e/ou de tratamento com uma molécula de ligação para reduzir ou eliminar o alvo reconhecido pela molécula de ligação. Por exemplo, em uma modalidade, o sujeito pode ser beneficiar de redução ou eliminação de uma molécula solúvel ou particulada da circulação ou soro (por exemplo, uma toxina ou patógeno) ou de redução ou eliminação de uma população de células expressando o alvo (por exemplo, células tumorais). Conforme descrito em mais detalhes aqui, neste Pedido de patente, a molécula de ligação pode ser usada em forma não-conjugada ou pode ser conjugada, por exemplo, para uma porção detectável, um fármaco, pró-farmaco, ou um isótopo.

O termo "receptor de TNF" ou "membro da família de receptores de TNF" se refere a qualquer receptor pertencente à superfamília de receptores de Fator de Necrose Tumoral ("TNF"). Membros da Superfamília de Receptores de TNF ("TNFRSF") são caracterizados por uma região extracelular com dois ou mais domínios ricos em cisteína (~40 aminoácidos cada) arranjados como nós de cisteína (vide Dempsey *et al., Cytokine Growth Factor Rev.* (2003). 14(3-4):193-209). Na ligação de seus ligantes de TNF cognatos, receptores de TNF transduzem sinais interagindo diretamente ou indiretamente com proteínas adaptadoras citoplasmáticas conhecidas como TRAFs (fatores associados com receptores de TNF). TRAFs podem induzir a ativação de várias cascatas de quinase que essencialmente levam à ativação de caminhos de transdução de sinal tais como NF-KappaB, JNK, ERK, p38 e PI3K, que por sua vez regulam processos celulares variando de fun-

ção imune e diferenciação de tecido até apoptose.

As seqüências de nucleotídeo e aminoácidos de vários membros da família de receptores de TNF são conhecidas na técnica e incluem no mínimo 29 genes humanos: TNFRSF1A (TNFR1, também conhecido como DR1, CD120a, TNF-R-I p55, TNF-R, TNFRI, TNFAR, TNF-R55, p55TNFR, p55R, ou TNFR60, GenBank GI Nº 4507575; vide também US 5,395,760)), TNFRSF1B (CD120b, também conhecido como p75, TNF-R, TNF-R-II, TNFR80, TNFR2,TNF-R75, TNFBR , ou p75TNFR; GenBank GI Nº 4507577), TNFRSF3 (β Receptor de Linfotoxina (LTβR), também conhecido como TNFR2-RP, CD18, TNFR-RP, TNFCR, ou TNF-R-III; GI Nos. 4505038 e 20072212), TNFRSF4 (OX40, também conhecido como antígeno ACT35, TXGP1L, ou CD134; GI Nos. 4507579 e 8926702), TNFRSF5 (CD40, também conhecido como p50 ou Bp50; GI Nos. 4507581 e 23312371), TNFRSF6 (FAS, também conhecido como FAS-R, DcR-2, DR2, CD95, APO-1, ou APT1; GenBank GI Nos. 4507583, 23510421, 23510423, 23510425, 23510427, 23510429, 23510431, e 23510434)), TNFRSF6B (DcR3, DR3; GenBank GI Nos. 4507569, 23200021, 23200023, 23200025, 23200027, 23200029, 23200031, 23200033, 23200035, 23200037, e 23200039), TNFRSF7 (CD27, também conhecido como Tp55 ou S152; GenBank GI Nº 4507587), TNFRSF8 (CD30, também conhecido como Ki-1, ou D1S166E; GenBank GI Nos. 4507589 e 23510437), TNFRSF9 (4-1-BB, também conhecido como CD137 ou ILA; GI Nos. 5730095 e 728738), TNFRSF10A (TRAIL-R1, também conhecido como DR4 ou Apo2; GenBank GI Nº 21361086), TNFRSF10B (TRAIL-R2, também conhecido como DR5, KILLER, TRICK2A, ou TRICKB; GenBank GI Nos. 22547116 e 22547119), TNFRSF10C (TRAIL-R3, também conhecido como DcR1, LIT, ou TRID; GenBank GI Nº 22547121), TNFRSF10D (TRAIL-R4, também conhecido como DcR2 ou TRUNDD), TNFRSF11A (RANK; GenBank GI Nº 4507565; vide as Patentes dos Estados Unidos Nos. 6.562.948; 6.537.763; 6.528.482; 6.479.635; 6.271.349; 6.017.729), TNFRSF11B (Osteoprotegerina (OPG), também conhecido como OCIF ou TR1; GI Nos. 38530116, 22547122 e 33878056), TNFRSF12 (Proteína de Membrana de Associação de cadeia de

Translocação (TRAMP), também conhecido como DR3, WSL-1, LARD, WSL-LR, DDR3, TR3, APO-3, Fn14, ou TWEAKR; GenBank GI Nº 7706186; Publicação do Pedido de patente dos Estados Unidos Nº 2004/0033225A1), TNFRSF12L (DR3L), TNFRSF13B (TACI; GI Nº 6912694), TNFRSF13C (BAFFR; GI Nº 16445027), TNFRSF14 (Mediador de Entrada do Herpes Vírus (HVEM), também conhecido como ATAR, TR2, LIGHTR, ou HVEA; GenBank GI Nos. 23200041, 12803895, e 3878821), TNFRSF16 (Receptor do Fator de Crescimento Nervoso de Baixa Afinidade (LNGFR), também conhecido como Receptor de Neurotrofina ou p75(NTR); GenBank GI Nos. 128156 e 4505393), TNFRSF17 (BCM, também conhecido como BCMA; GI Nº 23238192), TNFRSF18 (AITR, também conhecido como GITR; GenBank GI Nos. 4759246, 23238194 e 23238197), TNFRSF19 (Troy/Trade, também conhecido como TAJ; GenBank GI Nos. 23238202 e 23238204), TNFRSF20 (RELT, também conhecido como FLJ14993; GI Nos. 21361873 e 23238200), TNFRSF21 (DR6), TNFRSF22 (SOBa, também conhecido como Tnfrh2 ou 2810028K06Rik), e TNFRSF23 (mSOB, também conhecido como Tnfrh1). Outros membros da família de TNF incluem EDAR1 (Receptor de Ectodisplasina A, também conhecido como Downless (DL), ED3, ED5, ED1R, EDA3, EDA1R, EDA-A1R; GenBank GI Nº 11641231; Patente dos Estados Unidos Nº 6,355,782), XEDAR (também conhecido como EDA-A2R; GenBank GI Nº 11140823); e CD39 (GI Nos. 2135580 e 765256).

O termo "ligante de TNF" ou "membro da família de ligantes de TNF" se refere a um ligante pertencente à superfamília de Fator de Necrose Tumoral (TNF). Ligantes de TNF ligam a receptores distintos da superfamília de receptores de TNF e apresentam 15 a 25% de homologia de seqüência de aminoácidos um com o outro (Gaur *et al., Biochem. Pharmacol.* (2003), 66(8):1403-8). As seqüências de nucleotídeo e de aminoácidos de vários membros da superfamília de Receptores de TNF (Ligantes) ("TNFSF") são conhecidas na técnica e incluem no mínimo 16 genes humanos: TNFSF1 (também conhecido como Linfotoxina-α (LTA), TNFβ ou LT, GI Nº: 34444 e 6806893), TNFSF2 (também conhecido como TNF, TNFα, ou DIF; GI Nº 25952111), TNFSF3 (também conhecido como Linfotoxina-β (LTB), TNFC,

ou p33), TNFSF4 (também conhecido como OX-40L, gp34, CD134L, ou glicoproteína 1 ativada tax-transcripcionalmente, 34kD (TXGP1); GI Nº 4507603), TNFSF5 (também conhecido como CD40LG, IMD3, HIGM1, CD40L, hCD40L, TRAP, CD154, ou gp39; GI Nº 4557433), TNFSF6 (também conhecido como FasL ou APT1LG1; GenBank GI Nº 4557329), TNFSF7 (também conhecido como CD70, CD27L, ou CD27LG; GI Nº 4507605), TNFSF8 (também conhecido como CD30LG, CD30L, ou CD153; GI Nº 4507607), TNFSF9 (também conhecido como 4-1BB-L ou ILA ligand; GI Nº 4507609), TNFSF10 (também conhecido como TRAIL, Apo-2L, ou TL2; GI Nº 4507593), TNFSF11 (também conhecido como TRANCE, RANKL, OPGL, ou ODF; GI Nos. 4507595 e 14790152), TNFSF12 (também conhecido como Fn14L, TWEAK, DR3LG, ou APO3L; GI Nos. 4507597 e 23510441), TNFSF13 (também conhecido como APRIL), TNFSF14 (também conhecido como LIGHT, LTg, ou HVEM-L; GI Nos. 25952144 e 25952147), TNFSF15 (também conhecido como TL1 ou VEGI), ou TNFSF16 (também conhecido como AITRL, TL6, hGITRL, ou GITRL; GI Nº 4827034). Outros membros da família de ligante de TNF incluem ligante EDAR1 e XEDAR (ED1; GI Nº 4503449; Monreal *et al.* (1998) *Am J Hum Genet.* 63:380), ligante Troy/Trade, BAFF (também conhecido como TALL1; GI Nº 5730097), e ligantes NGF (por exemplo, NGF-β (GI Nº 4505391), NGF-2/NTF3; GI Nº 4505469), NTF5 (GI Nº 5453808)), BDNF (GI Nºs 25306267, 25306235, 25306253, 25306257, 25306261, 25306264; IFRD1 (GI Nº 4504607)).

O termo "Tm", também referido como a "temperatura de transição", é a temperatura na qual 50% de uma macromolécula, por exemplo, molécula de ligação, se torna desnaturada, e é considerada como sendo o parâmetro padrão para descrever a estabilidade térmica de uma proteína.

Conforme usado aqui, neste Pedido de patente, o termo "resíduo de formação de *estruturação*" se refere a resíduos aminoácidos ou posições de resíduos que não estão em uma interface (por exemplo, a interface VH/VL) mas que são importantes para manter a interface. Estes resíduos aminoácidos não interagem fisicamente com os resíduos da interface sobre o domínio oposto ou contribuem a área superficial para a interface, mas não

obstante são importantes para proporcionar contexto estrutural adequado para resíduos da interface. Os resíduos aminoácidos referidos *estrutura* a interação entre VH e VL.

Diz-se de duas ou mais posições de resíduos aminoácidos dentro de uma seqüência de polipeptídeo candidata que normalmente ocorrem juntos que "covariam" ("posições de resíduos covariáveis" ou "posições de resíduos covariantes"). Covariância entre duas ou mais posições de aminoácido é observada quando o tipo de aminoácido encontrado em uma primeira posição de aminoácido é dependente do tipo de aminoácido encontrado em outra posição de aminoácido. Isto é, quando um aminoácido particular é encontrado em uma primeira posição dentro de uma seqüência, um segundo aminoácido particular é geralmente encontrado em uma segunda posição dentro da seqüência.

Conforme usado aqui, neste Pedido de patente, o termo "dobra de Ig" inclui um domínio de proteína encontrado em proteínas pertencentes à superfamília de proteínas das imunoglobulinas. Conforme é de conhecimento geral na técnica, a dobra de Ig é uma característica peculiar da superfamília das imunoglobulinas (vide, por exemplo, Bork, P., Holm, L. & Sander, C. 1994. *The Immunoglobulin Fold. J. Mol. Biol.* 242, 309-320). Estruturas típicas para cada classe de dobra de Ig são representadas na Figura 73.

II. Moléculas scFv Estabilizadas

Em uma modalidade, uma molécula de ligação da invenção é uma molécula scFv estabilizada. As moléculas scFv estabilizadas da invenção podem compreender um ligante de scFv interposto entre um domínio V_H e um domínio V_L, em que os domínios VH e V_L são encadeados por uma ligação dissulfeto entre um aminoácido no domínio V_H e um aminoácido no domínio V_L. Em outras modalidades, as moléculas scFv estabilizadas da invenção compreendem um ligante de scFv tendo uma extensão ou composição otimizada. Em ainda outras modalidades, as moléculas scFv estabilizadas da invenção compreendem um domínio V_H ou V_L tendo no mínimo uma ou mais substituições de aminoácido estabilizantes. Em ainda outra modalidade, uma molécula scFv estabilizada da invenção compreende no mínimo

duas das características estabilizantes listadas acima.

As moléculas scFv estabilizadas da invenção têm aprimorada estabilidade. Em uma modalidade, populações das moléculas scFv estabilizadas da invenção ou polipeptídeos compreendendo as mesmas são expressados como uma proteína solúvel monomérica da qual é não mais de 10% em forma dimérica, tetramérica, ou agregada de modo diverso. Em outra modalidade, populações das moléculas scFv estabilizadas da invenção têm domínios VH e VL com valores de Tm de mais de 55°C. Em outra modalidade, populações das moléculas scFv estabilizadas da invenção têm uma T50 de mais de 49°C. Em outra modalidade, populações das moléculas scFv estabilizadas da invenção têm uma T50 de mais de 40, 41, 42, 43, 44, 45, 46, 47, ou 48°C. Em ainda outra modalidade, populações das moléculas scFv estabilizadas da invenção têm uma T50 de mais de 50, 51, 52, 53, 54, 55, 56, 57, 58, ou 49°C.

As moléculas scFv da invenção ligam a uma molécula alvo de interesse. Os domínios VH e VL usados para fabricar um scFv podem ser derivados do mesmo anticorpo ou de anticorpos diferentes. Em outra modalidade, um VH ou VL para uso em um scFv estabilizado da invenção pode compreender um ou mais CDRs as quais ligam a um alvo de interesse, enquanto o restante do domínio VH ou VL é derivado de um anticorpo diferente ou é sintético. Em uma modalidade preferencial, uma molécula de ligação da invenção compreende no mínimo uma CDR de um anticorpo, por exemplo, um anticorpo conhecido na técnica para ligar a um alvo de interesse. Em outra modalidade, uma molécula de ligação da invenção compreende no mínimo duas CDRs de um dado anticorpo. Em outra modalidade, uma molécula de ligação da invenção compreende no mínimo três CDRs de um dado anticorpo. Em outra modalidade, uma molécula de ligação da invenção compreende no mínimo quatro CDRs de um dado anticorpo. Em outra modalidade, uma molécula de ligação da invenção compreende no mínimo cinco CDRs de um dado anticorpo. Em outra modalidade, uma molécula de ligação da invenção compreende no mínimo seis CDRs de um dado anticorpo. Em uma modalidade preferencial, uma molécula de ligação da invenção com-

preende no mínimo um domínio VH de um anticorpo, por exemplo, um anticorpo conhecidas na técnica para ligar a um alvo de interesse. Em uma modalidade preferencial, uma molécula de ligação da invenção compreende no mínimo um domínio VL de um dado anticorpo. Em outra modalidade preferencial, uma molécula de ligação da invenção compreende no mínimo um domínio VH e um domínio VL de um anticorpo conhecido na técnica para ligar um alvo de interesse. Moléculas scFv podem ser construídas em uma orientação VH-ligante-VL ou orientação VL-ligante-VH.

A estabilidade de moléculas scFv da invenção ou de proteínas de fusão compreendendo as mesmas pode ser avaliada em referência às propriedades biofísicas (por exemplo, estabilidade térmica) de uma molécula scFv convencional (não-estabilizada) ou uma molécula de ligação compreendendo uma molécula scFv convencional. Em uma modalidade, as moléculas de ligação da invenção têm uma estabilidade térmica que é de mais de cerca de 0,1, cerca de 0,25, cerca de 0,5, cerca de 0,75, cerca de 1, cerca de 1,25, cerca de 1,5, cerca de 1,75, cerca de 2, cerca de 3, cerca de 4, cerca de 5, cerca de 6, cerca de 7, cerca de 8, cerca de 9, ou cerca de 10 graus Celsius do que uma molécula de ligação controle (por exemplo, uma molécula scFv convencional). Moléculas scFv estabilizadas da invenção incluem aquelas identificadas usando métodos da invenção conforme descrito na Seção VIII abaixo.

Em outras modalidades, as moléculas scFv estabilizadas da invenção compreendem um ligante de scFv com uma extensão e/ou composição de aminoácidos otimizadas. Ligantes preferenciais de scFv da invenção melhoram a estabilidade térmica de uma molécula de ligação da invenção por no mínimo cerca de 2°C ou 3°C comparada com uma molécula de ligação convencional. Em uma modalidade, uma molécula de ligação da invenção tem uma estabilidade térmica aprimorada de 1 °C comparada com uma molécula de ligação convencional. Em outra modalidade, uma molécula de ligação da invenção tem uma estabilidade térmica aprimorada de 2 °C comparada com uma molécula de ligação convencional. Em outra modalidade, uma molécula de ligação da invenção tem uma estabilidade térmica aprimo-

rada de 4, 5, 6 °C comparada com uma molécula de ligação convencional. Podem ser feitas comparações, por exemplo, entre as moléculas scFv da invenção e moléculas scFv preparadas usando métodos da técnica anterior ou entre moléculas scFv e fragmentos fab de um anticorpo a partir do qual foram derivados os VH e VL de scFv. A estabilidade térmica pode ser medida usando métodos conhecidos na técnica. Por exemplo, em uma modalidade, pode ser medida a Tm. Métodos para medir Tm e outros métodos para determinar a estabilidade protéica são descritos em mais detalhes abaixo.

Em uma modalidade, o ligante de scFv consiste da seqüência de aminoácidos $(Gly_4Ser)_4$ ou compreende uma seqüência $(Gly_4Ser)_4$. Outros ligantes exemplares compreendem ou consistem em seqüências $(Gly_4Ser)_3$ e $(Gly_4Ser)_5$. Ligantes de scFv da invenção podem ser de extensões variáveis. Em uma modalidade, um ligante de scFv da invenção tem a partir de cerca de 5 a cerca de 50 aminoácidos de extensão. Em outra modalidade, um ligante de scFv da invenção tem a partir de cerca de 10 a cerca de 40 aminoácidos de extensão. Em outra modalidade, um ligante de scFv da invenção tem a partir de cerca de 15 a cerca de 30 aminoácidos de extensão. Em outra modalidade, um ligante de scFv da invenção tem a partir de cerca de 17 a cerca de 28 aminoácidos de extensão. Em outra modalidade, um ligante de scFv da invenção tem a partir de cerca de 19 a cerca de 26 aminoácidos de extensão. Em outra modalidade, um ligante de scFv da invenção tem a partir de cerca de 21 a cerca de 24 aminoácidos de extensão.

Ligantes scFv podem ser introduzidos em seqüências de polipeptídeos usando técnicas conhecidas na técnica. Por exemplo, em uma modalidade, pode ser usada mutagênese de PCR. Modificações podem ser confirmadas por análise de seqüência de DNA. DNA de plasmídeo pode ser usado para transformar células hospedeiras para produção dos polipeptídeos estáveis produzidos.

Em algumas modalidades, as moléculas scFv estabilizadas da invenção compreendem no mínimo uma ligação dissulfeto a qual liga um aminoácido no domínio VL com um aminoácido no domínio VH. Resíduos

cisteína são necessários para proporcionar ligações dissulfeto. Ligações dissulfeto podem ser incluídsa em uma molécula scFv da invenção, por exemplo, para conectar FR4 de VL e FR2 de VH ou para conectar FR2 de VL e FR4 de VH. Posições típicas para ligação dissulfeto incluem: 43, 44, 45, 46, 47, 103, 104, 105, e 106 de VH e 42, 43, 44, 45, 46, 98, 99, 100, e 101 de VL, numeração de Kabat. Combinações típicas de posições de aminoácidos os quais são mutados para resíduos cisteína incluem: VH44-VL100, VH105-VL43, VH105-VL42, VH44-VL101, VH106-VL43, VH104-VL43, VH44-VL99, VH45-VL98, VH46-VL98, VH103-VL43, VH103-VL44, e VH103-VL45.

Em uma modalidade, uma ligação dissulfeto liga o aminoácido 44 de V_H e o aminoácido 100 de V_L.

Modificações dos genes os quais codificam os domínios VH e VL podem ser realizadas usando técnicas conhecidas na técnica, por exemplo, mutagênese sítio-dirigida.

Em uma modalidade, uma molécula scFv estabilizada da invenção compreende um ligante de scFv tendo a seqüência de aminoácidos $(Gly_4\ Ser)_4$ interposto entre um domínio V_H e um domínio V_L, em que os domínios VH e V_L são encadeados por uma ligação dissulfeto entre um aminoácido no domínio V_H na posição do aminoácido 44 e um aminoácido no domínio V_L na posição do aminoácido 100.

Em outras modalidades as moléculas scFv estabilizadas da invenção compreendem uma ou mais mutações estabilizantes dentro de um domínio variável (VH ou VL) do scFv. Em uma modalidade, a mutação estabilizante é selecionada entre o grupo consistindo em:

a) substituição de um aminoácido (por exemplo, glutamina) na posição de Kabat 3 de VL, por exemplo, com uma alanina, uma serina, uma valina, um ácido aspártico, ou uma glicina;

b) substituição de um aminoácido (por exemplo, serina) na posição de Kabat 46 de VL, por exemplo, com leucina;

c) substituição de um aminoácido (por exemplo, serina) na posição de Kabat 49 de VL, por exemplo, com tirosina ou serina;

d) substituição de um aminoácido (por exemplo, serina ou vali-

na) na posição de Kabat 50 de VL, por exemplo, com serina, treonina, e arginina, ácido aspártico, glicina, ou lisina;

e) substituições de aminoácidos (por exemplo, serina) na posição de Kabat 49 e (por exemplo, serina) na posição de Kabat 50 de VL, respectivamente com tirosina e serina; tirosina e treonina; tirosina e arginina; tirosina e glicina; serina e arginina; ou serina e lisina;

f) substituição de um aminoácido (por exemplo, valina) na posição de Kabat 75 de VL, por exemplo, com isoleucina;

g) substituição de um aminoácido (por exemplo, prolina) na posição de Kabat 80 de VL, por exemplo, com serina ou glicina;

h) substituição de um aminoácido (por exemplo, fenilalanina) na posição de Kabat 83 de VL, por exemplo, com serina, alanina, glicina, ou treonina;

i) substituição de um aminoácido (por exemplo, ácido glutâmico) na posição de Kabat 6 de VH, por exemplo, com glutamina;

j) substituição de um aminoácido (por exemplo, lisina) na posição de Kabat 13 de VH, por exemplo, com glutamato;

k) substituição de um aminoácido (por exemplo, serina) na posição de Kabat 16 de VH, por exemplo, com glutamato ou glutamina;

l) substituição de um aminoácido (por exemplo, valina) na posição de Kabat 20 de VH, por exemplo, com uma isoleucina;

m) substituição de um aminoácido (por exemplo, asparagina) na posição de Kabat 32 de VH, por exemplo, com serina;

n) substituição de um aminoácido (por exemplo, glutamina) na posição de Kabat 43 de VH, por exemplo, com lisina ou arginina;

o) substituição de um aminoácido (por exemplo, metionina) na posição de Kabat 48 de VH, por exemplo, com uma isoleucina ou uma glicina;

p) substituição de um aminoácido (por exemplo, serina) na posição de Kabat 49 de VH, por exemplo, com glicina ou alanina;

q) substituição de um aminoácido (por exemplo, valina) na posição de Kabat 55 de VH, por exemplo, com uma glicina;

r) substituição de um aminoácido (por exemplo, valina) na posição de Kabat 67 de VH, por exemplo, com uma isoleucina ou uma leucina;

s) substituição de um aminoácido (por exemplo, ácido glutâmico) na posição de Kabat 72 de VH, por exemplo, com aspartato ou asparagina;

t) substituição de um aminoácido (por exemplo, fenilalanina) na posição de Kabat 79 de VH, por exemplo, com serina, valina, ou tirosina; e

u) substituição de um aminoácido (por exemplo, prolina) na posição de Kabat 101 de VH, por exemplo, com um ácido aspártico.

Em uma modalidade exemplar, uma molécula scFv estabilizada da invenção compreende duas ou mais das mutações estabilizantes descritas em a) a u) acima. Em uma modalidade exemplar, uma molécula scFv estabilizada da invenção compreende uma substituição de um aminoácido na posição de Kabat 49 de VL, por exemplo, com tirosina ou serina e uma substituição de um aminoácido na posição de Kabat 50 de VL, por exemplo, com, treonina, e arginina, ácido aspártico, glicina, ou lisina. Em outra modalidade exemplar, uma molécula scFv estabilizada da invenção compreende uma substituição de um aminoácido na posição de Kabat 16 de VH, por exemplo, com um glutamato ou glutamina e uma substituição de um aminoácido na posição de Kabat 46 de VL, por exemplo, com uma lisina. Em outra modalidade uma molécula de ligação da invenção compreende uma substituição de um aminoácido na posição de Kabat 16 de VH, por exemplo, com um glutamato ou glutamina; uma substituição de um aminoácido na posição de Kabat 46 de VL, por exemplo, com uma lisina; e uma substituição de um aminoácido na posição de Kabat 55 de VH, por exemplo, com uma glicina. Em outra modalidade uma molécula de ligação da invenção compreende uma substituição de um aminoácido na posição de Kabat 16 de VH, por exemplo, com um glutamato ou glutamina; uma substituição de um aminoácido na posição de Kabat 46 de VL, por exemplo, com uma lisina; uma substituição de um aminoácido na posição de Kabat 55 de VH, por exemplo, com uma glicina; e substituição de um aminoácido na posição de Kabat 101 de VH, por exemplo, com um ácido aspártico. Em outra modalidade, uma molé-

cula de ligação da invenção compreende uma substituição de um aminoácido na posição de Kabat 6 de VH, por exemplo, com glutamina. Em outra modalidade, uma molécula de ligação da invenção compreende uma substituição de um aminoácido na posição de Kabat 13 de VH, por exemplo, com glutamato. Em outra modalidade, uma molécula de ligação da invenção compreende uma substituição de um aminoácido na posição de Kabat 16 de VH, por exemplo, com glutamato ou glutamina. Em outra modalidade, uma molécula de ligação da invenção compreende uma substituição de um aminoácido na posição de Kabat 20 de VH, por exemplo, com uma isoleucina; Em outra modalidade, uma molécula de ligação da invenção compreende uma substituição de um aminoácido na posição de Kabat 32 de VH, por exemplo, com serina. Em outra modalidade, uma molécula de ligação da invenção compreende uma substituição de um aminoácido na posição de Kabat 43 de VH, por exemplo, com lisina ou arginina. Em outra modalidade, uma molécula de ligação da invenção compreende uma substituição de um aminoácido na posição de Kabat 48 de VH, por exemplo, com uma isoleucina ou uma glicina. Em outra modalidade, uma molécula de ligação da invenção compreende uma substituição de um aminoácido na posição de Kabat substituição 49 de VH, por exemplo, com glicina ou alanina. Em outra modalidade, uma molécula de ligação da invenção compreende uma substituição de um aminoácido na posição de Kabat 55 de VH, por exemplo, com uma glicina. Em outra modalidade, uma molécula de ligação da invenção compreende uma substituição de um aminoácido na posição de Kabat 67 de VH, por exemplo, com uma isoleucina ou uma leucina. Em outra modalidade, uma molécula de ligação da invenção compreende uma substituição de um aminoácido na posição de Kabat 72 de VH, por exemplo, com aspartato ou asparaginas. Em outra modalidade, uma molécula de ligação da invenção compreende uma substituição de um aminoácido na posição de Kabat 79 de VH, por exemplo, com serina, valina, ou tirosina. Em outra modalidade, uma molécula de ligação da invenção compreende uma substituição de um aminoácido na posição de Kabat 101 de VH, por exemplo, com um ácido aspártico.

Em outra modalidade exemplar, uma molécula scFv estabilizada da invenção compreende uma ou mais das substituições de aminoácidos estabilizantes descritas aqui, neste Pedido de patente, e um ligante de scFv com uma extensão ou composição otimizada (por exemplo, $(Gly_4Ser)_4$). Em outra modalidade exemplar, uma molécula scFv estabilizada da invenção compreende uma ou mais das substituições de aminoácidos descritas aqui, neste Pedido de patente, e uma ligação dissulfeto a qual liga um aminoácido no domínio VL com um aminoácido no domínio VH (por exemplo, VH44-VL100). Em ainda outra modalidade exemplar, uma molécula scFv estabilizada da invenção compreende uma ou mais das substituições de aminoácidos descritas aqui, neste Pedido de patente, um ligante de scFv com uma extensão ou composição otimizada (por exemplo, $(Gly_4Ser)_4$), e uma ligação dissulfeto a qual liga um aminoácido no domínio VL com um aminoácido no domínio VH (por exemplo, VH44-VL100).

Moléculas estabilizadas scFv podem ser expressadas usando técnicas reconhecidas na técnica. Por exemplo, em uma modalidade, moléculas similares podem ser expressadas usando um vetor de expressão apropriado para expressão em um sistema de expressão celular, por exemplo, um sistema de expressão bacteriano ou mamífero.

Em uma modalidade, moléculas scFv podem ser expressadas em E. coli, por exemplo, usando um vetor apropriado para expressão periplásmica. Seqüências adicionais podem ser incluídas para otimizar expressão, por exemplo, uma seqüência de sinal e/ou uma etiqueta para facilitar purificação e/ou detecção do scFv.

Em uma modalidade, aditivos, tais como 1 a 2% de triton (por exemplo, triton x-100) ou 1 a 2% de glicina ou uma combinação dos mesmos (por exemplo, 1% de glicina e 1% de triton) podem ser adicionados para facilitar secreção do periplasma para o meio.

III. Métodos para Prever / Determinar Estabilidade de Proteína

Em alguns aspectos, a invenção proporciona métodos para prever, *a priori*, problemas biofísicos potenciais com proteínas selecionadas para expressão em grande escala, por exemplo, proteínas terapêuticas ou en-

zimas industriais. Em alguns aspectos típicos, os métodos da invenção possibilitam que uma pessoa versada na técnica evite expressar seqüências de proteínas que são previstas para terem estabilidade inerentemente pobre quando expressadas de modo recombinante, por exemplo, em células mamíferas. Em aspectos alternativos, os métodos da invenção podem ser empregados para identificar uma variante de uma seqüência de proteína que é prevista para ter propriedades biofísicas aprimoradas, incluindo, mas não limitadas a aprimorada estabilidade, aprimorada estabilidade para alterações no pH, e função bioquímica reforçada.

Em alguns aspectos, a invenção proporciona métodos computacionais para prever uma propriedade biofísica de (por exemplo, estabilidade térmica) algumas proteínas candidatas, ou uma variante de seqüência ou homólogo da mesma, com base na seqüência de polipeptídeo (por exemplo, seqüência de aminoácidos) da proteína candidata. Em outros aspectos, os métodos da invenção empregam modelagem estrutural para prever a estabilidade da proteína candidata ou para identificar homólogos ou variantes da proteína candidata com estabilidade típica conhecida que podem ser usados para melhorar a estabilidade de uma proteína candidata, por exemplo, uma molécula scFv.

a. Análise da Covariação

Em alguns aspectos típicos, a invenção proporciona um método computacional aprimorado denominado "análise de covariação" para prever uma propriedade biofísica (por exemplo, estabilidade protéica). Conforme usado aqui, neste Pedido de patente, "análise de covariação" se refere a um método computacional para identificar duas ou mais posições de resíduos aminoácidos dentro de uma seqüência de polipeptídeo candidata que normalmente ocorrem juntos, ou covariam, em homólogos da seqüência candidata (aqui, neste Pedido de patente, "posições de resíduos covariáveis" ou "posições de resíduos covariantes"). Covariância entre duas ou mais posições de aminoácido é observada quando o tipo de aminoácido encontrado em uma primeira posição de aminoácido é dependente do tipo de aminoácido encontrado em outra posição de aminoácido. Isto é, quando um aminoá-

cido em particular é encontrado em uma primeira posição dentro de uma seqüência, um segundo aminoácido em particular é geralmente encontrado em uma segunda posição dentro da seqüência.

Como resíduos covariantes são prováveis de terem coevoluídos, uma observação de covariação indica que a compatibilidade de resíduos aminoácidos presentes em posições normalmente covariáveis é provável que seja importante por razões funcionais ou estruturais. Por conseguinte, o método de covariação da invenção pode ser usado para identificar um ou mais resíduos covariáveis ou grupos de resíduos (por exemplo, pares de resíduos covariáveis) dentro de uma seqüência de polipeptídeo (por exemplo, uma seqüência de polipeptídeo candidata). Conforme usado aqui, neste Pedido de patente, um "resíduo covariável" (também denominado um "resíduo encadeado") é um aminoácido o qual é estatisticamente prevalente em uma posição de resíduo covariável dentro de uma seqüência de polipeptídeo.

Em algumas modalidades, os métodos de covariação da invenção podem ser usados para identificar importantes resíduos funcionais em uma seqüência candidata. Além disso, como o número de resíduos covariáveis dentro de um polipeptídeo candidato é preditivo de sua estabilidade, o método de covariação da invenção pode ser usado para prever a estabilidade da proteína candidata.

Em algumas outras modalidades, os métodos de covariação da invenção podem ser empregados para orientar designs de proteínas com êxito. Em uma modalidade típica, o método de covariação da invenção pode ser usado para identificar covariáveis aminoácidos dentro de uma seqüência candidata. Caso presentes, os resíduos covariáveis são preferencialmente retidos durante engenharia subseqüente da seqüência de proteína. Em outra modalidade exemplar, os métodos de covariação da invenção podem ser usado para identificar as posições de aminoácido de resíduos não-covariáveis dentro de uma seqüência de polipeptídeo candidata em que aminoácidos correspondentes em outras proteínas (por exemplo, proteínas correspondentes às seqüências de uma série de referência) são normalmen-

te aminoácidos covariáveis. Uma vez identificados, resíduos não-covariáveis podem ser substituídos (por exemplo, usando metodologia de DNA recombinante) com resíduos covariáveis correspondentes para reforçar a função ou melhorar as propriedades biofísicas (por exemplo, estabilidade) da seqüência de polipeptídeo candidata.

Posições de resíduos covariantes podem ser identificadas por uma análise estatística (por exemplo, uma análise de correlação) de posições de resíduos dentro de um banco de dados de seqüências de polipeptídeos relacionadas (por exemplo, uma série de referência alinhada ou alinhamento de múltiplas seqüências). Em modalidades preferenciais, a análise de covariação envolve uma comparação estatística de seqüências de polipeptídeos candidatas contra uma banco de dados curados de diversas seqüências de polipeptídeos tendo a mesma dobra estrutural (por exemplo, uma dobra de Ig).

Na nova análise de covariação descrita aqui, neste Pedido de patente, as seguintes etapas básicas são tomadas para prever a estabilidade de um polipeptídeo candidato:

1. Proporcionar uma série de seqüências homólogas (aqui, neste Pedido de patente, uma "série de referência") correspondentes a uma seqüência de um polipeptídeo candidato ou um domínio ou porção da mesma (aqui, neste Pedido de patente, uma "seqüência de domínio de ensaio").

2. Alinhar a seqüências da série de referência para gerar uma série alinhada de seqüência homóloga (aqui, neste Pedido de patente, uma "série alinhada).

3. Determinar covariação entre duas ou mais posições de resíduos (por exemplo, no mínimo um par de aminoácidos) dentro da série de referência para gerar dados de covariação ou uma série de dados de covariação ("série de dados de covariação").

4. Usar os dados de covariação para prever a estabilidade do polipeptídeo candidato.

Etapa 1: Proporcionar uma Série de Referência

Os métodos de covariação da invenção empregam uma série de de seqüências (uma "série de referência") as quais são homólogas a (isto é, relacionadas com) uma seqüência de interesse ou "seqüência de ensaio". A série de referência pode compreender uma série de seqüências homólogas tendo grau moderado a elevado de similaridade com a seqüência de ensaio. Em algumas modalidades, as seqüências homólogas da série de referência têm um grau moderado a elevado de similaridade das seqüências de aminoácidos. Em modalidades mais preferenciais, as seqüências homólogas da série de referência têm um grau moderado a elevado de similaridade estrutural. Em modalidades ainda mais preferenciais, as seqüências homólogas da série de referência têm um grau elevado de similaridade estrutural (por exemplo, elas partilham no mínimo um domínio ou dobra de proteína). A série de seqüências homólogas não precisa ser grande, na medida que seja obtida uma seleção razoavelmente não tendenciosa.

Em modalidades exemplares, o método de covariação da invenção emprega uma série de referência curada. Conforme usado aqui, neste Pedido de patente, o termo "série de referência curada" se refere a uma série de referência na qual seqüências de componentes foram eliminadas ou seqüências adicionais adicionadas de acordo com alguns critérios de seleção. Por conseguinte, enquanto uma referência não curada é gerada por seleção não tendenciosa de seqüências de componentes, um banco de dados curado é gerado por seleção tendenciosa. Por exemplo, a série de referência de seqüências homólogas pode ser selecionada para eliminar algumas seqüências, por exemplo, seqüência redundante. Alternativamente, um banco de dados curado pode ser expandido, por exemplo, para incluir um número adequado de seqüências não-redundantes ou não-idênticas.

Em algumas modalidades, o banco de dados curado compreende no mínimo cem seqüências (por exemplo, 100, 150, 200, 250, 300, 400, 500, 750, ou mais seqüências). Em uma modalidade, o banco de dados curado compreende no mínimo mil seqüências (por exemplo, 1000, 2000, 3000, 4000, 5000, 6000, 7000, 8000, 9000, ou mais seqüências). Em outra

modalidade, o banco de dados curado compreende no mínimo dez mil seqüências (por exemplo, 10.000, 20.000, 50.000, 100.000, 250.000, 500.000, 750.000 ou mais seqüência. Em uma modalidade o banco de dados curado compreende no mínimo 1 milhão de seqüências individuais (por exemplo, 1 milhão, 2 milhões, 5 milhões, 10 milhões, ou mais seqüências).

Em modalidades preferenciais, o método de covariação da invenção emprega uma série de referência curada tendo no mínimo 50% de diversidade (por exemplo, no mínimo 60%, no mínimo 70%, no mínimo 75%, no mínimo 80%, no mínimo 90%, no mínimo 95% de diversidade, no mínimo 97%, no mínimo 98%, no mínimo 99% ou mais de diversidade) de modo a minimizar covariações artificiais ou não-informativas. Conforme usado aqui, neste Pedido de patente, o termo "% de diversidade" se refere à percentagem de seqüências na série de referência que são não-redundantes ou não-idênticas. Conforme usado aqui, neste Pedido de patente, o termo "não redundante" se refere a uma seqüência a qual tem menos de 100% de identidade de seqüência com cada outra seqüência na série de referência, isto é, uma seqüência a qual difere no mínimo em uma posição de aminoácido de cada outra seqüência na série de referência. Em algumas modalidades, uma seqüência da série de referência tem menos de 99%, menos de 95%, menos de 90%, ou menos de 85% de identidade de seqüência com cada outra seqüência na série de referência. Em modalidades preferenciais, uma seqüência da série de referência tem menos de 80% de identidade de seqüência com cada outra seqüência na série de referência (por exemplo, menos de 75%, menos de 70%, menos de 65%, menos de 60%, ou menos de 50% de identidade de seqüência). Em modalidades particularmente preferenciais, cada seqüência da série de referência tem menos de de 80% de identidade de seqüência com cada outra seqüência na série de referência. Em modalidades exemplares, as seqüências da série de referência são curadas para a diversidade preferencial usando uma ferramenta de filtração a qual seleciona seqüências contendo tipos de seqüências mais comuns (por exemplo, uma ferramenta de classificação de Henikoff). Quando duas seqüências têm extensão igual, classificação pr peso de Henikoff assegura que seqüências

com tipos de resíduos raros são retidos enquanto seqüências com tipos de resíduos mais comuns são descartados (Henikoff *et al., J. Mol. Biol.*, 243: 574-578 (1994)).

Em outras modalidades preferenciais, a série de referência compreende seqüências codificadas por genes os quais são não evolucionarmente próximos de genes codificando outras seqüências no banco de dados. Por exemplo, para aumentar a diversidade, a série de referência pode compreender uma ou mais seqüências ortólogas, que é uma seqüência a qual é de uma espécie diferente (por exemplo, espécie mamífera diferente)) mas a qual tem a mesma função bioquímica ou similar. Em uma modalidade típica, a série de referência pode compreender uma seqüência humana e no mínimo uma seqüência não humana (por exemplo, uma seqüência de mamífero não-humano). Em outra modalidade exemplar, a série de referência pode compreender no mínimo uma seqüência para uma espécie mamífera (por exemplo, um ser humano, chimpanzé, cão, vaca, porco, gato, rato, ou camundongo) e no mínimo uma seqüência não-mamífera (por exemplo, uma seqüência de vertebrado não-mamífero (por exemplo, uma seqüência de teleósteo ou avícola), uma seqüência de invertebrado (por exemplo, uma seqüência de inseto (por exemplo, Drosophila) ou de nematódeo (por exemplo, C. elegans)), ou uma seqüência fúngica, de planta, bacteriana, ou viral). Em outras modalidades, a série de referência pode compreender uma ou mais seqüências parálogas, que é seqüências da mesma espécie que uma primeira seqüência, mas a qual tem a mesma função bioquímica ou similar.

Em ainda outras modalidades, as seqüências da série de referência têm acima de um determinado limiar de extensão. Preferencialmente, as seqüências da série de referência têm no mínimo 20 resíduos de extensão (por exemplo, 25, 30, ou 40 resíduos de extensão). Mais preferencialmente, as seqüências da série de referência têm no mínimo 50 resíduos de extensão (por exemplo, 60, 65, 70, 75, 80, 85, 90, ou 95 resíduos de extensão). Ainda mais preferenciais são seqüências de no mínimo 100 resíduos de extensão (por exemplo, 110, 120, 130, 140, 150, 160, 170, 180, ou mais resíduos de extensão). Em modalidades exemplares, as seqüências da série

de referência são curadas para a extensão preferencial por filtração usando uma ferramena de filtração (por exemplo, uma contagem de resíduo sem intervalos). Classificação por contagem de resíduo sem intervalos decrescente assegura que seqüências mais curtas (por exemplo, seqüências com intervalo) sejam filtradas sobre seqüências mais longas.

As seqüências podem ser obtidas através de estudos de homologia baseada em seqüência para quaiquer dos bancos de dados de seqüências conhecidos na técnica (por exemplo, bancos de dados NCBI, TIGR). Em uma modalidade exemplar, um estudo BLAST de rotina com parâmetros padrão pode ser realizado com a seqüência de ensaio para recuperar seqüências homólogas de interesse. Seqüências com uma percentagem mínima de identidade de seqüência (por exemplo, mais de 25%, 30%, 35%, 40%, 45% ou 50% de identidade de seqüência, preferencialmente mais de 30% de identidade de seqüência) com a seqüência de ensaio podem ser selecionadas para inclusão na série de referência.

Em algumas modalidades preferenciais, a série de referência é curada para compreender seqüências de proteínas tendo a mesma classe de dobra de proteína (por exemplo, proteínas tendo SH3, TPR, GPCR, serina protease, aspartil protease, globina, imunoglobulina, ou outras dobras). Proteínas similares ou estruturas podem ser identificadas e coletadas por meio de técnicas de bioinformática conhecidas na técnica (por exemplo, ser pesquisadas à base de estrutura usando o banco de dados de domínio conservado dos Institutos Nacionais de Saúde dos Estados Unidos (CDD, *Conserved Domain Database*, *National Institutes of Health*), ou pesquisando o banco de dados ASTRAL ou o banco de dados de proteína RCSB (PDB) usando uma classificação SCOP desejada). Preferencialmente, as seqüências são derivadas de proteínas cujas estruturas tridimensionais foram solvidas em alta resolução, por exemplo, por cristalografia de raios X.

Em modalidades preferenciais, seqüências correspondentes a estruturas coletadas são filtradas usando técnicas reconhecidas na técnica para remover estruturas categorizadas erroneamente, incompletas, redundantes, ou de domínio permutado. Em uma modalidade, as estruturas são

inspecionadas visualmente (por exemplo, usando o visor SwissPDB Viewer) para rupturas na seqüência correspondente devido a densidades não resolvidas ou permuta de domínio. Os arquivos PDB de estruturas imperfeitas são removidos manualmente da série de estruturas correspondentes. Em outra modalidade, seqüências (por exemplo, seqüências FASTA) correspondentes às estruturas coletadas são filtradas para remover quaisquer seqüências que sejam 100% idênticas, ou subcadeias de combinações perfeitas de seqüências remanescentes. Em ainda outras modalidades, estruturas PDB com seqüências de aminoácidos aberrantemente longas ou curtas (um marco de categorização estrutural errônea) são selecionadas das séries de dados de estruturas. Os critérios de corte de extensão podem ser determinados, por exemplo, examinando um histograma de todas as extensões de seqüência de seqüências da série de referência. Em ainda outras modalidades, as estruturas são inspecionadas visualmente (por exemplo, usando o visor Swiss PDB) para dobramento incorreto. Quaisquer estruturas que não se conformam com a topologia padrão da dobra de proteína são preferencialmente descartadas.

Em modalidades exemplares, a série de referência curada compreende seqüências correspondentes a uma dobra do tipo de imunoglobulina ("dobra de Ig") de uma proteína pertencente à superfamília de proteínas das imunoglobulinas. Conforme é de conhecimento geral na técnica, a dobra de Ig é uma característica diferencial da superfamília das imunoglobulinas (vide, por exemplo, Bork, P., Holm, L. & Sander, C. 1994. The Immunoglobulin Fold. *J. Mol. Biol.* 242, 309-320). A dobra de Ig aparece freqüentemente em proteínas de mamíferos e serve como uma plataforma para várias funções – particularmente interações proteína-proteína. Por exemplo, todas as imunoglobulinas e a maioria dos receptores de imunoglobulina são compostos de múltiplos domínios de dobra de Ig. Dobras de Ig podem ser subdivididas em várias subfamílias incluindo C1, C2, I, e V. Embora membros de todas as subfamílias compreendam uma topologia de chave gregade duas-β folha, estas subfamílias diferem pelo número de filamentos β sobre cada folha e pelas conexões através das β-sheets (vide, por exemplo, AF Williams,

Immunology Today, 8, (1987)). Estruturas típicas para cada classe de dobra de Ig são representadas na Figura 73.

Em uma modalidade, a série de referência compreende uma seqüência de dobra de Ig de uma proteína da superfamília das imunoglobulinas selecionada entre o grupo consistindo em uma proteína de adesão celular, uma integrina, um alergeno, um receptor de células T, uma proteína do complexo de histocompatibilidade principal (por exemplo, uma proteína MHC Classe I ou MHC Classe II), um receptor de imunoglobulina (por exemplo, um Fc γ receptor (por exemplo, FcγRI, FcγRIIa)), e uma imunoglobulina (por exemplo, uma imunoglobulina IgG, IgM, IgA, ou IgE).

Em algumas modalidades preferenciais, todas as seqüências da série de referência correspondem a uma dobra de Ig ou porção da mesma. Em uma modalidade, a dobra de Ig é uma dobra C1. Em outra modalidade, a dobra de Ig é uma dobra C2. Em outra modalidade, a dobra de Ig é uma dobra V. Em ainda outra modalidade, a dobra de Ig é uma dobra I. Em outra modalidade, todas as seqüências da série de referência correspondem a uma dobra C1. Em outra modalidade, todas as seqüências da série de referência correspondem a uma dobra C2. Em outra modalidade, todas as seqüências da série de referência correspondem a uma dobra I. Em algumas modalidades preferenciais, as seqüências da série de referência de dobra de Ig têm dentro de cerca de 75 a cerca de 150 resíduos de extensão.

Etapa 2: Alinhar a Série de Referência

Na segunda etapa, as seqüências da série de referência são alinhadas precisamente para gerar uma série alinhada de seqüências (ou alinhamento). Em uma modalidade, as seqüências da série de referência são alinhadas usando algoritmos de alinhamento de seqüência (por exemplo, LALIGN ou BLAST e outros alinhamentos reconhecidos na técnica descrita acima) para criar uma série alinhada de seqüências (aqui, neste Pedido de patente, um "alinhamento de seqüências à base de seqüência"). Em outra modalidade, estruturas correspondentes as seqüências da série de referência são alinhadas usando um algoritmo de alinhamento estrutural (por exemplo, combinação de estrutura secundária (*Secondary Structure Mat-*

ching, SSM) usando o pacote de alinhamento de estrutura *Schrödinger structalign*) para criar uma série alinhada de estruturas (aqui, neste Pedido de patente, um "alinhamento de estrutura"). Preferencialmente, o algoritmo de alinhamento estrutural assegura que regiões de núcleo (por exemplo, filamentos β) de cada estrutura sejam alinhadas, ao invés de alças intervenientes. Em ainda outra modalidade, as seqüências da série de referência são alinhadas por métodos à base de estrutura (por exemplo, combinando aminoácidos de uma estrutura sobreposta aos de outra estrutura sobreposta com base na distância mais curta entre os carbonos α das espinhas dorsais do polipeptídeo, por exemplo, usando o pacote Schrödinger).

Em outra modalidade, as seqüências da série de referência são alinhadas usando tanto algoritmos de alinhamento à base de seqüência quanto à base de estrutura. Preferencialmente, as estruturas correspondentes às seqüências da série de referência são primeiro alinhadas usando um algoritmo de alinhamento à base de estrutura ou outras ferramentas de de alinhamento orientadas por estrutura, seguidas por alinhamento das seqüências correspondentes com um algoritmo à base de seqüência. Mais preferencialmente, as estruturas correspondentes às seqüências da série de referência são primeiro alinhadas, seguido por alinhamento das seqüências usando um alinhamento de seqüência à base de estrutura.

Em algumas modalidades opcionais, a série alinhada de seqüências é adicionalmente curada. Por exemplo, a série alinhada de seqüências pode ser expandida (por exemplo, por classificação não-intervalo ou por classificação de Henikoff) para aumentar sua diversidade (por exemplo, para obter séries muito grandes de seqüências homólogas tendo a mesma dobra). Em outras modalidades, a série alinhada de seqüências pode ser adicionalmente submetida a várias rodadas de curação e, opcionalmente, alinhamento adicional (por exemplo, alinhamento à base de estrutura).

Em algumas modalidades preferenciais, a invenção proporciona métodos para curar uma série alinhada contendo um número relativamente pequeno de seqüências de modo a reforçar sua diversidade. Aumentando o

número de seqüência no alinhamento, a robustez da análise de covariação é aumentada. Em uma modalidade, os métodos podem empregar um modelo ou perfil derivado heuristicamente para pesquisar seqüências homólogas adicionais de um banco de dados de seqüências grande disponível publicamente (por exemplo, o banco de dados NR mantido pela NCBI). Os modelos derivados heuristicamente podem ser produzidos usando um ou mais algoritmos à base de regressão selecionados entre, por exemplo, uma regressão de quadrados mínimos parcial, uma regressão linear múltipla, uma regressão de quadrados mínimos inversa, uma regressão de componente principal, uma importância variável para projeção, ou similares. Em outras modalidades, o modelo derivado heuristicamente é produzido usando um ou mais algoritmos à base de padrão selecionados entre, por exemplo, um modelo de Markov oculto, um algoritmo de Smith Waterman, rede neural, uma árvore de classificação e regressão, um spline de regressão adaptiva multivariada, ou similares. Em uma modalidade exemplar, o algoritmo de reconhecimento de padrão é um Modelo de Markov Oculto. Um HMM é um modelo estatístico onde o sistema sendo modelado contém parâmetros ocultos os quais podem ser extraídos usando parâmetros observáveis e empregado para análise de reconhecimento de padrão. Por exemplo, HMMs reconhecidos na técnica podem empregar estruturas secundárias previstas para pesquisar anticorpos da mesma dobra que um dado anticorpo, ao invés de anticorpos tendo relação puramente baseada em seqüência (vide por exemplo, *Proteins* 36(1), 68-76). MetaFam (Silverstein et al. *Nucleic Acids Res* 29(1), 49-51 (2001)), Interpro (Apweiler, *et al.*, *Nucleic Acids Res*, 29(1), 37-40, (2001)), e HMMER (Bateman *et al, Nucleic Acids Res* 27(1), 260-2 (1999)) são ferramentas típicas de reconhecimento de padrão à base de HMM.

Em algumas modalidades adicionais, seqüências extraídas com um algoritmo de HMM podem ser validadas para sua atribuição de classe estrutural adequada usando software de bioinformática reconhecido na técnica. Uma ferramenta de validação típica é a ferramenta de classificação PFAM mantida pelo Wellcome Trust Sanger Institute e está disponível publicamente na Internet. Por exemplo, a ferramenta PFAM 'pfamverify' pode ser

aplicada a cada seqüência extraída para confirmar que foi classificada corretamente pelo HMM classe-específico criado a partir do alinhamento de estrutura à base de estrutura (Finn *et al., Nucleic Acids Res*, 24 (Publicação de banco de dados), D247-51, (2006). Perfis de HMM correspondentes a um grupo de PFAM (ou família de proteínas relacionadas, por exemplo, um grupo de dobra de Ig) podem ser baixados do website da PFAM para facilitar classificação das seqüências extraídas. Seqüências extraídas cujos escores se situaram abaixo de um corte recomendado são preferencialmente removidas da série de referência.

Em outras modalidades, seqüências extraídas com o algoritmo HMM também podem ser alinhadas usando o algoritmo HMMC. Como estes algoritmos HMM se baseiam em cuidadosos alinhamentos estruturais, este processo assegura o alinhamento orientado por estrutura das seqüências extraídas adicionais. Por exemplo, o pacote HMMER pode ser utilizado para gerar alinhamentos 'mapali' em formato de saíde FASTA.

Em modalidades preferenciais, o método da invenção proporciona um novo HMM o qual não somente encontra seqüências similares às da série alinhada, mas alinha as novas seqüências para assegurar o alinhamento correto da região de núcleo. Por exemplo, o novo algoritmo de HMM da invenção pode pesquisar domínios separados (por exemplo, domínios de dobra de Ig) dentro de uma única proteína (por exemplo, um membro da superfamília das imunoglobulina), extrair cada domínio separadamente, e adicionar este à série alinhada de seqüências. Algoritmos ou perfis de HMM podem ser construídos a partir de alinhamentos de seqüência à base de estrutura (por exemplo, usando o pacote de software HMMER o qual está disponível publicamente na Internet, por exemplo, em www.psc.edu). Para cada pesquisa de HMM estrutura-específica, seqüências alcançadas acima de um limiar de escore de critério são preferencialmente retidas como membros candidatos da classe estrutural cujo HMM foi usado. Para as regiões de alcence que são subseqüências da seqüência alcançada, o alcence de subseqüência exato pode ser extraído da seqüência inteira.

Em outras modalidades exemplares, a invenção proporciona

métodos para curar a série alinhada de seqüências de modo a reduzir sua redundância. Por exemplo, enquanto o grande número de seqüências em uma série alinhada estruturalmente (por exemplo, uma série alinhada de dobras de Ig) representa uma rica fonte de seqüências, algumas subfamílias (por exemplo, algumas subclasses de dobras de Ig) podem ser altamente super-representadas, enquanto outras são altamente sub-representadas. Super- ou sub- representação resulta em preferência significativa devido a íntima linhagem comum e limita a utilidade global de análises de covariação. Por conseguinte, em modalidades preferenciais, o alinhamento é submetido a etapas de filtração adicionais para reduzir a redundância do alinhamento. Em modalidades preferenciais, no mínimo uma seqüência com mais de 90% de identidade com outra seqüência (por exemplo, 91%, 92%, 93%, 95%, 96%, 97%, 98%, 99% ou mais de identidade de seqüência) é removida do alinhamento. Em modalidades particularmente preferenciais, no mínimo uma seqüência partilhando mais de 80% de identidade de seqüência com outra seqüência (por exemplo, 81%, 82%, 83%, 84%, 87%, 88%, 89%, ou 90% de identidade de seqüência) é removida do alinhamento.

Em modalidades exemplares, a invenção proporciona um novo método para reduzir a redundância dentro do alinhamento. Os métodos referidos compreendem uso de um algoritmo heurístico para desconbrir e classificar seqüências de corte de identidade de seqüência desejado. Os métodos referidos compreendem uma ou mais das seguintes etapas seqüenciais: (1) calcular percentagem de identidades para todos os pares de seqüências no alinhamento; (2) agrupar valores de identidade em um ou mais grupos de percentagem de identidade; (3) classificar as seqüências de cada grupo por contagem decrescente de resíduo sem intervalos e/ou por peso de seqüência de Henikoff; e (4) remover seqüências redundantes de acordo com um critério de corte (por exemplo, um nível de corte de % de identidade).

A remoção das seqüências redundantes pode ser realizada em uma série de modos. Em uma modalidade, seqüências classificadas são agrupadas em múltiplos grupos de identidade de seqüência (por exemplo, um grupo de 99%, um grupo de 98%, um grupo de 97%, e etc.). As seqüências

são então sistematicamente removidas por classificação (por exemplo, seqüências com classificação mais alta primeiro) do maior grupo de % de identidade (isto é, o grupo de 99%, seguido pelo grupo de 98%, seguido pelo grupo de 97%, e etc.). As etapas de cálculo de identidade, agrupamento e/ou classificação (Etapas (1) a (3)) são opcionalmente repetidas depois de cada etapa de remoção até o grupo final que satisfaz o critério de corte ser eliminado. Em outra modalidade, um único grupo de seqüências classificadas é criado no critério de corte e seqüências são sistematicamente removidas por classificação do grupo. As etapas de cálculo de identidade e/ou classificação (Etapa (1) e (3)) são opcionalmente repetidas depois de cada etapa de remoção.

Em modalidades preferenciais, o critério de corte é 90% ou maior (por exemplo, 91%, 92%, 93%, 95%, 96%, 97%, 98%, 99% ou mais de identidade de seqüência). Em modalidades particularmente preferenciais, o critério de corte é 80% ou maior (por exemplo, 81%, 82%, 83%, 84%, 87%, 88%, 89%, ou 90% de identidade de seqüência).

Em outra etapa opcional, as extensões de seqüências na série de referência alinhada podem ser truncadas. Em uma modalidade, regiões com intervalo dentro do alinhamento podem ser removidas para evitar cálculo sobre estas regiões menos informativas. Por exemplo, colunas que não são estados de combinação em um perfil de HMM usado para descobrir as seqüências podem ser removidas. Em outra modalidade, seqüências dentro do alinhamento as quais sobrepõem a extensão de consenso do alinhamento são cortadas para remover a porção de projeção da seqüência. Em outra modalidade, seqüências sobrepostas são removidas do alinhamento.

Etapa 3: Análise de Covariação da Série Alinhada

Uma vez que as seqüências da série alinhada são compiladas, análise de covariação é realizada sobre um ou mais pares de resíduos aminoácidos na série de alinhamento. A análise de covariação resulta na geração de uma nova série de dados (aqui, neste Pedido de patente, uma "série de dados de covariação") descrevendo a importância estatística de resíduos correlacionados dentro dos alinhamentos. Em algumas modalidades, a série

de dados de covariação lista correlações entre cada possível par de resíduos dentro do alinhamento.

Em modalidades preferenciais, computação de covariação é um processo implementado por computador (por exemplo, um processo executável em Java). A computação de covariação pode resultar em vários parâmetros estatísticos reconhecidos na técnica. Em uma modalidade, a importância estatística (por exemplo, um valor de χ^2) da covariação é calculada usando uma análise Qui-quadrada. Em outra modalidade, a força estatística (por exemplo, valor de ϕ) da covariação é calculada. Por exemplo, um valor de ϕnegativo pode ser preditivo de uma correlação negativa. Isto é, a presença de um aminoácido em uma primeira posição no alinhamento favorece a ausência de outro aminoácido específico em uma segunda posição. Ao invés, um valor de ϕpositivo pode ser preditivo de uma correlação positiva, deste modo indicando que o aminoácido na primeira posição favorece a presença do aminoácido na segunda posição.

Em uma modalidade, um valor de ϕ é calculado usando a fórmula I:

$$\phi(a_i b_j) = \frac{(a_i b_j * \bar{a}_i \bar{b}_j) - (a_i \bar{b}_j * \bar{a}_i b_j)}{\sqrt{(a_i b_j + \bar{a}_i b_j) * (a_i \bar{b}_j + \bar{a}_i \bar{b}_j) * (a_i b_j + a_i \bar{b}_j) * (\bar{a}_i b_j + \bar{a}_i \bar{b}_j)}} \quad \text{(I)}$$

em que

$a_i b_j$ é o número de vezes que resíduos do tipo "a" ou "b" são encontrados na mesma seqüência nas posições i e j, respectivamente;

$\bar{a}_i b_j$ é o número de vezes que ambos os tipos de resíduos estão ausentes da mesma seqüência;

$a_i \bar{b}_j$ é o número de vezes que a é encontrado presente enquanto b está ausente; e

$\bar{a}_i b_j$ é o número de vezes que a está ausente enquanto b está presente.

Em uma modalidade, um valor de ϕ é calculado usando a seguinte fórmula II:

$$\phi(a_i b_j) = \frac{(c * f) - (d * e)}{\sqrt{ghij}} \quad \text{(II)}$$

onde c a j são definidos pela matriz:

	a_i	$\bar{a}_i$	Total
b_j	c	d	g
$\bar{b}_j$	e	f	h
Total	i	j	

e

$c = a_i b_j$, $d = \bar{a}_i b_j$, $e = a_i \bar{b}_j$, $f = \bar{a}_i \bar{b}_j$, $g = c + d$, $h = e + f$, $i = c + e$, e $j = d + f$.

Em algumas modalidades, um valor de χ^2 pode ser calculado usando uma fórmula à base de "freqüência-de-ocorrência". Em outras modalidades, um valor de χ^2 é calculado usando uma fórmula "à base de evento" (isto é, número de ocorrências). Uma fórmula à base de evento exemplar é determinada como Fórmula III:

$$\chi^2 = \frac{[c(i,j) - (p(i) \cdot p(j) \cdot c(t))]^2}{p(i) \cdot p(j) \cdot c(t)} \quad \text{(III)}$$

em que

p(i) e p(j) são as freqüências de resíduos de quaisquer dois tipos de resíduos de interesse nas posições i e j, respectivamente, na série de seqüências alinhadas;

c(i,j) é o número de vezes p(i) e p(j) são observadas na mesma seqüência; e

c(t) é o número total de seqüências no alinhamento;

e em que freqüências de resíduos são definidas como o número de vezes um tipo de resíduo é observado em uma posição específico em um alinhamento dividido pelo número total de seqüências no alinhamento.

Em algumas modalidades, somente um resíduo aminoácido natural é considerado um "tipo de resíduo" para os fins de um cálculo de covariação. Em modalidades preferenciais, no entanto, um tipo de resíduo também pode incluir um intervalo como um tipo de resíduo distinto, uma vez que intervalos (especialmente em porção de alça) freqüentemente ajudam a discriminar um motivo de outro.

Em algumas modalidades, o cálculo de covariação pode incluir

uma função de ponderação de diversidade (por exemplo, um esquema de ponderação de Henikoff).

Em outras modalidades, Identidades Médias de Seqüências (SAIs) podem ser usadas para filtrar pares covariantes. Por exemplo, SAIs podem ser usadas para filtrar covariação de pares com mais de identidade média de seqüência de modo a remover covariações artefatuais potenciais originárias de seqüências intimamente relacionadas. Em modalidades onde o alinhamento foi submetido a curação por um alto limiar de identidade (por exemplo, 80% de identidade como critério de corte), cálculos de SAI são preferencialmente prescindidos.

Em algumas modalidades, pares covariantes não são reportados pelo cálculo a menos que sejam observados um número mínimo de vezes (aqui, neste Pedido de patente, um "corte de evento"). Em uma modalidade, o corte de evento é 10 ou mais eventos. Em uma modalidade mais preferencial, o corte de evento cerca de 2 ou mais, e menos de 10, eventos (por exemplo, 9, 8, 7, 6, ou 5 eventos).

Etapa 4: Usar Dados de Covariação para Prever Estabilidade Protéica

Em algumas modalidades, as covariações estatisticamente significantes dentro da série de dados de covariação são usadas para perquisar covariações correspondentes dentro da seqüência candidata ou de ensaio. Em particular, a conservação de um ou mais aminoácidos covariantes da série de referência em posições de aminoácidos correspondentes dentro da seqüência candidata ou de ensaio indica que estes resíduos são funcionalmente importantes e são preditivos de estabilidade de proteína favorável. Por conseguinte, o número de aminoácidos covariantes que são conservados dentro de uma seqüência candidata ou de ensaio pode ser usado para prever a estabilidade da seqüência.

Em alguns aspectos típicos, as covariações relevantes podem ser visualizadas usando uma interface gráfica do usuário (GUI) da invenção (por exemplo, a ferramenta NAPMAP descrita na Seção V abaixo) para análise da série de dados de covariação. Como a interpretação de dados na série de dados de covariação pode ser trabalhosa, a interface gráfica do usuá-

rio pode facilitar a análise de dados e facilitar rápidas análises de designs e funcionais de proteína.

Em modalidades exemplares, a presença ou ausência de resíduos covariantes dentro da série de referência ou alinhamento que estão presentes ou ausentes dentro da seqüência candidata pode ser usada para estabelecer um escore o qual é correlativo com a estabilidade da proteína. Por exemplo, onde um par covariante (por exemplo, positivamente covariante) de resíduos é considerado como sendo retido dentro da seqüência candidata, um primeiro escore (por exemplo, um escore positivo) pode ser atribuído à seqüência candidata. Ao invés, onde um par covariante (por exemplo, positivamente covariante) de resíduos está faltando na seqüência candidata, um segundo escore de sinal oposto ao primeiro escore (por exemplo, um escore negativo) pode ser atribuído à seqüência candidata. Em uma modalidade preferencial, à ausência de um par de resíduos negativamente covariante dentro de uma seqüência de ensaio é atribuído um escore positivo, ao passo que à presença do par de resíduos negativamente covariante é atribuído um escore negativo. Em outra modalidade preferencial, à presença de um par de resíduos positivamente covariante de uma seqüência de ensaio é atribuído um escore positivo, ao passo que à ausência de alguns resíduos da seqüência de ensaio de par positivamente covariante é atribuído um escore negativo.

Em algumas modalidades, escores de covariação são gerados somente para covariações que satisfazem um nível de limiar de importância estatística. Em uma modalidade, escores de covariação são gerados somente para covariações acima (ou abaixo) de um determinado valor de χ^2-valor ou valor de ϕ. Por exemplo, o corte para designação de escore de covariação negativo pode ser uma covariação com um coeficiente de associação phi (Φ) de menos de +0,2 (cerca de +0,2 a cerca de -1,0), e mais preferencial de menos de -0,2. Em outra modalidade exemplar, a uma covariação é atribuído um escore de covariação positivo se tiver um coeficiente de associação phi de mais de -0,2, preferencialmente mais de +0,2. Em modalidades preferenciais, um escore de covariação negativo é atribuído a uma cova-

riação com um coeficiente de associação phi (Φ) de menos de -0,5 (por exemplo, -0,5, -0,6, -0,7, -0,8, -0,9, ou -1,0), e um escore de covariação positivo é atribuído a uma covariação com um coeficiente de associação phi (Φ) de mais de +0,5 (por exemplo, +0,5, +0,6, +0,7, +0,8, +0,9, ou +1,0), ao passo que todas as outras covariações são mascaradas atribuindo um escore neutro (por exemplo, zero).

Em algumas modalidades, os escore de covariação negativos ou positivos podem ser ponderados para refletir a importância estatística ou força da covariação correspondente. Em uma modalidade típica, a uma covariação positiva pode ser atribuído um alto escore de covariação positivo se seu coeficiente de associação phi for também elevado ou um escore de covariação positivo menor se seu valor de ϕ for baixo. Por exemplo, a uma seqüência candidata tendo primeira e segunda covariações com valores de ϕ respecitivos de +0,5 e +1,0 pode ser atribuída um baixo escore de covariação positivo (por exemplo, +1) para a primeira covariação e um escore de covariação positivo maior para a segunda covariação (por exemplo, +2).

Em outras modalidades, o escore de covariação é atribuído somente às covariações as quais são validadas por um meio diferente de significância estatística calculada. Vários critérios não estatísticos podem ser usados para validar que as covariações calculadas são importantes e biologicamente significativas. Em uma modalidade, a covariação é validada examinando se há correlação ou tendência entre resíduos que covariam uns com os outros e sua promixidade uns com os outros dentro de um modelo estrutural da proteína correspondente. Se os resíduos referidos estiverem em íntima proximidade (por exemplo, 30 Å ou menos, preferencialmente 10 Å ou menos) uns com os outros na estrutura, a covariação prevista é validada como uma covariação verdadeira. Em outra modalidade, a covariação é validada determinando se os aminoácidos covariáveis formam uma conexão a qual já se sabe que existe dentro de um subgrupo de proteínas correspondentes a seqüências da série de referência (por exemplo, uma ligação dissulfeto conhecida). Por exemplo, resíduos 6-10 no N-término de sobras de cadeia pesada variável de IgG humana ou murina são conhecidos por adotar

conformações muito específicas com base na conservação de pares covariantes de aminoácidos (Ewert *et al., Methods,* 34: 184-199 (2004)).

Em ainda outras modalidades, um escore de covariação de um aminoácido em uma seqüência candidata pode ser ponderado para refletir o número de covariações dentro da série de referência ou alinhamento que satisfaz (ou viola). Em uma modalidade, à seqüência candidata é atribuído um escore de covariação positivo para cada covariação dentro da série de referência ou alinhamento que satisfaz, e um escore de covariação negativo para cada covariação dentro da série de referência que viola. Em modalidades preferenciais, o escore de covariação para um aminoácido particular de uma seqüência de ensaio é a soma total do escores de covariação positivo dividido pela soma total dos escores de covariação negativos. Por exemplo, onde 2 covariações negativas e 9 covariações positivas são observadas dentro da série de referência em posições de aminoácido correspondentes ao aminoácido da seqüência candidata, ao aminoácido da seqüência candidata pode ser atribuído um escore de covariação total de 7.

Em algumas modalidades, alguns escores de covariação total são somados para todos os aminoácidos em uma seqüência candidata (ou porção do mesmo) para obter a "seqüência escore de covariação". Este escore de covariação de seqüência pode ser empregado para prever a estabilidade da seqüência. De modo geral, um escore de covariação de seqüência negativo é preditivo de baixa estabilidade protéica, ao passo que um escore de covariação de seqüência positivo é preditivo de alta estabilidade protéica.

b. Classificação de Consenso

Em outras modalidades, os métodos computacionais da invenção empregam uma nova técnica de classificação denominada "classificação à base de consenso" a qual compara seqüências de polipeptídeos candidatas contra um banco de dados de seqüências de polipeptídeos relacionadas para identificar proteínas com estabilidade potencialmente baixa. Conforme usado aqui, neste Pedido de patente, "classificação à base de consenso" se refere a cálculo do número de aminoácidos não-consenso dentro de uma proteína (por exemplo, uma proteína de ensaio) baseado na informação dis-

ponível a partir de compilações de seqüências de proteínas relacionadas. O cálculo resulta em uma "classificação de consenso" para a proteína. Conforme mostrado nos Exemplos abaixo, a classificação de consenso de uma proteína é altamente correlativa com sua estabilidade protéica determinada empiricamente. Por conseguinte, um objeto da invenção é empregar classificação à base de consenso para prever estabilidade protéica. Quanto maior o desvio de uma seqüência de proteína de ensaio de uma seqüência de consenso, maior a classificação de consenso, e mais provavelmente a proteína de ensaio deve conter aminoácidos prejudiciais para sua estabilidade.

Na nova abordagem de classificação de consenso descrita aqui, neste Pedido de patente, as seguintes etapas básicas são empreendidas para prever a estabilidade de uma proteína candidata ou de ensaio:

1. Proporcionar uma série de seqüências homólogas (aqui, neste Pedido de patente, uma "série de referência") correspondentes a uma seqüência de um domínio (ou porção do mesmo) da proteína candidata (aqui, neste Pedido de patente, uma "seqüência de domínio de ensaio").

2. Determinar a freqüência de resíduos em cada posição de resíduo dentro da seqüência de domínio de ensaio (ou porção do mesmo) para obter uma classificação de consenso.

3. Usar a classificação de consenso para prever a estabilidade da proteína candidata ou de ensaio.

Em algumas modalidades exemplares, as seguintes etapas opcionais são empregadas para aumentar o valor preditivo da classificação de consenso:

4. Determinar a freqüência de resíduos em cada posição de resíduo correspondente dentro de cada seqüência da série de referência (ou um subgrupo da mesma) para obter um escore médio.

5. Comparar a classificação de consenso com a classificação de consenso média para determinar um escore de seqüência.

6. Usar o escore de seqüência para prever a estabilidade da proteína candidata ou de ensaio.

Etapa 1: Proporcionar uma Série de Referência

Os métodos à base de classificação de consenso da invenção empregam uma série de seqüências (uma "série de referência") as quais são homólogas a (isto é, relacionadas com) uma seqüência de interesse ou "seqüência de ensaio". A série de referência pode compreender uma série de seqüências homólogas tendo graus de similaridade moderado a elevado com a seqüência de ensaio.

A série de seqüências homólogas pode ser determinada por uma pessoa versada na técnica para incluir um número adequado de seqüências não-redundantes. A série de seqüências homólogas não precisa ser grande, contanto que seja obtida uma seleção razoavelmente não tendenciosa. Não é o número de seqüências, mas a proporção de freqüêcias que é importante.

As seqüências referidas podem ser obtidas através de pesquisas à base de homologia de quaisquer dos bancos de dados de seqüências conhecidos na técnica (por exemplo, bancos de dados NCBI e TIGR). Em uma modalidade exemplar, uma pesquisa BLAST de rotina com parâmetros padrão pode ser realizada com a seqüência de ensaio para recuperar seqüências homólogas de interesse. Seqüências com uma percentagem mínima de identidade de seqüência (por exemplo, mais de 25%, 30%, 35%, 40%, 45% ou 50% de identidade de seqüência, preferencialmente mais de 30% de identidade de seqüência) podem ser selecionadas para inclusão na série de referência. Em algumas modalidades, as seqüências com um alto grau de identidade de seqüência (por exemplo, mais de 85%, 90%, 95%, ou 98% de identidade de seqüência, preferencialmente mais de 95% de identidade de seqüência) são excluídas da série de referência para evitar preferência.

Em algumas modalidades, a série de referência pode ser curada para incluir somente as seqüências as quais satisfazem um ou mais critérios para inclusão na série. Por exemplo, em algumas modalidades, a série de referência pode compreender seqüências ortólogas (por exemplo, seqüências as quais são de diferentes espécies (por exemplo, espécies de mamífe-

ros diferentes) mas as quais têm a mesma função bioquímica). Em outras modalidades, a série de referência compreende seqüências humanas. Em uma modalidade exemplar, a série de seqüências homólogas inclui somente seqüências de linha germinal mamífera. Em modalidades preferenciais, a série de referência inclui seqüências de proteínas tendo a mesma classe de dobra de proteína (por exemplo, um domínio de proteína similar) que a proteína de ensaio (por exemplo, proteínas tendo uma dobra do tipo de imunoglobulina). As proteínas referidas podem ser identificadas por meio de técnicas de bioinformática conhecidas na técnica (por exemplo, pesquisando o banco de dados de domínios conservados (CDD) do Instituto Nacional de Saúde dos Estados Unidos.

Etapa 2: Determinar o Escore de Consenso da Seqüência de Ensaio

Uma vez que as seqüências da série de referência são compiladas, a seqüência de consenso da série de referência é determinada de modo a facilitar o método de classificação de consenso da invenção. Conforme usado aqui, neste Pedido de patente, o termo "seqüência de consenso" se refere a uma seqüência em que o resíduo (por exemplo, resíduo aminoácido) em cada posição dentro da seqüência corresponde ao resíduo mais comum nesta posição em uma série alinhada de seqüências relacionadas (isto é, a série de referência).

Para determinar a seqüência de consenso, as seqüências dentro da série de referência podem primeiro ser alinhadas usando algoritmos de alinhamento de seqüência conhecidos na técnica. Um exemplo preferencial e não-limitante de um algoritmo de alinhamento local utilizado para a comparação de seqüências é o algoritmo de Karlin e Altschul (1990) *Proc. Natl. Acad. Sci. USA* 87:2264-68, modificado como em Karlin e Altschul (1993) *Proc. Natl. Acad. Sci. USA* 90:5858-77. Um algoritmo similar é incorporado nos programas BLAST (versão 2.0) de Altschul, *et al.* (1990) *J. Mol. Biol.* 215:403-10. Em outra modalidade, o alinhamento é otimizado introduzindo intervalos apropriados e a percentagem de identidade é determinada sobre a extensão das seqüências alinhadas (isto é, um alinhamento com intervalos). Para obter alinhamentos com intervalos para fins de compara-

ção, Gapped BLAST pode ser utilizado conforme descrito em Altschul *et al.*, (1997) *Nucleic Acids Res.* 25(17):3389-3402. Em outra modalidade, o alinhamento é otimizado introduzindo intervalos apropriados e a percentagem de identidade é determinada sobre toda a extensão das seqüências alinhadas (isto é, um alinhamento global). Um exemplo preferencial e não-limitante de um algoritmo matemático utilizado para a comparação de seqüências global é o algoritmo de Myers e Miller, *CABIOS* (1989). Um algoritmo similar é incorporado no programa ALIGN (versão 2.0) o qual é parte do pacote do software de alinhamento de seqüência GCG. Ao utilizar o programa ALIGN para comparar seqüências de aminoácidos, uma tabela de resíduo de peso PAM120, uma penalidade de extensão de intervalo de 12, e pode ser usada uma penalidade de intervalo de 4.

Para determinar a seqüência de consenso da série alinhada de referências, é determinado o resíduo que ocorre mais freqüentemente em cada posição dentro da série alinhada de referências. A seqüência de consenso pode ser determinada visualmente ou pode ser determinada por análise computacional usando um programa de bioinformática disponível publicamente (por exemplo, a ferramenta EMBOSS CONS disponível no Helix Systems do National Institute of Health).

A seqüência de consenso é empregada nos métodos da invenção para determinar uma classificação de consenso. Em algumas modalidades, uma classificação de consenso é um valor numérico o qual equivale à freqüência ou ocorrência de resíduo dentro da série de referência do resíduo de consenso em cada posição de aminoácido na seqüência de consenso (isto é, a "freqüência de resíduo de consenso") com a freqüência ou ocorrência de resíduo dentro da série de referência do aminoácido na posição correspondente na seqüência de ensaio (isto é, a "freqüência de resíduo de ensaio"). Freqüências de resíduo em uma posição dentro da seqüência de ensaio ou de consenso são calculadas somando o número de vezes que o aminoácido nesta posição está presente na posição correspondente na série de referência alinhada, e dividindo o valor somado pelo número total de seqüências dentro da série de referência. Em uma modalidade exemplar, a

classificação de consenso em cada posição de aminoácido é calculada dividindo a freqüência de resíduo de ensaio pela freqüência de resíduo de consenso para dar um escore não-zero de entre 1 e 0. Por exemplo, a um resíduo o qual é idêntico ao aminoácido de consenso naquela posição é atribuído um escore perfeito de 1, ao passo que uma classificação de consenso se aproxima de zero, o menos comum o aminoácido é naquela posição.

Em algumas modalidades preferenciais, a classificação de consenso pode ser somada para cada posição da seqüência de ensaio (ou uma porção da mesma) para obter uma classificação de consenso total. Por exemplo, uma classificação de consenso total perfeita é o número de aminoácidos na seqüência de ensaio, indicando que a seqüência de ensaio é idêntica a sua seqüência de consenso correspondente.

Uma fórmula típica para calcular a classificação de consenso total de uma proteína é determinada abaixo:

$$\text{escore} = \sum_i \frac{h_i(r)}{c_i(r)}$$

em que

$c_i(r)$ equivale à freqüência de resíduo de consenso em uma posição de aminoácido da seqüência de consenso,

$h_i(r)$ equivale à freqüência de resíduo de ensaio de uma posição de aminoácido da seqüência de ensaio, e

i equivale ao número de posições de aminoácido dentro da seqüência de ensaio.

Etapa 3: Usar o Escore de Consenso para Prever a Estabilidade Protéica

A classificação de consenso pode ser então empregada para prever a estabilidade da proteína de ensaio. Em geral, quanto maior o valror da classificação de consenso, maior a probabilidade de que a proteína de ensaio terá estabilidade adequada para seu uso pretendido. Em algumas modalidades, a classificação de consenso da proteína de ensaio pode ser comparada com a classificação de consenso de uma seqüência controle. Em uma modalidade típica, a seqüência controle é de uma proteína a qual é conhecida por ter propriedades de estabilidade pobres ou indesejadas (aqui,

neste Pedido de patente, um "controle negativo"). Por conseguinte, é previsto que a proteína de ensaio tenha estabilidade aprimorada se sua classificação de consenso for maior do que a do controle negativo. Em outra modalidade exemplar, a seqüência controle é de uma proteína a qual é conhecida por ter propriedades de estabilidade desejáveis (aqui, neste Pedido de patente, um "controle positivo"). Por conseguinte, é previsto que a proteína de ensaio tenha estabilidade aprimorada se sua classificação de consenso for substancialmente similar ou maior do que a do controle positivo. Em outras modalidades, a classificação de consenso da seqüência de ensaio pode ser comparada com sua classificação de consenso perfeita hipotética, um valor o qual corresponde ao número total de aminoácidos na seqüência de ensaio.

Etapa 4: Determinar o Escore de Consenso Médio da Série de Referência

Em algumas modalidades, classificações de consenso médio em parte ou todas as posições de aminoácido das seqüências dentro da série de referência são determinadas para comparação com as classificações de consenso em posições correspondentes dentro da seqüência de ensaio. Em modalidades preferenciais, a classificação de consenso total médio para a série de referência é determinada para comparação com a classificação de consenso total da seqüência de ensaio. A classificação de consenso total média para a série de referência é a soma das freqüências de resíduos para todas as posições de aminoácidos de uma seqüência dentro da série de referência, dividida pelo número total de seqüências nesta série de referência.

Etapa 5: Determinar o Escore da Seqüência

Em algumas modalidades, o escore de seqüência é determinado comparando a classificação de consenso médio da série de referência com a classificação de consenso da seqüência de ensaio. O diferencial destes valores é o "escore de seqüência" da seqüência de ensaio (também referido aqui, neste Pedido de patente, como o "escore Δ"). O escore de seqüência pode ser determinado subtraindo a classificação de consenso da classificação de consenso médio.

Etapa 6: Usar o Escore de Seqüência para Prever a Estabilidade Protéica

O escore de seqüência proporciona uma medida do desvio da

classificação de consenso da seqüência de ensaio da classificação de consenso médio da série de referência. Os Exemplos abaixo demonstram que o escore de seqüência de uma proteína é altamente correlacionado com a estabilidade (por exemplo, estabilidade térmica) da proteína. Por conseguinte, em algumas modalidades, o escore de seqüência também pode ser usado para prever a estabilidade da proteína. Se o escore Δ da proteína for um valor negativo (isto é, a classificação de consenso for menor do que a classificação de consenso médio), a previsão seria que a estabilidade da proteína de interesse terá uma estabilidade menor do que a maioria das proteínas dentro da série de referência.

Em algumas modalidades, as seqüências de ensaio empregadas nos métodos da invenção são proteínas de domínio único. Conforme usado aqui, neste Pedido de patente, uma "proteína de domínio único" é uma proteína compreendendo um ou mais domínios, em que cada um dos domínios são do mesmo tipo (por exemplo, um domínio pesado variável (VH)).

Em outras modalidades, as seqüências de ensaio empregadas nos métodos da invenção são proteínas multidomínio. Conforme usado aqui, neste Pedido de patente, uma "proteína multidomínio" é uma proteína compreendendo dois ou mais domínios, onde no mínimo dois dos domínios são de um tipo diferente. Por exemplo, anticorpos são proteínas compreendendo tipicamente seis domínios (isto é, VH, VL, CL, CH1, CH2, e CH3). Em modalidades preferenciais, a estabilidade do domínio menos estável de uma proteína multidomínio é prevista usando os métodos da invenção. Os Exemplos abaixo demonstram que os métodos da invenção são particularmente eficazes para prever a estabilidade térmica quando o domínio menos estável (ou porção do mesmo) de uma proteína multidomínio (por exemplo, o domínio VH de um anticorpo) é empregado como uma seqüência de ensaio.

Se o domínio menos estável não for conhecido *a priori*, pode ser determinado usando métodos experimentais que são conhecidos na técnica (por exemplo, usando quaisquer dos métodos descritos acima, tais como calorimetria de varredura diferencial (DSC), dicroismo circular dependente

da temperatura (CD), ou medições da fluorescência). Os métodos referidos permitem a determinação de múltiplas transições de desdobramento térmico onde o domínio menos estável ou desdobra primeiro (vide a Figura 69A) ou limita o limiar de estabilidade global de uma unidade multidomínio que desdobra cooperativamente (isto é uma proteína multidomínio a qual apresenta uma transição de desdobramento único). O domínio menos estável pode ser identificado em uma série de modos adicionais. Mutagênese pode ser realizada para sondar qual domínio limita a estabilidade global. Adicionalmente, resistência a protease de uma proteína multidomínio pode ser realizada sob condições onde o domínio menos estável é conhecido por ser intrinsicamente desdobrado através de DSC ou outros métodos espectroscópicos (Fontana, *et al.*, *Fold. Des.*, 2: R17-26, 1997). Uma vez que o domínio menos estável seja identificado, a seqüência codificando este domínio (ou uma porção do mesmo) pode ser empregada como uma seqüência de ensaio nos métodos da invenção.

Em algumas modalidades exemplares, as proteínas multidomínio avaliadas usando os métodos da invenção são anticorpos. Os métodos da invenção possibilitam que uma pessoa versada na técnica exclua a seleção de domínios de região variável inadequados de entre o vasto conjunto de seqüências de região variável criadas pelos mecanismos de diversidade do sistema imune. Em modalidades exemplares, por exemplo, moléculas scFv ou moléculas de anticorpos compreendendo domínios de região variável (Fv) com uma estabilidade relativamente elevada (por exemplo, estabilidade térmica) podem ser identificadas usando os métodos da invenção, e selecionadas. Em outras modalidades exemplares, domínios de região variável (Fv) de imunoglobulina humana tendo estabilidade aceitável podem ser identificados usando os métodos da invenção e selecionados para uso como cadeias de imunoglobulina aceitadoras na humanização de um anticorpo. Em outras modalidades exemplares, os métodos da invenção podem ser empregados para triar seqüências variáveis de anticorpos engenheirados candidatas (por exemplo, seqüências VL ou VH humanizadas) para estabilidade apropriada antes de prosseguir com a síntese de uma molécula de li-

gação compreendendo a seqüência candidata.

Conforme é de conhecimento na técnica, a estabilidade protéica varia de um anticorpo para outro com base nas muitgas variações possíveis em anticorpos naturais ou engenheirados. Por exemplo, a pressão seletiva do sistema imune para criar diversidade dentro dos domínios variáveis de anticorpo natural pode afetar negativamente sua estabilidade ou dobrabilidade em sistemas de expressão recombinante (Knappik e Pluckthun, *Protein Engng.*, 8: 81-89, (1995); Wall e Pluckthun, *Protein Engng,* 12: 605-11 (1999); Knappik *et al.*, *J. Mol. Biol.,* 296: 57-86, (2000); Ewert *et al.*, *J. Mol. Biol.,* 3325: 531-33, (2003)). Anticorpos consistindo no mesmo isótipo e subclasse, IgG1 humana sendo o mais comumente usado para terapêutica de anticorpo, variam em estabilidade de um para o outro essencialmente com base em diferenças em seus domínios variáveis uma vez que o restante de suas seqüências são idênticas. O desdobramento térmico de fragmentos de anticorpos Fab predominantemente ocorre em uma única transição aparente. Na maioria dos casos, a região Fv de anticorpo limita a termoestabilidade global do Fab exceto nos raros casos onde a estabilidade de Fv (ou único V_H ou V_L) é extremamente baixa ou extremamente elevada e a transição de desdobramento de Fv é desacoplada daquela da região C_H1/C_L (Shimba, et al., 1995; Röthlisberger et al., 2005).

Em algumas modalidades, os métodos da invenção empregam uma seqüência de região variável de um anticorpo como uma seqüência de ensaio. Em algumas modalidades preferenciais, os métodos da invenção empregam o domínio variável de cadeia pesada (VH) (ou uma porção do mesmo) como uma seqüência de ensaio. Os Exemplos abaixo demonstram que o domínio VH é altamente preditivo da estabilidade de um anticorpo compreendendo o referido domínio.

Em algumas modalidades, a seqüência de ensaio de domínio variável é comparada com uma série de referência compreendendo, ou consistindo somente de, seqüências de domínio variável. As referidas seqüências de domínio variável podem ser compiladas a partir do banco de dados Kabat de seqüências de imunoglobulina (Kabat *et al.*, "*Distribution Files of*

the Fifth Edition of Sequences of Proteins of Immunological Interest", 1992). Em algumas modalidades preferenciais, a série de referência pode compreender ou consistir em seqüências de imunoglobulina humana.

Em outras modalidades preferenciais, a série de referência pode compreender, ou consistir somente em, seqüências de linha germinal humana. Em uma modalidade, a referência compreende seqüências de domínio variável da mesma subclasse Kabat de seqüências de anticorpos.

Em outras modalidades, uma porção de domínio variável (por exemplo, domínio VH) é usada como uma seqüência de ensaio nos métodos da invenção. Uma porção de domínio variável típica é uma seqüência compreendendo menos do que todas as regiões de framework ou CDRs da seqüência de região variável (por exemplo, uma seqüência VH que é truncada para remover CDR3 e FR4). Em uma modalidade, as seqüências de domínios VH são da subclasse de linha germinal de Kabat V_H1. Em outra modalidade, as seqüências de domínio variável são da subclasse de linha germinal de Kabat V_H2. Em outra modalidade, as seqüências de domínio variável são da subclasse de linha germinal de Kabat V_H3. Em outra modalidade, as seqüências de domínio variável são da subclasse de linha germinal de Kabat V_H4. Em outra modalidade, as seqüências de domínio variável são da subclasse de linha germinal de Kabat V_H5. Em outra modalidade, as seqüências de domínio variável são da subclasse de linha germinal de Kabat V_H7.

Os métodos da invenção podem ser empregados com proteínas conhecidas na técnica. Em uma modalidade, a proteína é um anticorpo ou porção do mesmo. Em algumas modalidades, o anticorpo é um anticorpo humanizado. Em outras modalidades, o anticorpo é um anticorpo humano. Em outras modalidades, o anticorpo é um anticorpo não-humano (por exemplo, anticorpo monoclonal de camundongo). Em outras modalidades, o anticorpo é um anticorpo de domínio único (por exemplo, camelídeo, de tubarão, humano). Em ainda outras modalidades, o anticorpo é um anticorpo quimérico. Outros anticorpos modificados típicos para uso nos métodos da invenção incluem anticorpos com domínio eliminado e moléculas de ligação biespecíficas. Outros anticorpos típicos incluem os anticorpos contendo scFv ou anti-

corpos modificados descritos acima. Em modalidades preferenciais, é avaliada a estabilidade das proteínas de ligação mencionadas acima usando uma seqüência de cadeia pesada variável das mesmas (ou porção das mesmas) como a seqüência de ensaio nos métodos da invenção.

Em uma modalidade, da invenção, os dados obtidos para um dado polipeptídeo são armazenados ou fornecidos os resultados como uma medida da estabilidade do polipeptídeo. Em uma modalidade, os métodos descritos aqui, neste Pedido de patente, podem ser repetidos para uma pluralidade de polipeptídeos. Em uma modalidade, uma seleção de um polipeptídeo pode ser feita com base nos dados obtidos. Em outra modalidade, um polipeptídeo selecionado pode ser formulado para uso terapêutico.

IV. Métodos para Desenhar / Produzir Proteínas com Propriedades Biofísicas Reforçadas (por exemplo, Estabilidade Protéica)

(a) Design de Covariação

Em outro aspecto, a invenção proporciona métodos para criar variantes de proteínas com uma propriedade biofísica reforçada em relação à molécula parental da qual é derivada usando os resultados de uma análise de covariação. Há muitos tipos de esforços de engenharia de proteína onde dados de covariação podem ser usados para facilitar o projeto, inclusive mas não limitado a projetos de estabilidade protéica reforçada, modificação de perfis de desdobramento com pH, remoção estável de glicosilação, e alteração da função bioquímica.

Em algumas modalidades, a invenção proporciona métodos para criar proteínas com projetos de estabilidade reforçada com base na identificação de escores de covariação (ou as próprias covariações) que são preditivos de estabilidade protéica reforçada ou reduzida de acordo com o método da Seção III acima. Por exemplo, se uma seqüência pertencente à proteína de interesse está faltando ou viola várias covariações chave, podem ser feitos projetos de proteína para corrigir isto. Também podem ser incluídas modificações as quais aumentam o número calculado de fortes covariações observadas dentro de uma única seqüência.

Foi desenvolvida uma série de projetos para estabilizar os do-

mínios VH e V_L de BHA10 scFv. Em particular, resultados iniciais de triagem de biblioteca para dois domínios scFv BHA10 validam a capacidade da Ferramenta de Análise de Covariação (*Covariation Analysis Tool*) para projeto da estabilidade protéica. O primeiro exemplo envolve a previsão com sucesso de uma mutação estabilizante (S46L) dentro do domínio V_L BHA10 que confere um significativo aprimoramento na estabilidade térmica (□T_{50} de +10 °C) conforme medido por uma prova de ensaio térmico. Mutação de Ser para Leu na posição 46 (sistema de numeração de Kabat) leva a uma conexão positivo com um Tyr existente no resíduo 36 o qual a Ferramenta mostra que covaria fortemente com Leu 46 (Figura 78). Um segundo exemplo da utilidade da Ferramenta de Análise de covariação é a correspondênia analítica dos efeitos negativos de uma mutação de BHA10 V_H Q6E. Análise de freqüência de resíduos únicos sugeriu que mutação para o Glu muito mais comumente observado nesta posição levaria a um aumento em estabilidade. No entanto, a Ferramenta de Análise de Covariação indicou que mutação única para Glu viola várias covariações existentes presentes dentro da seqüência de BHA10 V_H. Para obter um aprimoramento na estabilidade deve-se colocar vários aminoácidos que preferencialmente estabilizam Glu nesta posição.

Outro exemplo do valor preditivo da Ferramenta de Análise de Covariação é mostrado na Figura 79. Met 80 foi mutado para Leu como parte de um projeto de biblioteca de resíduo único. Imaginava-se que esta única mutação seria altamente estabilizante, uma vez que Leu é o aminoácido mais freqüente observado nesta posição dentro do banco de dados de seqüências. No entanto, análises de covariação indicam que dois outros aminoácidos devem ser mutados (V67F e T70S) de modo a obter harmonia de covariação.

(b) Design da Interface

Em ainda outros aspectos, a invenção proporciona métodos para reforçar uma propriedade biofísica (por exemplo, estabilidade protéica) de uma segunda proteína ou proteína candidata (por exemplo, um anticorpo) usando uma primeira proteína como uma proteína modelo para projeto racional.

Conforme usado aqui, neste Pedido de patente, o termo "proteína modelo", se refere a uma proteína com uma propriedade biofísica desejável (por exemplo, alta estabilidade protéica térmica) a qual é usada para modelar a mesma propriedade em uma segunda proteína para um aprimoramento da propriedade biofísica também é desejável. Em modalidades preferenciais, a proteína modelo é da mesma classe estrutural que a segunda proteína (por exemplo, se a segunda proteína for um anticorpo, um anticorpo é empregado como uma proteína modelo).

Em algumas modalidades, a proteína modelo é identificada como tendo propriedade biofísica desejável (por exemplo, estabilidade protéica) de acordo com um método da invenção. Em uma modalidade típica, a proteína modelo com propriedades de estabilidade desejáveis pode ter sido identificada entre um grupo de proteínas (Tqal como BIIB1-4, descrito na Tabela 20, acima) usando métodos experimentais tais como os descritos no Exemplo 15. Esta proteína modelo serviria então como uma plataforma de seqüência para projeto. Em uma modalidade típica, a proteína modelo é prevista para ter uma propriedade biofísica aprimorada (por exemplo, estabilidade protéica) por um método de análise de covariação da invenção resumido acima. Em outras modalidades exemplares, a proteína modelo é prevista para ser uma proteína desejável (por exemplo, uma proteína estabilizada) de acordo com um método de classificação de consenso conforme resumido acima. Em ainda outra modalidade exemplar, a proteína modelo é identificada de acordo com um método de triagem da invenção.

Em outras modalidades, a proteína modelo é uma proteína criada de acordo com um método da invenção. Em uma modalidade típica, a proteína modelo é designada para ter uma propriedade biofísica aprimorada (por exemplo, aumentada estabilidade protéica) de acordo com um método de covariação da invenção. Em outra modalidade exemplar, a proteína modelo é designada para ter uma propriedade biofísica aprimorada de acordo com um método de classificação de consenso.

Os métodos deste aspecto da invenção envolvem as seguintes etapas:

(1) proporcionar modelos estruturais da proteína modelo e da proteína candidata;

(2) identificar resíduos na proteína modelo que são importantes para estabilidade; e

(3) substituir resíduos na proteína candidata que correspondem aos resíduos importantes na proteína modelo.

Em uma modalidade, a proteína modelo modelo estrutural é uma estrutura de cristal (por exemplo, uma estrutura de cristal de raios X) e a proteína candidata é modelada por modelagem por homologia usando o modelo estrutural da proteína modelo. Em outra modalidade, um resíduo importante é um resíduo da interface VH/VL. Conforme usado aqui, neste Pedido de patente, um "resíduo da interface" é um resíduo que participa em interações domínio-domínio dentro de ou entre proteínas. Preferencialmente, um resíduo da interface se situa no limite, isto é, a "interface" entre dois domínios da proteína e contribui sua área superficial para no mínimo um dos domínios (preferencialmente ambos). Os resíduos que encobrem significativamente mais área superficial do que outros na interface, isto é, os resíduos os quais não expõem significativamente mais de sua área superficial para a superfície da proteína dobrada, são consideradas como sendo mais cruciais para a estabilidade da interface. A área superficial que cada resíduo contribui para a interface pode ser determinada usando técnicas reconhecidas na técnica (por exemplo, análise usando o software MOLMOL). Em algumas modalidades, o resíduo da interface encobre no mínimo 10 $Å^2$ de área superficial (por exemplo, 10 $Å^2$, 15 $Å^2$, 20 $Å^2$, 25 $Å^2$ ou mais de área superficial encoberta).

Resíduos da proteína candidata localizados em posições de aminoácidos correspondentes às dos resíduos importantes (por exemplo, resíduos da interface encobertos) da proteína modelo são preferencialmente substituídos se são de um tipo não-idêntico. Em algumas modalidades, somente são feitas substituições não-conservativas. Em outras modalidades, somente são feitas substituições conservativas.

Em modalidades exemplares, as proteínas modelo e candidatas

são moléculas de ligação (por exemplo, anticorpos ou variantes dos mesmos (por exemplo, moléculas scFv)). Em uma modalidade, a interface que é examinada é aquela entre um domínio Fc (por exemplo, articulação, CH2, e CH3) e outro domínio Fc. Em outra modalidade, a interface é aquela formada entre o domínio V_H e o domínio V_L.

Em modalidades preferenciais, os projetos são focalizados sobre a interface entre os domínios VH e VL de um anticorpo. Dados biofísicos (vide dados de DSC – Exemplo 15) sugerem que a afinidade de VH e VL um para o outro é marginal no formato de scFv (isto é na ausência dos domínios CH1 e CL do Fab). Em particular, interações de VH e VL não são sempre suficientemente fortes (isto é, não têm suficiente afinidade para levar a uma entidade única cooperativamente dobrada. Portanto, aumentar a estabilidade da interface VH/VL (isto é, aumentar a afinidade entre os domínios VH e VL) pode ser uma via preferencial para reforçar a estabilidade de anticorpo / scFv. Estabilizando a interface, a transição de dobramento de anticorpo pode ser convertida em um único evento, aumentar a afinidade dos domínios um pelo outro, e possibilitar flutuações menos dinâmicas que podem levar a área superficial hidrofobica exposta (conforme observado por experimentos de ligação de ANS).

Em ainda outras modalidades, os projetos podem ser validados usando experimentos de calorimetria de varredura difierencial (DSC) os quais possibilitam que uma pessoa versada na técnica observe não quantitativamente aumentos de afinidade real entre o VH e VL.

(c) Covariação Combinada e Design de Interface

Em ainda outros aspectos da invenção, duas ou mais das ferramentas de criação à base de seqüência descritas aqui, neste Pedido de patente, podem ser empregadas para a criação de variantes de proteínas otimizadas. Em uma modalidade típica, design de interface é empregado junto com análise de covariação para determinar resíduos ou posições de resíduos os quais são importantes para a estrutura e função da proteína, incluindo, por exemplo, estabilidade protéica. Em uma modalidade típica, análise de covariação é usada para determinar resíduos ou posições de resíduos

fora de uma interface (aqui, neste Pedido de patente, denominada "resíduos de *estruturação*") que não interagem fisicamente com os resíduos da interface sobre o domínio oposto ou contribuem com área superficial para a interface, mas não obstante são importantes para proporcionar contexto estrutural apropriado para resíduos da interface. Por exemplo, análise de covariação pode ser conduzida sobre resíduos localizados dentro da interface para identificar resíduos de *estruturação* os quais covariam com os resíduos da interface. Preferencialmente, resíduos da interface os quais encobrem um alto grau de área superficial (por exemplo, resíduos que encobrem no mínimo 40 $Å^2$ de área superficial) são selecionados para análise de covariação. Resíduos de *estruturação* os quais covariam com os resíduos da interface selecionados são então identificados usando os métodos de covariação da invenção. Preferencialmente, são selecionados resíduos dos quais covariam fortemente (por exemplo, valor de phi de no mínimo 0,25). Em outra modalidade preferencial, são selecionados resíduos de *estruturação* se eles covariam com no mínimo dois resíduos da interface (por exemplo, 2, 3, 4, 5, 6, 7, 8, 9, 10 ou mais resíduos da interface). Quanto maior o número de covariações, maior o escore de covariação que pode ser atribuído ao resíduo de *estruturação*, uma vez que o número de covariações é preditivo da contribuição feita pelo resíduo de estruturação para estabilizaçaõ da interface VH/VL.

Os métodos deste aspecto da invenção envolvem as seguintes etapas:

(1) proporcionar modelos estruturais de uma proteína modelo e uma proteína candidata;

(2) identificar resíduos da interface (por exemplo, VH/VL) na proteína modelo que são importantes para estabilidade;

(3) identificar resíduos de estruturação os quais covariam com os resíduos da interface da etapa (2), por exemplo, usando a Ferramenta de Análise de Covariação descrita aqui, neste Pedido de patente;

(4) substituir um ou mais resíduos da interface ou resíduos de estruturação na proteína candidata com os resíduos correspondente ou resíduos de estruturação identificados nas etapas (2) e (3).

Resíduos da proteína candidata localizados em posições de aminoácido correspondentes as dos resíduos importantes (por exemplo, resíduos de interface ou de estruturação) da proteína modelo são preferencialmente substituídos se eles são de um tipo não -idêntico. Em algumas modalidades, somente são feitas substituições não-conservativas. Em outras modalidades, somente são feitas substituições conservativas.

Em uma modalidade, da invenção, a seqüência aprimorada de um dado polipeptídeo é armazenada ou fornecida como resultado. Em outra modalidade, um polipeptídeo selecionado pode ser formulado para uso terapêutico.

V. Sistemas, Interfaces, e Métodos Implementados por Computador

Um aspecto da invenção pertence a métodos (por exemplo, métodos implementados por computador), equipamentos, e software para implementar os métodos de seleção ou de criação da invenção. Em modalidades exemplares, a invenção proporciona um método e equipamento implementado por computador para identificar resíduos aminoácidos (por exemplo, resíduos aminoácidos covariantes) em uma seqüência candidata. Em outra modalidade, a invenção proporciona métodos e equipamento implementados por computador para classificar uma ou mais seqüências de ensaio ou candidatas por um escore (por exemplo, uma classificação de consenso ou escore de covariação) o qual é preditivo de uma propriedade biofísica (por exemplo, estabilidade protéica, atividade catalítica, atividade terapêutica, resistência a um patógeno ou toxina, toxicidade, e etc).

O aspecto do métodos implementados por computador pode ser descrito pela seguinte seqüência de operações: (a) receber dados (por exemplo, dados de seqüência ou estruturais) caracterizando uma série de referência e uma seqüência de ensaio; (b) a partir dos dados, calcular um escore (por exemplo, um escore de covariação ou classificação de consenso); (c) classificar os escores para identificar uma ou mais seqüências ou resíduos aminoácidos de interesse.

Em algumas modalidades, os escores são classificados para identificar um ou mais resíduos aminoácidos que devem permanecer fixados

na seqüência de ensaio ou candidata. Em outras modalidades, os escores são classificados para identificar um ou mais resíduos aminoácidos os quais são candidatos para substituição. Em uma modalidade típica, o método implementado por computador classifica posições de resíduos (ou resíduos específicos em algumas posições) na ordem do escore de covariação. Em outra modalidade exemplar, o método implementado por computador classifica seqüências candidatas na ordem da classificação de consenso.

Ainda outro aspecto da invenção pertence a equipamento e meios legíveis à máquina nos quais são proporcionados instruções de programa e/ou arranjos de dados para implementar os métodos descritos acima. Freqüentemente, as instruções do programa são proporcionadas como código para realizar algumas operações do método. Dados, caso empregados para implementar características desta invenção, podem ser proporcionados como estruturas de dados, tabelas de banco de dados, objetos de dados, ou outros arranjos apropriados de informação especificada. Quaisquer dos métodos ou sistemas desta invenção podem ser representados, no todo ou em parte, como instruções de programa similares e/ou dados proporcionados sobre meios legíveis por máquina.

Os métodos implementados por computador da invenção pode incluir um dispositivo de saída que apresenta as informações a um usuário (por exemplo, um display de CRT, um LCD, uma impressora, um dispositivo de comunicação tal como um modem, saída de áudio, e similares). Além disso, instruções (por exemplo, an algoritmo) para realizar o cálculo, no todo ou em parte, podem ser conferidas a um meio adequado para uso em um dispositivo eletrônico para realizar as instruções. Portanto, os métodos da invenção são suscetíveis a uma abordagem de alta produtividade compreendendo software (por exemplo, instruções legíveis por computador) e hardware (por exemplo, computadores, robótica, e chips). O processo implementado por computador não é limitado a uma plataforma de computador particular, processador particular, ou linguagem de programação de alto nível particular.

Em algumas modalidades, os cálculos usados nos métodos da

invenção são realizados por um algoritmo de computador adequado para uso com uma linguagem de programação de computador (por exemplo, PERL). Em modalidades exemplares, o algoritmo de computador é projetado para reconhecer (i) um alinhamento de seqüência de uma série de referência e/ou (ii) uma ou mais seqüências de ensaio como input para o algoritmo.

(i) Classificação de Consenso Implementada por Computador

Em um aspecto da invenção pertence a método implementado por computador, equipamento, e software para conduzir o método de classificação de consenso da invenção descrito acima usando um algoritmo implementado por computador. Em modalidades exemplares, o algoritmo é projetado para realizar uma ou mais das seguintes etapas: (1) cortar cada posição de resíduo do alinhamento como uma coluna e calcular a freqüência de resíduo de consenso nesta posição; (2) cortar cada posição de resíduo da uma ou mais seqüências de ensaio e calcular a freqüência de resíduo de ensaio nesta posição, e (3) dividir a freqüência de resíduo de ensaio pela freqüência de resíduo de consenso para dar uma classificação de consenso; e/ou (4) somar a classificação de consenso em cada posição para dar uma classificação de consenso total. Em outras modalidades, o algoritmo pode ser projetado para realizar uma ou mais das seguintes etapas adicionais: (5) calcular a classificação de consenso média da série de referência, e/ou (6) o escore de seqüência da seqüência de ensaio.

Estes aspectos da invenção também são incorporados em um sistema para conduzir o método de classificação de consenso da invenção. O sistema inclui (a) um computador que inclui um banco de dados capaz de armazenar no mínimo uma população de séries de referência, e (b) software de sistema. O software de sistema inclui uma ou mais instruções de lógica para conduzir quaisquer das etapas (1) a (6) do método de classificação de consenso implementado por computador.

A invenção também proporciona um produto de programa de computador para conduzir o método de classificação de consenso da invenção. O produto de programa de computador inclui um meio legível por com-

putador tendo um ou mais instruções de lógica para conduzir quaisquer das etapas (1) a (6).

(ii) Análise de Covariação Implementada por Computador

Outro aspecto da invenção pertence a método implementado por computador para conduzir a método de análise de covariação da invenção conforme descrito acima usando um algoritmo implementado por computador. Em modalidades exemplares, os métodos realizam uma ou mais das seguintes etapas: (1) cortar cada posição de resíduo da série alinhada como uma coluna; (2) arranjar colunas em uma matriz; (3) calcular a correlação entre várias colunas na matriz para derivar um termo correlativo; (4) adicionar termos correlativos a termos lineares, os quais correspondem a resíduos aminoácidos, paa gerar uma matriz preditora expandida; (5) gerar um modelo derivado heuristicamente (por exemplo, um modelo de Markov oculto) a partir da matriz preditora expandida para identificar correlações importantes.

Estes aspectos da invenção também são incorporados em um sistema para conduzir o método de análise de covariação da invenção. O sistema inclui (a) um computador que inclui um banco de dados capaz de armazenar no mínimo uma população de bibliotecas de cadeias de qualidades características, e (b) software de sistema. O software de sistema inclui uma ou mais instruções de lógica para conduzir quaisquer das etapas (1) a (5) do método de classificação de covariação de método implementado por computador.

A invenção também proporciona um produto de programa de computador para conduzir o método de classificação de covariação da invenção. O produto de programa de computador inclui um meio legível por computador tendo uma ou mais instruções de lógica para conduzir quaisquer das etapas (1) a (5).

(iii) Interfaces Gráficas do Usuário para Análise de Covariação

Em alguns aspectos típicos, a invenção proporciona uma nova interface gráfica do usuário (GUI) para uso com os métodos de covariação da invenção. Esta ferramenta de análise de covariação, denominada NAP-

MAP, auxiliar o usuário a visualizar covariações no contexto de série alinhada de seqüências. A nova interface gráfica do usuário é implementada por computador usando linguagens de programação reconhecidas na técnica (por exemplo, em Java 1.4.2 utilizando a Processing Library em processing.org).

(a) Layouts de Grade

Em um aspecto, a interface gráfica do usuário da invenção compreende uma representação gráfica de um alinhamento (por exemplo, um alinhamento de seqüência criado de acordo com os métodos da invenção acima) ou dados correspondentes do mesmo. Uma simples forma deste display é o layout de grade. Conforme usado aqui, neste Pedido de patente, o termo "layout de grade" (também conhecido como uma plataforma de alinhamento) se refere a uma matriz na qual as propriedades de um alinhamento são representados. Em uma modalidade, o layout de grade graficamente representa posições de resíduos das seqüências de um alinhamento. Em outra modalidade, o layout de grade graficamente representa tipos de resíduo de seqüências dentro do alinhamento. Em outra modalidade, o layout de grade graficamente representa a freqüência de tipos de resíduo (ou "freqüência de uso de resíduo") dentro do alinhamento. Em ainda outras modalidades, o layout de grade representa um ou mais propriedades de resíduos dentro do alinhamento (por exemplo, hidrofobicidade, carga, tamanho, pKa, e etc.).

Em algumas modalidades exemplares, o layout de grade compreende um display representativo de (i) posições contíguas de resíduos correspondentes a uma seqüência do alinhamento; e (ii) freqüência de uso de resíduos para tipos de resíduo em cada posição de resíduo do alinhamento de seqüência. Em uma modalidade, a freqüência de uso de resíduo é representada simbolicamente (por exemplo, por um círculo ou quadrado). Em outras modalidades, todos os 20 aminoácidos naturais, intervalos, e resíduos ambíguos são representados. Em algumas modalidades, as posições de resíduos são apresentadas em linhas e a freqüência de uso de resíduos são apresentadas em colunas. Em algumas modalidades, as posições de

resíduos são apresentadas em colunas e a freqüência de uso de resíduos são apresentados em linhas.

Em uma modalidade, a freqüência de uso de resíduos se correlaciona (por exemplo, correlacionados positivamente ou negativamente) com o tamanho do símbolo (por exemplo, tamanho do círculo), de modo que resíduos mais conservados podem ser discriminados de resíduos observados menos freqüentemente. Por exemplo, a freqüência de resíduos pode ser representada como círculos de vários tamanhos com área de círculo proporcional à freqüência de seu uso. O raio de cada círculo (R) pode ser usado para desenhar o nó usando uma função de elipse calculada de acordo com a seguinte fórmula (IV)

$$R = \sqrt{\frac{A \cdot \frac{N}{M}}{\pi}} \quad \text{(IV)}$$

em que:

A = área média desejada de todos os nós

N = número de seqüências usando este resíduo

M = o número médio de seqüências usando cada resíduo

Em uma modalidade exemplar, o layout de grade compreende uma matriz de colunas e linhas contíguas em que (i) colunas contíguas representam graficamente posições contíguas correspondentes de resíduos do alinhamento; (ii) cada uma das linhas contíguas representam graficamente um tipo de aminoácido específico; e (iii) a freqüência de uso de resíduo em cada posição de é representada graficamente em cada célula da matriz. Preferencialmente todos os tipos de resíduo (por exemplo, todos os 20 aminoácidos e/ou intervalos possíveis) são representados por uma linha separada no layout de grade. Mais preferencialmente, a freqüência de uso de resíduo em cada posição de resíduo do alinhamento é representada graficamente em cada célula da matriz com um símbolo (por exemplo, um círculo) cujo tamanho se correlaciona (por exemplo, correlacionado positivamente) com a freqüência de uso do resíduo. Um layout de grade típico é representado na Figura 76.

(b) Revestimentos de Covariação

Em outros aspectos, a interface gráfica do usuário da invenção compreende uma representação gráfica de uma ou mais covariações (por exemplo, covariações identificadas de acordo com os métodos da invenção acima) ou dados correspondentes às mesmas.

Em algumas modalidades, a interface gráfica do usuário da invenção compreende representação gráfica tanto de (i) um layout de grade; quanto de (ii) uma ou mais covariações. Nestas modalidades, o layout de grade pode servir como plataforma de alinhamento para o display de dados de covariação. Dados de covariação podem ser apresentados, por exemplo, revestindo o layout de grade com dados de covariação. Dados de covariação arranjados deste modo são denominados um "revestimento de covariação". Por exemplo, covariações podem ser representadas no revestimento de covariação como linhas ou rede de linhas conectando dois ou mais aminoácidos covariáveis que são representados graficamente dentro do layout de grade.

As linhas do revestimento de covariação podem ser de várias espessuras (por exemplo, linhas finas ou barras) ou continuidades (por exemplo, tracejadas ou sólidas, retas ou dentadas), na medida que a conexão entre dois ou mais aminoácidos covariáveis seja aparenta para o usuário (vide, por exemplo, a Figura 77, em que covariações são represenadas como barras semitransparentes que conectam dois aminoácidos covariáveis). Em algumas modalidades exemplares, a representação gráfica da covariação pode representar adicionalmente a importância estatística da covariação (por exemplo, seu valor de ϕ ou de χ^2). Em uma modalidade típica, a espessura de uma linha dentro do revestimento da covariação pode ser proporcional à importância estatística da covariação. Por exemplo, a espessura da linha do revestimento de covariação pode ser desenhada em proporção ao valor de ϕ de acordo com a fórmula V:

$$W = 1 + (M \cdot \Phi^N) \quad \text{(V)}$$

em que:

W = espessura de linha

Φ = phi (coeficiente de correlação)

N = multiplicador de amplificação

M = espessura de linha máx

Em outra modalidade exemplar, covariações positivas e negativas podem ser discriminadas graficamente umas das outras (por exemplo, por linhas coloridas ou espessuras diferentes). Em ainda outras modalidades exemplares, somente covariações acima de um certo nível de limiar de importância estatística são apresentadas ao usuário. Por exemplo, em algumas modalidades, podem ser aplicados cortes do valor de ϕ para visualizar somente covariações com forças de correlação desejadas.

(c) Traços de Seqüência

Em outros aspectos opcionais, a interface gráfica do usuário compreende uma representação gráfica de uma seqüência de interesse para auxiliar com possíveis esforços de criação de proteína. A seqüência de interesse pode ser a seqüência de ensaio ou candidata (isto é, uma seqüência de entrada ou pode ser uma seqüência na qual covariações desejáveis são integradas (isto é, uma seqüência de saída). A seqüência de interesse pode ser representada graficamente como uma linha.

Em uma modalidade, a interface gráfica compreende tanto (i) uma representação de uma seqüência de interesse; quanto (ii) um layout de grade (por exemplo, a seqüência de interesse é sobreposta sobre o layout de grade). Em outra modalidade, a interface gráfica compreende tanto (i) uma representação de uma seqüência de interesse; quanto (ii) um revestimento de covariação (por exemplo, a seqüência de interesse é sobreposta sobre o revestimento de covariação). Em ainda outras modalidades, a interface gráfica compreende (i) uma representação de uma seqüência de interesse; (ii) um revestimento de covariação; e (iii) um layout de grade (por exemplo, a seqüência de interesse é sobreposta (por exemplo, acompanhamento) sobre o revestimento de covariação o qual é por sua vez sobreposto sobre o layout de grade).

Uma seqüência de interesse pode ser apresentada sobre um layout de grade, por exemplo, como uma curva conectando aminoácidos

adjacentes que são representados graficamente dentro do layout de grade. Uma seqüência de interesse apresentada deste modo é denominada um "traço de seqüência". Em modalidades preferenciais, o traço de seqüência deve passar ou atravessar através de somente uma célula (ou respresentação gráfica de um tipo de aminoácido) para uma linha ou coluna em particular do layout de grade que representa uma posição de resíduo em particular. Mais preferencialmente, o traço de seqüência passa através da célula a qual representa graficamente o tipo de resíduo ocorrendo em uma posição de resíduo particular dentro da seqüência de interesse. O traço de seqüência pode passar através de todas as células do layout de grade que representam um resíduo dentro da seqüência de interesse, ou pode passar através de somente as células que representam uma porção da seqüência de interesse. Um traço de seqüência típico de uma seqüência de interesse é representado na Figura 77. Traços de seqüências representados na mesma (*splines* de *catmull-rom*) foram desenhados usando uma função de Bezier. Dados dois nós para desenhar uma curva de Bezier de transição de seqüência, os nós dos pares de nós anterior e seguinte são usados como pontos de âncora e os próprios dois nós são usados como o primeiro e segundo pontos de controle. Para os da primeira coluna (representando os primeiros resíduos de seqüências de aminoácidos), o primeiro ponto de controle dobra como seu ponto de âncora; para os da última coluna, seu segundo ponto de controle dobra como seu ponto de âncora.

Em modalidades preferenciais, o gráfico do traço de seqüência é tornado interativo para manipulação pelo usuário. Por exemplo, as células dentro do layout de grade (representando talvez tornadas interativas para seleção pelo usuário. Em uma modalidade, a célula selecionada pode ser atribuído um estado de seleção particular (por exemplo, um estado de seleção positivo, neutro, ou negativo). Em uma modalidade típica, uma célula representando um resíduo particular pode ser selecionada positivamente pelo usuário de modo que o traço de seqüência pode ser retraçado para passar através da célula selecionada. Por exemplo, se uma célula for selecionada positivamente pelo usuário, múltiplos traços de seqüências podem

ser gerados para todas as seqüências de interesse (por exemplo, todas as seqüências dentro da série de referência ou alinhamento) que empregam o resíduo representado pela célula selecionada positivamente. Em uma modalidade alternativa, uma célula representando um resíduo particular pode ser selecionada negativamente pelo usuário de modo que o traço de seqüência não passa através da célula selecionada. Por conseguinte, traços de seqüências podem ser gerados para cada seqüência de interesse (por exemplo, todas as seqüências dentro da série de referência ou alinhamento) que exclui o resíduo representado pela célula selecionada negativamente.

Múltiplas seqüências de interesse podem ser apresentadas graficamente, na medida que possam ser distinguidas pelo usuário (por exemplo, usando diferentes espessuras ou continuidades de linhas). Por exemplo, para comparar diferentes alinhamentos de seqüência, traços de seqüências representando múltiplas seqüências de interesse de dois ou mais alinhamentos podem ser revestidos sobre o layout de grade (por exemplo, com cores diferentes, preferencialmente cores opostas) para permitir a visualização de uso de resíduo diferencial em uma escala global.

(d) Modos de Display

A interface gráfica do usuário pode facilitar qualquer funcionalidade reconhecida na técnica. Por exemplo, em algumas modalidades, a interface gráfica do usuário pode ser capaz de girar em movimento panorâmico ou dar um zoom sobre diferentes porções do layout de grade. Por exemplo, o tamanho da janela de visualização do layout de grade pode ser alterado dependendo da resolução de tela do computador do usuário e a visualização pode sair de *close-up* ou ser girada em movimento panorâmico para visualizar um alinhamento que não se ajusta na janela de visualização inicialmente.

Em modalidades preferenciais, é empregado mascaramento de dados, por exemplo, para enfatizar os pares de covariações mais importantes e/ou redes através de vários pares. A interace gráfica do usuário da invenção pode empregar um ou mais modos de display para a visualização de covariações (por exemplo, covariações estatisticamente signicativas). Prefe-

rencialmente, os modos de display referidos estão disponíveis em um gráfico interativo ao visualizar um traço de seqüência de uma seqüência de interesse. Viewermodes típicos são como se segue:

Modo de visualização nº 1: Somente são apresentadas covariações entre resíduos que também estão presentees na seqüência de interesse. Estas covariações podem ser consideradas como interações resíduo-resíduo concluídas por hipótese (por exemplo, aquelas encontradas por análise estatística da freqüência de pares de resíduos) que mantêm a molécula de proteína correspondente junto e podem ser importantes para função.

Modo de visualização nº 2: As únicas covariações apresentadas são aquelas entre pares de resíduos tendo somente um resíduo do par de resíduos presente na seqüência de interesse. Estas covariações são interações resíduo-resíduo que tendem a ser conservadas mas não estão presentees na seqüência de interesse, e cuja ausência pode ser prejudicial à estabilidade protéica. Este modo de visualização é particularmente preferencial quando a Ferramenta de Análise de Covariação é empregada para criação de proteína (vide aSeção IV acima).

Modo de visualização nº 3: As únicas covariações apresentadas são aquelas entre pares de resíduos não tendo nehum membro aminoácido do par covariante presente na seqüência de interesse.

Em algumas modalidades, todos os viewermodes são apresentados ao usuário mas cada um em um único formato de display. Em outra modalidade, todos os viewermodes são apresentados separadamente (por exemplo, em janelas de display separadas).

A Figura 89 representa um ambiente típico adequado para praticar uma modalidade da invenção. Um dispositivo de computação 8900 suporta um processo de compilação de seqüências inicial 8902 e facilidade de análise 8904. O dispositivo de computação 8900 pode ser uma estação de trabalho, um servidor, um laptop, um *mainframe*, um PDA ou outro dispositivo de computação equipado com um ou mais processadores e capaz de suportar o processo de compilação de seqüências inicial 8902 e a facilidade de análise 8904. O dispositivo de computação 8900 pode ter um único proces-

sador ou múltiplos processadores e cada um dos processadores pode ter um núcleo ou múltiplos núcleos. A facilidade de análise 8904 é implementada em software e analisa programaticamente covariação entre pares de resíduos dentro do alinhamento para obter dados de covariação. O dispositivo de computação 8900 está em comunicação com, e fornece dados para, um dispositivo de display de saída 8910. O dispositivo de display de Saída 8910 apresenta uma interface gráfica do usuário (GUI) 8912 gerada pela facilidade de análise 8904. A interface gráfica do usuário 8912 pode incluir um layout de grade 8914 e um revestimento de covariação 8916. O dispositivo de computação 8900 também pode estar em comunicação sobre uma rede 8920 com uma ou mais localizações de armazenamento 8930 e 8940 que encerram respectivamente dados de seqüências 8932 e 8942. O processo de compilação de seqüências inicial 8902 recupera programaticamente seqüências candidatas dos dados de seqüências 8932 e 8942 com base em parâmetros supridos pelo usuário e/ou predeterminados. Os dados de seqüências recuperados são analisados pela facilidade de análise 8904. Será reconhecido por aqueles versados na técnica que embora o processo de compilação de seqüências inicial 8902 e facilidade de análise 8904 sejam reporesentados separadamente na Figura 89, podem ser implementados como parte de uma aplicação, processo ou plug-in integrados. Similarmente, deve ser reconhecido que o processo de compilação de seqüências inicial 8902 e facilidade de análise podem cada um separadamente ser uma tarefa, linha, processo, aplicação ou plug-in. O dispositivo de computação 8900 pode comunicar sobre a rede 8920 com as localizações de armazenamento 8930 e 8940 usando uma série de meios e configurações diferentes. Por exemplo, a rede 8920 podem ser arranjada como uma rede de área local LAN (Local Area Network), uma rede de área ampla WAN (Wide Area Network), uma intranet, a Internet, e/ou pode ser uma rede wireless e/ou uma linha de telefone, ou algum outro tipo de rede possibilitando ao dispositivo de computação 8900 se comunicar com as localizações de armazenamento 8930 e 8940. As localizações de armazenamento 8930 e 8940 podem ser bancos de dados ou algum outro tipo de estrutura de dados de armazena-

mento hospedado pelos dispositivos de computação que estão acessíveis sobre a rede 8920. Deve ser observado que embora componentes da presente invenção sejam representados em uma configuração particular na Figura 89, outars configurações de componentes também são possíveis dentro do âmbito da presente invenção. Por exemplo, o processo de compilação de seqüências inicial 8902 e facilidade de análise 8904 podem estar localizados sobre dispositivos de computação separados e/ou os dados de seqüências 8932 e 8942 podem estar localizados sobre um ou mais bancos de dados hospedados pelo dispositivo de computação 8900.

A Figura 90 é um fluxograma de uma seqüência típica de etapas que podem ser seguidas por uma modalidade da invenção de modo a determinar um escore de covariação de seqüência que pode ser usado como uma medida da estabilidade de um polipeptídeo. A seqüência se inicia com o processo de compilação de seqüências inicial 8902 proporcionando alinhamento de uma série de referência curada de seqüências correspondentes a uma dobra de Ig de um polipeptídeo (etapa 9000). A facilidade de análise 8904 então calcula covariação entre resíduos da seqüência do alinhamento conforme descrito aqui, neste Pedido de patente, para gerar dados de covariação (etapa 9002). A facilidade de análise 8904 usa os dados de covariação para determinar um escore de covariação de seqüência para uma seqüência candidata entre covariações dentro dos dados de covariação (etapa 9004). O escore de covariação de seqüência determinado proporciona um indicador da medida de estabilidade do polipeptídeo e pode ser armazenado e/ou fornecido a um usuário através da interface gráfica do usuário 8912 ou em alguma outra maneira (etapa 9006).

Conforme observado acima, a facilidade de análise 8904 também pode ser usada para determinar uma classificação de consenso. A Figura 91 é um fluxograma de uma seqüência típica de etapas que podem ser seguidas por uma modalidade da invenção de modo a determinar e usar uma classificação de consenso para prever a estabilidade de uma proteína candidata. A seqüência se inicia usando o processo de compilação de seqüências inicial 8902 para proporcionar uma série de referência de seqüên-

cias correspondentes a uma seqüência de domínio de ensaio de uma proteína candidata (etapa 9100). A facilidade de análise 8904 então determina freqüências de resíduos em posições de aminoácido dentro da seqüência de domínio de ensaio para obter uma classificação de consenso (etapa 9102). A classificação de consenso determinada pode ser então armazenada e/ou fornecida a um usuário para proporcionar uma previsão da estabilidade da proteína candidata (etapa 9104).

A facilidade de análise 8904 também pode ser usada para determinar uma classificação de consenso média. A Figura 92 é um fluxograma de uma seqüência típica de etapas que podem ser seguidas por uma modalidade da invenção de modo a determinar e usar uma classificação de consenso média como uma medida da estabilidade de uma proteína candidata. A seqüência se inicia usando o processo de compilação de seqüências inicial 8902 para proporcionar uma série de referência de seqüências correspondentes a uma seqüência de domínio de ensaio de uma proteína candidata (etapa 9200). A facilidade de análise 8904 então determina freqüências de resíduos em posições de aminoácido dentro da seqüência de domínio de ensaio para obter uma classificação de consenso (etapa 9202). A facilidade de análise 8904 também determina freqüências de resíduos dentro das seqüências da série de referência para determinar uma classificação de consenso média (etapa 9204). A classificação de consenso é comparada com a classificação de consenso média de modo a determinar um escore de seqüência (etapa 9206) e o escore de seqüência determinado pode ser então armazenado e/ou fornecida a um usuário para proporcionar uma previsão da estabilidade da proteína candidata (etapa 9206).

A Figura 93 é um fluxograma de uma série típica de etapas que podem ser seguidas por uma modalidade da invenção de modo a identificar e substituir um resíduo aminoácido que falha em satisfazer uma covariação com o resíduo covariável encontrado em uma posição correspondente no alinhamento para determinar uma seqüência aprimorada de aminoácidos do polipeptídeo. A seqüência se inicia pelo processo de compilação de seqüências inicial 8902 proporcionando alinhamento de uma série de referência cu-

rada de seqüências correspondentes a uma dobra de Ig de um polipeptídeo (etapa 9300). A facilidade de análise 8904 então calcula covariação entre resíduos da seqüência do alinhamento para identificar resíduos covariáveis (etapa 9302). A facilidade de análise 8904 pode identificar um resíduo particular do polipeptídeo que falha em satisfazer uma covariação e pode substituir um resíduo covariável encontrado em uma posição correspondente no alinhamento (etapa 9304) de modo a melhorar a seqüência do polipeptídeo. A seqüência aprimorada pode ser então armazenada ou fornecida a um usuário (etapa 9306).

A presente invenção pode ser proporcionada como um ou mais programas legíveis por computador incorporados sobre ou em um ou mais meios. Os meios podem ser um disquete, um disco rígido, um CD, um disco versátil digital, um cartão de memória, uma PROM, uma MRAM, uma RAM, uma ROM, ou uma fita magnética. Em geral, os programas legíveis por computador podem ser implementados em qualquer linguagem de programação. Alguns exemplos de linguagens que podem ser usados incluem FORTRAN, C, C++, C#, Python ou Java. Os programas de software podem ser armazenados sobre ou em um ou mais meios como código de objeto. Aceleração de hardware pode ser usada e todo ou uma porção do código pode rodar em um FPGA, um ASIP, ou um ASIC. O código pode rodar em um ambiente virtualizado tal como em uma máquina virtual. Múltiplas máquinas virtuais rodando o código podem estar presentes sobre um único processador.

VI. Métodos para Avaliar a Estabilidade Protéica

As propriedades de estabilidade das composições da invenção podem ser analisadas usando métodos conhecidos na técnica. Podem ser empregados parâmetros de estabilidade aceitáveis por aqueles na técnica. Parâmetros típicos são descritos em mais detalhes abaixo. Em modalidades exemplares, é avaliada a estabilidade térmica. Em modalidades preferenciais, são avaliados os níveis de expressão (por exemplo, medidos por % de produção) das composições da invenção. Em outras modalidades preferenciais, são avaliados os níveis de agregração das composições da invenção.

Em algumas modalidades, as propriedades de estabilidade de uma composição de uma invenção são comparadas com as de um controle adequado. Controles típicos incluem molécula scFv convencional. Um controle particularmente preferencial é uma molécula scFv $(Gly_4Ser)_3$.

Em uma modalidade, são medidos um ou mais parâmetros descritos abaixo. Em uma modalidade, um ou mais destes parâmetros é medido depois de expressão em uma célula mamífera. Em uma modalidade, um ou mais parâmetros descritos abaixo são medidos sob condições de fabricação em larga escala (por exemplo, expressão de scFvs ou moléculas compreendendo scFvs em um biorreator).

a) Estabilidade Térmica

A estabilidade térmica das composições da invenção pode ser analisada usando uma série de técnicas biofísicas ou bioquímicas não-limitantes conhecidas na técnica. Em algumas modalidades, a estabilidade térmica é avaliada por espectroscopia analítica.

Um típica método de espectroscopia analítica é Calorimetria de Varredura Diferencial (DSC). DSC emprega um calorímetro o qual é sensível às absorvências de calor que acompanham o desdobramento da maioria dsa proteínas ou domínios protéicos (vide, por exemplo, Sanchez-Ruiz, et al., *Biochemistry*, 27: 1648-52, 1988). Para determinar a estabilidade térmica de uma proteína, uma amostra da proteína é inserida no calorímetro e a temperatura é aumentada até o Fac ou scFv se desdobrar. A temperatura na qual a proteína se desdobra é indicativa da estabilidade protéica global.

Outro método de espectroscopia analítica típico é espectroscopia por Dicroismo Circular (CD). A espectroscopia por CD mede a atividade ótica de uma composição como uma função da temperatura crescente. A espectroscopia por dicroismo circular (CD) mede diferenças na absorção de luz polarizada levógira versus luz polarizada dextrógira a qual surge devido a assimetria estrutural. Uma estrutura desordenada ou desdobrada resulta em um espectro de CD muito diferente daquele de uma estrutura ordenada ou dobrada. O espectro de CD reflete a sensibilidade das proteínas aos efeitos desnaturantes de temperatura crescente e portanto é indicativo da estabili-

dade térmica de uma proteína (vide van Mierlo e Steemsma, *J. Biotechnol.*, 79(3):281-98, 2000).

Outro método típico de espectroscopia analítica para medir a estabilidade térmica é Espectroscopia por Emissão de Fluorescência (vide van Mierlo e Steemsma, acima). Aionda outro método típico de de espectroscopia analítica para medir a estabilidade térmica é espectroscopia por Ressonância Magnética Nuclear (NMR) (vide, por exemplo, van Mierlo e Steemsma, acima).

Em outras modalidades, a estabilidade térmica de uma composição da invenção é medida bioquimicamente. Um método bioquímico típico para avaliar a estabilidade térmica é uma prova de ensaio térmico. Em uma "prova de ensaio térmico", uma composição da invenção é submetida a uma faixa de temperaturas elevadsa por um período de tempo determinado. Por exemplo, em uma modalidade, moléculas scFv de ensaio ou moléculas compreendendo moléculas scFv são submetidas a uma faixa de temperaturas crescentes, por exemplo, por 1 a 1,5 horas. A atividade da proteína é então avaliada por um ensaio bioquímico relevante. Por exemplo, se a proteína for uma proteína de ligação (por exemplo, um scFv ou um polipeptídeo contendo scFv da invenção) a atividade de ligação da proteína de ligação pode ser determinada por um ELISA funcional ou quantitativo.

Em uma modalidade, um ensaio similar pode ser feito em um formato de alta produtividade. Em outra modalidade, uma biblioteca de scFv variantes pode ser criada usando métodos conhecidos na técnica. A expressão de scFv pode ser induzida alguns scFvs podem ser submetidos a estímulo térmico. As amostras de ensaio estimuladas podem ser avaliadas para ligação e os scFvs os quais são estáveis podem ser aumentados em escala e adicionalmente caracterizados.

Em algumas modalidades, a estabilidade térmica é avaliada medindo a temperatura de fusão (Tm) de uma composição da invenção usando quaisquer das técnicas acima (por exemplo, técnicas de espectroscopia analítica techniques). A temperatura de fusão é a temperatura na qual o ponto central de uma curva de transição térmica em que 50% das moléculas

de uma composição estão em um estado dobrado.

Em outras modalidades, estabilidade térmica é avaliada medindo o calor específico ou capacidade térmica (Cp) de uma composição da invenção usando uma técnica calorimétrica analítica (por exemplo, DSC). O calor específico de uma composição é a energia (por exemplo, em kcal/mol) requerida para aumentar por 1°C, a temperatura de 1 mol de água. Um Cp grande é um marco de uma composição de proteína desnaturada ou inativa. Em algumas modalidades, a mudança na capacidade térmica (ΔCp) de uma composição é medida determinando o calor específico de uma composição antes e depois de sua transição térmica. Em outras modalidades, a estabilidade térmica pode ser avaliada medindo ou determinando outros parâmetros da estabilidade termodinâmica incluindo energia de desdobramento livre de Gibbs (ΔG), entalpia de desdobramento (ΔH), ou entropia de desdobramento (ΔS).

Em outras modalidades, uma ou mais das provas bioquímicas acima (por exemplo, uma prova de ensaio térmico) são usadas para determinar a temperatura (isto é, o valor da T_C) na qual 50% da composição conserva sua atividade (por exemplo, atividade de ligação).

b) <u>% de Agregação</u>

Em algumas modalidades, a estabilidade de uma composição da invenção é determinada medindo sua propensão para agregar. A agregação pode ser medida por uma série de técnicas bioquímico ou biofísicas não-limitantes. Por exemplo, a agregação de uma composição da invenção pode ser avaliada usando cromatografia, por exemplo, Cromatografia de Exclusão de Tamanho (SEC). SEC separa moléculas com base no tamanho. Uma coluna é preenchida com contas semi-sólidas de um gel polimérico que admitirá que íons e moléculas pequenas penetrem no interior mas não moléculas grandes. Quando uma composição de proteína é aplicada no topo da coluna, as proteínas dobradas compactas (isto é, proteínas não-agregadas) são distribuídas através de um maior volume de solvente do que é disponível para os agregados de proteínas grandes. Conseqüentemente, os agregados grandes se movem mais rapidamente através da coluna, e deste modo a

mistura pode ser separada ou fracionada em seus componentes. Cada fração pode ser quantificada separadamente (por exemplo, por dispersão luminosa) à medida que elutria do gel. Por conseguinte, a % de agregação de uma composição da invenção pode ser determinada comparando a concentração de uma fração com a concentração total de proteína aplicada ao gel. Composições estáveis elutriam da coluna como essencialmente uma única fração e aparecem como essencialmente um único pico no perfil de eluição ou cromatograma.

Em modalidades preferenciais, SEC é usado em combinação com dispersão luminosa in-line (por exemplo, dispersão luminosa clássica ou dinâmica) para determinar a % de agregação de uma composição. Em algumas modalidades preferenciais, dispersão luminosa estática é empregada para medir a massa de cada fração ou pico, independente do formato molecular ou da posição de eluição. Em outras modalidades preferenciais, dispersão luminosa dinâmica é empregada para medir o tamanho hidrodinâmico de uma composição. Outros métodos típicos para avaliar a estabilidade protéica incluem SEC de Alta Velocidade (High-Speed SEC, vide, por exemplo, Corbett *et al.*, Biochemistry. 23(8):1888-94, 1984).

Em uma modalidade preferencial, a % de agregação é determinada medindo os agregados de fração de proteína dentro da amostra de proteína. Em uma modalidade preferencial, a % de agregação de uma composição é medida determinando a fração de proteína dobrada dentro da amostra de proteína.

c) % de Rendimento

Em outras modalidades, a estabilidade de uma composição da invenção é avaliada medindo a quantidade de proteína que é recuperada (aqui, neste Pedido de patente, a "% de produção") depois de expressão (por exemplo, expressão recombinante) da proteína. Por exemplo, a "% de produção" pode ser medida determinando miligramas de proteína recuperados para cada ml de meios de cultura hospedeiro (isto é, mg/ml de proteína). Em uma modalidade preferencial a % de produção é avaliada depois de expressão em uma célula hospedeira mamífera (por exemplo, uma célula

CHO).

d) % de Perda

Em ainda outras modalidades, a estabilidade de uma composição da invenção é avaliada monitorando a perda de proteína em uma faixa de temperaturas (por exemplo, de -80 a 25 °C) depois de armazenamento por um período de tempo definido. A quantidade ou concentração de proteína recuperada pode ser determinada usando qualquer método de quantificação de proteína conhecido na técnica, e comparada com a concentração de proteínainicial. Métodos de quantificação de proteína típicos incluem análise SDS-PAGE ou a prova de Bradford para (Bradford, *et al., Anal. Biochem.* 72, 248, (1976)). Um preferred método para avaliar a % de perda emprega quaisquer dos métodos de SEC analíticos descritos acima. Será reconhecido que podem ser determinadas medições da % de Perda sob qualquer condição de armazenamento ou formulação de armazenamento desejada, incluindo, por exemplo, preparações de proteínas liofilizadas.

e) % de Proteólise

Em ainda outras modalidades, a estabilidade de uma composição da invenção é avaliada determinando a quantidade de proteína que é proteolisada depois de armazenamento sob condições de rotina. Em uma modalidade exemplar, a proteólise é determinada por SDS-PAGE de uma amostra da proteína em que a quantidade de proteína intacta é comparada com a quantidade de fragmentos de baixo peso molecular os quais aparecem sobre o gel do SDS-PAGE. Em outra modalidade exemplar, a proteólise é determinada por Espectrometria de Massa (MS), em que a quantidade de proteína do peso molecular esperado é comparada com a quantidade de fragmentos de proteína de baixo peso molecular dentro da amostra.

f) Afinidade de Ligação

Em ainda outras modalidades, a estabilidade de uma composição da invenção pode ser avaliada determinando sua afinidade de ligação alvo. Uma ampla variedade de métodos para determinar a afinidade de ligação são conhecidos na técnica. Um método típico para determinar a afinidade de ligação emprega ressonância de plasmon superficial. Ressonância de

plasmon superficial é um fenômeno ótico que possibilita a análise de interações bioespecíficas em tempo real por detecção de alterações nas concentrações de proteínas dentro de uma matriz de biosensor, por exemplo usando o sistema BIAcore (Pharmacia Biosensor AB, Uppsala, Sweden e Piscataway, NJ). Para descrições adicionais, vide Jönsson, U., *et al.* (1993) *Ann. Biol. Clin.* 51:19-26; Jönsson, U., *et al.* (1991) *Biotechniques* 11:620-627; Johnsson, B., *et al.* (1995) *J. Mol. Recognit.* 8:125-131; e Johnnson, B., *et al.* (1991) *Anal. Biochem.* 198:268-277.

g) Outros Estudos de Ligação

Em ainda outras modalidades, a estabilidade de uma composição da invenção pode ser avaliada quantificando a ligação de um composto marcado a porções desnaturadas ou desdobradas de uma molécula de ligação. As moléculas referidas são preferencialmente hidrofóbicas, uma vez que preferencialmente ligam ou interagem com grandes fragmentos hidrofóbicos de aminoácidos que são normalmente enterrados no interior da proteína nativa, mas os quais são expostos em uma molécula de ligação desnaturada ou desdobrada. Um composto marcado típico é o corante fluorescente hidrofóbico, sulfonato de 1-anilino-8-naftalina (ANS).

VII. Métodos para Selecionar Proteínas Estáveis

Os métodos descritos acima para a previsão de estabilidade protéica podem ser empregados para selecionar uma proteína candidata (por exemplo, anticorpos) para uso posterior. Em algumas modalidades, os métodos da invenção são empregados para selecionar uma proteína para expressão. Em outras modalidades, os métodos da invenção são usados para selecionar uma proteína candidata para modificação. Em modalidades exemplares, os métodos de previsão da invenção podem ser empregados na humanização de um anticorpo de doador não-humano (por exemplo, para selecionar uma imunoglobulina aceitadora).

Uma proteína candidata pode ser selecionada com base em sua classificação de consenso total ou escore de seqüência conforme determinado usando os métodos da invenção. Em algumas modalidades, a proteína candidata é selecionada se sua classificação de consenso for mais de um

controle negativo adequado (por exemplo, mais de 5%, preferencialmente mais de 10%, mais preferencialmente mais de 20%). Em outras modalidades, a proteína candidata é selecionada se sua classificação de consenso for substancialmente similar a (por exemplo, dentro de 20%, 10%, ou 5%) ou mais de (por exemplo, mais de 5%, preferencialmente mais de 10%, mais preferencialmente mais de 20%) um controle positivo adequado.

Em outras modalidades, uma proteína candidata é selecionada se sua classificação de consenso for substancialmente similar a sua classificação de consenso ideal ou perfeita (por exemplo, dentro de 30%, preferencialmente dentro de 20%, mais preferencialmente dentro de 10%).

Em outras modalidades, uma proteína candidata é selecionada se seu escore de seqüência for mais de zero. Em uma modalidade, a proteína candidata é selecionada se seu escore de seqüência for mais de 0,5. Em outra modalidade, a proteína candidata é selecionada se seu escore de seqüência for mais de 1. Em outra modalidade, a proteína candidata é selecionada se seu escore de seqüência for mais de 2. Em uma modalidade preferencial, a proteína candidata é selecionada se seu escore de seqüência for mais de 3.

Em outras modalidades, uma proteína candidata é selecionada se seu Δ escore for mínimo -3, no mínimo -2, ou no mínimo -1, preferencialmente no mínimo 0, mais preferencialmente no mínimo 1.

a. <u>Selecionar uma Imunoglobulina Aceitadora para Humanização de Anticorpos</u>

Anticorpos humanizados podem ser produzidos usando tecnologia de DNA recombinante, vide por exemplo, por exemplo, Queen *et al.*, *Proc. Natl. Acad. Sci. USA,* (1989), 86:10029-10033; Jones *et al.*, *Nature,* (1986), 321:522-25; Riechmann *et al.*, *Nature,* (1988), 332:323-27; Verhoeyen *et al.*, *Science,* (1988), 239:1534-36; Orlandi *et al.*, *Proc. Natl. Acad. Sci. USA*, (1989), 86:3833-37; as Patentes dos Estados Unidos Nos. US 5.225.539; 5.530.101; 5.585.089; 5.693.761; 5.693.762; 6.180.370. Quando um anticorpo de doador não-humano preferencial foi selecionado para humanização, pode ser obtido um anticorpo aceitador humano apropriado, por exemplo, de bancos de dados de seqüências de genes de anticorpos huma-

nos expressados, de seqüências de Ig de linha germinal ou uma seqüência de consenso de vários anticorpos humanos. É mais provável que a substituição de CDRs não humanas em uma framework de domínio variável humano resulte em retenção de sua orientação espacial correta se a framework de domínio variável humano adoptar a mesma conformação ou similar conformação à framework variável não humana da qual se originaram as CDRs. Isto é realizado obtendo os domínios variáveis humanos de anticorpos aceitadores humanos cujas seqüências de framework aprsentam um alto grau de identidade de seqüência com os domínios de framework variável não humana da qual as CDRs foram derivadas. As regiões de framework variável de cadeia pesada e leve podem ser derivadas das mesmas ou de diferentes seqüências de anticorpos humanos. Preferencialmente o anticorpo aceitador humano retém os resíduos canônicos e da interface do anticorpo doador. Adicionalmente, o anticorpo aceitador humano preferencialmente tem substancial similaridade na extensão de alças de CDR. Vide Kettleborough *et al.*, *Protein Engineering* 4:758 (1991); Kolbinger *et al.*, *Protein Engineering* 6:971 (1993) e Carter *et al.*, patente internacional Nº WO 92/22653.

Tendo identificado as CDRs do anticorpo doador e do anticorpo aceitador humano apropriado, a próxima etapa é determinar quais, se houver, resíduos destes componentes devem ser substituídos para otimizar as propriedades do anticorpo humanizado resultante. Tipicamente, alguns ou todos os aminoácidos da cadeia leve ou pesada de imunoglobulina de doador não-humano que são requeridos para ligação de antígeno (por exemplo, uma ou mais CDRs) são usados para substituto para os aminoácidos correspondentes da cadeia leve ou pesada do anticorpo aceitador humano. O anticorpo aceitador humano retém alguns ou todos os aminoácidos que não são necessários para ligação de antígeno. Quando necessário, um ou mais resíduos nas regiões de framework humanas podem ser trocados para resíduos nas posições correspondentes no anticorpo murino de modo a preservar a afinidade de ligação do anticorpo humanizado para o antígeno. Esta mudança algumas vezes é denominada "retro mutação". Alguns aminoácidos dos resíduos de framework de região variável humana são selecionados

para retro mutação com base em sua possível influência sobre a conformação da CDR e/ou ligação a antígeno.

Em alguns casos, no entanto, o processo de humanização resulta em mudanças não naturais para o domínio variável o que confere indesejável instabilidade ao anticorpo humanizado. Por exemplo, a imunoglobulina aceitadora pode ter raras variações de seqüências que conferem baixa estabilidade. Isto é particularmente verdadeiro para algumas seqüências de linha germinal as quais não podem ser usadas com grande freqüência pelo sistema imune.

Os métodos da invenção proporcionam métodos aprimorados para humanização. Em particular, os métodos da invenção possibilitam que o técnico versado preveja qual seqüência de ensaio (por exemplo, seqüência de linha germinal de ensaio) com uma série de referência de seqüências aceitadoras candidatas (por exemplo, seqüências de linha germinal) são convenientemente estáveis para uso na humanização da seqüência aceitadora. Seqüências aceitadoras de ensaio (por exemplo, seqüências de linha germinal) as quais têm um escore que é preditivo de estabilidade aceitável (por exemplo, têm uma alta classificação de consenso comparada com a classificação de consenso média da série de referência) podem ser selecionada\s como uma imunoglobulina aceitadora.

VIII. Métodos à Base de Biblioteca para Identificar Moléculas scFv Estabilizadas

Em outro aspecto, a presente invenção proporciona métodos para identificar moléculas scFv com aprimorada estabilidade protéica, por exemplo, aprimorada estabilidade térmica. Os métodos da invenção compreendem (i) proporcionar uma biblioteca compreendendo moléculas scFv candidatas; e (ii) triar a biblioteca para identificar moléculas scFv candidatas com aprimorada estabilidade protéica relativa a um controle adequado (por exemplo, uma molécula scFv controle). Conforme usado aqui, neste Pedido de patente, uma "molécula scFv candidata" é uma molécula scFv formada introduzindo uma ou mais mutações estabilizantes candidatas na molécula scFv, em que o efeito da mutação estabilizante candidata sobre a estabilida-

de de a molécula scFv não é conhecido *a priori*, isto é, uma molécula scFv na qual uma mutação foi introduzida e a qual deve ser avaliada para determinar se a mutação resultaa em uma molécula scFv com aumentada estabilidade. Em uma modalidade, uma molécula scFv candidata é uma molécula scFv formada introduzindo uma ou mais mutações estabilizantes candidatas em uma molécula scFv convencional (isto é, não-stabilizada). Em outra modalidade, uma molécula scFv candidata é uma molécula scFv formada introduzindo uma ou mais mutações estabilizantes candidatas em uma das moléculas scFv estabilizadas descritas na Seção II, acima.

Por uma "biblioteca de moléculas scFv candidatas" aqui, neste Pedido de patente, se indica no mínimo duas moléculas scFv candidatas não-redundantes, com no mínimo cerca de 10 sendo preferencial, no mínimo cerca de 100 sendo particularmente preferencial, e no mínimo cerca de 1000 sendo especialmente preferencial (por exemplo, no mínimo cerca de 10^4, 10^5, 10^6, 10^7, ou 10^8 moléculas scFv). Em uma modalidade, a biblioteca é randomizada, com mutações não especificadas geradas aleatoriamente em qualquer posição. Em outra modalidade, a biblioteca de scFv é uma parcialmente randomizada ou criada. Isto é, um resíduo aminoácido específico ou classe de resíduos aminoácidos é introduzido aleatoriamente em qualquer posição do scFv, ou uma posição de aminoácido específica é selecionada para mutagênese com um aminoácido não específico, uma classe particular de aminoácidos, ou um amino aminoácido específico. Em uma modalidade preferencial, uma biblioteca de scFv é criada substituindo um resíduo selecionado de uma molécula scFv com um aminoácido de uma classe definida, por exemplo, um aminoácido hidrofóbico, um aminoácido básico, um aminoácido hidrofílico, um aminoácido carregado, um aminoácido estericamente tendencioso (quer pequeno ou grande), ou uma cisteína capaz de formação de ligação dissulfeto.

Em algumas modalidades, a biblioteca de scFv compreende moléculas scFv candidatas com uma ou mais mutações estabilizantes candidatas que são introduzidas dentro de uma região variável (VL e/ou VH) da molécula scFv. Em outras modalidades, a biblioteca de scFv compreende molé-

culas scFv candidatas com uma ou mais mutações estabilizantes candidatas que são introduzidas dentro de um ligante de scFv da molécula scFv. Em outras modalidades, a biblioteca de scFv compreende moléculas scFv candidatas com uma ou mais mutações estabilizantes candidatas que são introduzidas dentro de um ligante de scFv da molécula scFv e uma ou mais mutações estabilizantes candidatas que são introduzidas dentro de uma região variável (VL e/ou VH).

a. Design de Bibliotecas de scFv

Métodos para desenhar a biblioteca de scFv podem ser auxiliados por modelagem molecular ou computacional. Em algumas modalidades, design de biblioteca pode ser realizado *in silico*. Em algumas modalidades, a biblioteca de scFv pode ser criada pelo seguinte método de duas etapas:

1) identificar resíduos aminoácidos alvo no scFv que quando alterados por uma mutação (por exemplo, por substituição de aminoácido), são previstos para resultar em aprimorada estabilidade de scFv (aqui, neste Pedido de patente, "resíduos desestabilizantes candidatos"); e

2) substituir os resíduos aminoácidos alvo com resíduos candidatos estabilizantes de aminoácidos.

Etapa n° *1:* Identificar Resíduos Desestabilizantes Candidatos

Em algumas modalidades, resíduos aminoácidos desestabilizantes candidatos podem ser identificado por análise à base de seqüência. Por exemplo, resíduos aminoácidos desestabilizantes candidatos podem ser identificado comparando uma seqüência de região variável de uma scFv com uma série de referência de seqüências de região variável, por exemplo, seqüências de região variável de anticorpos humanos que ocorrem naturalmente, e selecionando os resíduos aminoácidos de região variável do scFv os quais são incomuns ou raros em suas posições de aminoácidos correspondentes dentro da série de referência. Em modalidades preferenciais, somente as regiões de framework de uma seqüência de região variável da scFv são analisadas, ao passo que as regiões determinantes de complementaridade (CDRs) da região variável são conservadas de modo a evitar disrupção da atividade de ligação da molécula scFv.

Em algumas modalidades, um resíduo aminoácido desestabilizante candidato é um resíduo aminoácido o qual está ausente ou encontrado em uma baixa freqüência em uma posição correspondente dentro de uma série de referência de seqüências homólogas de região variável. Métodos para compilar séries de referência são descritos nas Seções III - V, acima. Em modalidades preferenciais, o resíduo aminoácido desestabilizante candidato é um resíduo aminoácido o qual está presente dentro de menos de 10% das seqüências em posições correspondentes dentro da série de referência. Em modalidades mais preferenciais, o resíduo aminoácido desestabilizante candidato é um resíduo aminoácido o qual está presente dentro de menos de 5% das seqüências em posições correspondentes dentro da série de referência. Em modalidades particularmente preferenciais, o resíduo aminoácido desestabilizante candidato é um resíduo aminoácido o qual está presente dentro de menos de 2% (por exemplo, 0,5%, 0,75%, ou 1%) das seqüências em posições correspondentes dentro da série de referência.

Em outra modalidade, um resíduo aminoácido desestabilizante candidato é um resíduo aminoácido o qual difere do aminoácido presente na posição correspondente na seqüência de consenso da série de referência de seqüências de região variável (isto é, o resíduo aminoácido de consenso). Métodos para determinar a seqüência de consenso da série de referência são descritos na Seção III e V, acima.

Em outra modalidade, um resíduo aminoácido desestabilizante candidato é um resíduo aminoácido o qual tem uma baixa classificação de consenso. Métodos para determinar uma classificação de consenso de um aminoácido são descritos na Seção III e V, acima. Em uma modalidade, o resíduo aminoácido desestabilizante candidato é um resíduo aminoácido com uma classificação de consenso de menos de 0,5 (por exemplo, menos de 0,4, menos de 0,3, menos de 0,2, ou menos de 0,1). Em uma modalidade preferencial, o resíduo aminoácido desestabilizante candidato é um resíduo aminoácido com uma classificação de consenso de menos de 0,3.

Etapa n° *2:* Seleção de Mutações Estabilizantes Candidatas

Tendo identificado um ou mais aminoácidos desestabilizantes

candidatos, a próxima etapa é selecionar uma ou mais mutações estabilizantes candidatas para cada aminoácido desestabilizante. A biblioteca de scFv pode ser então criada para incluir uma molécula scFv candidata representativa para cada mutação estabilizante candidata que é selecionada.

Em algumas modalidades, cada variante de aminoácido natural de um aminoácido desestabilizante em particular é selecionada como uma mutação estabilizante candidata (isto é, 19 mutações estabilizantes candidatas para cada aminoácido desestabilizante).

Em modalidades mais preferenciais, um subgrupo das variantes de aminoácidos naturais de um aminoácido desestabilizante em particular são selecionados como mutações estabilizantes candidatas (isto é, 1 a 18 aminoácidos estabilizantes candidatos para cada aminoácido desestabilizante). Em uma modalidade, o subgrupo de mutações estabilizantes candidatas compreende substituições com um aminoácido de uma classe definida, por exemplo, aminoácidos hidrofóbicos, aminoácidos básicos, aminoácidos hidrofílicos, aminoácidos carregado, ou aminoácidos estericamente tendenciosos (quer pequenos ou grandes).

Em uma modalidade, um subgrupo de substituições de mutações estabilizantes candidatas com aminoácidos as quais estão presentes em altas freqüências em uma posição correspondente à do aminoácido desestabilizante dentro de uma série de referência de seqüências homólogas de região variável. Em modalidades preferenciais, uma mutação estabilizante candidata compreende substituição com um aminoácido que está presente dentro do banco de dados em uma freqüência de mais de 10%. Em modalidades mais preferenciais, o aminoácido está presente em uma freqüência mais de 15%. Em modalidades ainda mais preferenciais, o aminoácido está presente em uma freqüência de mais de 20%. Em modalidades ainda mais preferenciais, o aminoácido está presente em uma freqüência de mais de 25%.

Em outra modalidade, uma mutação estabilizante candidata é uma substituição com o aminoácido de consenso (isto é, o resíduo mais freqüente) encontrado na posição do aminoácido desestabilizante dentro da

série de referência.

Em outra modalidade, um subgrupo de mutações estabilizantes candidatas compreende substituições com todo aminoácido que é encontrado dentro de uma série de referência de seqüências homólogas de região variável na posição do aminoácido desestabilizante. Por conseguinte, uma biblioteca de scFv pode ser então criada para incluir uma molécula scFv candidata típica para cada mutação estabilizante candidata que é representada na série de referência.

Em outras modalidades, o subgrupo de mutações estabilizantes candidatas pode ser identificado ou priorizado para triagem por uma análise (por exemplo, inspeção visual ou análise computacional) de uma estrutura tridimensional ou modelo de uma região variável da molécula scFv. A estrutura tridimensional de um polipeptídeo influencia sua estabilidade e atividade biológica, e a estrutura pode ser determinada ou prevista em uma série de modos. Estrutura terciária pode ser prevista usando construção modelo de estruturas tridimensionais de uma ou mais proteínas homólogas (ou complexos de proteínas) que têm uma estrutura tridimensional conhecida. Cristalografia de raios X é talvez o melhor modo conhecido para determinar a estrutura protéica (por conseguinte, o termo "estrutura de cristal" pode ser usado ao invés do termo "estrutura"), mas também podem ser feitas estimativas usando dicroismo circular, dispersão luminosa, ou medindo a absorção e a emissão de energia radiante. Outras técnicas úteis incluem difração de neutrons, ressonância magnética nuclear (NMR), e modelagem por homologia. Todos estes métodos são conhecidos por aqueles com conhecimento regular da técnia, e foram bem descritas em livros textos de rotina (vide, por exemplo, *Physical Chemistry*, 4th Ed., W.J. Moore, Prentiss-Hall, N.J., 1972, ou *Physical Biochemistry*, K.E. Van Holde, Prentiss-Hall, N.J., 1971)) e numerosas publicações. Quaisquer destas técnicas podem ser realizadas para determinar a estrutura de uma molécula compreendendo uma região variável de uma molécula scFv (por exemplo, um anticorpo, um Fab, ou a própria molécula scFv).

A estrutura da região variável pode ser modelada *in silico*. Por

exemplo, a compatabilidade de uma mutação estabilizante candidata com a estrutura tridimensional pode ser analisada modelando computacionalmente a substituição de uma mutação desestabilizante com uma mutação estabilizante candidata. A mutação estabilizante candidata pode ser selecionada para inclusão na biblioteca de scFv se for compatível com a estrutura total da molécula scFv. Em uma modalidade, a mutação estabilizante candidata pode ser selecionada se não perturbar o dobramento natural ou conformação da região variável da molécula scFv ou uma ou mais regiões determinantes de complementaridade (CDRs) da mesma. Em outra modalidade, a mutação estabilizante candidata pode ser selecionada se não perturbar a capacidade da região variável para formar uma interface VL/VH nativa.

Em algumas modalidades, uma mutação estabilizante candidata pode ser selecionada aplicando uma técnica de recondensação de cadeia lateral a uma estrutura (por exemplo, a estrutura de cristal ou modelo) da região variável. Em um cálculo de recondensação de cadeia lateral, os resíduos estabilizantes candidatos podem ser modificados computacionalmente, e a estabilidade dos mutantes resultantes é avaliada computacionalmente. O cálculo de recondensação de cadeia lateral gera uma lista classificada dos mutantes que têm estabilidade alterada (isto é, energia intramolecular alterada). O número de proteínas mutantes que é avaliado computacionalmente pode ser muito grande, uma vez que toda posição de aminoácido variável pode ser mutada em todos os 20 aminoácidos de rotina. Algoritmos computacionais típicos usados para classificar os resultados da análise computacional incluem eliminação de ponta morta e verdadeiros algoritmos de pesquisa (vide por exemplo, Lasters *et al.* (*Protein Eng.* 8:815-822, 1995), Looger e Hellinga (*J. Mol. Biol.* 307:429-445, 2001), e Dahiyat e Mayo (*Protein Sci.* 5:895-903, 1996)). Por conseguinte, uma biblioteca de scFv pode ser então criada para incluir uma molécula scFv candidata típica para cada uma das mutações estabilizantes candidatas de maior classificação na lista classificada de mutações geradas pelo cálculo de recondensação de cadeia lateral. Em algumas modalidades, no mínimo a mutação de maior classificação é selecionada (por exemplo, são selecionadas as mutações de maior classi-

ficação, as duas de maior classificação, as três de maior classificação, as quatro de maior classificação, ou as cinco de maior classificação).

b. Construção de Bibliotecas de scFv

Tendo determinado as mutações estabilizantes candidatas para incluir na biblioteca de scFv, pode-se usar qualquer um de uma variedade de métodos disponíveis para produzir moléculas scFv candidatas compreendendo as mutações. Os polipeptídeos reefridos podem ser produzidos, por exemplo, por meio de métodos recombinantes. Além disso, devido à degeneração do codigo genético, uma variedade de seqüências de ácido nucléico podem ser usadas para codificar cada scFv desejado.

Métodos reconhecidos na técnica típicos para preparar uma molécula de ácido nucléico codificando uma molécula scFv candidata incluem, mas não estão limitados a, preparação por mutagênese sítio-dirigida (ou mediada por oligonucleotídeo) , mutagênese por PCR, e mutagênese de cassete de um DNA preparado anteriormente codificando a scFv candidata.

Mutagênese sítio-dirigida é um método preferencial para preparar variantes de substituição. Esta técnica é de conhecimento geral na técnica (vide, por exemplo, Carter et al. Nucleic Acids Res. 13:4431-4443 (1985) e Kunkel et al., Proc. Natl. Acad. Sci. USA 82:488 (1987)). Em resumo, ao realizar mutagênese sítio-dirigida de DNA, o DNA parental é alterado primeiro hibridizando um oligonucleotídeo codificando a mutação desejada para um único filamento de similar DNA parental. Depois de hibridização, uma DNA polimerase é usada para sintetizar um segundo filamento inteiro, usando o oligonucleotídeo hibridizado como um iniciador, e usando o único filamento do DNA parental como um modelo. Portanto, o oligonucleotídeo codificando a mutação desejada é incorporado no DNA de filamento duplo resultante, o qual pode ser então ligado em um plasmídeo ou outro vetor adequado.

Mutagênese por PCR também é adequada para preparar moléculas scFv candidatas. Vide Higuchi, in PCR Protocols, pp.177-183 (Academic Press, 1990); e Vallette et al., Nuc. Acids Res. 17:723-583 (1989). Em resumo, quando pequenas quantidades de DNA modelo são usadas como

material de partida em uma reação de PCR, iniciadores que diferem ligeiramente em seqüência da região correspondente em uma DNA modelo podem ser usados para gerar quantidades relativamente grandes de um fragmento de DNA específico que difere da seqüência modelo somente nas posições onde os iniciadores diferem do modelo. Os iniciadores de PCR também podem ser criados para incorporar sítios de restrição, de tal modo que o produto de DNA da reação de PCR pode ser então diretamente ligado em um plasmídeo ou outro vetor adequado.

Outro método para preparar variantes, mutagênese de cassete, se baseia na técnica descrita por Wells et al., Gene 34:315-323 (1985). O material de partida é o plasmídeo (ou outro vetor) compreendendo o DNA do polipeptídeo de partida a ser mutado. O um ou mais códons no DNA parental a ser mutado são identificados. Deve haver um único sítio de endonuclease de restrição sobre cada lado do um ou mais sítios de mutação identificados. Se não existirem similares sítios de restrição, estes podem ser gerados usando o método de mutagênese mediada por oligonucleotídeo descrito acima para introduzir os mesmos em localizações apropriadas no DNA de polipeptídeo de partida. O DNA de plasmídeo é cortado nestes sítios para linearizá-lo. Um oligonucleotídeo de filamento duplo codificando a seqüência do DNA entre os sítios de restrição mas contendo a uma ou mais mutações desejadas é sintetizado usando procedimentos de rotina, em que os dois filamentos do oligonucleotídeo são sintetizados separadamente e em seguida hibridizados juntos usando técnicas de rotina. Este oligonucleotídeo de filamento duplo é referido como o cassete. Este cassete é criado para ter extremidades 5' e 3' que são compatíveis com as extremidades do plasmídeo linearizado, de tal modo que pode ser diretamente ligado ao plasmídeo.

Ácidos nucléicos típicos para cda uma das moléculas scFv candidatas podem ser geradas pelos métodos acima. Os ácidos nucléicos podem ser então clonados em vetores de expressão para formar uma biblioteca de vetores de expressão. Células hospedeiras podem ser então transformadas com a biblioteca de vetores resultante, e as células hospedeiras cultivadsa sob as condições apropriadas de modo a expressar cada molécula

scFv candidata.

c. Métodos de Triagem

Uma biblioteca de scFv a invenção pode ser triada em uma prova (por exemplo, uma prova de alta produtividade total) para identificar moléculas scFv candidatas com estabilidade protéica desejada. Uma prova similar pode empregar quaisquer dos métodos para avaliar estabilidade protéica descritos na Seção VI, acima. Um método particularmente preferencial é uma prova de ensaio térmico. Os métodos de ensaio referidos geralmente envolvem comparar a estabilidade térmica de uma molécula scFv candidata com a de um controle adequado e selecionar a molécula scFv candidata se a estabilidade térmica for mais do que a do controle. Controles adequados típicos incluem moléculas scFv convencionais, por exemplo, uma molécula scFv (Gly4Ser)3. Moléculas scFv candidatas podem ser selecionadas se tiverem uma estabilidade térmica que é de mais de cerca de 0,1, cerca de 0,25, cerca de 0,5, cerca de 0,75, cerca de 1, cerca de 1,25, cerca de 1,5, cerca de 1,75, cerca de 2, cerca de 3, cerca de 4, cerca de 5, cerca de 6, cerca de 7, cerca de 8, cerca de 9, ou cerca de 10 graus Celsius do que a do controle. Em uma modalidade exemplar, uma molécula scFv candidata é selecionada se tiver uma estabilidade térmica que é de mais de cerca de 3 graus Celsius do que a do controle.

Bibliotecas de scFv podem ser apresentadas em diferentes formatos de ensaio. Por exemplo, moléculas scFv podem ser apresentadas em solução (por exemplo, Houghten (1992) *Biotechniques* 13:412-421), sobre contas (Lam (1991) *Nature* 354:82-84), chips (Fodor (1993) *Nature* 364:555-556), bactérias (Ladner USP 5,223,409), esporos (Ladner USP '409), ou sobre fago (Scott e Smith (1990) *Science* 249:386-390); (Devlin (1990) *Science* 249:404-406); (Cwirla *et al.* (1990) *Proc. Natl. Acad. Sci.* 87:6378-6382); (Felici (1991) *J. Mol. Biol.* 222:301-310); (Ladner acima.).

Em algumas modalidades exemplares, as moléculas scFv candidatas são avaliadas em um formato de solução. Em uma modalidade, cada amostra compreende uma alíquouta de uma solução tendo moléculas scFv candidatas com a mesma mutação ou ou mutações estabilizantes candida-

tas. Uma solução similar pode ser gerada isolando colônias de células hospedeiras individuais de uma biblioteca de células hospedeiras transformadas com biblioteca de plasmídeo de expressão e cultivar a colônia de células hospedeiras em um vaso apropriado sob condições as quais facilitam expressão da molécula scFv. Em uma modalidade, a molécula scFv candidata podem ser purificada a partir da célula hospedeira e ressolubilizada em uma solução de ensaio apropriada. Em uma modalidade mais preferencial, a molécula scFv candidata é fundida a uma seqüência de peptídeo de sinal clivável de tal modo que a molécula scFv é secretada pela célula hospedeira no meio de cultura de células hospedeiras. Em ainda outras modalidades preferenciais, a célula hospedeira podem ser cultivada sob condições tais que molécula scFv candidata é liberada no meio, junto com proteínas de células hospedeiras.

Pode ser desejável automatizar quaisquer dos formatos de ensaio acima. Por exemplo, pode ser empregada robótica para isolar cada membro da biblioteca de scFv (por exemplo, como colônias de células hospedeiras individuais) em rápida sucessão de modo que podem ser avaliadas em vasos separados. Exemplos de vasos de ensaio incluem lâminas de microtítulo (por exemplo, lâminas de microtítulo de 96 cavidades), tubos de ensaio, e tubos de microcentrífuga.

d. Otimização Adicional de Moléculas scFv Estabilizadas

Uma molécula scFv estável identificada pelos métodos de triagem acima pode ser remodelada e adicionalmente otimizada para aumentar adicionalmente sua estabilidade protéica. Portanto, as etapas descritas acima podem ser repetidas com uma molécula scFv estável identificada em uma rodada de otimização inicial. Alternativamente, as mutações estabilizantes de duas ou mais moléculas scFv estabilizadas podem ser combinadas em uma única molécula scFv para aumentar adicionalmente a estabilidade protéica.

Em algumas modalidades, a molécula scFv estável identificada pelos métodos da invenção pode ser adicionalmente otimizada. Por exemplo, podem ser feitas uma ou mais das alterações adicionais seguintes. Em

uma modalidade, uma molécula scFv estabilizada pode ser adicionalmente estabilizada introduzindo uma ligação dissulfeto a qual liga um aminoácido no domínio VL com um aminoácido no domínio VH. Ligações dissulfeto típicas incluem quaisquer das ligações dissulfeto descritas na Seção II acima. Uma ligação dissulfeto particularmente preferencial é VH44-VL100.

Em outras modalidades, a molécula scFv estável identificada pelos métodos da invenção é adicionalmente otimizada introduzindo um ligante de scFv com uma extensão ou composição otimizada. Ligantes de scFv típicos são descritos na Seção II acima. Um ligante de scFv particularmente preferencial é $(Gly_4Ser)_4$.

Em outra modalidade, a molécula scFv estável identificada pelos métodos da invenção é adicionalmente otimizada introduzindo uma mutação estabilizante em no mínimo um entre o domínio VH ou VL.

IX. Métodos para Estabilizar Moléculas de Ligação

Em outros aspectos, a invenção proporciona métodos para melhorar as propriedades de estabilidade de moléculas de ligação. Estes métodos geralmente envolvem incorporar ou anexar uma molécula scFv estabilizada da invenção à molécula de ligação. Surpreendentemente, conforme mostrado nos exemplos de trabalho aqui, neste Pedido de patente, moléculas scFv da invenção não são somente estáveis por si só, também conferem aprimorada estabilidade a moléculas de ligação nas quais são incorporadas. Por conseguinte, os métodos da invenção proporcionam um meio conveniente e confiável para melhorar a estabilidade de moléculas de ligação comercialmente valiosas para as quais a fabricação em larga escala freqüentemente é limitada por estabilidade protéica pobre (por exemplo, anticorpos multiespecíficos (por exemplo, anticorpos biespecíficos) e outros anticorpos modificados).

Moléculas scFv estabilizadas podem ser incorporadas em moléculas de ligação usando metodologia de conjugação de proteína que é conhecida na técnica. Em uma modalidade, a scFv estabilizada é fundida diretamente a um N- ou C-término de um polipeptídeo, por exemplo, uma molécula de anticorpo. Em outra modalidade, um ligante não-peptídico é empre-

gado para ligar a scFv estabilizada a um N- ou C-término de um polipeptídeo. Em ainda outras modalidades, um peptídeo de ligação é usado para ligar a scFv estabilizada a um polipeptídeo. Em uma modalidade exemplar, o peptídeo de ligação é um peptídeo rico em gly/ser curto. Peptídeos riscos em Gly/Ser típicos são listados na Tabela 1 abaixo. Outros peptídeos de ligação típicos são conhecidos na técnica (vide, por exemplo, os Pedidos de patente Internacional Nos.WO 2005/000898 e WO 2005/000899). Em uma modalidade, um scFv estabilizado da invenção é encadeado à extremidade de C-terminal de uma molécula de ligação, por exemplo, uma molécula de anticorpo, usando um ligante $S(G_4S)_3$. Em outra modalidade, um scFv estabilizado da invenção é encadeado à extremidade de N-terminal de uma molécula de ligação, por exemplo, uma molécula de anticorpo, usando um ligante $(G_4S)_5$.

Tabela 1: Peptídeos De Ligação

Ligante	SEQ ID NO	Seqüência de DNA ou de Aminoácido
$(Gly_4Ser)_5$	132	5'-GGCGGTGGAGGGTCCGGTGGAGGGGGCTCTGGA GGGGGCGGTTCAGGGGGCGGTGGATCGGGGGG AGGTGGCTCC-3'
	133	GGGGSGGGGSGGGGSGGGGSGGGGS
$Ser(Gly_4Ser)_3$	42	5'-TCCGGCGGGGGTGGATCCGGTGGAGGGGGCTCC GGCGGTGGCGGGTCC-3'
	43	SGGGGSGGGGSGGGGS

Em uma modalidade, no mínimo uma molécula scFv estabilizada é anexado a uma molécula de anticorpo para produzir uma molécula biespecífica. Em outra modalidade, duas moléculas scFv estabilizadas são anexadas a uma molécula de anticorpo para produzir uma molécula biespecífica.

Em algumas modalidades, moléculas de ligação estabilizadas da invenção resultam em produção aumentada comparadas com moléculas scFv convencionais ou moléculas de ligação compreendendo moléculas scFv convencionais. Métodos para avaliar rendimento são descritos na Se-

ção VI, acima. Em uma modalidade, uma molécula de ligação estabilizada produzido pelo métodos da invenção tem um aumento na produção de no mínimo 1% em relação à molécula de ligação não estabilizada. Em outras modalidades, a molécula de ligação estabilizada tem um aumento na produção de no mínimo 2%, no mínimo 5%, no mínimo 10%, no mínimo 20%, no mínimo 30%, no mínimo 50%, no mínimo 75%, ou no mínimo 100%, em relação à molécula de ligação não estabilizada.

Em algumas modalidades, moléculas de ligação da invenção resultam em reduzida agregação comparadas com moléculas scFv convencionais ou moléculas de ligação compreendendo moléculas scFv convencionais. Métodos para avaliar agregação são descritos na Seção VI, acima. Em uma modalidade, uma molécula de ligação estabilizada produzida pelo métodos da invenção tem uma redução na agregação de no mínimo 1% em relação à molécula de ligação não estabilizada. Em outras modalidades, a molécula de ligação estabilizada tem uma redução na agregação de no mínimo 2%, no mínimo 5%, no mínimo 10%, no mínimo 20%, no mínimo 30%, no mínimo 50%, no mínimo 75%, ou no mínimo 100%, em relação à molécula de ligação não estabilizada.

Em outras modalidades, moléculas de ligação da invenção resultam em aumentada estabilidade de longo termo ou vida de armazenamento comparada com moléculas scFv convencionais ou moléculas de ligação compreendendo moléculas scFv convencionais. Métodos para avaliar a vida de armazenamento incluem % de perda ou % de proteólise conforme descrito na Seção VI, acima. Em uma modalidade, uma molécula de ligação estabilizada produzida pelo métodos da invenção tem um aumento em vida de armazenamentode no mínimo 1 dia em relação à molécula de ligação não estabilizada. Isto significa que a preparação de moléculas de ligação tem substancialmente a mesma quantidade de moléculas de ligação estáveis que a presente no dia anterior. Em outras modalidades, a molécula de ligação estabilizada tem um aumento na vida de armazenamento de no mínimo 2 dias, no mínimo 5 dias, no mínimo 1 semana, no mínimo 2 semanas, no mínimo 1 mês, no mínimo 2 meses, no mínimo 6 meses, ou no mínimo 1

ano, em relação à molécula de ligação não estabilizada.

Em outras modalidades, moléculas de ligação da invenção resultam em aprimorada estabilidade por exemplo, quando expressadas em um tipo de célula hospedeira em particular, comparadas com moléculas scFv convencionais ou moléculas de ligação compreendendo moléculas scFv convencionais. Em modalidades exemplares, os métodos da invenção resultam na produção de moléculas de ligação as quais têm aumentada estabilidade (por exemplo, aumentado rendimento) quando a molécula de ligação é expressada em uma célula hospedeira, por exemplo, uma célula hospedeira bacteriana ou eucariótica (por exemplo, levedura ou mamífera). Células hospedeiras mamíferas típicas incluem células de ovário de hamster chinês (CHO), células HELA (carcinoma cervical humano), células CVI (linhagem renal de macaco) células COS (um derivado de CVI com antígeno SV40 T), células R1610 (fibroblasto de hamster chinês), células BALBC/3T3 (fibroblasto de camundongo), células HAK (linhagem renal de hamster), células SP2/O (mieloma de camundongo), células BFA-1c1BPT (células endoteliais bovinas), células RAJI (linfócito humano) e células 293 (rim humano). Em uma modalidade preferencial, duas moléculas scFv estabilizadas são anexadas a uma molécula de anticorpo para criar uma molécula biespecífica estabilizada para secreção em células CHO.

Em outras modalidades, células hospedeiras capazes de expressar moléculas de ligação estabilizadas podem ser triadas para selecionar isolados celulares únicos que são capazes de expressar altos níveis de moléculas de ligação estabilizadas solubilizadas e adequadamente dobradas (por exemplo, moléculas de ligação apresentando menos de 10% de agregação). Os métodos referidos podem empregar técnicas de classificação celular ativada por fluorescência (FACS) (vide, por exemplo, Brezinky *et al.*, J Immunol Meth (2003). 277:141-155). Em uma modalidade, o isolado celular único é adaptado para condições livres de soro para estabelecer uma linhagem de células produtoras estáveis. A linhagem de células produtoras estáveis pode ser então cultivada para facilitar fabricação em larga escala de uma molécula de ligação estabilizada da invenção.

Em outras modalidades, os métodos da invenção são empregados para melhorar a estabilidade de uma molécula de ligação que é expressada a partir de uma célula hospedeira em um grande volume de meios de cultura. Em modalidades exemplares, os métodos da invenção resultam em uma estabilidade aumentada (por exemplo, aumentado rendimento) quando a molécula de ligação é expressada em no mínimo 1 litro de meio de cultura. Em outras modalidades, os métodos da invenção são usados para produzir uma molécula de ligação estabilizada a qual tem uma aumentada estabilidade (por exemplo, rendimento) quando expressada a partir de uma célula hospedeira em no mínimo 2 litros, no mínimo 10 litros, no mínimo 20 litros, no mínimo 50 litros, no mínimo 75 litros, no mínimo 100 litros, no mínimo 200 litros, ou no mínimo 500 litros de meio de cultura. Em uma modalidade exemplar, os métodos da invenção são usados para produzir no mínimo 10 mg de uma molécula de ligação estabilizada para cada litro de meio de cultura.

X. Moléculas de Ligação Estabilizadas Compreendendo Moléculas scFv Estabilizadas

Em uma modalidade, uma molécula de ligação estabilizada da invenção é uma proteína de fusão. Por exemplo, uma molécula scFv estabilizada da invenção pode ser encadeada a uma segunda molécula scFv ou uma molécula não-scFv. Em uma modalidade, uma molécula não-scFv à qual uma molécula scFv estabilizada da invenção pode ser encadeada proporciona no mínimo um sítio de ligação adicional. Por exemplo, sítios de ligação típicos que podem ser incluídos em uma molécula de ligação da invenção incluem: a porção de ligação de receptor de um ligante, a porção de ligação de ligante de um receptor, a porção de ligação de substrato de uma enzima, a porção de ligação de enzima de um substrato, ou uma ou mais porções de ligação de antígeno de um anticorpo. Moléculas scFv podem ser encadeadas, por exemplo, a moléculas de anticorpos para formar moléculas de anticorpo modificado ou a outros polipeptídeos para formar proteínas de fusão. Alguns exemplos são descritos abaixo.

A. Moléculas de Anticorpos Modificados

Em uma modalidade, uma molécula scFv estabilizada da invenção é encadeada a um anticorpo ou fragmento do mesmo para formar uma proteína de ligação estabilizada a qual é um anticorpo modificado. Em outra modalidade, um scFv estabilizado da invenção é encadeado a um anticorpo modificado, isto é, molécula de anticorpo que não ocorre naturalmente, para formar uma proteína de ligação estabilizada. Construtos de anticorpos modificados preferenciais são descritos em mais detalhes abaixo. Conforme usado aqui, neste Pedido de patente, o termo "anticorpo modificado" inclui formas sintéticas de anticorpos os quais são alterados de tal modo que não ocorem naturalmente, por exemplo, anticorpos que compreendem no mínimo duas porções de cadeia pesada mas não duas cadeias pesadas completas (tais como, anticorpos ou minibodies de domínio eliminado); formas multiespecíficas de anticorpos (por exemplo, biespecíficos, triespecíficos, e etc.) alteradas para ligar a dois ou mais antígenos diferentes ou a epítopes diferentes sobre um único antígeno). Além disso, o termo "anticorpo modificado" inclui formas multivalentes de anticorpos (por exemplo, trivalentes, tetravalentes, e etc., anticorpos que ligam a três ou mais cópias do mesmo antígeno).

Será entendido que ao discutir as moléculas de ligação da invenção, as especificidades de ligação típicas descritas aqui, neste Pedido de patente, podem ser conferidas por uma molécula scFv estabilizada da invenção, uma molécula de ligação compreendendo uma molécula scFv da invenção ou ambas.

Em uma modalidade, as proteínas de ligação estabilizadas da presente invenção podem ser imunorreativas com um ou mais antígenos tumorais ou antígenos associados com distúrbios imunes. Por exemplo, para distúrbios neoplásicos, o sítio de ligação (isto é, a região variável ou fragmento imunorreativo ou recombinante do mesmo) das moléculas de ligação relevadas liga a um antígeno associado a tumor selecionado no sítio da malignidade. Similarmente, em uma modalidade, uma molécula de ligação pode ligar a no mínimo um marcador selecionado presente sobre células imunes.

Dado o número de antígenos reportados associados com neoplasias e distúrbios imunes, e o número de antígenos relacionados, aqueles versados na técnica reconhecerão que as moléculas de ligação presentemente descritas podem portanto ser derivadas de qualquer um de uma série de anticorpos inteiros. Mais geralmente, polipeptídeos úteis na presente invenção podem ser obtidos ou derivados de qualquer anticorpo (incluindo aqueles previamente reportados na literatura) que reage com uma molécula ou marcador associado com a condição selecionada. Além disso, o anticorpo parental ou precursor, ou fragmento do mesmo, usado para produzir as moléculas de ligação estabilizadas da invenção pode ser murino, humano, quimérico, humanizado, de primata não-humano ou primatizado.

Conforme usado aqui, neste Pedido de patente, "antígenos associados com tumor" inclui antígenos os quais são geralmente associados com células tumorais, por exemplo, expressados sobre células tumorais. Mais geralmente, antígenos associados com tumor compreendem antígenos que proporcionam a localização de anticorpos imunorreativos em uma células neoplásica independente de sua expressão sobre células não-malignas. Os antígenos referidos podem ser relativamente tumor específicos, por exemplo, limitados em expressão à superfície de células malignas. Alternativamente, os antígenos referidos podem ser encontrados tanto sobre células malignas quanto não-malignas. Por exemplo, CD20 é um antígeno pan B que é encontrado sobre a superfície tanto de células B malignas quanto de células não-malignas que se comprovaram como sendo um alvo extremamente eficaz para anticorpos imunoterapêuticos para o tratamento de linfoma não-Hodgkin. A este respeio, antígenos de células pan T tais como CD2, CD3, CD5, CD6 e CD7 também compreendem antígenos associados com tumor dentro do significado da presente invenção. Ainda outros antígenos associados com tumor típicos compreendem mas não estão limitados a MAGE-1, MAGE-3, MUC-1, HPV 16, HPV E6 & E7, TAG-72, CEA, L6-Antigen, CD19, CD22, CD37, CD52, HLA-DR, EGF Receptor e HER2 Receptor. Em muitos casos anticorpos imunorreativos para cada um destes antígenos foram reportados na literatura. Aqueles versados na técnica reco-

nhecerão que cada um destes anticorpos pode servir como um precursor para anticorpos da invenção de acordo com a presente invenção.

Em uma modalidade, as moléculas de ligação estabilizadas da presente invenção preferencialmente associam com, e ligam a, antígenos associados com tumor ou imunes conforme descrito acima. Por conseguinte, conforme será discutido em alguns detalhes abaixo as moléculas de ligação estabilizadas da presente invenção podem ser derivadas, geradas ou fabricadas a partir de qualquer uma de uma série de anticorpos que reagem com antígenos associados com tumor. Em algumas modalidades as moléculas de ligação estabilizadas da invenção são anticorpos com domínio eliminado que são derivados usando técnicas comuns de engenharia genética por meio das quais no mínimo uma porção de um ou mais domínios de região constante são eliminados ou alterados de modo a proporcionar as características bioquímicas desejadas tais como meia-vida sérica reduzida. Mais particularmente, uma pessoa versada na técnica pode prontamente isolar a seqüência genética correspondente às regiões variáveis e/ou constantes da molécula de ligação estabilizada em questão e eliminar ou alterar os nucleotídeos apropriados para proporcionar polipeptídeos da invenção para uso como subunidades monoméricas de acordo com a presente invenção.

Anticorpos previamente reportados que reagem com antígenos associados com tumor podem ser estabilizados conforme descrito aqui, neste Pedido de patente, para proporcionar as moléculas de ligação estabilizadas da presente invenção. Anticorpos típicos que podem ser usados para proporcionar regiões de ligação de antígeno para gerar ou derivar as moléculas de ligação estabilizadas reveladas incluem, mas não estão limitados a 2B8 e C2B8 (Zevalin® e Rituxan®, IDEC Pharmaceuticals Corp., San Diego), Lym 1 e Lym 2 (Techniclone), LL2 (Immunomedics Corp., New Jersey), HER2 (Herceptin®, Genentech Inc., South San Francisco), B1 (Bexxar®, Coulter Pharm., San Francisco), Campath® (Millennium Pharmaceuticals, Cambridge), abagovomab (Menarini, Italy), CEA-Scan™ (Immunomedics, Morris Plains, NJ), capromab (Prostascint®, Cytogen Corp.), edrecolomab (Panorex®, Johnson & Johnson, New Brunswick, NJ), igovomab (CIS Bio

Intl., France), mitumomab (BEC2, Imclone Systems, Somerville, NJ), nofetumomab (Verluma®, Boehringer Ingleheim, Ridgefield, CT), OvaRex (Altarex Corp., Waltham, MA), satumomab (Onoscint®, Cytogen Corp.), cetuximab (Erbitux®, Imclone Systems, New York, NY), bevacizumab (AVASTIN®, Genentech Inc., S. San Francisco, CA), apolizumab (REMITOGEN ™, Protein Design Labs, Fremont, CA), labetuzumab (CEACIDE ™, Immunomedics Inc., Morris Plains, NJ) , pertuzumab (OMNITARG ™, Genentech Inc., S. San Francisco, CA), MB1, BH3, B4, B72.3 (Cytogen Corp.), CC49 (National Cancer Institute) e 5E10 (University of Iowa). Outros sítios de ligação que podem ser incorporados nas moléculas de ligação em questão incluem os encontrados em: Orthoclone OKT3 (CD3), ReoPro (GpIIb/gIIa), Zenapax (C25), Remicade (TNF-a), Simulect (CD25), Synagis (RSV), Mylotarg (CD33), e Campath (CD52). Em modalidades preferenciais, as moléculas de ligação estabilizadas da presente invenção ligarão aos mesmos antígenos associados com tumor que os anticorpos enumerados imediatamente acima. Em modalidades particularmente preferenciais, as moléculas de ligação estabilizadas serão derivadas de ou ligar os mesmos antígenos que 2B8, C2B8, CC49 e C5E10 e, ainda mais preferencialmente, carecerão de todo ou parte de um domínio CH2.

Em uma modalidade, uma molécula de ligação da invenção pode ter um ou mais sítios de ligação derivados de um ou mais dos seguintes anticorpos. tositumomab (BEXXAR®), muromonab (ORTHOCLONE®) e ibritumomab (ZEVALIN®), cetuximab (ERBITUX™), rituximab (MABTHERA® / RITUXAN®), infliximab (REMICADE®), abciximab (REOPRO®) e basiliximab (SIMULECT®), efalizumab (RAPTIVA®, bevacizumab (AVASTIN®), alemtuzumab (CAMPATH®), trastuzumab (HERCEPTIN®), gemtuzumab (MYLOTARG®), palivizumab (SYNAGIS®), omalizumab (XOLAIR®), daclizumab (ZENAPAX®), natalizumab (TYSABRI®) e ranibizumab (LUVENTIS®), adalimumab (HUMIRA®) e panitumumab (VECTIBIX®).

Em uma modalidade, uma molécula de ligação estabilizada da invenção liga a CD23 (Patente dos Estados Unidos Nº 6,011,138). Em uma modalidade preferencial, uma molécula de ligação estabilizada da invenção

liga ao mesmo epítope que o anticorpo 5E8. Em outra modalidade, uma molécula de ligação da invenção compreende no mínimo um CDR (por exemplo, no mínimo 1, 2, 3, 4, 5, ou 6 CDRs) de um anticorpo anti-CD23, por exemplo, o anticorpo 5E8.

Em uma modalidade, uma molécula de ligação estabilizada da invenção liga a um receptor de TNF. Em uma modalidade típica, uma molécula de ligação estabilizada da invenção liga a um LTβR. Em outra modalidade exemplar, uma molécula de ligação estabilizada da invenção liga a um TRAIL receptor. Em outra modalidade, uma molécula de ligação estabilizada da invenção compreende no mínimo um CDR (por exemplo, no mínimo 1, 2, 3, 4, 5, ou 6 CDRs) de um anticorpo anti-TRAIL-R2 (por exemplo, 14A2 murino ou quimérico). Em outra modalidade, uma molécula de ligação estabilizada da invenção compreende no mínimo um CDR de um anticorpo anti-LTβR. Exemplos de anticorpos anti-LTβR incluem BKA11, CDH10, BCG6, AGH1, BDA8, CBE11 e BHA10.

Em uma modalidade, uma molécula de ligação estabilizada da invenção liga ao antígeno CRIPTO-I (Pedido de patente internacional Nº WO02/088170A2 ou Pedido de patente internacional Nº WO03/083041A2). Em uma modalidade preferencial, uma molécula de ligação estabilizada da invenção liga ao mesmo epítope que o anticorpo B3F6. Em outra modalidade, uma molécula de ligação estabilizada da invenção compreende no mínimo um CDR de um anticorpo anti-CRIPTO-I, por exemplo, o anticorpo B3F6.

Em uma modalidade, a molécula de ligação estabilizada ligará ao mesmo antígeno associado com tumor que Rituxan®. Rituxan® (também conhecido como, rituximab, IDEC-C2B8 e C2B8) foi o primeiro anticorpo monoclonal aprovado pela agência americana FDA para tratamento de linfoma de células B humano (vide as Patentes dos Estados Unidos Nos. 5.843.439; 5.776.456 e 5.736.137 cada uma das quais é incorporada aqui, a este Pedido de patente, por meio de referência). Y2B8 (2B8 marcado com 90Y; Zevalin®; ibritumomab tiuxetan) é a origem murina de C2B8. Rituxan® é um anticorpo monoclonal anti-CD20 quimérico o qual é inibidor de crescimento e supostamente sensibiliza algumas linhagens celulares de linfoma

para apoptose por agentes quimioterápicos in vitro. O anticorpo liga de modo eficaz complemento humano, tem forte ligação de FcR, e pode destruir de modo eficaz linfócitos humanos in vitro através tanto de mecanismos dependentes de complemento (CDC) quanto dependentes de anticorpo (ADCC) (Reff et al., Blood 83: 435-445 (1994)). Aqueles versados na técnica reconhecerão que variantes diméricas (homodímeros ou heterodímeros) de C2B8 ou 2B8, modificadas de acordo com a presente descoberta, podem ser usadas em formas conjugadas ou não-conjugadas para tratar de modo eficaz pacientes se apresentando com malignidades de CD20+. Mais geralmente, deve ser reiterado que a molécula de ligação estabilizada revelada aqui, neste Pedido de patente, pode ser usada ou em um estado "nu" ou não conjugado ou conjugado a um agente citotóxico para tratar de modo eficaz qualquer um de uma série de distúrbios.

Em outras modalidades preferenciais da presente invenção, a molécula de ligação estabilizada da invenção será derivada de, ou ligará a, o mesmo antígeno associado com tumor que CC49. CC49 liga antígeno associado com tumor humano TAG-72 o qual é associado com a superfície de algumas células tumorais de origem humana, especificamente a linhagem de células tumorais LS174T. LS174T [American Type Culture Collection (aqui, neste Pedido de patente, ATCC) Nº CL 188] é uma variante da linhagem de adenocarcinoma de cólon LS180 (ATCC Nº CL 187). Será adicionalmente reconhecido que foram desenvolvidos numerosos anticorpos monoclonais murinos os quais têm especificidade de ligação para TAG-72. Um destes anticorpos monoclonais, designados B72.3, é uma IgGl murina produzida por hibridoma B72.3 (ATCC Nº HB-8108). B72.3 é um anticorpo monoclonal de primeira geração desenvolvido usando um extrato de carcinoma de mama humano como o imunogênio (vide Colcher et al., Proc. Natl. Acad. Sci. (USA), 78: 3199-3203 (1981); e Patentes dos Estados Unidos Nos. 4.522.918 e 4.612.282 cada uma das quais é incorporada aqui, a este Pedido de patente, por meio de referência). Outros anticorpos monoclonais dirigidos contra TAG-72 são designados "CC" (para câncer de cólon). Conforme descrito por Schlom et al. (Patente dos Estados Unidos Nº 5.512.443 a qual é incorpora-

da aqui, a este Pedido de patente, por meio de referência) anticorpos monoclonais CC são uma família de anticorpos monoclonais murinos de segunda geração que foram preparados usando TAG-72 purificado com B72.3. Devido a suas afinidades de ligação relativamente boas para TAG-72, os seguintes anticorpos CC foram depositados na ATCC, com acesso restrito tendo sido solicitado: CC49 (ATCC Nº HB 9459); CC 83 (ATCC Nº HB 9453); CC46 (ATCC Nº HB 9458); CC92 (ATTCC Nº HB 9454); CC30 (ATCC Nº HB 9457); CC11 (ATCC Nº 9455); e CC15 (ATCC Nº HB 9460). U.S.P.N. 5.512.443 ensina adicionalmente que os anticorpos revelados podem ser alterados em sua forma quimérica substituindo, por exemplo, domínios de regiões constantes humanas (Fc) para regiões constantes de camundongo por técnicas de DNA recombinante conhecidas na técnica. Além de revelar anticorpos anti-TAG-72 murinos e quiméricos, Schlom et al. também produziram variantes de um anticorpo CC49 humanizado conforme revelado na publicação de patente Nº PCT/US99/25552 e construtos de cadeia única conforme revelado na Patente dos Estados Unidos. Nº 5.892.019 cada uma das quais também é incorporada aqui, a este Pedido de patente, por meio de referência. Aqueles versados na técnica reconhecerão que cada um dos anticorpos, construtos ou recombinantes precedentes, e variações dos mesmos, podem ser modificados e usados para proporcionar polipeptídeos de acordo com a presente invenção.

Além dos anticorpos anti-TAG-72 discutidos acima, vários grupos também reportaram a construção e caracterização parcial de anticorpos CC49 e B72.3 com domínio eliminado (por exemplo, Calvo et al. Cancer Biotherapy, 8(1): 95-109 (1993), Slavin-Chiorini et al. Int. J. Cancer 53:97-103 (1993) e Slavin-Chiorini et al. Cancer. Res. 55:5957-5967 (1995).

Ainda outras modalidades preferenciais da presente invenção compreendem sítios de ligação que são derivados de ou ligam ao mesmo antígeno associado com tumor que C5E10. Conforme determinada no requerimento co-pendente Nº 09/104,717, C5E10 é um anticorpo que reconhece um determinante de glicoproteína de aproximadamente 115 kDa que parece ser específico para linhagens de células tumorais de próstata (por

exemplo, DU145, PC3, ou ND1). Portanto, em combinação com a presente invenção, moléculas de ligação estabilizadas (por exemplo, anticorpos com domínio CH2 eliminado) que ligam especificamente ao mesmo antígeno associado com tumor reconhecido por anticorpos C5E10 podem ser produzidos e usados em uma foram conjugada ou não-conjugada para o tratamento de distúrbios neoplásicos. Em modalidades particularmente preferenciais, a molécula de ligação estabilizada será derivada ou compreenderá toda ou parte da região de ligação de antígeno do anticorpo C5E10 conforme secretado a partir da linhagem de células de hibridoma tendo número de acesso da ATCC PTA-865. A molécula de ligação estabilizada resultante pode ser então conjugada a um radionuclídeo conforme descrito abaixo e administrada a um paciente sofrendo de câncer de próstata de acordo com os métodos aqui, neste Pedido de patente.

Em algumas modalidades, as moléculas de ligação estabilizadas da invenção têm no mínimo uma das especificidades de ligação descritas aqui, neste Pedido de patente, por exemplo, na Seção C abaixo. Em outra modalidade, uma molécula de ligação estabilizada da invenção pode ligar a uma molécula alvo de interesse, por exemplo, descrita aqui, neste Pedido de patente, por exemplo, na seção B ou C, abaixo.

Anticorpos previamente reportados que reagem com antígenos associados com distúrbios de células imunes (por exemplo, distúrbios de células B) podem ser estabilizados conforme descrito aqui, neste Pedido de patente, para proporcionar as moléculas de ligação estabilizadas da presente invenção. Anticorpos típicos que podem ser usados para proporcionar regiões de ligação de antígeno para gerar ou derivar as moléculas de ligação estabilizadas reveladas incluem, mas não estão limitadas a um anticorpo anti-TNFα (por exemplo, infliximab (Remicade®, Centocor, Horsham, PA); MAK195-F (Abbott Labs., Abbott Park, IL); adalimumab (Humira®, Abbott Labs, Abbott Park, IL); um anticorpo anti-CD3 (por exemplo, Orthoclone (OKT3®, OrthoBiotech, Bridgewater, NJ); MEDI-500 (Medimmune, Gaithersburg, MD); visilizumab (NUVION®, Protein Design Labs (Fremont, CA, USA)), um anticorpo anti-IgE (por exemplo, omalizumab, XOLAIR® , Genente-

ch, South San Francisco, CA), um anticorpo anti-VLA-4 (por exemplo, TYSABRI®, BiogenIdec, Cambridge, MA), um anticorpo anti-CD147 (por exemplo, ABX-CBL (Abgenix, Fremont, CA)), um anticorpo anti-CD25 (por exemplo, basiliximab, Simulect® (East Hanover, NJ); Inolimomab (OPI, France), um anticorpo anti-CD18 (por exemplo, odulimomab, Antilfa®, Pateur Meriuex, France), anti-NCA90 (por exemplo, sulesomab, Leukoscan®, Immunomedics, Morris Plains, NJ), um anticorpo anti-GpIIb/gIIa (por exemplo, abciximab, ReoPro®, Centocor, Horsham, PA), um anticorpo anti-C25 (por exemplo, Zenapax), um anticorpo anti-CD33 (por exemplo, Mylotarg), e um anticorpo anti-CD25 (por exemplo, alemtuzumab, Campath® (Milleneum Pharmaceuticals, Cambridge, MA). Em modalidades preferenciais, as moléculas de ligação estabilizadas da presente invenção ligarão aos mesmo antígenos associados a células imunes que os anticorpos enumerados imediatamente acima.

Em uma modalidade, o termo, "anticorpo modificado" de acordo com a presente invenção inclui imunoglobulinas, anticorpos, ou fragmentos imunorreativos ou recombinantes dos mesmos, nos quais no mínimo uma fração de um ou mais dos domínios de região constante foi eliminado ou alterado de modo diverso de modo a proporcionar características bioquímicas desejadas tais como a capacidade para dimerizar de modo não-covalente, aumentada capacidade para localizar no sítio de um tumor, ou reduzida meia-vida sérica quando comparado com um anticorpo inteiro, inalterado de aproximadamente a mesma imunogenicidade. Em uma modalidade preferencial, os polipeptídeos da presente invenção são anticorpos com domínio eliminado os quais compreendem uma cadeia de polipeptídeo similar a uma cadeia pesada de imunoglobulina, mas os quais carecem de no mínimo uma porção de um ou mais domínios de cadeia pesada. Em uma modalidade, um domínio inteiro da região constante da proteína de ligação estabilizada será eliminado. Em outra modalidade, todo ou parte do domínio CH2 será eliminado.

Em uma modalidade, as proteínas de ligação estabilizadas da invenção são minibodies. Minibodies são moléculas diméricas compostas de

duas cadeias de polipeptídeo cada uma compreendendo uma molécula scFv estabilizada (um único polipeptídeo compreendendo um ou mais sítios de ligação antigênica, por exemplo, um domínio VL encadeado por um ligante flexível a um domínio VH fundido a um domínio CH3 através de um peptídeo de ligação.

Minibodies podem ser preparados construindo um componente de scFv e componente de peptídeo de ligação-CH3 usando métodos descritos na técnica (vide, por exemplo, a patente dos Estados Unidos Nº 5.837.821 ou o Pedido de patente internacional Nº WO 94/09817A1). Estes componentes podem ser isolados dos plasmídeos separados como fragmentos de restrição e em seguida ligados e reclonados em um vetor apropriado. Montagem apropriada pode ser verificada por digestão de restrição e análise de seqüência de DNA.

Em outra modalidade, pode ser construído um "minibody" tetravalente. Minibodies tetravalentes podem ser construídos na mesma maneira que minibodies, exceto que duas moléculas scFv são encadeadas usando um ligante flexível, por exemplo, tendo uma seqüência de aminoácidos $(G_4S)_4G_3AS$.

Em uma modalidade, anticorpos tetravalentes podem ser produzidos combinando uma seqüência de DNA codificando um anticorpo com uma molécula scFv. Por exemplo, em uma modalidade, estas seqüências são combinadas de tal modo que a molécula scFv é encadeado em seu N-término ao domínio do anticorpo CH3 através de um ligante flexível (por exemplo, um ligante gly/ser tal como $(Gly_4Ser)_3$.

Em outra modalidade um anticorpo tetravalente pode ser produzido fundindo uma molécula scFv estabilizada a um peptídeo de ligação, o qual é fundido a um domínio CH1 para construir uma molécula tetravalente scFv estabilizada – Fab (Coloma e Morrison. 1997. Nature Biotechnology. 15:159; WO 95/09917).

Em uma modalidade uma molécula de ligação estabilizada da invenção compreende um anticorpo tetravalente tetravalente ou biespecífico dentro de um scFv anexado ao N-término da cadeia leve. Em outra modali-

dade da invenção, uma molécula de ligação compreende um anticorpo CH2 tetravalente ou tetravalente biespecífico com domínio eliminado dentro de um scFv anexado ao N-término da cadeia pesada. Em uma modalidade, a fixação da scFv ao N-término resulta em reduzida agregação das moléculas comparadas com moléculas nas quais a scFv é anexada ao término carbóxi.

Anticorpos ou fragmentos dos mesmos para uso em uma molécula de ligação estabilizada da invenção podem ser obtidos usando protocolos reconhecidos na técnica, por exemplo, anticorpos são preferencialmente cultivados em mamíferos por múltiplas injeções subcutâneas ou intraperitoneais do antígeno relevante (por exemplo, antígenos associados com tumor purificados ou células ou extratos celulares compreendendo os antígenos referidos) e um adjuvante. Esta imunização tipicamente provoca uma resposta imune que compreende produção de anticorpos antígeno-reativos a partir de esplenócitos ou linfócitos ativados. Apesar dos anticorpos resultantes poderem ser colhidos do soro animal para proporcionar preparações policlonais, freqüentemente é desejável isolar linfócitos individuais do baço, dos linfonodos ou do sangue periférico para proporcionar preparações homogêneas de anticorpos monoclonais (MAbs). Preferencialmente, os linfócitos são obtidos do baço.

Neste processo bem conhecido (Kohler et al., *Nature*, 256:495 (1975)) os linfócitos de vida relativamente curta, ou mortais, de um mamífero os quais foram injetados com antígeno são fundidos com uma linhagem de células tumorais imortais (por exemplo, linhagem de células de mieloma), deste modo, produzindo células híbridas ou "hibridomas" as quais são tanto imortais quanto capazes de produzir o anticorpo geneticamente codificado da célula B. Os híbridos resultantes são segregados em cepas genéticas únicas por seleção, diluição, e recrescimento com cada cepa individual compreendendo genes específicos para a formação de um único anticorpo. Produzem anticorpos os quais são homogêneos contra um antígeno desejado e, em referência a sua ascendência genética pura, são denominados "monoclonais."

Células de hibridoma preparadas deste modo são semeadas e

cultivadas em un meio de cultura adequado que preferencialmente contém uma ou mais substâncias que inibem o crescimento ou sobrevida das células de mieloma parentais não fundidas. Aqueles versados na técnica reconhecerão que reagentes, linhagens celulares e meios para a formação, seleção e crescimento de hibridomas estão disponíveis comercialmente a partir de uma série de fontes e protocolos padronizados estão bem estabelecidos. Geralmente, o meio de cultura no qual as células de hibridoma estão crescendo é avaliado para produção de anticorpos monoclonais contra o antígeno desejado. Preferencialmente, a especificidade de ligação dos anticorpos monoclonais produzidos por células de hibridoma é determinado por imunoprecipitação ou por um ensaio *in vitro*, tal como um radioimunoensaio (RIA) ou ensaio imunoabsorvente ligado a enzima (ELISA). Depois de serem identificadas células de hibridoma que produzem anticorpos da especificidade desejada, afinidade e/ou atividade, os clones podem ser subclonados por procedimentos de diluição limitante e cultivados por métodos de rotina (Goding, *Monoclonal Antibodies: Principles e Practice*, pp 59-103 (Academic Press, 1986)). Será adicionalmente reconhecido que os anticorpos monoclonais secretados pelos subclones podem ser separados do meio de cultura, fluido de ascite ou soro por procedimentos de purificação convencionais tais como, por exemplo, proteína-A, cromatografia por hidroxilapatita, eletroforese de gel, diálise ou cromatografia por afinidade.

Em outra modalidade, DNA codificando anticorpos monoclonais desejados podem ser prontamente isolados e seqüenciados usando procedimentos convencionais (por exemplo, usando sondas de oligonucleotídeos que são capazes de ligar especificamente a genes codificando as cadeias pesadas e leves de anticorpos murinos). As células de hibridoma isoladas e subclonadas servem como uma fonte preferencial de similar DNA. Uma vez isolado, o DNA pode ser colocado em vetores de expressão, os quais são então transfectados em células hospedeiras procarióticas ou eucarióticas tais como células de *E. coli*, células COS de símios, células de Ovário de Hamster Chinês (CHO) ou células de mieloma que não produzem imunoglobulinas de modo diverso. Mais particularmente, o DNA isolado (o qual pode

ser sintético conforme descrito aqui, neste Pedido de patente) pode ser usado para clonar seqüências de região constante e variável para a fabricação de anticorpos conforme descrito em Newman *et al.*, Patente dos Estafos Unidos Nº. 5.658.570, arquivada em 25 de janeiro de 1995, a qual é incorporada por meio de referência aqui, neste Pedido de patente. Essencialmente, isto acarreta extração de RNA das células selecionadas, conversão para cDNA, e amplificação por PCR usando iniciadores Ig específicos. Iniciadores adequados para estes fins também são descritos na Patente dos Estados Unidos Nº 5.658.570. Conforme será discutido em mais detalhes abaixo, células transformadas expressando o anticorpo desejado podem ser cultivadas em quantidades relativamente grandes para proporcionar suprimentos clínicos e comerciais da imunoglobulina.

Aqueles versados na técnica também reconhecerão que DNA codificando anticorpos ou fragmentos de anticorpos (por exemplo, sítios de ligação antigênica) também podem ser derivados de bibliotecas de fagos de anticorpos, por exemplo, usando tecnologia de fago pd ou fagemídeo Fd. Métodos típicos são determinados, por exemplo, na patente européia Nº EP 368 684 B1; na patente dos Estados Unidos Nº 5.969.108, Hoogenboom, H.R. e Chames. 2000. *Immunol. Today* 21:371; Nagy et al. 2002. *Nat. Med.* 8:801; Huie et al. 2001. *Proc. Natl. Acad. Sci. USA* 98:2682; Lui et al. 2002. *J. Mol. Biol.* 315:1063, cada uma das quais é incorporada aqui, a este Pedido de patente, por meio de referência. Várias publicações (por exemplo, Marks et al. *Bio/Technology* 10:779-783 (1992)) descreveram a produção de anticorpos humanos de alta afinidade por troca de cadeia, bem como infecção combinatorial e recombinação *in vivo* como uma estratégia para construir bibliotecas de fagos grandes. Em outra modalidade, pode ser usado display ribossômico para substituir bacteriófago como a plataforma de display (vide, por exemplo, Hanes et al. 2000. *Nat. Biotechnol.* 18:1287; Wilson et al. 2001. *Proc. Natl. Acad. Sci. USA* 98:3750; ou Irving et al. 2001 *J. Immunol. Methods* 248:31. Em ainda outra modalidade, bibliotecas superficiais celulares podem ser triadas para anticorpos (Boder et al. 2000. Proc. Natl. Acad. Sci. USA 97:10701; Daugherty et al. 2000 *J. Immunol. Methods* 243:211. Os

procedimentos referidos proporcionam alternativas para técnicas de hibridoma tradicionais para o isolamento e subseqüente clonagem de anticorpos monoclonais.

Em outra modalidade da presente invenção um sítio de ligação de uma molécula de ligação da invenção pode ser proporcionado por um anticorpo humano ou substancialmente humano. Anticorpos humanos ou substancialmente humanos podem ser fabricados em animais transgênicos (por exemplo, camundongos) que são incapazes de produção de imunoglobulina endógena (vide por exemplo, as Patentes dos Estados Unidos Nos. 6.075.181, 5.939.598, 5.591.669 e 5.589.369 cada uma das quais é incorporada aqui, a este Pedido de patente, por meio de referência). Por exemplo, foi descrito que a eliminação homozigótica da região de ligação de cadeia pesada do anticorpo em camundongos quiméricos e germ-line mutante resulta em completa inibição de produção de anticorpos endógenos. A transferência de um array genético de imunoglobulina humana a similares camundongos de linhagem germinativa mutante resultará na produção de anticorpos humano no estímulo antigênico. Outro meio preferencial para produzir anticorpos humanos usando camundongos SCID é revelado na Patente dos Estados Unidos Nº 5.811.524 a qual é incorporada aqui, a este Pedido de patente, por meio de referência. Será reconhecido que o material genético associado com estes anticorpos humanos também pode ser isolado e manipulado conforme descrito aqui, neste Pedido de patente.

Ainda outro meio altamente eficiente para produzir anticorpos recombinantes é revelado por Newman, *Biotechnology*, 10: 1455-1460 (1992). Especificamente, esta técnica resulta na produção de anticorpos primatizados que contêm domínios variáveis de macaco e seqüências constantes humanas. Esta referência é incorporada por meio de referência em sua totalidade aqui, neste Pedido de patente. Além disso, esta técnica também é descrita nas Patentes dos Estados Unidos Nos. 5.658.570, 5.693.780 e 5.756.096 comumente atribuídas, cada uma das quais é incorporada aqui, a este Pedido de patente, por meio de referência.

Em outra modalidade, linfócitos podem ser selecionados por

micromanipulação e os genes variáveis isolados. Por exemplo, células mononucleares do sangue periférico podem ser isoladas de um mamífero imunizado e cultivadas por cerca de 7 dias in vitro. As culturas podem ser triadas para IgGs específicas que satisfazem os critérios de triagem. Células de cavidades positivas podem ser isoladas. Células B produtoras de Ig individuais podem ser isoladas por FACS ou identificando as mesmas em uma prova de placa hemolítica mediada por complemento. Células B produtoras de Ig podem ser micromanipuladas para dentro de um tubo e os genes VH e VL podem ser amplificados usando, por exemplo, RT-PCR. Os genes VH e VL podem ser clonados em um vetor de expressão de anticorpos e transfectados para células (por exemplo, células eucarióticas ou procaróticas) para expressão.

Além disso, seqüências genéticas úteis para produzir as moléculas de ligação da presente invenção podem ser obtidas a partir de uma série de fonte diferentes. Por exemplo, conforme discutido extensivamente acima, uma variedade de genes de anticorpos humanos estão disponíveis sob a forma de depósitos publicamente acessíveis. Muitas seqüências de anticorpos e genes codificando anticorpos foram publicadas e genes de anticorpos adequados podem ser quimicamente sintetizados a partir destas seqüências usando técnicas reconhecidas na técnica. Técnicas de síntese de oligonucleotídeos compatíveis com este aspecto da invenção são de conhecimento geral do técnico versado e podem ser realizadas usando quaisquer dos vários sintetizadores automáticos disponíveis comercialmente. Além disso, seqüências de DNA codificando vários tipos de cadeias pesadas e leves determinadas aqui, neste Pedido de patente, podem ser obtidas através dos serviços de vendedores de síntese de DNA comercial. O material genético obtido usando quaisquer dos métodos precedentes pode ser então alterado ou sintético proporcionando a obtenção dos polipeptídeos da presente invenção.

Alternativamente, linhagens celulares produtoras de anticorpos podem ser selecionadas e cultivadas usando técnicas de conhecimento geral do técnico versado. As técnicas referidas são descritas em uma varieda-

de de manuais de laboratório e publicações primárias. A este respeito, técnicas adequadas para uso na invenção conforme descrito abaixo são descritas em *Current Protocols in Immunology*, Coligan et al., Eds., Green Publishing Associates e Wiley-Interscience, John Wiley e Sons, New York (1991) o qual é aqui, a este Pedido de patente, incorporado por meio de referência em sua totalidade, incluindo suplementos.

Será adicionalmente reconhecido que o âmbito desta invenção engloba moléculas de ligação estabilizadas compreendendo alelos, variantes e mutações de seqüências de DNA ligando antígeno reconhecidas na técnica e uma molécula scFv estabilizada da invenção.

Conforme é de conhecimento geral, RNA pode ser isolado a partir das células de hibridoma originais ou a partir de outras células transformadas por standard técnicas, tais como extração e precipitação de isotiocianato de guanidínio seguida por centrifugação ou cromatografia. Onde desejável, mRNA podem ser isolado a partir de RNA total por técnicas de rotina tais como cromatografia sobre oligo dT celulose. Técnicas adequadas são familiares na técnica.

Em uma modalidade, cDNAs que codificam as cadeias leve e pesada do anticorpo podem ser preparados, ou simultaneamente ou separadamente, usando transcriptase reversa e DNA polimerase de acordo com métodos de conhecimento geral. Reação de PCR pode ser iniciada por iniciadores de região constante de consenso ou por iniciadores mais específicos com base nas seqüências publicadas de DNA e de aminoácidos de cadeia pesada e leve. Conforme discutido acima, PCR também pode ser usado para isolar clones de DNA codificando as cadeias leve e pesada de anticorpos. Neste caso as bibliotecas podem ser triadas por iniciadores de consenso ou sondas homólogas maiores, tais como sondas de região constante de camundongo.

DNA, tipicamente DNA de plasmídeo, podem ser isolados das células usando técnicas conhecidas na técnica, mapeados por restrição e seqüenciados de acordo com técnicas de rotina de conhecimento geral determinadas em detalhes, por exemplo, nas referências precedentes relativas

a técnicas de DNA recombinante. Logicamente, o DNA pode ser sintético de acordo com a presente invenção em qualquer ponto durante o processo de isolamento ou análise subseqüente. Anticorpos típicos ou fragmentos dos mesmos para uso nas moléculas de ligação da invenção incluem anticorpos que reconhecem os alvos determinados aqui, neste Pedido de patente.

Em algumas modalidades, fragmentos de anticorpos de ligação de antígeno podem ser produzidos usando técnicas de conhecimento geral na técnica.

Em uma modalidade, uma molécula de ligação da invenção pode compreender uma molécula de anticorpo completa e uma molécula scFv estabilizada. Em outra modalidade, uma molécula de ligação da invenção pode compreender uma porção de uma molécula de anticorpo e uma molécula scFv estabilizada.

Em uma modalidade, uma molécula de ligação da invenção compreende um fragmento ou porção de um anticorpo e uma molécula scFv estabilizada. Por exemplo, em uma modalidade, uma molécula de ligação da invenção pode compreender um anticorpo com domínio eliminado e uma molécula scFv estabilizada. Anticorpos com domínio eliminado são moléculas de anticorpos nas quais um ou mais domínios são parcialmente ou inteiramente eliminados. Ainda moléculas de ligação estabilizadas compatíveis com modalidades especialmente preferenciais compreenderão construtos com domínio eliminado ou variantes em que todo o domínio CH2 foi removido (construtos ΔCH2). Para outras modalidades preferenciais um peptídeo de ligação curto pode ser substituído para o domínio eliminado para proporcionar flexibilidade e liberdade de movimento para a região variável. Aqueles versados na técnica reconhecerão que similares construtos são particularmente preferenciais devido às propriedades regulatórias do domínio CH2 sobre a taxa catabólica do anticorpo.

Construtos com domínio eliminado podem ser derivados usando um vetor (por exemplo, da IDEC Pharmaceuticals, San Diego) codificando um domínio constante de IgG_1 humana (vide, por exemplo, o Pedido de patente internacional Nº WO 02/060955A2 e o Pedido de patente internacional

Nº WO 02/096948A2). Será observado que estes construtos típicos foram engenheirados para fundir o domínio CH3 diretamente a uma região de articulação dos polipeptídeos respectivos da invenção. Em outros construtos pode ser desejável proporcionar um espaçador peptídico entre a região de articulação e os domínios CH2 e/ou CH3 sintéticos. Por exemplo, podem ser expressados construtos compatíveis em que o domínio CH2 foi eliminado e o domínio CH3 remanescente (sintético ou não sintético) é ligado à região de articulação com um espaçador de 5 a 20 aminoácidos. Um espaçador similar pode ser adicionado, por exemplo, para assegurar que os elementos regulatórios do domínio constante permaneçam livres e acessíveis ou que a região de articulação permaneça flexível. Por exemplo, pode ser usado um construto B3F6 com domínio eliminado tendo um espaçador de aminoácidos curto substituído para o domínio CH2 e a região de articulação menor (B3F6.ΔCH2 [gly/ser]). Outros peptídeos típicos são conhecidos na técnica (vide, por exemplo, Pedido de patente Internacional PCT Nos. WO 2005/000898 e WO 2005/000899). Estes peptídeos de ligação podem ser usados em conexão com as moléculas de ligação da invenção. Preferencialmente, os peptídeos de ligação são usados com um polipeptídeo carecendo de um domínio de cadeia pesada CH2. Preferencialmente, qualquer peptídeo de ligação compatível com a presente invenção será relativamente não-imunogênico e não inibirá a associação não-covalente dos polipeptídeos da invenção.

Em uma modalidade, uma molécula de ligação da invenção compreende uma cadeia pesada de imunoglobulina tendo eliminação ou substituição de uns poucos ou mesmo um único aminoácido na medida que permita a associação covalente ou não-covalente desejada entre as subunidades monoméricas. Por exemplo, a mutação de um único aminoácido em áreas selecionadas do domínio CH2 pode ser suficiente para reduzir substancialmente a ligação de Fc e deste modo aumentar a localização tumoral. Similarmente, pode ser desejável simplesmente eliminar a parte de um ou mais domínios de região constante que controlam a função efetora (por exemplo, ligação de complemento) a ser modulada. As eliminações parciais

referidas das regiões constantes podem melhorar características selecionadas do anticorpo (meia-vida sérica) ao mesmo tempo que deixando outras funções desejáveis associadas com o domínio de região constante em questão intacto. Além disso, conforme aludido acima, as regiões constantes dos anticorpos revelados podem ser sintéticas através da mutação ou substituição de um ou mais aminoácidos que reforçam o perfil do construto resultante. A este respeito pode ser possível romper a atividade proporcionada por um sítio de ligação conservado (por exemplo, ligação de Fc) enquanto mantendo substancialmente a configuração e o perfil imunogênico da molécula de ligação estabilizada. Ainda outras modalidades preferenciais podem compreender a adição de um ou mais aminoácidos à região constante para reforçar características desejáveis tais como função efetora ou proporcionar mais fixação de citotoxina ou carboidrato. Em similares modalidades pode ser desejável inserir ou replicar seqüências específicas derivadas de domínios de região constante selecionados.

É de conhecimento na técnica que a região constante media várias funções efetoras. Por exemplo, ligação do componente C1 do complemento a anticorpos ativa o sistema do complemento. A ativação do complemento é importante na opsonização e lise de patógenos celulares. A ativação do complemento também estimula a reação inflamatória e também pode estar envolvida em hipersensibilidade auto-imune. além disso, anticorpos ligam a células através da região de Fc, com um sítio receptor de Fc sobre a região Fc de anticorpo ligando a um receptor de Fc (FcR) sobre uma célula. Há uma série de receptores de Fc os quais são específicos para diferentes classes de anticorpo, incluindo IgG (receptores γ), IgE (receptores épsilon), IgA (receptores α) e IgM (receptores mu). A ligação de anticorpo a receptores de Fc sobre as superfícies celulares dispara uma série de reações biológicas importantes e diversas incluindo engolfamento e destruição de partículas revestidas de anticorpo, *clearance* de complexos imunes, lise de células alvo revestidas de anticorpos por células assassinas (denominada citotoxicidade mediada por células dependentes de anticorpo, ou ADCC), liberação de mediadores inflamatórios, transferência placentária e controle

da produção de imunoglobulina.

Em uma modalidade, funções efetoras podem ser eliminadas ou reduzidas usando uma região constante de um anticorpo de IgG4, o qual é considerado incapaz de eliminar células alvo, ou fabricar variantes de Fc, em que resíduos na região Fc crucial para uma ou mais funções efetoras são mutadas usando técnicas conhecidas na técnica, por exemplo, na Patente dos Estados Unidos Nº 5.585.097. Por exemplo, a eliminação ou inativação (através de mutações ponto ou outros meios) de um domínio de região constante pode reduzir a ligação de receptor de Fc da molécula de ligação estabilizada circulante deste modo aumentando a localização tumoral. Em outros casos pode ser que modificações de região constante consistentes com a presente invenção moderem a ligação de complemento e deste modo reduzam a meia-vida sérica e a associação inespecífica de uma citotoxina conjugada. Ainda outras modificações de região constante podem ser usadas para modificar ligações dissulfeto ou moléculas de oligossacarídeo que possibilitam aumentada localização devido a aumento da especificidade de antígeno ou flexibilidade de anticorpo. Mais geralmente, aqueles versados na técnica reconhecerão que anticorpos modificados conforme descrito aqui, neste Pedido de patente, podem exercer uma série de efeitos sutis que podem ser ou não prontamente reconhecidos. No entanto o perfil fisiológico resultante, biodisponibilidade e outros efeitos bioquímicos das modificações, tais como localização tumoral, biodistribuição e meia-vida sérica, podem ser facilmente medidos e quantificados usando técnicas imunológicas de conhecimento geral sem indevida experimentação.

Em uma modalidade, formas modificadas de anticorpos podem ser preparadas a partir de um precursor inteiro ou anticorpo parental usando técnicas conhecidas na técnica. Técnicas típicas são discutidas em mais detalhes aqui, neste Pedido de patente.

B. Proteínas de Fusão Modificadas

Em algumas modalidades, uma molécula de ligação estabilizada da invenção é uma proteína de fusão modificada. Em uma modalidade típica, uma molécula de ligação da invenção é uma proteína de fusão compre-

endendo uma molécula scFv estabilizada encadeada à região de ligação de ligante de um receptor, uma molécula de adesão, um ligante, ou uma enzima. Em outra modalidade exemplar, uma molécula de ligação da invenção é proteína de fusão compreendendo uma molécula scFv estabilizada encadeada a uma porção de ligante de ligação de receptor. Por exemplo uma molécula de ligação da invenção é proteína de fusão compreendendo uma molécula scFv estabilizada a uma ou mais das seguintes moléculas:

Citocinas e Receptores de Citocinas

Citocinas têm efeitos pleiotrópicos sobre a proliferação, diferenciação, e ativação funcional de linfócitos. Várias citocinas, ou porções das mesmas de ligação de receptores, podem ser utilizadas nas proteínas de fusão da invenção. Citocinas típicas incluem as interleucinas (por exemplo, IL-1, IL-2, IL-3, IL-4, IL-5, IL-6, IL-7, IL-8, IL-10, IL-11, IL-12, IL-13, e IL-18), os fatores estimulantes de colônias (CSFs) (por exemplo, CSF de granulócitos CSF (G-CSF), CSF de granulócitos-macrófagos (GM-CSF), e CSF de monócito macrófago CSF (M-CSF)), fator de necrose tumoral (TNF) α e β, e interferons tais como interferon-α, β, ou γ (Patentes dos Estados Unidos Nos. 4.925.793 e 4.929.554).

Receptores de citocina tipicamente consistem em um cadeia α ligante-específica e uma cadeia β comum. Receptores de citocina típicos incluem aqueles para GM-CSF, IL-3 (Patente dos Estados Unidos Nº 5.639.605), IL-4 (Patente dos Estados Unidos Nº 5.599.905), IL-5 (Patente dos Estados Unidos Nº 5.453.491), IFNγ (patente européia Nº E P0240975), e a família de receptores TNF (por exemplo, TNFα (por exemplo, β receptor de linfotoxina TNFR-1 (patente européia Nº EP 417, 563), TNFR-2 (patente européia Nº EP 417.014)).

Em outra modalidade, uma molécula scFv da invenção pode ligar a uma citocina ou a um receptor de citocina.

Proteínas de Adesão

Moléculas de adesão são proteínas ligadas a membrana que possibilitam que as células interajam umas com as outras. Várias proteínas de adesão, incluindo receptores residentes de leucócitos e moléculas de a-

desão celular, de porções de ligação de receptores dos mesmos, podem ser incorporadas em uma molécula de ligação da invenção. Receptores residentes de leucócitos são expressados sobre supefícies de células de leucocitos durante inflamação e incluem as β-1 integrinas (por exemplo, VLA-1, 2, 3, 4, 5, e 6) as quais mediaam ligação a componentes da matriz extracelular, e as β2-integrinas (por exemplo, LFA-1, LPAM-1, CR3, e CR4) as quais ligam moléculas de adesão celular (CAMs) sobre o endotélio vascular. CAMs típicas incluem ICAM-1, ICAM-2, VCAM-1, e MAdCAM-1. Outras CAMs incluem as da família das selectinas incluindo E-selectina, L-selectina, e P-selectina.

Em outra modalidade, uma molécula scFv da invenção pode ligar a uma proteína de adesão ou um receptor de proteína de adesão.

Quimocinas

Quimocinas, proteínas quimiotácticas as quais estimulam a migração de leucócitos para um sítio de infecção ou porções de ligação de receptores de quimocina dos mesmos, também podem ser incorporadas em uma molécula de ligação da invenção. Quimocinas típicas incluem proteínas inflamatórias de macrófagos (MIP-1-α e MIP-1-β), fator quimotáctico de neutrófilos, e RANTES (regulado na ativação de célula T expressada e secretada normalmente).

Em outra modalidade, uma molécula scFv da invenção pode ligar a uma quimocina ou um receptor.

Fatores de Crescimento e Receptores de Fator de Crescimento

Fatores de crescimento ou seus receptores (ou porções de ligação de receptor ou de ligação de ligante dos mesmos) ou moléculas as quais ligam aos mesmos podem ser incorporados na molécula de ligação da invenção. Fatores de crescimento típicos incluem angiopoietina, Fator de Crescimento Endotelial Vascular (VEGF) e suas isoformas (Patente dos Estados Unidos Nº 5.194.596); Fatores de Crescimento Epidérmico (EGFs); Fatores de Crescimento Fibroblástico (FGF), incluindo aFGF e bFGF; fator natriurético atrial (ANF); fatores de crescimento hepático (HGFs; Patentes dos Estados Unidos Nos. 5.227.158 e 6.099.841), fatores neurotróficos tais como fator neurotrófico ósteo-derivado (BDNF), neurotrofina-3, -4, -5, ou -6

(NT-3, NT-4, NT-5, ou NT-6), ou um fator de crescimento nervoso tal como fator de crescimento derivado de plaquetas NGF-ß (PDGF) (Patentes dos Estados Unidos Nos. 4.889.919, 4.845.075, 5.910.574, e 5.877.016); fatores de crescimento transformantes (TGF) tais como TGF-α e TGF-β (WO 90/14359), fatores osteoindutivos incluindo proteína morfogenética óssea (BMP); fatores de crescimento similares a insulina-I e -II (IGF-I e IGF-II; Patentes dos Estados Unidos Nos. 6.403.764 e 6.506.874); Eritropoietina (EPO); fator de células-tronco (SCF), trombopoietina (c-*Mpl* ligante), e os polipeptídeos Wnt (Patente dos Estados Unidos Nº 6.159.462).

Receptores de fatores de crescimento típicos os quais podem ser usados incluem EGF receptores (EGFRs); VEGF receptores (por exemplo, Flt1 ou Flk1/KDR), PDGF receptores (publicação de patente internacional Nº WO 90/14425); HGF receptores (Patentes dos Estados Unidos Nos. 5.648.273, e 5.686.292); IGF receptores (por exemplo, IGFR1 e IGFR2) e receptores neurotróficos incluindo o receptor de baixa afinidade (LNGFR), também denominado como $p75^{NTR}$ ou p75, o qual liga NGF, BDNF, e NT-3, e receptores de alta afinidade que são membros da família trk das tirosina quinases receptoras (por exemplo, trkA, trkB (EP 455,460), trkC (EP 522,530)). Em outra modalidade, tanto IGFR1 e VEGF são direcionados. Em ainda outra modalidade, VLA4 e VEGF são direcionados. Em outra modalidade, tanto LFA1 quanto VLA4 são direcionados.

Outros receptores da superfície celular e/ou seus ligantes também podem ser direcionados (por exemplo, os receptores da família de TNF ou seus ligantes (conforme descrito em mais detalhes aqui, neste Pedido de patente).

Em outra modalidade, uma molécula scFv da invenção pode ligar a um fator de crescimento ou receptor de fator de crescimento.

Hormônios

Hormônios de crescimento típicos ou moléculas as quais ligam aos mesmos para uso como agentes de direcionamento em uma molécula de ligação da invenção incluem renina, hormônio de crescimento humano (HGH; Patente dos Estados Unidos Nº 5.834.598), hormônio de crescimento

humano N-metionil; hormônio de crescimento bovino; fator de liberação de hormônio de crescimento; hormônio paratiróide (PTH); hormônio estimulante da tiróide (TSH); tiroxina; pró-insulina e insulina (Patentes dos Estados Unidos Nos. 5.157.021 e 6.576.608); hormônio folículo estimulante (FSH), calcitonina, hormônio luteinizante (LH), leptina, glucagons; bombesina; somatropina; substância de inibição muleriana; relaxina e prorrelaxina; peptídeo associado a gonadotropina; prolactina; lactogênio placentário; OB proteína; ou substância de inibição muleriana.

Em outra modalidade, uma molécula scFv da invenção pode ligar a um hormônio ou um receptor de hormônio.

Fatores de Coagulação

Fatores de coagulação sangüínea típicos para uso como agentes de direcionamento nas moléculas de ligação da invenção incluem os fatores de coagulação (por exemplo, fatores V, VII, VIII, X, IX, XI, XII e XIII, fator de von Willebrand); fator tecidual (Patentes dos Estados Unidos Nos. 5.346.991, 5.349.991, 5.726.147, e 6.596.84); trombina e protrombina; fibrina e fibrinogênio; plasmina e plasminogênio; ativadores de plasminogênio, tais como ativador de plasminogênio tecidual (t-PA) ou uroquinase ou urina humana.

Em outra modalidade, uma molécula scFv da invenção pode ligar a um fator de coagulação.

Em outras modalidades, proteína de ligação estabilizada da invenção é uma imunoadesina modificada. Conforme é de conhecimento geral na técnica, uma imunoadesina é uma proteína de fusão que combina a região de ligação alvo de um receptor, uma molécula de adesão, um ligante, ou uma enzima, com a região Fc de um anticorpo. Imunoadesinas típicas são descritas, por exemplo, nas Patentes dos Estados Unidos Nos. 5.116.964; 5.428.130; 5.714.147; e 6.406.697, cada uma das quais é incorporada por meio de referência aqui, a este Pedido de patente. Em uma modalidade, uma imunoadesina modificada da invenção é formada encadeando uma molécula scFv estabilizada a uma imunoadesina.

Imunoadesinas previamente reportadas podem ser estabilizadas

conforme descrito aqui, neste Pedido de patente, para proporcionar as imunoadesinas modificadas da presente invenção. Imunoadesinas típicas que podem ser usadas para gerar ou derivar a imunoadesina modificada revelada incluem, mas não estão limitadas a, abatecept (Orencia®, Bristol-Meyers Squibb, Princeton, NJ), alefacept (Amevive®, BiogenIdec, Cambridge, MA), etanercept (Enbrel®, Amgen, Thousand Oaks, CA), SMART ™ Anti-γ Interferon (Protein Design Labs, Fremont, CA), SMART ™ Anti-L-Selectina (Protein Design Labs. Fremont, CA), rilonacept (Regeneron Pharmaceuticals Inc., Tarrytown, NY), regavirumab (TI-23, Teijin America, New York, NY), R24 (National Cancer Institute (Bethesda, MD), Oprelvekin (NEUMEGA® , Genetics Institute, Cambridge, MA), ONCOLYSIN B, ONCOLYSIN CD6, ONCOLYSIN M, e ONCOLYSIN S (todas da ImmunoGen Inc., Cambridge, MA, USA), ONCOLYM ™ 131 (Techniclone Corp., Tustin, CA), ImmuRAIT-LL2 (Immunomedics Inc., Morris Plains, NJ), IL-4 RA (BAY 16-9996; Bayer Corp., Berkeley, CA), IC14 (ICOS Corporation, Bothell, WA), CYT-356-Y-90 (ONCORAD®PR, Cytogen Corp., Princeton, NJ), COTARA ™ (Techniclone Corp., Tustin, CA), CMB-401 (Wyeth Pharmaceuticals, Madison, NJ), conjugado AVICIDIN® (NeoRx Corp., Seattle, WA) e antiimunoadesina CD18 (Genentech Inc., S. San Francisco, CA).

Outra molécula típica que pode ser incluída em uma molécula de ligação da invenção é membro 9 da superfamília de imunoglobulinas (IGSF9; Genomics. 2002. 79:663-70).

C. Especificidade de Ligação

Em uma modalidade, uma molécula de ligação da invenção liga a uma molécula alvo que está presente sobre a superfície de uma célula ou que é solúvel.

Em uma modalidade, no mínimo uma especificidade de ligação de uma molécula de ligação da invenção é catalítica. Especificidades de ligação catalíticas podem ser feitas usando técnicas reconhecidas na técnica (vide, por exemplo, a Patente dos Estados Unidos Nº 6.590.080, e a a Patente dos Estados Unidos Nº 5.658.753). Especificidades de ligação catalíticas podem trabalhar por uma série de mecanismos básicos similares aque-

les identificados para enzimas para estabilizar o estado de transição, deste modo reduzindo a energia livre de ativação. Por exemplo, resíduos ácidos e básicos gerais podem ser otimamente posicionados para participação em catálise dentro de sítios ativos catalíticos; intermediários de substratos enzimáticos covalentes podem ser formados; anticorpos catalíticos também podem estar em orientação apropriada para reação e aumentar a concentração eficaz de reagentes por no mínimo sete ordens de magnitude (Fersht, A. R., et al., Am. Chem. Soc. 90 (1968): 5833) e deste modo reduzir grandemente a entropia de uma reação química. Finalmente, anticorpos catalíticos podem converter a energia obtida na ligação de substrato para distorcer a reação para uma estrutura se assemelhando ao estado de transição.

Em uma modalidade, resíduos de ácido ou base podem ser trazidos para o sítio de ligação usando uma molécula complementar carregada como um imunogênio. Esta técnica se comprovou de êxito para produção de anticorpos com um hapteno contendo um íon amônio carregado positivamente (Shokat, et al., Chem. Int. Ed. Engl. 27 (1988): 269-271).

Em outra abordagem, podem ser produzidos anticorpos para compostos estáveis que se assemelham ao tamanho, à forma, e à carga do estado de transição de uma reação desejada (isto é, análogos do estado de transição). Vide a a Patente dos Estados Unidos Nº 4.792.446 e a a Patente dos Estados Unidos Nº 4.963.355 as quais descrevem o uso de análogos do estado de transição para imunizar animais e a produção de anticorpos catalíticos. Ambas estas patentes são por este incorporadas por meio de referência. Em uma modalidade, similares moléculas podem ser administrada como parte de um imunoconjugado, por exemplo, com uma molécula de veículo imunogênica, tal como KLH.

Especificidades de ligação catalítica típicas podem ter, por exemplo, atividade de esterase (envolvendo um estado de transição carregado cujas características eletrostáticas e de formato podem ser simuladas por uma estrutura de fosfonato; Jacobs, et al., J. Am. Chem. Soc. 109 (1987): 2174-2176; Durfor, et al., J. Am. Chem. Soc. 110 (1988): 8713-8714; Tramontano, et al., J. Am. Chem. Soc. 110 (1988): 2282; Pollack, et al., J. Am.

Chem. Soc. 111 (1989): 5961-5962); atividade de peptidase ou amidase (Janda, et al., Science 241 (1988): 1188-1191; Iverson, et al., Science 243 (1989): 1184-1188; Paul, et al., Science 244 (1989): 1158-1162); redisposição de Claisen (Jackson, et al., J. Am. Chem. Soc. 110 (1988): 4841-4842; Hilvert, et al., Proc. Natl. Acad. Sci. USA 85 (1988): 4953-4955; Hilvert, et al., J. Am. Chem. Soc. 110 (1988): 5593-5594); reações redox (Shokat, et al., Angew. Chem. Int. Ed. Engl. 27 (1989): 269-271); clivagem fotoquímica de um dímero de timina (Cochran, et al., J. Am. Chem. Soc. 110 (1988): 7888-7890); redisposições de transesterificação estereoespecífica (Napper, et al., Science 237 (1987): 1041-1043); ou uma síntese de amida bimolecular (Benkovic, et al., Proc. Natl. Acad. Sci. USA 85 (1988): 5355-5358; Janda, et al., Science 241 (1988): 1188-1191).

Em outra abordagem, especificidades de ligação convencionais podem ser mutadas para as tornar catalíticas.

Métodos para triagem de atividade de anticorpo catalítico são de conhecimento geral na técnica (por exemplo, Reymond, J.L. 2002. Journal of Immunological Methods 269:125; Mouratou et al. 2002. J. of Immunological Methods. 269:147. Em ainda outra modalidade, células B catalíticas podem ser seleciondas, por exemplo, conforme descrito na patente dos Estados Unidos Nº 6.590.080 usando uma molécula pode ser construída a qual facilita seleção de células B catalíticas.

Em outra modalidade, especificidades de ligação catalíticas podem ser desenvolidas como parte de um processo de duas etapas. Anticorpos catalíticos podem ser selecionados somente se apresentando as seguintes características de ligação: ligar tanto o substrato quanto um grupamento reativo de tal modo que os dois grupamentos estão em uma posição reativa um para o outro. Segundo, os anticorpos selecionados podem ser quimicamente engenheirados ligando de modo covalente um grupamento reativo na cavidade de ligação do anticorpo. J Immunol Methods. 2002. 269:81-98.

Em uma modalidade, uma especificidade de ligação catalítica é específica para um pró-farmaco. Uma especificidade de ligação similar pode ser usada para catalisar a conversão de um pró-farmaco em um fármaco o

qual é eficaz in vivo. Preferencialmente, a reação catalisada é uma reação que não pode ser realizada por enzimas naturais in vivo. Exemplos de ativação de pró-farmaco por anticorpos são conhecidas na técnica (vide, por exemplo, Miyashita et al. 1993. Proc. Natl. Acad. Sci. USA 90:5337).

Em uma modalidade, uma molécula de ligação da invenção compreende no mínimo uma especificidade de ligação para uma célula alvo e no mínimo uma especificidade de ligação para um pró-farmaco. Por exemplo, em uma modalidade preferencial, uma molécula de ligação estabilizada da invenção compreende no mínimo uma especificidade de ligação para uma célula tumoral e no mínimo uma especificidade de ligação para um pró-farmaco a qual pode ser convertida para fármaco citotóxico. Em um exemplo, uma molécula de ligação estabilizada da invenção compreende uma especificidade de ligação para um pró-farmaco de carbamato: 4-[N,N,-bis(2-cloroetil)]aminofenil-N-[(1S-(1,3-dicarbóxi)propil]carbamato e gera a mostarda de nitrogênio citotóxica correspondente (Wentworth et al. 1996. Proc Natl. Acad. Sci. USA. 93:799).

Em uma modalidade, molécula de ligação é administrada antes de administração do pró-farmaco para possibilitar a acumulação no sítio da célula alvo. Pró-farmacos típicos são conhecidos na técnica. Pró-farmacos também podem ser sintetizados incorporando uma porção projetada para ser liberada por ação catalítica, por exemplo, por reações seqüenciais de retro-aldol / retro-Michael catalisadas por um anticorpo com atividade de aldolase. (Shabat et al. 2001. Proc. Natl. Acad. Sci. USA 98:7428). As porções de marcaramento de fármaco referidas podem ser preparadas, por exemplo, por modificação de grupamentos oxidrila ou tiol de fármacos.

Em uma modalidade, uma molécula de ligação da invenção é multiespecífica, isto é, tem no mínimo um sítio de ligação que liga a uma primeira molécula alvo ou epítope da molécula alvo e no mínimo um segundo sítio de ligação que liga a uma segunda molécula alvo diferente ou a um segundo epítope diferente da primeira molécula alvo. Em algumas modalidades, moléculas de ligação multiespecífica da invenção (por exemplo, moléculas de ligação biespecífica) compreendem no mínimo um sítio de ligação

de qualquer um dos anticorpos descritos aqui, neste Pedido de patente, por exemplo, na Seção A, acima.

Em uma modalidade, uma molécula de ligação da invenção é biespecífica. Moléculas biespecíficas podem ligar a dois sítios alvo diferentes, por exemplo, sobre a mesma molécula alvo ou sobre moléculas alvo diferentes. Por exemplo, no caso de anticorpos, moléculas biespecíficas podem ligar a dois epítopes diferentes, por exemplo, sobre o mesmo antígeno ou sobre dois antígenos diferentes. Moléculas biespecíficas podem ser usadas, por exemplo, em aplicações de diagnóstico e terapêuticas. Por exemplo, podem ser usadas para imobilizar enzimas para uso em imunoensaio. Também podem ser usadas em diagnóstico e tratamento de câncer, por exemplo, ligando tanto uma molécula associada a tumor quanto um marcador detectável (por exemplo, um quelador o qual liga firmemente um radionuclídeo. Moléculas biespecíficas também podem ser usadas para terapia humana, por exemplo, dirigindo a citotoxicidade para um alvo específico (por exemplo, ligando a um patógeno ou célula tumoral e a uma molécula de gatilho citotóxico, tal como o receptor de células T ou o Fcγ receptor. Anticorpos biespecíficos também podem ser usados, por exemplo, como agentes fibrinolíticos ou adjuvantes de vacina.

Em uma modalidade, as moléculas de ligação multiespecífica da invenção incluem aquelas com no mínimo um ramo (isto é, sítio de ligação) dirigido contra uma molécula da superfície celular, e no mínimo um ramo dirigido contra uma molécula solúvel. Em outra modalidade, um anticorpo multiespecífico da invenção tem dois sítios de ligação que ligam a moléculas solúveis. Em outra modalidade, um anticorpo multiespecífico da invenção tem dois sítios de ligação que ligam a moléculas da superfície celular.

As moléculas de ligação multiespecífica da invenção podem ser monovalentes para cada especificidade ou multivalentes para cada especificidade. Em uma modalidade, uma molécula de ligação biespecífica da invenção pode compreender um sítio de ligação que reage com uma primeira molécula alvo e um sítio de ligação que reage com uma segunda molécula alvo (por exemplo, um anticorpo biespecífico molécula, proteína de fusão, ou

"minibody"). Em outra modalidade, uma molécula de ligação biespecífica da invenção pode compreender dois sítios de ligação que reagem com uma primeira molécula alvo e dois sítios de ligação que reagem com uma segunda molécula alvo (por exemplo, um anticorpo tetravalente scFv2 biespecífica, "minibody" tetravalente, ou diabody).

Em algumas modalidades, no mínimo um sítio de ligação de uma molécula de ligação multiespecífica da invenção é uma região de ligação de antígeno de um anticorpo ou um fragmento de ligação de antígeno do mesmo.

Em uma modalidade, as moléculas de ligação multiespecífica da invenção são anticorpos bivalentes ou variantes de anticorpo com um ramo contendo no mínimo uma scFv estabilizada dirigida para uma primeira molécula alvo e um segundo ramo contendo no mínimo uma scFv estabilizada dirigida para uma segunda molécula alvo.

Em uma modalidade, as moléculas de ligação multiespecífica da invenção compreendem no mínimo uma scFv estabilizada (por exemplo, 2, 3, ou 4 scFvs) encadeada ao C-término de uma cadeia pesada, em que as scFvs têm a mesma especificidade de ligação ou diferentes. Uma molécula de ligação típica deste tipo (um anticorpo "C-Hercules") é mostrada na Figura 13. Em outra modalidade, as moléculas de ligação multiespecífica da invenção compreendem no mínimo uma scFv estabilizada (por exemplo, 2, 3, ou 4 scFvs) encadeadas ao N-término de uma cadeia pesada , em que as scFvs têm a mesma especificidade de ligação ou diferentes. Uma molécula de ligação típica deste tipo (um anticorpo "N_H-Hercules") é mostradaa na Figura 13. Em outra modalidade, as moléculas de ligação multiespecífica da invenção compreendem no mínimo uma scFv estabilizada (por exemplo, 2, 3, ou 4 scFvs) encadeada ao N-término de uma cadeia leve, em que as scFvs têm a mesma especificidade de ligação ou diferentes. Uma molécula de ligação típica deste tipo (um anticorpo "N_L-Hercules") é mostrada na Figura 13. Em outra modalidade, as moléculas de ligação multiespecífica da invenção compreendem no mínimo uma scFv estabilizada (por exemplo, 2, 3, ou 4 scFvs) encadeada ao N-término da cadeia pesada ou cadeia leve e no

mínimo uma scFv estabilizada (por exemplo, 2, 3, ou 4 scFvs) encadeada ao C-término da cadeia pesada, em que as scFvs têm a mesma especificidade de ligação ou diferentes.

Em uma modalidade, as moléculas de ligação multiespecífica da invenção são minibodies bivalentes com um ramo contendo no mínimo um fragmento de scFv dirigido para uma primeira molécula alvo e um segundo ramo contendo no mínimo um scFv dirigido para uma segunda molécula alvo em que no mínimo uma das moléculas scFv é estabilizada. Um construto de "minibody" bivalente biespecífico típico é mostrado na Figura 42. Na Figura 42 um domínio CH3 é fundido em seu N-término a um peptídeo de ligação o qual é fundido em seu N-término a um domínio VH o qual é fundido através de seu N-término a um ligante flexível (Gly4Ser)n o qual é fundido em seu N-término a um domínio VL.

Em outra modalidade, as moléculas de ligação multiespecífica da invenção são minibodies tetravalentes scFv, com cada "minibody" tetravalente da porção da scFv de cadeia pesada contendo primeiro e segundo fragmentos de scFv em que em elsat uma das moléculas scFv é estabilizada. O segundo fragmento de scFv referido pode ser encadeado ao N-término do primeiro fragmento de scFv (por exemplo, minibodies tetravalentes N_H scFv biespecíficos ou minibodies tetravalentes N_L scFv biespecíficos). Um exemplo de um "minibody" tetravalente N-scFv biespecífico é mostrado na Figura 43. Alternativamente, o segundo fragmento de scFv pode ser encadeado ao C-término da referida porção de cadeia pesada contendo o primeiro fragmento de scFv referido (por exemplo, minibodies tetravalentes C-scFv biespecíficos). Um exemplo de um "minibody" tetravalente C-scFv biespecífico é mostrado na Figura 44. Em uma modalidade, o primeiro e o segundo fragmentos de scFv podem ligar a mesma molécula alvo ou diferentes. Onde o primeiro e o segundo fragmentos de scFv de uma primeira porção de cadeia pesada de um "minibody" tetravalente biespecífico ligam a mesma molécula alvo, no mínimo um dos primeiro e segundo fragmentos de scFv da segunda porção de cadeia pesada do "minibody" tetravalente biespecífico liga uma molécula alvo diferente.

Em outra modalidade, as moléculas de ligação multiespecífica da invenção são diabodies biespecíficos, com cada ramo do diabody compreendendo fragmentos tandem de scFv no mínimo um dos quais é estabilizado. Em uma modalidade, um diabody biespecífico pode compreender um primeiro ramo com uma primeira especificidade de ligação e um segundo ramo com uma segunda especificidade de ligação (vide, por exemplo, a Figura 45). Em outra modalidade, cada ramo do diabody pode compreender um primeiro fragmento de scFv com uma primeira especificidade de ligação e um segundo fragmento de scFv com uma segunda especificidade de ligação.

Em outra modalidade, as moléculas de ligação multiespecífica da invenção são anticorpos tetravalentes scFv2 com cada porção de cadeia pesada do anticorpo tetravalente scFv2 contendo uma molécula scFv, em que no mínimo uma das moléculas scFv são estabilizadas. Os fragmentos de scFv podem ser encadeados aos N-términos de uma região variável das porções de cadeia pesada (por exemplo, anticorpos tetravalentes N_H scFv2 biespecíficos ou anticorpos tetravalentes N_L scFv2 biespecíficos). Alternativamente, os fragmentos de scFv podem ser encadeados aos C-términos das porções de cadeia pesada do anticorpo tetravalente scFv2 (por exemplo, anticorpos tetravalentes C-scFv2 biespecíficos, vide, por exemplo, a Figura 46). Cada porção de cadeia pesada do anticorpo tetravalente scFv2 pode ter regiões variáveis e fragmentos de scFv que ligam as mesmas ou diferentes moléculas alvo. Onde o fragmento de scFv e região variável de uma primeira porção de cadeia pesada de um anticorpo tetravalente scFc2 biespecífico ligam a mesma molécula alvo, no mínimo um dos primeiro e segundo fragmentos de scFv da segunda porção de cadeia pesada do "minibody" tetravalente biespecífico liga uma molécula alvo diferente.

Em outra modalidade, as moléculas de ligação multiespecífica da invenção são anticorpos com domínio eliminado tetravalentes scFv2 com cada porção de cadeia pesada do anticorpo tetravalente scFv2 contendo um fragmento de scFv no mínimo um dos quais é estabilizado. Os fragmentos de scFv podem ser encadeados aos N-términos de uma região variável das

porções de cadeia pesada (por exemplo, anticorpos com domínio eliminado tetravalentes N_H scFv2 biespecíficos (vide a Figura 48) ou anticorpos tetravalentes N_L scFv2 biespecíficos (vide a Figura 49). Alternativamente, os fragmentos de scFv podem ser encadeados aos C-términos das porções de cadeia pesada do anticorpo tetravalente com domínio eliminado scFv2 (por exemplo, anticorpos com domínio eliminado tetravalentes C-scFv2 biespecíficos, vide por exemplo, a Figura 47).

Moléculas da superfície celular típicas às quais uma molécula de ligação da invenção pode ligar incluem receptores ou antígenos de células tumorais que são superexpressadas sobre a superfície de uma célula tumoral ou neoplásica, bem como quaisquer dos receptores de citocina, moléculas de adesão, ou receptores de fator de crescimento descritos aqui, neste Pedido de patente, por exemplo, na seção B acima. Moléculas solúveis típicas incluem agentes antitumorais (por exemplo, toxinas, quimioterápicos, e pró-farmacos dos mesmos), enzimas solúveis (por exemplo, enzimas conversoras de pró-farmacos), citocinas, quimocinas, hormônios, fatores de crescimento, ou fatores de coagulação, por exemplo, conforme descrito aqui, neste Pedido de patente, por exemplo, na seção A acima.

Moléculas biespecíficas as quais ligam tanto a antígenos de células tumorais quanto a agentes antitumorais ou enzimas solúveis podem portanto localizar o agente anticancerígeno para uma célula tumoral expressando o referido antígeno de células tumorais, deste modo maximizando os efeitos tóxicos do agente anticancerígeno sobre uma célula tumoral e minimizando um efeito tóxico do anticancerígeno sobre células normais.

Moléculas de ligação biespecífica típicas com no mínimo um sítio de ligação para um antígeno tumoral e no mínimo um sítio de ligação para uma toxina incluem anti-saporin / anti-Id-1, anti-CD22 / anti-saporin, anti-CD7 / anti-saporin, anti-CD38 / anti-saporina, anti-CEA / anti-cadeia A de ricina, antiinterferon-.α.(IFN-.α.) / antiidiotipo de hibridoma, anti-CEA / antivinca alcalóide). Moléculas de ligação biespecífica típicas com no mínimo um sítio de ligação para uma molécula da superfície celular e no mínimo um sítio de ligação para um pró-farmaco convertendo enzima incluem, por e-

xemplo, anti-CD30 / antifosfatase alcalina (a qual catalisa conversão de pró-farmaco de fosfato de mitomicina para o álcool de mitomicina quimioterápico).

Em outras modalidades, as moléculas de ligação biespecífica ligam tanto a antígenos de células tumorais quanto a agentes de diagnóstico, deste modo localizando o referido agente de diagnóstico para uma célula tumoral expressando o referido antígeno de células tumorais e facilitando a detecção de tumor *in vitro* ou *in vivo*. Moléculas de ligação biespecíficas típicas incluem anti-CEA / anti-EOTUBE, anti-CEA / anti-DPTA, anti-CEA / anti-.β.-galactosidase, e anti-p185HER2 / anti- -hapteno).

Em outras modalidades, moléculas de ligação biespecífica da invenção ligam tanto a moléculas solúveis (por exemplo, antígenos solúveis) e moléculas da superfície celular sobre células não tumorais (por exemplo, células imunes). Por exemplo, podem ser usado para dirigir complexos imunes solúveis para receptores da superfície celular sobre células imunes, deste modo facilitando seu *clearance* do corpo por mecanismos imunes mediados por células. Moléculas biespecíficas típicas deste tipo incluem antilipoproteína de baixa densidade (LDL) / anti-Fc receptor (por exemplo, Fc. γ.RI, Fc. γ.RII ou Fc. γ.RIII)) e moléculas de ligação biespecífica para uso em terapia de doenças infecciosas (tais como anti-CD3 / anti-herpes simplex vírus (HSV), anti-receptor de células T:CD3 complexo / antiinfluenza, anti-Fc.gamma.R / anti-HIV.

Em outras modalidades, moléculas de ligação biespecífica da invenção são capazes de ligar tanto a receptores da superfície celular quanto a ligantes solúveis dos mesmos. Em uma modalidade, o ligante é o ligante cognato de um receptor da família de TNF.

Receptores da superfície celular típicos aos quais as moléculas de ligação biespecífica podem ligar são antígenos de células tumorais ou receptores de células imunes. Receptores da superfície celular típicos também incluem receptores de citocina, moléculas de adesão, ou receptores de fatores de crescimento, por exemplo, conforme descrito aqui, neste Pedido de patente, por exemplo, na seção B acima. Ligantes solúveis típicos inclu-

em citocinas, quimocinas, hormônios, fatores de crescimento, ou fatores de coagulação, por exemplo, conforme descrito na seção B acima.

Moléculas de ligação biespecífica típicas incluem anti-VLA4 / anti-Mac-1, anti-VLA4 / anti-VEGF, anti-VLA4 / antiangiopoietina, anti-VLA4 / anti-TNFα, anti-IGFR1 / anti-VEGF, anti-IGFR1 / antiangiopoietina, anti-IGFR1 / anti-EGFR, anti-HGF-SF / anti-VEGF, anti-HGF-SF / antiangiopoietina, e HGF-SF / qualquer segundo antígeno (Vide, por exemplo, Cao et al. Proc. Natl. Acad. Sci 2001. 98:7443; Lu et al. 2004. J. Biol. Chem. 279:2856).

Em outras modalidades, as moléculas de ligação biespecífica da invenção incluem aquelas com no mínimo um ramo (isto é, sítio de ligação) dirigido contra uma primeira molécula solúvel (por exemplo, ligante solúvel), e no mínimo um ramo dirigido contra uma segunda molécula solúvel (por exemplo, ligante solúvel). As moléculas de ligação biespecífica referidas podem ser empregadas como ferramentas de diagnóstico (por exemplo, anti-IgG de coelho / antiferritina, antiperoxidase de raiz-forte (HRP) / anti-hormônio, anti-somatostatina / anti-substância P, anti-HRP / anti-FITC (vide Nolan *et al.*, acima)) ou agentes fibrinolíticos (por exemplo, anti-fibrina / antiativador de plasminogênio tecidual (tPA), anti-fibrina / antiativador de plasminogênio do tipo uroquinase (uPA)).

Em uma modalidade preferencial, a molécula solúvel à uma molécula de ligação biespecífica da invenção liga é um ligante solúvel da família de TNF. Exemplos de ligantes da família de TNF incluem, mas não estão limitados a, LTA (o qual liga TNFR1 / TNFRSF1A), TNF (o qual liga CD120b / TNFRSF1B), LTB (o qual liga LTBR / TNFRSF3), OX40L (o qual liga OX40 / TNFRSF4), CD40L (o qual liga CD40 / TNFRSF5), (o qual liga Fas / TNFRSF6 e DcR3 / TNFRSF6B), CD27L (o qual liga CD27 / TNFRSF7), CD30L (o qual liga CD30 / TNFRSF8), 4-1-BB-L (o qual liga 4-1-BB / TNFRSF9), TRAIL (o qual liga TRAIL-R1 / TNFRSF10A, TRAIL-R2/TNFRSF10B, TRAIL-R3 / TNFRSF10C, e TRAIL-R4 / TNFRSF10D), RANKL (o qual liga RANK / TNFRSF11A e Osteoprotegrina / TNFRSF11B), APO-3L (o qual liga APO-3 / TNFRSF12 e DR3L / TNFRSF12L), APRIL (o qual liga TACI / TNFRSF13B),

BAFF (o qual liga BAFFR / TNFRSF13A), LIGHT (o qual liga HVEM / TNFRSF14), NGF ligands (o qual liga LNGFR, por exemplo, NGF-β, NGF-2 / NTF3, NTF5, BDNF, IFRD1), GITRL (o qual liga GITR / TNFRSF18), ligante EDAR1 & XEDAR, ligante Fn14, e ligante Troy/Trade.

Em outras modalidades exemplares, as moléculas de ligação biespecífica da invenção têm no mínimo um sítio de ligação para uma primeira molécula da superfície celular e no mínimo um sítio de ligação para uma segunda molécula da superfície celular. Em uma modalidade, a primeira e segunda moléculas da superfície celular estão localizadsa sobre células diferentes (por exemplo, tipos celulares diferentes). Por exemplo, moléculas biespecíficas podem ter no mínimo um ramo dirigido contra um antígeno de célula tumoral e no mínimo um ramo diridigo contra receptor da superfície celular sobre uma célula não tumoral (por exemplo, uma célula imune). Moléculas de ligação biespecífica típicas deste tipo incluem aquelas tendo no mínimo um sítio de ligação para um antígeno de células tumorais e no mínimo um sítio de ligação dirigido contra uma molécula de gatilho citotóxico de uma célula efetora imune (tal como anti-Fc. γ.RI / anti-CD15, anti-p185.sup.HER2 / Fc. γ.RIII (CD16), anti-p185.sup.HER2 / anti-VEGF, anti-CD3 / anti-célula B maligna (1D10), anti-CD3 / anti-p185.sup.HER2, anti-CD3 / anti-p97, anti-CD3 / anticarcinoma de células renais, anti-CD3 / anti-OVCAR-3, anti-CD3 / L-D1 (anticarcinoma do cólon), anti-CD3 / antianálogo de hormônio estimulante de melanócito, anti-EGF receptor / anti-CD3, anti-CD3 / anti-CAMA1, anti-CD3 / anti-CD19, anti-CD3 / MoV18, antimolécula de adesão de células neurais (NCAM) / anti-CD3, antiproteína de ligação de folato (FBP) / anti-CD3, e antiantígeno associado a pan carcinoma (AMOC-31) / anti-CD3)). Moléculas biespecíficas as quais ligam tanto a antígenos de células tumorais quanto molécula de gatilho citotóxico são capazes de justapor de modo eficaz uma célula tumoral com uma célula efetora imune, deste modo ativando a célula efetora para destruir a célula tumoral por meio de mecanismos imunes mediados por células.

Em outra modalidade, a primeira e segunda moléculas da superfície celular às quais uma molécula de ligação biespecífica é capaz de ligar

estão localizadas sobre a mesma célula ou tipo de célula. Por reticulação dos primeiro e segundo receptores sobre a mesma célula, as moléculas de ligação biespecífica da invenção podem inibir ou reforçar uma atividade (por exemplo, atividade de transdução de sinal) associada com um ou ambos o primeiro e segundo receptores. Em uma modalidade a primeira e segunda moléculas da superfície celular são do mesmo tipo (por exemplo, são da mesma família de moléculas). Em outra modalidade a primeira e segunda moléculas da superfície celular referidas são de tipos distintos (por exemplo, são de diferentes famílias de moléculas). Receptores da superfície celular típicos aos quais as moléculas de ligação biespecífica ligam incluem antígenos de células tumorais ou receptores de células imunes. Receptores da superfície celular típicos incluem quaisquer dos receptores de citocina, moléculas de adesão, ou receptores de fator de crescimento descritos na seção B acima.

Em uma modalidade, moléculas alvo típicas às quais uma molécula de ligação da invenção liga incluem um ou mais epítopes de, por exemplo, sulfato de heparina, fatores de crescimento ou seus receptores (por exemplo, receptor de fator de crescimento epidérmico, receptor de fator de crescimento similar a insulina, receptor de fator de crescimento de hepatócitos (HGF/SF) (Vide, por exemplo, Cao et al. Proc. Natl. Acad. Sci. 2001. 98:7443; Lu et al. 2004. J. Biol. Chem. 279:2856).

Em uma modalidade exemplar, no mínimo uma das moléculas às quais uma molécula de ligação biespecífica da invenção liga é um membro da família de receptores de TNF (TNFR). Em outra modalidade exemplar, a primeira e segunda moléculas alvo às quais uma molécula de ligação biespecífica da invenção liga são ambas membros da família de TNFR. Em outra modalidade, uma molécula de ligação da invenção liga a um ligante da família de TNFR. Em ainda outra modalidade, uma molécula de ligação da invenção liga a um membro da família de TNFR e um antígeno expressado sobre a superfície de uma célula tumoral, por exemplo, preferencialmente expressada sobre uma célula tumoral. O fator limitante no tratamento de tumores com moléculas de ligação TNFR monoespecíficas é que freqüente-

mente somente um subgrupo de tumores parece ser sensível a similares terapias. Moléculas de ligação TNFR multiespecíficas podem especificamente ativar TNFRs, e reforçar sinalização de receptor, por exemplo, trazendo os TNFRs em íntima proximidade. A invenção proporciona moléculas de ligação de TNFR biespecíficas aprimoradas as quais podem direcionar mais de um TNFR ou tipo de TNFR e reforçar sinalização, deste modo proporcionando um método aprimorado para tratar câncer. Em uma modalidade, a molécula de ligação de TNFR biespecífica aumenta a força do sinal ligando a dois ou mais TNFRs do mesmo tipo aumentando o número de TNFRs sendo trazidos juntos. Em outra modalidade mais preferencial, a molécula de ligação de TNFR biespecífica é capaz de ligar a dois receptores da família de TNF diferentes.

Em uma modalidade, no mínimo um dos TNFRs aos quais a molécula de ligação de TNFR biespecífica liga contém um domínio de morte. O termo "domínio de morte" se refere a uma região citoplasmática de um receptor da família de TNF o qual está envolvido em morte celular ou sinalização apoptótica mediadas por TNF e indução de citotoxicidade celular mediada por estes receptores. Esta região acopla o receptor para ativação de caspase através de proteínas adaptoras resultando em ativação do caminho de morte extrínsico.

Exemplos de receptores de TNF os quais contêm domínios de morte incluem, mas não estão limitados a, TNFR1 (TNFRSF1A), Fas (TNFRSF6), DR-3 (TNFRSF6B), LNGFR (TNFRSF16) TRAIL-R1 (TNFRSF10A), TRAIL-R2 (TNFRSF10B) e DR6 (TNFRSF21). A sinalização apoptótica destes receptores é modulada na ligação de um ligante cognato e formação de quaisquer um dos seguintes pares de receptor-ligante: TNFR1/TNFα, Fas/FasL, DR-3/DR-3LG, TRAIL-R1/TRAIL, ou TRAIL-R2/TRAIL.

Moléculas de ligação de TNFR biespecíficas que direcionam receptores da família de TNF alvo contendo domínios de morte são úteis para o tratamento de câncer uma vez que os TNFRs deste tipo são freqüentemente superexpressados sobre células tumorais e estimulação do receptor pode ativar apoptose de células tumorais. Em modalidades preferenciais, o

TNFR contendo domínio de morte ao qual a molécula de ligação de TNFR biespecífica da invenção liga é TRAIL-R2. TRAIL-R2 é preferencial para terapia de tumor humano uma vez que sua ativação não dispara apoptose de hepatócitos e portanto pode ter reduzida toxicidade.

Apesar da ativação de alguns receptores contendo domínio de morte, por exemplo, TNFR1 ou Fas, ter sido tóxica em aplicações *in vivo*, é provável que amarrando estes receptores a outros receptores de TNF possa diminuir a toxicidade e deste modo tornar um anticorpo tóxico menos tóxico.

Em uma modalidade, uma molécula de ligação de TNFR biespecífica da invenção compreende no mínimo um sítio de ligação dirigido para um TNFR contendo um domínio de morte e no mínimo um sítio de ligação dirigido para um TNFR carecendo de um domínio de morte.

Em algumas modalidades exemplares, TNFRs carecendo de um domínio de morte incluem TNFRs envolvidas em diferenciação tecidual. Exemplos de receptores de TNFR envolvidos em diferenciação tecidual incluem LTβR, RANK, EDAR1, XEDAR, Fn14, Troy/Trade, e NGFR. TNFRs envolvidos em diferenciação tecidual podem influenciar diferenciação tecidual depois de ligação de um ligante cognato. Moléculas de ligação de TNFR que direcionam TNFRs envolvidos em diferenciação tecidual podem afetar tumores de vários modos. Primeiro, têm o potencial de retardar diretamente o crescimento tumoral alterando a progressão do ciclo celular. Segundo, diferenciação tecidual no contexto de transformação de células tumorais pode levar ao conflito do ciclo celular e falha de apoptose. Terceiro, similar input conflitante pode tornar uma célula mais sensível a quimioterapia.

Em algumas modalidades preferenciais, TNFR envolvido em diferenciação tecidual é β receptor de linfotoxina (LTβR). LTβR está envolvido no controle do estado de maturação de várias céulas estromais especializadsa no sistema imune e desempenha um papel crítico durante o desenvolvimento dos elementos estromais do *anlagen* de linfonodo. Foi proposto que a ativação de um programa de desenvolvimento em células epiteliais ou fibroblastóides no contexto de uma célula transformada é prejudicial para sua sobrevida e esta ação pode ser responsável por parte da atividade anti-

tumoral de ativação de LTβR. Estes receptores também podem iniciar programas inflamatórios que envolvem liberação de quimocina ou promovem reações imunológicas antitumorais. A liberação referida pode afetar o estado inflamatório do tumor e/ou invocar infiltração de elementos linfóides promovendo uma reação imunológica ao tumor. Portanto, moléculas de ligação de TNFR biespecíficas as quais ligam LTβR, sozinho ou em combinação com receptores de TNF contendo domínios de morte (por exemplo, TRAIL-R2), são englobadas pela invenção.

Em algumas modalidades exemplares, os TNFRs carecendo de um domínio de morte incluem TNFRs envolvidos em regulação imune. Os receptores referidos incluem TNFR2, HVEM, CD27, CD30, CD40, 4-1BB, OX40, GITR, TACI, BAFF-R, BCMA, e RELT. Receptores da família de TNF adicionais envolvidos em regulação imune incluem TRAIL-R3 e TRAIL-R4.

Outros receptores da família de TNF alvo com um papel na formação do tumor podem ser identificados usando bancos de dados de RNA existentes de expressão de receptores em vários tipos de células os quais possibilitam que uma pessoa defina receptores da família de TNF que estão presentes ou são idealmente superexpressados sobre vários tumores. Além disso, bancos de dados de RNA existentes proporcionam uma vantagem adicional em que o par de receptores da família de TNF ao qual uma molécula de ligação de TNFR biespecífica da invenção liga pode ser otimizado identificando os pares de receptores que são mais unicamente expressados sobre um tipo tumoral ou subgrupo de tumores mas não são abundantes sobre tecidos normais, especialmente fígado e vasculatura. Em uma maneira similar são identificados pares de receptores (ou mais) que podem liberar um sinal potente para o tumor e evitar tecidos normais.

Métodos para produzir moléculas multiespecíficas são de conhecimento geral na técnica. Por exemplo, tecnologia recombinante pode ser usada para produzir moléculas multiespecíficas, por exemplo, diabodies, diabodies de cadeia única, scFvs tandem, e etc. Técnicas típicas para produzir moléculas multiespecíficas são conhecidas na técnica (por exemplo, Kontermann et al. Methods in Molecular Biology Vol. 248: Antibody Enginee-

ring: Methods e Protocols. Pp 227-242 Publicação do Pedido de patente dos Estados Unidos Nº US 2003/0207346 A1 e as referências citadas na mesma). Em uma modalidade, algumas moléculas multiespecíficas multiméricas são preparadas usando métodos tais como os descritos por exemplo, na publicação do Pedido de patente dos Estados Unidos Nº US 2003/0207346 A1 ou na publicação de patente dos Estados Unidos Nº 5.821.333, ou na publicação de patente dos Estados Unidos Nº US2004/0058400.

Em outra modalidade, uma molécula de ligação multiespecífica da invenção é uma proteína de fusão multiespecífica. Conforme usado aqui, neste Pedido de patente, a expressão "proteína de fusão multiespecífica" designa proteínas de fusão (conforme definido nas partes que se seguem) tendo no mínimo duas especificidades de ligação (isto é, combinar alguns domínios de ligação de um ligante ou receptor). Proteínas de fusão multiespecíficas podem ser montadas, por exemplo, como heterodímeros, heterotrímeros ou heterotetrâmeros, essencialmente conforme descrito na publicação de patente internacional Nº WO 89/02922 (publicada em 6 de abril de 1989), na patente européia Nº EP 314, 317 (publicada em 3 de maio de 1989), e na Patente dos Estados Unidos Nº 5.116.964 emitida em 2 de maio de 1992. Proteínas de fusão multiespecíficas preferenciais são biespecíficas. Exemplos de proteínas de fusão biespecíficas incluem CD4-scFv / TNF receptor-IgG e CD4-scFv/L-selectina-IgG. A última molécula mencionada combina a função de ligação de linfonodo do receptor residente de linfócitos (L-HR, L-selectina), e a função de ligação de HIV de CD4, e encontra aplicação potencial na prevenção ou no tratamento de infecção por HIV, condições relacionadas, ou como um diagnóstico.

Em outra modalidade, a invenção pertence a moléculas de ligação multiespecífica estabilizadas, por exemplo, moléculas de ligação biespecífica, por exemplo, anticorpos, os quais incorporam no mínimo um sítio de ligação que liga a um alvo conhecido e no mínimo um sítio de ligação o qual reconhece um alvo desconhecido (por exemplo, em uma modalidade, a molécula biespecífica incorpora sítios de ligação selecionados entre uma biblioteca de display de fago de anticorpos semi-sintéticos) e um scFv estabilizado

da invenção.

Em uma modalidade da invenção, uma pessoa com conhecimento regular da técnica pode iniciar com um anticorpo de cadeia única de conhecida especificidade e construir uma biblioteca de Fab usando técnicas conhecidas na técnica ou, alternativamente, o técnico versado pode iniciar com um fragmento Fab de conhecida especificidade e construir uma biblioteca de cadeia única estabilizada usando técnicas conhecidas na técnica. É de conhecimento na técnica que bibliotecas de fontes não imunizadas e preparadas por recombinação sintética de seqüências de V-gene (preferencialmente recombinação de VH com, DH e JH, e VL com seqüências JL) podem ser usadas para isolar anticorpos para qualquer antígeno. Por exemplo, o Pedido de patente internacional Nº WO92/01047 ensina que fragmentos de anticorpos podem ser apresentados sobre a superfície de bacteriófago e que ligarão antígeno. Fragmentos de anticorpos (por exemplo, Fab, Fv, scFv e VH) podem ser selecionados diretamente usando esta característica. Outros métodos conhecidos na técnica incluem os ensinados, por exemplo, nas Patentes dos Estados Unidos Nos. 5.698.426; 6.291.159; 5.658.727; 5.667.988; e 5.969.108.

Em outra modalidade, scFv o qual reconhece um alvo conhecido pode ser dimerizado com scFv isolado de uma biblioteca de display de anticorpo de fago humano semi-sintético. (vide, por exemplo, Kruif e Logtenberg 1996. J. Biol. Chem. 271:7630).

Em uma modalidade, a molécula multiespecífica em questão é expressada em um sistema de expressão usado para expressar moléculas de anticorpos, por exemplo células de mamíferos, levedura tais como Picchia, E. coli, Bacculovirus, e etc. Em uma modalidade, a molécula biespecífica em questão é expressada no sistema de vetor NEOSPLA (vide, por exemplo, a Patente dos Estados Unidos Nº 6.159.730). Este vetor contém o promotor / reforçador de citomegalovírus, o promotor principal de β globina de camundongo, a origem de replicação SV40, a seqüência de poliadenilação de hormônio de crescimento bovino, éxon 1 e éxon 2 de fosfotransferase de neomicina, o gene de diidrofolato redutase e seqüência líder.

Em uma modalidade, as moléculas multiespecíficas em questão compreendem um peptídeo de ligação sintético.

Estas moléculas multiespecíficas têm um ou mais sítios de ligação para um alvo conhecido e expressam uma biblioteca em um ou mais sítios de ligação. As moléculas multiespecíficas referidas podem ser usadas, por exemplo, para identificar moléculas em íntima proximidade com ou associadas com o alvo conhecido. Por exemplo, o técnico versado pode usar as moléculas multiespecíficas em questão em um ensaio para selecionar aquelas que induzem uma resposta particular, por exemplo, apoptose ou ativação celular, usando métodos de triagem de conhecimento geral na técnica. A molécula biespecífica identificada como produzindo a resposta triaeda pode ser então identificada e sua especificidade determinada. Usando os métodos referidos é possível identificar moléculas em íntima associação com alvos de interesse em particular, por exemplo, marcadores de células T ou outras moléculas de sinalização (tais como CRIPTO-I, moléculas de domínio de morte, ou moléculas envolvida em apoptose). A proximidade do alvo conhecido e a molécula recém identificada como um "vizinho mais próximo" pode ser confirmada usando imunoprecipitação ou outras técnicas de conhecimento daqueles versados na técnica. Usando estes métodos é possível identificar moléculas como alvos para modular uma reação celular em particular.

Especificidades de ligação compreendendo sítios de reconhecimento de antígeno ou regiões variáveis inteiras de molécula de ligação multiespecífica, em particular anticorpos multiespecíficos ou variantes de anticorpos da invenção podem ser derivadas de um ou mais anticorpos parentais. Os anticorpos parentais podem incluir anticorpos que ocorrem naturalmente ou fragmentos de anticorpos, anticorpos ou fragmentos de anticorpos adaptados de anticorpos que ocorrem naturalmente, anticorpos construídos *de novo* usando seqüências de anticorpos ou fragmentos de anticorpos conhecidos por serem uma molécula alvo específica. Seqüências que podem ser derivadas de anticorpos parentais incluem regiões variáveis de cadeia pesada e/ou leve e/ou CDRs, regiões de framework ou outras porções das mesmas.

Em uma modalidade exemplar da invenção, os anticorpos parentais usados para construir uma molécula de ligação TNFR multiespecífica são um anticorpo anti-TRAIL-R2, por exemplo 14A2, e um anticorpo anti-LTβR, por exemplo CBE11 ou BHA10. Anticorpos multivalentes, multiespecíficos podem conter uma cadeia pesada compreendendo duas ou mais regiões variáveis e/ou uma cadeia leve compreendendo uma ou mais regiões variáveis em que no mínimo duas das regiões variáveis reconhecem diferentes epítopes de LTβR.

Moléculas de ligação de TNFR multispecíficas, por exemplo, biespecíficas podem sr construídas em uma variedade de diferentes modos usando uma variedade de diferentes seqüências derivadas de anticorpos anti-LTβR parentais, incluindo BHA10 murino ou humanizado (Browning et al., *J. Immunol.* 154: 33 (1995); Browning et al. *J. Exp. Med.* 183:867 (1996)), CBE11 murino ou humanizado (U.S. Patent 6,312,691 e WO 02/30986, respectivamente), e/ou anti-TRAIL-R2 parental murino ou quimérico 14A2. Exemplos de anticorpos anti-LTβR os quais podem ser usados para as moléculas de ligação da invenção de TNFR biespecíficas incluem: BKA11, CDH10, BCG6, AGH1, BDA8, CBE11 e BHA10 ou BHA10. As seguintes linhagens de células de hibridoma produzindo anticorpos anti- LT-β-R monoclonais podem ser usadas para produzir anticorpos anti-LTβR dos quais derivar seqüências de construtos de anticorpos, os quais foram previamente depositados com a American Type Culture Collection (ATCC) de acordo com as provisões do Tratado de Budapest e foram atribuídas os números de acesso ATCC indicados:

Linhagem Celular	Nome mAb	N° de Acesso
a) AG.H1.5.1	AGH1	HB 11796
b) BD.A8.AB9	BDA8	HB 11798
c) BC.G6.AF5	BCG6	B 11794
d) BH.A10	BHA10	B 11795
e) BK.A11.AC10	BKA11	B 11799
f) CB.E11.1	CBE11	B 11793
g) CD.H10.1	CDH10	B 11797

Outros exemplos de anticorpos receptores anti-TNF os quais podem ser usados nas moléculas de ligação de TNFR multiespecíficas da invenção incluem anticorpos dirigidos para receptores de TNF contendo um domínio de morte. Uma série de anticorpos foram gerados para domínio de morte contendo receptores de TNF e são de conhecimento geral na técnica. Os anticorpos referidos incluem anticorpos monoclonais anti-TNF-R1 (sistemas R&D anti-TNF-R1; Tularik mAb nº 985, Patentes dos Estados Unidos Nos. 6.110.690; 6.437.113), anti-Fas receptor mAb CH-11 (Patente dos Estados Unidos nº 6.312.691; publicação de patente internacional nº WO 95/10540), anticorpos anti-DR3 (Patente dos Estados Unidos nº 5.985.547; Johnson, et al. (1984) ImmunoBiology of HLA, ed. Dupont, B.O., Springer, New York; Patentes dos Estados Unidos Nos. 6.462.176; 6.469.166), e anticorpos anti-TRAIL-R (Patentes dos Estados Unidos Nos. 5.763.223; 6.072.047; 6.284.236; 6.521.228; 6.569.642; 6.642.358; e Patente dos Estados Unidos nº 6.417.328).

Uma série de anticorpos também foram cultivados para receptores de TNF envolvidos em diferenciação tecidual e são conhecidos na técnica. Exemplos de anticorpos anti-receptores de TNF específicos para receptores de TNF envolvidos em diferenciação tecidual incluem: anticorpos monoclonais anti-RANK (Immunex – Patentes dos Estados Unidos Nos. 6.562.948; 6.537.763; 6.528.482; 6.479.635; 6.271.349; 6.017.729; Komed – WO 03/080671), anticorpos anti-EDAR policlonais (anti-humanos) e monoclonais (anticamundongo) (R&D Systems – MAB745, BAF157; Elomaa *et al.* (2001) Human Molecular Genetics. 10:953), anticorpos anti-XEDAR monoclonais e policlonais (R&D Systems – MAB1093 e AF1093), anticorpos anti-Fn14 monoclonais (Nakayama *et al.* (2003) J. Immunology 170:341; clones ITEM-1, ITEM-2, e ITEM-4 disponíveis na eBioscience), anticorpo anti-TROY (T3323 da Sigma-Aldrich), e anticorpos anti-NGFR (anti-roedor) (Chemicon USA).

Uma série de anticorpos também foram cultivados para receptores de TNF envolvidos em regulação imune e são conhecidos na técnica. Exemplos de anticorpos anti-receptores de TNF específicos para receptores

de TNF envolvidos em regulação imune incluem: anticorpos anti-HVEM (HGSI – publicação de patente internacional Nº WO 03/086301), anticorpos anti-CD40 (Biogen – publicação de patente internacional Nº WO 97/20063; Chiron – Patentes dos Estados Unidos Nºs 5.677.165; 5.874.082; 6.004.552; 6.056.959; 6.315.998; Publicação do Pedido de patente dos Estados Unidos Nº 2002/0106371; Publicação do Pedido de patente dos Estados Unidos Nos. 2003/0059427; US20030118588A1; 2003/0211100A1; US2002020142358A1; Patentes dos Estados Unidos Nos. US 6312693; US 6051228; Fanslow et al. – Patente dos Estados Unidos Nº US 5801227), anti-4-1BB (publicação de patente internacional Nº PCT WO 03/084999; Patente Européia Nº EP 0948353; Patente dos Estados Unidos Nº 6210669; Genecraft – publicação de patente internacional Nº WO 03/083069), e anticorpos anti-BAFF-R (policlonal de coelho – catálogo ProSci nº 3097), entre muitos outros anticorpos cultivados para receptores de regulação imune.

Uma variedade de outros construtos de anticorpos multivalentes podem ser desenvolvidos por uma pessoa versada na técnica usando técnicas de DNA recombinante de rotina, por exemplo conforme descrito no Pedido de patente Internacional Nº PCT/US86/02269; no Pedido de patente Européia Nº 184,187; no Pedido de patente Européia Nº 171,496; no Pedido de patente Européia Nº 173,494; na Publicação de Patente Internacional PCT Nº WO 86/01533; na Patente dos Estados Unidos Nº 4.816.567; no Pedido de patente Européia Nº 125,023; Better *et al.* (1988) *Science* 240:1041-1043; Liu *et al.* (1987) *Proc. Natl. Acad. Sci. USA* 84:3439-3443; Liu *et al.* (1987) *J. Immunol.* 139:3521-3526; Sun *et al.* (1987) *Proc. Natl. Acad. Sci. USA* 84:214-218; Nishimura *et al.* (1987) *Cancer Res.* 47:999-1005; Wood *et al.* (1985) *Nature* 314:446-449; Shaw *et al.* (1988) *J. Natl. Cancer Inst.* 80:1553-1559); Morrison (1985) *Science* 229:1202-1207; Oi *et al.* (1986) *BioTechniques* 4:214; na Patente dos Estados Unidos Nº 5,225,539; Jones *et al.* (1986) *Nature* 321:552-525; Verhoeyan *et al.* (1988) *Science* 239:1534; Beidler *et al.* (1988) *J. Immunol.* 141:4053-4060; e Winter e Milstein, Nature, 349, pp. 293-99 (1991)). Preferencialmente anticorpos não -humanos são "humanizados" ligando o domínio de ligação de antígeno

não-humano com um domínio constante humano (por exemplo, Cabilly et al., Patente dos Estados Unidos Nº 4.816.567; Morrison et al., Proc. Natl. Acad. Sci. U.S.A., 81, pp. 6851-55 (1984)).

Outros métodos os quais podem ser usados para preparar construtos de anticorpos multivalentes são descritos nas seguintes publicações: Ghetie, Maria-Ana et al. (2001) *Blood* 97:1392-1398; Wolff, Edith A. et al. (1993) *Cancer Research* 53:2560-2565; Ghetie, Maria-Ana et al. (1997) *Proc. Natl. Acad. Sci. 94:7509-7514; Kim, J.C. et al. (2002) Int. J. Cancer* 97(4): 542-547; Todorovska, Aneta et al. (2001) *Journal of Immunological Methods* 248:47-66; Coloma M.J. et al. (1997) *Nature Biotechnology* 15:159-163; Zuo, Zhuang et al. (2000) *Protein Engineering (Suppl.)* 13(5): 361-367; Santos A.D., et al. (1999) *Clinical Cancer Research 5:3118s-3123s; Presta, Leonard G.* (2002) *Current Pharmaceutical Biotechnology* 3:237-256; van Spriel, Annemiek et al., (2000) *Review Immunology Today* 21(8) 391-397.

XI. Expressão de Moléculas de Ligação

Depois de manipulação do material genético isolado proporcionar polipeptídeos da invenção conforme determinado acima, os genes são tipicamente inseridos em um vetor de expressão para introdução em células hospedeiras que podem ser usadas para produzir a quantidade desejada de polipeptídeo que, por sua vez, proporciona as moléculas de ligação reivindicadas.

O termo "vetor" ou "vetor de expressão" é usado aqui, neste Pedido de patente, para os fins da relatório descritivo e reivindicações, para indicar vetores usados de acordo com a presente invenção como um veículo para introduzir em e expressar um gene desejado em uma célula. Conforme de conhecimento daqueles versados na técnica, similares vetores podem ser facimente selecionados entre o grupo consistindo em plasmídeos, fagos, vírus e retrovírus. Em geral, vetores compatíveis com a presente invenção compreenderão um marcador de seleção, sítios de restrição apropriados para facilitar clonagem do gene desejado e a capacidade para penetrar e/ou replicar em células eucarióticas ou procarióticas.

Para os fins desta invenção, podem ser empregados numerosos

sistemas de vetor de expressão. Por exemplo, uma classe de vetor utiliza elementos de DNA os quais são derivados de vírus animais tais como papiloma vírus bovino, polioma vírus, adenovírus, vírus da varíola bovina, baculovírus, retrovírus (RSV, MMTV ou MOMLV) ou vírus SV40. Outros envolvem o uso de sistemas policistrônicos com sítios de ligação de ribossoma internos. Adicionalmente, células as quais têm integrado o DNA em seus cromossomos podem ser selecionadas introduzindo um ou mais marcadores os quais possibilitam seleção de células hospedeiras transfectadas. O marcador pode proporcionar prototrofia para um hospedeiro auxotrófico, resistência a biocida (por exemplo, antibióticos) ou resistência a metais pesados tais como cobre. O gene marcador selecionável pode ser ou diretamente encadeado às seqüências de DNA para ser expressado, ou introduzido na mesma célula por cotransformação. Elementos adicionais também podem ser necessários para síntese otimizada de mRNA. Estes elementos podem incluir seqüências de sinal, sinais de junção, bem como promotores transcricionais, reforçadores, e sinais de terminação. Em modalidades particularmente preferenciais os genes de região variável clonados são inseridos em um vetor de expressão junto com os genes de região constante de cadeia pesada e leve (preferencialmente humanos) sintéticos conforme discutido acima. Preferencialmente, isto é efetuado usando um vetor de expressão proprietário da IDEC, Inc., referido como NEOSPLA (Patente dos Estados Unidos Nº 6.159.730). Este vetor contém o promotor / reforçador de citomegalovírus, o promotor principal de β globina de camundongo, a origem de replicação SV40, a seqüência de poliadenilação de hormônio de crescimento bovino, neomicina fosfotransferase éxon 1 e éxon 2, o gene diidrofolato redutase e seqüência líder. Conforme visto nos exemplos abaixo, foi visto que este vetor resulta em nivel de expressão de anticorpos muito elevado na incorporação de genes de regiões variável e constante, transfecção em células CHO, seguida por seleção em meio contendo G418 e amplificação por metotrexato. Sistemas vetoriais também são ensinadas nas Patentes dos Estados Unidos Nos. 5.736.137 e 5.658.570, cada uma das quais é incorporada por meio de referência em sua totalidade aqui, neste Pedido de patente. Este

sistema proporciona altos níveis de expressão, por exemplo, > 30 pg/célula/dia. Outros sistemas vetoriais típicos são revelados, por exemplo, na Patente dos Estados Unidos Nº 6.413.777.

Em outras modalidades preferenciais os polipeptídeos da invenção da presente invenção podem ser expressados usando construtos policistrônicos tais como aqueles revelados no Pedido de patente provisória dos Estados Unidos copendente Nº 60/331.481 depositado em 16 de novembro de 2001 e incorporado aqui, a este Pedido de patente, em sua totalidade. Nestes novos sistemas de expressão, múltiplos produtos genéticos de interesse tais como cadeias pesadas e leves de anticorpos podem ser produzidos a partir de um único construto policistrônico. Estes sistemas usam vantajosamente um sítio de ingresso de ribossoma interno (IRES) proporcionando níveis relativamente elevados de polipeptídeos da invenção em células hospedeiras eucarióticas. Seqüências IRES compatíveis são reveladas na Patente dos Estados Unidos Nº 6.193.980 a qual também é incorporada aqui, a este Pedido de patente. Aqueles versados na técnica reconhecerão que os sistemas de expressão referidos podem ser usados para produzir de modo eficaz toda a γ de polipeptídeos revelados no presente requerimento.

Mais geralmente, uma vez que o vetor ou seqüência de DNA codificando uma subunidade monomérica da molécula de ligação (por exemplo, um anticorpo modificado) tenha sido preparado, o vetor de expressão pode ser introduzido em uma célula hospedeira apropriada. Isto é, as células hospedeiras podem ser transformadas. A iIntrodução do plasmídeo na célula hospedeira pode ser realizada por várias técnicas de conhecimento geral daqueles versados na técnica. Estas incluem, mas não estão limitadas a, transfecção (incluindo eletroforese e eletroporação), fusão de protoplasto, precipitação de fosfato de cálcio, fusão celular com DNA envelopado, microinjeção, e infecção com vírus intacto. Vide, Ridgway, A. A. G. "*Mammalian Expression Vectors*" Chapter 24.2, pp. 470-472 Vectors, Rodriguez e Denhardt, Eds. (Butterworths, Boston, Mass. 1988). O mais preferencialmente, a introdução de plasmídeo no hospedeiro é através de eletroporação. As células transformadas são cultivadas sob condições apropriadas para a pro-

dução das cadeias leves e cadeias pesadas, e avaliadas para síntese de proteína de cadeia pesada e/ou leve. Técnicas de ensaio típicas incluem ensaio imunossorvente ligado a enzima (ELISA), radioimunoensaio (RIA), ou análise por classificador celular ativada por flourescência (FACS), imunohistoquímica e similares.

Conforme usado aqui, neste Pedido de patente, o termo "transformação" será usado em um sentido amplo para se referir à introdução de DNA em uma célula hospedeira receptora que muda o genótipo e conseqüentemente resulta em uma mudança na célula receptora.

Ao longo destas mesmas linhagens, "células hospedeiras" se refere a células que tenham sido transformadas com vetores construídos usando recombinant DNA técnicas e codificando no mínimo um gene heterólogo. Em descrições de processos para isolamento de polipeptídeos a partir de hospedeiros recombinantes, os termos "célula" e "cultura celular" são usados de modo intercambiável para denotar a fonte de anticorpo a menos que seja claramente especificado de modo diverso. Em outras palavras, a recuperação de polipeptídeo a partir das "células" pode significar ou a partir de células inteiras centrifugadas, ou a partir da cultura celular contendo tanto o meio quanto as células suspendidas.

Em uma modalidade, a linhagem de células hospedeiras usadas para expressão de proteína (por exemplo, de moléculas multivalentes de ligação) é de origem de mamífero; os versados na técnica têm a capacidade de determinar preferencialmente linhagens de células hospedeiras particulares as quais são melhor adequadas para o produto genético desejado a ser expressado nas mesmas. Linhagens de células hospedeiras incluem, mas não estão limitados a, DG44 e DUXB11 (linhagens de ovário de hamster chinês CHO, menos DHFR), HELA (carcinoma cervical humano), CVI (linhagem de rim de macado), COS (um derivado de CVI com SV40 T antígeno), R1610 (fibroblasto de hamster chinês) BALBC/3T3 (fibroblasto de camundongo), HAK (linhagem de rim de hamster), SP2/O (mieloma de camundongo), P3.times.63-Ag3.653 (mieloma de camundongo), BFA-1c1BPT (células endoteliais bovinas), RAJI (linfócito humano) e 293 (rim humano). Em uma

modalidade podem ser usadas células NS0. Células CHO são particularmente preferenciais. Linhagens de células hospedeiras estão tipicamente disponíveis em serviços comerciais, a American Tissue Culture Collection ou a partir da literatura publicada.

A produção *in vitro* possibilita aumento em escala para dar grandes quantidades dos polipeptídeos desejados. Técnicas para cultura de células mamíferas sob condições de cultura tecidual são conhecidas na técnica e incluem cultura de suspensão homogênea, por exemplo, em um reator de erguimento de ar ou em um reator de agitador contínuo, ou cultura de células imobilizadas ou atraídas, por exemplo, em fibras ocas, microcápsulas, sobre microcontas de agarose ou cápsulas de cerâmica. Caso necessário e/ou desejado, as soluções de polipeptídeos podem ser purificadas pelos métodos de cromatografia habituais, por exemplo, filtração por gel, cromatografia de permuta iônica, cromatografia sobre DEAE-celulose ou cromatografia de (imuno-)afinidade, por exemplo, depois de biosíntese preferencial de um polipeptídeo da região de articulação sintético ou antes de ou subseqüente à etapa de cromatografia HIC descrita aqui, neste Pedido de patente.

Genes codificando o polipeptídeo da invenção também podem ser expressados em células não-mamíferas tais como células de bactérias ou leveduras ou plantas. A este respeito será reconhecido que vários microorganismos não-mamíferos unicelulares tais como bactérias também podem ser transformados; isto é aqueles capazes de serem cultivados em culturas ou fermentação. Bactérias, as quais são suscetíveis a transformação, incluem membros das enterobacteriaceae, tais como cepas de Escherichia coli ou Salmonella; Bacillaceae, tais como Bacillus subtilis; Pneumococcus; Streptococcus, e Haemophilus influenzae. Será adicionalmente reconhecido que, quando expressados em bactérias, os polipeptídeos tipicamente se tornam parte de corpos de inclusão. Os polipeptídeos devem ser isolados, purificados e em seguida montados em moléculas funcionais. Onde formas tetravalentes de anticorpos são desejadas, as subunidads serão então automontadas em anticorpos tetravalentes (publicação do Pedido de patente internacional Nº WO 02/096948A2).

Além de procariotas, micróbios eucarióticos também podem ser usados. Saccharomyces cerevisiae, ou levedura de baker comum, é o mais comumente usado entre microorganismos eucarióticos embora uma série de outras cepas estejam disponíveis comumente. Para expressão em Saccharomyces, o plasmídeo YRp7, por exemplo, (Stinchcomb et al., Nature, 282:39 (1979); Kingsman et al., Gene, 7:141 (1979); Tschemper et al., Gene, 10:157 (1980)) é comumente usado. Este plasmídeo já contém o gene TRP1 o qual proporciona um marcador de seleção para uma cepa mutante de levedura carecendo da capacidade para cultivar em triptofano, por exemplo ATCC Nº 44076 ou PEP4-1 (Jones, Genetics, 85:12 (1977)). A presença da lesão trpl como uma característica do genoma da célula hospedeira de levedura então proporciona um ambiente eficaz poara detectar transformação por crescimento na ausência de triptofano.

XII. Marcação ou Conjugação de Moléculas de Ligação

As moléculas de ligação da presente invenção podem ser usadas em forma não-conjugada ou podem ser conjugadas a no mínimo um de uma variedade de moléculas efetoras, isto é, funcionais, por exemplo, para facilitar a detecção alvo ou para estudo por imagens ou terapia do paciente. Os polipeptídeos da invenção podem ser marcados ou conjugados quer antes ou depois da purificação, quando é realizada purificação. Em particular, os polipeptídeos da presente invenção podem ser conjugados a citotoxinas (tais como radioisótopos, fármacos citotóxicas, ou toxinas) agentes terapêuticos, agentes citostáticos, toxinas biológicas, pró-farmacos, peptídeos, proteínas, enzimas, vírus, lipídeos, modificadores da resposta biológica, agentes farmacêuticos, ligantes imunologicamente ativos (por exemplo, linfocinas ou outros anticorpos em que a molécula resultante liga tanto à célula neoplásica quanto a uma célula efetora tal como uma célula T), PEG, ou moléculas detectáveis úteis em estudo por imagens. Em outra modalidade, um polipeptídeo da invenção pode ser conjugado a uma molécula que reduz a vascularização de tumores. Em outras modalidades, as composições reveladas podem compreender polipeptídeos da invenção acoplados a fármacos ou pró-farmacos. Ainda outras modalidades da presente invenção compre-

endem o uso de polipeptídeos da invenção conjugados a biotoxinas específicas ou seus fragmentos citotóxicos tais como ricina, gelonina, exotoxina de pseudomonas ou toxina diftérica. A seleção de qual polipeptídeo conjugado ou não conjugado usar dependerá do tipo e do estágio do câncer, do uso de tratamento auxiliar (por exemplo, quimioterapia ou radiação externa) e da condição do paciente. Será reconhecido que uma pessoa versada na técnica pode prontamente fazer uma seleção similar em vista dos ensinamentos aqui, neste Pedido de patente.

Será reconhecido que, em estudos prévios, anticorpos antitumorais marcados com isótopos têm sido usados com sucesso para destruir células em tumores sólidos bem como linfomas / leucemias em modelos animais, e em alguns casos em seres humanos. Radioisótopos típicos incluem: ^{90}Y, ^{125}I, ^{131}I, ^{123}I, ^{111}In, ^{105}Rh, ^{153}Sm, ^{67}Cu, ^{67}Ga, ^{166}Ho, ^{177}Lu, ^{186}Re e ^{188}Re. Os radionuclídeoos agem produzindo radiação ionizante a qual causa múltiplas rupturas de filamentos no DNA nuclear, levando à morte celular. Os isótopos usados para produzir conjugados terapêuticos tipicamente produzem α- ou β-partículas de alta energia as quais têm uma extensão de trajetória curta. Os radionuclídeos referidos destróem células à quais estão em íntima proximidade, por exemplo, células neoplásicas às quais o conjugado se anexou ou nas quais penetrou. Têm pouco ou nenhum efeito sobre células não-localizadas. Radionuclídeos são essencialmente não -imunogênicos.

Com respeito ao uso de conjugados radiomarcados em combinação com a presente invenção, polipeptídeos da invenção podem ser diretamente marcados (tais como através de iodação) ou podem ser marcados indiretamente através do uso de um agente quelante. Conforme usado aqui, neste Pedido de patente, as expressões "marcação indireta" e "abordagem de marcação indireta" ambas significam que um agente quelante é fixado de modo covalente a uma molécula de ligação e no mínimo um radionuclídeo é associado com o agente quelante. Os agentes quelantes referidos são tipicamente referidos como agentes quelantes bifuncionais uma vez que ligam tanto o polipeptídeo quanto o radioisótopo. Agentes quelantes particularmente preferenciais compreendem derivados de ácido 1-isotiocicmatobenzil-3-

metildioteleno triaminapentaacético ("MX-DTPA") e ácido ciclohexil dietilenotriamina pentaacético ("CHX-DTPA"). Outros agentes quelantes compreendem derivados de P-DOTA e EDTA. Radionuclídeos particularmente preferenciais para marcação indireta incluem ^{111}In e ^{90}Y.

Conforme usado aqui, neste Pedido de patente, as expressões "marcação direta" e "abordagem de marcação direta" ambas significam que um radionuclídeo é fixado de modo covalente diretamente a um polipeptídeo (tipicamente através de um resíduo aminoácido). Mais especificamente, estas tecnologias de ligação incluem marcação aleatória e marcação sítio-dirigida. No último caso, a marcação é dirigida em sítios específicos sobre o polipeptídeo, tais como os resíduos de açúcar N-encadeados presentes somente sobre a porção Fc dos conjugados. Além disso, várias técnicas e protocolos de marcação direta são compatíveis com a presente invenção. Por exemplo, polipeptídeos marcados com Tecnécio-99m podem ser preparados por processos de permuta de ligante, reduzindo pertecnato (TcO_4^-) com solução de íons estanoso, quelando o tecnécio reduzido sohre uma coluna Sephadex e aplicando os polipeptídeos a esta coluna, ou por técnicas de marcação em lote, por exemplo, incubando pertecnato, um agente redutor tal como $SnCl_2$, uma solução tampão tal como uma solução de ftalato de sódio-potássio, e os anticorpos. Em qualquer evento, radionuclídeos preferenciais para marcar anticorpos diretamente são de conhecimento geral na técnica e um radionuclídeo particularmente preferencial para marcação direta é ^{131}I anexado de modo covalente através de resíduos tirosina. Polipeptídeos de acordo com a invenção podem ser derivados, por exemplo, com iodeto de potássio ou sódio radioativo e um agente oxidante químico, tais como hipoclorito de sódio, cloramina T ou similares, ou um agente oxidante enzimático, tal como lactoperoxidase, glicose oxidase e glicose. No entanto, para os fins da presente invenção, a abordagem de marcação indireta é particularmente preferencial.

Patentes relativas a queladores e conjugados de queladores são conhecidas na técnica. Por exemplo, a Patente dos Estados Unidos Nº 4.831.175 de Gansow é dirigida para quelatos de ácido dietilenotriaminapen-

taacético polissubstituídos e conjugados protéicos contendo os mesmos, e métodos para sua preparação. As Patentes dos Estados Unidos Nos. 5.099.069, 5.246.692, 5.286.850, 5.434.287 e 5.124.471 de Gansow também se referem a quelatos de DTPA polissubstituídos. Estas patentes são incorporadas aqui, a este Pedido de patente, em sua totalidade. Outros exemplos de queladores de metais compatíveis são ácido etilenodiaminatetraacético (EDTA), ácido dietilenotriaminapentaacético (DPTA), 1,4,8,11-tetraazatetradecano, ácido 1,4,8,11-tetraazatetradecano-1,4,8,11-tetraacético, ácido 1-oxa-4,7,12,15-tetraazaheptadecano-4,7,12,15-tetraacético, ou similares. Ciclohexil-DTPA ou CHX-DTPA é particularmente preferencial e é exemplificado extensivamente abaixo. Ainda outros queladores compatíveis, incluindo aqueles ainda a serem descobertos, podem ser facilmente discernidos por um técnico versado e estão claramente dentro do âmbito da presente invenção.

Queladores compatíveis, incluindo o quelador bifuncional específico usado para facilitar quelação no Pedido de patente co-pendente Serial Nos. 08/475,813, 08/475,815 e 08/478,967, são preferencialmente selecionados proporcionando alta afinidade para metais trivalentes, apresentam aumentadas proporções de tumor-para-não-tumor e reduzida captação óssea bem como maior retenção *in vivo* de radionuclídeo em sítios alvo, isto é, sítios tumorais de linfoma de células B. No entanto, outros queladores bifuncionais que podem possuir ou não todas estas características são conhecidos na técnica e também podem ser benéficos em terapia tumoral.

Também será reconhecido que, de acordo com os ensinamentos aqui, neste Pedido de patente, polipeptídeos podem ser conjugados a diferentes radioetiquetas para fins de diagnóstico e terapêutico. Até este momento os Pedidos de patente co-pendentes supracitados, aqui, neste Pedido de patente, incorporados por meio de referência em sua totalidade, revelam conjugados terapêuticos radiomarcados para "estudo por imagens" para diagnóstico de tumores antes da administração de anticorpo terapêutico. Conjugado "In2B8" compreende um anticorpo monoclonal murino, 2B8, específico para antígeno CD20 humano, que é anexado a ^{111}In através de um quelador bifuncional, isto é, MX-DTPA (ácido dietilenotriaminapentaacé-

tico), o qual compreende uma mistura a 1:1 de 1-isotiocianatobenzil-3-metil-DTPA e 1-metil-3-isotiocianatobenzil-DTPA. ^{111}In é particularmente preferencial como um radionuclídeo diagnóstico porque entre cerca de 1 a cerca de 10 mCi pode ser administrado com segurança sem toxicidade detectável; e os dados de estudo por imagens são geralmente preditivos de distribuição de anticorpos marcados com ^{90}Y subseqüentes. A maioria dos estudos por imagens utilizam 5 mCi de anticorpo marcado com ^{111}In, porque esta dose é tanto segura quanto tem aumentada eficiência de estudo por imagens comparada com menores doses, com estudo otimizado por imagens ocorrendo em três a seis dias depois da administração de anticorpos. Vide, por exemplo, Murray, *J. Nuc. Med.* 26: 3328 (1985) e Carraguillo *et al.*, *J. Nuc. Med.* 26: 67 (1985).

Conforme indicado acima, uma variedade de radionuclídeos são aplicáveis à presente invenção e aqueles versados na técnica podem prontamente determinar qual radionuclídeo é o mais apropriado sob várias circunstâncias. Por exemplo, ^{131}I é um radionuclídeo de conhecimento geral usado para imunoterapia orientada. No entanto, a utilidade clínica de ^{131}I pode ser limitada por vários fatores incluindo: meia-vida física de oito dias; dealogenação de anticorpo iodado tanto no sangue quanto em sítios tumorais; e características de emissão (por exemplo, grande componente γ) as quais podem ser subótimas para deposição de dose localizada no tumor. Com o advento de agentes quelantes superiores, a oportunidade de fixar grupamentos quelantes metálicos a proteínas aumentou as oportunidades para utilizar outros radionuclídeos tais como ^{111}In e ^{90}Y. ^{90}Y proporciona vários benefícios para utilização em aplicações radioimunoterapêuticas: a meia-vida de 64 horas de ^{90}Y é suficientemente longa para possibilitar a acumulação de anticorpos pelo tumor e, diferentemente de, por exemplo, ^{131}I, ^{90}Y é um puro β emissor de alta energia sem irradiação γ associada em sua decomposição, com um alcance no tecido de 100 a 1.000 diâmetros celulares. Além disso, a quantidade mínima de radiação penetrante possibilita administração ambulatorial de anticorpos marcados com ^{90}Y. Adicionalmente, a internalização de anticorpo marcado não é necessária para destruição celu-

lar, e a emissão local de radiação ionizante pode ser letal para células tumorais adjacentes carecendo da molécula alvo.

Aqueles versados na técnica reconhecerão que estes conjugados não-radioativos também podem ser montados usando uma variedade de técnicas dependendo do agente selecionado a ser conjugado. Por exemplo, conjugados com biotina são preparados, por exemplo, reagindo os polipeptídeos com um éster ativado de biotina tal como o éster de N-hidroxissuccinimida de biotina. Similarmente, conjugados com um marcador fluorescente podem ser preparados na presença de um agente de ligação, por exemplo, os listados acima, ou por reação com um isotiocianato, preferencialmente fluoresceína-isotiocianato. Conjugados dos polipeptídeos da invenção com substâncias citostáticas / citotóxicas e quelatos de metal são preparados em uma maneira análoga.

Muitas moléculas efetoras carecem de grupamentos funcionais adequados aos quais anticorpos podem ser encadeados. Em uma modalidade, uma molécula efetora, por exemplo, um fármaco ou pró-farmaco é fixado ao anticorpo através de uma molécula de ligação. Em uma modalidade, a molécula de ligação contém uma ligação química que possibilita a ativação de citotoxicidade em um sítio particular. Ligações químcas adequadas são de conhecimento geral na técnica e incluem ligações dissulfeto, ligações lábeis em ácido, ligações fotolábeis, ligações lábeis em peptidase, ligações tioéter formadas entre grupamentos sulfidrila e maleimida, e ligações lábeis em esterase. O mais preferencialmente, a molécula de ligação compreende uma ligação dissulfeto ou uma ligação tioéter. De acordo com a invenção, a molécula de ligação preferencialmente compreende um grupamento químico reativo. Grupamentos químicos reativos particularmente preferenciais são ésteres de N-succinimidila e ésteres de N-sulfossuccinimidila. Em uma modalidade preferencial, o grupamento químico reativo pode ser ligado de modo covalente ao efetor através de ligação dissulfeto entre grupamentos tiol. Em uma modalidade uma molécula efetora é modificada para compreender um grupamento tiol. Uma pessoa versada na técnica reconhecerá que um grupamento tiol contém um átomo de enxofre ligado a um átomo de hidrogê-

nio e também é tipicamente referido na técnica como um grupamento sulfidrila, o qual pode ser denotado como "--SH" ou "RSH."

Em uma modalidade, uma molécula de ligação pode ser usada para unir a molécula efetora com a molécula de ligação. A molécula de ligação da invenção pode ser clivável ou não-clivável. Em uma modalidade, a molécula de ligação clivável é uma molécula de ligação redox-clivável, de tal modo que a molécula de ligação é clivável em ambientes com um menor potencial redox, tal como o citoplasma e outras regiões com maiores concentrações de moléculas com grupamentos sulfidrila livres. Exemplos de moléculas de ligação que podem ser clivadas devido a uma alteração no potencial redox incluem aquelas contendo dissulfetos. O estímulo clivante pode ser proporcionado na captação intracelular da proteína de ligação da invenção onde o menor potencial redox do citoplasma facilita clivagem da molécula de ligação. Em outra modalidade, uma redução no pH dispara a liberação da carga de maitansinóide para dentro da célula alvo. A redução no pH está implicada em muitos processos fisiológicos e patológicos, tais como tráfego de endossoma, crescimento tumoral, inflamação, e isquemia do miocárdio. O pH cai de um pH fisiológico 7.4 para 5 a 6 em endossomas ou 4 a 5 em lisossomas. Exemplos de moléculas de ligação sensíveis a ácido as quais podem ser usadas para ter por alvo lisossomas ou endossomas de células de câncer, incluem moléculas com ligações cliváveis por ácido tais como as encontradas em acetais, cetais, ortoésteres, hidrazonas, tritilas, cis-aconitilas, ou tiocarbamoílas (vide por exemplo, Willner *et al.*, (1993), *Bioconj. Chem.*, 4: 521-7; Patente dos Estados Unidos Nos. 4.569.789, 4.631.190, 5.306.809, e 5.665.358). Outras moléculas de ligação sensíveis a ácido típicas compreendem seqüências dipeptídicas Phe-Lys e Val-Lys (King *et al.*, (2002), *J. Med. Chem.*, 45: 4336-43). O estímulo clivante pode ser proporcionado no tráfego de captação intracelular para compartimentos endossômicos de baixo pH (por exemplo, lisossomas). Outras moléculas de ligação cliváveis por ácido típicas são as moléculas que contêm duas ou mais ligações cliváveis por ácido para fixação de dois ou mais maitansinóides (King *et al.*, (1999), *Bioconj. Chem.*, 10: 279-88; WO 98/19705).

Moléculas de ligação clivagens podem ser sensíveis a agentes de clivagem supridos biologicamente que estão associados com uma célula alvo particular, por exemplo, enzimas lisossômicas ou associadas com tumor. Exemplos de moléculas de ligação que podem ser clivadas enzimaticamente incluem, mas não estão limitadas a, peptídeos e ésteres. Moléculas de ligação clivagens por enzima típicas incluem aquelas que são sensíveis a proteases associadas com tumor tais como Catepsina B ou plasmina (Dubowchik *et al.,* (1999), *Pharm. Ther.*, 83: 67-123; Dubowchik *et al.,* (1998), *Bioorg. Med. Chem. Lett.*, 8: 3341-52; de Groot *et al.,* (2000), *J. Med. Chem.*, 43: 3093-102; de Groot *et al.,* (1999)m 42: 5277-83). Sítios cliváveis por catepsina B incluem as seqüências dipeptídicas valina-citrulina e fenilalanina-lisina (Doronina *et al.,* (2003), *Nat. Biotech.*, 21(7): 778-84); Dubowchik *et al.,* (2002), *Bioconjug. Chem.,* 13: 855-69). Outros sítios cliváveis por enzima típicos incluem aqueles formados por seqüências de oligopeptídeos de 4 a 16 aminoácidos (por exemplo, Suc-β-Ala-Leu-Ala-Leu) as quais são reconhecidas por proteases trouse tais como Thimet Oliogopeptidase (TOP), uma enzima que é preferencialmente liberada por neutrófilos, macrófagos, e outros granulócitos.

Em uma modalidade adicional, a molécula de ligação é formada reagindo uma molécula de ligação da invenção com uma molécula de ligação da fórmula:

X-Y-Z

em que:

X é uma molécula de fixação;

Y é uma molécula espaçadora; e

Z é uma porção de fixação efetora.

O termo "molécula de fixação de molécula de ligação" inclui moléculas as quais possibilitam a fixação covalente do peptídeo de ligação a uma molécula de ligação da invenção.

A molécula de fixação pode compreender, por exemplo, uma cadeia covalente de 1 a 60 átomos de carbono, oxigênio, nitrogênio, enxofre, opcionalmente substituídos com átomos de hidrogênio e outros substituintes

os quais possibilitam que a molécula de ligação realize sua função pretendida. A molécula de fixação pode compreender grupamentos funcionais peptídico, éster, alquila, alquenila, alquinila, arila, éter, tioéter, e etc. Preferencialmente, a molécula de fixação é selecionada de tal modo que é capaz de reagir com um grupamento funcional reativo sobre um polipeptídeo compreendendo no mínimo uma sítio de ligação antigênica, para formar uma molécula de ligação da invenção. Exemplos de moléculas de fixação incluem, por exemplo, moléculas de fixação de amino, carboxilato, e tiol.

Moléculas de fixação de amino incluem moléculas as quais reagem com grupamentos amino sobre um polipeptídeo, de tal modo que uma molécula de ligação da invenção é formada. Moléculas de fixação de amino são conhecidas na técnica. Exemplos de moléculas de fixação de amino incluem carbamidas ativadas (por exemplo, as quais podem reagir com um grupamento amino sobre uma molécula de ligação para formar uma molécula de ligação a qual compreende grupamento uréia), aldeídos (por exemplo, os quais podem reagir com grupamentos amino sobre uma molécula de ligação), e isocianatos ativados (os quais podem reagir com um grupamento amino sobre uma molécula de ligação para formar uma molécula de ligação a qual compreende um grupamento uréia). Exemplos de moléculas de fixação de amino incluem, mas não estão limitados a, molécula de N-succinimidila, N-sulfossuccinimidila, N-ftalimidila, N-sulfoftalimidila, 2-nitrofenila, 4-nitrofenila, 2,4-dinitrofenila, 3-sulfonila-4-nitrofenila, ou 3-carbóxi-4-nitrofenila.

Moléculas de fixação de carboxilato incluem moléculas as quais reagem com grupamentos carboxilato sobre um polipeptídeo, de tal modo que se forma uma molécula de ligação da invenção. Moléculas de fixação de carboxilato são conhecidas na técnica. Exemplos de moléculas de fixação de carboxilato incluem, mas não estão limitadas a intermediários éster ativados e intermediários carbonila ativados, os quais podem reagir com um grupamento COOH sobre uma molécula de ligação para formar uma molécula de ligação a qual compreende um grupamento éster, tioéster, ou amida.

Moléculas de fixação de tiol incluem moléculas as quais reagem com grupamentos tiol presentes sobre um polipeptídeo, de tal modo que se

forma uma molécula de ligação da invenção. Moléculas de fixação de tiol são conhecidas na técnica. Exemplos de moléculas de fixação de tiol incluem grupamentos acila ativados (os quais podem reagir com uma sulfidrila sobre uma molécula de ligação para formar uma molécula de ligação a qual compreende um tioéster), grupamentos alquila ativados (os quais podem reagir com uma sulfidrila sobre uma molécula de ligação para formar uma molécula de ligação a qual compreende uma molécula tioéster), aceitadores de Michael tais como grupamentos maleimida ou acrílicos (os quais podem reagir com uma sulfidrila sobre uma molécula de ligação para formar um produto de adição do tipo de Michael), grupamentos os quais reagem com grupamentos sulfidrila através de reações redox, grupamentos dissulfeto ativados (os quais podem reagir com um grupamento sulfidrila sobre uma molécula de ligação para formar, por exemplo, uma molécula de ligação a qual compreende uma molécula de dissulfeto). Outras moléculas de fixação de tiol incluem compostos de acrilamidas, α-iodoacetamidas, e ciclopropan-1,1-dicarbonila. Além disso, a molécula de fixação de tiol pode compreender uma molécula a qual modifica um tiol sobre a molécula de ligação para formar outra espécie reativa à qual a molécula de ligação pode ser fixada para formar uma molécula de ligação da invenção.

A molécula espaçador, Y, é uma ligação covalente ou uma cadeia covalente de átomos a qual pode conter um ou mais resíduos aminoácidos. Também pode compreender 0 a 60 átomos de carbono, oxigênio, enxofre ou nitrogênio opcionalmente substituído com hidrogênio ou outros substituintes os quais possibilitam que a molécula de ligação resultante realize sua função pretendida. Em uma modalidade, Y compreende uma molécula de alquila, alquenila, alquinila, éster, éter, carbonila, ou amida.

Em outra modalidade, um grupamento tiol sobre a molécula de ligação é convertido em um grupamento reativo, tal como um grupamento carbonila reativo, tal como uma cetona ou aldeído. A molécula de fixação é em seguida reagida com a cetona ou aldeído para formar o composto desejado da invenção. Exemplos de moléculas de fixação de carbonila reativo incluem, mas não estão limitadas a, hidrazinas, hidrazidas, hidroxilaminas O-

substituídas, cetonas α-β-insaturadas, e H_2C=CH-CO-NH-NH_2. Outros exemplos de moléculas de fixação e métodos para modificar moléculas tiol as quais podem ser usadas para formar moléculas de ligação da invenção são descritos em Pratt, M. L. *et al.* J Am Chem Soc. 2003 May 21; 125 (20): 6149-59; e Saxon, E. Science. 2000 Mar 17; 287 (5460): 2007-10.

A molécula de ligação pode ser uma molécula a qual é capaz de reagir com uma molécula efetora ou um derivado da mesma para formar uma molécula de ligação da invenção. Por exemplo, a molécula efetora pode ser encadeada às porções remanescentes da molécula através de uma ligação dissulfeto. Em similares casos, a molécula de ligação é selecionada de tal modo que seja capaz de reagir com um derivado da porção efetora apropriada de tal modo que a molécula efetora seja fixada à molécula de ligação da invenção. Conforme descrito acima, a molécula de ligação e/ou o peptídeo de ligação como um todo podem ser selecionados de modo que o peptídeo de ligação seja clivado em um ambiente apropriado.

Moléculas de peptídeos de ligação particularmente preferencial incluem, por exemplo, N-succinimidil 3-(2-piridilditio)propionato (SPDP) (vide, por exemplo, Carlsson et al., Biochem. J., 173, 723-737 (1978)), N-succinimidil 4-(2-piridilditio)butanoato (SPDB) (vide, por exemplo, a Patente dos Estados Unidos Nº 4.563.304), N-succinimidil 4-(2-piridilditio)pentanoato (SPP) (vide, por exemplo, CAS Número de registro 341498-08-6), N-succinimidil 4-(N-maleimidometil)ciclohexano-1-carboxilato (SMCC) (vide, por exemplo, Yoshitake et al., Eur. J. Biochem., 101, 395-399 (1979)), e N-succinimidil 4-metil-4-[2-(5-nitro-piridil)-- ditio]pentanoato (SMNP) (vide, por exemplo, a Patente dos Estados Unidos Nº 4.563.304) As moléculas de peptídeos de ligação mais preferenciais para uso na composição da invenção são SPP, SMCC, e SPDB. Em uma modalidade preferencial, SPDB é usado para encadear uma molécula efetora a uma molécula de ligação da invenção.

Moléculas efetoras citotóxicas preferenciais para uso na presente invenção são fármacos citotóxicos, particularmente aqueles que são usados para terapia do câncer. Conforme usado aqui, neste Pedido de patente, "uma citotoxina ou agente citotóxico" significa qualquer agente que é preju-

dicial para o crescimento e a proliferação de células e pode agir para reduzir, inibir ou destruir uma célula ou malignidade. Citotoxinas típicas incluem, mas não estão limitadas a, radionuclídeos, biotoxinas, toxinas enzimaticamente ativas, citostáticos ou citotóxicos agentes terapêuticos, pró-farmacos, ligantes imunologicamente ativos e modificadores da resposta biológica tais como citocinas. Qualquer citotoxina que age para retardar ou tornar mais lento o crescimento de células imunorreativas ou células malignas está dentro do âmbito da presente invenção.

Citotoxinas típicas incluem, em geral, agentes citostáticos, agentes alquilantes, antimetabólitos, agentes antiproliferativos, agentes de ligação de tubulina, hormônios e antagonistas hormonais, e similares. típicas Citotoxinas típicas que são compatíveis com a presente invenção incluem substâncias alquilantes, tais como mecloretamina, trietilenofosforamida, ciclofosfamida, ifosfamida, clorambucil, bussulfan, melfalan ou triaziquona, também compostos de nitrosouréia, tais como carmustina, lomustina, ou semustina.

Moléculas particularmente preferenciais para conjugação são maitansinóides. Maitansinóides foram isolados originalmente do arbusto da África Oriental pertencente ao gênero Maytenus, mas subseqüentemente também foram descobertos como sendo metabólitos de bactérias do solo, tais como Actinosynnema pretiosum (vide, por exemplo, a Patente dos Estados Unidos Nº 3.896.111). Maitansinóides são conhecidos na técnica por incluírem maitansina, maitansinol, C-3 ésteres de maitansinol, e outros análogos e derivados de maitansinol (vide, por exemplo, as Patentes dos Estados Unidos Nos. 5.208.020 e 6.441.163). C-3 Ésteres de maitansinol podem ser naturais ou derivados sinteticamente. Além disso, tanto C-3 ésteres de maitansinol que ocorrem naturalmente quanto sintéticos podem ser classificados como um C-3 éster com ácidos carboxílicos simples, ou um C-3 éster com derivados de N-metil-L-alanina, o último sendo mais citotóxico do que o precedente. Análogos de maitansinóides sintéticos também são conhecidos na técnica e são descritos, por exemplo, em Kupchan et al., J. Med. Chem., 21, 31-37 (1978). Métodos para produzir maitansinol e análogos e derivados

dos mesmos são descritos, por exemplo, na Patente dos Estados Unidos Nº 4.151.042.

Maitansinóides adequados para uso como conjugados de anticorpos podem ser isolados a partir de fontes naturais, produzidos sinteticamente, ou produzidos semi-sinteticamente usando métodos conhecidos na técnica. Além disso, os maitansinóides podem ser modificados em qualquer maneira adequada, contanto que seja preservada citotoxicidade suficiente na molécula do conjugado final.

Maitansinóides particularmente preferenciais compreendendo uma molécula de ligação que contém um grupamento químico reativo são C-3 ésteres de maitansinol e seus análogos onde a molécula de ligação contém uma ligação dissulfeto e a molécula de fixação compreende um éter N-succinimidílico ou N-sulfossuccinimidílico. Muitas posições sobre maitansinóides podem servir como a posição para ligar quimicamente a molécula de ligação, por exemplo, através de uma molécula de fixação efetora. Por exemplo, também são úteis a posição C-3 tendo um grupamento oxidrila, a posição C-14 modificada com hidroximetil, a posição C-15 modificada com hidróxi e a posição C-20 tendo um grupamento hidróxi. A molécula de ligação mais preferencialmente é encadeada à posição C-3 de maitansinol. O mais preferencialmente, o maitansinóide usado em conexão com a composição da invenção é N.sup.2'-deacetil-N.sup.2'-(- 3-mercapto-1-oxopropil)-maitansina (DM1) ou N.sup.2'-deacetil-N.sup.2'-(4-- mercapto-4-metil-1-oxopentil)-maitansina (DM4).

Moléculas de ligação com outras ligações químicas também podem ser usadas no contexto da invenção, assim como outros maitansinóides. Exemplos específicos de outras ligações químicas as quais podem ser incorporadas nas moléculas de ligação incluem aquelas descritas acima, tais como, por exemplo ligação lábil em ácido, ligações tioéter, ligações fotolábeis, ligações lábeis em peptidase e ligações lábeis em esterase. Métodos para produzir maitansinóides com moléculas de ligação e/ou moléculas efetoras de fixação são descritos, por exemplo, nas Patentes dos Estados Unidos Nos. 5.208.020, 5.416.064, e 6.333.410.

A molécula de ligação (e/ou a molécula de fixação efetora) de um maitansinóide tipicamente e preferencialmente é parte de uma molécula de peptídeo de ligação maior que é usada para unir o anticorpo ao maitansinóide. Qualquer molécula de peptídeo de ligação adequada pode ser usada em conexão com a invenção, na medida que a molécula de ligação proporcione retenção das características de citotoxicidade e direcionamento do maitansinóide e do anticorpo, respectivamente. A molécula de ligação une o maitansinóide ao anticorpo através de ligações químicas (conforme descrito acima), de tal modo que o maitansinóide e o anticorpo são acoplados quimicamente (por exemplo, ligado de modo covalente) um ao outro. Desejavelmente, a molécula de ligação liga quimicamente o maitansinóide ao anticorpo através de ligações dissulfeto ou ligações tioéter. O mais preferencialmente, o anticorpo é acoplado quimicamente ao maitansinóide através de ligações dissulfeto.

Outras classes preferenciais de agentes citotóxicos incluem, por exemplo, a família de fármacos de antraciclina, os fármacos vinca, as mitomicinas, as bleomicinas, os nucleosídeos citotóxicos, a família de fármacos de pteridina, diinenos, e as podofilotoxinas. Membros particularmente úteis destas classes incluem, por exemplo, adriamicina, carminomicina, daunorubicina (daunomicina), doxorubicina, aminopterina, metotrexato, metopterina, mitramicina, estreptonigrina, diclorometotrexato, mitomicina C, actinomicina-D, porfiromicina, 5-fluorouracil, floxuridina, ftorafur, 6-mercaptopurina, citarabina, arabinosídeo de citosina, podofilotoxina, ou podofilotoxina derivados tais como etoposídeo ou fosfato de etoposídeo, melfalan, vinblastina, vincristina, leurosidina, vindesina, leurosina e similares. Ainda outras citotoxias que são compatíveis com os ensinamentos aqui, neste Pedido de patente, incluem taxol, taxano, citocalasina B, gramicidina D, brometo de etídio, emetina, tenoposídeo, colchicina, diidróxi antracina diona, mitoxantrona, procaína, tetracaína, lidocaína, propranolol, e puromicina e análogos ou homólogos dos mesmos. Hormônios e antagonistas hormonais, tais como corticosteróides, por exemplo, prednisona, progestinas, por exemplo, hidroxiprogesterona ou medroprogesterona, estrogênios, por exemplo, dietilestilbestrol, anties-

trogênios, por exemplo, tamoxifeno, androgênios, por exemplo, testosterona, e inibidores de aromatase, por exemplo, aminoglutetimida também são compatíveis com os ensinamentos aqui, neste Pedido de patente. Conforme observado previamente, uma pessoa versada na técnica pode fazer modificações químicas para o composto desejado de modo a produzir reações deste composto mais convenientes para fins de preparar conjugados da invenção.

Um exemplo de citotoxinas particularmente preferenciais compreende membros ou derivados da família enediina de antibióticos antitumorais, incluindo caliqueamicina, esperamicinas ou dinemicinas. Estas toxinas são extremamente potentes e agem clivando DNA nuclear, levando a morte celular. Diferentemente de toxinas de proteína as quais podem ser clivadas *in vivo* para dar muitos fragmentos de polipeptídeos inativos porém imunogênicos, toxinas tais como caliqueamicina, esperamicinas e outras enediinas são moléculas pequenas as quais são essencialmente não-imunogênicas. Estas toxinas não-peptídicas são encadeadas quimicamente aos dímeros ou tetrâmeros por técnicas as quais foram previamente usadas para marcar anticorpos monoclonais e outras moléculas. Estas tecnologias de ligação incluem ligação sítio-específica através dos resíduos de açúcar N-encadeados presentes somente sobr a porção Fc dos construtos. Os métodos de ligação sítio-dirigida têm a vantagem de reduzir os possíveis efeitos de ligação sobre as propriedades de ligação dos construtos.

Entre outras citotoxinas, Será reconhecido que polipeptídeos também podem ser associados com uma biotoxina tal como ricina subunidade A, abrina, toxina diftérica, botulinum, cianginosinas, saxitoxina, shigatoxina, tétano, tetrodotoxina, tricoteceno, verrucologen ou uma enzima tóxica. Preferencialmente, os referidos construtos serão preparados usando técnicas de engenharia genética que possiblitam expressão direta do construto de molécula de ligação-toxina. Outros modificadores da resposta biológica que podem ser associados com os polipeptídeos da invenção da presente invenção compreendem citocinas tais como linfocinas e interferons. Em vista da presente descoberta sugere-se que uma pessoa versada na técnica pode prontamente formar os construtos referidos usando técnicas convencionais.

Outra classe de citotoxinas compatíveis que podem ser usadas em associação com os polipeptídeos revelados são fármacos radiossensibilizantes que podem ser dirigidos de modo eficaz a células imunorreativas ou tumorais. Os fármacos referidos reforçam a sensibilidade a radiação ionizante, deste modo aumentando a eficácia da radioterapia. Um conjugado internalizado pela célula tumoral liberaria o radiossensibilizador mais próximo do núcleo onde a radiossensibilização seria máxima. Os polipeptídeos encadeados a radiossensibilizador não ligado da invenção seria mais rapidamente clareado do sngue, localizando o agente de radiossensibilização remanescente no tumor alvo e proporcionando mínima captação em tecidos normais. Depois de rápido *clearance* do sangue, radioterapia auxiliar seria administrada em um de três modos: 1.) radiação de feixe externo dirigida especificamente para o tumor, 2.) radioatividade implantada diretamente no tumor ou 3.) radioimunoterapia sistêmica com a mesma molécula de ligação. Uma variação potencialmente atrativa desta abordagem seria a fixação de um radioisótopo terapêutico ao imunoconjugado radiossensibilizado, deste modo proporcionando a conveniência de administrar ao paciente um único fármaco.

Em uma modalidade, uma molécula que reforça a estabilidade ou a eficácia do polipeptídeo pode ser conjugada. Por exemplo, em uma modalidade, PEG pode ser conjugado aos polipeptídeos da invenção para aumentar sua meia-vida in vivo. Leong, S.R., et al. 2001. *Cytokine* 16:106; 2002; *Adv. in Drug Deliv. Rev.* 54:531; ou Weir et al. 2002. Biochem. Soc. Transactions 30:512.

Conforme previamente aludido, citotoxinas compatíveis podem compreender um pró-farmaco. Conforme usado aqui, neste Pedido de patente, o termo "pró-farmaco" se refere a um precursor ou forma derivada de uma substância farmaceuticamente ativa que é menos citotóxica para células tumorais comparado com o fármaco de origem e é capaz de seer ativado enzimaticamente ou convertido na forma de origem mais ativa. Pró-farmacos compatíveis com a invenção incluem, mas não estão limitadas a, pró-farmacos contendo fosfato, pró-farmacos contendo tiofosfato, pró-farmacos

contendo sulfato, pró-farmacos contendo peptídeo, pró-farmacos contendo β-lactama, pró-farmacos contendo fenoxiacetamida opcionalmente substituída ou pró-farmacos contendo fenilacetamida opcionalmente substituída, 5-fluorocitosina e outras pró-farmacos de 5-fluorouridina que podem ser convertidas para o fármaco livre de citotóxicos mais ativos. Em uma modalidade, um agente citotóxico, tal como um maitansinóide, é administrado como um pró-farmaco o qual é liberado pela hidrólise de ligações dissulfeto. Exemplos adicionais de fármacos citotóxicos que podem ser derivados em uma forma de pró-farmaco para uso na presente invenção compreendem os agentes quimioterápicos descritos acima.

XIII. Administração de Moléculas de Ligação

Métodos para preparar e administrar moléculas de ligação da invenção a um sujeito são de conhecimento geral ou são prontamente determinados por aqueles versados na técnica. A via de administração das moléculas de ligação da invenção pode ser oral, parenteral, por inalação ou tópica. O termo parenteral conforme usado aqui, neste Pedido de patente, inclui administração intravenosa, intra-arterial, intraperitoneal, intramuscular, subcutânea, rectal ou vaginal. As formas intravenosa, intra-arterial, subcutânea e intramuscular de administração parenteral são geralmente preferenciais. Apesar de todas estas formas de administração serem claramente contemplatadas como estando dentro do âmbito da invenção, uma forma para administração seria uma solução para injeção, em particular para injeção ou gotejamento intravenoso ou intra-arterial. Geralmente, uma composição farmacêutica adequada para injeção pode compreender um tampão (por exemplo, tampão de acetato, fosfato ou citrato), um tensoativo (por exemplo, polissorbato), opcionalmente um agente estabilizante (por exemplo, albumina humana), e etc. No entanto, em outros métodos compatíveis com os ensinamentos aqui, neste Pedido de patente, os polipeptídeos podem ser liberados diretamente no sítio da população celular adversa deste modo aumentando a exposição do tecido doente ao agente terapêutico.

Preparações para administração parenteral incluem soluções, suspensões, e emulsões estéreis aquosas ou não-aquosas. Exemplos de

solventes não-aquosos são propileno glicol, polietileno glicol, óleos vegetais tais como azeite de oliva, e ésteres orgânicos injetáveis tais como oleato de etila. Veículos aquosos incluem água, soluções alcoólicas / aquosas, emulsões ou suspensões, incluindo salina e meios tamponados. Na invenção em questão, veículos farmaceuticamente aceitáveis incluem, mas não estão limitados a, 0,01 a 0,1 M e preferencialmente 0,05 M de tampão de fosfato ou 0,8% de salina. Outros veículos parenterais comuns incluem soluções de fosfato de sódio, dextrose de Ringer, dextrose e cloreto de sódio, solução de Ringer lactada, ou óleos não voláteis. Veículos intravenosos incluem repositores de fluidos e nutrientes, repositores de eletrólitos, tais como aqueles baseados em dextrose de Ringer, e similares. Preservantes e outros aditivos também podem estar presentes tais como, por exemplo, antimicrobianos, antioxidantes, agentes quelantes, e gases inertes e similares.

Mais particularmente, composições farmacêuticas adequadas para uso injetável incluem soluções aquosas estéreis (onde solúvel em água) ou dispersões e pós estéreis para a preparação extemporânea de soluções ou dispersões injetáveis estéreis. Em similares casos, a composição deve ser estéril e deve ser fluida na medida que exista fácil seringabilidade. Deve ser estável sob as condições de fabricação e armazenamento e preferencialmente será preservada contra a ação contaminante de microorganismos, tais como bactérias e fungos. O veículo pode ser um solvente ou meio de dispersão contendo, por exemplo, água, etanol, poliol (por exemplo, glicerol, propileno glicol, e polietileno glicol líquido, e similares), e misturas adequadas dos mesmos. A fluidez apropriada pode ser mantida, por exemplo, pelo uso de um revestimento tal como lecitina, pela manutenção do tamanho de partícula requerido no caso de dispersão e pelo uso de tensoativos.

Prevenção da ação de microorganismos pode ser otida por vários agentes antibacterianos e antifúngicos, por exemplo, parabenos, clorobutanol, fenol, ácido ascórbico, timerosal e similares. Em muitos casos, será preferencial incluir agentes isotônicos, por exemplo, açúcares, poliálcoois, tais como manitol, sorbitol, ou cloreto de sódio na composição. Absorção prolongada das composições injetáveis pode ser realizada incluindo na

composição um agente o qual retarda a absorção, por exemplo, monoestearato de alumínio e gelatina.

Em qualquer caso, soluções injetáveis estéreis podem ser preparadas incorporando um composto ativo (por exemplo, um polipeptídeo por si só ou em combinação com outros agentes ativos) na quantidade requerida em um solvente apropriado com um ou uma combinação de ingredientes enumerados aqui, neste Pedido de patente, conforme requerido, seguida por esterilização filtrada. Geralmente, são preparadas dispersões incorporando o composto ativo em um veículo estéril, o qual contém um meio de dispersão básico e os outros ingredientes requeridos entre aqueles enumerados acima. No caso de pós estéreis para a preparação de soluções injetáveis estéreis, os métodos de preparação preferenciais são secagem a vácuo e secagem por congelamento, a qual produz um pó de um ingrediente ativo mais qualquer ingrediente desejado adicional de uma solução previamente filtrada estéril da mesma. As preparações para injeções são processadas, preenchidas em recipientes tais como ampolas, bolsas, garrafas, seringas ou frascos, e seladas sob condições assépticas de acordo com métodos conhecidos na técnica. Adicionalmente, as preparações podem ser embaladas e vendidas sob a forma de um kit tal como aqueles descritos nas publicações de patente co-pendentes U.S.S.N. 09/259.337 e U.S.S.N. 09/259.338 cada uma das quais é incorporada aqui, a este Pedido de patente, por meio de referência. Os artigos de fabricação referidos preferencialmente terão rótulos ou bulas indicando que as composições associadas são úteis para tratar um sujeito sofrendo de, ou predisposto a distúrbios autoimunes ou neoplásicos.

Doses eficazes das moléculas de ligação estabilizadas da presente invenção, para o tratamento das condições descritas acima variam dependendo de muitos fatores diferentes, incluindo meios de administração, sítio alvo, estado fisiológico do paciente, se o paciente é um ser humano ou um animal, outras medicações administradas, e se o tratamento é profilático ou terapêutico. Geralmente, o paciente é um ser humano, porém mamíferos não-humanos incluindo mamíferos transgênicos também podem ser tratados. As dosagens de tratamento podem ser tituladas usando métodos de

rotina de conhecimento daqueles versados na técnica para otimizar a segurança e a eficácia.

Para imunização passiva com uma molécula de ligação da invenção, a dosagem pode variar, por exemplo, a partir de cerca de 0,0001 a 100 mg/kg, e mais geralmente 0,01 a 5 mg/kg (por exemplo, 0,02 mg/kg, 0,25 mg/kg, 0,5 mg/kg, 0,75 mg/kg, 1 mg/kg, 2 mg/kg, e etc.), do peso corporal do hospedeiro. Por exemplo, as dosagens podem ser 1 mg/kg de peso corporal ou 10 mg/kg de peso corporal ou dentro da faixa de 1 a 10 mg/kg, preferencialmente no mínimo 1 mg/kg. Doses intermediárias nas faixas acima também são pretendidas para estarem dentro do âmbito da invenção.

Aos sujeitos podem ser administradas as doses referidas diariamente, em dias alternados, semanalmente ou de acordo com qualquer outro esquema determinado por análise empírica. Um tratamento típico acarreta administração em múltiplas dosagens durante um período prolongado, por exemplo, de no mínimo seis meses. Regimes de tratamento típicos adicionais acarretam administração uma vez por cada duas semanas ou uma vez ao mês ou uma vez a cada 3 a 6 meses. Esquemas de dosagem típicos incluem 1 a 10 mg/kg ou 15 mg/kg em dias consecutivos, 30 mg/kg em dias alternados ou 60 mg/kg semanalmente. Em alguns métodos, dois ou mais anticorpos monoclonais com diferentes especificidades de ligação são administrados simultaneamente, em cujo caso a dosagem de cada anticorpo administrado pode estar dentro das faixas indicadas.

As moléculas de ligação da invenção podem ser administrada em múltiplas ocasiões. Os intervalos entre dosagens únicas podem ser, por exemplo, diariamente, semanalmente, mensalmente ou anualmente. Os intervalos também podem ser irregulares conforme indicado medindo os níveis sangüíneos de polipeptídeo ou molécula alvo no paciente. Em alguns métodos, a dosagem é ajustada para obter uma determinada concentração plasmática de molécula de ligação ou toxina, por exemplo, 1 a 1000 µg/ml ou 25 a 300 µg/ml. Alternativamente, moléculas de ligação podem ser administradas como uma formulação de liberação gradual, em cujo caso é nexessária administração menos freqüente. A dosagem e a freqüência variam depen-

dendo da meia-vida do anticorpo no paciente. Em geral, anticorpos humanizados mostram a meia-vida mais longa, seguidos por anticorpos quiméricos e anticorpos não-humanos. Em uma modalidade, as moléculas de ligação da invenção podem ser administradas em forma não-conjugada, Em outra modalidade, os polipeptídeos da invenção podem ser administrados múltiplas vezes em forma conjugada. Em ainda outra modalidade, as moléculas de ligação da invenção podem ser administrada em forma não-conjugada, em seguida em forma conjugada, ou vice versa.

A dosagem e a freqüência de administração podem variar dependendo de se o tratamento é profilático ou terapêutico. Em aplicações profiláticas, composições contendo os presentes anticorpos ou um coquetel dos mesmos são administradas a um paciente ainda não no estado de doença para reforçar a resistência do paciente. Uma quantidade similar é definida como sendo uma "dose eficaz profilática." Neste uso, as quantidades precisas novamente dependem do estado de saúde e da imunidade geral do paciente, mas geralmente variam de 0,1 a 25 mg por dose, especialmente de 0,5 a 2,5 mg por dose. Uma dosagem relativamente baixa é administrada em intervalos relativamente infreqüentes durante um período de tempo longo. Alguns pacientes continuam a receber tratamento pelo resto de suas vidas.

Em aplicações terapêuticas, uma dosagem relativamente elevada (por exemplo, a partir de cerca de 1 a 400 mg/kg de molécula de ligação, por exemplo, anticorpo por dose, com dosagens de a partir de 5 a 25 mg sendo mais comumente usadas para radioimunoconjugados e doses maiores para moléculas conjugadas a fármaco de citotoxina) em intervalos relativamente curtos é algumas vezes necessária até a progressão da doença ser reduzida ou terminada, e preferencialmente até o paciente mostrar melhora parcial ou completa dos sintomas da doença. Depois disso, pode ser administrado ao paciente um regime profilático.

Em uma modalidade, um sujeito pode ser tratado com uma molécula de ácido nucléico codificando um polipeptídeo da invenção (por exemplo, em um vetor). Doses para ácidos nucléicos codificando polipeptídeos variam a partir de cerca de 10 ng a 1 g, 100 ng a 100 mg, 1 µg a

10 mg, ou 30 a 300 µg de DNA por paciente. Doses para vetores virais infecciosos variam de 10 a 100, ou mais, vírions por dose.

Agentes terapêuticos podem ser administrados por meios parenterais, tópicos, intravenosos, orais, subcutâneos, intra-arteriais, intracraniais, intraperitoneais, intranasais ou intramusculares para tratamento profilático e/ou terapêutico. Injeção intramuscular ou infusão intravenosa são preferenciais para administração de uma molécula de ligação da invenção. Em alguns métodos, moléculas de ligação terapêuticas particulares são injetadas diretamente dentro do crânio. Em alguns métodos, moléculas de ligação são administrada como uma composição ou dispositivo de liberação gradual, tal como um dispositivo Medipad™.

Agentes da invenção podem ser administrados opcionalmente em combinação com outros agentes que são eficazes no tratamento dos distúrbio ou condições que necessitem de tratamento (por exemplo, profilático ou terapêutico). Agentes adicionais preferenciais são aqueles os quais são reconhecidos na técnica e são administrados padronizadamente para um distúrbio particular.

As dosagens de tratamento único eficaz (isto é, quantidades terapeuticamente eficazes) de polipeptídeos marcados com ^{90}Y da invenção variam a partir de entre cerca de 5 e cerca de 75 mCi, mais preferencialmente entre cerca de 10 e cerca de 40 mCi. Dosagens não-ablativas da medula de tratamento único eficaz de anticorpos marcados com ^{131}I variam a partir de entre cerca de 5 e cerca de 70 mCi, mais preferencialmente entre cerca de 5 e cerca de 40 mCi. Dosagens ablativas de tratamento único eficaz (isto é, podem requerer transplantação de medula óssea autóloga) de anticorpos marcados com ^{131}I variam a partir de entre cerca de 30 e cerca de 600 mCi, mais preferencialmente entre cerca de 50 e menos de cerca de 500 mCi. Em associação com um anticorpo quimérico modificado, devido à meia-vida circulante mais longa vis-à-vis anticorpos murinos, um tratamento único eficaz de dosagens não-ablativas da marrow de anticorpos quiméricos marcados com Iodo ^{131} variam a partir de entre cerca de 5 e cerca de 40 mCi, mais preferencialmente menos de cerca de 30 mCi. Critérios de estudo por imagens

para, por exemplo, a etiqueta de ^{111}In, são tipicamente menos de cerca de 5 mCi.

Apesar de uma grande experiência clínica ter sido obtida com ^{131}I e ^{90}Y, outros radiomarcadores são conhecidas na técnica e têm sido usados para fins similares. Ainda outros radioisótopos são usados para estudo por imagens. Por exemplo, radioisótopos adicionais os quais são compatíveis com o âmbito da presente invenção incluem, mas não estão limitados a, ^{123}I, ^{125}I, ^{32}P, ^{57}Co, ^{64}Cu, ^{67}Cu, ^{77}Br, ^{81}Rb, ^{81}Kr, ^{87}Sr, ^{113}In, ^{127}Cs, ^{129}Cs, ^{132}I, ^{197}Hg, ^{203}Pb, ^{206}Bi, ^{177}Lu, ^{186}Re, ^{212}Pb, ^{212}Bi, ^{47}Sc, ^{105}Rh, ^{109}Pd, ^{153}Sm, ^{188}Re, ^{199}Au, ^{225}Ac, ^{211}At, e ^{213}Bi. A este respeito α, γ e β emissores são todos compatíveis com a presente invenção. Além disso, em vista da presente descoberta sugere-se que uma pessoa versada na técnica pode determinar prontamente quais radionuclídeos são compatíveis com um curso de tratamento selecionado sem indevida experimentação. Até o momento, radionuclídeos adicionais os quais já foram usados em diagnóstico clínico incluem ^{125}I, ^{123}I, ^{99}Tc, ^{43}K, ^{52}Fe, ^{67}Ga, ^{68}Ga, bem como ^{111}In. Anticorpos também têm sido marcados com uma variedade de radionuclídeos para uso potencial em imunoterapia dirigida (Peirersz et al. *Immunol. Cell Biol.* 65: 111-125 (1987)). Estes radionuclídeos incluem ^{188}Re e ^{186}Re bem como ^{199}Au e ^{67}Cu em uma menos extensão. A Patente dos Estados Unidos Nº 5.460.785 proporciona dados adicionais referentes a similares radioisótopos e é incorporada aqui, a este Pedido de patente, por meio de referência.

Quer ou não as moléculas de ligação da invenção são usadas em uma forma conjugada ou não-conjugada. Será reconhecido que a principal vantagem da presente invenção é a capacidade para usar estes polipeptídeos em pacientes mielossuprimidos, especialmente aqueles que estão sendo submetidos, ou sofreram, terapias auxiliares tais como radioterapia ou quimioterapia. Em outras modalidades preferenciais, os polipeptídeos (novamente em uma forma conjugada ou não-conjugada) podem ser usados em um regime terapêutico combinado com agentes quimioterápicos. Aqueles versados na técnica reconhecerão que similares regimes terapêuticos podem compreender a administração seqüencial, simultânea, concorrente ou

coextensiva dos anticorpos revelados e um ou mais agentes quimioterápicos. Modalidades particularmente preferenciais deste aspecto da invenção compreenderão a administração de uma molécula de ligação conjugada a toxina, por exemplo, conjugada a um maitansinóide tal como um D4 maitansinóide.

Apesar das moléculas de ligação poderem ser administradas conforme descrito imediatamente acima, deve ser enfatizado que em outras modalidades polipeptídeos conjugados e não conjugados podem ser administrados a pacientes saudáveis de modo diverso como um agente terapêutico de primeira linha. Em similares modalidades os polipeptídeos podem ser administrados a pacientes tendo reservas de medula vermelha normal ou média e/ou a pacientes que não têm, e não estão, sofrendo terapias auxiliares tais como radiação de feixe externo ou quimioterapia.

No entanto, conforme discutido acima, modalidades selecionadas da invenção compreendem a administração de moléculas de ligação a pacientes mielossuprimidos ou em combinação ou associação com uma ou mais terapias auxiliares tais como radioterapia ou quimioterapia (isto é um regime terapêutico combinado). Conforme usado aqui, neste Pedido de patente, a administração de polipeptídeos em associação ou combinação com uma terapia auxiliar significa a administração ou aplicação seqüencial, simultânea, coextensiva, concorrente, concomitante ou contemporânea da terapia e das moléculas de ligação reveladas. Aqueles versados na técnica reconhecerão que a administração ou aplicação dos vários componentes do regime terapêutico combinado pode ser regulada para reforçar a eficácia global do tratamento. Por exemplo, agentes quimioterápicos podem ser administrados em cursos de tratamento padrão de conhecimento geral seguidos dentro de umas poucas semanas por radioimunoconjugados da presente invenção. Inversamente, polipeptídeos associados a citotoxina podem ser administrados por via intravenosa seguidos por radiação de feixe externo localizada no tumor. Em ainda outras modalidades, o polipeptídeo pode ser administrado concorrentemente com um ou mais agentes quimioterápicos selecionados em uma única visita ao consultório. Um técnico versado (por

exemplo, um oncologista experiente) seria prontamente capaz de discernir regimes terapêuticos combinados eficazes sem indevida experimentação com base na terapia auxiliar selecionada e nos ensinamentos da presente relatório descritivo.

A este respeito será reconhecido que a combinação das moléculas de ligação (quer conjugadas ou não-conjugadas) e o agente quimioterápico pode ser administrada em qualquer ordem e dentro de qualquer frame de tempo que proporcione um benefício terapêutico para o paciente. Isto é, o agente quimioterápico e o polipeptídeo podem ser administrados em qualquer ordem ou concorrentemente. Moléculas de ligação e agentes quimioterápicos podem ser administrados separadamente ou podem ser administrados sob a forma de uma composição. Em modalidades selecionadas os polipeptídeos da presente invenção serão administrados a pacientes que foram submetidos previamente a quimioterapia. Em ainda outras modalidades, os polipeptídeos e o tratamento quimioterápico serão administrados substancialmente simultaneamente ou concorrentemente. Por exemplo, ao paciente pode ser administrada a molécula de ligação enquanto estiver sendo submetido a um tratamento de quimioterapia. Em modalidades preferenciais a molécula de ligação será administrada dentro de 1 ano de qualquer agente ou tratamento quimioterápico. Em outras modalidades preferenciais o polipeptídeo será administrado dentro de 10, 8, 6, 4, ou 2 meses de qualquer agente ou tratamento quimioterápico. Em ainda outras modalidades preferenciais o polipeptídeo será administrado dentro de 4, 3, 2 ou 1 semana de qualquer agente ou tratamento quimioterápico. Em ainda outras modalidades o polipeptídeo será administrado dentro de 5, 4, 3, 2 ou 1 dias do qualquer agente ou tratamento quimioterápico selecionado. Será adicionalmente reconhecido que os dois agentes ou tratamentos podem ser administrados ao paciente dentro de uma questão de horas ou minutos (isto é substancialmente simultaneamente).

Além disso, de acordo com a presente invenção um paciente mielossuprimido deverá significar qualquer paciente apresentando contagens sangüíneas reduzidas. Aqueles versados na técnica reconhecerão que

há vários parâmetros de contagem sangüínea usados convencionalmente como indicadores clínicos de mielossupressão e pode-ser facilmente medir a extensão na qual está ocorrendo mielossupressão em um paciente. Exemplos de medições de mielossupressão aceitas na técnica são a Contagem de Neutrófilos Absoluta (ANC) ou contagem de plaquetas. A mielossupressão referida ou mieloablação parcial podem ser a conseqüência de vários distúrbios bioquímicos ou doenças ou, mais provavelmente, como a conseqüência de quimioterapia ou radioterapia anterior. A este respeito, os versados na técnica reconhecerão que pacientes que sofreram quimioterapia tradicional tipicamente apresentam reduzidas reservas de medula vermelha. Conforme discutido acima, os sujeitos referidos freqüentemente não podem ser tratados usando níveis ótimos de citotoxina (isto é radionuclídeos) devido a efeitos colaterais inaceitáveis tais como anemia ou imunossupressão que resultam em aumentada mortalidade ou morbidez.

Mais especificamente polipeptídeos conjugados ou não conjugados da presente invenção podem ser usados para tratar de modo eficaz pacientes tendo ANCs menores do que cerca de 2000/mm^3 ou contagem de plaquetas menor do que cerca de 150,000/ mm^3. Mais preferencialmente os polipeptídeos da presente invenção podem ser usados para tratar pacientes tendo ANCs de menos de cerca de 1500/ mm^3, menos de cerca de 1000/mm^3 ou ainda mais preferencialmente menos de cerca de 500/ mm^3. Similarmente, os polipeptídeos da presente invenção podem ser usados para tratar pacientes tendo uma contagem de plaquetas de menos de cerca de 75.000/mm^3, menos de cerca de 50.000/mm^3 ou ainda menos de cerca de 10.000/mm^3. Em um sentido mais geral, os versados na técnica serão facilmente capazes de determinar quando um paciente é mielossuprimido usando diretrizes e procedimentos implementados pelo governo.

Conforme indicado acima, muitos pacientes mielossuprimidos sofreram cursos de tratamento incluindo quimioterapia, radioterapia por implante ou radioterapia de feixe externo. No caso do último, uma fonte de radiação externa é para irradiação local de uma malignidade. Para métodos de implantação de radioterapia, reagentes radioativos estão localizados cirurgi-

camente dentro da malignidade, deste modo irradiando seletivamente o sítio da doença. Em qualquer evento, os polipeptídeos revelados podem ser usados para tratar distúrbios em pacientes apresentando mielossupressão independente da causa.

A este respeito será adicionalmente reconhecido que os polipeptídeos da presente invenção podem ser usados em associação ou em combinação com qualquer agente ou agentes quimioterápicos (por exemplo, proporcionando um regime terapêutico combinado) que elimina, reduz, inibe ou controla o crescimento de células neoplásicas *in vivo*. Conforme discutido, os agentes referidos freqüentemente resultam na redução de reservas de medula vermelha. Esta redução pode ser compensada, no todo ou em parte, pela reduzida mielotoxicidade dos compostos da presente invenção que vantajosamente possibilitam o tratamento agressivo de neoplasias em similares pacientes. Em outras modalidades preferenciais os imunoconjugados radiomarcados revelados aqui, neste Pedido de patente, podem ser usados de modo eficaz com radiossensibilizadores que aumentam a suscetibilidade das células neoplásicas a radionuclídeos. Por exemplo, compostos radiossensibilizantes podem ser administrados depois da molécula de ligação radiomarcada terem sido largamente clareados da corrente sangüínea mas ainda permanecem em níveis terapeuticamente eficazes no sítio do tumor ou tumores.

Com respeito a estes aspectos da invenção, agentes quimioterápicos típicos que são compatíveis com a presente invenção incluem agentes alquilantes, vinca alcalóides (por exemplo, vincristina e vinblastina), procarbazina, metotrexato e prednisona. A combinação de quatro fármacos MOPP (mecletamina (mostarda de nitrogênio), vincristina (Oncovin), procarbazina e prednisona) é muito eficaz para tratar vários tipos de linfoma e compreende uma modalidade preferencial da presente invenção. Em pacientes resistentes a MOPP, podem ser usadas as combinações ABVD (por exemplo, adriamicina, bleomicina, vinblastina e dacarbazina), ChlVPP (clorambucil, vinblastina, procarbazina e prednisona), CABS (lomustina, doxorubicina, bleomicina e estreptozotocina), MOPP mais ABVD, MOPP mais ABV

(doxorubicina, bleomicina e vinblastina) ou BCVPP (carmustina, ciclofosfamida, vinblastina, procarbazina e prednisona). Arnold S. Freedman e Lee M. Nadler, *Malignant Lymphomas, em* HARRISON'S PRINCIPLES OF INTERNAL MEDICINE 1774-1788 (Kurt J. Isselbacher *et al.*, eds., 13th ed. 1994) e V. T. DeVita *et al.,* (1997) e as referências citadas no mesmo para dosagem de rotina e planejamento. Estas terapias podem ser usadas não modificadas, ou alteradas conforme necessário para um paciente em particular, em combinação com um ou mais polipeptídeos da invenção conforme descrito aqui, neste Pedido de patente.

Regimes adicionais que são úteis no contexto da presente invenção incluem uso de agentes alquilantes únicos tais como ciclofosfamida ou clorambucil, ou combinações tais como CVP (ciclofosfamida, vincristina e prednisona), CHOP (CVP e doxorubicina), C-MOPP (ciclofosfamida, vincristina, prednisona e procarbazina), CAP-BOP (CHOP mais procarbazina e bleomicina), m-BACOD (CHOP mais metotrexato, bleomicina e leucovorin), ProMACE-MOPP (prednisona, metotrexato, doxorubicina, ciclofosfamida, etoposídeo e leucovorin mais MOPP padrão), ProMACE-CytaBOM (prednisona, doxorubicina, ciclofosfamida, etoposídeo, citarabina, bleomicina, vincristina, metotrexato e leucovorin) e MACOP-B (metotrexato, doxorubicina, ciclofosfamida, vincristina, dose fixa de prednisona, bleomicina e leucovorin). Aqueles versados na técnica serão prontamente capazes de determinar dosagens de rotina e planejamento para cada um destes regimes. CHOP também foi combinado com bleomicina, metotrexato, procarbazina, mostarda de nitrogênio, arabinosídeo de citosina e etoposídeo. Outros agentes quimioterápicos compatíveis incluem, mas não estão limitadas a, 2-clorodeoxiadenosina (2-CDA), 2'-deoxicoformicina e fludarabina.

Para pacientes com grau intermediário e elevado de NHL, que não tenham êxito em obter remissão ou recidiva, é usada terapia de recuperação. Terapias de recuperação empregam fármacos tais como arabinosídeo de citosina, cisplatina, etoposídeo e ifosfamida administrados sozinhos ou em combinação. Em formas recidivantes ou agressivas de alguns distúrbios neoplásicos são freqüentemente usados os seguintes protocolos: IMVP-16

(ifosfamida, metotrexato e etoposídeo), MIME (enxerto de metila (*metil-gag*), ifosfamida, metotrexato e etoposídeo), DHAP (dexametasona, alta dose de citarabina e cisplatina), ESHAP (etoposídeo, metilpredisolona, HD citarabina, cisplatina), CEPP(B) (ciclofosfamida, etoposídeo, procarbazina, prednisona e bleomicina) e CAMP (lomustina, mitoxantrona, citarabina e prednisona) cada um com taxas de dosagem e esquemas de conhecimento geral.

A quantidade de agente quimioterápico a ser usado em combinação com os polipeptídeos da presente invenção pode variar por sujeito ou pode ser administrada de acordo com a que é conhecida na técnica. Vide, por exemplo, Bruce A Chabner *et al.*, *Antineoplastic Agents, in* GOODMAN & GILMAN'S THE PHARMACOLOGICAL BASIS OF TERAPÊUTICOS 1233-1287 ((Joel G. Hardman *et al.*, eds., 9th ed. 1996).

Em uma modalidade, uma molécula de ligação da invenção pode ser administrada a um sujeito que foi submetido, está sofrendo, ou sofrerá um procedimento cirúrgico, por exemplo, para remover um tumor primário, uma metástase ou tecido ou crescimento pré-canceroso como uma terapia preventiva.

Em outra modalidade, uma molécula de ligação da invenção é administrada em associação com um biológico. Biológicos úteis no tratamento de cânceres são conhecidos na técnica e uma molécula de ligação da invenção pode ser administrada, por exemplo, em associação com os referidos biológicos conhecidos.

Por exemplo, a agência americada FDA aprovou os seguintes bbiológicos para o tratamento de câncer de mama: Herceptin® (trastuzumab, Genentech Inc., South San Francisco, CA; um anticorpo monoclonal humanizado que tem atividade antitumoral em câncer de mama HER2-positivo); Faslodex® (fulvestrant, AstraZeneca Pharmaceuticals, LP, Wilmington, DE; um antagonista receptor de estrogênio usado para tratar câncer de mama); Arimidex® (anastrozole, AstraZeneca Pharmaceuticals, LP; um inibidor de aromatase não-esteroidal o qual bloqueia aromatase, uma enzima necessária para produzir estrogênio); Aromasin® (exemestane, Pfizer Inc., New York, NY; um inativador de aromatase esteróide irreversível usado no tratamento

de câncer de mama); Femara® (letrozole, Novartis Pharmaceuticals, East Hanover, NJ; um inibidor de aromatase não-esteróide aprovado pela agência americana FDA para tratar câncer de mama); e Nolvadex® (tamoxifeno , AstraZeneca Pharmaceuticals, LP; um antiestrogênio não-esteróide aprovado pela agência americana FDA para tratar câncer de mama). Outros biológicos com os quais as moléculas de ligação da invenção podem ser combinadas incluemj: Avastin™ (bevacizumab, Genentech Inc.; a primeira terapia aprovada pela FDA projetada para inibir a angiogênese); e Zevalin® (ibritumomab tiuxetan, Biogen Idec, Cambridge, MA; um anticorpo monoclonal radiomarcado aprovado atualmente para o tratamento de linfomas de células B).

Além disso, a FDA aprovou os seguintes biológicos para o tratamento de câncer colorretal: Avastin™ ;Erbitux™ (cetuximab, ImClone Systems Inc., New York, NY, e Bristol-Myers Squibb, New York, NY; é um anticorpo monoclonal dirigido contra o receptor de fator de crescimento epidérmico (EGFR)); Gleevec® (mesilato de imatinib; uma inibidor da quinase protéica); e Ergamisol® (cloridrato de levamisol, Janssen Pharmaceutica Products, LP, Titusville, NJ; um imunomodulador aprovado pela FDA em 1990 como um tratamento adjuvante em combinação com 5-fluorouracil depois de resseção cirúrgica em pacientes com câncer de cólon Estágio C de Dukes).

Para uso no tratamento de Linfomas Não-Hodgkin terapias atualmente aprovadas incluem: Bexxar® (tositumomab e iodo I-131 tositumomab, GlaxoSmithKline, Research Triangle Park, NC; um tratamento multietapas envolvendo um anticorpo monoclonal de camundongo (tositumomab) encadeado a uma molécula radioativa (iodo I-131)); Intron® A (interferon α-2b, Schering Corporation, Kenilworth, NJ; um tipo de interferon aprovado para o tratamento de lingoma não Hodgkin folicular em associação com quimioterapia de combinação contendo antraciclina (por exemplo, ciclofosfamida, doxorubicina, vincristina, e prednisona [CHOP])); Rituxan® (rituximab, Genentech Inc., South San Francisco, CA, e Biogen Idec, Cambridge, MA; um anticorpo monoclonal aprovado para o tratamento de linfoma não Hodgkin; Ontak® (denileukin diftitox, Ligand Pharmaceuticals Inc., San Diego, CA; uma proteína de fusão consistindo em um fragmento de toxina diftérica fun-

dida geneticamente a interleucina-2); e Zevalin® (ibritumomab tiuxetan, Biogen Idec; um anticorpo monoclonal radiomarcado aprovado pela FDA para o tratamento de linfomas Não-Hodgkin de células B).

Para tratamento de Leucemia, biológicos típicos os quais podem ser usados em combinação com as moléculas de ligação da invenção incluem Gleevec®; Campath®-1H (alemtuzumab, Berlex Laboratories, Richmond, CA; um tipo de anticorpo monoclonal usado no tratamento de leucemia linfocítica crônica). Além disso, Genasense (oblimersen, Genta Corporation, Berkley Heights, NJ; pode ser usada uma terapia anti-sentido de BCL-2 em desenvolvimento para tratar leucemia (por exemplo, sozinho ou em combinação com um ou mais fármacos de quimioterapia, tais como fludarabina e ciclofosfamida) podem ser administrados com as moléculas de ligação reivindicadas.

Para o tratamento de câncer pumonar, biológicos típicos incluem Tarceva™ (erlotinib HCL, OSI Pharmaceuticals Inc., Melville, NY; uma molécula pequena projetada para orientar o caminho do receptor de fator de crescimento epidérmico humano 1 (HER1)).

Para o tratamento de mieloma múltiplo, biológicos típicos incluem Velcade® Velcade (bortezomib, Millennium Pharmaceuticals, Cambridge MA; um inibidor de proteassoma). Biológicos adicionais incluem Thalidomid® (thalidomide, Clegene Corporation, Warren, NJ; um agente imunomodulador e parece ter múltiplas ações, incluindo a capacidade para inibir o crescimento e a sobrevida de células de mieloma e antiangiogênese).

Outros biológicos típicos incluem o MOAB IMC-C225, desenvolvido pela ImClone Systems, Inc., New York, NY.

Além disso, as moléculas de ligação reivindicadas podem ser administradas em associação com vacinas ou outros agentes (por exemplo, citocinas) para modular reações imunes anticancerígenas. Por exemplo, Melacine® (Corixa Corporation, Seattle, WA) é uma vacina tumoral alogênica que foi reportada por ter resultados promissores no tratamento de melanoma ressecado T3N0M0. GMK® (Progenics Pharmaceutical, Inc., Tarrytown, NY) é um antígeno de gangliosídeo administradp como um agente adjuvante fase

III em pacientes que estão em alto risco para recurrência de melanoma. A vacina terapêutica antigastrina (Anti-gastrin therapeutic vaccine®,Aphton Corporation, Miami, FL) neutraliza hormônios G17 e *glyextened* e está em estudos clínicos fase III para pacientes com cânceres colorretais, pancreáticos, e estomacais. CeaVac® (Titan Pharmaceuticals, Inc., South San Francisco, CA) é um vacina de anticorpo antiidiotipo sendo estudado em câncer colorretal. Finalmente, Theratope® (Biomira Inc., Edmonton, Alberta, Canada) e uma vacina terapêutica de carboidrato sintética sendo investigada como um agente fase III em pacientes com câncer de mama metastático (Pharmaceutical Research e Manufacturers of America, 2000).

Em outra modalidade, uma molécula de ligação da invenção pode ser administrada em associação com um agente antiangiogênico, por exemplo, Endostatin (um inibidor endotelial-específico, endógeno, derivado de tumor, que detém a produção de células endoteliais microvasculares); anticorpo anti-VEGF; talidomida; ou inibidores de metaloproteinase da matriz inibem a síntese e degradação da membrana de base de vasos sangüíneos).

Conforme previamente discutido, as moléculas de ligação da presente invenção, fragmentos imunorreativos ou recombinantes dos mesmos podem ser administrados em uma quantidade farmaceuticamente eficaz para o tratamento *in vivo* de distúrbios de mamíferos. A este respeito, será reconhecido que as moléculas de ligação reveladas serão formuladas de modo a facilitar administração e promover estabilidade do agente ativo. Preferencialmente, farmacêuticas composições de acordo com a presente invenção compreendem um veículo estéril não-tóxico e farmaceuticamente aceitável, tal como salina fisiológica, tampãos não-tóxicos, preservantes e similares. Para os fins do presente requerimento, uma quantidade farmaceuticamente eficaz de uma molécula de ligação da invenção, conjugada ou não-conjugada a um agente terapêutico, deve significar uma quantidade suficiente para obter ligação eficaz a um alvo e para obtr um benefício, por exemplo, para melhorar sintomas de uma doença ou um distúrbio ou para detectar uma substância ou uma célula. No caso de células tumorais, o poli-

peptídeo preferencialmente será capaz de interagir com antígenos imunorreativos selecionados sobre células neoplásicas ou imunorreativas e proporcionam um aumento na morte destas células. Logicamente, as composições farmacêuticas da presente invenção podem ser administradas em doses únicas ou múltiplas proporcionando uma quantidade farmaceuticamente eficaz do polipeptídeo.

Dentro do âmbito da presente descoberta, os polipeptídeos da invenção podem ser administrados a um ser humano ou outro animal de acordo com os métodos de tratamento supracitados em uma quantidade suficiente para produzir um efeito terapêutico ou profilático. Os polipeptídeos da invenção podem ser administrados a similar ser humano ou outro animal em uma forma de dosagem convencional preparada combinando a molécula de ligação da invenção com um veículo ou diluente farmaceuticamente aceitável convencional de acordo com técnicas conhecidas. Será reconhecido por uma pessoa versada na técnica que a forma e a qualidade característica do veículo ou diluente farmaceuticamente aceitável são ditadas pela quantidade de ingrediente ativo com o qual deve ser combinado, a via de administração e outroas variáveis de conhecimento geral. Aqueles versados na técnica reconhecerão adicionalmente que um coquetel compreendendo uma ou mais espécies de polipeptídeos de acordo com a presente invenção pode se comprovar particularmente eficaz.

XIV. Métodos de Uso

As moléculas da invenção podem ser usadas em circunstâncias onde é desejável usar moléculas scFv estabilizadas ou composições compreendendo similares moléculas scFv, por exemplo, para fins de diagnóstico ou terapêuticos. Modalidades preferenciais da presente invenção proporcionam compostos, composições, kits e métodos para o diagnóstico e/ou tratamento de distúrbios que se beneficiariam de administração de uma molécula de ligação da invenção, por exemplo, distúrbios neoplásicos em um sujeito mamífero que necessite de similar tratamento. Preferencialmente, o sujeito é um ser humano.

Em uma modalidade, as moléculas de ligação em questão po-

dem ser usadas em um ensaio para detectar um antígeno tumoral in vitro, por exemplo, usando um ensaio ELISA. Ensaios típicos são conhecidos na técnica, vide, por exemplo, o Requerimnto de Patente dos Estados Unidos Número 20040077025.

Em outra modalidade, as moléculas de ligação em questão são úteis para detectar a presença de células portando antígeno tumoral usando tecnologia de setudos por imagens. Para similares aplicações, pode ser desejável conjugar a molécula de ligação a uma molécula detectável, por exemplo, uma radioetiqueta, conforme descrito adicionalmente abaixo.

Em outra modalidade, as moléculas de ligação em questão são úteis para reduzir ou eliminar células (por exemplo, por apoptose) portando um epítope (por exemplo, um epítope de Cripto ou um epítope de um membro da família de receptores de TNF, por exemplo, TRAIL-R2 ou LTβR) reconhecido por uma molécula de ligação da invenção. Em outra modalidade, as moléculas de ligação em questão são eficazes para reduzir a concentração de ou eliminar moléculas alvo solúveis na circulação.

Em outra modalidade, uma molécula de ligação da invenção reduz o tamanho do tumor, inibe o crescimento do tumor e/ou prolonga o tempo de sobrevida de um sujeito portando tumor. Por conseguinte, esta invenção também se refere a um método para tratar tumores em um ser humano ou outro animal administrando a um humano ou animal similar uma quantidade eficaz e não-tóxica de polipeptídeo. Uma pessoa versada na técnica seria capaz, por experimentação de rotina, de determinar qual quantidade eficaz e não-tóxica de polipeptídeo seria para os fins de tratar malignidades. Por exemplo, uma quantidade terapeuticoamente ativa de um polipeptídeo pode variar de acordo com fatores tais como o estágio da doença (por exemplo, estágio I versus estágio IV), a idade, sexo, complicações médicas (por exemplo, condições ou doenças de imunossupressão) e peso do sujeito, e a capacidade da molécula de ligação para provocar uma reação desejada no sujeito. O regime de dosagem pode ser ajustado proporcionando a resposta terapêutica ótima. Por exemplo, várias doses divididas podem ser administradas diariamente, ou a dose pode ser reduzida proporcional-

mente conforme indicado pelas exigências da situação terapêutica. De modo geral, no entanto, espera-se que uma dosagem eficaz esteja dentro da faixa de cerca de 0,05 a 100 miligramas por quilograma de peso corporal por dia e mais preferencialmente a partir de cerca de 0,5 a 10, miligramas por quilograma de peso corporal por dia.

Para fins de esclarecimento "mamífero" se refere a qualquer animal classificado como um mamífero, incluindo humanos, animais domésticos e de criação, e animais de zoológico, de competição, ou de estimação, tais como cães, cavalor, gatos, vacas, e etc. Preferencialmente, o mamífero é humano. "Tratamento" se refere tanto a tratamento terapêutico quanto profilático ou medidas preventivas. Aqueles que necessitam de tratamento incluem aqueles já com uma doença ou distúrbio bem como aqueles nos quais a doença ou distúrbio deve ser evitado. Portanto, o mamífero pode ter sido diagnosticado como tendo a doença ou distúrbio ou pode ser predisposto ou suscetível à doença.

Em geral, as composições descritas podem ser usadas profilaticamente ou terapeuticamente. Por exemplo, um neoplasma compreendendo um marcador que possibilite o direcionamento das células cancerosas pela molécula de ligação pode ser detectado ou inibido (por exemplo, destruído) usando uma molécula de ligação da invenção. Em uma modalidade preferencial, as moléculas de ligação da invenção são usadas para tratar tumores sólidos. Cânceres típicos que podem ser tratados incluem, mas não estão limitados a, carcinomas de próstata, gástricos tais como câncer de cólon, de pele, de mana, ovariano, pulmonar e pancreático. Em outra modalidade, os anticorpos da presente invenção podem ser usados para tratar sarcoma de Kaposi, neoplasias do sistema nervoso central (hemangioblastomas capilares, meningiomas e metástases cerebrais), melanoma, sarcomas gastrointestinal e renal, rabdomiossarcoma, glioblastoma (preferencialmente glioblastoma multiforme), leiomiossarcoma, retinoblastoma, cistadenocarcinoma papilar do ovário, tumor de Wilms ou carcinoma pulmonar de células pequenas. Será reconhecido que polipeptídeos apropriados podem ser derivados para moléculas associadas a tumor relacionadas com cada uma das neopla-

sias precedentes sem indevida experimentação em vista da presente descrição.

Malignidades hematológicas típicas que são suscetíveis a tratamento com a invenção revelada incluem linfoma de Hodgkins e não-Hodgkins bem como leucemias, incluindo ALL-L3 (leucemia do tipo de Burkitt), leucemia linfocítica crônica (CLL) e leucemias celulares monocíticas. Será reconhecido que os compostos e métodos da presente invenção são particularmente eficazes para tratar uma variedade de linfomas de células B, incluindo linfoma não-Hodgkin de baixo grau / folicular (NHL), linfoma celular (FCC), linfoma de células do manto (MCL), linfoma de células grandes difusas (DLCL), (SL) NHL linfocítico pequeno, NHL de grau intermediário / folicular, NHL difuso de grau intermediário, NHL imunoblástico de alto grau, NHL linfoblástico de alto grau, NHL de células pequenas não clivadas de alto grau, doença maciça de NHL e Macroglobulinemia de Waldenstrom. Deve ser evidente para os versados na técnica que estes linfomas freqüentemente terão diferentes nomes devido a sistemas de classificação em mudança, e que pacientes tendo linfomas classificados sob diferentes nomes também podem se beneficiar dos regimes terapêuticos combinados da presente invenção. Além dos distúrbios neoplásicos mencionados acima, será reconhecido que a invenção revelada pode ser usada vantajosamente para tratar malignidades adicionais portando moléculas associadas com tumores compatíveis.

Em uma modalidade, uma molécula de ligação da invenção é capaz de ligar especificamente a um antígeno de células tumorais e inibir o crescimento de células tumorais em um paciente. Em algumas modalidades, as células tumorais são células tumorais do cérebro, cabeça, pescoço, próstata, mama, testicular, cólon, pulmão, ovário, bexiga, uterinas, cervicais, pancreáticas e estômago. Em outras modalidades, uma molécula de ligação da invenção liga especificamente ao antígeno de células tumorais e inibe o crescimento de células tumorais as quais superexpressam o antígeno. Em uma modalidade, as células tumorais são linhagens celulares as quais superexpressam o antígeno, tais como linhagens celulares derivadas de cân-

ceres do cérebro, mama, testicular, cólon, pulmão, ovário, bexiga, uterino, cervical, pancreático e do estômago.

Em ainda outras modalidades as moléculas de ligação da presente invenção podem ser usadas para tratar distúrbios imunes que incluem, mas não estão limitados a, aspergilose broncopulmonar alérgica; rinite alérgica; anemia hemolítica auto-imune; Acanthosis nigricans; dermatite de contato alérgica; doença de Addison; dermatite atópica; Alopecia areata; Alopecia universalis; Amiloidose; púrpura anafilactóide; reação anafilactóide; anemia aplásica; Angioedema, hereditário; Angioedema, idiopático; espondilite anquilosante; Arterite, cranial; Arterite, células gigantes; Arterite, de Takayasu; Arterite, temporal; Asma; Ataxia-telangiectasia; ooforite auto-imune; orquite auto-imune; falência poliendócrina auto-imune; doença de Behcet; doença de Berger; doença de Buerger; bronquite; pênfigo bolhoso; Candidíase, mucocutânea crônica; síndrome de Caplan; síndrome de enfarte pós-miocardial; síndrome pós-pericardiotomia; Cardite; sprue celíaco; doença de Chagas; síndrome de Chediak-Higashi; doença de Churg-Strauss; síndrome de Cogan; doença de aglutinina do frio; síndrome de CREST; doença de Crohn; Crioglobulinemia; alveolite fibrosante criptogênica; Dermatite herpetifome; Dermatomiosite; Diabetes mellitus; síndrome de Diamond-Blackfan; síndrome de DiGeorge; lúpus eritematoso discóide; fasciite eosinofílica; Episclerite; Drythema elevatum diutinum; Eritema marginatum; Eritema multiforme; Eritema nodoso; febre familial do mediterrâneo; síndrome de Felty; Fibrose pulmonar; Glomerulonefrite, anafilactóide; Glomerulonefrite, auto-imune; Glomerulonefrite, pós-estreptocóccica; Glomerulonefrite, pós-transplantação; Glomerulopatia, membranosa; síndrome de Goodpasture; Granulocitopenia, imunomediada; Granuloma anular; Granulomatose, alérgica; miosite granulomatosa; doença de Grave; tiroidite de Hashimoto; doença hemolítica do recém-nascido; Hemocromatose, idiopática; púrpura de Henoch-Schoenlein; Hepatite, crônica ativa e crônica progressiva; Histiocitose X; síndrome hipereosinofílica; purpura trombocitopênica idiopática; síndrome de Job; dermatomiosite juvenil; artrite reumatóide juvenil (artrite crônica juvenil); doença de Kawasaki; Ceratite; Ceratoconjuntivite seca; síndrome de Landry-

Guillain-Barre-Strohl; Lepra, lepromatosa; síndrome de Loeffler; lúpus; síndrome de Lyell; doença de Lyme; granulomatose linfomatóide; Mastocitose, sistêmica; doença do tecido conjuntivo misto; Mononeurite múltipla; síndrome de Muckle-Wells; síndrome de linfonodos mucocutâneos; síndrome de linfonodos mucocutâneos; reticulohistiocitose multicêntrica; esclerose múltipla; miastenia gravis; micose fungóide; vasculite necrotizante, sistêmica; síndrome nefrótica; síndrome de sobreposição; Paniculite; hemoglobinúria do resfriado paroxísmico; hemoglobinúria noturna paroxísmica; Penfigóide; Pênfigo; Pênfigo eritematoso; Pênfigo foliaceus; Pênfigo vulgaris; doença do criador de pombos; Pneumonite, hipersensibilidade; Poliarterite nodosa; Polimialgia reumática; Polimiosite; Polineurite, idiopática; polineuropatias familiares portuguesas; Pré-eclâmpsia / eclâmpsia; cirrose biliar primária; esclerose sistêmica progressiva (escleroderma); Psoríase; artrite psoriática; proteinose alveolar pulmonar; fibrose pulmonar, fenômeno / síndrome de Raynaud; tiroidite de Reidel; síndrome de Reiter, policrondrite relapsante; febre reumática; artrite reumatóide; sarcoidose; esclerite; colangite esclerosante; doença do soro; síndrome de Sezary; síndrome de Sjogren; síndrome de Stevens-Johnson; doença de Still; panencefalite esclerosante subaguda; oftalmia simpática; lúpus eritematoso sistêmico; rejeição de transplante; colite ulcerativa; doença de tecido conjuntivo indiferenciada; Urticária, crônica; Urticária, resfriado; Uveite; Vitiligo; doença de Weber-Christian; granulomatose de Wegener e síndrome de Wiskott-Aldrich.

Em outra modalidade, as moléculas de ligação da invenção podem ser usadas para aplicações de pré-direcionamento. Por exemplo, as mesmas vantagens serão evidentes em aplicações de para iberação de fármacos quimioterápicos.

Por exemplo, em pré-direcionamento um tumor é pré-direcionado com um construto de ligação que tem afinidade para o antígeno associado ao tumor por um lado e, por exemplo, para um hapteno radiomarcado pelo outro. O hapteno radiomarcado é administrado posteriormente, preferencialmente depois do construto de ligação que tem afinidade para o antígeno associado ao tumor ter clareado (vide, por exemplo, Boerman et al.

2003. J. Nuclear Med. 44:400). Em outro exemplo, um anticorpo o qual é não-tóxico, mas foi derivado para reagir com um fármaco ou pró-farmaco que é tóxico somente quando ligado pela molécula de ligação. Dados os dados de biodistribuição nos presentes exemplos, as moléculas de ligação da invenção são bem adequadas para uso em aplicações de pré-direcionamento. Em uma modalidade, um agente de clareamento pode ser eliminado da metodologia de pré-direcionamento usando as presentes moléculas de ligação.

Esta invenção é adicionalmente ilustrada pelos seguintes exemplos os quais não devem ser considerados como limitantes. Os conteúdos de todas as referências, patentes e Pedidos de patente publicados citados do início ao fim deste requerimento são incorporados aqui, a este Pedido de patente, por meio de referência.

EXEMPLOS

Do início ao fim dos exemplos, foram usados os seguintes materiais e métodos a menos que determinado de modo diverso.

Materiais e Métodos Gerais

Em geral, a prática da presente invenção emprega, a menos que indicado de modo diverso, técnicas convencionais de química, biofísica, biologia molecular, tecnologia de DNA recombinante, imunologia (especialmente, por exemplo, tecnologia de anticorpo), e técnicas de rotina em eletroforese. Vide, por exemplo, Sambrook, Fritsch e Maniatis, Molecular Cloning: Cold Spring Harbor Laboratory Press (1989); Antibody Engineering Protocols (Methods in Molecular Biology), 510, Paul, S., Humana Pr (1996); Antibody Engineering: A Practical Approach (Practical Approach Series, 169), McCafferty, Ed., Irl Pr (1996); Antibodies: A Laboratory Manual, Harlow *et al.*, C.S.H.L. Press, Pub. (1999); e Current Protocols in Molecular Biology, eds. Ausubel *et al.*, John Wiley & Sons (1992).

Construtos de Expressão

Em geral, a menos que indicado de modo diverso, os construtos de expressão para scFvs nos seguintes Exemplos incluíram um peptídeo de sinal de Gene III de N-terminal bem como um peptídeo de purificação de C-

terminal compreendendo as etiquetas Myc e His e um sítio de clivagem de Enteroquinase. Seqüência de DNA para cada peptídeo são determinadas abaixo:

Seqüência de DNA de peptídeo de sinal de Gene III de N-terminal ATGAAAAAACTGCTGTTCGCGATTCCGCTGGTGGTGCCGTTCTATAGCCATA GT (SEQ ID NO: 25)

Seqüência de DNA de peptídeo de sinal de Gene III de N-terminal MKKLLFAIPLVVPFYSHS (SEQ ID NO: 26)

Seqüência de DNA de peptídeo de purificação de C-terminal GACGACGACGACAAAAGCTTTCTAGAACAAAAACTCATCTCAGAAGAGGAT CTGAATAGCGCCGTCGACCATCATCATCATCATTGA (SEQ ID NO: 27)

Seqüência de DNA de peptídeo de purificação de C-terminal DDDDKSFLEQKLISEEDLNSAVDHHHHHH (SEQ ID NO: 58)

Anticorpos

Anticorpos BIIB usados em alguns Exemplos (BIIB1-BIIB18) são uma coleção de anticorpos terapêuticos de várias especificidades. Dezessete dos 18 anticorpos foram expressados a partir de linhagens de células CHO volumosas ou clonais estáveis e 1 dos anticorpos (BIIB13) foi expressado usando transfecção transitória em células HEK293E. Todos os 18 anticorpos contêm cadeias leves kappa. A maioria dos anticorpos foram IgG1 humana; no entanto, BIIB2, BIIB5 e BIIB11 foram IgG4 humana. Dezessete dos 18 anticorpos foram humanos ou humanizados. BIIB15 foi PRIMATIZED® (Nakamura et al., 2000). A identidade da linha germinal humana para cada anticorpo foi avaliada por alinhamento ClustalW (ClustalW WWW Service at the European Bioinformatics Institute, Thompson et al., 1994) das seqüências BIIB V_H ou V_K contra as linhas germinais humanas disponíveis publicamente (Lefranc et al., 1999).

Quinze dos 18 anticorpos totais foram da subclasse IgG1 e os 3 restantes foram IgG4 (BIIB2, BIIB5 e BIIB11). As seqüências C_H1 de IgG1 e IgG4 têm diferenças de 10 aminoácidos (6 são conservativas) e um padrão de ligação dissulfeto alternado com a cadeia leve.

Um construto IgG1 e e IgG4 com uma região Fv em duplicata

estava disponível para investigar o efeito da subclasse de IgG sobre a estabilidade de Fab. Foram criados dois construtos os quais continham a região V_H de BIIB7 enxertaa ou a uma região constante de cadeia pesada IgG1 ou a uma IgG4 .

Exemplo 1. Preparação de Proteínas Convencionais BHA10 scFv e Fab

BHA10 scFv foi subclonada a partir do plasmídeo pXWU034, usando a reação de cadeia polimerase (PCR) com iniciadores de oligonucleotídeos mostrados na Tabela 2 abaixo. O iniciador forward BHA10-01F contém um único sítio de endonuclease de restrição *Sph* I (seqüência sublinhada) seguido por 18 bases de seqüência complementar ao gene de domínio variável de cadeia pesada de N-terminal BHA10. O iniciador reverso, BHA10-01R, contém 24 bases de seqüência complementar ao gene de domínio variável de cadeia leve de C-terminal BHA10, 15 bases de seqüência complementar a um sítio de enteroquinase, e sítios de endonuclease de restrição adjacentes *Hind* III e *Xba* I (sítios de endonuclease são sublinhados). Depois de amplificação por PCR, um produto de PCR correspondente ao tamanho esperado foi resolvido por eletroforese por gel agarose, excisado, e purificado usando o kit de extração Ultrafree-DA da Millipore de acordo com as instruções do fabricante (Millipore; Bedford, MA). O produto de PCR purificado foi digerido com *Sph* I, tornado de ponta cega digerindo com DNA Polymerase I na presença de dNTPs, e em seguida digerido com *Hind* III. O produto de PCR blunt-ended/ digerido com *Hind* III foi ligado a pKJS216 digerido por *Sca* I/ *Hind* III. pKJS216 é um vetor de *E. coli* que dirige a expressão de proteína recombinante sob o controle de um promotor ara C indutível. Uma porção da mistura de ligação foi usada para cepa de *E. coli* transformada XL1-Blue. Colônias resistente a fármaco de ampicilina foram triadas e análise da seqüência de DNA confirmou a seqüência correta do construto pIEH003 final. Seqüências de DNA e de aminoácidos de BHA10 scFv são mostradas nas Figuras 1A e 1B, respectivamente.

Tabela 2. Oligonucleotídeos para amplificação por PCR de um BHA10 scFv convencional.

Iniciadores	Seqüência
BHA10-01F (SEQ ID NO: 1)	5'- CAGTAGCATGCAGGTCCAACTGGTGCAG -3'
BHA10-02R (SEQ ID NO: 2)	5'-GTTCTAGAAAGCTTTTGTCGTCGTCGTCTTTGATCTC CACCTTGGTACCCTG -3'

Para expressão de BHA10 scFv, colônias recém isoladas de *E. coli* cepa W3110 (ATCC, Manassas, Va. Cat. nº 27325) transformadas com plasmídeo pIEH003 foram cultivadas em 4 x 250 ml de meio SB (Teknova, Half Moon Bay, Ca. Cat. nº S0140) contendo 50 μg/ml de carbenicilina em frascos *baffled* de 1 L para $OD_{600} \cong 0,8$, induzidas adicionando a 0,02% de arabinose, e cultivadas de um dia para o outro. Bactérias foram coletadas por centrifugação. Os péletes foram solubilizados e lisados usando 40 mL de reagente de extração de proteína B-PER (Cat nº 78243, Pierce). scFv solubilizado foi aplicado a uma coluna de 5 mL de Ni-NTA-Superflow (Cat nº 30410, Qiagen). scFv ligado foi lavado com 60 mM de imidazol, pH 8,0 e eluído com 300 mM de imidazol, pH 8,0. scFv eluído foi carregado sobre uma coluna de 6 mL de agarose Protein L (Cat nº 20510, Pierce). Proteína ligada foi lavada com salina tamponada com fosfato (PBS) e eluída com 0,1 M de glicina, pH 3,0. scFvs purificadas foram dialisadas contra PBS e armazenadas a –20 ºC. As concentrações de proteína foram determinadas usando um $\varepsilon_{280\ nm} = 2,1\ ml\ mg^{-1}\ cm^{-1}$.

Para preparação enzimática de BHA10 Fab, BHA10 IgG foi misturada com 4 μl de um estoque de papaína concentrado (0,3 mg/mL, 30 Unidades/mg – Cat nº 108014, Roche) em 8,3 ml de solução contendo 2,4 mg/mL de BHA10 IgG1, 100 mM de Tris-HCl, 20 mM de EDTA a pH 7.0. A reação foi deixada para prosseguir por 90 minutos a 25 ºC. A solução do digesto foi diluída a 1:5 com 20 mM de acetato, pH 5.0 e carregada sobre uma coluna de 6 ml SP-Sepharose FF equilibrada com tampão de diluição. A coluna foi lavada com 2 volumes de coluna de tampão de diluição. Fragmentos Fab brutos foram eluídos com 30 volumes de coluna de um gradiente linear de 0 a200 mM de NaCl. Um amplo pico centrando a 140 mM de NaCl (total de ~24 mL) foi coletado e continha a maior parte do material de IgG

digerida. O volume elutriado foi reduzido para 2 mL por concentração e carregado sobre uma coluna de preparativo G300 SW Tosohaas SEC (109 mL) equilibrada com PBS. Fab foi elutriado a 0,8 volumes de coluna sobre 15 mL. Fab purificado foi concentrado para entre 2 a 11 mg/mL. As concentrações de foram determinadas usando um $\varepsilon_{280\,nm}$ = 1,5 mL mg^{-1} cm^{-1}.

Exemplo 2. Estabilidade Térmica de Moléculas Convencionais BHA10 scFv

Calorimetria de Varredura Diferencial (DSC) foi usada para testar se uma BHA10 scFv isolada é intrinsicamente menos estável do que seu complemento Fab. Foram realizadas varreduras usando um calorímetro de varredura diferencial capilar automatizado (capDSC, MicroCal, LLC). Soluções de proteína e de referência foram amostradas automaticamente de lâminas de 96 cavidades usando a fixação robótica. Antes de cada varredura protéica, foram realizadas 2 varreduras com tampão na célula da amostra e usadas para subtração de fundo. Uma única varredura de limpeza foi realizada usando 5% de Liquinox depois de cada varredura de proteína. Depois de cada varredura, o instrumento enxaguou automaticamente tanto as células de referência quanto da amostra três vezes com 2 ml de H_2O deionizada destilada contendo 0,01% de azida de sódio. Foram realizadas varreduras a 1 ºC/min usando o modo de feedback de meio para reforçada resolução de pico. A faixa de varredura foi de 20 a 95 ºC. Todas as lâminas de 96 cavidades contendo proteína foram armazenadas dentro do instrumento a 6 ºC.

A Figura 2 mostra medições por DSC com fragmentos de anticorpos purificados BHA10 Fab e scFv. Dentro do calorímetro, a temperatura é aumentada até o Fab ou scFv desdobrar. A temperatura na qual cada proteína desdobra (isto é, o valor da T_M) pode ser indicativa da estabilidade global. Todos os quatro domínios do BHA10 Fab (V_H, V_L, C_H1 e C_L) desdobram cooperativamente a 78 ºC (Figura 2A). O domínio scFv carece dos domínios C_H1 e C_L. Sem este estruturação, os domínios da scFv desdobram em temperaturas muito menores do que o Fab e há uma redução significativa na entalpia calorimétrica dos domínios indicando a perda de interações estabilizante. O domínio V_L desdobra com uma T_M de 68 ºC, ao passo que o domínio V_H desdobra a 58 ºC, 20 ºC menos do que o que foi observado para a

transição de desdobramento do BHA10 Fab (Figura 2B). Adicionalmente, há uma dependência da taxa de varredura da T_M, sugerindo que está ocorrendo agregação de proteína durante a fase de aquecimento a qual reduz artificialmente a T_M da scFv à medida que a taxa de varredura é retardada (Sánchez-Ruiz et al., *Biochemistry*, 27: 1648-1652, 1988). A baixa estabilidade aparente e a propensão para agregar podem ser fatores determinantes para a incapacidade de células CHO para produzir quantidades significativas de material estável, não –agregado o qual contém scFvs tais como as moléculas de anticorpos biespecíficos Hercules convencionais.

Exemplo 3. Construção de Moléculas BHA10 scFv com Aprimorada Estabilidade Térmica

Sabendo que o domínio BHA10 scFv, conforme evidenciado no Exemplo 2, é intrinsicamente instável, foi sugerida a hipótese de que produzindo o scFv através do uso de tecnologia de DNA recombinante para produzir um scFv modificado que é termodinamicamente ou funcionalmente equivalente a um Fab sob condições de estímulo térmico deve resultar em um domínio scFv que é útil para construir um anticorpo biespecífico. Além disso, também foi sugerida a hipótese de que produzindo o domínio scFv isolado sozinho deve conferir quaisquer propriedades biofísicas benéficas que são obtidas quando reintroduzido como um componente de uma molécula biespecífica inteira. Até o momento, foi iniciado um esforço para melhorar a estabilidade biofísica do domínio BHA10 scFv usando um sistema de expressão de *E. coli* e melhoras monitoradas na estabilidade medindo ligação de domínios scFv termicamente resistentes a ligante em uma prova de ensaio térmico.

Para estabilizar os domínios scFv foram aplicados dois métodos: 1) introduzir uma ligação dissulfeto entre o domínio V_H e V_L do BHA10 scFv; e 2) otimizando a extensão do $(Gly_4Ser)_n$ ligante que conecta os domínios VH e V_L do BHA10 scFv.

A. Construção de BHA10 scFvs estabilizados por dissulfeto

O vetor de expressão bacteriana produtor BHA10 scFv, pIEH003, foi utilizado como o vetor parental. O kit QuickChange Site-Directed

Mutagenesis Kit (Stratagene; La Jolla, CA) foi usado, de acordo com as instruções do fabricante, para introduzir dois resíduos cisteína, um em V_H e um segundo em V_L que pode participar na formação de uma ligação dissulfeto estabilizante. Pares de iniciadores VH44-F e VH44-R (Tabela 3) foram usados para mutagenizar o resíduo Gly (GGA) na posição 44 (sistema de numeração de Kabat) de cadeia pesada variável BHA10 para um resíduo Cys (TGC). O produto de mutagênese foi digerido com enzima sensível a metilação *Dpn* I de acordo com o protocolo do kit e transformado na cepa de *E. coli* XL10-GOLD® (Stratagene; La Jolla, CA). Colônias de *E. coli* transformadas para resistência a fármaco de ampicilina foram triadas para a mutação de seqüência correta por análise de seqüência de DNA. O plasmídeo resultante, pIEH004, foi utilizado para uma reação subseqüente para mutar o resíduo Gln (CAG) na posição 100 (sistema de numeração de Kabat) de cadeia leve variável BHA10 para um resíduo Cys (TGC) usando pares de iniciadores VL100-F e VL100-R (Tabela 3). Iniciadores forward (VH44-F) e reverso (VH44-R) mutam Gly (GGA) para um Cys (TGC) na posição 44 de V_H (TGC indicado por seqüência sublinhada). Iniciadores forward (VL100-F) e reverso (VL100-R) mutam Gln (CAG) para um Cys (TGC) na posição 100 de V_L (TGC indicado por seqüência sublinhaad).

Colônias de *E. coli* XL10-GOLD® transformadas para resistência a fármaco de ampicilina foram triadas para a mutação de seqüência correta por análise de seqüência de DNA e plasmídeo pIEH006 foi identificada como contendo as duplas mutações de cisteína nas posições V_H44 e V_L100. Seqüências de DNA e de aminoácidos de VH44/VL100 dissulfeto-estabilizada BHA10 scFv são mostradas nas Figuras 3A e 3B, respectivamente.

Tabela 3. Oligonucleotídeos para construção de VH44/VL100 dissulfeto-estabilizado BHA10 scFv.

Iniciadores	Seqüência
VH44-F (SEQ ID NO: 5)	5'- GCAGGCCCCTGGACAGTGCCTTGAGTGGATGGGATG -3'
VH44-R (SEQ ID NO: 6)	5'- CATCCCATCCACTCAAGGCACTGTCCAGGGGCCTGC -3'
VL100-F (SEQ ID NO: 7)	5'- CCTATCCATTCACGTTCGGCTGCGGTACCAAGGTGG AGATC -3'
VL100-R (SEQ ID NO: 8)	5'- GATCTCCACCTTGGTACCGCAGCCGAACGTGAATGG ATAGG -3'

B. Construção de BHA10 scFv com ligantes $(Gly_4Ser)_n$ alternativos

Plasmídeo pIEH003 codificando huBHA10 scFv com o $(Gly_4Ser)_3$ ligante convencional foi modificado para conter um $(Gly_4Ser)_4$ ou $(Gly_4Ser)_5$ ligante por amplificação por PCR usando os iniciadores de oligonucleotídeos descritos na Tabela 4.

BHA10 scFv $(Gly_4Ser)_4$ foi montado usando um iniciador de PCR 5' forward designado pXWU002-F1 e um iniciador de PCR 3' reverso designado XWU002-R. O iniciador de PCR 5' VH XW002-F1 incluiu um sítio de endonuclease de restrição *Btg* I (seqüência sublinhada) localizado no término carboxila de BHA10 VH seguido por seqüência codificando um $(Gly_4Ser)_4$ ligante. O iniciador de PCR 3' VL XW002-R incluiu um sítio *Xba* I e um sítio de Enteroquinase parcial. As regiões parciais BHA10 VH + $(Gly_4Ser)_4$ ligante + as BHA10 VL foram amplificadas em uma reação de PCR usando a série de iniciadores de PCR XW002-F1/ XW002-R do DNA de plasmídeo pIEH003 (descritas no Exemplo 1). O fragmento genético BHA10 scFv- $(Gly_4Ser)_4$ ligante parcial correspondente ao tamanho esperado foi resolvido por eletroforese de gel agarose, excisado, e purificado usando o kit de extração Ultrafree-DA da Millipore de acordo com as instruções do fabricante (Millipore; Bedford, MA). O produto de PCR purificado foi digerido e clonado no vetor pIEH003 digerido por *Btg* I/*Xba* I resultando no plasmídeo pXWU002 codificando BHA10 scFv contendo um $(Gly_4Ser)_4$ ligante. BHA10 scFv contendo o $(Gly_4Ser)_5$ ligante foi construído em modo similar usando iniciadores de PCR XW003-F e XW002-R para produzir o plasmídeo pXWU003. O iniciador de

PCR 5' forward (XWU003-F) continha um sítio *Btg* I (seqüência sublinhada) seguido por seqüência codificando uns poucos aminoácidos do término carboxila de BHA10 VH e seqüência codificando um $(Gly_4Ser)_5$ ligante parcial. Seqüências corretas foram confirmadas por análise de seqüência de DNA. Seqüências de DNA e de aminoácidos de BHA10 scFv contendo o $(Gly_4Ser)_4$ ligante são mostradas na Figuras 4A e 4B, respectivamente. Seqüências de DNA e de aminoácidos de BHA10 scFv contendo o $(Gly_4Ser)_5$ ligante são mostradas nas Figuras 5A e 5B, respectivamente.

Tabela 4. Oligonucleotídeos para construção de BHA10 scFv com ligantes $(Gly_4Ser)_4$ ou $(Gly_4Ser)_5$.

Iniciadores	Seqüência
XW002-F1 (SEQ ID NO: 11)	5'- AAGGGACCACGGTCACCGTCTCCTCAGGCGGTGGAGGGTCCG GTGGGGGCGGATCTGGGGGCGGCGGATCCGGTGGTGGTGGTAG-3'
XW002-R (SEQ ID NO: 12)	5'- TTTTGTTCTAGAAAACTTTTGTCGTCG-3'
XW003-F (SEQ ID NO: 13)	5'- AAGGGACCACGGTCACCGTCTCCTCAGGAGGGGGCGGTTCAG GCGGTGGAGGGTCCGGTGGGGGCGGATCTGGGGGCGGCGGATC-3'

Exemplo 4. Caracterização de Moléculas BHA10 scFv com Aprimorada Estabilidade Térmica

A. Expressão e Análise Western blot de BHA10 scFvs Engenheirados

Para expressão de BHA10 scFvs engenheirados, a cepa de *E. coli* W3110 (ATCC, Manassas, Va. Cat. nº 27325) foi transformada com plasmídeos pIEH003, pXWU002, pXWU003 e pIEH006 e colônias resistentes a ampicilina selecionadas e cultivadas em 10 ml de meio SB (Teknova, Half Moon Bay, Ca. Cat. nº S0140) contendo 50 µg/ml de carbenicilina em um tubo de centrífuga cônico de 50 ml para $OD_{600} \cong 0,8$, foram induzidas adicionando a 0,02% arabinose, e cultivadas de um dia para o outro. Bacté-

rias foram coletadas por centrifugação e os péletes foram ressuspensos em volume a 1/20 de uma solução gelada iso-osmótica de 50 mM de Tris-HCl, pH 8.0, 1 mM de EDTA, e 20% de sucrose (peso/ vol) e esfriado sobre gelo. Volume igual de 50 mM de Tris-HCl, pH 8.0, 1 mM de EDTA, e 20% de sucrose (peso/ vol) contendo 2 mg/ml de Lysozyme (Sigma) foi adicionado à suspensão bacteriana e incubada sobre gelo com mistura ocasional por 10 minutos. A suspensão bacteriana foi centrifugada por 10 minutos a 8000 x *g*, 4 °C e a fração periplásmica foi retida.

As amostras foram misturadas com tampão de amostra nativa ou tampão de amostra contendo o agente redutor ditiotreitol e aquecido a 90°C por 3 minutos. Amostras reduzidas e não reduzidas foram eletroforesadas sobre um gel de Tris-glicina poliacrilamida SDS-PAGE e transferido eletroforeticamente sobre uma membrana de nitrocelulose (Invitrogen, Life Technologies, Carlsbad, CA). A membrana foi bloqueada com PBS contendo 5% (peso/ vol) de leite sem gordura e 0,1% de Triton X-100 e incubada com um anticorpo kappa anti-humano (Roche Applied Science, Indianapolis, IN). A membrana foi lavada e em seguida incubada com um anticorpo HRP anti-coelho (Amersham Biosciences, Piscataway, NJ). Complexos imunes foram detectados usando o sistema de análise Western Blotting ECL de acordo com o fabricante (Amersham Biosciences, Piscataway, NJ).

Foi visto que um dos três pares de dissulfeto testados, isto é, V_H44:V_L100, produziram quantidades adequadas de proteína quando expressadas em E. coli (Figura 6, alameda 2). (Os dissulfetos V_H 44:V_L 105 e V_H 106:V_L43 testadas não produziram tanto scFv intacto). Similarmente, prolongando a extensão do $(Gly_4Ser)_n$ ligante para n = 4 (alameda 3) ou n = 5 (não mostrado) também produziu quantidades adequadas de proteína em *E. coli*. Combinando tanto as modificações V_H44:V_L100 e $(Gly_4Ser)_4$ ligante em BHA10 scFv usando os métodos decritos no Exemplo 3 também levou a quantidades adequadas de proteína expressada (Figura 6, alameda 4). Seqüências de DNA e de aminoácidos de BHA10 scFv contendo a combinação de modificações VH44:V_L100 e $(Gly_4Ser)_4$ ligante são mostradas nas Figuras 7A e 7B, respectivamente. Em ambos os casos onde o BHA10 scFv contém

as mutações V_H44:V_L100, foi visto que os scFvs migram com aumentada mobilidade no gel de poliacrilamida desnaturante, não-redutor, ainda migraram similarmente a BHA10 scFv convencional sob condições desnaturantes, redutoras (Figura 6, alamedas 2 e 4). Esta análise sugere que as variantes de BHA10 scFv contendo as mutações V_H44:V_L100 provavelmente formam ligações dissulfeto intactas e talvez obtendo uma estrutura mais compacta.

B. Prova de desnaturação térmica.

As atividades dos BHA10 scFvs convencionais e engenheirados foram então comparadas em uma prova de ensaio térmico a qual pode ser usada para determinar a temperatura na qual 50% das moléculas scFv retém sua atividade de ligação antigênia depois de um evento de estímulo térmico. O valor numérico correspondente a esta temperatura é referido como o valor "T_{50}" e as unidades são em °C. Neste ensaio, os scFvs foram submetidos a uma faixa de temperaturas que engloba a temperatura de transição térmica de BHA10 scFv convencional.

A cepa de *E. coli* W3110 (ATCC, Manassas, Va. Cat. nº 27325) foi transformada com plasmídeos codificando os BHA10 scFvs convencionais e engenheirados sob o controle de um promotor ara C indutível. Transformantes foram cultivados de um dia para o outro em meio de expressão consistindo em SB (Teknova, Half Moon Bay, Ca. Cat. nº S0140) suplementado com 1% de glicina, 1% de Triton X100, 0,02% de arabinose, e 50 µg/ml de carbenicilina ou em 37°C ou em 32°C. Bactérias foram peletizadas por centrifugação e os sobrenadantes foram colhidos para tratamento adicional. Incluindo glicina e Triton X-100 no meio resulta na liberação de conteúdos periplásmicos (proteína de *E. coli* nativa e scFv) no meio (Yang *et al.* Applied e Enviromental Microbiology. (1998) 64:2869-2874). A presença de proteínas de *E. coli* no sobrenadante é essencial para a performance deste ensaio porque as proteínas termicamente desnaturadas agem como um "sorvedouro", capturando moléculas scFv desdobradas transitoriamente em agregados inativos irreversíveis.

Cada biblioteca foi triada em duplicata usando uma prova de ensaio térmico com sobrenadante de uma replicata submetido a condições

de tratamento e o segundo sobrenadante servindo como referência não-tratada. Provas de desnaturação térmica podem ser realizadas em uma única temperatura ou faixa de temperaturas para medir a estabilidade. Máquinas Thermocycler capazes de gerar gradientes térmicos estáveis foram usados para tratar sobrenadantes da amostra (iCycler, Bio-Rad, Gaithersburg, Md).

A temperatura de estímulo variou dependendo das propriedades do variante de BHA10 scFv parental e foi geralmente dois a três graus Celsius maior do que o valor T_{50} determinado experimentalmente. Lâminas Frozen Master foram degeladas e usadas para inocular lâminas de microtítulo de cavidades profundas contendo 250 µl de meio de expressão por cavidade, e culturas cultivadas de um dia para o outro a 32°C. Como um controle, culturas contendo o plasmídeo parental foram cultivadas sob as mesmas condições e processadas simultaneamente à biblioteca. Bactérias foram peletizadas e alíquotas de 50 a 100 µl de sobrenadante de ensaio foram colocadas ou tubos de faixa de PCR (Applied Biosystems, Foster City, Ca, Cat. nº N801-535) ou lâminas de 96 cavidades (Applied Biosystems, Foster City, Ca, Cat. nº N801-560) e as amostras foram aquecidas por 60 a 90 minutos. As amostras foram transferidas para lâminas de 96 cavidades de fundo em v (Corning, Corning, NY, Cat. nº 3357) e centrifugadas em uma centrífuga clínica refrigerada (IEC modelo 8R, Thermo Electron, Waltham, Ma) por 30 minutos, e 100 µl do sobrenadante foi transferido para lâminas de microtítulo de rotina (Corning, Corning, NY, Cat. nº 3357). Uma alíquota do sobrenadante foi reservada para o DELFIA de referência. Para a maioria das bibliotecas, as lâminas contendo o restante dos sobrenadantes foram vedados (Nalge Nunc, Rochester, NY, Cat nº 235205) e colocados em um incubador ajustado para a temperatura de estímulo apropriada (Echo Therm, Torrey Pines Scientific, San Marcos, Ca) por 90 minutos. Para traiagens requerendo múltiplas temperaturas de estímulo (ou para temperaturas acima de de 75°C) os sobrenadantes foram transferidos para lâminas de PCR de 96 cavidades (Applied Biosystems, Foster City, Ca, Cat. nº N801-560) e incubadas por 90 minutos na temperatura desejada.

Depois de estímulo térmico, as amostras foram centrifugadas a 2.000 RPM para remover o material agregado. Amostras de BHA10 scFv solúveis remanescentes no sobrenadante clareado e tratado foram avaliadas para ligação a antígeno LTβR Ig cognato por prova DELFIA. Lâminas de 96 cavidades (MaxiSorp, Nalge Nunc, Rochester, NY, Cat. nº 437111) foram revestidas com proteína de fusão consistindo no ectodomínio do LTβ receptor (LTβR) fundido a uma região Fc humana a 1 μg/ml em 0,1 M de tampão de carbonato de sódio, pH 9.5. As lâminas foram revestidas de um dia para o outro a 4°C, e bloqueadas com tampão de prova DELFIA (DAB, 10 mM de Tris HCl, 150 mM de NaCl, 20 μM de EDTA, 0,5 % de BSA, 0,02% de Tween 20, 0,01% de NaN_3, pH 7.4) por uma hora com agitação em temperatura ambiente. As lâminas foram lavadas 3 vezes com DAB sem BSA (tampão de lavagem), e amostras de ensaio diluídas em DAB foram adicionadas às lâminas em um volume final de 100 μl. As lâminas foram incubadas por uma hora com agitação em temperatura ambiente, e em seguida lavadas 3 vezes com tampão de lavagem para remover moléculas scFv não ligadas e funcionalmente inativadas. BHA10 scFv ligado foi detectado por adição de 100 μl por cavidade de DAB contendo 40 ng/ml de anticorpo anti-His_6 marcado com-Eu (Perkin Elmer, Boston, MA, Cat. nº AD0109) e incubado em temperatura ambiente com agitação por uma hora. As lâminas foram lavadas 3 vezes com tampão de lavagem, e 100 μl de solução de reforço DELFIA (Perkin Elmer, Boston, MA, Cat. nº 4001-0010) foi adicionado por cavidade. Depois de incubação por 15 minutos, as lâminas foram lidas usando o método do Európio sobre um Victor 2 (Perkin Elmer, Boston, MA).

Dados de ensaio foram processados usando o software Spotfire DecisionSite (Spotfire, Somerville, Ma.) e expressados como a proporção das contagens DELFIA observadas em temperatura de estímulo para a temperatura de referência para cada clone. Clones que deram reprodutivelmente proporções maiores do que ou iguais a duas vezes a que foi observada para o plasmídeo parental foram considerados alcences. DNAs de plasmídeos destes clones positivos foram isolados por mini-prep (Wizard Plus, Promega, Madison, WI) e retransformados de volta em *E. coli* W3110 para provas de

ensaio térmico secundário de confirmação.

Para gradientes térmicos, os dados foram analisados usando o software Prism 4 (GraphPad Software, San Diego, Ca) usando uma reação de dose sigmóide com inclinação variável como o modelo. Os valores obtidos para o ponto central das curvas de desnaturação térmica são referidos como valores de T_{50}, e não são considerados como sendo equivalentes a valores de Tm biofisicamente derivados.

Os resultados da prova de ensaio térmico são representados na Figura 8. Conforme representado na Figura 8, todas as moléculas scFv estabilizadas da invenção resultaram em aprimoramentos na atividade de ligação (T_{50} > 49°C) comparadas com a scFv convencional. Em particular, os valores de T_{50} de BHA10 posição de biblioteca V_L46 scFv (S46L), posição de biblioteca V_H16 scFv (S16E e S16Q), e posições de biblioteca V_L49: V_L50 scFv apresentaram aumentos na estabilidade térmica variando de + 3°C a + 12°C em relação ao BHA10 scFv convencional. Além disso, os valores de T_{50} de BHA10 posição de biblioteca V_L3 scFv (Q3A, Q3G, Q3S, Q3V, e Q3D), posição de biblioteca V_H67 scFv (V67I e V67L), posição de biblioteca V_H48 scFv (M48I e M48G), posição de biblioteca V_H20 scFv (V20I) e posição de biblioteca V_H101 scFv (P101D) apresentaram aumentos na estabilidade térmica variando de + 4° C a + 18°C em relação ao BHA10 scFv convencional. Uma das mutações estabilizantes (V_L K13E) acidentalmente resultou de um erro de PCR. Foi visto que a incorporação de uma destas mutações estabilizantes em pIEH009 aumenta adicionalmente a estabilidade térmica e mesmo excede a de BHA10 Fab sob estas condições. De modo importante, a V_L46 não-covalente, mutação V_H101, e mutações V_H55 derivadas de um ou mais dos quatro métodos de criação (modelagem por homologia da interface V_L/V_H, classificação de consenso, modelagem computacional, e análise de covariação) resultaram em um aprimoramento na estabilidade térmica de scFv quase se aproximando daquela observada com as mutações dissulfeto e validando a utilidade e inovação destas ferramentas de criação. Em particular, a T_{50} do construto pIEH006, BHA10 V_H44:V_L100 scFv (59°C), diferiu de BHA10 Fab (62°C) por somente 3 °C, ao passo que foi visto que o cons-

truto pIEH009, BHA10 scFv V_H44:V_L100/$(Gly_4Ser)_4$ ligante, e funcionalmente equivalente a BHA10 Fab sob estas condições. Estes resultados demonstram que as scFvs estabilizadas da invenção têm atividade aprimorada depois de um evento de estímulo térmico.

c. Medições de Afinidade

Calorimetria de titulação isotérmica (ITC) foi usada para medir a afinidade de sLTβR para o BHA10 Fab derivado usando clivagem enzimática da molécula BHA10 IgG1. sLTβR foi preparado conforme descrito previamente (Eldredge *et al.*, *Biochemistry,* (2006)). O Fab e sLTβR foram concentrados para 6,0 e 2,0 mg/mL, respectivamente, usando dispositivos de filtro de ultracentrífuga Amicon (MWCO 10.000). As soluções de estoque concentradas foram simultaneamente dialisadas contra PBS antes das medições por ITC. Foi realizada ITC em uma unidade VP-ITC (MicroCal LLC, Northampton, MA) ajustada para 30 ºC. Aproximadamente 500 µL de uma solução a 7 µM de sLTβR foi colocada na célula da amostra e dialisado de PBS foi colocado na célula de referência. Um total de 234 µL de BHA10 Fab a 70 µM foi titulado para a célula da amostra em 7 x 10 µL, 12 x 7 µL, seguido por 8 x injeções de 10 µL. A estequiometria da reação foi 1:1. Curvas de ITC foram analisadas usando o programa Origin Software fornecido pelo fabricante.

Medições K_D de preparações de BHA10 scFv convencional, BHA10 $(Gly_4Ser)_4$ scFv, e BHA10 V_H44:V_L100/$(Gly_4Ser)_4$ scFv conforme descrito nos Exemplos 3 e expressadas e purificadas usando os métodos descritos no Exemplo 1 foram realizadas usando ressonância de plasmon superficial (SPR) em um instrumento Biacore 3000 (Biacore Inc., Piscataway, NJ). Todos os experimentos foram realizados em tampão de HBS-EP, pH 7.4. Anticorpo PENTA-His biotinilado (20 µg/mL, Cat nº 34440, Qiagen) foi imobilizado sobre um chip CM5 revestido com estreptavidina em uma taxa de fluxo de 10 µl/min por aproximadamente 1 minuto. Soluções de ensaio a 0,1 µM de scFvs foram injetadas sobre o chip em uma taxa de fluxo de 5 µl/min por 10 minutos e capturadas sobre a superfície através do anticorpo PENTA-His imobilizado. Para investigar ligação entre sLTβR e scFv captura-

do, uma série de concentração de sLTβR (1, 2, 5, 10, 25, 50, 100, e 200 nM) foi duplamente injetada em uma taxa de fluxo de 30 µl/min sobre a superfície revestida com scFv. O fundo foi subtraído dos sensorgramas da amostra de ensaio usando dados de sensorgrama de uma célula de fluxo contendo tampão somente e dados de sensorgrama da superfície PENTA-His onde scFv foi injetado seguido por uma injeção de tampão ao invés de sLTβR. As curvas foram analisadas usando o software de instruções do BiaEval 3.0. Valores de K_D foram calculados ajustando as curvas de associação e dissociação cinética poara um modelo de ligação de Langmuir a 1:1. As superfícies dos chips foram regeneradas por duas injeções de 10 µL consecutivas de 0,1 M de glicina, pH 3.0.

A Tabela 5 mostra os resultados da prova de afinidade Biacore. A afinidade de ligação de BHA10 scFv produzido de modo recombinante é essencialmente a mesmo que o BHA10 Fab preparado conforme descrito no Exemplo 1. Adicionalmente, introdução do $(G_4S)_4$ ligante somente ou na presença do dissulfeto (V_H44:V_L100) estabilizante, ou a combinação dos dois não resultou em qualquer perda de afinidade significativa para o antígeno.

Tabela 5. Prova de afinidade Biacore -ligação a LTβR.

Construto	Kd (nM)
BHA10 scFv	3,7
BHA10 $(Gly_4Ser)_4$ scFv	2,5
BHA10 V_H44:V_L100/$(Gly_4Ser)_4$ scFv	4,3
BHA10 Fab (CHO)	1,9 ± 1,0*

*Medido por calorimetria de titulação isotérmica

D. Estudos de Calorimetria de Varredura Diferencial

Análises por Calorimetria de Varredura Diferencial (DSC) foram realizadas com preparações de BHA10 scFv, BHA10 $(Gly_4Ser)_4$ scFv, e

BHA10 V_H44:V_L100/$(Gly_4Ser)_4$ scFv. Os experimentos foram realizados conforme descrito acima para as comparações iniciais do scFv convencional contra o Fab produzido enzimaticamente exceto que fora realizados em uma taxa de varredura de 4 ºC/min. BHA10 V_H44:V_L100/$(Gly_4Ser)_4$ scFv demonstrou propriedades superiores de termoestabilidade comparado com o scFv selvagem. Em particular, o menos estável dos dois domínios, o domínio V_H de BHA10 V_H44:V_L100/$(Gly_4Ser)_4$ scFv, desnaturou em aproximadamente 5 ºC acima da de BHA10 scFv convencional (Figura 9). Aumentos na "temperatura de fusão" ou T_M podem ser devido a dois fatores separados, (i) um aumento na termoestabilidade do estado de equilíbrio dobrado ou (ii) uma tendência reduzida para agregar. A diferença de temperatura ~ 5 ºC entre o BHA10 scFv convencional e BHA10 V_H44:V_L100/$(Gly_4Ser)_4$ scFv foi relativamente independente da taxa de varredura (entre 0,2 e 4,0 ºC/min) sugerindo a mudança de T_M observada foi devida a estabilização do estado de equilíbrio dobrado.

E. Estudos de Ligação de ANS

A capacidade para ligar o corante fluorescente hidrofóbico sulfonato de 1-anilino-8-naftalina (ANS) é freqüentemente um marco de interações com regiões hidrofóbicas significativamente maiores ou em um estado nativo da proteína ou em um estado parcialmente desdobrado que pode ocorrer em tratamento com temperatura aumentada. A fluorescência intrínsica do corante é enxtinta em solvente e significativamente aumenta quando exposta a grandes áreas superficiais hidrofóbicas. O BHA10 scFv convencional intrinsicamente liga ANS em um nível muito mais forte do que o BHA10 $(Gly_4Ser)_4$ scFv, ou BHA10 V_H44:V_L100/$(Gly_4Ser)_4$ scFv indicando a presença de exposição hidrofóbica forçada no scFv por um ligante de tamanho inadequado. A adição do ligante mais longo, especificamente $(Gly_4Ser)_4$, pareceu mediar este efeito. A exposição aparente de uma superfície hidrofóbica pelo BHA10 scFv convencional pode levar a um nível aumentado de agregação na presença de outras proteínas ou moléculas, ou mesmo na presença do próprio em isolamento. Aquecendo BHA10 scFv acima de sua T_M pareceu induzir ligação de ANS para os BHA10 scFvs convencionais e engenhei-

rados. De modo interessante, o BHA10 scFv convencional mostrou um aumento gradual na ligação de ANS, como uma função de temperatura crescente, sugerindo possíveis aumentos na exposição hidrofóbica em temperaturas abaixo da T_M – tais como temperaturas usadas em cultura de células bacterianas e de mamíferos (isto é, 37 ºC) (Figura 10). Em contraste, tanto BHA10 $(Gly_4Ser)_4$ e BHA10 $V_H44:V_L100/(Gly_4Ser)_4$ scFvs não apresentam esta propriedade sugerindo que as mutações estabilizantes descritas nesta invenção reduzem a tendênciar para auto-associar ou associar com outros componentes protéicos da cultura celular com base em reduzida exposição de áreas superficiais hidrofóbicas.

Exemplo 5. Identificação de Mutações Estabilizantes as quais conferem Aprimorada Estabilidade Protéica Intrínseca

(i) Identificação de Mutações Estabilizantes

Uma variedade de métodos à base de seqüência (por exemplo, classificação de consenso, análise de covariação, modelagem por homologia de interface VH/VL) foram usadas para identificar mutações estabilizantes as quais conferiram aprimorada estabilidade protéica a moléculas de ligação de interesse. Estas mutações estabilizantes foram usadas na criação e construção de bibliotecas de expressão de região variável de anticorpo.

A. Análise de covariação

Uma série de projetos foram desenvolvidos para estabilizar os domínios VH e V_L de BHA10 scFv com base na forte covariação exibida por dois ou mais resíduos dentro de uma única seqüência. Foi realizada análise de covariação sobre os domínios VH e V_L de BHA10 scFv usando métodos similares aos descritos no Exemplo 17 abaixo de modo a identificar covariações ausentes ou violativas de tal modo que mutações estabilizantes possam ser previstas e incluídas em uma biblioteca de expressão de região variável de anticorpo para triagem experimental.

Em um primeiro exemplo, mutação de Ser para Leu na posição 46 (S46L; sistema de numeração de Kabat) dentro da V_L de BHA10 exibiu uma conexão positiva com um Tyr existente no resíduo 36 o qual a Ferramenta de Análise de Covariação mostrou que covaria fortemente com Leu

46 (vide a Figura 78). Esta mutação também foi prevista como sendo estabilizante por análise de freqüência de resíduos.

Além de S46L, uma segunda mutação prevista por covariação como sendo estabilizante foi uma mutação V55G dentro do domínio V_H de BHA10. Apesar da freqüência de resíduo ter indicado que Val é observado infreqüentemente nesta posição, esta posição é embutida dentro de CDR2 e é variável. Portanto, sem informação adicional, não foram tentadas previamente alterações nesta posição. Na inspeção por análise de covariação, no entanto, os dados de covariação sugeriram que esta posição foi fortemente correlacionada com no mínimo 10 aminoácidos diferentes que já existem dentro do domínio V_H de BHA10. Mutação para Gly na posição 55 satisfez todas as 10 covariações.

Outro exemplo da utilidade da Ferramenta de Análise de Covariação foi o efeito negativo previsto de uma mutação Q6E de V_H de BHA10. Análise de freqüência de resíduo único sugeriu que mutação para o Glu muito mais comumente observado nesta posição levaria a um aumento em estabilidade. No entanto, a Ferramenta de Análise de Covariação indicou que mutação única para Glu viola várias covariações existentes presentes dentro da seqüência V_H de BHA10. Para obter um aprimoramento na estabilidade, deve-se substituir vários aminoácidos que preferencialmente estabilizam Glu nesta posição.

Ainda, outro exemplo do valor preditivo da Ferramenta de Análise de Covariação é mostrado na Figura 79. Met 80 foi mutada para Leu como parte de um projeto de biblioteca de resíduo único. Imaginava-se que esta única mutação seria altamente estabilizante, uma vez que Leu é o aminoácido mais freqüente observado nesta posição dentro do banco de dados de seqüências. No entanto, análises de covariação indicam que dois outros aminoácidos devem ser mutados (V67F e T70S) de modo a obter harmonia de covariação.

B. Classificação de Consenso

Classificação de Consenso foi utilizada como um método para identificar resíduos aminoácidos dentro das regiões V_H e V_L de scFv para

mutagênese para aumentar a estabilidade intrínseca da scFv. A classificação avalia a tendência relativa de seqüências V_H e V_K de consenso devido a mutações hipersomáticas e variações de linhas germinais evolucionárias. Informação derivada desta análise foi então usada para projetar uma biblioteca para triagem para variantes de scFv com aprimorada estabilidade.

A série de referência de seqüências V_H e V_K kappa de mamífero usadas para derivar a seqüência de consenso para classificação e as freqüências de aminoácidos individuais em cada posição de resíduo foram compiladas, classificadas, e selecionadas conforme descrito previamente (Demarest, *et al., J. Mol. Biol.* 335, 41-48 (2004); Demarest, *et al., Protein Eng. Des. Sel. In Press,* (2006). As séries de referência de mamífero foram construídas de forma simples (*naively*) para incluir V-genes de vários mamíferos de modo a obter diversidade através da tendência evolucionária entre espécies. A série de referência de V_H de mamífero contém 61 seqüências V_H essencialmente de NCBI e TIGR representando um total de 17 espécies de mamíferos diferentes. A série de referência de V_K de mamífero contém 53 seqüências V_K de 13 espécies de mamíferos diferentes.

Análises estatísticas das V_H e V_L de BHA10 foram realizadas usando bancos de dados de IgG designados usuais e um script PERL modificado (Demarest *et al.*, 2004; Demarest *et al.*, 2006). A freqüência de aminoácidos de cada resíduo dentro das V_H e V_L de BHA10 foi calculada a partir do banco de dados usual. A freqüência de resíduos de cada aminoácido dentro da VH e VL de BHA10, $S_i(r)$, para cada posição, i, em uma seqüência individual foi calculada pelo número de vezes que um tipo de resíduo em particular (r = A,C,D...V,W,Y) é observado dentro da série de dados dividida pelo número total de seqüências. A Figura 11 mostra as freqüências de resíduos de V_H e V_L de BHA10 dividida pela freqüência de resíduos dos resíduos de consenso do banco de dados. Dividindo pela freqüência de resíduos de consenso, um corte estringente para a criação de bibliotecas em posições de resíduos onde o aminoácido BHA10 é raramente observado entre seqüências de V_H ou V_L comuns é obtido. Posições de bibliotecas foram determinadas por aquelas cuja freqüência de resíduo dividida pela freqüência de resí-

duo mais freqüente ($S_i(r)/MFR_i(r)$) foi < 0,3. Os resíduos listados à direita dos cálculos de freqüência de resíduo são aqueles encontrados comumente em seqüências humanas conforme publicado (Chothia, *et al., J. Mol. Biol.* 278, 457-479, (1998)) e podem ser especificamente direcionados em um formato de biblioteca para triagem para o aminoácido mais estabilizante. As CDRs das V_H e V_L de BHA10 não foram consideradas para otimização de estabilidade devido à potencial interrupção da interação com LTβR, mas pode ser considerada para projetos de segunda geração.

C. Modelagem por Homologia de Interface VH/VL

Conforme descritas no Exemplo 15 abaixo, análises de calorimetria de varredura diferencial foram realizadas sobre 17 anticorpos humanos. Os candidatos superiores, BIIB1-4 tinham todos excelentes propriedades de estabilidade (isto é, muito elevada e desejada) . Estes anticorpos altamente estáveis podem ser usados como uma plataforma para aumentar a estabilidade de scFvs ou domínios de anticorpos com baixa estabilidade intrínseca. Em particular, foi posta ênfase sobre a interface entre V_H e V_L proporcionando um nível potencialmente maior de propriedades de estabilidade.

Convertendo o Fab de BHA10 para um formato de scFv resultou em uma alteração no termograma por DSC indicativa de um evento de desdobramento totalmente cooperativo, conforme observado por uma única transição de desdobramento de temperatura maior (Fab), para um evento de desdobramento largamente não cooperativo, conforme observado por fuas transições de desdobramento de temperatura menores individuais (scFv). Em outras palavras, no formato Fab, os contatos do interdomínio foram suficientemente fortes para prender todos os domínios juntos em uma única transição de desdobramento de temperatura maior. Assim que os domínios C_H1 e C_L foram removidos, os domínios VH e V_L da scFv não tiveram mais a capacidade de orientar a transição de desdobramento em um evento cooperativo. É postulado que ao estabilizar a interface, pode-se não somente estabilizar ambos os domínios VH e V_L simultaneamente, mas também promover a aderência da interface VH/VL e impedir agregação.

Os 4 anticorpos superiores BIIB mais estáveis (BIIB1-4) foram

usados para projetar mutações estabilizantes na scFv de BHA10 através de um procedimento de duas etapas. Primeiro, uma estrutura de cristal do Fab de BHA10 humanizado foi utilizada para análise estrutural. Especificamente, a estrutura de cristal foi empregada para identificar todos os resíduos na interface entre V_H e V_L bem como a quantidade de área superficial que cada resíduo contribu para a interface (usando o software MOLMOL). Os resíduos que encobrem significativamente mais área superficial do que outros na interface são considerados como sendo mais cruciais para a interface. Como uma priorização de limite arbitrário, aqueles que encobrem entre 30 e 40 $Å^2$ são considerados importantes. Aqueles que encobrem >40 $Å^2$ são considerados cruciais. A estas duas categorias foi dada a maior prioridade em termos de modelagem por homologia e mutagênese. A quantidade de área superficial que cada resíduo encobre na interface é listada na Tabela 6. Em segundo lugar, os tipos de aminoácidos na interface entre V_H e V_L foram comparados com os aminoácidos que existem nas mesmas posições em BIIB1-4 (isto é os anticorpos extremamente estáveis). Este método possibilitou identificação de uma série de mutações importantes para estabilizar a scFv de BHA10.

Duas mutações, S46L na V_L e P101D na V_H foram validadas experimentalmente conforme descrito no Exemplo 8. De fato, estas duas mutações foram as mutações estabilizantes únicas máximas testadas na prova de ensaio térmico (por exemplo, vide a Figura 50A). Ambas as mutações estabilizam ambos os domínios VH e V_L simultaneamente, sugerindo que ajudam a construir cooperatividade entre os domínios.

Tabela 6. Análise da Área Superficial V_HV_L de BHA10.

Posição de Resíduo de Interface VH/VL (nº Kabat)	Região	Tipo de BHA10 AA (Kabat)	Área Superficial Enterrada ($Å^2$)	Freqüência de Resíduo dentro de Banco de Dados de Memífero	AA Correspondentes de Ab Estável (BIIB)
VH 35	FR2	His	16,4		His
VH 37	FR2	Val	6,5		Val

Posição de Resíduo de Interface VH/VL (nº Kabat)	Região	Tipo de BHA10 AA (Kabat)	Área Superficial Enterrada ($Å^2$)	Freqüência de Resíduo dentro de Banco de Dados de Memífero	AA Correspondentes de Ab Estável (BIIB)
VH 39	FR2	Gln	37,4	98%	Gln
VH 44	FR2	Gly	13,1		
VH 45	FR2	Leu	97,8	97%	Leu
VH 47	CDR2	Trp	78,5	97%	Trp
VH 50	CDR2	Trp	24,1		
VH 59	CDR2	Gln	28,8		
VH 61	CDR2	Asn	26,4		
VH 90	FR3	Tyr	37,3		Tyr
VH 100I	CDR3	Glu	101,8	Altamente Variável	Gly (BIIB1), Tyr (BIIB2)
VH 100J	CDR3	Gly	36,4	Ala, Gly (MCR); Tyr (incomum);	Phe (BIIB1), Ala (BIIB2)
VH 100K	CDR3	Phe	66,7	Met (MCR); Phe (SMCR)	Phe (BIIB1); Met (BIIB2)
VH 101	CDR3	Pro	43,3	Asp (V. Cons); Ala (incomum)	Asp* (BIIB1, BIIB2)
VH 103	CDR3	Trp	78,9	98%	Trp
VH 104	FR4	Gly	6,1		
VH 108	Fr4	Gln	11,5		
VL 36	FR2	Tyr	44,2	86%	Tyr (BIIB1, BIIB2)
VL 38	FR2	Gln	34,5	92%	Gln (BIIB1, BIIB2)
VL 42	FR2	Lys			

Posição de Resíduo de Interface VH/VL (nº Kabat)	Região	Tipo de BHA10 AA (Kabat)	Área Superficial Enterrada (Å^2)	Freqüência de Resíduo dentro de Banco de Dados de Memífero	AA Correspondentes de Ab Estável (BIIB)
VL 43	FR2	Ala	47,7	Ser = 50%; Ala = 35%; Pro = 10%	Ala (BIIB1, BIIB2)
VL 44	FR2	Pro	69,7	100%	Pro (BIIB1, BIIB2)
VL 45	FR2	Lys	3,5		
VL 46	FR2	Ser	29,7	Leu = 76%; Arg = 8%; Ser = 2%	Leu* (BIIB1, BIIB2)
VL 49	FR2	Ser	14,8	Tyr = 88%; Glu = 4%; Ser = 2%	Ser (BIIB1)
VL 50	FR2	Ser	9,7		Gly (BIIB1)
VL 55	CDR2	Tyr	59,6	Ala = 34%; Glu = 28%; Tyr = 6%	Glu* (BIIB1)
VL 87	FR3	Phe	35,1	Tyr = 84%; Phe = 15%	Tyr* (BIIB1)
VL 89	CDR3	Gln	14,7		Gln (BIIB1)
VL 91	CDR3	Tyr	36,3		Phe (BIIB1)
VL 94	CDR3	Tyr	68,4	Thr = 20%; Tyr = 12%	Ala (BIIB1); Leu (BIIB2)
VL 95	CDR3	Pro	38,8	Pro = 90%	Pro (BIIB1); Leu (BIIB2)
VL 96	FR4	Phe	76,6	Tyr = 12%; Phe = 6%	Tyr* (BIIB1); Trp* (BIIB2)
VL 98	FR4	Phe	102,5	100%	Phe

Posição de Resíduo de Interface VH/VL (nº Kabat)	Região	Tipo de BHA10 AA (Kabat)	Área Superficial Enterrada ($Å^2$)	Freqüência de Resíduo dentro de Banco de Dados de Memífero	AA Correspondentes de Ab Estável (BIIB)
VL 99	FR4	Gly	5,4		
VL 105	FR4	Glu	6,2		

(Posições de resíduos marcadas com '*' indicam boas oportunidades para design de estabilidade – através de mutação para resíduos encontrados em BIIB1,2,3, ou 4).

D. Análise Computacional

Métodos computacionais foram usados para analisar BHA10 para fazer recomendações sobre posições por meio do qual substituições de aminoácidos podem aumentar a estabilidade. Estes métodos foram compostos de duas etapas: análise à base de seqüência (etapa 1); e análise à base de estrutura (etapa 2). Durante a primeira etapa, foi usado um banco de dados de seqüências de domínios variáveis de anticorpos. Aminoácidos presentes em baixas freqüências em suas posições respectivas (menos do que em 10% das seqüências do banco de dados) ou aqueles que não combinam o aminoácido de consenso correspondente foram selecionados para substituições para aminoácidos de alta freqüência ou de consenso (mutações candidatas). Durante a segunda etapa uma estrutura tridimensional ou modelo de um fragmento Fab do anticorpo foi usado. As mutações candidatas foram avaliadas em seu contexto estrutural e foram priorizadas para prova experimental com base em propriedades estruturais, compatibilidade com regiões determinantes de complementaridade (CDRs) conformações, papel no tamponamento (*packing*) da interface V_L/V_H, e dobramento de cadeias pesadas e leves.

BHA10 é incomum em tendo uma serina na posição 46 sobre a cadeia leve. No banco de dados, aproximadamente 1% das seqüências VL (subtipo kappa) possuem S46, ao passo que aproximadamente 79% das seqüências VL (subtipo kappa) têm L46. Além disso, KV1 de consenso hu-

mano tem L46. Uma vez identificado como uma mutação candidata, a substituição S46L foi avaliada no contexto de um modelo tridimensional de Fab de BHA10. Foi visto que a posição 46 está presente na interface V_L/V_H fazendo contatos com Y101 e W103 de VH e Y36 e Y55 de VL. Análise à base de estrutura revelou que substituição S46L foi compatível com a integridade da interface V_L/V_H, conformações de CDR, e dobramento de V_L.

Serina é um aminoácido infreqüente na posição 16 da cadeia pesada. No banco de dados, aproximadamente 5% das seqüências VH possuem S16, 25% têm A16, 21% têm G16, 11% têm E16, e aproximadamente 10% são Q16. Também HV1 de consenso humana tem A16. Uma vez identificadas como mutações candidatas, as substituições S16A, G, E, e Q foram avaliadas no contexto de modelo tridimensional de Fab de BHA10. Foi visto que a posição 16 está presente na alça proximal à interface V_H/C_H1 fazendo contatos com K13 e S16 de VH. Análise à base de estrutura revelou que as substituições S16A, G, E, e Q foram compatíveis com a integridade da interface V_L/V_H, conformações de CDR, e dobramento de V_H. Também foi concluído que seriam preferenciais cadeias laterais hidrofílicas longas de E16 e Q16. Por conseguinte, as posições V_L 46 e V_H 16 foram representadas na criação da biblioteca.

Uma abordagem complementar a análise computacional fez uso de técnicas de engenharia de proteína à base de estrutura, tais como Rosetta (Kuhlman B. et al., PNAS 200198(19): 10687-91) e DEEK (Hanf K.J., Ph.D. Thesis, MIT, 2002). Dada uma estrutura tridimensional de uma interface protéica, estes métodos podem produzir mutações em seqüência de aminoácidos que levam a aprimoradas $\Delta\Delta G$ para ligação de VH a VL (diferença entre estados ligado e não ligado) e/ou $\Delta\Delta G$ para dobramento de VH e VL (diferença entre estados dobrado e desdobrado de referência). O uso destes métodos revelaram que S46L é uma mutação que aumenta a ligação de VH a VL, ao passo que S16E é uma mutação que aumenta o dobramento da VH.

E. Abordagens de Design de Interface e Covariação Combinadas

Análise de covariação foi usada para determinar redes de resí-

duos importantes para proporcionar e suportar a interface entre VH e VL e mantendo uma forte afinidade entre $\underline{V_H/V_L}$, deste modo resultando em aumentada estabilidade de scFv. Resíduos diretamente envolvidos na interface entre VH e VL da scFv de BHA10 foram identificados conforme descrito no Exemplo 5C e são listados na Tabela 6 acima. Resíduos que covariam fortemente com os resíduos mais altamente enterrados listados na Tabela 6 acima foram calculados usando a metodologia de covariação descrita na Seção III (a) acima. O HMM usado para análise de covariação elimina a capacidade para investigar resíduos em CDR2 e CDR3 de ambos os domínios VH e VL que são enterrados na interface. Não obstante, acredita-se que não surgiriam fortes covariações a partir destas posições de resíduos uma vez que são altamente variáveis devido a intensas pressões seletivas que existem para anticorpos para ser capaz de reconhecer antígenos com alta afinidade. Portanto, os resíduos que foram usados para determinar se existe uma rede de covariação para suportar a interface entre tanto VH e VL foram:

VH (nº kabat)	VL (nº kabat)
V37/I37	Y36
Q39	Q38
G44	A43
L45	P44
W47	L46
W50	Y49
Y91	Y95
W103	F98

Resíduos cujo aparecimento no alinhamento de seqüências classe V se correlacionou com os resíduos da interface listados acima foram determinados usando um corte de valor fi > 0,25. Nenhum resíduo foi considerado tão importante para manter uma forte interface entre VH e VL se não tiver sido correlacionado com no mínimo 2 dos resíduos da interface listados diretamente acima. As Tabelas 7 e 8 listam os resíduos dentro dos domínios VH e VL, respectivamente, que demonstram 2 ou mais correlações com resíduos da interface. Quanto maior o número de links, maior o impacto que se

concebe que cada uma destas posições de aminoácido tenha para estabilizar a interface entre VH e VL.

Tabela 7 – Resíduos com VH com múltiplas correlações fortes para resíduos estruturais observados na interface da estrutura de cristal de raios X BHA10.

nº de Covar	Aminoácido	nº de raios X de Bha10	nº Kabat	nº de Links
67	W	47	47	6
48	V(I**)	37	37	5
61	L	45	45	5
66	*E*	*46*	*46*	*5*
71	I	51	51	5
85	*Y*	*60*	*59*	*5*
142	L	112	109	5
14	P	14	14	4
27	G	26	26	4
29	F	27	27	4
52	*A*	*40*	*40*	*4*
61	L	45	45	4
102	I	70	69	4
135	W	106	103	4
31	F	29	29	3
49	*R*	*38*	*38*	*3*
60	G	44	44	3
69	*G*	*49*	*49*	*3*
78	G	56	55	3
93	L	64	63	3
97	G	55	54	3
101	T	69	68	3
106	D	73	72	3
109	S	75	74	3
115	Y	80	79	3
120	T(S*)	85	82b	3

nº de Covar	Aminoácido	nº de raios X de Bha10	nº Kabat	nº de Links
6	E	8	8	2
43	Y	32	32	2
50	Q	39	39	2
53	*P*	*41*	*41*	*2*
92	S	63	62	2
124	A	89	85	2
131	Y	95	91	2
24	A,V,G	24	24	3*
89	A,N,P	61	60	3*

Resíduos em negrito encobrem área superficial na interface. Resíduos em itálico são aqueles diretamente adjacentes na seqüência primária a resíduos que encobrem superfície na interface.

*Característica de subclasse distintiva

**Covariações com muitos dos mesmos resíduos que o resíduo primário, mas em um menor nível de correlação devido a freqüência de resíduo mais fraca.

Tabela 8. Resíduos com VL com múltiplas fortes correlações para resíduos estruturais observados na interface da estrutura de cristal de raios X BHA10.

nº de Covar	Aminoácido	nº de raios X de Bha10	nº Kabat	nº de Links
49	*Q*	*37*	*37*	*7*
48	Y	36	36	6
61	P	44	44	6
93	P	59	59	6
91	G	57	57	6
102	G	64	64	6
67	L	46	46	4
69	*I*	*48*	*48*	*4*
101	S	63	63	4
107	S	67	67	4
108	G	68	68	4

nº de Covar	Aminoácido	nº de raios X de Bha10	nº Kabat	nº de Links
52	K	39	39	3
66	*K*	*45*	*45*	*3*
85	R	54	54	3
90	S	56	56	3
122	Q	79	79	3
129	D	85	85	3
135	F	98	98	3
68	*L*	*47*	*47*	*2*
92	V	58	58	2
117	T	74	74	2
120	G	77	77	2
126	E	83	83	2
133	*Q*	*89*	*89*	*2*
143	L	104	104	2

Resíduos em negrito encobrem área superficial na interface. Resíduos em itálico são aqueles diretamente adjacentes na seqüência primária a resíduos que encobrem superfície na interface.

*Q50, A89, Y135 são conservados tanto em cadeias pesadas quanto leves e covariações não parecem se correlacionar com posições de cadeia leve.

Os resíduos das Tabelas 7 e 8 foram mapeados para as superfícies das VH e VL de BHA10, respectivamente, e comparados com os resíduos reais que fazem contato direto na interface entre os dois domínios (vide as Figuras 87 e 88). Dois corroborados a partir desta análise. Primeiro, covariações não proporcionam informação (neste estágio) sobre redes envolvendo resíduos CDR2 e CDR3 na interface. Segundo, covariações sugerem que não somente são resíduos que fazem contatos diretos dentro da interface importantes para sua manutenção, ma que muitos outros resíduos fora dos resíduos diretos da interface são importantes para estruturação e suportando as posições dos resíduos da interface. Este ponto é ilustrado pelos resíduos que são ilustrados sobre a superfície da rede de covariação mas estão au-

sentes sobre a superfície obtida simplesmente usando as estruturas para calcular os resíduos que encobrem diretamente superfície na interface.

Foi visto que existem violações de algumas poucas destas redes de covariação de interface dentro das scFvs de BHA10 ou p5E8 (descritas abaixo no Exemplo 21): seqüências VH ou VL. Foi mostrado que mutação dos aminoácidos nativos / não-ideais nestas posições para os resíduos aminoácidos da interface idealmente suportiva descritas nas Tabelas 7 e 8 be é altamente estabilizante ou para as BHA10 ou as p5E8 scFvs: (a) BHA10 VL S46L, e (b) Idec152 VH S49G (nº kabat e de raios X) E72D (73 para raios X). Um destes resíduos ($S46L_{VL}$) estava diretamente na interface, um ($S49G_{VH}$) estava diretamente adjacente à interface, e um estava distal à interface ($E72D_{VH}$). No entanto, todos os três foram previstos com base na Análise de covariação para serem estabilizantes para a interface. A estabilização da interface pode ser observada usando medições por DSC como um aumento na estabilidade térmica tanto da V_H e quanto da V_L.

ii) Construção e Triagem de Bibliotecas de BHA10 scFv com Aprimorada Estabilidade Térmica

A) Construção de Bibliotecas de scFv

Bibliotecas projetadas para conter as substituições de aminoácidos desejadas na BHA10 scFv convencional (pXWU002) usando os métodos descritos nos Exemplos 4, 5, e 6 foram criadas usando o kit de mutagênese sítio-dirigida QuikChange II (QuikChange II Site-Directed Mutagenesis Kit) seguindo as instruções proporcionadas pelo fabricante (Stratagene, La Jolla, Ca.) usando oligonucleotídeos listados na Tabela 9. A seqüência de DNA da scFv convencional é representada na Figura 1.

Colônias transformadas individuais foram selecionadas para lâminas de cavidade profunda de 96 cavidades (Corning, Corning, NY, Cat. nº 3960) contendo 400 µl/cavidade de LB mais 50 mg/ml de carbenicilina e cultivadas de um dia para o outro a 37°C. Lâminas master foram criadas adicionando um volume igual de LB contendo 20% de glicerol para cada cavidade das lâminas de cavidade profunda de 96 cavidades e transferindo alíquotas de 50 µl da suspensão bacteriana para lâminas de microtítulo esté-

reis (Corning, Corning, NY, Cat. nº 3359) e congelando para armazenamento a –80°C.

Tabela 9. Oligonucleotídeos para engenharia de estabilidade de scFv de BHA10.

Posição da Bibliio-teca de BHA10	Seqüência†	Método(s) de Design
V_L 49 e 50 (SEQ ID NO: 123)	5'-GTACCGGTAGGAGGCMNNR KAAATCAGTGATTTAGG - 3'	Consenso
V_L 46 (SEQ ID NO: 124)	5'- GGGAAGGCTCCTAAATTACTG ATTTCCTCGGCC - 3'	Consenso, Computação, Interface V_L/V_H, Covariação
V_H 16 (SEQ ID NO: 125)	5'- GGACACCTTCACTGACBNCCC AGGCTTCTTCAC - 3'	Consenso, Computação
V_H 101 (SEQ ID NO: 126)	5'- GATCCTGGGAAGGTTTTGACT ACTGGGGCCAAGGGAC -3'	Interface V_L/V_H,
V_H 20 (SEQ ID NO: 127)	5'- GGGTCCTCAGTGAAGWTRTCC TGCAAGGCTTCTG -3'	Consenso
V_H 48 (SEQ ID NO: 128)	5'- CAGGGACTTGAGTGGVKKGG A TGGATTTATCCTG -3'	Consenso
V_H 67 (SEQ ID NO: 129)	5'- GAAGTTCAAGGGCAGGNYCA CAATCACTGCAGAC -3'	Consenso
V_H 55 (SEQ ID NO: 130)	5'- GGATGGATTTATCCTGGAAA TGGTCATGCTCAGTACAATGA G -3'	Covariação
V_L 3 (SEQ ID NO: 131)	5'- GGTGGTAGTGACATTVNSATG ACCCAGTCTCCTAGC -3'	Consenso

†Posições alvo para mutagênese são indicadas por sublinhado. Bases am-

bíguas são abreviadas como se segue: W = A ou T, V = A ou C ou G, Y = C ou T, S = C ou G, M = A ou C, N = A ou C ou G ou T, R = A ou G, K = G ou T, B = C ou G ou T (*J Biol Chem.* 261(1): 13-7 (1986)).

B. Ensaio de Estímulo Térmico

As atividades das scFvs BHA10 convencionais e produzidas por engenharia foram então comparadas em uma prova de ensaio térmico a qual pode ser usada para determinar a temperatura na qual 50% das moléculas scFv conservam sua atividade de ligação antigênica depois de um evento de estímulo térmico. O valor numérico correspondente a esta temperatura é referido como o valor de T_{50} e as unidades são em °C. Nesta prova, as scFvs foram submetidas a uma faixa de temperaturas que englobam a temperatura de transição térmica e BHA10 scFv convencional.

E. coli cepa W3110 (ATCC, Manassas, Va. Cat. nº 27325) foi transformada com plasmídeos codificando BHA10 scFvs convencionais e produzidas por engenharia sob o controle de um promotor ara C indutível. Transformantes foram cultivados de um dia para o outro em meio de expressão consistindo em SB (Teknova, Half Moon Bay, Ca. Cat. nº S0140) suplementado com 1% de glicina, 1% de Triton X100, 0,02% de arabinose, e 50 μg/ml de carbenicilina ou em 37°C ou em 32°C. Bactérias foram peletizadas por centrifugação e sobrenadantes foram colhidos para tratamento posterior. Incluindo glicina e Triton X-100 no meio resultou na liberação de conteúdos periplásmicos (proteína de *E. coli* nativa e scFv) no meio. A presença de proteínas de *E. coli* no sobrenadante é essencial para a performance deste ensaio porque as proteínas termicamente desnaturadas agem como um "sorvedouro", capturando moléculas scFv desdobradas transitoriamente em agregados irreversíveis inativos.

Cada biblioteca foi triada em duplicata usando uma prova de ensaio térmico com sobrenadante de uma replicata submetida a condições de tratamento e o segundo sobrenadante servindo como referência não tratada. Provas de estímulo térmico podem ser realizadas em uma única temperatura ou faixa de temperaturas para medir a estabilidade. Máquinas de Thermocycler capazes de gerar gradientes térmicos estáveis são usadas

para tratar sobrenadantes da amostra (iCycler, Bio-Rad, Gaithersburg, Md).

A temperatura de estímulo variou dependendo das propriedades da variante de BHA10 scFv parental e foi geralmente dois a três graus Celsius maior do que o valor da T_{50} determinado experimentalmente. Lâminas Master congeladas foram degeladas e usadas para inocular lâminas de microtítulo de cavidades profundas contendo 250 µl de meio de expressão por cavidade, e culturas cultivadas de um dia para o outro a 32°C. Como um controle, culturas contendo o plasmídeo parental foram cultivadas sob as mesmas condições e processadas simultaneamente à biblioteca. Bactérias foram peletizadas nas lâminas de cavidades profundas por centrifugação a 2000 rpm (IEC modelo 8R, Thermo Electron, Waltham, Ma) por 30 minutos, e 100 µl do sobrenadante foi transferido para lâminas de microtítulo de rotina (Corning, Corning, NY, Cat. nº 3357). Uma alíquota do sobrenadante foi reservada para o DELFIA de referência. Para a maioria das bibliotecas, as lâminas contendo o restante dos sobrenadantes foram seladas (Nalge Nunc, Rochester, NY, Cat nº 235205) e colocadas em uma incubadora ajustada para a temperatura de estímulo apropriada (Echo Therm, Torrey Pines Scientific, San Marcos, Ca) por 90 minutos. Para triagens requerendo múltiplas temperaturas de estímulo (ou por temperaturas de mais de 75°C) os sobrenadantes foram transferidos para lâminas de PCR de 96 cavidades (Applied Biosystems, Foster City, Ca, Cat. nº N801-560) e incubadas por 90 minutos na temperatura desejada.

Depois de estímulo térmico, o material agregado foi removido por centrifugação e amostras de BHA10 scFv solúveis remanescentes no sobrenadante clareado tratado foram avaliadas para ligação a antígeno LTβR Ig cognato por tete DELFIA. Lâminas de 96 cavidades (MaxiSorp, Nalge Nunc, Rochester, NY, Cat. nº 437111) foram revestidas com proteína de fusão consistindo no ectodomínio do LTβ receptor (LTβR) fundido a uma região Fc humana a 1 µg/ml em 0,1M de tampão de carbonato de sódio, pH 9.5. As lâminas foram revestidas de um dia para o outro a 4°C, e bloqueadas com tampão de ensaio DELFIA (DAB, 10 mM de Tris HCl, 150 mM de NaCl, 20 µM de EDTA, 0,5 % de BSA, 0,02% de Tween 20, 0,01% de NaN_3, pH

7.4) por uma hora com agitação em temperatura ambiente. As lâminas foram lavadas 3 vezes com DAB sem BSA (tampão de lavagem), e amostras de ensaio diluídas em DAB foram adicionadas às lâminas em um volume final de 100 µl. As lâminas foram incubadas por uma hora com agitação em temperatura ambiente, e em seguida lavadas 3 vezes com tampão de lavagem para remover moléculas scFv não ligadas e funcionalmente inativadas. BHA10 scFv ligado foi detectado por adição de 100 µl por cavidade de DAB contendo 40 ng/ml de anticorpo anti-His_6 marcado com Eu(Perkin Elmer, Boston, MA, Cat. nº AD0109) e incubado em temperatura ambiente com agitação por uma hora. As lâminas foram lavadas 3 vezes com tampão de lavagem, e 100 µl de solução de reforço DELFIA (Perkin Elmer, Boston, MA, Cat. nº 4001-0010) foi adicionado por cavidade. Depois de incubação por 15 minutos, as lâminas foram lidas usando o método de Európio em um Victor 2 (Perkin Elmer, Boston, MA).

Dados de ensaio foram processados usando o software Spotfire DecisionSite (Spotfire, Somerville, Ma.) e expressados como a proporção das contagens DELFIA observdas em temperatura de estímulo para a temperatura de referência para cada clone. Clones que deram reprodutivelmente proporções maiores do que ou iguais a duas vezes a que foi observada para o plasmídeo parental foram considerados alcences. DNAs de plasmídeos destes clones positivos foram isolados por mini-prep (Wizard Plus, Promega, Madison, WI) e retransformados de volta em *E. coli* W3110 para provas de ensaio térmico secundário de confirmação.

Para gradientes térmicos, os dados foram analisados usando o software Prism 4 (GraphPad Software, San Diego, Ca) usando uma reação de dose sigmoidal com inclinação variável como o modelo. Os valores obtidos para o ponto central das curvas de desnaturação térmica são referidos como valores T_{50}, e não são considerados como sendo equivalentes a valores de Tm biofisicamente derivados.

Resultados primários e confirmatórios destes ensaios são mostrados na Tabela 10. Várias das moléculas scFv estabilizadas da invenção resultaram em aprimoramentos na atividade de ligação (T_{50} > 49°C) compa-

radas com o scFv convencional. Em particular, os valores de T_{50} de BHA10 posição da biblioteca V_L46 scFv (S46L), posição da biblioteca V_H16 scFv (S16E e S16Q), e posições da biblioteca V_L49: V_L50 scFv apresentaram aumentos na estabilidade térmica variando de + 3°C a + 12°C em relação ao BHA10 scFv convencional. Além disso, os valores de T_{50} de BHA10 posição da biblioteca V_L3 scFv (Q3A, Q3G, Q3S, Q3V, e Q3D), posição da biblioteca V_H67 scFv (V67I e V67L), posição da biblioteca V_H48 scFv (M48I e M48G), posição da biblioteca V_H20 scFv (V20I) e posição da biblioteca V_H101 scFv (P101D) apresentaram aumentos na estabilidade térmica variando de + 4°C a + 18°C em relação ao BHA10 scFv convencional. Uma das mutações estabilizantes (V_L K13E) acidentalmente resultou de um erro de PCR. Foi visto que a incorporação de uma destas mutações estabilizantes em pIEH009 aumenta adicionalmente a estabilidade térmica e mesmo excede a de BHA10 Fab sob estas condições (Figura 12). De modo importante, a mutação V_L46 não-covalente e mutações V_H55 derivadas de um ou mais dos quatro métodos de design (modelagem por homologia da interface V_L/V_H, classificação de consenso, modelagem computacional, e análise de covariação) resultaram em um aprimoramento na estabilidade térmica de scFv quase se aproximando da observada com as mutações de dissulfeto e validando a utilidade e inovação destas ferramentas de criação.

A Figura 1B mostra as seqüências de aminoácidos de BHA10 scFv convencional (SEQ ID NO: 4) e as Figuras 94A e B mostram a seqüência de aminoácidos de BHA10 scFVs estabilizados contendo a mutação estabilizante S46L(VL) (SEQ ID NO: 137), e a mutação estabilizante V55G (VH) (SEQ ID NO: 138), respectivamente. A seqüência de DNA de BHA10 scFv selvagem (SEQ ID NO: 3) é representada na Figura 1A. A mutação estabilizante é indicada pelo resíduo envolvido por caixa. A seqüência lider, peptídeo de ligação gly/ser, e domínio CH1 são indicadas pelos resíduos sublinhados, em negrito, e em itálico, respectivamente.

Tabela 10. Posições da biblioteca de BHA10 VH e VL, composição da biblioteca, e resultados de triagem.

Posição	Biblioteca	Alcence Seq. Observada	□T_{50} °C
VL49 e 50	49 (Y ou S) 50 (all AA's)	YT, YS, YR, YG, SR, SK	+3-4
VH16	Q, K, E, R, W, G, P, S, T, A	S16E S16Q	+8 +4
VL46	L	S46L	+10
VH13	na	K13E	+3
VH101	na	P101D	+18
VH20	F, M, L, I	V20I	+7
VH48	R, S, M, I, L, V, G	M48I M48G	+3-4
VL3	N, Q, K, H, E, D, R, W, G, P, S, T, A	Q3A, Q3S, Q3V, Q3D, Q3G,	+3-5
VH55	na	VH V55G	+12
VH67	T, P, I, L, V, A	V67I V67L	+5 +8

na = não aplicável

A Tabela 11 mostra os resultados de uma análise de estabilidade térmica completa das várias mutações estabilizantes individuais e combinadas introduzidas em um scFv convencional. Estes resultados demonstram que os aprimoramentos na atividade são aditivos e que os métodos descritos nesta invenção são capazes de melhorar as propriedades de estabilidade térmica de scFvs mesmo além das de Fabs nativos. Mesmo na ausência de uma ligação dissulfeto covalente nas posições V_H44-V_L100, foram identificadas mutações estabilizantes que na combinação apresentaram aumentos na estabilidade térmica variando de + 19°C a + 33°C em relação ao BHA10 scFv convencional.

Tabela 11. Características de construtos BHA10 usados para produzir proteínas variantes e resultados de T_{50} de prova de ensaio térmico.

Plasmídeo	Dissulfeto	Extensão do Ligante (aa)	Outra Mutação	T_{50} °C
pIEH003	no	15	na	49
pIEH006	VH44-VL100	15	na	59
pXWU001	Fab	na	Fab	61
pXWU002	no	20	na	51
pXWU003	no	25	na	51
pIEH009	VH44-VL100	20	na	61
pIEH029	no	20	VL S46L	61
pIEH031	VH44-VL100	20	VL S46L	71
pIEH032	no	20	VH S16E	60
pIEH034	no	20	VH S16Q	56
pIEH058	no	20	VH P101D	67
pIEH059	no	20	VH V20I	56
pIEH060	no	20	VH M48G	54
pIEH061	no	20	VH M48I	53
pIEH062	no	20	VL Q3A	53
pIEH063	no	20	VL Q3S	53
pIEH065	no	20	VL Q3V	53
pIEH066	no	20	VL Q3D	54
pIEH067	no	20	VL Q3G	54
pIEH068	no	20	VH V67I	55
pIEH069	no	20	VH V67L	58
pIEH070	no	20	VH V55G	64
pIEH076	no	20	VH S16E + VL S46L	71
pIEH078	no	20	VH S16Q + VL S46L	68
pIEH080	VH44-VL100	20	VH S16E + VL S46L	75

Plasmídeo	Dissulfeto	Extensão do Ligante (aa)	Outra Mutação	T_{50} °C
pIEH081	VH44-VL100	20	VH S16Q + VL S46L	72
pIEH087	no	20	VH S16E, V55G + VL S46L	75
pIEH094	no	20	VH S16E, V55G, P101D + VL S46L	82

na = não aplicável

aa = aminoácidos

Sumário

Três mutações estabilizantes (V_L_S46L, V_H_V55G, e V_H_P101D) foram experimentalmente validadas por prova de estímulo térmico (T_{50}).

Tanto as mutações V_L_S46L como V_H_V55G foram previstas para serem estabilizantes para os domínios individuais VH e VL com base nas análises de covariação descritas no Exemplo 3 acima. Conforme representado na Figura 50A e B, ambas estas mutações levam a aumentos significativos na T_{50} do BHA10 scFv. Em particular, a mutação VL_S46L estabiliza o scFv por ~7-8 °C ao passo que a mutação VH_V55G estabiliza o scFv por ~12 °C. Ambas estas mutações também encolhem significativamente o intervalo da T_M entre V_H e V_L conforme determinado por DSC (vide a Figura 51). Os dados de DSC mostram que ambas as mutações levam a aumentos significativos tanto na T_M (ponto central das transições de desdobramento térmico) de ambos os domínios VH e VL e na entalpia calorimétrica (isto é, área sob a curva) para o scFv. Aumentos nestes valores proporcionam validação experimental de que estas mutações estabilizantes previstas são de fato estabilizantes para o scFv. Além disso, como estas duas mutações estabilizam ambos os domínios VH e V_L simultaneamente, provavelmente são importante para construir cooperatividade entre os domínios VH e V_L . Por conseguinte, espera-se que a intensificação da interface entre V_H e V_L possa não somente ajudar a aumentar a estabilidade, mas reduzir a tendência de formar agregados também.

A mutação estabilizante V_L_S46L e uma terceira mutação esta-

bilizante (VH_P101D) também foram previstas para serem estabilizantes com base em modelagem por homologia de interface VH/VL descrita no Exemplo 4 acima. A mutação estabilizante VH_P101D estabilizou o scFv por ~15 °C conforme medido por prova de ensaio térmico (Figura 50C). Além disso, DSC (vide a Figura 52) mostra que esta mutação estabilizante leva a aumentos significativos na T_M tanto dos domínios VH e VL e na entalpia calorimétrica, indicando que é provável que esta mutação também estabilize cooperativamente o scFv como um todo através da interface VH/VL.

Exemplo 6. Caracterização Biofísica de BHA10 scFvs Estabilizados Compreendendo Mutações Estabilizantes

Seqüências genéticas de região V de vários plasmídeos listados na Tabela 11 foram subclonadas em um vetor de expressão de *E. coli* modificado para dirigir expressão de proteína recombinante sob o controle de um promotor ara C indutível. BHA10 scFvs variantes foram expressados e purificados usando métodos descritos acima. Um sítio de clivagem aparente no N-término do domínio V_L levou a um baixo nível de VL na maioria das preparações de scFv conforme avaliado por análise SDS-Page (Figura 53).

As propriedades hidrodinâmicas de cada scFv foram investigadas por HPLC de exclusão de tamanho (Agilent Technologies) com detectores de dispersão luminosa estática e do índice refrativo (MiniDAWN/ReX, Wyatt Technology). Foi visto que cada scFv era predominantemente monomérico; no entanto, uns poucos dos scFvs, incluindo o scFv (não estabilizado) selvagem, apresentou um baixo nível de material dimérico (Tabela 12). Nenhum dos scFvs purificados teve níveis detectáveis de oligômeros maiores do que dímero.

Tabela 12. Resultados de SEC / Dispersão Luminosa de scFvs

Amostra	Rendimento Expressado (mg/L)[a]	% de Monômero	% de Dímero	% de Agregados Maiores
WT	n.d.[b]	87	13	0
GS4	0,6 ± 0,1	97	3	0
009 (ss)	n.d.	99	1[c]	0
072 V_L_S46L	2,0	85	15	0

Amostra	Rendimento Expressado (mg/L)[a]	% de Monômero	% de Dímero	% de Agregados Maiores
073 V_L_S46L + ss	*0,15	78	22	0
074 V_H_S16E	2,4 ±,1,7	94	6	0
075 V_H_S16Q	1,1 ± 0,5	96	4	0
076 V_L_S46L + V_H_S16E	1,8 ± 0,4	80	20	0
077 V_H_P101D	2,5	93	7	0
078 V_L_S46L + V_H_S16Q	1,0	85	15	0
079 V_H_V55G	2,0	92	8	0
080 V_L_S46L + V_H_S16E + ss	*0,21	76	24	0
087 V_L_S46L + V_H_S16E, V55G	2,8	76	24	0
094 V_L_S46L + V_H_S16E,V55G,P101D	2,5	79	21	0

[a]Proteínas sem erro padrão foram expressadas uma vez.

[b]n.d., não determinado.

[c]Estes valores foram determinados *DEPOIS* de SEC do preparativo a ser realizada para reduzir agregados

A meta final do experimento foi investigar se as mutações designadas reforçam a estabilidade e levam a uma menor propensão para agregação. A termoestabilidade de cada scFv foi avaliada usando dois métodos separados. Primeiro, foi analisado o desdobramento térmico dos scFvs usando DSC (Figura 54A). Em segundo lugar, a termoestabilidade de cada scFv foi investigada aquecendo as proteínas na presença de um corante hidrofóbico fluorescente, 1-anilino-8-sulfonato (ANS, Figura 54B). ANS geralmente liga a proteínas parcialmente desdobradas ou proteínas desnaturadas por calor em estados desdobrados compactos. Na associação com áreas superficiais hidrofóbicas expostas, a fluorescência de ANS aumenta significativamente presumivelmente devido a seqüestração dos efeitos de extinção do solvente.

As curvas de ANS dependente da temperatura foram ajustadas para uma transição de desdobramento protéico de dois estados conforme descrito abaixo. O tampão foi PBS e a concentração de scFv usado em cada

experimento foi 750 nM. Foram realizadas medições da fluorecência sobre um espectropolarímetro de Dicroismo Circular JASCO modelo 812 equipado com um dispositivo de aquecimento de Peltier e banho de água externo. A fluorecência foi compilada usando um acessório contendo um tubo fotomultiplier perpendicular à trajetória da luz. O acessório é equipado com um monocrômetro ajustável ajustado para 480 nm. A sensibilidade foi ajustada para 600 V. Foi realizado aquecimento em uma taxa contínua de 120 °C/min. O monocrômetro da excitação foi ajustado para 370 nm.

Desdobramento térmico de cada scFv neste estudo levou a um aumento na fluorescência de ANS. O ponto central do desdobramento térmico (isto é, a T_M) de cada domínio protéico scFv foi determinado usando DSC ajustando cada pico de desdobramento à equação de Gibbs-Helmholtz usando o software Origin 7.0 fornecido pelo fabricante (MicroCal, Inc). T_Ms também foram derivadas usando fluorescência de ANS por incorporação da equação de Gibbs-Helmholtz na rotina de ajuste da curva não-linear em KaleighdaGraph™:

$$\Delta G_U°(T) = \Delta H_U°(T) - T\Delta S_U°(T), \qquad (1)$$

onde $\Delta G_U°(T)$ é a alteração dependente da temperatura na energia livre de Gibbs no desdobramento, $\Delta H_U°(T)$ é a alteração da entalpia associada com desdobramento, e $\Delta S_U°(T)$ é a entropia associada com desdobrado. A equação pode ser expandida para:

$$\Delta G_U°(T) = \Delta H_U°(1-\frac{T}{T_M}) - \Delta C_P°\left[(T_M - T) + T\ln\left(\frac{T}{T_M}\right)\right], \qquad (2)$$

onde T_M é o ponto central da curva de desdobramento e $\Delta C_P°$ é a alteração na capacidade térmica entre os estados dobrado e desdobrado. Esta equação pode ser usada para derivar a dependência da temperatura da intensidade de fluorescência de ANS presumindo que a intensidade de fluorescência é a soma da fluorescência de ANS intrínseca na presença de scFv dobrado e desdobrado,

$$i_T(T) = i_F f_F + i_U f_U \ , \qquad (3)$$

onde $i_T(T)$ é o sinal de fluorescência total dependente da temperatura, i_F e i_U são as intensidades de fluorescência dos estados dobrado e desdobrado,

respectivamente, e f_F e f_U são as frações de scFv dobrado e desdobrado, respectivamente na temperatura dada. As frações dobradas estão relacionadas com a constante de desdobramento de equilíbrio e energia de desdobramento livre:

$$K_U(T) = \frac{f_U}{f_F} = \exp(\frac{-\Delta G_U°(T)}{RT}). \qquad (4)$$

Fatorando $K_U(T)$ e $\Delta G_U°(T)$ na equação (3) e presumindo que as intensidades de fluorescência de ANS na presença de scFv dobrado e desdobrado são linearmente dependente da temperatura (isto é $i_F = i_1 + i_2 T$ e $i_U = i_3 + i_4 T$) foi obtida a seguinte equação:

$$i_T(T) = \frac{i_F + i_U K_U}{1 + K_U} = \frac{i_1 + i_2(T) + (i_3 + i_4 T)\exp\left[-\frac{\Delta G_U°(T)}{RT}\right]}{1 + \exp\left[-\frac{\Delta G_U°(T)}{RT}\right]}. \qquad (5)$$

Para obter a equação final utilizada para ajustar os dados, a relação para $\Delta G_U°(T)$ na equação (2) é substituída na equação (5) assumindo que $\Delta C_P°$ é independente da temperatura e proporcional à diferença na área superficial exposta a solvente entre os estados dobrado e desdobrado (Haynie & Friere, *Proteins: Struct. Funct. Genet.* 16: 115-140, (1993); Myers *et al., Protein Sci.* 4: 2138-2148 (1995)).

As medições da termoestabilidade para todos os scFvs derivados de experimentos de DSC e de ligação de ANS foram comparáveis. O desdobramento de cda scFv foi irreversível; portanto, a agregação teve um efeito sobre a T_M absoluta medida por DSC ou ligação de ANS. Como as duas técnicas aqueceram as amostras diferentemente e em taxas diferentes, não se espera que as T_Ms sejam idênticas nos formatos experimentais (Sánchez-Ruiz *et al., Biochemistry,* 27:1648-1652 (1988)). No entanto, a tendência observada para cada técnica experimental foi idêntica (vide Tabela 13). Os experimentos de DSC prontamente discriminaram entre as transições dos desdobramentos de V_H e V_L. Os experimentos de ligação de ANS não foram capazes de discriminar precisamente 2 transições (isto é desdobramento V_H vs. V_L); portanto, somente a T_M *aparente* foi proporcionada para os experimentos de ligação de ANS. A T_M *aparente* observada por ligação de

ANS pareceu se correlacionar bem com a T_M do *último* domínio para desdobrar, quer V_H ou V_L dependendo da mutação, conforme determinado por DSC. Devido a agregação de material desdobrado em ambos os formatos de ensaio, as T_Ms foram usadas como um guia para classificar em ordem os reforços de estabilidade proporcionados por cada mutação sem interpretação adicional de energias livres de desdobramento de scFv, e etc.

Tabela 13. Medições da termoestabilidade de cada scFv.

Construto	V_H T_M (ºC DSC)	V_L T_M (ºC DSC)	V_H T_M (ºC ANS)	scFv T_M (ºC DSC)[a]	scFv K_D (M) $*10^{-9}$ (Biacore)
Tipo Selvagem, $(G_4S)*3$ encadeador[b]	55,4	68,7	65,6	-	2,7
Tipo Selvagem, $(G_4S)*4$ ligante	57,7	67,4	67,0	-	2,1
V_H_S16E	60,7	68,1	66, 76	-	2,0
V_H_S16Q	59,4	68,4	69[c]	-	1,9
V_L_S46L	65,6	74,2	71,0	-	4,0
V_H_V55G	-	-	75,4	68,4	2,2
V_H_P101D	-	-	-	71,9	5,2
V_H_S16E, V_L_S46L	71[d]	74,5[d]	-	72,2	3,4
V_H_S16Q, V_L_S46L	67,9[d]	74,9[d]	75,4	-	4,2
V_H_S16E, V_H_V55G, V_L_S46L	-	-	79,3	77,7	3,8
V_H_S16E, V_H_V55G, V_H_P101D, V_L_S46L	84,1	77	81,3	-	n.d.
Selvagem, $(G_4S)*4$ ligante w/ V_H44/V_L100 dissulfeto	61	68	n.d.	-	3,3
V_H_S16E, V_L_S46L	69	77	n.d.	-	4,0
V_H_S16Q, V_L_S46L	n.d.	n.d.	n.d.	n.d.	3,3

[a]Estas mutações particulares levam a eventos de desdobramento V_H/V_L único, cooperativo.

[b]Todos os scFvs subseqüentes foram construídos com o $(G_4S)*4$ ligante de

20 aminoácidos.

[c]Análise complicada de múltiplas transições.

[d]Incapaz de discriminar V_H vs. V_L.

Sumário

Todas as mutações únicas projetadas selecionadas das triagens de biblioteca: V_H S16E, S16Q, V55G, P101D, e V_L S46L estabilizaram significativamente o domínio V_H e em alguns casos, o domínio V_L também (Figura 53). O fundamento lógico por trás da experimentação destas posições para reforços da estabilidade freqüentemente vem de múltiplas formas de análise. Métodos de consenso previram que todas as mutações V_H S16E, S16Q, V55G, P101D e V_L S46L estabilizariam o BHA10 scFv (Exemplo 3). No entanto, V_H V55G e P101D (incidentalmente, as duas mutações estabilizantes principais) estão posicionadas dentro de CDR2 e CDR3 de VH, respectivamente, e não foram consideradas para mutagênese até ser compilada evidência preditiva adicional que sugeriu que mutação nestas duas posições pode levar a eventos estabilizantes. Tanto V_H V55G quanto V_L S46L também foram previstos para serem estabilizantes com base em Análises de Covariação. Finalmente, V_H P101D e V_L S46L foram previstos para serem potencialmente estabilizantes para a interface V_H/V_L com base na composição da interface de anticorpos humanos altamente estáveis (Exemplo 3).

Mutações únicas na interface V_H/V_L aumentaram a estabilidade de ambos os domínios, ao passo que mutações V_H fora da interface somente estabilizaram o próprio domínio V_H. V_H S16E, S16Q, e V55G estavam fora da interface e aumentaram a T_M aparente do BHA10 V_H por 3, 2, e 11 ºC, respectivamente. A estabilidade de V_L foi relativamente não afetada por esas mutações. A mutação V_H V55G pareceu aumentar a estabilidade da V_H para combinar com a do domínio V_L, sem levar a aumentos detectáveis na estabilidade de V_L. Ambas as mutações na interface, V_H P101D e V_L S46L aumentaram significativamente a estabilidade de ambos os domínios. A mutação P101D em particular leva a desdobramento totalmente cooperativo dos domínios VH e V_L sugerindo que a interface e a cooperatividade de dobramento entre V_H e V_L foi significativamente reforçada. Portanto, mutações na inter-

face V_H/V_L podem proporcionar os meios mais eficazes para formar uma região Fv estabilizada e proporcionar um fundamento lógico para priorizar designs de estabilidade para a interface V_H/V_L de scFvs sobre designs dirigidos para resíduos fora da interface Fv.

Combinando mutações de V_H eventualmente leva a uma desassociação entre a cooperatividade do desdobramento VH/V_L (Tabela 13). À medida que a T_M de V_H subiu acima de 75 ºC (o resultado de combinações de mutações de VH), a T_M da V_L começa a se mover para temperaturas menores. Iso sugere hiperestabilização do domínio V_H sem formar mutações de compensação dentro do domínio V_L podem levar a reduzida cooperatividade de dobramento e uma interação enfraquecida entre os dois domínios. Apesar das T_Ms dos domínios VH e V_L (incluindo a T_M do BHA10 scFv selvagem) serem freqüentemente muito diferentes, evidência de que mutações na interface entre os dois domínios podem estabilizar ambos os domínios simultaneamente sugere que os dois domínios estão interagindo em algum nível um com o outro – embora não se espere que a aparente afinidade entre os domínios isolados seja extremamente elevada, exceto talvez na variante estabilizada V_H P101D (Brandts & Lin, *Biochemistry,* 29: 6927-6940, (1990)).

Aumentando a estabilidade da scFv levou a reduzida ligação intrínseca do corante fluorescente hidrofóbico ANS em temperatura ambiente (15 ºC, Figura 55). O scFv selvagem liga fracamente a ANS sugerindo que a proteína pode expor permanentemente ou transitoriamente área superficial hidrofóbica a solvente. scFvs estabilizados pareceram perder completamente a capacidade de ligar ANS sob condições ambiente – a fluorescência de ANS na presença das scFvs mais estabilizadas não foi de mais do que a de ANS somente no solvente. Gráficos de VH_T_M (medida por DSC ou experimentos de ligação de ANS dependente da temperatura) vs. ANS-fluorescência na presença de proteína dobrada a 15 ºC demonstram que fluorescência de ANS scFv-dependente diminui à medida que a estabilidade de scFv aumenta (Figura 55). A redução de ligação de ANS por estabilização de scFv pode indicar que menos área superficial hidrofóbica fica exposta a solvente e que scFvs estabilizadas podem ter uma menor propensão in-

trínseca para agregação.

Medições de K_D de várias mutações BHA10 scFv individuais e combinadas estabilizantes foram realizadas usando ressonância de plasmon superficial (SPR) em um instrumento Biacore 3000. A Tabela 13 mostra os resultados da prova de afinidade Biacore. A afinidade de ligação dos BHA10 scFvs variantes variou de 1X a 2,5X do BHA10 scFv original.

Exemplo 7. Produção de Anticorpos "Hercules" Biespecíficos Estabilizados

Tanto BHA10 scFv convencional quanto os BHA10 scFvs estabilizados da invenção foram usados para construir um painel de anticorpos biespecíficos referidos como "Hercules". Anticorpos biespecíficos Hercules compreendem uma fusão de anticorpo quimérico 14A2 IgG que liga a TRAIL R2 receptor com um BHA10 scFv que liga a LTβR. Os anticorpos Hercules foram construídos tanto como fusões de BHA10 scFv de N- terminal e C-terminal (vide a Figura 13). Fusões de N- terminal scFv podem ser produzidas por engenharia como fusões de cadeia leve e/ou cadeia pesada ("N_L-Hercules" ou "N_H-Hercules", respectivamente). A decisão quanto a qual região amino terminal V (pesada ou leve) é selecionada para unir o scFv é essencialmente orientada por qual cadeia se considera que tolera um scFv fundido o qual é capaz de reconhecer o primeiro antígeno alvo ao mesmo tempo que não interferindo de modo apreciável com a ligação do domínio Fab ao segundo antígeno alvo.

É também possível usar os métodos descritos nesta invenção para produzir anticorpos tetravalentes ou biespecíficos consistindo somente de scFvs estabilizados fundidos diretamente a regiões de articulação de anticorpo ou a domínios CH2 ou CH3 conforme representado, por exemplo, na Figuras 42 a 49. Os anticorpos referidos podem compreender regiões Fc de extensão total (vide, por exemplo, a Figura 46) ou regiões Fc com domínio CH2 eliminado (vide, por exemplo, a Figura 42). Em outras modalidades exemplares, dois ou mais domínios scFv estabilizados podem ser fundidos ao mesmo término de uma cadeia pesada ou leve (vide, por exemplo, a Figura 43).

A. Construção de "N_H- Hercules" com BHA10 convencional, $(G_4S)_4$, V_H44:V_L100, e V_H44:V_L100/$(Gly_4Ser)_4$ scFvs

Quatro projetos de anticorpos biespecíficos anti-LTβR (BHA10) x anti-TRAIL R2 (chi14A2) foram baseados em anexar os BHA10 scFvs convencional e variante ao término amino da cadeia pesada de anticorpo anti-TRAIL R2. Os DNAs de BHA10 scFvs descritos no Exemplo 3 foram usados para construir um painel de anticorpos biespecíficos N_H- Hercules por amplificação por PCR usando os iniciadores de oligonucleotídeos descritos na Tabela 14. Um $(Gly_4Ser)_5$ ligante foi usado para conectar os BHA10 scFvs ao término amino maduro de cadeia pesada chi14A2. O iniciador por PCR 5' VH forward (scFvBHA10-F1) inclui um sítio de restrição de endonuclease *Mlu* I (seqüência sublinhada) para clonagem seguida por seqüência codificando os últimos três aminoácidos do peptídeo de sinal de cadeia pesada e o término amino de BHA10 VH. Dois iniciadores por PCR 3' reversos foram usados para gerar os produtos de PCR com iniciador reverso interno XWU005-R codificando o término carboxila de BHA10 VL seguido pelo $(Gly_4Ser)_5$ ligante, e iniciador reverso scFvBHA10-R1 codificando uma região anti-TRAIL R2 VH parcial e um sítio *Bgl* II (seqüência sublinhada) para clonagem.

Tabela 14. Oligonucleotídeos para amplificação por PCR de N- Hercules com BHA10 scFvs convencionais, $(Gly_4Ser)_4$, V_H44:V_L100 e V_H44:V_L100/$(Gly_4Ser)_4$

scFvBHA10-F1 (SEQ ID NO: 134)	5'-AGAGAGACGCGTGTCCTGTCCCAGGTCCAACTGGTGCAG-3'
XWU005-R (SEQ ID NO: 135)	5'-AGCCACCTCCCCCCGATCCACCGCCCCCTGAACCGCCCCCTCCAGAGCCCCCTCCACCGGACCCTCCACCGCCTTTGATCTCCACCTTG -3'
scFvBHA10-R1 (SEQ ID NO: 136)	5'-AGAGAGAGATCTTGACTGTCTCTCCAGGCTTCTTCAGCTCAGGTCCAGACTGCACCAACTGGATCTGGGAGCCACCTCCCCCCGATCCAC-3'

As seqüências genéticas BHA10 scFv + $(Gly_4Ser)_5$ ligante + anti-TRAIL R2 VH parcial foram amplificadas em duas reações por PCR seqüenciais através das seqüências de sobreposição comuns codificando o $(Gly_4Ser)_5$ ligante conforme representado na Figura 14A para preparar BHA10 scFv convencional a partir do DNA de plasmídeo pXWU034, BHA10 scFv $(Gly_4Ser)_4$ do plasmídeo pXWU002, BHA10 scFv $V_H44:V_L100$ do plasmídeo pIEH006, e BHA10 scFv $V_H44:V_L100/(Gly_4Ser)_4$ do plasmídeo pIEH009. Os produtos de PCR do painel de BHA10 scFvs amplificados foram purificador por eletroforese por gel agarose usando o kit de extração Millipore Ultrafree-DA de acordo com instruções do fabricante (Millipore; Bedford, MA). Os produtos de PCR purificados foram digeridos com endonucleases de restrição *Mlu* I/*Bgl* II e ligados em um vetor pN5KG1 digerido por *Mlu* I/*Bgl* II contendo a chi14A2 IgG1 previamente modificada para remover um sítio *Bgl* II interno presente na seqüência codificante chi14A2 IgG1. O vetor de expressão mamífero pN5KG1 contém um gene de neomicina fosfotransferase modificado (contendo íntron), de translação prejudicada, para selecionar eventos de integração transcricionalmente ativos, e um gene de diidrofolato redutase murino para permitir amplificação com metotrexato (Barnett, *et al.,* Antibody Expression e Engineering. (Imanaka, H. Y. W. a. T., ed), pp. 27-40, Oxford University Press, New York, NY, (1995)).

O painel resultante de construtos formam proteínas de fusão dos BHA10 scFvs variantes ao término amino do domínio VH de anticorpo R2 anti-TRAIL através do $(Gly_4Ser)_5$ ligante de 25 aminoácidos. Fusão da seqüência genética de BHA10 scFv convencional ao término amino da seqüência genética de cadeia pesada do anticorpo anti-TRAIL R2 produziu o plasmídeo pXWU005. Fusão da seqüência genética de BHA10 scFv $(Gly_4Ser)_4$ ao término amino da seqüência genética de cadeia pesada do anticorpo anti-TRAIL R2 produziu o plasmídeo pXWU026. Fusão da seqüência genética de BHA10 scFv $V_H44:V_L100$ ao término amino da seqüência genética de cadeia pesada do anticorpo anti-TRAIL R2 produziu o plasmídeo pXWU027. Fusão da seqüência genética de BHA10 scFv $V_H44:V_L100/(Gly_4Ser)_4$ ao término amino da seqüência genética de cadeia pesada do anticorpo anti-TRAIL R2 produ-

ziu o plasmídeo pXWU028. As misturas de ligação foram usadas para transformar células competentes de *E. coli* cepa TOP 10 (Invitrogen Corporation, Carlsbad, CA). Colônias de *E. coli* transformadas para resistência a fármaco de ampicillna foram triadas para presença de insertos. Análise de seqüência de DNA confirmou a seqüência correta dos construtos finais.

A cadeia leve quimérica 14A2 usada é comum entre todos os anticorpos biespecíficos N- e C-Hercules e as seqüências de DNA (SEQ ID NO: 28) e de aminoácidos (SEQ ID NO: 29) são mostradas nas Figuras 15A e 15B. A cadeia leve quimérica 14A2 é expressada com um peptídeo de sinal no N-término tendo as seguintes seqüências de DNA e de aminoácidos:

ATGGCCTGGACTCCTCTCTTCTTCTTTGTTCTTCATTGCTCAGGGT CTTTCTCC (SEQ ID NO: 59)

MAWTPLFFFFVLHCSGSFS (SEQ ID NO: 60)

As seqüências de DNA (SEQ ID NO: 30)_e de aminoácidos (SEQ ID NO: 31) de cadeia pesada para BHA10 scFv N_H- Hercules convencionais são mostradas nas Figuras 16 e 17, respectivamente. A seqüência de DNA (SEQ ID NO: 32) e de aminoácido (SEQ ID NO: 33) de cadeia pesada para BHA10 scFv $(Gly_4Ser)_4$ N_H- Hercules são mostradas nas Figuras 18 e 19, respectivamente. Seqüências de DNA (SEQ ID NO: 34) e de aminoácido (SEQ ID NO: 35) de cadeia pesada para BHA10 scFv V_H44:V_L100 N- Hercules são mostradas nas Figuras 20 e 21, respectivamente. Seqüências de DNA (SEQ ID NO: 36) e de aminoácido (SEQ ID NO: 37) de cadeia pesada para BHA10 scFv V_H44:V_L100/$(Gly_4Ser)_4$ N- Hercules são mostradas nas Figuras 22 e 23, respectivamente. Cada uma das cadeias pesadas foi expressada com um peptídeo de sinal no N-término tendo as seguintes seqüências de DNA e de aminoácidos:

ATGGGTTGGAGCCTCATCTTGCTCTTCCTTGTCGCTGTTGCTACGCGT GTCCTGTCC (SEQ ID NO: 61)

MGWSLILLFLVAVATRVLS (SEQ ID NO: 62)

B. Construção de C-scFv "Hercules" com BHA10 convencional, $(G_4S)_4$, V_H44:V_L100, e V_H44:V_L100/$(Gly_4Ser)_4$ scFvs

Quatro projetos de anticorpo biespecífico anti-LTβR (BHA10) x

anti-TRAIL R2 (chi14A2) foram baseados em anexar os BHA10 scFvs convencional e variantes ao término carboxila da cadeia pesada do anticorpo anti-TRAIL R2. Os DNAs de BHA10 scFvs descritos no Exemplo 3 foram usados para construir um painel de anticorpos biespecíficos C- Hercules por amplificação por PCR usando os iniciadores de oligonucleotídeos descritos na Tabela 15. Um $Ser(Gly_4Ser)_3$ ligante foi usado para conectar os BHA10 scFvs ao término carboxila de cadeia pesada chi14A2. O iniciador de PCR forward 5' VH (XWU006-F1) inclui um sítio de endonuclease de restrição *Bam*H I (seqüência sublinhada) para clonagem seguido por seqüência codificando uma porção do peptídeo ligante $Ser(Gly_4Ser)_3$ e o término amino de BHA10 VH. O iniciador de PCR reverso 3' VL (XWU006-R1) *primes* cadeia leve de BHA10 scFv e inclui um códon de parada seguido por um sítio *Bam*H I (seqüência sublinhada) para clonagem. Iniciador de PCR de sobreposição forward 5' interno (XWU006-F2) inclui seqüência codificando o $(Gly_4Ser)_3$ ligante e contém uma mutação silenciosa indicada em tipo em negrito. Iniciador de PCR de sobreposição reverso 3' interno (XWU006-R2) inclui seqüência codificando o $(Gly_4Ser)_3$ ligante e contém uma mutação silenciosa indicada em tipo em negrito para remover um sítio *Bam*HI localizado na região de BHA10 scFv $(Gly_4Ser)_3$ ligante.

Tabela 15. Oligonucleotídeos para amplificação por PCR de C- Hercules com BHA10 convencional, $(G_4S)_4$, $V_H44:V_L100$, e $V_H44:V_L100/(Gly_4Ser)_4$ scFvs.

XWU006-F1 (SEQ ID NO: 38)	5'- GGGGGTGGATCCGGTGGAGGGGGCTCCGGCGGTGGCGG GTCCCAGGTCCAACTGGTGCAGTCTG -3'
XWU006-F2 (SEQ ID NO: 39)	5'- TGGGGGCGGCGGGTCCGGTGGTGGTGGTAG -3'
XWU006-R2 (SEQ ID NO: 40)	5'- TACCACCACCACCGGACCCGCCGCCCCCAG -3'
XWU006-R1 (SEQ ID NO: 41)	5'- GTTAACGGATCCTCATTTGATCTCCACCTTGG -3'

Conforme representado na Figura 14B, seqüências genéticas BHA10 scFv foram amplificadas em uma reação de PCR de duas etapas usando a série de iniciadores de PCR 5' VH XWU006-F1 + 3' VL XWU006-

R1 e uma série de iniciadores de PCR de sobreposição interna (XWU006-F2 e XWU006-R2) designada para eliminar um sítio *Bam*H I dentro da região do BHA10 scFv $(Gly_4Ser)_3$ ligante. Estas condições de PCR foram usadas para preparar BHA10 scFv convencional do plasmídeo pXWU034, BHA10 scFv $(Gly_4Ser)_4$ do plasmídeo pXWU002, BHA10 scFv $V_H44:V_L100$ do plasmídeo pIEH006, e BHA10 scFv $V_H44:V_L100/(Gly_4Ser)_4$ do plasmídeo pIEH009. Os produtos de PCR do painel de BHA10 scFvs amplificados foram purificados por eletroforese por gel agarose usando o kit de extração Ultrafree-DA da Millipore de acordo com instruções do fabricante (Millipore; Bedford, MA). Os produtos de PCR purificados foram digeridos com endonuclease de restrição *Bam*H I e ligados em um único sítio *Bam*H I do vetor pN5KG1 previamente engenheirado para conter várias modificações. Em resumo, o vetor pN5KG1 contendo chi14A2 IgG1 foi modificado para remover o códon de parada na extremidade 3' do gene de cadeia pesada e introduzir nucleotídeos codificando para a seqüência de aminoácidos Ser-Gly-Gly-Gly imediatamente seguidos por um sítio de endonuclease de restrição *Bam*H I (codificando para Gly-Ser) para clonagem. Por último, um sítio de endonuclease de restrição *Bam*H I indesejado interno na região chi14A2 VL também foi eliminado.

O painel de construtos resultantes forma proteínas de fusão dos BHA10 scFvs variantes para o término carboxila da cadeia pesada de anticorpo anti-TRAIL R2 através do $Ser(Gly_4Ser)_3$ ligante de 16 aminoácido. Fusão da seqüência genética BHA10 scFv convencional ao término carboxila da seqüência genética de cadeia pesada de anticorpo anti-TRAIL R2 produziu o plasmídeo pXWU006. Fusão da seqüência genética BHA10 scFv $(Gly_4Ser)_4$ ao término carboxila da seqüência genética de cadeia pesada de anticorpo anti-TRAIL R2 produziu o plasmídeo pXWU034. Fusão da seqüência genética BHA10 scFv $V_H44:V_L100$ ao término carboxila da seqüência genética de cadeia pesada de anticorpo anti-TRAIL R2 produziu o plasmídeo pXWU035. Fusão da seqüência genética BHA10 scFv $V_H44:V_L100/(Gly_4Ser)_4$ ao término carboxila da seqüência genética de cadeia pesada de anticorpo anti-TRAIL R2 produziu o plasmídeo pXWU036.

As misturas da ligação foram usadas para transformar *E. coli*

cepa TOP 10 células competentes (Invitrogen Corporation, Carlsbad, CA). Colônias de *E. coli* transformadas para resistência a fármaco de ampicilina foram triadas para presença de insertos. Análise de seqüência de DNA confirmou a seqüência correta dos construtos finais. A cadeia leve quimérica 14A2 usada é comum entre todos os anticorpos biespecíficos N- e C- Hercules e seqüências de DNA e de aminoácidos são mostradas nas Figuras 15A e 15B. Seqüências de DNA e de aminoácidos de cadeia pesada para BHA10 scFv C- Hercules convencional são mostradas nas Figuras 24 e 25, respectivamente. Seqüências de DNA e de aminoácidos de cadeia pesada para BHA10 scFv $(Gly_4Ser)_4$ C- Hercules são mostradas nas Figuras 26 e 27, respectivamente. Seqüências de DNA e de aminoácidos de cadeia pesada para BHA10 scFv $V_H44:V_L100$ C- Hercules são mostradas nas Figuras 28 e 29, respectivamente. Seqüências de DNA e de aminoácidos de cadeia pesada para BHA10 scFv $V_H44:V_L100/(Gly_4Ser)_4$ C- Hercules são mostradas nas Figuras 30 e 31, respectivamente. Um sumário dos construtos Hercules biespecíficos de C- terminal é encontrado na Tabela 16.

Tabela 16. Plasmídeos intermediários e de expressão codificando 'Hercules'

Vetor	Composição	Anticorpo
pXWU002	pIEH003 + BHA10 scFv $(G_4S)_4$ ligante	BHA10 scFv com $(G_4S)_4$ ligante
pXWU005	pN5KG1 + BHA10 scFv N- Hercules	BHA10 scFv N- Hercules
pXWU006	pN5KG1 + BHA10 scFv C- Hercules	BHA10 scFv C- Hercules
pXWU026	pXWU005 + $(G_4S)_4$ ligante	BHA10 scFv $(G_4S)_4$ N- Hercules
pXWU027	pXWU005 + $V_H44:V_L100$	BHA10 scFv $V_H44:V_L100$ N- Hercules
pXWU028	pXWU005 + $V_H44:V_L100/(Gly_4Ser)_4$	BHA10 scFv $V_H44:V_L100/(Gly_4Ser)_4$ N- Hercules
pXWU033	Modified pXWU006	*Bam* HI site removed in 14A2 VL of C- Hercules
pXWU034	pXWU033 + $(G_4S)_4$ ligante	BHA10scFv $(G_4S)_4$ ligante C- Hercules
pXWU035	pXWU033 + $V_H44:V_L100$	BHA10 scFv $V_H44:V_L100$ C- Hercules
pXWU036	pXWU033 + $V_H44:V_L100/(Gly_4Ser)_4$	BHA10 scFv $V_H44:V_L100/(Gly_4Ser)_4$ C- Hercules

C. Expressão Transitória de Anticorpos Biespecíficos em Células CHO

DNAs de plasmídeo pXWU005, pXWU026, pXWU027, pXWU028; e pXWU006, pXWU034, pXWU035, e pXWU036 (Tabela 16) foram usados

para transformar células CHO DG44 para produção transitória de proteína de anticorpo. Cada 20 ug de plasmídeo de DNA foi combinado com 4 x 10^6 células em um volume de 0,4 ml de 1XPBS. A mistura foi adicionada a uma cuvette de 0,4 cm (BioRad) e colocada sobre gelo por 15 min. As células foram eletroporadsa a 600 uF e 350 volts com um eletroporador Gene Pulser (BioRad). As células foram colocadas no meio CHO-SSFM II contendo 100 uM de Hipoxantina e 16 uM de Timidina dentro de um frasco T-25 e incubadas a 37° por 4 dias.

Sobrenadantes contendo proteínas Hercules produzidas por este sistema de expressão CHO transitório foram coletados e avaliados por Western Blot. A Figura 32 mostra que os anticorpos Hercules contendo ou o ligante V_H44:V_L100 (ds = dissulfeto) ou V_H44:V_L100/$(Gly_4Ser)_4$ BHA10 scFvs estabilizados aprimoraram dramaticamente as produções de expressão independente de se a scFv foi fundida ao N_H- ou C- término (alamedas 3, 4, 7, 8). Surpreendentemente, não foi detectada proteína secretada com BHA10 scFv selvagem nem fusões $(Gly_4Ser)_4$ ligante BHA10 scFv N-terminal indicando o benefício de estabilização de scFv sobre expressão (alamedas 6 e 7). Além disso, fusões convencionais BHA10 scFv e $(Gly_4Ser)_4$ ligante BHA10 scFv C-terminal apresentaram uma quantidade significativa de um subproduto de ~55-60 de peso molecular a qual é substancialmente reduzida nos construtos estabilizados sugerindo que a estabilização de scFv pode melhorar a qualidade do produto.

D. Ensaio ELISA de Ligação Biespecífica

Os sobrenadantes foram testados para atividade de ligação individual a TRAIL R2 e LTβR receptores produzidos de modo recombinante em ensaios ELISA. Em ambos os ensaios, o receptor foi imobilizado sobre lâminas e as amostras de ensaio foram incubadas para permitir ligação a receptor. Amostras ligadas foram detectadas com anticorpo marcado.

Lâminas de microtítulo de 96 cavidades Immulon II (Fisher, Cat nº 14245-61) foram revestidas com 100 µl/ cavidade de 2 µg/ml de LTβR-Ig em tampão Na_2CO_3 / $NaHCO_3$ pH 9.5, de um dia para o outro a 4 ºC e bloqueadas com 200 µl de tampão de diluição (0,5% de leite em pó sem gordu-

ra Nonfat Dry Milk em PBS mais 0,01% de Thimerosal) por 1 h a 37 ºC. Na etapa seguinte, 100 µl de sobrenadante 'Hercules' individual ou proteína purificada em tampão de diluição foi adicionado a cavidades em duplicata e incubado por 1 h a 37 ºC. Depois de lavar com água corrente, 100 µl de 100 ng/ml de proteína de fusão TRAIL-R2 Fc 6X-His marcada (R&D Systems, Minneapolis, MN) foi adicionado às cavidades e incubado por 1 h a 37 ºC. Depois de lavar, 100 µl de uma diluição a 1/2000 de Conjugado Penta-His HRP (QIAGEN, Cat nº 34460) foi adicionado a cada cavidade e incubado por 1 h a 37 ºC. Depois de lavar, 100 µl/ cavidade de um Substrato HRPO combinado a TMB Peroxidase Substrato /Peroxidase Solução B (Kirdgaard e Perry Labs, Cat. 50-76-00) foi adicionado. A reação foi interrompida com 100 µl de 2 M de H_2SO_4 depois de 5 a 10 min. O OD foi medido a 450 nm e 540 nm usando uma leitora de lâminas Molecular Devices e foram geradas curvas de ligação.

As Figuras 33 e 34 mostram que tanto os anticorpos Hercules de N- terminal (Figura 33A, 34A) e quanto de C- terminal (Figura 33B, 34B) contendo ou a BHA10 scFv estabilizada por V_H44:V_L100 (ds = dissulfeto) ou por ligante V_H44:V_L100/$(Gly_4Ser)_4$ mostram aprimorada ligação a LTβR (Figura 33) e TRAIL R2 (Figura 34), respectivamente, comparados com Hercules contendo um BHA10 scFv convencional (wt).

<u>E. Expressão Estável de Anticorpos Biespecíficos em Células CHO, Purificação de Anticorpos, e Caracterização.</u>

DNAs de plasmídeo pXWU027, pXWU028, e pXWU006, pXWU035, e pXWU036 (Tabela 16) foram usados para transformar células CHO DG44 deficientes de DHFR para produção estável de proteína de anticorpo. Células transfectadas foram cultivadas em meio α menos MEM contendo 2 mM de glutamina suplementada com 10% de soro fetal bovino dialisado (Invitrogen Corporation) e enriquecido como um conjunto de cultura de carga estável usando anticorpos marcados com fluorescente e classificação celular ativada por fluorescente reiterativa (FACS) (Brezinsky, *et al. J Immunol Methods.* 277(1-2): 141-55 (2003)). FACS também foi usada para gerar linhagens celulares individuais. Conjuntos de células ou linhagens celulares

foram adaptados para condições livres de soro e aumentados em escala para produção de anticorpos.

Os sobrenadantes de conjuntos de células CHO transfectadas ou linhagens celulares expressando 1) Hercules de C –terminal com a BHA10 scFv estabilizada por $V_H44:V_L100$ e 2) Hercules de C- terminal com a BHA10 scFv estabilizada por ligante $V_H44:V_L100/(Gly_4Ser)_4$, bem como linhagens de células CHO clonais expressando 3) Hercules de N- terminal com a BHA10 scFv estabilizada por $V_H44:V_L100$ e 4) Hercules de N- terminal com a BHA10 scFv estabilizada por ligante $V_H44:V_L100/(Gly_4Ser)_4$ foram coletadas e purificadas usando coluna de Proteína A Sepharose FF (6 mL) usando um PBS, 10 mM de tampão de fluxo EDTA. Proteínas Hercules foram elutriadas usando 0,1 M de glicina, pH 3.0 e neutralizadas imediatamente para pH 7.5-8.5 usando base Tris. Além disso, C-Hercules contendo a BHA10 scFv convencional foi expressado e purificao por cromatografia de Proteína A para comparação. Proteína A elutria de C-Hercules contendo a BHA10 scFv convencional, Hercules de C-terminal com a $V_H44:V_L100$ BHA10 scFv, e Hercules de C-terminal com a $V_H44:V_L100/(Gly_4Ser)_4$ BHA10 scFv foram examinados para a presença de agregados por cromatografia de exclusão de tamanho analítica (Figura 35). O perfil do cromatograma de C-Hercules contendo a BHA10 scFv convencional apresentou ~ 40% de agregados. Em contraste, Hercules de C-terminal com a $V_H44:V_L100$ BHA10 scFv reduziu o nível de agregados para 20% e estabilização adicional atingida através de Hercules C-terminal com a $V_H44:V_L100/(Gly_4Ser)_4$ BHA10 scFv reduziu o nível de agregados para 10%. O nível de agregação pareceu depender das propriedades da scFv e não se foi anexado ao N_H- ou C-término da IgG 14A2 (dados não mostrados). Os eluentes de proteína A Hercules foram adicionalmente purificados por diálise em 0,1 M de acetato, pH 5.0, e purificação por cromatografia de permuta catiônica MonoS (GE Healthcare) usando um tampão de fluxo idêntico ao dialisado. Proteínas Hercules foram elutriadas usando uma etapa de gradiente para 0,1 M de acetato, pH 5.0, 0,5 M de NaCl. Os eluentes MonoS foram coletados e passados através de uma coluna SEC de preparativo TosoHaas para remover

agregados.

As Figuras 36A e 36B mostram géis SDS-PAGE de Hercules de N- terminal purificado com o BHA10 scFv estabilizado por V_H44:V_L100 e Hercules de N_H- terminal com o BHA10 scFv estabilizado por ligante V_H44:V_L100/$(Gly_4Ser)_4$ e Hercules de C- terminal com o BHA10 scFv estabilizado por V_H44:V_L100 e Hercules de C- terminal com o BHA10 scFv estabilizado por ligante V_H44:V_L100/$(Gly_4Ser)_4$, respectivamente. As alamedas reduzidas mostram os tamanhos esperados das proteínas de cadeia pesada e leve. De modo importante, não há nível significativo de subprodutos de peso menor molecular degradados ou indesejados que tenha sido freqüentemente observado com Hercules contendo domínios scFv selvagens.

A Figura 37 (Painéis A e C) mostra os perfis de eluição SEC analítica de Hercules de N_H- terminal com a BHA10 scFv estabilizada por V_H44:V_L100 e Hercules de N_H- terminal com a BHA10 scFv estabilizada por ligante V_H44:V_L100/$(Gly_4Ser)_4$, respectivamente, subseqüente à etapa de purificação de proteína A inicial. A Figura 37 (Panels B e D) também mostra os perfis de eluição SEC analítica de Hercules de N_H-terminal com a BHA10 scFv estabilizada por V_H44:V_L100 e Hercules de N_H- terminal com a BHA10 scFv estabilizada por ligante V_H44:V_L100/$(Gly_4Ser)_4$, respectivamente, depois de remoção por SEC do preparativo de contaminantes de proteína não-monomérica residual. Similarmente, Figura 38 (Painéis A e C) mostra os perfis de eluição SEC analítica de Hercules de C- terminal com a BHA10 scFv estabilizada por V_H44:V_L100 e Hercules de C- terminal com a BHA10 scFv estabilizada por ligante V_H44:V_L100/$(Gly_4Ser)_4$, respectivamente, subseqüente à etapa de purificação de proteína A inicial. A Figura 38 (Painéis B e D) mostra os perfis de eluição SEC analítica de Hercules de C-terminal com a BHA10 scFv estabilizada por V_H44:V_L100 e Hercules C- terminal com a BHA10 scFv estabilizada por ligante V_H44:V_L100/$(Gly_4Ser)_4$, respectivamente, depois de remoção por SEC do preparativo de contaminantes de proteína não-monomérica residual. Estes estudos demonstram que a estabilização da BHA10 scFv por adição ou do V_H44:V_L100 dissulfeto ou do V_H44:V_L100 dissulfeto e $(Gly_4Ser)_4$ ligante resulta em quantidades de preparativo de >98%

pure, anticorpo biespecífico Hercules monomérico que é essencialmente livre de espécies de peso molecular de superior.

Estudos de DSC realizados sob condições idênticas demonstram que C-Hercules $V_H44:V_L100/(Gly_4Ser)_4$ BHA10 scFv é mais termoestável do que C-Hercules contendo convencional BHA10 scFv (Figura 39). O perfil de desnaturação para C-Hercules contendo BHA10 scFv convencional inicia ~5 ºC temperatura menor do que o perfil observado para o C-Hercules $V_H44:V_L100/(Gly_4Ser)_4$ BHA10 scFv. Estes resultados sugerem que a scFv pode limitar a termoestabilidade global das moléculas Hercules. Isto se correlaciona com um aumento de 5 ºC da T_M observada para a $V_H44:V_L100/(Gly_4Ser)_4$ BHA10 scFv em relação à BHA10 scFv convencional.

Versões purificadas dos construtos Hercules estabilizados foram então testados para atividade de ligação biespecífica a receptores TRAIL-R2 e LTβR produzidos de modo recombinante em uma prova ELISA. Nesta prova LTβR Ig receptor foi imobilizado sobre lâminas e amostras de ensaio foram em seguida incubadas para permitir ligação a receptor. Amostras não ligadas foram removidas por lavagem seguida por uma segunda etapa de incubação com TRAIL R2- $(His)_6$. Depois de uma etapa de lavagem, complexos duplamente ligados foram detectados com um anticorpo anti- $(His)_6$ marcado. Os resultados deste estudo e a concentração eficaz (EC50, em μg/ml) na qual cada construto são representados na Figura 40. A Figura 40 mostra que tanto os anticorpos Hercules de N- terminal quanto de C- terminal com ou a BHA10 scFv estabilizada por $V_H44:V_L100$ (ds=dissulfeto) ou por ligante $V_H44:V_L100/(Gly_4Ser)_4$ demonstram clara ligação biespecífica tanto a LTβR quanto a TRAIL R2. Uma versão multivalente (pentamérica) de um de anticorpo ligação de LTβR (CBE11) foi usado como um controle para demonstrar que reconhecimento único de LTβR somente não induz sinal na prova de ligação.

Exemplo 8. Aumento em Escala da Produção de Anticorpos "Hercules" Biespecíficos Estabilizados em Células CHO

Linhagens de células CHO DHFR-deficientes transfectadas estavelmente com DNAs de plasmídeos pXWU028 e pXWU036 (Tabela 16)

foram triadas para selecionar isolados celulares únicos que são capazes de expressar altos níveis de moléculas Hercules solubilizadas e adequadamente dobradas que foram estabilizadas usando os métodos da invenção. Os métodos de triagem celular empregaram a análise de classificação celular ativada por fluorescência (FACS) de Brezinky et al. (Brezinky *et al.*, J Immunol Meth (2003). 277:141-155). Em resumo, anticorpos anti-Hercules com etiqueta fluorescente foram usados para marcar células CHO apresentando expressão transitória de anticorpos Hercules sobre sua superfície. Céulas apresentando uma assinatura de intensidade de fluorescência foram então selecionadas adequando a comutação do equipamento de classificação celular à assinatura. Células únicas apresentando altos níveis de produtividade foram selecionadas deste modo e adaptadas para condições livres de soro para estabelecer linhagens celulares produtoras estáveis. As linhagens celulares produtoras foram subseqüentemente aumentadas em escala para produção e purificação da proteína de anticorpo biespecífico.

80 L de sobrenadante de Hercules de N- terminal contendo $V_H44:V_L100/(Gly_4Ser)_4$ BHA10 scFv (XWU028) de uma rodada de biorreator de 11 dias foi colhido e pré-clareado por ultrafiltração. O anticorpo biespecífico foi capturado por cromatografia de Proteína A e elutriado em volume de 1218 ml. A fração de Proteína A foi adicionalmente purificada em dois lotes separados por cromatografia de permuta aniônica seguida por cromatografia de exclusão de tamanho do preparativo. O primeiro lote produziu 825,5 mg em uma concentração de 4,85 mg/ml em PBS com 98,9% de pureza e uma carga de endotoxina de 0,13 EU/mg. O segundo lote produziu 318 mg em uma concentração de 10,3 mg/ml com uma pureza de 99,1% e uma carga de endotoxina de 2,37 EU/mg. Um total de 1143,5 mg de Hercules de N-terminal monomérico altamente purificado contendo $V_H44:V_L100/(Gly_4Ser)_4$ BHA10 scFv foi recuperado.

24 L de sobrenadante de Hercules de C- terminal contendo $V_H44:V_L100/(Gly_4Ser)_4$ BHA10 scFv (XWU036) de uma rodada de biorreator de 11 dias foi colhido e pré-clareado por ultrafiltração. O anticorpo biespecífico foi purificado em dois lotes separados conforme descrito acima usando

cromatografia de Proteína A para capturar produto bruto seguida por cromatografia de permuta aniônica e cromatografia de exclusão de tamanho do preparativo. O primeiro lote produziu 985,5 mg em uma concentração de 13 mg/ml com 98,5% de pureza e uma carga de endotoxina de 0,01 EU/mg. O segundo lote produziu 570 mg em uma concentração de 11,42 mg/ml com uma pureza de 99 % e uma carga de endotoxina de 1,91 EU/mg. Um total de 1555,5 mg de Hercules de C- terminal monomérico altamente purificado contendo V_H44:V_L100/$(Gly_4Ser)_4$ BHA10 scFv foi recuperado.

Uma linhagem de células CHO produtoras do anticorpo biespecífico XWU054 foi isolada e foi visto que produzem 21,5 mg/L. Isto está dentro da faixa esperada para linhagens de células CHO não amplificadas em grau de pesquisa e nós preveríamos produtividade muito maior se fôssemos gerar uma Linhagem de Células de Produção.

Estes estudos de aumento em escala da produção foram conduzidos com linhagens de células CHO não amplificadas e ainda resultaram em produções excedentes de 1 gm de anticorpos biespecíficos biofisicamente estáveis e de alta qualidade. Presumivelmente uma estratégia de amplificação de linhagem celular aumentaria muito a produtividade celular das linhagens de células CHO transfectadas possibilitando o desenvolvimento de processos adequados para aplicações comerciais. Estes estudos exemplificam a utilidade dos métodos descritos nesta invenção para possibilitar o aumento em escala da produção de anticorpos biespecíficos estáveis em uma linhagem celular (por exemplo, CHO) adequada para fabricação.

Exemplo 9. Atividade Biológica de Anticorpos "Hercules" Biespecíficos

Linhagens de células tumorais WiDr, (ATCC CCL-218) uma linhagem celular de carcinoma de cólon humano, Me180, (ATCC HTB 33) uma linhagem celular de carcinoma epitelial cervical humano, e MDA231, (Dr. Dajun Yang, University of Michigan) uma linhagem celular de carcinoma de mama humano foram cultivadas em MEM-Earles com 10% de FCS, 2 mM de L-Glutamina,1X aminoácidos não-essenciais, 0,5 mM de piruvato de sódio, e Penicilina / Estreptomicina. Linhagens de células tumorais foram enxaguadas uma vez com PBS e as células foram liberadsa por digestão com

tripsina. As células foram coletadas por centrifugação, ressuspendida em meio completo, contadas e semeadas em lâminas de cultura de tecido de 96 cavidades a 5000 células/ cavidade para WiDr e Me180 e 1500 células/ cavidade para MDA231. IFN□ humano (Biogen Idec, Corp) é adicionado às suspensões de células para resultar em uma concentração final de citocina de 80 U/ml para WiDr e MDA231 e 50 U/ml para Me180. 50 µl da suspensão de células tumorais / IFNγ foram misturadas com 50 µl de 2X diluições seriais de 3 dobras concentradas de anticorpos de ensaio preparados em meio completo. As concentrações finais de anticorpos de ensaio tipicamente variaram de 5000 pM a 0,07 pM. As células foram cultivadas por 4 dias (WiDr & Me180) ou 3 dias (MDA231) a 37 °C em uma câmara umidificada com 5% de CO_2 e destruição celular avaliada pela adição de 20 µL/ cavidade de reagente Promega CellTiter 96 AQueous One Solution Cell Proliferation Assay (Promega Corporation, Madison, WI). As lâminas foram lidas em uma leitora de lâminas de microtítulo a 490 nM (Spectromax Plus, Molecular Devices, Sunnyvale CA). Foi feito gráfico dos dados usando Microsoft Excel (Microsoft Inc, WA).

Os resultados destes estudos são mostrados nas Figuras 41A-41D. A Figura 41A mostra que o anticorpo 14A2 IgG teve atividade modesta para inibir o crescimento de células tumorais WiDr e em combinação com BHA10 IgG apresentou um ligeiro aumento em atividade anticélulas tumorais comparado com BHA10 IgG somente. Em contraste, ambos os anticorpos Hercules biespecíficos apresentaram aumento da destruição de células tumorais das células WiDr. A Figura 41B mostra que tanto os anticorpos 14A2 IgG quanto BHA10 IgG tiveram atividade modesta, se não desprezível, na inibição do crescimento de células tumorais Me180 como agentes únicos ou em combinação. Em contraste, ambos os anticorpos Hercules biespecíficos apresentaram destruição de células tumorais das células Me180 cells. A Figura 41C mostra que tanto os anticorpos 14A2 IgG quanto BHA10 IgG tiveram atividade desprezível na inibição do crescimento de células tumorais MDA231 como agentes únicos e talvez alguma atividade quando usados em combinação. Em contraste, o anticorpo Hercules biespecífico XWU036 (Her-

cules C- terminal com o BHA10 scFv estabilizado pelo ligante $V_H44:V_L100/(Gly_4Ser)_4$) apresentou destruição de células tumorais das células MDA231. A Figura 41D mostra que nenhum dos anticorpos demonstra qualquer atividade para células endoteliais da veia umbilical humana cultivadas de controle (HUVEC) demonstrando que a atividade de destruição celular dos anticorpos biespecíficos não é indiscriminada. HUVEC mostra coloração positiva para LTβR e TRAIL-R2 por análise FACS (dados não mostrados) indicando presença de receptores e que a atividade do anticorpo biespecífico pode ser dependente de gatilho específico ou componentes de caminho único em células tumorais.

Exemplo 10. Estudos da Estabilidade de Anticorpos "Hercules" Biespecíficos

Foram conduzidos estudos da estabilidade em tempo real das amostras de anticorpos biespecíficos de N- e C-terminal XWU028 (BHA10 scFv $V_H44:V_L100/(Gly_4Ser)_4$ N- Hercules) e XWU036 (BHA10 scFv $V_H44:V_L100/(Gly_4Ser)_4$ C-Hercules) armazenados em 2 a 8 ºC por três meses. A qualidade da proteína foi avaliada para (1) agregação, (2) precipitação, (3) clivagem de polipeptídeo ou proteólise, e (4) modificações pós-translacionais tais como deamidação ou oxidação.

As amostras de anticorpos biespecíficos de N- e C-terminal de alta e baixa concentração usadas para os estudos foram XWU028 a 1,8 mg/mL (Baixa) e 10,3 mg/mL (Alta) e XWU036 a 5,0 mg/mL (Baixa) e 11,4 mg/mL (Alta). Os anticorpos foram formulados em PBS. Para o estudo da estabilidade, amostras dos pontos de tempo inicial (T = 0), intermediário (T = 1 semana, 2 semanas, 1 mês, 2 meses) e final (T = 3 meses) foram analisadas imediatamente depois da coleta da amostra. Além disso, amostras iniciais (T = 0) foram congeladas e armazenadas a –70 ºC até serem degeladas para análises secundárias ao final do estudo de 3 meses. IgG BHA10 (8,7 mg/mL) foi usada como controle e foi manuseada de modo similar.

A. Análise de agregação e precipitação de proteína

A agregação ou precipitação de proteína foi monitorada usando cromatografia de exclusão de tamanho analítica (SEC) encadeada a dispersão luminosa inline e detectores de do índice refrativo (Wyatt Technologies,

MiniDawn e rEX, respectivamente). A análise SEC / de dispersão luminosa não mostrou evidência de agregação ou perda de material que possa ocorrer através de precipitação. Os perfis de eluição SEC dos pontos de tempo inicial, T = 0, e final, T = 3 meses, para XWU028, XWU036, e o anticorpo BHA10 foram quase idênticos (Figuras 56A, 56B, e 56C, respectivamente). As amostras XWU028 e XWU036 acumularam ambas ~1% de material agregado por volta do completamento do estudo; no entanto a detecção destes agregados foi próxima do menor limite de detecção determinado para este método (Tabela 17). A formação de agregados foi independente da concentração de proteína. Com base nas massas moleculares determinadas usando detecção de dispersão luminosa, é possível que os baixos níveis de agregados nas amostras de ensaio sejam provavelmente um produto de espécies monomérica\s convertendo para dímeros. Nenhuma das amostras de anticorpos biespecíficos, XWU028 ou XWU036, acumulou níveis detectáveis de agregados de alta ordem, muito embora este tipo de agregado pudesse ter sido facilmente observado usando estes métodos. O anticorpo BHA10 não demonstrou aumento em agregados durante o curso de 3 meses do estudo (Tabela 17).

Tabela 17. Percentagem de monômero detectado usando cromatografia de exclusão de tamanho analítica.

	T = 0	T = 1 semana	T = 2 semanas	T = 1 mês	T = 2 meses	T = 3 meses
SEQ ID nº 49 Baixa	99,0	98,7	98,4	98,3	98,2	97,9
SEQ ID nº 49 Alta	99,1	99,0	98,9	98,6	98,2	98,1
SEQ ID nº 51 Baixa	98,8	97,5	97,5	97,4	97,2	96,7
SEQ ID nº 51 Alta	98,9	98,8	98,7	98,5	98,1	96,7
BHA10 IgG	97,4	97,5	97,5	97,5	97,3	97,5

B. Análises de proteólise e de modificação Pós-translacional

A proteólise foi monitorada usando SDS-PAGE e análise de massa intacta por cromatografia líquida / espectroscopia de massa (LC/MS) (Agilent LC/MSD TOF acoplado a um sistema Agilent 1100 LC através de uma interface de electrospray). Para análise SDS-PAGE, 5 ug de proteína foi

carregado por alameda. Amostras reduzidas foram preparadas em 1X tampão de amostra de Tris-glicina contendo 5% de β-mercaptoetanol. XWU028 e XWU036 não apresentaram evidência de proteólise durante o período de armazenamento de 3 meses em 2 a 8 ºC conforme determinado por análise SDS-PAGE não-redutora e redutora (Figuras 57A e B). Similarmente, XWU028 e XWU036 não apresentaram evidência de produtos proteolíticos de peso molecular inferior durante o período de armazenamento de 3 meses em 2 a 8 ºC conforme determinado por análise LC/MS (Figura 58).

Modificações pós-translacionais foram monitoradas usando LC/MS. Amostras de T = 0, T = 1 mês, e T = 3 meses foram analisadas tanto na forma não-reduzida quanto reduzida. Amostras reduzidas foram preparadas tratando em 50 mM de DTT / 4 M de cloridrato de guanidina por 1 hora a 37°C. Tampão A de HPLC consiste de 0,03% de TFA em água e tampão B de HPLC contém 0,025% de TFA em acetonitrila. A taxa de fluxo foi mantida constante a 100 µl por minuto. 7,5 µg de cada amostra (reduzida e não-reduzida) foi injetado sobre uma coluna C4 de 2,1 x 50 mm e analisado por Agilent ESI-TOF. Um método de ligar-e-eluir foi usado para amostras não-reduzidas ao passo que um método de gradiente foi usado para amostras reduzidas. Foram obtidos espectros usando o pacote de software Analyst e desconvolutos usando o pacote de software MaxEnt1 incluídos com a instrumentação. Espectros de massa desconvolutos de XWU028 e XWU036 (altas concentrações) em T = 0, T = 1 mês, e T = 3 meses são mostrados na Figura 58. Nem XWU028, nem XWU036, demonstraram quaisquer alterações detectáveis em seus espectros de massa, indicando ausência de modificação pós-translacional que potencialmente pode afetar de modo adverso a função ou estabilidade de anticorpos biespecíficos.

Tomados juntos, os dados indicam que ambos os anticorpos biespecíficos de N- e C-terminal contendo domínios scFv estabilizados são estáveis sob condições de armazenamento prolongado tais como as requeridas para produtos de fármacos biológicos. Estes resultados são particularmente estimulantes porque estes estudos de estabilidade foram conduzidos com os anticorpos biespecíficos de N- e C-terminal preparados em um tam-

pão simples (PBS) e não com soluções preparadas com formulações ótimas.

Exemplo 11. Atividade Farmacocinética In Vivo e Estabilidade Sérica de Anticorpos "Hercules" Biespecíficos

Uma injeção de bólus único de 10 mg/kg (1 mg/ml) de Hercules de N- terminal (XWU028) ou Hercules de C- terminal (XWU036) diluído em salina tamponada com fosfato (PBS) foi administrada por via intraperitoneal em camundongos machos CB17- *scid*. Os camundongos foram sacrificados em 0, 0.5, 2, 6, 24, 48, 72, 96, 168, 240, e 336 horas pós-injeção usando três camundongos por ponto de tempo para cada anticorpo biespecífico. Amostras de soro foram preparadas para análise por ensaio ELISA para quantificar os níveis dos anticorpos biespecíficas. Lâminas de ELISA foram revestidas com IgG de cabra-anti-humana, bloqueada com PBS / 1% de BSA, e diluições de soro contendo os anticorpos biespecíficos foram diluídas serialmente em PBS / 1% de BSA, adicionadas às lâminas e incubadas. Anticorpos capturados foram detectados com um anticorpo ligado a HRP de dadeia kappa de cabra-anti-humana. Os resultados do estudo farmacocinético são mostrados na Tabela 18. Hercules de N- terminal (XWU028) tem uma meia-vida de eliminação ($t_{½}$) de 10,3 dias com uma concentração sérica máxima de 85,67 µg/ml. Hercules de C- terminal (XWU036) tem uma meia-vida de eliminação mais longa ($t_{½}$) de 15,1 dias com uma concentração sérica máxima de 105,67 µg/ml. Ambas as moléculas têm volumes similares de distribuição (Vd) embora estes valores não tenham sido ajustados para biodisponibilidade.

Tabela 18. Parâmetros Farmacocinéticos

Parâmetro	Unidades	Hercules N-terminal (XWU028) Valor	Hercules C-terminal (XWU036) Valor
T1/2	hora	247,9386	362,3607
Cmax	ug/ml	85,67	105,67
Vd *	L/kg	0,1707	0,1748
Cl *	ml/min/kg	0,008	0,0056

* não ajustado para biodisponibilidade

Amostras de soro também foram analisadas no ELISA de ligação biespecífica descrito no Exemplo 7. Neste ensaio, as amostras de soro contendo Hercules N- terminal (XWU028) ou Hercules C- terminal (XWU036) descritas acima foram testadsa para atividade de ligação biespecífica a receptores TRAIL-R2 e LTβR produzidos de modo recombinante em um ensaio ELISA. As Figuras 59A e B mostram que amostras de soro de camundongos tratados com Hercules de N- e C- terminal contêm anticorpos que ligam de modo biespecífico tanto a TRAIL-R2 quanto a LTβR e são intimamente paralelos aos perfis de eliminação observados a partir do estudo farmacocinético indicando que os anticorpos biespecíficos estão permanecendo intactos sob condições fisiológicas por períodos de tempo prolongado.

Exemplo 12. Atividade Biológica In Vivo de Anticorpos "Hercules" Biespecíficos

(A) Atividade Biológica In Vivo de Anticorpos "Hercules" Biespecíficos em um Modelo de Tumor de Câncer do Cólon

Células de carcinoma de cólon humano WiDr (2x10E6 células por camundongo) foram implantadas por via subcutânea em 125 camundongos nus atímicos. Os tumores foram cultivados até atingirem aproximadamente 100 mg, ponto no qual um total de 70 camundongos foram selecionados para o estudo, divididos em 7 grupos de 10 camundongos. Foram administrados tratamentos IP, iniciando no dia 13 pós-implantação, como se segue: Grupo 1 = PBS sem pirogênio; Grupo 2 = CBE11, 2 mg/kg, 1x/ semana; Grupo 3 = hBHA10, 2 mg/kg, 2x/ semana; Grupo 4 = ch14A2, 2 mg/kg, 2x/ semana; Grupo 5 = Hercules-II XWU028, 2 mg/kg, 1x/ semana; Grupo 6 = Hercules-II XWU036, 2 mg/kg, 1x/ semana; Grupo 7 = hBHA10 + ch14A2, 1 mg/kg each, 2x/ semana. Os tamanhos dos tumores e os pesos corporais foram registrados bissemanalmente. O estudo foi terminado quando o tamanho médio do tumor do grupo de veículo atingiu aproximadamente 2000 mg. O volume do tumor foi calculado usando a fórmula: ($L \times W^2/2$). Análises interinas de camundongos tratados com o XWU028 e XWU036 apresentaram atividade antitumoral significativa ($p < 0,001$) para ambos os anticorpos biespecíficos comparados com veículo controle de PBS (Figura 60).

Para XWU028, foram observadas reações antitumorais significativas ($p < 0,05$) comparadas com reações com tratamento único com mAb hBHA10 ou ch14A2. Para XWU036, foi observada uma reação antitumoral significativa ($p < 0,05$) comparada com a reação com tratamento com mAb ch14A2 e, notavelmente, uma maior reação significativa ($p < 0,01$) comparada com mAb hBHA10. De modo importante, os anticorpos biespecíficos demonstraram superior atividade antitumoral *in vivo* embora administrados somente uma vez por semana sugerindo que os reforços da estabilidade descritos nesta invenção resultam em aprimoradas propriedades de anticorpo e estabilidade física sob condições fisiológicas.

(B) Atividade Biológica In Vivo de Anticorpos "Hercules" Biespecíficos em um Modelo de Tumor de Câncer do Mama

Células de carcinoma de mama humano MDA-MB-231 foram implantadas por via subcutânea em 135 camundongos nu atímicos (2x10E6 células por camundongo). Os tumores foram cultivados até o dia 13 ponto no qual 75 camundongos portando tumor com um tamanho médio de aproximadamente 168 mg foram atribuídos a grupos de tratamento (N = 10) e de controle de veículo (N = 15). Os camundongos receberam anticorpos e veículo IP iniciando no dia 13. Grupos exemplares são mostrados como se segue: Grupo 1 = PBS sem pirogênio; 1x/ semana; Grupo 2 = hBHA10, 2 mg/kg, 2x/ semana; Grupo 3 = ch14A2, 2 mg/kg, 2x/ semana; Grupo 4 = Hercules-II XWU036, 2 mg/kg, 1x/ semana; Grupo 5 = hBHA10 + ch14A2, 1 mg/kg cada, 2x/ semana. Os tamanhos dos tumores e os pesos corporais foram registrados bi-semanalmente. O estudo foi terminado quando o tamanho médio do tumor do grupo de veículo atingiu aproximadamente 2800 mg. O volume do tumor foi calculado usando a fórmula: $(L \times W^2/2)$. XWU036 demonstrou atividade antitumoral estatisticamente significativa ($p < 0,001$) comparado com qualquer tratamento único com hBHA10 ou ch14A2 mAbs ou com tratamento com uma mistura dos dois anticorpos. De modo importante, o anticorpo biespecífico XWU036 demonstrou boa atividade antitumoral *in vivo* administrado em um esquema de dosagem de uma vez por semana sugerindo que os reforços da estabilidade descritos nesta invenção resultam em aprimora-

das propriedades de anticorpo e estabilidade física sob condições fisiológicas (vide a Figura 86).

Exemplo 13. Produção de Anticorpos "Hercules" Biespecíficos Estabilizados Contendo Mutações scFv Estabilizantes Não-covalentes

BHA10 scFvs da invenção contendo mutações estabilizantes não-covalentes somente e em combinação com uma ligação dissulfeto V_H44-V_L100 foram usados para construir anticorpos C-Hercules biespecíficos estabilizados compreendendo uma fusão de anticorpo de IgG 14A2 quimérico que liga a TRAIL R2 receptor com um BHA10 scFv que liga a LTβR. Os anticorpos biespecíficos foram construídos como fusões de BHA10 scFv de C-terminal usando métodos essencialmente conforme descrito no Exemplo 7.

A. Construção de Anticorpo Biespecífico C-Hercules com BHA10 V_H S16E + V_L S46L e V_H44-V_L100/V_H S16E + V_L S46L scFvs

PCR foi usado para amplificar fragmentos genéticos de BHA10 scFv variantes de DNAs de plasmídeo pIEH-050 (plasmídeo parental do plasmídeo de elevada expressão pIEH076) e pIEH-052 (plasmídeo parental do plasmídeo de elevada expressão pIEH080) contendo os BHA10 V_H S16E + V_L S46L e V_H44-V_L100/V_H S16E + V_L S46L scFvs, respectivamente usando os iniciadores de oligonucleotídeo descritos na Tabela 19. Os fragmentos genéticos de BHA10 scFv variantes foram isolados em gel. Devido a um sítio BamHI na região de ligante de plasmídeos pIEH-050 e pIEH-052, os fragmentos genéticos foram digeridos com endonucleases de restrição pPuM I e Kpn I e ligados ao plasmídeo modificado pN5KG1 digerido com as mesmas endonucleases de restrição resultando em um produto de fusão dos scFvs anti-LTBR estabilizados (BHA10) ao término carboxila do domínio CH3 do anticorpo anti-TRAILR2 (14A2) através de um ligante de 16 aminoácidos $Ser(Gly_4Ser)_3$. Seqüências corretas foram confirmadas por análise de seqüência de DNA.

Tabela 19. Oligonucleotídeos para amplificação por PCR de BHA10 V_H S16E + V_L S46L e V_H44-V_L100/V_H S16E + V_L S46L scFvs.

pXWU006-F1 (SEQ ID NO: 56)	5'- GGGGGTGGATCCGGTGGAGGGGGCTCCGGCGGTGGCGG GTCC CAGGTCCAACTGGTGCAGTCTG -3'
XWU006-R1 (SEQ ID NO: 57)	5'- GTTAACGGATCCTCATTTGATCTCCACCTTGG -3'

Fusão da seqüência genética de BHA10 scFv V_H S16E + V_L S46L ao término carboxila da seqüência genética de cadeia pesada do anticorpo anti-TRAIL R2 produziu o plasmídeo pXWU054. Fusão da seqüência genética de BHA10 scFv V_H44-V_L100/V_H S16E + V_L S46L ao término carboxila da seqüência genética de cadeia pesada do anticorpo anti-TRAIL R2 produziu o plasmídeo pXWU055.

As misturas de ligação foram usadas para transformar 10 células competentes de *E. coli* cepa TOP (Invitrogen Corporation, Carlsbad, CA). Colônias de *E. coli* transformadas para resistência ao fármaco ampicilina foram triadas para presença de inserções. Análise de seqüência de DNA confirmou a seqüência correta dos construtos finais. O DNA de cadeia leve 14A2 quimérico (SEQ ID NO: 28) e seqüências de aminoácidos (SEQ ID NO: 29) são mostrados nas Figuras 15A e 15B. O DNA de cadeia pesada (SEQ ID NO: 52) e seqüência de aminoácidos (SEQ ID NO: 53) para anticorpo biespecífico C-Hercules BHA10 scFv V_H S16E + V_L S46L são mostrados nas Figuras 61 e 62, respectivamente. O DNA de cadeia pesada (SEQ ID NO: 54) e seqüência de aminoácidos (SEQ ID NO: 55) para anticorpo biespecífico C-Hercules BHA10 scFv V_H44-V_L100/V_H S16E + V_L S46L são mostrados nas Figuras 63 e 64, respectivamente.

As cadeias pesadas empregaram o mesmo peptídeo de sinal usado previamente.

B. Expressão Estável de Anticorpos Biespecíficos em Células CHO, Purificação de Anticorpos, e Caracterização.

Linhagens de células CHO deficientes de DHFR transfectadas de modo estável com DNAs de plasmídeo pXWU054 e pXWU055 e adaptadas para condições livre de soro foram escaladas para produção de proteína de anticorpo biespecífico e proteínas purificadas conforme descrito no Exemplo 8. A proteína A elutria de sobrenadantes contendo C-Hercules com o

anticorpo biespecífico estabilizado V_H S16E + V_L S46L BHA10 scFv e C-com o V_H44-V_L100/V_H S16E + V_L S46L BHA10 scFv estabilizado foram examinados para a presença de agregados por cromatografia de exclusão de tamanho analítica (Figura 65). O perfil do cromatograma de C-Hercules contendo o BHA10 scFv convencional mostrou ~ 40% de agregados. Em contraste, Hercules de C-terminal com o V_H S16E + V_L S46L BHA10 scFv estabilizado reduziu significativamente agregados para níveis comparáveis aos observados com IgGs de rotina.

As Figuras 66A e 66B mostram géis SDS-PAGE de Hercules de C-terminal purificado com os V_H S16E + V_L S46L BHA10 scFv e Hercules de C-terminal com os V_H44:V_L100/V_H S16E + V_L S46L BHA10 scFv, respectivamente. As alamedas reduzidas mostram os tamanhos esperados das proteínas de cadeia pesada e leve. De modo importante, não há nível significativo de subprodutos de peso molecular inferior degradados ou indesejados que tenham freqüentemente sido observados com Hercules contendo domínios scFv selvagens .

As Figuras 67A e 67B mostram os perfis de eluição SEC analíticos de Hercules de C-terminal com os V_H S16E + V_L S46L BHA10 scFv e Hercules de C-terminal com os V_H44:V_L100/V_H S16E + V_L S46L BHA10 scFv, respectivamente, subseqüente à etapa de purificação de proteína A inicial. Estes estudos demonstram que a estabilização do BHA10 scFv por adição de mutações VH S16E + V_L S46L somente e em combinação com o dissulfeto V_H44:V_L100 resulta em quantidades de preparativo >98% puro, anticorpo biespecífico Hercules monomérico que é essencialmente livre de espécies de peso molecular de ordem superior.

Exemplo 14. Atividade Biológica de Anticorpos "Hercules" Biespecíficos Contendo Mutações scFv Estabilizantes Não-covalentes

A atividade biológica *in vitro* de Hercules de C-terminal com os V_H S16E + V_L S46L BHA10 scFv (XWU054) e Hercules de C-terminal com os V_H44:V_L100/V_H S16E + V_L S46L BHA10 scFv (XWU055) foram testadas na prova de proliferação de células tumorais conforme descrito no Exemplo 7.

Resultados destes estudos são mostrados na Figura 68. A Figura 68B mostra que o anticorpo IgG 14A2 teve moderada atividade na inibição do crescimento de células tumorais WiDr e em combinação com IgG BHA10 apresentaram um ligeiro aumento na atividade anticélulas tumorais comparados com IgG BHA10 somente. Em contraste, ambos os anticorpos Hercules biespecíficos (XWU054 e XWU055) apresentaram aumentada destruição de células tumorais das células WiDr comparável à observada com a molécula XWU036 (Hercules de C- terminal com o BHA10 scFv estabilizado pelo ligante $V_H44:V_L100/(Gly_4Ser)_4$) descrita no Exemplo 10. A Figura 68A mostra que ambos os anticorpos 14A2 IgG e BHA10 IgG tiveram atividade desprezível na inibição do crescimento de células tumorais MDA231 como agentes únicos bem como quando usados em combinação. Em contraste, os anticorpos Hercules biespecíficos (XWU054 e XWU055) apresentaram aumentada destruição de células tumoraios das células MDA231 comparável à observada com a molécula XWU036 (Hercules de C- terminal com o BHA10 scFv estabilizado pelo ligante $V_H44:V_L100/(Gly_4Ser)_4$) descrita no Exemplo 10.

Exemplo 15: Utilização de Calorimetria de Varredura Diferencial (DSC) para Medir a Estabilidade Térmica de Seqüências de Anticorpos Humanos ou Humanizados

Para avaliar a faixa de estabilidades aparentes de anticorpos, foi realizada calorimetria de varredura diferencial (DSC) sobre uma série de 17 anticorpos BIIB humanos ou humanizados.

Os anticorpos BIIB foram dialisados exaustivamente contra um tampão de 20 mM de citrato de sódio, 150 mM de cloreto de sódio a pH 6.0. Dialisados foram usados dentro da célula de referência de um calorímetro para definir a linha basal de cada varredura de anticorpos. Subseqüente à diálise, as concentrações de BIIB1-17 foram medidas por 280 nm de absorvência (Pace et al., *Protein Sci.* 4, 2411-2423, 1995). Soluções de anticorpos usadas para DSC foram preparadas universalmente a 1 mg/mL diluindo os estoques concentrados com seus dialisados.

Foram realizadas varreduras usando um calorímetro de varredura diferencial capilar automático (capDSC, MicroCal, LLC). Soluções de pro-

teína e de referência foram amostradas automaticamente de lâminas de 96 cavidades usando a fixação robótica. Antes de cada varredura de proteína, foram realizadas 2 varreduras com tampão na célula da amostra e usadas para subtração de fundo. Uma única varredura de limpeza foi realizada usando 5% de Liquinox depois de cada varredura de proteína. Depois de cada varredura, o instrumento enxaguou automaticamente tanto as células de referência quanto de amostra com 3X 2 mL de H_2O deionizada destilada contendo 0,01% de azida de sódio. Foram realizadas varreduras em 1 ºC/min usando o modo de feedback de meio para aumento da resolução máxima. A faixa de varredura foi de 20 a 95 ºC. Todas as lâminas de 96 cavidades contendo proteína foram armazenadas dentro do instrumento a 6 ºC.

As varreduras foram analisadas usando o software Origin suprido pelo fabricante. Subseqüente à subtração do fundo, linhas basais não zero foram corrigidas usando uma terceira ordem polinomial. As transições de desdobramento de cada domínio de IgG1 foram desconvolutas ajustando as curvas multi-pico para 3 transições separadas. Domínios de IgG1 e IgG4 Fcγ recombinantes foram utilizados para identificar os picos de C_H2 e C_H3 dentro de cada transição. Presumiu-se que a uma ou mais transições restantes correspondem ao desdobramento da porção do anticorpo Fab.

A maioria das porções Fab dos 17 anticorpos desdobraram em transições únicas aparentes com pontos centrais de desdobramento térmico (T_Ms) variando de 57 a 82 °C dependendo das propriedades únicas de cada Fv individual. Somente três anticorpos, BIIB15, BIIB16, e BIIB17 apresentaram 4 transições. Um traço de DSC de IgG1 típico é mostrado na Figura 69A. A C_{Pmax} (isto é a altura de um pico de DSC) para porções Fab é geralmente significativamente mais do que a do domínios C_H2 ou C_H3. Isto é razoável considerando que a massa de proteína Fab dentro de um anticorpo é 4 dobras maior do que a de ou os domínios C_H2 ou os C_H3, e a porção Fab do anticorpo também contém área superficial enterrada entre 4 domínios ao contrário dos domínios C_H2 e C_H3 cujas transições envolvem cada a disrupção de somente uma única interface homodimérica. Em geral, as transições

observadas para os domínios C_H2 e C_H3 de cada anticorpo IgG1 se sobrepuseram bem uma sobre a outra. O mesmo é verdadeiro para os domínios C_H2 e C_H3 das 3 IgG4s. Por estas razões, as transições de Fab de cada molécula de IgG1 ou IgG4 foram prontamente identificadas. Um anticorpo (BIIB18) não foi estudado por DSC. BIIB18 expressou pobremente e nunca resultou em material solúvel / não-agregado. Este anticorpo foi incluída nas análises de seqüências descritas no Exemplo 16.

Um dos anticorpos, BIIB7, foi analisado tanto no formato de IgG1 quanto de IgG4 humanas. As aparentes termoestabilidades de Fab de BIIB7 no formato de IgG1 e IgG4 conforme medido pelso pontos centrais de suas transições de desdobramento térmico (T_M) foram muito similares (vide a Figura 69B). Variação na subclasse de domínio constante não afetou a T_M de Fab ou a entalpia do desdobramento calorimétrico de Fab.

As transições de sobreposição correspondentes aos domínios C_H2 e C_H3 do IgG4 Fc ocorrem em temperaturas significativamente menores do que as de IgG1 e fazem artificialmente as transições de Fab parecerem ~1-2 °C menores para IgG4. Uma vez que as transições de IgG sejam completamente desconvolutas, as T_Ms de Fab estão muito próximas do que parecem nos traços de DSC ($\Delta T_M < 1$ °C). O fato de que as curvas por DSC de Fab de BIIB7 nos formatos de IgG1 e IgG4 serem altamente similares sugere que a termoestabilidade dos Fabs de IgG1 e dos Fabs de IgG4 podem ser diretamente comparadas.

Medições por DSC com 17 dos anticorpos BIIB humanizados indicam que as regiões Fv influenciaram extremamente a aparente estabilidade de seus Fabs respectivos. Por exemplo, a Figura 70 demonstra as curvas de DSC muito diferentes obtidas para BIIB1, BIIB4, BIIB6 e BIIB16. Todos os quatro anticorpos são IgG1, e a únicas diferenças de seqüências entre estes se situam na região de Fv. A posição e a magnitude das transições por DSC de C_H2 e C_H3 de todos os 17 anticorpos são independentes do Fab ao qual estão fixados (vide a Figura 70). A faixa de T_Ms de Fab para os 17 anticorpos da parte de cima na Tabela 20 foi entre 57,2 e 81,6 ºC (±24,4 ºC), com BIIB1 apresentando a maior estabilidade de Fab aparente e BIIB17 a

menor.

Os três anticorpos com os menores valores de T_M de Fab mensuráveis, BIIB15, BIIB16, e BIIB17, demonstram uma disconexão entre transições de desdobramento de domínio dentro do Fab (Vide a Figura 70 para curva de DSC de BIIB16). O pico de Fab esperado foi dividido em duas transições separadas. As menores transições de T_M foram listadas na Tabela 20, mas uma segunda transição, a qual pode representar o desdobramento da unidade estrutural C_H1/C_L, ocorreu com T_Ms de 74, 72 e 70 °C para BIIB15, BIIB16 e BIIB17, respectivamente. BIIB18 foi incluído no estudo como um exemplo de um anticorpo o qual não expressariam e cuja seqüência de V_H divergiu significativamente da de consenso (vide o Exemplo 16).

As análises por DSC mostraram claramente que as únicas propriedades de regiões Fv individuais podem atenuar altamente a T_Ms aparente de uma porção Fab inteira de um anticorpo. Nos casos extremos de BIIB15, BIIB16 e BIIB17, as estabilidades de Fv aparentes foram tão baixas que as transições de desdobramento dos domínios dentro da porção Fab se tornaram desacopladas uma da outra. De modo a manter uma única transição aparente para o desdobramento de todos os quatro domínios do Fab, pode haver um limite de T_M total mínimo (>70 °C para IgG1s sob as condições de varredura descritas aqui), caso contário aparecem múltiplas transições. Seria conveniente atribuir as transições de desdobramento de baixa temperatura de BIIB15, BIIB16, e BIIB17 ao Fv e as segundas transições de maior temperatura, ao domínio C_H1/C_L. No entanto, Ewert e colaboradores mostraram que o desdobramento do domínio V_H/V_L não é sempre acoplado e pode depender da subclasse, da estabilidade V_H/V_L, e da complementaridade V_H/V_L (Ewert et al., *.J. Mol. Biol.* 325: 531-553, 2003). De fato, muito poucos de seus construtos de scFv V_H/V_L com estrutura derivados por consenso apresentaram desdobramento cooperativo. Portanto, é provável que, à medida que as transições desacoplam, o cenário de desdobramento possa ser mais complicado do que uma única divisão em duas transições representando a região Fv e as regiões C_H1/C_L.

As estabilidades aparentes conforme medido por DSC para os

Fabs de IgG1 e IgG4 de BIIB7 foram altamente similares muito embora haja diferenças inerentes substanciais em seqüência C_H1 e ligação dissulfeto as duas subclasses. Há muitas explicações possíveis para isto tais como (1) as propriedades biofísicas do Fv determinam o teto de termoestabilidade do Fab de BIIB7 independente de se a C_H1 é IgG1 ou IgG4, (2) a aparente estabilidade da região C_H1/C_L é idêntica entre IgG1 e IgG4, ou (3) a unidade estrutural C_H1/C_L não desdobra dentro do frame de tempo do experimento de DSC. O fato de que o Fab BIIB7 desdobrou em uma única T_M e com uma entalpia calorimétrica similar tanto nos formatos de IgG1 quanto de IgG4 sugere uma potencial dependência de estabilidade sobre o Fv e não as regiões constantes.

Tabela 20. Valores de T_M de Fab de IgG BIIB, escores e linhas germinais V_H e V_k.

IgG	T_M de Fab (ºC)	sub-classe V_H	linha germinal V_H[a]	escore de V_H	escore da subclasse V_H	Δesco-re	sub-classe V_k	linha germinal V_k[a]	escore de V_k	escore da subclasse V_k	Δescore
BIIB1	81,6	VH1	X92340VH1	58,8	70,1+/-3,1	+3,7	VK1	M64855VK1	78,2	80,3+/-3,5	-2,1
BIIB2	78,5	VH1	X92340VH1	72,2	70,1+/-3,1	+2,1	VK1	M64855VK1	72,8	80,3+/-3,5	-8,5
BIIB3	78,2	VH1	AB019438VH1	72,6	70,1+/-3,1	+2,5	VK2	X12684VK2	82,4	85,3+/-2,1	-2,9
BIIB4	77,7	VH3	M99649VH3	96,1	92,2+/-4,5	+3,9	VK2	X12684VK2	83,9	85,3+/-2,1	-1,4
BIIB5	77,1	VH2	X56365VH4	69,8	68,1+/-5,5	+1,7	VK1	X59316VK1	83,0	80,3+/-3,5	2,7
BIIB6	76,8	VH3	M99649VH3	97,7	92,2+/-4,5	+5,6	VK1	M64855VK1	77,4	80,3+/-3,5	-2,9
BIIB7	75,9	VH1	AB019438VH1	72,7	70,1+/-3,1	+2,6	VK3	X72812VK3	83,2	84,1+/-2,4	-0,9
BIIB8	75,6	VH1	Z14300VH1	58,0	70,1+/-3,1	+2,9	VK1	X59316VK1	79,8	80,3+/-3,5	-0,5
BIIB9	74,7	VH3	Z12358VH3	95,2	92,2+/-4,5	+3,0	VK4	Z00023VK4	88,8	92,6+/-2,6	-3,8
BIIB10	74,7	VH4	X56365VH4	69,3	68,1+/-5,5	+0,8	VK1	X59316VK1	84,0	80,3+/-3,5	3,7
BIIB11	58,1	VH1	AB019438VH1	71,3	70,1+/-3,1	+1,2	VK4	Z00023VK4	94,4	92,6+/-2,6	1,8
BIIB12	71,2	VH1	AB019438VH1	67,9	70,1+/-3,1	-2,2	VK4	Z00023VK4	92,5	92,6+/-2,6	-0,1
BIIB13	70,8	VH3	M99657VH3	91,1	92,2+/-4,5	-1,3	VK4	Z00023VK4	93,1	92,6+/-2,6	0,5
BIIB14	70,6	VH7	L10057VH7	65,1	70,2+/-3,2	-5,1	VK-	-	56,6	-	
BIIB15	68,5	VH3	M99660VH3	92,4	92,2+/-4,5	+0,2[b]	VK1	M64858VK1	81,6	80,3+/-3,5	1,3
BIIB16	68,0	VH3	M99649VH3	95,6	92,2+/-4,5	+3,3[b]	VK3	X72812VK3	82,0	84,1+/-2,4	-2,1
BIIB17	57,2	VH3	J00239VH3	89,4	92,2+/-4,5	-3,0	VK2	X63397VK2	84,8	85,3+/-2,1	-0,5
BIIB18	-[c]	VH3	M99400VH3	80,0	92,2+/-4,5	-12,3[b]	VK2	X63397VK2	84,5	85,3+/-2,1	-0,8

[a]Lefranc et al., 1999; [b]Inserção / eliminação incomum em CDR1 ou CDR2.[c]Não expressa

Exemplo 16: Utilização de Classificação de Consenso para identificar Região variável de anticorpo com estabilidade sub-ótima

Classificação de consenso foi utilizada como um método para identificar quais dos anticorpos BIIB continham números significativamente maiores de aminoácidos não-ótimos dentro de suas regiões Fv. A classificação avalia a tendência relativa de seqüências V_H e V_κ de consenso devido a mutações hipersomáticas e variações de linhas germinais evolucionárias. As medições de DSC para os 17 anticorpos foram usadas para qualificar a capacidade da abordagem de classificação de consenso para prever pobre estabilidade de anticorpos.

A série de referência de seqüências V_H e V_κ kappa mamíferas usadas para derivar a seqüência de consenso para classificação e as freqüências de aminoácidos individuais em cada posição de resíduo foram coletadas, armazenadas, e selecionadas conforme descrito previamente (Demarest et al., *J. Mol. Biol.* 335: 41-48, 2004). As séries de referência de mamíferos foram construídas de forma simples (*naively*) para incluir V-genes de vários mamíferos de modo a obter diversidade através da tendência evolucionária entre espécies. A série de referência de V_H de mamífero contém 61 seqüências V_H essencialmente de NCBI e TIGR representando um total de 17 espécies de mamíferos diferentes. A série de referência V_κ de mamífero contém 53 seqüências V_κ de 13 espécies de mamíferos diferentes.

Seqüências V_H e V_κ humanas para construir uma série de referência (ou para uso como seqüências de ensaio) foram reunidas do banco de dados NCBI usando os critérios de pesquisa de "cadeia pesada variável de anticorpo de homo sapiens" e "cadeia leve variável kappa de anticorpo de homo sapiens", respectivamente. Um total de 182 seqüências de V-gene humanas, 114 V_H e 68 V_κ, foram selecionadas semi-aleatoriamente do banco de dados NCBI para comparação contra bancos de dados de V_H e V_κ de mamíferos. A distribuição de subclasse de seqüências V_H e V_κ humanas obtidas por seleção do tipo "*coleta*" semi-aleatória do banco de dados NCBI produziu a representação de subclasse esperada baseada em uso de V_H ou V_κ de linha germinal natural (Guigou et al., 1990) sugerindo pouca preferên-

cia no modo que as seqüências foram escolhidas do NCBI.

As seqüências foram então categorizadas em suas subclasses de domínios variáveis individuais (V_H1-V_H7 e V_κ 1- V_κ 4) por alinhamento de ClustalW contra as subclasse de seqüências de consenso obtidas do website Aho's Amazing Atlas of Antibody Anatomy. As subclasses foram adicionalmente confirmadas por alinhamento contra as seqüências de linha germinal humana disponíveis publicamente (Lefranc et al., *Nucleic Acids Res.* 27, 209-212, 1999). Não mais de 5 seqüências humanas de qualquer uma submissão de NCBI multisseqüência foram incluídas dentro das séries de referência para reduzir preferência potencial introduzida por séries de iniciadores ou uso de V-gene à base de antígeno. A reunião das seqüências foi realizada singelamento (*naively*); portanto, houve uma abundância natural de seqüências VH3 e $V_\kappa1$ depois do agrupamento das subclasses. O escore médio e desvio padrão (descrito abaixo e na Tabela 20) foram calculados para cada subclasse usando 25 seqüências humanas V_H1, 3 V_H2, 44 V_H3, 34 V_H4, 5 V_H5, 3 V_H7, 29 V_κ 1, 4 V_κ 2, 19 V_κ 3 e 16 $V_\kappa4$. Somente 2 dos 18 anticorpos (BIIB5 e BIIB14) continham domínios VH com populações de seqüências de subgrupo mínimo (isto é V_H2, V_H5 ou V_H7). Quatro dos 18 anticorpos (BIIB3, BIIB4, BIIB17 e BIIB18) continham domínios variáveis da subclasse $V_\kappa2$ cuja subclasse foi subrepresentada no espaço de seqüência e cuja precisão de classificação foi mal definida. Não foram encontradas seqüências V_H6 naturais dentro da série de referência de seqüência V_H humana. Isto não foi uma preocupação uma vez que nenhum dos construtos de anticorpos BIIB continha um domínio V_H6.

As seqüências de domínio variável dos 18 anticorpos BIIB e as 182 seqüências humanas V_H e V_κ obtidas do NCBI foram classificadas contra a série de referência de seqüências de anticorpos de mamíferos. Todas as seqüências V_H e V_κ foram truncadas antes de CDR3 (dois resíduos depois dos resíduos YYC de consenso) para reduzir a imprevisibilidade introduzida pela natureza hipervariável da região de união V-(D-)J-C. Freqüências de resíduos ($p_i(r)$), em cada posição dentro do banco de dados de mamíferos foram calculadas somando o número de vezes que cada aminoácido

(A,C,D...V,W,Y,-) ocorre em cada posição de resíduo dentro do banco de dados dividido pelo número total de seqüências. Os escores foram avaliados para cada seqüência de ensaio V_H e V_K humana usando a seguinte fórmula:

$$\text{escore} = \sum_i \frac{h_i(r)}{c_i(r)}$$

em que $c_i(r)$ equivale à freqüência de resíduo de consenso em cada posição do bancos de dados de domínio variável de mamífero e $h_i(r)$ equivale à freqüência de resíduo de ensaio de cada aminoácido da seqüência de ensaio de domínio variável humano. Escores de consenso perfeito para as seqüências de consenso dos bancos de dados de mamíferos foram 104 para V_H e 100 para V_K.

A. Classificação Aleatória de Seqüências V_H e V_K Usando um Bnaco de Dados de V-Gene de Mamífero

Seqüências de consenso das subclasses V_H e V_K foram classificadas contra o consenso do bancos de dados de mamíferos e plotadas nas Figuras 71A e 71C, respectivamente. O consenso da subclasse V_H3 classificou somente 1 ponto abaixo do escore "perfeito" hipotético do consenso do bancos de dados de mamíferos e diferiu do consenso do bancos de dados de mamíferos por um total de três resíduos. Portanto, é evidente que o banco de dados de mamíferos é tendencioso para seqüências similares a V_H3. Todas as outras seqüências de consenso da subclasse V_H humanas tiveram escore significativamente menor. Uma preferência para seqüências V_K4 foi observada para o banco de dados de V_K de mamíferos (Figura 71C). Os escores das seqüências BIIB humanas ou humanizadas foram analisados com base na performance de seqüências do NCBI humanas de subclasse comparável. Distribuições dos escores de seqüências V_H individuais são mostradas na Figura 71B, e distribuições dos escores de seqüências V_K individuais são mostradas na Figura 71D.

B. Resultados de Classificação para BIIB1 a BIIB18

Escores de V_H e V_K para todos os 18 anticorpos BIIB foram calculados e comparados com os números relativos derivados da série de referência de seqüências de V-gene humano da NCBI (Tabela 20). Também in-

cluída na Tabela 20 é a diferença entre o escore do V-gene de BIIB e o escore de subclasse médio. É a diferença entre estes escores que foi usada para examinar tendências de estabilidade potenciais. BIIB1 a 18 foram marcados em ordem descendente de suas estabilidades de Fab medidas. A subclasse de V-gene e a linha germinal humana individual mais próxima para os genes BIIB V_H e V_K foram determinadas por análise ClustalW.

Escores à base de consenso foram derivados para os 18 anticorpos BIIB e comparados com os perfis de desdobramento térmico de Fab determinados experimentalmente. Anticorpos BIIB com maior diferença entre o escore de BIIB V-gene e seu escore de subclasse médio (isto é, menores escores de seqüência) foi correlacionado com a baixa T_Ms de Fab medida no Exemplo 15 (Tabela 20, Figura 72A). Não houve tendência aparente correlacionando os escores V_K com a T_Ms medida por DSC (Figura 72B). De fato, a seqüência V_K de menor classificação pertenceu a BIIB2 o qual tevew a segunda maior estabilidade de Fab aparente.

Inserções e eliminações de aminoácidos de CDR incomuns foram encontradas em BIIB15 e BIIB16, bem como em BIIB18, que teve o pior escore de V_H de todos os anticorpos BIIB e nunca foi expressado em níveis apreciáveis. A classificação de V_H não sugeriu problemas de estabilidade potenciais para BIIB15 e BIIB16; no entanto, suas T_Ms de Fab foram duas das menores testadas (Figura 72A). BIIB15 continha uma CDR1 incomumente longa (13 aminoácidos de extensão e 2 mais longo que qualquer domínio similar a V_H3 em nosso banco de dados de mamíferos ou domínio V_H3 em série de seqüências humanas) e BIIB16 continha uma CDR2 incomumente curta (15 aminoácidos de extensão e 1 aminoácido mais curto do que a maioria significativa das outras CDR2s de V_H3 tanto em nosso banco de dados de V_H de mamíferos quanto em nossa série de referência de V_Hs humanos).

<u>C. Utilidade de Bancos de Dados de Seqüências de Mamíferos Pequenas para Análise da Freqüência de Resíduos e Previsão da Estabilidade dos Fabs BIIB</u>

Classificação dos domínios VH BIIB contra o banco de dados de

seqüências VH de mamíferos se comprovou uma ferramenta útil para determinar a abundância de aminoácidos não-ótimos dentro de cada domínio V_H, os quais podem por sua vez limitar a estabilidade total de um Fab. Os escores foram úteis para selecionar Fabs com baixa estabilidade e o potencial para reduzida longevidade *in vitro* e aumentadas taxas de agregação. Com base nas preferências de subclasses observadas para os bancos de dados de V-gene de mamíferos, pode-se questionar se seria melhor usar bancos de dados de V-gene estritamente humanos para realizar a classificação de consenso em oposição ao banco de dados de mamífero usado aqui. No entanto, com base no número significativo de variáveis de determinação de estabilidade que a classificação não reflete, tais como as posições de resíduos exatos de aminoácidos não consenso, a força da interação V_H/V_L, a natureza da união V-(D-)J-C e a incidência de inserções ou eliminações hipersomáticas, é difícil acreditar que um banco de dados de V_H apenas de Rum proporcionaria uma melhora significativa. Um banco de dados estritamente humano usado para classificação também pode requerer um número significativamente maior de seqüências para combinar com a diversidade evolucionária de forma simples (*naively*) incluída dentro do banco de dados de mamíferos. Uma parte dos dados que indica que a abordagem do banco de dados de mamíferos funciona em é a das seqüências V_H de consenso de subclasse humana classificadas em ordem ascendente de sua abundância de subclasse genômica e uso de subclasse aparente (Guigou et al., 1990; Teale e Medina, 1992; Tomlinson *et al.,* 1992). O banco de dados de V_H de mamíferos parece pesar as contribuções totais das subclasses V_H humanas adequadamente. O fato de que cada consenso de seqüência das subclasses V_H e V_k geralmente classificadas acima da grande maioria das seqüências individuais e humanas únicas selecionadas entre o NCBI sugere que o banco de dados de mamíferos está capturando informação de consenso para todas as subclasses, muito embora exista uma preferência para V_H3. Portanto, devido à exisência de outros fatores influencindo a estabilidade que a classificação não incorpora, é impropável que a tendência observada na Figura 72A aumente por ajuste fino do banco de dados de seqüências V_H usa-

das para classificação.

Inserções ou eliminações incomuns em domínios VH posrecem ter efeitos significativos sobre a estabilidade não refletidos por classificação de consenso. A CDR1 incomumente longa (13 aminoácidos) de BIIB15 e a CDR2 incomumente curta (15 aminoácidos) de BIIB16 provavelmente são sobras de humanização da seqüência de camundongo original. De fato, a seqüência BIIB16 V_H alinha intimamente com duas seqüências V_H de camundongo no banco de dados de mamífero que também continha uma CDR2 de 15 aminoácidos. Forçar a CDR incomumente pequena sobre a framework humana pode ter afetado de modo adverso a estabilidade deste anticorpo. BIIB18 não somente apresentou o escore de menor V_H de todas as seqüências BIIB mas também continha uma CDR2 extraordinariamente longa (22 resíduos comparada com o máximo de 19 aminoácidos usual). Portanto, não é surpreendente considerar os múltiplos problemas possíveis inerentes dentro da seqüência de BIIB18 que nunca atingiram níveis de expressão possibilitando nada menos do que ensaios biológicos extremamente brutos com sobrenadantes não purificados. Os resultados para BIIB15, BIIB16 e BIIB18 sugere que inserções ou eliminações de V_H incomuns podem ter um maior efeito sobre a termoestabilidade do anticorpo do que a maioria das mutações de ponto único. O comportamento excepcionalmente pobre de BIIB18 não foi previsto com base em uma visão superficial sobre sua seqüência uma vez que seu Fv continha todos os aminoácidos "essenciais" estritamente conservados dentro das dobras do gene V. Ter a capacidade de selecionar anticorpos inoportunos usando um pente um tanto mais fino é a real utilidade da abordagem de classificação.

Não foram evidentes tendências ao comparar as estabilidades de Fab medidas contra os escores de V_k. Combinar os escores de V_k com os escores de V_H simplesmente reduziu a propensão observada para os escores de V_H somente. Diferentemente dos resultados de classificação de V_H, as classificações de consenso de V_k humano não estavam em linha com a abundância natural de linhas germinais V_k e uso de gene (Guigou et al., 1990; Meindl et al., 1990; Cox et al., 1994); muito embora as próprias classifica-

ções de consenso todas tenham classificado melhor do que suas seqüências de subclasses kappa humanas individuais (Figura 72B). É somente especulação que estes desvios da ordem esperada de classificações de consenso V_k podem deter a capacidade dos bancos de dados de mamíferos para prever com precisão a domínios V_k de baixa estabilidade. Além disso, não houve seqüências V_k2 suficientes para definir o escore médio da subclasse V_k2. Depois das distribuições de subclasses serem determinadas para as 68 seqüências V_k humanas, somente 4 foram V_k2 (o que reflete muito bem o uso de V_k2). Mesmo com estas desvantagens, os resultados de V_k ainda sugerem que a maior diversidade de seqüência de genes VH pode, mais freqüentemente do que não, tornar domínios VH o componente determinante de estabilidade de Fabs (Demarest et al., 2006); embora estejam disponíveis exemplos do oposto na literatura (Röthlisberger et al., 2006).

D. Dependência da Estabilidade na Linha germinal V_H ou V_L e pareamento V_H/V_L

Muitos dos 18 anticorpos continham V-genes de linhas germinais de sobreposição cujas CDRs e mutações hipersomáticas variaram e cujo pareamento V_H ou V_k foi o mesmo ou diferente. Isto possibilitou uma comparação de estabilidades de Fab de anticorpos contra as linhas germinais individuais das quais foram derivados. Os dois Fabs mais estáveis, BIIB1 e BIIB2, têm as mesmas combinações de linha germinal V_H1 e V_k1. Os pareamentos V_H e V_k não parecem ter claras preferências intrínsecas com base em um estudo de Winter e colaboradores, mas se acredita que sejam "orientados por receptor" (de Wildt et al., 1999b). BIIB1 e BIIB2 ligam antígenos relativamente dissimilares, CCL2 e VLA4; portanto, a combinação ou ocorreu aleatoriamente ou foi a conseqüência de humanização. Houve uma nítida abundância de aminoácidos não-ótimos dentro dos domínios VH dos Fabs menos estáveis, BIIB17 e BII18 (isto é baixos escores de seqüência e BIIB18 continha uma inserção incomum). Ambos estes anticorpos também continham a mesma linha germinal V_k2, sugerindo potencialmente problemas de estabilidade protéica de cadeia leve no topo dos problemas à base de V_H identificados pela classificação. Um estudo futuro interessante seria

determinar se este V_k2 particular foi um contribuinte para o comportamento biofísico pobre destes dois anticorpos. A subclasse V_k2 em general não foi associada com estabilidades de Fab pobres uma vez que BIIB3 e BIIB4 também contêm um gene da subclasse V_k2. A linha germinal AB019438V_H1 (Lefranc et al., 1999) cortado quatro vezes. Os valores da T_M para estes Fabs (BIIB12, BIIB11, BIIB7 e BIIB3) variou de 71,2 a 78,2. Interessantemente, seus escores de seqüência V_H se correlacionaram com suas estabilidades de Fab aparentes (BIIB12 < BIIB11 < BIIB7 ≈ BIIB3). Uma correlação de classificação de V_H positiva também foi vista nos anticorpos BIIB1 e BIIB2 descritos acima com a mesma linha germinal V_H1. A aparente imprecisão da classificação de V_H, no entanto, foi enfatizada por uma correlação de estabilidade negativa observada para BIIB4 e BIIB6 as quais ambas continham a linha germinal M99649V_H3.

Apesar do uso de famílias da linha germinal VH em adultos humanos parecer ser razoavelmente aleatório, a família V_H3 contém a maioria dos membros em seres humanos. Genes V_H3 ou similares a V_H3 estocasticamente apareceram mais freqüentemente na série de referência humana e no banco de dados de mamíferos. Os resultados de Ewert e Plückthun demonstraram que uma seqüência V_H3 humana derivada por consenso foi o mais estável dos domínios VH derivados por consenso (Ewert e Pluckthun, 2003). Este resultado pode não ser inteiramente inesperado considerando que a subclasse V_H3 tem mais seqüências de linha germinal para contribuir para a criação de um consenso ótimo comparada com as outras subclasses. Os dados de estabilidade mostram que IgGs contendo linhas germinais V_H3 não apresentam em geral valores de T_M de Fab maiores. Há genes das subclasses V_H3 e V_H1 tanto na parte de cima quanto na parte de baixo da lista de estabilidade. Apesar de muitos outros fatores contribuírem para a estabilidade global do Fab, especialmente as propriedades do complemento V_k dentro do Fv, pode-se esperar uma tendência para maiores estabilidades de Fab para aqueles contendo linhas germinais da subclasse V_H3 se os domínios V_H3 em geral forem mais estáveis do que as outras subclasses. Este não parece ser o caso dos estudos de estabilidade limitada descritos aqui.

Exemplo 17. Análise de Covariação de Polipeptídeos de dobra de Ig

A. Coletar e Filtrar Estructuras de dobra de Ig

Estruturas de proteínas de Dobra de Ig ou domínios de Dobra de Ig de proteínas multidomínio foram reunidas do banco de dados ASTRAL, o qual contém estruturas de domínio combinando a hierarquia SCOP. Domínios de Ig específicos de imunoglobulina foram vistos sob as seguintes classificações dentro da hierarquia SCOP: "All β proteínas" --> "β-sanduíche similar a Imunoglobulina" --> "Imunoglobulina". Sob imunoglobulina, estavam disponíveis quatro séries de estruturas: "série V," "série C1," "série C2," e "série I." Quatro scripts de download personalizado foram rodados separadamente para obter arquivos pdb de "classe V", "classe C1", "classe I", e "classe C2".

Uma vez que os arquivos da estrutura para cada subfamília foram descarregados, foi realizada alguma filtração para remover estruturas categorizadas erroneamente, incompletas, redundantes, ou com domínio permutado (Liu Y, *et al., Protein Sci.* (2002), 1285-1299). Houve 4 etapas de filtração principas, descritas abaixo:

Etapa nº *1*. Remoção de seqüências que têm rupturas

Estruturas de cada subclasse foram inspecionadas visualmente, usando o visor SwissPDB Viewer, para rupturas em seqüência (quer densidades não resolvidas ou seções faltando devido a permuta de domínio). Os arquivos PDB de estruturas imperfeitas foram removidos manualmente das séries de dados de estruturas classe V, C1, C2, e I.

Etapa nº *2*. Remoção de seqüências 100% idênticas

Estruturas de formato PDB foram convertidas para seqüências de aminoácidos de formato FASTA, e filtradas para remover quaisquer seqüências que foram ou 100% idênticas a, ou subcadeias de combinação perfeita de, seqüências restantes.

Etapa nº *3*. Remoção de seqüências de extensão aberrante

Estruturas de PDB com seqüências de aminoácidos aberrantemente longas ou curtas foram removidas das séries de dados de estruturas. Os critérios de corte de extensão foram determinados examinando o histo-

grama de todas as extensões de seqüências (Figura 74). Os histogramas pareceram um tanto distribuídos normalmente. Seqüências fora de dois desvios padrão da média foram removidas de cada série de dados de subfamília. O número médio global de resíduos para a superfamília de imunoglobulinas foi 106,10 e o desvio padrão foi 12,19. Conseqüentemente, os cortes globais usados foram <= 81 e >=131 resíduos. Uma decomposição das extensões médias e desvios padrão é mostrada na Tabela 21.

Tabela 21. Critérios de Extensão de Seqüência de Séries de Dados de Estruturas

Subclasse	Extensão Média	Desvio Padrão
C1	99,44	8,79
C2	88,07	15,14
I	95,5	4,94
V	111,95	10,62
Global	106,10	12,19

Etapa nº 4. Remoção de seqüências mal dobradas

As estruturas foram inspecionadas visualmente, usando o visor SwissPDB Viewer, para dobramento incorreto. Foram descartadas quaisquer estruturas que não não se conformam à topologia de sanduíche de duas β-folha. Como somente cinco domínios classe C2 foram obtidos da SCOP, Dobras de Ig subclasse C2 não foram perseguidas adicionalmente. Nas partes que se seguem, classe C1 é referida como classe C.

B. Obtenção de Alinhamentos de Seqüências de Alinhamentos de Estrutura

Para cada classe separada, as estruturas de Dobra de Ig foram sobrepostas uma à outra usando Combinação de Estrutura Secundária SSM (*Secondary Structure Matching*) dentro do pacote *Schrödinger structalign* (Vide a documentação do programa Schrödinger Prime para instruções sobre como criar sobreposições). Sobreposições foram realizadas em uma base 'all-to-all' ou uma base 'all-to-one'. Cada algoritmo levou a alinhamentos de qualidade similares, de modo que 'all-to-one' foi escolhido para as sobreposições. Algumas sobreposições foram corrigidas para assegurar que as regiões de núcleo de cada estrutura foram sobrepostas ao invés das alças.

As sobreposições foram então usadas para gerar alinhamentos de seqüência à base de estrutura de todas as seqüências classe V, classe I, e classe C dentro de cada alinhamento estrutural. O alinhamento de cada seqüência sobre outra foi realizado usando o pacote Schrödinger combinando aminoácidos de uma seqüência aos da segunda seqüência com base na distância mais curta entre carbonos α das espinhas dorsais do polipeptídeo.

C. Construção de um Banco de Dados de Seqüências de Dobra de Ig Curadas

Uma série de etapas definidas foram criadas para gerar uma série de dados de Dobras de Ig curadas para calcular estatísticas de covariação robusta:

Etapa nº 1. Construção de perfis de HMM

Três perfis de Modelo de Markov Oculto HMM (*Hidden Markov Model*) foram construídos, cada um baseado nos alinhamentos de seqüências à base de estrutura para uma das três classes de Dobras de Ig. Os perfis foram criados com o pacote de software HMMER (versão 1.8), usando os comandos "hmmbuild" e "hmmcalibrate" com opções padrão. Estes HMMs foram usados para detectar e alinhar seqüências de Dobras de Ig adicionais no banco de dados NR mantido na NCBI.

Etapa nº 2. Pesquisar NR usando os novos perfis de HMM classe V, I, e C.

Os três HMMs classe-específicos foram usados para pesquisar seqüências similares em um banco de dados NR local. O banco de dados NR é um arquivo grande contendo ~3 milhões de seqüências de proteínas não-redundantes. Para cada um dos HMMs classe V, I, e C, o comando de HMMER "hmmsearch" foi usado para pesquisar NR. Cada resultado classificou seqüências de NR por seus escores relativos ao HMM usado, e proporcionou informação sobre o número de regiões de alcence pelo HMM e as posições das regiões de alcence dentro de cada seqüência de NR. Para cada pesquisa HMM classe-específica de Dobra de Ig, seqüências alcançadas de NR acima de um limiar de escore do critério recomendado foram retidas como membros candidatos da classe de Dobra de Ig cujo HMM foi usado. Para estas regiões de alcence que foram subseqüências de uma seqüência

de NR, o alcence de subseqüência de NR exato foi extraído da seqüência de NR inteira usando um programa personalizado.

Etapa nº 3. Validação de atribuição de classe de Dobra de Ig usando PFAM.

PFAM é uma ferramenta de classificação de domínio e família de proteína, criada e mantida no Wellcome Trust Sanger Institute (Fin, *et al., Nuc. Acids. Res.,* (2006), 34: D247-D251), que pode ser aplicada a seqüências de proteínas individuais. Uma ferramenta Pfam 'pfamverify' foi aplicada a cada seqüência candidata de Dobra de Ig obtida na etapa 2 acima, para confirmar que foi classificada corretamente pelos HMMs de Dobra de Ig classe-específicos criados a partir dos alinhamentos de seqüências à base de estrutura. Perfis de HMM Ig-clan da PFAM (incluindo HMMs de Ig V-, I-, C1-, C2-, e menos específicas) foram baixados do website PFAM. Cada seqüência de NR foi classificada pelos HMMs Ig-clan, revelando quão bem cada seqüência se conforma a cada PFAM HMM. Seqüências cujos escores se situam abaixo de cortes recomendados (TC1 – definido no website PFAM) para as classes V, I, ou C foram removidas das séries de seqüências respectivas. Portanto, as seqüências sw Dobra de Ig candidatas classe-específicas encontradas na etapa 2 foram retidas somente se seus escores PFAM validaram suas atribuições de classe de Dobra de Ig.

Etapa nº 4. Alinhar as novas seqüências de Dobra de Ig.

As seqüências de Dobras de Ig que foram extraídas de NR com estudos de HMM e que sobreviveram à etapa 3 acima, foram em seguida alinhadas a nossos HMMs personalizados de sua classe atribuída. Uma vez que estes HMMs tinham se baseado em cuidadosos alinhamentos estruturais, este processo assegurou um alinhamento das novas seqüências orientado por estrutura. O pacote HMMER foi utilizado para gerar alinhamentos 'mapali' em formato de *fasta output*, a ser usado na Ferramenta de Covariação de Seqüência descrita abaixo. Os alinhamentos também foram fornecidos em formato de output de 'Estocolmo', e inspecionados para seqüências aberrantes ou mal alinhadas, as quais foram omitidas manualmente. Os alinhamentos resultantes, consistindo das seqüências estruturalmente-alinhadas originais e as seqüências HMM-alinhadas adicionadas de NR,

continham os seguintes números de seqüências: ~50.000 classe V; ~10.000 classe C; ~10.000 classe I.

Etapa nº 5. Remover seqüências redundantes ou similares.

Espera-se que estes alinhamentos sejam tendenciosos para seqüências freqüentemente observadas, com subrepresentação de seqüências raras e um conseqüente mascaramento da diversidade de seqüência. Por exemplo, espera-se uma grande preferência no alinhamento de seqüência de Dobra de Ig classe V para domínios variáveis de imunoglobulina murina e humana. Uma vez que a super-representação de tipos de seqüência particulares limita a utilidade de análises de covariação, uma ferramenta de filtração foi criada e aplicada para reduzir a redundância do alinhamento. A ferramento usou um novo algoritmo heurístico para encontrar seqüências com >80% de identidade uma com a outra, e removeu as mesmas do alinhamentos. Em resumo, percentagem de identidades foram calculadas para todos os pares de seqüências. Os valores de identidade foram então agrupados em grupos de percentagem de identidade (isto é grupo de 99%, grupo de 98%, grupo de 97%, e etc.). Durante a redução de cada grupo, as seqüências de cada grupo foram classificadas por contagem decrescente de resíduo não-intervalo, e em seguida por seu peso de seqüência de Henikoff (Henikoff S and Henikoff JG, *J. Mol. Biol.*, (1994), 243: 184-199). Depois desta etapa, o número de seqüências restantes em cada alinhamento foram: 2.786 classe V; 1.587 classe I; e 518 classe C.

Etapa nº 6. Remoção de intervalos para criar os alinhamentos finais

Nos alinhamentos finais, foram removidas colunas que não foram estados de combinação no perfil de HMM usado para encontrar estas seqüências (vide o manual HMMER). Portanto, enquanto muitas nas seqüências contêm mais aminoácidos do que a extensão do alinhamento final, as extensões foram truncadas paa evitar cálculos sobre as regiões de intervalo menos informativas do alinhamento. Os números finais de resíduos (incluindo intervalos) dentro do alinhamentos foram 144 para a classe V, 60 para a classe I, e 111 para a classe C.

D. Descrição geral dos alinhamentos finais

a) Classe V

Depois do filtro de 80% final ser aplicado, as 2.786 seqüências classe V podem ser divididas em três categorias: (1) Gene variável de imunoglobulina classe, 49%, as quais incluem ambos os domínios VH e V_L de diversas imunoglobulinas; (2) Genes classe V de Receptor de Células T, 16%, os quais continham os domínios hipervariáveis; e (3) Outros genes classe V, 35%. Os genes classe V de Imunoglobulina vêm de uma imensa variedade de espécies variando de peixe cartilaginoso a primatas. Houve uma preferência para seqüências classe V humanas (537 das 1272 seqüências). No entanto, outras espécies de vertebrados foram somente minimamente representadas: camundongo (33), vaca (5), camelo (23), lhama (31), símio do gênero *Macacus* (17), galinha (6), e etc. A série de seqüências de receptores de células T foi quase tão diversa variando de peixe a primata e não continha uma preferência para seqüências humanas. A "Outra" categoria compreendeu um grande número de diversas seqüências sem subcategorias reais compreendendo mais de 1 a 2% das seqüências restantes. A "Outra" categoria continha uma série ainda mais ampla de espécies incluindo cordatos (essencialmente), peixe cartilaginoso, cefalópodes, e insetos.

b) Classe C

Depois do filtro de 80% final ser aplicado, as 518 seqüências classe C podem ser divididas em três categorias: (1) domínios constantes de Imunoglobulina, 44%; (2) Dobras de Ig tipo MHC, 21%; e (3) Outra Dobras de Ig classe C, 35%. A categoria de domínio constante de Imunoglobulina continha diversas seqüências variando de peixe cartilaginoso a primatas. A subcategoria de imunoglobulina classe C não foi tendenciosa para primata ou ainda seqüências de mamífero (somente 3 seqüências humanas permaneceram não filtradas: IgE-C_H4, um C_L kappa , e um C_L lâmbda). A subcategoria MHC também variou de peixe cartilaginoso a primatas e também apresentou pouca preferência para um grupo evolucionário. A "Outra" categoria classe C continha muitas proteínas desconhecidas, uma subcategoria muito pequena de várias β-2-microglobulinas, e muitas outras proteínas que não

foram observadas com uma freqüência digna de uma subcategoria.

c) Classe I

A classe I não teve quaisquer subcategorias claras. Moléculas de titânio ou similares a titânio cortadas comumente como moléculas de adesão celular ou similares a adesão celular.

E. Cálculo da Estatística de Covariação

A força de correlação (em termos de um valor de ϕ) e significância (em termos de um valor de χ^2) de covariações entre todos os pares possíveis de aminoácidos nos alinhamentos foi calculada com base em métodos reconhecidos na técnica, mas incluíram as seguintes variações: 1) Intervalos foram incluídos como um tipo de resíduo distinto; (2) o esquema de ponderação de Henikoff não foi realizado; (3) Identidades Médias de Seqüência (SAI) não foram usadas para filtrar pares covariáveis; (4) valores de χ^2 foram calculados usando contagens à base de evento (número de ocorrências), ao invés de freqüências; e (5) pares covariantes não foram reportados pelo programa a menos que foram observados um número mínimo de vezes. Foram criadas dois conjuntos de estatística, o primeiro (e menor série) com um corte de evento de 10 e um segundo (série muito maior) com um corte de evento de 6 eventos. A capacidade para calcular estes parâmetros sobre qualquer alinhamento de seqüência dado foi codificada em um Java executável e testado com Java Runtime Engine (JRE) versão 1.4.2.

Com base no número de posições dentro do alinhamento e um total de 23 resíduos possíveis em cada posição (incluindo intervalos, resíduos ambíguos B, e Z), houve 2,4 milhões de correlações calculadas para a classe C, 700 mil calculadas paa a classe I, e 4,1 milhões calculados para a classe V. A Figura 75 mostra os valores ϕ calculados para a classe V com alta significância de χ^2. Os valores de ϕ variam de -1 a $+1$. Os valores de ϕ positivos se afastam de 0 (isto é para +1), a força de correlação (isto é a covariação) entre dois aminoácidos em posições definidas dentro dos alinhamentos aumenta. Correlações negativas (isto é, a presença de um aminoácido em um sítio no alinhamento forçando a ausência de outro aminoácido específico em um segundo sítio) aumentam enquanto os valores de ϕ se

movem para -1. Cortes para o que é uma covariação real são um tanto arbitrários, mas aqueles com um valor de $\phi > 0,5$ são fortemente correlacionados uns com os outros. Portanto, de acordo com a Figura 75, somente 370 das 2,4 milhões de correlações possíveis dentro da classe C são fortes. No entanto, correlações acima 0,3 foram reportadas como importantes e aumentariam o número de correlações positivas observadas para todas as três classes por ~10 dobras. Estas correlações representam a parte muito pequena, porém importante da série de dados que é usada para análises posteriores, gerando designs de proteína, ou previsão de posições de resíduos funcionais.

F. Validação da Estatística de Covariação

Houve vários critérios usados para validar as fortes covariações que foram derivadas do banco de dados de Dobras de Ig foram significativos e expressivos à parte das forças estatísticas com base nos cálculos. O primeiro foi para analisar se houve uma tendência entre resíduos que covariaram um com o outro e sua proximidade dentro de Dobras de Ig de estrutura conhecida. Estudos prévios na literatura demonstraram que existe uma correlação entre posições covariáveis e sua proximidade um com o outro – embora sejam muito fracas – de acordo com esses resultados. Um segundo critério viu se os cálculos de covariação de Dobra de Ig geraram conexões conhecidas entre posições de aminoácido que já tenha sido reportado, onde exista dentro de subgrupos conhecidos de Dobras de Ig. Um caso de ensaio claro foi em N-término de dobras de cadeia pesada variável de IgG humana / murina (parte da classe V). Os resíduos 6 a 10 são conhecidos por adotar conformações muito específicas com base na conservação de pares covariantes de aminoácidos (Ewert S, *et al., Methods,* (2004), 24: 184-199). Foi visto que dois dos três pares reportados de aminoácidos covariáveis têm valores de ϕ acima de 0,3 – bem acima do fundo aparente da série de dados.

Exemplo 18. Preparação de um Antcorpo Tetravalente p5E8 PRIMATIZED® Compreendendo um p5E8 scFv Convencional

p5E8G1 PRIMATIZED® é um anticorpo monoclonal de símio do

gênero *Macacus* / humano quimérico (PRIMATIZED®) contendo regiões variáveis pesadas e leves de símio do gênero *Macacus* fundidas a regiões constantes humanas γ 1 e kappa , respectivamente. p5E8G1 PRIMATIZED® liga a CD23 humano, o receptor de baixa afinidade para IgE (FcεRII) (Mavromatis and Cheson. 2003. J. Clin. Oncol. 21:1874; Pedido de patente dos Estados Unidos Nº 20030059424). Um anticorpo p5E8 PRIMATIZED® tetravalente foi construído usando uma estratégia similar àquela descrita nos Exemplos 1 a 8. O p5E8 scFv PRIMATIZED® usado para construir o anticorpo tetravalente consiste de seqüências de regiões VL e VH de p5E8 amarradas por um ligante curto na orientação VL → $(Gly_4Ser)_3$ ligante → VH e é descrito em maiores detalhes no Exemplo 19. Seqüências corretas foram confirmadas por análise de seqüência de DNA. DNA de plasmídeo foi usado para transformar células CHO DG44 para produção estável de proteína de anticorpo. A Figura 80 mostra a seqüência de DNA de anticorpo PRIMATIZED® p5E8 tetravalente de C-terminal de cadeia pesada compreendendo um scFv convencional. A Figura 81 mostra a seqüência de aminoácidos de anticorpo PRIMATIZED® p5E8 tetravalente de C-terminal de cadeia pesada compreendendo um scFv convencional. A Figura 82A mostra a seqüência de DNA de cadeia leve de p5E8 PRIMATIZED®. A Figura 82B mostra a seqüência de aminoácidos de cadeia leve de p5E8 PRIMATIZED®.

Exemplo 19. Preparação de p5E8 scFv PRIMATIZED® e Proteínas Fab

scFvs de p5E8 PRIMATIZED® em ambas as orientações (VL → $(Gly_4Ser)_3$ ligante → VH (VL/VH) e VH → $(Gly_4Ser)_3$ ligante → VL (VH/VL)) foram subclonados por amplificação por PCR de plasmídeos descritos no Pedido de Patente dos Estados Unidos Nº 20050163782. Oligonucleotídeos usados na construção são mostrados na Tabela 22.

scFvs de p5E8 PRIMATIZED (VL/VH) foi construído por PCR usando o iniciador forward P5E8-VL01F o qual contém 29 bases codificando parte da seqüência líder gpIII seguida por 15 bases de seqüência complementar ao gene de domínio variável leve de N-terminal p5E8 e o iniciador reverso, P5E8-VH01R, o qual contém 15 bases de seqüência complementar ao domínio variável pesado de C-terminal p5E8 seguido por um único sítio

de endonuclease *Sal* I adjacente (o sítio de endonuclease é sublinhado). Similarmente, scFvs de p5E8 PRIMATIZED (VH/VL) foi construído por PCR usando o iniciador forward P5E8-VH01F o qual contém 29 bases codificando parte da seqüência líder gpIII seguida por 18 bases de seqüência complementar ao gene de domínio pesado variável de N-terminal p5E8 e o iniciador reverso, P5E8-VL01R, o qual contém 12 bases de seqüência complementar ao gene de domínio leve variável de C-terminal p5E8 seguido por um único sítio de endonuclease *Sal* I adjacente (o sítio de endonuclease é sublinhado).

Tabela 22. Oligonucleotídeos para amplificação de PCR de alguns PRIMATIZED® p5E8 scFvs convencionais.

Iniciadores	Seqüência
P5E8-VL01F (SEQ ID NO: 71)	5'- CGCTGGTGGTGCCGTTCTATAGCCATAGTGAC ATCCAGATGACC -3'
P5E8-VL01R (SEQ ID NO: 72)	5'- GTGGTCGACTTTGATTTCCAC -3'
P5E8-VH01F (SEQ ID NO: 73)	5'- CGCTGGTGGTGCCGTTCTATAGCCATAGTGAGG TGCAGCTGGTGGAG -3'
P5E8-VH01R (SEQ ID NO: 74)	5'- GTGGTCGACTGAGGAGACGGTGAC -3'
P5E8-Leader01 (SEQ ID NO: 75)	5'- GGCATATGAAAAAACTGCTGTTCGCGATTCCGCTG GTGGTGCCGTTCTATAG -3'

Depois de amplificação por PCR, iniciador P5E8-Leader01 foi adicionado a ambas as reações para uma segunda reação de PCR. O iniciador P5E8-Leader01 contém um único sítio de endonuclease *Nde* I seguido por 25 bases codificando a porção N-terminal da seqüência lider gpIII, seguida por 22 bases complementares às extremidades 5' de P5E8-VL01F e P5E8-VH01F e amplificados por PCR de novo. Produtos de PCR correspondentes aos tamanhos esperados foram resolvidos por eletroforese por gel agarose, excisados, e purificados usando o kit de extração Millipore Ultrafree-DA de acordo com instruções do fabricante (Millipore; Bedford, MA). Os

produtos de PCR purificados foram em seguida digeridos com *Nde* I e *Sal* I e clonados nos sítios *Nde* I/ *Sal* I de um vetor de expressão de *E. coli* modificado designado para orientar expressão de proteína recombinante sob o controle de um promotor ara C indutível. O vetor de expressão continha uma modificação codificando um único sítio *Nde* I sobrepondo o códon de partida do BHA10 scFv. Reações de ligação individuais foram realizadas com cada um dos produtos de PCR purificados por gel e o vetor de expressão digerido e uma porção de cada uma das misturas de ligação foram usados para transformar cepa XL1-Blue de *E. coli*. Colônias resistentes a fármaco de ampicilina foram triadas e análise de seqüência de DNA confirmou a seqüência correta do p5E8 (VH/VL) final codificando pIEH-162 e p5E8 (VL/VH) codificando construtos pIEH-163. Seqüências de DNA e de aminoácidos de p5E8 (VL/VH) scFv são mostradas nas Figuras 83A e 83B, respectivamente. Seqüências de DNA e de aminoácidos de p5E8 (VH/VL) scFv são mostradas nas Figuras 84A e 84B, respectivamente.

Para expressão de scFvs de p5E8 convencionais, colônias recém isoladas de *E. coli* cepa W3110 (ATCC, Manassas, Va. Cat. nº 27325) transformadas com plasmídeos pIEH-162 e pIEH-163 foram cultivadas e foram preparadas ou sobrenadantes da cultura ou extratos de periplasma conforme descrito no Exemplo 4.

Fab p5E8 PRIMATIZED® foi preparado por digestão enzimática de IgG de p5E8 PRIMATIZED® seguindo métodos descritos no Exemplo 1. Fab purificado foi concentrado para entre 2 a 11 mg/mL. As concentrações de Fab foram determinadas usando um $\varepsilon_{280\,nm} = 1{,}5\ mL\ mg^{-1}\ cm^{-1}$.

Exemplo 20. Estabilidade Térmica de Anticorpos p5E8 scFv Convencionais

Como foi observada uma excelente correlação entre valores de T50 obtidos usando a prova de ensaio térmico descrita nos Exemplos 4 e 5 e valores de Tm calculados por análise DSC ($r^2 = 0{,}93$), uma prova de ensaio térmico foi empregada para avaliar as estabilidades térmicas relativas dos scFvs p5E8 (VL/VH) e p5E8 (VH/VL). A prova de ensaio térmico foi utilizada conforme descrito nos Exemplos 4 e 5 exceto que antígeno solúvel CD23 em 1 ug/ml foi usado para revestir as lâminas e a concentração do anticorpo

anti-6 His marcado com Eu (Perkin Elmer, Boston, MA, Cat. nº AD0109) foi aumentada para 250 ng/ml. Resultados deste ensaio determinaram que o valor de T50 de p5E8 (VL/VH) é 38 ºC e p5E8 (VH/VL) é ligeiramente menor a 34 ºC (Figura 85). Dados os valores de T50 notavelmente baixos de ambos os scFvs e a observação de que dos dois p5E8 (VL/VH) foi ligeiramente mais termicamente estável, p5E8 (VL/VH) foi selecionado para engenharia de estabilidade posterior.

Exemplo 21. Construção de Moléculas p5E8 scFv com Aprimorada Estabilidade Térmica

Variantes individuais e bibliotecas projetadas para conter as substituições de aminoácidos desejadas no p5E8 (VL/VH) scFv convencional usando os métodos descritos nos Exemplos 3 e 5 foram criadas conforme previamente descrito usando oligonucleotídeos listados na Tabela 23.

Na Tabela 23, cada nome de oligonucleotídeo dá referência a substituições de aminoácidos em uma ou mais posições em VH ou VL de acordo com o sistema de numeração de Kabat. "Fundamento lógico" se refere ao método de criação empregado. Os métodos referidos são descritos em detalhes no Exemplo 5 acima. Resíduos mutagênicos são mostrados como letras maiúsculas. Pares de oligonucleotídeos para introduzir dissulfetos VH/VL estão dentro de caixas. Abreviações são: "COMP"- Análise Computacional, "COVAR"- Análise de Covariação, "CONS" – Classificação de Consenso, Design de Interface "INTER-VH/VL", "SS"- ligação dissulfeto VH/VL, e "TBD"- a ser determinado.

Tabela 23. Oligonucleotídeos e fundamento lógico para construção de p5E8 (VL/VH) scFvs variantes.

Oligo_name	Fundamento lógico	Seqüência†	SEQ ID NO:
VH_E6Q	COMP	gaggtgcagctggtgCagtctgggggcggcttg	76
VH_L11SDG	COMP	gagtctgggggcggcRRCgcaaagcctgggggg	77
VH_A12VK_K13QER	COVAR	gtctgggggcggcttgRHGMAGcctgggggtccctg	78
VH_N32S	CONS	gttcaggttcaccttcAGCaactactacatggac	79
VH_D35bHSN	INTER	caataactactacatgMRCtgggtccgccaggctc	80

Oligo_name	Fundamento lógico	Seqüência†	SEQ ID NO:
VH_Q43KR	COVAR	cgccaggctccagggARGgggctggagtgggtc	81
VH_S49GA	COVAR	gggctggagtgggtcGSCcgtattagtagtagtg	82
VH_I51M	COMP	gagtgggtctcacgtatGagtagtagtggtgatc	83
VL_Q37A	COMP	ggtattatttaaattggtatGCCcagaaaccaggaaaag	84
VH_D55GS	CONS	gtattagtagtagtggtRGCcccacatggtacgcag	85
VH_P56STD	CONS	gtagtagtggtgatRVCacatggtacgcagac	86
VH_W58YN	INTER	gtggtgatcccacaWACtacgcagactccgtg	87
VH_E72DN	COVAR	gattcaccatctccagaRACaacgccaagaacacac	88
VH_A74S	COVAR	catctccagagagaacAGCaagaacacactgtttc	89
VH_F79YSV	COVAR	gccaagaacacactgKHCcttcaaatgaacagc	90
VH_Q81E	COMP	gaacacactgtttcttGaaatgaacagcctgagag	91
VH_A84G	COMP	caaatgaacagcctgagaGGCgaggacacggctgtc	92
VH_V89AGT	COMP	gagctgaggacacggctRSCtattactgtgcg	93
VH_S94RK	CONS	gtctattactgtgcgARGttgactacagggtctg	94
VH_V107T	COVAR	TBD	
VH_L108AGS	COMP	TBD	
VH_T110VS	COVAR	TBD	
VL_L11G	COVAR	cagtctccatcttccGGCtctgcatctgtaggg	95
VL_V15AGS	COMP	ccctgtctgcatctRSCggggacagagtcacc	96
VL_V19L	COVAR	ctgtaggggacagaCTGaccatcacttgcagg	97
VL_I21L	COVAR	ggggacagagtcaccCTGacttgcagggcaag	98
VL_T22S	COMP	gacagagtcaccatcAGCtgcagggcaagtcag	99
VL_D28SG	CONS	gcagggcaagtcagRGCattaggtattatttaaattg	100
VL_R30SL	CONS	gcaagtcaggacattMKCtattatttaaattgg	101
VL_K39A	COVAR	aattggtatcagcagGCCccaggaaaagctcc	102
VL_K45ER	INTER	ccaggaaaagctcctSRCctcctgatctatgttg	103
VL_V50n	INTER	ctaagctcctgatctatNNKgcatccagtttgcaaag	104
VL_L54R	CONS	ctatgttgcatccagtCGCcaaagtggggtccc	105

Oligo_name	Fundamento lógico	Seqüência†	SEQ ID NO:
VL_V58S	COVAR	cagtttgcaaagtgggTCCccatcaaggttcagc	106
VL_E70D	COVAR	cagtggatctgggacaGACttcactctcaccgtc	107
VL_V75I	COMP	gagttcactctcaccATCagcagcctgcagcc	108
VL_P80AG	COMP	cagcagcctgcagRGCgaagattttgcgac	109
VL_F83AGST	COMP	ctgcagcctgaagatRSCgcgacttattactg	110
VL_T85D	COVAR	cctgaagattttgcgGACtattactgtctacag	111
VL_L89AQ	INTER	gcgacttattactgtSMRcaggtttatagtacc	112
VL_R96LY	INTER	gtttatagtacccctMWAacgttcggccaaggg	113
VL_F98W	INTER	gtacccctcggacgTGGggccaagggaccaag	114
VL_I106AGS	COMP	TBD	
VH_L45C	SS	gctccagggcaggggTGCgagtgggtctcacg	115
VL_F98C	SS	gtacccctcggacgTGCggccaagggaccaag	116
VH_D101C	SS	cttgactacagggtctTGCtcctggggccagggag	117
VL_L46C	SS	ggaaaagctcctaagTGCctgatctatgttgc	118
VH_S102C	SS	gactacagggtctgacTGCtggggccagggagtc	119
VL_L46C	*SS*	ggaaaagctcctaagTGCctgatctatgttgc	120
VH_G44C	SS	caggctccagggcagTGCctggagtgggtctcac	121
VL_Q100C	SS	cctcggacgttcggcTGCgggaccaaggtggaaatc	122

†Posições alvo para mutagênese são indicadas por sublinhado. Bases ambíguas são abreviadas como se segue: W = A ou T, V = A ou C ou G, Y = C ou T, S = C ou G, M = A ou C, N = A ou C ou G ou T, R = A ou G, K = G ou T, B = C ou G ou T (J Biol Chem. 261(1):13-7 (1986)).

Colônias transformadas individuais foram selecionadas em placas de cavidades profundas de 96 cavidades, processadas, e triadas de acordo com os métodos detalhados no Exemplo 5. Transformantes foram cultivados de um dia para o outro em meio de expressão consistindo em SB (Teknova, Half Moon Bay, Ca. Cat. nº S0140) suplementado com 0,6 % de glicina, 0,6 % de Triton X100, 0,02% de arabinose, e 50 µg/ml de carbenicili-

na em ou 37°C ou 32°C.

Depois de estímulo térmico, o material agregado foi removido por centrifugação e testado no CD23 DELFIA solúvel conforme descrito no Exemplo 3. Dados de ensaio foram processados usando o software Spotfire DecisionSite (Spotfire, Somerville, Ma.) e expressados como a taxa das contagens DELFIA observadas na temperatura de estímulo para a temperatura de referência para cada clone. Clones que dão de modo reprodutível taxas maiores ou iguais a duas vezes a observada para o plasmídeo parental foram considerados alcences. DNAs de plasmídeos destes clones positivos foram isolador por mini-prep (Wizard Plus, Promega, Madison, WI) e retransformado de volta em *E. coli* W3110 para provas de ensaio térmico secundário de confirmação bem como para determinação da seqüência de DNA.

Resultados primáros e confirmatórios destes ensaios são mostrados na Tabela 24. Várias das moléculas scFv estabilizadas da invenção resultaram em aprimoramentos na atividade de ligação (T_{50} >38°C) comparadas com o p5E8 scFv convencional. Em particular, os valores de T_{50} da posição da biblioteca de p5E8 V_H6 (E6Q), posição da biblioteca V_H49 (S49G e S49A), posição da biblioteca V_H43 (Q43K), posições da biblioteca V_H72 (E72D e E72N), e posição da biblioteca V_H79 (F79S), apresentaram aumentos na estabilidade térmica variando de + 4°C a + 5°C em relação ao p5E8 scFv convencional. Os valores da T_{50} de posição da biblioteca de p5E8 V_L50 (V50D e V50S), posição da biblioteca V_L75 (V75I), posição da biblioteca V_L80 (P80S), e posições da biblioteca V_L83 (F83A, F83G, F83S e F83T), apresentaram aumentos na estabilidade térmica variando de + 3°C a + 7°C em relação ao p5E8 scFv convencional.

Além disso, os valores de T_{50} da posição da biblioteca de p5E8 V_H32 (N32S) e posição da biblioteca V_H79 (F79Y) apresentaram aumentos na estabilidade térmica de + 2°C em relação ao p5E8 scFv convencional.

Mutações estabilizantes foram identificadas usando um sortimento dos métodos de criação descritos no Exemplo 5- as mutações V_H6 (E6Q), V_L75 (V75I), V_L80 (P80S) e V_L83 (F83A, F83G, F83S e F83T) foram derivadas usando Análise Computacional, a mutação V_H32 (N32S) foi deri-

vada usando Classificação de Consenso, e as quatro mutações V_H49 (S49G e S49A), V_H43 (Q43K), V_H72 (E72D e E72N), as mutações V_H79 (F79S e F79Y) foram derivadas usando Análise de Covariação, e as mutações V_L50 (V50D e V50S) foram derivadas usando análise da interface VH/VL, validando a utilidade e a novidade destas ferramentas de criação. Combinando as mutações V_H43Q e V_H32S com mutações estabilizantes V_H49G, V_H72D, ou V_H49A reforçou a estabilidade térmica (T_{50}) de p5E8 até 53°C, um aumento de 15°C em relação ao p5E8 scFv convencional.

Tabela 24. Posições da biblioteca de p5E8 VH e VL, composição da biblioteca, e resultados da triagem.

Posição	Biblioteca	Seq. Alcence Observada	ΔT_{50} °C
VH6	Q	E6Q	+4
VH32	S	N32S	+2
VH49	S, A	S49G, S49A	+5, +5
VH43	K, R	Q43K	+4
VH72	D, N	E72D, E72N	+5, +4
VH79	S, V, Y	F79S, F79Y	+4, +2
VL50	todos os aminoácidos	V50D, V50S	+4, +3
VL75	I	V75I	+5
VL80	S, G	P80S	+4
VL83	S, A, G, T	F83S, F83A, F83G, F83T	+4, +6, +7, +6

A Tabela 25 mostra os resultados de uma análise extensiva da estabilidade térmica as várias mutações estabilizantes individuais e combinadas introduzidas em um scFv convencional. Foram identificadas mutações estabilizantes que na combinação apresentaram aumentos na estabilidade térmica variando de + 10°C a + 15°C com relação ao p5E8 scFv convencional. Estes resultados demonstram, como mostrado de modo similar no Exemplo 5, que melhoras na atividade são aditivas e que os métodos descritos nesta invenção são capazes de melhorar as propriedades de estabilidade térmica de um segundo scFv de ensaio demonstrando a utilidade geral dos métodos.

Tabela 25. Características de construtos p5E8 usadas para produzir proteínas variantes e resultados de T_{50} de prova de ensaio térmico.

Plasmídeo	Dissulfeto	Extensão de Ligante (aa)	Outra Mutação	T_{50} °C
pIEH162	não	15	na	34
pIEH163	não	15	na	38
pIEH164	não	20	na	37
pIEH165	não	20	na	-
pIEH171	não	15	VH E6Q	42
pIEH172	não	15	VH N32S	40
pIEH173	não	15	VH S49G	43
pIEH174	não	15	VH E72D	43
pIEH175	não	15	VH Q43K	42
pIEH176	não	15	VH S49A	45
pIEH177	não	15	VH E72N	42
pIEH178	não	15	VH F79S	42
pIEH179	não	15	VH F79Y	40
pIEH187	não	15	VL V75I	43
pIEH188	não	15	VL P80S	42
pIEH189	não	15	VL F83S	42
pIEH190	não	15	VL F83A	44
pIEH191	não	15	VL F83G	45
pIEH192	não	15	VL F83T	44
pIEH182	VH45-VL98	15	na	TBD
pIEH183	VH101-VL46	15	na	TBD
pIEH184	VH102-VL46	15	na	TBD
pIEH193	VH44-VL100	15	na	TBD
pIEH195	não	15	VL V50D	43
pIEH196	não	15	VL V50S	41
pIEH197	não	15	VH E6Q, S49G	51
pIEH198	não	15	VH E6Q, N32S, S49G	53
pIEH199	não	15	VH E6Q, E72N	48

Plasmídeo	Dissulfeto	Extensão de Ligante (aa)	Outra Mutação	T_{50} °C
pIEH-200	não	15	VH E6Q, N32S, E72N	50
pIEH-201	não	15	VH E6Q, N32S, E72D	50
pIEH-202	não	15	VH E6Q, E72D	48
pIEH-203	não	15	VH E6Q, N32S, S49A	51
pIEH-204	não	15	VH E6Q, S49A	48

na = não aplicável

aa = aminoácidos

TBD = a ser determinado

Exemplo 22. Produção de Anticorpos Tetravalentes p5E8 Estabilizados

p5E8 scFvs estabilizados da invenção são usados para construir anticorpos tetravalentes tanto como fusões de scFv de N- terminal quanto de C- terminal similares em configuração conforme mostrado na Figura 13.

A. Construção de Antcorpo Tetravalente N_H- p5E8

Os DNAs de p5E8 scFvs descritos no Exemplo 21 são usados para construir um anticorpo tetravalente N_H- p5E8 similar ao descrito no Exemplo 7 acima. Um ligante $(Gly_4Ser)_5$ é usado para conectar o p5E8 scFv ao término amino maduro de cadeia pesada de IgG p5E8 PRIMATIZED®. A organização da molécula é um p5E8 scFv- (Gly4Ser)5 ligante- cadeia pesada de IgG p5E8 PRIMATIZED®. Esta molécula se assemelha à cadeia leve para formar o anticorpo tetravalente. A seqüência correta seria confirmada por análise de seqüência de DNA. O construto pode ser usado para transfectar células CHO conforme descrito no Exemplo 8 para aumento em escala da produção de um Antcorpo Tetravalente N_H- p5E8. Células CHO transfectadas podem ser triadas para ligação a antígeno CD23 imobilizado por ensaio DELFIA. Lâminas de 96 cavidades, por exemplo, (MaxiSorp, Nalge Nunc, Rochester, NY, Cat. nº 437111) podem ser revestidas com antígeno CD23 solúvel a 1 ug/ml em 0,1 M de tampão de carbonato de sódio, pH 9.5. As lâminas são revestidas de um dia para o outro a 4oC, e bloqueadas com tampão de ensaio DELFIA (DAB, 10 mM de Tris HCl, 150 mM de NaCl, 20 mM de EDTA, 0,5 % de BSA, 0,02% de Tween 20, 0,01% de NaN3, pH 7.4)

por uma hora com agitação em temperatura ambiente. As lâminas são lavadas 3 vezes com DAB sem BSA (tampão de lavagem), e amostras de ensaio diluídas em DAB são adicionadas às lâminas em um volume final de 100 ul. As lâminas são incubadas por uma hora com agitação em temperatura ambiente, e em seguida lavadas 3 vezes com tampão de lavagem para remover material não ligado. Anticorpo ligado é detectado por adição de 100 ul por cavidade de DAB contendo 250 ng/ml de anticorpo anti-humano marcado com Eu (Perkin Elmer, Boston, MA, Cat. nº 1244-330) e incubado em temperatura ambiente com agitação por uma hora. As lâminas são lavadas 3 vezes com tampão de lavagem, e 100 ul de solução de reforço DELFIA (Perkin Elmer, Boston, MA, Cat. nº 4001-0010) é adicionado por cavidade. Depois de incubação por 15 minutos, as lâminas são lidas usando o método de Európio em uma leitora de lâminas Victor 2 (Perkin Elmer, Boston, MA).

B. Construção de C- p5E8 Antcorpo Tetravalente

Os DNAs de p5E8 scFvs descritos no Exemplo 21 acima são usados para construir um anticorpo tetravalente C- p5E8 conforme descrito no Exemplo 7 acima. Um peptídeo ligante $Ser(Gly_4Ser)_3$ é usado para conectar o p5E8 scFv ao término carbóxi de cadeia pesada de IgG p5E8 PRIMATIZED®. A organização da molécula é uma cadeia pesada de IgG p5E8 PRIMATIZED® - $Ser(Gly_4Ser)_3$ ligante- p5E8 scFv. Esta molécula se assemelha com a cadeia leve para formar o anticorpo tetravalente. A seqüência correta seria confirmada por análise da seqüência de DNA. O construto pode ser usada para transfectar células CHO conforme descrito no Exemplo 8 para aumento em escala da produção de um Antcorpo Tetravalente C- p5E8. Células CHO transfectadas podem ser triadas conforme descrito acima.

Equivalentes

Os versados na técnica reconhecerão, ou serão capazes de determinar usando não mais do que experimentação de rotina, muitos equivalentes das modalidades específicas da invenção descritas aqui, neste Pedido de patente. Pretende-se que os equivalentes referidos sejam englobados pelas seguintes reivindicações.

LISTAGEM DE SEQÜÊNCIAS

<110> BIOGEN IDEC MA INC.

<120> **COMPOSIÇÕES DE POLIPEPTÍDEOS ESTABILIZADOS**

<130> BGN-A242PC

<140> PCT/US2007/006883

<141> 2007-03-19

<150> 11/725.970

<151> 2007-03-19

<150> 60/783.622

<151> 2006-03-17

<150> 60/812.688

<151> 2006-06-09

<150> 60/873.502

<151> 2006-12-08

<150> 60/873.996

<151> 2006-12-08

<160> 161

<170> PatentIn version 3.3

<210> 1

<211> 28

<212> DNA

<213> Seqüência Artificial

<220>

<223> Descrição de Seqüência Artificial: Iniciador sintético

<400> 1

cagtagcatg caggtccaac tggtgcag 28

<210> 2

<211> 52

<212> DNA

<213> Seqüência Artificial

<220>

<223> Descrição de Seqüência Artificial: Iniciador sintético

<400> 2

gttctagaaa gcttttgtcg tcgtcgtctt tgatctccac cttggtaccc tg 52

<210> 3

<211> 714

<212> DNA

<213> Seqüência Artificial

<220>

<223> Descrição de Seqüência Artificial: Construto sintético

<400> 3

caggtccaac tggtgcagtc tggagctgag gtgaagaagc ctgggtcctc agtgaaggtg 60

tcctgcaagg cttctggcta cactttcaca acctactatt tgcactgggt gaggcaggcc 120

cctggacagg gacttgagtg gatgggatgg atttatcctg gaaatgttca tgctcagtac 180

aatgagaagt tcaagggcag ggtcacaatc actgcagaca aatccaccag cacagcctac 240

atggagctca gcagcctgag gtctgaagat actgcggtct attactgtgc aagatcctgg 300

gaaggttttc cttactgggg ccaagggacc acggtcaccg tctcctcagg tgggggcgga 360

tctgggggcg gcggatccgg tggtggtggt agtgacattc agatgaccca gtctcctagc 420

tccctgtccg cctcagtagg agacagggtc accatcacct gcaaggccag tcagaatgtg 480

ggtattaatg tagcctggta tcaacagaaa ccagggaagg ctcctaaatc actgatttcc 540

tcggcctcct accggtacag tggagtccct tccagattca gcggcagtgg atctgggaca 600

gatttcactc tcaccatcag cagcctccag cctgaagact tcgcaaccta tttctgtcag 660

caatatgaca cctatccatt cacgttcggc cagggtacca aggtggagat caaa 714

<210> 4

<211> 238

<212> PRT

<213> Seqüência Artificial

<220>

<223> Descrição de Seqüência Artificial: Construto sintético

<400> 4

Gln Val Gln Leu Val Gln Ser Gly Ala Glu Val Lys Lys Pro Gly Ser

1 5 10 15

Ser Val Lys Val Ser Cys Lys Ala Ser Gly Tyr Thr Phe Thr Thr Tyr
20 25 30

Tyr Leu His Trp Val Arg Gln Ala Pro Gly Gln Gly Leu Glu Trp Met
35 40 45

Gly Trp Ile Tyr Pro Gly Asn Val His Ala Gln Tyr Asn Glu Lys Phe
50 55 60

Lys Gly Arg Val Thr Ile Thr Ala Asp Lys Ser Thr Ser Thr Ala Tyr
65 70 75 80

Met Glu Leu Ser Ser Leu Arg Ser Glu Asp Thr Ala Val Tyr Tyr Cys
85 90 95

Ala Arg Ser Trp Glu Gly Phe Pro Tyr Trp Gly Gln Gly Thr Thr Val
100 105 110

Thr Val Ser Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly Gly
115 120 125

Gly Gly Ser Asp Ile Gln Met Thr Gln Ser Pro Ser Ser Leu Ser Ala
130 135 140

Ser Val Gly Asp Arg Val Thr Ile Thr Cys Lys Ala Ser Gln Asn Val
145 150 155 160

Gly Ile Asn Val Ala Trp Tyr Gln Gln Lys Pro Gly Lys Ala Pro Lys
165 170 175

Ser Leu Ile Ser Ser Ala Ser Tyr Arg Tyr Ser Gly Val Pro Ser Arg
180 185 190

Phe Ser Gly Ser Gly Ser Gly Thr Asp Phe Thr Leu Thr Ile Ser Ser
195 200 205

Leu Gln Pro Glu Asp Phe Ala Thr Tyr Phe Cys Gln Gln Tyr Asp Thr
210 215 220

Tyr Pro Phe Thr Phe Gly Gln Gly Thr Lys Val Glu Ile Lys
225 230 235

<210> 5
<211> 36
<212> DNA
<213> Seqüência Artificial
<220>
<223> Descrição de Seqüência Artificial: Iniciador sintético
<400> 5
gcaggcccct ggacagtgcc ttgagtggat gggatg 36

<210> 6
<211> 36
<212> DNA
<213> Seqüência Artificial
<220>
<223> Descrição de Seqüência Artificial: Iniciador sintético
<400> 6
catcccatcc actcaaggca ctgtccaggg gcctgc 36

<210> 7
<211> 41

<212> DNA

<213> Seqüência Artificial

<220>

<223> Descrição de Seqüência Artificial: Iniciador sintético

<400> 7

cctatccatt cacgttcggc tgcggtacca aggtggagat c 41

<210> 8

<211> 41

<212> DNA

<213> Seqüência Artificial

<220>

<223> Descrição de Seqüência Artificial: Iniciador sintético

<400> 8

gatctccacc ttggtaccgc agccgaacgt gaatggatag g 41

<210> 9

<211> 714

<212> DNA

<213> Seqüência Artificial

<220>

<223> Descrição de Seqüência Artificial: Construto sintético

<400> 9

caggtccaac tggtgcagtc tggagctgag gtgaagaagc ctgggtcctc agtgaaggtg 60

tcctgcaagg cttctggcta cactttcaca acctactatt tgcactgggt gaggcaggcc 120

cctggacagt gccttgagtg gatgggatgg atttatcctg gaaatgttca tgctcagtac 180

aatgagaagt tcaagggcag ggtcacaatc actgcagaca aatccaccag cacagcctac 240

atggagctca gcagcctgag gtctgaagat actgcggtct attactgtgc aagatcctgg 300

gaaggttttc cttactgggg ccaagggacc acggtcaccg tctcctcagg tgggggcgga 360

tctgggggcg gcggatccgg tggtggtggt agtgacattc agatgaccca gtctcctagc 420

tccctgtccg cctcagtagg agacagggtc accatcacct gcaaggccag tcagaatgtg 480

ggtattaatg tagcctggta tcaacagaaa ccagggaagg ctcctaaatc actgatttcc 540

tcggcctcct accggtacag tggagtccct tccagattca gcggcagtgg atctgggaca 600

gatttcactc tcaccatcag cagcctccag cctgaagact tcgcaaccta tttctgtcag 660

caatatgaca cctatccatt cacgttcggc tgcggtacca aggtggagat caaa 714

<210> 10

<211> 238

<212> PRT

<213> Seqüência Artificial

<220>

<223> Descrição de Seqüência Artificial: Construto sintético

<400> 10

Gln Val Gln Leu Val Gln Ser Gly Ala Glu Val Lys Lys Pro Gly Ser

1 5 10 15

Ser Val Lys Val Ser Cys Lys Ala Ser Gly Tyr Thr Phe Thr Thr Tyr

20 25 30

Tyr Leu His Trp Val Arg Gln Ala Pro Gly Gln Cys Leu Glu Trp Met

35 40 45

Gly Trp Ile Tyr Pro Gly Asn Val His Ala Gln Tyr Asn Glu Lys Phe

50 55 60

Lys Gly Arg Val Thr Ile Thr Ala Asp Lys Ser Thr Ser Thr Ala Tyr

65 70 75 80

Met Glu Leu Ser Ser Leu Arg Ser Glu Asp Thr Ala Val Tyr Tyr Cys

85 90 95

Ala Arg Ser Trp Glu Gly Phe Pro Tyr Trp Gly Gln Gly Thr Thr Val

100 105 110

Thr Val Ser Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly Gly
115 120 125

Gly Gly Ser Asp Ile Gln Met Thr Gln Ser Pro Ser Ser Leu Ser Ala
130 135 140

Ser Val Gly Asp Arg Val Thr Ile Thr Cys Lys Ala Ser Gln Asn Val
145 150 155 160

Gly Ile Asn Val Ala Trp Tyr Gln Gln Lys Pro Gly Lys Ala Pro Lys
165 170 175

Ser Leu Ile Ser Ser Ala Ser Tyr Arg Tyr Ser Gly Val Pro Ser Arg
180 185 190

Phe Ser Gly Ser Gly Ser Gly Thr Asp Phe Thr Leu Thr Ile Ser Ser
195 200 205

Leu Gln Pro Glu Asp Phe Ala Thr Tyr Phe Cys Gln Gln Tyr Asp Thr
210 215 220

Tyr Pro Phe Thr Phe Gly Cys Gly Thr Lys Val Glu Ile Lys
225 230 235

<210> 11

<211> 85

<212> DNA

<213> Seqüência Artificial

<220>

<223> Descrição de Seqüência Artificial: Iniciador sintético

<400> 11

aagggaccac ggtcaccgtc tcctcaggcg gtggagggtc cggtgggggc ggatctgggg 60

gcggcggatc cggtggtggt ggtag 85

<210> 12

<211> 27

<212> DNA

<213> Seqüência Artificial

<220>

<223> Descrição de Seqüência Artificial: Iniciador sintético

<400> 12

ttttgttcta gaaaactttt gtcgtcg 27

<210> 13

<211> 85

<212> DNA

<213> Seqüência Artificial

<220>

<223> Descrição de Seqüência Artificial: Iniciador sintético

<400> 13

aagggaccac ggtcaccgtc tcctcaggag ggggcggttc aggcggtgga gggtccggtg 60

ggggcggatc tgggggcggc ggatc 85

<210> 14

<211> 729

<212> DNA

<213> Seqüência Artificial

<220>

<223> Descrição de Seqüência Artificial: Construto sintético

<400> 14

caggtccaac tggtgcagtc tggagctgag gtgaagaagc ctgggtcctc agtgaaggtg 60

tcctgcaagg cttctggcta cactttcaca acctactatt tgcactgggt gaggcaggcc 120

cctggacagg gacttgagtg gatgggatgg atttatcctg gaaatgttca tgctcagtac 180

aatgagaagt tcaagggcag ggtcacaatc actgcagaca aatccaccag cacagcctac 240

atggagctca gcagcctgag gtctgaagat actgcggtct attactgtgc aagatcctgg 300

gaaggttttc cttactgggg ccaagggacc acggtcaccg tctcctcagg cggtggaggg 360

tccggtgggg gcggatctgg gggcggcgga tccggtggtg gtggtagtga cattcagatg 420

acccagtctc ctagctccct gtccgcctca gtaggagaca gggtcaccat cacctgcaag 480

gccagtcaga atgtgggtat taatgtagcc tggtatcaac agaaaccagg gaaggctcct 540

aaatcactga tttcctcggc ctcctaccgg tacagtggag tcccttccag attcagcggc 600

agtggatctg ggacagattt cactctcacc atcagcagcc tccagcctga agacttcgca 660

acctatttct gtcagcaata tgacacctat ccattcacgt tcggccaggg taccaaggtg 720

gagatcaaa 729

<210> 15

<211> 243

<212> PRT

<213> Seqüência Artificial

<220>

<223> Descrição de Seqüência Artificial: Construto sintético

<400> 15

Gln Val Gln Leu Val Gln Ser Gly Ala Glu Val Lys Lys Pro Gly Ser

1 5 10 15

Ser Val Lys Val Ser Cys Lys Ala Ser Gly Tyr Thr Phe Thr Thr Tyr

20 25 30

Tyr Leu His Trp Val Arg Gln Ala Pro Gly Gln Gly Leu Glu Trp Met

35 40 45

Gly Trp Ile Tyr Pro Gly Asn Val His Ala Gln Tyr Asn Glu Lys Phe

50 55 60

Lys Gly Arg Val Thr Ile Thr Ala Asp Lys Ser Thr Ser Thr Ala Tyr

65 70 75 80

Met Glu Leu Ser Ser Leu Arg Ser Glu Asp Thr Ala Val Tyr Tyr Cys
85 90 95

Ala Arg Ser Trp Glu Gly Phe Pro Tyr Trp Gly Gln Gly Thr Thr Val
100 105 110

Thr Val Ser Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly Gly
115 120 125

Gly Gly Ser Gly Gly Gly Gly Ser Asp Ile Gln Met Thr Gln Ser Pro
130 135 140

Ser Ser Leu Ser Ala Ser Val Gly Asp Arg Val Thr Ile Thr Cys Lys
145 150 155 160

Ala Ser Gln Asn Val Gly Ile Asn Val Ala Trp Tyr Gln Gln Lys Pro
165 170 175

Gly Lys Ala Pro Lys Ser Leu Ile Ser Ser Ala Ser Tyr Arg Tyr Ser
180 185 190

Gly Val Pro Ser Arg Phe Ser Gly Ser Gly Ser Gly Thr Asp Phe Thr
195 200 205

Leu Thr Ile Ser Ser Leu Gln Pro Glu Asp Phe Ala Thr Tyr Phe Cys
210 215 220

Gln Gln Tyr Asp Thr Tyr Pro Phe Thr Phe Gly Gln Gly Thr Lys Val
225 230 235 240

Glu Ile Lys

<210> 16

<211> 744

<212> DNA

<213> Seqüência Artificial

<220>

<223> Descrição de Seqüência Artificial: Construto sintético

<400> 16

caggtccaac tggtgcagtc tggagctgag gtgaagaagc ctgggtcctc agtgaaggtg 60

tcctgcaagg cttctggcta cactttcaca acctactatt tgcactgggt gaggcaggcc 120

cctggacagg gacttgagtg gatgggatgg atttatcctg gaaatgttca tgctcagtac 180

aatgagaagt tcaagggcag ggtcacaatc actgcagaca aatccaccag cacagcctac 240

atggagctca gcagcctgag gtctgaagat actgcggtct attactgtgc aagatcctgg 300

gaaggttttc cttactgggg ccaagggacc acggtcaccg tctcctcagg agggggcggt 360

tcaggcggtg gagggtccgg tgggggcgga tctgggggcg gcggatccgg tggtggtggt 420

agtgacattc agatgaccca gtctcctagc tccctgtccg cctcagtagg agacagggtc 480

accatcacct gcaaggccag tcagaatgtg ggtattaatg tagcctggta tcaacagaaa 540

ccagggaagg ctcctaaatc actgatttcc tcggcctcct accggtacag tggagtccct 600

tccagattca gcggcagtgg atctgggaca gatttcactc tcaccatcag cagcctccag 660

cctgaagact tcgcaaccta tttctgtcag caatatgaca cctatccatt cacgttcggc 720

cagggtacca aggtggagat caaa 744

<210> 17

<211> 248

<212> PRT

<213> Seqüência Artificial

<220>

<223> Descrição de Seqüência Artificial: Construto sintético

<400> 17

Gln Val Gln Leu Val Gln Ser Gly Ala Glu Val Lys Lys Pro Gly Ser

1 5 10 15

Ser Val Lys Val Ser Cys Lys Ala Ser Gly Tyr Thr Phe Thr Thr Tyr
20 25 30

Tyr Leu His Trp Val Arg Gln Ala Pro Gly Gln Gly Leu Glu Trp Met
35 40 45

Gly Trp Ile Tyr Pro Gly Asn Val His Ala Gln Tyr Asn Glu Lys Phe
50 55 60

Lys Gly Arg Val Thr Ile Thr Ala Asp Lys Ser Thr Ser Thr Ala Tyr
65 70 75 80

Met Glu Leu Ser Ser Leu Arg Ser Glu Asp Thr Ala Val Tyr Tyr Cys
85 90 95

Ala Arg Ser Trp Glu Gly Phe Pro Tyr Trp Gly Gln Gly Thr Thr Val
100 105 110

Thr Val Ser Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly Gly
115 120 125

Gly Gly Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Asp Ile Gln
130 135 140

Met Thr Gln Ser Pro Ser Ser Leu Ser Ala Ser Val Gly Asp Arg Val
145 150 155 160

Thr Ile Thr Cys Lys Ala Ser Gln Asn Val Gly Ile Asn Val Ala Trp
165 170 175

Tyr Gln Gln Lys Pro Gly Lys Ala Pro Lys Ser Leu Ile Ser Ser Ala
180 185 190

Ser Tyr Arg Tyr Ser Gly Val Pro Ser Arg Phe Ser Gly Ser Gly Ser
195 200 205

Gly Thr Asp Phe Thr Leu Thr Ile Ser Ser Leu Gln Pro Glu Asp Phe
210 215 220

Ala Thr Tyr Phe Cys Gln Gln Tyr Asp Thr Tyr Pro Phe Thr Phe Gly
225 230 235 240

Gln Gly Thr Lys Val Glu Ile Lys
245

<210> 18
<211> 729
<212> DNA
<213> Seqüência Artificial
<220>
<223> Descrição de Seqüência Artificial: Construto sintético
<400> 18
caggtccaac tggtgcagtc tggagctgag gtgaagaagc ctgggtcctc agtgaaggtg 60
tcctgcaagg cttctggcta cactttcaca acctactatt tgcactgggt gaggcaggcc 120
cctggacagt gccttgagtg gatgggatgg atttatcctg gaaatgttca tgctcagtac 180
aatgagaagt tcaagggcag ggtcacaatc actgcagaca aatccaccag cacagcctac 240
atggagctca gcagcctgag gtctgaagat actgcggtct attactgtgc aagatcctgg 300
gaaggttttc cttactgggg ccaagggacc acggtcaccg tctcctcagg cggtggaggg 360
tccggtgggg gcggatctgg gggcggcgga tccggtggtg gtggtagtga cattcagatg 420
acccagtctc ctagctccct gtccgcctca gtaggagaca gggtcaccat cacctgcaag 480
gccagtcaga atgtgggtat taatgtagcc tggtatcaac agaaaccagg gaaggctcct 540
aaatcactga tttcctcggc ctcctaccgg tacagtggag tcccttccag attcagcggc 600
agtggatctg ggacagattt cactctcacc atcagcagcc tccagcctga agacttcgca 660
acctatttct gtcagcaata tgacacctat ccattcacgt tcggctgcgg taccaaggtg 720

gagatcaaa 729

<210> 19

<211> 243

<212> PRT

<213> Seqüência Artificial

<220>

<223> Descrição de Seqüência Artificial: Construto sintético

<400> 19

Gln Val Gln Leu Val Gln Ser Gly Ala Glu Val Lys Lys Pro Gly Ser
1 5 10 15

Ser Val Lys Val Ser Cys Lys Ala Ser Gly Tyr Thr Phe Thr Thr Tyr
20 25 30

Tyr Leu His Trp Val Arg Gln Ala Pro Gly Gln Cys Leu Glu Trp Met
35 40 45

Gly Trp Ile Tyr Pro Gly Asn Val His Ala Gln Tyr Asn Glu Lys Phe
50 55 60

Lys Gly Arg Val Thr Ile Thr Ala Asp Lys Ser Thr Ser Thr Ala Tyr
65 70 75 80

Met Glu Leu Ser Ser Leu Arg Ser Glu Asp Thr Ala Val Tyr Tyr Cys
85 90 95

Ala Arg Ser Trp Glu Gly Phe Pro Tyr Trp Gly Gln Gly Thr Thr Val
100 105 110

Thr Val Ser Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly Gly
115 120 125

Gly Gly Ser Gly Gly Gly Gly Ser Asp Ile Gln Met Thr Gln Ser Pro
130 135 140

Ser Ser Leu Ser Ala Ser Val Gly Asp Arg Val Thr Ile Thr Cys Lys
145 150 155 160

Ala Ser Gln Asn Val Gly Ile Asn Val Ala Trp Tyr Gln Gln Lys Pro
165 170 175

Gly Lys Ala Pro Lys Ser Leu Ile Ser Ser Ala Ser Tyr Arg Tyr Ser
180 185 190

Gly Val Pro Ser Arg Phe Ser Gly Ser Gly Ser Gly Thr Asp Phe Thr
195 200 205

Leu Thr Ile Ser Ser Leu Gln Pro Glu Asp Phe Ala Thr Tyr Phe Cys
210 215 220

Gln Gln Tyr Asp Thr Tyr Pro Phe Thr Phe Gly Cys Gly Thr Lys Val
225 230 235 240

Glu Ile Lys

<210> 20
<211> 20
<212> PRT
<213> Seqüência Artificial
<220>
<223> Descrição de Seqüência Artificial: Peptídeo sintético
<400> 20

Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly

1 5 10 15

Gly Gly Gly Ser

20

<210> 21

<211> 15

<212> PRT

<213> Seqüência Artificial

<220>

<223> Descrição de Seqüência Artificial: Peptídeo sintético

<400> 21

Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser

1 5 10 15

<210> 22

<211> 6

<212> PRT

<213> Seqüência Artificial

<220>

<223> Descrição de Seqüência Artificial: Etiqueta 6xHis sintética

<400> 22

His His His His His His

1 5

<210> 23

<211> 25

<212> PRT

<213> Seqüência Artificial

<220>

<223> Descrição de Seqüência Artificial: Peptídeo sintético

<400> 23

Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly

1 5 10 15

Gly Gly Gly Ser Gly Gly Gly Ala Ser

20 25

<210> 24

<211> 5

<212> PRT

<213> Seqüência Artificial

<220>

<223> Descrição de Seqüência Artificial: Peptídeo sintético

<400> 24

Gly Gly Gly Gly Ser

1 5

<210> 25

<211> 54

<212> DNA

<213> Seqüência Artificial

<220>

<223> Descrição de Seqüência Artificial: Oligonucleotídeo sintético

<400> 25

atgaaaaaac tgctgttcgc gattccgctg gtggtgccgt tctatagcca tagt 54

<210> 26

<211> 18

<212> PRT

<213> Seqüência Artificial

<220>

<223> Descrição de Seqüência Artificial: Peptídeo sintético

<400> 26

Met Lys Lys Leu Leu Phe Ala Ile Pro Leu Val Val Pro Phe Tyr Ser

1 5 10 15

His Ser

<210> 27

<211> 90

<212> DNA

<213> Seqüência Artificial

<220>

<223> Descrição de Seqüência Artificial: Oligonucleotídeo sintético

<400> 27

gacgacgacg acaaaagctt tctagaacaa aaactcatct cagaagagga tctgaatagc 60

gccgtcgacc atcatcatca tcatcattga 90

<210> 28

<211> 669

<212> DNA

<213> Seqüência Artificial

<220>

<223> Descrição de Seqüência Artificial: Construto sintético

<400> 28

caacttgtgc tcactcagtc atcttcagtc tctttctccc tgggagcctc agcaaaactc 60

acgtgcacct tgagtagtca gcacagtacg tacaccattg aatggtatca gcaacagccc 120

ctcaagcctc ctaagtatgt gatggagctt aagaaagatg gaagccacag cacaggtgat 180

gggattcctg atcgcttctc tggatccagc tctggtgctg atcgctacct tagcatttcc 240

aacatccagc ctgaagatga agcaatatac atctgtggtg tgggtgatac aattaaggaa 300

caatttgtgt atgttttcgg cggtggaacc aaggtcgaaa tcaaacgtac ggtggctgca 360

ccatctgtct tcatcttccc gccatctgat gagcagttga aatctggaac tgcctctgtt 420

gtgtgcctgc tgaataactt ctatcccaga gaggccaaag tacagtggaa ggtggataac 480

gccctccaat cgggtaactc ccaggagagt gtcacagagc aggacagcaa ggacagcacc 540

tacagcctca gcagcaccct gacgctgagc aaagcagact acgagaaaca caaagtctac 600

gcctgcgaag tcacccatca gggcctgagc tcgcccgtca caaagagctt caacagggga 660

gagtgttga 669

<210> 29

<211> 222

<212> PRT

<213> Seqüência Artificial

<220>

<223> Descrição de Seqüência Artificial: Construto sintético

<400> 29

Gln Leu Val Leu Thr Gln Ser Ser Ser Val Ser Phe Ser Leu Gly Ala

1 5 10 15

Ser Ala Lys Leu Thr Cys Thr Leu Ser Ser Gln His Ser Thr Tyr Thr

20 25 30

Ile Glu Trp Tyr Gln Gln Gln Pro Leu Lys Pro Pro Lys Tyr Val Met

35 40 45

Glu Leu Lys Lys Asp Gly Ser His Ser Thr Gly Asp Gly Ile Pro Asp

50 55 60

Arg Phe Ser Gly Ser Ser Ser Gly Ala Asp Arg Tyr Leu Ser Ile Ser

65 70 75 80

Asn Ile Gln Pro Glu Asp Glu Ala Ile Tyr Ile Cys Gly Val Gly Asp

85 90 95

Thr Ile Lys Glu Gln Phe Val Tyr Val Phe Gly Gly Gly Thr Lys Val

100 105 110

Glu Ile Lys Arg Thr Val Ala Ala Pro Ser Val Phe Ile Phe Pro Pro
115 120 125

Ser Asp Glu Gln Leu Lys Ser Gly Thr Ala Ser Val Val Cys Leu Leu
130 135 140

Asn Asn Phe Tyr Pro Arg Glu Ala Lys Val Gln Trp Lys Val Asp Asn
145 150 155 160

Ala Leu Gln Ser Gly Asn Ser Gln Glu Ser Val Thr Glu Gln Asp Ser
165 170 175

Lys Asp Ser Thr Tyr Ser Leu Ser Ser Thr Leu Thr Leu Ser Lys Ala
180 185 190

Asp Tyr Glu Lys His Lys Val Tyr Ala Cys Glu Val Thr His Gln Gly
195 200 205

Leu Ser Ser Pro Val Thr Lys Ser Phe Asn Arg Gly Glu Cys
210 215 220

<210> 30
<211> 2139
<212> DNA
<213> Seqüência Artificial
<220>
<223> Descrição de Seqüência Artificial: Construto sintético
<400> 30
caggtccaac tggtgcagtc tggagctgag gtgaagaagc ctgggtcctc agtgaaggtg 60
tcctgcaagg cttctggcta cactttcaca acctactatt tgcactgggt gaggcaggcc 120
cctggacagg gacttgagtg gatgggatgg atttatcctg gaaatgttca tgctcagtac 180

aatgagaagt tcaagggcag ggtcacaatc actgcagaca aatccaccag cacagcctac 240

atggagctca gcagcctgag gtctgaagat actgcggtct attactgtgc aagatcctgg 300

gaaggttttc cttactgggg ccaagggacc acggtcaccg tctcctcagg tgggggcgga 360

tctgggggcg gcggatccgg tggtggtggt agtgacattc agatgaccca gtctcctagc 420

tccctgtccg cctcagtagg agacagggtc accatcacct gcaaggccag tcagaatgtg 480

ggtattaatg tagcctggta tcaacagaaa ccagggaagg ctcctaaatc actgatttcc 540

tcggcctcct accggtacag tggagtccct tccagattca gcggcagtgg atctgggaca 600

gatttcactc tcaccatcag cagcctccag cctgaagact tcgcaaccta tttctgtcag 660

caatatgaca cctatccatt cacgttcggc cagggtacca aggtggagat caaaggcggt 720

ggagggtccg gtggaggggg ctctggaggg ggcggttcag ggggcggtgg atcgggggga 780

ggtggctccc agatccagtt ggtgcagtct ggacctgagc tgaagaagcc tggagagaca 840

gtcaagatct cctgcaaggc ttctggtttt accttcacag actattcaat acactgggtg 900

aaacaggctc caggaaaggg tttaaagtgg atgggctgga taaacactga gactggtgag 960

ccaacatata cagatgactt caagggacga tttgccttct ctttggtgac ctctgccacc 1020

actgcctatt tgcagatcaa caacctcaac aatgaggaca cggctacatt tttctgtgct 1080

agattcatct atgatcctta ttgggggttt gcttactggg gccaggggac tctggtcact 1140

gtctccgcag ctagcaccaa gggcccatcg gtcttccccc tggcaccctc ctccaagagc 1200

acctctgggg gcacagcggc cctgggctgc ctggtcaagg actacttccc cgaaccggtg 1260

acggtgtcgt ggaactcagg cgccctgacc agcggcgtgc acaccttccc ggctgtccta 1320

cagtcctcag gactctactc cctcagcagc gtggtgaccg tgccctccag cagcttgggc 1380

acccagacct acatctgcaa cgtgaatcac aagcccagca acaccaaggt ggacaagaaa 1440

gttgagccca aatcttgtga caaaactcac acatgcccac cgtgcccagc acctgaactc 1500

ctgggggggac cgtcagtctt cctcttcccc ccaaaaccca aggacaccct catgatctcc 1560

cggacccctg aggtcacatg cgtggtggtg gacgtgagcc acgaagaccc tgaggtcaag 1620

ttcaactggt acgtggacgg cgtggaggtg cataatgcca agacaaagcc gcgggaggag 1680

cagtacaaca gcacgtaccg tgtggtcagc gtcctcaccg tcctgcacca ggactggctg 1740

aatggcaagg agtacaagtg caaggtctcc aacaaagccc tcccagcccc catcgagaaa 1800

accatctcca aagccaaagg gcagccccga gaaccacagg tgtacaccct gcccccatcc 1860

cgggatgagc tgaccaagaa ccaggtcagc ctgacctgcc tggtcaaagg cttctatccc 1920

agcgacatcg ccgtggagtg ggagagcaat gggcagccgg agaacaacta caagaccacg 1980

cctcccgtgc tggactccga cggctccttc ttcctctaca gcaagctcac cgtggacaag 2040

agcaggtggc agcaggggaa cgtcttctca tgctccgtga tgcatgaggc tctgcacaac 2100

cactacacgc agaagagcct ctccctgtct ccgggttga 2139

<210> 31

<211> 712

<212> PRT

<213> Seqüência Artificial

<220>

<223> Descrição de Seqüência Artificial: Construto sintético

<400> 31

Gln Val Gln Leu Val Gln Ser Gly Ala Glu Val Lys Lys Pro Gly Ser
1 5 10 15

Ser Val Lys Val Ser Cys Lys Ala Ser Gly Tyr Thr Phe Thr Thr Tyr
20 25 30

Tyr Leu His Trp Val Arg Gln Ala Pro Gly Gln Gly Leu Glu Trp Met
35 40 45

Gly Trp Ile Tyr Pro Gly Asn Val His Ala Gln Tyr Asn Glu Lys Phe
50 55 60

Lys Gly Arg Val Thr Ile Thr Ala Asp Lys Ser Thr Ser Thr Ala Tyr
65 70 75 80

Met Glu Leu Ser Ser Leu Arg Ser Glu Asp Thr Ala Val Tyr Tyr Cys
85 90 95

Ala Arg Ser Trp Glu Gly Phe Pro Tyr Trp Gly Gln Gly Thr Thr Val
100 105 110

Thr Val Ser Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly Gly
115 120 125

Gly Gly Ser Asp Ile Gln Met Thr Gln Ser Pro Ser Ser Leu Ser Ala
130 135 140

Ser Val Gly Asp Arg Val Thr Ile Thr Cys Lys Ala Ser Gln Asn Val
145 150 155 160

Gly Ile Asn Val Ala Trp Tyr Gln Gln Lys Pro Gly Lys Ala Pro Lys
165 170 175

Ser Leu Ile Ser Ser Ala Ser Tyr Arg Tyr Ser Gly Val Pro Ser Arg
180 185 190

Phe Ser Gly Ser Gly Ser Gly Thr Asp Phe Thr Leu Thr Ile Ser Ser
195 200 205

Leu Gln Pro Glu Asp Phe Ala Thr Tyr Phe Cys Gln Gln Tyr Asp Thr
210 215 220

Tyr Pro Phe Thr Phe Gly Gln Gly Thr Lys Val Glu Ile Lys Gly Gly
225 230 235 240

Gly Gly Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly Gly Gly
245 250 255

Gly Ser Gly Gly Gly Gly Ser Gln Ile Gln Leu Val Gln Ser Gly Pro
260 265 270

Glu Leu Lys Lys Pro Gly Glu Thr Val Lys Ile Ser Cys Lys Ala Ser
275 280 285

Gly Phe Thr Phe Thr Asp Tyr Ser Ile His Trp Val Lys Gln Ala Pro

290 295 300

Gly Lys Gly Leu Lys Trp Met Gly Trp Ile Asn Thr Glu Thr Gly Glu
305 310 315 320

Pro Thr Tyr Thr Asp Asp Phe Lys Gly Arg Phe Ala Phe Ser Leu Val
325 330 335

Thr Ser Ala Thr Thr Ala Tyr Leu Gln Ile Asn Asn Leu Asn Asn Glu
340 345 350

Asp Thr Ala Thr Phe Phe Cys Ala Arg Phe Ile Tyr Asp Pro Tyr Trp
355 360 365

Gly Phe Ala Tyr Trp Gly Gln Gly Thr Leu Val Thr Val Ser Ala Ala
370 375 380

Ser Thr Lys Gly Pro Ser Val Phe Pro Leu Ala Pro Ser Ser Lys Ser
385 390 395 400

Thr Ser Gly Gly Thr Ala Ala Leu Gly Cys Leu Val Lys Asp Tyr Phe
405 410 415

Pro Glu Pro Val Thr Val Ser Trp Asn Ser Gly Ala Leu Thr Ser Gly
420 425 430

Val His Thr Phe Pro Ala Val Leu Gln Ser Ser Gly Leu Tyr Ser Leu
435 440 445

Ser Ser Val Val Thr Val Pro Ser Ser Ser Leu Gly Thr Gln Thr Tyr
450 455 460

Ile Cys Asn Val Asn His Lys Pro Ser Asn Thr Lys Val Asp Lys Lys
465 470 475 480

Val Glu Pro Lys Ser Cys Asp Lys Thr His Thr Cys Pro Pro Cys Pro
485 490 495

Ala Pro Glu Leu Leu Gly Gly Pro Ser Val Phe Leu Phe Pro Pro Lys
500 505 510

Pro Lys Asp Thr Leu Met Ile Ser Arg Thr Pro Glu Val Thr Cys Val
515 520 525

Val Val Asp Val Ser His Glu Asp Pro Glu Val Lys Phe Asn Trp Tyr
530 535 540

Val Asp Gly Val Glu Val His Asn Ala Lys Thr Lys Pro Arg Glu Glu
545 550 555 560

Gln Tyr Asn Ser Thr Tyr Arg Val Val Ser Val Leu Thr Val Leu His
565 570 575

Gln Asp Trp Leu Asn Gly Lys Glu Tyr Lys Cys Lys Val Ser Asn Lys
580 585 590

Ala Leu Pro Ala Pro Ile Glu Lys Thr Ile Ser Lys Ala Lys Gly Gln
595 600 605

Pro Arg Glu Pro Gln Val Tyr Thr Leu Pro Pro Ser Arg Asp Glu Leu
610 615 620

Thr Lys Asn Gln Val Ser Leu Thr Cys Leu Val Lys Gly Phe Tyr Pro
625 630 635 640

Ser Asp Ile Ala Val Glu Trp Glu Ser Asn Gly Gln Pro Glu Asn Asn
645 650 655

Tyr Lys Thr Thr Pro Pro Val Leu Asp Ser Asp Gly Ser Phe Phe Leu
660 665 670

Tyr Ser Lys Leu Thr Val Asp Lys Ser Arg Trp Gln Gln Gly Asn Val
675 680 685

Phe Ser Cys Ser Val Met His Glu Ala Leu His Asn His Tyr Thr Gln
690 695 700

Lys Ser Leu Ser Leu Ser Pro Gly
705 710

<210> 32
<211> 2154
<212> DNA
<213> Seqüência Artificial
<220>
<223> Descrição de Seqüência Artificial: Construto sintético
<400> 32
caggtccaac tggtgcagtc tggagctgag gtgaagaagc ctgggtcctc agtgaaggtg 60
tcctgcaagg cttctggcta cactttcaca acctactatt tgcactgggt gaggcaggcc 120
cctggacagg gacttgagtg gatgggatgg atttatcctg gaaatgttca tgctcagtac 180
aatgagaagt tcaagggcag ggtcacaatc actgcagaca aatccaccag cacagcctac 240
atggagctca gcagcctgag gtctgaagat actgcggtct attactgtgc aagatcctgg 300
gaaggttttc cttactgggg ccaagggacc acggtcaccg tctcctcagg cggtggaggg 360
tccggtgggg gcggatctgg gggcggcgga tccggtggtg gtggtagtga cattcagatg 420
acccagtctc ctagctccct gtccgcctca gtaggagaca gggtcaccat cacctgcaag 480
gccagtcaga atgtgggtat taatgtagcc tggtatcaac agaaaccagg gaaggctcct 540

aaatcactga tttcctcggc ctcctaccgg tacagtggag tcccttccag attcagcggc 600

agtggatctg ggacagattt cactctcacc atcagcagcc tccagcctga agacttcgca 660

acctatttct gtcagcaata tgacacctat ccattcacgt tcggccaggg taccaaggtg 720

gagatcaaag gcggtggagg gtccggtgga gggggctctg gaggggggcgg ttcaggggggc 780

ggtggatcgg ggggaggtgg ctcccagatc cagttggtgc agtctggacc tgagctgaag 840

aagcctggag agacagtcaa gatctcctgc aaggcttctg gttttacctt cacagactat 900

tcaatacact gggtgaaaca ggctccagga aagggtttaa agtggatggg ctggataaac 960

actgagactg gtgagccaac atatacagat gacttcaagg gacgatttgc cttctctttg 1020

gtgacctctg ccaccactgc ctatttgcag atcaacaacc tcaacaatga ggacacggct 1080

acatttttct gtgctagatt catctatgat ccttattggg ggtttgctta ctggggccag 1140

gggactctgg tcactgtctc cgcagctagc accaagggcc catcggtctt ccccctggca 1200

ccctcctcca agagcacctc tgggggcaca gcggccctgg gctgcctggt caaggactac 1260

ttccccgaac cggtgacggt gtcgtggaac tcaggcgccc tgaccagcgg cgtgcacacc 1320

ttcccggctg tcctacagtc ctcaggactc tactccctca gcagcgtggt gaccgtgccc 1380

tccagcagct tgggcaccca gacctacatc tgcaacgtga atcacaagcc cagcaacacc 1440

aaggtggaca agaaagttga gcccaaatct tgtgacaaaa ctcacacatg cccaccgtgc 1500

ccagcacctg aactcctggg gggaccgtca gtcttcctct tccccccaaa acccaaggac 1560

accctcatga tctcccggac ccctgaggtc acatgcgtgg tggtggacgt gagccacgaa 1620

gaccctgagg tcaagttcaa ctggtacgtg gacggcgtgg aggtgcataa tgccaagaca 1680

aagccgcggg aggagcagta caacagcacg taccgtgtgg tcagcgtcct caccgtcctg 1740

caccaggact ggctgaatgg caaggagtac aagtgcaagg tctccaacaa agccctccca 1800

gcccccatcg agaaaaccat ctccaaagcc aaagggcagc cccgagaacc acaggtgtac 1860

accctgcccc catcccggga tgagctgacc aagaaccagg tcagcctgac ctgcctggtc 1920

aaaggcttct atcccagcga catcgccgtg gagtgggaga gcaatgggca gccggagaac 1980

aactacaaga ccacgcctcc cgtgctggac tccgacggct ccttcttcct ctacagcaag 2040

ctcaccgtgg acaagagcag gtggcagcag gggaacgtct tctcatgctc cgtgatgcat 2100

gaggctctgc acaaccacta cacgcagaag agcctctccc tgtctccggg ttga 2154

<210> 33

<211> 717

<212> PRT

<213> Seqüência Artificial

<220>

<223> Descrição de Seqüência Artificial: Construto sintético

<400> 33

Gln Val Gln Leu Val Gln Ser Gly Ala Glu Val Lys Lys Pro Gly Ser
1 5 10 15

Ser Val Lys Val Ser Cys Lys Ala Ser Gly Tyr Thr Phe Thr Thr Tyr
20 25 30

Tyr Leu His Trp Val Arg Gln Ala Pro Gly Gln Gly Leu Glu Trp Met
35 40 45

Gly Trp Ile Tyr Pro Gly Asn Val His Ala Gln Tyr Asn Glu Lys Phe
50 55 60

Lys Gly Arg Val Thr Ile Thr Ala Asp Lys Ser Thr Ser Thr Ala Tyr
65 70 75 80

Met Glu Leu Ser Ser Leu Arg Ser Glu Asp Thr Ala Val Tyr Tyr Cys
85 90 95

Ala Arg Ser Trp Glu Gly Phe Pro Tyr Trp Gly Gln Gly Thr Thr Val
100 105 110

Thr Val Ser Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly Gly
115 120 125

Gly Gly Ser Gly Gly Gly Gly Ser Asp Ile Gln Met Thr Gln Ser Pro
130 135 140

Ser Ser Leu Ser Ala Ser Val Gly Asp Arg Val Thr Ile Thr Cys Lys
145 150 155 160

Ala Ser Gln Asn Val Gly Ile Asn Val Ala Trp Tyr Gln Gln Lys Pro
165 170 175

Gly Lys Ala Pro Lys Ser Leu Ile Ser Ser Ala Ser Tyr Arg Tyr Ser
180 185 190

Gly Val Pro Ser Arg Phe Ser Gly Ser Gly Ser Gly Thr Asp Phe Thr
195 200 205

Leu Thr Ile Ser Ser Leu Gln Pro Glu Asp Phe Ala Thr Tyr Phe Cys
210 215 220

Gln Gln Tyr Asp Thr Tyr Pro Phe Thr Phe Gly Gln Gly Thr Lys Val
225 230 235 240

Glu Ile Lys Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly Gly Gly
245 250 255

Gly Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Gln Ile Gln Leu
260 265 270

Val Gln Ser Gly Pro Glu Leu Lys Lys Pro Gly Glu Thr Val Lys Ile
275 280 285

Ser Cys Lys Ala Ser Gly Phe Thr Phe Thr Asp Tyr Ser Ile His Trp
290 295 300

Val Lys Gln Ala Pro Gly Lys Gly Leu Lys Trp Met Gly Trp Ile Asn
305 310 315 320

Thr Glu Thr Gly Glu Pro Thr Tyr Thr Asp Asp Phe Lys Gly Arg Phe

325 330 335

Ala Phe Ser Leu Val Thr Ser Ala Thr Thr Ala Tyr Leu Gln Ile Asn
340 345 350

Asn Leu Asn Asn Glu Asp Thr Ala Thr Phe Phe Cys Ala Arg Phe Ile
355 360 365

Tyr Asp Pro Tyr Trp Gly Phe Ala Tyr Trp Gly Gln Gly Thr Leu Val
370 375 380

Thr Val Ser Ala Ala Ser Thr Lys Gly Pro Ser Val Phe Pro Leu Ala
385 390 395 400

Pro Ser Ser Lys Ser Thr Ser Gly Gly Thr Ala Ala Leu Gly Cys Leu
405 410 415

Val Lys Asp Tyr Phe Pro Glu Pro Val Thr Val Ser Trp Asn Ser Gly
420 425 430

Ala Leu Thr Ser Gly Val His Thr Phe Pro Ala Val Leu Gln Ser Ser
435 440 445

Gly Leu Tyr Ser Leu Ser Ser Val Val Thr Val Pro Ser Ser Ser Leu
450 455 460

Gly Thr Gln Thr Tyr Ile Cys Asn Val Asn His Lys Pro Ser Asn Thr
465 470 475 480

Lys Val Asp Lys Lys Val Glu Pro Lys Ser Cys Asp Lys Thr His Thr
485 490 495

Cys Pro Pro Cys Pro Ala Pro Glu Leu Leu Gly Gly Pro Ser Val Phe
500 505 510

Leu Phe Pro Pro Lys Pro Lys Asp Thr Leu Met Ile Ser Arg Thr Pro
515 520 525

Glu Val Thr Cys Val Val Val Asp Val Ser His Glu Asp Pro Glu Val
530 535 540

Lys Phe Asn Trp Tyr Val Asp Gly Val Glu Val His Asn Ala Lys Thr
545 550 555 560

Lys Pro Arg Glu Glu Gln Tyr Asn Ser Thr Tyr Arg Val Val Ser Val
565 570 575

Leu Thr Val Leu His Gln Asp Trp Leu Asn Gly Lys Glu Tyr Lys Cys
580 585 590

Lys Val Ser Asn Lys Ala Leu Pro Ala Pro Ile Glu Lys Thr Ile Ser
595 600 605

Lys Ala Lys Gly Gln Pro Arg Glu Pro Gln Val Tyr Thr Leu Pro Pro
610 615 620

Ser Arg Asp Glu Leu Thr Lys Asn Gln Val Ser Leu Thr Cys Leu Val
625 630 635 640

Lys Gly Phe Tyr Pro Ser Asp Ile Ala Val Glu Trp Glu Ser Asn Gly
645 650 655

Gln Pro Glu Asn Asn Tyr Lys Thr Thr Pro Pro Val Leu Asp Ser Asp
660 665 670

Gly Ser Phe Phe Leu Tyr Ser Lys Leu Thr Val Asp Lys Ser Arg Trp
675 680 685

Gln Gln Gly Asn Val Phe Ser Cys Ser Val Met His Glu Ala Leu His
690 695 700

Asn His Tyr Thr Gln Lys Ser Leu Ser Leu Ser Pro Gly
705 710 715

<210> 34
<211> 1773
<212> DNA
<213> Seqüência Artificial
<220>
<223> Descrição de Seqüência Artificial: Construto sintético
<400> 34

caggtccaac tggtgcagtc tggagctgag gtgaagaagc ctgggtcctc agtgaaggtg 60
tcctgcaagg cttctggcta cactttcaca acctactatt tgcactgggt gaggcaggcc 120
cctggacagt gccttgagtg gatgggatgg atttatcctg gaaatgttca tgctcagtac 180
aatgagaagt tcaagggcag ggtcacaatc actgcagaca aatccaccag cacagcctac 240
atggagctca gcagcctgag gtctgaagat actgcggtct attactgtgc aagatcctgg 300
gaaggttttc cttactgggg ccaagggacc acggtcaccg tctcctcagg tgggggcgga 360
tctgggggcg gcggatccgg tggtggtggt agtgacattc agatgaccca gtctcctagc 420
tccctgtccg cctcagtagg agacagggtc accatcacct gcaaggccag tcagaatgtg 480
ggtattaatg tagcctggta tcaacagaaa ccagggaagg ctcctaaatc actgatttcc 540
tcggcctcct accggtacag tggagtccct tccagattca gcggcagtgg atctgggaca 600
gatttcactc tcaccatcag cagcctccag cctgaagact tcgcaaccta tttctgtcag 660
caatatgaca cctatccatt cacgttcggc tgcggtacca aggtggagat caaaggcggt 720
ggagggtccg gtggaggggg ctctggaggg ggcggttcag ggggcggtgg atcgggggga 780
ggtggctccc agatccagtt ggtgcagtct ggacctgagc tgaagaagcc tggagagaca 840
gtcaagatct cctgcaaggc ttctggtttt accttcacag actattcaat acactgggtg 900

aaacaggctc caggaaaggg tttaaagtgg atgggctgga taaacactga gactggtgag 960

ccaacatata cagatgactt caagggacga tttgccttct ctttggtgac ctctgccacc 1020

actgcctatt tgcagatcaa caacctcaac aatgaggaca cggctacatt tttctgtgct 1080

agattcatct atgatcctta ttgggggttt gcttactggg gccaggggac tctggtcact 1140

gtctccgcag ctagcaccaa gggcccatcg gtcttccccc tggcaccctc ctccaagagc 1200

acctctgggg gcacagcggc cctgggctgc ctggtcaagg actacttccc cgaaccggtg 1260

acggtgtcgt ggaactcagg cgccctgacc agcggcgtgc acaccttccc ggctgtccta 1320

cagtcctcag gactctactc cctcagcagc gtggtgaccg tgccctccag cagcttgggc 1380

acccagacct acatctgcaa cgtgaatcac aagcccagca acaccaaggt ggacaagaaa 1440

gttgagccca aatcttgtga caaaactcac acatgcccac cgtgcccagc acctgaactc 1500

ctgggggggac cgtcagtctt cctcttcccc ccaaaaccca aggacaccct catgatctcc 1560

cggacccctg aggtcacatg cgtggtggtg gacgtgagcc acgaagaccc tgaggtcaag 1620

ttcaactggt acgtggacgg cgtggaggtg cataatgcca agacaaagcc gcgggaggag 1680

cagtacaaca gcacgtaccg tgtggtcagc gtcctcaccg tcctgcacca ggactggctg 1740

aatggcaagg agtacaagtg caaggtctcc aac 1773

<210> 35

<211> 712

<212> PRT

<213> Seqüência Artificial

<220>

<223> Descrição de Seqüência Artificial: Construto sintético

<400> 35

Gln Val Gln Leu Val Gln Ser Gly Ala Glu Val Lys Lys Pro Gly Ser

1 5 10 15

Ser Val Lys Val Ser Cys Lys Ala Ser Gly Tyr Thr Phe Thr Thr Tyr

20 25 30

Tyr Leu His Trp Val Arg Gln Ala Pro Gly Gln Cys Leu Glu Trp Met

35 40 45

Gly Trp Ile Tyr Pro Gly Asn Val His Ala Gln Tyr Asn Glu Lys Phe
50 55 60

Lys Gly Arg Val Thr Ile Thr Ala Asp Lys Ser Thr Ser Thr Ala Tyr
65 70 75 80

Met Glu Leu Ser Ser Leu Arg Ser Glu Asp Thr Ala Val Tyr Tyr Cys
85 90 95

Ala Arg Ser Trp Glu Gly Phe Pro Tyr Trp Gly Gln Gly Thr Thr Val
100 105 110

Thr Val Ser Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly Gly
115 120 125

Gly Gly Ser Asp Ile Gln Met Thr Gln Ser Pro Ser Ser Leu Ser Ala
130 135 140

Ser Val Gly Asp Arg Val Thr Ile Thr Cys Lys Ala Ser Gln Asn Val
145 150 155 160

Gly Ile Asn Val Ala Trp Tyr Gln Gln Lys Pro Gly Lys Ala Pro Lys
165 170 175

Ser Leu Ile Ser Ser Ala Ser Tyr Arg Tyr Ser Gly Val Pro Ser Arg
180 185 190

Phe Ser Gly Ser Gly Ser Gly Thr Asp Phe Thr Leu Thr Ile Ser Ser
195 200 205

Leu Gln Pro Glu Asp Phe Ala Thr Tyr Phe Cys Gln Gln Tyr Asp Thr
210 215 220

Tyr Pro Phe Thr Phe Gly Cys Gly Thr Lys Val Glu Ile Lys Gly Gly
225 230 235 240

Gly Gly Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly Gly Gly
245 250 255

Gly Ser Gly Gly Gly Gly Ser Gln Ile Gln Leu Val Gln Ser Gly Pro
260 265 270

Glu Leu Lys Lys Pro Gly Glu Thr Val Lys Ile Ser Cys Lys Ala Ser
275 280 285

Gly Phe Thr Phe Thr Asp Tyr Ser Ile His Trp Val Lys Gln Ala Pro
290 295 300

Gly Lys Gly Leu Lys Trp Met Gly Trp Ile Asn Thr Glu Thr Gly Glu
305 310 315 320

Pro Thr Tyr Thr Asp Asp Phe Lys Gly Arg Phe Ala Phe Ser Leu Val
325 330 335

Thr Ser Ala Thr Thr Ala Tyr Leu Gln Ile Asn Asn Leu Asn Asn Glu
340 345 350

Asp Thr Ala Thr Phe Phe Cys Ala Arg Phe Ile Tyr Asp Pro Tyr Trp
355 360 365

Gly Phe Ala Tyr Trp Gly Gln Gly Thr Leu Val Thr Val Ser Ala Ala
370 375 380

Ser Thr Lys Gly Pro Ser Val Phe Pro Leu Ala Pro Ser Ser Lys Ser

385 390 395 400

Thr Ser Gly Gly Thr Ala Ala Leu Gly Cys Leu Val Lys Asp Tyr Phe
405 410 415

Pro Glu Pro Val Thr Val Ser Trp Asn Ser Gly Ala Leu Thr Ser Gly
420 425 430

Val His Thr Phe Pro Ala Val Leu Gln Ser Ser Gly Leu Tyr Ser Leu
435 440 445

Ser Ser Val Val Thr Val Pro Ser Ser Ser Leu Gly Thr Gln Thr Tyr
450 455 460

Ile Cys Asn Val Asn His Lys Pro Ser Asn Thr Lys Val Asp Lys Lys
465 470 475 480

Val Glu Pro Lys Ser Cys Asp Lys Thr His Thr Cys Pro Pro Cys Pro
485 490 495

Ala Pro Glu Leu Leu Gly Gly Pro Ser Val Phe Leu Phe Pro Pro Lys
500 505 510

Pro Lys Asp Thr Leu Met Ile Ser Arg Thr Pro Glu Val Thr Cys Val
515 520 525

Val Val Asp Val Ser His Glu Asp Pro Glu Val Lys Phe Asn Trp Tyr
530 535 540

Val Asp Gly Val Glu Val His Asn Ala Lys Thr Lys Pro Arg Glu Glu
545 550 555 560

Gln Tyr Asn Ser Thr Tyr Arg Val Val Ser Val Leu Thr Val Leu His
565 570 575

Gln Asp Trp Leu Asn Gly Lys Glu Tyr Lys Cys Lys Val Ser Asn Lys
580 585 590

Ala Leu Pro Ala Pro Ile Glu Lys Thr Ile Ser Lys Ala Lys Gly Gln
595 600 605

Pro Arg Glu Pro Gln Val Tyr Thr Leu Pro Pro Ser Arg Asp Glu Leu
610 615 620

Thr Lys Asn Gln Val Ser Leu Thr Cys Leu Val Lys Gly Phe Tyr Pro
625 630 635 640

Ser Asp Ile Ala Val Glu Trp Glu Ser Asn Gly Gln Pro Glu Asn Asn
645 650 655

Tyr Lys Thr Thr Pro Pro Val Leu Asp Ser Asp Gly Ser Phe Phe Leu
660 665 670

Tyr Ser Lys Leu Thr Val Asp Lys Ser Arg Trp Gln Gln Gly Asn Val
675 680 685

Phe Ser Cys Ser Val Met His Glu Ala Leu His Asn His Tyr Thr Gln
690 695 700

Lys Ser Leu Ser Leu Ser Pro Gly
705 710

<210> 36

<211> 2154

<212> DNA

<213> Seqüência Artificial

<220>

<223> Descrição de Seqüência Artificial: Construto sintético

<400> 36

caggtccaac tggtgcagtc tggagctgag gtgaagaagc ctgggtcctc agtgaaggtg 60

tcctgcaagg cttctggcta cactttcaca acctactatt tgcactgggt gaggcaggcc 120

cctggacagt gccttgagtg gatgggatgg atttatcctg gaaatgttca tgctcagtac 180

aatgagaagt tcaagggcag ggtcacaatc actgcagaca aatccaccag cacagcctac 240

atggagctca gcagcctgag gtctgaagat actgcggtct attactgtgc aagatcctgg 300

gaaggttttc cttactgggg ccaagggacc acggtcaccg tctcctcagg cggtggaggg 360

tccggtgggg gcggatctgg gggcggcgga tccggtggtg gtggtagtga cattcagatg 420

acccagtctc ctagctccct gtccgcctca gtaggagaca gggtcaccat cacctgcaag 480

gccagtcaga atgtgggtat taatgtagcc tggtatcaac agaaaccagg gaaggctcct 540

aaatcactga tttcctcggc ctcctaccgg tacagtggag tcccttccag attcagcggc 600

agtggatctg ggacagattt cactctcacc atcagcagcc tccagcctga agacttcgca 660

acctatttct gtcagcaata tgacacctat ccattcacgt tcggctgcgg taccaaggtg 720

gagatcaaag gcggtggagg gtccggtgga gggggctctg gagggggcgg ttcaggggc 780

ggtggatcgg ggggaggtgg ctcccagatc cagttggtgc agtctggacc tgagctgaag 840

aagcctggag agacagtcaa gatctcctgc aaggcttctg gttttacctt cacagactat 900

tcaatacact gggtgaaaca ggctccagga aagggtttaa agtggatggg ctggataaac 960

actgagactg gtgagccaac atatacagat gacttcaagg gacgatttgc cttctctttg 1020

gtgacctctg ccaccactgc ctatttgcag atcaacaacc tcaacaatga ggacacggct 1080

acatttttct gtgctagatt catctatgat ccttattggg ggtttgctta ctggggccag 1140

gggactctgg tcactgtctc cgcagctagc accaagggcc catcggtctt ccccctggca 1200

ccctcctcca agagcacctc tgggggcaca gcggccctgg gctgcctggt caaggactac 1260

ttccccgaac cggtgacggt gtcgtggaac tcaggcgccc tgaccagcgg cgtgcacacc 1320

ttcccggctg tcctacagtc ctcaggactc tactccctca gcagcgtggt gaccgtgccc 1380

tccagcagct tgggcaccca gacctacatc tgcaacgtga atcacaagcc cagcaacacc 1440

aaggtggaca agaaagttga gcccaaatct tgtgacaaaa ctcacacatg cccaccgtgc 1500

ccagcacctg aactcctggg gggaccgtca gtcttcctct tccccccaaa acccaaggac 1560

accctcatga tctcccggac ccctgaggtc acatgcgtgg tggtggacgt gagccacgaa 1620

gaccctgagg tcaagttcaa ctggtacgtg gacggcgtgg aggtgcataa tgccaagaca 1680

aagccgcggg aggagcagta caacagcacg taccgtgtgg tcagcgtcct caccgtcctg 1740

caccaggact ggctgaatgg caaggagtac aagtgcaagg tctccaacaa agccctccca 1800

gcccccatcg agaaaaccat ctccaaagcc aaagggcagc cccgagaacc acaggtgtac 1860

accctgcccc catcccggga tgagctgacc aagaaccagg tcagcctgac ctgcctggtc 1920

aaaggcttct atcccagcga catcgccgtg gagtgggaga gcaatgggca gccggagaac 1980

aactacaaga ccacgcctcc cgtgctggac tccgacggct ccttcttcct ctacagcaag 2040

ctcaccgtgg acaagagcag gtggcagcag gggaacgtct tctcatgctc cgtgatgcat 2100

gaggctctgc acaaccacta cacgcagaag agcctctccc tgtctccggg ttga 2154

<210> 37

<211> 717

<212> PRT

<213> Seqüência Artificial

<220>

<223> Descrição de Seqüência Artificial: Construto sintético

<400> 37

Gln Val Gln Leu Val Gln Ser Gly Ala Glu Val Lys Lys Pro Gly Ser

1 5 10 15

Ser Val Lys Val Ser Cys Lys Ala Ser Gly Tyr Thr Phe Thr Thr Tyr

20 25 30

Tyr Leu His Trp Val Arg Gln Ala Pro Gly Gln Cys Leu Glu Trp Met

35 40 45

Gly Trp Ile Tyr Pro Gly Asn Val His Ala Gln Tyr Asn Glu Lys Phe

50 55 60

Lys Gly Arg Val Thr Ile Thr Ala Asp Lys Ser Thr Ser Thr Ala Tyr

65 70 75 80

Met Glu Leu Ser Ser Leu Arg Ser Glu Asp Thr Ala Val Tyr Tyr Cys
85 90 95

Ala Arg Ser Trp Glu Gly Phe Pro Tyr Trp Gly Gln Gly Thr Thr Val
100 105 110

Thr Val Ser Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly Gly
115 120 125

Gly Gly Ser Gly Gly Gly Gly Ser Asp Ile Gln Met Thr Gln Ser Pro
130 135 140

Ser Ser Leu Ser Ala Ser Val Gly Asp Arg Val Thr Ile Thr Cys Lys
145 150 155 160

Ala Ser Gln Asn Val Gly Ile Asn Val Ala Trp Tyr Gln Gln Lys Pro
165 170 175

Gly Lys Ala Pro Lys Ser Leu Ile Ser Ser Ala Ser Tyr Arg Tyr Ser
180 185 190

Gly Val Pro Ser Arg Phe Ser Gly Ser Gly Ser Gly Thr Asp Phe Thr
195 200 205

Leu Thr Ile Ser Ser Leu Gln Pro Glu Asp Phe Ala Thr Tyr Phe Cys
210 215 220

Gln Gln Tyr Asp Thr Tyr Pro Phe Thr Phe Gly Cys Gly Thr Lys Val
225 230 235 240

Glu Ile Lys Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly Gly Gly
245 250 255

Gly Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Gln Ile Gln Leu
260 265 270

Val Gln Ser Gly Pro Glu Leu Lys Lys Pro Gly Glu Thr Val Lys Ile
275 280 285

Ser Cys Lys Ala Ser Gly Phe Thr Phe Thr Asp Tyr Ser Ile His Trp
290 295 300

Val Lys Gln Ala Pro Gly Lys Gly Leu Lys Trp Met Gly Trp Ile Asn
305 310 315 320

Thr Glu Thr Gly Glu Pro Thr Tyr Thr Asp Asp Phe Lys Gly Arg Phe
325 330 335

Ala Phe Ser Leu Val Thr Ser Ala Thr Thr Ala Tyr Leu Gln Ile Asn
340 345 350

Asn Leu Asn Asn Glu Asp Thr Ala Thr Phe Phe Cys Ala Arg Phe Ile
355 360 365

Tyr Asp Pro Tyr Trp Gly Phe Ala Tyr Trp Gly Gln Gly Thr Leu Val
370 375 380

Thr Val Ser Ala Ala Ser Thr Lys Gly Pro Ser Val Phe Pro Leu Ala
385 390 395 400

Pro Ser Ser Lys Ser Thr Ser Gly Gly Thr Ala Ala Leu Gly Cys Leu
405 410 415

Val Lys Asp Tyr Phe Pro Glu Pro Val Thr Val Ser Trp Asn Ser Gly

420 425 430

Ala Leu Thr Ser Gly Val His Thr Phe Pro Ala Val Leu Gln Ser Ser
435 440 445

Gly Leu Tyr Ser Leu Ser Ser Val Val Thr Val Pro Ser Ser Ser Leu
450 455 460

Gly Thr Gln Thr Tyr Ile Cys Asn Val Asn His Lys Pro Ser Asn Thr
465 470 475 480

Lys Val Asp Lys Lys Val Glu Pro Lys Ser Cys Asp Lys Thr His Thr
485 490 495

Cys Pro Pro Cys Pro Ala Pro Glu Leu Leu Gly Gly Pro Ser Val Phe
500 505 510

Leu Phe Pro Pro Lys Pro Lys Asp Thr Leu Met Ile Ser Arg Thr Pro
515 520 525

Glu Val Thr Cys Val Val Val Asp Val Ser His Glu Asp Pro Glu Val
530 535 540

Lys Phe Asn Trp Tyr Val Asp Gly Val Glu Val His Asn Ala Lys Thr
545 550 555 560

Lys Pro Arg Glu Glu Gln Tyr Asn Ser Thr Tyr Arg Val Val Ser Val
565 570 575

Leu Thr Val Leu His Gln Asp Trp Leu Asn Gly Lys Glu Tyr Lys Cys
580 585 590

Lys Val Ser Asn Lys Ala Leu Pro Ala Pro Ile Glu Lys Thr Ile Ser
595 600 605

Lys Ala Lys Gly Gln Pro Arg Glu Pro Gln Val Tyr Thr Leu Pro Pro
610 615 620

Ser Arg Asp Glu Leu Thr Lys Asn Gln Val Ser Leu Thr Cys Leu Val
625 630 635 640

Lys Gly Phe Tyr Pro Ser Asp Ile Ala Val Glu Trp Glu Ser Asn Gly
645 650 655

Gln Pro Glu Asn Asn Tyr Lys Thr Thr Pro Pro Val Leu Asp Ser Asp
660 665 670

Gly Ser Phe Phe Leu Tyr Ser Lys Leu Thr Val Asp Lys Ser Arg Trp
675 680 685

Gln Gln Gly Asn Val Phe Ser Cys Ser Val Met His Glu Ala Leu His
690 695 700

Asn His Tyr Thr Gln Lys Ser Leu Ser Leu Ser Pro Gly
705 710 715

<210> 38

<211> 64

<212> DNA

<213> Seqüência Artificial

<220>

<223> Descrição de Seqüência Artificial: Oligonucleotídeo sintético

<400> 38

ggggggtggat ccggtggagg gggctccggc ggtggcgggt cccaggtcca actggtgcag 60

tctg 64

<210> 39

<211> 30

<212> DNA

<213> Seqüência Artificial

<220>

<223> Descrição de Seqüência Artificial: Oligonucleotídeo sintético

<400> 39

tgggggcggc gggtccggtg gtggtggtag 30

<210> 40

<211> 30

<212> DNA

<213> Seqüência Artificial

<220>

<223> Descrição de Seqüência Artificial: Oligonucleotídeo sintético

<400> 40

taccaccacc accggacccg ccgcccccag 30

<210> 41

<211> 32

<212> DNA

<213> Seqüência Artificial

<220>

<223> Descrição de Seqüência Artificial: Oligonucleotídeo sintético

<400> 41

gttaacggat cctcatttga tctccacctt gg 32

<210> 42

<211> 48

<212> DNA

<213> Seqüência Artificial

<220>

<223> Descrição de Seqüência Artificial: Oligonucleotídeo sintético

<400> 42

tccggcgggg gtggatccgg tggagggggc tccggcggtg gcgggtcc 48

<210> 43

<211> 16

<212> PRT

<213> Seqüência Artificial

<220>

<223> Descrição de Seqüência Artificial: Peptídeo sintético

<400> 43

Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser

1 5 10 15

<210> 44

<211> 2115

<212> DNA

<213> Seqüência Artificial

<220>

<223> Descrição de Seqüência Artificial: Construto sintético

<400> 44

cagatccagt tggtgcagtc tggacctgag ctgaagaagc ctggagagac agtcaagatc 60

tcctgcaagg cttctggttt taccttcaca gactattcaa tacactgggt gaaacaggct 120

ccaggaaagg gtttaaagtg gatgggctgg ataaacactg agactggtga gccaacatat 180

acagatgact tcaagggacg atttgccttc tctttggtga cctctgccac cactgcctat 240

ttgcagatca acaacctcaa caatgaggac acggctacat ttttctgtgc tagattcatc 300

tatgatcctt attgggggtt tgcttactgg ggccagggga ctctggtcac tgtctccgca 360

gctagcacca agggcccatc ggtcttcccc ctggcaccct cctccaagag cacctctggg 420

ggcacagcgg ccctgggctg cctggtcaag gactacttcc ccgaaccggt gacggtgtcg 480

tggaactcag gcgccctgac cagcggcgtg cacaccttcc cggctgtcct acagtcctca 540

ggactctact ccctcagcag cgtggtgacc gtgccctcca gcagcttggg cacccagacc 600

tacatctgca acgtgaatca caagcccagc aacaccaagg tggacaagaa agttgagccc 660

aaatcttgtg acaaaactca cacatgccca ccgtgcccag cacctgaact cctgggggga 720

ccgtcagtct tcctcttccc cccaaaaccc aaggacaccc tcatgatctc ccggacccct 780

gaggtcacat gcgtggtggt ggacgtgagc cacgaagacc ctgaggtcaa gttcaactgg 840

tacgtggacg gcgtggaggt gcataatgcc aagacaaagc cgcgggagga gcagtacaac 900

agcacgtacc gtgtggtcag cgtcctcacc gtcctgcacc aggactggct gaatggcaag 960

gagtacaagt gcaaggtctc caacaaagcc ctcccagccc ccatcgagaa aaccatctcc 1020

aaagccaaag ggcagccccg agaaccacag gtgtacaccc tgcccccatc ccgggatgag 1080

ctgaccaaga accaggtcag cctgacctgc ctggtcaaag gcttctatcc cagcgacatc 1140

gccgtggagt gggagagcaa tgggcagccg gagaacaact acaagaccac gcctcccgtg 1200

ctggactccg acggctcctt cttcctctac agcaagctca ccgtggacaa gagcaggtgg 1260

cagcagggga acgtcttctc atgctccgtg atgcatgagg ctctgcacaa ccactacacg 1320

cagaagagcc tctccctgtc tccgggtaaa tccggcgggg gtggatccgg tggaggggggc 1380

tccggcggtg gcgggtccca ggtccaactg gtgcagtctg gagctgaggt gaagaagcct 1440

gggtcctcag tgaaggtgtc ctgcaaggct tctggctaca ctttcacaac ctactatttg 1500

cactgggtga ggcaggcccc tggacaggga cttgagtgga tgggatggat ttatcctgga 1560

aatgttcatg ctcagtacaa tgagaagttc aagggcaggg tcacaatcac tgcagacaaa 1620

tccaccagca cagcctacat ggagctcagc agcctgaggt ctgaagatac tgcggtctat 1680

tactgtgcaa gatcctggga aggttttcct tactggggcc aagggaccac ggtcaccgtc 1740

tcctcaggtg ggggcggatc tgggggcggc gggtccggtg gtggtggtag tgacattcag 1800

atgacccagt ctcctagctc cctgtccgcc tcagtaggag acagggtcac catcacctgc 1860

aaggccagtc agaatgtggg tattaatgta gcctggtatc aacagaaacc agggaaggct 1920

cctaaatcac tgatttcctc ggcctcctac cggtacagtg gagtcccttc cagattcagc 1980

ggcagtggat ctgggacaga tttcactctc accatcagca gcctccagcc tgaagacttc 2040

gcaacctatt tctgtcagca atatgacacc tatccattca cgttcggcca gggtaccaag 2100

gtggagatca aatga 2115

<210> 45

<211> 704

<212> PRT

<213> Seqüência Artificial

<220>

<223> Descrição de Seqüência Artificial: Construto sintético

<400> 45

Gln Ile Gln Leu Val Gln Ser Gly Pro Glu Leu Lys Lys Pro Gly Glu
1 5 10 15

Thr Val Lys Ile Ser Cys Lys Ala Ser Gly Phe Thr Phe Thr Asp Tyr
20 25 30

Ser Ile His Trp Val Lys Gln Ala Pro Gly Lys Gly Leu Lys Trp Met
35 40 45

Gly Trp Ile Asn Thr Glu Thr Gly Glu Pro Thr Tyr Thr Asp Asp Phe
50 55 60

Lys Gly Arg Phe Ala Phe Ser Leu Val Thr Ser Ala Thr Thr Ala Tyr
65 70 75 80

Leu Gln Ile Asn Asn Leu Asn Asn Glu Asp Thr Ala Thr Phe Phe Cys
85 90 95

Ala Arg Phe Ile Tyr Asp Pro Tyr Trp Gly Phe Ala Tyr Trp Gly Gln
100 105 110

Gly Thr Leu Val Thr Val Ser Ala Ala Ser Thr Lys Gly Pro Ser Val
115 120 125

Phe Pro Leu Ala Pro Ser Ser Lys Ser Thr Ser Gly Gly Thr Ala Ala
130 135 140

Leu Gly Cys Leu Val Lys Asp Tyr Phe Pro Glu Pro Val Thr Val Ser
145 150 155 160

Trp Asn Ser Gly Ala Leu Thr Ser Gly Val His Thr Phe Pro Ala Val
165 170 175

Leu Gln Ser Ser Gly Leu Tyr Ser Leu Ser Ser Val Val Thr Val Pro
180 185 190

Ser Ser Ser Leu Gly Thr Gln Thr Tyr Ile Cys Asn Val Asn His Lys
195 200 205

Pro Ser Asn Thr Lys Val Asp Lys Lys Val Glu Pro Lys Ser Cys Asp
210 215 220

Lys Thr His Thr Cys Pro Pro Cys Pro Ala Pro Glu Leu Leu Gly Gly
225 230 235 240

Pro Ser Val Phe Leu Phe Pro Pro Lys Pro Lys Asp Thr Leu Met Ile
245 250 255

Ser Arg Thr Pro Glu Val Thr Cys Val Val Val Asp Val Ser His Glu
260 265 270

Asp Pro Glu Val Lys Phe Asn Trp Tyr Val Asp Gly Val Glu Val His
275 280 285

Asn Ala Lys Thr Lys Pro Arg Glu Glu Gln Tyr Asn Ser Thr Tyr Arg
290 295 300

Val Val Ser Val Leu Thr Val Leu His Gln Asp Trp Leu Asn Gly Lys
305 310 315 320

Glu Tyr Lys Cys Lys Val Ser Asn Lys Ala Leu Pro Ala Pro Ile Glu

325 330 335

Lys Thr Ile Ser Lys Ala Lys Gly Gln Pro Arg Glu Pro Gln Val Tyr
340 345 350

Thr Leu Pro Pro Ser Arg Asp Glu Leu Thr Lys Asn Gln Val Ser Leu
355 360 365

Thr Cys Leu Val Lys Gly Phe Tyr Pro Ser Asp Ile Ala Val Glu Trp
370 375 380

Glu Ser Asn Gly Gln Pro Glu Asn Asn Tyr Lys Thr Thr Pro Pro Val
385 390 395 400

Leu Asp Ser Asp Gly Ser Phe Phe Leu Tyr Ser Lys Leu Thr Val Asp
405 410 415

Lys Ser Arg Trp Gln Gln Gly Asn Val Phe Ser Cys Ser Val Met His
420 425 430

Glu Ala Leu His Asn His Tyr Thr Gln Lys Ser Leu Ser Leu Ser Pro
435 440 445

Gly Lys Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly Gly Gly
450 455 460

Gly Ser Gln Val Gln Leu Val Gln Ser Gly Ala Glu Val Lys Lys Pro
465 470 475 480

Gly Ser Ser Val Lys Val Ser Cys Lys Ala Ser Gly Tyr Thr Phe Thr
485 490 495

Thr Tyr Tyr Leu His Trp Val Arg Gln Ala Pro Gly Gln Gly Leu Glu
500 505 510

Trp Met Gly Trp Ile Tyr Pro Gly Asn Val His Ala Gln Tyr Asn Glu
515 520 525

Lys Phe Lys Gly Arg Val Thr Ile Thr Ala Asp Lys Ser Thr Ser Thr
530 535 540

Ala Tyr Met Glu Leu Ser Ser Leu Arg Ser Glu Asp Thr Ala Val Tyr
545 550 555 560

Tyr Cys Ala Arg Ser Trp Glu Gly Phe Pro Tyr Trp Gly Gln Gly Thr
565 570 575

Thr Val Thr Val Ser Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser
580 585 590

Gly Gly Gly Gly Ser Asp Ile Gln Met Thr Gln Ser Pro Ser Ser Leu
595 600 605

Ser Ala Ser Val Gly Asp Arg Val Thr Ile Thr Cys Lys Ala Ser Gln
610 615 620

Asn Val Gly Ile Asn Val Ala Trp Tyr Gln Gln Lys Pro Gly Lys Ala
625 630 635 640

Pro Lys Ser Leu Ile Ser Ser Ala Ser Tyr Arg Tyr Ser Gly Val Pro
645 650 655

Ser Arg Phe Ser Gly Ser Gly Ser Gly Thr Asp Phe Thr Leu Thr Ile
660 665 670

Ser Ser Leu Gln Pro Glu Asp Phe Ala Thr Tyr Phe Cys Gln Gln Tyr
675 680 685

Asp Thr Tyr Pro Phe Thr Phe Gly Gln Gly Thr Lys Val Glu Ile Lys
690 695 700

<210> 46

<211> 2130

<212> DNA

<213> Seqüência Artificial

<220>

<223> Descrição de Seqüência Artificial: Construto sintético

<400> 46

cagatccagt tggtgcagtc tggacctgag ctgaagaagc ctggagagac agtcaagatc 60

tcctgcaagg cttctggttt taccttcaca gactattcaa tacactgggt gaaacaggct 120

ccaggaaagg gtttaaagtg gatgggctgg ataaacactg agactggtga gccaacatat 180

acagatgact tcaagggacg atttgccttc tctttggtga cctctgccac cactgcctat 240

ttgcagatca acaacctcaa caatgaggac acggctacat ttttctgtgc tagattcatc 300

tatgatcctt attgggggtt tgcttactgg ggccagggga ctctggtcac tgtctccgca 360

gctagcacca agggcccatc ggtcttcccc ctggcaccct cctccaagag cacctctggg 420

ggcacagcgg ccctgggctg cctggtcaag gactacttcc ccgaaccggt gacggtgtcg 480

tggaactcag gcgccctgac cagcggcgtg cacaccttcc cggctgtcct acagtcctca 540

ggactctact ccctcagcag cgtggtgacc gtgccctcca gcagcttggg cacccagacc 600

tacatctgca acgtgaatca caagcccagc aacaccaagg tggacaagaa agttgagccc 660

aaatcttgtg acaaaactca cacatgccca ccgtgcccag cacctgaact cctgggggga 720

ccgtcagtct tcctcttccc cccaaaaccc aaggacaccc tcatgatctc ccggacccct 780

gaggtcacat gcgtggtggt ggacgtgagc cacgaagacc ctgaggtcaa gttcaactgg 840

tacgtggacg gcgtggaggt gcataatgcc aagacaaagc cgcgggagga gcagtacaac 900

agcacgtacc gtgtggtcag cgtcctcacc gtcctgcacc aggactggct gaatggcaag 960

gagtacaagt gcaaggtctc caacaaagcc ctcccagccc ccatcgagaa aaccatctcc 1020

aaagccaaag ggcagccccg agaaccacag gtgtacaccc tgcccccatc ccgggatgag 1080

ctgaccaaga accaggtcag cctgacctgc ctggtcaaag gcttctatcc cagcgacatc 1140

gccgtggagt gggagagcaa tgggcagccg gagaacaact acaagaccac gcctcccgtg 1200

ctggactccg acggctcctt cttcctctac agcaagctca ccgtggacaa gagcaggtgg 1260

cagcagggga acgtcttctc atgctccgtg atgcatgagg ctctgcacaa ccactacacg 1320

cagaagagcc tctccctgtc tccgggtaaa tccggcgggg gtggatccgg tggaggggggc 1380

tccggcggtg gcgggtccca ggtccaactg gtgcagtctg gagctgaggt gaagaagcct 1440

gggtcctcag tgaaggtgtc ctgcaaggct tctggctaca ctttcacaac ctactatttg 1500

cactgggtga ggcaggcccc tggacaggga cttgagtgga tgggatggat ttatcctgga 1560

aatgttcatg ctcagtacaa tgagaagttc aagggcaggg tcacaatcac tgcagacaaa 1620

tccaccagca cagcctacat ggagctcagc agcctgaggt ctgaagatac tgcggtctat 1680

tactgtgcaa gatcctggga aggttttcct tactggggcc aagggaccac ggtcaccgtc 1740

tcctcaggcg gtggagggtc cggtgggggc ggatctgggg gcggcgggtc cggtggtggt 1800

ggtagtgaca ttcagatgac ccagtctcct agctccctgt ccgcctcagt aggagacagg 1860

gtcaccatca cctgcaaggc cagtcagaat gtgggtatta atgtagcctg gtatcaacag 1920

aaaccaggga aggctcctaa atcactgatt tcctcggcct cctaccggta cagtggagtc 1980

ccttccagat tcagcggcag tggatctggg acagatttca ctctcaccat cagcagcctc 2040

cagcctgaag acttcgcaac ctatttctgt cagcaatatg acacctatcc attcacgttc 2100

ggccagggta ccaaggtgga gatcaaatga 2130

<210> 47

<211> 709

<212> PRT

<213> Seqüência Artificial

<220>

<223> Descrição de Seqüência Artificial: Construto sintético

<400> 47

Gln Ile Gln Leu Val Gln Ser Gly Pro Glu Leu Lys Lys Pro Gly Glu

1 5 10 15

Thr Val Lys Ile Ser Cys Lys Ala Ser Gly Phe Thr Phe Thr Asp Tyr

20 25 30

Ser Ile His Trp Val Lys Gln Ala Pro Gly Lys Gly Leu Lys Trp Met
35 40 45

Gly Trp Ile Asn Thr Glu Thr Gly Glu Pro Thr Tyr Thr Asp Asp Phe
50 55 60

Lys Gly Arg Phe Ala Phe Ser Leu Val Thr Ser Ala Thr Thr Ala Tyr
65 70 75 80

Leu Gln Ile Asn Asn Leu Asn Asn Glu Asp Thr Ala Thr Phe Phe Cys
85 90 95

Ala Arg Phe Ile Tyr Asp Pro Tyr Trp Gly Phe Ala Tyr Trp Gly Gln
100 105 110

Gly Thr Leu Val Thr Val Ser Ala Ala Ser Thr Lys Gly Pro Ser Val
115 120 125

Phe Pro Leu Ala Pro Ser Ser Lys Ser Thr Ser Gly Gly Thr Ala Ala
130 135 140

Leu Gly Cys Leu Val Lys Asp Tyr Phe Pro Glu Pro Val Thr Val Ser
145 150 155 160

Trp Asn Ser Gly Ala Leu Thr Ser Gly Val His Thr Phe Pro Ala Val
165 170 175

Leu Gln Ser Ser Gly Leu Tyr Ser Leu Ser Ser Val Val Thr Val Pro
180 185 190

Ser Ser Ser Leu Gly Thr Gln Thr Tyr Ile Cys Asn Val Asn His Lys
195 200 205

Pro Ser Asn Thr Lys Val Asp Lys Lys Val Glu Pro Lys Ser Cys Asp
210 215 220

Lys Thr His Thr Cys Pro Pro Cys Pro Ala Pro Glu Leu Leu Gly Gly
225 230 235 240

Pro Ser Val Phe Leu Phe Pro Pro Lys Pro Lys Asp Thr Leu Met Ile
245 250 255

Ser Arg Thr Pro Glu Val Thr Cys Val Val Val Asp Val Ser His Glu
260 265 270

Asp Pro Glu Val Lys Phe Asn Trp Tyr Val Asp Gly Val Glu Val His
275 280 285

Asn Ala Lys Thr Lys Pro Arg Glu Glu Gln Tyr Asn Ser Thr Tyr Arg
290 295 300

Val Val Ser Val Leu Thr Val Leu His Gln Asp Trp Leu Asn Gly Lys
305 310 315 320

Glu Tyr Lys Cys Lys Val Ser Asn Lys Ala Leu Pro Ala Pro Ile Glu
325 330 335

Lys Thr Ile Ser Lys Ala Lys Gly Gln Pro Arg Glu Pro Gln Val Tyr
340 345 350

Thr Leu Pro Pro Ser Arg Asp Glu Leu Thr Lys Asn Gln Val Ser Leu
355 360 365

Thr Cys Leu Val Lys Gly Phe Tyr Pro Ser Asp Ile Ala Val Glu Trp

370 375 380

Glu Ser Asn Gly Gln Pro Glu Asn Asn Tyr Lys Thr Thr Pro Pro Val
385 390 395 400

Leu Asp Ser Asp Gly Ser Phe Phe Leu Tyr Ser Lys Leu Thr Val Asp
405 410 415

Lys Ser Arg Trp Gln Gln Gly Asn Val Phe Ser Cys Ser Val Met His
420 425 430

Glu Ala Leu His Asn His Tyr Thr Gln Lys Ser Leu Ser Leu Ser Pro
435 440 445

Gly Lys Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly Gly Gly
450 455 460

Gly Ser Gln Val Gln Leu Val Gln Ser Gly Ala Glu Val Lys Lys Pro
465 470 475 480

Gly Ser Ser Val Lys Val Ser Cys Lys Ala Ser Gly Tyr Thr Phe Thr
485 490 495

Thr Tyr Tyr Leu His Trp Val Arg Gln Ala Pro Gly Gln Gly Leu Glu
500 505 510

Trp Met Gly Trp Ile Tyr Pro Gly Asn Val His Ala Gln Tyr Asn Glu
515 520 525

Lys Phe Lys Gly Arg Val Thr Ile Thr Ala Asp Lys Ser Thr Ser Thr
530 535 540

Ala Tyr Met Glu Leu Ser Ser Leu Arg Ser Glu Asp Thr Ala Val Tyr
545 550 555 560

Tyr Cys Ala Arg Ser Trp Glu Gly Phe Pro Tyr Trp Gly Gln Gly Thr
565 570 575

Thr Val Thr Val Ser Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser
580 585 590

Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Asp Ile Gln Met Thr Gln
595 600 605

Ser Pro Ser Ser Leu Ser Ala Ser Val Gly Asp Arg Val Thr Ile Thr
610 615 620

Cys Lys Ala Ser Gln Asn Val Gly Ile Asn Val Ala Trp Tyr Gln Gln
625 630 635 640

Lys Pro Gly Lys Ala Pro Lys Ser Leu Ile Ser Ser Ala Ser Tyr Arg
645 650 655

Tyr Ser Gly Val Pro Ser Arg Phe Ser Gly Ser Gly Ser Gly Thr Asp
660 665 670

Phe Thr Leu Thr Ile Ser Ser Leu Gln Pro Glu Asp Phe Ala Thr Tyr
675 680 685

Phe Cys Gln Gln Tyr Asp Thr Tyr Pro Phe Thr Phe Gly Gln Gly Thr
690 695 700

Lys Val Glu Ile Lys
705

<210> 48

<211> 2115

<212> DNA

<213> Seqüência Artificial

<220>

<223> Descrição de Seqüência Artificial: Construto sintético

<400> 48

cagatccagt tggtgcagtc tggacctgag ctgaagaagc ctggagagac agtcaagatc 60

tcctgcaagg cttctggttt taccttcaca gactattcaa tacactgggt gaaacaggct 120

ccaggaaagg gtttaaagtg gatgggctgg ataaacactg agactggtga gccaacatat 180

acagatgact tcaagggacg atttgccttc tctttggtga cctctgccac cactgcctat 240

ttgcagatca acaacctcaa caatgaggac acggctacat ttttctgtgc tagattcatc 300

tatgatcctt attgggggtt tgcttactgg ggccagggga ctctggtcac tgtctccgca 360

gctagcacca agggcccatc ggtcttcccc ctggcaccct cctccaagag cacctctggg 420

ggcacagcgg ccctgggctg cctggtcaag gactacttcc ccgaaccggt gacggtgtcg 480

tggaactcag gcgccctgac cagcggcgtg cacaccttcc cggctgtcct acagtcctca 540

ggactctact ccctcagcag cgtggtgacc gtgccctcca gcagcttggg cacccagacc 600

tacatctgca acgtgaatca caagcccagc aacaccaagg tggacaagaa agttgagccc 660

aaatcttgtg acaaaactca cacatgccca ccgtgcccag cacctgaact cctgggggga 720

ccgtcagtct tcctcttccc cccaaaaccc aaggacaccc tcatgatctc ccggacccct 780

gaggtcacat gcgtggtggt ggacgtgagc cacgaagacc ctgaggtcaa gttcaactgg 840

tacgtggacg gcgtggaggt gcataatgcc aagacaaagc cgcgggagga gcagtacaac 900

agcacgtacc gtgtggtcag cgtcctcacc gtcctgcacc aggactggct gaatggcaag 960

gagtacaagt gcaaggtctc caacaaagcc ctcccagccc ccatcgagaa aaccatctcc 1020

aaagccaaag ggcagccccg agaaccacag gtgtacaccc tgcccccatc ccgggatgag 1080

ctgaccaaga accaggtcag cctgacctgc ctggtcaaag gcttctatcc cagcgacatc 1140

gccgtggagt gggagagcaa tgggcagccg gagaacaact acaagaccac gcctcccgtg 1200

ctggactccg acggctcctt cttcctctac agcaagctca ccgtggacaa gagcaggtgg 1260

cagcagggga acgtcttctc atgctccgtg atgcatgagg ctctgcacaa ccactacacg 1320

cagaagagcc tctccctgtc tccgggtaaa tccggcgggg gtggatccgg tggaggggggc 1380

tccggcggtg gcgggtccca ggtccaactg gtgcagtctg gagctgaggt gaagaagcct 1440

gggtcctcag tgaaggtgtc ctgcaaggct tctggctaca ctttcacaac ctactatttg 1500

cactgggtga ggcaggcccc tggacagtgc cttgagtgga tgggatggat ttatcctgga 1560

aatgttcatg ctcagtacaa tgagaagttc aagggcaggg tcacaatcac tgcagacaaa 1620

tccaccagca cagcctacat ggagctcagc agcctgaggt ctgaagatac tgcggtctat 1680

tactgtgcaa gatcctggga aggttttcct tactggggcc aagggaccac ggtcaccgtc 1740

tcctcaggtg ggggcggatc tgggggcggc gggtccggtg gtggtggtag tgacattcag 1800

atgacccagt ctcctagctc cctgtccgcc tcagtaggag acagggtcac catcacctgc 1860

aaggccagtc agaatgtggg tattaatgta gcctggtatc aacagaaacc agggaaggct 1920

cctaaatcac tgatttcctc ggcctcctac cggtacagtg gagtcccttc cagattcagc 1980

ggcagtggat ctgggacaga tttcactctc accatcagca gcctccagcc tgaagacttc 2040

gcaacctatt tctgtcagca atatgacacc tatccattca cgttcggctg cggtaccaag 2100

gtggagatca aatga 2115

<210> 49

<211> 704

<212> PRT

<213> Seqüência Artificial

<220>

<223> Descrição de Seqüência Artificial: Construto sintético

<400> 49

Gln Ile Gln Leu Val Gln Ser Gly Pro Glu Leu Lys Lys Pro Gly Glu

1 5 10 15

Thr Val Lys Ile Ser Cys Lys Ala Ser Gly Phe Thr Phe Thr Asp Tyr

20 25 30

Ser Ile His Trp Val Lys Gln Ala Pro Gly Lys Gly Leu Lys Trp Met

35 40 45

Gly Trp Ile Asn Thr Glu Thr Gly Glu Pro Thr Tyr Thr Asp Asp Phe

50 55 60

Lys Gly Arg Phe Ala Phe Ser Leu Val Thr Ser Ala Thr Thr Ala Tyr
65 70 75 80

Leu Gln Ile Asn Asn Leu Asn Asn Glu Asp Thr Ala Thr Phe Phe Cys
85 90 95

Ala Arg Phe Ile Tyr Asp Pro Tyr Trp Gly Phe Ala Tyr Trp Gly Gln
100 105 110

Gly Thr Leu Val Thr Val Ser Ala Ala Ser Thr Lys Gly Pro Ser Val
115 120 125

Phe Pro Leu Ala Pro Ser Ser Lys Ser Thr Ser Gly Gly Thr Ala Ala
130 135 140

Leu Gly Cys Leu Val Lys Asp Tyr Phe Pro Glu Pro Val Thr Val Ser
145 150 155 160

Trp Asn Ser Gly Ala Leu Thr Ser Gly Val His Thr Phe Pro Ala Val
165 170 175

Leu Gln Ser Ser Gly Leu Tyr Ser Leu Ser Ser Val Val Thr Val Pro
180 185 190

Ser Ser Ser Leu Gly Thr Gln Thr Tyr Ile Cys Asn Val Asn His Lys
195 200 205

Pro Ser Asn Thr Lys Val Asp Lys Lys Val Glu Pro Lys Ser Cys Asp
210 215 220

Lys Thr His Thr Cys Pro Pro Cys Pro Ala Pro Glu Leu Leu Gly Gly
225 230 235 240

Pro Ser Val Phe Leu Phe Pro Pro Lys Pro Lys Asp Thr Leu Met Ile
245 250 255

Ser Arg Thr Pro Glu Val Thr Cys Val Val Val Asp Val Ser His Glu
260 265 270

Asp Pro Glu Val Lys Phe Asn Trp Tyr Val Asp Gly Val Glu Val His
275 280 285

Asn Ala Lys Thr Lys Pro Arg Glu Glu Gln Tyr Asn Ser Thr Tyr Arg
290 295 300

Val Val Ser Val Leu Thr Val Leu His Gln Asp Trp Leu Asn Gly Lys
305 310 315 320

Glu Tyr Lys Cys Lys Val Ser Asn Lys Ala Leu Pro Ala Pro Ile Glu
325 330 335

Lys Thr Ile Ser Lys Ala Lys Gly Gln Pro Arg Glu Pro Gln Val Tyr
340 345 350

Thr Leu Pro Pro Ser Arg Asp Glu Leu Thr Lys Asn Gln Val Ser Leu
355 360 365

Thr Cys Leu Val Lys Gly Phe Tyr Pro Ser Asp Ile Ala Val Glu Trp
370 375 380

Glu Ser Asn Gly Gln Pro Glu Asn Asn Tyr Lys Thr Thr Pro Pro Val
385 390 395 400

Leu Asp Ser Asp Gly Ser Phe Phe Leu Tyr Ser Lys Leu Thr Val Asp

405 410 415

Lys Ser Arg Trp Gln Gln Gly Asn Val Phe Ser Cys Ser Val Met His
420 425 430

Glu Ala Leu His Asn His Tyr Thr Gln Lys Ser Leu Ser Leu Ser Pro
435 440 445

Gly Lys Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly Gly Gly
450 455 460

Gly Ser Gln Val Gln Leu Val Gln Ser Gly Ala Glu Val Lys Lys Pro
465 470 475 480

Gly Ser Ser Val Lys Val Ser Cys Lys Ala Ser Gly Tyr Thr Phe Thr
485 490 495

Thr Tyr Tyr Leu His Trp Val Arg Gln Ala Pro Gly Gln Cys Leu Glu
500 505 510

Trp Met Gly Trp Ile Tyr Pro Gly Asn Val His Ala Gln Tyr Asn Glu
515 520 525

Lys Phe Lys Gly Arg Val Thr Ile Thr Ala Asp Lys Ser Thr Ser Thr
530 535 540

Ala Tyr Met Glu Leu Ser Ser Leu Arg Ser Glu Asp Thr Ala Val Tyr
545 550 555 560

Tyr Cys Ala Arg Ser Trp Glu Gly Phe Pro Tyr Trp Gly Gln Gly Thr
565 570 575

Thr Val Thr Val Ser Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser
580 585 590

Gly Gly Gly Gly Ser Asp Ile Gln Met Thr Gln Ser Pro Ser Ser Leu
595 600 605

Ser Ala Ser Val Gly Asp Arg Val Thr Ile Thr Cys Lys Ala Ser Gln
610 615 620

Asn Val Gly Ile Asn Val Ala Trp Tyr Gln Gln Lys Pro Gly Lys Ala
625 630 635 640

Pro Lys Ser Leu Ile Ser Ser Ala Ser Tyr Arg Tyr Ser Gly Val Pro
645 650 655

Ser Arg Phe Ser Gly Ser Gly Ser Gly Thr Asp Phe Thr Leu Thr Ile
660 665 670

Ser Ser Leu Gln Pro Glu Asp Phe Ala Thr Tyr Phe Cys Gln Gln Tyr
675 680 685

Asp Thr Tyr Pro Phe Thr Phe Gly Cys Gly Thr Lys Val Glu Ile Lys
690 695 700

<210> 50
<211> 2130
<212> DNA
<213> Seqüência Artificial
<220>
<223> Descrição de Seqüência Artificial: Construto sintético
<400> 50
cagatccagt tggtgcagtc tggacctgag ctgaagaagc ctggagagac agtcaagatc 60

tcctgcaagg cttctggttt taccttcaca gactattcaa tacactgggt gaaacaggct 120

ccaggaaagg gtttaaagtg gatgggctgg ataaacactg agactggtga gccaacatat 180

acagatgact tcaagggacg atttgccttc tctttggtga cctctgccac cactgcctat 240

ttgcagatca acaacctcaa caatgaggac acggctacat ttttctgtgc tagattcatc 300

tatgatcctt attgggggtt tgcttactgg ggccagggga ctctggtcac tgtctccgca 360

gctagcacca agggcccatc ggtcttcccc ctggcaccct cctccaagag cacctctggg 420

ggcacagcgg ccctgggctg cctggtcaag gactacttcc ccgaaccggt gacggtgtcg 480

tggaactcag gcgccctgac cagcggcgtg cacaccttcc cggctgtcct acagtcctca 540

ggactctact ccctcagcag cgtggtgacc gtgccctcca gcagcttggg cacccagacc 600

tacatctgca acgtgaatca caagcccagc aacaccaagg tggacaagaa agttgagccc 660

aaatcttgtg acaaaactca cacatgccca ccgtgcccag cacctgaact cctgggggga 720

ccgtcagtct tcctcttccc cccaaaaccc aaggacaccc tcatgatctc ccggaccccct 780

gaggtcacat gcgtggtggt ggacgtgagc cacgaagacc ctgaggtcaa gttcaactgg 840

tacgtggacg gcgtggaggt gcataatgcc aagacaaagc cgcgggagga gcagtacaac 900

agcacgtacc gtgtggtcag cgtcctcacc gtcctgcacc aggactggct gaatggcaag 960

gagtacaagt gcaaggtctc caacaaagcc ctcccagccc ccatcgagaa aaccatctcc 1020

aaagccaaag ggcagccccg agaaccacag gtgtacaccc tgcccccatc ccgggatgag 1080

ctgaccaaga accaggtcag cctgacctgc ctggtcaaag gcttctatcc cagcgacatc 1140

gccgtggagt gggagagcaa tgggcagccg gagaacaact acaagaccac gcctcccgtg 1200

ctggactccg acggctcctt cttcctctac agcaagctca ccgtggacaa gagcaggtgg 1260

cagcagggga acgtcttctc atgctccgtg atgcatgagg ctctgcacaa ccactacacg 1320

cagaagagcc tctccctgtc tccgggtaaa tccggcgggg gtggatccgg tggagggggc 1380

tccggcggtg gcgggtccca ggtccaactg gtgcagtctg gagctgaggt gaagaagcct 1440

gggtcctcag tgaaggtgtc ctgcaaggct tctggctaca ctttcacaac ctactatttg 1500

cactgggtga ggcaggcccc tggacagtgc cttgagtgga tgggatggat ttatcctgga 1560

aatgttcatg ctcagtacaa tgagaagttc aagggcaggg tcacaatcac tgcagacaaa 1620

tccaccagca cagcctacat ggagctcagc agcctgaggt ctgaagatac tgcggtctat 1680

tactgtgcaa gatcctggga aggttttcct tactggggcc aagggaccac ggtcaccgtc 1740

tcctcaggcg gtggagggtc cggtgggggc ggatctgggg gcggcgggtc cggtggtggt 1800

ggtagtgaca ttcagatgac ccagtctcct agctccctgt ccgcctcagt aggagacagg 1860

gtcaccatca cctgcaaggc cagtcagaat gtgggtatta atgtagcctg gtatcaacag 1920

aaaccaggga aggctcctaa atcactgatt tcctcggcct cctaccggta cagtggagtc 1980

ccttccagat tcagcggcag tggatctggg acagatttca ctctcaccat cagcagcctc 2040

cagcctgaag acttcgcaac ctatttctgt cagcaatatg acacctatcc attcacgttc 2100

ggctgcggta ccaaggtgga gatcaaatga 2130

<210> 51

<211> 709

<212> PRT

<213> Seqüência Artificial

<220>

<223> Descrição de Seqüência Artificial: Construto sintético

<400> 51

Gln Ile Gln Leu Val Gln Ser Gly Pro Glu Leu Lys Lys Pro Gly Glu

1 5 10 15

Thr Val Lys Ile Ser Cys Lys Ala Ser Gly Phe Thr Phe Thr Asp Tyr

20 25 30

Ser Ile His Trp Val Lys Gln Ala Pro Gly Lys Gly Leu Lys Trp Met

35 40 45

Gly Trp Ile Asn Thr Glu Thr Gly Glu Pro Thr Tyr Thr Asp Asp Phe

50 55 60

Lys Gly Arg Phe Ala Phe Ser Leu Val Thr Ser Ala Thr Thr Ala Tyr

65 70 75 80

Leu Gln Ile Asn Asn Leu Asn Asn Glu Asp Thr Ala Thr Phe Phe Cys

85 90 95

Ala Arg Phe Ile Tyr Asp Pro Tyr Trp Gly Phe Ala Tyr Trp Gly Gln

100 105 110

Gly Thr Leu Val Thr Val Ser Ala Ala Ser Thr Lys Gly Pro Ser Val
115 120 125

Phe Pro Leu Ala Pro Ser Ser Lys Ser Thr Ser Gly Gly Thr Ala Ala
130 135 140

Leu Gly Cys Leu Val Lys Asp Tyr Phe Pro Glu Pro Val Thr Val Ser
145 150 155 160

Trp Asn Ser Gly Ala Leu Thr Ser Gly Val His Thr Phe Pro Ala Val
165 170 175

Leu Gln Ser Ser Gly Leu Tyr Ser Leu Ser Ser Val Val Thr Val Pro
180 185 190

Ser Ser Ser Leu Gly Thr Gln Thr Tyr Ile Cys Asn Val Asn His Lys
195 200 205

Pro Ser Asn Thr Lys Val Asp Lys Lys Val Glu Pro Lys Ser Cys Asp
210 215 220

Lys Thr His Thr Cys Pro Pro Cys Pro Ala Pro Glu Leu Leu Gly Gly
225 230 235 240

Pro Ser Val Phe Leu Phe Pro Pro Lys Pro Lys Asp Thr Leu Met Ile
245 250 255

Ser Arg Thr Pro Glu Val Thr Cys Val Val Val Asp Val Ser His Glu
260 265 270

Asp Pro Glu Val Lys Phe Asn Trp Tyr Val Asp Gly Val Glu Val His
275 280 285

Asn Ala Lys Thr Lys Pro Arg Glu Glu Gln Tyr Asn Ser Thr Tyr Arg
290 295 300

Val Val Ser Val Leu Thr Val Leu His Gln Asp Trp Leu Asn Gly Lys
305 310 315 320

Glu Tyr Lys Cys Lys Val Ser Asn Lys Ala Leu Pro Ala Pro Ile Glu
325 330 335

Lys Thr Ile Ser Lys Ala Lys Gly Gln Pro Arg Glu Pro Gln Val Tyr
340 345 350

Thr Leu Pro Pro Ser Arg Asp Glu Leu Thr Lys Asn Gln Val Ser Leu
355 360 365

Thr Cys Leu Val Lys Gly Phe Tyr Pro Ser Asp Ile Ala Val Glu Trp
370 375 380

Glu Ser Asn Gly Gln Pro Glu Asn Asn Tyr Lys Thr Thr Pro Pro Val
385 390 395 400

Leu Asp Ser Asp Gly Ser Phe Phe Leu Tyr Ser Lys Leu Thr Val Asp
405 410 415

Lys Ser Arg Trp Gln Gln Gly Asn Val Phe Ser Cys Ser Val Met His
420 425 430

Glu Ala Leu His Asn His Tyr Thr Gln Lys Ser Leu Ser Leu Ser Pro
435 440 445

Gly Lys Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly Gly Gly

450 455 460

Gly Ser Gln Val Gln Leu Val Gln Ser Gly Ala Glu Val Lys Lys Pro
465 470 475 480

Gly Ser Ser Val Lys Val Ser Cys Lys Ala Ser Gly Tyr Thr Phe Thr
485 490 495

Thr Tyr Tyr Leu His Trp Val Arg Gln Ala Pro Gly Gln Cys Leu Glu
500 505 510

Trp Met Gly Trp Ile Tyr Pro Gly Asn Val His Ala Gln Tyr Asn Glu
515 520 525

Lys Phe Lys Gly Arg Val Thr Ile Thr Ala Asp Lys Ser Thr Ser Thr
530 535 540

Ala Tyr Met Glu Leu Ser Ser Leu Arg Ser Glu Asp Thr Ala Val Tyr
545 550 555 560

Tyr Cys Ala Arg Ser Trp Glu Gly Phe Pro Tyr Trp Gly Gln Gly Thr
565 570 575

Thr Val Thr Val Ser Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser
580 585 590

Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Asp Ile Gln Met Thr Gln
595 600 605

Ser Pro Ser Ser Leu Ser Ala Ser Val Gly Asp Arg Val Thr Ile Thr
610 615 620

Cys Lys Ala Ser Gln Asn Val Gly Ile Asn Val Ala Trp Tyr Gln Gln
625 630 635 640

Lys Pro Gly Lys Ala Pro Lys Ser Leu Ile Ser Ser Ala Ser Tyr Arg
645 650 655

Tyr Ser Gly Val Pro Ser Arg Phe Ser Gly Ser Gly Ser Gly Thr Asp
660 665 670

Phe Thr Leu Thr Ile Ser Ser Leu Gln Pro Glu Asp Phe Ala Thr Tyr
675 680 685

Phe Cys Gln Gln Tyr Asp Thr Tyr Pro Phe Thr Phe Gly Cys Gly Thr
690 695 700

Lys Val Glu Ile Lys
705

<210> 52
<211> 2130
<212> DNA
<213> Seqüência Artificial
<220>
<223> Descrição de Seqüência Artificial: Construto sintético
<400> 52
cagatccagt tggtgcagtc tggacctgag ctgaagaagc ctggagagac agtcaagatc 60
tcctgcaagg cttctggttt taccttcaca gactattcaa tacactgggt gaaacaggct 120
ccaggaaagg gtttaaagtg gatgggctgg ataaacactg agactggtga gccaacatat 180
acagatgact tcaagggacg atttgccttc tctttggtga cctctgccac cactgcctat 240
ttgcagatca acaacctcaa caatgaggac acggctacat ttttctgtgc tagattcatc 300
tatgatcctt attgggggtt tgcttactgg ggccagggga ctctggtcac tgtctccgca 360
gctagcacca agggcccatc ggtcttcccc ctggcaccct cctccaagag cacctctggg 420

ggcacagcgg ccctgggctg cctggtcaag gactacttcc ccgaaccggt gacggtgtcg 480

tggaactcag gcgccctgac cagcggcgtg cacaccttcc cggctgtcct acagtcctca 540

ggactctact ccctcagcag cgtggtgacc gtgccctcca gcagcttggg cacccagacc 600

tacatctgca acgtgaatca caagcccagc aacaccaagg tggacaagaa agttgagccc 660

aaatcttgtg acaaaactca cacatgccca ccgtgcccag cacctgaact cctgggggga 720

ccgtcagtct tcctcttccc cccaaaaccc aaggacaccc tcatgatctc ccggacccct 780

gaggtcacat gcgtggtggt ggacgtgagc cacgaagacc ctgaggtcaa gttcaactgg 840

tacgtggacg gcgtggaggt gcataatgcc aagacaaagc cgcgggagga gcagtacaac 900

agcacgtacc gtgtggtcag cgtcctcacc gtcctgcacc aggactggct gaatggcaag 960

gagtacaagt gcaaggtctc caacaaagcc ctcccagccc ccatcgagaa aaccatctcc 1020

aaagccaaag ggcagccccg agaaccacag gtgtacaccc tgcccccatc ccgggatgag 1080

ctgaccaaga accaggtcag cctgacctgc ctggtcaaag gcttctatcc cagcgacatc 1140

gccgtggagt gggagagcaa tgggcagccg gagaacaact acaagaccac gcctcccgtg 1200

ctggactccg acggctcctt cttcctctac agcaagctca ccgtggacaa gagcaggtgg 1260

cagcagggga acgtcttctc atgctccgtg atgcatgagg ctctgcacaa ccactacacg 1320

cagaagagcc tctccctgtc tccgggtaaa tccggcgggg gtggatccgg tggagggggc 1380

tccggcggtg gcgggtccca ggtccaactg gtgcagtctg gagctgaggt gaagaagcct 1440

ggggagtcag tgaaggtgtc ctgcaaggct tctggctaca ctttcacaac ctactatttg 1500

cactgggtga ggcaggcccc tggacaggga cttgagtgga tgggatggat ttatcctgga 1560

aatgttcatg ctcagtacaa tgagaagttc aagggcaggg tcacaatcac tgcagacaaa 1620

tccaccagca cagcctacat ggagctcagc agcctgaggt ctgaagatac tgcggtctat 1680

tactgtgcaa gatcctggga aggttttcct tactggggcc aagggaccac ggtcaccgtc 1740

tcctcaggcg gtggagggtc cggtgggggc ggatctgggg gcggcgggtc cggtggtggt 1800

ggtagtgaca ttcagatgac ccagtctcct agctccctgt ccgcctcagt aggagacagg 1860

gtcaccatca cctgcaaggc cagtcagaat gtgggtatta atgtagcctg gtatcaacag 1920

aaaccaggga aggctcctaa attactgatt tcctcggcct cctaccggta cagtggagtc 1980

ccttccagat tcagcggcag tggatctggg acagatttca ctctcaccat cagcagcctc 2040

cagcctgaag acttcgcaac ctatttctgt cagcaatatg acacctatcc attcacgttc 2100

ggccagggta ccaaggtgga gatcaaatga 2130

<210> 53

<211> 709

<212> PRT

<213> Seqüência Artificial

<220>

<223> Descrição de Seqüência Artificial: Construto sintético

<400> 53

Gln Ile Gln Leu Val Gln Ser Gly Pro Glu Leu Lys Lys Pro Gly Glu
1 5 10 15

Thr Val Lys Ile Ser Cys Lys Ala Ser Gly Phe Thr Phe Thr Asp Tyr
20 25 30

Ser Ile His Trp Val Lys Gln Ala Pro Gly Lys Gly Leu Lys Trp Met
35 40 45

Gly Trp Ile Asn Thr Glu Thr Gly Glu Pro Thr Tyr Thr Asp Asp Phe
50 55 60

Lys Gly Arg Phe Ala Phe Ser Leu Val Thr Ser Ala Thr Thr Ala Tyr
65 70 75 80

Leu Gln Ile Asn Asn Leu Asn Asn Glu Asp Thr Ala Thr Phe Phe Cys
85 90 95

Ala Arg Phe Ile Tyr Asp Pro Tyr Trp Gly Phe Ala Tyr Trp Gly Gln
100 105 110

Gly Thr Leu Val Thr Val Ser Ala Ala Ser Thr Lys Gly Pro Ser Val
115 120 125

Phe Pro Leu Ala Pro Ser Ser Lys Ser Thr Ser Gly Gly Thr Ala Ala
130 135 140

Leu Gly Cys Leu Val Lys Asp Tyr Phe Pro Glu Pro Val Thr Val Ser
145 150 155 160

Trp Asn Ser Gly Ala Leu Thr Ser Gly Val His Thr Phe Pro Ala Val
165 170 175

Leu Gln Ser Ser Gly Leu Tyr Ser Leu Ser Ser Val Val Thr Val Pro
180 185 190

Ser Ser Ser Leu Gly Thr Gln Thr Tyr Ile Cys Asn Val Asn His Lys
195 200 205

Pro Ser Asn Thr Lys Val Asp Lys Lys Val Glu Pro Lys Ser Cys Asp
210 215 220

Lys Thr His Thr Cys Pro Pro Cys Pro Ala Pro Glu Leu Leu Gly Gly
225 230 235 240

Pro Ser Val Phe Leu Phe Pro Pro Lys Pro Lys Asp Thr Leu Met Ile
245 250 255

Ser Arg Thr Pro Glu Val Thr Cys Val Val Val Asp Val Ser His Glu
260 265 270

Asp Pro Glu Val Lys Phe Asn Trp Tyr Val Asp Gly Val Glu Val His
275 280 285

Asn Ala Lys Thr Lys Pro Arg Glu Glu Gln Tyr Asn Ser Thr Tyr Arg
290 295 300

Val Val Ser Val Leu Thr Val Leu His Gln Asp Trp Leu Asn Gly Lys
305 310 315 320

Glu Tyr Lys Cys Lys Val Ser Asn Lys Ala Leu Pro Ala Pro Ile Glu
325 330 335

Lys Thr Ile Ser Lys Ala Lys Gly Gln Pro Arg Glu Pro Gln Val Tyr
340 345 350

Thr Leu Pro Pro Ser Arg Asp Glu Leu Thr Lys Asn Gln Val Ser Leu
355 360 365

Thr Cys Leu Val Lys Gly Phe Tyr Pro Ser Asp Ile Ala Val Glu Trp
370 375 380

Glu Ser Asn Gly Gln Pro Glu Asn Asn Tyr Lys Thr Thr Pro Pro Val
385 390 395 400

Leu Asp Ser Asp Gly Ser Phe Phe Leu Tyr Ser Lys Leu Thr Val Asp
405 410 415

Lys Ser Arg Trp Gln Gln Gly Asn Val Phe Ser Cys Ser Val Met His
420 425 430

Glu Ala Leu His Asn His Tyr Thr Gln Lys Ser Leu Ser Leu Ser Pro
435 440 445

Gly Lys Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly Gly Gly
450 455 460

Gly Ser Gln Val Gln Leu Val Gln Ser Gly Ala Glu Val Lys Lys Pro
465 470 475 480

Gly Glu Ser Val Lys Val Ser Cys Lys Ala Ser Gly Tyr Thr Phe Thr

485 490 495

Thr Tyr Tyr Leu His Trp Val Arg Gln Ala Pro Gly Gln Gly Leu Glu
500 505 510

Trp Met Gly Trp Ile Tyr Pro Gly Asn Val His Ala Gln Tyr Asn Glu
515 520 525

Lys Phe Lys Gly Arg Val Thr Ile Thr Ala Asp Lys Ser Thr Ser Thr
530 535 540

Ala Tyr Met Glu Leu Ser Ser Leu Arg Ser Glu Asp Thr Ala Val Tyr
545 550 555 560

Tyr Cys Ala Arg Ser Trp Glu Gly Phe Pro Tyr Trp Gly Gln Gly Thr
565 570 575

Thr Val Thr Val Ser Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser
580 585 590

Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Asp Ile Gln Met Thr Gln
595 600 605

Ser Pro Ser Ser Leu Ser Ala Ser Val Gly Asp Arg Val Thr Ile Thr
610 615 620

Cys Lys Ala Ser Gln Asn Val Gly Ile Asn Val Ala Trp Tyr Gln Gln
625 630 635 640

Lys Pro Gly Lys Ala Pro Lys Leu Leu Ile Ser Ser Ala Ser Tyr Arg
645 650 655

Tyr Ser Gly Val Pro Ser Arg Phe Ser Gly Ser Gly Ser Gly Thr Asp

660 665 670

Phe Thr Leu Thr Ile Ser Ser Leu Gln Pro Glu Asp Phe Ala Thr Tyr

675 680 685

Phe Cys Gln Gln Tyr Asp Thr Tyr Pro Phe Thr Phe Gly Gln Gly Thr

690 695 700

Lys Val Glu Ile Lys

705

<210> 54

<211> 2130

<212> DNA

<213> Seqüência Artificial

<220>

<223> Descrição de Seqüência Artificial: Construto sintético

<400> 54

cagatccagt tggtgcagtc tggacctgag ctgaagaagc ctggagagac agtcaagatc 60

tcctgcaagg cttctggttt taccttcaca gactattcaa tacactgggt gaaacaggct 120

ccaggaaagg gtttaaagtg gatgggctgg ataaacactg agactggtga gccaacatat 180

acagatgact tcaagggacg atttgccttc tctttggtga cctctgccac cactgcctat 240

ttgcagatca acaacctcaa caatgaggac acggctacat ttttctgtgc tagattcatc 300

tatgatcctt attgggggtt tgcttactgg ggccagggga ctctggtcac tgtctccgca 360

gctagcacca agggcccatc ggtcttcccc ctggcaccct cctccaagag cacctctggg 420

ggcacagcgg ccctgggctg cctggtcaag gactacttcc ccgaaccggt gacggtgtcg 480

tggaactcag gcgccctgac cagcggcgtg cacaccttcc cggctgtcct acagtcctca 540

ggactctact ccctcagcag cgtggtgacc gtgccctcca gcagcttggg cacccagacc 600

tacatctgca acgtgaatca caagcccagc aacaccaagg tggacaagaa agttgagccc 660

aaatcttgtg acaaaactca cacatgccca ccgtgcccag cacctgaact cctgggggga 720

ccgtcagtct tcctcttccc cccaaaaccc aaggacaccc tcatgatctc ccggaccccct 780

gaggtcacat gcgtggtggt ggacgtgagc cacgaagacc ctgaggtcaa gttcaactgg 840

tacgtggacg gcgtggaggt gcataatgcc aagacaaagc cgcgggagga gcagtacaac 900

agcacgtacc gtgtggtcag cgtcctcacc gtcctgcacc aggactggct gaatggcaag 960

gagtacaagt gcaaggtctc caacaaagcc ctcccagccc ccatcgagaa aaccatctcc 1020

aaagccaaag ggcagccccg agaaccacag gtgtacaccc tgcccccatc ccgggatgag 1080

ctgaccaaga accaggtcag cctgacctgc ctggtcaaag gcttctatcc cagcgacatc 1140

gccgtggagt gggagagcaa tgggcagccg gagaacaact acaagaccac gcctcccgtg 1200

ctggactccg acggctcctt cttcctctac agcaagctca ccgtggacaa gagcaggtgg 1260

cagcagggga acgtcttctc atgctccgtg atgcatgagg ctctgcacaa ccactacacg 1320

cagaagagcc tctccctgtc tccgggtaaa tccggcgggg gtggatccgg tggaggggggc 1380

tccggcggtg gcgggtccca ggtccaactg gtgcagtctg gagctgaggt gaagaagcct 1440

ggggagtcag tgaaggtgtc ctgcaaggct tctggctaca ctttcacaac ctactatttg 1500

cactgggtga ggcaggcccc tggacagtgc cttgagtgga tgggatggat ttatcctgga 1560

aatgttcatg ctcagtacaa tgagaagttc aagggcaggg tcacaatcac tgcagacaaa 1620

tccaccagca cagcctacat ggagctcagc agcctgaggt ctgaagatac tgcggtctat 1680

tactgtgcaa gatcctggga aggttttcct tactggggcc aagggaccac ggtcaccgtc 1740

tcctcaggcg gtggagggtc cggtgggggc ggatctgggg gcggcgggtc cggtggtggt 1800

ggtagtgaca ttcagatgac ccagtctcct agctccctgt ccgcctcagt aggagacagg 1860

gtcaccatca cctgcaaggc cagtcagaat gtgggtatta atgtagcctg gtatcaacag 1920

aaaccaggga aggctcctaa attactgatt tcctcggcct cctaccggta cagtggagtc 1980

ccttccagat tcagcggcag tggatctggg acagatttca ctctcaccat cagcagcctc 2040

cagcctgaag acttcgcaac ctatttctgt cagcaatatg acacctatcc attcacgttc 2100

ggctgcggta ccaaggtgga gatcaaatga 2130

<210> 55

<211> 709

<212> PRT

<213> Seqüência Artificial

<220>

<223> Descrição de Seqüência Artificial: Construto sintético

<400> 55

Gln Ile Gln Leu Val Gln Ser Gly Pro Glu Leu Lys Lys Pro Gly Glu

1 5 10 15

Thr Val Lys Ile Ser Cys Lys Ala Ser Gly Phe Thr Phe Thr Asp Tyr
20 25 30

Ser Ile His Trp Val Lys Gln Ala Pro Gly Lys Gly Leu Lys Trp Met
35 40 45

Gly Trp Ile Asn Thr Glu Thr Gly Glu Pro Thr Tyr Thr Asp Asp Phe
50 55 60

Lys Gly Arg Phe Ala Phe Ser Leu Val Thr Ser Ala Thr Thr Ala Tyr
65 70 75 80

Leu Gln Ile Asn Asn Leu Asn Asn Glu Asp Thr Ala Thr Phe Phe Cys
85 90 95

Ala Arg Phe Ile Tyr Asp Pro Tyr Trp Gly Phe Ala Tyr Trp Gly Gln
100 105 110

Gly Thr Leu Val Thr Val Ser Ala Ala Ser Thr Lys Gly Pro Ser Val
115 120 125

Phe Pro Leu Ala Pro Ser Ser Lys Ser Thr Ser Gly Gly Thr Ala Ala
130 135 140

Leu Gly Cys Leu Val Lys Asp Tyr Phe Pro Glu Pro Val Thr Val Ser
145 150 155 160

Trp Asn Ser Gly Ala Leu Thr Ser Gly Val His Thr Phe Pro Ala Val
165 170 175

Leu Gln Ser Ser Gly Leu Tyr Ser Leu Ser Ser Val Val Thr Val Pro
180 185 190

Ser Ser Ser Leu Gly Thr Gln Thr Tyr Ile Cys Asn Val Asn His Lys
195 200 205

Pro Ser Asn Thr Lys Val Asp Lys Lys Val Glu Pro Lys Ser Cys Asp
210 215 220

Lys Thr His Thr Cys Pro Pro Cys Pro Ala Pro Glu Leu Leu Gly Gly
225 230 235 240

Pro Ser Val Phe Leu Phe Pro Pro Lys Pro Lys Asp Thr Leu Met Ile
245 250 255

Ser Arg Thr Pro Glu Val Thr Cys Val Val Val Asp Val Ser His Glu
260 265 270

Asp Pro Glu Val Lys Phe Asn Trp Tyr Val Asp Gly Val Glu Val His
275 280 285

Asn Ala Lys Thr Lys Pro Arg Glu Glu Gln Tyr Asn Ser Thr Tyr Arg
290 295 300

Val Val Ser Val Leu Thr Val Leu His Gln Asp Trp Leu Asn Gly Lys
305 310 315 320

Glu Tyr Lys Cys Lys Val Ser Asn Lys Ala Leu Pro Ala Pro Ile Glu
325 330 335

Lys Thr Ile Ser Lys Ala Lys Gly Gln Pro Arg Glu Pro Gln Val Tyr
340 345 350

Thr Leu Pro Pro Ser Arg Asp Glu Leu Thr Lys Asn Gln Val Ser Leu
355 360 365

Thr Cys Leu Val Lys Gly Phe Tyr Pro Ser Asp Ile Ala Val Glu Trp
370 375 380

Glu Ser Asn Gly Gln Pro Glu Asn Asn Tyr Lys Thr Thr Pro Pro Val
385 390 395 400

Leu Asp Ser Asp Gly Ser Phe Phe Leu Tyr Ser Lys Leu Thr Val Asp
405 410 415

Lys Ser Arg Trp Gln Gln Gly Asn Val Phe Ser Cys Ser Val Met His
420 425 430

Glu Ala Leu His Asn His Tyr Thr Gln Lys Ser Leu Ser Leu Ser Pro
435 440 445

Gly Lys Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly Gly Gly
450 455 460

Gly Ser Gln Val Gln Leu Val Gln Ser Gly Ala Glu Val Lys Lys Pro
465 470 475 480

Gly Glu Ser Val Lys Val Ser Cys Lys Ala Ser Gly Tyr Thr Phe Thr
485 490 495

Thr Tyr Tyr Leu His Trp Val Arg Gln Ala Pro Gly Gln Cys Leu Glu
500 505 510

Trp Met Gly Trp Ile Tyr Pro Gly Asn Val His Ala Gln Tyr Asn Glu

515 520 525

Lys Phe Lys Gly Arg Val Thr Ile Thr Ala Asp Lys Ser Thr Ser Thr
530 535 540

Ala Tyr Met Glu Leu Ser Ser Leu Arg Ser Glu Asp Thr Ala Val Tyr
545 550 555 560

Tyr Cys Ala Arg Ser Trp Glu Gly Phe Pro Tyr Trp Gly Gln Gly Thr
565 570 575

Thr Val Thr Val Ser Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser
580 585 590

Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Asp Ile Gln Met Thr Gln
595 600 605

Ser Pro Ser Ser Leu Ser Ala Ser Val Gly Asp Arg Val Thr Ile Thr
610 615 620

Cys Lys Ala Ser Gln Asn Val Gly Ile Asn Val Ala Trp Tyr Gln Gln
625 630 635 640

Lys Pro Gly Lys Ala Pro Lys Leu Leu Ile Ser Ser Ala Ser Tyr Arg
645 650 655

Tyr Ser Gly Val Pro Ser Arg Phe Ser Gly Ser Gly Ser Gly Thr Asp
660 665 670

Phe Thr Leu Thr Ile Ser Ser Leu Gln Pro Glu Asp Phe Ala Thr Tyr
675 680 685

Phe Cys Gln Gln Tyr Asp Thr Tyr Pro Phe Thr Phe Gly Cys Gly Thr
690 695 700

Lys Val Glu Ile Lys
705

<210> 56
<211> 64
<212> DNA
<213> Seqüência Artificial
<220>
<223> Descrição de Seqüência Artificial: Oligonucleotídeo sintético
<400> 56
gggggtggat ccggtggagg gggctccggc ggtggcgggt cccaggtcca actggtgcag 60
tctg 64

<210> 57
<211> 32
<212> DNA
<213> Seqüência Artificial
<220>
<223> Descrição de Seqüência Artificial: Oligonucleotídeo sintético
<400> 57
gttaacggat cctcatttga tctccacctt gg 32

<210> 58
<211> 29
<212> PRT
<213> Seqüência Artificial
<220>
<223> Descrição de Seqüência Artificial: Peptídeo sintético
<400> 58

Asp Asp Asp Asp Lys Ser Phe Leu Glu Gln Lys Leu Ile Ser Glu Glu

1 5 10 15

Asp Leu Asn Ser Ala Val Asp His His His His His His

20 25

<210> 59

<211> 57

<212> DNA

<213> Seqüência Artificial

<220>

<223> Descrição de Seqüência Artificial: Oligonucleotídeo sintético

<400> 59

atggcctgga ctcctctctt cttcttcttt gttcttcatt gctcagggtc tttctcc 57

<210> 60

<211> 19

<212> PRT

<213> Seqüência Artificial

<220>

<223> Descrição de Seqüência Artificial: Peptídeo sintético

<400> 60

Met Ala Trp Thr Pro Leu Phe Phe Phe Phe Val Leu His Cys Ser Gly

1 5 10 15

Ser Phe Ser

<210> 61

<211> 57

<212> DNA

<213> Seqüência Artificial

<220>

<223> Descrição de Seqüência Artificial: Oligonucleotídeo sintético

<400> 61

atgggttgga gcctcatctt gctcttcctt gtcgctgttg ctacgcgtgt cctgtcc 57

<210> 62

<211> 19

<212> PRT

<213> Seqüência Artificial

<220>

<223> Descrição de Seqüência Artificial: Peptídeo sintético

<400> 62

Met Gly Trp Ser Leu Ile Leu Leu Phe Leu Val Ala Val Ala Thr Arg

1 5 10 15

Val Leu Ser

<210> 63

<211> 2115

<212> DNA

<213> Seqüência Artificial

<220>

<223> Descrição de Seqüência Artificial: Construto sintético

<400> 63

gaggtgcagc tggtggagtc tgggggcggc ttggcaaagc ctgggggtc cctgagactc 60

tcctgcgcag cctccgggtt caggttcacc ttcaataact actacatgga ctgggtccgc 120

caggctccag ggcaggggct ggagtgggtc tcacgtatta gtagtagtgg tgatcccaca 180

tggtacgcag actccgtgaa gggcagattc accatctcca gagagaacgc caagaacaca 240

ctgtttcttc aaatgaacag cctgagagct gaggacacgg ctgtctatta ctgtgcgagc 300

ttgactacag ggtctgactc ctggggccag ggagtcctgg tcaccgtctc ctcagctagc 360

accaagggcc catcggtctt ccccctggca ccctcctcca agagcacctc tgggggcaca 420

gcggccctgg gctgcctggt caaggactac ttccccgaac cggtgacggt gtcgtggaac 480

tcaggcgccc tgaccagcgg cgtgcacacc ttcccggctg tcctacagtc ctcaggactc 540

tactccctca gcagcgtggt gaccgtgccc tccagcagct tgggcaccca gacctacatc 600

tgcaacgtga atcacaagcc cagcaacacc aaggtggaca agaaagttga gcccaaatct 660

tgtgacaaaa ctcacacatg cccaccgtgc ccagcacctg aactcctggg gggaccgtca 720

gtcttcctct tccccccaaa acccaaggac accctcatga tctcccggac ccctgaggtc 780

acatgcgtgg tggtggacgt gagccacgaa gaccctgagg tcaagttcaa ctggtacgtg 840

gacggcgtgg aggtgcataa tgccaagaca aagccgcggg aggagcagta caacagcacg 900

taccgtgtgg tcagcgtcct caccgtcctg caccaggact ggctgaatgg caaggagtac 960

aagtgcaagg tctccaacaa agccctccca gcccccatcg agaaaaccat ctccaaagcc 1020

aaagggcagc cccgagaacc acaggtgtac accctgcccc catcccggga tgagctgacc 1080

aagaaccagg tcagcctgac ctgcctggtc aaaggcttct atcccagcga catcgccgtg 1140

gagtgggaga gcaatgggca gccggagaac aactacaaga ccacgcctcc cgtgctggac 1200

tccgacggct ccttcttcct ctacagcaag ctcaccgtgg acaagagcag gtggcagcag 1260

gggaacgtct tctcatgctc cgtgatgcat gaggctctgc acaaccacta cacgcagaag 1320

agcctctccc tgtctccggg taaatccggc gggggtggat ccggtggagg gggctccggc 1380

ggtggcgggt ccgacatcca gatgacccag tctccatctt ccctgtctgc atctgtaggg 1440

gacagagtca ccatcacttg cagggcaagt caggacatta ggtattattt aaattggtat 1500

cagcagaaac caggaaaagc tcctaagctc ctgatctatg ttgcatccag tttgcaaagt 1560

ggggtcccat caaggttcag cggcagtgga tctgggacag agttcactct caccgtcagc 1620

agcctgcagc ctgaagattt tgcgacttat tactgtctac aggtttatag taccctcgg 1680

acgttcggcc aagggaccaa ggtggaaatc aaaggtgggg gcggatctgg gggcggcggg 1740

tccggtggtg gtggtagtga ggtgcagctg gtggagtctg ggggcggctt ggcaaagcct 1800

ggggggtccc tgagactctc ctgcgcagcc tccgggttca ggttcacctt caataactac 1860

tacatggact gggtccgcca ggctccaggg caggggctgg agtgggtctc acgtattagt 1920

agtagtggtg atcccacatg gtacgcagac tccgtgaagg gcagattcac catctccaga 1980

gagaacgcca agaacacact gtttcttcaa atgaacagcc tgagagctga ggacacggct 2040

gtctattact gtgcgagctt gactacaggg tctgactcct ggggccaggg agtcctggtc 2100

accgtctcct catga 2115

<210> 64

<211> 704

<212> PRT

<213> Seqüência Artificial

<220>

<223> Descrição de Seqüência Artificial: Construto sintético

<400> 64

Glu Val Gln Leu Val Glu Ser Gly Gly Gly Leu Ala Lys Pro Gly Gly
1 5 10 15

Ser Leu Arg Leu Ser Cys Ala Ala Ser Gly Phe Arg Phe Thr Phe Asn
20 25 30

Asn Tyr Tyr Met Asp Trp Val Arg Gln Ala Pro Gly Gln Gly Leu Glu
35 40 45

Trp Val Ser Arg Ile Ser Ser Ser Gly Asp Pro Thr Trp Tyr Ala Asp
50 55 60

Ser Val Lys Gly Arg Phe Thr Ile Ser Arg Glu Asn Ala Lys Asn Thr
65 70 75 80

Leu Phe Leu Gln Met Asn Ser Leu Arg Ala Glu Asp Thr Ala Val Tyr
85 90 95

Tyr Cys Ala Ser Leu Thr Thr Gly Ser Asp Ser Trp Gly Gln Gly Val
100 105 110

Leu Val Thr Val Ser Ser Ala Ser Thr Lys Gly Pro Ser Val Phe Pro
115 120 125

Leu Ala Pro Ser Ser Lys Ser Thr Ser Gly Gly Thr Ala Ala Leu Gly
130 135 140

Cys Leu Val Lys Asp Tyr Phe Pro Glu Pro Val Thr Val Ser Trp Asn
145 150 155 160

Ser Gly Ala Leu Thr Ser Gly Val His Thr Phe Pro Ala Val Leu Gln
165 170 175

Ser Ser Gly Leu Tyr Ser Leu Ser Ser Val Val Thr Val Pro Ser Ser
180 185 190

Ser Leu Gly Thr Gln Thr Tyr Ile Cys Asn Val Asn His Lys Pro Ser
195 200 205

Asn Thr Lys Val Asp Lys Lys Val Glu Pro Lys Ser Cys Asp Lys Thr
210 215 220

His Thr Cys Pro Pro Cys Pro Ala Pro Glu Leu Leu Gly Gly Pro Ser
225 230 235 240

Val Phe Leu Phe Pro Pro Lys Pro Lys Asp Thr Leu Met Ile Ser Arg
245 250 255

Thr Pro Glu Val Thr Cys Val Val Val Asp Val Ser His Glu Asp Pro
260 265 270

Glu Val Lys Phe Asn Trp Tyr Val Asp Gly Val Glu Val His Asn Ala
275 280 285

Lys Thr Lys Pro Arg Glu Glu Gln Tyr Asn Ser Thr Tyr Arg Val Val
290 295 300

Ser Val Leu Thr Val Leu His Gln Asp Trp Leu Asn Gly Lys Glu Tyr
305 310 315 320

Lys Cys Lys Val Ser Asn Lys Ala Leu Pro Ala Pro Ile Glu Lys Thr
325 330 335

Ile Ser Lys Ala Lys Gly Gln Pro Arg Glu Pro Gln Val Tyr Thr Leu
340 345 350

Pro Pro Ser Arg Asp Glu Leu Thr Lys Asn Gln Val Ser Leu Thr Cys
355 360 365

Leu Val Lys Gly Phe Tyr Pro Ser Asp Ile Ala Val Glu Trp Glu Ser
370 375 380

Asn Gly Gln Pro Glu Asn Asn Tyr Lys Thr Thr Pro Pro Val Leu Asp
385 390 395 400

Ser Asp Gly Ser Phe Phe Leu Tyr Ser Lys Leu Thr Val Asp Lys Ser
405 410 415

Arg Trp Gln Gln Gly Asn Val Phe Ser Cys Ser Val Met His Glu Ala
420 425 430

Leu His Asn His Tyr Thr Gln Lys Ser Leu Ser Leu Ser Pro Gly Lys
435 440 445

Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser
450 455 460

Asp Ile Gln Met Thr Gln Ser Pro Ser Ser Leu Ser Ala Ser Val Gly
465 470 475 480

Asp Arg Val Thr Ile Thr Cys Arg Ala Ser Gln Asp Ile Arg Tyr Tyr

485 490 495

Leu Asn Trp Tyr Gln Gln Lys Pro Gly Lys Ala Pro Lys Leu Leu Ile
500 505 510

Tyr Val Ala Ser Ser Leu Gln Ser Gly Val Pro Ser Arg Phe Ser Gly
515 520 525

Ser Gly Ser Gly Thr Glu Phe Thr Leu Thr Val Ser Ser Leu Gln Pro
530 535 540

Glu Asp Phe Ala Thr Tyr Tyr Cys Leu Gln Val Tyr Ser Thr Pro Arg
545 550 555 560

Thr Phe Gly Gln Gly Thr Lys Val Glu Ile Lys Gly Gly Gly Gly Ser
565 570 575

Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Glu Val Gln Leu Val Glu
580 585 590

Ser Gly Gly Gly Leu Ala Lys Pro Gly Gly Ser Leu Arg Leu Ser Cys
595 600 605

Ala Ala Ser Gly Phe Arg Phe Thr Phe Asn Asn Tyr Tyr Met Asp Trp
610 615 620

Val Arg Gln Ala Pro Gly Gln Gly Leu Glu Trp Val Ser Arg Ile Ser
625 630 635 640

Ser Ser Gly Asp Pro Thr Trp Tyr Ala Asp Ser Val Lys Gly Arg Phe
645 650 655

Thr Ile Ser Arg Glu Asn Ala Lys Asn Thr Leu Phe Leu Gln Met Asn
660 665 670

Ser Leu Arg Ala Glu Asp Thr Ala Val Tyr Tyr Cys Ala Ser Leu Thr
675 680 685

Thr Gly Ser Asp Ser Trp Gly Gln Gly Val Leu Val Thr Val Ser Ser
690 695 700

<210> 65
<211> 645
<212> DNA
<213> Seqüência Artificial
<220>
<223> Descrição de Seqüência Artificial: Construto sintético
<400> 65
gacatccaga tgacccagtc tccatcttcc ctgtctgcat ctgtagggga cagagtcacc 60
atcacttgca gggcaagtca ggacattagg tattatttaa attggtatca gcagaaacca 120
ggaaaagctc ctaagctcct gatctatgtt gcatccagtt tgcaaagtgg ggtcccatca 180
aggttcagcg gcagtggatc tgggacagag ttcactctca ccgtcagcag cctgcagcct 240
gaagattttg cgacttatta ctgtctacag gtttatagta cccctcggac gttcggccaa 300
gggaccaagg tggaaatcaa acgtacggtg gctgcaccat ctgtcttcat cttcccgcca 360
tctgatgagc agttgaaatc tggaactgcc tctgttgtgt gcctgctgaa taacttctat 420
cccagagagg ccaaagtaca gtggaaggtg gataacgccc tccaatcggg taactcccag 480
gagagtgtca cagagcagga cagcaaggac agcacctaca gcctcagcag caccctgacg 540
ctgagcaaag cagactacga gaaacacaaa gtctacgcct gcgaagtcac ccatcagggc 600
ctgagctcgc ccgtcacaaa gagcttcaac aggggagagt gttga 645

<210> 66
<211> 214
<212> PRT
<213> Seqüência Artificial

<220>

<223> Descrição de Seqüência Artificial: Construto sintético

<400> 66

Asp Ile Gln Met Thr Gln Ser Pro Ser Ser Leu Ser Ala Ser Val Gly
1 5 10 15

Asp Arg Val Thr Ile Thr Cys Arg Ala Ser Gln Asp Ile Arg Tyr Tyr
20 25 30

Leu Asn Trp Tyr Gln Gln Lys Pro Gly Lys Ala Pro Lys Leu Leu Ile
35 40 45

Tyr Val Ala Ser Ser Leu Gln Ser Gly Val Pro Ser Arg Phe Ser Gly
50 55 60

Ser Gly Ser Gly Thr Glu Phe Thr Leu Thr Val Ser Ser Leu Gln Pro
65 70 75 80

Glu Asp Phe Ala Thr Tyr Tyr Cys Leu Gln Val Tyr Ser Thr Pro Arg
85 90 95

Thr Phe Gly Gln Gly Thr Lys Val Glu Ile Lys Arg Thr Val Ala Ala
100 105 110

Pro Ser Val Phe Ile Phe Pro Pro Ser Asp Glu Gln Leu Lys Ser Gly
115 120 125

Thr Ala Ser Val Val Cys Leu Leu Asn Asn Phe Tyr Pro Arg Glu Ala
130 135 140

Lys Val Gln Trp Lys Val Asp Asn Ala Leu Gln Ser Gly Asn Ser Gln
145 150 155 160

Glu Ser Val Thr Glu Gln Asp Ser Lys Asp Ser Thr Tyr Ser Leu Ser
165 170 175

Ser Thr Leu Thr Leu Ser Lys Ala Asp Tyr Glu Lys His Lys Val Tyr
180 185 190

Ala Cys Glu Val Thr His Gln Gly Leu Ser Ser Pro Val Thr Lys Ser
195 200 205

Phe Asn Arg Gly Glu Cys
210

<210> 67
<211> 723
<212> DNA
<213> Seqüência Artificial
<220>
<223> Descrição de Seqüência Artificial: Construto sintético
<400> 67
gacatccaga tgacccagtc tccatcttcc ctgtctgcat ctgtagggga cagagtcacc 60
atcacttgca gggcaagtca ggacattagg tattatttaa attggtatca gcagaaacca 120
ggaaaagctc ctaagctcct gatctatgtt gcatccagtt tgcaaagtgg ggtcccatca 180
aggttcagcg gcagtggatc tgggacagag ttcactctca ccgtcagcag cctgcagcct 240
gaagattttg cgacttatta ctgtctacag gtttatagta cccctcggac gttcggccaa 300
gggaccaagg tggaaatcaa aggcggtggc gggtccggtg ggggtggctc cgggggcggt 360
ggctccgagg tgcagctggt ggagtctggg ggcggcttgg caaagcctgg ggggtccctg 420
agactctcct gcgcagcctc cgggttcagg ttcaccttca ataactacta catggactgg 480
gtccgccagg ctccagggca ggggctggag tgggtctcac gtattagtag tagtggtgat 540
cccacatggt acgcagactc cgtgaagggc agattcacca tctccagaga gaacgccaag 600
aacacactgt ttcttcaaat gaacagcctg agagctgagg acacggctgt ctattactgt 660
gcgagcttga ctacagggtc tgactcctgg ggccagggag tcctggtcac cgtctcctca 720

tga 723

<210> 68

<211> 240

<212> PRT

<213> Seqüência Artificial

<220>

<223> Descrição de Seqüência Artificial: Construto sintético

<400> 68

Asp Ile Gln Met Thr Gln Ser Pro Ser Ser Leu Ser Ala Ser Val Gly
1 5 10 15

Asp Arg Val Thr Ile Thr Cys Arg Ala Ser Gln Asp Ile Arg Tyr Tyr
20 25 30

Leu Asn Trp Tyr Gln Gln Lys Pro Gly Lys Ala Pro Lys Leu Leu Ile
35 40 45

Tyr Val Ala Ser Ser Leu Gln Ser Gly Val Pro Ser Arg Phe Ser Gly
50 55 60

Ser Gly Ser Gly Thr Glu Phe Thr Leu Thr Val Ser Ser Leu Gln Pro
65 70 75 80

Glu Asp Phe Ala Thr Tyr Tyr Cys Leu Gln Val Tyr Ser Thr Pro Arg
85 90 95

Thr Phe Gly Gln Gly Thr Lys Val Glu Ile Lys Gly Gly Gly Gly Ser
100 105 110

Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Glu Val Gln Leu Val Glu
115 120 125

Ser Gly Gly Gly Leu Ala Lys Pro Gly Gly Ser Leu Arg Leu Ser Cys
130 135 140

Ala Ala Ser Gly Phe Arg Phe Thr Phe Asn Asn Tyr Tyr Met Asp Trp
145 150 155 160

Val Arg Gln Ala Pro Gly Gln Gly Leu Glu Trp Val Ser Arg Ile Ser
165 170 175

Ser Ser Gly Asp Pro Thr Trp Tyr Ala Asp Ser Val Lys Gly Arg Phe
180 185 190

Thr Ile Ser Arg Glu Asn Ala Lys Asn Thr Leu Phe Leu Gln Met Asn
195 200 205

Ser Leu Arg Ala Glu Asp Thr Ala Val Tyr Tyr Cys Ala Ser Leu Thr
210 215 220

Thr Gly Ser Asp Ser Trp Gly Gln Gly Val Leu Val Thr Val Ser Ser
225 230 235 240

<210> 69

<211> 723

<212> DNA

<213> Seqüência Artificial

<220>

<223> Descrição de Seqüência Artificial: Construto sintético

<400> 69

gaggtgcagc tggtggagtc tgggggcggc ttggcaaagc ctgggggtc cctgagactc 60

tcctgcgcag cctccgggtt caggttcacc ttcaataact actacatgga ctgggtccgc 120

caggctccag ggcaggggct ggagtgggtc tcacgtatta gtagtagtgg tgatcccaca 180

tggtacgcag actccgtgaa gggcagattc accatctcca gagagaacgc caagaacaca 240

ctgtttcttc aaatgaacag cctgagagct gaggacacgg ctgtctatta ctgtgcgagc 300

ttgactacag ggtctgactc ctggggccag ggagtcctgg tcaccgtctc ctcaggcggt 360

ggcgggtccg gtgggggtgg ctccgggggc ggtggctccg acatccagat gacccagtct 420

ccatcttccc tgtctgcatc tgtaggggac agagtcacca tcacttgcag ggcaagtcag 480

gacattaggt attatttaaa ttggtatcag cagaaaccag gaaaagctcc taagctcctg 540

atctatgttg catccagttt gcaaagtggg gtcccatcaa ggttcagcgg cagtggatct 600

gggacagagt tcactctcac cgtcagcagc ctgcagcctg aagattttgc gacttattac 660

tgtctacagg tttatagtac ccctcggacg ttcggccaag ggaccaaggt ggaaatcaaa 720

tga 723

<210> 70

<211> 240

<212> PRT

<213> Seqüência Artificial

<220>

<223> Descrição de Seqüência Artificial: Construto sintético

<400> 70

Glu Val Gln Leu Val Glu Ser Gly Gly Gly Leu Ala Lys Pro Gly Gly

1 5 10 15

Ser Leu Arg Leu Ser Cys Ala Ala Ser Gly Phe Arg Phe Thr Phe Asn

20 25 30

Asn Tyr Tyr Met Asp Trp Val Arg Gln Ala Pro Gly Gln Gly Leu Glu

35 40 45

Trp Val Ser Arg Ile Ser Ser Ser Gly Asp Pro Thr Trp Tyr Ala Asp

50 55 60

Ser Val Lys Gly Arg Phe Thr Ile Ser Arg Glu Asn Ala Lys Asn Thr

65 70 75 80

Leu Phe Leu Gln Met Asn Ser Leu Arg Ala Glu Asp Thr Ala Val Tyr
85 90 95

Tyr Cys Ala Ser Leu Thr Thr Gly Ser Asp Ser Trp Gly Gln Gly Val
100 105 110

Leu Val Thr Val Ser Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser
115 120 125

Gly Gly Gly Gly Ser Asp Ile Gln Met Thr Gln Ser Pro Ser Ser Leu
130 135 140

Ser Ala Ser Val Gly Asp Arg Val Thr Ile Thr Cys Arg Ala Ser Gln
145 150 155 160

Asp Ile Arg Tyr Tyr Leu Asn Trp Tyr Gln Gln Lys Pro Gly Lys Ala
165 170 175

Pro Lys Leu Leu Ile Tyr Val Ala Ser Ser Leu Gln Ser Gly Val Pro
180 185 190

Ser Arg Phe Ser Gly Ser Gly Ser Gly Thr Glu Phe Thr Leu Thr Val
195 200 205

Ser Ser Leu Gln Pro Glu Asp Phe Ala Thr Tyr Tyr Cys Leu Gln Val
210 215 220

Tyr Ser Thr Pro Arg Thr Phe Gly Gln Gly Thr Lys Val Glu Ile Lys
225 230 235 240

<210> 71

<211> 44

<212> DNA

<213> Seqüência Artificial

<220>

<223> Descrição de Seqüência Artificial: Iniciador sintético

<400> 71

cgctggtggt gccgttctat agccatagtg acatccagat gacc 44

<210> 72

<211> 21

<212> DNA

<213> Seqüência Artificial

<220>

<223> Descrição de Seqüência Artificial: Iniciador sintético

<400> 72

gtggtcgact ttgatttcca c 21

<210> 73

<211> 47

<212> DNA

<213> Seqüência Artificial

<220>

<223> Descrição de Seqüência Artificial: Iniciador sintético

<400> 73

cgctggtggt gccgttctat agccatagtg aggtgcagct ggtggag 47

<210> 74

<211> 24

<212> DNA

<213> Seqüência Artificial

<220>

<223> Descrição de Seqüência Artificial: Iniciador sintético

<400> 74

gtggtcgact gaggagacgg tgac 24

<210> 75

<211> 52

<212> DNA

<213> Seqüência Artificial

<220>

<223> Descrição de Seqüência Artificial: Iniciador sintético

<400> 75

ggcatatgaa aaaactgctg ttcgcgattc cgctggtggt gccgttctat ag 52

<210> 76

<211> 33

<212> DNA

<213> Seqüência Artificial

<220>

<223> Descrição de Seqüência Artificial: Oligonucleotídeo sintético

<400> 76

gaggtgcagc tggtgcagtc tgggggcggc ttg 33

<210> 77

<211> 33

<212> DNA

<213> Seqüência Artificial

<220>

<223> Descrição de Seqüência Artificial: Oligonucleotídeo sintético

<400> 77

gagtctgggg gcggcrrcgc aaagcctggg ggg 33

<210> 78

<211> 37

<212> DNA

<213> Seqüência Artificial

<220>

<223> Descrição de Seqüência Artificial: Oligonucleotídeo sintético

<400> 78

gtctgggggc ggcttgrhgm agcctggggg gtccctg 37

<210> 79

<211> 34

<212> DNA

<213> Seqüência Artificial

<220>

<223> Descrição de Seqüência Artificial: Oligonucleotídeo sintético

<400> 79

gttcaggttc accttcagca actactacat ggac 34

<210> 80

<211> 35

<212> DNA

<213> Seqüência Artificial

<220>

<223> Descrição de Seqüência Artificial: Oligonucleotídeo sintético

<400> 80

caataactac tacatgmrct gggtccgcca ggctc 35

<210> 81

<211> 33

<212> DNA

<213> Seqüência Artificial

<220>

<223> Descrição de Seqüência Artificial: Oligonucleotídeo sintético

<400> 81

cgccaggctc cagggarggg gctggagtgg gtc 33

<210> 82

<211> 34

<212> DNA

<213> Seqüência Artificial

<220>

<223> Descrição de Seqüência Artificial: Oligonucleotídeo sintético

<400> 82

gggctggagt gggtcgsccg tattagtagt agtg 34

<210> 83

<211> 34

<212> DNA

<213> Seqüência Artificial

<220>

<223> Descrição de Seqüência Artificial: Oligonucleotídeo sintético

<400> 83

gagtgggtct cacgtatgag tagtagtggt gatc 34

<210> 84

<211> 39

<212> DNA

<213> Seqüência Artificial

<220>

<223> Descrição de Seqüência Artificial: Oligonucleotídeo sintético

<400> 84

ggtattattt aaattggtat gcccagaaac caggaaaag 39

<210> 85

<211> 36

<212> DNA

<213> Seqüência Artificial

<220>

<223> Descrição de Seqüência Artificial: Oligonucleotídeo sintético

<400> 85

gtattagtag tagtggtrgc cccacatggt acgcag 36

<210> 86

<211> 32

<212> DNA

<213> Seqüência Artificial

<220>

<223> Descrição de Seqüência Artificial: Oligonucleotídeo sintético

<400> 86

gtagtagtgg tgatrvcaca tggtacgcag ac 32

<210> 87

<211> 32

<212> DNA

<213> Seqüência Artificial

<220>

<223> Descrição de Seqüência Artificial: Oligonucleotídeo sintético

<400> 87

gtggtgatcc cacawactac gcagactccg tg 32

<210> 88

<211> 36

<212> DNA

<213> Seqüência Artificial

<220>

<223> Descrição de Seqüência Artificial: Oligonucleotídeo sintético

<400> 88

gattcaccat ctccagarac aacgccaaga acacac 36

<210> 89

<211> 35

<212> DNA

<213> Seqüência Artificial

<220>

<223> Descrição de Seqüência Artificial: Oligonucleotídeo sintético

<400> 89

catctccaga gagaacagca agaacacact gtttc 35

<210> 90

<211> 33

<212> DNA

<213> Seqüência Artificial

<220>

<223> Descrição de Seqüência Artificial: Oligonucleotídeo sintético

<400> 90

gccaagaaca cactgkhcct tcaaatgaac agc 33

<210> 91

<211> 35

<212> DNA

<213> Seqüência Artificial

<220>

<223> Descrição de Seqüência Artificial: Oligonucleotídeo sintético

<400> 91

gaacacactg tttcttgaaa tgaacagcct gagag 35

<210> 92

<211> 36

<212> DNA

<213> Seqüência Artificial

<220>

<223> Descrição de Seqüência Artificial: Oligonucleotídeo sintético

<400> 92

caaatgaaca gcctgagagg cgaggacacg gctgtc 36

<210> 93

<211> 32

<212> DNA

<213> Seqüência Artificial

<220>

<223> Descrição de Seqüência Artificial: Oligonucleotídeo sintético

<400> 93

gagctgagga cacggctrsc tattactgtg cg 32

<210> 94

<211> 34

<212> DNA

<213> Seqüência Artificial

<220>

<223> Descrição de Seqüência Artificial: Oligonucleotídeo sintético

<400> 94

gtctattact gtgcgargtt gactacaggg tctg 34

<210> 95

<211> 33

<212> DNA

<213> Seqüência Artificial

<220>

<223> Descrição de Seqüência Artificial: Oligonucleotídeo sintético

<400> 95

cagtctccat cttccggctc tgcatctgta ggg 33

<210> 96

<211> 32

<212> DNA

<213> Seqüência Artificial

<220>

<223> Descrição de Seqüência Artificial: Oligonucleotídeo sintético

<400> 96

ccctgtctgc atctrscggg gacagagtca cc 32

<210> 97

<211> 32

<212> DNA

<213> Seqüência Artificial

<220>

<223> Descrição de Seqüência Artificial: Oligonucleotídeo sintético

<400> 97

ctgtagggga cagactgacc atcacttgca gg 32

<210> 98

<211> 32

<212> DNA

<213> Seqüência Artificial

<220>

<223> Descrição de Seqüência Artificial: Oligonucleotídeo sintético

<400> 98

ggggacagag tcaccctgac ttgcagggca ag 32

<210> 99

<211> 33

<212> DNA

<213> Seqüência Artificial

<220>

<223> Descrição de Seqüência Artificial: Oligonucleotídeo sintético

<400> 99

gacagagtca ccatcagctg cagggcaagt cag 33

<210> 100

<211> 37

<212> DNA

<213> Seqüência Artificial

<220>

<223> Descrição de Seqüência Artificial: Oligonucleotídeo sintético

<400> 100

gcagggcaag tcagrgcatt aggtattatt taaattg 37

<210> 101

<211> 33

<212> DNA

<213> Seqüência Artificial

<220>

<223> Descrição de Seqüência Artificial: Oligonucleotídeo sintético

<400> 101

gcaagtcagg acattmkcta ttatttaaat tgg 33

<210> 102

<211> 32

<212> DNA

<213> Seqüência Artificial

<220>

<223> Descrição de Seqüência Artificial: Oligonucleotídeo sintético

<400> 102

aattggtatc agcaggcccc aggaaaagct cc 32

<210> 103

<211> 34

<212> DNA

<213> Seqüência Artificial

<220>

<223> Descrição de Seqüência Artificial: Oligonucleotídeo sintético

<400> 103

ccaggaaaag ctcctsrcct cctgatctat gttg 34

<210> 104

<211> 37

<212> DNA

<213> Seqüência Artificial

<220>

<223> Descrição de Seqüência Artificial: Oligonucleotídeo sintético

<220>

<221> modified_base

<222> (18)..(19)

<223> a, c, g or t

<400> 104

ctaagctcct gatctatnnk gcatccagtt tgcaaag 37

<210> 105

<211> 33

<212> DNA

<213> Seqüência Artificial

<220>

<223> Descrição de Seqüência Artificial: Oligonucleotídeo sintético

<400> 105

ctatgttgca tccagtcgcc aaagtggggt ccc 33

<210> 106

<211> 34

<212> DNA

<213> Seqüência Artificial

<220>

<223> Descrição de Seqüência Artificial: Oligonucleotídeo sintético

<400> 106

cagtttgcaa agtgggtccc catcaaggtt cagc 34

<210> 107

<211> 34

<212> DNA

<213> Seqüência Artificial

<220>

<223> Descrição de Seqüência Artificial: Oligonucleotídeo sintético

<400> 107

cagtggatct gggacagact tcactctcac cgtc 34

<210> 108

<211> 32

<212> DNA

<213> Seqüência Artificial

<220>

<223> Descrição de Seqüência Artificial: Oligonucleotídeo sintético

<400> 108

gagttcactc tcaccatcag cagcctgcag cc 32

<210> 109

<211> 30

<212> DNA

<213> Seqüência Artificial

<220>

<223> Descrição de Seqüência Artificial: Oligonucleotídeo sintético

<400> 109

cagcagcctg cagrgcgaag attttgcgac 30

<210> 110

<211> 32

<212> DNA

<213> Seqüência Artificial

<220>

<223> Descrição de Seqüência Artificial: Oligonucleotídeo sintético

<400> 110

ctgcagcctg aagatrscgc gacttattac tg 32

<210> 111

<211> 33

<212> DNA

<213> Seqüência Artificial

<220>

<223> Descrição de Seqüência Artificial: Oligonucleotídeo sintético

<400> 111

cctgaagatt ttgcggacta ttactgtcta cag 33

<210> 112

<211> 33

<212> DNA

<213> Seqüência Artificial

<220>

<223> Descrição de Seqüência Artificial: Oligonucleotídeo sintético

<400> 112

gcgacttatt actgtsmrca ggtttatagt acc 33

<210> 113

<211> 33

<212> DNA

<213> Seqüência Artificial

<220>

<223> Descrição de Seqüência Artificial: Oligonucleotídeo sintético

<400> 113

gtttatagta cccctmwaac gttcggccaa ggg 33

<210> 114

<211> 32

<212> DNA

<213> Seqüência Artificial

<220>

<223> Descrição de Seqüência Artificial: Oligonucleotídeo sintético

<400> 114

gtacccctcg gacgtggggc caagggacca ag 32

<210> 115

<211> 32

<212> DNA

<213> Seqüência Artificial

<220>

<223> Descrição de Seqüência Artificial: Oligonucleotídeo sintético

<400> 115

gctccagggc aggggtgcga gtgggtctca cg 32

<210> 116

<211> 32

<212> DNA

<213> Seqüência Artificial

<220>

<223> Descrição de Seqüência Artificial: Oligonucleotídeo sintético

<400> 116

gtacccctcg gacgtgcggc caagggacca ag 32

<210> 117

<211> 35

<212> DNA

<213> Seqüência Artificial

<220>

<223> Descrição de Seqüência Artificial: Oligonucleotídeo sintético

<400> 117

cttgactaca gggtcttgct cctggggcca gggag 35

<210> 118

<211> 32

<212> DNA

<213> Seqüência Artificial

<220>

<223> Descrição de Seqüência Artificial: Oligonucleotídeo sintético

<400> 118

ggaaaagctc ctaagtgcct gatctatgtt gc 32

<210> 119

<211> 34

<212> DNA

<213> Seqüência Artificial

<220>

<223> Descrição de Seqüência Artificial: Oligonucleotídeo sintético

<400> 119

gactacaggg tctgactgct ggggccaggg agtc 34

<210> 120

<211> 32

<212> DNA

<213> Seqüência Artificial

<220>

<223> Descrição de Seqüência Artificial: Oligonucleotídeo sintético

<400> 120

ggaaaagctc ctaagtgcct gatctatgtt gc 32

<210> 121

<211> 34

<212> DNA

<213> Seqüência Artificial

<220>

<223> Descrição de Seqüência Artificial: Oligonucleotídeo sintético

<400> 121

caggctccag ggcagtgcct ggagtgggtc tcac 34

<210> 122

<211> 36

<212> DNA

<213> Seqüência Artificial

<220>

<223> Descrição de Seqüência Artificial: Oligonucleotídeo sintético

<400> 122

cctcggacgt tcggctgcgg gaccaaggtg gaaatc 36

<210> 123

<211> 36

<212> DNA

<213> Seqüência Artificial

<220>

<223> Descrição de Seqüência Artificial: Oligonucleotídeo sintético

<220>

<221> modified_base

<222> (17)..(18)

<223> a, c, g or t

<400> 123

gtaccggtag gaggcmnnrk aaatcagtga tttagg 36

<210> 124

<211> 33

<212> DNA

<213> Seqüência Artificial

<220>

<223> Descrição de Seqüência Artificial: Oligonucleotídeo sintético

<400> 124

gggaaggctc ctaaattact gatttcctcg gcc 33

<210> 125

<211> 33

<212> DNA

<213> Seqüência Artificial

<220>

<223> Descrição de Seqüência Artificial: Oligonucleotídeo sintético

<220>

<221> modified_base

<222> (18)

<223> a, c, g or t

<400> 125

ggacaccttc actgacbncc caggcttctt cac 33

<210> 126

<211> 37

<212> DNA

<213> Seqüência Artificial

<220>

<223> Descrição de Seqüência Artificial: Oligonucleotídeo sintético

<400> 126

gatcctggga aggttttgac tactggggcc aagggac 37

<210> 127

<211> 34

<212> DNA

<213> Seqüência Artificial

<220>

<223> Descrição de Seqüência Artificial: Oligonucleotídeo sintético

<400> 127

gggtcctcag tgaagwtrtc ctgcaaggct tctg 34

<210> 128

<211> 34

<212> DNA

<213> Seqüência Artificial

<220>

<223> Descrição de Seqüência Artificial: Oligonucleotídeo sintético

<400> 128

cagggacttg agtggvkkgg atggatttat cctg 34

<210> 129

<211> 34

<212> DNA

<213> Seqüência Artificial

<220>

<223> Descrição de Seqüência Artificial: Oligonucleotídeo sintético

<220>

<221> modified_base

<222> (17)

<223> a, c, g or t

<400> 129

gaagttcaag ggcaggnyca caatcactgc agac 34

<210> 130

<211> 42

<212> DNA

<213> Seqüência Artificial

<220>

<223> Descrição de Seqüência Artificial: Oligonucleotídeo sintético

<400> 130

ggatggattt atcctggaaa tggtcatgct cagtacaatg ag 42

<210> 131

<211> 36

<212> DNA

<213> Seqüência Artificial

<220>

<223> Descrição de Seqüência Artificial: Oligonucleotídeo sintético

<220>

<221> modified_base

<222> (17)

<223> a, c, g or t

<400> 131

ggtggtagtg acattvnsat gacccagtct cctagc 36

<210> 132

<211> 75

<212> DNA

<213> Seqüência Artificial

<220>

<223> Descrição de Seqüência Artificial: Oligonucleotídeo sintético

<400> 132

ggcggtggag ggtccggtgg agggggctct ggagggggcg gttcaggggg cggtggatcg 60

ggggggaggtg gctcc 75

<210> 133

<211> 25

<212> PRT

<213> Seqüência Artificial

<220>

<223> Descrição de Seqüência Artificial: Peptídeo sintético

<400> 133

Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly

1 5 10 15

Gly Gly Gly Ser Gly Gly Gly Gly Ser

20 25

<210> 134

<211> 39

<212> DNA

<213> Seqüência Artificial

<220>

<223> Descrição de Seqüência Artificial: Oligonucleotídeo sintético

<400> 134

agagagacgc gtgtcctgtc ccaggtccaa ctggtgcag 39

<210> 135

<211> 89

<212> DNA

<213> Seqüência Artificial

<220>

<223> Descrição de Seqüência Artificial: Oligonucleotídeo sintético

<400> 135

agccacctcc ccccgatcca ccgccccctg aaccgccccc tccagagccc cctccaccgg 60

accctccacc gcctttgatc tccaccttg 89

<210> 136

<211> 90

<212> DNA

<213> Seqüência Artificial

<220>

<223> Descrição de Seqüência Artificial: Oligonucleotídeo sintético

<400> 136

agagagagat cttgactgtc tctccaggct tcttcagctc aggtccagac tgcaccaact 60

ggatctggga gccacctccc cccgatccac 90

<210> 137

<211> 290

<212> PRT

<213> Seqüência Artificial

<220>

<223> Descrição de Seqüência Artificial: Construto sintético

<400> 137

Met Lys Lys Leu Leu Phe Ala Ile Pro Leu Val Val Pro Phe Tyr Ser

1 5 10 15

His Ser Gln Val Gln Leu Val Gln Ser Gly Ala Glu Val Lys Lys Pro

20 25 30

Gly Ser Ser Val Lys Val Ser Cys Lys Ala Ser Gly Tyr Thr Phe Thr

35 40 45

Thr Tyr Tyr Leu His Trp Val Arg Gln Ala Pro Gly Gln Gly Leu Glu

50 55 60

Trp Met Gly Trp Ile Tyr Pro Gly Asn Val His Ala Gln Tyr Asn Glu

65 70 75 80

Lys Phe Lys Gly Arg Val Thr Ile Thr Ala Asp Lys Ser Thr Ser Thr
85 90 95

Ala Tyr Met Glu Leu Ser Ser Leu Arg Ser Glu Asp Thr Ala Val Tyr
100 105 110

Tyr Cys Ala Arg Ser Trp Glu Gly Phe Pro Tyr Trp Gly Gln Gly Thr
115 120 125

Thr Val Thr Val Ser Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser
130 135 140

Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Asp Ile Gln Met Thr Gln
145 150 155 160

Ser Pro Ser Ser Leu Ser Ala Ser Val Gly Asp Arg Val Thr Ile Thr
165 170 175

Cys Lys Ala Ser Gln Asn Val Gly Ile Asn Val Ala Trp Tyr Gln Gln
180 185 190

Lys Pro Gly Lys Ala Pro Lys Ser Leu Ile Ser Ser Ala Ser Tyr Arg
195 200 205

Tyr Ser Gly Val Pro Ser Arg Phe Ser Gly Ser Gly Ser Gly Thr Asp
210 215 220

Phe Thr Leu Thr Ile Ser Ser Leu Gln Pro Glu Asp Phe Ala Thr Tyr
225 230 235 240

Phe Cys Gln Gln Tyr Asp Thr Tyr Pro Phe Thr Phe Gly Gln Gly Thr

245 250 255

Lys Val Glu Ile Lys Asp Asp Asp Asp Lys Ser Phe Leu Glu Gln Lys

260 265 270

Leu Ile Ser Glu Glu Asp Leu Asn Ser Ala Val Asp His His His His

275 280 285

His His

290

<210> 138

<211> 290

<212> PRT

<213> Seqüência Artificial

<220>

<223> Descrição de Seqüência Artificial: Construto sintético

<400> 138

Met Lys Lys Leu Leu Phe Ala Ile Pro Leu Val Val Pro Phe Tyr Ser

1 5 10 15

His Ser Gln Val Gln Leu Val Gln Ser Gly Ala Glu Val Lys Lys Pro

20 25 30

Gly Ser Ser Val Lys Val Ser Cys Lys Ala Ser Gly Tyr Thr Phe Thr

35 40 45

Thr Tyr Tyr Leu His Trp Val Arg Gln Ala Pro Gly Gln Gly Leu Glu

50 55 60

Trp Met Gly Trp Ile Tyr Pro Gly Asn Gly His Ala Gln Tyr Asn Glu

65 70 75 80

Lys Phe Lys Gly Arg Val Thr Ile Thr Ala Asp Lys Ser Thr Ser Thr
85 90 95

Ala Tyr Met Glu Leu Ser Ser Leu Arg Ser Glu Asp Thr Ala Val Tyr
100 105 110

Tyr Cys Ala Arg Ser Trp Glu Gly Phe Pro Tyr Trp Gly Gln Gly Thr
115 120 125

Thr Val Thr Val Ser Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser
130 135 140

Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Asp Ile Gln Met Thr Gln
145 150 155 160

Ser Pro Ser Ser Leu Ser Ala Ser Val Gly Asp Arg Val Thr Ile Thr
165 170 175

Cys Lys Ala Ser Gln Asn Val Gly Ile Asn Val Ala Trp Tyr Gln Gln
180 185 190

Lys Pro Gly Lys Ala Pro Lys Ser Leu Ile Ser Ser Ala Ser Tyr Arg
195 200 205

Tyr Ser Gly Val Pro Ser Arg Phe Ser Gly Ser Gly Ser Gly Thr Asp
210 215 220

Phe Thr Leu Thr Ile Ser Ser Leu Gln Pro Glu Asp Phe Ala Thr Tyr
225 230 235 240

Phe Cys Gln Gln Tyr Asp Thr Tyr Pro Phe Thr Phe Gly Gln Gly Thr

245 250 255

Lys Val Glu Ile Lys Asp Asp Asp Asp Lys Ser Phe Leu Glu Gln Lys

260 265 270

Leu Ile Ser Glu Glu Asp Leu Asn Ser Ala Val Asp His His His His

275 280 285

His His

290

<210> 139

<211> 33

<212> PRT

<213> Seqüência Artificial

<220>

<223> Descrição de Seqüência Artificial: Peptídeo sintético

<400> 139

Val Lys Leu Thr Val Ser Gln Gly Gln Pro Val Lys Leu Asn Cys Ser

1 5 10 15

Val Glu Gly Glu Glu Pro Asp Ile Gln Trp Val Lys Asp Gly Ala Val

20 25 30

Val

<210> 140

<211> 44

<212> PRT

<213> Seqüência Artificial

<220>

<223> Descrição de Seqüência Artificial: Peptídeo sintético

<400> 140

Gln Val Leu Gln Ser Val Ala Gly Gln Thr Leu Thr Val Arg Cys Gln
1 5 10 15

Tyr Pro Pro Thr Gly Ser Leu Tyr Glu Lys Lys Gly Trp Cys Lys Glu
20 25 30

Ala Ser Ala Leu Val Cys Ile Arg Leu Val Thr Ser
35 40

<210> 141

<211> 42

<212> PRT

<213> Seqüência Artificial

<220>

<223> Descrição de Seqüência Artificial: Peptídeo sintético

<400> 141

Leu Lys Ile Gln Ala Tyr Phe Asn Glu Thr Ala Asp Leu Pro Cys Gln
1 5 10 15

Ala Asn Ser Gln Asn Gln Ser Leu Ser Glu Leu Val Val Phe Trp Gln
20 25 30

Asp Gln Glu Asn Leu Val Leu Asn Glu Val
35 40

<210> 142

<211> 51

<212> PRT

<213> Seqüência Artificial

<220>

<223> Descrição de Seqüência Artificial: Peptídeo sintético

<400> 142

Val Val Arg Val Pro Thr Ala Thr Leu Val Arg Val Val Gly Thr Glu
1 5 10 15

Leu Val Ile Pro Cys Asn Val Ser Asp Tyr Asp Gly Pro Ser Glu Gln
20 25 30

Asn Phe Asp Trp Ser Phe Ser Ser Leu Gly Ser Ser Phe Val Glu Leu
35 40 45

Ala Ser Thr
50

<210> 143

<211> 44

<212> PRT

<213> Seqüência Artificial

<220>

<223> Descrição de Seqüência Artificial: Peptídeo sintético

<400> 143

Ala Leu Thr Gln Gln Asp Gly Gly Met Ser Val Ala Glu Gly Gln Ala
1 5 10 15

Leu Leu Leu Pro Cys Gly Ala Thr Leu Gly Gly Pro Asn Leu Ser Val
20 25 30

Val Trp Val Ser Pro Arg Gln Glu Asp Leu Val Ala
35 40

<210> 144

<211> 45

<212> PRT

<213> Seqüência Artificial

<220>

<223> Descrição de Seqüência Artificial: Peptídeo sintético

<400> 144

Ser Leu Arg Val Ser Pro Lys Asn Pro Arg Val Pro Lys Ser His Thr
1 5 10 15

Leu Glu Leu His Cys Glu Ser Gln Cys Asp Ser His Leu Lys Tyr Ser
20 25 30

Leu Lys Leu Ser Trp Ser Lys Asp Gly Glu Ala Phe Glu
35 40 45

<210> 145

<211> 44

<212> PRT

<213> Seqüência Artificial

<220>

<223> Descrição de Seqüência Artificial: Peptídeo sintético

<400> 145

Val Val Pro Thr Val His Arg Asn Val Thr Val Ser Arg Gly Glu Pro
1 5 10 15

Val Thr Leu Asn Cys Ser Ile Asn Met Thr Asn Ile Thr Gln Ile Ser
20 25 30

Trp Arg Lys Asp Thr Leu Ile Phe Val Tyr Leu Ala
35 40

<210> 146

<211> 39

<212> PRT

<213> Seqüência Artificial

<220>

<223> Descrição de Seqüência Artificial: Peptídeo sintético

<400> 146

His Gln Gln Gln Ile Gly Val Glu Gly Lys Glu Val Ile Leu Asn Cys
1 5 10 15

Lys Thr His Asp Lys Asp Val Thr Trp Lys Tyr Lys Tyr Asp Thr Gly
20 25 30

Ser Ala Ile Ile Ile Ile Gln
35

<210> 147

<211> 51

<212> PRT

<213> Seqüência Artificial

<220>

<223> Descrição de Seqüência Artificial: Peptídeo sintético

<400> 147

Val Ala Ser Gln Gly Gly His Val Glu Ser Val Tyr Leu Tyr Gly Thr
1 5 10 15

Val Glu Leu Pro Cys Glu Phe Pro Phe Val Thr Gly Pro Gln Asp Leu
20 25 30

Val Met Thr Trp Leu Lys Val Val Gln Asp Ser Glu Asn Val Val Val
35 40 45

His Ser Tyr
50

<210> 148

<211> 51

<212> PRT

<213> Seqüência Artificial

<220>

<223> Descrição de Seqüência Artificial: Peptídeo sintético

<400> 148

Thr Ile Thr Ala Pro Lys Asp Leu Tyr Val Val Glu Tyr Gly Ser Asn
1 5 10 15

Val Thr Met Glu Cys Arg Phe Pro Val Glu Arg Glu Asp Leu Leu Ala
20 25 30

Leu Val Val Tyr Trp Glu Lys Glu Asp Glu Gln Val Ile Gln Phe Val
35 40 45

Ala Gly Glu
50

<210> 149

<211> 50

<212> PRT

<213> Seqüência Artificial

<220>

<223> Descrição de Seqüência Artificial: Peptídeo sintético

<400> 149

Lys Ile Glu Ala Pro Val Asn Val Ile Val Gln Arg Gly Leu Cys Val
1 5 10 15

Leu Ile Pro Cys Asn Phe Thr Val Gly Pro Lys Tyr Asn Leu Thr Lys
20 25 30

Asp Ala Ile Gly Ile Trp Tyr Lys Gly His Pro Asn Asp Pro Val Ala
35 40 45

Ala Ser
50

<210> 150
<211> 49
<212> PRT
<213> Seqüência Artificial
<220>
<223> Descrição de Seqüência Artificial: Peptídeo sintético
<400> 150
Gln Gly Ala Val Pro Glu Glu Leu His Lys His Pro Gly Gln Thr Leu
1 5 10 15

Leu Leu Gln Cys Gln Tyr Ser Pro Lys Arg Gly Pro Tyr Gln Pro Lys
20 25 30

Ser Trp Cys Gln Gln Thr Ser Pro Ser Arg Cys Thr Leu Leu Val Thr
35 40 45

Ser

<210> 151
<211> 46
<212> PRT
<213> Seqüência Artificial
<220>
<223> Descrição de Seqüência Artificial: Peptídeo sintético

<400> 151

Ser Leu Ser Asp Ser Val Ser Ser Ala Val Leu Gly Ser Ser Ala Lys
1 5 10 15

Leu Ile Cys Gln Phe Thr Pro Val Val Asp Ala Lys Asn Val Glu Ile
20 25 30

Arg Trp Phe Thr Arg Ser Phe Arg Pro Tyr Val His Gln Tyr
35 40 45

<210> 152

<211> 49

<212> PRT

<213> Seqüência Artificial

<220>

<223> Descrição de Seqüência Artificial: Peptídeo sintético

<400> 152

Val Gly Thr Gln Arg Pro Ile Glu Leu His Leu Gly Glu Leu Arg Gly
1 5 10 15

Ser Ile Thr Ile Pro Cys Pro Pro Gly Asp Ser Trp Gly Gly Thr Arg
20 25 30

Arg Leu Trp Cys Arg Val Gly Arg Ser Arg Cys Ile Leu Ile Ala Asp
35 40 45

Thr

<210> 153

<211> 39

<212> PRT

<213> Seqüência Artificial

<220>

<223> Descrição de Seqüência Artificial: Peptídeo sintético

<400> 153

Gln Leu Val Pro Val Pro Asn Val Thr Val Ser Pro Gly Glu Thr Ala
1 5 10 15

Ile Leu Ser Cys Leu Val Leu Ser Glu Ala Pro Tyr Asn Leu Thr Trp
20 25 30

Val Arg Asp Trp Arg Val Leu
35

<210> 154

<211> 47

<212> PRT

<213> Seqüência Artificial

<220>

<223> Descrição de Seqüência Artificial: Peptídeo sintético

<400> 154

Met Val Glu Gln Met Asp Asn Lys Thr Val Val Ser Gly Lys Pro Val
1 5 10 15

Thr Phe Leu Cys Asn Tyr Ile Leu Ser Met Arg Val Arg Gln Val Leu
20 25 30

Trp Lys Lys Thr Ala Glu Gln Gly Asp Thr Ala Ile Val Ala Ser
35 40 45

<210> 155

<211> 42

<212> PRT

<213> Seqüência Artificial

<220>

<223> Descrição de Seqüência Artificial: Peptídeo sintético

<400> 155

Ser Gln Met Val Ser Glu Pro Gly Lys Asn Ile Thr Leu Thr Cys Glu
1 5 10 15

Pro Arg Glu Asn His Leu Leu Glu Glu Val Ser Trp Glu Lys Ile Gln
20 25 30

Pro Gln Gln Ile Asp Leu Leu Ile Ser Cys
35 40

<210> 156

<211> 30

<212> PRT

<213> Seqüência Artificial

<220>

<223> Descrição de Seqüência Artificial: Peptídeo sintético

<220>

<221> MOD_RES

<222> (6)..(10)

<223> Região pode estar ou não presente

<220>

<221> MOD_RES

<222> (11)..(15)

<223> Região pode estar ou não presente

<220>

<221> MOD_RES

<222> (16)..(20)

<223> Região pode estar ou não presente

<220>

<221> MOD_RES

<222> (21)..(25)

<223> Região pode estar ou não presente

<220>

<221> MOD_RES

<222> (26)..(30)

<223> Região pode estar ou não presente

<400> 156

Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly

1 5 10 15

Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser

20 25 30

<210> 157

<211> 31

<212> PRT

<213> Seqüência Artificial

<220>

<223> Descrição de Seqüência Artificial: Peptídeo sintético

<220>

<221> MOD_RES

<222> (7)..(11)

<223> Região pode estar ou não presente

<220>

<221> MOD_RES

<222> (12)..(16)

<223> Região pode estar ou não presente

<220>

<221> MOD_RES

<222> (17)..(21)

<223> Região pode estar ou não presente

<220>

<221> MOD_RES

<222> (22)..(26)

<223> Região pode estar ou não presente

<220>

<221> MOD_RES

<222> (27)..(31)

<223> Região pode estar ou não presente

<400> 157

Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser

1 5 10 15

Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser

20 25 30

<210> 158

<211> 4

<212> PRT

<213> Seqüência Artificial

<220>

<223> Descrição de Seqüência Artificial: Sintética

<220>

<221> MOD_RES

<222> (1)

<223> Suc-β-Ala

<400> 158

Ala Leu Ala Leu

1

<210> 159

<211> 5

<212> PRT

<213> Seqüência Artificial

<220>

<223> Descrição de Seqüência Artificial: Etiqueta 5xHis sintética

<400> 159

His His His His His

1 5

<210> 160

<211> 25

<212> PRT

<213> Seqüência Artificial

<220>

<223> Descrição de Seqüência Artificial: Peptídeo sintético

<220>

<221> MOD_RES

<222> (16)..(20)

<223> Região pode estar ou não presente

<220>

<221> MOD_RES

<222> (21)..(25)

<223> Região pode estar ou não presente

<400> 160

Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly

1 5 10 15

Gly Gly Gly Ser Gly Gly Gly Gly Ser

20 25

<210> 161

<211> 4

<212> PRT

<213> Seqüência Artificial

<220>

<223> Descrição de Seqüência Artificial: Peptídeo sintético

<400> 161

Ser Gly Gly Gly

1

REIVINDICAÇÕES

1. População de moléculas scFv estabilizadas, em que as moléculas scFv estabilizadas compreendem

i) um ligante de scFv interposto entre um domínio V_H e um domínio V_L, em que os domínios V_H e V_L são encadeados por uma ligação dissulfeto entre um aminoácido no domínio V_H e um aminoácido no domínio V_L, e em que os domínios V_H e V_L das moléculas scFv têm valores de Tm de mais de 55°C; ou

ii) um ligante de scFv interposto entre um domínio V_H e um domínio V_L, em que os domínios V_H e V_L são encadeados por uma ligação dissulfeto entre um aminoácido no domínio V_H e um aminoácido no domínio V_L, e em que moléculas scFv têm uma T_{50} de mais de 49°C.

2. Molécula scFv estabilizada em que a molécula scFv estabilizada compreende

i) um ligante de scFv tendo a seqüência de aminoácidos $(Gly_4\ Ser)_n$ interposto entre um domínio V_H e um domínio V_L, em que os domínios V_H e V_L são ligados por uma ligação dissulfeto entre o aminoácido 44 de V_H e o aminoácido 100 de V_L;

ii) um domínio V_H e um domínio V_L em que a molécula tem no mínimo uma substituição selecionada entre um grupo consistindo em:

a) substituição de um aminoácido na posição de Kabat 13 de VH;
b) substituição de um aminoácido na posição de Kabat 16 de VH;
c) substituição de um aminoácido na posição de Kabat 46 de VL;
d) substituição de um aminoácido na posição de Kabat 49 de VL;
e) substituição de um aminoácido na posição de Kabat 50 de VL;
f) substituição de aminoácidos nas posições de Kabat 49 e 50 de VL;
g) substituição de aminoácido na posição de Kabat 101 de VH;
h) substituição de aminoácido na posição de Kabat 20 de VH;
i) substituição de aminoácido na posição de Kabat 48 de VH;
j) substituição de aminoácido na posição de Kabat 3 de VL;
k) substituição de aminoácido na posição de Kabat 55;

l) substituição de aminoácido na posição de Kabat 67 de VH;
m) substituição de aminoácido na posição de Kabat 6 de VH;
n) substituição de aminoácido na posição de Kabat 32 de VH ;
o) substituição de aminoácido na posição de Kabat 49 de VH;
p) substituição de aminoácido na posição de Kabat 43 de VH;
q) substituição de aminoácido na posição de Kabat 72 de VH;
r) substituição de aminoácido na posição de Kabat 79 de VH;
s) substituição de aminoácido na posição de Kabat 50 de VL;
t) substituição de aminoácido na posição de Kabat 75 de VL;
u) substituição de aminoácido na posição de Kabat 80 de VH ;
v) substituição de aminoácido na posição de Kabat 83 de VH ; ou

iii) um domínio V_H e um domínio V_L com uma interface estabilizada entre VH e VL da scFv em que a molécula tem no mínimo uma substituição em uma posição de aminoácido diretamente dentro da interface e/ou em uma posição de aminoácido que *estruturas* a interação entre VH e VL, em que a no mínimo uma posição de aminoácido é selecionada entre um grupo consistindo em: 47H, 37H, 45H, 46H, 51H, 59H, 109H, 14H, 26H, 27H, 40H, 69H, 103H, 29H, 38H, 44H, 49H, 55H, 63H, 54H, 68H, 72H, 74H, 79H, 82bH, 8H, 32H, 39H, 41H, 62H, 85H, 91H, 24H, 60H, 37L, 36L, 44L, 59L, 57L, 64L, 46L, 48L, 63L, 67L, 68L, 39L, 45L, 47L, 54L, 56L, 79L, 85L, 98L, 58L, 74L, 77L, 83L, 89L, e 104L, de acordo com a numeração de Kabat, e em que a molécula scFv estabilizada tem uma T_{50} mais do que a de uma molécula scFv carecendo da substituição.

3. Molécula scFv estabilizada de acordo com a reivindicação 2, em que a molécula scFv estabilizada tem uma estabilidade equivalente à de um fragmento Fab convencional sob condições de ensaio térmico.

4. Proteína de fusão compreendendo uma molécula scFv estabilizada como definida na reivindicação 2.

5. População de moléculas de ligação de antígenos multivalentes estabilizadas, em que as moléculas de ligação estabilizadas compreendem no mínimo uma molécula scFv estabilizada compreendendo um ligante de scFv interposto entre um domínio V_H e um domínio V_L, em que os domí-

nios V_H e V_L são encadeados por uma ligação dissulfeto entre um aminoácido no domínio V_H e um aminoácido no domínio V_L, e em que a população de moléculas de ligação estabilizadas compreendem proteínas solúveis monoméricas das quais não mais de 10% está presente em forma aggregada.

6. População a reivindicação 5, em que as moléculas de ligação de antígenos multivalentes estabilizadas são expressadas em células CHO ou NS0.

7. População de moléculas de ligação de antígenos multivalentes estabilizadas, em que cada molécula de ligação de antígeno multivalente compreende

i) no mínimo uma molécula scFv compreendendo um ligante de scFv interposto entre um domínio V_H e um domínio V_L, em que os domínios V_H e V_L são encadeados por uma ligação dissulfeto entre um aminoácido no domínio V_H e um aminoácido no domínio V_L, e em que os domínios V_H e V_L da no mínimo uma molécula scFv têm uma Tm de mais de 55°C; ou

ii) no mínimo uma molécula scFv compreendendo um ligante de scFv interposto entre um domínio V_H e um domínio V_L, em que os domínios V_H e V_L são encadeados por uma ligação dissulfeto entre um aminoácido no domínio V_H e um aminoácido no domínio V_L, e em que a no mínimo uma molécula scFv da molécula de ligação de antígeno multivalente tem uma T_{50} de mais de 49°C.

8. Molécula de ligação de antígenos multivalentes estabilizada compreendendo

i) no mínimo uma molécula scFv estabilizada, em que a molécula scFv estabilizada compreende um ligante de scFv $(Gly_4\ Ser)_n$ interposto entre um domínio V_H e um domínio V_L e em que os domínios V_H e V_L são encadeados por uma ligação dissulfeto;

ii) no mínimo uma molécula scFv estabilizada compreendendo um ligante de scFv tendo a seqüência de aminoácidos $(Gly_4\ Ser)_n$ interposto entre um domínio V_H e um domínio V_L, em que os domínios V_H e V_L são encadeados por uma ligação dissulfeto entre o aminoácido 44 de V_H e o aminoácido 100 de V_L;

iii) no mínimo uma molécula scFv estabilizada, em que a molécula scFv estabilizada compreende no mínimo uma substituição selecionada entre um grupo consistindo em:

a) substituição de um aminoácido na posição de Kabat 13 de VH;
b) substituição de um aminoácido na posição de Kabat 16 de VH;
c) substituição de um aminoácido na posição de Kabat 46 de VL;
d) substituição de um aminoácido na posição de Kabat 49 de VL;
e) substituição de um aminoácido na posição de Kabat 50 de VL;
f) substituição de aminoácidos nas posições de Kabat 49 e 50 de VL;
g) substituição de aminoácido na posição de Kabat 101 de VH;
h) substituição de aminoácido na posição de Kabat 20 de VH;
i) substituição de aminoácido na posição de Kabat 48 de VH;
j) substituição de aminoácido na posição de Kabat 3 de VL;
k) substituição de aminoácido na posição de Kabat 55 de VH;
l) substituição de aminoácido na posição de Kabat 67 de VH;
m) substituição de aminoácido na posição de Kabat 6 de VH;
n) substituição de aminoácido na posição de Kabat 32 de VH;
o) substituição de aminoácido na posição de Kabat 49 de VH;
p) substituição de aminoácido na posição de Kabat 43 de VH;
q) substituição de aminoácido na posição de Kabat 72 de VH;
r) substituição de aminoácido na posição de Kabat 79 de VH;
s) substituição de aminoácido na posição de Kabat 50 de VL;
t) substituição de aminoácido na posição de Kabat 75 de VL;
u) substituição de aminoácido na posição de Kabat 80 de VH;
v) substituição de aminoácido na posição de Kabat 83 de VH;

ou

iv) um domínio V_H e um domínio V_L com uma interface estabilizada entre VH e VL da scFv em que a molécula tem no mínimo uma substituição em uma posição de aminoácido diretamente dentro da interface e/ou em uma posição de aminoácido que *estruturas* a interação entre VH e VL, em que a no mínimo uma posição de aminoácido é selecionada entre

um grupo consistindo em: 47H, 37H, 45H, 46H, 51H, 59H, 109H, 14H, 26H, 27H, 40H, 69H, 103H, 29H, 38H, 44H, 49H, 55H, 63H, 54H, 68H, 72H, 74H, 79H, 82bH, 8H, 32H, 39H, 41H, 62H, 85H, 91H, 24H, 60H, 37L, 36L, 44L, 59L, 57L, 64L, 46L, 48L, 63L, 67L, 68L, 39L, 45L, 47L, 54L, 56L, 79L, 85L, 98L, 58L, 74L, 77L, 83L, 89L, e 104L, de acordo com a numeração de Kabat,

9. Molécula de ligação de antígeno multivalente a reivindicação 8, em que a no mínimo uma molécula scFv estabilizada é geneticamente fundida a um anticorpo.

10. Molécula de ligação de antígeno multivalente a reivindicação 8, em que duas moléculas scFv estabilizadas são geneticamente fundidas a um anticorpo.

11. Molécula de ligação de antígeno multivalente a reivindicação 9, em que a no mínimo uma molécula scFv estabilizada é geneticamente fundida ao término carbóxi da cadeia leve ou pesada do anticorpo.

12. Molécula de ligação de antígeno multivalente a reivindicação 9, em que a no mínimo uma molécula scFv estabilizada é geneticamente fundida ao término amino de uma cadeia leve ou cadeia pesada do anticorpo.

13. Molécula de ligação de antígeno multivalente de acordo com qualquer uma das reivindicações 8 a 12, a qual compreende no mínimo um sítio de ligação que liga a uma molécula preferencialmente expressada sobre células de câncer.

14. Molécula de ligação de antígeno multivalente de acordo com qualquer uma das reivindicações 8 a 12, em que a molécula é selecionada entre o grupo consistindo em HER1, HER3, CD80, CD86, PD-1, CTLA4, B7-H4, RON, CD200, CD4, BAF R, EGFR, IGFR, VEGFR, um membro da família de receptores TNF, um Tie receptor, MET, IGF1, IGF2, TNF, um ligante de TNF, IL-6, TWEAK, Fn14, CD20, CD23, CRIPTO, HGF, $\alpha 4\beta 1$ integrina, $\alpha 5\beta 1$ integrina, $\alpha 6\beta 4$ integrina,, e $\alpha V\beta 6$ integrina.

15. Molécula de ligação de antígeno multivalente de acordo com qualquer uma das reivindicações 8 a 12, a qual compreende no mínimo um sítio de ligação que liga a uma molécula envolvida em modulação de rea-

ções imunes.

16. Molécula de ligação de antígeno multivalente de acordo com qualquer uma das reivindicações 8 a 12, a qual compreende no mínimo um sítio de ligação que liga a uma molécula envolvida em modulação de angiogênese.

17. Molécula de ligação de antígeno multivalente de acordo com qualquer uma das reivindicações 18 a 12, a qual compreende no mínimo um sítio de ligação que liga a um alvo neurológico.

18. Molécula de ligação de antígeno multivalente de acordo com qualquer uma das reivindicações 8 a 12, a qual é multiespecífica.

19. Molécula de ligação de antígeno multivalente de qualquer uma das reivindicações 18 a 12, a qual é biespecífica.

20. Molécula de ligação de antígeno multiespecífica da reivindicação 19, a qual liga a dois membros da família de receptores de TNF.

21. Molécula de ácido nucléico compreendendo uma seqüência de nucleotídeos a qual codifica a molécula scFv estabilizada como definida em qualquer uma das reivindicações 2 a 3 ou o complemento da mesma.

22. Molécula de ácido nucléico de acordo com a reivindicação 21, a qual está em um vetor.

23. Célula hospedeira compreendendo o vetor da reivindicação 22.

24. Célula hospedeira de acordo com a reivindicação 23, em que a célula é uma célula CHO ou uma célula NS0.

25. Célula hospedeira de acordo com a reivindicação 23, a qual expressa uma população de moléculas de ligação compreendendo moléculas scFv, a população tendo no mínimo 10% menos agregados relativos a uma célula hospedeira a qual expressa uma população de moléculas de ligação compreendendo moléculas scFv convencionais.

26. Método para produzir uma molécula de ligação estabilizada, compreendendo cultivar a célula hospedeira como definida na reivindicação 23 sob condições tais que a molécula de ligação estabilizada seja produzida.

27. Método de acordo com a reivindicação 26, em que no míni-

mo 10 mg da molécula de ligação estabilizada seja produzida para cada litro do meio de cultura de células hospedeiras e em que não mais de 10% da molécula de ligação está presente em forma de agregado.

28. Método para tratar um sujeito que se beneficiaria de tratamento com uma molécula de ligação como definida em qualquer uma das reivindicações 8 a 20 compreendendo administrar a molécula de ligação ao sujeito de tal modo que ocorre tratamento.

29. Método de acordo com a reivindicação 28, em que o sujeito está sofrendo de uma doença ou um distúrbio selecionados entre o grupo consistindo em câncer, uma doença ou distúrbio auto-imune, e uma doença ou distúrbio neurológico.

30. Método para estabilizar uma molécula scFv compreendendo fundir geneticamente um domínio V_H e um domínio V_L usando um $(Gly_4\ Ser)_n$ ligante de scFv e produzir o domínio V_H para compreender uma cisteína no aminoácido 44 e produzir o domínio V_L para compreender uma cisteína no aminoácido 100 para formar uma molécula scFv estabilizada.

31. Método para produzir um anticorpo multivalente estabilizado compreendendo uma molécula scFv estabilizada, o método compreendendo fundir geneticamente uma molécula scFv estabilizada ao término amino ou ao término carbóxi de uma cadeia leve ou pesada de um anticorpo molécula.

32. Método para produzir uma molécula scFv estabilizada compreendendo substituir no mínimo um aminoácido na interface VH/VL da molécula scFv e/ou no mínimo um aminoácido que *estruturas* uma ou mais substituições na interface VH/VL da molécula scFv simultaneamente melhoram a estabilidade térmica de ambos os domínios V_H e V_L da molécula scFv comparada com uma molécula scFv convencional.

33. Método para produzir uma molécula scFv estabilizada compreendendo substituir no mínimo um aminoácido no domínio VH ou no domínio VL que co-varia com dois ou mais aminoácidos na interface entre os domínios VH e VL.

34. Método para produzir uma proteína de fusão estabilizada compreendendo uma molécula scFv compreendendo substituir no mínimo

um aminoácido na interface VH/VL da molécula scFv e/ou no mínimo um aminoácido no *estrutura* VH/VL da molécula scFv e fundir geneticamente a molécula scFv a um polipeptídeo para deste modo produzir uma proteína de fusão estabilizada.

35. Método para melhorar a estabilidade de uma molécula multivalente compreendendo no mínimo uma molécula scFv, o método compreendendo introduzir no mínimo uma mutação estabilizante em no mínimo uma molécula scFv para deste modo melhorar a estabilidade da molécula multivalente.

36. Método para fabricação em larga escala de uma proteína de fusão estabilizada, o método compreendendo:

(a) conduzir um ensaio de estabilidade térmica de uma molécula scFv para determinar a estabilidade térmica de uma molécula scFv candidata;

(b) comparar a estabilidade térmica da molécula scFv candidata com um controle adequado para identificar moléculas scFv estabilizadas tendo aumentada estabilidade térmica em relação ao controle,

(c) selecionar as moléculas scFv estabilizadas identificado na etapa (b);

(d) fundir geneticamente no mínimo uma molécula scFv estabilizada a proteína para formar uma proteína de fusão estabilizada;

(e) transfectar uma célula hospedeira mamífera com uma molécula de ácido nucléico codificando a proteína de fusão estabilizada,

(f) cultivar a célula hospedeira da etapa (e) sob condições tais que a proteína de fusão estabilizada é expressada;

em que a proteína de fusão estabilizada é expressada com não mais de 10% de agregação de proteína quando cultivada em 10 L ou mais de meio de cultura.

37. Interface gráfica do usuário compreendendo um display gráfico de:

(i) um layout de grade compreendendo uma coleção de dados de freqüência de uso de resíduos para uma série de referência de

seqüências de polipeptídeos; e

(ii) um revestimento de co-variação.

38. Método para produzir um polipeptídeo da superfamília de imunoglobulinas (Ig) com uma propriedade biofísica aprimorada, o referido método compreendendo as etapas de:

a) proporcionar um alinhamento de uma série de referência curada de seqüências correspondentes a uma Ig *fold* do polipeptídeo;

b) calcular a covariação entre resíduos das seqüências do alinhamento para identificar resíduos covariáveis;

c) substituir um resíduo do polipeptídeo que falha em satisfazer uma covariação com o resíduo covariável encontrado em uma posição correspondente no alinhamento.

39. Método de acordo com a reivindicação 38, em que a propriedade biofísica é selecionada entre o grupo consistindo em estabilidade térmica, perfil de desdobramento com pH (N.T.: *pH unfolding*), remoção estável de glicosilação, solubilidade, função bioquímica, e combinações das mesmas.

40. Método de acordo com a reivindicação 39, em que a função bioquímica é selecionada entre o grupo consistindo em reconhecimento de um alvo protéico ou uma porção química, cuja porção química é selecionada entre o grupo consistindo em fosfato, metila, acetila, lipídeo, detergente, íon metálico, halogênio, RNA, DNA, oligonucleotídeo, e oligonucleosídeo.

41. Método de acordo com a reivindicação 38, em que o polipeptídeo da superfamília das Ig é derivado de uma proteína selecionada entre o grupo consistindo em um anticorpo ou fragmento de ligação de antígeno do mesmo, um receptor de imunoglobulina, uma proteína de adesão celular, uma integrina, um alergeno, um receptor de células T, e uma molécula de complexo de histocompatibilidade maior (MHC).

42. Método de acordo com a reivindicação 38, em que a Ig *fold* é selecionada entre o grupo consistindo em uma *fold* classe V, uma *fold* classe I, uma *fold* classe C1, e uma *fold* classe C2.

43. Método de acordo com a reivindicação 38, em que o polipep-

tídeo da superfamília das Ig é um anticorpo modificado selecionado entre o grupo consistindo em um anticorpo de domínio, um anticorpo humanizado, um anticorpo humano, um anticorpo monoclonal não-humano, um anticorpo quimérico, um anticorpo biespecífico, um anticorpo contendo scFv, e um anticorpo com domínio eliminado.

44. Método da reivindicação 38, em que os resíduos covariáveis são parte de uma característica estrutural selecionada entre o grupo consistindo em uma ligação dissulfeto, uma ponte de sal, uma porção de uma superfície ou bolsa de ligação de ligante, e uma rede de van Der Waals, ligação hidrogênio, e/ou interações carga-carga.

45. Método para produzir um anticorpo ou anticorpo modificado com uma aprimorada estabilidade, o referido método compreendendo as etapas de:

a) proporcionar um modelo estrutural de alta resolução de um anticorpo padrão que é experimentalmente validado como estável;

b) usar o modelo estrutural para identificar uma interface VH/VL e/ou resíduo de *estruturação* de interface no anticorpo template;

c) usar um modelo de homologia para identificar um resíduo da interface VH/VL correspondente do anticorpo ou anticorpo modificado que é importante para estabilizar a interface; e

d) substituir o resíduo da interface VH/VL correspondente no anticorpo ou anticorpo modificado com o resíduo da interface da proteína template.

46. Método da reivindicação 45, em que o resíduo da interface da proteína padrão encobre no mínimo 10 $Å^2$ de área superficial na interface.

47. Método da reivindicação 45, em que o anticorpo modificado é selecionado entre o grupo consistindo em um anticorpo de domínio único, um anticorpo humanizado, um anticorpo humano, um anticorpo monoclonal não-humano, um anticorpo quimérico, um anticorpo biespecífico, um anticorpo contendo scFv, e um anticorpo com domínio eliminado.

48. Polipeptídeo produzido pelo método como definido em qualquer uma das reivindicações 38 a 47.

49. Polipeptídeo de acordo com a reivindicação 48, em que o polipeptídeo apresenta um aprimoramento em uma propriedade biofísica selecionada entre o grupo consistindo em estabilidade térmica, perfil de desdobramento com pH, remoção estável de glicosilação, solubilidade, função bioquímica, e combinações das mesmas.

50. Polipeptídeo de acordo com a reivindicação 49, em que função bioquímica é afinidade ou especificidade de ligação de ligante.

51. Polipeptídeo de acordo com a reivindicação 48, o qual derivado de uma proteína selecionada entre o grupo consistindo em um anticorpo ou fragmento de ligação de antígeno do mesmo, um receptor de imunoglobulina, uma proteína de adesão celular, uma integrina, um alérgeno, um receptor de células T, e um complexo de histocompatibilidade maior (MHC) molécula.

52. Polipeptídeo de acordo com a reivindicação 48, o qual é selecionado entre o grupo consistindo em uma região variável de cadeia pesada (V_H), uma região variável de cadeia leve (V_L), e um anticorpo de cadeia única (sFv).

53. Polipeptídeo de acordo com a reivindicação 48, em que a Ig *fold* é selecionada entre o grupo consistindo em uma *fold* classe V, uma *fold* classe I, uma *fold* classe C1, e uma *fold* classe C2.

54. Polipeptídeo de acordo com a reivindicação 48, o qual é um anticorpo modificado selecionado entre o grupo consistindo em um anticorpo de domínio único, um anticorpo humanizado, um anticorpo humano, um anticorpo monoclonal não-humano, um anticorpo quimérico, um anticorpo biespecífico, um anticorpo contendo scFv, e um anticorpo com domínio eliminado.

55. Método para produzir uma biblioteca compreendendo moléculas scFv estabilizadas, o método compreendendo:

(a) proporcionar uma série de referência de seqüências correspondentes a uma seqüência de domínio variável de uma molécula scFv;

(b) identificar um aminoácido dentro do domínio variável o qual está ausente ou é encontrado em baixas freqüências em uma posição correspondente na série de referência;

(c) combinar esta informação com dados de covariação sugerindo que modificar o aminoácido identificado com baixa freqüência na série de referência pode satisfazer e de covariação existentes dentro do domínio variável; e

(d) substituir o aminoácido da etapa (b) com um aminoácido estabilizante candidato para deste modo produzir uma biblioteca compreendendo moléculas scFv estabilizadas.

56. Método de acordo com a reivindicação 55, em que o aminoácido da etapa (b) tem uma classificação de consenso de menos de 0,5.

57. Método para produzir uma molécula scFv estabilizada, o método compreendendo:

(a) proporcionar uma biblioteca de scFv designada de acordo com o método de acordo com a reivindicação 56,

(b) triar a biblioteca de scFv com uma prova de ensaio térmico para identificar moléculas scFv candidatas,

(c) comparar a estabilidade térmica de uma molécula scFv candidata da biblioteca de scFv com um controle adequado, por meio do qual um aumento na estabilidade térmica de uma molécula scFv candidata em relação ao controle identifica a molécula scFv candidata como uma molécula scFv estabilizada.

58. Molécula scFv estabilizada identificado usando o método de acordo com a reivindicação 57.

59. Proteína de fusão compreendendo a molécula scFv estabilizada de acordo com a reivindicação 58.

60. Método para prever a estabilidade de uma proteína candidata compreendendo uma seqüência de aminoácidos tendo seqüências de domínios candidatos, o método compreendendo:

(a) proporcionar uma série de referência de seqüências de aminoácidos correspondentes a uma seqüência de domínio candidato da proteína candidata;

(b) determinar freqüências de resíduos em posições de aminoácidos individuais dentro da seqüência de domínio de ensaio para obter uma

classificação de consenso; e

(c) usando a classificação de consenso para prever a estabilidade da proteína candidata, em que a classificação de consenso se correlaciona com a estabilidade da proteína candidata.

CAGGTCC AACTG GTGCAGT CTGGA GCTGAGG TGAAG AAGCCTG GGTC
CTCAGTG AAGGT GTCCTGC AAGGC TTCTGGC TACAC TTTCACA ACCT
ACTATTT GCACT GGGTGAG GCAGG CCCCTGG ACAGG GACTTGA GTGG
ATGGGAT GGATT TATCCTG GAAAT GTTCATG CTCAG TACAATG AGAA
GTTCAAG GGCAG GGTCACA ATCAC TGCAGAC AAATC CACCAGC ACAG
CCTACAT GGAGC TCAGCAG CCTGA GGTCTGA AGATA CTGCGGT CTAT
TACTGTG CAAGA TCCTGGG AAGGT TTTCCTT ACTGG GGCCAAG GGAC
CACGGTC ACCGT CTCCTCA **GGTGGGGGCGGATCTGGGGGCGGCGGAT**
CCGGTGGTGGTGGTAGTGA CATTC AGATGAC CCAGT CTCCTAG CTCC
CTGTCCG CCTCA GTAGGAG ACAGG GTCACCA TCACC TGCAAGG CCAG
TCAGAAT GTGGG TATTAAT GTAGC CTGGTAT CAACA GAAACCA GGGA
AGGCTCC TAAAT CACTGAT TTCCT CGGCCTC CTACC GGTACAG TGGA
GTCCCTT CCAGA TTCAGCG GCAGT GGATCTG GGACA GATTTCA CTCT
CACCATC AGCAG CCTCCAG CCTGA AGACTTC GCAAC CTATTTC TGTC
AGCAATA TGACA CCTATCC ATTCA CGTTCGG CCAGG GTACCAA GGTG
GAGATCA AA (SEQ ID NO: 3)

Fig. 1A

QVQLVQS GAEVK KPGSSVK VSCKA SGYTFTT YYLHW VRQAPGQ G
LEWMGWI YPGNV HAQYNEK FKGRV TITADKS TSTAY MELSSLR S
EDTAVYY CARSW EGFPYWG QGTTV TVSS**GGG GSGGGGSGGGGS**D
IQMTQSP SSLSA SVGDRVT ITCKA SQNVGIN VAWYQ QKPGKAP K
SLISSAS YRYSG VPSRFSG SGSGT DFTLTIS SLQPE DFATYFC Q
QYDTYPF TFGQG TKVEIK (SEQ ID NO: 4)

Fig. 1B

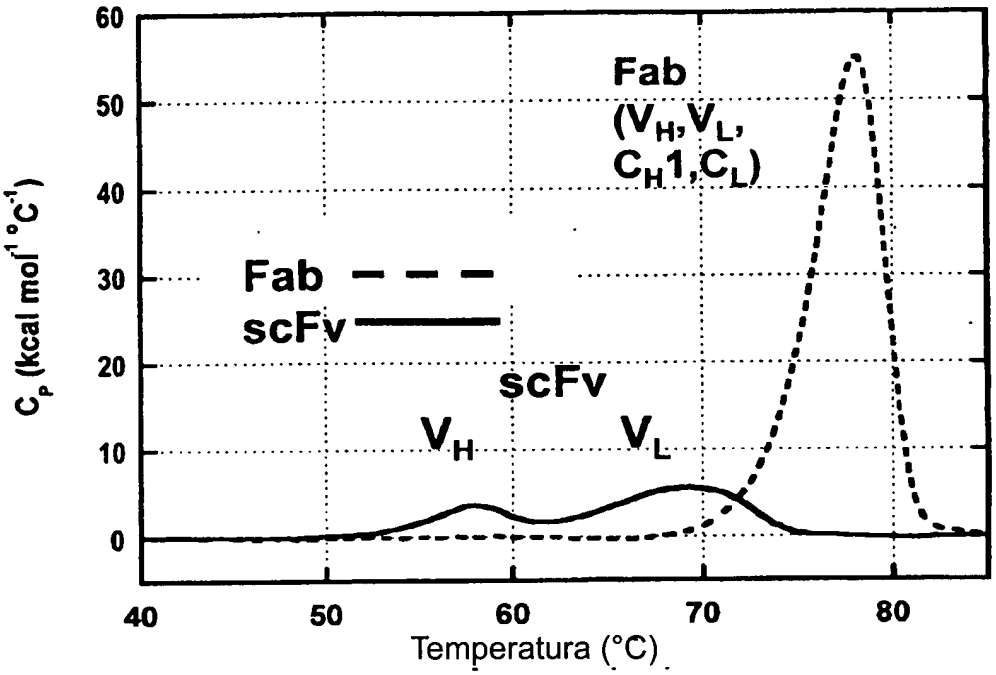

Fab
(V_H,V_L,
C_H1,C_L)
Fab
scFv
scFv
V_H
V_L
C_p (kcal mol^{-1} °C^{-1})
Temperatura (°C)

Fig. 2A

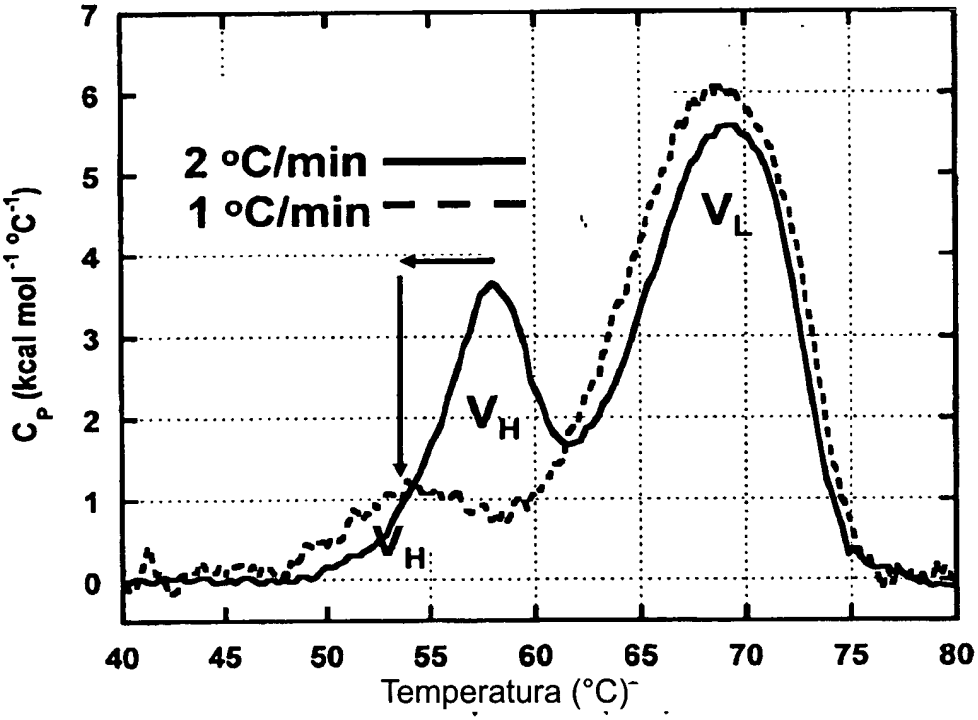

2 °C/min
1 °C/min
V_L
V_H
V_H
C_p (kcal mol^{-1} °C^{-1})
Temperatura (°C)

Fig. 2B

CAGGTCC AACTG GTGCAGT CTGGA GCTGAGG TGAAG AAGCCTG GGTC
CTCAGTG AAGGT GTCCTGC AAGGC TTCTGGC TACAC TTTCACA ACCT
ACTATTT GCACT GGGTGAG GCAGG CCCCTGG ACAGT GCCTTGA GTGG
ATGGGAT GGATT TATCCTG GAAAT GTTCATG CTCAG TACAATG AGAA
GTTCAAG GGCAG GGTCACA ATCAC TGCAGAC AAATC CACCAGC ACAG
CCTACAT GGAGC TCAGCAG CCTGA GGTCTGA AGATA CTGCGGT CTAT
TACTGTG CAAGA TCCTGGG AAGGT TTTCCTT ACTGG GGCCAAG GGAC
CACGGTC ACCGT CTCCTCA **GGTGGGGGCGGATCTGGGGGCGGCGGAT**
CCGGTGGTGGTGGTAGTGACATTCAGATGACCCAGTCTCCTAGCTCC
CTGTCCG CCTCA GTAGGAG ACAGG GTCACCA TCACC TGCAAGG CCAG
TCAGAAT GTGGG TATTAAT GTAGC CTGGTAT CAACA GAAACCA GGGA
AGGCTCC TAAAT CACTGAT TTCCT CGGCCTC CTACC GGTACAG TGGA
GTCCCTT CCAGA TTCAGCG GCAGT GGATCTG GGACA GATTTCA CTCT
CACCATC AGCAG CCTCCAG CCTGA AGACTTC GCAAC CTATTTC TGTC
AGCAATA TGACA CCTATCC ATTCA CGTTCGG CTGCG GTACCAA GGTG
GAGATCA AA (SEQ ID NO: 9)

Fig. 3A

QVQLVQS GAEVK KPGSSVK VSCKA SGYTFTT YYLHW VRQAPGQ *C*LE
WMGWIYP GNVHA QYNEKFK GRVTI TADKSTS TAYME LSSLRSE DTA
VYYCARS WEGFP YWGQGTT VTVSS **GGGGSGGGGSGGGGS**DIQM TQS
PSSLSAS VGDRV TIT *C*KASQNVGINVAWYQQKPGKAPKSLISSASY
RYSGVPS RFSGS GSGTDFT LTISS LQPEDFA TYFCQ QYDTYPF TFG
CGTKVEI K (SEQ ID NO:10)

Fig. 3B

CAGGTCC AACTG GTGCAGT CTGGA GCTGAGG TGAAG AAGCCTG GGTC
CTCAGTG AAGGT GTCCTGC AAGGC TTCTGGC TACAC TTTCACA ACCT
ACTATTT GCACT GGGTGAG GCAGG CCCCTGG ACAGG GACTTGA GTGG
ATGGGAT GGATT TATCCTG GAAAT GTTCATG CTCAG TACAATG AGAA
GTTCAAG GGCAG GGTCACA ATCAC TGCAGAC AAATC CACCAGC ACAG
CCTACAT GGAGC TCAGCAG CCTGA GGTCTGA AGATA CTGCGGT CTAT
TACTGTG CAAGA TCCTGGG AAGGT TTTCCTT ACTGG GGCCAAG GGAC
CACGGTC ACCGT CTCCTCA **GGCGGTGGAGGGTCCGGTGGGGGCGGAT**
CTGGGGGCGGCGGATCCGGTGGTGGTGGTAGTGACA TTCAGAT GACC
CAGTCTC CTAGC TCCCTGT CCGCC TCAGTAG GAGAC AGGGTCA CCAT
CACCTGC AAGGC CAGTCAG AATGT GGGTATT AATGT AGCCTGG TATC
AACAGAA ACCAG GGAAGGC TCCTA AATCACT GATTT CCTCGGC CTCC
TACCGGT ACAGT GGAGTCC CTTCC AGATTCA GCGGC AGTGGAT CTGG
GACAGAT TTCAC TCTCACC ATCAG CAGCCTC CAGCC TGAAGAC TTCG
CAACCTA TTTCT GTCAGCA ATATG ACACCTA TCCAT TCACGTT CGGC
CAGGGTA CCAAG GTGGAGA TCAAA (SEQ ID NO:14)

Fig. 4A

QVQLVQS GAEVK KPGSSVK VSCKA SGYTFTT YYLHW VRQAPGQ GL
EWMGWIY PGNVH AQYNEKF KGRVT ITADKST STAYM ELSSLRS ED
TAVYYCA RSWEG FPYWGQG TTVTV SS**GGGGS GGGGSGGGGSGGGG**
SDIQMTQ SPSSL SASVGDR VTITC KASQNVG INVAW YQQKPGK AP
KSLISSA SYRYS GVPSRFS GSGSG TDFTLTI SSLQP EDFATYF CQ
QYDTYPF TFGQG TKVEIK (SEQ ID NO:15)

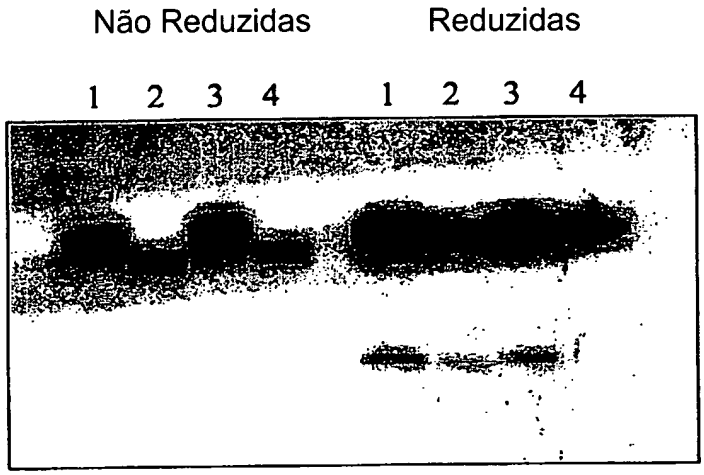
Fig. 4B

CAGGTCC AACTG GTGCAGT CTGGA GCTGAGG TGAAG AAGCCTG GGTCC
TCAGTGA AGGTG TCCTGCA AGGCT TCTGGCT ACACT TTCACAA CCTAC
TATTTGC ACTGG GTGAGGC AGGCC CCTGGAC AGGGA CTTGAGT GGATG
GGATGGA TTTAT CCTGGAA ATGTT CATGCTC AGTAC AATGAGA AGTTC
AAGGGCA GGGTC ACAATCA CTGCA GACAAAT CCACC AGCACAG CCTAC
ATGGAGC TCAGC AGCCTGA GGTCT GAAGATA CTGCG GTCTATT ACTGT
GCAAGAT CCTGG GAAGGTT TTCCT TACTGGG GCCAA GGGACCA CGGTC
ACCGTCT CCTCA **GGAGGGGGCGGTTCAGGCGGTGGAGGGTCCGGTGGG**
GGCGGATCTGGGGGCGGCGGATCCGGTGGTGGTGGTAGTGACA TTCAG
ATGACCC AGTCT CCTAGCT CCCTG TCCGCCT CAGTA GGAGACA GGGTC
ACCATCA CCTGC AAGGCCA GTCAG AATGTGG GTATT AATGTAG CCTGG
TATCAAC AGAAA CCAGGGA AGGCT CCTAAAT CACTG ATTTCCT CGGCC
TCCTACC GGTAC AGTGGAG TCCCT TCCAGAT TCAGC GGCAGTG GATCT
GGGACAG ATTTC ACTCTCA CCATC AGCAGCC TCCAG CCTGAAG ACTTC
GCAACCT ATTTC TGTCAGC AATAT GACACCT ATCCA TTCACGT TCGGC
CAGGGTA CCAAG GTGGAGA TCAAA (SEQ ID NO:16)

Fig. 5A

QVQLVQSGAEVKKPGSSVKVSCKASGYTFTTYYLHWVRQAPGQGLEWMGWIYPGN
VHAQYNEKFKGRVTITADKSTSTAYMELSSLRSEDTAVYYCARSWEGFPYWGQGTT
VTVSS**GGGGSGGGGSGGGGSGGGGSGGGGS**DIQMTQSPSSLSASVGDRVTITCKA
SQNVGINVAWYQQKPGKAPKSLISSASYRYSGVPSRFSGSGSGTDFTLTISSLQPEDFA
TYFCQQYDTYPFTFGQGTKVEIK (SEQ ID NO:17)

Fig. 5B



Alameda

1 – BHA10 scFv convencional
2 – BHA10 scFv V_H44/V_L100
3 – BHA10 scFv encadeador $(Gly_4Ser)_4$
4 – BHA10 scFv V_H44/V_L100 e encadeador $(Gly_4Ser)_4$

Fig. 6

CAGGTCC AACTG GTGCAGT CTGGA GCTGAGG TGAAG AAGCCTG GGTC
CTCAGTG AAGGT GTCCTGC AAGGC TTCTGGC TACAC TTTCACA ACCT
ACTATTT GCACT GGGTGAG GCAGG CCCCTGG ACAGT GCCTTGA GTGG
ATGGGAT GGATT TATCCTG GAAAT GTTCATG CTCAG TACAATG AGAA
GTTCAAG GGCAG GGTCACA ATCAC TGCAGAC AAATC CACCAGC ACAG
CCTACAT GGAGC TCAGCAG CCTGA GGTCTGA AGATA CTGCGGT CTAT
TACTGTG CAAGA TCCTGGG AAGGT TTTCCTT ACTGG GGCCAAG GGAC
CACGGTC ACCGT CTCCTCA **GGCGGTGGAGGGTCCGGTGGGGGCGGAT**
CTGGGGGCGGCGGATCCGGTGGTGGTGGTAGTGACA TTCAGAT GACC
CAGTCTC CTAGC TCCCTGT CCGCC TCAGTAG GAGAC AGGGTCA CCAT
CACCTGC AAGGC CAGTCAG AATGT GGGTATT AATGT AGCCTGG TATC
AACAGAA ACCAG GGAAGGC TCCTA AATCACT GATTT CCTCGGC CTCC
TACCGGT ACAGT GGAGTCC CTTCC AGATTCA GCGGC AGTGGAT CTGG
GACAGAT TTCAC TCTCACC ATCAG CAGCCTC CAGCC TGAAGAC TTCG
CAACCTA TTTCT GTCAGCA ATATG ACACCTA TCCAT TCACGTT CGGC
TGCGGTA CCAAG GTGGAGA TCAAA (SEQ ID NO:18)

Fig. 7A

QVQLVQS GAEVK KPGSSVK VSCKA SGYTFTT YYLHW VRQAPGQ ***C***L
EWMGWIY PGNVH AQYNEKF KGRVT ITADKST STAYM ELSSLRS ED
TAVYYCA RSWEG FPYWGQG TTVTV SS**GGGGS GGGGSGGGGSGGGG**
SDIQMTQ SPSSL SASVGDR VTIT ***C***KASQNVGINVAWYQQKPGKAP
KSLISSA SYRYS GVPSRFS GSGSG TDFTLTI SSLQP EDFATYF CQ
QYDTYPF TFGCG TKVEIK (SEQ ID NO:19)

Fig. 7B

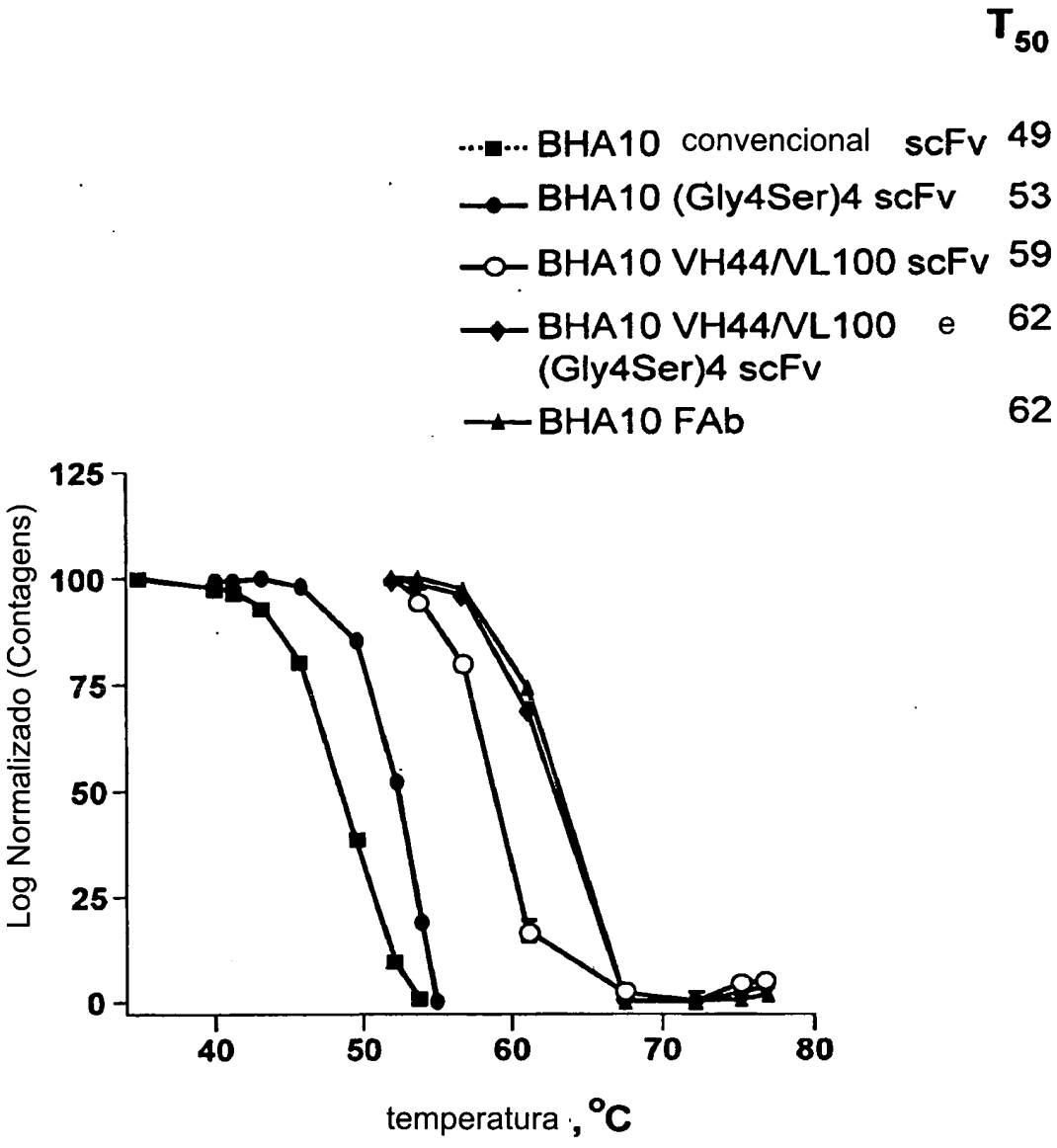

T_{50}
BHA10 convencional scFv 49
BHA10 (Gly4Ser)4 scFv 53
BHA10 VH44/VL100 scFv 59
BHA10 VH44/VL100 e (Gly4Ser)4 scFv 62
BHA10 FAb 62
Log Normalizado (Contagens)
125
100
75
50
25
0
40
50
60
70
80
temperatura , °C

Fig. 8

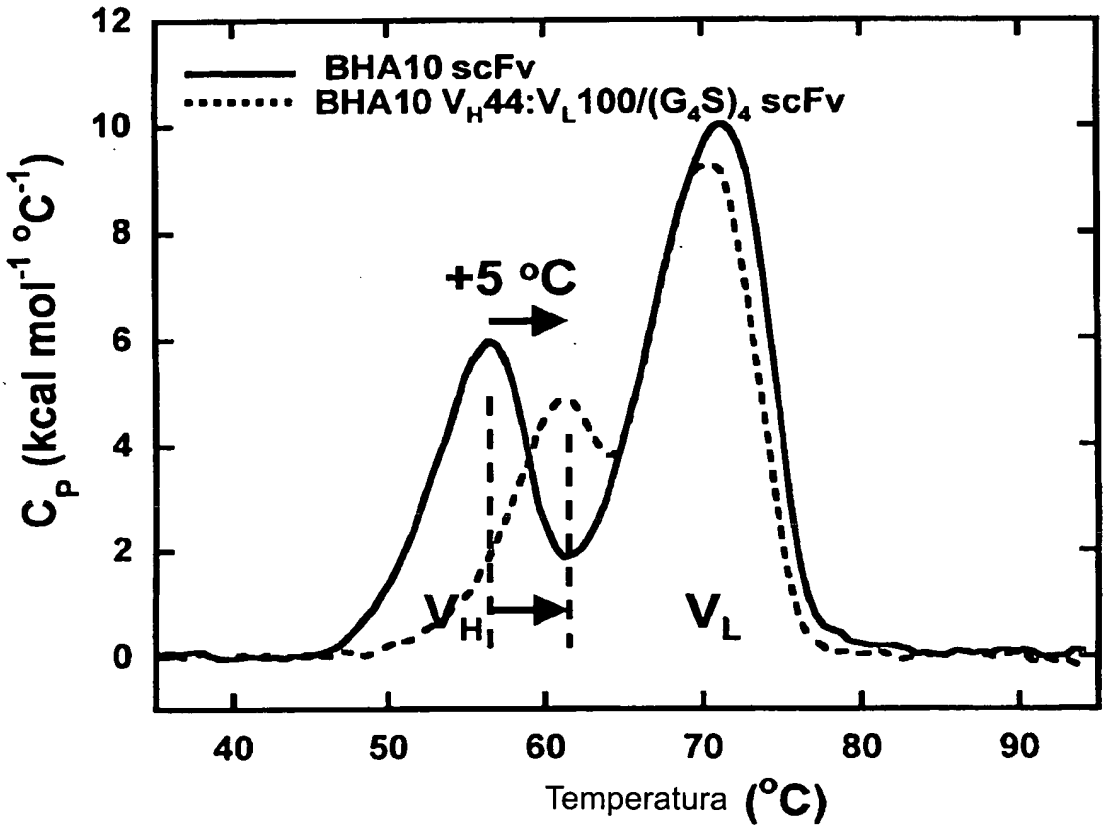

BHA10 scFv
BHA10 V_H44:V_L100/$(G_4S)_4$ scFv
+5 °C
V_{H_I}
V_L
C_p (kcal mol^{-1} °C^{-1})
Temperatura (°C)
0
2
4
6
8
10
12
40
50
60
70
80
90

Fig. 9

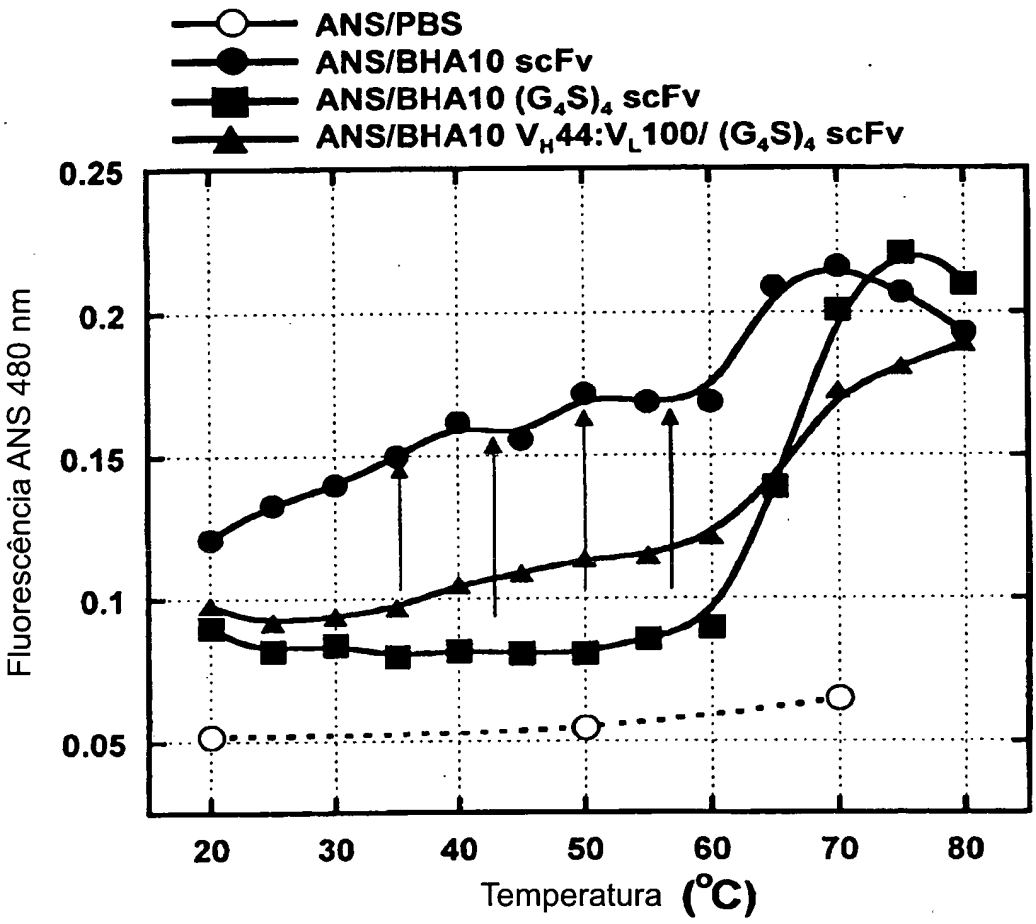

ANS/PBS
ANS/BHA10 scFv
ANS/BHA10 (G4S)4 scFv
ANS/BHA10 VH44:VL100/ (G4S)4 scFv
Fluorescência ANS 480 nm
0.25
0.2
0.15
0.1
0.05
20
30
40
50
60
70
80
Temperatura (°C)

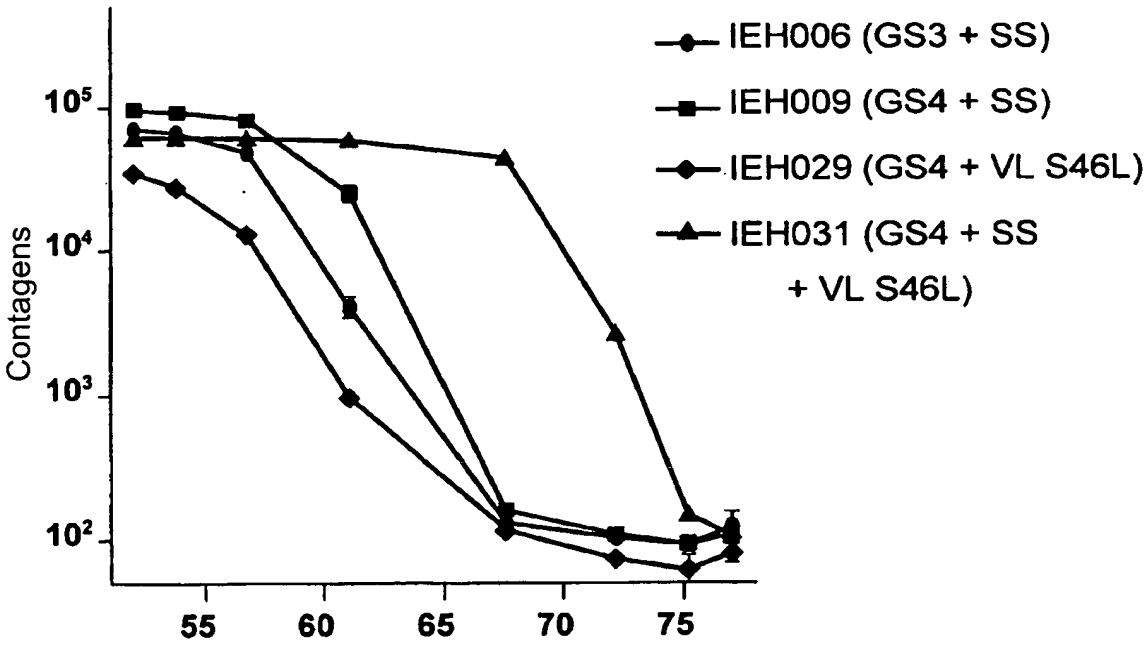
Fig. 10

Pos Nº	MFR	P(i)	SMFR	P(i)	SR	S(i)	SR/MFR	Kabat Nº	VH1cons	Chothia
9	G	0.593	P	0.220	A	0.169	0.285	9	A	sem dados
11	L	0.745	V	0.186	V	0.186	0.25	11	V	P(7),A(7),T(3), S(32),V(10),L(41)
12	V	0.796	K	0.101	K	0.101	0.127	12	K	T(3), G(3),K(6),A(22),V(57),L(4)
16	G	0.491	Q	0.152	S	0.033	0.068	16	S	D(15), R(2),E(26),Q(17),G(19),A(14)
18	L	0.728	V	0.186	V	0.186	0.255	18	V	A(14),I(2),S(4),T(2), V(52),L(29)
20	L	0.813	V	0.101	V	0.101	0.125	20	V	V(4),L(42),I(38),M(14)
45	K	0.830	Q	0.135	Q	0.135	0.163	43	Q	S(3),T(7), R(6),K(38),Q(33),N(2),H(3),G(3)
50	V	0.559	I	0.186	M	0.084	0.151	48	M	W(7), V(15),L(41),I(30),M(9)
74	F	0.610	A	0.118	V	0.118	0.194	67	V	I(4), A(17),T(4),V(8),L(4),F(61)
78	R	0.559	K	0.169	A	0.118	0.212	71	A	G(34),A(9),S(3), R(22),K(8),V(16),L(4)
80	N	0.627	T	0.203	K	0.118	0.189	73	K	G(42), K(17),D(3),N(15),T(18)
82	K	0.661	S	0.118	T	0.033	0.051	75	T	sem dados
87	L	0.847	M	0.118	M	0.118	0.14	80	M	L(76),M(20),F(3)
88	Q	0.694	E	0.067	E	0.067	0.097	81	E	E(7),Q(40),K(13),N(2),S(2),T(33)

Fig. 11A

Pos Nº	MFR	P(i)	SMFR	P(i)	SR	S(i)	SR/MFR	Kabat Nº	Vk1cons	Chothia
3	V	0.72	Q	0.16	Q	0.16	0.222	3	Q	sem dados
13	V	0.64	A	0.16	A	0.16	0.25	13	A	K(6),G(4), A(22),T(3),V(51),L(4)
48	Q	0.82	K	0.16	K	0.16	0.195	42	K	R(6),H(3), K(40), (Q33),G(3),T(7)
52	L	0.76	R	0.08	S	0.02	0.026	46	L	Y(2),W(53), R(6),G(2),T(4),V(15),L(32)
55	Y	0.88	E	0.04	S	0.02	0.022	49	Y	sem dados
56	W	0.3	L	0.12	S	0.12	0.4	50	A	sem dados
70	D	0.68	S	0.18	S	0.18	0.264	60	S	G(34), D(22),N(2),A(12),S(27),V(2)
95	V	0.58	T	0.16	T	0.16	0.275	85	T	L(3), D(6),T(27),V(44),I(8),M(9)
97	Y	0.84	F	0.14	F	0.14	0.166	87	Y	Y(79),F(19)

Fig. 11B



IEH006	IEH009	IEH029	IEH031
61.54	63.30	60.26	72.33

Fig. 12

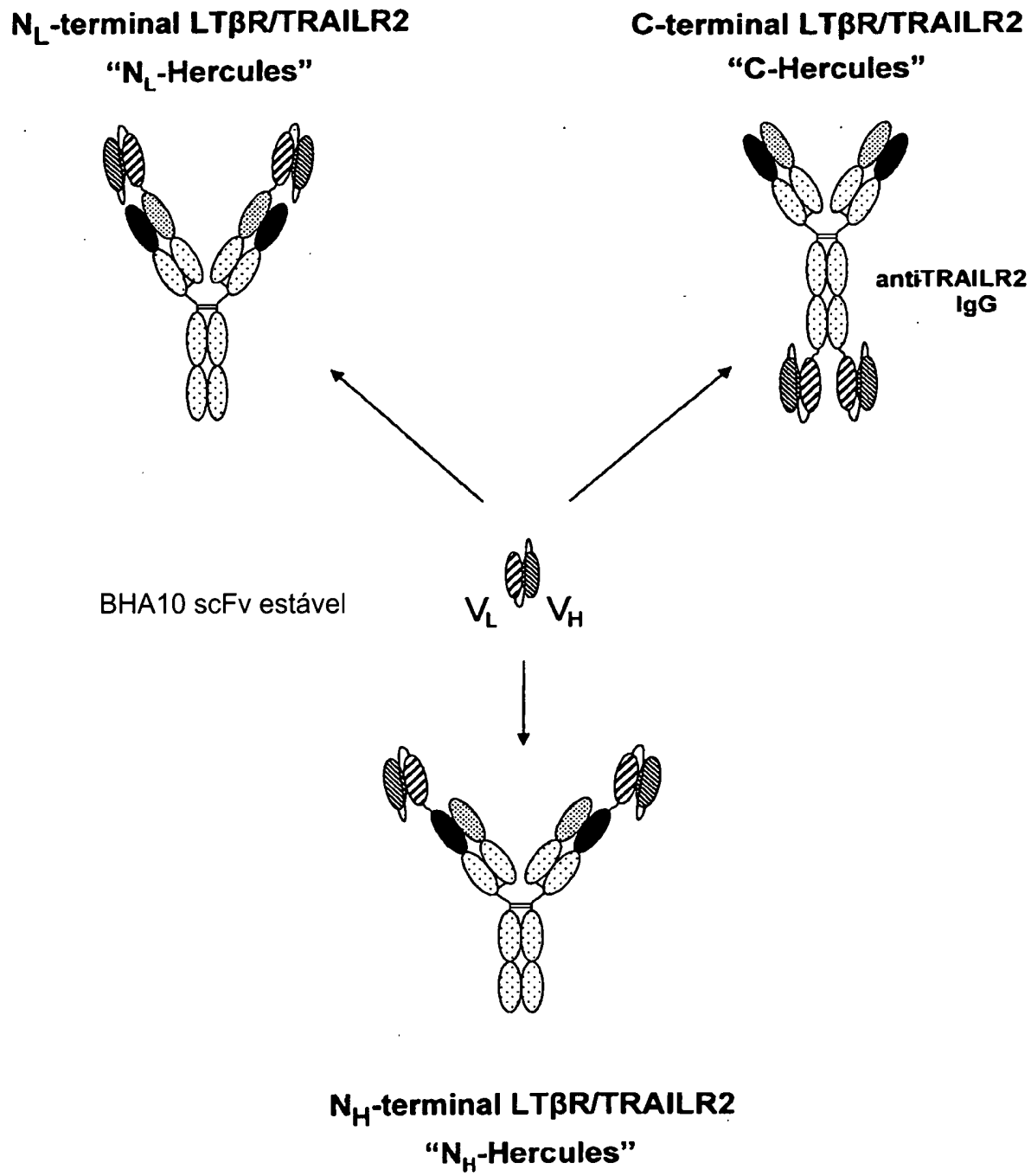

N_L-terminal LTβR/TRAILR2
"N_L-Hercules"
C-terminal LTβR/TRAILR2
"C-Hercules"
antiTRAILR2
IgG
BHA10 scFv estável
V_L V_H
N_H-terminal LTβR/TRAILR2
"N_H-Hercules"

Fig. 13

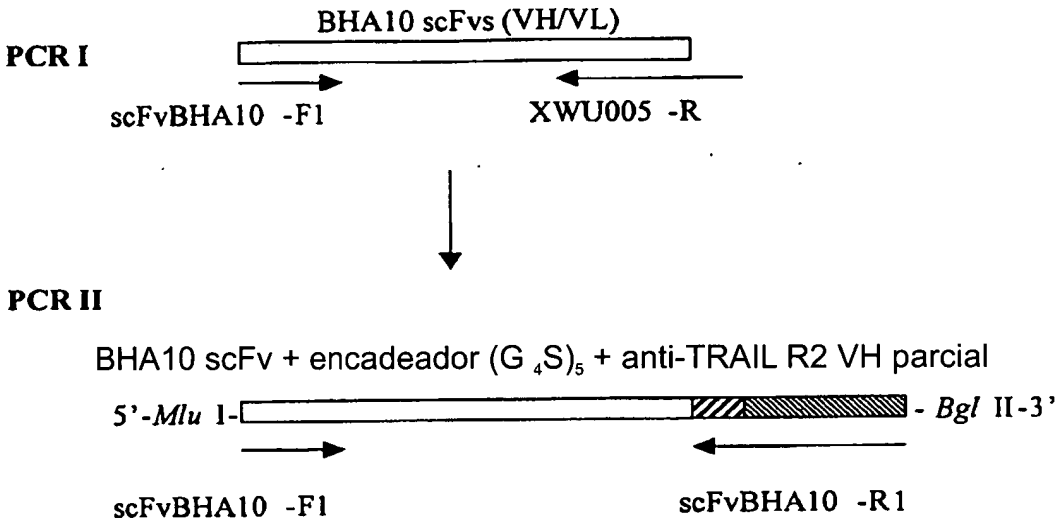


Fig. 14A

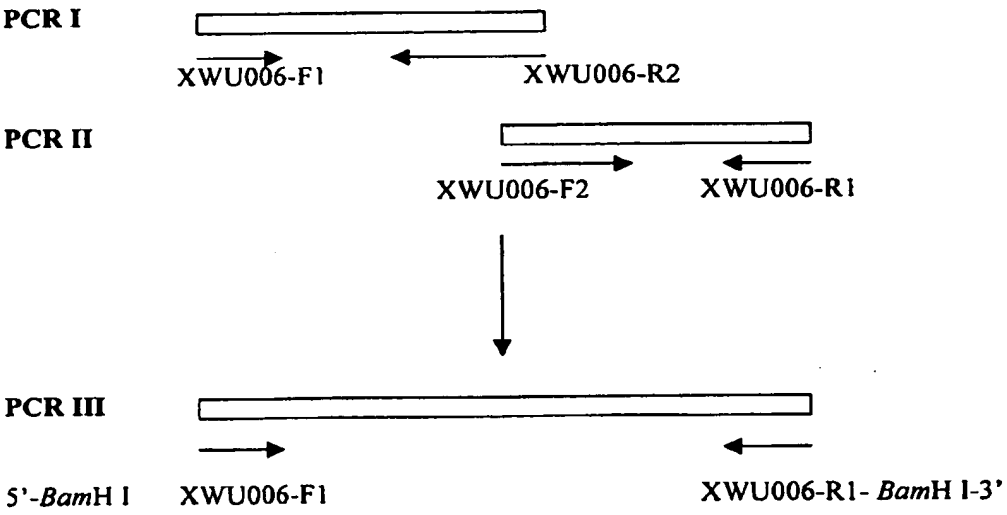


Fig. 14B

CAACTTG TGCTC ACTCAGT CATCT TCAGTCT CTTTC TCCCTGG GAGC
CTCAGCA AAACT CACGTGC ACCTT GAGTAGT CAGCA CAGTACG TACA
CCATTGA ATGGT ATCAGCA ACAGC CCCTCAA GCCTC CTAAGTA TGTG
ATGGAGC TTAAG AAAGATG GAAGC CACAGCA CAGGT GATGGGA TTCC
TGATCGC TTCTC TGGATCC AGCTC TGGTGCT GATCG CTACCTT AGCA
TTTCCAA CATCC AGCCTGA AGATG AAGCAAT ATACA TCTGTGG TGTG
GGTGATA CAATT AAGGAAC AATTT GTGTATG TTTTC GGCGGTG GAAC
CAAGGTC GAAAT CAAACGT ACGGT GGCTGCA CCATC TGTCTTC ATCT
TCCCGCC ATCTG ATGAGCA GTTGA AATCTGG AACTG CCTCTGT TGTG
TGCCTGC TGAAT AACTTCT ATCCC AGAGAGG CCAAA GTACAGT GGAA
GGTGGAT AACGC CCTCCAA TCGGG TAACTCC CAGGA GAGTGTC ACAG
AGCAGGA CAGCA AGGACAG CACCT ACAGCCT CAGCA GCACCCT GACG
CTGAGCA AAGCA GACTACG AGAAA CACAAAG TCTAC GCCTGCG AAGT
CACCCAT CAGGG CCTGAGC TCGCC CGTCACA AAGAG CTTCAAC AGGG
GAGAGTG TTGA (SEQ ID NO:28)

Fig. 15A

QLVLTQS SSVSF SLGASAK LTCTL SSQHSTY TIEWY QQQPLKP PK
YVMELKK DGSHS TGDGIPD RFSGS SSGADRY LSISN IQPEDEA IY
ICGVGDT IKEQF VYVFGGG TKVEI KRTVAAP SVFIF PPSDEQL KS
GTASVVC LLNNF YPREAKV QWKVD NALQSGN SQESV TEQDSKD ST
YSLSSTL TLSKA DYEKHKV YACEV THQGLSS PVTKS FNRGEC
(SEQ ID NO:29)

Fig. 15B

CAGGTCCAACTGGTGCAGTCTGGAGCTGAGGTGAAGAAGCCTGGGTCCTCAGTGAAGGT
GTCCTGCAAGGCTTCTGGCTACACTTTCACAACCTACTATTTGCACTGGGTGAGGCAGG
CCCCTGGACAGGGACTTGAGTGGATGGGATGGATTTATCCTGGAAATGTTCATGCTCAG
TACAATGAGAAGTTCAAGGGCAGGGTCACAATCACTGCAGACAAATCCACCAGCACAGC
CTACATGGAGCTCAGCAGCCTGAGGTCTGAAGATACTGCGGTCTATTACTGTGCAAGAT
CCTGGGAAGGTTTTCCTTACTGGGGCCAAGGGACCACGGTCACCGTCTCCTCAGGTGGG
GGCGGATCTGGGGGCGGCGGATCCGGTGGTGGTGGTAGTGACATTCAGATGACCCAGTC
TCCTAGCTCCCTGTCCGCCTCAGTAGGAGACAGGGTCACCATCACCTGCAAGGCCAGTC
AGAATGTGGGTATTAATGTAGCCTGGTATCAACAGAAACCAGGGAAGGCTCCTAAATCA
CTGATTTCCTCGGCCTCCTACCGGTACAGTGGAGTCCCTTCCAGATTCAGCGGCAGTGG
ATCTGGGACAGATTTCACTCTCACCATCAGCAGCCTCCAGCCTGAAGACTTCGCAACCT
ATTTCTGTCAGCAATATGACACCTATCCATTCACGTTCGGCCAGGGTACCAAGGTGGAG
ATCAAAGGCGGTGGAGGGTCCGGTGGAGGGGGCTCTGGAGGGGGCGGTTCAGGGGGCGG
TGGATCGGGGGGAGGTGGCTCCCAGATCCAGTTGGTGCAGTCTGGACCTGAGCTGAAGA
AGCCTGGAGAGACAGTCAAGATCTCCTGCAAGGCTTCTGGTTTTACCTTCACAGACTAT
TCAATACACTGGGTGAAACAGGCTCCAGGAAAGGGTTTAAAGTGGATGGGCTGGATAAA
CACTGAGACTGGTGAGCCAACATATACAGATGACTTCAAGGGACGATTTGCCTTCTCTT
TGGTGACCTCTGCCACCACTGCCTATTTGCAGATCAACAACCTCAACAATGAGGACACG
GCTACATTTTTCTGTGCTAGATTCATCTATGATCCTTATTGGGGGTTTGCTTACTGGGG
CCAGGGGACTCTGGTCACTGTCTCCGCAGCTAGCACCAAGGGCCCATCGGTCTTCCCCC
TGGCACCCTCCTCCAAGAGCACCTCTGGGGGCACAGCGGCCCTGGGCTGCCTGGTCAAG
GACTACTTCCCCGAACCGGTGACGGTGTCGTGGAACTCAGGCGCCCTGACCAGCGGCGT
GCACACCTTCCCGGCTGTCCTACAGTCCTCAGGACTCTACTCCCTCAGCAGCGTGGTGA
CCGTGCCCTCCAGCAGCTTGGGCACCCAGACCTACATCTGCAACGTGAATCACAAGCCC
AGCAACACCAAGGTGGACAAGAAAGTTGAGCCCAAATCTTGTGACAAAACTCACACATG
CCCACCGTGCCCAGCACCTGAACTCCTGGGGGGACCGTCAGTCTTCCTCTTCCCCCCAA
AACCCAAGGACACCCTCATGATCTCCCGGACCCCTGAGGTCACATGCGTGGTGGTGGAC
GTGAGCCACGAAGACCCTGAGGTCAAGTTCAACTGGTACGTGGACGGCGTGGAGGTGCA
TAATGCCAAGACAAAGCCGCGGGAGGAGCAGTACAACAGCACGTACCGTGTGGTCAGCG
TCCTCACCGTCCTGCACCAGGACTGGCTGAATGGCAAGGAGTACAAGTGCAAGGTCTCC
AACAAAGCCCTCCCAGCCCCCATCGAGAAAACCATCTCCAAAGCCAAAGGGCAGCCCCG
AGAACCACAGGTGTACACCCTGCCCCCATCCCGGGATGAGCTGACCAAGAACCAGGTCA
GCCTGACCTGCCTGGTCAAAGGCTTCTATCCCAGCGACATCGCCGTGGAGTGGGAGAGC
AATGGGCAGCCGGAGAACAACTACAAGACCACGCCTCCCGTGCTGGACTCCGACGGCTC
CTTCTTCCTCTACAGCAAGCTCACCGTGGACAAGAGCAGGTGGCAGCAGGGGAACGTCT
TCTCATGCTCCGTGATGCATGAGGCTCTGCACAACCACTACACGCAGAAGAGCCTCTCC
CTGTCTCCGGGTTGA (SEQ ID NO:30)

Fig. 16

QVQLVQS GAEVK KPGSSVK VSCKA SGYTFTT YYLHW VRQAPGQ GL
EWMGWIY PGNVH AQYNEKF KGRVT ITADKST STAYM ELSSLRS ED
TAVYYCA RSWEG FPYWGQG TTVTV SSGGGGS GGGGS GGGGSDI QM
TQSPSSL SASVG DRVTITC KASQN VGINVAW YQQKP GKAPKSL IS
SASYRYS GVPSR FSGSGSG TDFTL TISSLQP EDFAT YFCQQYD TY
PFTFGQG TKVEI KGGGGSG GGGSG GGGSGGG GSGGG GSQIQLV QS
GPELKKP GETVK ISCKASG FTFTD YSIHWVK QAPGK GLKWMGW IN
TETGEPT YTDDF KGRFAFS LVTSA TTAYLQI NNLNN EDTATFF CA
RFIYDPY WGFAY WGQGTLV TVSAA STKGPSV FPLAP SSKSTSG GT
AALGCLV KDYFP EPVTVSW NSGAL TSGVHTF PAVLQ SSGLYSL SS
VVTVPSS SLGTQ TYICNVN HKPSN TKVDKKV EPKSC DKTHTCP PC
PAPELLG GPSVF LFPPKPK DTLMI SRTPEVT CVVVD VSHEDPE VK
FNWYVDG VEVHN AKTKPRE EQYNS TYRVVSV LTVLH QDWLNGK EY
KCKVSNK ALPAP IEKTISK AKGQP REPQVYT LPPSR DELTKNQ VS
LTCLVKG FYPSD IAVEWES NGQPE NNYKTTP PVLDS DGSFFLY SK
LTVDKSR WQQGN VFSCSVM HEALH NHYTQKS LSLSP G (SEQ ID
NO:31)

Fig. 17

CAGGTCCAACTGGTGCAGTCTGGAGCTGAGGTGAAGAAGCCTGGGTCCTCAGTGAAGGTG
TCCTGCAAGGCTTCTGGCTACACTTTCACAACCTACTATTTGCACTGGGTGAGGCAGGCC
CCTGGACAGGGACTTGAGTGGATGGGATGGATTTATCCTGGAAATGTTCATGCTCAGTAC
AATGAGAAGTTCAAGGGCAGGGTCACAATCACTGCAGACAAATCCACCAGCACAGCCTAC
ATGGAGCTCAGCAGCCTGAGGTCTGAAGATACTGCGGTCTATTACTGTGCAAGATCCTGG
GAAGGTTTTCCTTACTGGGGCCAAGGGACCACGGTCACCGTCTCCTCAGGCGGTGGAGGG
TCCGGTGGGGGCGGATCTGGGGGCGGCGGATCCGGTGGTGGTGGTAGTGACATTCAGATG
ACCCAGTCTCCTAGCTCCCTGTCCGCCTCAGTAGGAGACAGGGTCACCATCACCTGCAAG
GCCAGTCAGAATGTGGGTATTAATGTAGCCTGGTATCAACAGAAACCAGGGAAGGCTCCT
AAATCACTGATTTCCTCGGCCTCCTACCGGTACAGTGGAGTCCCTTCCAGATTCAGCGGC
AGTGGATCTGGGACAGATTTCACTCTCACCATCAGCAGCCTCCAGCCTGAAGACTTCGCA
ACCTATTTCTGTCAGCAATATGACACCTATCCATTCACGTTCGGCCAGGGTACCAAGGTG
GAGATCAAAGGCGGTGGAGGGTCCGGTGGAGGGGGCTCTGGAGGGGGCGGTTCAGGGGGC
GGTGGATCGGGGGGGAGGTGGCTCCCAGATCCAGTTGGTGCAGTCTGGACCTGAGCTGAAG
AAGCCTGGAGAGACAGTCAAGATCTCCTGCAAGGCTTCTGGTTTTACCTTCACAGACTAT
TCAATACACTGGGTGAAACAGGCTCCAGGAAAGGGTTTAAAGTGGATGGGCTGGATAAAC
ACTGAGACTGGTGAGCCAACATATACAGATGACTTCAAGGGACGATTTGCCTTCTCTTTG
GTGACCTCTGCCACCACTGCCTATTTGCAGATCAACAACCTCAACAATGAGGACACGGCT
ACATTTTTCTGTGCTAGATTCATCTATGATCCTTATTGGGGGTTTGCTTACTGGGGCCAG
GGGACTCTGGTCACTGTCTCCGCAGCTAGCACCAAGGGCCCATCGGTCTTCCCCCTGGCA
CCCTCCTCCAAGAGCACCTCTGGGGGCACAGCGGCCCTGGGCTGCCTGGTCAAGGACTAC
TTCCCCGAACCGGTGACGGTGTCGTGGAACTCAGGCGCCCTGACCAGCGGCGTGCACACC
TTCCCGGCTGTCCTACAGTCCTCAGGACTCTACTCCCTCAGCAGCGTGGTGACCGTGCCC
TCCAGCAGCTTGGGCACCCAGACCTACATCTGCAACGTGAATCACAAGCCCAGCAACACC
AAGGTGGACAAGAAAGTTGAGCCCAAATCTTGTGACAAAACTCACACATGCCCACCGTGC
CCAGCACCTGAACTCCTGGGGGGACCGTCAGTCTTCCTCTTCCCCCCAAAACCCAAGGAC
ACCCTCATGATCTCCCGGACCCCTGAGGTCACATGCGTGGTGGTGGACGTGAGCCACGAA
GACCCTGAGGTCAAGTTCAACTGGTACGTGGACGGCGTGGAGGTGCATAATGCCAAGACA
AAGCCGCGGGAGGAGCAGTACAACAGCACGTACCGTGTGGTCAGCGTCCTCACCGTCCTG
CACCAGGACTGGCTGAATGGCAAGGAGTACAAGTGCAAGGTCTCCAACAAAGCCCTCCCA
GCCCCCATCGAGAAAACCATCTCCAAAGCCAAAGGGCAGCCCCGAGAACCACAGGTGTAC
ACCCTGCCCCCATCCCGGGATGAGCTGACCAAGAACCAGGTCAGCCTGACCTGCCTGGTC
AAAGGCTTCTATCCCAGCGACATCGCCGTGGAGTGGGAGAGCAATGGGCAGCCGGAGAAC
AACTACAAGACCACGCCTCCCGTGCTGGACTCCGACGGCTCCTTCTTCCTCTACAGCAAG
CTCACCGTGGACAAGAGCAGGTGGCAGCAGGGGAACGTCTTCTCATGCTCCGTGATGCAT
GAGGCTCTGCACAACCACTACACGCAGAAGAGCCTCTCCCTGTCTCCGGGTTGA
(SEQ ID NO:32)

Fig. 18

QVQLVQS GAEVK KPGSSVK VSCKA SGYTFTT YYLHW VRQAPGQ GLEW
MGWIYPG NVHAQ YNEKFKG RVTIT ADKSTST AYMEL SSLRSED TAVY
YCARSWE GFPYW GQGTTVT VSSGG GGSGGGG SGGGG SGGGGSD IQMT
QSPSSLS ASVGD RVTITCK ASQNV GINVAWY QQKPG KAPKSLI SSAS
YRYSGVP SRFSG SGSGTDF TLTIS SLQPEDF ATYFC QQYDTYP FTFG
QGTKVEI KGGGG SGGGGSG GGGSG GGGSGGG GSQIQ LVQSGPE LKKP
GETVKIS CKASG FTFTDYS IHWVK QAPGKGL KWMGW INTETGE PTYT
DDFKGRF AFSLV TSATTAY LQINN LNNEDTA TFFCA RFIYDPY WGFA
YWGQGTL VTVSA ASTKGPS VFPLA PSSKSTS GGTAA LGCLVKD YFPE
PVTVSWN SGALT SGVHTFP AVLQS SGLYSLS SVVTV PSSSLGT QTYI
CNVNHKP SNTKV DKKVEPK SCDKT HTCPPCP APELL GGPSVFL FPPK
PKDTLMI SRTPE VTCVVVD VSHED PEVKFNW YVDGV EVHNAKT KPRE
EQYNSTY RVVSV LTVLHQD WLNGK EYKCKVS NKALP APIEKTI SKAK
GQPREPQ VYTLP PSRDELT KNQVS LTCLVKG FYPSD IAVEWES NGQP
ENNYKTT PPVLD SDGSFFL YSKLT VDKSRWQ QGNVF SCSVMHE ALHN
HYTQKSL SLSPG (SEQ ID NO:33)

Fig. 19

CAGGTCC AACT GGT GCAGTCT GGAG CTGAGGTGAA GAAG CCTGGGTCCT CAGTG
AAGGTGT CCTG CAAGGCTTCT GGCT ACACTTTCAC AACC TACTATTTGC ACTGG
GTGAGGC AGGC CCC TGGACAG TGCC TTGAGTGGAT GGGA TGGATTTATC CTGGA
AATGTTC ATGC TCAGTACAAT GAGA AGTTCAAGGG CAGG GTCACAATCA CTGCA
GACAAAT CCAC CAG CACAGCC TACA TGGAGCTCAG CAGC CTGAGGTCTG AAGAT
ACTGCGG TCTA TTA CTGTGCA AGAT CCTGGGAAGG TTTT CCTTACTGGG GCCAA
GGGACCA CGGT CAC CGTCTCC TCAG GTGGGGGCGG ATCT GGGGGCGGCG GATCC
GGTGGTG GTGG TAG TGACATT CAGA TGA CCCAGTC TCCT AGCTCCCTGT CCGCC
TCAGTAG GAGA CAGGGTCACC ATCA CCTGCAAGGC CAGT CAGAATGTGG GTATT
AATGTAG CCTG GTA TCAACAG AAAC CAGGGAAGGC TCCT AAATCACTGA TTTCC
TCGGCCT CCTA CCG GTACAGT GGAG TCC CTTCCAG ATTC AGCGGCAGTG GATCT
GGGACAG ATTT CACTCTCACC ATCA GCAGCCTCCA GCCT GAAGACTTCG CAACC
TATTTCT GTCA GCAATATGAC ACCT ATC CATTCAC GTTC GGCTGCGGTA CCAAG
GTGGAGA TCAA AGG CGGTGGA GGGT CCGGTGGAGG GGGC TCTGGAGGGG GCGGT
TCAGGGG GCGG TGGATCGGGG GGAG GTGGCTCCCA GATC CAGTTGGTGC AGTCT
GGACCTG AGCT GAAGAAGCCT GGAG AGA CAGTCAA GATC TCCTGCAAGG CTTCT
GGTTTTA CCTT CACAGACTAT TCAA TACACTGGGT GAAA CAGGCTCCAG GAAAG
GGTTTAA AGTG GATGGGCTGG ATAA ACA CTGAGAC TGGT GAGCCAACAT ATACA
GATGACT TCAA GGGACGATTT GCCT TCTCTTTGGT GACC TCTGCCACCA CTGCC
TATTTGC AGAT CAACAACCTC AACA ATGAGGACAC GGCT ACATTTTTCT GTGCT
AGATTCA TCTA TGATCCTTAT TGGG GGTTTGCTTA CTGG GGC CAGGGGA CTCTG
GTCACTG TCTC CGCAGCTAGC ACCA AGGGCCCATC GGTC TTCCCCCTGG CACCC
TCCTCCA AGAG CAC CTCTGGG GGCA CAG CGGCCCT GGGC TGC CTGGTCA AGGAC
TACTTCC CCGA ACC GGTGACG GTGT CGTGGAACTC AGGC GCC CTGACCA GCGGC
GTGCACA CCTT CCC GGCTGTC CTAC AGT CCTCAGG ACTC TACTCCCTCA GCAGC
GTGGTGA CCGT GCC CTCCAGC AGCT TGGGCACCCA GACC TACATCTGCA ACGTG
AATCACA AGCC CAG CAACACC AAGG TGGACAAGAA AGTT GAGCCCAAAT CTTGT
GACAAAA CTCA CACATGCCCA CCGT GCC CAGCACC TGAA CTCCTGGGGG GACCG
TCAGTCT TCCT CTTCCCCCCA AAAC CCAAGGACAC CCTCATGATCTCCC GGACC
CCTGAGG TCAC ATG CGTGGTG GTGGACGTGAGCCA CGAA GACCCTGAGG TCAAG
TTCAACT GGTA CGTGGACGGC GTGGAGGTGCATAA TGCC AAGACAAAGC CGCGG
GAGGAGC AGTA CAA CAGCACG TACC GTGTGGTCAG CGTC CTCACCGTCC TGCAC
CAGGACT GGCT GAATGGCAAG GAGT ACAAGTGCAA GGTC TCCAAC
(SEQ ID NO:34)

Fig. 20

QVQLVQS GAEVK KPGSSVK VSCKA SGYTFTT YYLHW VRQAPGQ CLE
WMGWIYP GNVHA QYNEKFK GRVTI TADKSTS TAYME LSSLRSE DTA
VYYCARS WEGFP YWGQGTT VTVSS GGGGSGG GGSGG GGSDIQM TQS
PSSLSAS VGDRV TITCKAS QNVGI NVAWYQQ KPGKA PKSLISS ASY
RYSGVPS RFSGS GSGTDFT LTISS LQPEDFA TYFCQ QYDTYPF TFG
CGTKVEI KGGGG SGGGGSG GGGSG GGGSGGG GSQIQ LVQSGPE LKK
PGETVKI SCKAS GFTFTDY SIHWV KQAPGKG LKWMG WINTETG EPT
YTDDFKG RFAFS LVTSATT AYLQI NNLNNED TATFF CARFIYD PYW
GFAYWGQ GTLVT VSAASTK GPSVF PLAPSSK STSGG TAALGCL VKD
YFPEPVT VSWNS GALTSGV HTFPA VLQSSGL YSLSS VVTVPSS SLG
TQTYICN VNHKP SNTKVDK KVEPK SCDKTHT CPPCP APELLGG PSV
FLFPPKP KDTLM ISRTPEV TCVVV DVSHEDP EVKFN WYVDGVE VHN
AKTKPRE EQYNS TYRVVSV LTVLH QDWLNGK EYKCK VSNKALP API
EKTISKA KGQPR EPQVYTL PPSRD ELTKNQV SLTCL VKGFYPS DIA
VEWESNG QPENN YKTTPPV LDSDG SFFLYSK LTVDK SRWQQGN VFS
CSVMHEA LHNHY TQKSLSL SPG

(SEQ ID NO:35)

Fig. 21

CAGGTCCAACTGGTGCAGTCTGGAGCTGAGGTGAAGAAGCCTGGGTCCTCAGTGAAGGTGT
CCTGCAAGGCTTCTGGCTACACTTTCACAACCTACTATTTGCACTGGGTGAGGCAGGCCCC
TGGACAGTGCCTTGAGTGGATGGGATGGATTTATCCTGGAAATGTTCATGCTCAGTACAAT
GAGAAGTTCAAGGGCAGGGTCACAATCACTGCAGACAAATCCACCAGCACAGCCTACATGG
AGCTCAGCAGCCTGAGGTCTGAAGATACTGCGGTCTATTACTGTGCAAGATCCTGGGAAGG
TTTTCCTTACTGGGGCCAAGGGACCACGGTCACCGTCTCCTCAGGCGGTGGAGGGTCCGGT
GGGGGCGGATCTGGGGGCGGCGGATCCGGTGGTGGTGGTAGTGACATTCAGATGACCCAGT
CTCCTAGCTCCCTGTCCGCCTCAGTAGGAGACAGGGTCACCATCACCTGCAAGGCCAGTCA
GAATGTGGGTATTAATGTAGCCTGGTATCAACAGAAACCAGGGAAGGCTCCTAAATCACTG
ATTTCCTCGGCCTCCTACCGGTACAGTGGAGTCCCTTCCAGATTCAGCGGCAGTGGATCTG
GGACAGATTTCACTCTCACCATCAGCAGCCTCCAGCCTGAAGACTTCGCAACCTATTTCTG
TCAGCAATATGACACCTATCCATTCACGTTCGGCTGCGGTACCAAGGTGGAGATCAAAGGC
GGTGGAGGGTCCGGTGGAGGGGGCTCTGGAGGGGGCGGTTCAGGGGGCGGTGGATCGGGGG
GAGGTGGCTCCCAGATCCAGTTGGTGCAGTCTGGACCTGAGCTGAAGAAGCCTGGAGAGAC
AGTCAAGATCTCCTGCAAGGCTTCTGGTTTTACCTTCACAGACTATTCAATACACTGGGTG
AAACAGGCTCCAGGAAAGGGTTTAAAGTGGATGGGCTGGATAAACACTGAGACTGGTGAGC
CAACATATACAGATGACTTCAAGGGACGATTTGCCTTCTCTTTGGTGACCTCTGCCACCAC
TGCCTATTTGCAGATCAACAACCTCAACAATGAGGACACGGCTACATTTTTCTGTGCTAGA
TTCATCTATGATCCTTATTGGGGGTTTGCTTACTGGGGCCAGGGGACTCTGGTCACTGTCT
CCGCAGCTAGCACCAAGGGCCCATCGGTCTTCCCCCTGGCACCCTCCTCCAAGAGCACCTC
TGGGGGCACAGCGGCCCTGGGCTGCCTGGTCAAGGACTACTTCCCCGAACCGGTGACGGTG
TCGTGGAACTCAGGCGCCCTGACCAGCGGCGTGCACACCTTCCCGGCTGTCCTACAGTCCT
CAGGACTCTACTCCCTCAGCAGCGTGGTGACCGTGCCCTCCAGCAGCTTGGGCACCCAGAC
CTACATCTGCAACGTGAATCACAAGCCCAGCAACACCAAGGTGGACAAGAAAGTTGAGCCC
AAATCTTGTGACAAAACTCACACATGCCCACCGTGCCCAGCACCTGAACTCCTGGGGGGAC
CGTCAGTCTTCCTCTTCCCCCCAAAACCCAAGGACACCCTCATGATCTCCCGGACCCCTGA
GGTCACATGCGTGGTGGTGGACGTGAGCCACGAAGACCCTGAGGTCAAGTTCAACTGGTAC
GTGGACGGCGTGGAGGTGCATAATGCCAAGACAAAGCCGCGGGAGGAGCAGTACAACAGCA
CGTACCGTGTGGTCAGCGTCCTCACCGTCCTGCACCAGGACTGGCTGAATGGCAAGGAGTA
CAAGTGCAAGGTCTCCAACAAAGCCCTCCCAGCCCCCATCGAGAAAACCATCTCCAAAGCC
AAAGGGCAGCCCCGAGAACCACAGGTGTACACCCTGCCCCCATCCCGGGATGAGCTGACCA
AGAACCAGGTCAGCCTGACCTGCCTGGTCAAAGGCTTCTATCCCAGCGACATCGCCGTGGA
GTGGGAGAGCAATGGGCAGCCGGAGAACAACTACAAGACCACGCCTCCCGTGCTGGACTCC
GACGGCTCCTTCTTCCTCTACAGCAAGCTCACCGTGGACAAGAGCAGGTGGCAGCAGGGGA
ACGTCTTCTCATGCTCCGTGATGCATGAGGCTCTGCACAACCACTACACGCAGAAGAGCCT
CTCCCTGTCTCCGGGTTGA (SEQ ID NO:36)

Fig. 22

QVQLVQS GAEVK KPGSSVK VSCKA SGYTFTT YYLHW VRQAPGQ CLEW
MGWIYPG NVHAQ YNEKFKG RVTIT ADKSTST AYMEL SSLRSED TAVY
YCARSWE GFPYW GQGTTVT VSSGG GGSGGGG SGGGG SGGGGSD IQMT
QSPSSLS ASVGD RVTITCK ASQNV GINVAWY QQKPG KAPKSLI SSAS
YRYSGVP SRFSG SGSGTDF TLTIS SLQPEDF ATYFC QQYDTYP FTFG
CGTKVEI KGGGG SGGGGSG GGGSG GGGSGGG GSQIQ LVQSGPE LKKP
GETVKIS CKASG FTFTDYS IHWVK QAPGKGL KWMGW INTETGE PTYT
DDFKGRF AFSLV TSATTAY LQINN LNNEDTA TFFCA RFIYDPY WGFA
YWGQGTL VTVSA ASTKGPS VFPLA PSSKSTS GGTAA LGCLVKD YFPE
PVTVSWN SGALT SGVHTFP AVLQS SGLYSLS SVVTV PSSSLGT QTYI
CNVNHKP SNTKV DKKVEPK SCDKT HTCPPCP APELL GGPSVFL FPPK
PKDTLMI SRTPE VTCVVVD VSHED PEVKFNW YVDGV EVHNAKT KPRE
EQYNSTY RVVSV LTVLHQD WLNGK EYKCKVS NKALP APIEKTI SKAK
GQPREPQ VYTLP PSRDELT KNQVS LTCLVKG FYPSD IAVEWES NGQP
ENNYKTT PPVLD SDGSFFL YSKLT VDKSRWQ QGNVF SCSVMHE ALHN
HYTQKSL SLSPG (SEQ ID NO:37)

Fig. 23

CAGATCCAGTTGGTGCAGTCTGGACCTGAGCTGAAGAAGCCTGGAGAGACAGTCAAG
ATCTCCTGCAAGGCTTCTGGTTTTACCTTCACAGACTATTCAATACACTGGGTGAAA
CAGGCTCCAGGAAAGGGTTTAAAGTGGATGGGCTGGATAAACACTGAGACTGGTGAG
CCAACATATACAGATGACTTCAAGGGACGATTTGCCTTCTCTTTGGTGACCTCTGCC
ACCACTGCCTATTTGCAGATCAACAACCTCAACAATGAGGACACGGCTACATTTTTC
TGTGCTAGATTCATCTATGATCCTTATTGGGGGTTTGCTTACTGGGGCCAGGGGACT
CTGGTCACTGTCTCCGCAGCTAGCACCAAGGGCCCATCGGTCTTCCCCCTGGCACCC
TCCTCCAAGAGCACCTCTGGGGGCACAGCGGCCCTGGGCTGCCTGGTCAAGGACTAC
TTCCCCGAACCGGTGACGGTGTCGTGGAACTCAGGCGCCCTGACCAGCGGCGTGCAC
ACCTTCCCGGCTGTCCTACAGTCCTCAGGACTCTACTCCCTCAGCAGCGTGGTGACC
GTGCCCTCCAGCAGCTTGGGCACCCAGACCTACATCTGCAACGTGAATCACAAGCCC
AGCAACACCAAGGTGGACAAGAAAGTTGAGCCCAAATCTTGTGACAAAACTCACACA
TGCCCACCGTGCCCAGCACCTGAACTCCTGGGGGGACCGTCAGTCTTCCTCTTCCCC
CCAAAACCCAAGGACACCCTCATGATCTCCCGGACCCCTGAGGTCACATGCGTGGTG
GTGGACGTGAGCCACGAAGACCCTGAGGTCAAGTTCAACTGGTACGTGGACGGCGTG
GAGGTGCATAATGCCAAGACAAAGCCGCGGGAGGAGCAGTACAACAGCACGTACCGT
GTGGTCAGCGTCCTCACCGTCCTGCACCAGGACTGGCTGAATGGCAAGGAGTACAAG
TGCAAGGTCTCCAACAAAGCCCTCCCAGCCCCCATCGAGAAAACCATCTCCAAAGCC
AAAGGGCAGCCCCGAGAACCACAGGTGTACACCCTGCCCCCATCCCGGGATGAGCTG
ACCAAGAACCAGGTCAGCCTGACCTGCCTGGTCAAAGGCTTCTATCCCAGCGACATC
GCCGTGGAGTGGGAGAGCAATGGGCAGCCGGAGAACAACTACAAGACCACGCCTCCC
GTGCTGGACTCCGACGGCTCCTTCTTCCTCTACAGCAAGCTCACCGTGGACAAGAGC
AGGTGGCAGCAGGGGAACGTCTTCTCATGCTCCGTGATGCATGAGGCTCTGCACAAC
CACTACACGCAGAAGAGCCTCTCCCTGTCTCCGGGTAAATCCGGCGGGGGTGGATCC
GGTGGAGGGGGCTCCGGCGGTGGCGGGTCCCAGGTCCAACTGGTGCAGTCTGGAGCT
GAGGTGAAGAAGCCTGGGTCCTCAGTGAAGGTGTCCTGCAAGGCTTCTGGCTACACT
TTCACAACCTACTATTTGCACTGGGTGAGGCAGGCCCCTGGACAGGGACTTGAGTGG
ATGGGATGGATTTATCCTGGAAATGTTCATGCTCAGTACAATGAGAAGTTCAAGGGC
AGGGTCACAATCACTGCAGACAAATCCACCAGCACAGCCTACATGGAGCTCAGCAGC
CTGAGGTCTGAAGATACTGCGGTCTATTACTGTGCAAGATCCTGGGAAGGTTTTCCT
TACTGGGGCCAAGGGACCACGGTCACCGTCTCCTCAGGTGGGGGCGGATCTGGGGGC
GGCGGGTCCGGTGGTGGTGGTAGTGACATTCAGATGACCCAGTCTCCTAGCTCCCTG
TCCGCCTCAGTAGGAGACAGGGTCACCATCACCTGCAAGGCCAGTCAGAATGTGGGT
ATTAATGTAGCCTGGTATCAACAGAAACCAGGGAAGGCTCCTAAATCACTGATTTCC
TCGGCCTCCTACCGGTACAGTGGAGTCCCTTCCAGATTCAGCGGCAGTGGATCTGGG
ACAGATTTCACTCTCACCATCAGCAGCCTCCAGCCTGAAGACTTCGCAACCTATTTC
TGTCAGCAATATGACACCTATCCATTCACGTTCGGCCAGGGTACCAAGGTGGAGATC
AAATGA (SEQ ID NO:44)

Fig. 24

QIQLVQS GPELK KPGETVK ISCKA SGFTFTD YSIHW VKQAPGK G
LKWMGWI NTETG EPTYTDD FKGRF AFSLVTS ATTAY LQINNLN N
EDTATFF CARFI YDPYWGF AYWGQ GTLVTVS AASTK GPSVFPL A
PSSKSTS GGTAA LGCLVKD YFPEP VTVSWNS GALTS GVHTFPA V
LQSSGLY SLSSV VTVPSSS LGTQT YICNVNH KPSNT KVDKKVE P
KSCDKTH TCPPC PAPELLG GPSVF LFPPKPK DTLMI SRTPEVT C
VVVDVSH EDPEV KFNWYVD GVEVH NAKTKPR EEQYN STYRVVS V
LTVLHQD WLNGK EYKCKVS NKALP APIEKTI SKAKG QPREPQV Y
TLPPSRD ELTKN QVSLTCL VKGFY PSDIAVE WESNG QPENNYK T
TPPVLDS DGSFF LYSKLTV DKSRW QQGNVFS CSVMH EALHNHY T
QKSLSLS PGKSG GGGSGGG GSGGG GSQVQLV QSGAE VKKPGSS V
KVSCKAS GYTFT TYYLHWV RQAPG QGLEWMG WIYPG NVHAQYN E
KFKGRVT ITADK STSTAYM ELSSL RSEDTAV YYCAR SWEGFPY W
GQGTTVT VSSGG GGSGGGG SGGGG SDIQMTQ SPSSL SASVGDR V
TITCKAS QNVGI NVAWYQQ KPGKA PKSLISS ASYRY SGVPSRF S
GSGSGTD FTLTI SSLQPED FATYF CQQYDTY PFTFG QGTKVEI K
(SEQ ID NO:45)

Fig. 25

CAGATCCAGTTGGTGCAGTCTGGACCTGAGCTGAAGAAGCCTGGAGAGACAGTCAAGATC
TCCTGCAAGGCTTCTGGTTTTACCTTCACAGACTATTCAATACACTGGGTGAAACAGGCT
CCAGGAAAGGGTTTAAAGTGGATGGGCTGGATAAACACTGAGACTGGTGAGCCAACATAT
ACAGATGACTTCAAGGGACGATTTGCCTTCTCTTTGGTGACCTCTGCCACCACTGCCTAT
TTGCAGATCAACAACCTCAACAATGAGGACACGGCTACATTTTTCTGTGCTAGATTCATC
TATGATCCTTATTGGGGGTTTGCTTACTGGGGCCAGGGGACTCTGGTCACTGTCTCCGCA
GCTAGCACCAAGGGCCCATCGGTCTTCCCCCTGGCACCCTCCTCCAAGAGCACCTCTGGG
GGCACAGCGGCCCTGGGCTGCCTGGTCAAGGACTACTTCCCCGAACCGGTGACGGTGTCG
TGGAACTCAGGCGCCCTGACCAGCGGCGTGCACACCTTCCCGGCTGTCCTACAGTCCTCA
GGACTCTACTCCCTCAGCAGCGTGGTGACCGTGCCCTCCAGCAGCTTGGGCACCCAGACC
TACATCTGCAACGTGAATCACAAGCCCAGCAACACCAAGGTGGACAAGAAAGTTGAGCCC
AAATCTTGTGACAAAACTCACACATGCCCACCGTGCCCAGCACCTGAACTCCTGGGGGGA
CCGTCAGTCTTCCTCTTCCCCCCAAAACCCAAGGACACCCTCATGATCTCCCGGACCCCT
GAGGTCACATGCGTGGTGGTGGACGTGAGCCACGAAGACCCTGAGGTCAAGTTCAACTGG
TACGTGGACGGCGTGGAGGTGCATAATGCCAAGACAAAGCCGCGGGAGGAGCAGTACAAC
AGCACGTACCGTGTGGTCAGCGTCCTCACCGTCCTGCACCAGGACTGGCTGAATGGCAAG
GAGTACAAGTGCAAGGTCTCCAACAAAGCCCTCCCAGCCCCCATCGAGAAAACCATCTCC
AAAGCCAAAGGGCAGCCCCGAGAACCACAGGTGTACACCCTGCCCCCATCCCGGGATGAG
CTGACCAAGAACCAGGTCAGCCTGACCTGCCTGGTCAAAGGCTTCTATCCCAGCGACATC
GCCGTGGAGTGGGAGAGCAATGGGCAGCCGGAGAACAACTACAAGACCACGCCTCCCGTG
CTGGACTCCGACGGCTCCTTCTTCCTCTACAGCAAGCTCACCGTGGACAAGAGCAGGTGG
CAGCAGGGGAACGTCTTCTCATGCTCCGTGATGCATGAGGCTCTGCACAACCACTACACG
CAGAAGAGCCTCTCCCTGTCTCCGGGTAAATCCGGCGGGGGTGGATCCGGTGGAGGGGGC
TCCGGCGGTGGCGGGTCCCAGGTCCAACTGGTGCAGTCTGGAGCTGAGGTGAAGAAGCCT
GGGTCCTCAGTGAAGGTGTCCTGCAAGGCTTCTGGCTACACTTTCACAACCTACTATTTG
CACTGGGTGAGGCAGGCCCCTGGACAGGGACTTGAGTGGATGGGATGGATTTATCCTGGA
AATGTTCATGCTCAGTACAATGAGAAGTTCAAGGGCAGGGTCACAATCACTGCAGACAAA
TCCACCAGCACAGCCTACATGGAGCTCAGCAGCCTGAGGTCTGAAGATACTGCGGTCTAT
TACTGTGCAAGATCCTGGGAAGGTTTTCCTTACTGGGGCCAAGGGACCACGGTCACCGTC
TCCTCAGGCGGTGGAGGGTCCGGTGGGGGCGGATCTGGGGGCGGCGGGTCCGGTGGTGGT
GGTAGTGACATTCAGATGACCCAGTCTCCTAGCTCCCTGTCCGCCTCAGTAGGAGACAGG
GTCACCATCACCTGCAAGGCCAGTCAGAATGTGGGTATTAATGTAGCCTGGTATCAACAG
AAACCAGGGAAGGCTCCTAAATCACTGATTTCCTCGGCCTCCTACCGGTACAGTGGAGTC
CCTTCCAGATTCAGCGGCAGTGGATCTGGGACAGATTTCACTCTCACCATCAGCAGCCTC
CAGCCTGAAGACTTCGCAACCTATTTCTGTCAGCAATATGACACCTATCCATTCACGTTC
GGCCAGGGTACCAAGGTGGAGATCAAATGA (SEQ ID NO:46)

Fig. 26

QIQLVQS GPELK KPGETVK ISCKA SGFTFTD YSIHW VKQAPGK GLKW
MGWINTE TGEPT YTDDFKG RFAFS LVTSATT AYLQI NNLNNED TATF
FCARFIY DPYWG FAYWGQG TLVTV SAASTKG PSVFP LAPSSKS TSGG
TAALGCL VKDYF PEPVTVS WNSGA LTSGVHT FPAVL QSSGLYS LSSV
VTVPSSS LGTQT YICNVNH KPSNT KVDKKVE PKSCD KTHTCPP CPAP
ELLGGPS VFLFP PKPKDTL MISRT PEVTCVV VDVSH EDPEVKF NWYV
DGVEVHN AKTKP REEQYNS TYRVV SVLTVLH QDWLN GKEYKCK VSNK
ALPAPIE KTISK AKGQPRE PQVYT LPPSRDE LTKNQ VSLTCLV KGFY
PSDIAVE WESNG QPENNYK TTPPV LDSDGSF FLYSK LTVDKSR WQQG
NVFSCSV MHEAL HNHYTQK SLSLS PGKSGGG GSGGG GSGGGGS QVQL
VQSGAEV KKPGS SVKVSCK ASGYT FTTYYLH WVRQA PGQGLEW MGWI
YPGNVHA QYNEK FKGRVTI TADKS TSTAYME LSSLR SEDTAVY YCAR
SWEGFPY WGQGT TVTVSSG GGGSG GGGSGGG GSGGG GSDIQMT QSPS
SLSASVG DRVTI TCKASQN VGINV AWYQQKP GKAPK SLISSAS YRYS
GVPSRFS GSGSG TDFTLTI SSLQP EDFATYF CQQYD TYPFTFG QGTK
VEIK (SEQ ID NO:47)

Fig. 27

CAGATCCAGTTGGTGCAGTCTGGACCTGAGCTGAAGAAGCCTGGAGAGACAGTCAAGAT
CTCCTGCAAGGCTTCTGGTTTTACCTTCACAGACTATTCAATACACTGGGTGAAACAGG
CTCCAGGAAAGGGTTTAAAGTGGATGGGCTGGATAAACACTGAGACTGGTGAGCCAACA
TATACAGATGACTTCAAGGGACGATTTGCCTTCTCTTTGGTGACCTCTGCCACCACTGC
CTATTTGCAGATCAACAACCTCAACAATGAGGACACGGCTACATTTTTCTGTGCTAGAT
TCATCTATGATCCTTATTGGGGGTTTGCTTACTGGGGCCAGGGGACTCTGGTCACTGTC
TCCGCAGCTAGCACCAAGGGCCCATCGGTCTTCCCCCTGGCACCCTCCTCCAAGAGCAC
CTCTGGGGGCACAGCGGCCCTGGGCTGCCTGGTCAAGGACTACTTCCCCGAACCGGTGA
CGGTGTCGTGGAACTCAGGCGCCCTGACCAGCGGCGTGCACACCTTCCCGGCTGTCCTA
CAGTCCTCAGGACTCTACTCCCTCAGCAGCGTGGTGACCGTGCCCTCCAGCAGCTTGGG
CACCCAGACCTACATCTGCAACGTGAATCACAAGCCCAGCAACACCAAGGTGGACAAGA
AAGTTGAGCCCAAATCTTGTGACAAAACTCACACATGCCCACCGTGCCCAGCACCTGAA
CTCCTGGGGGGACCGTCAGTCTTCCTCTTCCCCCCAAAACCCAAGGACACCCTCATGAT
CTCCCGGACCCCTGAGGTCACATGCGTGGTGGTGGACGTGAGCCACGAAGACCCTGAGG
TCAAGTTCAACTGGTACGTGGACGGCGTGGAGGTGCATAATGCCAAGACAAAGCCGCGG
GAGGAGCAGTACAACAGCACGTACCGTGTGGTCAGCGTCCTCACCGTCCTGCACCAGGA
CTGGCTGAATGGCAAGGAGTACAAGTGCAAGGTCTCCAACAAAGCCCTCCCAGCCCCCA
TCGAGAAAACCATCTCCAAAGCCAAAGGGCAGCCCCGAGAACCACAGGTGTACACCCTG
CCCCCATCCCGGGATGAGCTGACCAAGAACCAGGTCAGCCTGACCTGCCTGGTCAAAGG
CTTCTATCCCAGCGACATCGCCGTGGAGTGGGAGAGCAATGGGCAGCCGGAGAACAACT
ACAAGACCACGCCTCCCGTGCTGGACTCCGACGGCTCCTTCTTCCTCTACAGCAAGCTC
ACCGTGGACAAGAGCAGGTGGCAGCAGGGGAACGTCTTCTCATGCTCCGTGATGCATGA
GGCTCTGCACAACCACTACACGCAGAAGAGCCTCTCCCTGTCTCCGGGTAAATCCGGCG
GGGGTGGATCCGGTGGAGGGGGCTCCGGCGGTGGCGGGTCCCAGGTCCAACTGGTGCAG
TCTGGAGCTGAGGTGAAGAAGCCTGGGTCCTCAGTGAAGGTGTCCTGCAAGGCTTCTGG
CTACACTTTCACAACCTACTATTTGCACTGGGTGAGGCAGGCCCCTGGACAGTGCCTTG
AGTGGATGGGATGGATTTATCCTGGAAATGTTCATGCTCAGTACAATGAGAAGTTCAAG
GGCAGGGTCACAATCACTGCAGACAAATCCACCAGCACAGCCTACATGGAGCTCAGCAG
CCTGAGGTCTGAAGATACTGCGGTCTATTACTGTGCAAGATCCTGGGAAGGTTTTCCTT
ACTGGGGCCAAGGGACCACGGTCACCGTCTCCTCAGGTGGGGGCGGATCTGGGGGCGGC
GGGTCCGGTGGTGGTGGTAGTGACATTCAGATGACCCAGTCTCCTAGCTCCCTGTCCGC
CTCAGTAGGAGACAGGGTCACCATCACCTGCAAGGCCAGTCAGAATGTGGGTATTAATG
TAGCCTGGTATCAACAGAAACCAGGGAAGGCTCCTAAATCACTGATTTCCTCGGCCTCC
TACCGGTACAGTGGAGTCCCTTCCAGATTCAGCGGCAGTGGATCTGGGACAGATTTCAC
TCTCACCATCAGCAGCCTCCAGCCTGAAGACTTCGCAACCTATTTCTGTCAGCAATATG
ACACCTATCCATTCACGTTCGGCTGCGGTACCAAGGTGGAGATCAAATGA
(SEQ ID NO:48)

Fig. 28

QIQLVQS GPELK KPGETVK ISCKA SGFTFTD YSIHW VKQAPGK GLK
WMGWINT ETGEP TYTDDFK GRFAF SLVTSAT TAYLQ INNLNNE DTA
TFFCARF IYDPY WGFAYWG QGTLV TVSAAST KGPSV FPLAPSS KST
SGGTAAL GCLVK DYFPEPV TVSWN SGALTSG VHTFP AVLQSSG LYS
LSSVVTV PSSSL GTQTYIC NVNHK PSNTKVD KKVEP KSCDKTH TCP
PCPAPEL LGGPS VFLFPPK PKDTL MISRTPE VTCVV VDVSHED PEV
KFNWYVD GVEVH NAKTKPR EEQYN STYRVVS VLTVL HQDWLNG KEY
KCKVSNK ALPAP IEKTISK AKGQP REPQVYT LPPSR DELTKNQ VSL
TCLVKGF YPSDI AVEWESN GQPEN NYKTTPP VLDSD GSFFLYS KLT
VDKSRWQ QGNVF SCSVMHE ALHNH YTQKSLS LSPGK SGGGGSG GGG
SGGGGSQ VQLVQ SGAEVKK PGSSV KVSCKAS GYTFT TYYLHWV RQA
PGQCLEW MGWIY PGNVHAQ YNEKF KGRVTIT ADKST STAYMEL SSL
RSEDTAV YYCAR SWEGFPY WGQGT TVTVSSG GGGSG GGGSGGG GSD
IQMTQSP SSLSA SVGDRVT ITCKA SQNVGIN VAWYQ QKPGKAP KSL
ISSASYR YSGVP SRFSGSG SGTDF TLTISSL QPEDF ATYFCQQ YDT
YPFTFGC GTKVE IK (SEQ ID NO:49)

Fig. 29

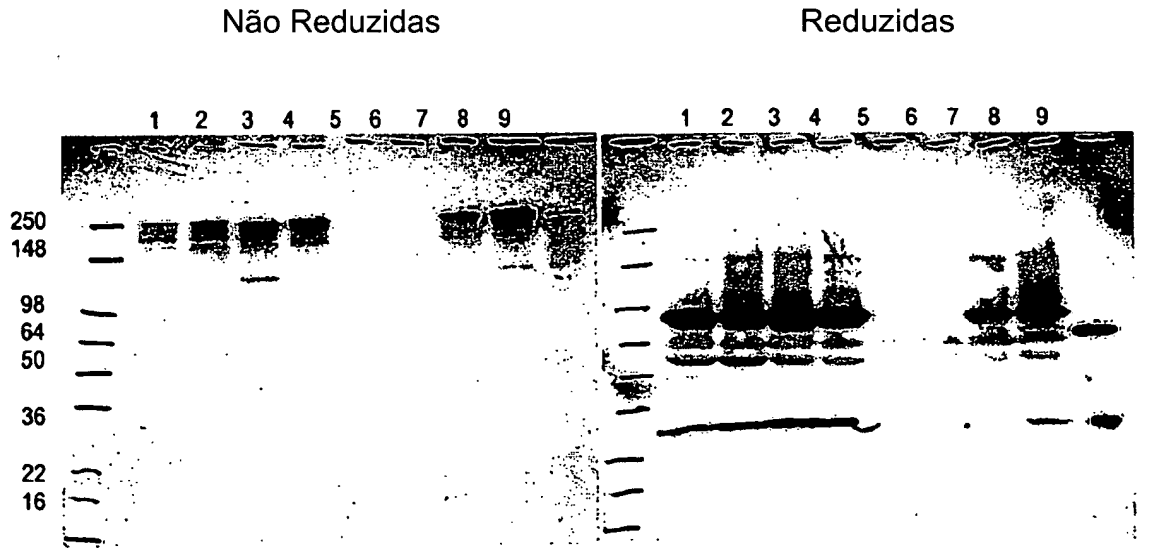
CAGATCCAGTTGGTGCAGTCTGGACCTGAGCTGAAGAAGCCTGGAGAGACAGTCAAG
ATCTCCTGCAAGGCTTCTGGTTTTACCTTCACAGACTATTCAATACACTGGGTGAAA
CAGGCTCCAGGAAAGGGTTTAAAGTGGATGGGCTGGATAAACACTGAGACTGGTGAG
CCAACATATACAGATGACTTCAAGGGACGATTTGCCTTCTCTTTGGTGACCTCTGCC
ACCACTGCCTATTTGCAGATCAACAACCTCAACAATGAGGACACGGCTACATTTTTC
TGTGCTAGATTCATCTATGATCCTTATTGGGGGTTTGCTTACTGGGGCCAGGGGACT
CTGGTCACTGTCTCCGCAGCTAGCACCAAGGGCCCATCGGTCTTCCCCCTGGCACCC
TCCTCCAAGAGCACCTCTGGGGGCACAGCGGCCCTGGGCTGCCTGGTCAAGGACTAC
TTCCCCGAACCGGTGACGGTGTCGTGGAACTCAGGCGCCCTGACCAGCGGCGTGCAC
ACCTTCCCGGCTGTCCTACAGTCCTCAGGACTCTACTCCCTCAGCAGCGTGGTGACC
GTGCCCTCCAGCAGCTTGGGCACCCAGACCTACATCTGCAACGTGAATCACAAGCCC
AGCAACACCAAGGTGGACAAGAAAGTTGAGCCCAAATCTTGTGACAAAACTCACACA
TGCCCACCGTGCCCAGCACCTGAACTCCTGGGGGGACCGTCAGTCTTCCTCTTCCCC
CCAAAACCCAAGGACACCCTCATGATCTCCCGGACCCCTGAGGTCACATGCGTGGTG
GTGGACGTGAGCCACGAAGACCCTGAGGTCAAGTTCAACTGGTACGTGGACGGCGTG
GAGGTGCATAATGCCAAGACAAAGCCGCGGGAGGAGCAGTACAACAGCACGTACCGT
GTGGTCAGCGTCCTCACCGTCCTGCACCAGGACTGGCTGAATGGCAAGGAGTACAAG
TGCAAGGTCTCCAACAAAGCCCTCCCAGCCCCCATCGAGAAAACCATCTCCAAAGCC
AAAGGGCAGCCCCGAGAACCACAGGTGTACACCCTGCCCCCATCCCGGGATGAGCTG
ACCAAGAACCAGGTCAGCCTGACCTGCCTGGTCAAAGGCTTCTATCCCAGCGACATC
GCCGTGGAGTGGGAGAGCAATGGGCAGCCGGAGAACAACTACAAGACCACGCCTCCC
GTGCTGGACTCCGACGGCTCCTTCTTCCTCTACAGCAAGCTCACCGTGGACAAGAGC
AGGTGGCAGCAGGGGAACGTCTTCTCATGCTCCGTGATGCATGAGGCTCTGCACAAC
CACTACACGCAGAAGAGCCTCTCCCTGTCTCCGGGTAAATCCGGCGGGGGTGGATCC
GGTGGAGGGGGCTCCGGCGGTGGCGGGTCCCAGGTCCAACTGGTGCAGTCTGGAGCT
GAGGTGAAGAAGCCTGGGTCCTCAGTGAAGGTGTCCTGCAAGGCTTCTGGCTACACT
TTCACAACCTACTATTTGCACTGGGTGAGGCAGGCCCCTGGACAGTGCCTTGAGTGG
ATGGGATGGATTTATCCTGGAAATGTTCATGCTCAGTACAATGAGAAGTTCAAGGGC
AGGGTCACAATCACTGCAGACAAATCCACCAGCACAGCCTACATGGAGCTCAGCAGC
CTGAGGTCTGAAGATACTGCGGTCTATTACTGTGCAAGATCCTGGGAAGGTTTTCCT
TACTGGGGCCAAGGGACCACGGTCACCGTCTCCTCAGGCGGTGGAGGGTCCGGTGGG
GGCGGATCTGGGGGCGGCGGGTCCGGTGGTGGTGGTAGTGACATTCAGATGACCCAG
TCTCCTAGCTCCCTGTCCGCCTCAGTAGGAGACAGGGTCACCATCACCTGCAAGGCC
AGTCAGAATGTGGGTATTAATGTAGCCTGGTATCAACAGAAACCAGGGAAGGCTCCT
AAATCACTGATTTCCTCGGCCTCCTACCGGTACAGTGGAGTCCCTTCCAGATTCAGC
GGCAGTGGATCTGGGACAGATTTCACTCTCACCATCAGCAGCCTCCAGCCTGAAGAC
TTCGCAACCTATTTCTGTCAGCAATATGACACCTATCCATTCACGTTCGGCTGCGGT
ACCAAGGTGGAGATCAAATGA (SEQ ID NO:50)

Fig. 30

QIQLVQS GPELK KPGETVK ISCKA SGFTFTD YSIHW VKQAPGK GLK
WMGWINT ETGEP TYTDDFK GRFAF SLVTSAT TAYLQ INNLNNE DTA
TFFCARF IYDPY WGFAYWG QGTLV TVSAAST KGPSV FPLAPSS KST
SGGTAAL GCLVK DYFPEPV TVSWN SGALTSG VHTFP AVLQSSG LYS
LSSVVTV PSSSL GTQTYIC NVNHK PSNTKVD KKVEP KSCDKTH TCP
PCPAPEL LGGPS VFLFPPK PKDTL MISRTPE VTCVV VDVSHED PEV
KFNWYVD GVEVH NAKTKPR EEQYN STYRVVS VLTVL HQDWLNG KEY
KCKVSNK ALPAP IEKTISK AKGQP REPQVYT LPPSR DELTKNQ VSL
TCLVKGF YPSDI AVEWESN GQPEN NYKTTPP VLDSD GSFFLYS KLT
VDKSRWQ QGNVF SCSVMHE ALHNH YTQKSLS LSPGK SGGGGSG GGG
SGGGGSQ VQLVQ SGAEVKK PGSSV KVSCKAS GYTFT TYYLHWV RQA
PGQCLEW MGWIY PGNVHAQ YNEKF KGRVTIT ADKST STAYMEL SSL
RSEDTAV YYCAR SWEGFPY WGQGT TVTVSSG GGGSG GGGSGGG GSG
GGGSDIQ MTQSP SSLSASV GDRVT ITCKASQ NVGIN VAWYQQK PGK
APKSLIS SASYR YSGVPSR FSGSG SGTDFTL TISSL QPEDFAT YFC
QQYDTYP FTFGC GTKVEIK

(SEQ ID NO:51)

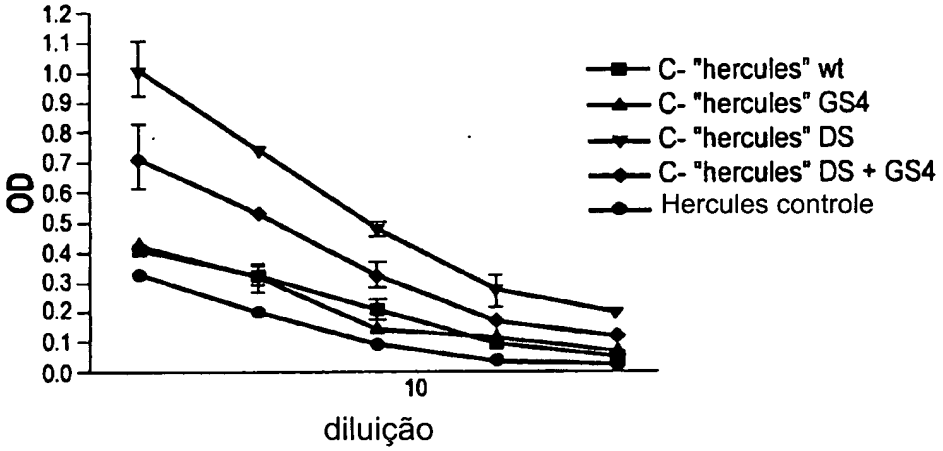
Fig. 31



1. C - C-Hercules com scFv convencional
2. C C-Hercules com $(G_4S)_4$ scFv
3. C C-Hercules com VH44:VL100 scFv
4. C - C-Hercules com VH44:VL100/$(G_4S)_4$ scFv
5. N - N-Hercules com scFv convencional
6. N - N-Hercules com $(G_4S)_4$ scFv
7. N - N-Hercules com VH44:VL100 scFv
8. N - N-Hercules com VH44:VL100/$(G_4S)_4$ scFv
9. Tetravalentcc49-A(40ng)

Fig. 32

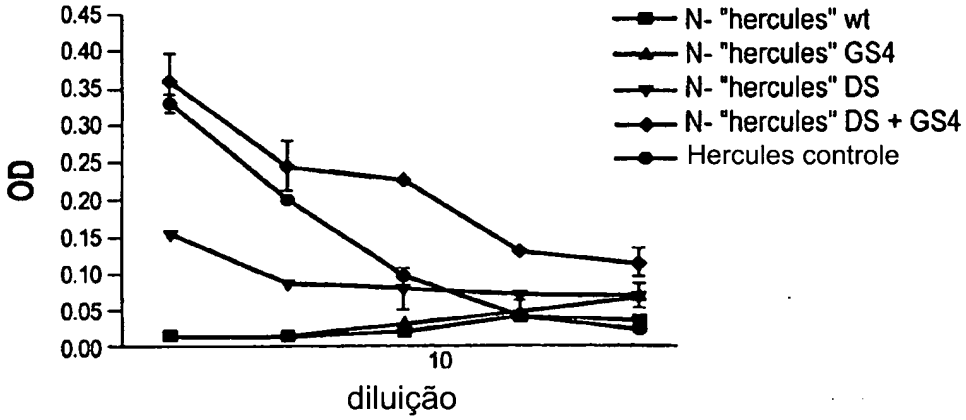
Atividade de Ligação de LTBR-Ig de constructos "Hercules" de C-ter



Lâmina revestida com 2ug/ml LTBR-Ig
2°-Ab anti-hu Kappa-HRP(10,000)

Fig. 33A

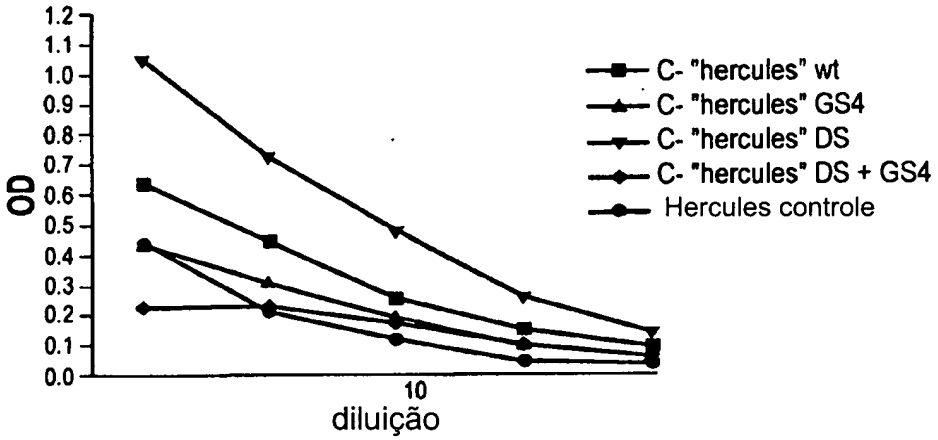
Atividade de Ligação de LTBR-Ig de constructos "Hercules" de N-ter



Lâmina revestida com 2ug/ml LTBR-Ig
2°-Ab anti-hu Kappa-HRP(10,000)

Fig. 33B

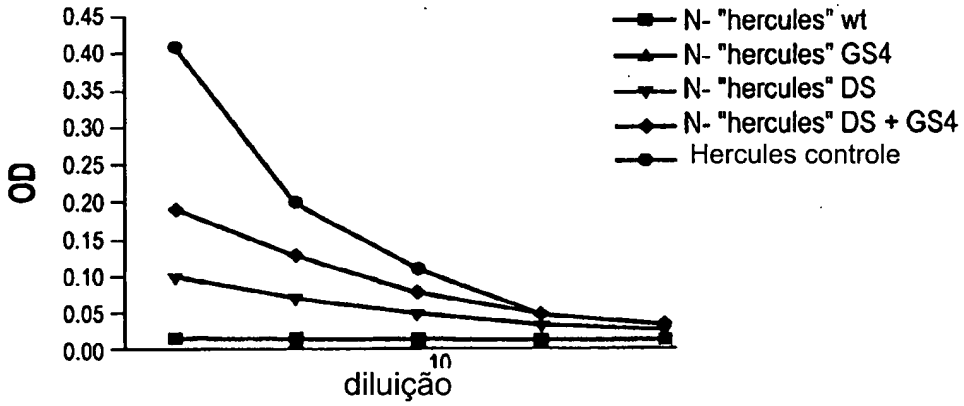
Atividade de Ligação de TRAIL-R2 de constructos "Hercules" de C-ter



Adicionado 100 ng/ml de TRAIL-R2/F c/His a lâmina Ni-NTA HisSorb 2°. –Ab anti-hu Kappa-HRP (10.000)

Fig. 34A

Atividade de Ligação de TRAIL-R2 de constructos "Hercules" de N-ter



Adicionado 100 ng/ml de TRAIL-R2/F c/His a lâmina Ni-NTA HisSorb 2°. –Ab anti-hu Kappa-HRP (10.000)

Fig. 34B

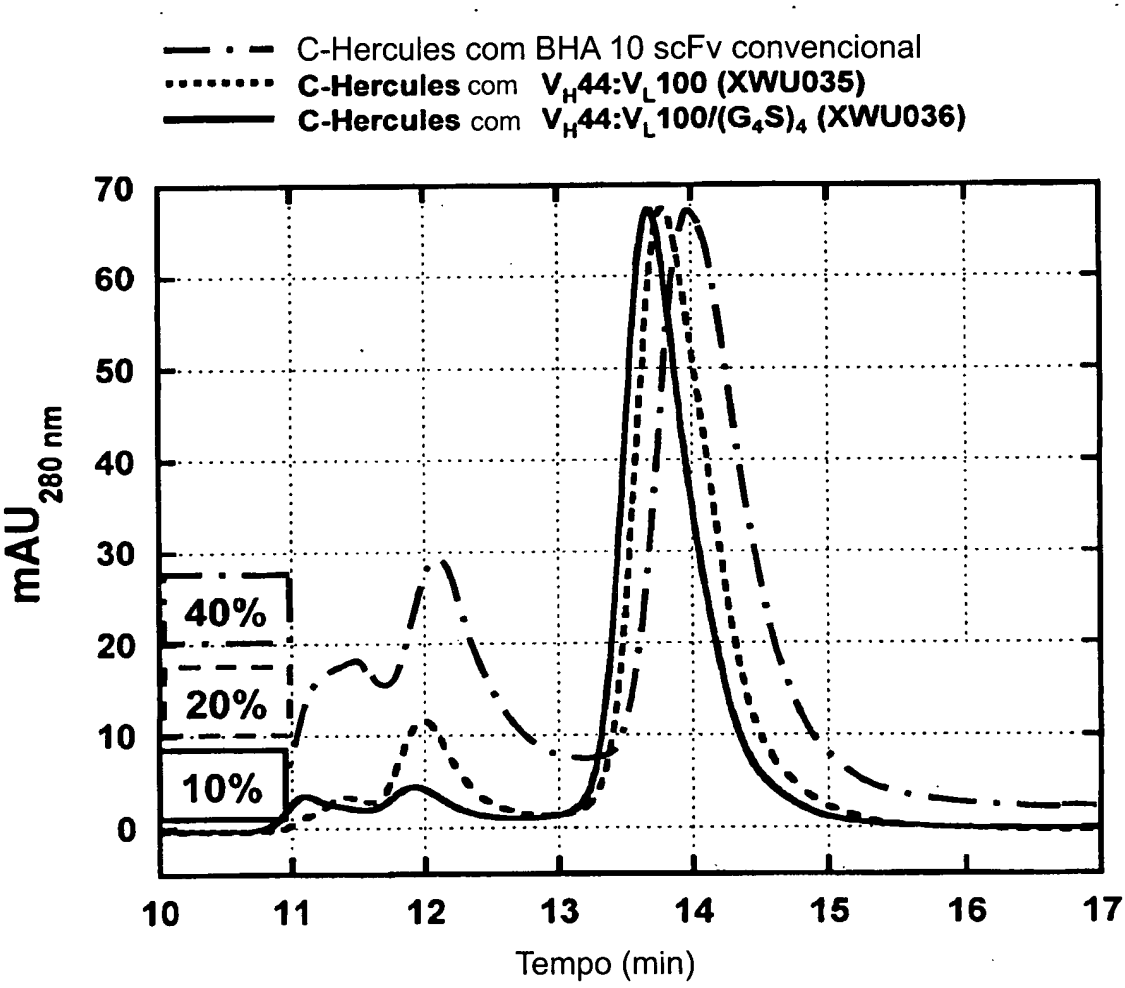
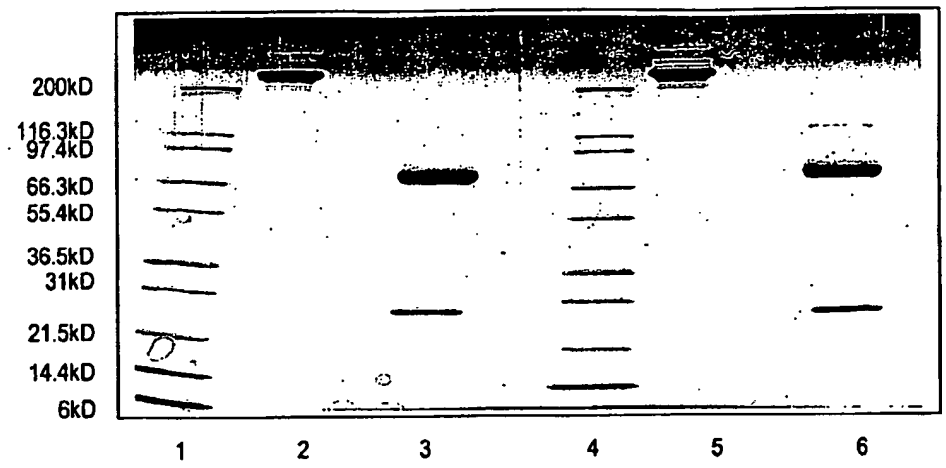
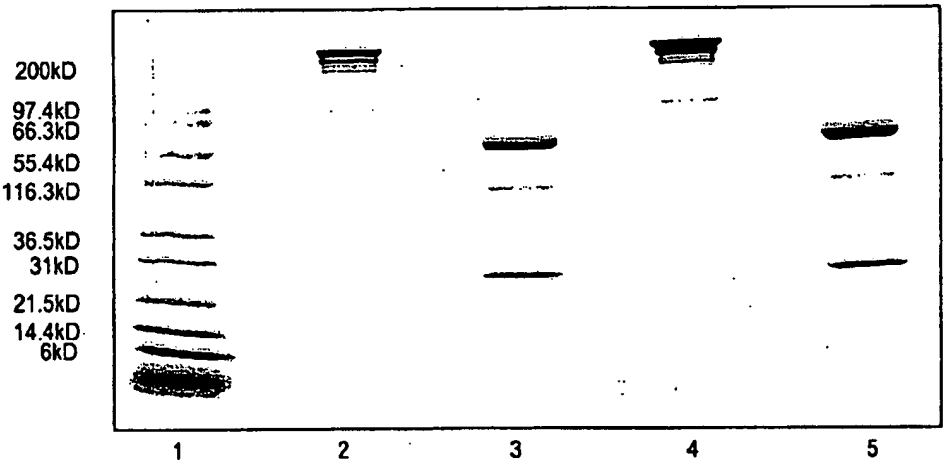

C-Hercules com BHA 10 scFv convencional
C-Hercules com $V_H44:V_L100$ (XWU035)
C-Hercules com $V_H44:V_L100/(G_4S)_4$ (XWU036)
mAU 280 nm
70
60
50
40
30
20
10
0
40%
20%
10%
10
11
12
13
14
15
16
17
Tempo (min)

Fig. 35



1. Mark 12 Marcadores
2. N- Hercules VH44:VL100 não reduzido
3. N- Hercules VH44:VL100 reduzido
4. Mark 12 Marcadores
5. N- Hercules VH44:VL100/$(G_4/S)_4$ não reduzido
6. N- Hercules VH44:VL100/$(G_4/S)_4$ reduzido

Fig. 36A



1. Mark 12 Marcadores
2. C-Hercules VH44:VL100 não reduzido
3. C-Hercules VH44:VL100 reduzido
4. C-Hercules VH44:VL100/$(G_4/S)_4$ não reduzido
5. C-Hercules VH44:VL100/$(G_4/S)_4$ reduzido

Fig. 36B

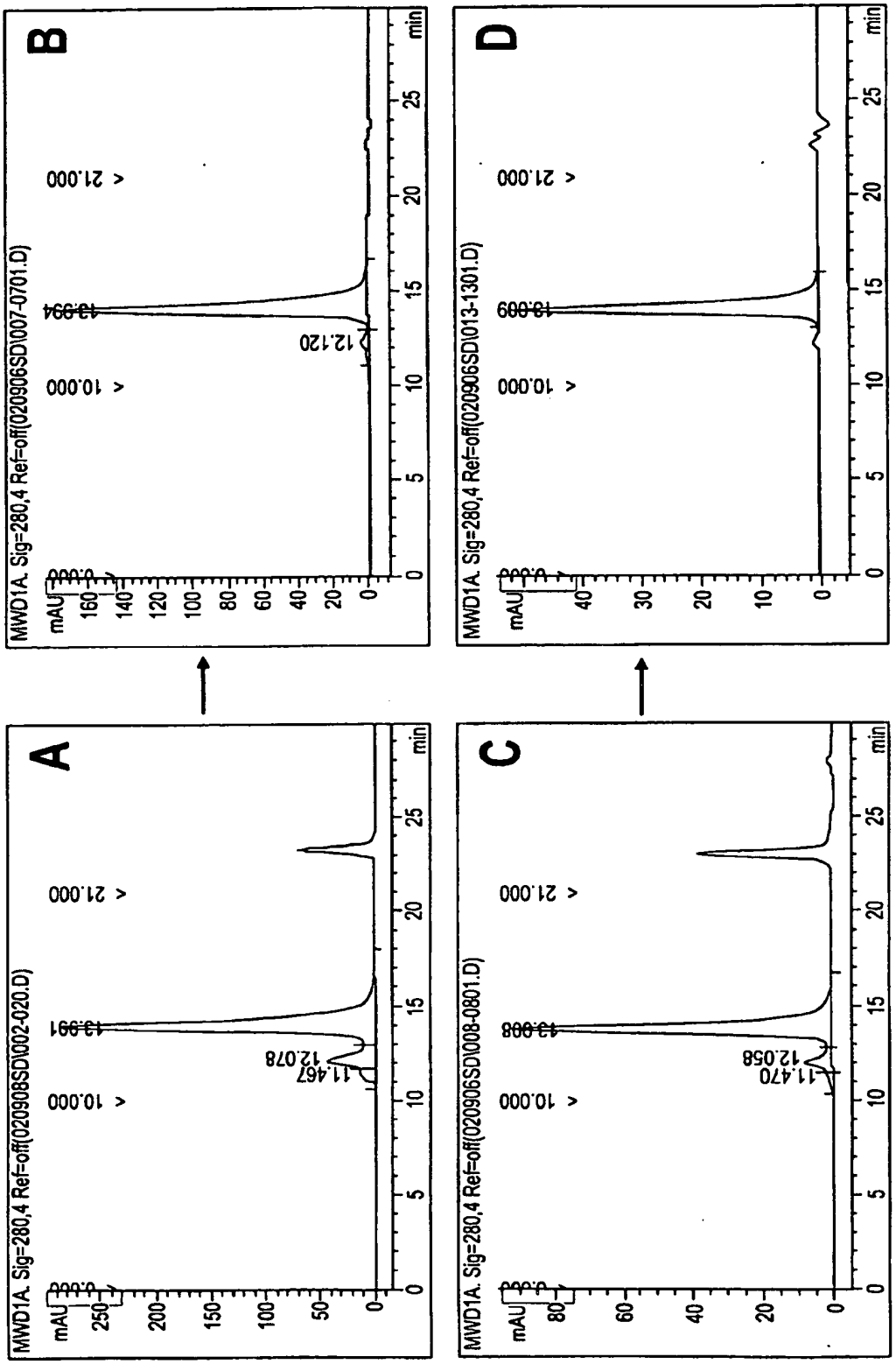

A
B
C
D
MWD1A. Sig=280,4 Ref=off(020908SD\002-020.D)
MWD1A. Sig=280,4 Ref=off(020906SD\007-0701.D)
MWD1A. Sig=280,4 Ref=off(020906SD\008-0801.D)
MWD1A. Sig=280,4 Ref=off(020906SD\013-1301.D)
mAU
min
< 10.000
< 21.000
11.467
12.078
13.991
12.120
13.994
11.470
12.058
13.988
13.989

Fig. 37

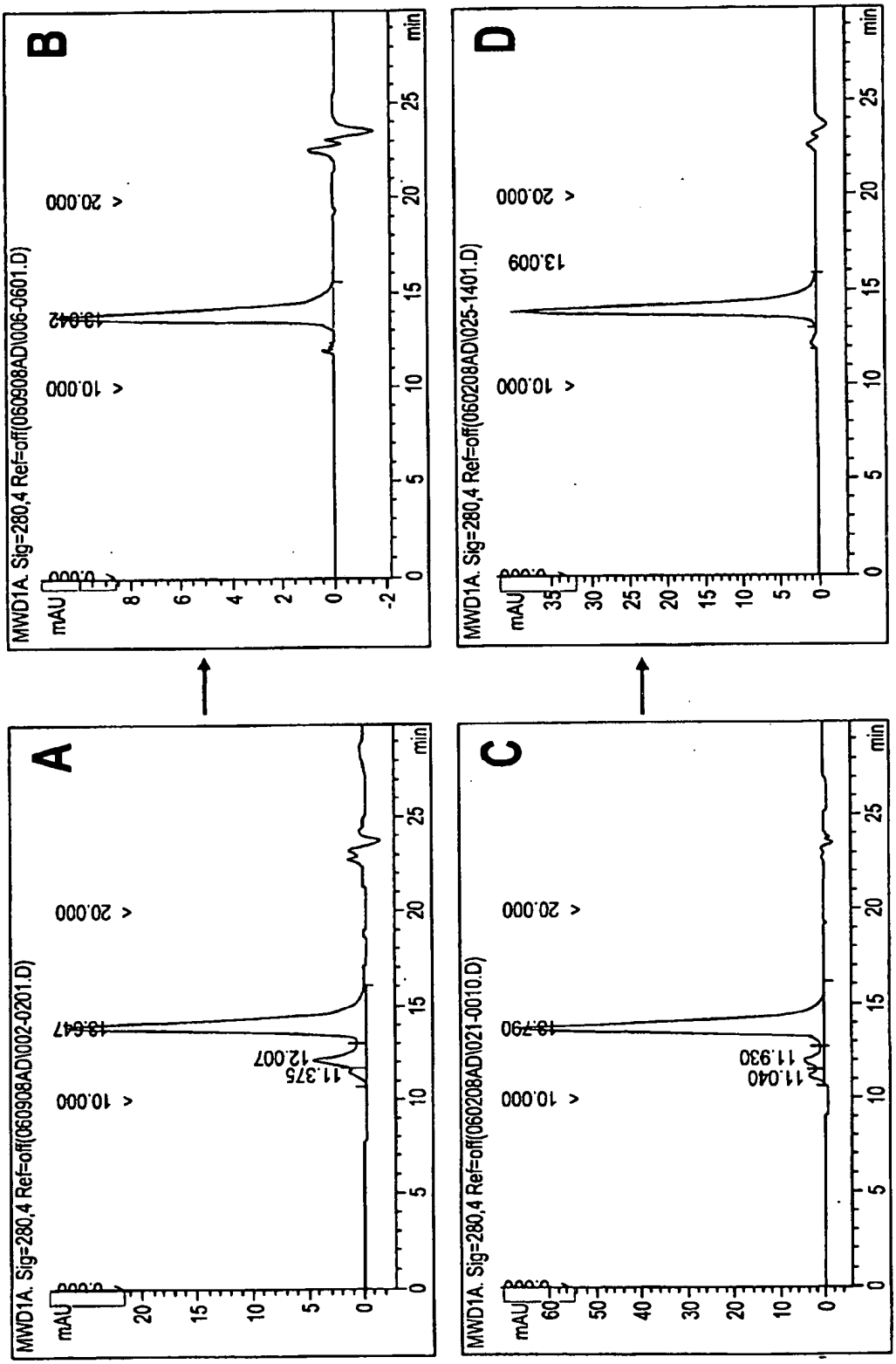

A
B
C
D
MWD1A. Sig=280,4 Ref=off(060908AD\002-0201.D)
MWD1A. Sig=280,4 Ref=off(060908AD\006-0601.D)
MWD1A. Sig=280,4 Ref=off(060208AD\021-0010.D)
MWD1A. Sig=280,4 Ref=off(060208AD\025-1401.D)
mAU
min
< 10.000
< 20.000
13.647
12.007
11.375
18.042
13.009
18.790
11.930
11.040

Fig. 38

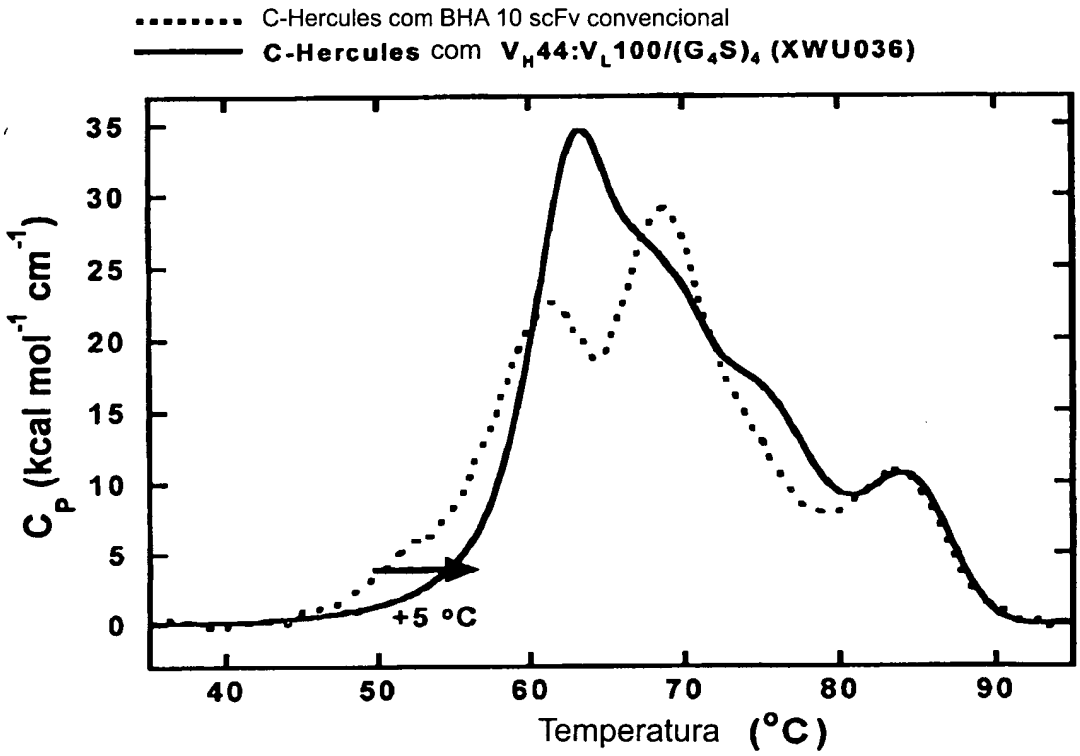

C-Hercules com BHA 10 scFv convencional
C-Hercules com $V_H44:V_L100/(G_4S)_4$ (XWU036)
C_p (kcal mol^{-1} cm^{-1})
35
30
25
20
15
10
5
0
+5 °C
40
50
60
70
80
90
Temperatura (°C)

Fig. 39

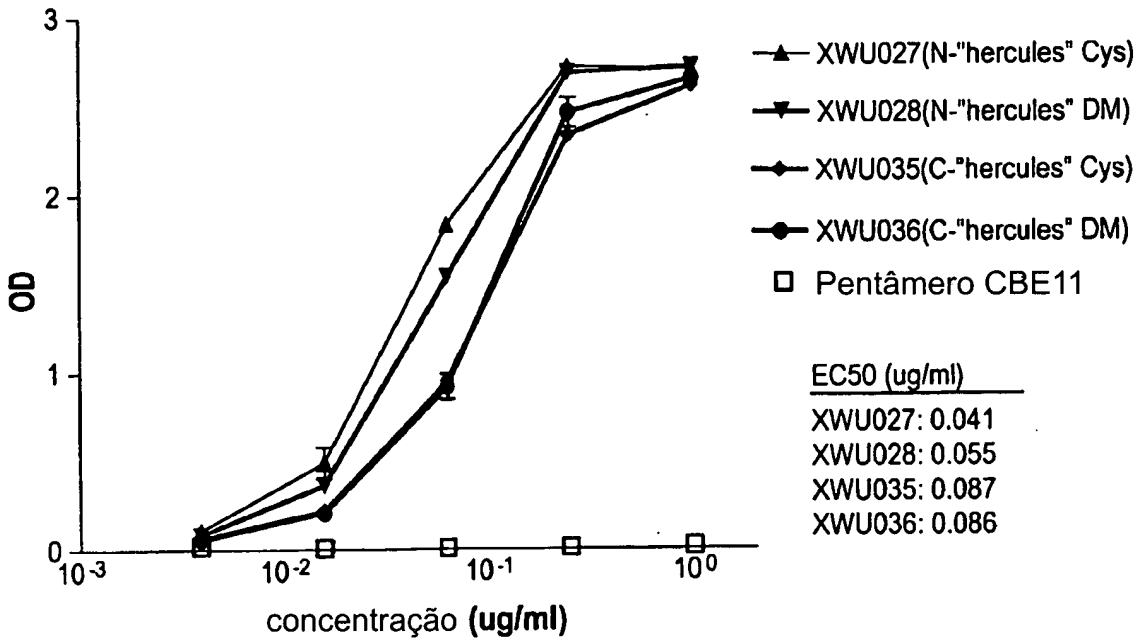

3
2
1
0
OD
10-3
10-2
10-1
100
concentração (ug/ml)
XWU027(N-"hercules" Cys)
XWU028(N-"hercules" DM)
XWU035(C-"hercules" Cys)
XWU036(C-"hercules" DM)
Pentâmero CBE11
EC50 (ug/ml)
XWU027: 0.041
XWU028: 0.055
XWU035: 0.087
XWU036: 0.086

Fig. 40

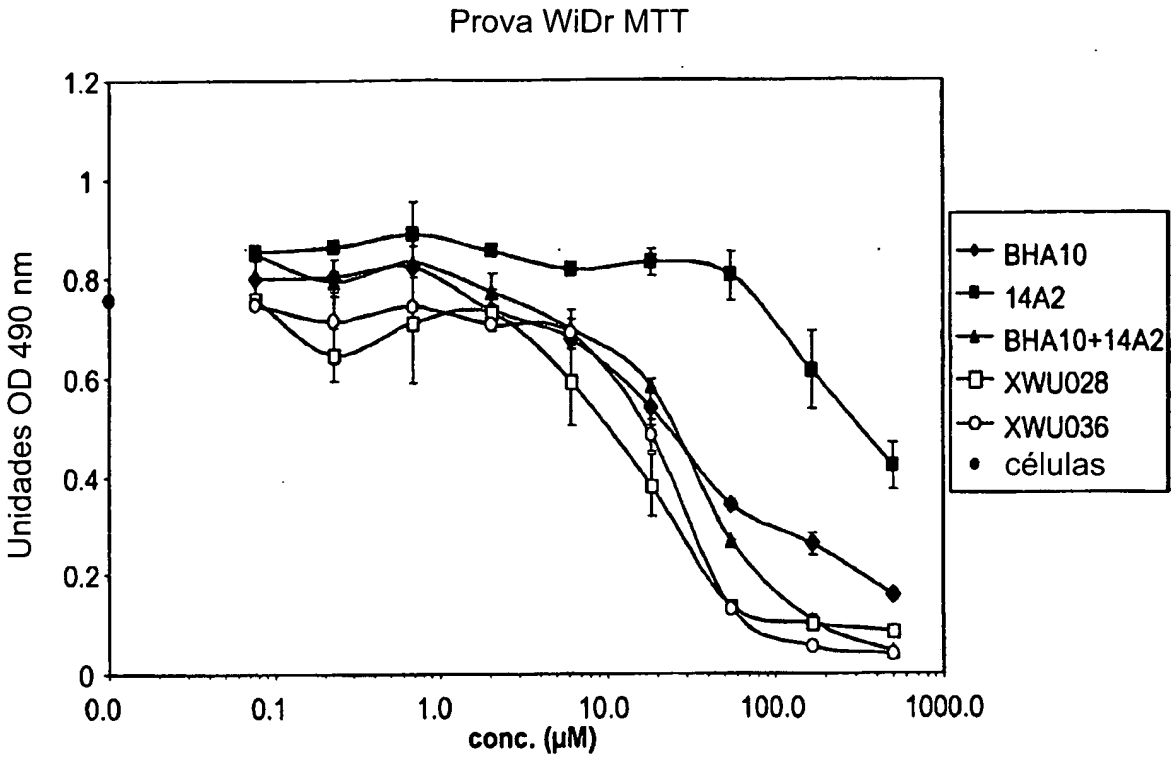


Fig. 41A

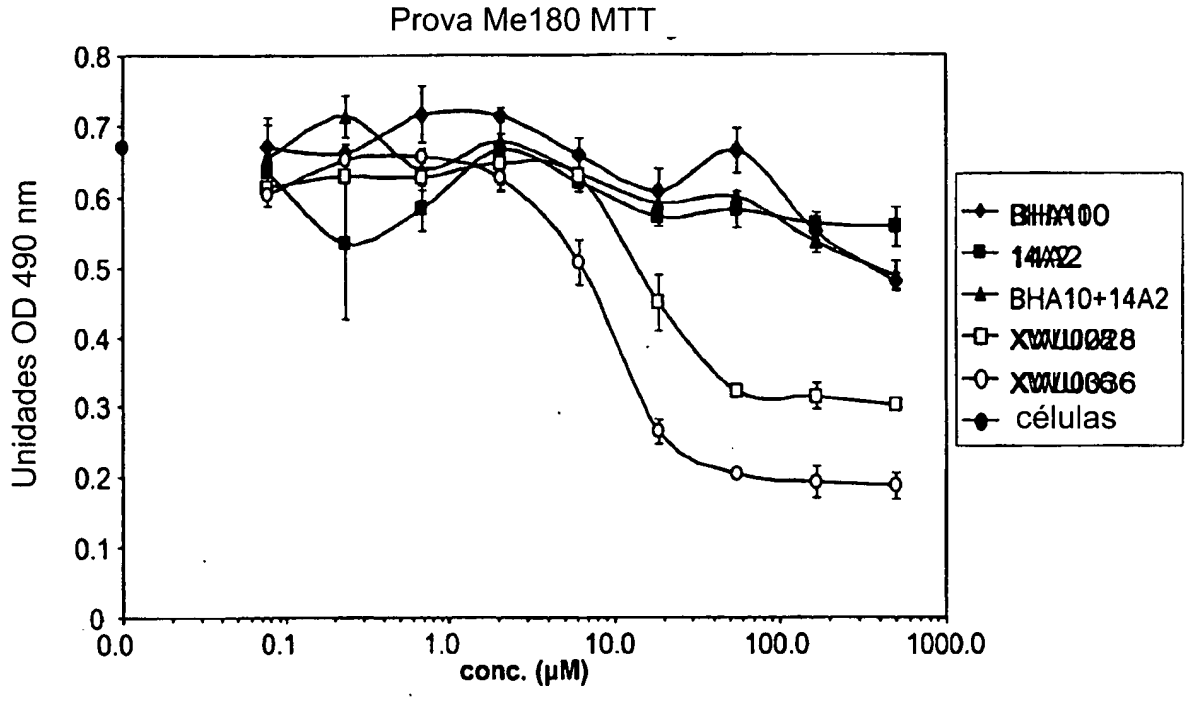


Fig. 41B

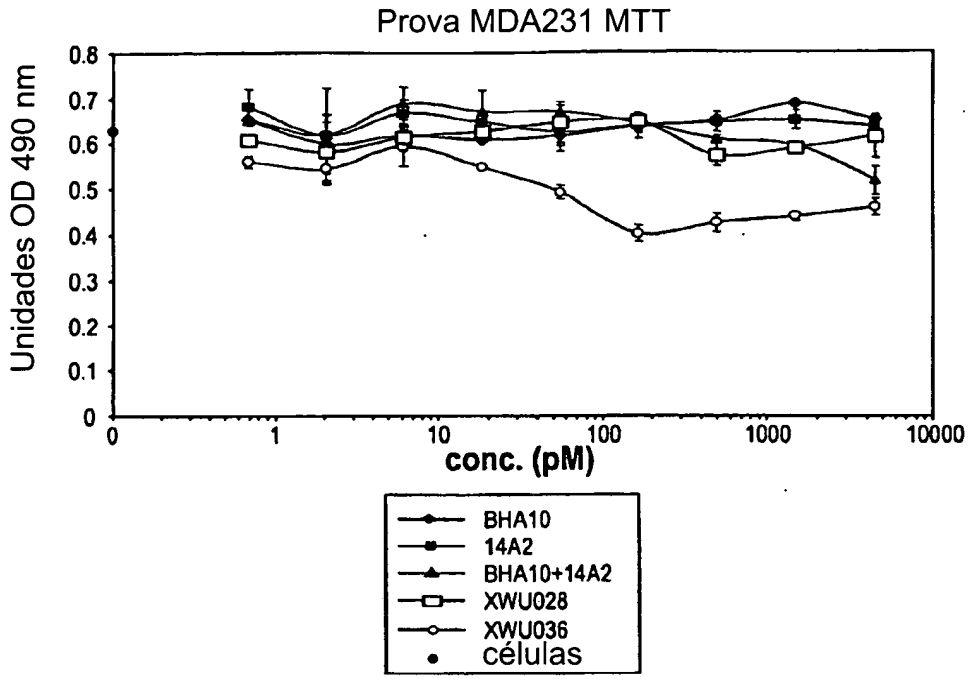


Fig. 41C

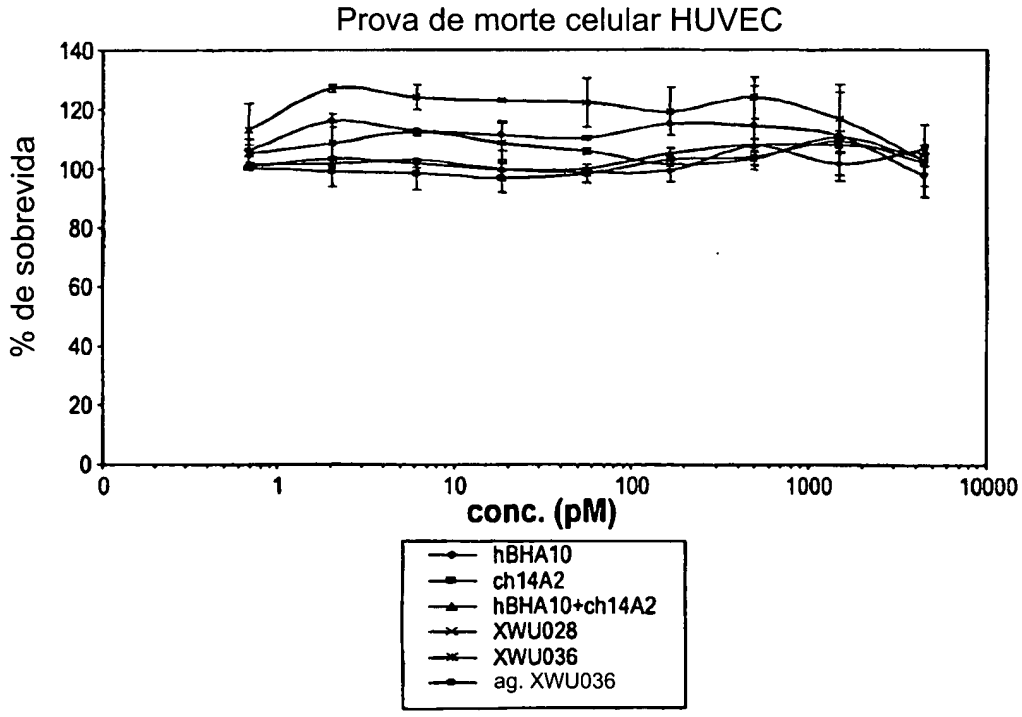
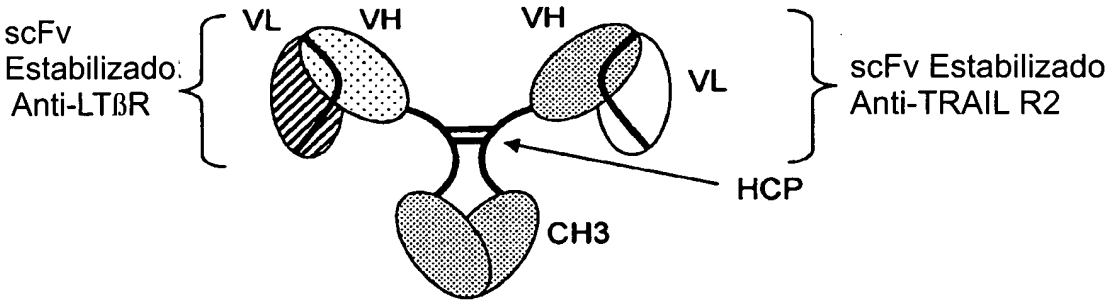


Fig. 41D

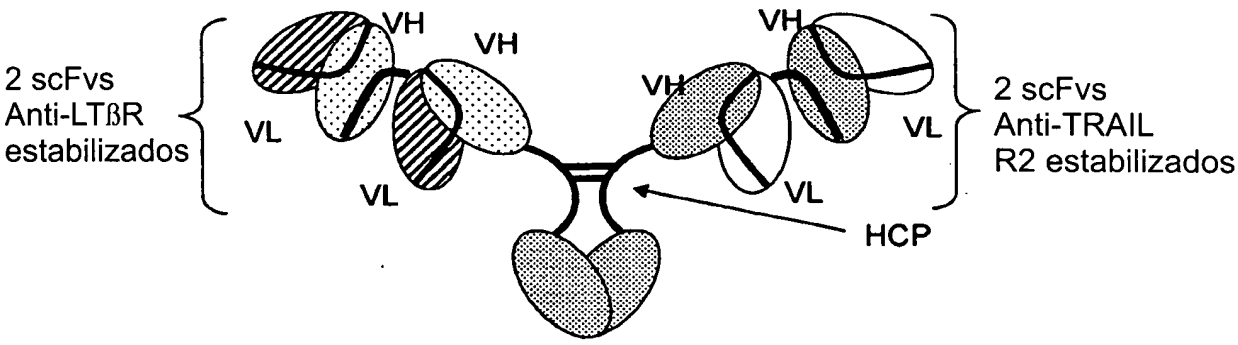


V = Variável
C = Constante
L = Leve
H = Pesado
HCP = Peptídeo de Ligação de Articulação

= encadeador (G_4S_4)

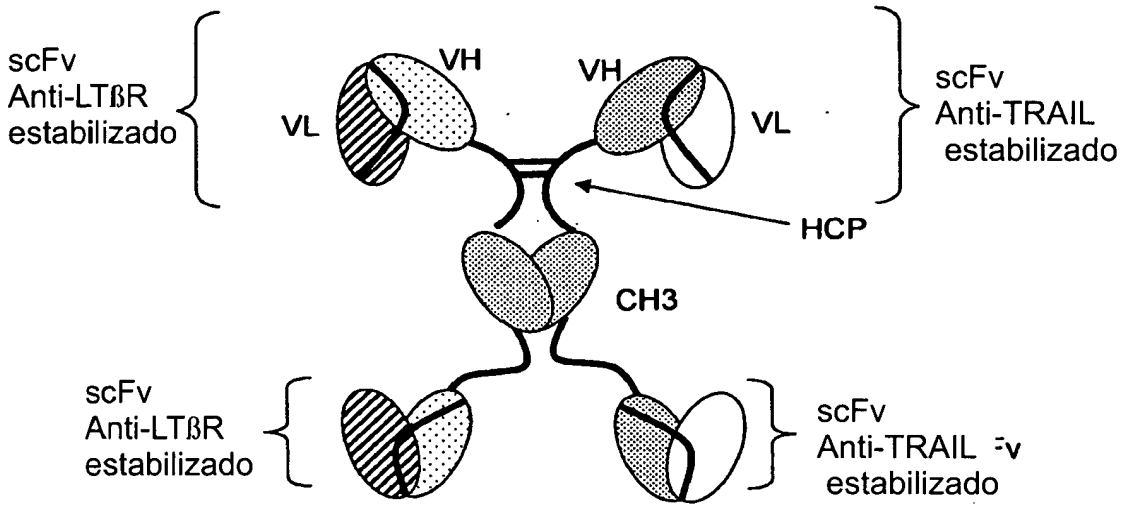
= encadeador flexível $(G_4S)_4G_3AS$

Fig. 42



Minibody Tetravalente Biespecífico 2 sc(Fv)2 Estabilizado
(Minibody Tetravalente N-scFv Biespecífico estabilizado)

Fig. 43



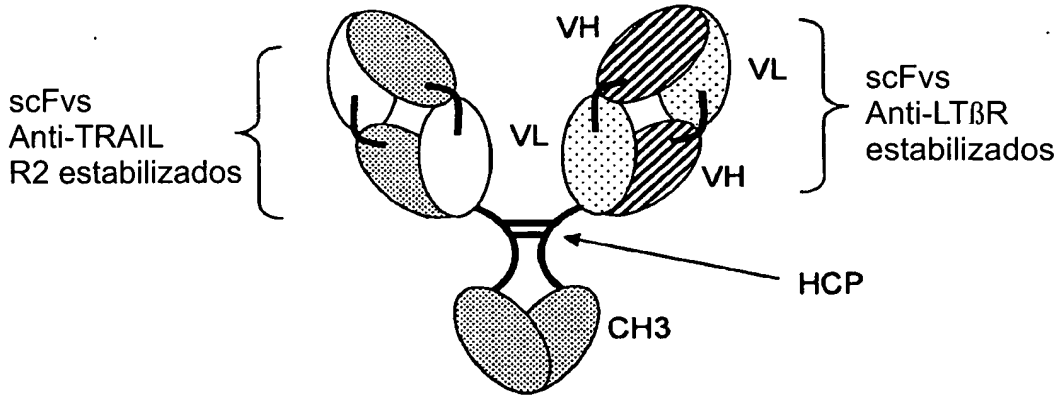
Minibody Tetravalente Biespecífico 2 sc(Fv)2 Estabilizado
(Minibody tetravalente C-scFv Biespecífico estabilizado)

V = Variável
C = Constante
L = Leve
H =Pesado
HCP = Peptídeo de Ligação de Articulação

= encadeador flexível (G_4S_4)

= encadeador flexível $(G_4S)_4G_3AS$

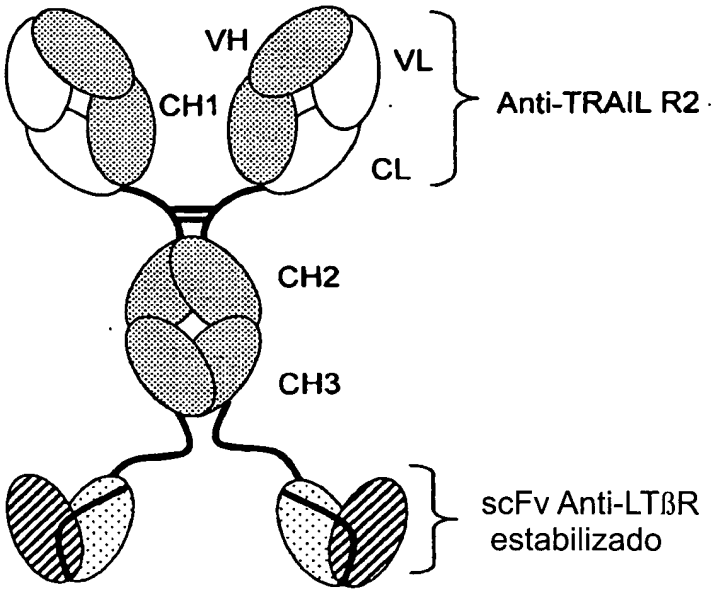
Fig. 44



Diabody Biespecífico
Estabilizado

V = Variável
C = Constante
L = Leve
H = Pesado
HCP = Peptídeo de Ligação de Articulação
J = encadeador (Gly_4Ser_4)

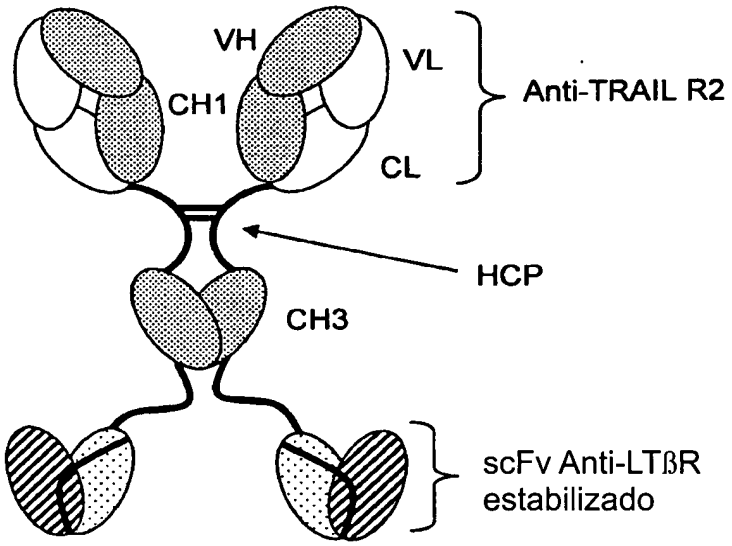
Fig. 45



Anticorpo Tetravalente sc(Fv)2 Biespecífico Estabilizado
(Anticorpo Tetravalente C-scFv Biespecífico Estabilizado)

V = Variável
C = Constante
L = Leve
H = Pesado
HCP = Peptídeo de Ligação de Articulação
= encadeador (Gly_4Ser_4)

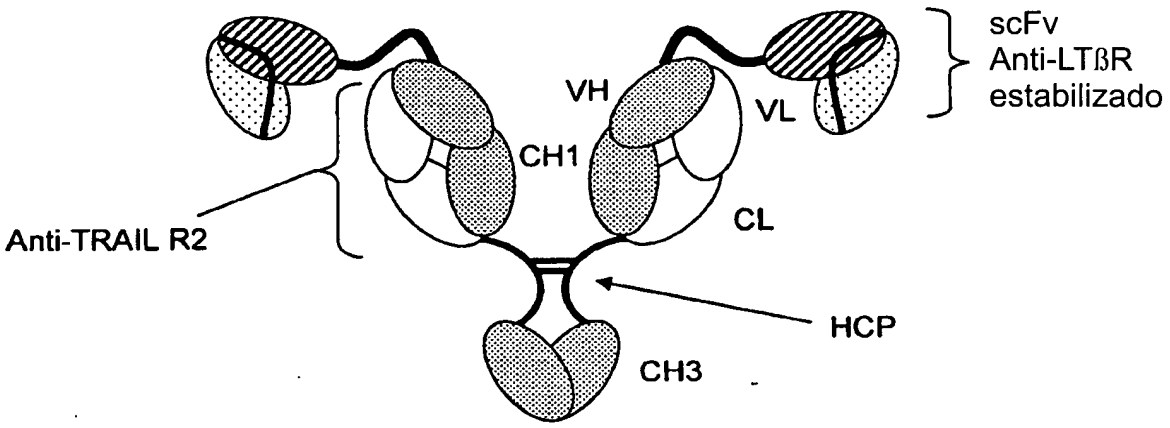
Fig. 46



Anticorpo Biespecífico Com Domínio CH2 Eliminado Tetravalente sc(Fv)2 Estabilizado
(Anticorpo Biespecífico Com Domínio CH2 Eliminado Tetravalente C-scFv Estabilizado)
anticorpo Com Domínio CH2 Eliminado Tetravalente C-scFv biespecífico)

V = Variável
C = Constante
L = Leve
H = Pesado
HCP = Peptídeo de Ligação de Articulação
= encadeador (Gly_4Ser_4)

Fig. 47



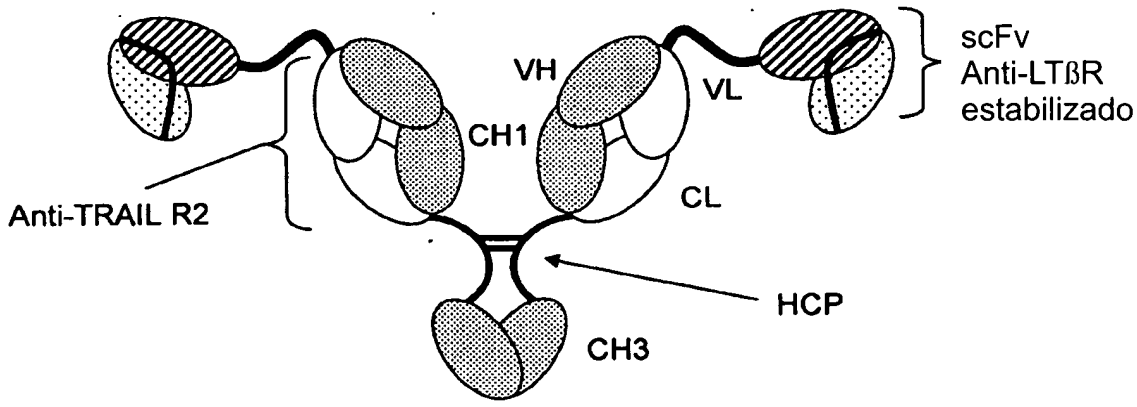
Anticorpo Biespecífico Com Domínio Ch2
Eliminado Tetravalente N_H-scFv Estabilizado

V = Variável
C = Constante
L = Leve
H = Pesado
HCP = Peptídeo de Ligação de Articulação

= encadeador $(G_4S)_4$

= encadeador flexível $(G_4S)_4G_3AS$

Fig. 48



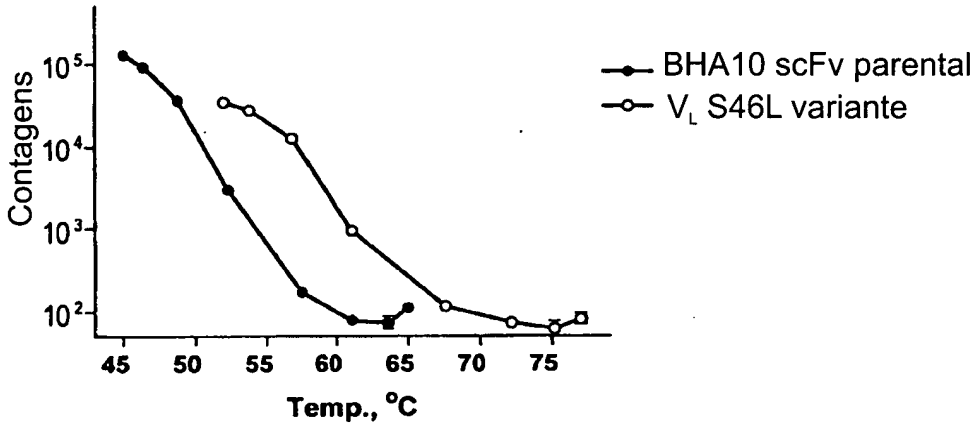
V = Variável
C = Constante
L = Leve
H = Pesado
HCP = Peptídeo de Ligação de Articulação

= encadeador $(G_4S)_4$

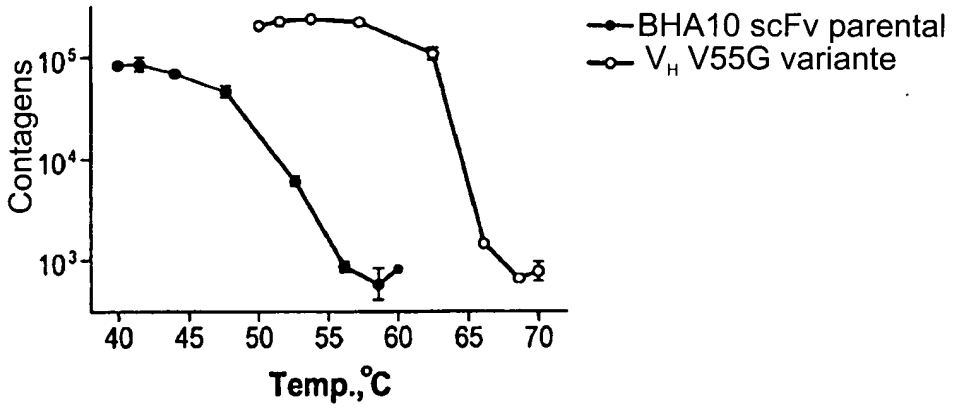
= encadeador flexível $(G_4S)_4G_3AS$

Fig. 49

A.


Contagens
10^5
10^4
10^3
10^2
45 50 55 60 65 70 75
Temp., °C
BHA10 scFv parental
V_L S46L variante

B.


Contagens
10^5
10^4
10^3
40 45 50 55 60 65 70
Temp.,°C
BHA10 scFv parental
V_H V55G variante

C.

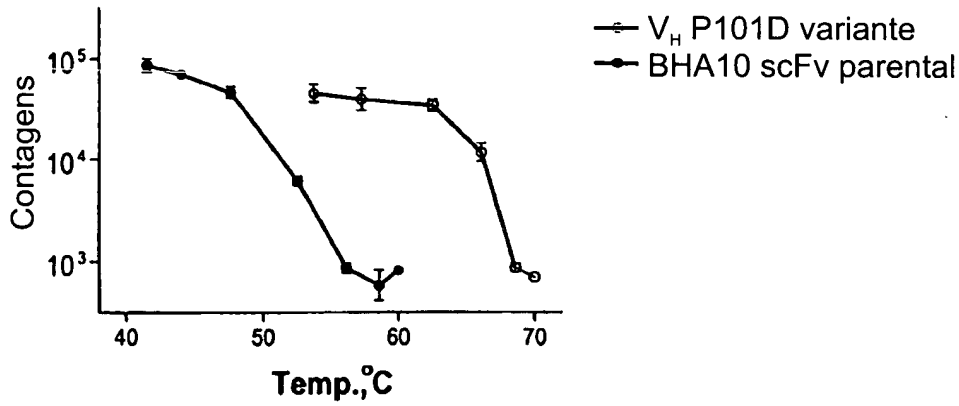

Contagens
10^5
10^4
10^3
40 50 60 70
Temp.,°C
V_H P101D variante
BHA10 scFv parental

Fig. 50

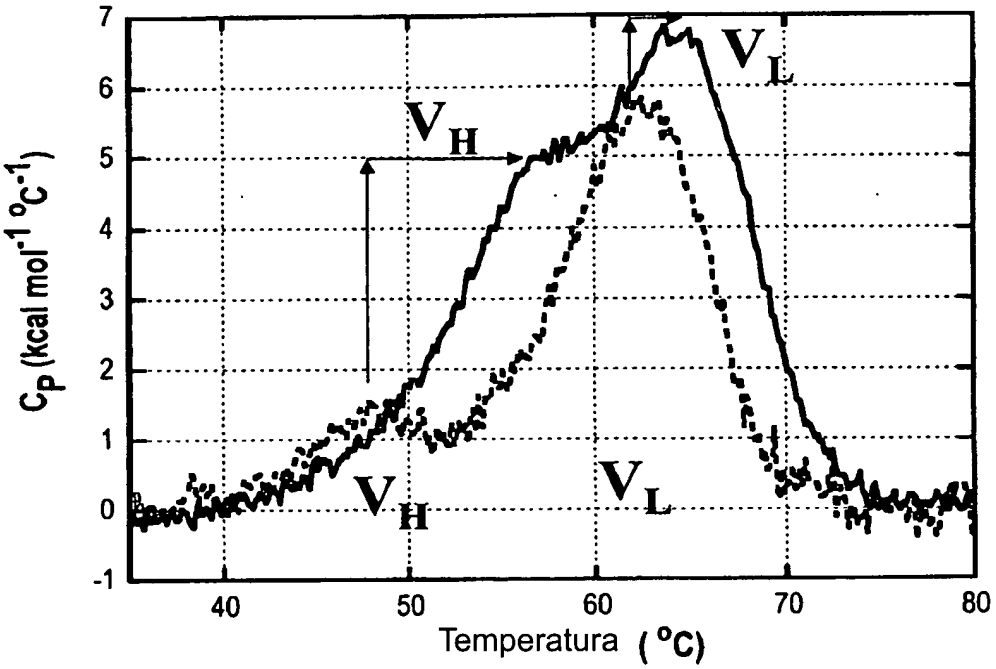


Fig. 51A

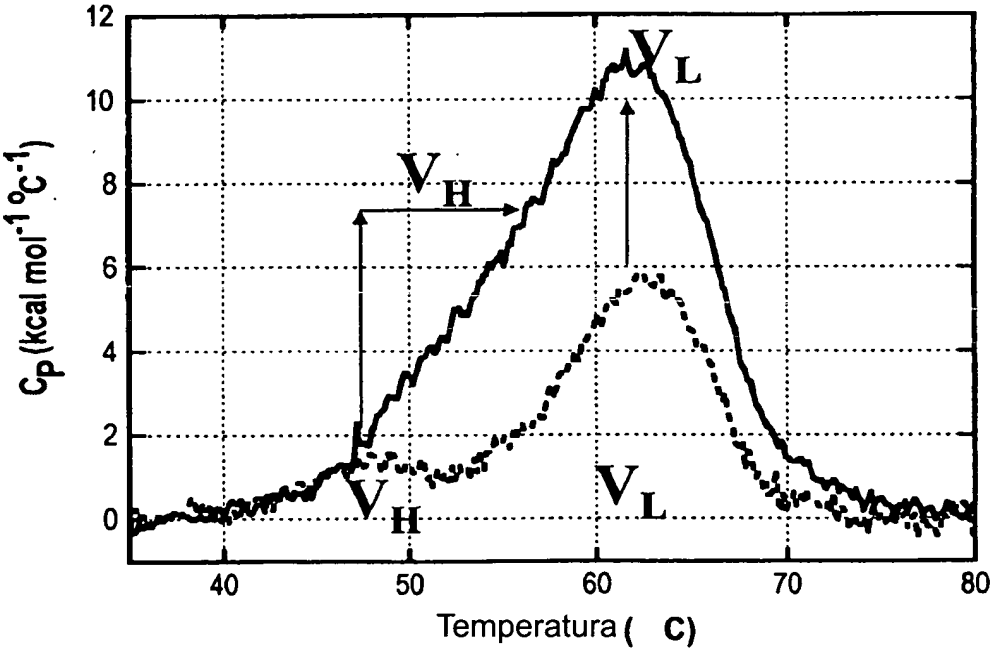


Fig. 51B

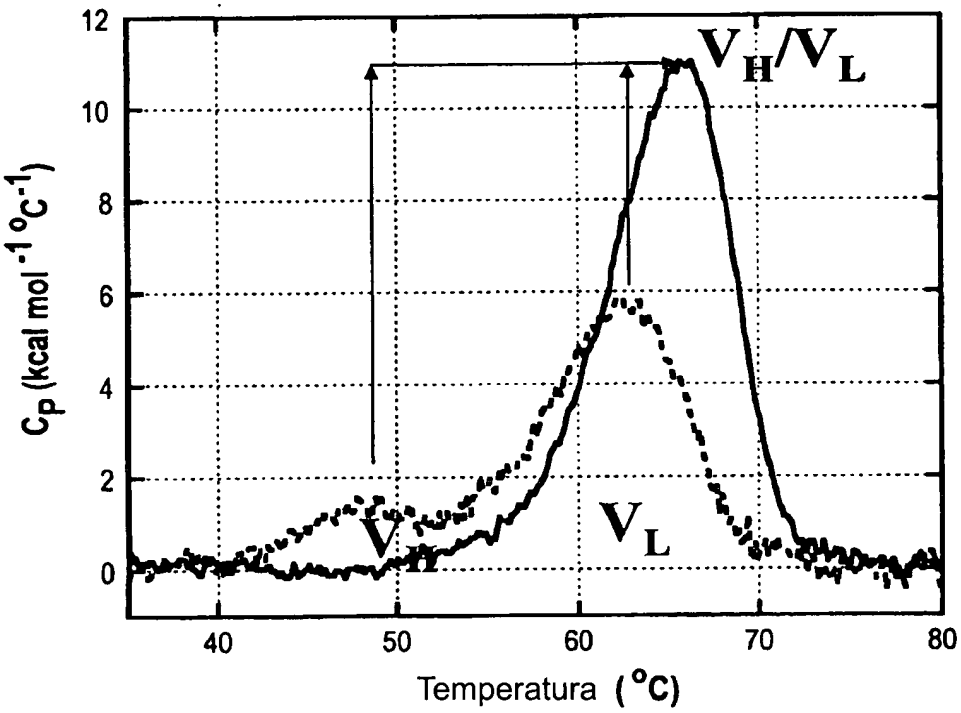

12
10
8
6
4
2
0
40
50
60
70
80
C_p (kcal mol^{-1} °C^{-1})
Temperatura (°C)
V_H/V_L
V_H
V_L

Fig. 52

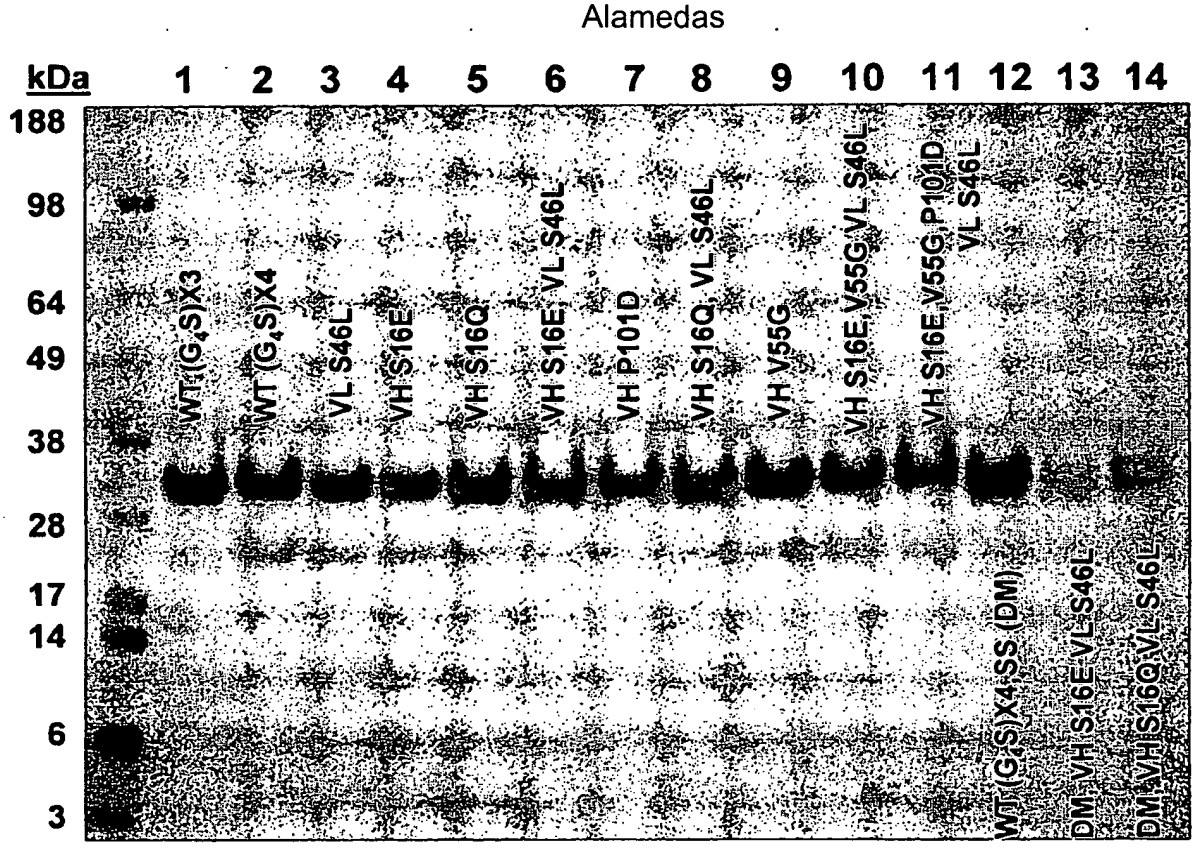

Alamedas
kDa
1 2 3 4 5 6 7 8 9 10 11 12 13 14
188
98
64
49
38
28
17
14
6
3
WT(G4S)X3
WT(G4S)X4
VL S46L
VH S16E
VH S16Q
VH S16E, VL S46L
VH P101D
VH S16Q, VL S46L
VH V55G
VH S16E,V55G,VL S46L
VH S16E,V55G,P101D
VL S46L
WT(G4S)X4SS (DM)
DM VH S16E VL S46L
DM VH S16Q VL S46L

Fig. 53

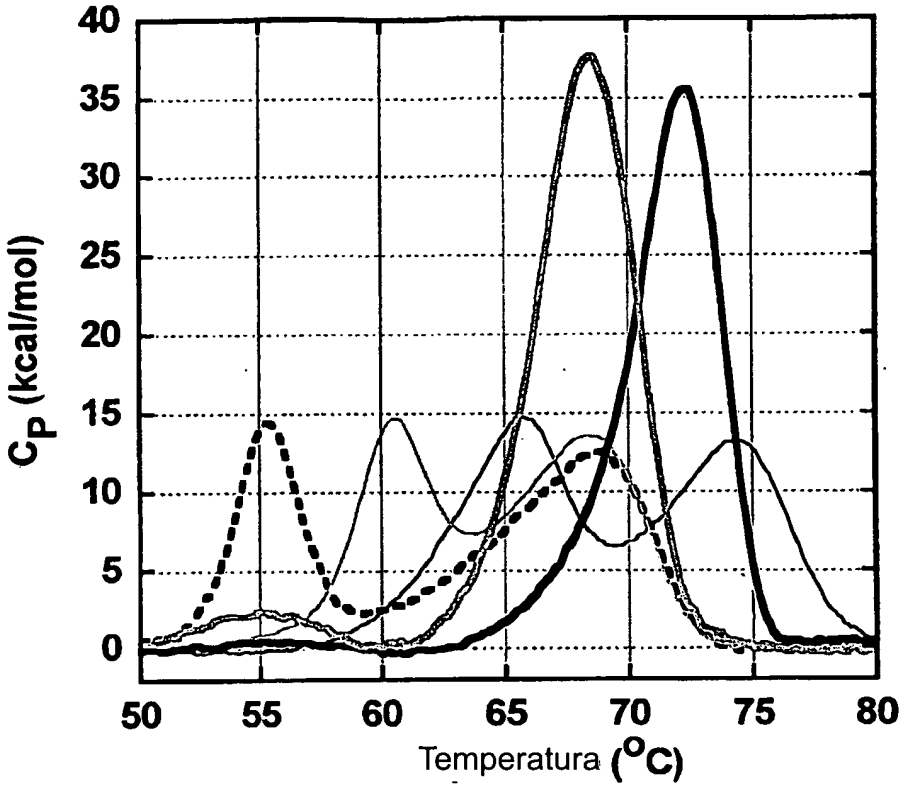


Fig. 54A

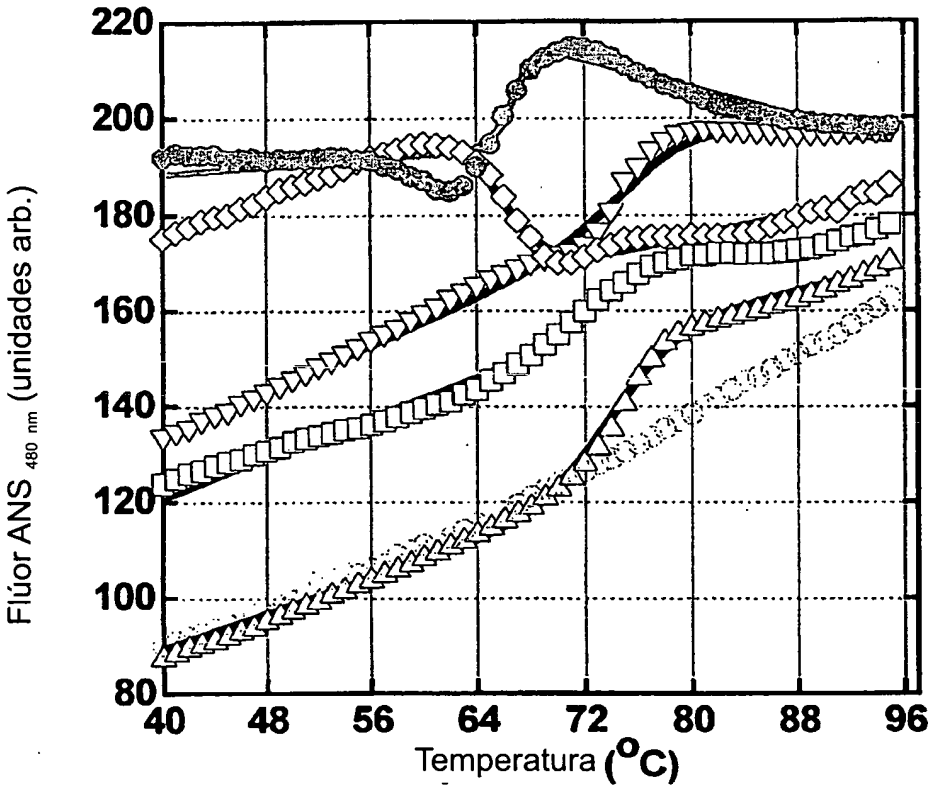


Fig. 54B

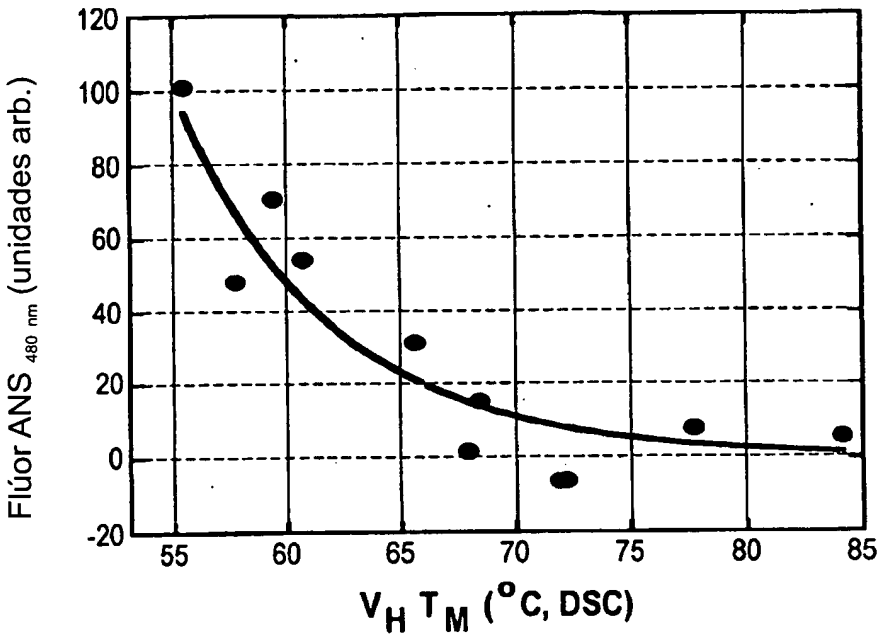


Fig. 55A

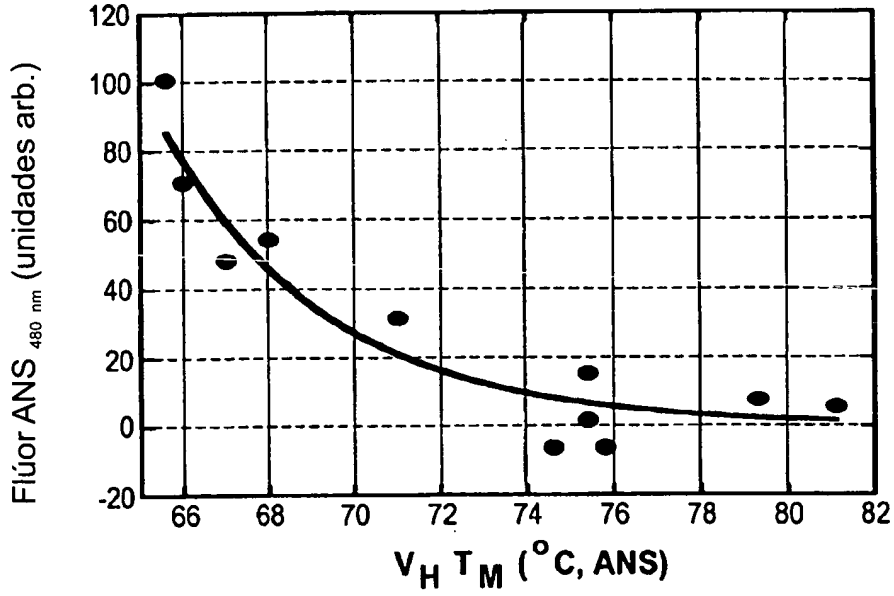


Fig. 55B

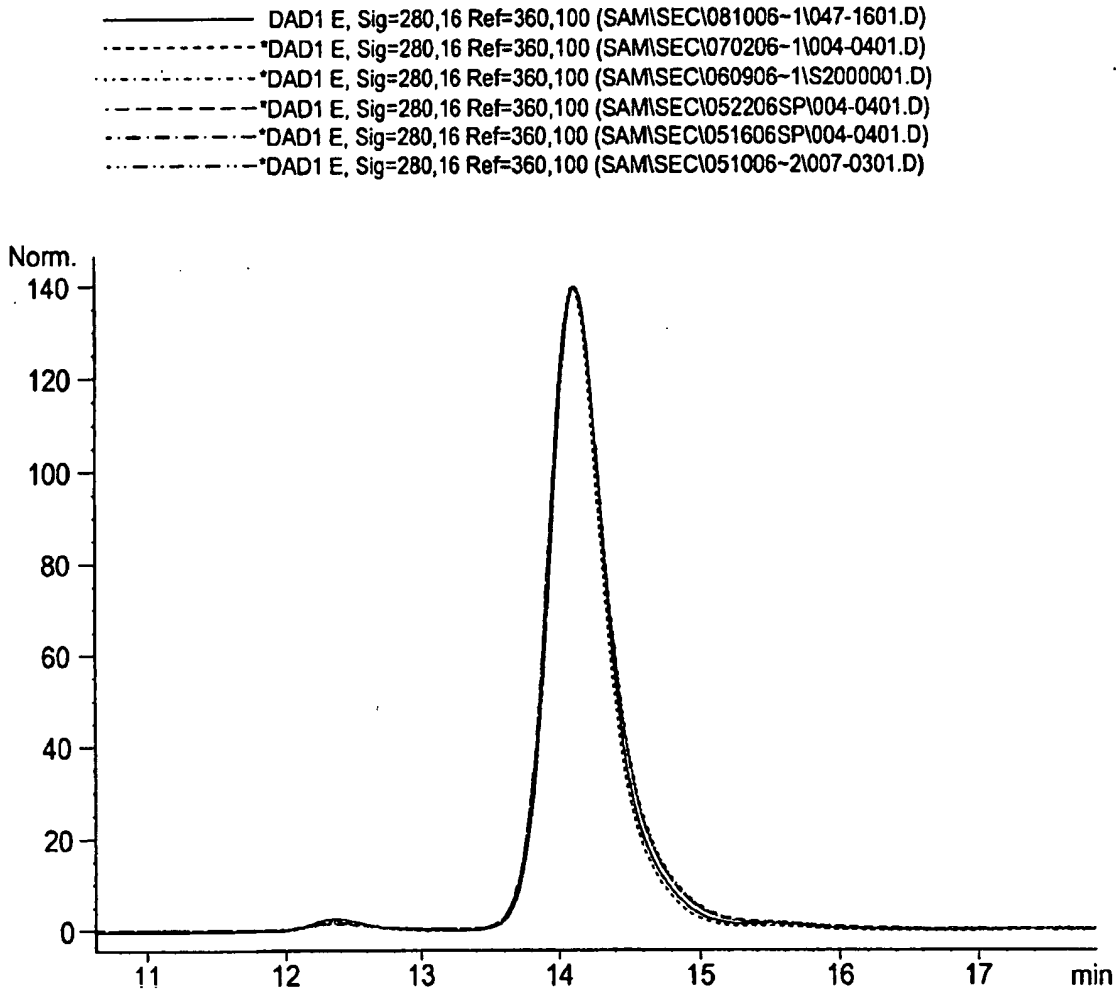

DAD1 E, Sig=280,16 Ref=360,100 (SAM\SEC\081006~1\047-1601.D)
*DAD1 E, Sig=280,16 Ref=360,100 (SAM\SEC\070206~1\004-0401.D)
*DAD1 E, Sig=280,16 Ref=360,100 (SAM\SEC\060906~1\S2000001.D)
*DAD1 E, Sig=280,16 Ref=360,100 (SAM\SEC\052206SP\004-0401.D)
*DAD1 E, Sig=280,16 Ref=360,100 (SAM\SEC\051606SP\004-0401.D)
*DAD1 E, Sig=280,16 Ref=360,100 (SAM\SEC\051006~2\007-0301.D)
Norm.
140
120
100
80
60
40
20
0
11
12
13
14
15
16
17
min

Fig. 56A

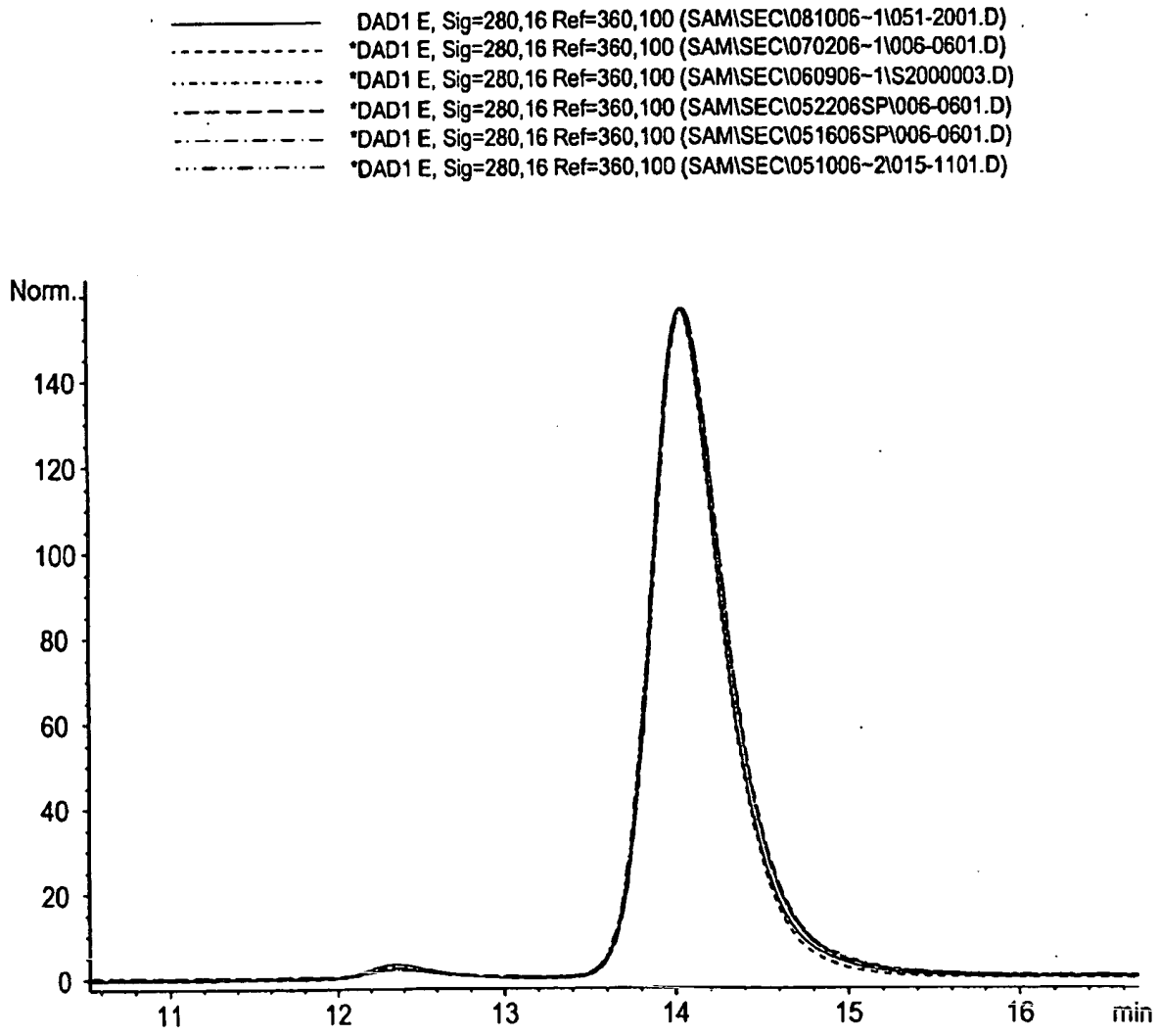

DAD1 E, Sig=280,16 Ref=360,100 (SAM\SEC\081006~1\051-2001.D)
*DAD1 E, Sig=280,16 Ref=360,100 (SAM\SEC\070206~1\006-0601.D)
*DAD1 E, Sig=280,16 Ref=360,100 (SAM\SEC\060906~1\S2000003.D)
*DAD1 E, Sig=280,16 Ref=360,100 (SAM\SEC\052206SP\006-0601.D)
*DAD1 E, Sig=280,16 Ref=360,100 (SAM\SEC\051606SP\006-0601.D)
*DAD1 E, Sig=280,16 Ref=360,100 (SAM\SEC\051006~2\015-1101.D)
Norm.
140
120
100
80
60
40
20
0
11
12
13
14
15
16
min

Fig. 56B

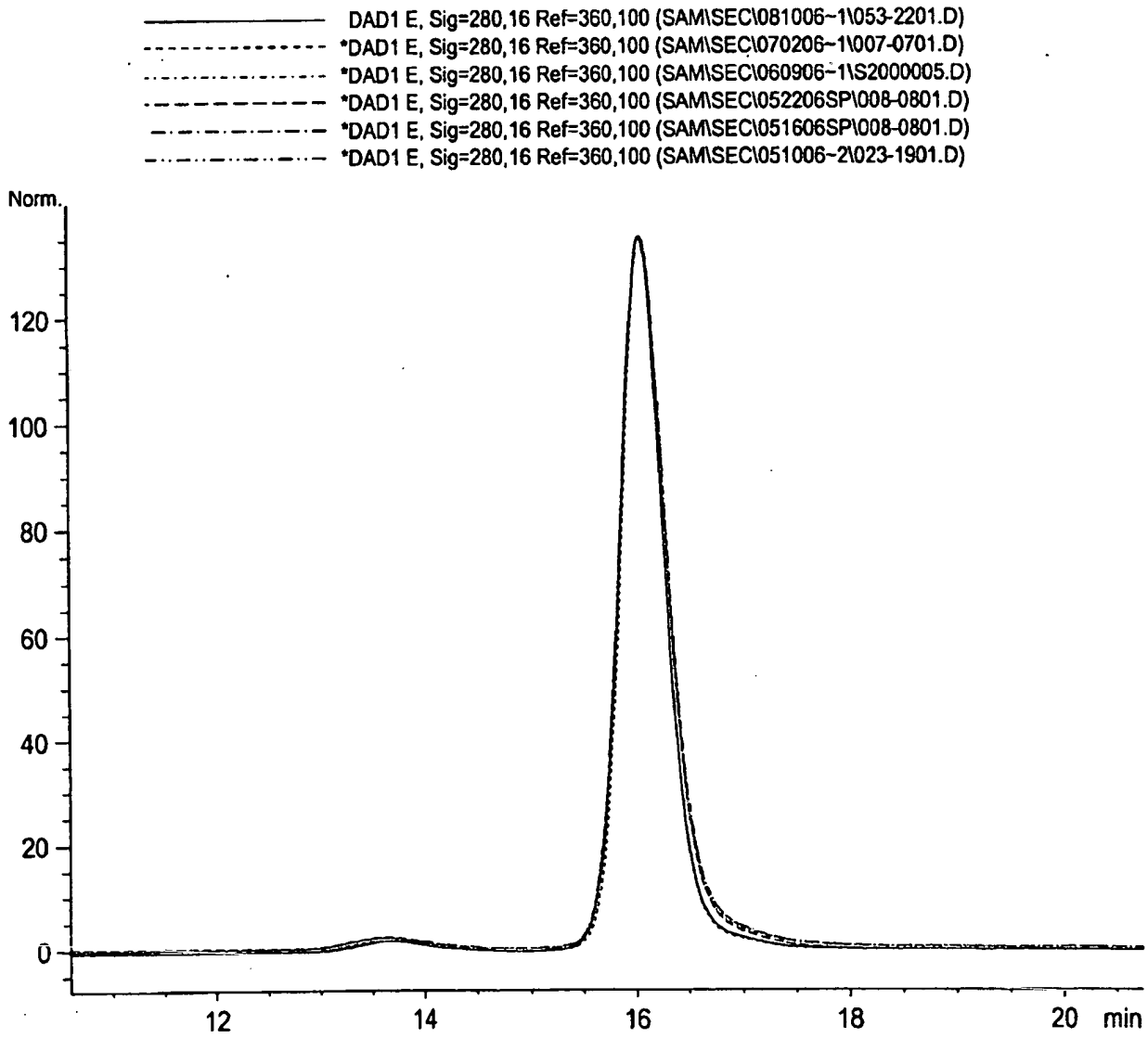

DAD1 E, Sig=280,16 Ref=360,100 (SAM\SEC\081006~1\053-2201.D)
*DAD1 E, Sig=280,16 Ref=360,100 (SAM\SEC\070206~1\007-0701.D)
*DAD1 E, Sig=280,16 Ref=360,100 (SAM\SEC\060906~1\S2000005.D)
*DAD1 E, Sig=280,16 Ref=360,100 (SAM\SEC\052206SP\008-0801.D)
*DAD1 E, Sig=280,16 Ref=360,100 (SAM\SEC\051606SP\008-0801.D)
*DAD1 E, Sig=280,16 Ref=360,100 (SAM\SEC\051006~2\023-1901.D)
Norm.
120
100
80
60
40
20
0
12
14
16
18
20
min

Fig. 56C

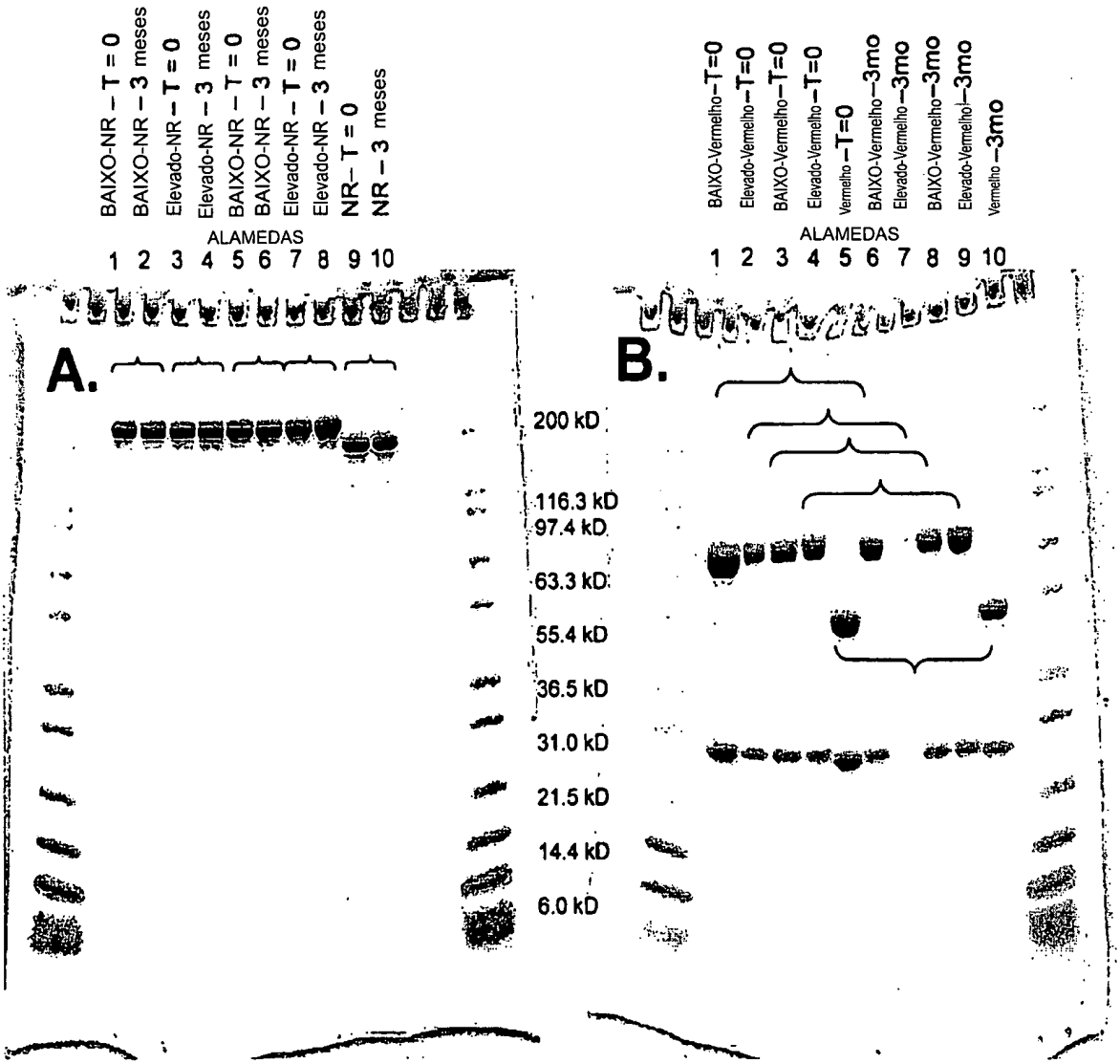

A.
BAIXO-NR – T = 0
BAIXO-NR – 3 meses
Elevado-NR – T = 0
Elevado-NR – 3 meses
BAIXO-NR – T = 0
BAIXO-NR – 3 meses
Elevado-NR – T = 0
Elevado-NR – 3 meses
NR– T = 0
NR – 3 meses
ALAMEDAS
1 2 3 4 5 6 7 8 9 10
200 kD
116.3 kD
97.4 kD
63.3 kD
55.4 kD
36.5 kD
31.0 kD
21.5 kD
14.4 kD
6.0 kD
B.
BAIXO-Vermelho–T=0
Elevado-Vermelho–T=0
BAIXO-Vermelho–T=0
Elevado-Vermelho–T=0
Vermelho –T=0
BAIXO-Vermelho–3mo
Elevado-Vermelho–3mo
BAIXO-Vermelho–3mo
Elevado-Vermelho–3mo
Vermelho –3mo
ALAMEDAS
1 2 3 4 5 6 7 8 9 10

Fig. 57

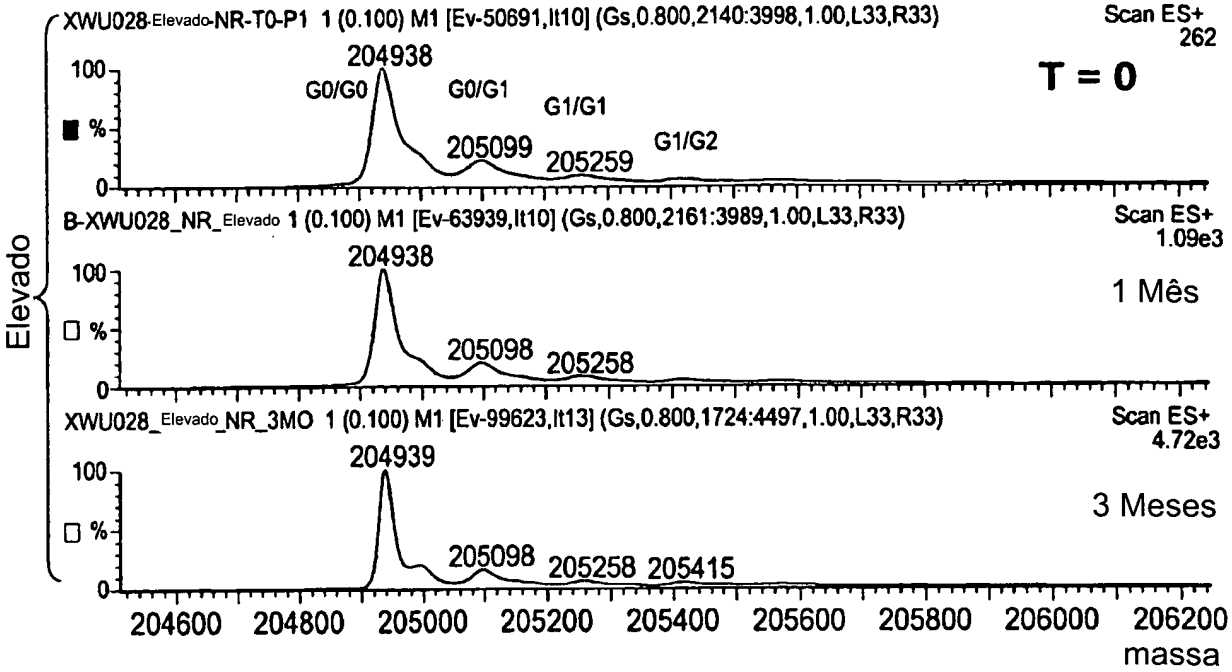


Fig. 58A

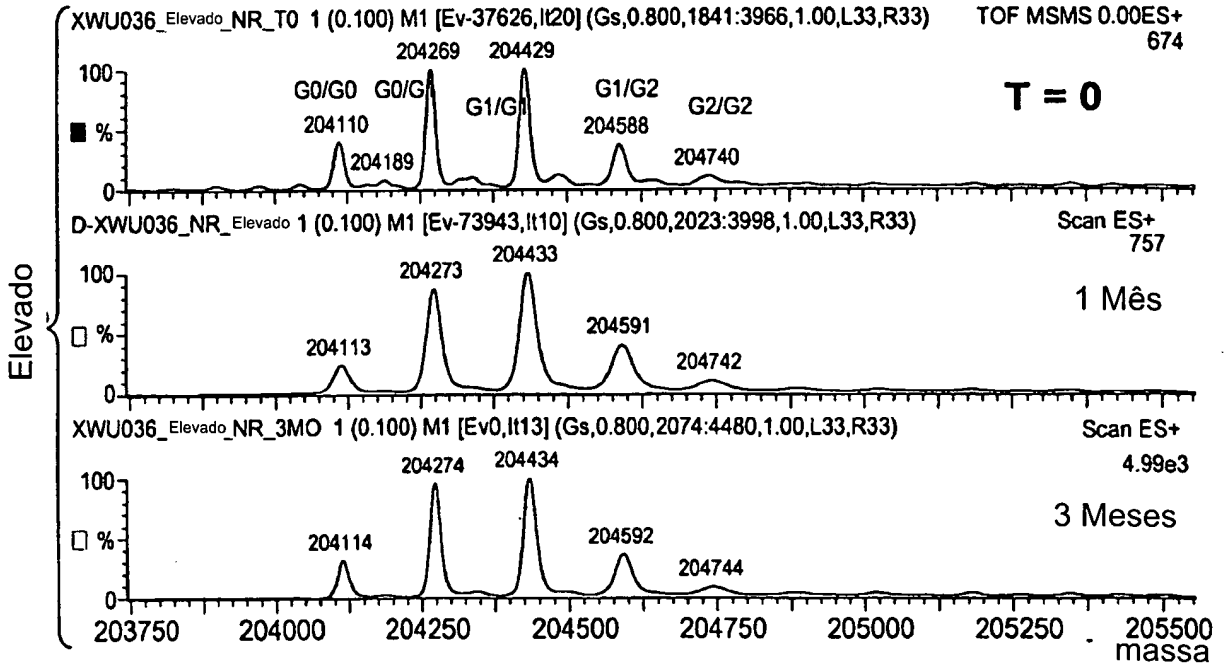


Fig. 58B

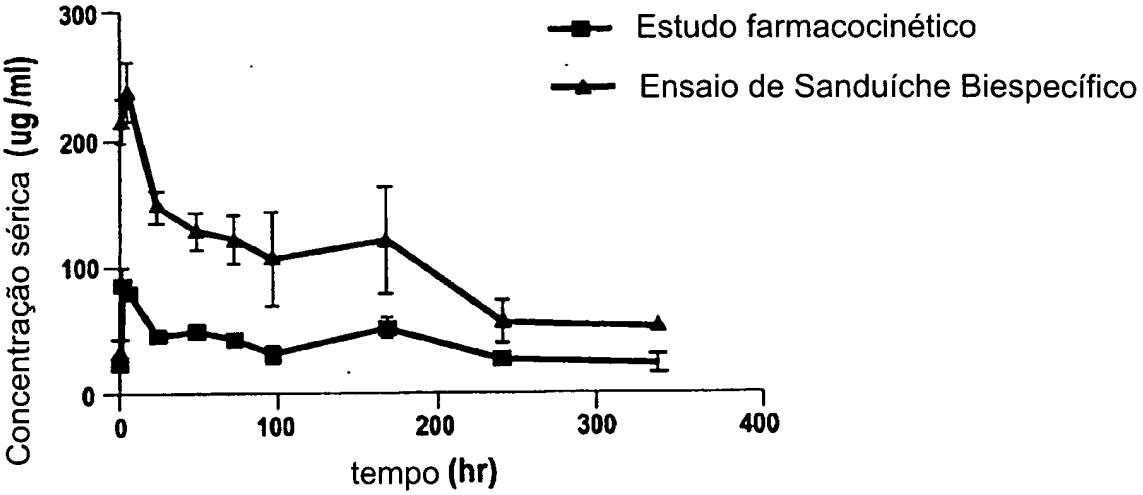


Fig. 59A

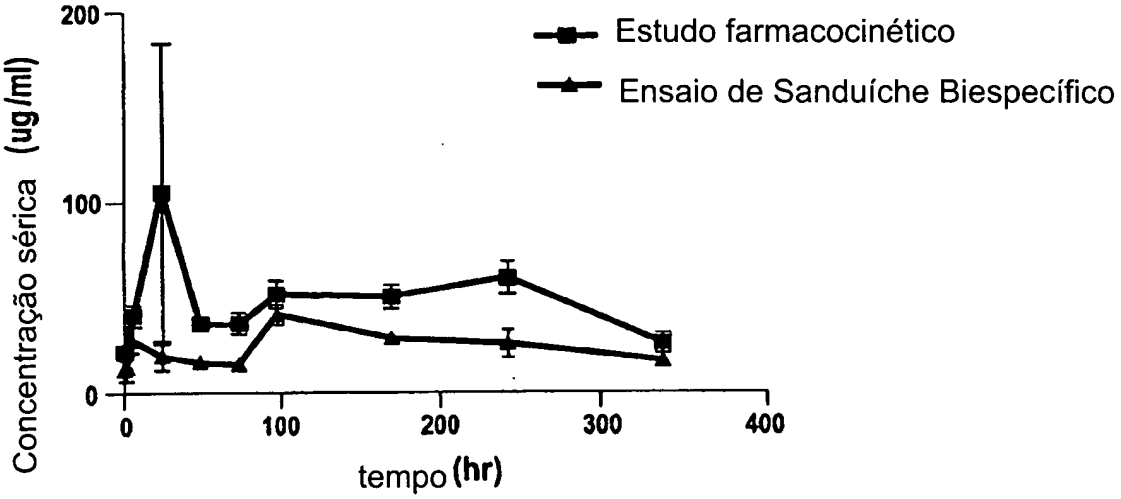


Fig. 59B

A Dosagem iniciou no Dia 13

hCBE11 vs VC: P<0.001
hBHA10 vs VC: P<0.001
ch14A2 vs VC: P<0.01
Hercules-II XWU028 vs V.C.:P<0.001
Hercules-II XWU028 vshBHA10.:P<0.05
Hercules-II XWU028 vs ch14A2.:P<0.05
Hercules-II XWU036 vs V.C.:P<0.001
Hercules-II XWU036 vs hBHA10:P<0.01
Hercules-II XWU036 vs ch14A2:P<0.05
Combo vs. V.C.:P<0.001

Veículo Controle (PBS sem pirogênio) I, **10 ml/kg/inj, IP,** 2x/semana
hCBE11, **2 mg/kg/inj, IP,** 1x/semana
hBHA10, **2 mg/kg/inj, IP,** 2x/semana
ch14A2, **2 mg/kg/inj, IP,** 2x/semana
Hercules-II XWU028, **2 mg/kg/inj, IP,** 1x/semana
Hercules-II XWU036, **2 mg/kg/inj, IP,** 1x/semana
hBHA10, **1 mg/kg/inj, IP,** 2x/semana
mais ch14A2, **1 mg/kg/inj, IP,** 2x/semana

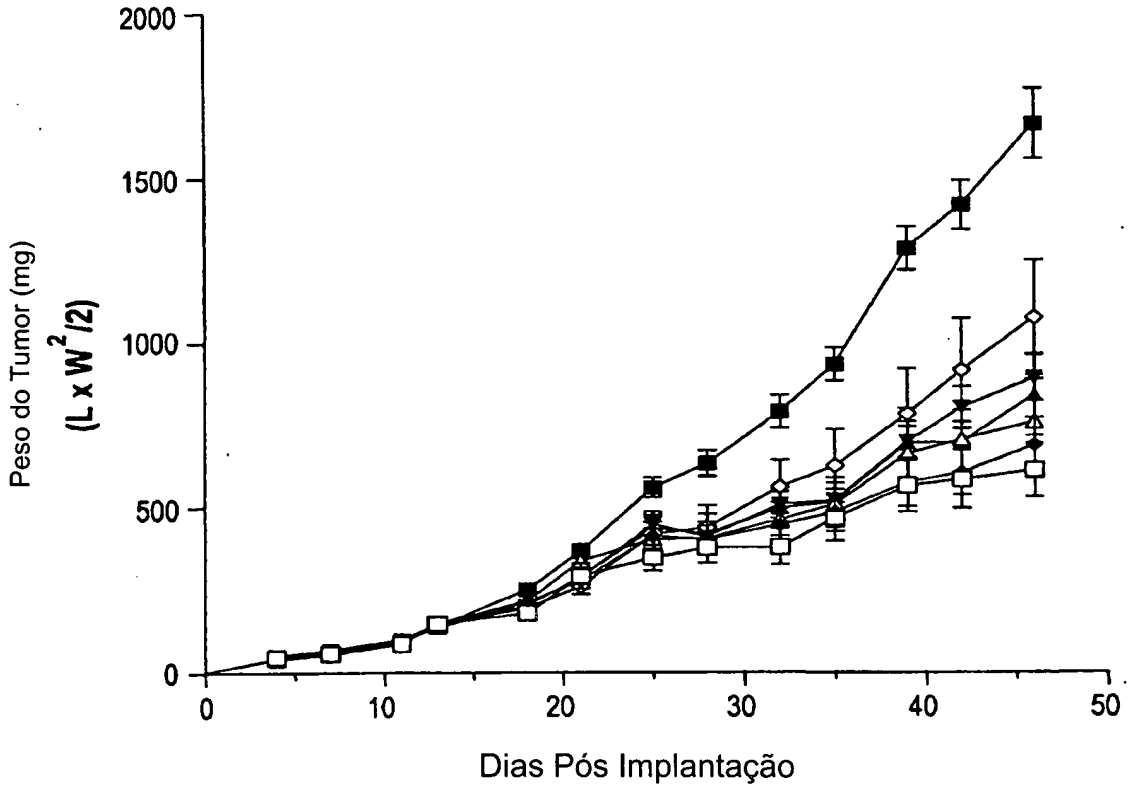


Fig. 60

CAGATCCAGTTGGTGCAGTCTGGACCTGAGCTGAAGAAGCCTGGAGAGACAGTCAAGAT
CTCCTGCAAGGCTTCTGGTTTTACCTTCACAGACTATTCAATACACTGGGTGAAACAGG
CTCCAGGAAAGGGTTTAAAGTGGATGGGCTGGATAAACACTGAGACTGGTGAGCCAACA
TATACAGATGACTTCAAGGGACGATTTGCCTTCTCTTTGGTGACCTCTGCCACCACTGC
CTATTTGCAGATCAACAACCTCAACAATGAGGACACGGCTACATTTTTCTGTGCTAGAT
TCATCTATGATCCTTATTGGGGGTTTGCTTACTGGGGCCAGGGGACTCTGGTCACTGTC
TCCGCAGCTAGCACCAAGGGCCCATCGGTCTTCCCCCTGGCACCCTCCTCCAAGAGCAC
CTCTGGGGGCACAGCGGCCCTGGGCTGCCTGGTCAAGGACTACTTCCCCGAACCGGTGA
CGGTGTCGTGGAACTCAGGCGCCCTGACCAGCGGCGTGCACACCTTCCCGGCTGTCCTA
CAGTCCTCAGGACTCTACTCCCTCAGCAGCGTGGTGACCGTGCCCTCCAGCAGCTTGGG
CACCCAGACCTACATCTGCAACGTGAATCACAAGCCCAGCAACACCAAGGTGGACAAGA
AAGTTGAGCCCAAATCTTGTGACAAAACTCACACATGCCCACCGTGCCCAGCACCTGAA
CTCCTGGGGGGACCGTCAGTCTTCCTCTTCCCCCCAAAACCCAAGGACACCCTCATGAT
CTCCCGGACCCCTGAGGTCACATGCGTGGTGGTGGACGTGAGCCACGAAGACCCTGAGG
TCAAGTTCAACTGGTACGTGGACGGCGTGGAGGTGCATAATGCCAAGACAAAGCCGCGG
GAGGAGCAGTACAACAGCACGTACCGTGTGGTCAGCGTCCTCACCGTCCTGCACCAGGA
CTGGCTGAATGGCAAGGAGTACAAGTGCAAGGTCTCCAACAAAGCCCTCCCAGCCCCCA
TCGAGAAAACCATCTCCAAAGCCAAAGGGCAGCCCCGAGAACCACAGGTGTACACCCTG
CCCCCATCCCGGGATGAGCTGACCAAGAACCAGGTCAGCCTGACCTGCCTGGTCAAAGG
CTTCTATCCCAGCGACATCGCCGTGGAGTGGGAGAGCAATGGGCAGCCGGAGAACAACT
ACAAGACCACGCCTCCCGTGCTGGACTCCGACGGCTCCTTCTTCCTCTACAGCAAGCTC
ACCGTGGACAAGAGCAGGTGGCAGCAGGGGAACGTCTTCTCATGCTCCGTGATGCATGA
GGCTCTGCACAACCACTACACGCAGAAGAGCCTCTCCCTGTCTCCGGGTAAATCCGGCG
GGGGTGGATCCGGTGGAGGGGGCTCCGGCGGTGGCGGGTCCCAGGTCCAACTGGTGCAG
TCTGGAGCTGAGGTGAAGAAGCCTGGGGAGTCAGTGAAGGTGTCCTGCAAGGCTTCTGG
CTACACTTTCACAACCTACTATTTGCACTGGGTGAGGCAGGCCCCTGGACAGGGACTTG
AGTGGATGGGATGGATTTATCCTGGAAATGTTCATGCTCAGTACAATGAGAAGTTCAAG
GGCAGGGTCACAATCACTGCAGACAAATCCACCAGCACAGCCTACATGGAGCTCAGCAG
CCTGAGGTCTGAAGATACTGCGGTCTATTACTGTGCAAGATCCTGGGAAGGTTTTCCTT
ACTGGGGCCAAGGGACCACGGTCACCGTCTCCTCAGGCGGTGGAGGGTCCGGTGGGGGC
GGATCTGGGGGCGGCGGGTCCGGTGGTGGTGGTAGTGACATTCAGATGACCCAGTCTCC
TAGCTCCCTGTCCGCCTCAGTAGGAGACAGGGTCACCATCACCTGCAAGGCCAGTCAGA
ATGTGGGTATTAATGTAGCCTGGTATCAACAGAAACCAGGGAAGGCTCCTAAATTACTG
ATTTCCTCGGCCTCCTACCGGTACAGTGGAGTCCCTTCCAGATTCAGCGGCAGTGGATC
TGGGACAGATTTCACTCTCACCATCAGCAGCCTCCAGCCTGAAGACTTCGCAACCTATT
TCTGTCAGCAATATGACACCTATCCATTCACGTTCGGCCAGGGTACCAAGGTGGAGATC
AAATGA **(SEQ ID NO: 52)**

Fig. 61

QIQLVQS GPELK KPGETVK ISCKA SGFTFTD YSIHW VKQAPGK G
LKWMGWI NTETG EPTYTDD FKGRF AFSLVTS ATTAY LQINNLN N
EDTATFF CARFI YDPYWGF AYWGQ GTLVTVS AASTK GPSVFPL A
PSSKSTS GGTAA LGCLVKD YFPEP VTVSWNS GALTS GVHTFPA V
LQSSGLY SLSSV VTVPSSS LGTQT YICNVNH KPSNT KVDKKVE P
KSCDKTH TCPPC PAPELLG GPSVF LFPPKPK DTLMI SRTPEVT C
VVVDVSH EDPEV KFNWYVD GVEVH NAKTKPR EEQYN STYRVVS V
LTVLHQD WLNGK EYKCKVS NKALP APIEKTI SKAKG QPREPQV Y
TLPPSRD ELTKN QVSLTCL VKGFY PSDIAVE WESNG QPENNYK T
TPPVLDS DGSFF LYSKLTV DKSRW QQGNVFS CSVMH EALHNHY T
QKSLSLS PGKSG GGGSGGG GSGGG GSQVQLV QSGAE VKKPGES V
KVSCKAS GYTFT TYYLHWV RQAPG QGLEWMG WIYPG NVHAQYN E
KFKGRVT ITADK STSTAYM ELSSL RSEDTAV YYCAR SWEGFPY W
GQGTTVT VSSGG GGSGGGG SGGGG SGGGGSD IQMTQ SPSSLSA S
VGDRVTI TCKAS QNVGINV AWYQQ KPGKAPK LLISS ASYRYSG V
PSRFSGS GSGTD FTLTISS LQPED FATYFCQ QYDTY PFTFGQG T
KVEIK **(SEQ ID NO: 53)**

Fig. 62

CAGATCCAGTTGGTGCAGTCTGGACCTGAGCTGAAGAAGCCTGGAGAGACAGTCAAG
ATCTCCTGCAAGGCTTCTGGTTTTACCTTCACAGACTATTCAATACACTGGGTGAAA
CAGGCTCCAGGAAAGGGTTTAAAGTGGATGGGCTGGATAAACACTGAGACTGGTGAG
CCAACATATACAGATGACTTCAAGGGACGATTTGCCTTCTCTTTGGTGACCTCTGCC
ACCACTGCCTATTTGCAGATCAACAACCTCAACAATGAGGACACGGCTACATTTTTC
TGTGCTAGATTCATCTATGATCCTTATTGGGGGTTTGCTTACTGGGGCCAGGGGACT
CTGGTCACTGTCTCCGCAGCTAGCACCAAGGGCCCATCGGTCTTCCCCCTGGCACCC
TCCTCCAAGAGCACCTCTGGGGGCACAGCGGCCCTGGGCTGCCTGGTCAAGGACTAC
TTCCCCGAACCGGTGACGGTGTCGTGGAACTCAGGCGCCCTGACCAGCGGCGTGCAC
ACCTTCCCGGCTGTCCTACAGTCCTCAGGACTCTACTCCCTCAGCAGCGTGGTGACC
GTGCCCTCCAGCAGCTTGGGCACCCAGACCTACATCTGCAACGTGAATCACAAGCCC
AGCAACACCAAGGTGGACAAGAAAGTTGAGCCCAAATCTTGTGACAAAACTCACACA
TGCCCACCGTGCCCAGCACCTGAACTCCTGGGGGGACCGTCAGTCTTCCTCTTCCCC
CCAAAACCCAAGGACACCCTCATGATCTCCCGGACCCCTGAGGTCACATGCGTGGTG
GTGGACGTGAGCCACGAAGACCCTGAGGTCAAGTTCAACTGGTACGTGGACGGCGTG
GAGGTGCATAATGCCAAGACAAAGCCGCGGGAGGAGCAGTACAACAGCACGTACCGT
GTGGTCAGCGTCCTCACCGTCCTGCACCAGGACTGGCTGAATGGCAAGGAGTACAAG
TGCAAGGTCTCCAACAAAGCCCTCCCAGCCCCCATCGAGAAAACCATCTCCAAAGCC
AAAGGGCAGCCCCGAGAACCACAGGTGTACACCCTGCCCCCATCCCGGGATGAGCTG
ACCAAGAACCAGGTCAGCCTGACCTGCCTGGTCAAAGGCTTCTATCCCAGCGACATC
GCCGTGGAGTGGGAGAGCAATGGGCAGCCGGAGAACAACTACAAGACCACGCCTCCC
GTGCTGGACTCCGACGGCTCCTTCTTCCTCTACAGCAAGCTCACCGTGGACAAGAGC
AGGTGGCAGCAGGGGAACGTCTTCTCATGCTCCGTGATGCATGAGGCTCTGCACAAC
CACTACACGCAGAAGAGCCTCTCCCTGTCTCCGGGTAAATCCGGCGGGGGTGGATCC
GGTGGAGGGGGCTCCGGCGGTGGCGGGTCCCAGGTCCAACTGGTGCAGTCTGGAGCT
GAGGTGAAGAAGCCTGGGGAGTCAGTGAAGGTGTCCTGCAAGGCTTCTGGCTACACT
TTCACAACCTACTATTTGCACTGGGTGAGGCAGGCCCCTGGACAGTGCCTTGAGTGG
ATGGGATGGATTTATCCTGGAAATGTTCATGCTCAGTACAATGAGAAGTTCAAGGGC
AGGGTCACAATCACTGCAGACAAATCCACCAGCACAGCCTACATGGAGCTCAGCAGC
CTGAGGTCTGAAGATACTGCGGTCTATTACTGTGCAAGATCCTGGGAAGGTTTTCCT
TACTGGGGCCAAGGGACCACGGTCACCGTCTCCTCAGGCGGTGGAGGGTCCGGTGGG
GGCGGATCTGGGGGCGGCGGGTCCGGTGGTGGTGGTAGTGACATTCAGATGACCCAG
TCTCCTAGCTCCCTGTCCGCCTCAGTAGGAGACAGGGTCACCATCACCTGCAAGGCC
AGTCAGAATGTGGGTATTAATGTAGCCTGGTATCAACAGAAACCAGGGAAGGCTCCT
AAATTACTGATTTCCTCGGCCTCCTACCGGTACAGTGGAGTCCCTTCCAGATTCAGC
GGCAGTGGATCTGGGACAGATTTCACTCTCACCATCAGCAGCCTCCAGCCTGAAGAC
TTCGCAACCTATTTCTGTCAGCAATATGACACCTATCCATTCACGTTCGGCTGCGGT
ACCAAGGTGGAGATCAAATGA **(SEQ ID NO: 54)**

Fig. 63

QIQLVQS GPELK KPGETVK ISCKA SGFTFTD YSIHW VKQAPGK GLK
WMGWINT ETGEP TYTDDFK GRFAF SLVTSAT TAYLQ INNLNNE DTA
TFFCARF IYDPY WGFAYWG QGTLV TVSAAST KGPSV FPLAPSS KST
SGGTAAL GCLVK DYFPEPV TVSWN SGALTSG VHTFP AVLQSSG LYS
LSSVVTV PSSSL GTQTYIC NVNHK PSNTKVD KKVEP KSCDKTH TCP
PCPAPEL LGGPS VFLFPPK PKDTL MISRTPE VTCVV VDVSHED PEV
KFNWYVD GVEVH NAKTKPR EEQYN STYRVVS VLTVL HQDWLNG KEY
KCKVSNK ALPAP IEKTISK AKGQP REPQVYT LPPSR DELTKNQ VSL
TCLVKGF YPSDI AVEWESN GQPEN NYKTTPP VLDSD GSFFLYS KLT
VDKSRWQ QGNVF SCSVMHE ALHNH YTQKSLS LSPGK SGGGGSG GGG
SGGGGSQ VQLVQ SGAEVKK PGESV KVSCKAS GYTFT TYYLHWV RQA
PGQCLEW MGWIY PGNVHAQ YNEKF KGRVTIT ADKST STAYMEL SSL
RSEDTAV YYCAR SWEGFPY WGQGT TVTVSSG GGGSG GGGSGGG GSG
GGGSDIQ MTQSP SSLSASV GDRVT ITCKASQ NVGIN VAWYQQK PGK
APKLLIS SASYR YSGVPSR FSGSG SGTDFTL TISSL QPEDFAT YFC
QQYDTYP FTFGC GTKVEIK

(SEQ ID NO: 55)

Fig. 64

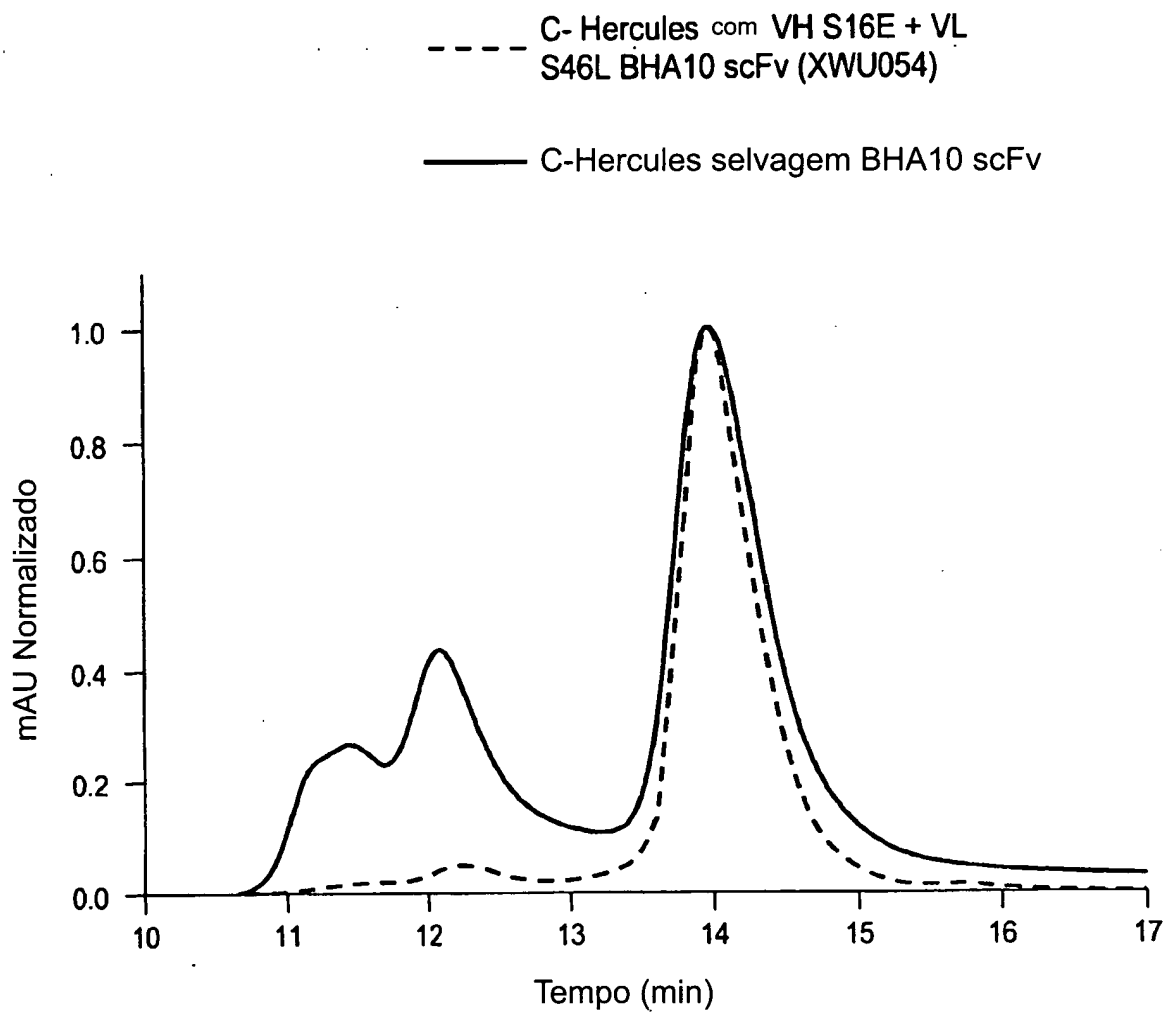

C- Hercules com VH S16E + VL S46L BHA10 scFv (XWU054)
C-Hercules selvagem BHA10 scFv
mAU Normalizado
1.0
0.8
0.6
0.4
0.2
0.0
10
11
12
13
14
15
16
17
Tempo (min)

Fig. 65

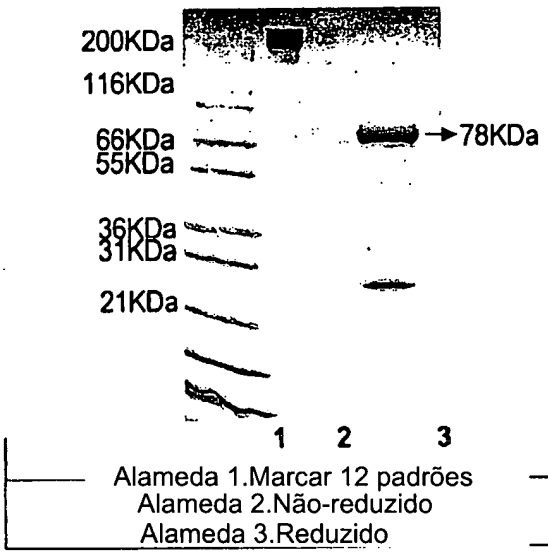


Fig. 66A

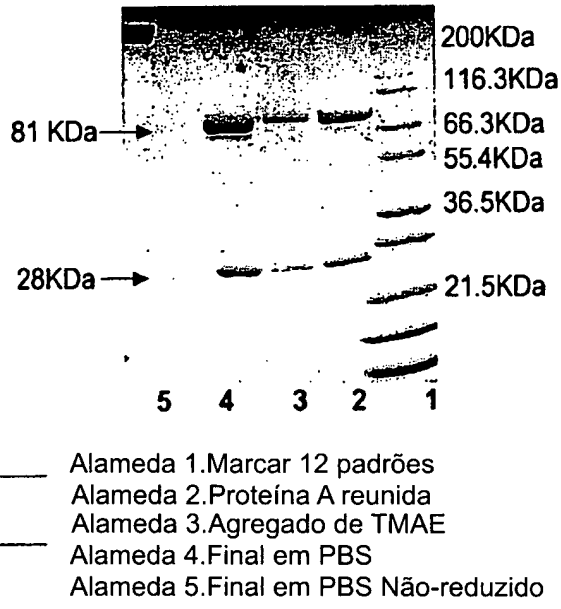


Fig. 66B

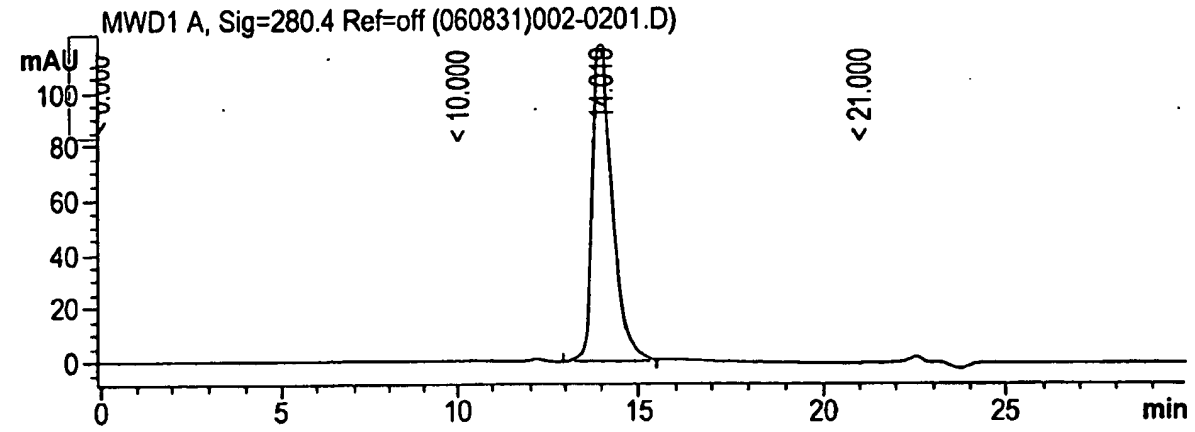


Fig. 67A

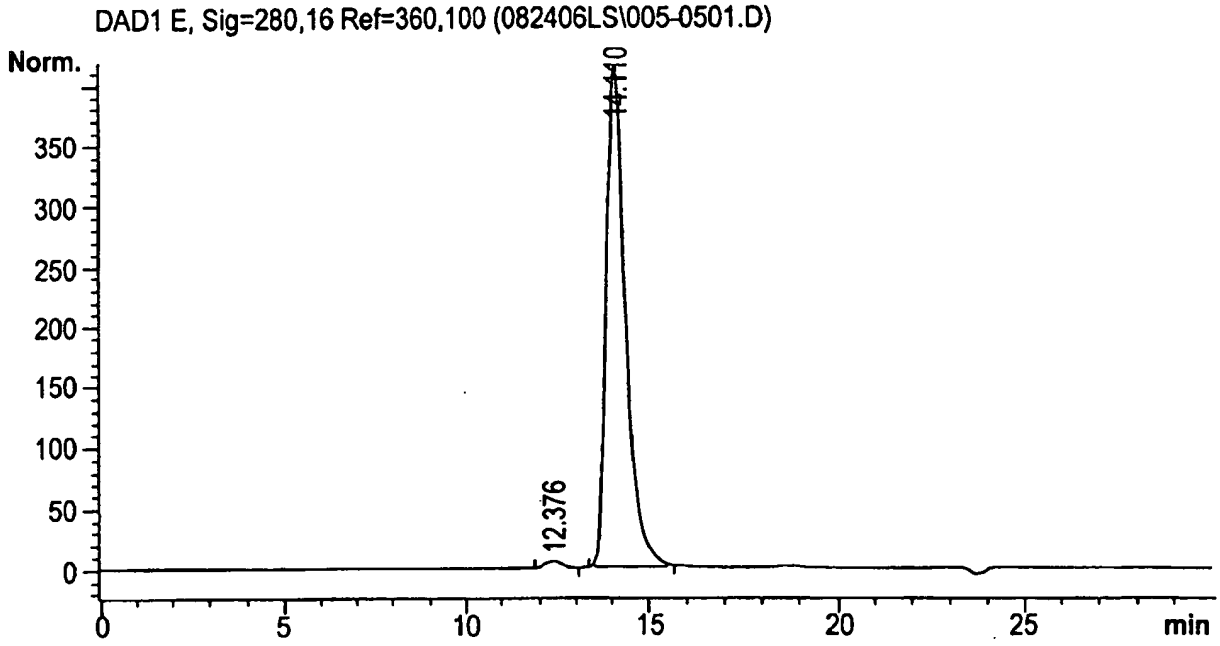


Fig. 67B

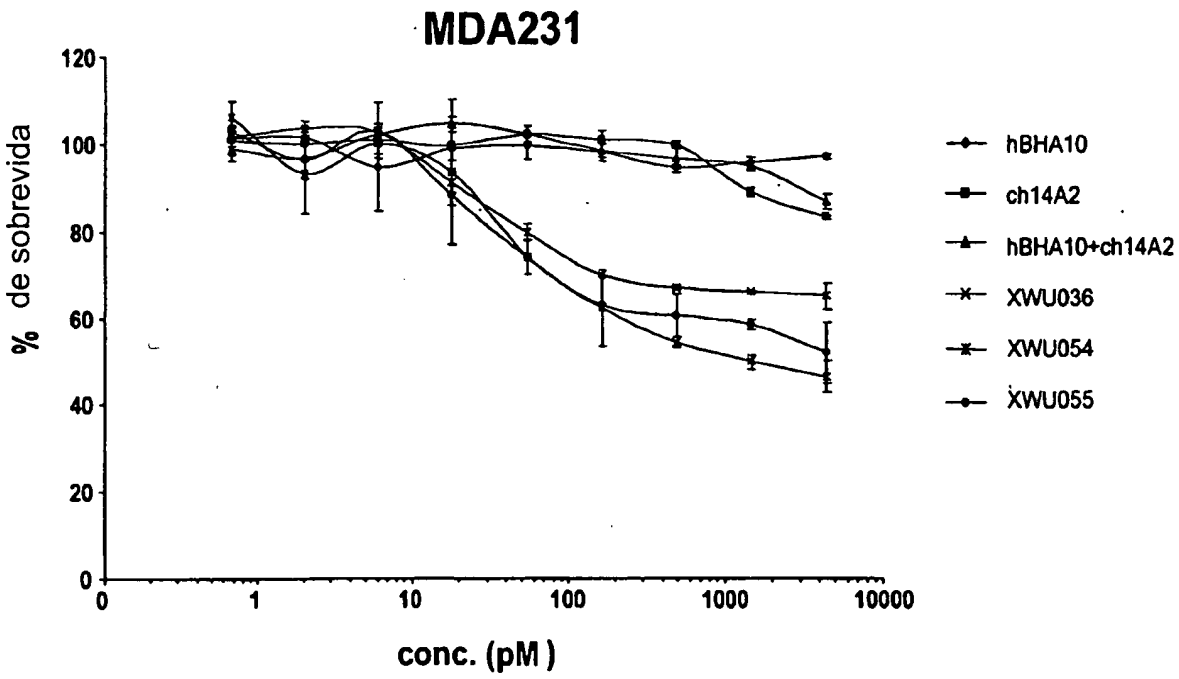


Fig. 68A

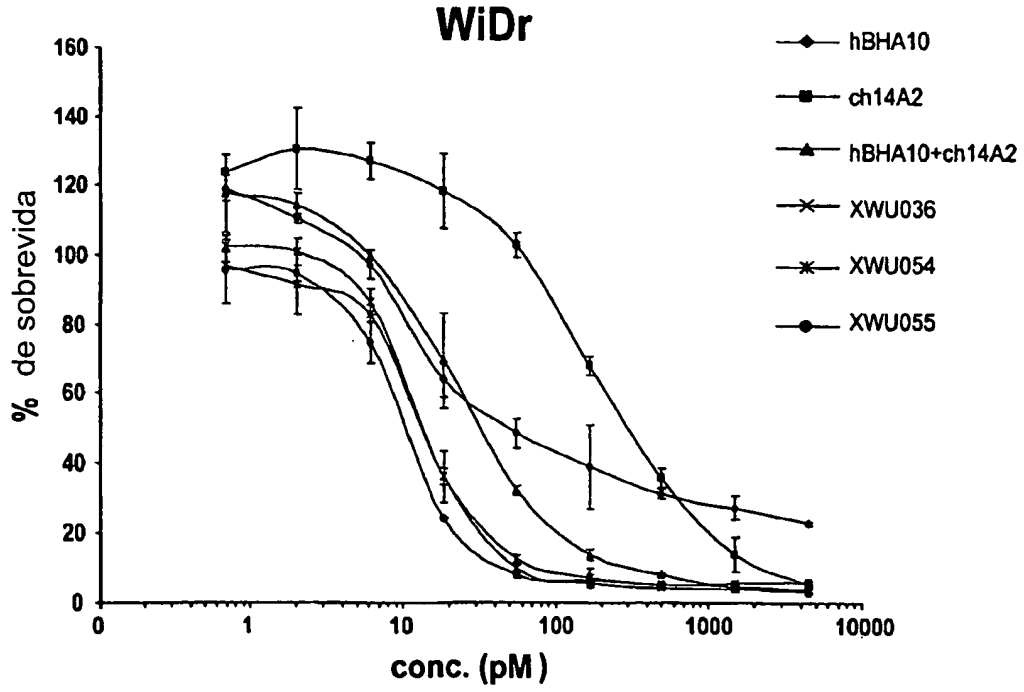


Fig. 68B

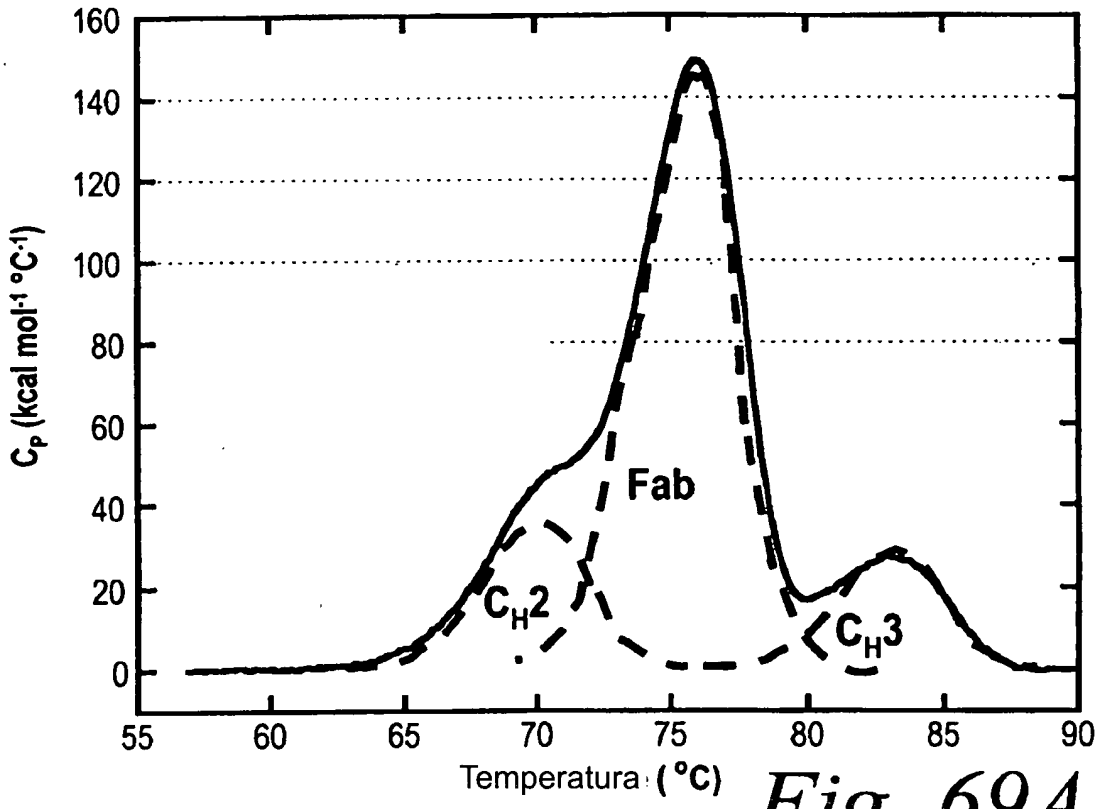

160
140
120
100
80
60
40
20
0
C_p (kcal mol^{-1} °C^{-1})
Fab
C_H2
C_H3
55
60
65
70
75
80
85
90
Temperatura (°C)

Fig. 69A

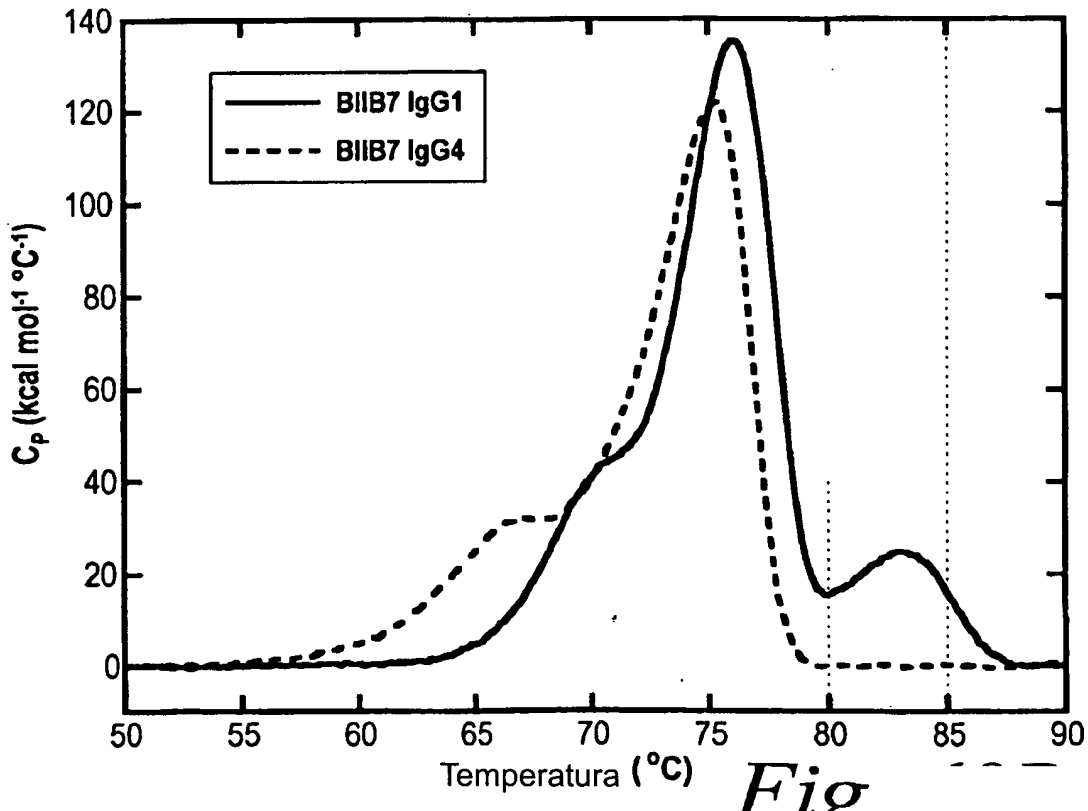

140
120
100
80
60
40
20
0
C_p (kcal mol^{-1} °C^{-1})
BIIB7 IgG1
BIIB7 IgG4
50
55
60
65
70
75
80
85
90
Temperatura (°C)

Fig.

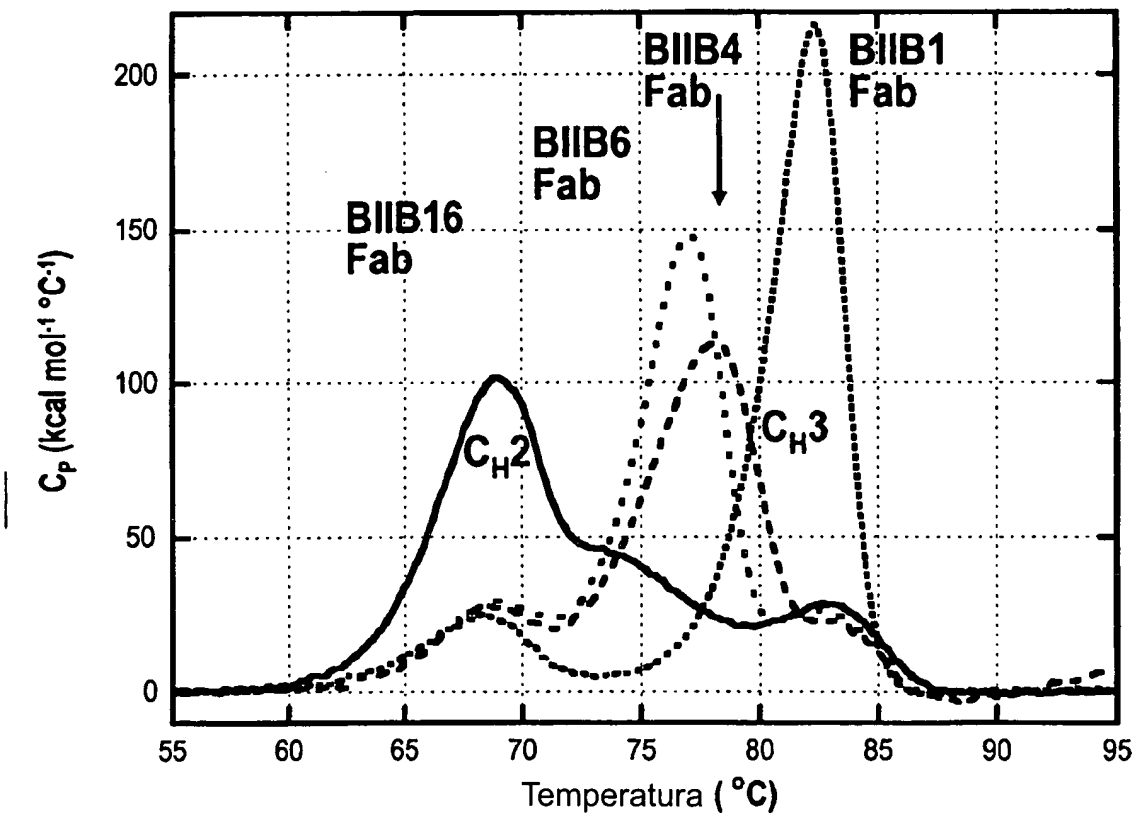
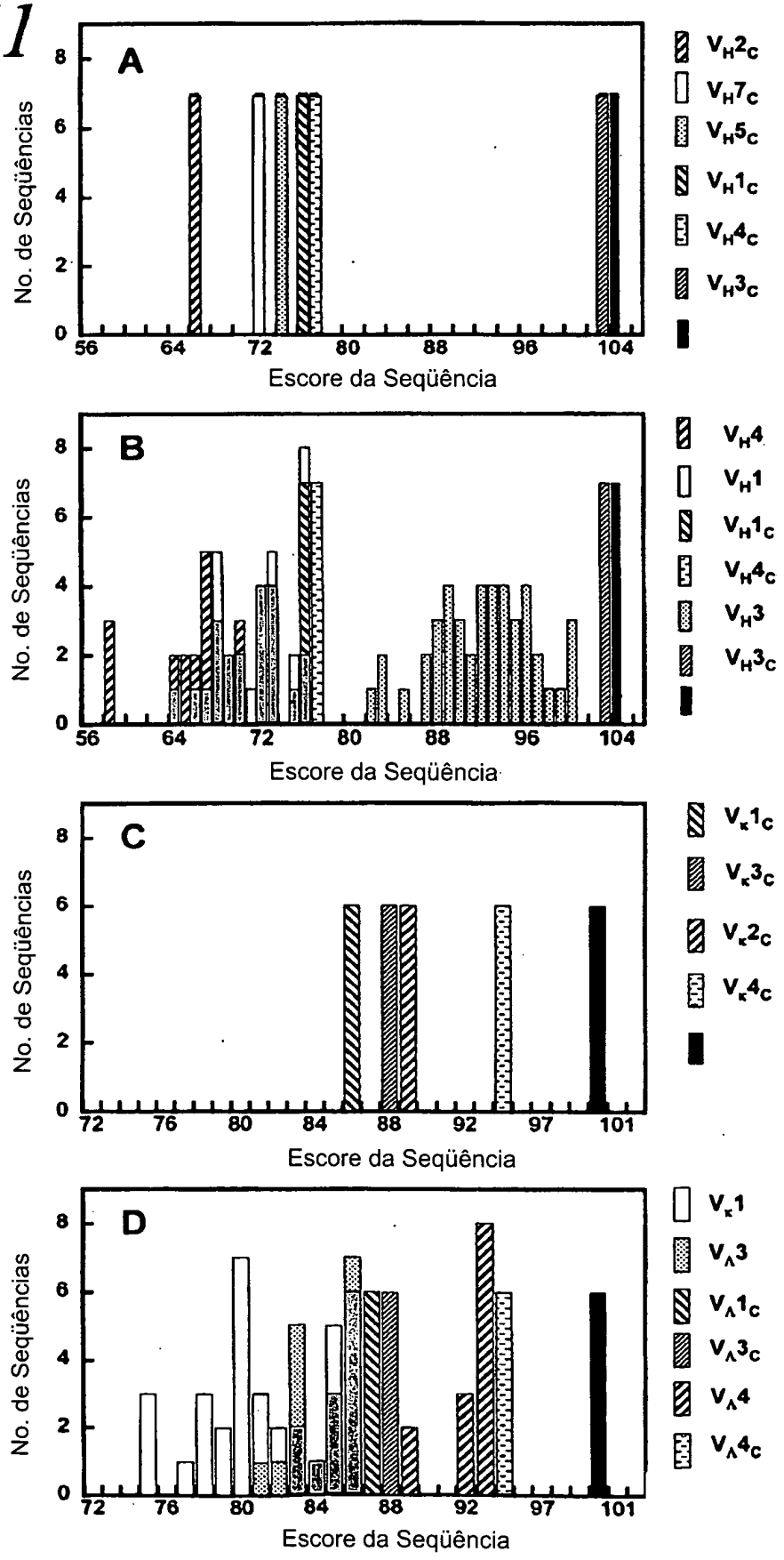

BIIB4
Fab
BIIB1
Fab
BIIB6
Fab
BIIB16
Fab
C_H2
C_H3
200
150
100
50
0
55
60
65
70
75
80
85
90
95
C_p (kcal mol^{-1} °C^{-1})
Temperatura (°C)

Fig. 70

Fig. 71


A
No. de Seqüências
Escore da Seqüência
V_H2_C
V_H7_C
V_H5_C
V_H1_C
V_H4_C
V_H3_C
B
No. de Seqüências
Escore da Seqüência
V_H4
V_H1
V_H1_C
V_H4_C
V_H3
V_H3_C
C
No. de Seqüências
Escore da Seqüência
$V_\kappa1_C$
$V_\kappa3_C$
$V_\kappa2_C$
$V_\kappa4_C$
D
No. de Seqüências
Escore da Seqüência
$V_\kappa1$
$V_\Lambda3$
$V_\Lambda1_C$
$V_\Lambda3_C$
$V_\Lambda4$
$V_\Lambda4_C$

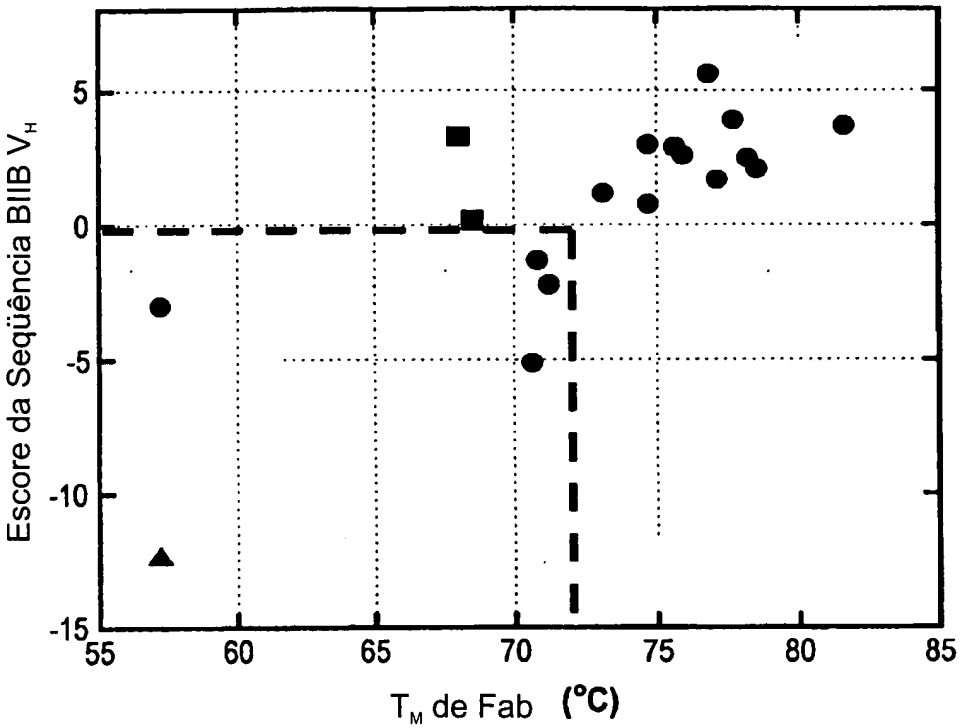


Fig. 72A

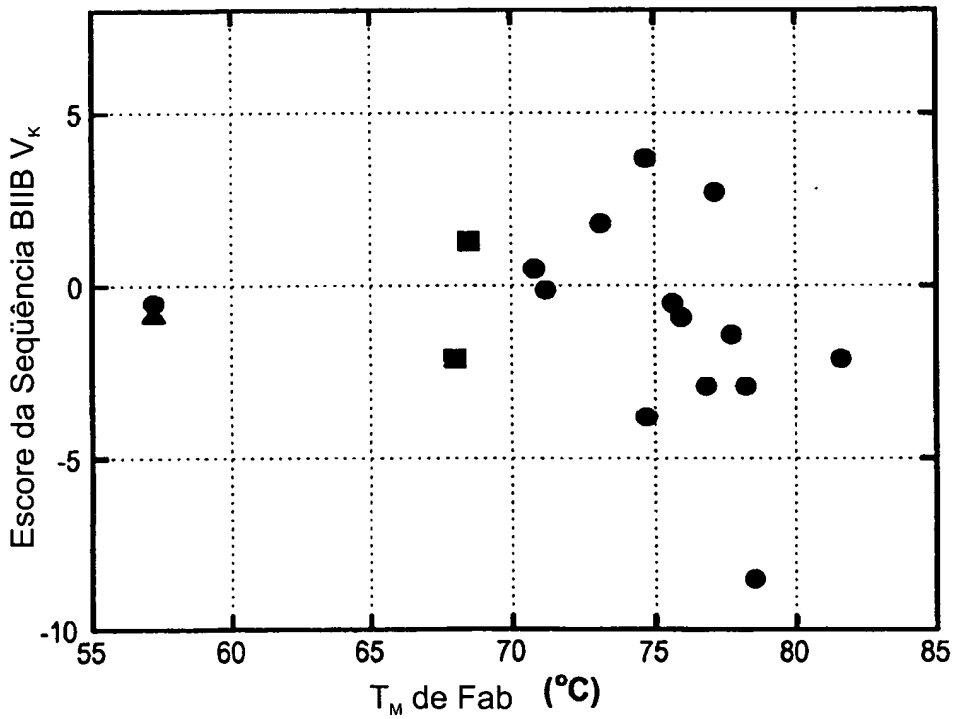


Fig. 72B

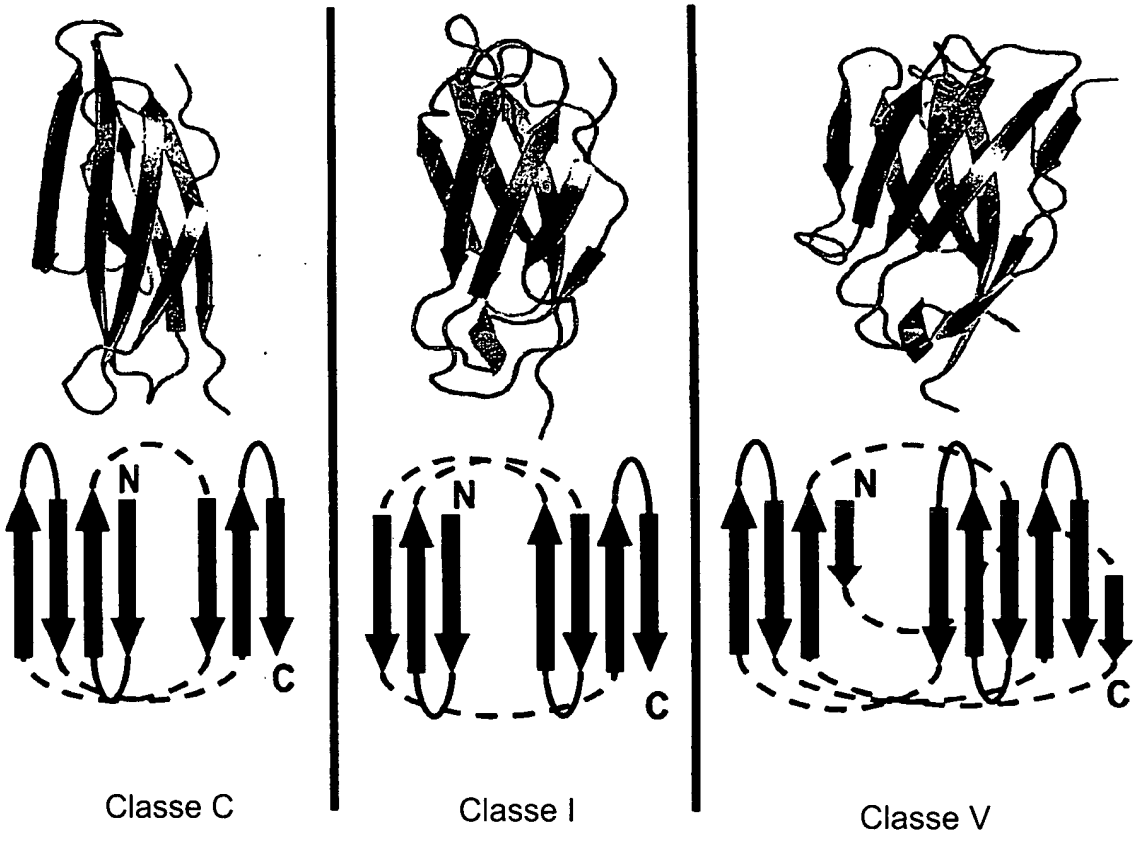


Fig. 73

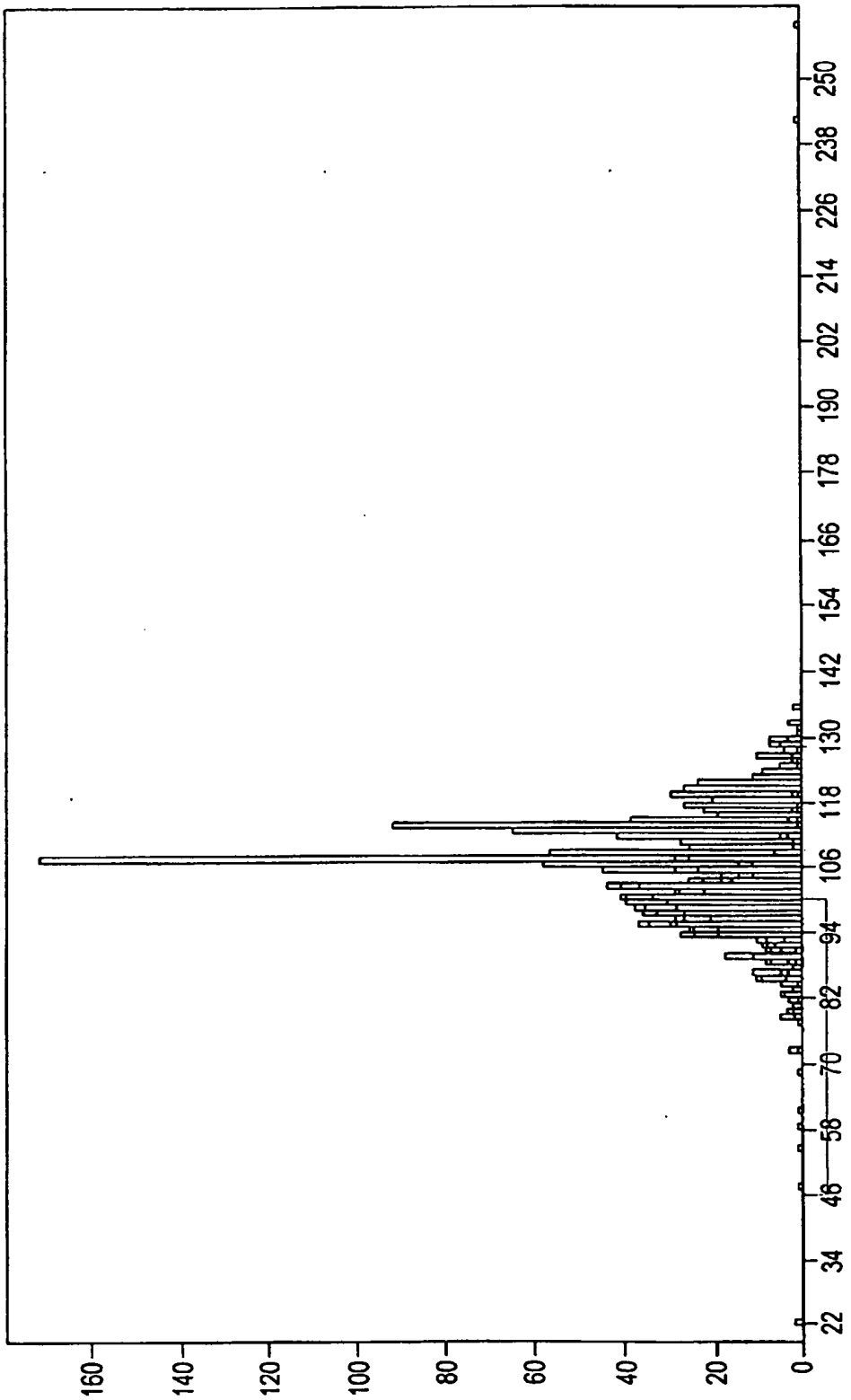

160
140
120
100
80
60
40
20
0
22
34
46
58
70
82
94
106
118
130
142
154
166
178
190
202
214
226
238
250

Fig. 74

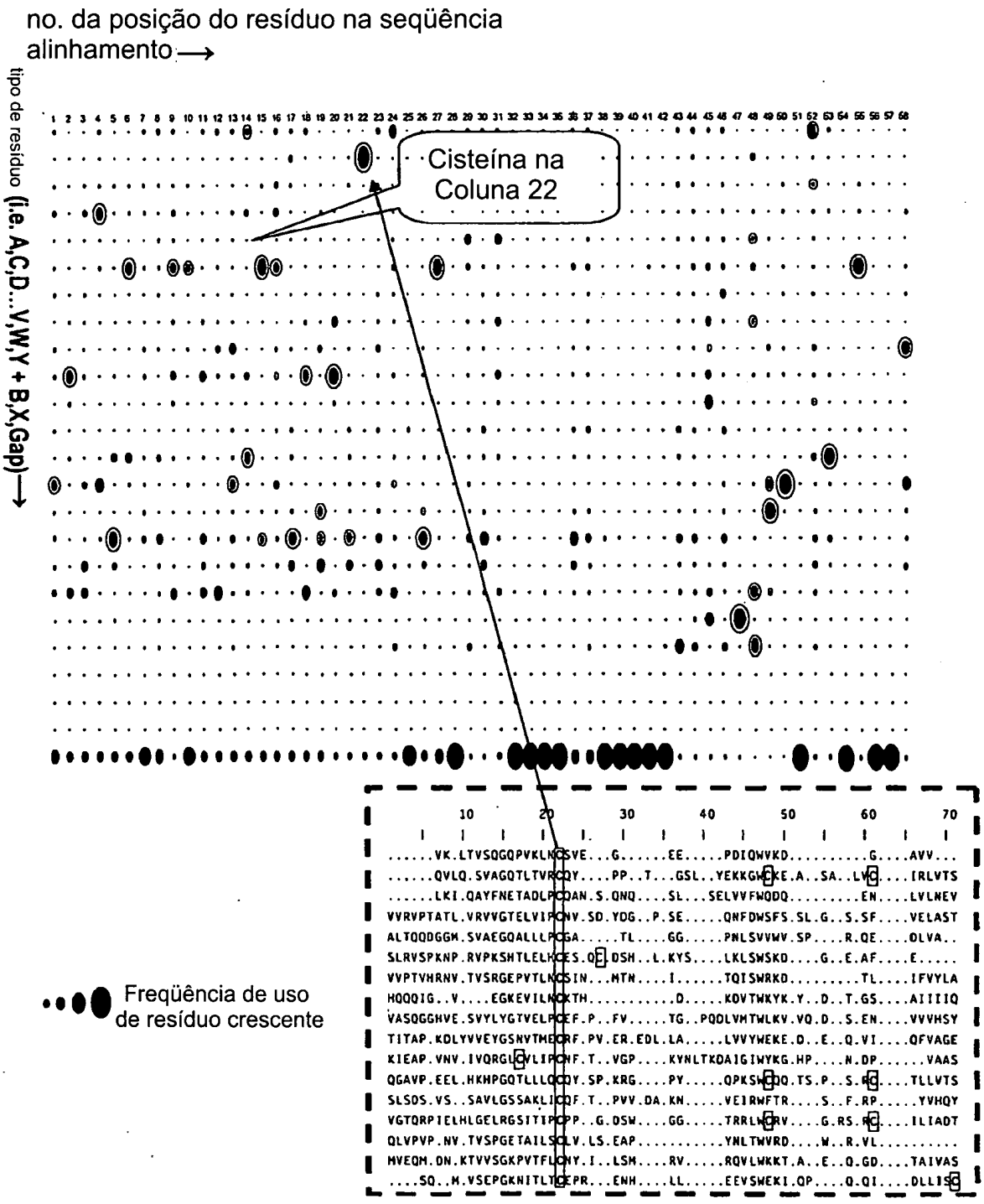

no. da posição do resíduo na seqüência alinhamento ⟶
tipo de resíduo (i.e. A,C,D...V,W,Y + B,X,Gap) ⟶
Cisteína na Coluna 22
Freqüência de uso de resíduo crescente
10
20
30
40
50
60
70

Fig. 76

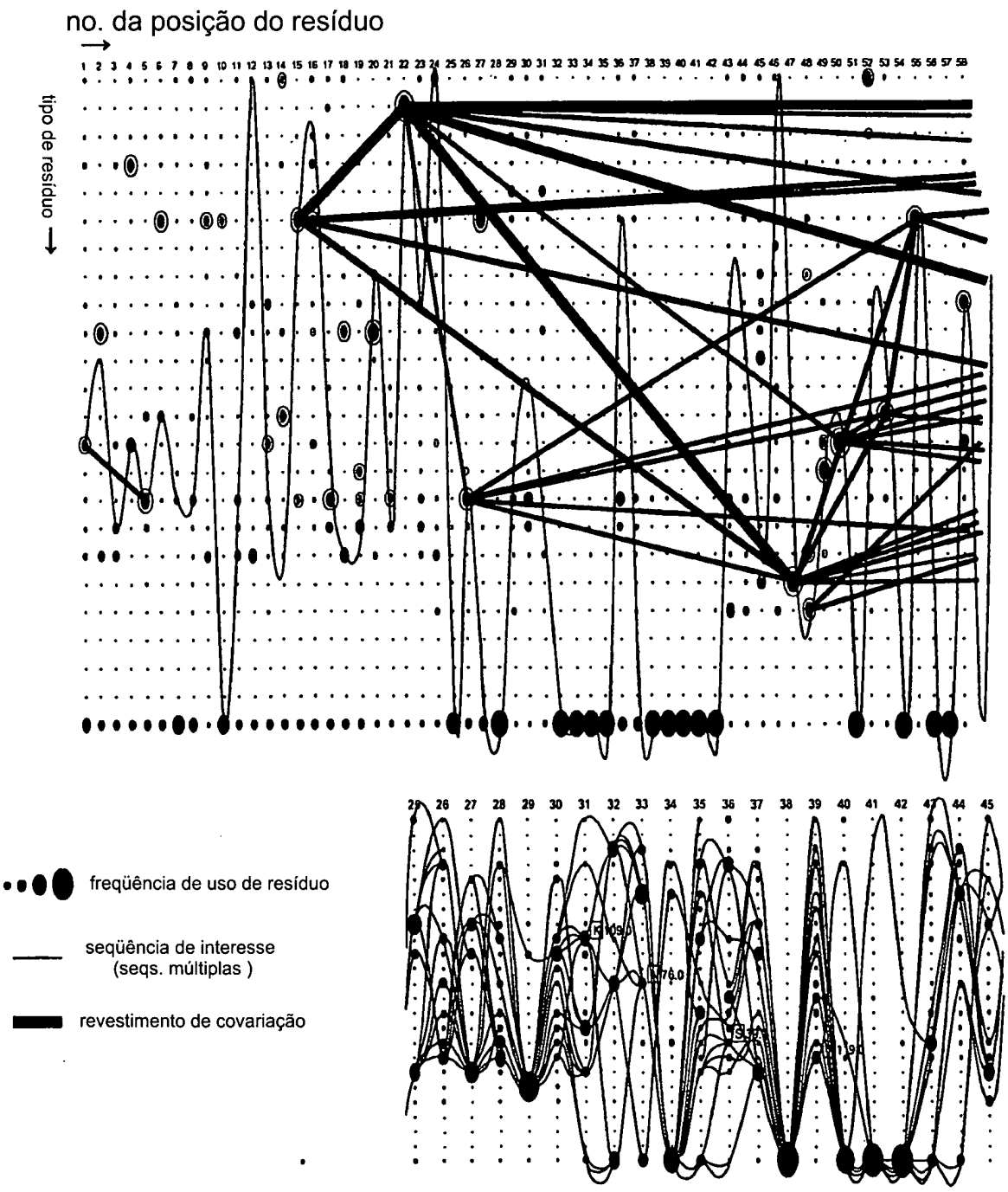

no. da posição do resíduo
tipo de resíduo
freqüência de uso de resíduo
seqüência de interesse
(seqs. múltiplas)
revestimento de covariação

Fig. 77

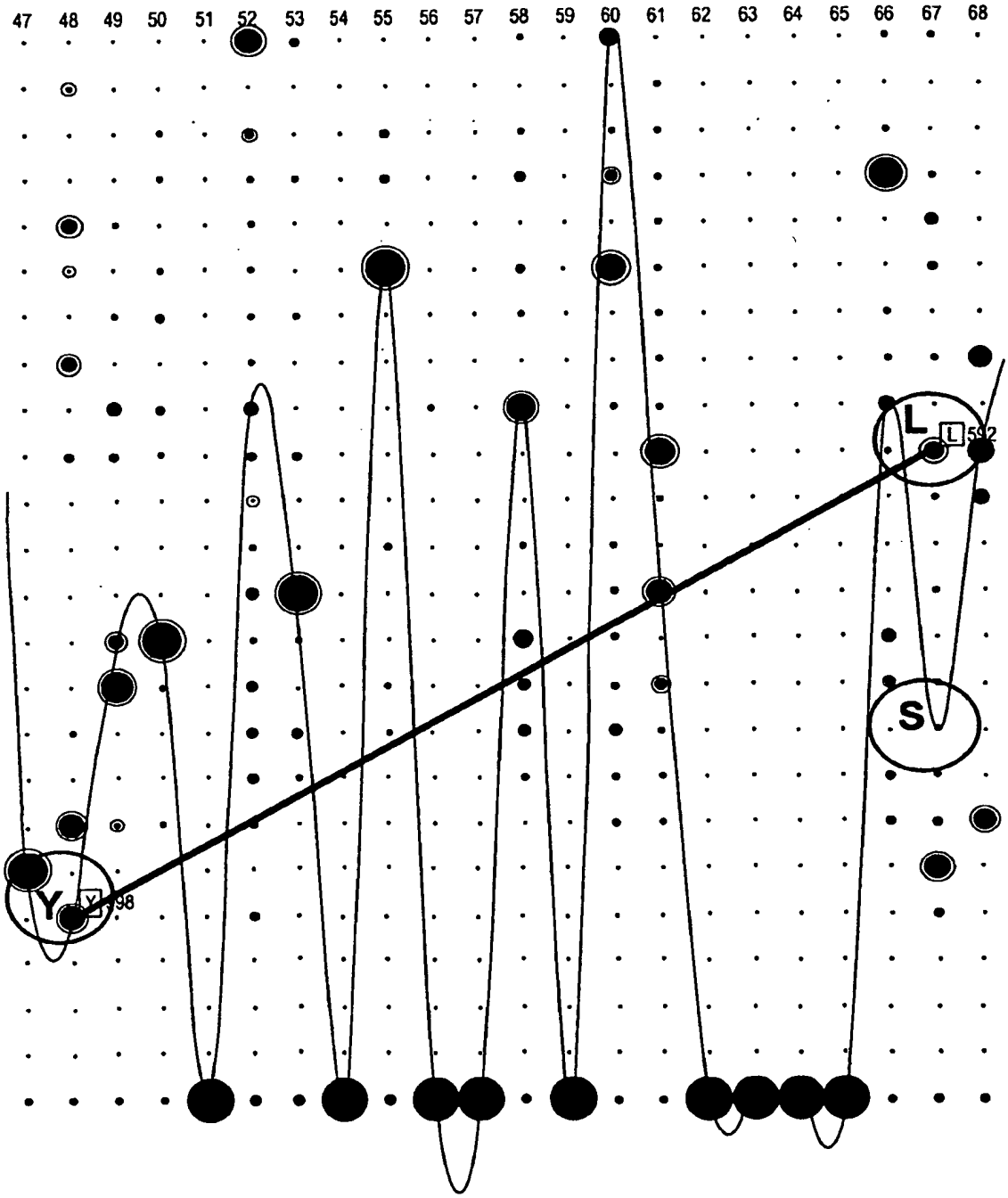
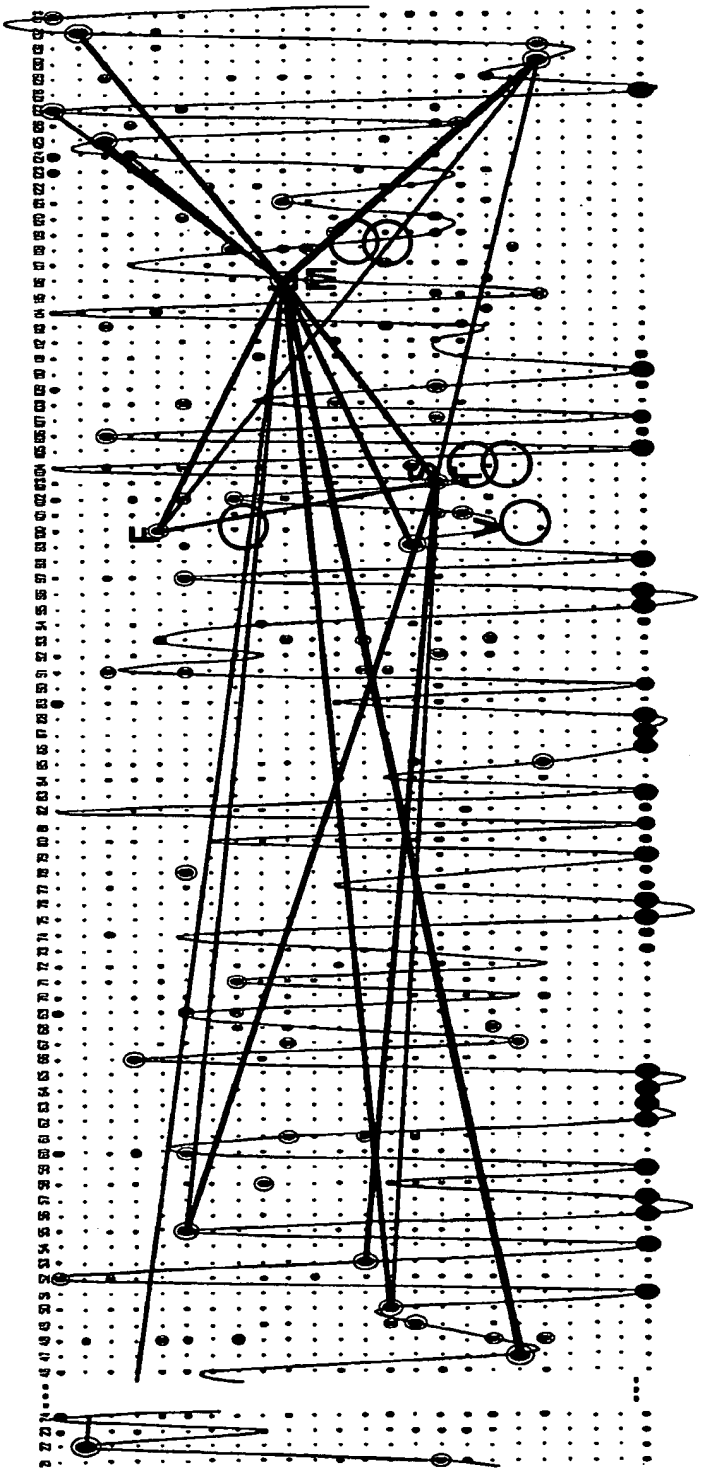

47 48 49 50 51 52 53 54 55 56 57 58 59 60 61 62 63 64 65 66 67 68
L
592
S
Y
598

Fig. 78



67 V ▲ 67 F

70 T ▲ 70 S

80 M ▲ 80 L

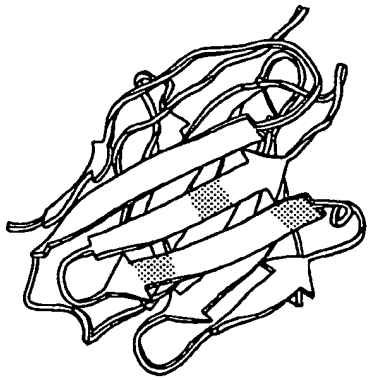


Fig. 79

GAGGTGCAGCTGGTGGAGTCTGGGGGCGGCTTGGCAAAGCCTGGGGGGTCCCTGAGACTCT
CCTGCGCAGCCTCCGGGTTCAGGTTCACCTTCAATAACTACTACATGGACTGGGTCCGCCA
GGCTCCAGGGCAGGGGCTGGAGTGGGTCTCACGTATTAGTAGTAGTGGTGATCCCACATGG
TACGCAGACTCCGTGAAGGGCAGATTCACCATCTCCAGAGAGAACGCCAAGAACACACTGT
TTCTTCAAATGAACAGCCTGAGAGCTGAGGACACGGCTGTCTATTACTGTGCGAGCTTGAC
TACAGGGTCTGACTCCTGGGGCCAGGGAGTCCTGGTCACCGTCTCCTCAGCTAGCACCAAG
GGCCCATCGGTCTTCCCCCTGGCACCCTCCTCCAAGAGCACCTCTGGGGGCACAGCGGCCC
TGGGCTGCCTGGTCAAGGACTACTTCCCCGAACCGGTGACGGTGTCGTGGAACTCAGGCGC
CCTGACCAGCGGCGTGCACACCTTCCCGGCTGTCCTACAGTCCTCAGGACTCTACTCCCTC
AGCAGCGTGGTGACCGTGCCCTCCAGCAGCTTGGGCACCCAGACCTACATCTGCAACGTGA
ATCACAAGCCCAGCAACACCAAGGTGGACAAGAAAGTTGAGCCCAAATCTTGTGACAAAAC
TCACACATGCCCACCGTGCCCAGCACCTGAACTCCTGGGGGGACCGTCAGTCTTCCTCTTC
CCCCCAAAACCCAAGGACACCCTCATGATCTCCCGGACCCCTGAGGTCACATGCGTGGTGG
TGGACGTGAGCCACGAAGACCCTGAGGTCAAGTTCAACTGGTACGTGGACGGCGTGGAGGT
GCATAATGCCAAGACAAAGCCGCGGGAGGAGCAGTACAACAGCACGTACCGTGTGGTCAGC
GTCCTCACCGTCCTGCACCAGGACTGGCTGAATGGCAAGGAGTACAAGTGCAAGGTCTCCA
ACAAAGCCCTCCCAGCCCCCATCGAGAAAACCATCTCCAAAGCCAAAGGGCAGCCCCGAGA
ACCACAGGTGTACACCCTGCCCCCATCCCGGGATGAGCTGACCAAGAACCAGGTCAGCCTG
ACCTGCCTGGTCAAAGGCTTCTATCCCAGCGACATCGCCGTGGAGTGGGAGAGCAATGGGC
AGCCGGAGAACAACTACAAGACCACGCCTCCCGTGCTGGACTCCGACGGCTCCTTCTTCCT
CTACAGCAAGCTCACCGTGGACAAGAGCAGGTGGCAGCAGGGGAACGTCTTCTCATGCTCC
GTGATGCATGAGGCTCTGCACAACCACTACACGCAGAAGAGCCTCTCCCTGTCTCCGGGTA
AATCCGGCGGGGGTGGATCCGGTGGAGGGGGCTCCGGCGGTGGCGGGTCCGACATCCAGAT
GACCCAGTCTCCATCTTCCCTGTCTGCATCTGTAGGGGACAGAGTCACCATCACTTGCAGG
GCAAGTCAGGACATTAGGTATTATTTAAATTGGTATCAGCAGAAACCAGGAAAAGCTCCTA
AGCTCCTGATCTATGTTGCATCCAGTTTGCAAAGTGGGGTCCCATCAAGGTTCAGCGGCAG
TGGATCTGGGACAGAGTTCACTCTCACCGTCAGCAGCCTGCAGCCTGAAGATTTTGCGACT
TATTACTGTCTACAGGTTTATAGTACCCCTCGGACGTTCGGCCAAGGGACCAAGGTGGAAA
TCAAAGGTGGGGGCGGATCTGGGGGCGGCGGGTCCGGTGGTGGTGGTAGTGAGGTGCAGCT
GGTGGAGTCTGGGGGCGGCTTGGCAAAGCCTGGGGGGTCCCTGAGACTCTCCTGCGCAGCC
TCCGGGTTCAGGTTCACCTTCAATAACTACTACATGGACTGGGTCCGCCAGGCTCCAGGGC
AGGGGCTGGAGTGGGTCTCACGTATTAGTAGTAGTGGTGATCCCACATGGTACGCAGACTC
CGTGAAGGGCAGATTCACCATCTCCAGAGAGAACGCCAAGAACACACTGTTTCTTCAAATG
AACAGCCTGAGAGCTGAGGACACGGCTGTCTATTACTGTGCGAGCTTGACTACAGGGTCTG
ACTCCTGGGGCCAGGGAGTCCTGGTCACCGTCTCCTCATGA (SEQ ID NO:63)

Fig. 80

EVQLVES GGGLAKPGGSLRLS CAAS GFRFTFNNYY MDWV RQA PGQGLEW VSRI S
SSGDPTW YADS VKGRFTISRE NAKN TLFLQMNSLR AEDT AVY YCASLTT GSDS W
GQGVLVT VSSA STKGPSVFPL APSS KSTSGGTAAL GCLV KDY FPEPVTV SWNS G
ALTSGVH TFPA VLQSSGLYSL SSVV TVP SSSLGTQ TYI C NVN HKPSNTK VDKK V
EPKSCDK THTC PPC PAPELLG GPSV FLF PPKPKDT LMIS RTP EVTCVVV DVSH E
DPEVKFN WYVD GVE VHNAKTK PREE QYN STYRVVS VLTV LHQ DWLNGKE YKCK V
SNKALPA PIEKTIS KAKGQPR EPQV YTL PPSRDEL TKNQ VSLTCLVKGF YPSD I
AVEWESN GQPE NNY KTTPPVL DSDG SFF LYSKLTV DKSR WQQ GNVFSCS VMHE A
LHNHYTQ KSLS LSPGKSGGGG SGGG GSGGGGSDIQ MTQS PSS LSASVGD RVTI T
CRASQDI RYYLNWY QQKPGKA PKLL IYV ASSLQSG VPSR FSG SGSGTEF TLTV S
SLQPEDF ATYY CLQVYSTPRT FGQG TKV EIKGGGG SGGG GSGGGGSEVQ LVES G
GGLAKPG GSLR LSCAASGFRF TFNN YYMDWVRQAP GQGL EWV SRISSSG DPTW Y
ADSVKGR FTIS RENAKNTLFL QMNS LRA EDTAVYY CASL TTG SDSWGQG VLVT V
SS* (SEQ ID NO:64)

Fig. 81

GACATCC AGATGACCCAGTCT CCATCTTCCCTGTC TGCATCTGTAGGGG AC
AGAGTCA CCATCACTTGCAGG GCAAGTCAGGACAT TAGGTATTATTTAA AT
TGGTATC AGCAGAAACCAGGA AAAGCTCCTAAGCT CCTGATCTATGTTG CA
TCCAGTT TGCAAAGTGGGGTC CCATCAAGGTTCAG CGGCAGTGGATCTG GG
ACAGAGT TCACTCTCACCGTC AGCAGCCTGCAGCC TGAAGATTTTGCGA CT
TATTACT GTCTACAGGTTTAT AGTACCCCTCGGAC GTTCGGCCAAGGGA CC
AAGGTGG AAATCAAACGTACG GTGGCTGCACCATC TGTCTTCATCTTCC CG
CCATCTG ATGAGCAGTTGAAA TCTGGAACTGCCTC TGTTGTGTGCCTGC TG
AATAACT TCTATCCCAGAGAG GCCAAAGTACAGTG GAAGGTGGATAACG CC
CTCCAAT CGGGTAACTCCCAG GAGAGTGTCACAGA GCAGGACAGCAAGG AC
AGCACCT ACAGCCTCAGCAGC ACCCTGACGCTGAG CAAAGCAGACTACG AG
AAACACA AAGTCTACGCCTGC GAAGTCACCCATCA GGGCCTGAGCTCGC CC
GTCACAA AGAGCTTCAACAGG GGAGAGTGTTGA (SEQ ID NO:65)

Fig. 82A

DIQMTQS PSSLSASVGDRVTI TCRASQDIRYYLNW YQQKPGKAPKLLIY VA
SSLQSGV PSRFSGSGSGTEFT LTVSSLQPEDFATY YCLQVYSTPRTFGQ GT
KVEIKRT VAAPSVFIFPPSDE QLKSGTASVVCLLN NFYPREAKVQWKVD NA
LQSGNSQ ESVTEQDSKDSTYS LSSTLTLSKADYEK HKVYACEVTHQGLS SP
VTKSFNR GEC (SEQ ID NO:66)

Fig. 82B

GACATCC AGAT GAC CCAGTCT CCAT CTT CCCTGTC TGCA TCTGTAGGGG AC
AGAGTCA CCAT CAC TTGCAGG GCAA GTCAGGACAT TAGG TAT TATTTAA AT
TGGTATC AGCA GAAACCAGGA AAAG CTC CTAAGCT CCTGATCTATGTTG CA
TCCAGTT TGCA AAG TGGGGTC CCAT CAAGGTTCAG CGGCAGTGGATCTG GG
ACAGAGT TCAC TCT CACCGTC AGCA GCC TGCAGCC TGAA GAT TTTGCGA CT
TATTACT GTCT ACA GGTTTAT AGTA CCC CTCGGAC GTTC GGC CAAGGGA CC
AAGGTGG AAAT CAAAGGCGGT GGCG GGT CCGGTGG GGGT GGC TCCGGGG GC
GGTGGCT CCGA GGT GCAGCTG GTGG AGT CTGGGGG CGGC TTGGCAAAGC CT
GGGGGGT CCCT GAGACTCTCC TGCG CAG CCTCCGG GTTCAGGTTCACCT TC
AATAACT ACTA CAT GGACTGG GTCC GCCAGGCTCC AGGG CAGGGGCTGG AG
TGGGTCT CACG TAT TAGTAGT AGTG GTGATCCCAC ATGG TACGCAGACT CC
GTGAAGG GCAG ATT CACCATC TCCA GAGAGAACGC CAAG AACACACTGT TT
CTTCAAA TGAA CAG CCTGAGA GCTG AGGACACGGC TGTC TATTACTGTG CG
AGCTTGA CTAC AGG GTCTGAC TCCT GGGGCCAGGG AGTC CTGGTCACCG TC
TCCTCAT GA (SEQ ID NO:67)

Fig. 83A

DIQMTQS PSSL SAS VGDRVTI TCRA SQD IRYYLNW YQQK PGKAPKLLIY V
ASSLQSG VPSR FSG SGSGTEF TLTV SSLQPEDFAT YYCL QVY STPRTFG Q
GTKVEIK GGGG SGG GGSGGGG SEVQ LVE SGGGLAK PGGS LRL SCAASGF R
FTFNNYY MDWV RQA PGQGLEW VSRI SSS GDPTWYA DSVK GRF TISRENA K
NTLFLQM NSLR AED TAVYYCA SLTT GSD SWGQGVL VTVS S*
(SEQ ID NO:68)

Fig. 83B

GAGGTGC AGCTGGTGGAGTCT GGGGGCGGCTTGGC AAAGCCTGGGGGGT CCCT
GAGACTC TCCTGCGCAGCCTC CGGGTTCAGGTTCA CCTTCAATAACTAC TACA
TGGACTG GGTCCGCCAGGCTC CAGGGCAGGGGCTG GAGTGGGTCTCACG TATT
AGTAGTA GTGGTGATCCCACA TGGTACGCAGACTC CGTGAAGGGCAGAT TCAC
CATCTCC AGAGAGAACGCCAA GAACACACTGTTTC TTCAAATGAACAGC CTGA
GAGCTGA GGACACGGCTGTCT ATTACTGTGCGAGC TTGACTACAGGGTC TGAC
TCCTGGG GCCAGGGAGTCCTG GTCACCGTCTCCTC AGGCGGTGGCGGGT CCGG
TGGGGGT GGCTCCGGGGGCGG TGGCTCCGACATCC AGATGACCCAGTCT CCAT
CTTCCCT GTCTGCATCTGTAG GGGACAGAGTCACC ATCACTTGCAGGGC AAGT
CAGGACA TTAGGTATTATTTA AATTGGTATCAGCA GAAACCAGGAAAAG CTCC
TAAGCTC CTGATCTATGTTGC ATCCAGTTTGCAAA GTGGGGTCCCATCA AGGT
TCAGCGG CAGTGGATCTGGGA CAGAGTTCACTCTC ACCGTCAGCAGCCT GCAG
CCTGAAG ATTTTGCGACTTAT TACTGTCTACAGGT TTATAGTACCCCTC GGAC
GTTCGGC CAAGGGACCAAGGT GGAAATCAAATGA (SEQ ID NO:69)

Fig. 84A

EVQLVES GGGLAKPGGSLRLS CAASGFRFTFNNYY MDWVRQAPGQGLEW VSRI S
SSGDPTW YADSVKGRFTISRE NAKNTLFLQMNSLR AEDTAVYYCASLTT GSDSW
GQGVLVT VSSGGGSGGGGSG GGGSDIQMTQSPSS LSASVGDRVTITCR ASQDI
RYYLNWY QQKPGKAPKLLIYV ASSLQSGVPSRFSG SGSGTEFTLTVSSL QPEDF
ATYYCLQ VYSTPRTFGQGTKV EIK*
(SEQ ID NO:70)

Fig. 84B

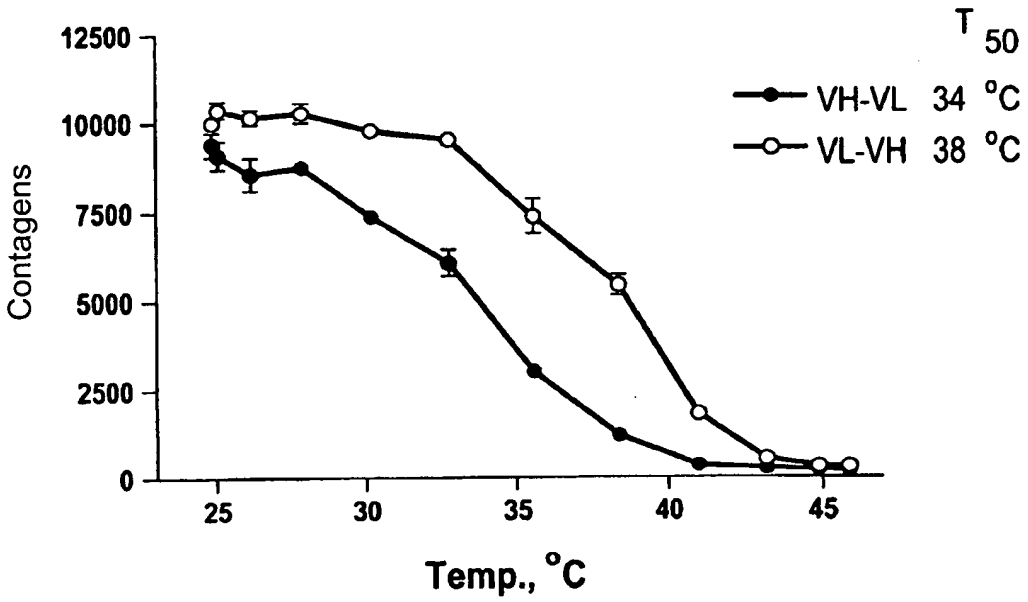

T 50
VH-VL 34 °C
VL-VH 38 °C
12500
10000
7500
5000
2500
0
Contagens
25
30
35
40
45
Temp., °C

Fig. 85

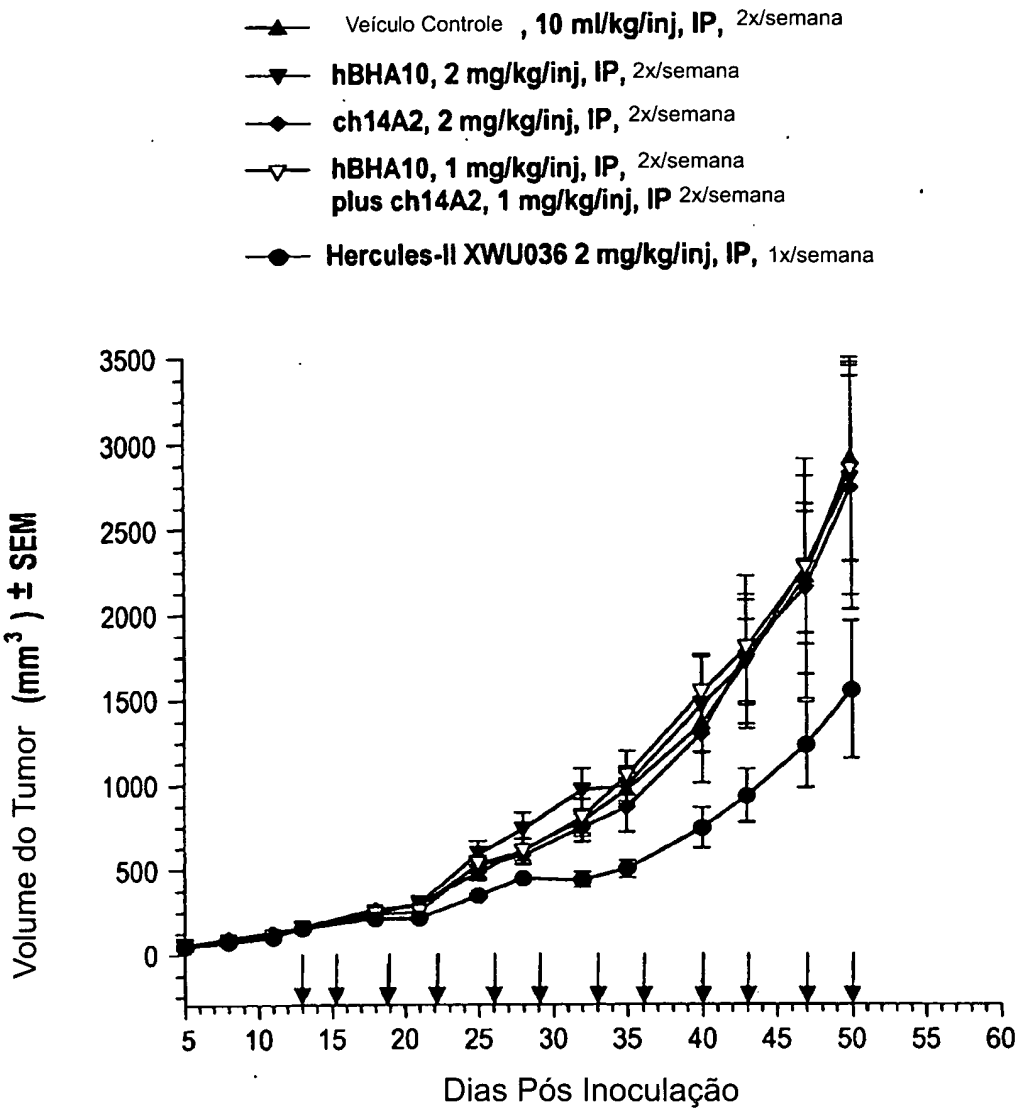

Veículo Controle , 10 ml/kg/inj, IP, 2x/semana
hBHA10, 2 mg/kg/inj, IP, 2x/semana
ch14A2, 2 mg/kg/inj, IP, 2x/semana
hBHA10, 1 mg/kg/inj, IP, 2x/semana
plus ch14A2, 1 mg/kg/inj, IP 2x/semana
Hercules-II XWU036 2 mg/kg/inj, IP, 1x/semana
Volume do Tumor (mm3) ± SEM
3500
3000
2500
2000
1500
1000
500
0
5
10
15
20
25
30
35
40
45
50
55
60
Dias Pós Inoculação

Fig. 86

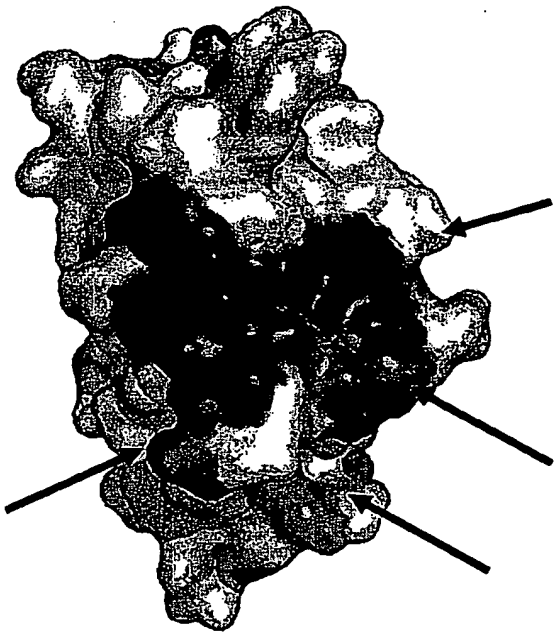


Fig. 87A

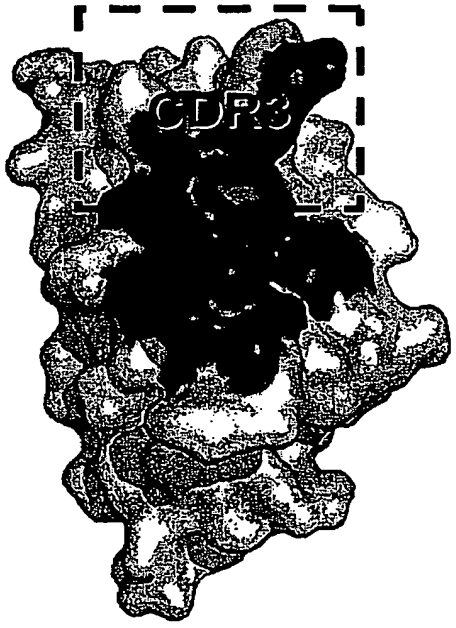


Fig. 87B

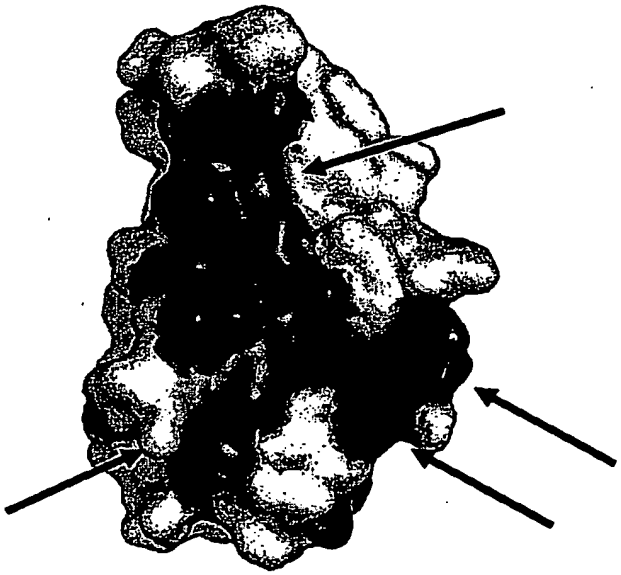


Fig. 88A

Fig. 88B

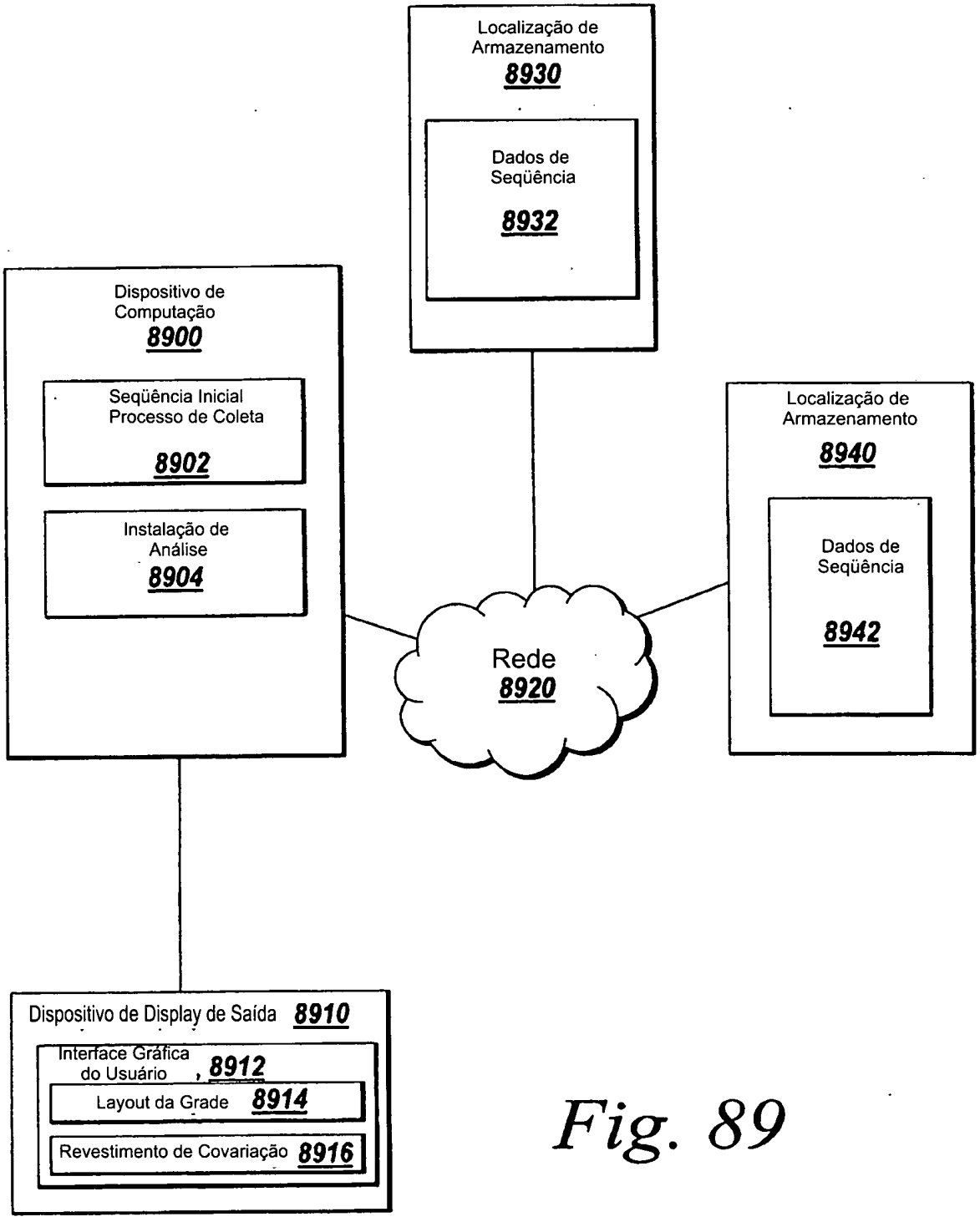

Localização de Armazenamento
8930
Dados de Seqüência
8932
Dispositivo de Computação
8900
Seqüência Inicial Processo de Coleta
8902
Instalação de Análise
8904
Rede
8920
Localização de Armazenamento
8940
Dados de Seqüência
8942
Dispositivo de Display de Saída 8910
Interface Gráfica do Usuário 8912
Layout da Grade 8914
Revestimento de Covariação 8916

Fig. 89

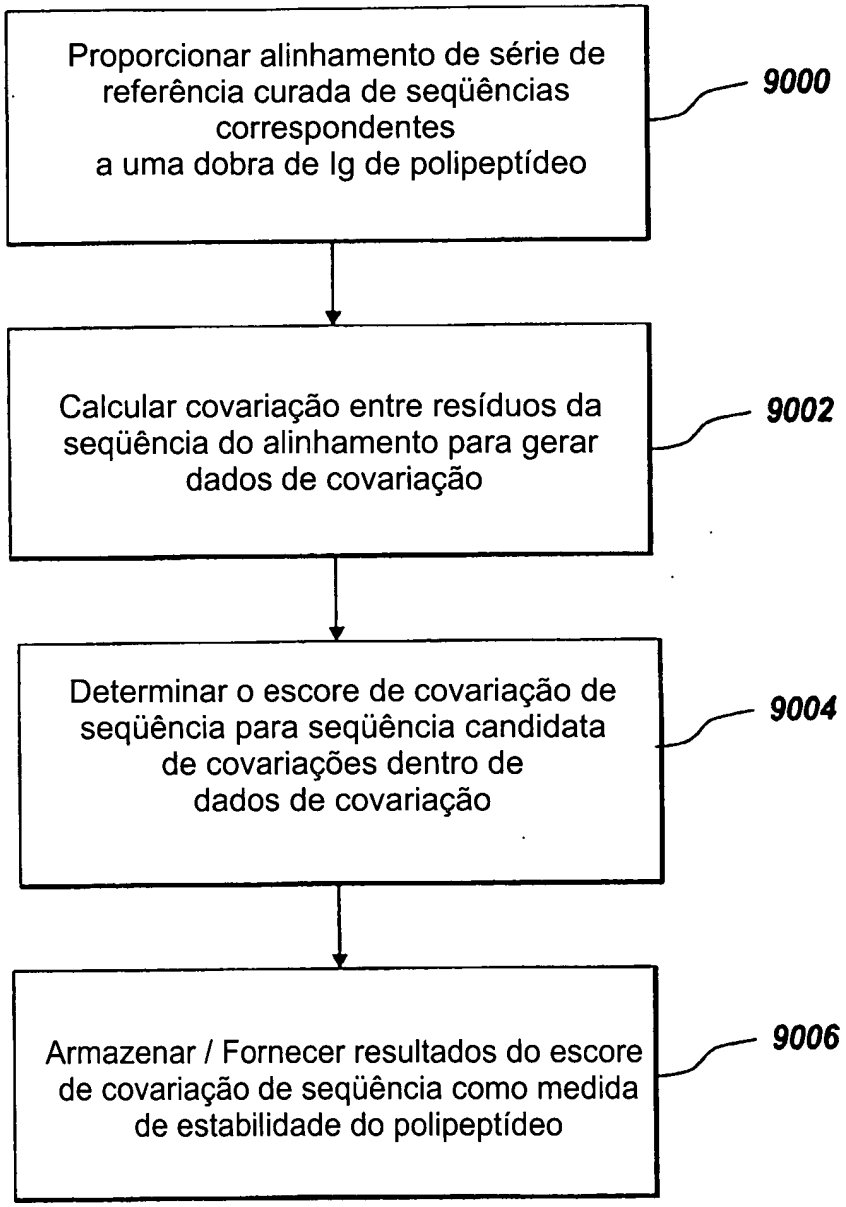

Proporcionar alinhamento de série de referência curada de seqüências correspondentes a uma dobra de Ig de polipeptídeo
9000
Calcular covariação entre resíduos da seqüência do alinhamento para gerar dados de covariação
9002
Determinar o escore de covariação de seqüência para seqüência candidata de covariações dentro de dados de covariação
9004
Armazenar / Fornecer resultados do escore de covariação de seqüência como medida de estabilidade do polipeptídeo
9006

Fig. 90

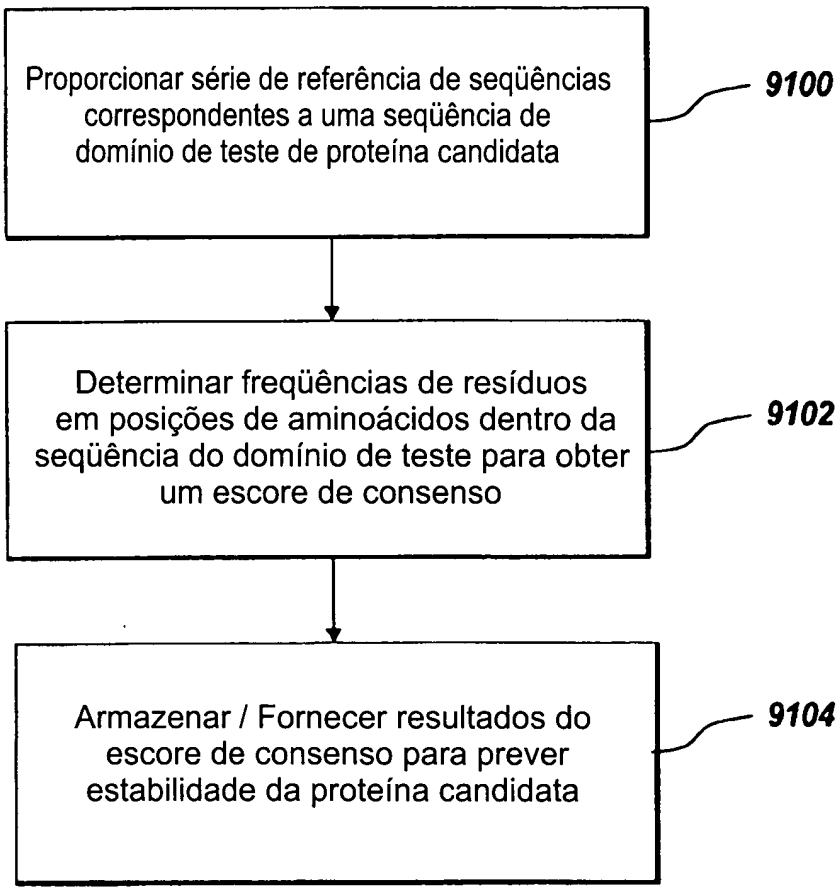

Proporcionar série de referência de seqüências correspondentes a uma seqüência de domínio de teste de proteína candidata
9100
Determinar freqüências de resíduos em posições de aminoácidos dentro da seqüência do domínio de teste para obter um escore de consenso
9102
Armazenar / Fornecer resultados do escore de consenso para prever estabilidade da proteína candidata
9104

Fig. 91

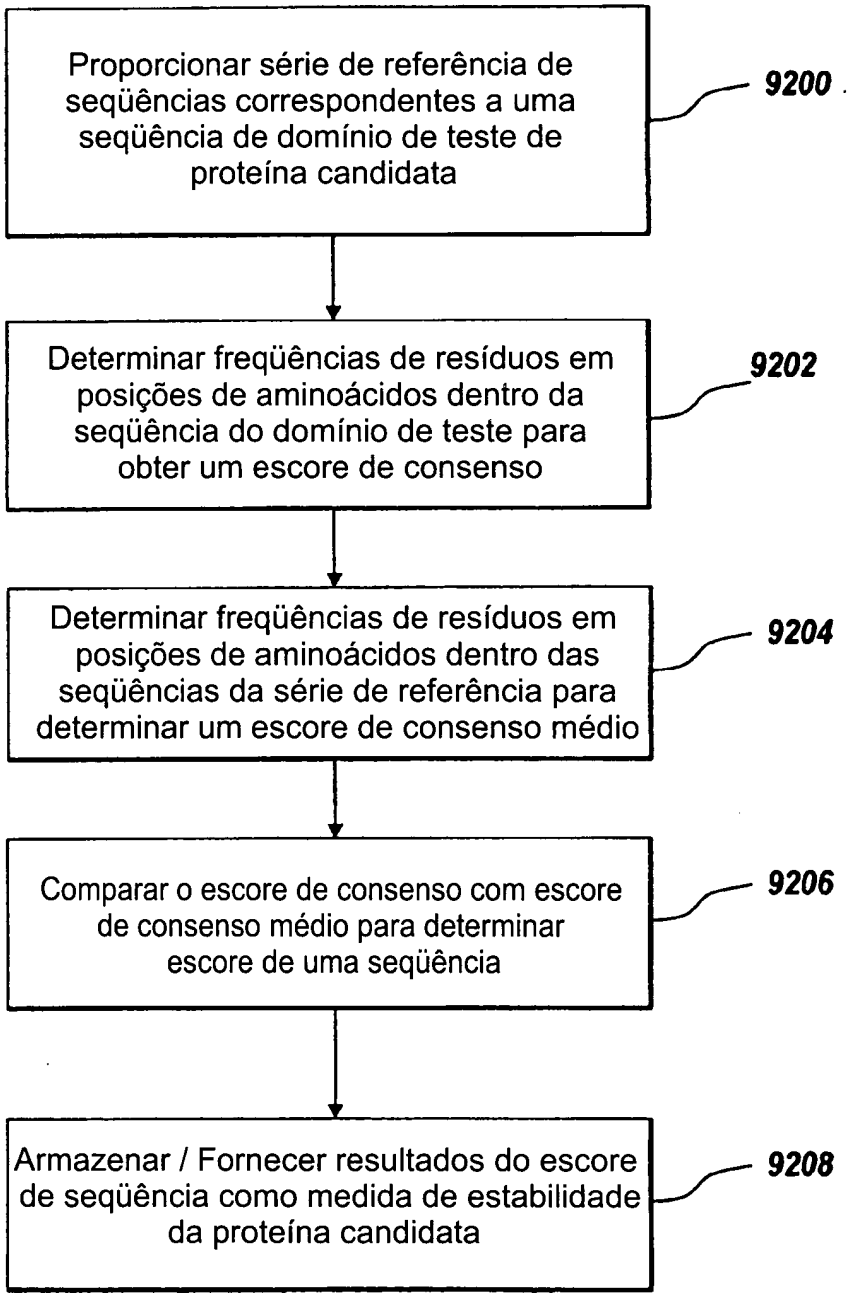

Proporcionar série de referência de seqüências correspondentes a uma seqüência de domínio de teste de proteína candidata
9200
Determinar freqüências de resíduos em posições de aminoácidos dentro da seqüência do domínio de teste para obter um escore de consenso
9202
Determinar freqüências de resíduos em posições de aminoácidos dentro das seqüências da série de referência para determinar um escore de consenso médio
9204
Comparar o escore de consenso com escore de consenso médio para determinar escore de uma seqüência
9206
Armazenar / Fornecer resultados do escore de seqüência como medida de estabilidade da proteína candidata
9208

Fig. 92

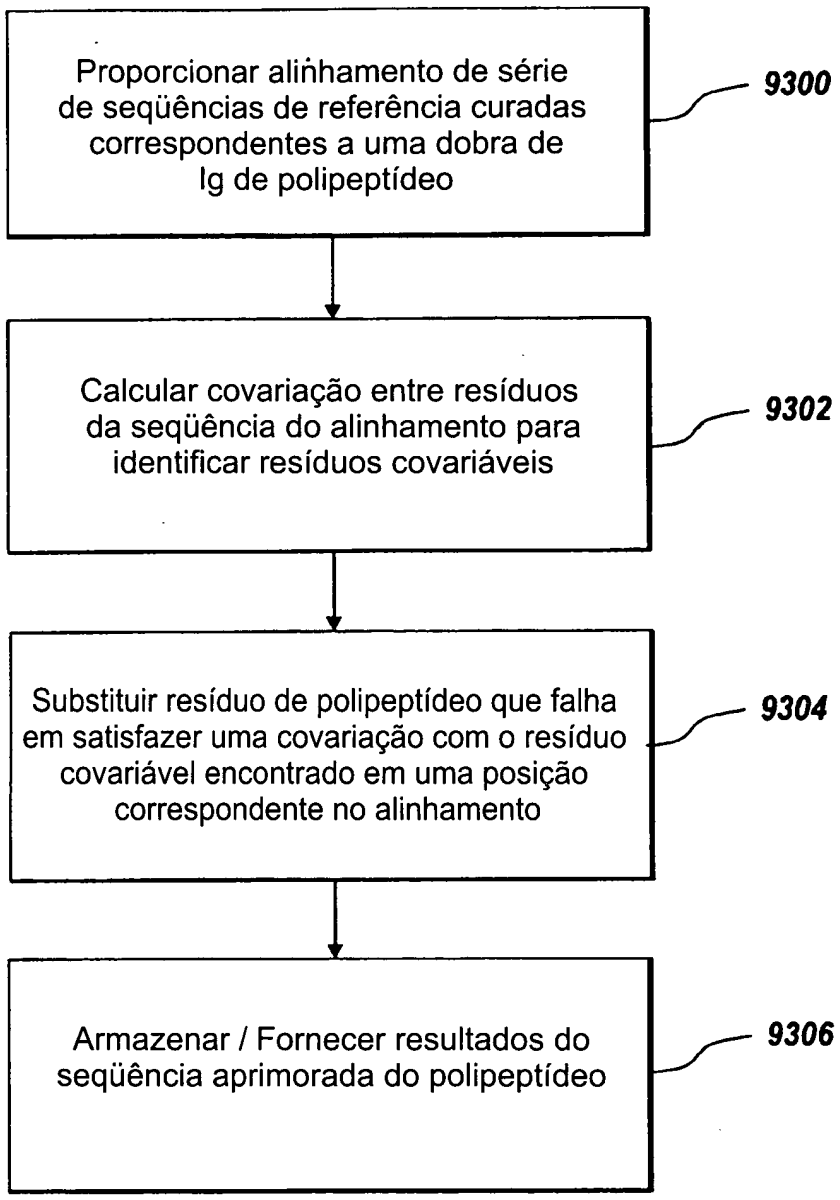

Proporcionar alinhamento de série de seqüências de referência curadas correspondentes a uma dobra de Ig de polipeptídeo
9300
Calcular covariação entre resíduos da seqüência do alinhamento para identificar resíduos covariáveis
9302
Substituir resíduo de polipeptídeo que falha em satisfazer uma covariação com o resíduo covariável encontrado em uma posição correspondente no alinhamento
9304
Armazenar / Fornecer resultados do seqüência aprimorada do polipeptídeo
9306

Fig. 93

A

MKKLLFAIPLVVPFYSHSQVQLVQSGAEVKKPGSSVKVSCKASGYTFT
TYYLHWVRQAPGQGLEWMGWIYPGNVHAQYNEKFKGRVTITADKSTST
AYMELSSLRSEDTAVYYCARSWEGFPYWGQGTTVTVSS**GGGGSGGGGS**
GGGGSGGGGSDIQMTQSPSSLSASVGDRVTITCKASQNVGINVAWYQQ
KPGKAPKSLISSASYRYSGVPSRFSGSGSGTDFTLTISSLQPEDFATY
FCQQYDTYPFTFGQGTKVEIK*DDDDKSFLEQKLISEEDLNSAVDHHHH*
*HH** **(SEQ ID NO: 3)**

B

MKKLLFAIPLVVPFYSHSQVQLVQSGAEVKKPGSSVKVSCKASGYTFT
TYYLHWVRQAPGQGLEWMGWIYPGNGHAQYNEKFKGRVTITADKSTST
AYMELSSLRSEDTAVYYCARSWEGFPYWGQGTTVTVSS**GGGGSGGGGS**
GGGGSGGGGSDIQMTQSPSSLSASVGDRVTITCKASQNVGINVAWYQQ
KPGKAPKSLISSASYRYSGVPSRFSGSGSGTDFTLTISSLQPEDFATY
FCQQYDTYPFTFGQGTKVEIK*DDDDKSFLEQKLISEEDLNSAVDHHHH*
*HH** **(SEQ ID NO: 4)**

Fig. 94

PI0709598-8

RESUMO

Patente de Invenção: "**COMPOSIÇÕES DE POLIPEPTÍDEOS ESTABILIZADOS**".

A presente invenção refere-se, no mínimo em parte, ao desenvolvimento de moléculas de ligação estabilizadas que consistem em ou compreendem uma scFv estabilizada e métodos para preparar similares moléculas estabilizadas.