DESCRIÇÃO

DA

PATENTE DE INVENÇÃO

N.º 97 612

REQUERENTE: SCRIPPS CLINIC AND RESEARCH FOUNDATION, norte-americana, com sede em 10666 North Torrey Pines Road, La Jolla, Califórnia 92037, Estados Unidos da América

EPÍGRAFE: "Processo de obtenção de proteínas ligantes de metais"

INVENTORES: Richard A. Lerner, Victoria N. Roberts, Elizabeth D. Getzoff, John A. Tainer e Stephen J. Benkovic

Reivindicação do direito de prioridade ao abrigo do artigo 4.º da Convenção de Paris de 20 de Março de 1883.

Estados Unidos da América em 8 de Maio de 1990 sob o nº 521 258; e em 18 de Junho de 1990 sob o nº 539 980

72 590
SCRF Case 206.1

Patente Nº 97 612

"Processo de obtenção de proteínas ligantes de metais"

para que

SCRIPPS CLINIC AND RESEARCH FOUNDATION, pretende obter privilégio de invenção em Portugal.

R E S U M O

O presente invento refere-se ao processo de obtenção de uma proteína de ligação a metais, capaz de formar um complexo de coordenação com catião metálico.

O processo compreende:

a) obter um gene codificando uma sequência de resíduos de aminoácidos que define um domínio variável de uma molécula de imunoglobulina;

b) determinar a sequência do gene para localizar as regiões que codificam as posições dos resíduos de aminoácidos adequadas para introduzir uma mutação do nucleótido para codificar um resíduo de aminoácido de contacto e produzir um local de ligação a metal;

c) modificar a sequência base de nucleótidos do gene, de modo a codificar resíduos de contacto metal-ligando em três das quatro posições dos resíduos de local de contacto metal-ligando, do domínio variável;

d) inserir a sequência base de nucleótidos modificados num vector de expressão;

e) expressar a proteína de ligação de metal a partir do gene modificado usando o vector de expressão num hospedeiro adequado; e

f) isolar a prroteína expressa do meio de expressão.

MEMÓRIA DESCRITIVA

Remissão a um pedido de patente relacionada

Esta é uma continuação em parte do pedido copendente de patente Nº de série 07/521 258, depositado em 8 de Maio de 1990. As suas descrições estão incorporadas no presente invento para referência.

Campo Técnico

O presente invento refere-se a proteínas ligantes de metais, particularmente a moléculas de anti-corpos que se ligam a catiões metálicos, na forma de metalo-anticorpos.

Antecedentes

As metaloproteínas, como uma classe, são proteínas possuindo um ião metálico complexado com a molécula da proteína no local de ligação de metal da proteína. O ião metálico contribui para a função da proteína através de uma variedade de mecanismos químicos, incluindo a estabilização da estrutura da proteína, facilitação da transferência electrónica em reacções de oxidação ou redução, e similares.

A estereoquímica da estrutura do complexo do ião metálico em associação com a proteína tem sido extensivamente caracterizada. Os estudos de metaloproteínas conhecidas resultaram na caracterização de muitos iões metálicos que participam em complexos metal-proteína, os quais ajudam a identificar a natureza do complexo metal-proteína. Estão disponíveis estruturas tridimensionais de metaloproteínas baseadas em dados cristalográficos de raio X para numerosas metaloproteínas e proporcionam um conhecimento adicional sobre a natureza do complexo metal-proteína.

Freeman et al., Adv. Protein Chem., 22: 257-424 (1967), descreveram a preparação de locais de ligação a metal, sintéticos, em polipéptidos. Descreveram-se alterações de metaloproteínas existentes nas quais os resíduos de aminoácido que participam no complexo metal-proteína foram substituídos, resultando em variações na especificidade do ião metálico da

apoproteína e na constante de ligação para o ião metálico. No entanto, nenhuma estrutura de proteína foi suficientemente caracterizada para permitir que a engenharia de proteínas reprodutíveis formasse um local de ligação ao metal numa proteína para proporcionar os benefícios de um cofactor metálico a uma proteína.

Breve Sumário do Invento

Verificou-se agora que as moléculas de imunoglobulina contêm regiões estruturais conservadas adjacentes ao local de combinação de anticorpo as quais podem ser modificadas reprodutivamente para formar um local de ligação ao metal. A incorporação de resíduos aminoácido seleccionados numa região estrutural conservada da sequência primária de aminoácidos da proteína proporciona os locais de contacto metal-ligando, necessários para formar um local de ligação ao metal na molécula de imunoglobulina, produzindo assim uma proteína ligante de metais.

Assim, numa concretização, o presente invento contempla uma proteína capaz de formar um complexo de coordenação com um catião metálico, i.e., uma metaloproteína. A metaloproteína caracteriza-se por 1) uma sequência de resíduos de aminoácido que define um domínio variável de uma imunoglobulina e 2) três resíduos de aminoácido de contacto no domínio variável que definem um local de ligação ao metal.

Nas concretizações preferidas, o presente invento contempla uma região variável de cadeia leve de uma imunoglobulina possuindo os três resíduos de aminoácido de contacto que formam um local de ligação ao metal.

Em outras concretizações preferidas, o presente invento contempla uma região variável de cadeia pesada de uma imunoglobulina possuindo três resíduos de aminoácido de contacto que formam um local de ligação ao metal.

Na concretização preferida, a metaloproteína está operativamente ligada a uma sequência de resíduos de aminoácido formando

uma região variável de cadeia pesada de imunoglobulina para formar um metalo-anticorpo. As moléculas de metalo-anticorpos preferidas são as moléculas Fv, Fab ou $F(ab')_2$, que ligam o Zn(II) e que são capazes de clivar cataliticamente uma ligação peptídica num local específico (sequência específica de resíduos de aminoácido).

As metaloproteínas do presente invento podem-se utilizar para a partição de catiões metálicos em solução por mistura de metaloproteínas com a solução contendo o catião metálico. Mantém-se então a mistura reaccional de ligação do catião metálico assim formada, durante um período de tempo predeterminado suficiente para a metaloproteína se ligar aos catiões metálicos e formar um complexo.

Breve Descrição das Figuras

Nas figuras, formando uma parte desta descrição:

A figura 1 ilustra a sequência de nucleótidos de um ADN que codifica uma proteína de ligação a antigénios de cadeia simples, possuindo um local de ligação ao metal, mostrada da esquerda para a direita e na direcção da extremidade 5' para a extremidade 3' utilizando o código de bases dos nucleótidos de uma única letra. O gene estrutural para a proteína de ligação ao antigénio madura começa na base 1 e termina na base 789, com o número da posição do último resíduo de base em cada fila indicado à direita da fila que mostra a sequência. O aminoácido para a proteína de ligação ao antigénio de cadeia simples está indicado pelo código de letra única acima da sequência de bases dos nucleótidos, com o número da posição para o primeiro resíduo em cada fila indicado à esquerda da fila que mostra a sequência dos resíduos de aminoácido. A estrutura de leitura está indicada pela posição da sequência de resíduos de aminoácido deduzida acima da sequência de nucleótidos de modo que a letra que representa cada aminoácido está localizada acima da base do meio no codão correspondente. A sequência de resíduos de aminoácido da proteína de ligação ao antigénio, madura, começa no resíduo 1 e termina no resíduo 263.

A figura 2 ilustra o vector de expressão plasmídico, pQM212, o qual contêm um gene estrutural que codifica a sequência de resíduos de aminoácido mostrada na figura 1 para a proteína de ligação ao antigénio, QM212, e possui a capacidade de dirigir a expressão de QM212. O vector contém também um gene de resistência à ampicilina (ampR) e a origem de replicação M13 (M13 Ori) e um promotor lambda (OL/PR) que proporciona um local de ligação para o repressor lambda sensível à temperatura para a expressão do gene QM212.

A figura 3 ilustra a capacidade do Cu(II) para interromper a fluorescência de resíduos de triptofano presentes na solução da proteína QM212. Os dados ilustram também que o Cu(II) não interrompe substancialmente a fluorescência associada à proteína 4-4-20/212. A interrupção da fluorescência está expressa na forma de uma razão F/Fo, onde 1,0 é uma unidade arbitrária de 0% de interrupção na solução não possuindo metal adicionado (Fo) e F é emissão medida quando se adiciona metal. Os dados mostram que a interrupção da fluorescência com Cu(II) da QM212 apresenta comportamento de saturação acima de cerca de 12-15 mM de Cu(II) com uma constante de ligação aparente de 2×10^5 M^{-1}.

A figura 4 ilustra a inibição da ligação ao Cu(II) da QM212 utilizando os iões metálicos competidores, zinco na forma Zn(II) (figura 4a) e cádmio na forma Cd(II) (figura 4b). A interrupção da fluorescência está expressa como descrito na legenda da figura 3. Antes da adição de Cu(II) à mistura de interrupção, adicionaram-se várias concentrações de Zn(II) ou Cd(II) à mistura. Os dados mostram que tanto o cádmio como o zinco ligam a QM212 de uma maneira dependente da dose suficiente para bloquear o Cu(II) e inibir a ligação do Cu(II) à QM212.

A figura 5 ilustra o espectro da excitação resultante da mistura de QM212 com fluoresceína utilizando quantidades variando entre 0 e 7,0 de QM212. As curvas do espectro de excitação estão expressas na forma de intensidade de fluorescência emitida após a excitação numa gama de comprimentos de onda entre 400 e 520 nm. Mediu-se a emissão a 530 nm. Obteve-se uma banda de 5 nm de

espessura para as medições de excitação e de emissão. Os dados mostram que à medida que se adiciona mais QM212 à solução de fluoresceína, reduz-se a quantidade de excitação, interrompendo a excitação de fluorescência por cerca de 80% (máximo a máximo). À medida que a QM212 se liga à fluoresceína, o máximo de excitação desloca-se de cerca de 493 nm para cerca de 507 nm.

A figura 6 ilustra os efeitos do Cu(II) sobre o espectro de excitação da fluoresceína medido com QM212 ligada à fluoresceína. Adicionaram-se várias quantidades de Cu(II), variando entre 0 e 30 mM, à solução contendo fosfato de sódio 15 nM, pH 6,0, 7,0 nM de fluoresceína e 7,7 mM de QM212. Mediram-se os espectros de excitação como descrito anteriormente para a figura 5.

A figura 7A ilustra o local de ligação ao zinco da anidrase carbónica (CAB) e o local de ligação ao metal introduzido no local L1-L3 do anticorpo HyHEL5. A comparação dos ligandos do zinco do CAB por [Kannan et al., Proc. Natl. Acad. Sci. USA, 72: 51-55 (1975)], no painel esquerdo com os resíduos do local L1-L3 na estrutura do HyHEL5 por [Sheriff et al., Proc. Natl. Acad. Sci. USA, 84: 8075-8079 (1987)] no painel direito mostra a similaridade de cadeias ß anti-paralelas das ligações do hidrogénio, direcções da ligação C_{α}-C_{β} e disposição das cadeias laterais. Todas as outras estruturas de anticorpos conhecidas partilhavam o arranjo visto no HyHEL5 nesta região. As setas indicam a direcção da cadeia polipeptídica.

A figura 7B ilustra o local do zinco do L1-L3 geral em combinação com uma biblioteca de cadeias pesada para formar uma variedade de locais de cofactor de zinco no anticorpo Fv. O modelo do esqueleto C_{α} do Fv HyHEL5 com o local de ligação ao zinco no CAB estabelecida no V_L combina-se com esqueletos C_{α} sobrepostos de domínios V_H de anticorpos HyHEL5, HyHEL10, Newm, J539, HEd10, Kol, D1,3 e McPC603. Os átomos C_{α} e C_{β} dos resíduos de ligação ao zinco no CAB, 94, 96 e 119, foram sobrepostos nos átomos correspondentes dos resíduos V_L de HyHEL5, 89, 91 e 34, respectivamente, com um desvio RMS entre as duas estruturas de 0,57 Å para os seis átomos. Os três ligandos histidina colocaram

o ião zinco (mostrado por uma esfera) no fundo da cavidade de ligação entre L3 e H3.

Descrição Detalhada do Invento

A. Definições

Resíduo de aminoácido: Prefere-se que os resíduos de aminoácido descritos no presente invento estejam na forma isomérica "L". No entanto, os resíduos na forma isomérica "D" podem substituir qualquer resíduo de L-aminoácido, desde que seja retida a propriedade funcional desejada, pelo polipéptido. NH_2 refere-se ao grupo amino livre presente na extremidade amino de um polipéptido. COOH refere-se ao grupo carboxilo livre presente na extremidade carboxilo de um polipéptido. De acordo com a nomenclatura padrão de polipéptidos, como descrita em J. Biol. Chem., 243: 3552-59 81969), mostram-se na tabela de correspondência seguinte as abreviaturas para os resíduos de aminoácido:

TABELA DE CORRESPONDÊNCIA

SÍMBOLO		AMINOÁCIDO
1 Letra	3 Letras	
Y	Tyr	Tirosina
G	Gly	Glicina
F	Phe	Fenilalanina
M	Met	Metionina
A	Ala	Alanina
S	Ser	Serina
I	Ile	Isoleucina
L	Leu	Leucina
T	Thr	Treonina
V	Val	Valina
P	Pro	Prolina
K	Lys	Lisina
H	His	Histidina
Q	Gln	Glutamina
E	Glu	Ácido glutâmico
W	Trp	Triptofano
R	Arg	Arginina
D	Asp	Ácido aspártico
N	Asn	Asparagina
C	Cys	Cisteína

Deve-se notar que todas as sequências de resíduos de aminoácido estão representadas no presente invento por fórmulas cuja orientação da esquerda para a direita está na direcção convencional da extremidade amino para a extremidade carboxilo. Além disto, deve-se notar que um travessão no início ou no final de uma sequência de resíduos aminoácido indica uma ligação peptídica a uma sequência adicional de um ou mais resíduos de aminoácido ou a um grupo NH_2 amino-terminal ou a um grupo COOH carboxi-terminal.

Polipéptido: refere-se a uma série linear de resíduos de aminoácido ligados uns aos outros por ligações peptídicas entre o grupo alfa-amino e o grupo carboxilo de resíduos de aminoácido contíguos.

Proteína: refere-se a uma série linear de mais de 50 resíduos de aminoácido ligados uns aos outros como num polipéptido.

Polipéptido sintético: refere-se a uma cadeia, produzida quimicamente, de resíduos de aminoácido ligados conjuntamente por ligações peptídicas que está livre de proteínas de ocorrência natural e de seus fragmentos.

Molécula de anticorpo: a frase "molécula de anticorpo" nas suas várias formas gramaticais, quando utilizada no presente invento, contempla tanto uma molécula de imunoglobulina intacta como uma porção imunologicamente activa de uma molécula de imunoglobulina, i.e., moléculas que contêm um local de combinação ao anticorpo ou paratopo.

Local de combinação ao anticorpo: um "local de combinação ao anticorpo" é a porção estrutural de uma molécula de um anticorpo que liga especificamente (imunorreage com) o antigénio e é compreendida por regiões variáveis e hipervariáveis de ambas as cadeias, pesada e leve. O termo "imunorreage" nas suas várias formas significa ligações específicas entre uma molécula contendo uma posição de combinação ao anticorpo tal como uma molécula de anticorpo completa ou uma sua porção.

Anticorpo: o termo anticorpo, nas suas várias formas gramaticais refere-se a uma composição contendo moléculas de anticorpos.

Exemplos de moléculas de anticorpos são as moléculas de imunoglobulina intactas, as moléculas de imunoglobulina substancialmente intactas, as porções de uma molécula de imunoglobulina que contêm o paratopo incluindo as porções conhecidas na arte como Fab, Fab', $F(ab')_2$ e F(v), e as proteínas de ligação a antigénios de cadeia simples.

As porções Fab e $F(ab')_2$ de moléculas de anticorpos são preparadas pela reacção proteolítica de papaína e pepsina, respectivamente, com moléculas de anticorpos substancialmente intactas, por processos que são bem conhecidos. Ver por exemplo, Patente dos E.U. Nº 4 342 566 de Theofilopolous e Dixan. As porções da molécula de anticorpo Fab' são também bem conhecidas e são produzidas a partir de porções $F(ab')_2$ seguidas por redução das ligações de dissulfureto que ligam as duas porções de cadeia pesada, como com mercaptanoetanol e seguida pela alquilação do mercaptano de proteína resultante com um reagente tal como a iodo-acetamida.

Proteína de ligação a antigénios de cadeia simples: um polipéptido, também referido como uma molécula de local de combinação de anticorpos de cadeia simples, composta por uma sequência de aminoácidos de região variável de cadeia leve de imunoglobulina (VL) ligada a uma sequência de aminoácido de região de cadeia pesada de imunoglobulina (VH) por um péptido que liga a extremidade carboxilo da sequência VL à extremidade amino da sequência VH. Ver por exemplo, Bird et al., Science, 242: 423-426 (1988). (Os ensinamentos das referências citadas no presente invento incorporam-se para referência).

Condições de imunorreacção: as condições de imunorreacção são as que mantêm a actividade imunológica de um metalo-anticorpo deste invento. Estas condições incluem uma gama de temperaturas de cerca de 4 graus C (4°C) a cerca de 45°C, preferivelmente cerca de 37°C, uma gama de valores de pH de cerca de 5 a cerca de

9, preferivelmente cerca de 7 e uma força iónica variando entre a da água destilada e a de cloreto de sódio cerca de um molar, preferivelmente cerca da da solução salina fisiológica. Os processos de optimização destas condições são bem conhecidos na arte.

Metal: O termo "metal" no contexto de um local de ligação ao metal refere-se aos catiões metálicos descritos no presente invento que formam um complexo de coordenação com uma proteína ligante de metais.

Anticorpo monoclonal: A frase "anticorpo monoclonal", nas suas várias formas gramaticais, refere-se a um anticorpo contendo apenas uma espécie de local de combinação a anticorpos capaz de imunorreagir com um antigénio particular. Assim, um anticorpo monoclonal apresenta, tipicamente, uma afinidade de ligação única para qualquer antigénio, com o qual imunorreage. Um anticorpo monoclonal pode portanto conter uma molécula de anticorpo possuindo uma pluralidade de locais de combinação de anticorpo, cada um imuno-específico para um antigénio diferente, p. ex., um anticorpo monoclonal bis-específico.

Nucleótido: uma unidade monomérica de ADN ou de ARN consistindo numa porção açúcar (pentose), um fosfato, e uma base azotada heterocíclica. A base está ligada à porção açúcar por meio do carbono glicosídico (carbono 1' da pentose) e esta combinação de base e açúcar é um nucleósido. Quando o nucleósido contém um grupo fosfato ligado na posição 3' ou 5' da pentose, refere-se como um nucleótido. Uma sequência de nucleótidos operativamente ligados refere-se tipicamente, no presente invento, como uma "sequência de bases" ou "sequência de nucleótidos", e representa-se, no presente invento, por uma fórmula cuja orientação da esquerda para a direita está na direcção convencional da extremidade 5' para a extremidade 3'.

Polinucleótido: uma molécula de ácido nucleíco compreendendo uma unidade polimérica de ADN ou ARN possuindo uma sequência de dois ou mais nucleótidos operativamente ligados que formam uma cadeia linear simples de nucleótidos, também referida como um

oligonucleótido.

Duplex: uma molécula de ácido nucleico de cadeia dupla consistindo de duas cadeias de polinucleótidos complementares hibridados em conjunto pela formação de uma ligação de hidrogénio entre cada um dos nucleótidos complementares presentes num par de bases do duplex. Como os nucleótidos que formam um par de bases podem ser ou uma base ribonucleotídica ou uma base desoxi--ribonucleotídica, o termo "duplex" refere-se tanto a um duplex ADN-ADN compreendendo duas cadeias de ADN, como a um duplex ARN-ADN compreendendo uma cadeia de ADN e uma de ARN.

Par de bases (pb): uma associação de adenina (A) com timina (T), ou de citosina (C) com guanina (G) num duplex de ADN de cadeia dupla.

Ácido nucleíco: o termo refere-se a qualquer uma de uma classe de moléculas que inclui ácido ribonucleíco (RNA), ácido desoxi-ribonucleíco (ADN) nas suas formas de cadeia simples ou dupla, e polinucleótidos.

Segmento de ADN: uma duplex ADN-ADN possuindo uma sequência de nucleótidos conservada, pré-seleccionada e uma sequência que codifica uma proteína ligante de metais do presente invento.

B. Proteína Ligante de Metal

O presente invento contempla uma proteína que possui a capacidade de se complexar (ligar) com um catião metálico através de ligandos (contactos) proporcionados por resíduos de aminoácido de contacto na proteína, que são apresentados numa geometria que coordena a complexação de um catião metálico.

A proteína liga o catião metálico e portanto refere-se como uma proteína ligante de metais. Um catião metálico é ligado pela proteína contemplada porque a proteína é capaz de formar um complexo de coordenação com o catião metálico de uma maneira análoga às formações dos complexos de coordenação encontrados em outras metalo-proteínas bem conhecidas, tais como a anidrase

carbónica, a superóxido dismutase e similares.

Uma proteína ligante de metais do presente invento é compreendida por uma sequência de resíduos de aminoácido que definem um domínio variável de uma imunoglobulina.

Uma imunoglobulina representa uma vasta família de moléculas que inclui vários tipos de moléculas, tais como IgD, IgG, IgA, IgM e IgE. Uma imunoglobulina compreende, tipicamente, duas cadeias pesadas (H) e duas leves (L) com um domínio variável (V) e um constante (C) presentes em cada cadeia. Várias regiões diferentes de uma imunoglobulina contêm sequências conservadas com base em estudos comparativos de sequências conhecidas de cadeias pesadas ou leves. Foram compilados dados extensivos de sequências de ácidos nucleícos e de aminoácidos apresentando exemplos de sequências conservadas, para moléculas de imunoglobulina, por Kabat et al., em Sequences of Proteins of Immunological Interest, 4th Ed. Nacional Institutes of Health, Bethesda, MD (1987).

O presente conhecimento dos locais de uma molécula de anticorpo responsáveis pela ligação antigénio-anticorpo, indicam que parte do local de combinação do anticorpo é formado por regiões hiper-variáveis de cadeia pesada e parte do local de combinação é formado pelas regiões hiper-variáveis de cadeia leve. Ver por exemplo, Getzoff et al., Adv. Immunol., 43: 1-98 (1988). As regiões hiper-variáveis são compreendidas por seis voltas do polipéptido; 3 voltas do domínio variável da cadeia leve (VL) e três voltas do domínio variável da cadeia pesada (VH), denominadas por L1, L2, L3 e H1, H2, H3, respectivamente. Ver por exemplo, Chothia et al., Nature, 342: 877-883 (1989). As regiões hiper-variáveis são também conhecidas como regiões determinantes da complementaridade ou CDR.

Os estudos comparativos das estruturas tridimensionais conhecidas de numerosas moléculas de anticorpos determinaram que cada região hiper-variável adopta uma de algumas conformações de cadeia principal ou estruturas canónicas. Utilizando homologias

de sequência, a sequência de resíduos de aminoácido de uma cadeia leve ou pesada, pode-se alinhar com a sequência de uma imunoglobulina conhecida de estrutura de cadeia leve ou pesada, respectivamente, podendo-se identificar reprodutivelmente as voltas das regiões hiper-variáveis e os locais dos resíduos de aminoácido específicos dentro das voltas, da cadeia leve ou pesada.

Assim, por exemplo, uma estrutura de voltas de L1, L2 ou L3 (CDR L1, CDR L2 ou CDR L3) pode-se identificar reprodutivelmente apenas com base nas homologias de sequência com outras cadeias leves da imunoglobulina, localizando-se assim a posição das posições dos resíduos críticos dentro da estrutura da volta. Devido à existência de variação no comprimento da cadeia para uma cadeia leve de imunoglobulina, particular, os números de posição dos resíduos de aminoácido referem-se, no presente invento, por um esquema de numeração que se baseia nos alinhamentos utilizando sequências homólogas como descrito por Kabat et al., em Sequences of Proteins of Immunological Interest, 4th, ed., U.S. Department of Health e Human Services, National Institute of Health, Bethesda, MD (1987). Assim, o número de resíduo específico referido às posições do resíduos de aminoácido para resíduos particulares será citado no presente invento como "número de posição Kabat" ou "número de posição do resíduo de aminoácido de Kabat" para conotar uma posição de resíduo identificável reprodutivelmente numa estrutura de voltas CDR reconhecida. Onde quer que se dêem os números de posição, eles referem-se à posição de Kabat.

O domínio variável de qualquer molécula de cadeia pesada ou leve de imunoglobulina possuindo as estruturas das voltas identificáveis é útil no presente invento para produzir um local de ligação a metais. Um domínio variável representativo de cadeia leve no qual se introduziu um local de ligação a metais pelo presente invento é o domínio variável definido pela sequência de resíduos de aminoácido mostrada na figura 1 do resíduo 22 ao resíduo 132.

Uma proteína ligante de metais possuindo o domínio variável anterior é a molécula do local de combinação a anticorpos, de

cadeia simples, mostrada na figura 1 do resíduo 1 ao resíduo 263. A estrutura de proteínas de ligação a antigénios, de cadeia simples, i.e., as moléculas de cadeia simples possuindo um local de combinação a anticorpos ou moléculas de local de combinação a anticorpos, de cadeia simples, foram descritas por Bird et al., Science, 224: 423-426 (1988), e a sua preparação foi também descrita na Patente dos E.U. Nº 4 704 692 por Ladner.

Uma proteína ligante de metais inclui ainda um local de ligação a metais na proteína quando a proteína está numa conformação dobrada, tal que a proteína forme, pela sua estrutura dobrada, um local de ligação a metais e seja capaz de se complexar com (ligar-se) um catião metálico pré-seleccionado. Um local de ligação a metais, forma-se pelo posicionamento geométrico de três ligandos de metais (posições de contacto) proporcionados pelos resíduos da cadeia lateral de três resíduos de aminoácido de contacto para formar ligandos de coordenação para complexar um catião metálico. Assim, o posicionamento de três resíduos de aminoácido de contacto na sequência de aminoácidos de um domínio variável de imunoglobulina define um local de ligação a metais.

A estrutura e a estereoquímica das interacções proteína-metal em metaloproteínas é geralmente bem entendida. Ver por exemplo, Freeman et al., Adv. Protein Chem., 22: 257-424 (1967); Kannan et al., Annals. NY Acad. Sci., 429: 49 (1984); e Trainer et al., J. Mol. Biol., 160: 181-217 (1982).

Os ligandos de metais (locais de contacto) para a ligação de um catião metálico a um local de ligação de metais numa proteína deste invento estão situados em três localizações para proporcionar três pontos de contacto de ligandos tipicamente necessários para um complexo de coordenação de um catião metálico. As geometrias de complexos de coordenação representativas para os ligandos de metais podem ser tetraédrica, quadrangular plana ou trigonal. No entanto, as geometrias tetraédricas são preferidas. Mostram-se os complexos de coordenação de metais representativos com a geometria de coordenação tetraédrica preferida através da

estrutura do complexo de Zinco(II) na enzima superóxido dismutase ou do complexo de Cobre(II) na anidrase carbónica. Ver Trainer et al., J. Mol. Biol., 160: 181-217 (1982); e Kannan et al., Annals, NY Acad. Sci., 429: 49 (1984). Nestes exemplos, embora um catião metálico apresente quatro locais de contacto potenciais, tipicamente, três participam no contacto metal-ligando e o quarto local do catião metálico está livre para participar em trocas de electrões ou na partilha de electrões com o solvente ou o soluto em solução que têm acesso ao metal complexado. similarmente, na presente proteína ligante de metais, três resíduos de aminoácido de contacto (locais de contacto do ligando) participam na complexação do catião metálico.

Um resíduo de aminoácido que ocupa uma das três posições de resíduos de aminoácido para proporcionar um local de contacto metal-ligando, refere-se como um resíduo de aminoácido de contacto. Os resíduos de aminoácido adequados para utilização como resíduos de aminoácido de contacto são conhecidos na arte da bioquímica de metaloproteínas e incluem a histidina, o ácido aspártico, a cisteína, o ácido glutâmico e resíduos similares conhecidos por proporcionarem um ligando para catiões metálicos em metaloproteínas. Numa concretização preferida, os três resíduos de aminoácido são duas histidinas e um terceiro resíduo seleccionado entre o grupo histidina, ácido aspártico, cisteína e ácido glutâmico. O uso de resíduos histidina é particularmente preferido para todos os três resíduos de aminoácido de contacto.

Determinou-se pelo presente invento que as moléculas de cadeia pesada e leve de imunoglobulina contêm, cada uma, locais para o posicionamento de três resíduos de aminoácido de contacto adequados para produzir um local de ligação do ligando ao metal. Estes locais podem estar localizados reprodutivelmente de modo a apresentarem os contactos catião metálico-ligando nas coordenadas adequadas para a formação de um local de ligação ao metal porque a localização da posição do resíduo de aminoácido crítica para a formação de regiões hiper-variáveis de ligação ao metal tanto de cadeia pesada como leve, podem-se identificar reprodutivelmente na estrutura de voltas CDR devido às suas características conser-

vadas. Ver Getzoff et al., Adv. Immunol., 43: 1-98 (1988); e Chithia et al., Nature, 342: 877-883 (1989). Portanto, pode-se modificar qualquer molécula de cadeia pesada ou leve de imunoglobulina de modo a conter um local de ligação ao metal de acordo com os preceitos destes invento, uma vez que as posições dos resíduos de aminoácido de contacto se identificam pelo número de posição de Kabat como se descreveu anteriormente.

Pelo presente invento verificou-se quais as posições dos resíduos de aminoácido de contacto que definem o local de ligação ao metal. Podem-se, portanto, projectar resíduos de aminoácido de contacto para um domínio variável de uma molécula de cadeia pesada ou leve de imunoglobulina para formar um local de ligação ao metal e sumarizam-se as suas posições na tabela 1.

Assim, numa concretização, as posições dos resíduos de aminoácido usadas para proporcionar três locais de contacto metal-ligando localizam-se em posições específicas da sequência de aminoácidos que define o domínio variável da molécula de cadeia leve de imunoglobulina.

Numa concretização de cadeia leve, o domínio variável inclui uma região L1 e uma região L3 e os três resíduos de aminoácido de contacto localizam-se em quaisquer três das quatro posições de resíduos de aminoácido 32, 34, 89 e 91.

Assim, um local de ligação ao metal preferido inclui três resíduos de aminoácido de contacto ocupando as posições de resíduos de aminoácido 32, 34 e 89; as posições 32, 34 e 91; as posições 32, 89 e 91; ou 34, 89 e 91, de um domínio variável de cadeia leve de imunoglobulina. Um local de ligação ao metal preferido e exemplificativo é formado utilizando histidina como resíduo de aminoácido de contacto nas posições 34, 89 e 91 e está descrito em maior detalhe no Exemplo 1a e na figura 1. Na proteína de ligação ao metal exemplificativa, cuja sequência de resíduos de aminoácido completa se mostra na figura 1, os resíduos de aminoácido de contacto possuindo os números de posição de kabat 34, 89 e 91 localizam-se nas posições de

resíduos de aminoácido 60, 115 e 117, respectivamente, da proteína de ligação ao antigénio, de cadeia simples, (QM212) mostrada na figura 1.

Várias características do domínio variável contendo CDR L1 e CDR L3, da cadeia leve, proporcionam um ambiente óptimo para o posicionamento de um local de ligação ao metal na proteína reivindicada, descrita no presente invento. As regiões L1 e L3 formam estruturas de cadeias principais anti-paralelas ligadas com ligações de hidrogénio entre as cadeias em vários pontos. Esta característica proporciona uma estrutura de folha beta que é reprodutível e prontamente identificável em todos os domínios variáveis de cadeia leve. Proporciona também estabilização para facilitar as distâncias atómicas adequadas de modo a que os locais pré-seleccionados para os resíduos de aminoácido de contacto estejam adequadamente localizadas para formar um arranjo tetraédrico para apresentação dos ligandos para complexar o catião metálico.

As ligações de hidrogénio encontram-se tipicamente entre os átomos da cadeia principal do par de resíduos 89 e 34, ou do par 91 e 32, que estabilizam as estruturas da cadeia principal L3 e L1, respectivamente.

Na concretização que utiliza os resíduos de aminoácido de contacto nas posições 34, 89 e 91 da cadeia leve, uma proteína de ligação ao metal do presente invento, inclui preferivelmente glicina ou um resíduo de aminoácido não polar na posição de resíduo de aminoácido 36 da sequência da região variável de cadeia leve, onde o resíduo não polar é um resíduo de aminoácido possuindo menos átomos na cadeia lateral que a tirosina e não é prolina. Os resíduos preferidos são leucina, valina e alanina. Nas concretizações particularmente preferidas, uma proteína de ligação ao metal possui um resíduo leucina na posição 36 no domínio variável da cadeia leve.

Noutra concretização de cadeia leve, o domínio variável inclui uma região L3 e os três resíduos de aminoácido de contacto

estão localizados nas posições de resíduo de aminoácido 90, 92 e 97. As ligações de hidrogénio localizam-se preferivelmente entre os átomos da cadeia principal do par de resíduos 90 e 99 para estabilizar a estrutura da cadeia principal L3.

Outra concretização de cadeia leve inclui uma região L3 no domínio variável contendo os três resíduos de aminoácido de contacto nas posições de resíduos de aminoácidos 89, 91 e 96.

Uma concretização de cadeia leve relacionada inclui uma região L2 no domínio variável contendo os três resíduos de aminoácido de contacto nas posições de resíduos de aminoácido 50, 53 e 55.

Outra concretização de cadeia leve inclui uma região L1 e uma região L3 de um domínio variável de cadeia leve e contém três resíduos de aminoácido de contacto nas posições de resíduos de aminoácido 27d, 29 e 93. As posições 27d e 29 localizam-se na região L1. A posição de kabat 27d denota uma posição de resíduo de aminoácido que está a quatro resíduos de aminoácido de distância da posição 29 na direcção da extremidade amino da cadeia leve.

Noutra concretização, as posições de resíduos de aminoácido utilizadas para proporcionar três locais de contacto metal-ligando, localizam-se em posições específicas da sequência de aminoácidos que define o domínio variável da molécula de cadeia pesada de imunoglobulina.

Numa concretização de cadeia pesada, o domínio variável inclui uma região H1 e uma região H3, e os três resíduos de aminoácido de contacto localizam-se nas posições de resíduos de aminoácido 3, 35 e 95. As ligações de hidrogénio localizam-se preferivelmente entre os átomos da cadeia principal do par de resíduos 33 e 95 para estabilizar as estruturas da cadeia principal H1 e H3.

Noutra concretização de cadeia pesada, o domínio variável

inclui uma região H1 e uma região H2 e os três resíduos de aminoácido de contacto localizam-se em quaisquer três das quatro posições de resíduos de aminoácido 33, 35, 50 e 52.

Outra concretização de cadeia pesada inclui uma região H1 e uma região H2 no domínio variável contendo os três resíduos de aminoácido de contacto nas posições de resíduos de aminoácido 31, 33 e 52.

Uma concretização de cadeia pesada relacionada inclui uma região H2 no domínio variável contendo três resíduos de aminoácido de contacto em quaisquer três das quatro posições de resíduos de aminoácido 50, 52, 56 e 58. Os pares com ligações de hidrogénio localizam-se preferivelmente entre os átomos da cadeia principal do par de resíduos 50 e 58 para estabilizar a estrutura da cadeia principal H2.

Noutra concretização de cadeia pesada, o domínio variável inclui uma região H2 e os três resíduos de aminoácido de contacto localizam-se nas posições de resíduos de aminoácido 50, 58 e 60. Os pares com ligações de hidrogénio localizam-se preferivelmente entre os átomos da cadeia principal do par de resíduos 50 e 58 para estabilizar a estrutura da cadeia principal H2.

Outra concretização de cadeia pesada inclui uma região H3 no domínio variável contendo os três resíduos de aminoácido de contacto nas posições de resíduos de aminoácido 95, 101x e 101x-2. A designação de posição "101x" denota posições alternativas do número de posição de Kabat 101 ou da posição de resíduo de aminoácido que precede a posição 101, i.e., a posição afastada um resíduo da posição 101 e na direcção da extremidade amino da cadeia pesada. A designação de posição "101x-2" denota posições alternativas que dependem da posição indicado pelo termo "101x". Se 101x é a posição 101, então 101x-2 é uma posição afastada dois resíduos da posição 101 na direcção da extremidade amino da cadeia pesada. Se 101x é a posição que antecede a 101 por um resíduo, então 101x-2 é uma posição afastada três resíduos da posição 101 na direcção da extremidade amino da cadeia pesada.

C. Processos de Ligação de Catiões Metálicos

O presente invento contempla um processo de separação de um catião metálico de uma amostra fluída contendo o catião metálico. Este processo de separação do metal inclui os passos seguintes:

1. Mistura de uma amostra fluída contendo catiões metálicos, com uma proteína ligante de metais para formar uma mistura de ligação.

2. Manutenção da mistura de ligação durante um período de tempo suficiente para o catião metálico ligar a proteína de ligação a metais e formar um complexo catião metálico-proteína.

3. Remoção do complexo catião metálico-proteína da mistura de ligação e portanto separação do catião metálico da amostra fluída.

A quantidade de tempo necessária para o catião metálico ligar a proteína ligante de metais dependerá pelo menos da proteína de ligação a metais particular empregue, da constante de ligação exibida pela proteína ligante de metais e da concentração do metal. O complexo catião metálico-proteína forma-se quando pelo menos um catião metálico se associa com a proteína ligante de metais e fica ligado a essa proteína para formar um complexo catião metálico-proteína. Remove-se então o complexo catião metálico-proteína da mistura de ligação, separando assim o catião metálico da amostra fluída. Pode-se remover o complexo utilizando processos bem conhecidos dos peritos na arte que incluem remoção mecânica, filtração, sedimentação e outros meios de separação.

Noutras concretizações preferidas, fixa-se a proteína ligante de metais num suporte sólido tal como plástico, nitrocelulose, membranas à base de nylon, de placa microtitulação de alvéolos. Quando se fixa a proteína ligante de metais num suporte sólido, o complexo catião metálico-proteína formado é imobilizado no suporte sólido e faz parte de uma fase sólida.

Fixa-se tipicamente a proteína ligante de metais na matriz

sólida por adsorção a partir de um meio aquoso embora se possam utilizar outros modos de fixação, bem conhecidos dos peritos na arte.

As matrizes sólidas (suportes sólidos) úteis são também bem conhecidos na arte. Estes materiais são insolúveis em água e incluem o dextrano reticulado disponível sob a marca registada SEPHADEX da Pharmacia Fine Chemicals (Piscataway, NY); a agarose; pérolas de poliestireno, pérolas de cerca de 1 micron a cerca de 5 milímetros de diâmetro estão disponíveis em Abbott Laboratories of North Chicago, IL; poli(cloreto de vinilo), poliestireno, poliacrilamida reticulada, tecidos baseados em nitrocelulose ou em nylon tais como folhas, fitas ou pás; ou tubos, placas ou alvéolos de uma placa de microtitulação tais como os feitos de poliestireno ou poli(cloreto de vinilo).

Nas concretizações preferidas, a proteína de ligante de metais, do presente invento, contém um local de combinação a anticorpos que é capaz de imunorreagir com um antigénio pré-seleccionado. Assim, um processo conveniente para a remoção de um complexo catião metálico-proteína, que pode imunorreagir com um antigénio pré-seleccionado, de uma mistura de ligação é (1) misturar a mistura de ligação com uma solução contendo o antigénio pré-seleccionado para formar uma mistura imunorreaccional, (2) manter a mistura imunorreaccional durante um período de tempo suficiente para que se forme um produto de imunorreacção, e (3) remover o produto da imunorreacção da mistura imunorreaccional, removendo portanto o complexo catião metálico-proteína da mistura de ligação. Nas concretizações preferidas, imobiliza-se o antigénio pré-seleccionado num suporte sólido para facilitar a remoção do produto da imunorreacção da mistura imunorreaccional (fase líquida). Os suportes sólidos úteis para a imobilização do antigénio pré-seleccionado são os mesmos suportes sólidos descritos para utilizar na imobilização da proteína ligante de metais.

D. Processo de Preparação de uma Proteína Ligante de Metais

Pode-se preparar uma proteína ligante de metais do presente invento pelos processos gerais de clonação molecular de moléculas de ADN recombinantes, manipulação das moléculas clonadas e expressão da proteína utilizando as moléculas de ADN manipuladas. Uma revisão do processo envolve os elementos seguintes:

1. Obtenção de um gene codificando uma sequência de resíduos de aminoácido que define um domínio variável de uma molécula de imunoglobulina.

2. Determinação da sequência do gene para localizar as regiões que codificam as posições dos resíduos de aminoácido descritas no presente invento que são adequadas para introduzir uma mutação do nucleótido para codificar um resíduo de aminoácido de contacto e produzir um local de ligação a metais.

3. Modificação a sequência base de nucleótidos do gene para codificar resíduos de contacto metal-ligando em três das quatro posições dos resíduos de local de contacto metal-ligando dentro do domínio variável.

4. Inserção da sequência base de nucleótidos modificada num vector de expressão.

5. Expressão da proteína ligante de metais a partir do gene modificado utilizando o vector de expressão num hospedeiro adequado.

6. Isolamento da proteína expressa do meio de expressão.

Pode-se obter um gene por vários meios. Pode-se preparar um segmento de ADN contendo uma sequência de nucleótidos pré-seleccionada por síntese química utilizando, por exemplo, o processo do fosfotriéster de Matteucci et al., J.Am. Chem. Soc., 103: 3185 (1981). Uma vez preparado, inclui-se o segmento de ADN num vector de ADN recombinante que é útil para manipular e ex-

pressar o gene codificando uma proteína ligante de metais.

Pode-se também montar um segmento de ADN contendo uma sequência de nucleótido pré-seleccionada a partir de um grande número de poli-nucleótidos quimicamente sintetizados designados para formar um segmento de ADN particular quando hibridados conjuntamente para formar um segmento de ADN de cadeia dupla. O segmento de ADN de cadeia dupla liga-se então conjuntamente utilizando ADN ligase T4 e insere-se num vector de expressão apropriado.

Alternativamente, pode-se obter um gene codificando moléculas de imunoglobulina, particularmente domínios variáveis de cadeia leve, por clonação molecular de ARN mensageiro (ARNm) ou ADN genómico.

Os processos para o isolamento de um gene são bem conhecidos na arte. Ver, por exemplo, "Guide of Molecular Cloning Techniques", em Methods in Enzymology, volume 152, Berger e Kimmel, eds. (1987); e Current Protocols in Molecular Biology, Ausubel et al., eds., John Wiley and Sons, NY (1987), cujas descrições se incorporam no presente invento para referência.

Os genes úteis para a prática deste invento incluem os genes codificando uma sequência de resíduos de aminoácido definindo um domínio variável contido em produtos de imunoglobulina, moléculas de imunoglobulina, fragmentos Fab, fragmentos F_v e abzimas. Os genes particularmente preferidos são os genes codificando domínios variáveis de cadeia leve de imunoglobulina intacta. Numa concretização, utilizam-se os genes codificando a região variável da cadeia leve de imunoglobulina de uma imunoglobulina capaz de ligar um antigénio pré-seleccionado. Estes genes isolam-se a partir de células obtidas de um vertebrado, preferivelmente de um mamífero, o qual tenha sido imunizado com um ligando antigénico (antigénio) contra o qual se procura actividade, i.e., um antigénio pré-seleccionado. A imunização pode-se realizar convencionalmente e pode-se monitorizar o título de anticorpos no animal para determinar o estado de imunização desejado, o qual cor-

responde à afinidade ou avidez desejada. Os animais parcialmente imunizados, tipicamente recebem apenas uma imunização e recolhem-se as células destes pouco depois de se detectar uma resposta. Os animais completamente imunizados apresentam um pico de título que se consegue com uma ou mais injecções repetidas de um antigénio no mamífero hospedeiro, normalmente com intervalos de duas a três semanas.

Usualmente, três a cinco dias após o último desafio, remove-se o baço e isolam-se os genes codificando cadeias pesadas de imunoglobulina e cadeias leves de imunoglobulina, a partir das células B rearranjadas presentes no baço, utilizando procedimentos padrão. Ver Current Protocols in Molecular Biology, Ausubel et al., eds., John Wiley and Sons, NY (1987); e Antibodies: A Laboratory Manual, Harlowe and Lane, eds., Cold Spring Harbor, NY (1988).

Podem-se derivar genes codifciando polipéptidos V_H e V_L a partir de células que produzem IgA, IgD, IgE, IgG ou IgM, mais preferivelmente de células que produzem IgM e IgG. Os processos de preparação de fragmentos de ADN genómico a partir dos quais se podem clonar genes de região variável de imunoglobulina são bem conhecidos na arte. Ver por exemplo, Herrmann et al., Methods in Enzymol., 152: 180-183 (1987); Frischauf, Methods in Enzymol., 152: 183-190; Frischauf, Methods in Enzymol., 152: 199-212 (1987); e DiLella et al., Methods in Enzymol., 152: 199-212 (1987).

Podem-se isolar genes codificando um domínio variável de uma cadeia leve de imunoglobulina a partir do ADN genómico contendo o gene descrito anteriormente ou a partir do ARNm que codifica o domínio variável. A dificuldade na utilização do ADN genómico está na justaposição das sequências codificando um polipéptido onde as sequências estão separadas por intrões. O(s) fragmento(s) de ADN contendo os exões adequados devem ser isolados, os intrões excisados e os exões unidos ("spliced") conjuntamente na ordem e direcção adequadas. Para a maior parte, isto será difícil, portanto a técnica alternativa, empregando o ARNm, será o

processo escolhido pois a sequência é contígua (livre de intrões) para todo o polipéptido. Os processos de isolamento de ARNm codificando péptidos ou proteínas são bem conhecidos na arte. Ver, por exemplo, Current Protocols in Molecular Biology, Ausubel et al., John Wiley and Sons, Ny (1987); Guide to Molecular Cloning Techniques, em Methods in Enzymology, volume 152, Berger and Kimmel, eds (1987), e Molecular Cloning: A Laboratory Manual, Second Edition, Maniatis et al., eds., Cold Spring Harbor, NY (1989).

Podem-se obter genes codificando ambas as regiões variáveis de cadeia pesada e de cadeia leve, a partir de ARNm e/ou ADN genómico, utilizando a reacção em cadeia de polimerase e vários iniciadores. Ver, pro exemplo, Orlandi et al., Proc. Natl. Acad. Sci. USA, 86: 3833-3837 (1989); Sastry et al., Proc. Natl. Acad. Sci. USA, 86: 5728-5732 (1989); Ward et al., Nature, 341: 544 (1989); Huse et al., Science, 246: 1275-1281 (1989) e PCR Protocol: A Guide to Methods and Applications, Innis et al., eds., Academic Press, London (1990).

Podem-se introduzir modificações na sequência de nucleótidos que codifica a proteína ligante de metais utilizando qualquer um dos processos bem conhecidos de mutagénese dirigida aos locais. Ver, por exemplo, Smith, Ann. Rev. Genet., 19: 423-463 (1985); Kunkel et al., Proc. Natl. Acad. Sci., USA, 82: 488-492 (1985); e Kunkel et al., Meth. Enzymol., 154: 367-382 (1987).

A sequência de nucleótidos modificada capaz de codificar uma protéina ligante de metais liga-se tipicamente a um vector de expressão. Utilizam-se vectores de expressão compatíveis com as células hospedeiras, preferivelmente os que são compatíveis com células hospedeiras eucariotas ou procariotas, para expressar os genes do presente invento. Os vectores de expressão compatíveis com células de mamíferos, de insectos, de plantas ou de bactérias, são particularmente úteis.

Nas concretizações preferidas, um vector de expressão contemplado pelo presente invento inclui um replicão procariota,

i.e., uma sequência de ADN possuindo a capacidade de dirigir a replicação autónoma e a manutenção da molécula de ADN recombinante extra-cromossomicamente numa célula hospedeira procariota, tal como uma célula hospedeira bacteriana, com isso transformada. Estes replicões são bem conhecidos na arte. Em adição, aquelas concretizações que incluem um replicão procariota, incluem também um gene cuja expressão confere resistência a drogas, a um hospedeiro bacteriano com ele transformado. Os genes de resistência a drogas bacterianas típicos são aqueles que conferem resistência à ampicilina ou à tetraciclina.

Aqueles vectores que incluem um replicão procariótico podem também incluir um promotor procariótico capaz de dirigir a expressão (transcrição e tradução) da sequência de nucleótidos da proteína ligante de metais numa célula hospedeira bacteriana, tal como a E. coli, com eles transformada. Um promotor é um elemento de controlo de expressão formado por uma sequência de ADN que permite a ligação da ARN polimerase e que ocorra a transcrição. As sequências do promotor compatíveis com hospedeiros bacterianos são tipicamente proporcionados em vectores plasmídicos contendo locais de restrição convenientes para a inserção de um segmento de ADN do presente invento. Os vectores plasmídicos típicos são pUC8, pUC9, pBR322 e pBR329, disponíveis em Biorad Laboratories, (Richmond, CA) e pPL e pKK223 disponíveis em Pharmacia, Piscataway, NY.

Os vectores de expressão compatíveis com células eucariotas, preferivelmente os compatíveis com células de vertebrados, podem-se também utilizar para expressar a sequência de nucleótidos codificando uma proteína ligante de metais do presente invento. Os vectores de expressão de células eucariotas são bem conhecidos na arte e estão disponíveis em várias fontes comerciais. Tipicamente, estes vectores são proporcionados contendo locais de restrição convenientes para a inserção do segmento de ADN desejado. Estes vectores típicos são pSVL e pKSV-10 (Pharmacia), pBPV--1pML2d (International Biotechnologies, Inc.), e pTDT1 (ATCC, # 31255).

Nas concretizações preferidas, os vectores de expressão de células eucariotas utilizadas para expressar a sequência de nucleótidos codificando uma proteína ligante de metais do presente invento, incluem um marcador de selecção que seja eficaz numa célula eucariota, preferivelmente um marcador de selecção de resistência a drogas. Um marcador de resistência a drogas preferido é o gene cuja expressão resulta na resistência à neomicina, i.e., o gene da neomicina fosfotransferase (neo). Southern et al., J.Mol. Appl. Genet., 1: 327-341 (1982).

A utilização de vectores de expressão retrovirais para expressar uma sequência de nucleótidos codificando uma proteína ligante de metais do presente invento, está também contemplada. Quando utilizado no presente invento, o termo "vector de expressão retroviral" refere-se a uma molécula de ADN que inclui uma sequência promotora derivada a partir da região de repetição terminal longa (LTR) de um genoma de retrovirus.

Nas concretizações preferidas, o vector de expressão é tipicamente um vector de expressão retroviral que é, preferivelmente, incompetente na replicação em células eucariotas. A construção e a utilização de vectores retrovirais foi descrita por Sorge, et al., Mol. Cell. Biol., 4: 1730-37 (1984).

Têm-se desenvolvido vários processos para ligar, operativamente, sequências de nucleótidos a vectores por meio de extremidades coesivas complementares. Por exemplo, podem-se adicionar sistemas de homopolímeros complementares ao segmento de ADN a ser inserido e ao vector de ADN. Juntam-se então o vector e o segmento de ADN por ligações de hidrogénio entre as caudas homopoliméricas complementares para formar moléculas de ADN recombinante.

Os ligadores sintéticos contendo um ou mais locais de restrição proporcionam um processo alternativo para juntar o segmento de nucleótidos aos vectores. O segmento de nucleótidos, gerado por digestão por endonuclease de restrição como previamente descrito, trata-se com ADN-polimerase de bacteriofago T4 da ADN-polimerase I da E. coli, enzimas que removem os terminais

protuberantes, de cadeia simples, 3', com as suas actividades exonucleotídicas 3'-5' e preenchem os extremos 3' cortados com as suas actividades de polimerização. A combinação destas actividades gera portanto segmentos de ADN com extremidades rombas. Os segmentos com extremidades rombas são então incubados com um grande excesso molar de moléculas de ligador na presença de uma enzima que é capaz de catalisar a ligação de moléculas de ADN com extremidades rombas, tal como a ADN-ligase de bacteriofago T4 . Assim, os produtos da reacção são segmentos de ADN possuindo sequências ligadoras poliméricas nas suas extremidades. Estes segmentos de ADN são então clivados com a enzima de restrição apropriada e ligadas a um vector de expressão que tenha sido clivado com uma enzima que produz extremidades compatíveis com as do segmento de ADN.

Os ligadores sintéticos contendo vários locais de endonuclease de restrição estão comercialmente disponíveis em várias fontes incluindo International Biotechnologies, Inc., New Haven, CAN.

Os vectores de expressão típicos úteis para a expressão de genes em plantas são bem conhecidos na arte e incluem vectores derivados do plasmídeo indutor de tumores (Ti) de Agrobacterium tumefaciens descritos por Rogers et al., Meth. in Enzynol., 153: 253-277 (1987). No entanto, conhecem-se vários outros sistemas de vectores de expressão que funcionam em plantas. Ver, por exemplo, Verma et al., PCT Publication Nº WO87/00551; e Cocking e Davey, Science, 236: 1259-1262 (1987).

Os vectores de expressão descritos anteriormente contêm elementos de controlo de expressão incluindo o promotor. A sequência de nucleótidos codificando a proteína ligante de metais está operativamente ligada ao vector de expressão para permitir que a sequência promotora dirija a ligação da ARN polimerase e a síntese do gene codificando o polipéptido desejado. São úteis na expressão das sequências de nucleótidos codificando a proteína ligante de metais os promotores que são indutíveis, virais, sintéticos, constitutivos, temporalmente regulados, espacialmente

regulados e espaço-temporalmente regulados.

A escolha de qual o vector de expressão e, finalmente a que promotor uma sequência de nucleótidos codificando a proteína ligante de metais (gene) está operativamente ligada, dependem directamente, como é bem conhecido na arte, das propriedades funcionais desejadas, por ex., da localização e ajustamento da expressão das proteínas, e da célula hospedeira a transformar, estas sendo limitações inerentes na arte da construção de moléculas de ADN recombinante. No entanto, um vector de expressão útil na prática do presente invento é, pelo menos, capaz de dirigir a replicação e, preferivelmente, também a expressão do gene codificando os polipéptidos incluído no segmento de ADN ao qual está operativamente ligado.

Nas concretizações preferidas, o vector de expressão utilizado para expressar o gene codificando a proteína ligante de metais, inclui um marcador de selecção que é eficaz numa célula da planta, preferivelmente um marcador de selecção de resistência a drogas. Um marcador de resistência a drogas preferido é o gene cuja expressão resulta em resistência à canamicina, i.e., o gene quimérico contendo o promotor de nopalina-sintase, Tn5 neomicina--fosfotransderase II e a região 3' não traduzida da nopalina sintase, descrito por Rogers et al., em Methods For Plant Molecular Biology, Weissbach e H. Weissbach, eds., Acadenic Press Inc., San Diego, CA (1988). Está comercialmente disponível um vector de expressão de plantas útil em Pharmacia, Piscataway, NY.

Têm-se desenvolvido vários processos para ligar operativamente o ADN a vectores por meio de extremidades coesivas complementares. Por exemplo, podem-se adicionar sistemas e homopolímeros complementares ao segmento de ADN a ser inserido e ao vector de ADN. Juntam-se então o vector e o segmento de ADN por meio de ligações de hidrogénio entre as caudas homopoliméricas complementares para formar moléculas de ADN recombinante.

Alternativamente, podem-se utilizar ligadores sintéticos

contendo um ou mais locais de endonuclease de restrição para juntar a sequência de nucleótidos ao vector de expressão. Os ligadores sintéticos ligam-se aos segmentos de ADN com extremidades rombas incubando os segmentos de ADN com extremidades rombas com um grande excesso de moléculas de ligadores sintéticos na presença de uma enzima que seja capaz de catalisar a ligação das moléculas de ADN com extremidades rombas tal como ADN-ligase de bacteriofago T4. Assim, os produtos da reacção são segmentos de ADN possuindo sequências de ligadores sintéticas nas suas extremidades. Estes segmentos de ADN são então clivados com a endonuclease de restrição apropriada e ligados num vector de expressão que tenha sido clivado com uma enzima que produz extremidades compatíveis com as do ligador sintético. Os ligadores sintéticos contendo vários locais de endonuclease de restrição estão comercialmente disponíveis em várias fontes incluindo a New England Biolabs, Beverley, MA.

Os processos para a introdução de genes codificando proteínas ligantes de metais em plantas incluem a transformação de plantas mediada pela Agrobacterium, a transformação de protoplastos, transferência do gene para polén, injecção em orgãos reprodutores e injecção em embriões imaturos. Cada um destes processos tem vantagens e desvantagens distintas. Assim, um processo particular de introdução de genes numa espécie de plantas particular pode não ser, necessariamente a mais eficaz para outras espécies de plantas.

O presente invento refere-se também a uma célula hospedeira transformada com uma molécula de ADN recombinante do presente invento, preferivelmente uma sequência de nucleótidos capaz de expressar uma forma solúvel de uma proteina ligante de metais. A célula hospedeira pode ser procariota ou eucariota. As células bacterianas são as células hospedeiras procariotas preferidas e são, tipicamente, uma estirpe de E. coli, tal como, por exemplo, a estirpe DH5 da E. coli, disponível em Bethesda Research Laboratories, Inc., Bethesda, M.D. As células hospedeiras eucariotas preferidas incluem células de leveduras e de mamíferos, preferivelmente células de vertebrados tais como as da

linha de células fibroblásticas do ratinho, da ratazana, do macaco ou do homem. As células hospedeiras eucariotas preferidas incluem as células do ovário de hamster chinês (CHO) disponíveis no ATCC como CCL61 e as células de embriões de ratinhos suiços NIH, NIH/3T3, disponíveis no ATCC como CRL 1658. A transformação de células hospedeiras apropriadas como uma molécula de ADN recombinante do presente invento consegue-se por processos bem conhecidos na arte que dependem tipicamente do tipo de vector utilizado. Em relação à transformação de células hospedeiras procariotas ver, por exemplo, Cohen et al., Proc. Natl. Acad. Sci. USA, 69: 2110 (1972); e Maniatis et al., Molecular Cloning, A Laboratory Mannual, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY (1982). Em relação à transformação de células de vertebrados com vectores retrovirais contendo ADNr ver, por exemplo, Sorge et al., Mol. Cell. Biol., 4: 1730-37 (1984); Graham et al., Virol, 52: 456 (1973); e Wigler et al., Proc. Natl. Acad. Sci. USA, 76: 1373-76 (1979).

Podem-se isolar as proteínas ligantes de metais produzidas em várias células hospedeiras por um vector de expressão, por vários meios bem conhecidos incluindo a homogeneização da célula hospedeira seguida por isolamento da proteína ligante de metais ou isolamento da proteína ligante de metais excretada para o meio no qual as células hospedeiras crescem.

Uma proteína isolada é uma proteína que está substancialmente livre de outras proteínas celulares encontradas com essa proteína na natureza. Por exemplo, uma proteína ligante de metais isolada está substancialmente livre de outras proteínas da célula hospedeira E. coli que estão presentes nas células hospedeiras E. coli que produziram a proteína de ligação a metais. Uma proteína ligante de metais isolada está também substancialmente livre de proteínas que não possuem uma sequência de resíduos de aminoácido idêntica.

E. Metaloproteína

Uma proteína ligante de metais do presente invento torna-se uma metaloproteína quando está complexada com um catião metálico.

Os catiões metálicos (catiões cofactor) adequados para a complexação com uma proteína ligante de metais deste invento são quaisquer dos metais de transição da tabela periódica, e dos metais que não são de transição, cálcio (Ca), zinco (Zn), cádmio (Cd), mercúrio (Hg), estrôncio (Sr) e bário (Ba), metais estes que possuem a capacidade de ocupar um estado de oxidação tetraédrico e portanto de se complexarem com a referida proteína através dos ligandos coordenados proporcionados pelos três resíduos de aminoácido de contacto na proteína ligante de metais. Os catiões metálicos preferidos para utilizar numa metaloproteína deste invento são catiões divalentes, preferivelmente, Cu(II), Zn(II), Ni(II), Co(II), Fe(II), Ag(II), Mn(II) ou Cd(II), e mais preferivelmente Cu(II).

As metaloproteínas formam-se preparando uma proteína ligante de metais do presente invento e, depois, expondo a proteína ligante de metais a um catião metálico seleccionado, preferivelmente num meio aquoso tamponado, durante um tempo suficiente para permitir que se forme um complexo de coordenação metal-proteína. A metaloproteína preparada e descrita nos exemplos é preferida.

Podem-se também formar metaloproteínas misturando uma proteína ligante de metais do presente invento com outra metalo-proteína tal como a calmodulina que está complexada com cálcio ou a metalotioneína complexada com cádmio. As metaloproteínas do presente invento podem então complexar com o catião doado pela outra metaloproteína.

F. Moléculas de Anticorpos Ligante de Metais e Metalo-anticorpos

O presente invento contempla também uma proteína ligante de metais deste invento que possua a capacidade de imunorreagir com um antigénio pré-seleccionado.

Uma proteína ligante de metais contém um domínio variável de cadeia leve de imunoglobulina como se descreveu no presente invento. Numa concretização, essa proteína ligante de metais pode

estar presente em associação com um domínio variável de cadeia pesada de modo a formar um local de combinação de anticorpos e portanto possuem a capacidade de imunorreagir com antigénios.

Por exemplo, pode-se isolar um par de genes codificando V_H e V_L, como se descreveu no presente invento para o V_L sozinho. Modifica-se então o gene V_L de modo a conter um local de ligação a metais, e co-expressam-se então os genes V_L e V_H para produzirem os produtos de gene V_H e V_L que se reúnem em complexos de proteína di- ou multi-méricos, os quais possuem a capacidade de formar um local de combinação de anticorpos e de imunorreagir com um antigénio pré-seleccionado.

Têm-se isolado repetidamente genes clonados codificando moléculas de imunoglobulina de cadeia pesada e leve e, depois introduzido num meio de expressão que produz moléculas de anticorpos montadas compreendendo tanto V_H como V_L, que são capazes de imunorreagir com a mesma especificidade que o anticorpo produzido pelo anticorpo que produz a célula a partir da qual se isolam os genes que codificam a proteína imunoglobulina. Ver, por exemplo, Roberts et al., Protein Engineering, 1: 59-65 (1986); Morrison, Science, 229: 1202-07 (1985); Patente dos E.U. Nº 4 474 893; e pedidos de patente publicados nº EP 0125023, Ep 0239400 e WO 89/00999.

Para o isolamento de um gene codificando uma cadeia pesada ou leve de imunoglobulina na preparação de um anticorpo ligante de metais, podem-se obter os genes a partir de qualquer célula que produza uma molécula de anticorpo que imunorreaja com um antigénio pré-seleccionado. Uma célula preferida é a célula de hibridoma ou uma célula de um baço de imunoresposta.

Como exemplos de processos de clonação de ADN recombinante gerais, ver Molecular Cloning, Maniatis et al., Cold Spring Harbor, NY, (1982); DNA Cloning, Glover, ed., IRL Press, McLean, VA (1985). Para a clonação genómica e expressão de genes de imunoglobulina em células linfóides, ver Neuberger et al., Nature, 312: 604-8 (1984); Ochi et al., Proc. Natl. Acad. Sci.

USA, 80: 6351-55 (1987); e Oi et al., Proc. Natl. Acad. Sci. USA, 80: 825-29 (1983). Para a clonação de genes de imunoglobulina de células de hibridoma, e sua expressão em oócitos de Xenopus, ver Roberts et al., Protein Engineering, 1: 59-65 (1986), e ver Wood et al., Nature, 314: 446-9 (1985) para a sua expressão em leveduras.

Pode-se também produzir um anticorpo ligante de metais a partir de um reportório imunológico, como descrito por Sastry et al., Proc. Natl. Acad. Sci., 86: 5728-5732 (1989); Huse et al., Science, 246: 1275-1281 (1989). Nesta concretização, pode-se modificar um gene clonado codificando uma proteína que inclui um domínio variável de cadeia leve de imunoglobulina (V_L) de modo a conter um local de ligação a metais, como descrito no presente invento. Consequentemente combina-se o gene V_L modificado com qualquer dos duplicatos do gene V_H presentes no reportório imunológico. Por este processo podem-se preparar um grande número de combinações de V_L (modificado) mais V_H (reportório) em que a especificidade antigénica variará baseada na proteína V_H codificada pelo reportório que é montada com a proteína V_L possuindo o local de ligação a metais. Seleccionam-se então as combinações quanto à capacidade para imunorreagir com um antigénio predeterminado na presença de um catião cofactor.

Estão também contempladas as combinações de uma região variável de cadeia pesada (V_H) que tenha sido modificada de modo a conter um local de ligação a metais do presente invento. Combina-se o V_H modificado com qualquer das duplicatas de V_L presentes no reportório imunológico. Este processo produz um grande número de combinações de V_H modificado (modificado) e V_H (reportório). Seleccionam-se então as combinações quanto à capacidade de imunorreagir com um antigénio predeterminado na presença de um catião cofactor.

Uma molécula de ligação a antigénios, de cadeia simples, é ainda uma concretização de um anticorpo ligante de metais deste invento. Nesta concretização, o V_H e o V_L estão operativamente ligados por um ligador peptídico como ainda se descreveu no

presente invento e o aspecto do V_L contém um local de ligação a metais. Um exemplo de um anticorpo ligante de metais, de cadeia simples, é a proteína QM212 descrita e preparada no exemplo 1.

As moléculas de anticorpos ligantes de metais descritas no presente invento são úteis para a partição selectiva e/ou remoção de catiões metálicos de soluções aquosas pelos processos de ligação ao catião metálico descritos no presente invento. Pode-se tomar vantagens da imunorreactividade do aspecto do local de combinação a anticorpos de um complexo metalo-anticorpo formado, de modo a que se possam utilizar antigénios pré-seleccionados, na fase sólida, por exemplo, para remover complexos metalo-anticorpo da fase líquida.

Noutra concretização, pode-se fazer imunorreagir um anticorpo ligante de metais com um antigénio que serve como indicador de formação do complexo metálico. Por exemplo, um antigénio pode ser sensível à presença de um catião metálico no micro-meio do local de combinação do anticorpo e indicar a complexação do local de ligação a metais com o catião metálico. Os antigénios cromofóbicos tais como a fluoresceína, por exemplo, produzem deslocamentos na fluorescência quando um catião metálico está ligado num local de ligação a metais deste invento adjacente ao local de combinação de anticorpos.

G. Metalo-anticorpo Catalítico

Noutra concretização, o invento contempla um metalo-anticorpo do presente invento que tenha a capacidade de promover uma reacção química predeterminada, i.e., um metalo-anticorpo catalítico. Nesta concretização, um antigénio pré-seleccionado com o qual o local de combinação de anticorpos imunorreage é também um substrato para uma reacção que é promovida pelo metalo-anticorpo.

Nas concretizações preferidas, o catião metálico de um metalo-anticorpo catalítico é Zn(II), e uma reacção química preferida é a hidrólise de uma ligação peptídica.

Os anticorpos catalíticos foram descritos por Tramontano et al., Science, 234: 1566-1570 (1986); Pollack et al., Science, 234: 1570-1573 (1986); Janda et al., Science, 241: 1188-1191 (1988); e Janda et al., Science, 244: 437-440 (1989).

EXEMPLOS

Os seguintes exemplos pretendem ilustrar, mas não limitar, o presente invento.

1. Preparação de uma molécula de anticorpo ligante de metais

A. Vector de expressão da QM212

O vector de expressão plasmídico, pQM212, contém um gene estrutural que codifica a sequência de resíduos de aminoácido mostrada na figura 1 para a proteína QM212 e possui a capacidade de dirigir a expressão da QM212 sob as condições seguidamente descritas. O vector contém também os aspectos mostrados na figura 2 incluindo um gene de resistência à ampicilina (ampR), uma origem de replicação M13 (M13 Ori), e um promotor lambda (OL/PR) proporcionando um local de ligação para o repressor lambda sensível à temperatura para a expressão do gene da QM212, em adição a conter a sequência de bases de nucleótidos mostrada na figura 1, da base 1 à base 789 que codifica a QM212.

A QM212 é uma proteína de ligação a antigénios de cadeia simples, possuindo um local de ligação a metais e é derivada da protéina 4-4-20/212 descrita por Bird et al., Science, 242: 423-426 (1988). A QM212 difere da 4-4-20/212 por substituições de aminoácidos nas posições dos resíduos 60, 62, 115 e 117, para proporcionar os resíduos de aminoácido de contacto, três dos quais formam um local de ligação a metais. Introduziram-se quatro mutações no gene codificando a molécula da proteína 4-4-20/212 pelas técnicas de mutagénese in vitro descritas por Sayels et al., Nucl. Acids. Res., 16: 791-802 (1988). A substituição de Leu na posição do resíduo 36 em vez de Tyr 36, embora não seja necessária para introduzir os resíduos de aminoácido de contacto, fez-se para remover interacções estereoquímicas, possíveis mas não verificadas, entre a His 89 e o resíduo hidroxilo da Tyr 36.

Assim, introduziram-se duas mutações no CDR L3, convertendo uma Ser numa His em ambas as posições de resíduos 115 e 117, utilizando um único oligonucleótido de 30 bases. Introduziram-se duas mutações no CDR L1 na posição do resíduo 60 convertendo uma Arg numa His e na posição do resíduo 62 convertendo uma Tyr numa Leu, utilizando um único oligonucleótido de 30 bases. Isolou-se o plasmídeo resultante contendo as quatro substituições e verificaram-se as substituições por sequenciação do ADN.

B. Expressão da proteína QM212

Introduziu-se o plasmídeo de expressão pQM212 nas células de E. coli, estirpe N99 (C1857), por transformação, e obtiveram-se células transformadas com sucesso e designaram-se por N99/pQM212. O plasmídeo pQM212 na Escherichia coli estirpe N99 (C1857), designado por QM212 # 7, depositou-se na American Type Culture Collection (ATCC, Bethesda, MD) em 8 de Maio de 1990 e atribui-se-lhe o número de acesso, 68 321. Inoculou-se cerca de 1 mililitro (ml) de cultura saturada com células N99/pQM212, em 400 ml de meio LB Amp (extrato de leveduras 10 g/l, peptona 10 g/l, NaCl 5 g/l, K_2HPO_4 10 g/l, ampicilina 100 mg/l, pH 7,0) num frasco tapado de 2 litros (l) e cultivou-se a 30 graus centígrados (30°C) para formar uma cultura bacteriana. Mediu-se periodicamente a densidade óptica (D.O.) da cultura bacteriana a 600 namómetros (nm) utilizando um meio LB Amp como linha base do zero de absorvância. Quando a D.O. da cultura bacteriana atingiu cerca de 3,0 unidades de D.O., após cerca de 5 horas, transferiu-se a cultura para um banho de água a 50°C e manteve-se a agitação para aumentar rapidamente a temperatura da cultura para cerca de 42-44°C. O aumento de temperatura permitiu a indução do repressor lambda dentro das células hospedeiras N99 (C1857), activando assim o promotor do operão esquerdo (OL/PR) presente no pQM212 quando da ligação do repressor lambda de modo a que a proteína QM212 seja expressa. Manteve-se então a cultura a 42°C durante cerca de 30 minutos para permitir que ocorresse a expressão da QM212 e que se formasse uma cultura contendo QM212 possuindo cerca de 75 a 150 miligramas (mg) de QM212 por 400 ml de cultura.

Centrifugou-se a cultura contendo QM212 (400 ml) a 10 000 rpm num rotor JAQ-10 Sorval (Beckonan Instruments, Fullerton CA) durante 10 minutos para formar uma pelota de células. Reteve-se a pelota de células e voltou-se a suspender em tampão de lise [50 milimolar (mM) de Tris-HCl, 1 mM de EDTA, 0,1 mM de PMSF, pH 8,0] a 4°C e numa concentração de 10 ml por mol de volume da pelota de células. Passou-se a pelota de células ressuspensas através de um homogeneizador do tipo "french press" (Matt-Guzlin) a 4°C, três vezes, para formar uma suspensão de células bacterianas lisadas, e depois centrifugou-se a suspensão a 23 400 x g durante 45 minutos a 6°C para formar uma pelota. Voltou-se a suspender a pelota em metade do seu volume do tampão de lise utilizado antes da homogeneização na "french press", e voltou-se a centrifugar a 23 400 x g durante 45 minutos a 6°C para formar uma pelota lavada. Recuperou-se a pelota lavada e lavou-se duas vezes adicionais com tampão de lise, voltou-se a suspender e a centrifugar para formar uma pelota lavada três vezes.

Voltou-se a suspender a pelota anterior lavada três vezes, em tampão Gu-HCl recentemente preparado (6 M em guanidina-HCl, 50 mM de Tris-HCl, 10 mM de $CaCl_2$, 100 mM de KCl, 0,1 mM de PMSF, pH 8,0) a 4°C e numa concentração de cerca de 9 ml por ml de volume da pelota lavada três vezes, e misturou-se completamente a suspensão para formar uma suspensão solubilizada. Depois disso, centrifugou-se a suspensão solubilizada a 23 400 x g durante 30 minutos a 6°C e recuperou-se o sobrenadante resultante. Mediu-se então a absorção a 280 nm (A280) do sobrenadante recuperado, e adicionou-se tampão Gu-HCl adicional para ajustar a A280 a cerca de 25 unidades, relativamente ao tampão Gu-HCl sozinho possuindo uma A280 de 0,0 unidades, para formar o extracto da proteína QM212 solubilizado.

Tratou-se então como se segue o extracto da proteína QM212 solubilizado contendo proteína QM212 desnaturada de guanidina para permitir que a proteína QM212 desnaturada se "redobrasse" numa conformação "dobrada" tri-dimensional. Preparou-se tampão de redobragem (50 mM de Tris-HCl, 10 mM de $CaCl_2$, 50 mM de KCl,

pH 8,0) e pré-arrefeceu-se a cerca de 4-7°C, adiconou-se PMSF até 0,1 mM e imediatamente depois misturou-se o tampão de redobragem contendo PMSF com o extracto da proteína QM212 solubilizado até uma razão de extrcato para o tampão de cerca de 1:200 a 1:500 para formar uma mistura de redobragem. Manteve-se a mistura de redobragem a cerca de 4°C durante 12 a 16 horas para formar uma solução contendo QM212 redobrada.

Reduziu-se em volume cerca de 1600 ml de solução contendo QM212 redobrada utilizando uma unidade de ultrafiltração Amicon equipada com um ultrafiltro de 0,3 micra Pharmacia Ortega possuindo um isolamento de massa molecular de 10 000, até um volume final de cerca de 10 a 50 ml, e dializou-se o volume reduzido contra o tampão de fosfato de sódio 10 mM, pH 6,0 a 4°C, utilizando quatro mudanças de cerca de 20 volumes de tampão de diálise. Aplicou-se o resultante da diálise e a solução concentrada numa coluna de permuta catiónica Monos (HR1010, Pharmacia, Piscataway, NY) pré-equilibrada com tampão de equilíbrio de 10 mM de tampão fosfato, pH 8,0. Lavou-se, depois, a coluna com o tampão de equilíbrio. Eluiu-se um único pico, contendo a proteína QM212, utilizando um gradiente linear de tampão variando de 10 mM de fosfato, pH 8,0 a 150 mM de NaCl e recolheu-se a fracção do pico. Dializou-se então extensivamente a fracção do pico recolhida por 5 diálises sucessivas de 500 ml cada a 4°C contra 1 mM de EDTA para remover quaisquer catiões metálicos que possam estar presentes. A diálise de EDTA foi seguida por diálise extensiva em tampão fosfato isento de metais para remover o EDTA, para formar uma solução da proteína QM212 livre de metais. A proteína QM212 isolada é capaz de formar um complexo de coordenação com um ião metálico e é compreendida de uma sequência de aminoácidos que define um domínio variável de uma imunoglobulina e três resíduos de aminoácido de contacto nesse domínio variável que definem um local de ligação a metais.

A solução de proteína QM212 isenta de metais, como analisada por SDS-PAGE utilizando um gel de poliacrilamida a 20 por cento, apresentou uma única banda homogénea com uma mobilidade electroforética de cerca de 26 500 daltons. Assim, a solução de

proteína QM212 isenta de metais é uma composição homogénea contendo a molécula ligante de metais QM212.

Preparou-se também a proteína 4-4-20/212 relecionada, pelos procedimentos anteriores, para se utilizar como um controlo nos estudos descritos em seguida. Expressou-se a proteína 4-4-20/212 a partir de um vector de expressão proporcionado por Dr. Syd Johnson (Genex Corp., MD) que difere da pQM212 apenas nas substituições descritas no exemplo 1a.

2. Ensaio de Ligação a Metais

Mediu-se a ligação de metais a uma proteína pela interrupção da fluorescência de resíduos de triptofano presentes na proteína, na presença ou ausência de iões metálicos adicionados. Diluiu-se uma solução de proteína 4-4-20/212 ou QM212 isenta de metais, preparada como no exemplo 1b, em fosfato de sódio a 15 mM, pH 6,0, até uma concentração de 150 nanomolar (nM), utilizando uma massa molecular de cerca de 26 500 daltons para cada proteína para calcular a molaridade, para formar uma mistura de interrupção. Determinou-se a fluorescência da mistura de interrupção utilizando um espectrofluorofotómetro RF500-U (Shimadzu, Japan) utilizando um comprimento de onda de excitação de 292 nm e detectando um espectro de emissão utilizando uma largura de banda de 5 nm tanto para a excitação como para a emissão. Detectou-se emissão a 343 nm quando se mediu a fluorescência de uma solução contendo a QM212 e detectou-se a 352 nm quando se mediu a solução contendo a 4-4-20/212 porque se determinou que estes comprimentos de onda representavam o máximo de emissão para cada proteína quando se mediram os espectros de emissão numa gama de 300 a 450 nm. Mediu-se a fluorescência em 1 ml de uma solução, com várias quantidades de cloreto de cobre (II) [Cu(II)] adicionado na gama de 0 a 30 micromolar (μM) à mistura de interrupção antes de se conduzirem as medições da excitação e da emissão. A interrupção da fluorescência é expressa na forma de uma razão F/F_o, onde 1,0 é uma unidade arbitrária de 0% de interrupção numa solução não possuindo metal adicionado (F_o) e F é a emissão quando se adicionou o metal.

A figura 3 mostra que o Cu(II) interrompeu a fluorescência dos resíduos de triptofano presentes na solução da proteína QM212. Em contraste, o Cu(II) não interrompe substancialmente a fluorescência associada com a proteína 4-4-20/212. os dados mostram também que a interrupção da fluorescência pelo Cu(II) da QM212 apresenta comportamento de saturação acima de cerca de 12-15 mM de Cu(II) com uma constante de ligação aparente de $2 \times 10^{5} M^{-1}$.

A estrutura tridimensional do presente anticorpo (4-4-20) com fluoresceína ligada, é conhecida (Herron et al, Proteins, 5: 271-286, 1989), e cinco dos seis resíduos de triptofano tanto na QM212 como no anticorpo progenitor, estão numa vizinhança próxima tanto do local de combinação de antigénio como do local de ligação a metais, produzidos pelo presente invento. Baseado na proximidade, antecipou-se e subsequentemente mostrou-se, pelos dados de interrupção anteriores, que os iões metálicos ligados interrompem eficazmente a fluorescência do triptofano de uma maneira análoga à mostrada nos estudos em sistemas de metaloproteínas modelo demonstrando que o Cu(II) é um interruptor particularmente eficaz do triptofano na região de ligação a metais, seguida pelo níquel na forma Ni(II) e pelo cobalto na forma Co(II). A natureza eficaz da interrupção com a QM212 pelo Cu(II) indica que o átomo de Cu(II) está relativamente perto da maioria dos resíduos de triptofano, uma observação consistente com a ocorrência de ligações nos locais de ligação a metais da QM212. Assim, os resultados da interrupção de uma interacção de ligação metal-proteína específica, indicando que a QM212 se liga especificamente a metais, i.e., que a QM212 é uma protéina de ligação a metais capaz de formar um complexo de coordenação com um catião metálico, enquanto que a 4-4-20/212 não se liga especificamente a metais.

Numa experiência relacionada, adicionou-se EDTA, em excesso molar da quantidade de Cu(II) presente, à mistura de interrupção, e mediu-se a emissão. O excesso de EDTA fez com que o nível da intensidade da fluorescência voltasse a subir até ao nível da proteína QM212 quando medida sem Cu(II) adicionado. Estes dados

indicam que aquela interrupção é o resultado de uma interacção metal-proteína reversível, convencional, encontrada em metaloproteínas, e não o resultado de variações irreversíveis no triptofano causadas por contacto com iões metálicos.

A capacidade da proteína ligante de metais, QM212, se ligar a outros catiões metálicos, mediu-se na forma de inibição da ligação do Cu(II) à QM212 por iões competitivos, zinco na forma Zn(II) e cádmio na forma Cd(II). Antes da adição de Cu(II) à mistura de interrupção descrita anteriormente, adicionou-se Zn(II) 100 mM ou Cd(II) 500 µM e mediu-se a interrupção da fluorescência como anteriormente. As figuras 4A e 4B, respectivamente, mostram que tanto o zinco como o cádmio ligam a QM212 de uma maneira dependente da dose suficiente para inibir a ligação do Cu(II) à QM212. Assim, muitos iões metálicos, incluindo Cu(II), Zn(II), Cd(II), Ni(II) e Co(II), que apresentam ligandos para formar complexos de coordenação tetraédricos, incluindo Cu(II), Zn(II), Cd(II), Ni(II) e Co(II) podem ligar-se a proteína ligante de metais do presente invento.

3. Ensaio de ligação a antigénio

Devido à QM212, como modelo de metalo-anticorpo, ser derivada de uma porção de cadeia leve de um anticorpo monoclonal anti-fluoresceína, a capacidade de a QM212 imunorreagir com antigénios pode-se determinar utilizando fluoresceína. A ligação da fluoresceína à QM212 mediu-se detectando o efeito da interrupção da fluorescência da QM212 sobre o espectro de excitação da fluoresceína. Ver, por exemplo, J.N. Herron em Fluorescein Hapten: An Immunological Probe, E.W. Voss, Jr., Ed. CRC Press, Boca raton, FL, 1984 pp. 49-76. Mediu-se o espectro de excitação de uma solução contendo fosfato de sódio 15 mM, pH 6,0, e fluoresceína 7,0 nm (grau padrão de espectrometria), na presença de várias quantidades de QM212 preparada como no exemplo 1b, utilizando um espectrofluorofotómetro (Shimadzu). A excitação a 400 - 520 nm foi seguida de medição da emissão a 530 nm utilizando uma largura de banda de 5 nm tanto para a excitação como para a emissão. Os dados dos espectros de excitação nas figuras 5 e 6 estão expressos na forma de intensidade de fluorescência que

se segue à excitação a de 400 a 520 nm.

Os dados do espectro de excitação resultantes da mistura de QM212 com fluoresceína em quantidades variando entre 0 e 7,7 µM mostram-se na figura 5. À medida que se adicionou mais QM212 à solução de fluoresceína, a quantidade de excitação reduziu-se, interrompendo a excitação da fluoresceína por cerca de 80 por cento (máximo para máximo). O máximo de excitação deslocou-se de cerca de 493 nm para cerca de 507 nm à medida que a fluoresceína se ligava à QM212. Observou-se um resultado similar quando se utilizou a proteína 4-4-20/212 em vez da QM212. Estes dados indicam que a fluoresceína se liga à QM212 de uma maneira similar à ligação da fluoresceína com a 4-4-20/212, tendo-se caracterizado esta última como uma imunorreacção antigénio-anticorpo. Bird et al., Science, 242: 423-426, 1988. Assim, a QM212 possui propriedades de uma molécula de anticorpo e portanto é um exemplo de uma molécula de anticorpo de ligação a metais.

Para estabelecer que a QM212 se pode ligar simultaneamente tanto a um metal como a um antigénio, e actuar como um metalo-anticorpo, mediu-se o efeito do metal sobre o espectro de excitação da fluoresceína quando a fluoresceína se ligou à QM212. Prepararam-se soluções contendo fosfato de sódio 15 nM, pH 6,0, 7,0 nM de fluoresceína e 7,7 µM de QM212. Depois disso, adicionaram-se várias quantidades de Cu(II) a cada solução variando entre 0 e 30 µM, e mediram-se os espectros de excitação. Os dados mostram-se na figura 6.

Sem Cu(II) adicionado, similarmente à condição mostrada na figura 5 quando a fluoresceína estava completamente ligada pela QM212, o espectro tinha um pico de intensidade de cerca de 21,60 no máximo a cerca de 508 nm. Com Cu(II) adicionado, a interrupção não foi invertida como seria de esperar se a fluoresceína fosse deslocada pelo Cu(II) na ligação à QM212. Em vez disso, o máximo do espectro de excitação deslocou-se de cerca de 508 nm a cerca de 494 nm e diminuiu à medida que se adicionou mais Cu(II). Observou-se que ocorreu saturação a cerca de 18 µM de Cu(II). Não se observaram variações adicionais no espectro com 18 a 21 µM de

Cu(II). Como se esperava, não se viu nenhuma perturbação dependente do Cu(II) da fluorescência da fluoresceína ligada quando se conduziu a experiência anterior utilizando a proteína 4-4-20/212 em vez da proteína QM212.

Estes dados indicam que a molécula de QM212 se pode ligar simultaneamente ao Cu(II) e à fluoresceína na molécula de QM212. A natureza perturbada do máximo de excitação da fluoresceína observado à medida que se adicionou Cu(II) indica que o catião metálico, Cu(II), se ligou na vizinhança próxima da fluoresceína, mostrando assim que o átomo de Cu(II) se ligou no local de ligação a metais que se projectou pelo presente invento para se localizar estritamente adjacente ao local de ligação a antigénio.

4. Identificação dos locais do domínio variável da imunoglobulina para a introdução de um ligando metálico

A estrutura dos domínios variáveis de anticorpos HyHEL5, HyHEL10, Newm, J539, HEd10, e Kol, disponíveis no Protein Data Bank descrito por Bernstein et al., J.Mol. Biol., 112: 535-542 (1977), e D1.3, descrita por Amit et al., Science, 233: 747-753 (1986) e fornecidos por Roberto Poljak, compararam-se por sobreposição dos domínios variáveis de cadeia leve (V_L) e dos domínios variáveis de cadeia pesada (V_H), individuais, sobre os correspondentes domínios do McPC603 descritos por Bernstein et al., J.Mol. Biol., 112: 535-542 (1977). Utilizou-se um conjunto ligeiramente modificado da estrutura em folha ß comum de Chothia et al., J.Mol. Biol., 196: 901-917 (1987), para definir a correspondência um a um da sobreposição. No esquema de numeração de Kabat et al., em Sequences of Proteins of Immunological Interest, 4th, Ed., U.S. Department of Health and Human Services, National Institute of Health, Bethesda, MD (1987), as correspondências definidas no presente invento são os resíduos de V_L 4-6, 9, 11-13, 19-25, 33-49, 53-55, 61-76, 84-90, e 97-107, e os resíduos de V_H 3-12, 17-25, 33-52, 56-60, 68-82, 88-95, e 102-112. O V_L 10 e o 106A foram excluídos porque estes resíduos são eliminados em algumas sequências. As coordenadas para a estrutura cristalográfica do anticorpo anti-fluoresceína 4-4-20 foram tornadas disponíveis por Edmundson et al., Proteins, 5:

271-280 (1989). Uma pesquisa no Protein Data Bank proporcionou as estruturas cristalográficas de cinco proteínas catalíticas contendo zinco: anidrase carbónica, por Kannan et al., Proc. Natl. Acad. Sci. USA, 72: 51-55 (1975); carboxipeptidase A, por Rees et al., J.Mol. Biol., 168: 367-387 (1983); superóxido--dismutase, por Tainer et al., Nature, 306: 284-287 (1983); álcool-desidrogenase, por Branden et al., em Enzymes, ed. Boyer, P.D. (Academic Press, New York), 3 ed Edition); e termolisina, por Holmes et al., J. Mol. Biol., 160: 623-939 (1982).

As sobreposições dos oito conjuntos de domínios variáveis de cadeia leve e pesada de anticorpos (V_L e V_H, respectivamente) identificaram locais estruturalmente conservados dentro das regiões determinantes da complementaridade das sequências variáveis (CDR). O modelo para locais de zinco, catalíticos, incluem dois ligandos próximos na sequência, um ligando distante na sequência, e uma ligação de hidrogénio na cadeia principal juntando dois ligandos. Tanto na V_L como na V_H, a estereoquímica de cinco locais conservados estruturalmente gerais para todas as estruturas de anticorpos conhecidas igualaram as dos ligandos do zinco da anidrase carbónica (CAB): três resíduos em duas cadeias ß, antiparalelas, ligadas por ligações de hidrogénio.

Examinaram-se os CDR dos anticorpos para todas as ligações de hidrogénio da cadeia principal, definidas por uma distância máxima de 3,5 Å entre os átomos de azoto e de oxigénio. Utilizou-se o programa INSIGHT (da Biosym Technologies, Inc.) para sobrepor os átomos Cα e Cß dos ligandos do zinco sobre os CDR dos anticorpos, para examinar e analisar as estruturas dos anticorpos sobrepostos e para criar as figuras 7a e 7B.

A sobreposição da estrutura da V_L e V_H de sete anticorpos sobre o anticorpo McPC603 descrito por Getzoff et al., Adv. Immunol., 43: 1-98 (1988), revelou que a estrutura em folha ß conservada da região do esqueleto continuou nos CDR em todas as estruturas dos anticorpos. Estas regiões CDR possuíam as características necessárias para a estrutura conservada do local do cofactor geral com variabilidade de sequências permitida, de

modo que se podem introduzir mutações sem perturbação da estrutura global. A primeira CDR de V_L (L1) é uma volta de chave grega juntando cadeias-ß não adjacentes através da extremidade do cilindro ß, L2 é uma cadeia prolongada e curva ß (excepto em Newm), e L3 é uma continuação de duas estruturas em folha ß adjacentes com uma volta de ligação variável. No V_H, H1 é uma cadeia-ß prolongada, H2 é duas cadeias-ß adjacentes ligadas com uma volta variável e H3 é uma continuação de duas cadeias-ß com uma volta de ligação altamente variável. As regiões dos CDR L1, L3, H2 e H3 descritas no presente invento mostraram a maior variabilidade conformacional e portanto proporcionaram os locais mais preferidos para a introdução de um local de ligação a metais.

Uma pesquisa de estruturas cristalográficas no Protein Data Bank descrito por Bernstein et al., J. Mol. Biol., 112: 535-542 (1977) revelou vários motivos de ligação de zinco, catalíticos. Três resíduos, predominantemente, histidina, mas também glutamato, aspartato e cisteína, coordenaram o ião metálico numa geometria grosseiramente tetraédrica. Geralmente, pelo menos dois dos ligandos nas proteínas de ligação ao zinco estavam próximas na sequência da proteína. Uma característica estrutural adicional importante de um local de ligação a metais, não mostrada pela análise da sequência, é a presença de ligações de hidrogénio na cadeia principal ligando dois dos resíduos de aminoácido de contacto que proporcionam um ligando de metal.

A estereoquímica dos ligandos de metal de uma proteína contendo metais pode incluir três motivos de estrutura secundária com cadeias principais ligadas por ligações de hidrogénio, nomeadamente, cadeias-ß adjacentes, curvas-ß e hélices-α. Os ligandos do zinco da anidrase (CAB) descritos por Kannan et al., Proc. Natl. Acad. Sci. USA, 72: 51-55 (1975), representando o motivo de cadeia-ß, eram três de quatro resíduos em dois pares ligados por ligação de hidrogénio adjacentes, ligando duas cadeias-ß antiparalelas. Os dois ligandos sobre uma cadeia-ß estavam separadas por um resíduo intermédio. Todos os quatro resíduos nos dois pares ligados por ligações de hidrogénio,

adjacentes, tinham as suas ligações C_{α}-C_{β} apontando na direcção do ião zinco e da posição de coordenação aberta, de modo que a substituição predeterminada de quaisquer três das quatro cadeias laterais cria um local de ligação a metais. Na carboxipeptidase A descrita por Rees et al., J. Mol. Biol., 168: 367-387 (1983) e na superóxido dismutase descrita por Tainer et al., Nature, 306: 284-287 (1983), dois dos ligandos do zinco, separados por dois resíduos intermédios, estavam nas posições 1 e 4 de uma cadeia-ß do tipo I. No motivo de hélice-α visto na termolisina descrita por Holmes et al., J. Mol. Biol., 160 623-639 (1982), na elastase descrita por Thayer et al., na sua tese Ph. D., University of Colorado (1989), e na protease neutra descrita por Pauptit et al., J. Mol. Biol., 199: 525-537 (1988), dois ligandos histidina, separados por três resíduos intermédios, estavam ligados por uma ligação de hidrogénio de cadeia principal, da hélice-α. O terceiro ligando estava sobre uma hélice antiparalela adjacente.

O constrangimento das ligações de hidrogénio da cadeia principal em todas as estruturas das enzimas anteriores provocou que duas das cadeias laterais ligantes se estendessem para longe do esqueleto, aproximadamente na mesma direcção, posicionando os átomos de ligação ao zinco próximos uns dos outros. Apenas a álcool desidrogenase, descrita por Branden et al., em Enzymes, ed. Boyer, P.D. Academic Press, New York, 3rd edition, que difere dos outros locais catalíticos tanto no tipo de ligando como do espaçamento sequencial, descrito por Valee et al., Proc. Natl. Acad. Sci. USA, 87: 220-224 (1990), continha um local de zinco, catalítico, que não tinham dois ligandos por ligações de hidrogénio. Assim, as estruturas de proteínas disponíveis com locais de zinco catalíticos mostraram que uma ligação de hidrogénio da cadeia principal era importante para a disposição de ligandos adequada, enquanto que a conformação do esqueleto principal, folha-ß, curva-ß, ou hélice ß era variável. Portanto, os CDR das estruturas de anticorpos disponíveis foram examinadas quanto aos resíduos ligados por ligações de hidrogénio à cadeia principal que poderiam formar dois dos três ligandos ao metal (resíduos de aminoácidos de contacto). Sobrepuseram-se então os locais de ligação do zinco da metalo-enzima sobre os locais do

anticorpo identificados para proporcionar modelos tridimensionais para os locais de ligação a metais em anticorpos.

Dentro dos CDR V_L e V_H assim analisados, identificaram-se 11 locais para a coordenação de metais, os quais, conjuntamente, envolviam todos os seis CDR. Os 11 locais estão sumarizados na Tabela 1. Apenas se encontrou um exemplo de cada de ligandos ligados por ligações de hidrogénio com estruturas secundárias em volta ou hélice-α. Um local de curva-β geral no L2 colocou o ião zinco entre L2 e H3. Uma hélice curta em L1 lambda V deve acomodar o local de ligação de zinco, colocando um local de coordenação aberto para uma molécula de água entre L1 e L3. No entanto, o motivo cadeia-β na CAB proporcionou o melhor modelo para a construção geral de locais do metalo-anticorpo pois imitou as características estruturais conservadas nos CDR no fundo da cavidade de ligação dos anticorpos. Para os oito anticorpos examinados, os CDR, que consistem de cerca de 50 a 60 resíduos por anticorpo, continham nove locais gerais com estereoquímica similar à da CAB. Cinco destes locais incluíam as importantes características das ligações de hidrogénio: dois colocavam o zinco no centro da cavidade de ligação (o local L1-L3 e o local H1-H3), um colocava o zinco entre o L1 e o L3 (um local L3) e dois colocavam o zinco entre o H2 e o L3 (os dois locais H2).

(Segue Tabela)

TABELA 1

Locais de Ligação a Metais nos CDR de Anticorpos

Local de ligante de metal[1]	Locação do catião[2]	Ligação de H[3]	Outros Ligandos
Cadeia-ß			
L1 (34), L3 (89,91)	Centro	34-89	L1 (32)
L3 (90,92,97)	L1 e L3	90-97	L3 (95)
L3 (89,91,96)	Centro	91:96	
H1 (33,35), H3 (95)	Centro	33-95	
H1 (33,35), H2 (50)	Centro	35:50	H2 (52)
H1 (31,33), H2 (52)	Centro	33:52	
H2 (61,52,58)	H2 e L3	50-58	H2 (56)
H2 (50,58,60)	H2 e L3	50-58	
H3 (95,101x, 101x-2)[4]	Centro	95:101x	
Curva ß			
L2 (50,53,55)	L2 e H3	50...53	
Hélice α			
L1 (27d,29), L3 (93)[5]	Local de cima	27d...29	

1. A numeração entre parênteses é para a posição do resíduo de aminoácido para um resíduo de aminoácido de contacto que contribua para o local de ligação a metais. Os números referem-se ao número de posição de Kabat. Indica-se também o CDR na qual está localizado o resíduo de contacto.

2. O catião estava posicionado no centro da cavidade de ligação entre L3 e H3 (centro), ou entre os dois CDR indicados.

3. Indicam-se os pares de resíduos possuindo ligações hidrogénio conservadas em todos os anticorpos por "-" e os que possuem distâncias N-O superiores a 3,5 Å em alguns anticorpos por "...". Os pares de resíduos ligados, por ligações diferentes das ligações de hidrogénio, que ocorrem nas cadeias-ß indicam-se por ":".

4. "101x" indica o resíduo 101 ou o resíduo precedente de H3 e "101x-2" indica o resíduo localizado a duas posições de resíduos de aminoácido do 101x em direcção à extremidade amino da cadeia pesada, como ainda descrito no presente invento.

5. O local encontra-se apenas no lambda V. A posição do resíduo de aminoácido 27d no CDR L1 localiza-se a quatro posições de resíduo de aminoácidos da posição 29 em direcção à extremidade amino da cadeia leve, como ainda descrito no presente invento.

Para um local tão geral, um modelo de anticorpo substituíndo a posição 34 no primeiro CDR da cadeia leve (L1) e as posições 89 e 91 em L3 com resíduos de aminoácido de contacto histidina (ligandos) formou um local de ligação a metais com uma posição de coordenação aberta no fundo da cavidade de ligação do anticorpo. Para o anticorpo anti-fluoresceína 4-4-20, modificado como se descreveu nos exemplos 1-3, este local L1-L3 colocou o ião zinco a cerca de 4 Å da fluoresceína ligada, um indicador para a ligação de metais.

O local L1-L3 satisfaz todos os requisitos de projecto descritos no presente invento: 1) uma região de sequência variável, estruturalmente conservada, comum a todos os anticorpos, 2) um par de resíduos ligados por ligações de hidrogénio e 3) três ligandos que posicionariam o metal e o seu local exposto ao solvente para interactuar com o antigénio ligado. A ligação de hidrogénio da cadeia principal entre o resíduo 34, o último resíduo de L1, e o 89, o primeiro resíduo de L3, continuou a estrutura antiparalela da estrutura em folha-ß nos CDR. As posições da cadeia principal e da cadeia lateral no local L1-L3 assemelham-se fortemente ao local de ligação do zinco do CAB, como se viu na figura 7A. Em todos os anticorpos, os átomos C_α e C_β do resíduo 89, 91 e 34 prolongaram-se para fora a partir de um lado da folha-ß com estereoquímica similar aos resíduos 94, 96 e 119 da enzima CAB. No anticorpo HyHEL5 mostrado na figura 7A e, similarmente, nos anticorpos D13, HyHEL10, J539, e McPC603, a estrutura-ß prolongava-se ainda até L1 e L3, proporcionando um segundo par ligado por ligações de hidrogénio

entre os resíduos 32 e 91, ainda imitando a CAB. Embora a cadeia lateral do resíduo 32 de L1 apontasse para fora do local de ligação a anticorpos em vez de para dentro, a energia ganha pela ligação do metal é suficiente para posicionar o resíduo 32 como um ligando de metal. Assim, o padrão estrutural conservado do local L1-13 observado em todas as estruturas de anticorpos, indica que a substituição de quaisquer três dos resíduos 32, 34, 89 e 91, por histidinas, cria um local de ligação a metais na cavidade de ligação de anticorpos. A sobreposição dos resíduos ligantes do zinco, His, 94, 96 e 119 da CAB descrita por Kannan et al., Proc. Natl. Acad. Sci. USA, 72: 51-55 (1975), sobre os resíduos Gln 89, Trp 91 e Tyr 34 do anticorpo HyHEL5, descrito por Sheriff et al., Proc. Natl. Acad. Sci. USA, 84: 8075-8079 (1987), revelou que o ião zinco estaria posicionado perto do centro do fundo da cavidade de ligação numa orientação para actuar como um cofactor catalítico, como mostrado na figura 7A. A substituição de resíduos e a ligação a metais não têm um efeito significante na dobragem dos CDR, porque as substituições dos aminoácidos não são incomuns para estas quatro posições de resíduos do anticorpo. Por exemplo, o Tyr 34 está substituído por uma histidina em ambos os anticorpos J539 e HyHEL10.

Um local de ligação a metais V_L posicionado no fundo da cavidade, tal como um local L1-L3, interactua com um substrato cujo local de ligação seja formado por outros resíduos de anticorpos, especialmente os de H3 altamente variáveis mostrados na figura 7B. Devido à variabilidade do H3 e ao seu envolvimento em muitas interacções antigénio-anticorpo, como sumarizado por Amit et al., Science, 233: 747-753 (1986); e Herron et al., Proteins, 5: 271-280 (1989), os anticorpos quiméricos compostos de cadeias pesadas específivas para substrato e uma cadeia leve de ligação a metais, são particularmente preferidos onde o objectivo seja ligar o substrato adjacente ao ião metálico como para facilitar a catálise.

5. Projecto de anticorpo catalítico

Os locais de ligação e catalíticos de enzimas são usualmente bem conservados, enquanto que os locais de ligação de anticorpos,

feitas por CDR, são inerentemente variáveis. O exame das estruturas dos locais de iões zinco catalíticos em proteínas indicou uma estereoquímica altamente conservada incluindo uma ligação de hidrogénio da cadeia principal específica entre dois dos três ligandos. O exame de estruturas de anticorpos mostrou que as regiões dos CDR têm conformações altamente variáveis determinadas pela sequência e pelo meio. Ambas as ligações a metais e a antigénios, podem induzir rearranjos conformacionais destas regiões. Ver, por exemplo, Getzoff et al., Science, 235: 1191-1196 (1987); e Getzoff et al., Adv. Immunol., 43: 1-98 (1988). No entanto, as regiões dos CDR estruturalmente conservadas, mas de sequência variável, que são identificadas no presente invento, toleram a mutação com perturbação conformacional significativa.

O conhecimento da sequência de aminoácidos primária dos CDR de anticorpos é suficiente para construir anticorpos de ligação a metais num único passo de projecto utilizando os processos descritos no presente invento. Assim, os presentes ensinamentos proporcionam uma abordagem geral para a remodelação de regiões variáveis de anticorpos, especificamente nos CDR, para produzir locais de ligação a metal, predeterminados.

A capacidade para projectar locais de ligação a metais em áreas estruturalmente conservadas dos CDR possui largas implicações para o projecto de anticorpos catalíticos. Por exemplo, os anticorpos monoclonais humanizados descritos por Riechmann et al., Nature, 332: 323-327 (1988) que se dirigem contra alvos medicamente importantes podem-se modificar pelos presentes ensinamentos retendo a ligação, para conseguir a clivagem da ligação peptídica ou outras reacções químicas. A capacidade para construir locais de coordenação do zinco em anticorpos mostrada aqui e para criar anticorpos anti-peptídicos que reajam de forma cruzada com regiões flexíveis, expostas, de proteínas mostrada anteriormente por Tainer et al., Nature, 312: 127-133 (1984), sugerem que as proteases dirigidas a locais, específicas de sequência, se podem obter adicionando um local de ligação ao zinco ao anticorpo apropriado. Também, combinando cadeias leves

de ligação a metais preparadas pelos presentes processos, com várias bibliotecas de cadeia pesada dirigidas para diferentes substratos utilizando os processos de reportórios combinatórios de Sastry et al., Proc. Natl. Acad. Sci. USA, 86: 5728-5732 (1989), Huse et al., Science, 246: 1275-1281 81989) e Ward et al., Nature, 341: 544 (1989), estende este processo a uma vasta variedade de reacções importantes que utilizam um cofactor metálico.

A discussão anterior, incluindo as concretizações e exemplos específicos, pretende ser ilustrativa do presente invento e não se deve tomar como limitante. Podem-se efectuar numerosas outras variações e modificações sem afastamento do verdadeiro espírito e âmbito do presente invento.

R E I V I N D I C A Ç Õ E S

1 - Processo de obtenção de uma proteína isolada capaz de formar um complexo de coordenação com um catião metálico, compreendendo uma sequência de resíduos de aminoácido que define um domínio variável de uma imunoglobulina e três resíduos de aminoácidos de contacto, no referido domínio variável, que definem um local de ligação ao metal, caracterizado por compreender:

a) obter um gene codificando uma sequência de resíduos de aminoácidos que define um domínio variável de uma molécula de imunoglobulina;

b) determinar a sequência do gene para localizar as regiões que codificam as posições dos resíduos de aminoácido adequadas para introduzir uma mutação do nucleótido para codificar um resíduo de aminoácido de contacto e produzir um local de ligação a metal;

c) modificar a sequência base de nucleótidos do gene de modo a codificar resíduos de contacto metal-ligando em três das quatro posições dos resíduos de local de contacto metal-ligando, do domínio variável;

d) inserir a sequência base de nucleótidos modificada num vector de expressão;

e) expressar a proteína de ligação de metal a partir do gene modificado usando o vector de expressão num hospedeiro adequado; e

f) isolar a proteína expressa do meio de expressão.

2 - Processo de acordo com a reivindicação 1, caracterizado por o referido domínio variável ser um domínio variável de cadeia leve, possuindo uma região L1 e uma região L3 e por os referidos três resíduos de aminoácido de contacto estarem localizados em quaisquer três das quatro posições de resíduos de aminoácido 32, 34, 89 e 91.

3 - Processo de acordo com a reivindicação 1, caracterizado por o referido domínio variável ser um domínio variável de cadeia leve, possuindo uma região L3 e por os três referidos resíduos de aminoácido de contacto estarem localizados em quaisquer

três das quatro posições dos resíduos de aminoácido 90, 92, 95 e 97.

4 - Processo de acordo com a reivindicação 1, caracterizado por o referido domínio variável ser um domínio variável de cadeia leve, possuindo uma região L3 e por os três referidos resíduos de aminoácido de contacto estarem localizados nas posições de resíduos de aminoácido 89, 91 e 96.

5 - Processo de acordo com a reivindicação 1, caracterizado por o referido domínio variável ser um domínio variável de cadeia leve, possuindo uma região L2 e por os três referidos resíduos de aminoácido de contacto estarem localizados nas posições de resíduos de aminoácido 50, 53 e 55.

6 - Processo de acordo com a reivindicação 1, caracterizado por o referido domínio variável ser um domínio variável de cadeia leve, possuindo uma região L1 e uma região L3 e por os referidos três resíduos de aminoácido de contacto estarem localizados nas posições de resíduos de aminoácido 27d, 29 e 93.

7 - Processo de acordo com a reivindicação 1, caracterizado por o referido domínio variável ser um domínio variável de cadeia pesada, possuindo uma região H1 e uma região H2 e por os referidos três resíduos de aminoácido de contacto estarem localizados nas posições de resíduos de aminoácido 33, 35 e 95.

8 - Processo de acordo com a reivindicação 1, caracterizado por o referido domínio variável ser um domínio variável de cadeia pesada, possuindo uma região H1 e uma região H3 e por os referidos três resíduos de aminoácido de contacto estarem localizados em quaisquer três das quatro posições de resíduos de aminoácido 33, 35, 50 e 52.

9 - Processo de acordo com a reivindicação 1, caracterizado por o referido domínio variável ser um domínio variável de cadeia pesada, possuindo uma região H1 e uma região H2 e por os referidos três resíduos de aminoácido de contacto estarem locali-

zados nas posições de resíduos de aminoácido 31, 33 e 52.

10 - Processo de acordo com a reivindicação 1, caracterizado por o referido domínio variável ser um domínio variável de cadeia pesada, possuindo uma região H2 e por os referidos três resíduos de aminoácido de contacto estarem localizados em quaisquer três das quatro posições de resíduos de aminoácido 50, 52, 56 e 58.

11 - Processo de acordo com a reivindicação 1, caracterizado por o referido domínio variável ser um domínio variável de cadeia pesada, possuindo uma região H2 e por os referidos três resíduos de aminoácido de contacto estarem localizados nas posições de resíduos de aminoácido 50, 58 e 60.

12 - Processo de acordo com a reivindicação 1, caracterizado por o referido domínio variável ser um domínio variável de cadeia pesada, possuindo uma região H3 e por os referidos três resíduos de aminoácido de contacto estarem localizados nas posições de resíduos de aminoácido 95, 101x e 101x-2.

13 - Processo de acordo com a reivindicação 1, caracterizado por os referidos resíduos de aminoácido de contacto serem seleccionados do grupo de resíduos de aminoácido de contacto que consiste em histidina, ácido aspártico, cisteína e ácido glutâmico.

14 - Processo de acordo com a reivindicação 13, caracterizado por os referidos três resíduos de aminoácido de contacto serem histidina.

15 - Processo de acordo com a reivindicação 1, caracterizado por os referidos três resíduos de aminoácido de contacto serem duas histidinas e um resíduo de aminoácido seleccionado do grupo consistindo em histidina, ácido glutâmico, ácido aspártico e cisteína.

16 - Processo de acordo com a reivindicação 1, caracteriza-

do por compreender ainda um catião metálico complexado com a referida proteína por meio de contactos coordenados tetraédricos fornecidos pelos referidos resíduos de amincácido de contacto.

17 - Processo de acordo com a reivindicação 1, caracterizado por o referido catião metálico ser seleccionado do grupo consistindo em Cu(II), Zn(II), Ni(II), Co(II) e Cd(II).

18 - Processo de acordo com a reivindicação 17, caracterizado por o referido catião metálico ser Cu(II).

19 - Processo de acordo com a reivindicação 16, caracterizado por a referida proteína ser uma molécula de anticorpo capaz de imunorreagir com um antigénio pré-seleccionado.

20 - Processo de acordo com a reivindicação 19, caracterizado por o referido domínio variável ser um domínio variável de cadeia leve e a referida proteína estar covalentemente ligada a uma sequência de resíduos de aminoácido que definem um domínio variável de uma cadeia pesada de imunoglobulina.

21 - Processo de acordo com a reivindicação 19, caracterizado por a referida proteína estar presente como uma parte de uma molécula de anticorpo heterodimérico constituída por duas moléculas de cadeia pesada de imunoglobulina e duas moléculas de cadeia leve de imunoglobulina.

22 - Processos de acordo com a reivindicação 19, caracterizado por a referida molécula de anticorpo ser uma molécula de anticorpo catalítica capaz de promover uma reacção química predeterminada.

23 - Processo de acordo com a reivindicação 22, caracterizado por a referida reacção química predeterminada ser a clivagem de uma ligação peptídica.

72 590

SCRF Case 206.1

Lisboa, -8 MAI 1991

Por SCRIPPS CLINIC AND RESEARCH FOUNDATION

- O AGENTE OFICIAL -

FIG. 1a

```
1     M  K  K  T  A  I  A  I  A  V  A  L  A  G  F  A  T  V  A  Q
      ATGAAAAAGACAGCTATCGCGATTGCAGTGGCACTGGCTGGTTTCGCTACCGTAGCGCAG  60

21    A  D  V  V  M  T  Q  T  P  L  S  L  P  V  S  L  G  D  Q  A
      GCCGACGTCGTTATGACTCAGACACCACTATCACTTCCTGTTAGTCTAGGTGATCAAGCC 120

41    S  I  S  C  R  S  S  Q  S  L  V  H  S  N  G  N  T  Y  L  H
      TCCATCTCTTGCAGATCTAGTCAGAGCCTTGTACACAGTAATGGAAACACCTATTTACAT 180

61    W  L  L  Q  K  P  G  Q  S  P  K  V  L  I  Y  K  V  S  N  R
      TGGTTACTGCAGAAGCCAGGCCAGTCTCCAAAGGTCCTGATCTACAAAGTTTCCAACCGA 240

81    F  S  G  V  P  D  R  F  S  G  S  G  S  G  T  D  F  T  L  K
      TTTTCTGGGGTCCCAGACAGGTTCAGTGGCAGTGGATCAGGGACAGATTTCACACTCAAG 300

101   I  S  R  V  E  A  E  D  L  G  V  Y  F  C  H  Q  H  T  H  V
      ATCAGCAGAGTGGAGGCTGAGGATCTGGGAGTTTATTTCTGCCATCAACATACACATGTT 360

121   P  W  T  F  G  G  G  T  K  L  E  I  K  G  S  T  S  G  S  G
      CCGTGGACGTTCGGTGGAGGCACCAAGCTTGAAATCAAAGGTTCTACCTCTGGTTCTGGT 420
```

```
141   K  S  S  E  G  K  G  E  K  L  D  E  T  G  G  G  L  V  Q  P
      AAATCTTCTGAAGGTAAAGGTGAAAAACTGGATGAGACTGGAGGAGGCTTGGTGCAACCT 480

161   G  R  P  M  K  L  S  C  V  A  S  G  F  T  F  S  D  Y  W  M
      GGGAGGCCCATGAAACTCTCCTGTGTTGCCTCTGGATTCACTTTTAGTGACTACTGGATG 540

181   N  W  V  R  Q  S  P  E  K  G  L  E  W  V  A  Q  I  R  N  K
      AACTGGGTCCGCCAGTCTCCAGAGAAAGGACTGGAGTGGGTAGCACAAATTAGAAACAAA 600

201   P  Y  N  Y  E  T  Y  Y  S  D  S  V  K  G  R  F  T  I  S  R
      CCTTATAATTATGAAACATATTATTCAGATTCTGTGAAAGGCAGATTCACCATCTCAAGA 660

221   D  D  S  K  S  S  V  Y  L  Q  M  N  N  L  R  V  E  D  M  G
      GATGATTCCAAAAGTAGTGTCTACCTGCAAATGAACAACTTAAGAGTTGAAGACATGGGT 720

241   I  Y  Y  C  T  G  S  Y  Y  G  M  D  Y  W  G  Q  G  T  S  V
      ATCTATTACTGTACGGGTTCTTACTATGGTATGGACTACTGGGGTCAAGGAACCTCAGTC 780

261   T  V  S
      ACCGTCTCCTAATAAGGATCCAG                                       803
```

FIG. Ib

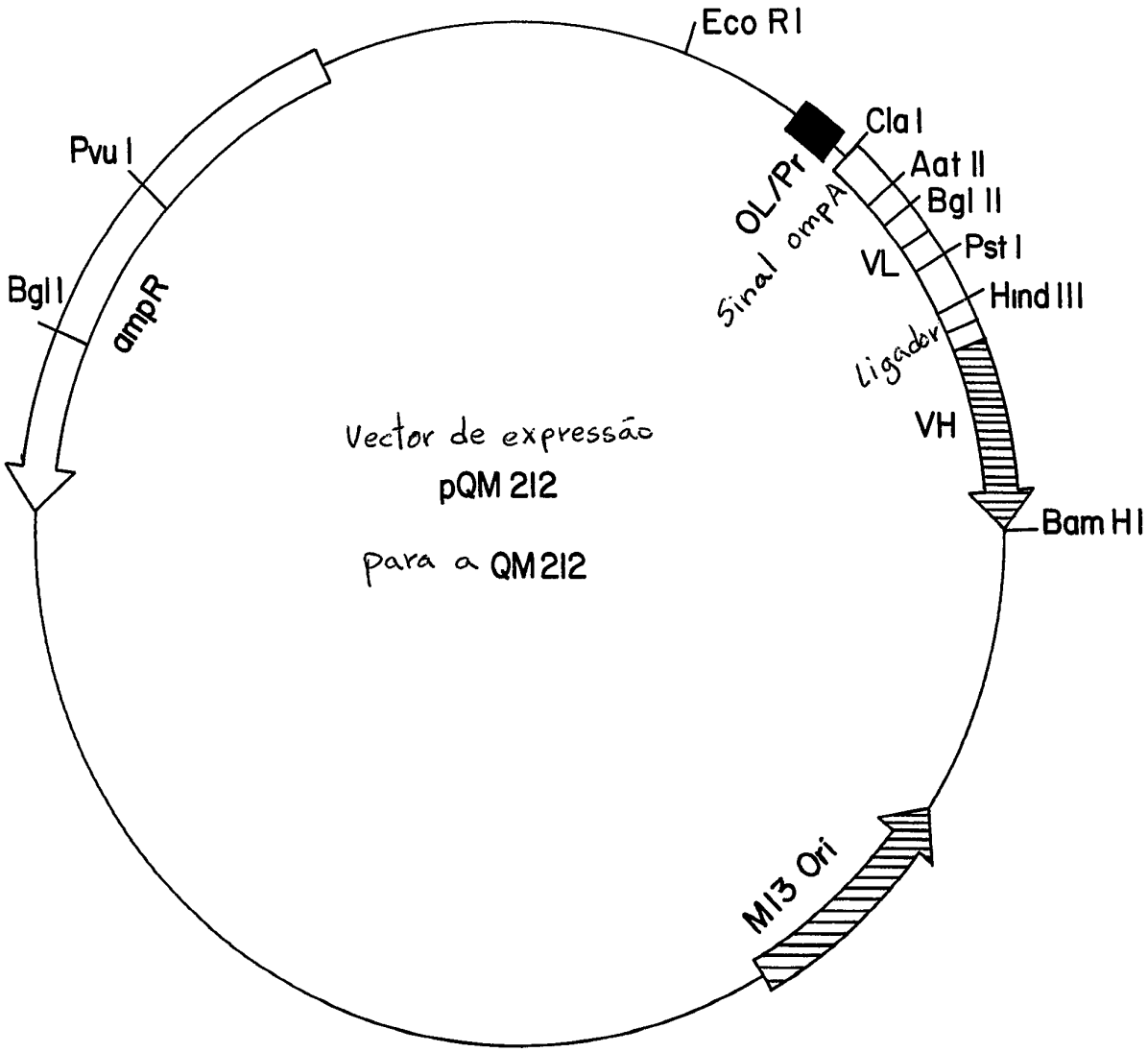
Eco RI
Cla I
Aat II
Bgl II
Pst I
Hind III
OL/Pr
Sinal ompA
VL
Ligador
VH
Bam HI
Pvu I
Bgl I
ampR
M13 Ori
Vector de expressão
pQM 212
para a QM212

FIG. 2

Interrupção da fluorescência do triptofano Cu (II)

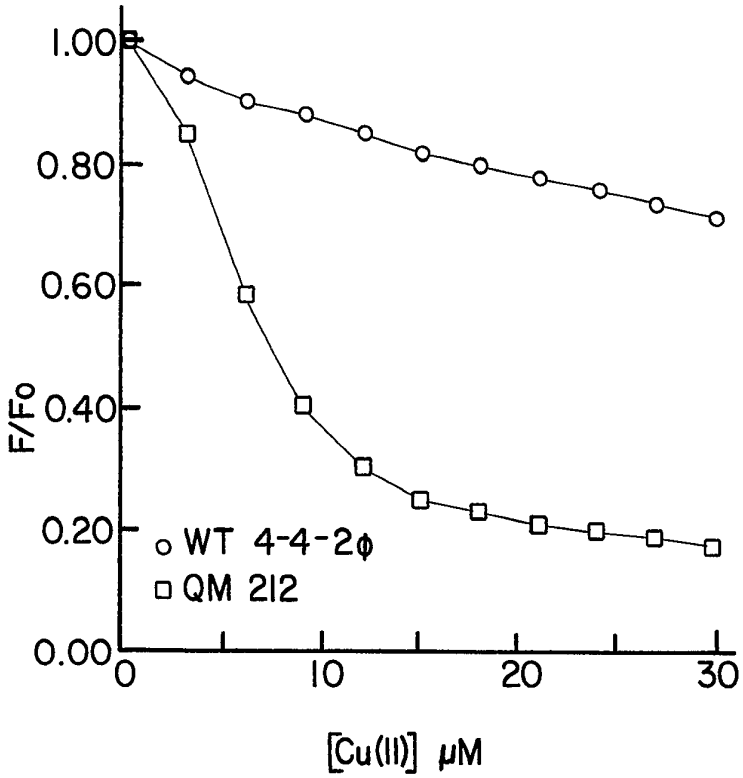

FIG. 3

Estudo da inibição do Zn (II)

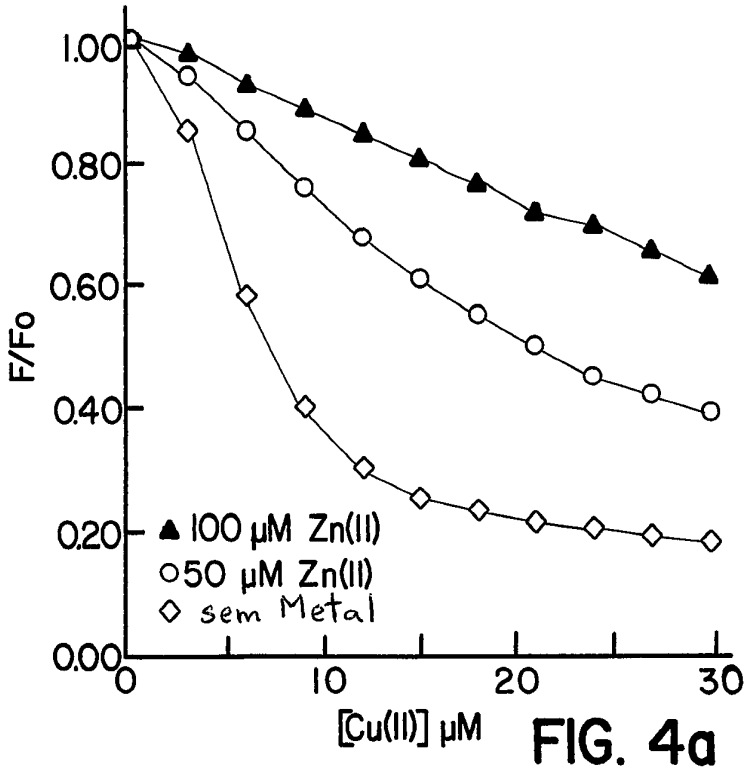

FIG. 4a

Estudo da inibição do Cd (II)

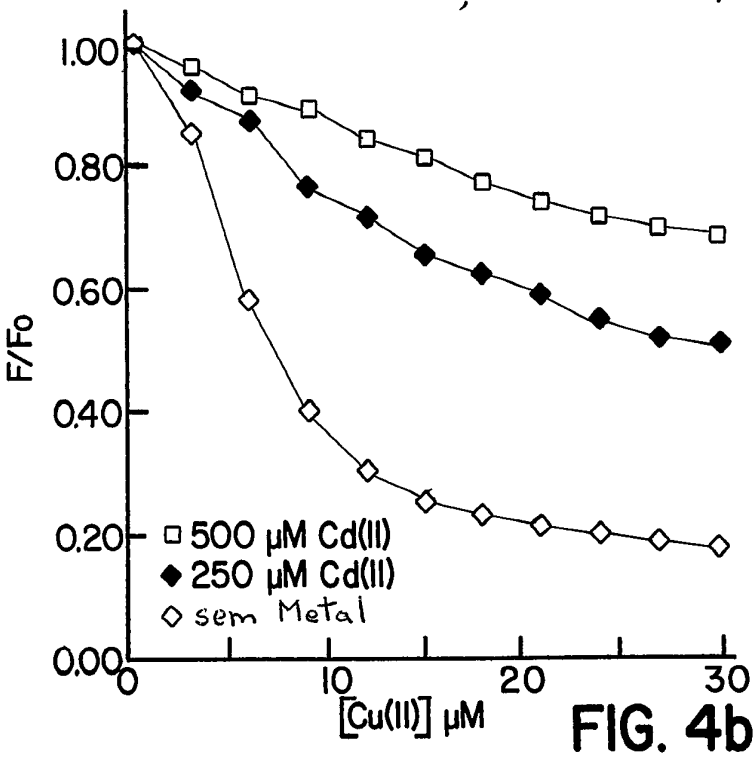

FIG. 4b

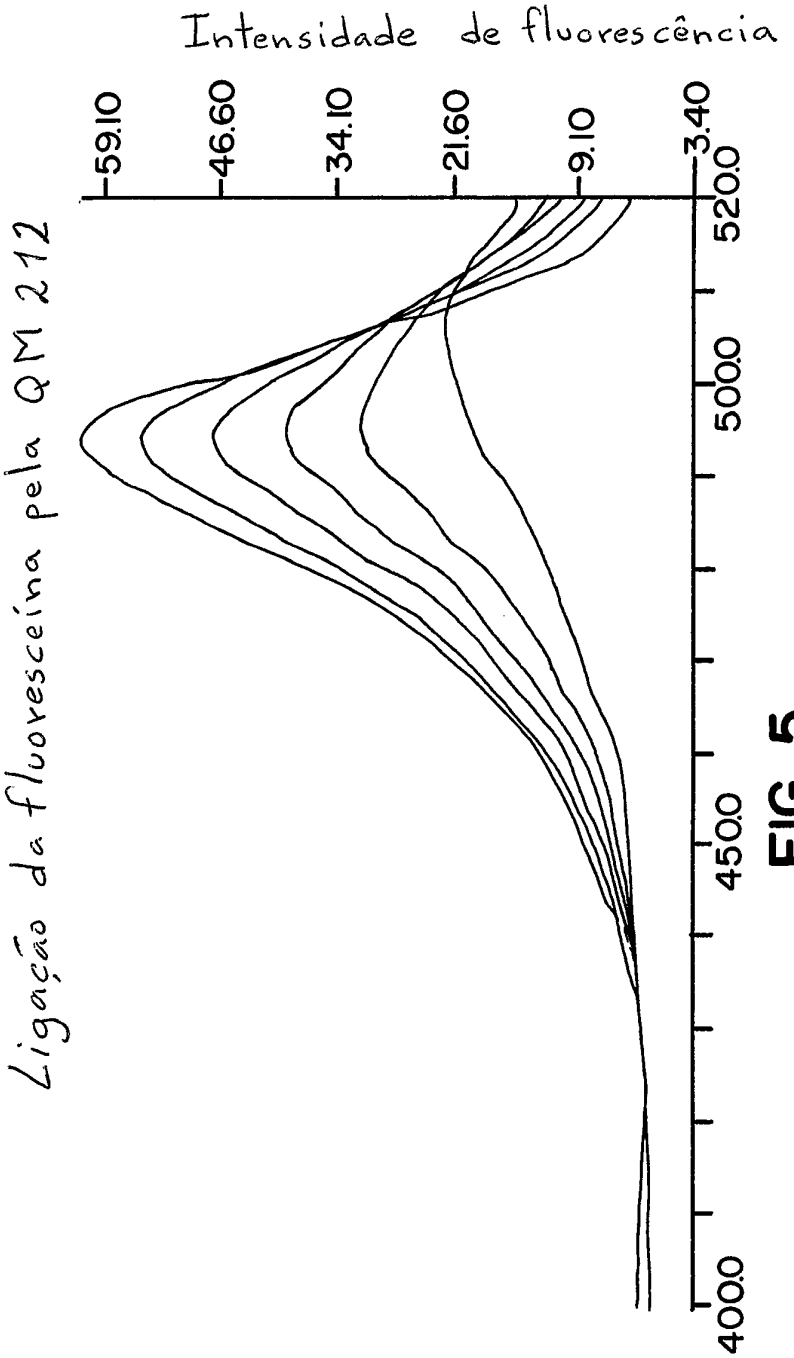
Intensidade de fluorescência
59.10
46.60
34.10
21.60
9.10
3.40
520.0
500.0
450.0
400.0
Ligação da fluoresceína pela QM 212

FIG. 5

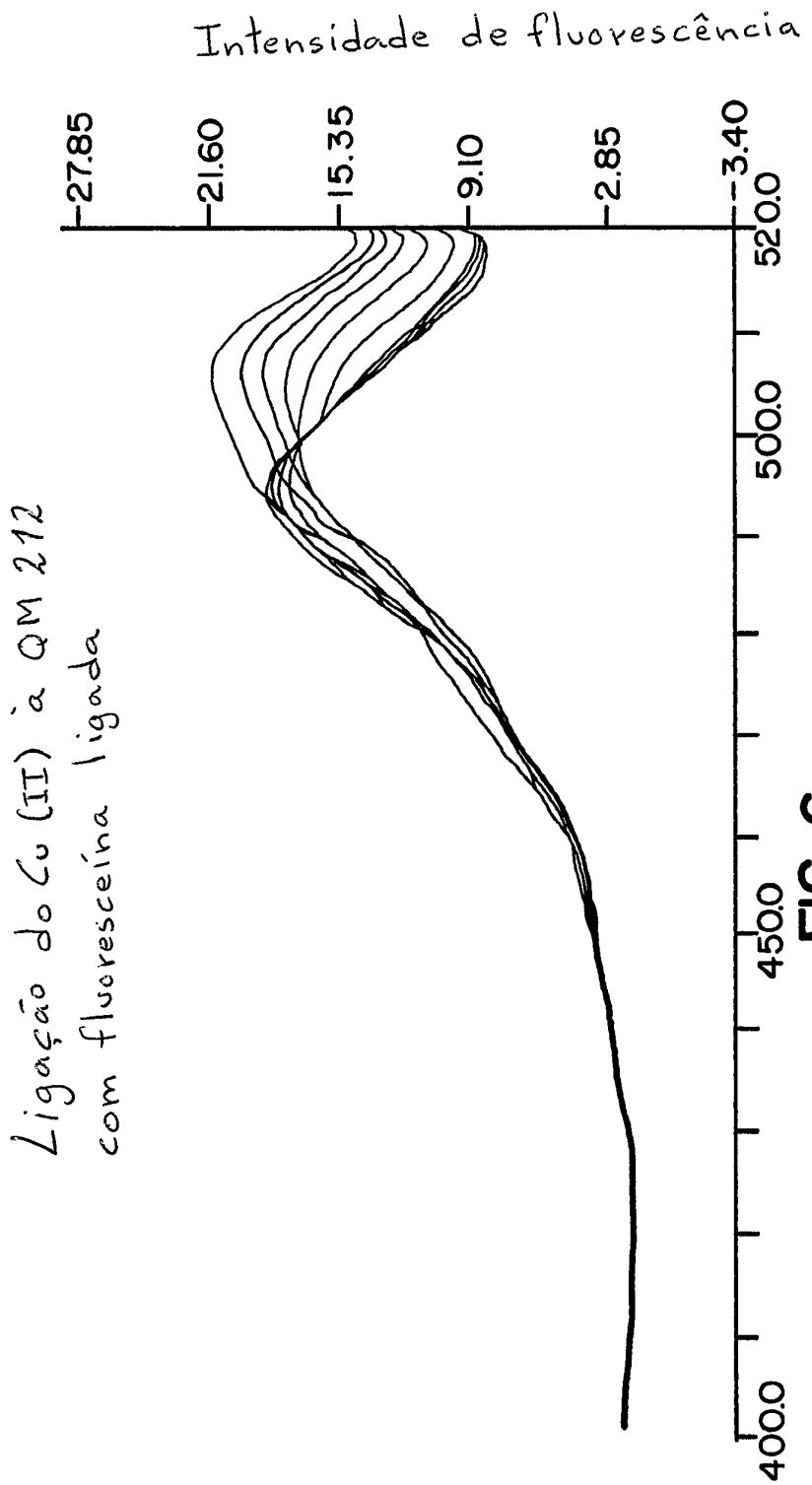
Intensidade de fluorescência
27.85
21.60
15.35
9.10
2.85
-3.40
520.0
500.0
450.0
400.0
Ligação do Cu (II) à QM 212
com fluoresceína ligada

FIG. 6

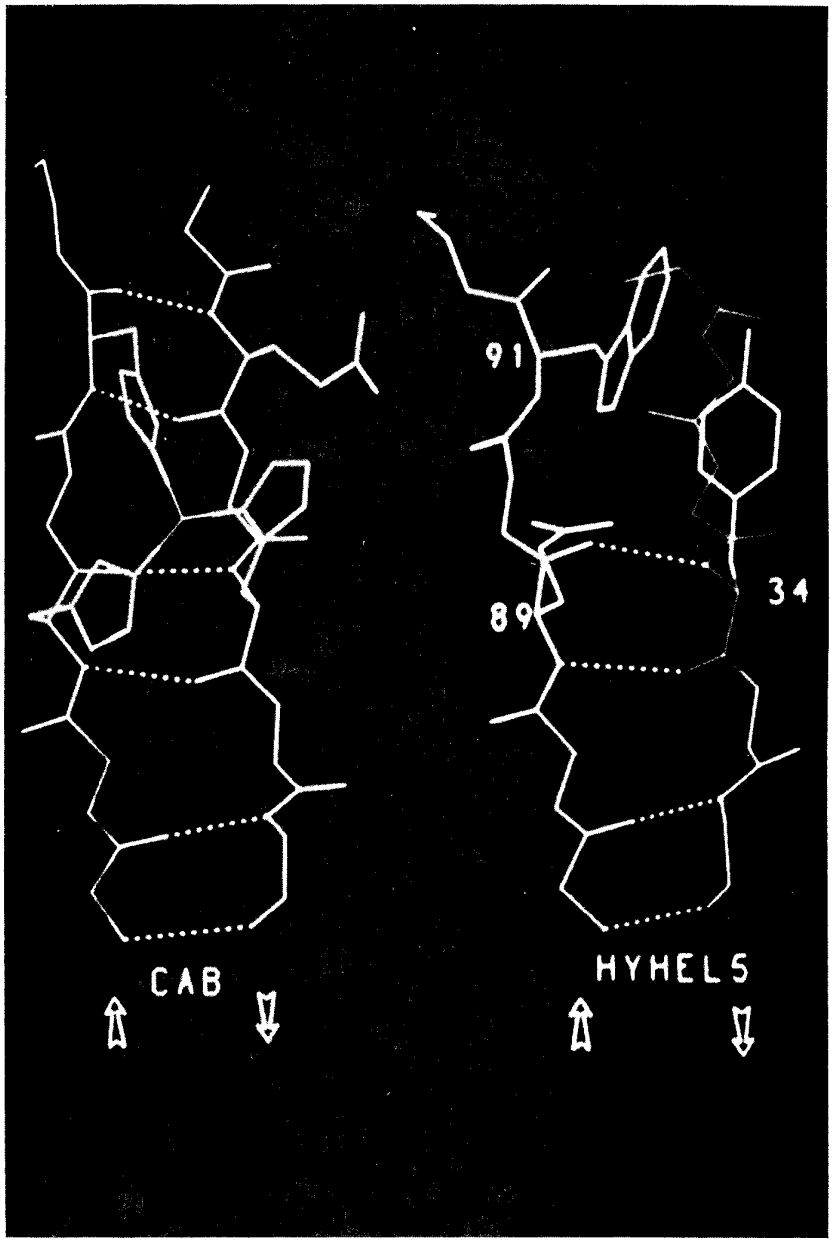

FIG. 7a



FIG. 7b