

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

(22) Data de pedido: <b>1993.08.18</b>	(73) Titular(es): <b>VRIJE UNIVERSITEIT BRUSSEL</b> <b>PLEINLAAN 2 B-1050 BRUSSELS</b> <b>BE</b>
(30) Prioridade(s): <b>1992.08.21 EP 92402326</b> <b>1993.05.21 EP 93401310</b>	
(43) Data de publicação do pedido: <b>2005.01.19</b>	(72) Inventor(es): <b>CÉCILE CASTERMAN</b> <b>BE</b> <b>RAYMOND HAMERS</b> <b>BE</b>
(45) Data e BPI da concessão: <b>2009.12.16</b> <b>055/2010</b>	(74) Mandatário: <b>MARIA SILVINA VIEIRA PEREIRA FERREIRA</b> <b>RUA CASTILHO, N.º 50, 5º - ANDAR 1269-163 LISBOA</b> <b>PT</b>

(54) Epígrafe: **IMUNOGLOBULINAS DESPROVIDAS DE CADEIAS LEVES**

(57) Resumo:

## RESUMO

### "IMUNOGLOBULINAS DESPROVIDAS DE CADEIAS LEVES"

A invenção refere-se a uma imunoglobulina modificada de 4 cadeias ou fragmento da mesma, cujas regiões  $V_H$  foram parcialmente substituídas por sequências específicas ou resíduos de aminoácidos de uma imunoglobulina (a afamada imunoglobulina de cadeia pesada) a qual compreende duas cadeias pesadas de polipéptidos, capazes de reconhecer e ligar um ou vários antígenos, sendo a referida cadeia pesada de imunoglobulina desprovida de cadeias leves.

A invenção refere-se especificamente a fragmentos modificados da região  $V_H$  de uma imunoglobulina de quatro cadeias compreendendo resíduos de aminoácidos específicos ou sequências de uma imunoglobulina (a afamada imunoglobulina de cadeia pesada) a qual compreende duas cadeias pesadas de polipéptidos capazes de reconhecer e ligar um ou mais antígenos, sendo a referida cadeia pesada de imunoglobulina desprovida de cadeias leves.

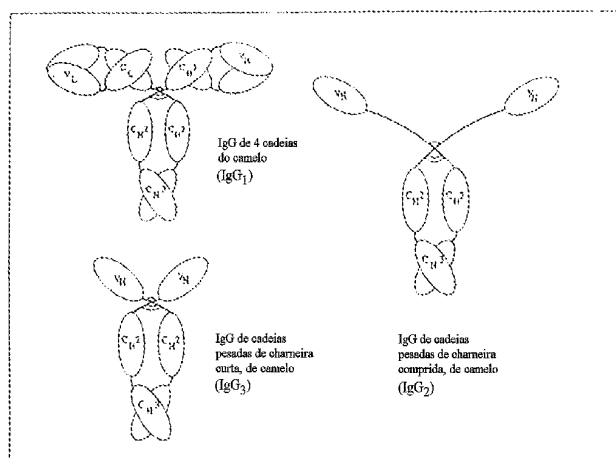


Figura 6: representação esquemática das imunoglobulinas IgG<sub>1</sub> e IgG<sub>2</sub> e IgG<sub>3</sub> putativas. A sequência (Pro-X)<sub>12</sub> da molécula IgG putativa pode ser modelada numa repetição 6 aa

### **DESCRIÇÃO**

#### **"IMUNOGLOBULINAS DESPROVIDAS DE CADEIAS LEVES"**

A invenção diz respeito a novas imunoglobulinas isoladas, as quais são desprovidas de cadeias polipeptídicas leves. Estas imunoglobulinas não são produtos de degradação de imunoglobulinas constituídas simultaneamente por cadeias polipeptídicas pesadas e cadeias polipeptídicas leves, sendo antes pelo contrário, tal como definido na invenção, um novo membro da família das imunoglobulinas, especialmente um novo tipo de moléculas capazes de participarem no reconhecimento imunitário. Tais imunoglobulinas podem ser utilizadas para diversos fins, em especial para diagnóstico e para fins terapêuticos, incluindo a protecção contra agentes patológicos ou para regular a expressão ou a actividade de proteínas.

Até ao momento presente, a estrutura proposta para as imunoglobulinas consiste de um modelo de quatro cadeias, respeitantes á presença de duas cadeias polipeptídicas leves idênticas (cadeias leves) e duas cadeias polipeptídicas pesadas idênticas (cadeias pesadas), ligadas entre si por pontes dissulfureto para formarem macromoléculas em forma de Y ou em forma de T. Estas cadeias são constituída por uma região constante e por uma região variável, sendo a região constante subdividida em vários domínios. As duas cadeias polipeptídicas pesadas estão ligadas normalmente por pontes dissulfureto numa chamada "região de charneira" situada entre o primeiro e o segundo domínios da região constante.

Entre as proteínas que formam a classe de imunoglobulinas, a maior parte delas são anticorpos e consequentemente

apresentam um local de ligação do antigénio ou vários locais de ligação dos antigénios.

De acordo com o modelo de quatro cadeias, o local de um anticorpo de ligação do antigénio está localizado nos domínios variáveis de cada uma das cadeias pesadas e leves e exige a associação dos domínios variáveis das cadeias pesadas e leves.

Para a definição destas imunoglobulinas de modelo de quatro cadeias, faz-se referência à obra de Roitt, L et al. (Immunology, segunda edição, Gower Medical Publishing USA, 1989). Merece especial referência a parte respeitante à definição das imunoglobulinas de quatro cadeias, às suas estruturas polipeptídicas e genéticas, à definição das suas regiões variáveis e constantes e à obtenção dos fragmentos produzidos por degradação enzimática, de acordo com técnicas bem conhecidas.

Os inventores concluíram surpreendentemente que é possível isolar moléculas diferentes a partir de animais que as produzem naturalmente, possuindo essas moléculas propriedades funcionais das imunoglobulinas, estando essas funções, em alguns casos, relacionadas com elementos estruturais que são distintos daqueles que estão implicados na função das imunoglobulinas de quatro cadeias, por exemplo, devido à ausência de cadeias leves.

A invenção diz respeito a imunoglobulinas segundo um modelo de duas cadeias, que não correspondem nem aos fragmentos obtidos, por exemplo, pela degradação, em particular a degradação enzimática, de uma imunoglobulina natural de modelo de quatro cadeias, nem correspondem à expressão, em células hospedeiras, do ADN que codifica a região constante ou a região variável de uma imunoglobulina natural de modelo de quatro cadeias, ou de uma parte dessas regiões,

nem corresponde aos anticorpos produzidos em linfopatias, por exemplo, em murganhos, ratos ou seres humanos.

E.S. Ward et al. (1) descreveram algumas experiências efectuadas sobre os domínios variáveis das cadeias polipeptídicas pesadas ( $V_H$ ) ou/e sobre as cadeias polipeptídicas leves ( $V_K/F_V$ ), para testar a capacidade destes domínios variáveis para se ligarem a antígenos específicas. Para o efeito, foi preparado um banco de genes  $V_H$  a partir de ADN genómico do baço de murganhos previamente imunizados com estes antígenos específicos.

Ward et al. descreveram, no trabalho que publicaram, que os domínios  $V_H$  são relativamente aderentes, presumivelmente devido à superfície hidrofóbica exposta, normalmente recoberta pelos domínios  $V_K$  ou  $V_L$ . Consequentemente admitiram que seria possível conceber domínios  $V_H$  que tivessem melhores propriedades e presumiram ainda que os domínios  $V_H$  com actividades ligantes poderiam servir de blocos de construção para a produção de fragmentos variáveis (fragmentos  $F_V$ ) ou anticorpos completos.

A publicação de Blier P.R. et al. (The Journal of Immunology, vol. 139, 3996-4006, nº 12, 15 de Dezembro de 1987) descreve sequências de nucleótidos incompletas obtidas a partir de hibridomas.

A invenção não parte da ideia de que os diferentes fragmentos (cadeias leves e pesadas) e os diferentes domínios destes fragmentos de imunoglobulina de modelo de quatro cadeias possam ser modificados para se definir novos ou melhores locais de ligação dos antígenos ou uma imunoglobulina de modelo de quatro cadeias.

Os inventores determinaram que as imunoglobulinas podem ter uma estrutura diferente daquela que está consagrada para o modelo conhecido de quatro cadeias e que essas imunoglobulinas diferentes proporcionam novos meios para a

preparação de reagentes de diagnóstico, agentes terapêuticos ou quaisquer outros reagentes utilizáveis para fins de investigação ou industriais.

Assim sendo, o presente pedido proporciona novas imunoglobulinas que são capazes de apresentar propriedades funcionais das imunoglobulinas do modelo de quatro cadeias, embora a sua estrutura pareça ser mais apropriada, em muitas circunstâncias, para a sua utilização, para a sua preparação e em alguns casos para a sua modificação. No entanto, estas moléculas podem ser consideradas como estruturas de primeiro plano para a modificação de outras imunoglobulinas. As vantagens proporcionadas por estas imunoglobulinas compreendem a possibilidade de as preparar mais facilmente.

As imunoglobulinas preparadas de acordo com a invenção são caracterizadas pelo facto de compreenderem duas cadeias polipeptídicas pesadas, suficientes para a formação de um local de ligação do antigénio completo ou de vários locais de ligação dos antígenos, sendo ainda essas imunoglobulinas desprovidas de cadeias polipeptídicas leves. Estas imunoglobulinas caracterizam-se ainda pelo facto de serem o produto da expressão, numa célula hospedeira procariótica ou eucariótica, de um ADN ou de um ADNc que possua a sequência de uma imunoglobulina desprovida de cadeias leves, susceptível de se obter a partir de linfócitos ou outras células de Camelídeos.

As imunoglobulinas descritas podem ser obtidas, por exemplo, a partir de sequências que se encontram descritas na figura 7.

As imunoglobulinas preparadas de acordo com a invenção, as quais são desprovidas de cadeias leves, são tais que os domínios variáveis das suas cadeias pesadas têm propriedades diferentes das observadas no domínio  $V_H$  das

imunoglobulinas do modelo de quatro cadeias. O domínio variável da imunoglobulina de cadeia pesada da presente invenção não tem nenhuns locais de interacções normais com os domínios  $V_L$ , ou com os domínios  $C_{H1}$  que não existem nas imunoglobulinas de cadeias pesadas. Trata-se pois de um novo fragmento em muitas das suas propriedades, tais como solubilidade e posição do local de ligação. Por razões de clareza designá-lo-emos por  $V_{HH}$  no presente texto para o distinguir do  $V_H$  clássico das imunoglobulinas de quatro cadeias.

A expressão "um local de ligação do antigénio completo", de acordo com a presente invenção, designa um local que irá permitir por si só o reconhecimento e a ligação completa de um antigénio. Isto poderia ser verificado por qualquer método conhecido respeitante aos testes de afinidade de ligação.

Estas imunoglobulinas, que podem ser preparadas pela técnica do ADN recombinante ou isoladas a partir de animais, serão designadas por vezes por "imunoglobulinas de cadeias pesadas" nas páginas subsequentes. Preferivelmente, estas imunoglobulinas encontram-se numa forma pura.

O presente pedido descreve as imunoglobulinas que podem ser obtidas em células procarióticas, especialmente em células de *E. coli*, por um processo que compreende os passos seguintes:

- a) efectuar a clonagem, num vector 'Bluescript', de uma sequência de ADN ou de ADNc que codifique o domínio VHE de uma imunoglobulina desprovida de cadeia leve, susceptível de ser obtida, por exemplo, a partir dos linfócitos de Camelídeos,
- b) recuperar o fragmento danado após a amplificação, utilizando um iniciador de 5' que contém um local *Xho* e um

iniciador de 3' que contém o local *Spe* que possui a sequência seguinte:

TC TTA ACT AGT GAG GAG ACG GTG ACC TG,

c) clonar o fragmento recuperado em fase no vector 'immuno PBS' após a digestão do vector com as enzimas de restrição *Xho* e *Spe*,

d) transformar células hospedeiras, especialmente de *E. coli*, por transfecção com o vector 'Immuno PBS' recombinante do passo c,

e) recuperar o produto de expressão da sequência de codificação  $V_{HH}$ , por exemplo, utilizando anticorpos criados contra o domínio  $V_{HH}$  dos dromedários.

O presente pedido descreve as imunoglobulinas que podem ser imunoglobulinas heteroespecíficas que podem ser obtidas por um processo que compreende os passos seguintes:

- obter uma primeira sequência de ADN ou de ADNc que codifique um domínio  $V_{HH}$  ou uma parte sua, que tenha uma especificidade determinada contra um determinado antigénio e que esteja compreendido entre os locais *Xho* e *Spe*,
- obter uma segunda sequência de ADN ou de ADNc que codifique um domínio  $V_{HH}$  ou uma parte sua, que possua uma especificidade determinada, diferente da especificidade da primeira sequência de ADN ou de ADNc, e que esteja compreendido entre os locais *Spe* e *EcoRI*,
- fazer digerir um vector 'immuno PBS' com as enzimas de restrição *EcoRI* e *XhoI*, ligar as sequências de ADN ou de ADNc obtidas, que codificam os domínios  $V_{HH}$ , de modo a que as sequências de ADN ou de ADNc sejam clonadas sequencialmente no vector,
- transformar uma célula hospedeira, em especial uma célula de *E. coli*, por transfecção e recuperar as imunoglobulinas obtidas.



O presente pedido descreve também as imunoglobulinas que podem ser obtidas por um processo que compreende os passos seguintes:

- obter uma sequência de ADN ou de ADNc que codifique um domínio V ou uma parte sua, que tenha um local de ligação do antígeno com uma especificidade determinada,
- amplificar o ADN ou o ADNc obtido, utilizando um iniciador de 5' que contenha um codão de iniciação e um local *HindIII*, e utilizando um iniciador de 3' que contenha um codão de terminação que possua um local *XhoI*,
- efectuar a recombinação do ADN ou do ADNc amplificado no interior dos locais *HindIII* (posição 2650) e *XhoI* (posição 4067) do plasmídeo pMM984,
- efectuar a transfecção de células permissivas, especialmente células NB-E, com o plasmídeo recombinante,
- recuperar os produtos obtidos.

A expressão efectuada com sucesso pode ser verificada com anticorpos dirigidos contra uma região de um domínio  $V_{HH}$ , especialmente por meio de um protocolo de EISLE (em inglês ELISA).

De acordo com outro aspecto particular deste processo, as imunoglobulinas são clonadas num parvovírus.

Segundo outro exemplo, estas imunoglobulinas descritas podem ser obtidas por um processo que compreende uma nova clonagem de uma segunda sequência de ADN ou de ADNc que possui outro local determinado de ligação do antígeno, no plasmídeo pMM984.

Uma tal imunoglobulina pode ainda ser caracterizada pelo facto de haver a possibilidade de a obter por um processo em que o vector é *Yep 52* e a célula recombinante transformada é uma levedura, em especial *S. cerevisiae*.

Há uma imunoglobulina descrita particular que se caracteriza pelo facto de possuir uma actividade

catalítica, em especial pelo facto de ser dirigida contra um antigénio que simula um estado activado de um determinado substrato. Estes anticorpos catalíticos podem ser modificados ao nível do seu local de ligação, por mutagénese aleatória ou dirigida, com a finalidade de aumentar ou de modificar a sua função catalítica. É possível referir aqui a obra publicada por Lerner *et al*, ('I IBS, Novembro de 1987, 427-430), a propósito da técnica geral para a preparação de tais imunoglobulinas catalíticas.

De acordo com uma variante descrita preferida, as imunoglobulinas são caracterizadas pelo facto de as suas regiões variáveis conterem, na posição 45, um aminoácido que é diferente de um resíduo leucina, prolina ou glutamina.

Além disso, as imunoglobulinas de cadeia pesada não são produtos característicos de linfócitos de animais nem de linfócitos de um paciente humano que padeça de linfopatias. Tais imunoglobulinas produzidas nas linfopatias são de origem monoclonal e resultam de mutações patogénicas ao nível genómico. Aparentemente não possuem nenhum local de ligação do antigénio.

As duas cadeias polipeptídicas pesadas destas imunoglobulinas podem ser ligadas por meio de uma região de charneira, de acordo com a definição de Roitt *et al*..

De acordo com uma variante particular da presente invenção, as imunoglobulinas correspondentes às moléculas anteriormente definidas são capazes de actuar como anticorpos.

O(s) local(is) das imunoglobulinas descritas, de ligação do(s) antigénio(s), estão localizado(s) na região variável da cadeia pesada.

Há um grupo particular destas imunoglobulinas em que cada cadeia polipeptídica pesada possui um local de ligação do antigénio na sua região variável e tais locais correspondem à mesma sequência de aminoácidos.

Noutra variante da presente invenção, as imunoglobulinas preparadas de acordo com a presente invenção caracterizam-se pelo facto de as suas cadeias polipeptídicas pesadas conterem uma região variável ( $V_{HH}$ ) e uma região constante ( $C_H$ ) de acordo com a definição de Roitt et al., mas serem desprovidos do primeiro domínio da sua região constante. Este primeiro domínio da região constante é designado por  $C_{H1}$ .

Estas imunoglobulinas, que não possuem nenhum domínio  $C_{H1}$ , são tais que a região variável das suas cadeias está directamente ligada à região de charneira na parte do terminal C da região variável.

As imunoglobulinas do tipo anteriormente aqui descrito podem englobar as imunoglobulinas de tipo G e em especial as imunoglobulinas que são definidas como imunoglobulinas de classe 2 (IgG2) ou imunoglobulinas de classe 3 (IgG3).

A ausência de cadeias leves e do primeiro domínio constante obriga a uma modificação da nomenclatura dos fragmentos de imunoglobulinas obtidos por digestão enzimática, de acordo com Roitt et al..

Os termos Fc e pFc, por um lado, Fc' e pFc', por outro lado, correspondentes respectivamente aos fragmentos de digestão com papaína e pepsina, são mantidos.

Os termos Fla,  $F(la)_2$ ,  $F(la')_2$ , Flac, Fd e Fv (em inglês respectivamente Fab,  $F(ab)_2$ ,  $F(ab')_2$ , Fabc, Fd e Fv) deixam de ser aplicáveis no seu sentido original, na medida em que tais fragmentos possuem em alternativa uma cadeia leve, a parte variável da cadeia leve ou o domínio  $C_{H1}$ .

Os fragmentos obtidos por digestão com papaína e constituídos pelo domínio VHE pela região de charneira passarão a ser designados por  $FV_{HH}h$  ou  $F(V_{HH}h)_2$ , dependendo disso do facto de continuarem ou não ligados por pontes dissulfureto.

Noutra variante da presente invenção, as imunoglobulinas que satisfazem às definições anteriormente aqui apresentadas podem ser obtidas a partir de animais, em especial a partir de animais da família dos Camelídeos. Os inventores concluíram que as imunoglobulinas de cadeia pesada que se encontram presentes nos Camelídeos não estão associadas a uma situação patológica que poderia induzir a produção de anticorpos anormais relativamente às imunoglobulinas de quatro cadeias. Com base num estudo comparativo entre Camelídeos do velho mundo (*Camelus bactrianus* e *Camelus dromedarius*) e Camelídeos do novo mundo (por exemplo, *Lama paccos*, *Lama glama* e *Lama vicugna*), os inventores demonstraram que as imunoglobulinas da invenção, que são desprovidas de cadeias polipeptídicas leves, existem em todas as espécies. Apesar disso, pode haver diferenças no peso molecular destas imunoglobulinas, dependendo disso dos animais. Em especial, o peso molecular de uma cadeia pesada existente nestas imunoglobulinas pode estar compreendido entre cerca de 43 kd e cerca de 47 kd, sendo em particular igual a 45 kd.

De forma vantajosa, as imunoglobulinas de cadeias pesadas da presente invenção são segregadas no sangue dos Camelídeos.

As imunoglobulinas de acordo com esta variante particular da presente invenção podem ser obtidas por purificação a partir de soro de Camelídeos, havendo um processo para a sua purificação que está descrito minuciosamente nos exemplos. No caso em que as imunoglobulinas são obtidas a

partir de Camelídeos, as imunoglobulinas não estão no seu ambiente biológico natural.

O presente pedido descreve a imunoglobulina IgG2, quando obtida por purificação a partir do soro de Camelídeos, pode ser caracterizada pelos factos seguintes:

- não é adsorvida por cromatografia na coluna de 'Proteína G-Sepharose', é adsorvida por cromatografia numa coluna de 'Proteína A-Sepharose',
- possui um peso molecular de cerca de 100 kd após a eluição com uma solução tampão a pH 4,5 (NaCl 0,15 M com ácido acético a 0,58%, ajustada para o valor de pH igual a 4,5 com NaOH),
- é constituída por cadeias polipeptídicas  $\gamma_2$  pesadas com um peso molecular de cerca de 46 kd e preferencialmente 45 kd após redução.

De acordo com o presente pedido, descreve-se um outro grupo de imunoglobulinas, correspondente á IgG3, que podem ser obtidas por purificação a partir do soro de Camelídeos, que se caracterizam pelo facto de uma dessas imunoglobulina ter as particularidades seguintes:

- é adsorvida por cromatografia numa coluna de Proteína A-Sepharose',
- tem um peso molecular de cerca de 100 kd após a eluição com uma solução tampão a pH 3,5 (NaCl 0,15 M com ácido acético a 0,58%),
- é adsorvida por cromatografia numa coluna de 'Proteína G-Sepharose' e a sua eluição tem lugar com uma solução tampão a pH 3,5 (NaCl 0,15 M e ácido acético a 0,58%),
- é constituída por cadeias polipeptídicas  $\gamma_3$  pesadas com um peso molecular de cerca de 45 kd e em particular compreendido entre 43 kd e 47 kd após redução.

Apesar disso, as imunoglobulinas preparadas de acordo com a presente invenção, que são desprovidas de cadeias leves,

possuem nas suas cadeias pesadas uma região constante e uma região variável. A região constante é constituída por diferentes domínios.

A região variável destas imunoglobulinas compreende armações estruturais (AE, em inglês FW) e regiões determinantes da complementaridade (RDC, em inglês CDR), em especial quatro armações estruturais e três regiões de complementaridade. Distinguem-se das imunoglobulinas de quatro cadeias especialmente pelo facto de esta região variável poder conter, por si própria, um ou vários locais de ligação dos antigénios, sem contribuição da região variável de uma cadeia leve que neste caso não existe.

As sequências de aminoácidos das armações estruturais 1 e 4, compreendem, entre outras, as sequências de aminoácidos que podem ser seleccionadas respectivamente entre as seguintes:

para o domínio da armação estrutural 1

```

G G S V Q T G G S L R L S C E I S G L T F D
G G S V Q T G G S L R L S C A V S G F S F S
G G S E Q G G G S L R L S C A I S G Y T Y G
G G S V Q P G G S L T L S C T V S G A T Y S
G G S V Q A G G S L R L S C T G S G F P Y S
G G S V Q A G G S L R L S C V A G F G T S
G G S V Q A G G S L R L S C V S F S P S S

```

para o domínio da armação estrutural 4

```

W G Q G T Q V T V S S
W G Q G T L V T V S S
W G Q G A Q V T V S S
W G Q G T Q V T A S S
R G Q G T Q V T V S L

```

para o domínio da RDC3

```

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

```

Conforme se disse antes, as imunoglobulinas preparadas de acordo com a presente invenção são preferencialmente desprovidas da totalidade do seu domínio C<sub>H1</sub>.

Tais imunoglobulinas possuem domínios C<sub>H2</sub> e C<sub>H3</sub> na região do terminal C relativa à região de charneira.

De acordo com uma variante descrita particular, a região constante das imunoglobulinas compreende os domínios C<sub>H2</sub> e C<sub>H3</sub> constituídos por uma sequência de aminoácidos seleccionada entre as seguintes:

para o domínio C<sub>H2</sub>:

APELLGGPTVFIFPPKPKDVLSITLTP  
 APELPGGPSVFVFPTKPKDVLSISGRP  
 APELPGGPSVFVFPPKPKDVLSISGRP  
 APELLGGPSVFIFPPKPKDVLSISGRP

para o domínio C<sub>H</sub>3:

GQTRPQVYTLA  
 GQTRPQVYTLAPXRLEL  
 GQPREPQVYTLPPSRDEL  
 GQPREPQVYTLPPSREEM  
 GQPREPQVYTLPPSQEEM

Os inventores demonstraram, de forma interessante, que a região de charneira das imunoglobulinas da invenção pode apresentar comprimentos variáveis. Quando estas imunoglobulinas actuam como anticorpos, o comprimento da região de charneira vai contribuir para a determinação da distância que separa os locais de ligação dos antigénios.

Uma imunoglobulina descrita aqui caracteriza-se, preferencialmente, pelo facto de a sua região de charneira possuir entre 0 e 50 aminoácidos.

As sequências particulares da região de charneira das imunoglobulinas descritas são as seguintes:

GTNEVCKCPKCP

ou

EPKIPQPKPQPKPQPKPQPKPQPKPEPECTCPKCP.

A região de charneira curta corresponde a uma molécula de IgG3 e a sequência de charneira comprida corresponde a uma molécula de IgG2.

A V<sub>HH</sub> isolada, obtida a partir de imunoglobulinas de cadeias pesadas, ou os bancos de V<sub>HH</sub>, correspondentes às imunoglobulinas de cadeias pesadas, podem distinguir-se da clonagem V<sub>HH</sub> de imunoglobulinas do modelo de quatro



cadeias, com base nas particularidades das sequências que caracterizam as imunoglobulinas de cadeias pesadas.

A região V da imunoglobulina de cadeias pesadas do camelo apresenta diversas diferenças em relação às regiões  $V_{HH}$  obtidas a partir de imunoglobulinas do modelo de quatro cadeias pertencentes a todas as espécies examinadas. Ao nível dos resíduos implicados nas interações de tipo  $V_{HH}/V_L$ , foi observada uma diferença importante na posição 45 (AE) que é praticamente sempre leucina nas imunoglobulinas de quatro cadeias (98%), sendo os outros aminoácidos nesta posição a prolina (1%) ou a glutamina (1%).

Na imunoglobulina de cadeias pesadas do camelo, nas sequências examinadas até ao momento presente, apenas se encontrou uma vez o resíduo leucina na posição 45. Poderia ter a sua origem numa imunoglobulina de quatro cadeias. Nos outros casos, este resíduo é substituído por um resíduo arginina, cisteína ou ácido glutâmico. A presença de aminoácidos com carga eléctrica nesta posição deve contribuir para fazer com que a região  $V_{HH}$  seja mais solúvel.

A substituição por resíduos específicos dos Camelídeos, tais como os da posição 45, parece ser interessante para a construção de regiões  $V_{HH}$  engendradas artificialmente, obtidas a partir do repertório de  $V_{HH}$  das imunoglobulinas de quatro cadeias.

Uma segunda particularidade específica do domínio  $v_{HH}$  dos camelídeos é a presença frequente de um resíduo cisteína da  $RDC_3$  associado a um resíduo cisteína na posição 31 ou 33 da RDC ou na posição 45 da região  $AS_2$  (em inglês  $FW_2$ ). A possibilidade de se estabelecer uma ponte dissulfureto entre a  $RDC_3$  e a parte restante do domínio variável poderia contribuir para a estabilidade e para o posicionamento do local de ligação.

Com a exceção de uma única proteína patogénica de mieloma (DAW), nunca uma tal ponte dissulfureto foi encontrada nas regiões V das imunoglobulinas, derivadas de imunoglobulinas de quatro cadeias.

As imunoglobulinas de cadeias pesadas preparadas de acordo com a presente invenção têm também a vantagem particular de não serem viscosas. Assim sendo, estando estas imunoglobulinas presentes no soro, agregam-se muito menos do que as cadeias pesadas isoladas das imunoglobulinas de quatro cadeias. As imunoglobulinas descritas são solúveis até uma concentração superior a 0,5 mg/mL, de preferência superior a 1 mg/mL e mais vantajosamente superior a 2 mg/mL.

Estas imunoglobulinas também são portadoras de um exaustivo repertório de ligação de antígenos e experimentam uma maturação de afinidade e de especificidade *in vivo*. Em consequência, permitem isolar e preparar anticorpos com uma especificidade definida, tendo em conta determinados antígenos.

Uma outra propriedade interessante das imunoglobulinas descritas é o facto de poderem ser modificadas e em especial poderem ser humanizadas. De uma forma peculiar, é possível substituir a totalidade ou uma parte da região constante destas imunoglobulinas pela totalidade ou por uma parte de uma região constante de um anticorpo humano. Por exemplo, os domínios C<sub>H</sub>2 e/ou C<sub>H</sub>3 da imunoglobulina podem ser substituídos pelos domínios C<sub>H</sub>2 e/ou C<sub>H</sub>3 da imunoglobulina IgG humana de cadeia  $\gamma$ 3.

Em tais anticorpos humanizados também é possível substituir uma parte da sequência variável, designadamente um ou vários dos resíduos da sequência estrutural que não interfiram no local de ligação, por resíduos humanos da

armação estrutural, ou por uma parte de um anticorpo humano.

Reciprocamente, de acordo com a presente invenção, poderiam ser introduzidas particularidades (em especial fragmentos peptídicos) das regiões VIM das imunoglobulinas de cadeia pesada, nas regiões  $V_H$  ou  $V_L$  obtidas a partir de imunoglobulinas de quatro cadeias, por exemplo, com a finalidade de se conseguir uma maior solubilidade das imunoglobulinas.

A invenção também diz respeito a um processo para a preparação de um fragmento prode uma imunoglobulina do tipo aqui anteriormente descrito, e em especial diz respeito de um fragmento seleccionado entre o conjunto seguinte e diz respeito a fragmentos assim obtidos:

- um fragmento correspondente a uma cadeia polipeptídica pesada de uma imunoglobulina desprovida de cadeias leves,
- fragmentos obtidos por digestão enzimática das imunoglobulinas da invenção, em especial aqueles que são obtidos por digestão parcial com papaína, dando origem ao fragmento Fc (fragmento constante) e dando origem ao fragmento  $FV_{HHh}$  (que contém locais das cadeias pesadas onde se ligam os antigénios) ou o seu dímero  $F(V_{HHh})_2$  ou um fragmento obtido por meio de outra digestão do fragmento Fc com papaína, dando origem ao fragmento pFc correspondente à parte do terminal C do fragmento Fc,
- fragmentos homólogos obtidos com outras enzimas proteolíticas,
- um fragmento que possua pelo menos 20 aminoácidos da região variável da imunoglobulina, ou a região variável completa, em especial um fragmento correspondente aos domínios  $V_{HH}$  isolados ou aos dímeros de  $V_{HH}$  ligados à ponte dissulfureto da charneira,

- um fragmento corresponde pelo menos a 10 e de preferência a 20 aminoácidos da região constante, ou que corresponda à região constante completa da imunoglobulina. Os fragmentos podem ser obtidos por degradação enzimática ou não enzimática das imunoglobulinas. Também podem ser obtidos por expressão, em células ou em organismos, da sequência de nucleótidos que codifica as imunoglobulinas, ou podem ser sintetizados por via química.

O presente pedido também descreve anticorpos anti-idiotípicos pertencentes às classes das imunoglobulinas de cadeias pesadas. Tais anticorpos anti-idiotípicos podem ser produzidos contra idiotipos humanos ou animais. Uma propriedade destes anticorpos anti-idiotípicos é o facto de poderem ser utilizados como vacinas idiotípicas, em particular para a vacinação contra glicoproteínas ou glicolípidos e se o hidrato de carbono determinar o epítipo.

O presente pedido também descreve anticorpos anti-idiotípicos capazes de reconhecerem os idiotipos de imunoglobulinas de cadeias pesadas.

Tais anticorpos anti-idiotípicos podem ser anticorpos singeneicos ou anticorpos alogeneicos ou anticorpos xenogeneicos.

O presente pedido também descreve sequências de nucleótidos que codificam a totalidade ou uma parte de uma proteína, compreendendo essa sequência de aminoácidos uma sequência de péptidos seleccionados entre os seguintes:

```

G G S V Q T G G S L R L S C E I S G L T F D
G G S V Q T G G S L R L S C A V S G F S F S

```

GGSEQGGGSLRLSCAISGYTYG  
 GGSVQPGGSLTLSCTVSGATYS  
 GGSVQAGGSLRLSCTGSGFPYS  
 GGSVQAGGSLRLSCVAGFGTS  
 GGSVQAGGSLRLSCVSFSPSS

WGQGTQVTVSS  
 WGQGTLVTVSS  
 WGQGAQVTVSS  
 WGQGTQVTASS  
 RGQGTQVTVSL

ALQFGG YCGYGX - - - - - CL  
 VSLMDRISQH - - - - - GC  
 VPAHLGPGAILDLKKY - - - - - KY  
 FCYSTAGDGGSGE - - - - - MY  
 ELSGGSCELPLLF - - - - - DY  
 DWKYWTCGAQTGGYF - - - - - GQ  
 RLTEMGACDARWATLATRTFAYNY  
 QKKDRTRWAEPREW - - - - - NN  
 GSRFSSPVGSTSRLES - SDY - - NY  
 ADPSIYYYSILXIEY - - - - - KY  
 DSPCYMPTMPAPPIRDSPFGW - - DD  
 TSSFYWYCTTAPY - - - - - NV  
 TEIEWYGCNLRRTTF - - - - - TR  
 KQLAGGWYLDPNYWLSVGAY - - AI  
 RLTEMGACDARWATLATRTFAYNY  
 DGWTRKEGGIGLPKSVQCEDGYNY  
 DSYPCHELL - - - - - DV  
 VEYPIADMCS - - - - - RY

APELLGGPSVFVFPPKPKDVLSISGXPK  
 APELPGGPSVFVFPTKPKDVLSISGRPK  
 APELPGGPSVFVFPPKPKDVLSISGRPK  
 APELLGGPSVFIFPPKPKDVLSISGRPK  
 GQPREPQVYTLAPXRLEL  
 GQPREPQVYTLPPSRDEL  
 GQPREPQVYTLPPSREEM  
  
 GQPREPQVYTLPPSQEEM  
 VTVSSGTNEVCKCPKCPAPELPGGPSVFVFP

ou

VTVSSEPKIPQPKPKQPKQPKQPKQPKQPKPEPECTCPKCPAPELLGGPSVFIFP  
 GTNEVCKCPKCP  
 APELPGGPSVFVFP  
 EPKIPQPKPKQPKQPKQPKQPKQPKPEPECTCPKCP  
 APELLGGPSVFIFP

Tais sequências de nucleótidos podem ser deduzidas das sequências de aminoácidos, tendo em consideração a degeneração do código genético. Podem ser sintetizadas ou isoladas a partir de células que produzam as imunoglobulinas da invenção.

Nos exemplos está descrito um procedimento para a obtenção de tais sequências de ADN.

O presente pedido também descreve sequências de ARN, em especial as sequências de ARNm correspondentes a essas sequências de ADN e também correspondentes às sequências de ADNC.

As sequências de nucleótidos ainda podem ser utilizadas para a preparação de iniciadores adequados para detectar ou para pesquisar em células bancos de ADN ou de ADNC para isolar sequências de nucleótidos que codifiquem as imunoglobulinas da invenção.

Tais sequências de nucleótidos podem ser utilizadas para a preparação de vectores recombinantes e para a expressão

dessas sequências, contidas nos vectores, por células hospedeiras, em especial células procarióticas, tais como células de bactérias, ou ainda por células eucarióticas e também, por exemplo, células de OCC (em inglês CHO), células de insectos, células de símios, tais como as células Vero, ou quaisquer outras células de mamíferos. O facto de as imunoglobulinas da invenção estarem desprovidas de cadeias leves permite, em especial, segregá-las em células eucarióticas, uma vez que não é necessário recorrer ao passo que consiste na formação da proteína BIP que é necessária nas imunoglobulinas de quatro cadeias.

As insuficiências dos métodos conhecidos para a produção de anticorpos monoclonais ou imunoglobulinas pela tecnologia do ADN recombinante resultam da necessidade, na grande maioria dos casos, de se clonar simultaneamente os domínios VH e VL correspondentes ao local de ligação específico das imunoglobulinas de quatro cadeias. Os animais, e especialmente os camelídeos que produzem imunoglobulinas de cadeias pesadas de acordo com a invenção, e possivelmente outras espécies de vertebrados, são capazes de produzir imunoglobulinas de cadeias pesadas nas quais o local de ligação está localizado exclusivamente no domínio  $V_{HH}$ . Ao contrário de algumas imunoglobulinas de cadeias pesadas produzidas noutras espécies por separação de cadeias ou por clonagem directa, as imunoglobulinas de cadeias pesadas dos Camelídeos foram já objecto de uma maturação exaustiva *in vivo*. Além disso, a sua região V evoluiu naturalmente para funcionar na ausência do domínio  $V_L$ . Por tal motivo são ideais para a produção de anticorpos monoclonais por meio da tecnologia do ADN recombinante. Uma vez que a obtenção de clones específicos de ligação dos antígenos não depende de um processo estocástico que necessite de um número muito

grande de células recombinantes, isto permite também um exame muito mais exaustivo do repertório.

Isto pode ser feito ao nível do repertório de  $V_{HH}$  não rearranjado, utilizando ADN proveniente de um tipo de células ou de tecidos arbitrariamente escolhidos, ou ao nível do repertório de  $V_{HH}$  rearranjado, utilizando o ADN obtido a partir dos linfócitos B. No entanto, é muito mais interessante transcrever o ARNm proveniente das células produtoras dos anticorpos e efectuar a clonagem de ADN, com ou sem amplificação prévia, num vector adequado. Daqui irá resultar a obtenção de anticorpos que tenham já sido objecto de maturação por afinidade.

O exame de um grande repertório deve demonstrar que é particularmente útil na pesquisa de anticorpos com actividades catalíticas.

Assim, o presente pedido descreve bancos que podem ser gerados de uma forma que inclua parte da sequência de charneira e a identificação é simples uma vez que a charneira está directamente ligada ao domínio  $V_{HH}$ .

Estes bancos podem ser obtidos por clonagem de ADN a partir de células linfóides, com ou sem amplificação prévia por RCP. Os iniciadores de RCP estão localizados nas sequências promotoras, líderes ou da armação estrutural do domínio V no caso do iniciador de 5' e nas regiões de charneira,  $C_H2$ ,  $C_H3$  e não traduzida em 3' ou no apêndice poli-A no caso do iniciador de 3'. Uma selecção dimensional do material amplificado permite construir um banco limitado às imunoglobulinas de cadeia pesada.

Como um exemplo particular, utilizou-se o iniciador de 3' a seguir apresentado, no qual foi construído um local *KpnI* e o qual corresponde aos aminoácidos 313 a 319 (CGC CAT CAA GGT AAC AGT TGA), em conjunto com os iniciadores de  $V_{HH}$  de murganho, descritos por Sestry e que contêm um local *Xho*



AG GTC CAG CTG CTC GAG TCT GG  
 AG CTC CAG CTG CTC GAG TCT GG  
 AG GTC CAG CTT CTC GAG TCT GG

local *Xho*I

Estes iniciadores geraram um banco de imunoglobulinas de cadeias pesadas de Camelídeos que compreendem a região VHE (relacionada com o subgrupo 111 de murganho ou de seres humanos), a charneira e uma secção de CH<sub>2</sub>.

De acordo com outro exemplo, o ADNc é poliadenilado na sua extremidade 5' e os iniciadores de V específicos de murganho são substituídos por um iniciador poli-T e um local *Xho*I permanente, ao nível do nucleótido 12:

CTCGAGT<sub>12</sub>.

Utiliza-se o mesmo iniciador de 3' com um local *Kpn*I.

Este método gera um banco que contém todos os subgrupos de imunoglobulinas.

Uma parte do interesse na clonagem de uma região que abranja a ligação charneira-CH<sub>2</sub> é o facto de haver nas duas cadeias  $\gamma_2$  e  $\gamma_3$  um local *Sac* imediatamente a seguir à charneira. Este local permite enxertar a sequência codificadora da região V<sub>HH</sub> e a charneira na região Fe de outras imunoglobulinas, em particular nas IgG1 e IgG3 humanas que têm a mesma sequência de aminoácidos neste local (Glu<sub>246</sub> Leu<sub>247</sub>).

Como exemplo, o presente pedido descreve um banco de ADNc constituído por sequências de nucleótidos que codificam uma imunoglobulina de cadeias pesadas, tal como a que se obtém executando os passos seguintes:

a) tratar uma amostra que contenha células linfóides, em especial linfócitos periféricos, esplenócitos, nodos linfáticos ou outros tecidos linfóides provenientes de um animal saudável, escolhido especialmente entre os

Camelídeos, com a finalidade de separar as células linfóides,

b) separar o ARN poliadenilado de outros ácidos nucleicos e de outros componentes das células,

c) fazer reagir o ARN assim obtido com uma transcriptase reversa, com a finalidade de se obter o ADNc correspondente,

d) fazer contactar o ADNc do passo c) com os iniciadores de 5' correspondentes ao domínio  $V_H$  do murganho, da imunoglobulina de quatro cadeias, possuindo esse iniciador um determinado local de restrição, por exemplo, o local *XhoI*, e com os iniciadores de 3' correspondentes à parte do terminal N de um domínio  $C_H2$  que contenha um local *KpnI*,

e) amplificar o ADN,

f) clonar num vector a sequência amplificada, em especial num vector 'Bluescript',

g) recuperar os clones que hibridem com uma sonda correspondente à sequência codificadora de um domínio constante de uma imunoglobulina de cadeias pesadas que tenha sido isolada.

Esta clonagem permite obter clones que contêm sequências de ADN que compreendem a sequência codificadora da charneira. Esta clonagem permite pois a caracterização da subclasse da imunoglobulina e o local *SacI* útil para enxertar o fragmento  $FV_{HH}h$  na região  $Fc$ .

Também é possível efectuar a recuperação das sequências que codificam as imunoglobulinas de cadeias pesadas, por meio da selecção de clones que contenham as sequências de ADN que possuem um local compatível com a falta do domínio  $C_H1$ . De acordo com outra variante do presente pedido, é possível acrescentar, entre os passos c) e d) do processo anterior, os passos seguintes:

- na presença de uma polimerase de ADN e de desoxirribonucleótido-trifosfatos, fazer contactar o referido ADNc com iniciadores oligonucleotídicos degenerados, sendo essas sequências capazes de codificar a região de charneira e o domínio V<sub>HH</sub> do terminal N de uma imunoglobulina, sendo os iniciadores capazes de hibridar com o ADNc e capazes de iniciar o prolongamento de uma sequência de ADN complementar do ADNc utilizado como matriz,
- recuperar o ADN amplificado.

Os clones podem ser expressos em diversos tipos de vectores de expressão. Como exemplo, utilizando um vector 'Immuno PBS' comercialmente disponível (Huse *et al.*: Science (1989) 246, 1275), os clones produzidos em 'Bluescriptf, de acordo com o procedimento anteriormente descrito, são recuperados por RCP utilizando o mesmo iniciador de 5' que contém XhoI e um novo iniciador de 3', correspondente aos resíduos 113-103 na armação estrutural das imunoglobulinas, onde foi construído um local Spe: TC TTA ACT AGT GAG GAG ACG GTG ACC TG. Este procedimento permite a clonagem da região V<sub>HH</sub> no local Xho/Spe do vector 'Immuno PBS'. No entanto, a extremidade 3' do gene não está em fase com o "marcador" de identificação nem com o codão de paragem do vector. Para se conseguir isto, corta-se o arquétipo com Spe e faz-se o preenchimento com os apêndices de 4 bases, utilizando o fragmento de Klenow, após o que o vector é novamente ligado. Uma outra operação de refinamento consiste em substituir o identificador ("marcador") com uma poli-histidina, de modo a que seja realizada a purificação de metais da região V<sub>ini</sub> clonada. Para se conseguir isto, constrói-se em primeiro lugar um oligonucleótido de cadeia dupla SpeI/EcoRI que codifica 6 resíduos histidina e um

codão de terminação, por síntese das duas cadeias, a que se segue o tratamento térmico e a recombinação:

CTA GTG CAC CAC CAT CAC CAT CAC TAA\* TAG\*

AC GTG GTG GTA GTG GTA GTG ATT ATC TTA A.

O vector que contém o segmento intercalar é então digerido com SpeI e EcoRI para a remoção da sequência "marcadora" residente a qual pode ser substituída por uma sequência terminal de poliHis. O domínio VHH produzido pode igualmente ser detectado através da utilização de anticorpos cultivados contra as regiões VHH do camelídeo. Sob condições laboratoriais, as regiões VHH são produzidas no vector Imuno PBS em quantidades em mg por litro.

O presente pedido também descreve um banco de ADN composto por sequências de nucleótidos que codificam uma imunoglobulina de cadeia pesada, tal como a obtida das células com genes de imunoglobulinas rearranjada.

Numa forma de realização do presente pedido, o banco é preparado a partir de células de um animal previamente imunizado contra um determinado antigénio. Isto permite a selecção de anticorpos com uma especificidade pré-seleccionada para o antigénio usado para imunização.

Noutra forma de realização do presente pedido, a amplificação de ADNc não é realizada antes da clonagem do ADNc.

A cadeia pesada das imunoglobulinas de quatro cadeias permanece sequestrada na célula por uma proteína adventícia (BIP) até que seja combinada com uma cadeia leve. O local de ligação para a proteína adventícia é o domínio CH1. Uma vez que este domínio está isento de imunoglobulinas de cadeia pesada, a sua secreção é independente da presença da proteína BIP ou da cadeia leve. Para além disso, os inventores mostraram que as imunoglobulinas obtidas não são

pegajosas e consequentemente não irão agregar-se de modo anormal.

O presente pedido também descreve um processo para a preparação de um anticorpo monoclonal contra um antígeno específico, em que o local de aglutinação do antígeno consiste em cadeias polipeptídicas pesadas e em que o anticorpo é ainda desprovido de cadeias polipeptídicas leves, o processo compreendendo:

- imortalização de linfócitos, obtidos por exemplo do sangue periférico de camelídeos previamente imunizados com um determinado antígeno, com uma célula imortal e de preferência com células mieloma, de modo a realizar um hibridoma,
- cultura das células imortalizadas (hibridoma) formadas e recuperação das células que produzem os anticorpos que possuem a especificidade desejada.

A preparação dos anticorpos pode também ser realizada sem uma imunização prévia dos camelídeos.

De acordo com outro processo para a preparação de anticorpos, não é necessário o recurso à técnica da célula de hibridoma.

De acordo com tal processo, os anticorpos são preparados *in vitro* e podem ser obtidos por um processo compreendendo as seguintes etapas:

- clonagem em vectores, especialmente em fagos e mais particularmente bacteriófagos filamentosos, sequências de ADN ou ADNc obtidas de linfócitos, especialmente PBLs ou camelídeos previamente imunizados com determinados antígenos,
- transformação de células procarióticas com os vectores mencionados em cima, em condições que permitam a produção de anticorpos,
- selecção dos anticorpos pela sua estrutura de cadeia

pesada e adicionalmente através da sua sujeição a selecção por afinidade com antigénio,

- recuperação dos anticorpos com a especificidade desejada,

Noutra forma de realização da invenção, a clonagem é realizada em vectores, especialmente em plasmídeos que codificam proteínas de membrana bacteriana.

As células procarióticas são então transformadas com os vectores mencionados em cima, em condições que permitam a expressão dos anticorpos nas suas membranas.

As células positivas são ainda escolhidas por selecção pela afinidade do antigénio.

Os anticorpos de cadeia pesada que não possuam o domínio  $C_H1$  apresentam uma vantagem distinta nesta matéria. Com efeito, o domínio  $C_H1$  liga-se às proteínas adventícias de tipo BIP presentes em vectores eucarióticos e as cadeias pesadas não são transportadas para fora do retículo endocitoplásmico a não ser que estejam presentes cadeias leves. Significa isto que em células eucarióticas, a clonagem eficiente de imunoglobulinas de quatro cadeias em células que não sejam de mamíferos, tais como as células de leveduras, pode depender das propriedades das proteínas adventícias de tipo BIP residentes e por tal motivo pode ser muito difícil conseguir essa clonagem. A propósito disto, os anticorpos de cadeias pesadas da invenção que não tenham o domínio  $C_H1$  representam uma vantagem distintiva.

Numa variante do presente pedido, é possível realizar a clonagem em leveduras, quer para a produção de anticorpos quer para a modificação do metabolismo das leveduras. Como exemplo, é possível utilizar o vector Yep 52. Este vector tem a origem de replicação (ORI)  $2\mu$ . da levedura conjuntamente com um marcador de selecção *Leu 2*.

O gene clonado está sob o controlo do promotor *gall* e consequentemente é induzível pela galactose. No entanto, a expressão pode ser reprimida pela glicose, o que permite obter concentrações muito elevadas de células antes da indução.

A clonagem entre locais *Bam*HI e *Sal*I, utilizando a mesma estratégia de produção de genes por RCP, tal como a estratégia anteriormente descrita, permite clonar genes de imunoglobulinas de Camelídeos em *E. coli*. Como exemplo de modulação metabólica que pode ser conseguida pelos anticorpos e proposta para a levedura, é possível indicar a clonagem de anticorpos dirigidos contra as ciclinas, isto é, proteínas implicadas na regulação do ciclo celular das leveduras (TIBS 16 430 J.D. McKinney, N. Heintz 1991). Outro exemplo é a introdução, por meio de técnicas da engenharia genética, de um anticorpo dirigido contra as CD<sub>28</sub>, anticorpo esse que poderia ser induzível (por exemplo, por *gall*), no interior do genoma da levedura. As CD<sub>28</sub> têm uma participação ao nível da iniciação da divisão celular e por tal motivo a expressão de anticorpos contra esta molécula poderia permitir um controlo eficiente da multiplicação das células e a optimização de métodos para a produção em biorreactores ou por meio de células imobilizadas.

Ainda de acordo com outra variante do presente pedido, o vector de clonagem é um plasmídeo ou um vector virai eucariótico e as células que irão ser transformadas são células eucarióticas, em especial células de leveduras, células de mamíferos, por exemplo, células de OCC (em inglês CHO) ou células de símios, tais como células Veto, células de insectos, células de plantas ou células de protozoários.

Para saber mais pormenores sobre o procedimento que deve ser aplicado num caso desses, faz-se referência à obra de Marks et al., J. Moi. Biol. 1991, 222:581-597,

Além disso, começando com as imunoglobulinas descritas ou com os seus fragmentos é possível preparar novas imunoglobulinas ou os seus derivados.

Posto isto, é possível preparar imunoglobulinas, que satisfaçam às definições anteriormente apresentadas, contra determinados antigénios. O presente pedido descreve especialmente anticorpos monoclonais ou policlonais desprovidos de cadeias polipeptídicas leves ou anti-soros que contenham tais anticorpos e que são dirigidos contra determinados antigénios, por exemplo, contra antigénios de agentes patológicos, tais como bactérias, vírus ou parasitas. Como exemplos de antigénios ou determinantes antigénicos contra os quais é possível preparar os anticorpos, é possível citar as glicoproteínas do invólucro de vírus ou os seus péptidos, tais como- a glicoproteína do invólucro externo do VTE-1 e o antigénio da superfície do vírus da hepatite B.

As imunoglobulinas descritas também podem ser dirigidas contra uma proteína, um hapteno, um hidrato de carbono ou um ácido nucleico.

Os anticorpos particulares são dirigidos contra o epítipo de galactosil- $\alpha$ -1-3-galactose.

As imunoglobulinas descritas permitem ainda preparar produtos combinados, tais como a combinação da imunoglobulina de cadeias pesadas, ou um seu fragmento, com uma toxina, uma enzima, um fármaco ou uma hormona.

Como exemplo, é possível preparar a combinação de uma imunoglobulina de cadeias pesadas, portadora de um local de ligação de um antigénio que reconheça um epítipo de imunoglobulina de mieloma, com a ab•ina ou a toxina da



lectina de visco-branco. Tal arquétipo pode ter as suas aplicações em terapias específicas de pacientes.

Outra combinação vantajosa é a que resulta da preparação entre uma imunoglobulina de cadeias pesadas, que reconheça um antigénio do intestino de insectos, com uma toxina específica para insectos, tal como uma das toxinas dos diferentes serotipos de *Bacillus thuringiensis* ou *Strigomonas sphaericus*. Tal arquétipo clonado em plantas pode ser utilizado para aumentar a especificidade ou a variedade de hospedeiros de toxinas bacterianas existentes.

O presente pedido também proporciona anticorpos que possuem diferentes especificidades em cada cadeia polipeptídica pesada. Estes anticorpos multifuncionais, especialmente bifuncionais, poderiam ser preparados combinando duas cadeias pesadas de imunoglobulinas da invenção ou uma cadeia pesada de uma imunoglobulina da invenção com um fragmento de uma imunoglobulina do modelo de quatro cadeias.

O presente pedido também descreve anticorpos heteroespecíficos que podem ser utilizados para encaminhar fármacos ou qualquer substância biológica, por exemplo, hormonas. Em particular, podem ser utilizados para encaminhar, de forma selectiva, hormonas ou citocinas para uma categoria limitada de células. Como exemplos refere-se uma combinação de um anticorpo de murinos ou de seres humanos, criado contra a interleucina 2 ( $IL_2$ ), e um anticorpo de cadeias pesadas criado contra as células  $CD_4$ . Estes anticorpos poderiam ser utilizados para reactivar células  $CD_4$  que tenham perdido o seu receptor de  $IL_2$ .

As imunoglobulinas de cadeias pesadas descritas também podem ser utilizadas para a preparação de anticorpos heteroespecíficos. Isto pode ser conseguido quer de acordo com o método anteriormente descrito, por redução das pontes

entre as diferentes cadeias e reoxidação, em conformidade com as técnicas habituais, de dois anticorpos que tenham especificidades diferentes, ou então pode ser conseguido efectuando a clonagem sequencial de dois anticorpos, por exemplo, no vector 'Immuno PBS'.

Se for esse o caso, prepara-se um primeiro gene correspondente ao domínio  $V_{HH}$  compreendido entre o local *Xho* e o local *Spe*, conforme descrito antes. Depois prepara-se um segundo gene por um método análogo utilizando como extremidade 5' um iniciador que possua um local *Spe* e utilizando como extremidade 3' um iniciador que possua um codão de terminação e um local *EcoRI*. Depois faz-se digerir o vector com *EcoRI* e *XhoI* e ainda se faz digerir os dois genes de  $V_{HH}$  com *Xho/Spe* e *Sp/EcoRI*.

Após a ligação, os dois genes de imunoglobulinas são clonados sequencialmente. O afastamento entre os dois genes pode ser aumentado mediante a introdução de codões de adição dentro do iniciador *SpeI* de 5'.

Numa variante descrita particular, a região de charneira das imunoglobulinas IgG2 é semi-rígida e por isso é adequada para o acoplamento de proteínas. Numa aplicação dessas, as proteínas ou os péptidos podem ser ligados a diversas substâncias, especialmente a ligandos, através da região de charneira utilizada como separador. De forma vantajosa, o fragmento compreende pelo menos 6 aminoácidos. De acordo com a revelação, é interessante, é utilizar uma sequência que compreenda uma sequência repetida de Pro-X em que o símbolo X designa qualquer aminoácido e de preferência Gln, Lis ou Glu, em especial um tiagmento constituído pelo menos por uma repetição tripla e de preferência por uma repetição dodécupla (12 vezes), para acoplar proteínas ao ligando ou para a montagem de diferentes domínios proteínicos.

Também é possível utilizar a região de charneira ou um seu fragmento para acoplar proteínas aos ligandos ou para a montagem de diferentes domínios proteínicos.

Para o acoplamento são adequadas as técnicas habituais e é possível fazer uma referência especial à técnica de construção de proteínas por montagem de sequências clonadas.

Os anticorpos descritos poderiam ser utilizados como reagentes para diagnóstico *in vitro* ou para a prática de técnicas de imagiologia. As imunoglobulinas poderiam ser marcadas com radioisótopos, marcadores químicos ou enzimáticos ou marcadores quimioluminescentes.

Como exemplo, especialmente no caso da detecção ou da observação com as imunoglobulinas, pelas técnicas de imatologia, é vantajoso utilizar um marcador, tal como o tecnécio, em especial o tecnécio 99. Este marcador pode ser utilizado para orientar a marcação, por meio de um procedimento de acoplamento com as imunoglobulinas ou com os seus fragmentos, ou para a marcação indirecta a seguir a um passo de preparação de um complexo com o tecnécio.

Por exemplo, há outros marcadores radioactivos interessantes, entre os quais se refere o índio e em especial o índio III, ou o iodo, em especial  $I^{131}$ ,  $I^{125}$  e  $I^{123}$ .

Para estudo destas técnicas faz-se referência ao pedido de patente de invenção fiñancesa publicada com o nº 2649488.

Nestas aplicações, o tamanho pequeno do fragmento V constitui uma vantagem definitiva para efeitos de penetração nos tecidos.

O presente pedido também descreve anticorpos monoclonais que reagem com anti-idiotipos dos anticorpos descritos antes.

O presente pedido também descreve células ou organismos em que as imunoglobulinas de cadeias pesadas tenham sido clonadas. Tais células ou organismos podem ser utilizados para a produção de imunoglobulinas de cadeias pesadas que possuam uma especificidade desejada pré-seleccionada, ou correspondentes a um repertório particular. Também podem ser produzidas para a modificação do metabolismo das células que as expressem. No caso de modificação do metabolismo de células transformadas com as sequências que codificam as imunoglobulinas de cadeias pesadas, estas imunoglobulinas de cadeias pesadas assim produzidas são utilizadas como ADN anti-sentido. O ADN anti-sentido está envolvido normalmente no bloqueio da expressão de determinados genes tais como, por exemplo, o antígeno de superfície variável de tripanossomas ou de outros agentes patogénicos. De igual modo, a produção ou a actividade de determinadas proteínas ou enzimas poderia ser inibida mediante a expressão de anticorpos contra essas proteínas ou enzimas no interior da mesma célula.

O presente pedido também descreve uma imunoglobulina de quatro cadeias modificada, ou seus fragmentos, cujas regiões  $V_H$  tenham sido parcialmente substituídas por sequências específicas ou aminoácidos de imunoglobulinas de cadeias pesadas, especialmente por sequências de domínio  $V_{HH}$ . Há um domínio  $V_H$  modificado particular de uma imunoglobulina de quatro cadeias de acordo com a invenção que se caracteriza pelo facto de a leucina, a prolina ou a glutamina na posição 45 das regiões  $V_1$  terem sido substituídas por outros aminoácidos e de preferência por arginina, ácido glutâmico ou cisteína.

Há um outro domínio  $V_H$  ou  $V_L$  modificado de uma imunoglobulina de quatro cadeias que se caracteriza pela ligação de elos das RDC (em inglês CDR) entre si ou às

regiões da AE mediante a introdução de cisteínas emparelhadas, sendo a RDC escolhida entre a  $RDC_1$  e a  $RDC_3$ , e em que a região da AE é a região  $AE_2$ , e em especial nos casos em que uma das cisteínas introduzidas está na posição 31 ou 33 da  $AE_2$  ou na posição 45 da  $RDC_2$  e a outra está na  $RDC_3$ .

Em especial, a introdução de cisteínas emparelhadas é tal que o elo da  $RDC_3$  está ligado à  $AE_2$  ou ao domínio  $RDC_1$  e mais especialmente a cisteína da  $RDC_3$  da região  $V_H$  está ligada a uma cisteína na posição 31 ou 33 da  $AE_2$  ou na posição 45 da  $RDC_2$ .

Noutra variante descrita, é possível modificar células de plantas por meio de imunoglobulinas de cadeias pesadas, com a finalidade de lhes conferir novas propriedades ou melhores propriedades.

As imunoglobulinas de cadeias pesadas podem ser utilizadas para a terapia génica do cancro, por exemplo, utilizando anticorpos dirigidos contra as proteínas presentes nas células do tumor.

Em tal caso, é possível conseguir a expressão de um ou dois genes de  $V_{HH}$  utilizando vectores obtidos a partir de parvovírus ou adenovírus. Os parvovírus caracterizam-se pelo facto de serem desprovidos de patogenicidade ou tle praticamente não serem patogénicos para as células humanas normais e pelo facto de serem capazes de se multiplicarem facilmente em células cancerosas (Russel S.J. 1990, Immunol. Today II, 196-200).

As imunoglobulinas de cadeias pesadas são clonadas, por exemplo, no interior de locais *HindIII/XbaI* do plasmídeo infeccioso do vírus MVM dos merinos (pMM984). (Merchlinsky et al., 1983, J. Virol. 47, 227-232) e depois são colocadas sob o controlo do promotor MVM38.

O gene do domínio V é amplificado por RCP utilizando um iniciador de 5' que contém um codão de iniciação e um local *HindIII*, possuindo o iniciador de 3' um codão de terminação e um local *XbaI*.

Este arquétipo é depois inserido entre as posições 2650 (*HindIII*) e 4067 (*XbaI*) do plasmídeo.

A eficiência da clonagem pode ser confirmada por transfecção. O vector que contém o anticorpo é depois introduzido em células permissivas (NB-E) por transfecção.

As células são recuperadas ao fim de 2 dias e determina-se a presença de regiões  $V_{HH}$  segundo um protocolo de EISLE, utilizando para tal anti-soro de coelho que reage com a parte de  $V_{HH}$ .

O presente pedido também diz respeito à preparação de anticorpos catalíticos por caminhos diferentes. A produção de anticorpos dirigidos contra componentes que simulam estados activados de substratos (por exemplo, o vanadato como componente que simula o estado activado do fosfato, com o intuito de se produzir as suas actividades de fosfoesterase, o fosfonato como composto que simula a ligação peptídica, com o intuito de se produzir proteases) permite obter anticorpos que têm uma função catalítica. Outro caminho para se obter tais anticorpos consiste em realizar uma mutagénese aleatória em clones de anticorpos, por exemplo, por RCP, introduzindo bases anormais durante a amplificação dos clones. Estes fragmentos amplificados, obtidos por RCP, são depois introduzidos num vector adequado para clonagem. A sua expressão à superfície das bactérias permite que o substrato detecte clones que possuem actividade enzimática. Estas duas metodologias podem ser combinadas, como é evidente. Finalmente, com base nos dados disponíveis sobre a estrutura, por exemplo, os dados obtidos por cristalografia aos raios X ou por RMN, é

possível comandar as modificações. Estas modificações podem ser realizadas por meio de técnicas habituais da engenharia genética ou por síntese completa. Uma vantagem da região  $V_{HH}$  das imunoglobulinas de cadeias pesadas da invenção é o facto de serem suficientemente solúveis.

As imunoglobulinas de cadeias pesadas também podem ser produzidas em células vegetais, em especial em plantas transgénicas. Como exemplo, é possível produzir imunoglobulinas de cadeias pesadas em plantas, utilizando para tal o plasmídeo pMon530 (Roger *et al.*, Meth. Enzym 153 1566 1987). O vector constitutivo de expressão de plantas foi já descrito para anticorpos clássicos do modelo de quatro cadeias (Hiat *et al.* Nature 342 76-78, 1989), mais uma vez utilizando os iniciadores adequados de RCP, conforme descrito *supra*, para se gerar um fragmento de ADN na fase correcta.

Há outras vantagens e características da invenção que se tornarão evidentes nos exemplos e na observação das figuras subsequentes.

#### FIGURAS

**Figura 1: características e purificação tia igG de camelo por cromatografia de afinidade sobre Sepharose com Proteína A e com Proteína G (Pharmacia)**

A figura 1A mostra, após redução, o perfil da proteína obtida por EGPA-DSS, das fracções adsorvida e não adsorvida de soro de *Camelus dromedarius*. A fracção adsorvida na Proteína A e eluída com NaCl 0,15 M e ácido acético a 0,58% revelou, após redução (pista c), três componentes de cadeia pesada respectivamente com 50, 46 e 43 kd e uma cadeia leve (IgG de coelho na pista a). Às fracções adsorvidas num derivado de 'Sepharose' com Proteína G (Pharmacia), que havia sido concebido artificialmente para suprimir a região de ligação da albumina (pista e) e cuja eluição foi feita

com  $\text{gly-HCl}$  0,1 M a pH 2,7, falta-lhes a cadeia pesada de 46 kd que é recuperada na fracção não adsorvida (pista f). Não há nenhum destes três componentes na fracção não adsorvida na Proteína A (pista d); a pista b contém os marcadores de pesos moleculares.

Figuras 1B e 1C. Por eluição diferencial é possível separar as fracções com Proteína G e depois lavou-se exaustivamente a coluna com tampão fosfato 20 mM a pH 7,0. Após a eluição com tampão a pH 3,5 ( $\text{NaCl}$  0,15 M e ácido acético a 0,58%), tem lugar a eluição de um componente de 100 kd que após redução dá origem a uma cadeia pesada de 43 kd (pista 1). Depois de a absorvência do eluente da coluna ter caído para o nível natural é possível efectuar a eluição de um segundo componente da imunoglobulina, com um peso molecular de 170 kd, utilizando para tal tampão a pH 2,7 ( $\text{glicina-HCl}$  0,1 M). Esta fracção, após redução, dá origem a uma cadeia pesada de 50 kd e a uma faixa de cadeia leve larga (pista 2). A fracção não adsorvida na Proteína G é depois aplicada a uma coluna de .5 mL de 'Sephrose' com Proteína A. Após a lavagem e eluição com tampão a pH 3,5 ( $\text{NaCl}$  0,15 M e ácido acético a 0,58%) é possível obter uma terceira imunoglobulina de 100 kd que é constituída exclusivamente pelas cadeias pesadas de 46 kd (pista 3).

**Figura 2: imunoglobulinas de *Camelus bactrianus*, *Lama vieugna*, *Lama glama* e *Lama patos* para a Proteína A (pistas A) e para a Proteína G (pistas G) analisadas por EGPA-DSS antes (A) e após redução (B)**

Juntou-se 10  $\mu\text{L}$  de soro, proveniente de diferentes espécies, a tubos de Eppendorf que continham 10 mg de 'Sephrose' com Proteína A ou com Proteína G, em suspensão em 400  $\mu\text{L}$ , de tampão de imunoprecipitação a pH 8,3 ( $\text{NaCl}$  0,2 M, Tris 0,01 M; EDTA 0,011\4, (Triton X100' a 1% e



ovalbumina a 0,1%). Manteve-se estes tubos a rodar lentamente durante 2 horas a 4°C. Após a centrifugação, as massas obtidas foram lavadas três vezes em tampão e uma vez em tampão em que não havia 'Triton' nem ovalbumina. As massas obtidas foram depois recolocadas em suspensão na solução de amostra de EGPA-DSS, sendo utilizados 70 µL por cada massa na presença ou na ausência de ditiotreitól enquanto agente redutor. Depois de se manter em ebulição durante 3 minutos a 100°C, efectuou-se a centrifugação dos tubos e fez-se uma análise dos sobrenadantes. Em todas as espécies examinadas as frações não reduzidas (A) continham também, para além das moléculas de aproximadamente 170 kd, componentes importantes mais pequenos de aproximadamente 100 kd. Na amostra reduzida (B) são detectadas as cadeias pesadas e leves constituintes. Há em todas as espécies um componente de cadeia pesada (marcado com um asterisco (\*)) que se encontra presente no material resultante da eluição da Proteína A, mas que não existe no material da eluição da Proteína G

**Figura 3: preparação de IgG<sub>1</sub>, IgG<sub>2</sub> e IgG<sub>3</sub> a partir de soro proveniente de *Camelus dromedarius* saudáveis ou infectados com *Trypanosoma evansi* (titulação CATT a 1/160 (3)) e análises feitas por imunoprecipitação ou por autorradiografias à Western para pesquisa da actividade anti-tripanosossómica**

Figura 3A. Acrescentou-se lisado de antigénios de *Trypanosoma evansi* marcados com <sup>35</sup>S-metionina (500 000 unidades de contagem) a tubos de Eppendorf que continham 10 µL de soro ou 20 µg de IgG<sub>1</sub>, IgG<sub>2</sub> e IgG<sub>3</sub> em 200 µL de tampão de imunoprecipitação a pH 8,3 que continha TLCK 0,1 M, enquanto inibidor das proteinases, e manteve-se os tubos a rodar lentamente a 4°C durante 1 hora. Os tubos

foram depois enriquecidos com 10 mg de 'Sepharose' com Proteína A em suspensão em 200 µL do mesmo tampão a pH 8,3 e ficaram a incubar a 4°C durante mais 1 hora.

Após a lavagem e a centrifugação a 15 000 r.p.m. durante 12 segundos, colocou-se cada massa obtida novamente em suspensão em 75 µL de solução de amostra de EGPA-DSS que continha DTT e aqueceu-se durante 3 minutos a 100°C. Após a centrifugação numa minicentrífugadora de Eppendorf a 15 000 r.p.m. durante 30 segundos, retirou-se 5 µL do sobrenadante para determinação da radioactividade e a parte restante foi analisada por EGPA-DSS e por fluorografia. O número de unidades de contagem de radioactividade por cada 5 µL, de amostra estão inscritos em cada linha.

Figura 3B. Separou-se 20 µg de IgG<sub>1</sub>, IgG<sub>2</sub> e IgG<sub>3</sub> provenientes de animais saudáveis e de animais infectados com tripanossomas, recorrendo ao protocolo de EGPA-DSS, sem redução prévia nem aquecimento. As amostras separadas foram então electrotransferidas para uma membrana de nitrocelulose, contrastou-se uma parte da membrana com o corante vermelho de Ponceau para se localizar o material proteínico e a parte restante ficou a incubar com ovalbumina a 1% em tampão TST (Tris 10 mM, NaCl 150 mM e Tween a 0,05%) para bloquear os locais de ligação das proteínas.

Após o bloqueio, lavou-se a membrana exaustivamente com tampão TST e manteve-se a incubar durante 2 horas com antigénio de tripanossomas marcado com <sup>35</sup>S. Depois de uma lavagem exaustiva, secou-se a membrana e analisou-se por autorradiografia. Para se evitar a ligação natural e de tipo não específico, o usado de tripanossomas marcado foi filtrado através de um fdtro de miliporos de 45 µm e manteve-se a incubar com imunoglobulina de camelo saudável e ovalbumina adsorvida sobre uma membrana de nitrocelulose.

**Figura 4: a IgG<sub>3</sub> de camelo purificada por cromatografia de afinidade sobre 'Sepharose' com Proteína A é parcialmente digerida com papaina e separada sobre 'Sepharose' com Proteína A**

Dissolveu-se 14 mg de IgG<sub>3</sub> purificada em tampão fosfato 0,1 M a pH 7,0 que continha EDTA 2 mM. Fez-se digerir soro de leite coalhado por incubação durante 1 hora a 37°C com a enzima mercuriopapaína (proporção de 1% entre enzima e proteína) activada por cisteína  $5 \times 10^{-4}$  M. Bloqueou-se a digestão por adição de um excesso de iodoacetamida ( $4 \times 10^{-2}$  M) (13). Após a centrifugação do produto de digestão numa centrífugadora de Eppendorf durante 5 minutos a 15 000 r.p.m., efectuou-se a separação dos fragmentos de papaína numa coluna de 'Sepharose' com Proteína A, em fracções ligantes (L) e não ligantes (NL). A eluição da fracção ligante teve lugar a partir da coluna com tampão de glicina-HCl 0,1 M a pH 1,7.

**Figura 5: representação esquemática de um modelo para as moléculas de IgG<sub>3</sub> desprovidas de cadeias leves.**

**Figura 6: • representação esquemática de imunoglobulinas que possuem cadeias polipeptídicas pesadas e desprovidas de cadeias leves, tendo em conta a imunoglobulina do modelo convencional de quatro cadeias;**

- representação de uma região de charneira.

**Figura 7: alinhamento de 17 sequências de ADN de Vim de imunoglobulinas de cadeias pesadas do camelo.**

**Figura 8: expressão e purificação da proteína V<sub>E111</sub>21 do camelo a partir de *E. coli*.**

#### I ANTICORPOS DE CADEIAS PESADAS NOS CAMELÍDEOS

Quando o soro de *Camelus dromedarius* é adsorvido em 'Sepharose' com Proteína G, há uma quantidade apreciável (25%-35%) de imunoglobulinas (Ig) que ficam em solução e que podem ser depois recuperadas por cromatografia de afinidade sobre 'Sepharose' com Proteína A (figura 1A). A fracção adsorvida na Proteína G pode ser objecto de eluição diferencial no interior de uma fracção coesamente ligada (25%), constituída por moléculas com um peso molecular aparente não reduzido (PM) igual a 170 kd e por uma fracção mais fracamente ligada (30%-45%) que possui um peso molecular aparente igual a 100 kd (figura 1B). O componente de 170 kd, quando reduzido, origina cadeias pesadas de 50 kd .e cadeias leves grandes de 30 kd. A fracção de 100 kd é totalmente desprovida de cadeias leves e parece ser constituída exclusivamente por cadeias pesadas que após redução têm um PM aparente igual a 43 kd (figura 1C). A fracção que não se liga à Proteína G pode ser purificada por afinidade e eluida a partir de uma coluna de Proteína A como um segundo componente de 100 kd que após redução parece ser constituído exclusivamente por cadeias pesadas de 46 kd.

As imunoglobulinas de cadeias pesadas desprovidas de cadeias leves perfazem até 75% das moléculas que se ligam à Proteína A.

Uma vez que a totalidade das três imunoglobulinas se liga à Proteína A, referir-nos-emos a elas como IgG: designadamente IgG<sub>1</sub> (γ1 (50 kd) de cadeias leves e de cadeias pesadas) que se liga à Proteína G, IgG<sub>2</sub> (γ2 (46 kd) de cadeias pesadas) que não se liga à Proteína G e IgG<sub>3</sub> (γ3

(43 kd) de cadeias pesadas) que se liga à Proteína G. Há a possibilidade de estas três subclasses ainda poderem ser subdivididas.

Um estudo comparativo entre os Camelídeos do velho mundo (*Camelus bactrianus* e *Camelus dromedarius*) e os Camelídeos do novo mundo (*Lama pacos*, *Lama glama* e *Lama vicugna*) comprovou que as imunoglobulinas de cadeias pesadas existem em todas as espécies examinadas, embora com diferenças irrelevantes na sua proporção e nos seus pesos moleculares aparentes. Os Camelídeos do novo mundo são diferentes dos Camelídeos do velho mundo pelo facto de possuírem uma molécula IgG<sub>3</sub> maior (imunoglobulina de cadeias pesadas que se liga à Proteína G) em que as cadeias pesadas constituintes têm um peso molecular aparente de 47 kd (figura 2).

A abundância de imunoglobulinas de cadeias pesadas no soro dos Camelídeos levanta a questão de se saber qual é o seu papel na resposta imunitária e em particular se são portadoras de especificidade de ligação dos antígenos e em caso afirmativo qual é a extensão do seu repertório. Esta questão pôde ser respondida examinando as imunoglobulinas de camelus (*Camelus dromedarius*) infectados com *Dypanasoma evansi*.

Para o efeito, preparou-se as fracções correspondentes de IgG<sub>1</sub>, IgG<sub>2</sub> e IgG<sub>3</sub> a partir de soro de um camelo saudável e a partir de soro de camelus com uma elevada concentração de anti-tripanosoma, medida em conformidade com o 'Card Agglutination Test' (3). No ensaio de radioimunoprecipitação, demonstrou-se que as IgG<sub>1</sub>, IgG<sub>2</sub> e IgG<sub>3</sub>, obtidas a partir de camelus infectados, indicando o repertório extenso de heterogeneidade e de complexidade (figura 3A), se ligam a um grande número de antígenos

presentes num lisado de tripanossomas marcado com  $^{35}\text{S}$ -metionina.

Em experiências de obtenção de autorradiografias, o lisado de tripanossomas marcado com  $^{35}\text{S}$ -metionina liga-se às IgG<sub>1</sub>, IgG<sub>2</sub> e IgG<sub>3</sub>, separadas por EGPA-DSS, provenientes de animais infectados (figura 3B).

Isto levou-nos a concluir que as IgG<sub>2</sub> e IgG<sub>3</sub> de cadeias pesadas dos Camelídeos são anticorpos que se ligam a antigénios genuínos.

Há um paradigma imunológico que estabelece que o repertório total de anticorpos é gerado pela combinação dos repertórios da região V variável de cadeias leves e pesadas (6). As imunoglobulinas de cadeias pesadas do camelo parecem contradizer este paradigma.

As imunoglobulinas são caracterizadas por um padrão de focagem isoeléctrica (FIE) complexo, que reflecte a sua heterogeneidade extrema. Para se determinar se as duas cadeias pesadas que constituem as IgG<sub>2</sub> e IgG<sub>3</sub> são idênticas ou não, efectuou-se uma observação do padrão de focagem isoeléctrica (FIE) antes e depois da separação das cadeias por redução e alquilação, utilizando como agente de alquilação a iodoacetamida.

Uma vez que este agente de alquilação não introduz cargas eléctricas adicionais na molécula, os monómeros resultantes da redução e da alquilação do homodímero de cadeias pesadas irá ter praticamente o mesmo ponto isoeléctrico do dímero, ao passo que se fossem derivados de um heterodímero de cadeias pesadas, os monómeros iriam diferir, na maior parte dos casos, suficientemente no ponto isoeléctrico para gerarem um padrão diferente no FIE.

Após a redução e a alquilação com iodoacetamida, o padrão observado não é modificado no caso das IgG<sub>2</sub> e IgG<sub>3</sub> de *Camelus dromedarius*, o que indica que cada uma destas

moléculas é constituída por duas cadeias pesadas idênticas que migram para a mesma posição para que migra a molécula não reduzida onde tiveram a sua origem.

Pelo contrário, o padrão de FIE da IgG<sub>1</sub> é completamente modificado após a redução, uma vez que o ponto isoelétrico de cada molécula é determinado pela combinação dos pontos isoelétricos das cadeias leves e pesadas, migrando qualquer uma delas, após a separação, para uma posição diferente. Estes resultados indicam que as cadeias pesadas por si sós podem gerar um repertório extenso e por em causa a contribuição da cadeia leve para o repertório de anticorpos úteis. Se esta necessidade for negada, fica por saber qual o papel que a cadeia leve desempenha.

Normalmente, a cadeia pesada isolada a partir de imunoglobulinas de mamíferos tem tendência para se agregar de forma considerável, mas apenas é solubilizada pelas cadeias leves (8, 9) que se ligam ao domínio C<sub>H</sub>1 da cadeia pesada.

Nos seres humanos e nos murganhos há diversos mielomas espontâneos ou induzidos que produzem uma imunoglobulina patológica constituída exclusivamente por cadeias pesadas (doença das cadeias pesadas). Estas cadeias pesadas das proteínas dos mielomas contêm supressões nos domínios C<sub>H</sub>1 e V<sub>HH</sub> (10). O motivo pelo qual as cadeias pesadas de comprimento completo não originam cadeias pesadas segregadas em tais imunoglobulinas patológicas parece ficar a dever-se ao facto de a síntese de Ig necessitar de uma proteína adventícia, a proteína de ligação de cadeias pesadas da imunoglobulina ou BIP (11), que é normalmente substituída pela cadeia leve (12). É possível que o papel primordial da cadeia leve nas imunoglobulinas do modelo de quatro cadeias seja o de uma proteína adventícia de cadeia

pesada comprometida e que a emergência dos repertórios de cadeias leves tenha sido mesmo um bônus evolutivo.

As cadeias  $\gamma_2$  e  $\gamma_3$  dos Camelídeos são consideravelmente mais curtas do que a cadeia  $\gamma$  normal dos mamíferos. Isto poderia sugerir a hipótese de terem ocorrido supressões no domínio  $C_H1$ . As diferenças dos tamanhos das imunoglobulinas  $\gamma_2$  e  $\gamma_3$  dos Camelídeos do velho mundo e do novo mundo sugerem que as supressões ocorreram em diversos passos evolutivos, especialmente no domínio  $C_H1$ .

## II FALTA O DOMÍNIO $C_H1$ ÀS IMUNOGLOBULINAS DE CADEIAS PESADAS DOS CAMELÍDEOS

A estratégia seguida na investigação da estrutura primária das imunoglobulinas de cadeias pesadas é uma combinação da sequenciação de proteínas e da sequenciação de ADN; a sequenciação de proteínas é necessária para identificar os segmentos de sequências característicos de cada imunoglobulina. Sendo o terminal N da imunoglobulina derivado da região variável da cadeia pesada, o repertório apenas fornece informação sobre os subgrupos  $V_{HH}$  (região variável da cadeia pesada) e não pode ser utilizado para identificação da classe ou da subclasse. Significa isto que é necessário obter dados das sequências a partir de locais internos de clivagem enzimática ou química.

Uma combinação de digestão com papaína e cromatografia por afinidade com Proteína A permitiu fazer a separação de diversos fragmentos que fornecem informação sobre a estrutura geral da  $IgG_3$ .

A  $IgG_3$  do camelo (*Camelus dromedarius*), purificada por cromatografia de afinidade sobre 'Sepharose' com Proteína A, foi parcialmente digerida com papaína e o produto da digestão foi separado sobre 'Sepharose' com Proteína A, tendo sido obtidas as fracções de ligação e de não ligação.



Estas fracções foram analisadas por EGPA-DSS em condições redutoras e não redutoras (figura 4).

A fracção ligada continha dois componentes, um de 28 kd e um de 14,4 kd, para além do material não clivado ou parcialmente clivado. Essas fracções foram bem separadas por electroforese em gel (a partir de geles preparativos de EGPA-DSS a 19%), em condições não redutoras, e foram ainda purificadas melhor por electroeluição (em bicarbonato de amónio 50 mM e DSS a 0,1% (p/v), utilizando um dispositivo de electroeluição da 'BioRad'). Após a liofilização destas fracções electroeluídas, eliminou-se a parte remanescente do DSS fazendo precipitar a proteína por adição de etanol a 90% e misturando e mantendo a mistura a incubar de um dia para o outro a -20°C (14). A proteína precipitada foi recolhida numa massa combinada por centrifugação (15 000 r.p.m. durante 15 minutos) que foi utilizada para a sequenciação proteínica. Efectuou-se a sequenciação do terminal N utilizando a química de Edman para um sequenciador automático de proteínas em fase líquida, em regime pulsatório, da 'Applied Biosystem 477A'. Os aminoácidos foram identificados pelos seus derivados de fenil-tio-hidantoína (PTH), tendo para tal sido utilizado o analisador de modelo 120 PTH da Applied Biosystem. Todos os produtos químicos e reagentes foram adquiridos á empresa Applied Biosystem A análise dos dados cromatográficos foi realizada utilizando a versão informática 1.61 da Applied Biosystems. Em todos os casos, a análise das sequências, auxiliada por computador, foi confirmada por inspecção directa dos cromatogramas resultantes do analisador de PTH. Dissolveu-se amostras da proteína sequenciada alternativamente em ácido trifluoroacético (ATF) a 50% (v/v) (fragmento de 28 kd) ou em ATF a 100% (fragmento de 14 kd). As amostras de equivalente proteínico dissolvido

para 2000 pmol (fragmento de 28 kd) ou para 500 pmol (fragmento de 14 kd) foram aplicadas sobre discos de fibras de vidro, tratados com ATF. Os discos de fibras de vidro foram recobertos com BioBerne (3 mg) e tratados em pré-ciclo uma vez antes da sua utilização.

A sequenciação do terminal N do fragmento de 28 kd gera uma sequência homóloga da parte do terminal N do domínio  $\gamma C_H2$  e portanto homóloga da extremidade do terminal N 46 do fragmento Pc. A sequência do terminal N do fragmento de 14,4 kd corresponde à última lisina de um domínio  $\gamma C_H2$  e à extremidade do terminal N de um domínio  $\gamma C_H3$  (quadro 1). O peso molecular (PM) dos fragmentos obtidos com papaína e a identificação das suas sequências do terminal N leva-nos a concluir que os domínios  $C_H2$  e  $C_H3$  das cadeias pesadas  $\gamma 3$  têm um tamanho normal e que a supressão deve ocorrer quer no domínio  $C_H1$  quer no domínio  $V_{HH}$  para gerar a cadeia  $\gamma 3$  mais curta. As fracções que não se ligam à Sepharose' com Proteína A contêm duas bandas de 34 kd e 17 kd que são mais difusas no EGPA-DSS, indicando isso que a sua origem está na parte variável do terminal N da molécula (figura 4).

Após a redução, observa-se a existência de uma única banda difusa de 17 kd, o que indica que a banda de 34 kd é um dímero do componente de 17 kd ligado por pontes dissulfureto. O fragmento de 34 kd contém aparentemente o domínio  $V_{HH}$  da charneira e do terminal N.

Os dados da sequência da proteína podem ser utilizados para a construção de iniciadores oligonucleotídicos degenerados que facilitam a amplificação de ADNc ou de ADN genómico por RCP.

Demonstrou-se que os esplenócitos de camelo marcam as células que reagiram com os soros de coelho e com os soros anti-imunoglobulina de camelo e que por tal motivo o baço é um local de síntese, pelo menos de uma classe de

imunoglobulinas. O ADNc foi pois sintetizado a partir de ARNm do baço de camelo. As condições para o isolamento do ARN foram as seguintes: isolou-se ARN total a partir de baço de dromedário, pelo método do isotiocianato de guanidínio (15). Purificou-se o ARNm com esférulas paramagnéticas-oligoT.

Consegue-se sintetizar o ADNc utilizando 1 µg de matriz de ARNm, um iniciador oligo-dT e transcriptase reversa (BOEHRINGER MAN). A segunda cadeia do ADNc é obtida utilizando RNase H e polimerase I do ADN de *E. coli*, em conformidade com as instruções do fornecedor.

As sequências relevantes foram amplificadas por RCP: amplificou-se 5 ng de ADNc por RCP numa mistura de reacção de 100 µL (Tris 10 mM-HCl a pH 8,3, KCl 80 mM, MgCl<sub>2</sub> 15 mM, gelatina a 0,01% (p/v), 200 µmol de cada um dos dNTP e 25 pmol de cada iniciador) recoberta com óleo mineral (Sigma).

Os iniciadores degenerados, contendo os locais *EcoRI* e *KpnI*, foram ainda clonados no interior do pUC 18. Após um tratamento de desnaturação e de recombinação (a 94°C durante 5 minutos e a 54°C durante 5 minutos), acrescentou-se 2 unidades de polimerase de ADN de Taq à mistura de reacção antes de a submeter a 35 ciclos de amplificação: 1 minuto a 94°C (desnaturar), 1 minuto a 54°C (recombinar) e 2 minutos a 72°C (alongar). Para se amplificar as sequências de ADN entre os domínios V<sub>HH</sub> e C<sub>H2</sub> (clones da série nº 72), efectuou-se a RCP nas mesmas condições com a excepção de se ter aumentado para 60°C a temperatura de recombinação.

Um dos clones examinados (#56/36) tinha uma sequência correspondente à parte do terminal N de um domínio C<sub>H2</sub> idêntico à sequência do fragmento de 28 kd. A disponibilidade de dados desta sequência permitiu construir

um iniciador de 3' exacto e a clonagem da região da extremidade do terminal N entre os domínios  $V_{HH}$  e  $C_{H2}$ .

Os iniciadores de 5', correspondentes ao domínio  $V_{HH}$  de murganho (16) e que contêm um local de restrição *XhoI*, foram utilizados em conjunto com o iniciador de 3' em que havia sido inserido um local *KpnI* e as sequências amplificadas foram clonadas no interior do vector 'pBluescript®'. Fez-se digerir o clone #56136, que havia revelado dois locais internos *HaeIII*, com esta enzima para se obter uma sonda para identificar os clones positivos na RCP.

Após a amplificação, os produtos da RCP foram confirmados sobre um gel de agarose a 1,2% (p/v). A limpeza dos produtos da RCP consistiu numa extracção com fenol-clorofórmio a que se seguiu outra purificação por CLER (coluna de tipo 'GEN-PAC FAX', Waters) e finalmente mediante a utilização do estojo 'MERMAID' ou 'GENECLEAN II', (BIO 101, Inc.). A seguir a estes passos de purificação, fez-se digerir então o ADNc amplificado, tendo a digestão sido feita com *EcoRI* e *KpnI* para os clones da série nº 56 e com *XhoI* e *KpnI* para os clones da série nº 72. Fez-se uma extracção final com fenol-clorofórmio a anteceder a ligação no interior do pUC18 (clones da série nº 56) ou no interior de 'pBluescripte (clones da série nº 72).

Todos os clones obtidos tinham menos do que os 860 pares de bases que seria de esperar se possuísem uma região completa  $V_{HH}$  e  $C_{H1}$ . Os dados parciais da sequência, correspondentes ao terminal N da região  $V_{HH}$ , revelam que em 20 clones havia 3 que eram idênticos e possivelmente não eram independentes. As sequências obtidas assemelhara-se ao subgrupo III humano e aos subgrupos lila e Illb dos murinos (quadro 2).

Foram obtidos clones correspondentes a dois conjuntos diferentes de sequências de proteínas de domínio C<sub>H</sub>2. Um primeiro conjunto de sequências (#72/41) possuía uma região C<sub>H</sub>2 do terminal N idêntica àquela que foi obtida por sequenciação dos fragmentos de 28 kd da cadeia pesada  $\gamma$ 3 das proteínas, obtidos com papaína, possuía uma região de charneira curta com 3 resíduos cisteína e uma região variável correspondente aos resíduos da armação estrutural (AE4, em inglês FR4) codificados pelos minigenes J adjacentes à charneira. Faltava-lhes totalmente o domínio C<sub>H</sub>1. Este ADNc corresponde à cadeia  $\gamma$ 3 (quadro 4).

Numa sequência estreitamente afim (#72/1) a prolina na posição 259 é substituída por treonina.

A sequência correspondente à região C<sub>H</sub>3 e a parte restante da região C<sub>H</sub>2 foram obtidas por RCP realizada com o ADNc, utilizando como iniciador de *Kpn*I um poli-T em cuja extremidade 5' foi inserido um local de restrição *Kpn*I. A sequência total d cadeia  $\gamma$ 3 corresponde a um peso molecular (PM) que concorda bastante bem com os dados obtidos com o protocolo de EGPA-DSS.

A sequência desta cadeia  $\gamma$ 3 apresenta semelhanças com outras cadeias  $\gamma$ 7, com a excepção de lhe faltar o domínio C<sub>H</sub>1, estando o domínio V<sub>HH</sub> adjacente à charneira.

Um, ou a totalidade dos três resíduos cisteína, pode provavelmente ser responsável por manter unidas as duas cadeias  $\gamma$ 3.

Estes resultados permitiram-nos definir um modelo para a molécula IgG<sub>3</sub>, com base na sequência e clivagem com papaína (figura 5). A papaína pode clivar a molécula em cada lado das pontes dissulfureto da charneira e também entre os domínios C<sub>H</sub>2 e C<sub>H</sub>3. Em condições não redutoras, os domínios V<sub>HH</sub> da IgG<sub>3</sub> podem ser isolados sob a forma de um dímero ligado por pontes dissulfureto ou sob a forma de um

monómero, consoante o local onde a papaína realize a clivagem.

Há um segundo conjunto de clones #72/29 que tem uma sequência ligeiramente diferente para o domínio C<sub>H</sub>2 e que foi caracterizado por uma charneira muito comprida precedida imediatamente pelo domínio variável. Esta região da charneira possui três resíduos cisteína na extremidade do seu terminal C numa sequência homóloga com a da charneira  $\gamma$ 3. Esse segundo conjunto de clones pode representar eventualmente a subclasse de IgG<sub>2</sub>. Para a parte constante da cadeia  $\gamma$ 3 e também para a cadeia  $\gamma$ 2 putativa, a maior parte dos clones são idênticos, revelando sequências específicas das cadeias  $\gamma$ 2 ou  $\gamma$ 3. No entanto, há alguns clones, tais como o #7211, que revelam diferenças menos significativas. Por exemplo, no caso dos clones #72/1, são detectadas duas diferenças nos nucleótidos.

Foram agora sequenciados total ou parcialmente vários ADNc de regiões V<sub>HH</sub>, com a excepção de um troço curto na extremidade do terminal N, que deriva do iniciador.

Após a tradução, a maior parte revela as sequências características Ser<sub>21</sub> Cis<sub>22</sub> e Tir<sub>90</sub> Tir<sub>91</sub> Cis<sub>92</sub> das cadeias pesadas da ponte dissulfureto da infra-região V<sub>HH</sub> que liga os resíduos 22 e 29. Todos estes clones têm uma sequência correspondente aos resíduos da armação estrutural 4 (AE4, em inglês FR4) da região variável que precede imediatamente a sequência de charneira postulada (quadro 3). Esta sequência é gerada pelos minigenes J e é, na maior parte dos casos, semelhante à sequência codificada pelos minigenes J dos seres humanos e dos murinos. O comprimento da sequência entre as regiões Cis<sub>92</sub> e a extremidade do terminal C das regiões V<sub>HH</sub> é variável e nas sequências determinadas varia entre 25 e 37 aminoácidos, conforme

seria de esperar dos rearranjos dos minigenes J e D a variar de comprimento.

Surgem várias questões importantes levantadas pela existência única destas imunoglobulinas de cadeias pesadas numa situação não patológica. Acima de tudo, serão anticorpos genuínos? As imunoglobulinas de cadeias pesadas, obtidas a partir de camelus infectados com tripanossomas, reagem com um grande número de antigénios parasitas, conforme se demonstra na parte 1 destes exemplos. Isto implica que o sistema imunitário dos Camelídeos gera um número grande de locais de ligação constituídos por domínios  $V_{HH}$  singulares. Isto é confirmado pela diversidade das regiões  $V_{HH}$  das imunoglobulinas de cadeias pesadas obtidas por RCP.

A segunda questão é "como são segregadas?". A segregação de cadeias pesadas das imunoglobulinas, que constituem as imunoglobulinas do modelo de quatro cadeias, não ocorre em condições normais. Há uma proteína adventícia, a proteína de ligação de cadeias pesadas, ou proteína BIP, que impede a segregação das cadeias pesadas. A segregação apenas pode ocorrer (13) quando a cadeia leve desloca a proteína BIP no retículo endoplasmático.

Ao dímero de cadeias pesadas encontrado no soro de seres humanos ou de murganhos, que padeçam da chamada "doença das cadeias longas", falta-lhes os domínios  $C_H1$  que se presume que acomodem o local BIP (14). Na ausência deste domínio, a proteína BIP deixa de poder ligar-se e de impedir o transporte das cadeias pesadas.

A presença de uma classe de IgG 1 nos camelus, constituída por cadeias pesadas e leves que representam uma quantidade compreendida entre 25% e 50% das moléculas totais de IgG, levanta também o problema do modo como a maturação e o desvio de classe ocorre e ainda sobre o papel desempenhado

pela cadeia leve. A cadeia leve dos Camelídeos parece ser invulgarmente grande e heterogênea quando examinada por EGPA-DSS.

A maior dimensão de um domínio isolado é de 40 Å e o intervalo máximo que se consegue entre os locais de ligação de uma IgG convencional com domínios  $C_{H1}$  e  $V_{HH}$  irá ser da ordem de 160 Å ( $2V_{HH} + 2C_{H1}$ ) (19). A supressão do domínio  $CH1$  nos dois tipos de anticorpos de cadeias pesadas desprovidos de cadeias leves, já sequenciados, tem consequentemente uma modificação do seu intervalo máximo (figura 6). Na  $IgG_3$  a distância máxima entre as extremidades das regiões  $V_{HH}$  irá ser da ordem de 80 Å ( $2V_{HH}$ ). Isto poderia ser uma grave limitação à aglutinação ou à interligação. Na  $IgG_2$  isto é compensado pelo troço extremamente longo da charneira, constituído por uma repetição, que ocorre 12 vezes, da sequência Pro-X (em que o símbolo X representa Gln, Lis ou Glu) e que fica localizada do lado do terminal N em relação às pontes dissulfureto da charneira. Pelo contrário, na  $IgG_3$  humana, a charneira muito comprida que aparentemente surge também como resultado de duplicação das sequências, não contribui para aumentar a distância existente entre os dois locais de ligação, uma vez que esta charneira está intercalada com pontes dissulfureto.

O domínio  $V_{HH}$  singular também poderia permitir provavelmente, de uma forma considerável, a liberdade rotacional do local de ligação em relação ao domínio  $Fc$ .

Ao contrário das cadeias pesadas dos mielomas, que resultam provavelmente da supressão do domínio  $C_{H1}$  numa célula produtora de um anticorpo singular, ou de anticorpos de cadeias pesadas produzidos por clonagem de expressão (15), os anticorpos de cadeias pesadas dos Camelídeos (desprovidos de cadeias leves) surgiram num ambiente



imunológico normal e presume-se que tenham experimentado um refinamento selectivo em termos da especificidade e da afinidade associadas à maturação das células B.

Expressão e purificação da proteína V<sub>HH</sub>21 do camelo (RD21 na figura 7), a partir de *E. Coli*

Os clones podem ser expressos em vários tipos de vectores de expressão. Como exemplo, utilizando um vector 'Immuno PBS' comercialmente disponível (Huse et al: Science (1989) 246, 1275) foram obtidos clones, produzidos em 'Bluescribe', em conformidade com o procedimento anteriormente descrito, por meio de RCP, utilizando o mesmo iniciador de 5' que contém o local *Xho*I e um novo iniciador de 3', correspondente aos resíduos 113-103 na armação estrutural das imunoglobulinas, onde foi construído um local *Spe*I: TC TTA ACT AGT GAG GAG ACG GTG ACC TU. Este procedimento permite a clonagem do domínio V<sub>HH</sub> no local *Xho*/ *Spe* do vector Immuno PBS'. No entanto, a extremidade 3' do gene não estava em fase com o "marcador" de identificação nem com o codão de paragem do vector. Para se conseguir isso, cortou-se o arquétipo com *Spe*I e as 4 bases sobressalientes foram preenchidas utilizando o fragmento de Klenow e depois disso ligou-se novamente o vector.

O vector de expressão plasmídico 'ipBS' ('immunopBS') (Stratagene) contém uma sequência líder pelB que é utilizada para a expressão das cadeias da imunoglobulina em *E. coli* sob o controlo do promotor pLAC, um local de ligação dos ribossomas e os codões de paragem. Além disso, contém uma sequência para um marcador decapeptídico do terminal C.

A estirpe JM 101 de *E. coli*, que acomoda o plasmídeo ipBS-V<sub>HH</sub>21, desenvolveu-se em 1 L de meio constituído por CT (em inglês TB) com ampicilina na concentração de 100 µg/mL e 0,1% de glicose a 32°C. A expressão foi induzida pela

adição de IPTG 1 mM (concentração final) para um valor da  $D_{0.550}$  igual a 1,0. Após a indução de um dia para o outro a 28°C, efectuou-se a colheita das células por centrifugação a 4000 x g durante 10 minutos (4°C), tendo sido recolocadas em suspensão em 10 mL de tampão TES (Tris 0,1 M-HCl a pH 8,0, EDTA 0,5 mM e sacarose 0,5 M). Manteve-se a suspensão em gelo durante 2 horas. Efectuou-se a remoção das proteínas periplásmicas por choque osmótico mediante a adição de 20 mL de tampão TES diluído a 1:4 v/v com água, manteve-se em gelo durante 1 hora e depois centrifugou-se a 12 000 x g durante 30 minutos a 4°C. Efectuou-se a diálise da fracção periplásmica do sobrenadante em presença de Tris-HCl a pH 8,8, NaCl 50 mM, aplicou-se a uma coluna de escoamento rápido de tipo 'Q-Sepharose (Pharmacia), lavou-se com o tampão anteriormente referido e realizou-se a eluição com um gradiente linear de NaCl a variar desde 50 mM até 1 M no tampão.

As fracções que continham a proteína  $V_{HH}$  ainda foram purificadas numa coluna de 'Superdex 75' (Pharmacia) equilibrada com tampão STP (fosfato 0,01 M a pH 7,2 e NaCl 0,15 M). A produção de proteína VHM purificada varia entre 2 e 5 mg/L de cultura de células.

As fracções foram analisadas por EGPA-DSS (1). A identificação positiva do fragmento do anticorpo de  $V_{HH}$  do camelo foi feita por análise das autonadrogafias de Westem, utilizando um anticorpo criado em coelhos na presença de IgG<sub>3</sub> de camelo purificada e um conjugado de anti-IgG do coelho/fosfatase alcalina (2).

Como proteínas padrão (Pharmacia), foram utilizadas proteínas periplásmicas preparadas a partir de 1 mL de JM101/ipBS- $V_{HH}21$  com a indução feita com IPTG. A figura 8 mostra o seguinte: C,D: fracções provenientes da cromatografia em coluna de escoamento rápido sobre 'S-

Sepharose' (C: eluição feita com NaCl 650 mM, D: eluição feita com NaCl 700 mM); E,F: fracções provenientes da cromatografia em coluna de 'Superdex 75'.

Conforme se pode ver, a impureza principal é eliminada por cromatografia de permuta iónica e a maior parte das impurezas restantes são eliminadas por filtração através de gel.



Iniciador derivado	10										20										30									
	G	G	S	V	Q	T	G	G	S	L	R	L	S	C	F	I	S	G	L	T	F	D								# 72/4
	G	G	S	V	Q	T	G	G	S	L	R	L	S	C	A	V	S	G	F	S	F	S								# 72/3
	G	G	S	E	Q	G	G	G	S	L	R	L	S	C	A	I	S	G	Y	T	Y	G								# 72/7
	G	G	S	V	Q	P	G	G	S	L	T	L	S	C	T	V	S	G	A	T	Y	S								# 72/17
	G	G	S	V	Q	A	G	G	S	L	R	L	S	C	T	G	S	G	F	P	Y	S								# 72/18
D V Q L V A S G G G S V G A G G S L R L S C T A S G D S F S																														# 72/2
E V K L V E S G G G L V E P G G S L R L S C A T S G F T F S																														V <sub>H</sub> IIIA demurganho
E V Q L L S G G G L V Q P G G S L R L S C A A S G F T F S																														V <sub>H</sub> III de ser humano

## Quadro 2

Comparação das regiões Fr 1 do terminal N da V<sub>HH</sub> do camelo com a proteína do subgrupo V<sub>H</sub>III dos seres humanos e com a proteína do subgrupo V<sub>H</sub>IIIA do murganho.

Os resíduos específicos do subgrupo invariável estão acidentados.

	Armação estrutural 4	Genes J
Ser humano	W G Q G T L V T V S S	J1, J4, J5
	W G R G T L V T V S S	J2
	W G Q G T T V T V S S	J6
	W G Q G T M V T V S S	J3
Murino	W G Q G T T L T V S S	J1
	W G Q G T L V T V S S	J2
	W G Q G T S V T V S A	J3
	W G A G T T V T V S S	J4
Clones de ADNc		
Camelo	W G Q G T Q V T V S S	Clones
	W G Q G T Q V T V S S	# 72/19 = # 72/3
	W G Q G T L V T V S S	Clone 1
	W G R G T Q V T V S S	# 72/24
	W G Q G T H V T V S S	# 72/21
	W G Q G I Q V T A S S	# 72/16

### Quadro 3

Comparação de alguns resíduos da armação estrutural 4 encontrados na região  $V_{HH}$  do camelo com os resíduos da armação estrutural 4 correspondentes à região de consenso dos minigenes J dos seres humanos e dos murganhos

	Sequência da VHB de murganho										Amostra aleatoria										18 regiões V <sub>HH</sub> diferentes de camelo										
	D	Y	Y	G	S	S	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
95	95																														
1	A	L	Q	P	M	H	L	G	A	I	S	Q	H	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
2	V	S	L	M	H	L	G	A	I	S	Q	H	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
3	V	P	A	S	T	A	C	E	L	P	L	G	A	I	S	Q	H	-	-	-	-	-	-	-	-	-	-	-	-	-	-
4	F	C	Y	S	G	T	C	C	G	A	C	D	A	E	S	I	L	P	A	P	Y	-	-	-	-	-	-	-	-	-	-
7	E	L	S	G	Y	H	T	C	C	G	A	C	D	A	E	S	I	L	P	A	P	Y	-	-	-	-	-	-	-	-	-
9	D	W	K	Y	H	T	C	C	G	A	C	D	A	E	S	I	L	P	A	P	Y	-	-	-	-	-	-	-	-	-	-
11	R	L	T	E	H	T	C	C	G	A	C	D	A	E	S	I	L	P	A	P	Y	-	-	-	-	-	-	-	-	-	-
13	Q	K	K	D	R	S	S	I	H	P	Y	C	N	L	D	P	H	Y	A	T	W	S	-	-	-	-	-	-	-	-	-
16	G	S	R	F	S	I	H	P	Y	C	N	L	D	P	H	Y	A	T	W	S	-	-	-	-	-	-	-	-	-	-	-
17	A	D	S	P	C	Y	H	P	Y	C	N	L	D	P	H	Y	A	T	W	S	-	-	-	-	-	-	-	-	-	-	-
18	D	S	S	F	Y	H	P	Y	C	N	L	D	P	H	Y	A	T	W	S	-	-	-	-	-	-	-	-	-	-	-	-
19	T	E	I	E	H	G	G	A	C	D	A	E	S	I	L	P	A	P	Y	-	-	-	-	-	-	-	-	-	-	-	-
20	T	E	I	E	H	G	G	A	C	D	A	E	S	I	L	P	A	P	Y	-	-	-	-	-	-	-	-	-	-	-	-
21	N	Q	L	A	G	G	A	C	D	A	E	S	I	L	P	A	P	Y	-	-	-	-	-	-	-	-	-	-	-	-	-
24	R	L	T	E	H	G	G	A	C	D	A	E	S	I	L	P	A	P	Y	-	-	-	-	-	-	-	-	-	-	-	-
25	D	G	W	T	R	K	E	L	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
27	D	S	Y	P	C	H	L	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
29	V	E	Y	P	I	A	D	M	C	S	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	

RDC3

Seres humanos e murganhos -

Camelo

amplitude 0-19 aminoácido

8-24 aminoácido

mais de 600 entradas

18 entradas

## QUADRO 4

10	20	40
EVQLVESGGG	LVQPGGSLRL SCAASG	CDR1: WVRQA PGKGLEWVS CDR2:
GG	SVQGGGSLRL SCAISG	CDR1: WFREG PGKEREGLIA CDR2:
GG	SVQAGGSLRL SCASSS	CDR1: WYRQA PGKEREFEVS CDR2:

70	80	90	110
RFTIS	RDNSKNTLYL QMNSLRAEDTAVY	YCAR CDR3:	WGQGTSLVT VSS
RFTIS	QDSTLKTHYL LMNNLXPEDTGTY	YCAA CDR3:	WGQGTQVT VSS
RFTIS	QDSAKNTVYL QMNSLKPEDTAMY	YCKI CDR3:	WGQGTQVT VSS

	V <sub>H</sub> H do camelo	charneira	C <sub>H</sub> <sup>2</sup>
	WGQGTQVT VSS	GTNEVCKCPKCP	APELPGG PSVFVFP
camelo	WGQGTQVT VSS	BPKIPQPQPKPQPKPQ	
		QFQPKPQ	
		KPEPECTCPKCP	APELLGG PSVFIFP

Quadro 5(1)

	C <sub>H</sub> 1 humano	charneira	C <sub>H</sub> <sup>2</sup>
γ3 humana	KVDKRV	ELKTPLGDTTHTCPRCP	
		EPKCSDTPPPCPRCP	
		EPKSCDTPPPCPRCP	APELLGG PSVFLFP
γ1 humana	KVDKK	AEPKSCDRTHTCPPCP	APELLGG PSVFLFP
γ2 humana	KVKVTV	ERKCCVECPPCP	APPVAG - PSVFLFP
γ4 humana	KVDKRV	ESKYGPCCPSCP	APEFLGG PSVFLFP

Quadro 5(2)



# REFERÊNCIAS

1. Ward, E.S., Gussow, D., Griffiths, A.D., Jones, P.T. e Winter, G, Nature 341, 544-546 (1989).
2. Ungar-Waron H., Eliase E., Glukman A. e Trainin Z. Isr. J. Vet. Med., 43, 198-203 (1987).
3. Bajyana Songa E. e Hamers R., Ann. Soc. Belge Med. Trop., 68, 233-240 (1988).
4. Edelman G.M., Olins D.E., Gally J.A. e Zinder N.D., Proc. Nat. Acad. Sci., 50, 753 (1963).
5. Franek F. e Nezlin R S., Bioldinniya, 28, 193 (1963).
6. Roitt I.M., Brostof J. e Male D.K., Immunology, Gower Med. Pub, Londres, Nova Iorque, p. 9.2. (1985).
7. Schifer M., Girling R.L., Ely K.R. e Edmundson B., Biochemistry, 12, 4620-4631 (1973).
8. Fleischman J.B., Pain R.H. e Porter R.R., Arch. Biochem. Biophys., Supl. 1, 174 (1962).
9. Roholt O., Onoue K. e Pressman D., PNAS 51, 173-178 (1964).
10. Seligmann M., Mihaesco E., Preud'homme J.L., Danon F. e Brouet J.C., Immunological Rev., 48, 145-167 (1979).
11. Hendershot L., Bole D., Köhler G. e Kearney J.F., The Journal of Cell Biology, 104, 761-767 (1987).
12. Hendershot L.M., The Journal of Cell Biology, 111, 829-837 (1990).
13. Hamers-Casterman, C., E. Wittouck, W. Van der Loo e R. Hamers, Journal of Immunogenetics, 6, 373-381 (1979).
14. Applied Biosystems - Ethanol Precipitation of Electro Eluted Electrodialysed Sample, Publicação nº 27.
15. Maniatis, T., E.F. Fritsch e J. Sambrook, Molecular Cloning: A Laboratory (1988).
16. Sastry et al., PNAS, 86, 5728 (1989).
17. Sanger, F., S. Nicklen e A.R. Coulson, Proc. Nati. Acad. Sci U.S.A., 74, 5463-5467 (1977).

18. Kabat E.A., Tai Te Wu, M. Reid-Miller, H.M. Peny e K.S. Gottesman, U.S. Dpt of Health and Human Services, Public Health Service, National Institutes of Health (1987).
19. Valentine, R.C. e N.M. Geen, J.M.B., 27, 615-617 (1967).

Lisboa, 15 de Março de 2010

### **REIVINDICAÇÕES**

1. Região  $V_H$  derivada de uma imunoglobulina com 4 cadeias, que contém um sítio de ligação ao antigénio ou vários sítios de ligação ao antigénio e em que os resíduos de aminoácidos foram parcialmente substituídos por resíduos de aminoácidos ou por sequências específicas de uma imunoglobulina (chamada imunoglobulina de cadeia pesada) que compreende duas cadeias polipeptídicas pesadas capazes de reconhecer e de se ligar a um ou a vários antigénios, sendo a referida imunoglobulina de cadeia pesada desprovida de cadeias leves e sendo obténível a partir de Camelídeos e em que na referida  $V_H$  derivada, a leucina, a prolina ou a glutamina na posição 45 da região  $V_H$  foi substituída por um resíduo de aminoácido carregado ou por um resíduo de cisteína.

2. Região  $V_H$  derivada de uma imunoglobulina com 4 cadeias, em que os resíduos de aminoácidos foram parcialmente substituídos por resíduos de aminoácidos ou por sequências específicas de uma imunoglobulina (chamada imunoglobulina de cadeia pesada) que compreende duas cadeias polipeptídicas pesadas capazes de reconhecer e de se ligar a um ou a vários antigénios, sendo a referida imunoglobulina de cadeia pesada desprovida de cadeia leves e sendo obténível a partir de Camelídeos e em que na referida  $V_H$  derivada, a leucina, a prolina ou a glutamina na posição 45 da região  $V_H$  foi substituída por um resíduo de aminoácido carregado ou por um

resíduo de cisteína e que tem maior solubilidade.

3. A região  $V_H$  de acordo com a reivindicação 1 ou 2, em que a leucina, a prolina ou a glutamina na posição 45 da região  $V_H$  foi substituída por um resíduo de arginina, ácido glutâmico ou cisteína.

4. Região  $V_H$  de acordo com as reivindicações 1, 2 ou 3, em que a referida  $V_H$  contém ela própria um sítio de ligação ao antígeno ou vários sítios de ligação ao antígeno e que funciona na ausência de  $V_L$ .

5. Região  $V_H$  de acordo com a reivindicação 1, 2 ou 3, em que a maior solubilidade da referida  $V_H$  é conseguida a uma concentração acima de 0,5 mg/mL, preferencialmente acima de 1 mg/mL e mais vantajosamente acima de 2 mg/mL.

6. A região  $V_H$  de acordo com qualquer uma das reivindicações 1 até 5 em que uma ligação dissulfureto é estabelecida na região variável, envolvendo resíduos de aminoácidos na região CDR3.

7. A região  $V_H$  de acordo com qualquer uma das reivindicações 1 até 5, na qual as ansas CDR da região estão ligadas a outras partes da região  $V_H$  por introdução de cisteínas emparelhadas, em particular na qual a ansa CDR<sub>3</sub> está ligada a FR<sub>2</sub> ou a CDR<sub>1</sub> e mais especialmente onde a cisteína da CDR<sub>3</sub> da  $V_H$  está ligada a uma cisteína na posição 31 ou 33 da CDR<sub>1</sub> ou na posição 45 da FR<sub>2</sub>.

8. A região  $V_H$  de acordo com qualquer uma das

reivindicações 1 até 5, em que as ansas CDR estão ligadas a outras partes de uma região V por introdução de cisteínas emparelhadas.

9. Processo para a preparação de uma imunoglobulina de cadeia pesada que é desprovida de cadeias leves e do primeiro domínio (CH1) da região constante das suas cadeias polipeptídicas pesadas ou um fragmento seu, sendo o referido fragmento seleccionado a partir de um fragmento correspondente a uma cadeia polipeptídica pesada, fragmentos obtidos por digestão enzimática, especialmente aqueles obtidos por digestão parcial com papaína conduzindo ao fragmento Fc, conduzindo ao fragmento  $FV_{HH}h$  ou ao seu dímero  $F(V_{HH}h)_2$ , ou a um fragmento obtido por digestão adicional com papaína do fragmento Fc, conduzindo ao frgmento pFc, fragmentos homólogos obtidos com outras enzimas proteolíticas, um fragmento com pelo menos 20 aminoácidos da região variável da imunoglobulina, ou o domínio variável completo, especialmente um fragmento correspondente aos domínios  $V_{HH}$  isolados ou aos dímeros  $V_{HH}$  ligados à articulação dissulfureto, ou um fragmento correspondente a pelo menos 10 preferencialmente 20 aminoácidos da região constante ou à região constante completa da imonuglobulina, o referido processo compreendendo os passos de:

- seleccionar imunoglobulinas tendo duas cadeias polipeptídicas pesadas e sendo desprovidas das cadeias polipeptídicas leves a partir de um animal da família dos Camelídeos e recuperar as mesmas.

10. Processo de acordo com a reivindicação 9 que compreende adicionalmente o passo de preparar fragmentos das imonuglobulinas seleccionadas.

11. Processo para a preparação da região  $V_H$  de acordo com qualquer uma das reivindicações 1 até 8, compreendendo os passos de:

- substituir resíduos de aminoácidos na região  $V_H$  derivada de uma imunoglobulina com 4 cadeias, por resíduos de aminoácidos ou sequências específicas de uma imunoglobulina (chamada de imunoglobulina de cadeia pesada) que compreende duas cadeias polipeptídicas pesadas capazes de reconhecer e de se ligar a um ou a vários antigénios, sendo a referida imunoglobulina de cadeia pesada desprovida de cadeias leves e sendo obtível a partir de Camelídeos,
- recuperar a referida região  $V_H$ .

12. Processo para a preparação da região  $V_H$  de acordo com a reivindicação 11, em que os resíduos de aminoácidos ou sequências específicas de uma imunoglobulina (chamada imonuglobulina de cadeia pesada) que compreende duas cadeias polipeptídicas pesadas capazes de reconhecer e de se ligar a um ou a vários antigénios, sendo a referida imunoglobulina de cadeia pesada desprovida de cadeias leves, que são inseridos na referida região  $V_H$  são fragmentos peptídicos da região  $V_{HH}$  da imunoglobulina de cadeia pesada.

Lisboa, 15 de Março de 2010

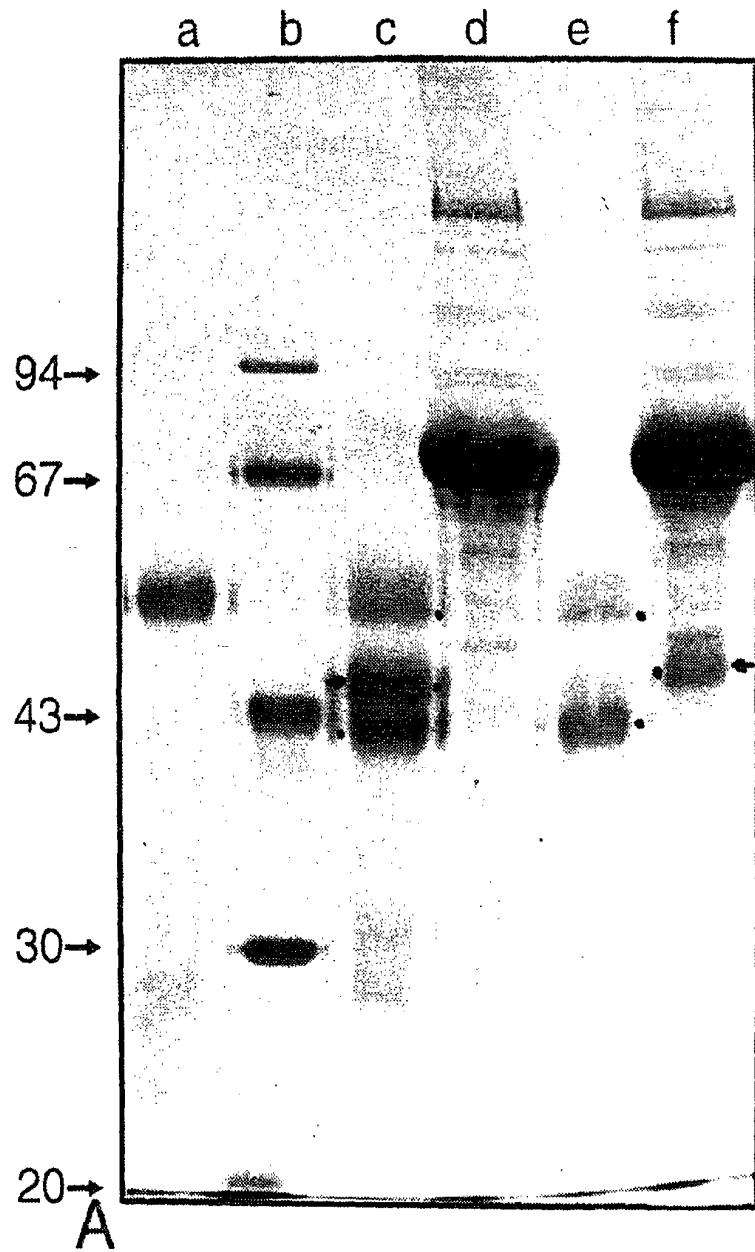


FIGURA 1A

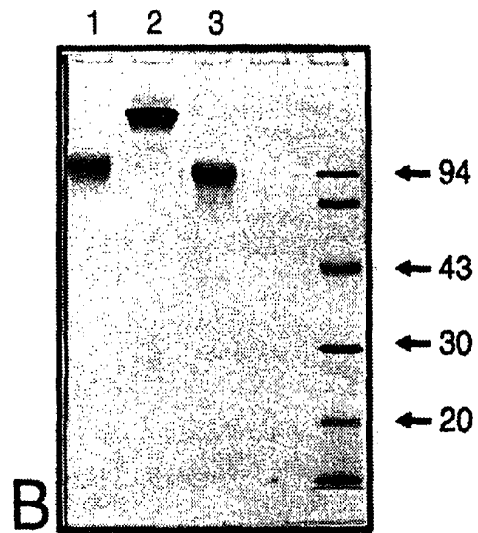


FIGURA 1B

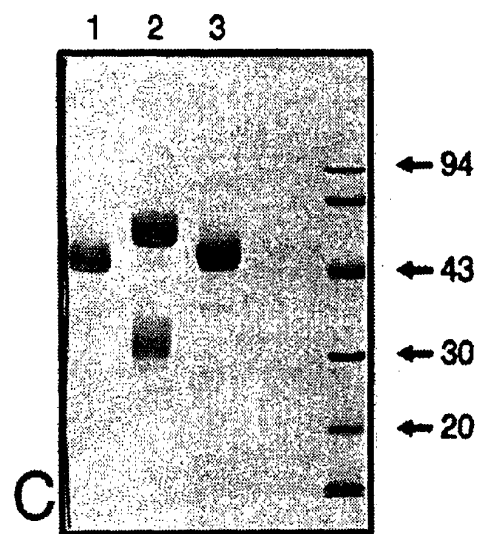


FIGURA 1C



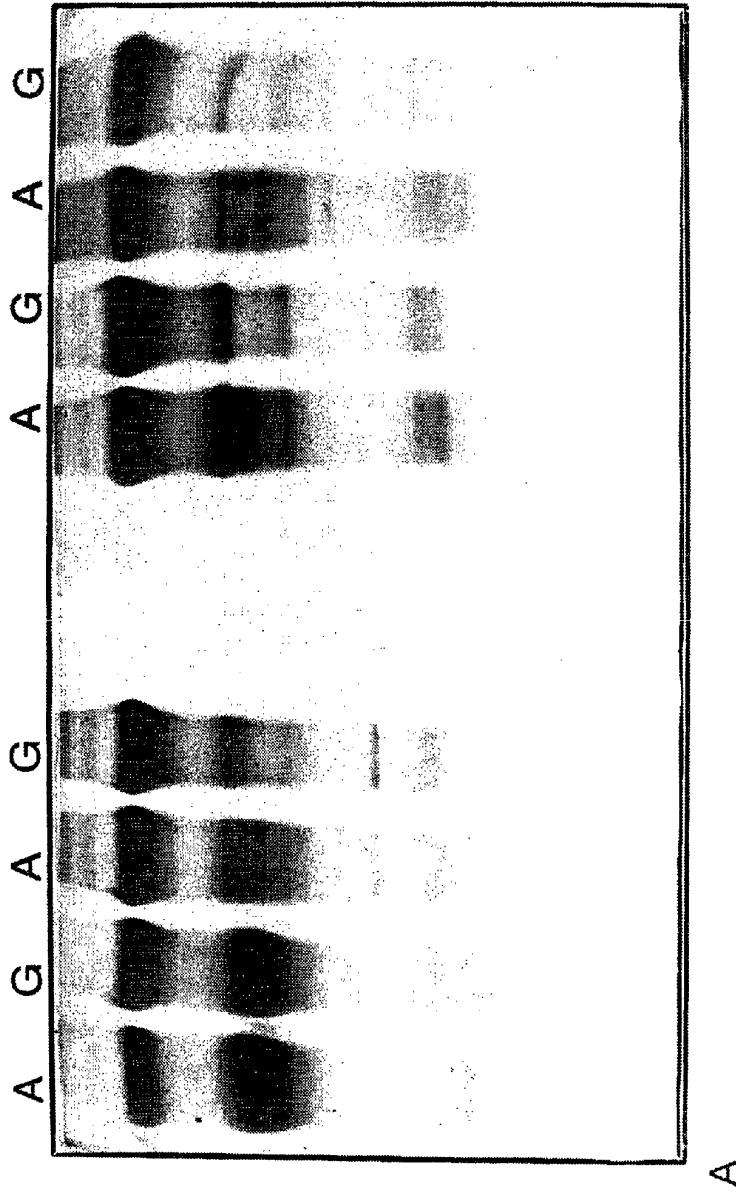


FIGURA 2A

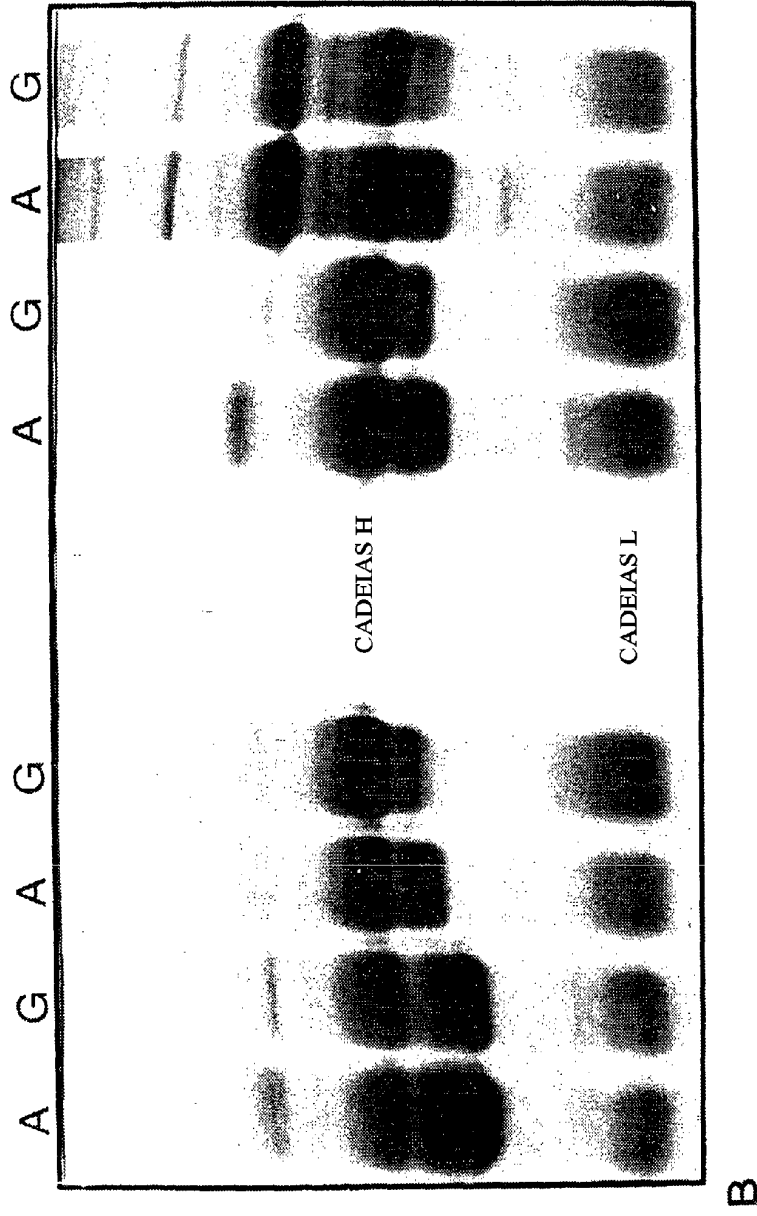


FIGURA 2B

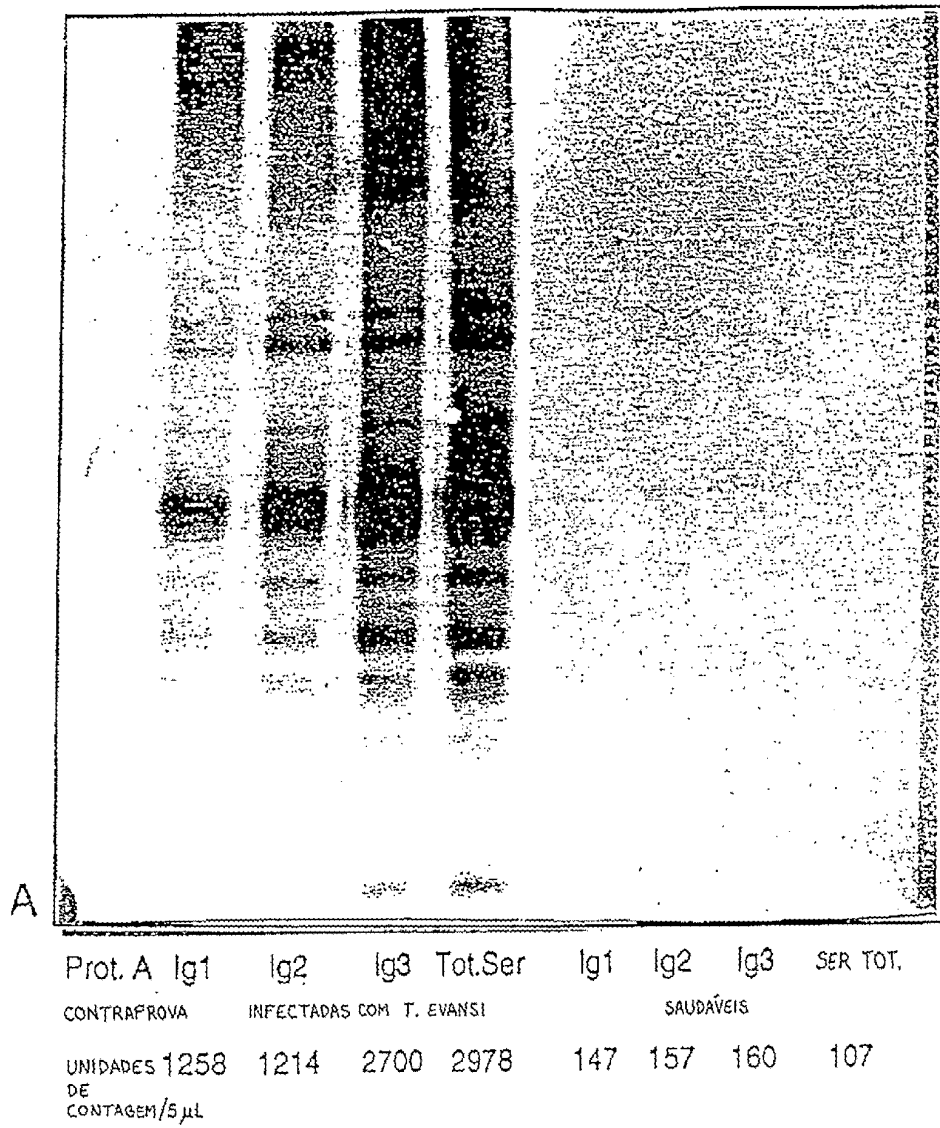


FIGURA 3A

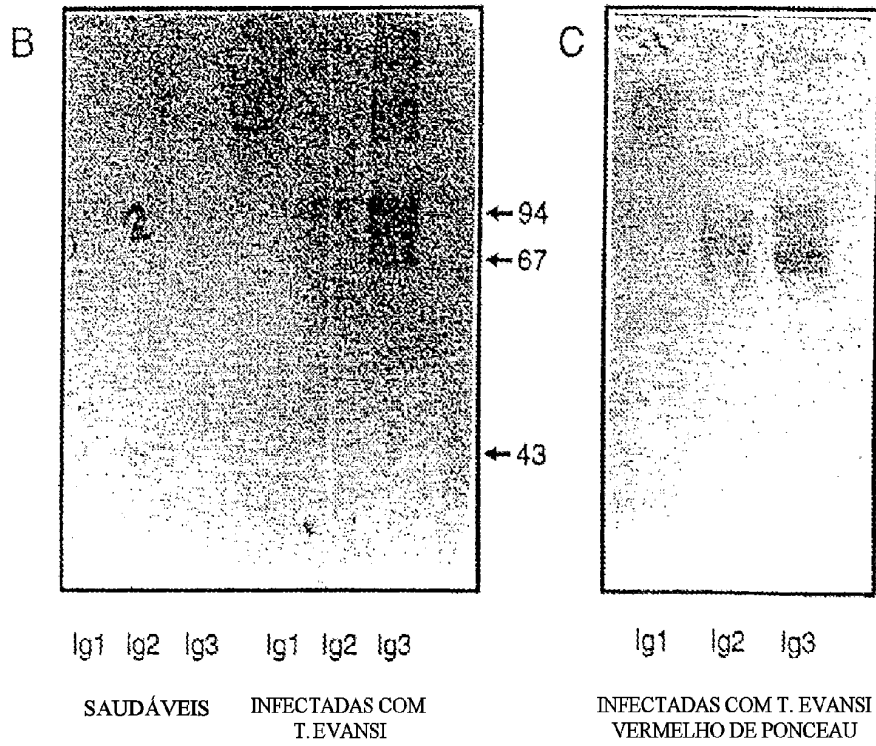


FIGURA 3B

FIGURA 3C

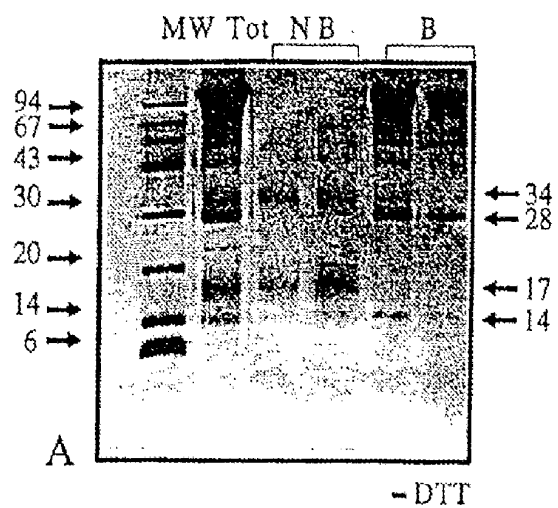


FIGURA 4A

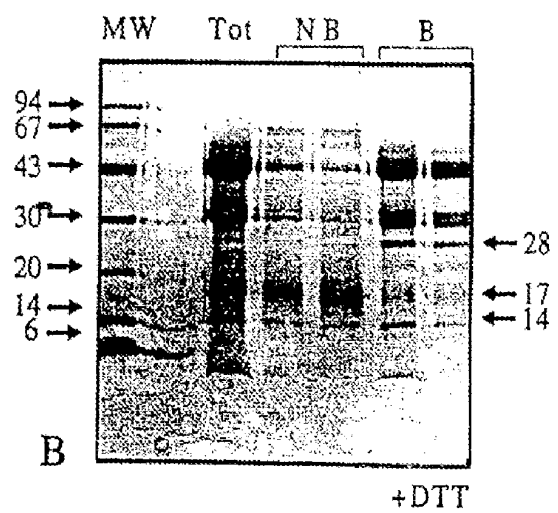


FIGURA 4B

ANÁLISE FEITA POR EGPA-DSS AOS  
FRAGMENTOS DE IgG<sub>3</sub> OBTIDOS COM PAPAÍNA

# Modello para a IgG3 do camelo

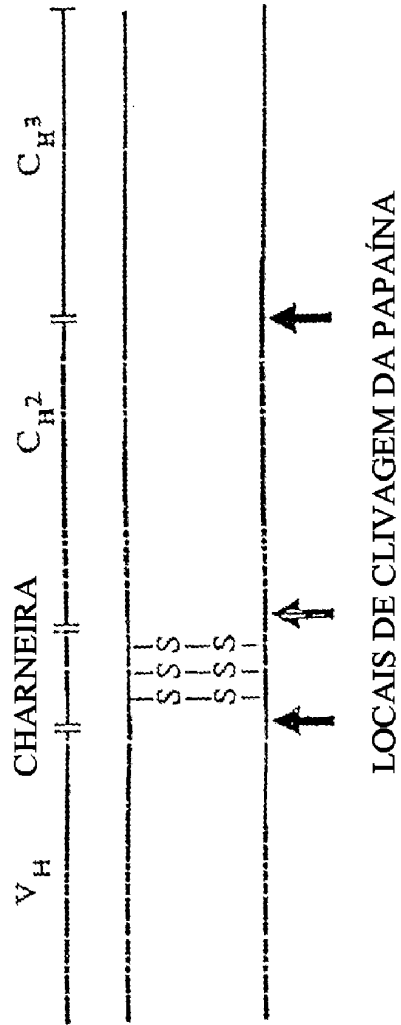


Fig. 5: Representação esquemática do modelo de IgG<sub>3</sub> do camelo

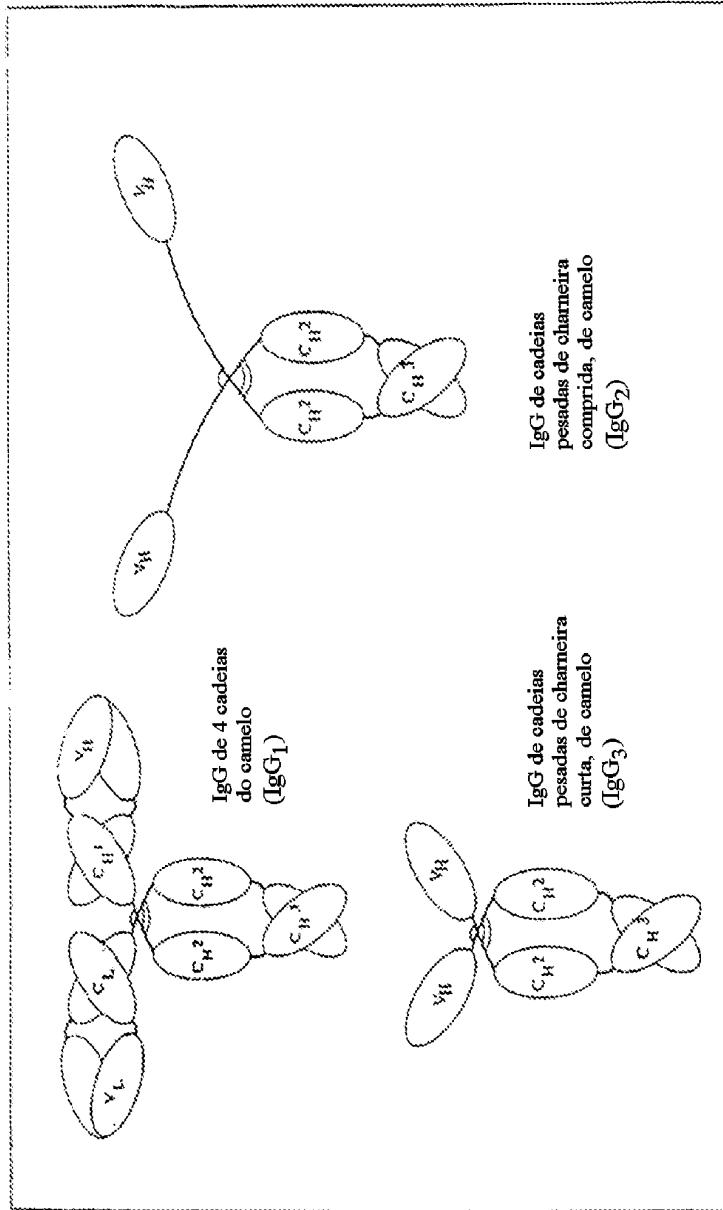


Figura 6: representação esquemática das imunoglobulinas IgG<sub>1</sub> e IgG<sub>2</sub> e IgG<sub>3</sub> putativas. A sequência (Pro-X)<sub>12</sub> da molécula IgG putativa pode ser modelada numa repetição 6 aa

```

DR01006 C-----TCGAG---TCTGGGGGAGG
DR27006 C-----TCGAG---TCTGGGGGAGG
DR03006 C-----AGGTGA-----AACTGCTCGAG---TCTGGAGGAGG
DR11006 C-----TCGAG---TCTGGGGGAGG
DR24006 C-----AGGTGA-----AACTGCTCGAG---TCTGGGGGAGG
DR16006 C-----TCGAG---TCTGGAGGAGG
DR19006 C-----TCGAG---TCTGGAGGAGG
DR07006 C-----TCGAG---TCTGGGGGAGG
DR16006 C-----TCGAG---TCTGGGGGAGG
DR20006 C-----TCGAG---TCAGGGGGAGG
DR25006 C-----TCGAG---TCTGGGGGAGG
DR20006 C-----TCGAG---TCTGGAGGAGG
DR21006 C-----TCGAG---TCTGGGGGAGG
DR09006 C-----AGGTGA-----AACTGCTCGAG---TCTGGGGGAGG
DR17006 C-----TCGAG---TCTGGGGGAGG
DR13006 C-----TCGAG---TCAGGGGGAGG
DR02006 CTCGAGTCAGGTGTCCGGTCTGATGTGCAGCTGGTGGCGTCTGGGGGAGG

DR01006 ATCGGTGCAGGCTGGAGGGTCTCTGAGACTCTC--GTGCG-CAGCCTCTG
DR27006 CTCGGTGCAGGCTGGAGGGTCTCTGAGACTCTCCTGTGCATCTTCTTCTA
DR03006 CTCGGTGCAGACTGGAGGATCTCTGAGACTCTCCTGTGCAGT--C-TCTG
DR11006 GTCGGTGCAGGCTGGAGGGTCTCTGAGACTCTCCTGTAATGT--C-TCTG
DR24006 GTCGGTGCAGGCTGGAGGGTCTCTGAGACTCTCCTGTAATGT--C-TCTG
DR16006 CTCGGGCGAGGCTGGAGGATCTCTGAGACTCTCCTGTGCAGC--CCACGG
DR19006 CTCGGTTCAGGCTGGAGGGTCCCTTAGACTCTCCTGTGCAGC--C-TCTG
DR07006 CTCGGTGCAGGGTGGAGGGTCTCTGAGACTCTCCTGTGCAA--TCTCTG
DR16006 CTCGGTGCAGGCTGGAGGGTCTCTGAGACTCTCCTGTACAG--GCTCTG
DR20006 CTCGGTACAGGTGGAGGGTCTCTGAGACTCTCCTGTGTAG--CCTCTA
DR25006 CTCGGTACAAACTGGAGGGTCTCTGAGACTCTCTGTGCG--AAATCTCTG
DR20006 CTCGGTGCAGGCTGGAGGGTCTCTGAGACTCTCCTGTG--TAGCCTCTG
DR21006 CTCGGTGCAGGTTGGAGGGTCTCTGAAACTCTCCTGTAAAT--CTCTG
DR09006 CTCGGTGCAGGCTGGAGGGTCTCTGACACTCTCTGTG--TATACAC--
DR17006 CTCGGTCCAACCTGGAGGATCTCTGACACTCTCCTGTACAGT--TCTG
DR13006 CTCGGTGGAGGCTGGAGGGTCTCTGAGACTCTCCTGTACAG--CCTCTG
DR02006 CTCGGTGCAGGCTGGAGGCTCTCTGAGACTCTCCTGTACAG--CCTCTG

DR01006 GA--TACAGTAATT--GTCCCCTCACTTG-GAGCTGGTATCGCCAGTTT
DR27006 AA--TATATGCCTT--GCACCTACGACAT-GACCTGGTACCGCCAGGCT
DR03006 GA--TTCTCCTTTA--GTACCACTTGTAT-GGCCTGGTTCGCCAGGCT
DR11006 GC--TCTCCAGTA--GTACTTATTGCCT-GGGCTGGTTCGCCAGGCT
DR24006 GC--TCTCCAGTA--GTACTTATTGCCT-GGGCTGGTTCGCCAGGCT
DR16006 GA--TTCCGC-TCA--ATGGTTACTACAT-CGCCCTGGTTCGGTCAAGCT
DR19006 AC--TACACCATCA--CTGATTATTGCAT-GGCCTGGTTCGCCAGGCT
DR07006 GA--TACACGTACG--GTAGCTTCTGTAT-GGGCTGGTTCGCCAGGCT
DR16006 GA--TTCCCTATA--GTACCTTCTGTCT-GGGGTGGTTCGCCAGGCT
DR20006 CT--CACACCGACA--GTAGCACCTGTAT-AGGCTGGTTCGCCAGGCT
DR25006 GA--TTGACTTTTG--ATGATTCTGACGT-GGGGTGGTACCGCCAGGCT
DR20006 GA--TTCAATTTTG--AAACTTCTCGTAT-GGCCTGGTTCGCCAGGCT
DR21006 GAGGTACCCAGATCGTGTTCCTAAATCTTTGGCCTGGTTCGCCAGGCT
DR09006 -----CAACGATACTGGGACCA-----TGGGATGGTTTCGCCAGGCT
DR17006 --GGGCCACCTACA--GTGACTACAGTATTGGA-TGGATCCGCCAGGCT
DR13006 G-----ATACGTAT-CCT-----CTATGGCCTGGTTCGCCAGGCT
DR02006 GAGA-----CAGTTTCAGTAGATT--TGCCATGTCTTGGTTCGCCAGGCT

```

FIG. 7A



```

DR01006 CCAGGAACGGAGCGCGAGTTCGTCTCCAGTATGGATCCGGATGGAAATAC
DR27006 CCAGGCAAGGAGCGCGAATTTGTCTCAAGTATAAATATTGATGGTAAGAC
DR03006 TCAGGAAAGCAGCGTGAGGGGTCGCGAGCCATTAAATAGTGCGCGTGGTAG
DR11006 CCAGGGAGGGAGCGTGAGGGGGTCACAGCGATTAA-----CACTGATGG
DR24006 CCAGGGAAGGAGCGTGAGGGGGTCACAGCGATTAA-----CACTGATGG
DR16006 CCTGGGAAGGGGCGTGAGGGGGTCGCAACAATTAATGGTGGTCG-----
DR19006 CCAGGGAAGGAGCGTGAAATGGTCGAGCGATTCAAGTTGTCCGTAGTGA
DR07006 CCAGGCAAGGAACGTGAGGGGATCGCAACTATTCTTAATGGTGGTACTAA
DR16006 CCAGGGAAGGAACGTGAGGGGGTCGCGGGTATTAAATAGTGCGAGGAGTAA
DR20006 CCAGGGAAGGAACGTGAGGGGGTCGCAAGTATATATTTTGGTGGTGGTGG
DR25006 CCAGGGCATGAGTGCAATTTGGTCTCAGGTATTCTGAGTGAATGGTACT-C
DR20006 CCAGGAAATGTGTGTGAGTTGGTCTCAAGTATTTACAGTGAATGG-----
DR21006 CCAGAGAAGGAGCGCGAAGGGGATCGCAGTTCTTTGACTAAGGATGGTAA
DR09006 CCAGGGAAGAGTGCAGAAAGGGTCGCGCATATTACGCTGATGGTATGA-
DR17006 CCAGGGAAGGACCGTGAAAGTAGTCGAGCCGCTAATACTGGTG-----
DR13006 CCAGGGCAGGAGCGCGAGGGGGTCGCGTTTGTCAAACGG-----
DR02006 CCAGGGAAGGAAGTGCAGATTGGTCTCAAGCATTCAAAGTAATGGAAGGAC
CAAGTACA-----CATACTCCGTGAAGGGCCGCTTCACC
AACATACG-----CAGACTCCGTGAAGGGCCGATTTCACC
DR03006 GACATACTA-CAACACATATGTCGCGGAGTCCGTGAAGGGCCGATTTCACC
DR11006 CAGTATCAT-ATACGCA-----GCCGACTCCGTGAAGGGCCGATTTCACC
DR24006 CAGTGTCTA-ATACGCA-----GCCGACTCCGTGAAGGGCCGATTTCACC
DR16006 -----CGA-CGTCACATACTACGCCGACTCCGTGAAGGGCCGATTTCACC
DR19006 TACT--CGC-C-TCACAGACTACGCCGACTCCGTGAAGGGACGATTTCACC
DR07006 -----CACATACTATGCCGACTCCGTGAAGGGCCGATTTCACC
DR16006 -----TACTTACTATGCCGACTCCGTGAAGGGCCGATTTCACC
DR20006 -----TACGAATTATCCGCACTCCGTGAAGGGCCGATTTCACC
DR25006 CATATACAAAGAGTGGAGACTATGCTGAGTCTGTGAGGGGGCCGGTTACC
DR20006 CA-AAACATACTACGTCGACC--GCA-----TGAAGGGCCGATTTCACC
DR21006 GA-----CATTCTATGCCGACTCCGTGAAGGGCCGATTTCACC
DR09006 -----CCTTCATTGATGAACCGTGAAGGGCCGATTTCACC
DR17006 -----CGACTAGTAAATTCTACGTCGACTTTGTGAAGGGCCGATTTCACC
DR13006 --CTGACAAT-AGTGCAATTATATGGCGACTCCGTGAAGGGCCGATTTCACC
DR02006 AACTGA-----GGCCGATTCCGTGCAAGGGCCGATTTCACC
ATGTCCCGAGGCGAGCACCAGTACACAGTATTTCTGCAATGGACAATCT
DR27006 ATCTCCCAAGACAGCGCCCAAGAACACGGTGTATCTGCAGATGAACAGCCT
DR03006 ATCTCCCAAGACAACGCCAAGACCACGGTATATCTTGATATGAACAACCT
DR11006 ATCTCCCAAGACACCGCCCAAGGAACGGTACATCTCCAGATGAACAACCT
DR24006 ATCTCCCAAGACACCGCCCAAGGAACGGTATATCTCCAGATGAACAACCT
DR16006 ATCTCCCGAGACAGCCCCAAGAAATACGGTGTATCTGCAGATGAACAGCCT
DR19006 ATCTCCCAAGGCAACACCAAGAACACAGTGAATCTGCAATGAACAGCCT
DR07006 ATCTCCCAAGACAGCACGTTGAAGACGATGTATCTGCTAATGAACAACCT
DR16006 ATCTCCCAAGGGAATGCCAAGAATACGGTGTCTGCAATGGATAACTT
DR20006 ATCTCCCAACTCAACGCCCAGAACACAGTGTATCTGCAATGAACAGCCT
DR25006 ATCTCCAGAGACAACGCCAAGAACATGATATACCTTCAATGAACGACCT
DR20006 ATTTCTAGAGAGAATGCCAAGAAATACATTGTATCTACAACTGAGCGGCCT
DR21006 ATCTTCTTAGATAATGACAAGACCACTTTCTCCTTACAACTTGATCGACT
DR09006 ATCTCCCGAGACAACGCCAAGAAACGTTGTCTTTGCGAATGAATAGTCT
DR17006 ATTTCCCAAGACAACGCCAAGAAATACGGTATATCTGCAATGAGCTTCT
DR13006 ATCTCCACGACAACGCCAAGAACACGGTGTATCTGCAATGCGCAACCT
DR02006 ATCTCCCGAGACAATTCAGGAACACAGTGTATCTGCAATGAACAGCCT

```

FIG. 7B

```

DR01006 GAAACCTGAGGACACGGCGATGTATTACTGTAAAAC-A---GCCCTAC--
DR27006 GAAACCTGAGGACACGGCGATGTATTACTGTAAAAT-A---GA---TTC--
DR03006 AACCCCTGAAGACACGGCTACGTATTACTGTGCGGCGG---TCCCAGCCC
DR11006 GCAACCTGAGGATACGGCCACCTATTACTGCGCGGCAA---GACTGACGG
DR24006 GCAACCTGAGGATACGGCCACCTATTACTGCGCGGCAA---GACTGACGG
DR16006 GAAACCTGAGGACACGGCCATCTACTTCTGTGCAGCAG---G---CTC
DR19006 GACACCTGAGGACACGGCCATCTACAGTTGTGCGGCAA---C---CAG
DR07006 GAAACCTGAAGACACGGGCACCTATTACTGTGCTG-CA---GAACTAAGT
DR16006 GAAACCTGAGGACACGGCCATCTATTACTGCGCGG-CG---GATAGTCCA
DR20006 GAAACCTGAGGACACGGCCATGTACTACTGTGCAATCA---CTGAAATTG
DR25006 GAAACCTGAGGACACGGCCATGTATTACTGCGCGGTAGATGGTTGGACCC
DR20006 CAAACCTGAGGACACGGCCATGTATTACTGTGCG-----CC
DR21006 GAACCCGGAGGACACTGCCGACTACTACTGCGCTGCAATCAATTAGC--
DR09006 GAGGCCCTGAGGACACGGCCGTGTATTACTGTGCGGCGAGATTG-----
DR17006 GAAACCTGAGGACACGGCCATCTATTACTGTGCGGCGAG---CGGACCC
DR13006 GCAACCTGACGACACTGCGGTGTACTACTGTGCGGCC-----CAA
DR02006 GAAACCCGAGGACACGGCCGTGTATTACTGTGGGGCAGT-----

DR01006 -----A-AC--CTGGGGGTTATTGTGGGTA-
DR27006 -----GTAC--CCGTGCCATCTCCTTGATG-
DR03006 ACTTGGGACCT-----GGCG-CCATT-----CTTGATTTG
DR11006 AGATGGGGGCTTGTGATGCGAGATGGGCGACCTTAGC--GACAAGGAC-G
DR24006 AGATGGGGGCTTGTGATGCGAGATGGGCGACCTTAGC--GACAAGGAC-G
DR16006 GCGTTTTT-CTAGTCTGTGGGAGCACTTC-TAGAC---TCGAAAGTAG-
DR19006 TAGTTTTTACTGGTACT-----GCAC-----C---ACG-----G
DR07006 GGTGGTAGTTGTGAATTGC---CTTTGC-----TATTTGACTA-----
DR16006 TGTACATGCCGACTATGC---CCGTCCCCGATACGAGACAGTTTTTG
DR20006 AGTGGTATGGGTGCAATTT---AAGGACTACTTTTACT---C-----G
DR25006 GGAAGGAAG--GGGGAATCGGGTTAC---CCTGGTGGTCCAAATGTGAA
DR20006 GGTGGA-----TATC-----CTATTGCAGAC--ATGTGTT
DR21006 ---TGGTGGCTGGTATT-----TGGACCCGAATTACTGG-CTCTCTGTG
DR09006 ---GAAATACTGGA---CTTGTGTTGC---CCAGA-CTGG-----AG
DR17006 AAGTATATATTATAATATC-----CTCCNNAT-----
DR13006 AAGAAGGATCGTA-----CTAGATGGGC-----CGAGCCT-----
DR02006 -----CTCCCTAA--TGGACCGAATTTC

DR01006 --TGGGTANTGCCTCTGGGGCCAGGGGACCCAGGTCAACCGTCTCCTCACT
DR27006 --T-----CTGGGGCCAGGGGACCCAGGTCAACCGTCTCCTCACT
DR03006 AAAAAGTATAAGTACTGGGGCCAGGGGACCCAGGTCAACCGTCTCCTCACT
DR11006 TTTGCGTATAACTACTGGGGCCGGGGGACCCAGGTCAACCGTCTCCTCACT
DR24006 TTTGCGTATAACTACTGGGGCCGGGGGACCCAGGTCAACCGTCTCCTCACT
DR16006 CGA-CT-ATAACTATTGGGGCCAGGGGATCCAGGTCAACCGTCACTCACT
DR19006 CGC-CTTATAACGTCTGGGGTCAAGGGGACCCAGGTCAACCGTCTCCTCACT
DR07006 CTGGG-----GCCAGGGGACCCAGGTCAACCGTCTCCTCACT
DR16006 CTGGGATGATTTT-----GGCCAGGGGACCCAGGTCAACCGTCTCCTCACT
DR20006 CTGGG-----GCCAGGGGACCCAGGTCAACCGTCTCCTCACT
DR25006 GATGGTTATAACTATTGGGGCCAGGGGACCCAGGTCAACCGTCTCCTCAC-
DR20006 CGAGAT---ACG---GCGACCCGGGGACCCAGGTCAACCGTCTCCTCAC-
DR21006 GGTGCATATGCCATCTGGGGCCAGGGGACCCAGGTCAACCGTCTCCTCAC-
DR09006 GATACTTCGGACAG-TGGGGTCAAGGGGACCCAGGTCAACCGTCTCCTCACT
DR17006 --TGAGTATAAGTACTGGGGCCAGGGGACCCAGGTCAACCGTCTCCTCA--
DR13006 CGAGAATGGAACAACCTGGGGCCAGGGGACCCAGGTCAACCGTCTCCTCA--
DR02006 CCAACATGGG--TGCCGGGGCCAGGGGACCCAGGTCAACCGTCTCCT----

```

FIG. 7C

DR01006	AG----	TTACCCGTACGACGTTCCGGACTACGGTTCTTAATAGAATTC
DR27006	AG----	TTACCCGTACGAGCTTCCGGACTACGGTTCTTAATAGAATTC
DR03006	AGCTAGTT	ACCCGTACGACGTTCCGGACTACGGTTCTTAATAGAATTC
DR11006	AG----	TTACCCGTACGACGTTCCGGACTACGGTTCTTAATAGAATTC
DR24006	AGCTAGTT	ACCCGTACGACGTTCCGGACTACGGTTCTTAATAGAATTC
DR16006	----	AGTTACCCGTACGACGTTCCGGACTACGGTTCTTAATAGAATTC
DR19006	----	AGTTACCCGTACGACGTTCCGGACTACGGTTCTTAATAGAATTC
DR07006	----	AGTTACCCGTACGACGTTCCGGACTACGGTTCTTAATAGAATTC
DR16006	----	AGTTACCCGTACGACGTTCCGGACTACGGTTCTTAATAGAATTC
DR20006	----	AGTTACCCGTACGACGTTCCGGACTACGGTTCTTAATAGAATTC
DR25006	----	AGTTACCCGTACGACGTTCCGGACTACGGTTCTTAATAGAATTC
DR20006	----	AGTTACCCGTACGACGTTCCGGACTACGGTTCTTAATAGAATTC
DR21006	----	AGTTACCCGTACGACGTTCCGGACTACGGTTCTTAATAGAATTC
DR09006	AGCTAGTT	ACCCGTACGACGTTCCGGACTACGGTTCTTAATAGAATTC
DR17006	-----	-----
DR13006	-----	-----
DR02006	-----	-----

-----TA

FIG. 7D

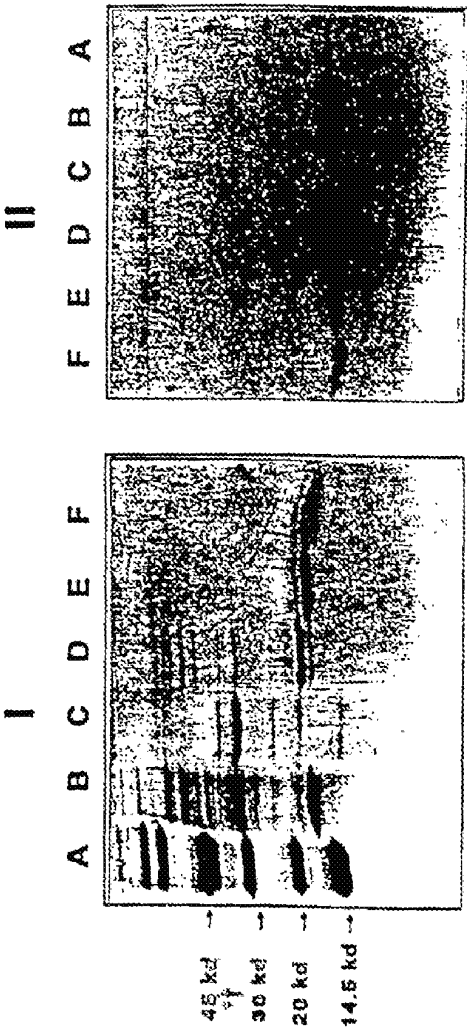


FIGURA 8A

FIGURA 8B