

(11) *Número de Publicação:* **PT 94066 B**

(51) *Classificação Internacional:* (Ed. 6)

C12N015/00 A

C12P019/34 B

(12) FASCÍCULO DE PATENTE DE INVENÇÃO

(22) <i>Data de depósito:</i> 1990.05.16 (30) <i>Prioridade:</i> 1989.05.16 US 352927 1989.05.17 US 352884 1989.05.17 US 353235 (43) <i>Data de publicação do pedido:</i> 1991.01.08 (45) <i>Data e BPI da concessão:</i> 10/96 1996.10.16	(73) <i>Titular(es):</i> SCRIPPS CLINIC AND RESEARCH FOUNDATION 10666 NORTH TORREY PINES ROAD, LA JOLLA CALIFÓRNIA 92037-1093 US (72) <i>Inventor(es):</i> WILLIAM D. HUSE US (74) <i>Mandatário(s):</i> ANTÓNIO JOÃO COIMBRA DA CUNHA FERREIRA RUA DAS FLORES 74 4/AND. 1294 LISBOA PT
--	--

(54) *Epígrafe:* PROCESSO PARA A PRODUÇÃO DE VECTORES DE ADN DE CADEIA DUPLA E DE UM SISTEMA DE CLONAÇÃO COMPREENDENDO OS REFERIDOS VECTORES

(57) *Resumo:*

VECTOR EXPRESSÃO; SISTEMA DE CLONAÇÃO; CLIVAGEM DE ADN

[Fig.]

MEMÓRIA DESCRITIVA
DA
PATENTE DE INVENÇÃO
Nº 94 066

NOME: SCRIPPS CLINIC AND RESEARCH FOUNDATION

EPÍGRAFE: "Processo para a produção de vectores de ADN de cadeia dupla e de um sistema de clonação compreendendo os referidos vectores"

INVENTORES: William D. Huse

Reivindicação do direito de prioridade (ao abrigo do artigo 4º da Convenção de Paris de 20 de Março de 1883): Estados Unidos da América em 16 de Maio de 1989 sob os nºs. 07/352,927, 17 de Maio de 1989 sob os nºs. 07/353,235 e 07/352,884, 20 de Setembro de 1989 sob os nºs. 07/353,241 e 07/410,888, 07/410,749 e 07/410,716, 21 de Setembro de 1989 sob o nº. 07/411,058, 4 de Dezembro de 1989 sob o nº. 07/446,333 e 20 de Março de 1990 sob o nº. 07/497,106



PATENTE Nº. 94 066

"Processo para a produção de vectores de ADN de cadeia dupla e de um sistema de clonação compreendendo os referidos vectores"

para que

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

R E S U M O

O presente invento refere-se a um processo para a produção de um vector de expressão de ADN de cadeia dupla, linear, contendo um promotor 5' terminal controlando a expressão de um primeiro e de um segundo polipéptidos, a partir de um primeiro e de um segundo genes, respectivos, sendo os referidos primeiro e segundo polipéptidos capazes de formar um receptor heterodimérico de especificidade predeterminada, caracterizado por compreender:

(a) o isolamento de uma primeira população diversa, de primeiros genes contendo um gene codificando o referido primeiro polipéptido, sendo a referida primeira população constituída por moléculas de ADN de cadeia dupla, linear, definindo um sítio de restrição reconhecido por uma endonuclease, estando o referido sítio de restrição localizado 5' terminal em relação aos referidos genes mas 3' terminal em relação a um promotor controlando a expressão dos referidos genes;

(b) o isolamento de uma segunda população diversa, de segundos genes contendo um gene codificando o referido segundo polipéptido, sendo a referida segunda população constituída por moléculas de ADN de cadeia dupla, linear, definindo o referido



-2-

sítio de restrição, estando o referido sítio de restrição localizado 3' terminal em relação aos referidos segundos genes, estando os referidos segundos genes localizados 3' terminais em relação a um promotor controlando a expressão dos referidos segundos genes;

(c) a clivagem das referidas moléculas de ADN de cadeia dupla, quando presentes nas referidas populações, com a referida endonuclease para produzir fragmentos de restrição possuindo um terminal coesivo e um dos referidos primeiros genes ou um dos referidos segundos genes; e

(d) a ligação aleatória dos referidos fragmentos de restrição uns aos outros, através do referido terminal coesivo de cada um, para produzir uma população diversa de vectores de expressão de ADN, linear, de cadeia dupla, possuindo um dos referidos primeiros genes e um dos referidos segundos genes, em tandem, na mesma cadeia de ADN e sob o controlo de um único promotor;

(e) o isolamento de um vector de expressão capaz de exprimir o referido receptor heterodimérico de especificidade predeterminada, a partir da referida população diversa de vectores de expressão de ADN de cadeia dupla.

O presente invento refere-se também a um processo de produção de um sistema de clonação compreendendo os referidos vectores.



MEMÓRIA DESCRITIVA

Campo da invenção

O presente invento refere-se a um processo para a produção de polímeros com actividade pré-seleccionada.

Antecedentes

Os fenómenos de ligação entre ligandos e receptores desempenham muitos papeis cruciais em sistemas biológicos. Exemplos destes fenómenos são, a ligação de moléculas de oxigénio à desoxi-hemoglobina, para formar a oxi-hemoglobina, e a ligação de um substrato a uma enzima, que actua no referido substrato, tal como a ligação entre uma proteína e uma protease, tal como a tripsina. Exemplos adicionais de fenómenos de ligação biológicos incluem, ainda, a ligação de um antigénio a um anticorpo e a ligação do componente de complemento C3 ao, assim chamado, receptor CR1.

Crê-se também que muitas drogas e outros agentes terapêuticos estão também dependentes de fenómenos de ligação. Por exemplo, referiu-se que os opiáceos, tais como a morfina, se ligam a receptores específicos no cérebro. Referiu-se que os agonistas e antagonistas de opiáceos competem com drogas, tais como a morfina, por esses sítios de ligação.

Os ligandos, como drogas produzidas pelo homem, tal como a morfina e os seus derivados, e os que estão naturalmente presentes em sistemas biológicos, tais como as endorfinas e as hormonas, ligam-se a receptores que estão, naturalmente, presentes em sistemas biológicos e serão aqui tratados em conjunto. Esta ligação pode conduzir a um número de fenómenos biológicos, incluindo, particularmente, a hidrólise de ligações amida e éster, quando as proteínas são hidrolisadas nos polipeptídeos constituintes por uma enzima, tal como a tripsina ou a papaína, ou quando uma gordura é clivada em glicerina e em três ácidos carboxílicos, respectivamente. Adicionalmente, esta ligação pode conduzir à formação de ligações amida e éster na formação de proteínas e de gorduras, bem como à formação de ligações



carbono-carbono e carbono-azoto.

Um exemplo de um sistema produtor de receptores em vertebrados é o sistema imunológico. O sistema imunológico de um mamífero é um dos sistemas biológicos mais versáteis onde, se podem produzir, provavelmente, mais do que $1,0 \times 10^7$ especificidades de receptor, na forma de anticorpos. De facto, grande parte da pesquisa biológica e médica, contemporânea, é dirigida no sentido de identificar este reportório. Durante a última década, tem-se verificado um aumento dramático na capacidade de aproveitar os efeitos de um vasto reportório imunológico. O desenvolvimento da metodologia do hibridoma, por Kohler e Milstein, tem tornado possível produzir anticorpos monoclonais, i.e., uma composição de moléculas de anticorpos de especificidade única a partir do reportório de anticorpos induzidos durante uma imuno-resposta.

Infelizmente, os métodos actuais para gerar anticorpos monoclonais não são capazes de pesquisar eficientemente toda a resposta de anticorpos induzida por um imunogénio particular. Num animal individual existem, pelo menos, 5-10000 clones de células B, diferentes, capazes de gerar anticorpos únicos, em relação a pequenos imunogénios relativamente rígidos, tais como, por exemplo, o dinitrofenol. Adicionalmente, devido ao processo de mutação somática, que se verifica durante a geração da diversidade de anticorpos, pode ser gerado um número essencialmente ilimitado de moléculas de anticorpos únicos. Em contraste com este vasto potencial para anticorpos diferentes, com as actuais metodologias de hibridoma obtêm-se apenas algumas centenas de anticorpos monoclonais, diferentes, por fusão.

Outras dificuldades na produção de anticorpos monoclonais, com a metodologia do hibridoma, incluem a instabilidade genética e a baixa capacidade de produção de culturas de hibridoma. Um dos meios pelo qual a arte tem tentado ultrapassar estes dois últimos problemas, tem sido clonar os genes produtores de imunoglobulina, de um hibridoma particular de interesse, num sistema de expressão procariótico. Ver, por exemplo, Robinson et al., Publicação PCT



nº. WO 89/0099; Winter et al., Patente Europeia nº. 0239400; Reading, Patente dos E.U.A. nº. 4 714 681; e Cabilly et al., Patente Europeia nº. 0125023.

Verificou-se, recentemente, que o reportório imunológico dos vertebrados contém genes que codificam imunoglobulinas com actividade catalítica. Tramontano et al., Sci., 234:1566-1570 (1986); Pollack et al., Sci., 234:1570-1573 (1986); Janda et al., Sci., 231-1188-1191 (1988); e Janda et al., Sci., 244:437-440 (1989). A presença de, ou a capacidade de induzir o reportório a produzir, moléculas de anticorpos capazes de catalisar uma reacção química, i.e., de actuarem como enzimas, tinha sido, previamente, postulada, quase há 20 anos, por W. P. Jencks em Catalysis in Chemistry and Enzymology, McGraw-Hill, N.Y. (1969).

Crê-se que, uma das razões pela qual a arte não conseguiu isolar, mais cedo, anticorpos catalíticos do reportório imunológico, e o facto de não conseguir isolar, até à data, muitos destes anticorpos, mesmo após as descobertas actuais, resulta da incapacidade de testar uma grande porção do reportório quando à actividade desejada. Crê-se que outra razão é a tendência das técnicas de ensaio actualmente disponíveis, tais como a técnica de hibridoma, em relação à produção de anticorpos de afinidade elevada concebidos para participação no processo de neutralização, em oposição à catálise.

Breve sumário da Invenção

O presente invento proporciona um novo método para o ensaio de uma maior porção, de um reportório de genes conservados, codificando receptores, quanto à presença de receptores com uma actividade pré-seleccionada, do que tem, até agora, sido possível, ultrapassando deste modo os inconvenientes anteriormente mencionados da técnica de hibridoma.

Numa concretização sintetiza-se uma biblioteca de genes conservados, codificando receptores contendo uma porção substancial do reportório de genes conservados, codificando receptores. Em concretizações preferidas, a biblioteca de genes conservados, codificando receptores, contém, pelo menos, cerca



de 10^3 , preferivelmente, pelo menos, cerca de 10^4 e, mais preferivelmente, pelo menos cerca de 10^5 genes, codificando receptores, diferentes.

A biblioteca de genes pode ser sintetizada por qualquer um de dois métodos, dependendo do material de partida.

Quando o material de partida é uma pluralidade de genes codificando receptores, o reportório é sujeito a duas reacções de extensão com iniciador, distintas. A primeira reacção de extensão com iniciador usa um primeiro iniciador de síntese de polinucleótido capaz de iniciar a primeira reacção, por hibridação com uma sequência de nucleótidos conservada (partilhada por uma pluralidade de genes) no reportório. A extensão do primeiro iniciador produz complementos de homólogos, codificando receptores, conservados, diferentes (cadeias de ácido nucleico, complementar dos genes no reportório).

A segunda reacção de extensão com iniciador produz, usando os complementos como modelo, uma pluralidade de homólogos de ADN conservados, codificando receptores, diferentes. A segunda reacção de extensão com iniciador usa um segundo iniciador de síntese de polinucleótido que é capaz de iniciar a segunda reacção por hibridação com uma sequência de nucleótidos conservada entre uma pluralidade de complementos.

Quando o material de partida é uma pluralidade de complementos dos genes conservados, codificando receptores, sujeita-se o reportório à segunda reacção de extensão com iniciador, descrita acima. Claro que, se estão presentes tanto um reportório de genes conservados codificando receptores, como os seus complementos, podem-se usar ambas as aproximações, em combinação.

Um homólogo de ADN conservado, codificando um receptor, i.e., um gene codificando um receptor capaz de se ligar a um ligando pré-seleccionado, é depois segregado da biblioteca, para produzir o gene isolado. Tipicamente, isto é realizado ligando operativamente, para expressão, uma pluralidade de homólogos de ADN conservados, diferentes, codificando um receptor, da



biblioteca, a um vector de expressão. Os vectores de expressão de receptor assim produzidos, são introduzidos numa população de células hospedeiras compatíveis, i.e., células capazes de exprimir um gene ligado, operativamente, ao vector para expressão. Os transformantes são cultivados sob condições adequadas para a expressão do receptor codificado pelo homólogo de ADN codificando o receptor. Os transformantes são clonados e testados quando à expressão de um receptor que se liga a um ligando pré-seleccionado. Podem-se usar quaisquer dos métodos adequados, conhecidos na arte, para detectar a ligação de um ligando a um receptor. Um transformante exprimindo a actividade desejada é, depois, segregado da população para produzir o gene isolado.

Noutra concretização, o presente invento contempla uma biblioteca de genes compreendendo uma mistura isolada de, pelo menos, cerca de 10^3 , preferivelmente, pelo menos, cerca de 10^4 e, mais preferivelmente, pelo menos 10^5 homólogos de ADN conservados, codificando receptores, uma pluralidade dos quais partilha um determinante antigénico conservado. Preferivelmente, os homólogos estão presentes num meio adequado para manipulação *in vitro*, tal como água; solução salina tamponada com fosfato e semelhantes, que mantêm a actividade biológica dos homólogos.

É também contemplado um receptor com uma actividade pré-seleccionada, preferivelmente, actividade catalítica, produzido por um processo do presente invento, preferivelmente, um monómero ou um dímero, tal como aqui se descreve.

Breve descrição dos Desenhos

Nos desenhos, que constituem parte da presente descrição:

Figura 1 Ilustra um diagrama esquemático da molécula de imunoglobulina mostrando as principais características estruturais. A área, dentro dum círculo, na cadeia pesada representa a região variável (V_H), um polipéptido contendo uma porção biologicamente activa (ligação a ligando) dessa região e um gene codificando esse polipéptido, produzidos pelos processos do presente invento.



Figura 2A É um esquema em diagrama de uma cadeia H da IgG humana (subclasse IgG1). A numeração é do terminal N, à esquerda, para o terminal C, à direita. É de notar a presença de quatro domínios, contendo, cada um, uma ligação dissulfureto (S-S) intracadeia, englobando aproximadamente 60 resíduos aminoácidos. O símbolo CHO representa carbo-hidrato. A região V da cadeia pesada (H) (V_H) assemelha-se à região V_L por possuir três CDR hipervariáveis (não mostrados).

Figura 2B É um esquema em diagrama de uma cadeia K humana (Painel 1). A numeração é do terminal N, à esquerda, para o terminal C, à direita. É de notar a ligação dissulfureto (S-S) intracadeia, englobando, aproximadamente, o mesmo número de resíduos aminoácidos nos domínios V_L e C_L . O Painel 2 mostra as localizações das regiões determinantes da complementaridade (CDR) no domínio V_L . Os segmentos fora das CDR são os segmentos da estrutura (framework).

Figura 3 Ilustra as sequências de aminoácidos das regiões V_H de 19 anticorpos monoclonais de ratinho, com especificidade para a fosforilcolina. A designação HP indica que a proteína é o produto de um hibridoma. O restante são proteínas de mieloma. (De Gearhart et al., Nature, 291:29, 1981).

Figura 4 Ilustra os resultados obtidos da amplificação, por PCR, de ARNm, obtido a partir do baço de um ratinho, imunizado com FITC. As faixas R17-R24 correspondem a reacções de amplificação com iniciadores 5' únicos (2-9, Tabela 1) e com o iniciador 3' (12, Tabela 1), R16 representa a reacção, por PCR, com o iniciador 5' contendo inosina (10, Tabela 1) e com o iniciador 3' (12, Tabela 1). Z e R9 são os controlos de amplificação; o controlo Z envolve a amplificação de V_H a partir de um plasmídeo (PLR2) e R9 representa a amplificação das regiões constantes de ARNm de baço, usando iniciadores 11 e 13 (Tabela 1)

Figura 5 As sequências de nucleótidos são clones da biblioteca de ADNc das regiões V_H , amplificadas por PCR, em Lambda ZAP. Estão aqui indicadas 110 bases N-terminais e os nucleótidos sublinhados representam a CDR1 (região determinante

de complementaridade). As sequências L03, L35, L47 e L48 não puderam ser classificadas em quaisquer subgrupos pré-definidos.

Figura 6 Mostra-se a sequência da inserção de ADN sintética inserida no Lambda ZAP para produzir os vectores de expressão Lambda Zap II V_H (Painel A) e Lambda Zap V_L (Painel B). As várias características para este vector exprimir os homólogos de ADN codificando V_H e V_L , incluem, o sítio de ligação de ribossoma Shine-Dalgarno, uma sequência líder para dirigir a proteína expressa para o periplasma, tal como é descrito por Mouva et al., J. Biol. Chem., 255:27, 1980, e vários sítios de enzimas de restrição usados para ligar operativamente os homólogos V_H e V_L ao vector de expressão. A sequência do vector de expressão de V_H contém, também, uma sequência curta de ácido nucleico que codifica aminoácidos que se encontram, tipicamente, na cadeia pesada das regiões variáveis (esqueleto de V_H). Este esqueleto de V_H está, imediatamente, a montante e na cadeia de leitura correcta, tal como os homólogos de ADN de V_H que estão ligados operativamente nos sítios Xho I e Spe I. Os homólogos de ADN de V_L estão ligados operativamente na sequência V_L (Painel B) nos sítios das enzimas de restrição Nco I e Spe I e, assim, a região de esqueleto de V_H é deleccionada quando os homólogos de ADN de V_L são ligados operativamente no vector de V_L .

Figura 7 Mostram-se as características principais do vector de expressão bacteriana Lambda Zap II V_H (vector de expressão de V_H). Mostra-se, no topo, a sequência de ADN sintética da Figura 6, conjuntamente com o promotor de polimerase T_3 do Lambda Zap II. Mostra-se a orientação da inserção no Lambda Zap II. Os homólogos de ADN de V_H estão inseridos nos sítios das enzimas de restrição Xho I e Spe I. O ADN de V_H está inserido nos sítios Xho I e Spe I e a leitura por transcrição produz o decapeptido epitopo (tag) que está localizado, imediatamente, 3' em relação aos sítios de clonação.

Figura 3 Mostram-se as características principais do vector de expressão bacteriana Lambda Zap II V_L (vector de expressão de V_L). Mostra-se, no topo, a sequência sintética da Figura 6,

conjuntamente com o promotor de polimerase T₃ do Lambda Zap II. Mostra-se a orientação da inserção no Lambda Zap II. Os homólogos de ADN de V_L estão inseridos no fagomideo que é produzido pelo protocolo de excisão in vivo descrito por Short et al., Nucleic Acids Res., 16:7583-7600, 1988. Os homólogos de ADN de V_L estão inseridos nos sítios de clonação Nco I e Spe I do fagomideo.

Figura 9 Representa-se um vector de expressão bacteriana modificado Lambda Zap II V_L II. Este vector é construído inserindo esta sequência de ADN sintética,

TGAATTCTAAACTAGTCGCCAAGGAGACAGTCATAATGAA
TCGAACTTAAGATTTGATCAGCGGTTCTCTGTCAGTATTACTT
ATACCTATTGCCTACGGCAGCCGCTGGATTGTTATTACTCGCTG
TATGGATAACGGATGCCGTCGGCGACCTAACAATAATGAGCGAC
CCCAACCAGCCATGGCCGAGCTCGTCAGTTCTAGAGTTAAGCGGCCG
GGGTTGGTCGGTACCGGCTCGAGCAGTCAAGATCTCAATTCGCCGGCAGCT

no Lambda Zap II que tinha sido digerido com as enzimas de restrição Sac I e Xho I. Esta sequência contém a sequência Shine-Dalgarno (sítio de ligação de ribossoma), a sequência líder para dirigir a proteína expressa para o periplasma e a sequência de ácido nucleico apropriada para permitir que os homólogos de ADN de V_L se liguem operativamente nos sítios das enzimas de restrição Sac I e Xba I proporcionados por este vector.

Figura 10 Representa-se a sequência do segmento de ADN sintético inserido no Lambda Zap II para produzir o vector de expressão lambda V_LII. Mostram-se as várias características e os vários sítios de reconhecimento de endonucleases de restrição.

Figura 11 Mostram-se os vectores para a expressão de V_H e de V_L, separadamente ou em combinação. Mostram-se os vários componentes essenciais destes vectores. O vector de cadeia leve, ou vector de expressão de V_L, pode ser combinado com o vector de expressão de V_H para produzir um vector de combinação contendo o V_H e o V_L, ligados, operativamente, para expressão, ao mesmo promotor.

Figura 12 Mostram-se as proteínas marcadas, imunoprecipitadas a partir de E. coli, contendo um homólogo de ADN de V_H e de V_L . Na faixa 1, mostram-se as proteínas residuais, imunoprecipitadas a partir de E. coli, que não contêm um homólogo de ADN de V_H ou de V_L . A faixa 2 contém a proteína V_H imunoprecipitada a partir de E. coli contendo apenas um homólogo de ADN de V_H . Nas faixas 3 e 4, mostra-se a co-migração de uma proteína V_H e de uma proteína V_L , imunoprecipitadas a partir de E. coli, contendo os homólogos de ADN de V_H e de V_L . Na faixa 5 demonstra-se a presença da proteína V_H e da proteína V_L , expressas a partir dos homólogos de ADN de V_H e de V_L , pelas duas espécies de proteína distinguíveis. A faixa 5 contém as proteínas residuais imunoprecipitadas por anticorpos anti-E. coli presentes no fluido ascítico do ratinho.

Figura 13 Mostram-se as fórmulas do análogo do estado de transição (fórmula 1) que induz os anticorpos para hidrólise do substrato de carboxamida (fórmula 2). O composto de fórmula 1, contendo um espaçador de glutarilo e um apêndice ligante de N-hidroxissuccinimida, é a forma usada para acoplar o hapteno (fórmula 1) aos transportadores de proteína KLH e BSA, enquanto que o composto de fórmula 3 é o inibidor. A funcionalidade de fosfonamidato é uma mimetização das características estereoeletrônicas do estado de transição da hidrólise da ligação amida.

Figura 14 Ilustra-se a amplificação, por PCR, das regiões Fd e kappa de ARNm de baço de um ratinho imunizado com NPN. A amplificação foi realizada como se descreve no Exemplo 17, usando híbridos de ADNc/ARN obtidos por transcrição inversa do ARNm com iniciador específico para amplificação das sequências de cadeia leve (Tabela 2) ou de cadeia pesada (Tabela 1). As faixas F1-F8 representam o produto de reacções de amplificação da cadeia pesada com cada um dos oito iniciadores 5' (iniciadores 2-9, Tabela 1) e com o iniciador 3' único (iniciador 15, Tabela 2). Nas faixas F9-F13 mostram-se as amplificações da cadeia leve (k) com os iniciadores 5' (iniciadores 3-6, e 12, respectivamente, Tabela 2) e com o iniciador 3' apropriado (iniciador 13, Tabela 2). Em todas as faixas pode-se observar uma banda de 700 pb

indicando a amplificação bem sucedida das regiões Fd e k.

Figura 15 Ilustra-se a pesquisa das bibliotecas de fago para determinar a ligação de antigénio, de acordo com o Exemplo 17C. Testaram-se colheitas de placas, em duplicado, de bibliotecas de expressão de Fab (filtros A, B), da cadeia pesada (filtros E, F) e da cadeia leve (filtros G, H) contra BSA marcado com ^{125}I conjugado com NPN para uma densidade de, aproximadamente, 30 000 placas por prato. Os filtros C e D ilustram a pesquisa secundária, em duplicado, de um teste positivo de um filtro primário A (setas) tal como se discute no texto.

A pesquisa utilizou métodos, de colheita das placas, convencionais: As células Azuis XL1 infectadas com fago, foram incubadas em pratos de 150 mm, durante 4 h a 37°C , e induziu-se a expressão de proteína cobrindo os pratos com filtros de nitrocelulose embebidos em tiogalactosido de isopropilo 10 mM (IPTG) e incubando os pratos a 25°C durante 8 h. Obtiveram-se filtros em duplicado durante uma segunda incubação empregando as mesmas condições. Os filtros foram então bloqueados numa solução de 1% de BSA em PBS, durante 1 h antes da incubação, com agitação, a 25°C durante 1 h, com uma solução de BSA, marcada com ^{125}I , conjugada com NPN (2×10^6 cpm ml^{-1} ; concentração de BSA 0,1 M; aproximadamente 15 NPN por molécula de BSA) em 1% de BSA em PBS. O fundo foi reduzido por pré-centrifugação de uma solução de reserva de BSA radiomarcada, a $100\,000 \times g$ durante 15 minutos, e pré-incubação de soluções com colheitas de placas de pratos contendo bactérias injectadas com um fago sem inserção. Após a marcação, lavaram-se os filtros, repetidamente com PBS/0,05% de Tween 20 antes do desenvolvimento de autorradiografias durante a noite.

Figura 16 Ilustra-se a especificidade da ligação de antigénio, tal como se mostra por inibição competitiva, de acordo com o Exemplo 17C. As colheitas dos filtros de placas positivas foram expostas a ^{125}I -BSA-NPN na presença de concentrações crescentes do NPN inibidor.

Neste estudo mancharam-se um número de fagos correlacionados



com a ligação de NPN, como na figura 15 (cerca de 100 partículas por mancha) directamente nos tapetes bacterianos. O prato foi depois coberto com um filtro embebido em IPTG e incubado durante 19 h a 25°C. O filtro foi depois bloqueado em 1% de BSA em PBS, antes da incubação em ^{125}I -BSA-NPN, como se descreveu anteriormente na Figura 15, com a excepção da inclusão de quantidades variáveis de NPN na solução de marcação. As outras condições e procedimentos foram como na Figura 15. Na figura mostram-se os resultados, em duplicado, para um fago de afinidade moderada. Para quatro outros fagos obtiveram-se resultados similares, com algumas diferenças nas gamas de concentração de inibidor eficazes.

Figura 17 Ilustra-se a caracterização de uma proteína de ligação de antígeno de acordo com o Exemplo 18D. O sobrenadante bacteriano, purificado parcialmente, concentrado de um clone, ligando NPN, foi separado por filtração em gel e aplicaram-se aliquotas de cada fracção a placas de microtitulo revestidas com BSA-NPN. A adição, quer dos anticorpos anti-decapeptido (---) quer dos anticorpos anti-cadeia kappa (—), conjugados com fosfatase alcalina, foi seguida pelo desenvolvimento de cor. A seta indica a posição de eluição de um fragmento Fab conhecido. Os resultados mostram que a ligação de antígeno é uma propriedade de uma proteína de 50 kD contendo ambas as cadeias pesada e leve.

Recolheram-se placas simples de dois clones NPN-positivos (Figura 15) e excisou-se o plasmideo contendo as inserções de cadeia leve e de cadeia pesada. Inocularam-se culturas de 500 ml em caldo L, com 3 ml de uma cultura saturada contendo os plasmideos excisados, e incubaram-se durante 4 h a 37°C. A síntese de proteínas foi induzida pela adição de IPTG até uma concentração final de 1 mM e as culturas foram incubadas durante 10 h a 25°C. Concentraram-se 200 ml de sobrenadante de células até 2 ml e aplicaram-se a uma coluna TSK-G4000. Testaram-se aliquotas de 50 µl das fracções eluídas por ELISA.

Para a análise ELISA, revestiram-se placas de microtitulo



com BSA-NPN, com 1 µg/ml, adicionaram-se amostras de 50 µl misturadas com 50 µl de PBS-Tween 20 (0,05%)-BSA (0,1 %) e incubaram-se as placas durante 2 h a 25°C. Após lavagem com PBS-Tween 20-BSA, adicionaram-se 50 µl de concentrações apropriadas de um anticorpo de ratinho anti-decapeptido (20) e um anticorpo de cabra anti-cadeia leve kappa de ratinho (Southern Biotech), conjugado com fosfatase alcalina e incubou-se durante 2 h a 25°C. Após lavagem adicionaram-se 50 µl de fosfato de p-nitrofenilo (1 mg/ml em tris 0,1M, pH 9,5 contendo MgCl₂ 50 mM) e incubaram-se as placas durante 15-30 minutos antes da leitura do valor de DO a 405 nm.

Descrição detalhada da Invenção

A. DEFINIÇÕES

Nucleótido: é uma unidade monomérica de ADN e de ARN consistindo de uma unidade de açúcar (pentose), um fosfato e uma base heterocíclica azotada. A base está ligada à unidade açúcar através do carbono glicosídico (carbono 1' da pentose) e essa combinação de base e de açúcar é um nucleósido. Quando o nucleósido contém um grupo fosfato ligado à posição 3' ou 5' da pentose, é designado por nucleótido.

Par de bases (pb): é uma associação de adenina (A) com timina (T), ou de citosina (C) com guanina (G) numa molécula de ADN de cadeia dupla. No ARN a timina é substituída por uracilo (U).

Ácido Nucleico: é um polímero de nucleótidos, quer de cadeia simples, quer de cadeia dupla.

Gene: é um ácido nucleico cuja sequência de nucleótidos codifica um ARN ou um polipeptido. Um gene pode ser, quer ARN quer ADN.

Bases complementares: são nucleótidos que, normalmente, se emparelham quando o ADN ou o ARN adota uma configuração de cadeia dupla.

Sequência de nucleótidos complementar: é uma sequência



de nucleótidos numa molécula de ADN ou de ARN de cadeia simples que é suficientemente complementar com a de outra cadeia simples para se hibridar especificamente com essa outra cadeia com consequente ligação de hidrogénio.

Conservada: uma sequência de nucleótidos é conservada em relação a uma sequência (de referência) pré-seleccionada se se hibridar não aleatoriamente com um complemento exacto da sequência pré-seleccionada.

Hibridação: é o emparelhamento de sequências de nucleótidos (cadeias de ácido nucleico), substancialmente complementares, para formar um duplex ou um heteroduplex pelo estabelecimento de ligações de hidrogénio entre pares de bases complementares. É uma interacção específica, i.e., não aleatória, entre dois polinucleótidos complementares que pode ser competitivamente inibida.

Análogo de Nucleótido: é um nucleótido de purina ou de pirimidina que difere estruturalmente de A, T, G, C ou U, mas que é suficientemente similar para substituir o nucleótido normal numa molécula de ácido nucleico.

Homólogo de ADN: é um ácido nucleico possuindo uma sequência de nucleótidos conservada, pré-seleccionada, e uma sequência codificando um receptor capaz de se ligar a um ligando pré-seleccionado.

Anticorpo: o termo anticorpo, nas suas várias formas gramaticais, é aqui usado para designar moléculas de imunoglobulina e porções imunologicamente activas de moléculas de imunoglobulina, i.e., moléculas que contêm um sítio de combinação de anticorpos ou paratopo. Exemplos de moléculas de anticorpos são as moléculas de imunoglobulina intactas, moléculas de imunoglobulina substancialmente intactas e porções de uma molécula de imunoglobulina, incluindo as porções conhecidas na arte por Fab, Fab', F(ab')₂ e F(v).

Sítio de combinação de anticorpos: um sítio de combinação de anticorpos é a porção estrutural de uma molécula de



anticorpo, constituída por uma região variável de cadeia pesada e de cadeia leve e por uma região hipervariável, que se liga especificamente (imunorreage com) um antigénio. O termo imunorreagir, nas suas várias formas, significa a ligação específica entre uma molécula contendo um determinante antigénico e uma molécula contendo um sitio de combinação de anticorpos, tal como uma molécula de anticorpo inteira ou uma sua porção.

Anticorpo monoclonal: a expressão anticorpo monoclonal, nas suas várias formas gramaticais, refere-se a uma população de moléculas de anticorpo que contém apenas uma espécie de sitio de combinação de anticorpos, capaz de imunorreagir com um antigénio particular. Um anticorpo monoclonal apresenta assim, tipicamente, uma afinidade de ligação única para qualquer antigénio com o qual imunorreage. Deste modo, um anticorpo monoclonal pode conter uma molécula de anticorpo com uma pluralidade de sitios de combinação de anticorpos, cada um imuno específico de um antigénio diferente, p.e., um anticorpo monoclonal biespecífico.

B. PROCESSOS

O presente invento contempla um processo de isolamento, de entre um reportório de genes conservados, de um gene codificando um receptor possuindo uma actividade pré-seleccionada, preferivelmente, uma actividade catalítica. O receptor pode ser um polipéptido, uma molécula de ARN, tal como um ARN de transferência, um ARN apresentando actividade enzimática, e semelhantes. Preferivelmente, o receptor será um polipéptido capaz de se ligar a um ligando, tal como uma enzima, uma molécula de anticorpo ou uma sua porção imunologicamente activa, um receptor celular ou uma proteína de adesão celular, codificada por um dos membros de uma família de genes conservados, i.e., genes contendo uma sequência de nucleótidos conservada de pelo menos cerca de 10 nucleótidos de comprimento.

Exemplos de famílias de genes conservados são os que codificam as imunoglobulinas, os antigénios, do complexo de histocompatibilidade principal, da classe I ou II, receptores de

linfocitos, integrinas e semelhantes.

Um gene pode ser identificado como pertencendo a um reportório de genes conservados usando vários métodos. Por exemplo, pode-se usar um gene isolado como sonda de hibridação, sob condições pouco rigorosas, para detectar outros membros do reportório de genes conservados, presentes no ADN genómico, usando os métodos descritos por Southern, J. Mol. Biol., 98:503 (1975). Se o gene usado como sonda de hibridação se hibrida com múltiplos fragmentos de endonucleases de restrição esse gene é um membro de uma família de genes conservados.

Imunoglobulinas

As imunoglobulinas, ou moléculas de anticorpos são uma grande família de moléculas que inclui vários tipos de moléculas tais como IgD, IgG, IgA, IgM e IgE. A molécula de anticorpo é constituída, tipicamente, por duas cadeias, pesada (H) e leve (L), com uma região variável (V) e uma região constante (C) presentes em cada cadeia. Várias regiões diferentes de uma imunoglobulina contêm sequências conservadas úteis para isolamento de um reportório de imunoglobulina. Resultados extensivos de sequenciação de aminoácidos e de ácidos nucleicos apresentando exemplos de sequências conservadas, foram compilados para moléculas de imunoglobulinas por Kabat et al., em Sequences of Proteins of Immunological Interest, National Institutes of Health, Bethesda, MD, 1987.

A região C, da cadeia H, define o tipo de imunoglobulina particular. Deste modo, a selecção de sequências conservadas, tal como aqui se definem, da região C da cadeia H, resulta na preparação de um reportório de genes de imunoglobulina possuindo membros do tipo de imunoglobulina da região C seleccionada.

A região V das cadeias H ou L compreende, tipicamente, quatro regiões de estrutura (framework) (FR), cada uma contendo graus relativamente menores de variabilidade, que inclui comprimentos de sequências conservadas. O uso de sequências conservadas das regiões de estrutura (framework) FR1 e FR4 (região J) da cadeia V_H é um exemplo de uma concretização



preferida e é aqui descrito nos Exemplos. As regiões de estrutura (framework) são, tipicamente, conservadas através de vários ou de todos os tipos de imunoglobulina e, assim, as sequências conservadas aí contidas são particularmente adequadas para a preparação de reportórios com vários tipos de imunoglobulina.

Complexo de Histocompatibilidade Principal

O complexo de histocompatibilidade principal (MHC) é um grande locus genético que codifica uma extensa família de proteínas que inclui várias classes de moléculas designadas por moléculas MHC da classe I, da classe II ou da classe III. Paul et al., em Fundamental Immunology, Raven Press, NY, pp. 303-378 (1984).

As moléculas MHC da classe I são um grupo polimórfico de antígenos de transplantação representando uma família conservada na qual o antígeno é constituído por uma cadeia pesada e por uma cadeia leve não codificada no MHC. A cadeia pesada inclui várias regiões, designadas por regiões N, C1, C2, de membrana e citoplásmica. As sequências conservadas úteis no presente invento encontram-se, principalmente, nas regiões N, C1 e C2 e são identificadas como sequências contínuas de "resíduos invariantes" em Kabat et al., supra.

As moléculas MHC da classe II compreendem uma família conservada de antígenos polimórficos que participam na imunoresposta e que compreendem uma cadeia alfa e uma cadeia beta. Os genes que codificam a cadeia alfa e beta incluem, cada um, várias regiões que contêm sequências conservadas, adequadas para a produção dos reportórios das cadeias alfa e beta da classe II de MHC. Exemplos de sequências de nucleótidos conservadas incluem as que codificam os resíduos aminoácidos 26-30 da região A1, os resíduos 161-170 da região A2 e os resíduos 195-206 da região de membrana, todas da cadeia alfa. Estão também presentes sequências conservadas nas regiões B1, B2 e de membrana, da cadeia beta, nas sequências de nucleótidos que codificam os resíduos aminoácidos 41-45, 150-162 e 200-209, respectivamente.



Receptores de Linfócito e Antígenios de Superfície Celular

Os linfócitos contêm várias famílias de proteínas nas suas superfícies celulares, incluindo o receptor de células T, o antígeno Thy-1 e numerosos antígenios de superfície de células T, incluindo os antígenios definidos pelos anticorpos monoclonais OKT4 (leu3), OKUT5/8 (leu 2), OKUT3, OKUT1 (leu 1), OKT 11 (leu5), OKT6 e OKT9. Paul, supra nas págs. 458-479.

A designação receptor de células T é um termo usado para uma família de moléculas de ligação de antígenios que se encontram na superfície das células T. O receptor de células T, como família, exibe especificidade de ligação polimórfica similar à das imunoglobulinas, na sua diversidade. O receptor de células T maduras é constituído por cadeias alfa e beta, cada uma possuindo uma região variável V e uma região constante (C). As semelhanças que o receptor de células T têm em relação às imunoglobulinas em organização genética e em função, mostra que o receptor de células T contém regiões de sequência conservada. Lai et al., Nature, 331:543-546 (1988).

Exemplos de sequências conservadas incluem as que codificam os resíduos aminoácidos 84-90 da cadeia alfa, os resíduos aminoácidos 107-115 da cadeia beta, e os resíduos aminoácidos 91-95 e 111-116 da cadeia gama. Kabat et al., supra, p. 279.

Integrinas e Adesinas

As proteínas adesivas envolvidas na ligação de células são membros de uma grande família de proteínas relacionadas, designadas por integrinas. As integrinas são heterodímeros constituídos por uma subunidade beta e por uma subunidade alfa. Membros da família da integrina incluem o receptor de plaquetas de glicoproteínas da superfície celular GpIIb-IIIa, o receptor de vibronectina (VnR), o receptor de fibronectina (FnR) e os receptores de adesão de leucócitos LFA-1, Mac-1, Mo-1 e 60.3. Roushahti et al., Science, 238:491-497 (1987). Os dados das sequências de ácidos nucleicos e de proteínas demonstram que existem regiões de sequências conservadas nos membros destas famílias, particularmente entre a cadeia beta de GpIIb-IIIa, VnR



-20-

e FnR e entre a subunidade alfa de VnR, Mac-1, LFA-1, FnR e GpIIb-IIIa. Suzuki et al., Proc. Natl. Acad. Sci. USA, 83:8614-8618, 1986; Ginsberg et al., J. Biol. Chem., 262:5437-5440, 1987.

A discussão seguinte ilustra o processo do presente invento aplicado ao isolamento de um gene codificando um receptor, conservado, do reportório do gene da imunoglobulina. Esta discussão não deve ser tomada como limitativa, mas sim como aplicação ilustrativa dos princípios que podem ser usados para isolar um gene de qualquer família de genes conservados, codificando receptores relacionados funcionalmente.

Geralmente, o processo combina os elementos seguintes:

1. Isolamento de ácidos nucleicos contendo uma porção substancial do reportório imunológico.
2. Preparação de iniciadores de polinucleótido para a clonação de segmentos de polinucleótido contendo os genes das regiões V_H e V_L de imunoglobulina.
3. Preparação de uma biblioteca de genes contendo uma pluralidade de genes V_H e V_L diferentes, a partir do reportório.
4. Expressão dos polipeptídeos V_H e V_L num hospedeiro adequado, incluindo hospedeiros procarióticos e eucarióticos, quer separadamente, quer na mesma célula e quer no mesmo quer em vectores de expressão diferentes.
5. Pesquisa dos polipeptídeos expressos quanto a uma actividade pré-seleccionada, e a segregação de uma combinação de genes codificando V_H e V_L , identificada pelo processo de pesquisa.

Um anticorpo produzido pelo presente invento assume uma conformação possuindo um sitio de ligação específico, tal como é evidenciado pela sua capacidade para ser inibido competitivamente, de um ligando pré-seleccionado ou predeterminado tal como um antigénio, um substrato enzimático e similares. Numa concretização, um anticorpo do presente invento forma um sitio de



ligação com antigénio que se liga, especificamente, a um antigénio pré-seleccionado para formar um produto de imuno-reacção (complexo), possuindo uma ligação suficientemente forte entre o antigénio e o sítio de ligação para o produto de imuno-reacção a ser isolado. Tipicamente, o anticorpo possui uma afinidade ou uma avidéz que, geralmente, é superior a 10^5 M^{-1} , mais usualmente, superior a 10^6 e, preferivelmente superior a 10^8 M^{-1} .

Numa outra concretização, um anticorpo do presente invento liga-se a um substrato e catalisa a formação de um produto a partir do substrato. Embora a topologia do sítio de ligação de ligando de um anticorpo catalítico seja provavelmente, mais importante, para a sua actividade pré-seleccionada, do que a sua afinidade (constante de associação ou pK_a) pelo substrato, os presentes anticorpos catalíticos têm uma constante de associação para o substrato pré-seleccionado geralmente superior a 10^3 M^{-1} , mais usualmente, superior a 10^5 M^{-1} ou 10^6 M^{-1} e, preferivelmente, superior a 10^7 M^{-1} .

Preferivelmente, o anticorpo produzido pelo presente invento, é heterodimérico e, deste modo, é constituído, normalmente, por duas cadeias diferentes de polipeptido, que em conjunto assumem uma conformação possuindo uma afinidade de ligação ou uma constante de associação para o anticorpo pré-seleccionado que é diferente, preferivelmente, superior à afinidade ou à constante de associação de qualquer dos polipeptidos sozinho, i.e., os monómeros. Uma ou ambas as cadeias de polipeptido diferentes são derivadas da região variável das cadeias leve e pesada de uma imunoglobulina. Tipicamente, os polipeptidos compreendendo as regiões variáveis leve (V_L) e pesada (V_H) são utilizados conjuntamente para ligação do anticorpo pré-seleccionado.

Um V_H ou um V_L produzido pelo presente invento pode ser activo na forma monomérica bem como em formas multiméricas, quer homoméricas, quer heteroméricas, preferivelmente, heterodiméricas. Um polipeptido V_H e V_L de ligação a ligandos produzido pelo



presente invento pode ser vantajosamente combinado num heterodimero (molécula de anticorpo) para modular a actividade de qualquer um dos polipéptidos ou para produzir uma actividade única do heterodimero. Os polipéptidos de ligação a ligando, individuais, serão designados por V_H e V_L e o heterodimero será designado por molécula de anticorpo.

Contudo, será de notar que um polipéptido de ligação V_H pode conter, para além do V_H , substancialmente toda ou uma porção da região constante da cadeia pesada. Um polipéptido de ligação V_L pode conter, para além do V_L , substancialmente toda ou uma porção da região constante da cadeia leve. Um heterodimero constituído por um polipéptido de ligação V_H , contendo uma porção da região constante da cadeia pesada, e por um polipéptido de ligação V_L , contendo substancialmente toda a região constante da cadeia leve, é designado por fragmento Fab. A produção de Fab pode ser vantajosa em algumas situações, em virtude das sequências adicionais da região constante contidas num Fab, em comparação com um F_y , poderem estabilizar a interacção V_H e V_L . Esta estabilização pode fazer com que o Fab tenha uma maior afinidade pelo antigénio. Adicionalmente, o Fab é mais vulgarmente usado na arte e, assim, existem mais anticorpos disponíveis, comercialmente, que reconhecem especificamente um Fab.

Os polipéptidos individuais V_H e V_L terão, geralmente, menos do que 125 resíduos aminoácidos, mais usualmente, menos do que cerca de 120 resíduos aminoácidos, enquanto que terão, normalmente, mais do que 60 resíduos aminoácidos, usualmente, mais do que cerca de 95 resíduos aminoácidos, mais usualmente, mais do que cerca de 100 resíduos aminoácidos. Preferivelmente, o V_H terá entre cerca de 110 e cerca de 125 resíduos aminoácidos de comprimento, enquanto que o V_L terá entre cerca de 95 e cerca de 115 resíduos aminoácidos de comprimento.

As sequências de resíduos aminoácidos variarão, grandemente, dependendo do idiótipo particular envolvido. Usualmente existirão, pelo menos, duas cisteínas separadas por de cerca de 60 a 75 resíduos aminoácidos e ligadas por uma ligação



dissulfureto. Os polipéptidos produzidos pelo presente invento serão normalmente, cópias substanciais de idiotipos das regiões variáveis das cadeias pesada e/ou leve de imunoglobulinas mas, em algumas situações, um polipéptido pode conter mutações aleatórias nas sequências dos resíduos aminoácidos de modo a melhorar vantajosamente a actividade desejada.

Em algumas situações, é desejável proporcionar uma ligação cruzada covalente dos polipéptidos V_H e V_L , que pode ser realizada proporcionando resíduos cisteína no terminal carboxilo. O polipéptido será preparado, normalmente, sem as regiões constantes da imunoglobulina, mas contudo, pode-se incluir uma pequena porção da região J como resultado da selecção vantajosa dos iniciadores de síntese de ADN. A região D será normalmente, incluída na cópia do V_H .

Noutras situações é desejável proporcionar um ligante de péptido para ligar o V_L e o V_H , para formar uma proteína de ligação a antigénio, de cadeia simples, contendo um V_H e um V_L . Esta proteína de ligação a antigénio, de cadeia simples, seria sintetizada como uma única cadeia de proteína. Estas proteínas de ligação a antigénio, de cadeia simples, foram descritas por Bird et al., Science, 242:423-426 (1988). A concepção de regiões de ligantes de péptido adequadas é descrita por Robert Landner na Patente dos E.U.A. nº. 4 704 692.

Este ligante de péptido pode ser concebido como parte das sequências de ácido nucleico contidas no vector de expressão. As sequências de ácido nucleico codificando um ligando de péptido estariam entre os homólogos de ADN de V_H e V_L e os sítios de endonuclease de restrição usados para ligar operativamente os homólogos de ADN de V_H e V_L ao vector de expressão.

Este ligante de péptido pode também ser codificado por sequências de ácido nucleico que são parte dos iniciadores de polinucleótido, usados para preparar as várias bibliotecas de genes. A sequência de ácido nucleico codificando o ligante de péptido pode ser construída por ácidos nucleicos ligados a um dos iniciadores ou a sequência de ácidos nucleicos codificando o



ligante de polipeptido pode ser obtido a partir de sequências de ácido nucleico que estão ligadas a vários iniciadores de polipeptido usados para criar as bibliotecas de genes.

Tipicamente, a região do terminal C dos polipeptidos V_H e V_L terá uma maior variedade de sequências do que a do terminal N e, com base na presente estratégia, pode ser ainda modificada para permitir uma variação das cadeias V_H e V_L de ocorrência natural. Pode-se usar um polinucleótido sintético para variar um ou mais aminoácidos numa região hipervariável.

1. Isolamento do Reportório

Para preparar uma composição de ácidos nucleicos contendo uma porção substancial do reportório de genes imunológico, é necessária uma fonte de genes codificando os polipeptidos V_H e/ou V_L . Preferivelmente, a fonte será uma população heterogênea de células produtoras de anticorpos, i.e., linfócitos B (células B), preferivelmente, células B rearranjadas tais como as que se encontram na circulação ou no baço de um vertebrado. (As células B rearranjadas são aquelas nas quais ocorreu a translocação do gene da imunoglobulina, i.e., o rearranjo, tal como é evidenciado pela presença, na célula, de ARNm com os transcritos das regiões V, D e J do gene da imunoglobulina aí localizados, de um modo adjacente).

Em alguns casos é desejável fazer tender o reportório para uma actividade pré-seleccionada, por exemplo, usando, como fonte de ácido nucleico, células (células fonte) de vertebrados em qualquer um de vários estados de idade, saúde e resposta imunológica. Por exemplo, a imunização repetida de um animal saudável antes da colheita de células B rearranjadas, resulta na obtenção de um reportório enriquecido em material genético produzindo um polipeptido, que se liga a ligandos, de elevada afinidade. Inversamente, a colheita de células B rearranjadas de um animal saudável, cujo sistema imunológico não foi recentemente desafiado, resulta na produção de um reportório que não está vocacionado em relação à produção de polipeptidos V_H e/ou V_L de elevada afinidade.



Será de notar que quanto maior for a heterogeneidade genética da população de células a partir das quais se obtêm os ácidos nucleicos, maior será a diversidade do repertório imunológico que ficará disponível para pesquisa, de acordo com o método do presente invento. Assim, as células de diferentes indivíduos, particularmente, de indivíduos possuindo uma diferença de idade imunologicamente significativa, e as células de indivíduos de diferentes estirpes, raças ou espécies, podem ser vantajosamente combinadas para aumentar a heterogeneidade do repertório.

Assim, numa concretização preferida, as células de fonte são obtidas a partir de um vertebrado, preferivelmente, de um mamífero, que foi imunizado ou parcialmente imunizado com um ligando antigénico (antigénio) contra o qual se procura actividade, i.e., um ligando pré-seleccionado. A imunização pode ser realizada de um modo convencional. Pode-se ir analisando o título de anticorpos no animal, para determinar o estado de imunização desejado, estado este que corresponde à quantidade de enriquecimento ou de inclinação do repertório desejado. Os animais parcialmente imunizados recebem apenas, tipicamente, uma imunização e as células são colhidas dos animais logo após se detectar uma resposta. Os animais totalmente imunizados apresentam um título de pico, que é atingido com uma ou mais injeções repetidas do antigénio no mamífero hospedeiro, normalmente com intervalos de 2 a 3 semanas. Usualmente, três a cinco dias após o último desafio, o baço é removido e o repertório genético dos esplenocitos, cerca de 90% dos quais são células B rearranjadas, é isolado usando procedimentos convencionais. Ver Current Protocols in Molecular Biology, Ausubel et al., eds., John Wiley & Sons, NY. Os ácidos nucleicos codificando os polipéptidos V_H e V_L podem ser obtidos a partir de células produzindo IgA, IgD, IgE, IgG ou IgM, muito preferivelmente, de células produzindo IgM e IgG.

Os processos para a preparação de fragmentos de ADN genómico, a partir dos quais os genes das regiões variáveis de imunoglobulina podem ser clonados como uma população diversa, são



bem conhecidos na arte. Ver, por exemplo, Herrmann et al., Methods In Enzymol., 152:180-183, (1987); Frischauf, Methods In Enzymol., 152:183-190 (1987); Frischauf, Methods In Enzymol., 152:190-199 (1987); e DiLella et al., Methods In Enzymol., 152:199-212 (1987). (Os ensinamentos das referências aqui citadas são aqui incorporados como referência).

O reportório de genes desejado pode ser isolado, quer a partir do material genômico contendo o gene exprimindo a região variável quer a partir do ARN mensageiro (ARNm) que representa um transcrito da região variável. A dificuldade em usar o ADN genômico de outras células, que não os linfócitos B não rearranjados, está em justapor as sequências codificando a região variável, quando as sequências estão separadas por intrões. O(s) fragmento(s) de ARN contendo os exões apropriados tem(têm) de ser isolado(s), os intrões excisados e os exões depois reunidos, na ordem adequada e na orientação adequada. Para a maior parte isto será difícil, pelo que a técnica alternativa usando células B rearranjadas será o método preferido, porque as regiões dos genes de imunoglobulina C, D e J foram translocadas para se tornarem adjacentes, de modo que a sequência é contínua (isenta de intrões) para as regiões variáveis completas.

Quando se usa ARNm as células serão lisadas sob condições de inibição da ribonuclease. Numa concretização, o primeiro passo consiste em isolar o ARNm celular total, por hibridação com uma coluna de oligo-dT-celulose. A presença de ARNm codificando os polipéptidos de cadeia pesada e/ou leve pode, depois, ser avaliada por hibridação, com cadeias simples de ADN, dos genes apropriados. Convenientemente, as sequências codificando a porção constante do V_H e do V_L podem ser usadas como sondas de polinucleótido, sequências estas que podem ser obtidas a partir de fontes disponíveis. Ver, por exemplo, Earlye Hood, Genetic Engineering, Setlow and Hollaender, eds., Vol. 3, Plenum Publishing Corporation, New York, (1981), pages 157-188; e Kabat et al., Sequences of Immunological Interest, National Institutes of Health, Bethesda, MD, (1987). Em concretizações preferidas, a preparação contendo o ARNm celular total é, primeiramente,



enriquecida em ARNm codificando V_H e/ou V_L . Tipicamente, o enriquecimento é realizado sujeitando a preparação de ARNm total, ou o seu produto de ARNm parcialmente purificado, a uma reacção de extensão de iniciador empregando um iniciador de síntese de polinucleótido do presente invento.

2. Preparação de iniciadores de polinucleótido

O termo "polinucleótido", tal como é aqui usado em referência a iniciadores, sondas e fragmentos ou segmentos de ácido nucleico, a serem sintetizados por extensão do iniciador, refere-se a uma molécula constituída por dois ou mais desoxirribonucleótidos ou ribonucleótidos, preferivelmente mais do que 3. O seu tamanho exacto dependerá de muitos factores, que por sua vez dependem das condições finais de uso.

O termo "iniciador", tal como é aqui usado, refere-se a um polinucleótido, quer purificado a partir de um digerido de restrição de ácido nucleico quer produzido sinteticamente, que é capaz de actuar como ponto de início de síntese quando colocado sob condições nas quais se induz a síntese de um produto de extensão de iniciador, que é complementar a uma cadeia de ácido nucleico, i.e., na presença de nucleótidos e de um agente para polimerização, tal como ADN-polimerase, transcriptase inversa e similares, e a uma temperatura e a um pH adequados. Preferivelmente, o iniciador é de cadeia simples, para uma eficiência máxima mas, alternativamente, pode ser de cadeia dupla. Se for de cadeia dupla, o iniciador é tratado, em primeiro lugar, para separar as suas cadeias, antes de ser usado para preparar produtos de extensão. Preferivelmente, o iniciador é um polidesoxirribonucleótido. O iniciador tem de ser suficientemente longo para iniciar a síntese de produtos de extensão na presença dos agentes para a polimerização. Os comprimentos exactos dos iniciadores dependerão de muitos factores, incluindo a temperatura e a fonte de iniciador. Por exemplo, dependendo da complexidade da sequência desejada, o iniciador de polinucleótido contém, tipicamente, 15 a 25, ou mais, polinucleótidos, ainda que possa conter menos nucleótidos. As moléculas iniciadoras curtas

requerem, geralmente, temperaturas mais frias para formar complexos híbridos suficientemente estáveis com o molde.

Os iniciadores aqui usados são seleccionados para serem "substancialmente" complementares com as diferentes cadeias de cada sequência específica a ser sintetizada ou amplificada. Isto significa que o iniciador tem de ser suficientemente complementar para se hibridar não aleatoriamente com a sua cadeia molde respectiva. Deste modo, a sequência de iniciador pode não reflectir a sequência exacta do molde. Por exemplo, um fragmento de nucleótidos não complementares pode estar ligado à extremidade 5' do iniciador, sendo o restante da sequência de iniciador substancialmente complementar com a cadeia. Estes fragmentos não complementares codificam, tipicamente, um sítio de restrição de endonuclease. Alternativamente, bases não complementares ou sequências mais compridas podem ser intercaladas no iniciador, desde que a sequência de iniciador tenha complementaridade suficiente com a sequência da cadeia a ser sintetizada ou amplificada, para com ela se hibridar não aleatoriamente e formar, assim, um produto de extensão nas condições de síntese de polinucleótidos.

Os iniciadores de polinucleótidos podem ser preparados usando qualquer método adequado, tal como, por exemplo, os métodos do fosfotriéster ou do fosfodiéster; ver Narang et al., Meth. Enzymol., 68:90, (1979); U.S. Patent No. 4 356 270; e Brown et al., Meth. Enzymol., 68:109, (1979).

A escolha de uma sequência de nucleótidos de iniciador depende de factores tais como a distância, no ácido nucleico da região codificando o receptor desejado, o seu sítio de hibridação no ácido nucleico, em relação a qualquer segundo iniciador a ser usado, o número de genes no reportório com os quais se vai hibridar e similares.

Por exemplo, para produzir homólogos de ADN, codificando V_H , por extensão de iniciador, a sequência de nucleótidos de um iniciador é seleccionada para se hibridar com uma pluralidade de genes da cadeia pesada de imunoglobulina num sítio substancial-



mente adjacente a uma região de codificação de V_H de modo a obter uma sequência de nucleótidos codificando um polipeptído funcional (capaz de se ligar). Para se hibridar com uma pluralidade de cadeias de ácido nucleico codificando V_H diferentes, o iniciador tem de ser um complemento substancial de uma sequência de nucleótidos conservada entre as diferentes cadeias. Estes sítios incluem sequências de nucleótidos na região constante, qualquer das regiões de estrutura (framework) da região variável, preferivelmente, a terceira região de estrutura, a região líder, a região promotora, a região J e similares.

Os iniciadores do presente invento podem também conter uma sequência de promotor de ARN-polimerase, dependente de ADN, ou o seu complemento. Ver, por exemplo, Krieg et al., Nucleic Acids Research, 12:7057-70 (1984); Studier et al., 189:113-130 (1986); e Molecular Cloning: A Laboratory Manual, Second Edition, Maniatis et al., eds. Cold Spring Harbor, NY (1989).

Quando se usa um iniciador contendo um promotor de ARN-polimerase, dependente de ADN, o iniciador é hibridado com a cadeia de ADN a ser amplificada e a segunda cadeia de polinucleótido do promotor de ARN-polimerase, dependente de ADN, é completada usando um agente indutor, tal como a ADN-polimerase I de E. coli ou o fragmento Klenow de ADN-polimerase I de E. coli. O polinucleótido de ARN complementar é, então, produzido adicionando uma ARN-polimerase, dependente de ARN. O polinucleótido de partida de um polinucleótido de ARN é amplificado alternando-se a produção de um polinucleótido de ARN e de um polinucleótido de ADN.

Se se pretende produzir os homólogos de ADN, codificando V_H e codificando V_L , por amplificação de reacção em cadeia com polimerase (PCR), têm de se usar dois iniciadores para cada cadeia de codificação de ácido nucleico a ser amplificada. O primeiro iniciador torna-se parte da cadeia nonsense (menos ou complementar) e hibrida-se com uma sequência de nucleótidos conservada entre as cadeias V_H (mais) no reportório. Para produzir homólogos de ADN codificando V_H , os primeiros



iniciadores são, deste modo, escolhidos para se hibridarem com (i.e., serem complementares com) regiões conservadas na região J, na região CH1, na região de articulação, na região CH2, ou na região CH3 de genes de imunoglobulina e similares. Para produzir um homólogo de ADN codificando V_L , os primeiros iniciadores são escolhidos para se hibridarem com (i.e., serem complementares com) uma região conservada na região J ou na região constante de genes de cadeia leve de imunoglobulina e similares. Os segundos iniciadores tornam-se parte da cadeia (mais) de codificação e hibridam-se com uma sequência de nucleótidos conservada entre as cadeias menos. Para produzir os homólogos de ADN codificando V_H , os segundos iniciadores são, deste modo, escolhidos para se hibridarem com uma sequência de nucleótidos conservada na extremidade 5' do gene de imunoglobulina codificando V_H , tal como na área codificando a região líder ou a primeira região de estrutura (framework). Será de notar que, na amplificação de ambos os homólogos de ADN codificando V_H e V_L , a sequência de nucleótidos 5' conservada do segundo iniciador, pode ser complementar de uma sequência adicionada exogenamente usando a desoxinucleotidilo-terminal-transferase tal como é descrito por Loh et al., Sci. Vol 243:217-220 (1989). Um ou ambos os primeiros e segundos iniciadores podem conter uma sequência de nucleótidos definindo um sítio de reconhecimento de endonuclease. O sítio pode ser heterólogo do gene de imunoglobulina a ser amplificado e, tipicamente, aparece na extremidade 5' ou próximo da extremidade 5' do iniciador.

3. Preparação de uma biblioteca de genes

A estratégia usada para a clonação i.e., para reproduzir, substancialmente, os genes V_H e/ou V_L contidos no reportório isolado, dependerá, como é bem conhecido na arte, do tipo, complexidade e pureza dos ácidos nucleicos que constituem o reportório. Outros factores dependerão de se pretender ou não amplificar ou mutar os genes.

Numa estratégia, o objectivo é clonar os genes codificando V_H e/ou V_L de um reportório constituído por cadeias de



codificação de polinucleótido, tais como ARNm e/ou a "cadeia de sentido" (sense strand) de ADN genómico. Se o reportório está na forma de ADN genómico de cadeia dupla, usualmente, é em primeiro lugar desnaturado, tipicamente, por fusão em cadeias simples. O reportório é sujeito a uma primeira reacção de extensão primária por tratamento (contacto) do reportório com um primeiro iniciador de síntese de polinucleótido possuindo uma sequência de nucleótidos pré-seleccionada. O primeiro iniciador é capaz de iniciar a primeira reacção de extensão de iniciador, por hibridação com uma sequência de nucleótidos, preferivelmente, com pelo menos cerca de 10 nucleótidos de comprimento e, mais preferivelmente, com pelo menos cerca de 20 nucleótidos de comprimento, conservada no reportório. O primeiro iniciador é por vezes aqui designado por "iniciador de sentido" (sense primer) porque se hibrida com a cadeia de codificação ou de sentido de um ácido nucleico. Adicionalmente, o segundo iniciador é por vezes aqui designado como "iniciador anti-sentido" (anti-sense primer) porque se hibrida com uma cadeia, de não codificação ou anti-sentido de um ácido nucleico, i.e., uma cadeia complementar de uma cadeia de codificação.

A extensão do primeiro iniciador é realizada misturando primeiro iniciador, preferivelmente, uma sua quantidade predeterminada, com os ácidos nucleicos do reportório, preferivelmente, com uma sua quantidade predeterminada, para formar uma mistura reaccional de extensão do primeiro iniciador. A mistura é mantida sob condições de síntese de polinucleótidos, por um período de tempo, que é, tipicamente, predeterminado, suficiente para a formação de um produto de reacção de extensão do primeiro iniciador, produzindo-se assim uma pluralidade de complementos de homólogos de ADN codificando V_H diferentes. Os complementos são então sujeitos a uma segunda reacção de extensão de iniciador por tratamento com um segundo iniciador de síntese de polinucleótido possuindo uma sequência de nucleótidos pré-seleccionada. O segundo iniciador é capaz de iniciar a segunda reacção por hibridação com uma sequência de nucleótidos, preferivelmente com pelo menos cerca de 10 nucleotidos de



comprimento e, mais preferivelmente, com pelo menos cerca de 20 nucleótidos de comprimento, conservada entre uma pluralidade de diferentes complementos de genes codificando V_H , tais como os produzidos, por exemplo, pela primeira reacção de extensão de iniciador. Isto é realizado misturando o segundo iniciador, preferivelmente uma sua quantidade predeterminada, com os ácidos nucleicos complementares, preferivelmente, com uma sua quantidade predeterminada, para formar uma mistura reaccional de extensão do segundo iniciador. A mistura é mantida sob condições de síntese de polinucleótidos por um período de tempo, que é, tipicamente, predeterminado, suficiente para a formação de um produto da reacção de extensão do primeiro iniciador, produzindo-se assim uma biblioteca de genes contendo uma pluralidade de homólogos de ADN, codificando V_H e/ou V_L , diferentes.

Pode-se usar, em cada amplificação, uma pluralidade de primeiros iniciadores e/ou uma pluralidade de segundos iniciadores, ou um par individual de primeiro e de segundo iniciador. Em qualquer dos casos, os produtos de amplificação de amplificações usando as mesmas ou diferentes combinações de primeiro e de segundo iniciadores, podem ser combinados para aumentar a diversidade da biblioteca de genes.

Numa outra estratégia, o objectivo é clonar os genes codificando V_H e/ou V_L de um reportório proporcionando um complemento de polinucleótido do reportório, tal como a cadeia anti-sentido de ADN genómico ou o polinucleótido produzido sujeitando o ARNm a uma reacção de transcriptase inversa. Os métodos para produzir estes complementos são bem conhecidos na arte. O complemento é sujeito a uma reacção de extensão de iniciador, similar à reacção de extensão do segundo iniciador, descrita acima, i.e., uma reacção de extensão de iniciador usando um iniciador de síntese de polinucleótido capaz de se hibridar com uma sequência de nucleótidos conservada entre uma pluralidade de comprimentos de genes codificando V_H diferentes.

A reacção de extensão de iniciador é realizada usando qualquer método adequado. Geralmente, realiza-se numa solução



aquosa tamponada, preferivelmente, a um pH de 7-9, muito preferivelmente, de cerca de 8. Preferivelmente, mistura-se um excesso molar (para ácido nucleico genómico, usualmente, cerca de $10^6:1$, de iniciador:molde) do iniciador com o tampão, contendo a cadeia de molde. Prefere-se um grande excesso molar para melhorar a eficiência do processo.

Os trifosfatos de desoxirribonucleótido dATP, dCTP, dGTP e dTTP são, também, misturados com a mistura reaccional de extensão de iniciador (síntese de polinucleótidos) em quantidades adequadas e a solução resultante é aquecida a cerca de 90°C-100°C, durante cerca de 1 a 10 minutos, preferivelmente, de 1 a 4 minutos. Após este período de aquecimento, deixa-se a solução arrefecer até à temperatura ambiente, que é preferível para a hibridação do iniciador. A mistura arrefecida adiciona-se um agente apropriado, para induzir ou catalisar a reacção de extensão do iniciador, e deixa-se a reacção ocorrer sob condições conhecidas na arte. A reacção de síntese pode ocorrer a uma temperatura compreendida entre a temperatura ambiente e uma temperatura acima da qual o agente indutor já não funciona eficazmente. Assim, por exemplo, se se usa ADN-polimerase como agente indutor, a temperatura não é, em geral superior a cerca de 40°C.

O agente indutor pode ser qualquer composto ou sistema que funcionará para realizar a síntese de produtos de extensão de iniciador, incluindo enzimas. Enzimas adequadas para este propósito incluem, por exemplo, a ADN-polimerase de E. coli, o fragmento Klenow de ADN-polimerase I de E. coli, a ADN-polimerase T4, outras ADN-polimerases disponiveis, a transcriptase inversa, e outras enzimas, incluindo enzimas estáveis ao calor, que facilitarão a combinação dos nucleótidos de uma maneira adequada para formar os produtos de extensão de iniciador que são complementares com cada cadeia de ácido nucleico. Geralmente, a síntese será iniciada na extremidade 3' de cada iniciador e prosseguirá na direcção 5', ao longo da cadeia molde, até a síntese terminar, produzindo-se moléculas de comprimentos diferentes. Podem existir agentes indutores, contudo, que iniciem



a síntese na extremidade 5' e que prossigam na direcção anterior, usando o mesmo processo como se descreveu acima.

O agente indutor pode ser também um composto ou sistema, que funcionará para realizar a síntese de produtos de extensão de iniciadores de ARN, incluindo as enzimas. Em concretizações preferidas, o agente indutor pode ser uma ARN-polimerase dependente de ADN, tal como ARN-polimerase T7, ARN-polimerase T3 ou ARN-polimerase SP6. Estas ARN-polimerases iniciam a síntese de um promotor contido num iniciador do presente invento. Estas polimerases produzem um polinucleótido de ARN complementar. A elevada velocidade de regeneração ("turn-over") da ARN polimerase amplifica o polinucleótido de partida, tal como foi descrito por Chamberlin et al., The Enzymes, ed., P. Boyer, pp 87-108, Academic Press, NY (1982). Outra vantagem da ARN-polimerase T7 consiste em se poderem introduzir mutações na síntese de polinucleótido, substituindo uma porção do ADNc por um ou mais oligodesoxinucleótidos (polinucleótidos) e transcrevendo directamente o molde, parcialmente desajustado tal como anteriormente foi descrito por Joyce et al., Nucleic Acid Research, 17:711-722 (1989).

Se o agente indutor é uma ARN-polimerase dependente de ADN e, deste modo, incorpora trifosfatos de ribonucleótido, misturam-se quantidades suficientes de ATP, CTP, GTP e UTP com a mistura reacção de extensão de iniciador e a solução resultante é tratada como se descreveu acima.

A cadeia acabada de sintetizar e a sua cadeia de ácido nucleico complementar formam uma molécula de cadeia dupla que pode ser usada nos passos seguintes do processo.

Pode-se usar, vantajosamente, a reacção de extensão do primeiro e/ou do segundo iniciador para incorporar, no receptor, um epitopo pré-seleccionado, útil na detecção e/ou isolamento, imunológicos, de um receptor. Isto é realizado usando um primeiro e/ou segundo iniciador de síntese de polinucleótido ou um vector de expressão, para incorporar uma sequência de resíduos aminoácidos predeterminada na sequência de resíduos aminoácidos do receptor.



Após produzir os homólogos de ADN codificando V_H e/ou V_L , para uma pluralidade de genes codificando V_H e/ou V_L diferentes no reportório, tipicamente, os homólogos são amplificados. Enquanto que os homólogos de ADN codificando V_H e/ou V_L podem ser amplificados por técnicas clássicas, tais como a incorporação num vector de replicação autónoma, prefere-se amplificar em primeiro lugar os homólogos de ADN, sujeitando-os a uma reacção em cadeia com polimerase (PCR) antes de os inserir num vector. De facto, em estratégias preferidas as reacções de extensão do primeiro e/ou segundo iniciador, usadas para produzir a biblioteca de genes, são as reacções de extensão do primeiro e do segundo iniciador na reacção em cadeia com polimerase.

A PCR é realizada, tipicamente, por ciclos i.e., realizando, simultaneamente, numa mistura, as reacções de extensão do primeiro e do segundo iniciador acima descritas, cada ciclo compreendendo a síntese de polinucleótidos seguida da desnaturação dos polinucleótidos de cadeia dupla formados. Os métodos e os sistemas para amplificar um homólogo de ADN são descritos nas Patentes dos E.U.A. nº. 4 683 195 e nº. 4 683 202, ambas de Mullis et al..

Em concretizações preferidas, para cada reacção de amplificação, apenas se usa um par de primeiro e de segundo iniciadores. Combinam-se depois os produtos de reacção de amplificação obtidos de uma pluralidade de amplificações diferentes, cada uma, usando uma pluralidade de pares de iniciadores diferentes.

Contudo, o presente invento contempla também a produção de um homólogo de ADN por co-amplificação (usando dois pares de iniciadores) e amplificação multiplex (usando até cerca de 8, 9 ou 10 pares de iniciadores).

Os homólogos de ADN, codificando V_H e V_L , produzidos por amplificação de PCR estão, tipicamente, na forma de cadeia dupla e possuem, numa posição contigua ou adjacente a cada um dos seus terminais, uma sequência de nucleótidos definindo um sitio de restrição de endonuclease. A digestão dos homólogos de ADN,



codificando V_H e V_L , possuindo sítios de restrição nos seus terminais ou próximo dos seus terminais, com uma ou mais endonucleases apropriados resulta na produção de homólogos com terminais coesivos de especificidade predeterminada. Em concretizações preferidas, o processo de PCR é usado, não apenas para amplificar os homólogos de ADN codificando V_H e/ou V_L da biblioteca, mas também, para induzir mutações na biblioteca e proporcionar assim uma biblioteca possuindo uma maior heterogeneidade. Em primeiro lugar, será de notar que o processo de PCR é, ele próprio, inerentemente, mutagénico, devido a uma variedade de factores bem conhecidos na arte. Em segundo lugar, para além da mutação, induzindo variações, descrita na Patente dos E.U.A. nº. 4 683 195 anteriormente referenciada, podem-se usar outras variações de PCR induzindo mutações. Por exemplo, a mistura reaccional de PCR, i.e., as misturas reaccionais combinadas de extensão do primeiro e do segundo iniciador, podem ser formadas com diferentes quantidades de um ou mais dos nucleótidos a serem incorporados no produto de extensão. Sob estas condições, a reacção de PCR prossegue para produzir substituições de nucleótidos no produto de extensão como resultado da escassez de uma base particular. Similarmente, podem-se incorporar quantidades molares dos nucleótidos aproximadamente iguais, na mistura reaccional de PCR inicial, numa quantidade para desempenhar eficientemente, um número X de ciclos e, depois, para submeter a mistura a um número de ciclos em excesso de X , tal como, por exemplo, $2X$. Alternativamente, as mutações podem ser induzidas durante a reacção de PCR incorporando, na mistura reaccional, derivados de nucleótido, tais como inosina, que não se encontram, normalmente, nos ácidos nucleicos do reportório que está a ser amplificado. Durante a amplificação subsequente in vivo, o derivado de nucleótido será substituído por um nucleótido substituído induzindo uma mutação pontual.

4. Expressão de homólogos de ADN de V_H e/ou de V_L

Os homólogos de ADN, codificando V_H e/ou V_L , contidos na biblioteca produzida pelo método acima descrito podem ser ligados



operativamente a um vector para amplificação e/ou expressão.

Tal como é aqui usado, o termo "vector" refere-se a uma molécula de ácido nucleico capaz de transportar, entre diferentes ambientes genéticos, outro ácido nucleico ao qual foi operativamente ligada. Um tipo de vector preferido é um epissona, i.e., uma molécula de ácido nucleico capaz de replicação extra-cromossómica. Os vectores preferidos são aqueles que são capazes de replicação autónoma e/ou expressão de ácidos nucleicos aos quais estão ligados. Os vectores capazes de dirigir a expressão de genes aos quais estão operativamente ligados são aqui designados por "vectores de expressão".

A escolha do vector ao qual o homólogo de ADN codificando V_H e/ou V_L está operativamente ligado depende, directamente, como é bem conhecido na arte, das propriedades funcionais desejadas, p.e., de replicação ou de expressão de proteína, e da célula hospedeira a ser transformada, sendo estas limitações inerentes à arte de construir moléculas de ADN recombinante.

Em concretizações preferidas, o vector utilizado inclui um replicação procariótico, i.e., uma sequência de ADN possuindo a capacidade de dirigir a replicação autónoma e a manutenção da molécula de ADN recombinante extra-cromossomicamente numa célula hospedeira procariótica, tal como uma célula hospedeira bacteriana, com ela transformada. Estes replicões são bem conhecidos na arte. Adicionalmente, essas concretizações que incluem um replicação procariótico incluem também um gene cuja expressão confere uma vantagem selectiva, tal como resistência a drogas, a um hospedeiro bacteriano com ele transformado. Genes bacterianos, de resistência a drogas, típicos, incluem os que conferem resistência à ampicilina ou à tetraciclina.

Os vectores que incluem um replicação procariótico podem também incluir um promotor procariótico capaz de dirigir a expressão (transcrição e tradução) dos homólogos codificando V_H e/ou V_L numa célula hospedeira bacteriana, tal como E. coli, com ele transformada. Um promotor é um elemento de controlo de expressão, formado por uma sequência de ADN, que permite que a



ligação de ARN-polimerase e a transcrição ocorram. As sequências de promotor compatíveis com hospedeiros bacterianos são, tipicamente, fornecidas em vetores plasmídeo, contendo sítios de restrição convenientes para inserção de um segmento de ADN, do presente invento. Exemplos típicos destes plasmídeos vetores incluem pUC8, pUC9, pBR322 e pBR329, disponíveis nos BioRad Laboratories, (Richmond, CA, E.U.A.) e pPL e pKK223, disponíveis de Pharmacia, (Piscataway, NJ, E.U.A.).

Podem-se também usar vetores de expressão compatíveis com células eucarióticas, preferivelmente, os compatíveis com células de vertebrados. Os vetores de expressão de células eucarióticas são bem conhecidos na arte e estão disponíveis em várias fontes comerciais. Tipicamente, estes vetores são fornecidos contendo os sítios de restrição convenientes para inserção dos homólogos de ADN desejados. Exemplos típicos destes vetores incluem pSVL e pKSV-10 (Pharmacia), pBPV-1/PML2d (International Biotechnologies, Inc.), e pTDT1 (ATCC, Nº. 31255).

Em concretizações preferidas, os vetores de expressão de células eucarióticas incluem um marcador de selecção que é eficaz numa célula eucariótica, preferivelmente, um marcador de selecção de resistência a drogas. Um marcador de resistência a drogas preferido é o gene cuja expressão resulta na resistência à neomicina, i.e., o gene da neomicina-fosfotransferase (neo). Southern et al., J. Mol. Appl. Genet., 1:327-341 (1982).

É também contemplado o uso de vetores de expressão retrovirais para exprimir os genes dos homólogos de ADN codificando V_H e/ou V_L . Tal como é aqui usado, o termo "vector de expressão retroviral" refere-se a uma molécula de ADN que inclui uma sequência de promotor derivada da região de repetição terminal longa (LTR) de um genoma de retrovirus.

Em concretizações preferidas o vector de expressão é, tipicamente, um vector de expressão retroviral que é, preferivelmente, incompetente na replicação em células eucarióticas. A construção e o uso de vetores retrovirais foi descrita por Sorge et al., Mol. Cel. Biol., 4:1730-1737 (1984).



Uma variedade de métodos foi desenvolvida para ligar operativamente ADN a vectores, via terminais coesivos complementares. Por exemplo, os terminais coesivos complementares podem ser ligados por engenharia genética aos homólogos de ADN codificando V_H e/ou V_L , durante a reacção de extensão de iniciador, pelo uso de um iniciador de síntese de polinucleótido apropriadamente concebido, como anteriormente se discutiu. O vector, e se necessário o homólogo de ADN, é clivado com uma endonuclease de restrição para produzir terminais complementares dos do homólogo de ADN. Os terminais coesivos, complementares do vector e do homólogo de ADN são, depois, ligados operativamente (unidos) para produzir uma molécula de ADN de cadeia dupla unitária.

Em concretizações preferidas, os homólogos de ADN codificando V_H e V_L de diversas bibliotecas são combinados aleatoriamente in vitro para expressão policistrónica a partir dos vectores individuais. Isto é, produz-se uma população diversa de vectores de expressão de ADN de cadeia dupla, na qual cada vector exprime, sob o controlo de um único promotor, um homólogo de ADN codificando V_H e um homólogo de ADN codificando V_L , sendo a diversidade da população o resultado de diferentes combinações de homólogos de ADN codificando V_H e/ou V_L . A combinação aleatória in vitro pode ser realizada usando dois vectores de expressão, distintos um do outro, pela localização, em cada um deles de um sítio de restrição comum a ambos. Preferivelmente, os vectores são ADN de cadeia dupla linear, tal como um vector derivado do Lambda Zap, tal como aqui se descreve. No primeiro vector, o sítio está localizado entre um promotor e um poli-ligante, i.e., 5' terminal (a montante em relação à direcção de expressão) em relação ao poli-ligante, mas 3' terminal (a jusante em relação à direcção de expressão). No segundo vector, o poli-ligante está localizado entre um promotor e o sítio de restrição, i.e., o sítio de restrição está localizado 3' terminal em relação ao poli-ligante, e o poli-ligante está localizado 3' terminal em relação ao promotor.

Em concretizações preferidas, cada um dos vectores define



uma sequência de nucleótidos codificando uma ligação de ribossoma e um líder, estando a sequência localizada entre o promotor e o poli-ligante, mas a jusante (3' terminal) do sitio de restrição partilhado, se esse sitio está entre o promotor e o poli-ligante. São também preferidos os vectores contendo um codão de paragem a jusante do poli-ligante, mas a montante de qualquer sitio de restrição partilhado, se esse sitio está a jusante do poli-ligante. O primeiro e/ou o segundo vector podem definir também uma sequência de nucleótidos codificando um péptido tag. A sequência tag está localizada, tipicamente a jusante do poli-ligante mas a montante de qualquer codão de paragem que possa estar presente. Em concretizações preferidas, os vectores contêm marcadores seleccionáveis de modo a que a presença de uma porção desse vector, i.e., de um braço lambda particular, possa ser seleccionado a favor ou seleccionado contra. Os marcadores seleccionáveis típicos são bem conhecidos dos peritos na arte. Exemplos destes marcadores incluem genes de resistência a antibióticos, marcadores seleccionáveis geneticamente, supressores de mutação, tais como supressores âmbar e similares. Os marcadores seleccionáveis estão localizados, tipicamente, a montante do promotor e/ou a jusante do segundo sitio de restrição. Em concretizações preferidas, um marcador seleccionável está localizado a montante do promotor no primeiro vector contendo os homólogos de ADN codificando V_H . Um segundo marcador seleccionável está localizado a jusante do segundo sitio de restrição no vector contendo os homólogos de ADN codificando V_L . Este segundo marcador seleccionável pode ser o mesmo ou diferente do primeiro, desde que, quando os vectores codificando V_H e os vectores codificando V_L são combinados aleatoriamente através do primeiro sitio de restrição, os vectores resultantes contendo V_H e V_L e ambos os marcadores seleccionáveis, possam ser seleccionados.

Tipicamente, o poli-ligante é uma sequência de nucleótidos que define um ou mais, preferivelmente, pelo menos dois, sitios de restrição, cada um único do vector e, preferivelmente, não partilhado pelo outro vector, i.e., se está presente no primeiro

vector, não está presente no segundo vector. Os sítios de restrição do poli-ligante estão orientados de modo a permitir a ligação de homólogos de ADN codificando V_H ou V_L no vector na mesma cadeia de leitura que qualquer sequência líder, tag ou de codão de paragem, presente.

A combinação aleatória é realizada ligando homólogos de ADN codificando V_H no primeiro vector, tipicamente, num sítio ou em sítios de restrição no poli-ligante. Similarmente, os homólogos de ADN codificando V_L são ligados no segundo vector, criando-se assim duas populações diversas de vectores de expressão. Não importa qual é o tipo de homólogo de ADN, i.e., V_H ou V_L que se liga a qual, vector. mas prefere-se, por exemplo, que todos os homólogos de ADN codificando V_H estejam ligados, quer ao primeiro, quer ao segundo vector e que todos os homólogos de ADN codificando V_L estejam ligados ao outro do primeiro ou do segundo vector. Os membros de ambas as populações são depois clivados com uma endonuclease no sítio de restrição partilhado, tipicamente, digerindo ambas as populações com a mesma enzima. O produto resultante é duas populações diversas de fragmentos de restrição cujos membros de uma têm terminais coesivos complementares com os terminais coesivos dos membros da outra. Os fragmentos de restrição das duas populações são ligados aleatoriamente uns aos outros, i.e., realiza-se uma ligação, interpopulação, aleatória, para produzir uma população diversa de vectores, possuindo, cada um, um homólogo de ADN codificando V_H e codificando V_L na mesma cadeia de leitura e sob o controlo do promotor do segundo vector. Claro que se podem efectuar recombinações subsequentes por clivagem no sítio de restrição partilhado, que é, tipicamente, formado de novo, por ligação de membros das duas populações, seguida de re-ligações subsequentes.

A construção resultante é depois introduzida num hospedeiro apropriado para proporcionar amplificação e/ou expressão dos homólogos de ADN codificando V_H e/ou V_L , quer separadamente, quer em combinação. Quando co-expressos no mesmo organismo, quer no mesmo, quer em vectores diferentes, produz-se um Fv funcionalmente activo. Quando se exprimem os polipéptidos V_H e V_L em



organismos diferentes, os polipéptidos respectivos são isolados e depois combinados num meio apropriado para formar um Fv. Os hospedeiros celulares nos quais se introduziu uma construção contendo homólogo de ADN codificando V_H e/ou V_L , são aqui referidos como tendo sido "transformados" ou como "transformantes".

A célula hospedeira pode ser quer procariótica quer eucariótica. As células bacterianas são células hospedeiras procarióticas preferidas e, tipicamente, são uma estirpe de E. coli tal como, por exemplo, E. coli da estirpe DH5 disponível nos Bethesda Research Laboratories, Inc., Bethesda, MD, E.U.A.. As células hospedeiras procarióticas preferidas incluem células de levedura e de mamífero, preferivelmente, células de vertebrado tais como as de linhas de células de ratinho, de ratazana, de macaco ou humanas.

A transformação de hospedeiros celulares apropriados como uma molécula de ADN recombinante do presente invento é realizada por métodos que, tipicamente, dependem do tipo de vector usado. Em relação à transformação de células hospedeiras procarióticas, ver, por exemplo, Cohen et al., Proceedings National Academy of Science, USA, Vol. 69, P. 2110 (1972); e Maniatis et al., Molecular Cloning, a Laboratory Manual, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, E.U.A. (1982). Em relação à transformação de células de vertebrado com vectores retrovirais contendo ADN_r, ver, por exemplo, Sorge et al., Mol. Cell. Biol., 4:1730-1737 (1984); Graham et al., Virology, 52:456 (1973); e Wigler et al., Proceedings National Academy of Sciences, USA, Vol. 76, P. 1373-1376 (1976).

5. Pesquisa para determinar a expressão de polipéptidos V_H e/ou V_L

As células transformadas com sucesso, i.e., as células contendo um homólogo de ADN codificando V_H e/ou V_L , ligado operativamente a um vector, podem ser identificadas por qualquer técnica adequada, bem conhecida para a detecção da ligação de um receptor a um ligando ou para detectar a presença de um

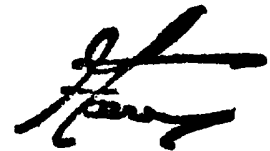
polinucleótido codificando o receptor, preferivelmente, o seu sitio activo. Os ensaios de detecção preferidos incluem aqueles onde a ligação do ligando pelo receptor produz um sinal detectável, quer directa, quer indirectamente. Estes sinais incluem, por exemplo, a produção de um complexo, a formação de um produto de reacção catalitica, a libertação ou a incorporação de energia e similares. Por exemplo, podem-se clonar células de uma população sujeita a transformação com um dado ADN para produzir colónias monoclonais. As células que formam essas colónias podem ser colhidas, lisadas e pode-se examinar o seu conteúdo em ADN quanto à presença do ADN usando um método tal como o descrito por Southern, J. Mol. Biol., 98:503 (1975) ou por Berent et al., Biotech. 3:208 (1985).

Para além de pesquisar directamente a presença de um homólogo de ADN codificando V_H e/ou V_L , a transformação bem sucedida pode ser confirmada por métodos imunológicos bem conhecidos, especialmente, quando os polipéptidos V_H e/ou V_L produzidos contêm um epitopo pré-seleccionado. Por exemplo, testam-se amostras de células, suspeitas de terem sido transformadas, quanto à presença do epitopo pré-seleccionado, usando um anticorpo contra o epitopo.

6. Bibliotecas de genes codificando V_H e/ou V_L

O presente invento contempla uma biblioteca de genes, preferivelmente, produzida por uma reacção de extensão de iniciador ou por uma combinação de reacções de extensão de iniciadores, tal como aqui se descreve, contendo pelo menos cerca de 10^3 , preferivelmente, pelo menos cerca de 10^4 e, mais preferivelmente, pelo menos, cerca de 10^5 homólogos de ADN diferentes codificando V_H e/ou V_L . Os homólogos estão, preferivelmente, numa forma isolada, isto é, substancialmente isenta de materiais tais como, por exemplo, agentes e/ou substratos de reacção de extensão, de iniciador, de segmentos de ADN genómico e similares.

Em concretizações preferidas, uma porção substancial dos homólogos presentes na biblioteca estão ligados operativamente a



um vector, preferivelmente, ligados operativamente para expressão a um vector de expressão.

Preferivelmente, os homólogos estão presentes num meio adequado para manipulação in vitro, tal como água, água contendo sais tamponantes e similares. O meio deverá ser compatível com a manutenção da actividade biológica dos homólogos. Adicionalmente, os homólogos deverão estar presentes numa concentração suficiente para permitir a transformação de uma célula hospedeira com eles compatível, com frequências razoáveis.

Prefere-se, ainda, que os homólogos estejam presentes em células hospedeiras transformadas com eles compatíveis.

D. Vectorios de expressão

O presente invento contempla também vários vectorios de expressão úteis na realização inter alia dos processos do presente invento. Cada um dos vectorios de expressão é um novo derivado do Lambda Zap.

1. Lambda Zap II

O Lambda Zap II é preparado substituindo o gene Lambda S do vector Lambda Zap pelo gene Lambda S do vector Lambda gt10, tal como se descreve no Exemplo 6.

2. Lambda Zap II V_H

O Lambda Zap II V_H é preparado inserindo as sequências de ADN sintético, ilustradas na Figura 6A, no vector Lambda Zap II acima descrito. A sequência de nucleótidos inserida proporciona, vantajosamente, um sítio de ligação de ribossoma (sequência de Shine-Dalgarno) para permitir a imitação adequada da tradução de ARNm em proteína, e uma sequência líder para dirigir eficazmente a proteína traduzida para o periplasma. A preparação do Lambda Zap II V_H é descrita, com mais detalhe, no Exemplo 9, e as suas características são ilustradas nas figuras 6A e 7.

3. Lambda Zap II V_L

O Lambda Zap II V_L é preparado como se descreve no



Exemplo 12 inserindo, no Lambda Zap II, a sequência de ADN sintético, ilustrada na Figura 6B. As características importantes do Lambda Zap II V_L estão ilustradas na Figura 8.

4. Lambda Zap II V_L II

O Lambda Zap II V_L II é preparado como se descreve no Exemplo 11, inserindo, no Lambda Zap II, a sequência de ADN sintético, ilustrada na Figura 10.

Os vectores acima descritos são compatíveis com hospedeiros E. coli, i.e., podem exprimir, para secreção para o periplasma, proteínas codificadas por genes aos quais foram ligados operativamente para expressão.

Exemplos

Pretende-se que os exemplos seguintes ilustrem mas não limitem o âmbito do presente invento.

1. Seleccção de polinucleótidos

As sequências de nucleótidos codificando CDR's de proteína de imunoglobulina são altamente variáveis. Contudo, existem várias regiões de sequências conservadas que flanqueiam os domínios V_H . Por exemplo, contêm sequências de nucleótidos substancialmente conservadas, i.e., sequências que se hibridarão com a mesma sequência de iniciador. Deste modo, construíram-se iniciadores de síntese de polinucleótidos (amplificação) que se hibridam com as sequências conservadas e que incorporam sítios de restrição, no homólogo de ADN produzido, que são adequados para ligar, operativamente, os fragmentos de ADN sintetizados a um vector. Mais especificamente, inseriram-se os homólogos de ADN no vector Lambda Zap II (Stratagene Cloning System, San Diego, CA, E.U.A.) nos sítios Xho I e EcoR I. Para amplificação dos domínios V_H , o iniciador 3' (iniciador 12 da Tabela 1) foi concebido para ser complementar com o ARNm na região J_H . Em todos os casos, os iniciadores 5' (iniciadores 1-10, Tabela 1) foram escolhidos para serem complementares com a primeira cadeia de ADNc na região do terminal N conservada (cadeia anti-sentido). Inicialmente a



amplificação foi realizada com uma mistura de 32 iniciadores (iniciador 1, Tabela 1) que eram degenerados em cinco posições. O ARNm de hibridoma pôde ser amplificado com iniciadores mistos, mas as tentativas iniciais para amplificar o ARNm de baço conduziram a resultados variáveis. Deste modo, compararam-se várias alternativas de amplificação usando os iniciadores 5' mistos.

A primeira alternativa consistiu em construir múltiplos iniciadores únicos, oito dos quais se mostram na Tabela 1, correspondentes aos membros individuais do lote de iniciadores mistos. Os iniciadores individuais 2-9 da Tabela 1 foram construídos incorporando qualquer dos dois nucleótidos possíveis em três das cinco posições degeneradas.

A segunda alternativa consistiu em construir um iniciador contendo inosina (iniciador 10, Tabela 1) em quatro das posições variáveis com base no trabalho, publicado, de Takahashi, et al., Proc. Natl. Acad. Sci. (U.S.A.), 82:1931-1935, (1985) e de Ohtsuka et al., J. Biol. Chem., 260: 2605-2608, (1985). Este iniciador tem a vantagem de não ser degenerado e, ao mesmo tempo, minimiza os efeitos negativos de não ajustamento nas posições não conservadas, tal como discutido por Martin et al., Nuc. Acids Res., 13:8927 (1985). Contudo, não se sabia se a presença de nucleótidos de inosina resultaria na incorporação de sequências não desejadas nas regiões V_H clonadas. Deste modo, a inosina não foi incluída na única posição que permanece nos fragmentos amplificados após a clivagem dos sítios de restrição. Como resultado, a inosina não estava presente na inserção clonada.

Conceberam-se iniciadores de amplificação V_H , adicionais, incluindo o iniciador 3' único para serem complementares com uma porção do primeiro domínio da região constante do ARNm de cadeia pesada gama 1 (iniciadores 15 e 16, Tabela 1). Estes iniciadores produzirão homólogos de ADN contendo polinucleótidos, codificando aminoácidos dos domínios V_H e da primeira região constante, da cadeia pesada. Estes homólogos de ADN podem, deste modo, ser usados para produzir fragmentos Fab em vez de um F_V .



Como controle para amplificação, a partir de ARNm de baço ou de hibridoma, construíram-se um conjunto de iniciadores que se hibridam com uma região altamente conservada no gene de cadeia pesada, da região constante de IgG. O iniciador 5' (iniciador 11, Tabela 1) é complementar do ADNc na região C_H2 enquanto que o iniciador 3' (iniciador 13, Tabela 1) é complementar com o ARNm na região C_H3 . Crê-se que não existiam desajustamentos entre estes iniciadores e os seus moldes.

As sequências de nucleótidos codificando as CDR de V_L são altamente variáveis. Contudo, existem várias regiões de sequências conservadas que flanqueiam os domínios CDR de V_L incluindo as regiões J_L , estrutura (framework) V_L e líder/promotor V_L . Deste modo, construíram-se iniciadores de amplificação que se hibridam com as sequências conservadas e incorporam sítios de restrição que permitem a clonação de fragmentos amplificados no vector pBluescript SK-, cortado com Nco I e Spe I. Para amplificação dos domínios CDR de V_L , o iniciador 3' (iniciador 14 na Tabela 1) foi concebido para ser complementar com o ARNm nas regiões J_L . O iniciador 5' (iniciador 15, Tabela 1) foi escolhido para ser complementar do ADN da primeira cadeia na região do terminal N conservada (cadeia anti-sentido).

Concebeu-se um segundo conjunto de iniciadores de amplificação para amplificação dos domínios CDR V_L , iniciadores 5' (iniciadores 1-8 na Tabela 2), para serem complementares com o ADNc da primeira cadeia na região do terminal N conservada. Estes iniciadores introduziram também um sítio de endonuclease de restrição Sac I para permitir que o homólogo de ADN de V_L fosse clonado no vector de expressão V_L II. O iniciador 3' de amplificação de V_L (iniciador 9 na Tabela 2) foi concebido para ser complementar com o ARNm nas regiões J_L e para introduzir o sítio de endonuclease de restrição Xba I requerido para inserir o homólogo de ADN de V_L no vector de expressão V_L II (Figura a).

Conceberam-se iniciadores 3' de amplificação de V_L adicionais, para se hibridarem com a região constante do ARNm



kappa ou lambda (iniciadores 10 e 11 na Tabela 2). Estes iniciadores permitem que se produza um homólogo de ADN, contendo sequências de polinucleótidos codificando os aminoácidos da região constante da cadeia kappa ou lambda. Estes iniciadores tornam possível produzir um fragmento Fab em vez de um F_y.

Pelo menos na Tabela 2, mostram-se os iniciadores usados para a amplificação de sequências de cadeia leve kappa, para construção de Fab. A amplificação com estes iniciadores foi realizada em 5 reacções separadas, cada uma contendo um dos iniciadores 5' (iniciadores 3-6 e 12) e um dos iniciadores 3' (iniciador 13). O iniciador 3' restante (iniciador 9) foi usado para construir fragmentos F_y. Os iniciadores 5' contêm um sítio de restrição Sac I e os iniciadores 3' contêm um sítio de restrição Xba I.

Pelo menos na Tabela 1, mostram-se os iniciadores usados para amplificação de fragmentos Fd de cadeia pesada, para a construção de Fab. A amplificação foi realizada em oito reacções separadas, cada uma contendo um dos iniciadores 5' (iniciadores 2-9) e um dos iniciadores 3' (iniciador 15). Os iniciadores 5' restantes foram usados para a amplificação numa única reacção e são, quer um iniciador degenerado (iniciador 1) quer um iniciador que incorpora inosina em quatro posições degeneradas (iniciador 10, Tabela 1, e iniciadores 17 e 18, Tabela 2). O iniciador 3' restante (iniciador 14, Tabela 2) foi usado para construir fragmentos F_y. Muitos dos iniciadores 5' incorporam um sítio Xho I, e os iniciadores 3' incluem um sítio de restrição Spe I.

Na Tabela 2 mostram-se também os iniciadores de amplificação de V_L concebidos para amplificar regiões variáveis, de cadeia leve, humanas, de ambos os isotipos lambda e kappa.

Todos os iniciadores e polinucleótidos sintéticos aqui usados e que se mostram nas Tabelas 1-4 ou foram adquiridos no Research Genetics em Huntsville, Alabama, E.U.A. ou foram sintetizados num sintetizador de ADN da Applied Biosystems, modelo 381A, usando as instruções do fabricante.

TABELA 1

(1)	5' AGGT(C/G) (C/A) A(G/A) CT(G/T) CTCGAGTC(T/A) GG 3'	Iniciador 5' degenerado para a amplificação da região de cadeia pesada variável (V_H)
(2)	5' AGGTCCAGCTGCTCGAGTCTGG 3'	Iniciador 5' único para a amplificação de V_H
(3)	5' AGGTCCAGCTGCTCGAGTCAGG 3'	"
(4)	5' AGGTCCAGCTTCTCGAGTCTGG 3'	"
(5)	5' AGGTCCAGCTTCTCGAGTCAGG 3'	"
(6)	5' AGGTCCAACTGCTCGAGTCTGG 3'	"
(7)	5' AGGTCCAACTGCTCGAGTCAGG 3'	"
(8)	5' AGGTCCAACTTCTCGAGTCTGG 3'	"
(9)	5' AGGTCCAACTTCTCGAGTCAGG 3'	"
(10)	5' AGGTIIAICTICTCGAGTC(T/A) 3'	Iniciador 5' degenerado contendo inosina em 4 posições degeneradas
(11)	5' GCCCAAGGATGTGCTCACC 3'	Iniciador 5' para amplificação na região C_H2 de IgG1 de ratinho
(12)	5' CTATTAGAAATTCACGGTAACAGTGGTGCCTTGGCCCCA 3'	Iniciador 3' para a amplificação de V_H
(12A)	5' CTATTAACTAGTAACGGTAACAGTGGTGCCTTGGCCCCA 3'	Iniciador 3' para a amplificação de V_H usando o sítio Spe I 3'
(13)	5' CTCAGTATGGTGTGTGC 3'	Iniciador 3' para amplificação na região C_H3 de IgG1 de ratinho
(14)	5' GCTACTAGTTTGTGATTTCCACCTTGG 3'	Iniciador 3' para a amplificação de V_L
(15)	5' CAGCCATGGCCGACATCCAGATG 3'	Iniciador 5' para a amplificação de V_L
(16)	5' AATTTTACTAGTCACCTTGGTGTGCTGGC 3'	Iniciador 3' único para a amplificação de V_H incluindo parte da primeira constante gama 1 de ratinho
(17)	5' TATGCCAACTAGTACAAACCAATCCCTGGGCACAATTTT 3'	Iniciador 3' único para a amplificação de V_L incluindo parte da primeira região constante gama 1 de ratinho e a região de articulação

TABELA 2

(1)	5' CCAGTTCGGAGCTCGTTGTGACTCAGGAATCT 3'	Iniciador 5' único para a amplificação de V_L
(2)	5' CCAGTTCGGAGCTCGTTGTGACGCGCCGCC 3'	"
(3)	5' CCAGTTCGGAGCTCGTTGCTCAGCCAGTCTCCA 3'	"
(4)	5' CCAGTTCGGAGCTCCAGATGACCCAGTCTCCA 3'	"
(5)	5' CCAGATGTGAGCTCGTGATGACCCAGACTCCA 3'	"
(6)	5' CCAGATGTGAGCTCGTCATGACCCAGTCTCCA 3'	"
(7)	5' CCAGATGTGAGCTCTTGATGACCCAACTCAA 3'	"
(8)	5' CCAGATGTGAGCTCGTGATAACCCAGGATGAA 3'	"
(9)	5' GCAGCATTCCTAGAGTTTCAGCTCCAGCTTGCC 3'	Iniciador 3' único para a amplificação de V_L
(10)	5' CCGCCGTCTAGAACACTCATTCCTGTTGAAGCT 3'	Iniciador 3' único para amplificação de V_L incluindo a região constante kappa
(11)	5' CCGCCGTCTAGAACATTCTGCAGGAGACAGACT 3'	Iniciador 3' único para a amplificação de V_L incluindo a região constante lambda
(12)	5' CCAGTTCGGAGCTCGTGATGACACAGTCTCCA 3'	Iniciador 5' único para a amplificação de V_L
(13)	5' GCGCCGTCTAGAAATTAACACTCATTCCTGTTGAA 3'	Iniciador 3' único para a amplificação de V_L
(14)	5' CTATTAACTAGTAACGGTAACAGTGGTGCCTTGCCCCA 3'	"



TABELA 2 (continuação)

(15)	5' AGGCTTACTAGTACAAATCCCTGGGCACAAT 3'	Iniciador 3' único para a amplificação de V_H
(16)	5' GCCGCTCTAGAACACTCATTCCTGTTGAA 3'	Iniciador 3' único para a amplificação de V_L
(17)	5' AGGTIIAICTICTCGAGTCTGC 3'	Iniciador 5' degenerado contendo inosi
(18)	5' AGGTIIAICTICTCGAGTCAGC 3'	na em 4 posições degeneradas
		" "
		" "



2. Produção de um relatório codificando V_H enriquecido em proteínas que se ligam ao FITC

O isotiocianato de fluoresceína (FITC) foi seleccionado como ligando para a ligação a receptor. Decidiu-se ainda enriquecer, por imunização, o relatório de genes imunológicos, i.e., reportórios de genes codificando V_H e V_L , em genes codificando receptores anti-FITC. Isto foi realizado ligando FITC a hemocianina de lapa Fissurela (KLH) usando as técnicas descritas em Antibodies A Laboratory Manual, Harlow and Lowe, eds., Cold Spring Harbor, New York, E.U.A. (1988). Resumidamente, adicionaram-se 10 miligramas (mg) de hemocianina de lapa Fissurela e 0,5 mg de FITC a 1 ml de tampão contendo carbonato de sódio 0,1 M a pH 9,6, e agitou-se durante 18 a 24 horas a 4 graus C (4°C). O FITC não ligado foi removido por filtração de gel através de Sephadex G-25.

Preparou-se o conjugado de KLH-FITC para injeção de ratinhos adicionando 100 µg do conjugado a 250 µl de solução salina tamponada com fosfato (PBS). Adicionou-se igual volume de adjuvante de Freund completo e emulsionou-se toda a solução durante 5 minutos. Injectou-se um ratinho 129 G_{IX}+ com 300 µl da emulsão. As injeções foram administradas subcutaneamente, em vários locais, usando uma agulha calibre 21. Duas semanas mais tarde administrou-se uma segunda imunização com KLH-FITC. Esta injeção foi preparada como se segue. Diluíram-se 50 µg de KLH-FITC em 250 µl de PBS e com esta solução de KLH-FITC misturou-se um volume igual de alumínio. O ratinho foi injectado intraperitonealmente com 500 µl da solução, usando uma agulha calibre 23. Um mês mais tarde, aos ratinhos, foi administrada uma injeção final de 50 µg do conjugado KLH-FITC diluídos até 200 µl em PBS. Esta injeção foi administrada intravenosamente, na veia lateral da cauda, usando uma agulha de calibre 30. Cinco dias após esta injeção final os ratinhos foram, sacrificados e isolou-se o ARN celular total dos seus baços.

Cultivou-se o hibridoma PCP BD11, produzindo um anticorpo imuno-específico para o éster de fosfonato, em meio DMEM (Gibco



Laboratories Grand Island, New York, E.U.A.) contendo 10 por cento de soro de vitelo fetal, suplementado com penicilina e estreptomicina. Recolheram-se cerca de 5×10^8 células de hibridoma e lavaram-se duas vezes em solução salina tamponada com fosfato. O ARN celular total foi preparado a partir destas células de hibridoma isoladas.

3. Preparação de um reportório de genes codificando V_H

O ARN celular total foi preparado a partir do baço de um único ratinho imunizado com KLH-FITC, como se descreveu no Exemplo 2, usando os métodos de preparação de ARN descritos por Chomczynski et al., Anal Biochem., 162:156-159 (1987) usando as instruções do fabricante e o conjunto experimental para isolamento de ARN, produzido por Stratagene Cloning Systems, La Jolla, CA. Resumidamente, imediatamente após a remoção do baço do ratinho imunizado, o tecido foi homogeneizado em 10 ml de uma solução desnaturante, contendo isotiocianato de guanina 4,0 M, citrato de sódio (0,25 M) a pH 7,0 e 2-mercaptoetanol 0,1 M, usando um homogeneizador de vidro. Com o baço homogeneizado misturou-se 1 ml de solução de acetato de sódio a uma concentração de 2 M a pH 4,0. Com a solução desnaturante, contendo o baço homogeneizado, misturou-se também um ml de fenol que tinha sido previamente saturado com H_2O . A este homogeneizado adicionaram-se 2 ml de uma mistura de clorofórmio:álcool isoamílico (24:1 v/v). Agitou-se vigorosamente o homogeneizado durante dez segundos e manteve-se em gelo durante 15 minutos. O homogeneizado foi depois transferido para um tubo de centrifuga de polipropileno de 50 ml, de parede espessa (Fisher Scientific Company, Pittsburg, PA, E.U.A.). A solução foi centrifugada a $10\,000 \times g$ durante 20 minutos a $40^\circ C$. A fase aquosa superior, contendo ARN, foi transferida para um tubo de centrifuga de polipropileno de 50 ml, novo, e misturada com um volume igual de álcool isopropílico. Esta solução foi mantida a $-20^\circ C$ durante pelo menos uma hora para precipitar o ARN. A solução contendo o ARN precipitado foi centrifugada a $10\,000 \times g$ durante vinte minutos a $40^\circ C$. O ARN celular total sedimentado foi recolhido e dissolvido em 3 ml da solução desnaturante descrita acima. Ao ARN celular total,



ressuspendido, adicionaram-se 3 ml de álcool isopropílico e agitou-se vigorosamente. Esta solução foi mantida a -20°C durante pelo menos 1 hora, para precipitar o ARN. A solução, contendo o ARN precipitado, foi centrifugada a $10\,000 \times g$ durante dez minutos a 4°C . O ARN sedimentado foi lavado uma vez com uma solução contendo etanol a 75%. O ARN sedimentado foi seco sob vácuo, durante 15 minutos, e depois suspenso de novo em H_2O tratada com pirocarbonato de dimetilo (DEPC), (H_2O -DEPC).

O ARN mensageiro (ARNm) enriquecido em sequências contendo tractos poli A longos foi preparado a partir do ARN celular total usando métodos descritos em Molecular Cloning A Laboratory Manual, Maniatis et al., eds., Cold Spring Harbor Laboratory, New York, E.U.A. (1982). Resumidamente, ressuspendeu-se metade do ARN total, isolado a partir do baço de um único ratinho imunizado, em um ml de H_2O -DEPC e manteve-se a 65°C durante cinco minutos. Ao ARN em suspensão adicionou-se um ml de tampão com uma carga de sais duas vezes mais elevada, consistindo de Tris-HCl 100 mM, cloreto de sódio 1 M, ácido etilenodiaminatetraacético dissódico (EDTA) 2,0 mM a pH 7,5 e 0,2% de dodecilsulfato de sódio (SDS) e deixou-se a mistura arrefecer até à temperatura ambiente. Depois, aplicou-se a mistura a uma coluna de oligo-dT (Tipo 2 ou Tipo 3 da Collaborative Research) que tinha sido previamente preparada lavando o oligo-dT com uma solução contendo hidróxido de sódio 0,1 M e EDTA 5 mM e, depois, equilibrando a coluna com H_2O -DEPC. O eluido foi recolhido num tubo esterilizado de polipropileno e aplicado de novo à mesma coluna após aquecimento do eluido durante 5 minutos a 65°C . A coluna de oligo-dT foi depois lavada com 2 ml de tampão com uma elevada carga de sais consistindo de Tris-HCl 50 mM a pH 7,5, cloreto de sódio 500 mM, EDTA 1 mM a pH 7,5 e 0,1% de SDS. A coluna de oligo-dT foi depois lavada com 2 ml de um tampão de médio teor em sais 1 X consistindo de Tris-HCl 50 mM a pH 7,5, cloreto de sódio 100 mM, EDTA 1 mM e 0,1% de SDS. O ARN mensageiro foi eluido da coluna de oligo-dT com 1 ml de tampão consistindo de Tris-HCl 10 mM a pH 7,5, EDTA 1 mM a pH 7,5 e 0,05% de SDS. O ARN mensageiro foi purificado por extracção desta solução com fenol/clorofórmio,

seguida por uma extração simples com clorofórmio a 100%. O ARN mensageiro foi concentrado, por precipitação em etanol, e ressuspendido em H₂O-DEPC.

O ARN mensageiro isolado pelo processo anterior contém uma pluralidade de polinucleótidos codificando V_H diferentes, i.e., mais do que cerca de 10⁴ genes codificando V_H diferentes.

4. Preparação de um polinucleótido codificando um único V_H

Os polinucleótidos codificando um único V_H foram isolados de acordo com o Exemplo 3 com a excepção de o ARN celular total ter sido extraído de células de hibridoma monoclonais preparadas no Exemplo 2. Os polinucleótidos isolados deste modo codificam um único V_H.

5. Preparação de homólogo de ADN

Na preparação para amplificação por PCR, o ARNm, preparado de acordo com os exemplos anteriores, foi usado como molde para a síntese de ADNc por uma reacção de extensão de iniciador. Tipicamente, numa mistura reaccional de transcrição de 50 µl, em primeiro lugar, hibridaram-se (temperaram-se) 5-10 µg de ARNm de baço ou de hibridoma em água, com 500 ng (50,0 pmole) do iniciador 3' de V_H (iniciador 12, Tabela 1) a 65°C, durante 5 minutos. Subsequentemente, ajustou-se a mistura até uma concentração 1,5 mM de dATP, dCTP, dGTP e dTTP, 40 mM de Tris-HCl a pH 8,0 8 mM de MgCl₂, 50 mM de NaCl e 2 mM de espermidina. Adicionou-se transcriptase inversa de virus de leucemia Moloney-Murina (Stratagene Cloning Systems), 26 unidades e manteve-se a solução durante 1 hora a 37°C.

A amplificação por PCR foi realizada num meio reaccional de 100 µl contendo os produtos de reacção da transcriptase inversa (aproximadamente, 5 µg do híbrido ADNc/ARN), 300 ng de iniciador 3' de V_H (iniciador 12 da Tabela 1), 300 ng de cada um dos iniciadores 5' de V_H (iniciadores 2-10 da Tabela 1), 200 mM de uma mistura de dNTP, 50 mM de KCl, 10 mM de Tris-HCl pH 8,3, 15 mM de MgCl₂, 0,1% de gelatina e 2 unidades de ADN-polimerase Taq. A mistura foi espalhada conjuntamente com óleo mineral e

submetida a 40 ciclos de amplificação. Cada ciclo de amplificação envolveu desnaturação a 92°C durante 1 minuto, t $\hat{e}$ mpera a 52°C durante 2 minutos e síntese de polinucleótidos por extensão de iniciador (elongação) a 72°C durante 1,5 minutos. As amostras contendo homólogo de ADN codificando V_H amplificado foram extractadas duas vezes com fenol/clorofórmio, uma vez com clorofórmio, precipitadas com etanol e armazenadas a -70°C num meio 10 mM em Tris-HCl, (pH 7,5) e 1 mM em EDTA.

Como se mostra na Figura 3, faixas R17-R24, usando iniciadores 5' únicos (2-9, Tabela 1) conseguiu-se a síntese e a amplificação, eficientes, do homólogo de ADN codificando V_H a partir de ARNm de baço. Observa-se o ADNc amplificado (homólogo de ADN codificando V_H) como uma banda principal com o tamanho esperado (360 pb). As intensidades do fragmento de polinucleótido, codificando V_H, amplificado, em cada reacção, parecem ser similares, indicando que todos estes iniciadores são, aproximadamente, igualmente eficientes na iniciação da amplificação. O rendimento e a qualidade da amplificação com estes iniciadores eram reprodutíveis. O iniciador contendo inosina sintetizou também homólogos de ADN, codificando V_H, amplificador, a partir de ARNm de baço, de um modo reprodutível, conduzindo à produção do fragmento com o tamanho esperado, de uma intensidade similar à de outros ADNc amplificados (Figura 4, faixa R16). Este resultado indicou que a presença de inosina permite também a síntese e a amplificação eficientes do homólogo de ADN, indicando claramente quão úteis são estes iniciadores na geração de uma pluralidade de homólogos de ADN codificando V_H. Os produtos de amplificação obtidos a partir dos iniciadores da região constante (iniciadores 11 e 13, Tabela 1) foram mais intensos indicando que a amplificação foi mais eficiente, possivelmente, em virtude de um maior grau de homologia entre o molde e os iniciadores (Figura 4, faixa R9). Com base nestes resultados, construiu-se uma biblioteca de genes codificando V_H a partir dos produtos de oito amplificações, cada uma realizada com um iniciador 5' diferente. Misturaram-se porções iguais dos produtos de cada reacção de extensão de iniciador e usou-se

depois o produto misto para gerar uma biblioteca de vectores contendo homólogos de ADN codificando V_H .

Os homólogos de ADN do V_L foram preparados a partir do ARNm purificado preparado como descrito acima. Na preparação para amplificação por PCR, o ARNm preparado, de acordo com os exemplos acima, foi usado como molde para a síntese de ADNc. Tipicamente, numa mistura reaccional de transcrição de 50 μ l, temperaram-se, em primeiro lugar 5-10 μ g de ARNm de hibridoma ou de baço, em água, com 300 ng (50,0 pmole) do iniciador 3' de V_L (iniciador 14, Tabela 1), a 65°C durante cinco minutos. Subsequentemente ajustou-se a mistura a uma concentração 1,5 mM de dATP, dCTP, dGTP e dTTP, 40 mM de Tris-HCl a pH 8,0, 8 mM de $MgCl_2$, 50 mM de NaCl e 2 mM de espermidina. Adicionaram-se 26 unidades de transcriptase inversa de virus de leucemia Moloney-Murina (Stratagene Cloning Systems) e manteve-se a solução durante 1 hora a 37°C. A amplificação por PCR foi realizada numa mistura reaccional de 100 μ l contendo, aproximadamente, 5 μ g do híbrido de ADNc/ARN, produzido como se descreveu acima, 300 ng do iniciador 3' de V_L (iniciador 14 da Tabela 1), 300 ng do iniciador 5' de V_L (iniciador 15 da Tabela 1), 200 mM de uma mistura de dNTP, 50 mM de KCl, 10 mM de Tris-HCl, pH 8,3, 15 mM de $MgCl_2$, 0,1% de gelatina e 2 unidades de ADN-polimerase Taq. A mistura reaccional foi espalhada conjuntamente com óleo mineral e submetida a 40 ciclos de amplificação. Cada ciclo de amplificação envolveu desnaturação a 92°C durante 1 minuto, tẽpera a 52°C durante 2 minutos e elongação a 72°C durante 1,5 minutos. As amostras amplificadas foram extractadas duas vezes com fenol/clorofõrmio, uma vez com clorofõrmio, precipitadas em etanol e foram armazenadas a -70°C num meio 10 mM em Tris-HCl a pH 7,5 e 1 mM em EDTA.

6. Inserção de homólogos de ADN em vectores

Na preparação para clonação de uma biblioteca enriquecida em sequências de V_H , digeriram-se produtos amplificados por PCR (2,5 mg/30 μ l de meio 150 mM em NaCl, 8 mM em Tris-HCl (pH 7,5) 6 mM em $MgSO_4$, 1 mM em DTT, 200 mg/ml de albumina de soro de bovino



(BSA)), a 37°C, com as enzimas de restrição Xho I (125 unidades) e EcoR I (10 U) e purificaram-se num gel de agarose a 1%. Em experiências de clonação, que requereram uma mistura dos produtos das reacções de amplificação, combinaram-se volumes iguais (50 µl, concentração de 1-10 µg) de cada mistura reaccional, após amplificação mas antes da digestão de restrição. Após electroforese em gel do ARNm de baço, amplificado por PCR, digerido, excisou-se a região do gel contendo fragmentos de ADN de, aproximadamente, 350 pb, electroeluiu-se para uma membrana de diálise, precipitou-se em etanol e ressuspendeu-se em meio 10 mM em Tris-HCl, pH 7,5 e 1 mM em EDTA, até uma concentração de 10 ng/µl. Ligaram-se, depois, quantidades equimolares do inserido, de um dia para o outro, a 50°C, a 1 µg de vector Lambda Zap(R) II (Stratagene Cloning Systems, La Jolla, CA, E.U.A.) previamente cortado por EcoR I e Xho I. Uma porção da mistura de ligação (1 µl) foi empacotada durante 2 horas à temperatura ambiente usando extracto de empacotamento Gigapack Gold (Stratagene Cloning Systems, La Jolla, CA, E.U.A.) e o material empacotado foi colocado, em placa, em células hospedeiras Azuis XL1. Determinou-se que a biblioteca consistia em 2×10^7 homólogos de V_H com menos do que 30% de produto residual não recombinante.

O vector usado acima, o Lambda Zap II, é um derivado do Lambda Zap original (ATCC # 40 298) que mantém todas as características do Lambda Zap original, incluindo 6 sítios de clonação únicos, expressão de proteínas de fusão, e a capacidade de excisar rapidamente o inserido na forma de um fagomídeo (Bluescript SK-) mas não tem a mutação SAM 100, permitindo o crescimento em muitas estirpes Não-Sup F, incluindo a Azul XL1. O Lambda Zap II foi construído como se descreve em Short et al., Nucleic Acids Res., 16:7583-7600, 1988, mas substituindo o gene Lambda S contido num fragmento de ADN de 4254 pares de bases (pb) produzido pela digestão do Lambda Zap com a enzima de restrição NcoI. Este fragmento de ADN, de 4254 pb, foi substituído pelo fragmento de ADN de 4254 pb contendo o gene Lambda S, isolado a partir do Lambda gt10 (ATCC # 40 179), após digestão do vector com a enzima de restrição NcoI. O fragmento de ADN de 4254 pb,



isolado a partir do lambda gt10, foi ligado no vector Lambda Zap original, usando ADN-ligase T4 e protocolos convencionais, para estes procedimentos, descritos em Current Protocols in Molecular Biology, Ausubel et al., eds., John Wiley and Sons, New York, E.U.A., 1987.

Na preparação de clonação de uma biblioteca enriquecida em sequências V_L , digeriram-se 2 μ g de produtos amplificados por PCR (2,5 mg/30 μ l de meio 150 mM em NaCl, 8 mM em Tris-HCl (pH 7,5), 6 mM em $MgSO_4$, 1 mM em DTT, 200 mg/ml de BSA) a 37°C, com as enzimas de restrição Nco I (30 unidades) e Spe I (45 unidades). Os produtos amplificados por PCR digeridos foram purificados num gel de agarose a 1%, usando a técnica de electroeluição convencional descrita em Molecular Cloning A Laboratory Manual, Maniatis et al., eds., Cold Spring Harbor, New York, E.U.A. (1982). Resumidamente, após electroeluição em gel de produto amplificado por PCR digerido, excisou-se a região do gel contendo o fragmento de ADN, codificando V_L , com o tamanho apropriado, electroeluiu-se para uma membrana de diálise, precipitou-se com etanol e ressuspendeu-se, até uma concentração final de 10 ng por ml, numa solução contendo 10 mM de Tris-HCl, a um pH de 7,5 e 1 mM de EDTA.

Ligou-se uma quantidade molar igual de ADN, representando uma pluralidade de homólogos de ADN, codificando V_L , diferentes, a um vector fagomideo pBluescript SK- que tinha sido previamente cortado com Nco I e Spe I. Uma porção da mistura de ligação foi transformada, usando as instruções do fabricante, em células competentes Epicurian Coli Azuis XL1 (Stratagene Cloning Systems), La Jolla, CA). Determinou-se que a biblioteca transformante consistia de $1,2 \times 10^3$ unidades formando colónias/ μ g de homólogos de V_L , com menos do que 3% de produto residual não recombinante.

7. Sequenciação de plasmideos da biblioteca de ADNc codificando V_H

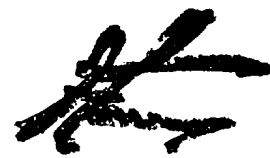
Para analisar os clones de fago Lambda Zap II, os clones foram excisados do Lambda Zap para plasmideos, de acordo com as instruções do fabricante (Stratagene Cloning Systems, La Jolla,

CA, E.U.A.). Resumidamente, as placas de fago foram nucleadas em pratos de agar e transferidas para tubos de microcentrifugadora estéreis, contendo, 500 μ l de um tampão 50 mM de Tris-HCl, a pH 7,5, 100 mM de NaCl, 10 mM de $MgSO_4$, e 0,01% de gelatina e 20 μ l de clorofórmio.

Para as excisões incubaram-se 200 μ l da reserva de fago, 200 μ l de células Azuis XL1 ($A_{600} = 1,00$) e 1 μ l de fago auxiliar R408 (1×10^{11} ufp/ml), a 37°C durante 15 minutos. Os plasmídeos excisados foram infectados em células Azuis XL1 e colocados em pratos de LB, contendo ampicilina. O ADN de cadeia dupla foi preparado a partir de células contendo o fagomídeo de acordo com os métodos descritos por Holmes et al., Anal. Biochem., 114:193, (1981). Os clones foram, em primeiro lugar, pesquisados, relativamente às inserções de ADN, por digestão de restrição, quer com Pvu II, quer com Bgl I e os clones contendo o inserido V_H putativo foram sequenciados, usando transcriptase inversa, de acordo com o método geral descrito por Sanger et al., Proc. Natl. Acad. Sci., USA, 74:5463-5467, (1977) e com as modificações, específicas deste método, fornecidas nas instruções do fabricante no conjunto experimental de sequenciação com transcriptase inversa, AMV, usando ^{35}S -dATP, de Stratagene Cloning Systems, La Jolla, CA, E.U.A..

8. Caracterização do reportório de V_H clonado

Os produtos amplificados que foram digeridos com Xho I e com EcoR I e clonados no Lambda ZAP, resultaram numa biblioteca de ADNc com $9,0 \times 10^5$ ufp. De modo a confirmar que a biblioteca consistia de uma população diversa de homólogos de ADN codificando V_H , as 120 bases do terminal N de 18 clones, seleccionados ao acaso da biblioteca, foram excisadas e sequenciadas (Figura 5). Para determinar se os clones tinham origem no gene de V_H , as sequências clonadas foram comparadas com sequências de V_H e com sequências de V_L , conhecidas. Os clones exibiram uma homologia de 80 a 90% com sequências com origem em cadeias pesadas conhecidas e uma pequena homologia com sequências com origem em cadeias leves, quando comparadas com as sequências



disponíveis em Sequences of Proteins of Immunological Interest by Kabot et al., 4th ed., U.S. Dept. of Health and Human Sciences, (1987). Isto demonstra que a biblioteca foi enriquecida na sequência V_H desejada, de preferência relativamente a outras sequências tais como as sequências de cadeia leve.

A diversidade da população foi determinada classificando os clones sequenciados em subgrupos pré-definidos (Figura 5). As sequências de V_H de ratinho são classificadas em onze subgrupos (Figura 5). As sequências de V_H de ratinho são classificadas em onze subgrupos [I (A,B), II (A,B,C), III (A,B,C,D) V (A,B)] com base nas sequências de aminoácidos da estrutura descritas em Sequences of Proteins of Immunological Interest por Kabot et al., 4th ed., U.S. Dept. of Health and Human Sciences, (1987); Dildrop, Immunology Today, 5:84, (1984); e Brodeur et al., Eur. J. Immunol., 14; 922, (1984). A classificação dos clones sequenciados demonstrou que a biblioteca de ADNc continha sequências de V_H de, pelo menos, 7 subgrupos diferentes. Adicionalmente, uma comparação, par a par, da homologia entre os clones sequenciados mostrou que não existiam duas sequências idênticas em todas as posições, sugerindo que a população é variada, na medida em que é possível caracterizá-la por análise e sequência.

Seis dos clones (L 36-50, Figura 5) pertencem à subclasse III B e tinham sequências de nucleótidos muito similares. Isto pode reflectir uma preponderância de ARNm derivado de um ou de vários genes variáveis relacionados em baço estimulado, mas os resultados não permitem pôr de parte a possibilidade de uma tendência no processo de amplificação.

9. Construção do vector de expressão de V_H

O critério principal usado na escolha de um sistema de vector foi a necessidade de gerar o maior número de fragmentos Fab que pudesse ser pesquisado directamente. O lambda de bacteriofago foi seleccionado como vector de expressão por três razões. A primeira, é que o empacotamento in vitro do ADN de fago é o método mais eficiente de re-introduzir ADN nas células de



hospedeiro. A segunda razão, é que é possível detectar a expressão de proteína ao nível de placas de fago simples. Por último, a pesquisa de bibliotecas de fago envolve, tipicamente, menos dificuldade com a ligação não específica. A alternativa, os vectores de clonação de plasmideo são apenas vantajosos na análise de clones após terem sido identificados. Esta vantagem não se perde no presente sistema devido ao uso do lambda Zap, permitindo, deste modo, a excisão de um plasmideo contendo as inserções exprimindo a cadeia pesada, a cadeia leve ou o Fab.

Para exprimir a pluralidade de homólogos de ADN codificando um V_H numa célula hospedeira de E. coli, construiu-se um vector que colocou os homólogos de ADN codificando o V_H na cadeia de leitura apropriada, forneceu um sítio de ligação a ribossomas tal como se descreve em Shine et al., Nature, 254:34, 1975, forneceu uma sequência líder para dirigir a proteína expressa para o espaço periplasmático, forneceu uma sequência de polinucleótidos que codificava um epitopo conhecido (epitopo tag) e forneceu também um polinucleótido que codificava uma proteína espaçadora entre o homólogo de ADN codificando V_H e o polinucleótido codificando o epitopo tag. Construiu-se uma sequência de ADN sintético, contendo todos os polinucleótidos e as características anteriores, concebendo segmentos de polinucleótidos de cadeia simples de 20-40 bases que se hibridam, umas com as outras, e formam a sequência de ADN sintético de cadeia dupla que se mostra na Figura 6. Na Tabela 3 mostram-se os polinucleótidos, de cadeia simples, individuais (N_1 - N_{12}).

Os polinucleótidos 2, 3, 9-4', 11, 10-5', 6, 7 e 8 foram quinaseados adicionando 1 μ l de cada polinucleótido (0,1 μ g/ μ l) e 20 unidades de polinucleótido-quinase T_4 a uma solução contendo 70 mM de Tris-HCl, a pH 7,6, 10 mM de $MgCl_2$, 5 mM de DTT, 10 mM de 2 ME e 500 microgramas por ml de BSA. A solução foi mantida a 37°C durante 30 minutos e a reacção foi parada mantendo a solução a 65°C durante 10 minutos. Adicionaram-se os dois polinucleótidos restantes, 20 ng de polinucleótidos N_1 e de polinucleótidos N_{12} , à solução reaccional de quinase, conjuntamente com um volume 1/10 de uma solução contendo 20,0 mM de Tris-HCl, a pH 7,4, 2,0 mM de

MgCl₂ e 50,0 mM de NaCl. Esta solução foi aquecida até 70°C durante 5 minutos e deixou-se arrefecer até a temperatura ambiente, aproximadamente, 25°C, durante 1,5 horas num copo de 500 ml de água. Durante este período de tempo todos os 10 polinucleótidos se ligaram para formar o inserido de ADN sintético, de cadeia dupla, que se mostra na Figura 6A. Os polinucleótidos individuais foram ligados covalentemente uns aos outros, para estabilizar o inserido de ADN sintético, adicionando 40 µl da mistura reaccional anterior a uma solução contendo 50 mM de Tris-HCl, a pH 7,5, 7 mM de MgCl₂, 1 mM de DTT, 1 mM de trifosfato de adenosina ATP e 10 unidades de ADN-ligase T4. Esta solução foi mantida a 37°C durante 30 minutos e, depois, a ADN-ligase T4 foi desactivada mantendo a solução a 65°C durante 10 minutos. Os polinucleótidos restantes foram quinaseados misturando 52 µl da mistura reaccional anterior, 4 µl de uma solução contendo 10 mM de ATP e 5 unidades de polinucleótido-quinase T4. Esta solução foi mantida a 37°C durante 30 minutos e, depois, a polinucleótido-quinase T4 foi desactivada mantendo a solução a 65°C durante 10 minutos. O inserido de ADN sintético, completo, foi ligado directamente num vector lambda Zap II que tinha sido previamente digerido com as enzimas de restrição Not I e Xho I. A mistura de ligação foi empacotada, de acordo com as instruções do fabricante, usando extracto de empacotamento Gigapack II Gold disponível dos Stratagene Cloning Systems, La Jolla, CA, E.U.A.. A mistura de ligação empacotada foi colocada em placa em células Azuis XL1 (Stratagene Cloning Systems, San Diego, CA, E.U.A.). As placas de lambda Zap II individuais foram nucleadas e os inseridos foram excisados de acordo com o protocolo de excisão in vivo fornecido pelo fabricante, Stratagene Cloning Systems, La Jolla, CA, E.U.A.. Este protocolo de excisão in vivo move o inserido clonado do vector lambda Zap II para um vector de plasmideo para permitir a fácil manipulação e sequenciação. A exactidão dos passos de clonagem anteriores foi confirmada por sequenciação do inserido usando o método do didesoxido de Sanger descrito em Sanger et al., Proc. Natl. Acad. Sci. USA, 74:5463-5467, (1977) e usando as instruções do fabricante no conjunto de sequenciação de ³⁵S-ATP de transcriptase inversa, AMV, de

Stratagene, Cloning Systems, La Jolla, CA, E.U.A.. A sequência do vector de expressão de V_H resultante é mostrada na Figura 6A e na Figura 7.

Tabela 3

N1)	5' GGCCGCAAATTCTATTTCAAGGAGACAGTCAT 3'
N2)	5' AATGAAATACCTATTGCCTACGGCAGCCGCTGGATT 3'
N3)	5' GTTATTACTCGCTGCCCAACCAGCCATGGCCC 3'
N4)	5' AGGTGAAACTGCTCGAGAATTCTAGACTAGGTTAATAG 3'
N5)	5' TCGACTATTAAGTAGTCTAGAATTCTCGAG 3'
N6)	5' CAGTTTCACCTGGGCCATGGCTGGTTGGG 3'
N7)	5' CAGCGAGTAATAACAATCCAGCGGCTGCCGTAGGCAATAG 3'
N8)	5' GTATTTTATTATGACTGTCTCCTTGAAATAGAATTTGC 3'
N9-4)	5' AGGTGAAACTGCTCGAGATTTCTAGACTAGTTACCCGTAC 3'
N11)	5' GACGTTCCGGACTACGGTTCTTAATAGAATTTCG 3'
N12)	5' TCGACGAATTCTATTAAGAACCGTAGTC 3'
N10-5)	5' CGGAACGTCGTACGGGTAAGTAGTCTAGAAATCTCGAG 3'

10. Construção do vector de expressão de V_L

Para exprimir a pluralidade de polinucleótidos codificando um V_L numa célula hospedeira de E. coli, construiu-se um vector, que colocou o polinucleótido codificando o V_L na cadeia de leitura apropriado, forneceu um sitio de ligação a ribossomas tal como se descreve em Shine et al., Nature, 254:34, (1975), forneceu uma sequência líder que dirige a proteína expressa para o espaço periplasmático e forneceu também um polinucleótido, que codificava uma proteína espaçadora, entre o polinucleótido de V_L e o polinucleótido codificando o epitopo tag. Construiu-se uma sequência de ADN sintética contendo todos os polinucleótidos e as características anteriores, concebendo segmentos de polinucleótidos de cadeia simples com 20-40 bases que se hibridam uns com os outros e que formam a sequência de ADN sintético, de cadeia dupla, que se mostra na Figura 6B. Na Tabela 3 mostram-se os polinucleótidos, de cadeia simples, individuais (N₁-N₈).

Os polinucleótidos N2, N3, N4, N6, N7 e N8 foram quinaseados adicionando 1 μ l de cada polinucleótido e 20 unidades de

polinucleótido-quinase T₄ a uma solução contendo 70 mM de Tris-HCl, a pH 7,6, 10 mM de MgCl₂, 5 mM de DTT, 10 mM de 2 ME e 500 microgramas por ml de BSA. A solução foi mantida a 37°C durante 30 minutos e a reacção foi parada mantendo a solução a 65°C durante 10 minutos. Os dois polinucleótidos restantes, 20 ng de polinucleótidos N1 e de polinucleótidos N5, foram adicionados à solução reaccional de quinase, conjuntamente com um volume 1/10 de uma solução contendo 20,0 mM de Tris-HCl, a pH 7,4, 2,0 mM de MgCl₂ e 50,0 mM de NaCl. Esta solução foi aquecida a 70°C durante 5 minutos e deixou-se arrefecer até à temperatura ambiente, aproximadamente 25°C, durante 1,5 horas, num copo de 500 ml de água. Durante este período de tempo todos os polinucleótidos se ligaram para formar o inserido de ADN sintético de cadeia dupla. Os polinucleótidos individuais foram ligados covalentemente, uns aos outros, para estabilizar o inserido de ADN sintético, adicionando 40 µl da mistura reaccional anterior a uma solução contendo 50 µM de Tris-HCl, a pH 7,5, 7 mM de MgCl₂, 1 mM de DTT, 1 mM de ATP e 10 unidades de ADN ligase T₄. Esta solução foi mantida a 37°C durante 30 minutos e, depois, a ADN-ligase T₄ foi desactivada mantendo a solução a 65°C durante 10 minutos. Os polinucleótidos restantes foram quinasados misturando 52 µl da mistura reaccional anterior, 4 µl de uma solução contendo 10 mM de ATP e 5 unidades de polinucleótido-quinase T₄. Esta solução foi mantida a 37°C durante 30 minutos e, depois, a polinucleótido-quinase T₄ foi desactivada mantendo a solução a 65°C durante 10 minutos. O inserido de ADN sintético, completo, foi ligado directamente num vector lambda Zap II que tinha sido previamente digerido com as enzimas de restrição Not I e Xho I. A mistura de ligação foi empacotada, de acordo com as instruções do fabricante, usando um extracto de empacotamento Gigapack II Gold, disponível no Stratagene Cloning Systems, La Jolla, CA, E.U.A.. A mistura de ligação empacotada foi colocada em placa em células XL1 Azuis (Stratagene Cloning Systems, La Jolla, CA, E.U.A.). As placas de lambda Zap II individuais foram nucleadas e os inseridos foram excisados de acordo com o protocolo de excisão, in vivo, fornecido pelo fabricante, Strategene Cloning Systems, La Jolla, CA, E.U.A., e descrito em Short et al., Nucleic Acids

Res., 16:7583-7600, 1988. Este protocolo de excisão in vivo move o inserido clonado do vector lambda Zap II para um vector de fagomideo, para permitir a fácil manipulação e sequenciação, e produz também a versão de fagomideo dos vectores de expressão de V_L . A exactidão dos passos de clonação anteriores foi confirmada por sequenciação do inserido usando o método do didesóxido de Sanger descrito por Sanger et al., Proc. Natl. Acad. Sci. USA, 74:5463-5467, (1977) e usando as instruções do fabricante no conjunto de sequenciação, de ^{35}S -dATP de transcriptase inversa, AMV, de Stratagene Cloning Systems, La Jolla, CA, E.U.A.. Na Figura 6 e na Figura 8 mostra-se a sequência do vector de expressão de V_L resultante.

O vector de expressão de V_L , usado para construir a biblioteca de V_L , foi o fagomideo, produzido para permitir a determinação do ADN do vector de expressão de V_L . O fagomideo foi produzido, como acima se descreve, detalhadamente, pelo processo de excisão in vivo a partir do vector de expressão Lambda Zap V_L (Figura 8). A versão de fagomideo deste vector foi usada devido ao sitio de enzima de restrição Nco I ser único, nesta versão, e poder, assim, ser usado para ligar operativamente os homólogos de ADN de V_L no vector de expressão.

11. Construção do vector de expressão V_L II

Para exprimir a pluralidade de homólogos de ADN, codificando V_L , numa célula de hospedeiro de E. coli, construiu-se um vector que colocou os homólogos de ADN, codificando V_L , na cadeia de leitura apropriada, forneceu um sitio de ligação a ribossomas, como é descrito por Shine et al., Nature, 254:34, 1975, forneceu a sequência líder do gene Pel B, que tinha sido usada previamente com sucesso para segregar fragmentos Fab em E. coli por Lei et al., J. Bac., 169:4379 (1987) e Better et al., Science, 240:1041 (1988), e forneceu, também, um polinucleótido contendo um sitio de endonuclease de restrição para clonação. Construiu-se uma sequência de ADN sintético, contendo todos os polinucleótidos e características anteriores, concebendo segmentos de polinucleótidos de cadeia simples de 20-60 bases que se hibridam



uns com os outros e formam a sequência de ADN sintética de cadeia dupla que se mostra na Figura 10. Na Tabela 4 mostra-se a sequência de cada um dos polinucleótidos, de cadeia simples individuais (01-08) na sequência de ADN sintético de cadeia dupla.

Os polinucleótidos 02, 03, 04, 05, 06 e 07 foram quinasados adicionando 1 μ l (0,1 μ g/ μ l) de cada polinucleótido e 20 unidades de polinucleótido-quinase T4 a uma solução contendo 70 mM de Tris-HCl, a pH 7,6, 10 mM de cloreto de magnésio ($MgCl_2$), 5 mM de ditionitrotoil (DTT), 10 mM de 2-mercaptoetanol (2 ME) e 500 microgramas por ml de albumina de soro de bovino. A solução foi mantida a 37°C durante 30 minutos e a reacção foi parada mantendo a solução a 65°C durante 10 minutos. Adicionaram-se 20 ng de cada um dos dois polinucleótidos restantes, 01 e 08, à solução reaccional de quinase anterior, conjuntamente com um volume 1/10 de uma solução contendo 20,0 mM de Tris-HCl, a pH 7,4, 2,0 mM de $MgCl_2$ e 15,0 mM de cloreto de sódio (NaCl). Esta solução foi aquecida até 70°C durante 5 minutos e deixou-se arrefecer até à temperatura ambiente, aproximadamente 25°C, durante 1,5 horas num copo de 500 ml de água. Durante este periodo de tempo todos os 8 polinucleótidos se ligaram para formar o inserido de ADN sintético, que se mostra na Figura 9. Os polinucleótidos individuais foram ligados, covalentemente, uns aos outros, para estabilizar o inserido de ADN sintético, adicionando 40 μ l da mistura reaccional anterior a uma solução contendo 50 mM de Tris-HCl, a pH 7,5, 7 mM de $MgCl_2$, 1 mM de DTT, 1 mM de ATP e 10 unidades de ADN-ligase T4. Esta solução foi mantida a 37°C durante 30 minutos e, depois, desactivou-se a ADN-ligase T4 mantendo a solução a 65°C durante 10 minutos. Os polipéptidos restantes foram quinasados misturando 52 μ l da mistura reaccional anterior, 4 μ l de uma solução contendo 10 mM de ATP e 5 unidades de polinucleótido-quinase T4. Esta solução foi mantida a 37°C durante 30 minutos e, depois, a polinucleótido-quinase T4 foi desactivada mantendo a solução a 65°C durante 10 minutos. O inserido de ADN sintético, completo, foi ligado directamente num vector Lambda Zap II que tinha sido previamente digerido com os



enzimas de restrição Not I e Xho I. A mistura de ligação foi empacotada, de acordo com as instruções do fabricante, usando um extracto de empacotamento Gigapack II Gold, disponível da Stratagene Cloning Systems, La Jolla, CA, E.U.A.. A mistura de ligação empacotada foi colocada em placa em células Azuis XL1 (Stratagene Cloning Systems, San Diego, CA, E.U.A.). As placas Lambda Zap II individuais foram nucleadas e os inseridos foram excisados de acordo com o protocolo de excisão in vivo, fornecido pelo fabricante, Stratagene Cloning Systems, La Jolla, CA, E.U.A.. Este protocolo de excisão in vivo move o inserido clonado do vector Lambda Zap II para um vector de plasmideo para permitir a fácil manipulação e sequenciação. A exactidão dos passos de clonação anteriores foi confirmada pela sequenciação do inserido, usando as instruções do fabricante, no conjunto de sequenciação de ^{35}S -dATP de transcriptase inversa, AMV, da Stratagene Cloning Systems, La Jolla, CA, E.U.A., Na Figura 9 e na Figura 11 mostra-se a sequência do vector de expressão de V_L II, resultante.

TABELA 4

- | | | | |
|-----|----|---|----|
| 01) | 5' | TGAATTCTAACTAGTCGCCAAGGAGACAGTCAT | 3' |
| 02) | 5' | AATGAAATACCTATTGCCTACGGCAGCCGCTGGATT | 3' |
| 03) | 5' | GTTATTACTCGCTGCCCAACCAGCCATGGCC | 3' |
| 04) | 5' | GAGCTCGTCAGTTCTAGAGTTAAGCGGCCG | 3' |
| 05) | 5' | GTATTTTATTATGACTGTCTCCTTGGCGACTAGTTTAGAA-
TTCAAGCT | 3' |
| 06) | 5' | CAGCGAGTAATAACAATCCAGCGGCTGCCGTAGGCAATAG | 3' |
| 07) | 5' | TGACGAGCTCGGCCATGGCTGGTTGGG | 3' |
| 08) | 5' | TCGACGGCCGCTTAACTCTAGAAC | 3' |

12. Construção da Biblioteca de $V_H + V_L$

Para preparar uma biblioteca de expressão enriquecida nas sequências de V_H , prepararam-se homólogos de ADN enriquecidos em sequências de V_H , de acordo com o Exemplo 6, usando o mesmo conjunto de iniciadores 5' mas com o iniciador 12A (Tabela 1) como iniciador 3'. Estes homólogos foram depois digeridos com os enzimas de restrição Xho I e Spe I e purificados num gel de



agarose a 1%, usando a técnica de electroeluição padrão descrita em Molecular Cloning A Laboratory Manual, Maniatis et al., eds., Cold Spring Harbor, New York, E.U.A. (1982). Estes homólogos de ADN de V_H preparados foram depois inseridos directamente no vector de expressão de V_H que tinha sido previamente digerido com Xho I e Spe I.

A mistura de ligação contendo os homólogos de ADN de V_H foi empacotada de acordo com as especificações do fabricante, usando o Extracto de Empacotamento Gigapack Gold II (Stratagene Cloning Systems, La Jolla, CA, E.U.A.). As bibliotecas de expressão estavam então prontas para serem colocadas em placa em células XL1 Azuis.

Para preparar uma biblioteca enriquecida em sequências de V_L , prepararam-se produtos, amplificados por PCR, enriquecidos em sequências de V_L , de acordo com o Exemplo 6. Estes homólogos de ADN de V_L foram digeridos com os enzimas de restrição Nco I e Spe I. Os homólogos de ADN de V_L digeridos foram purificados num gel de agarose a 1% usando técnicas de electroeluição convencionais descritas em Molecular Cloning A Laboratory Manual, Maniatis et al., eds., Cold Spring Harbor, NY, E.U.A. (1982). Os homólogos de ADN de V_L preparados foram inseridos, directamente, no vector de expressão de V_L que tinha sido previamente digerido com os enzimas de restrição Nco I e Spe I. Usou-se a mistura de ligação, contendo os homólogos de ADN de V_L , para transformar células competentes XL-1 azuis usando as instruções do fabricante (Stratagene Cloning Systems, La Jolla, CA, E.U.A.).

13. Inserção dos homólogos de ADN codificando V_L no vector de expressão de V_L

Na preparação, para clonação, de uma biblioteca enriquecida em sequências de V_L , digeriram-se os produtos amplificados por PCR (2,5 µg/30 µl de meio contendo 150 mM de NaCl, 8 mM de Tris-HCl (pH 7,5), 6 mM de MgSO₄, 1 mM de DTT, 200 µg/ml de BSA) a 37°C, com os enzimas de restrição Sac I (125 unidades) e Xba I (125 unidades) e purificaram-se num gel de agarose a 1%. Nas experiências de clonação que requereram uma mistura dos produtos



das reacções de amplificação, combinaram-se volumes iguais (50 μ l, concentração de 1-10 μ g) de cada mistura reaccional, após a amplificação mas antes da digestão de restrição. Após electroforese em gel do ARNm de baço, amplificado por PCR, digerido, a região do gel, contendo fragmentos de ADN de aproximadamente 350 pb, foi excisada, electroeluida para uma membrana de diálise, precipitada com etanol e ressuspendida numa solução de TE contendo 10 mM de Tris-HCl, pH 7,5, e 1 mM de EDTA, até uma concentração final de 50 ng/ μ l.

O vector de ADN de expressão de V_L II foi preparado para clonação, misturando 100 μ g deste ADN, com uma solução contendo 250 unidades de cada uma das endonucleases de restrição Sac I e Xba I (ambas da Boehringer Mannheim, Indianapolis, IN, E.U.A.) e um tampão recomendado pelo fabricante.

Esta solução foi mantida a 37°C durante 1,5 horas. A solução foi aquecida a 65°C, durante 15 minutos. para desactivar as endonucleases de restrição. A solução foi arrefecida até 30°C e misturaram-se, com a solução, 25 unidades de fosfatase (HK) desactivável pelo calor (Epicenter, Madison, WI) e $CaCl_2$, de acordo com as especificações do fabricante. Esta solução foi mantida a 30°C durante 1 hora. O ADN foi purificado, extraíndo a solução com uma mistura de fenol e de clorofórmio seguindo-se a precipitação com etanol. O vector de expressão V_L II estava agora pronto para ligação aos homólogos de ADN de V_L , preparados nos exemplos anteriores.

Os homólogos de ADN enriquecidos em sequências V_L , foram preparados de acordo com o Exemplo 5 mas usando um iniciador 5' de cadeia leve e um iniciador 3' de cadeia leve, que se mostram na Tabela 2. As reacções de amplificação individuais foram realizadas usando cada um dos iniciadores 5' de cadeia leve em combinação com o iniciador 3' de cadeia leve. Estas misturas reaccionais separadas, contendo homólogo de V_L , foram misturadas e digeridas com as endonucleases de restrição Sac I e Xba I, de acordo com o Exemplo 6. Os homólogos de V_L foram purificados, num gel de agarose a 1% usando a técnica de electroeluição

convencional descrito em Molecular Cloning A Laboratory Manual, Maniatis et al., eds., Cold Spring Harbor, New York, E.U.A. (1982). Estes homólogos de ADN de V_L preparados foram depois inseridos, directamente, no vector de expressão V_L II, clivado com Sac I - Xba I, que foi preparado anteriormente, ligando 3 moles de inseridos de homólogo de ADN de V_L , com cada mole do vector de expressão V_L II, de um dia para o outro, a 50°C. Obtiveram-se $3,0 \times 10^5$ unidades formando placas, após empacotamento do ADN com Gigapack II Bold (Stratagene Cloning Systems, La Jolla, CA, E.U.A.) e 50% eram recombinantes.

14. Combinação aleatória dos homólogos de ADN de V_H e de V_L , no mesmo vector de expressão

A biblioteca de expressão de V_L II, preparada no Exemplo 13, foi amplificada e prepararam-se 500 µg de ADN de fago, da biblioteca de expressão de V_L II, a partir da reserva de fago amplificado, usando os procedimentos descritos em Molecular Cloning: A Laboratory Manual, Maniatis et al., eds., Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, E.U.A. (1982). Mantiveram-se 50 µg deste ADN de fago, da biblioteca de expressão de V_L II, numa solução contendo 100 unidades da endonuclease de restrição MluI (Boehringer Mannheim, Indianapolis, IN, E.U.A.) em 200 µl de um tampão fornecido pelo fabricante de endonuclease, durante 1,5 horas a 37°C. A solução foi depois extraída com uma mistura de fenol e de clorofórmio. O ADN foi depois precipitado em etanol e re-suspendido em 100 µl de água. Misturou-se esta solução com 100 unidades da endonuclease de restrição EcoR I (Boehringer Mannheim, Indianapolis, IN, E.U.A.), num volume final de 200 µl de tampão, contendo os componentes especificados pelo fabricante. Esta solução foi mantida a 37°C durante 1,5 horas e a solução foi depois extraída com uma mistura de fenol e de clorofórmio. O ADN foi precipitado em etanol e re-suspendido em TE.

A biblioteca de expressão de V_H preparada no Exemplo 12 foi amplificada e prepararam-se 500 µg de ADN de fago, da biblioteca de expressão de V_H , usando os métodos descritos acima.



Mantiveram-se 50 µg do ADN de fago, da biblioteca de expressão de V_H , numa solução contendo 100 unidades da endonuclease de restrição Hind III (Boehringer Mannheim, Indianapolis, IN, E.U.A.), em 200 µl de um tampão, fornecido pelo fabricante de endonuclease, durante 1,5 horas a 37°C. A solução foi depois extraída com uma mistura de fenol e de clorofórmio saturada com Tris-HCl 0,1 M a pH 7,5. O ADN foi depois precipitado em etanol e re-suspendido em 100 µl de água. Esta solução foi misturada com 100 unidades da endonuclease de restrição EcoR I (Boehringer Mannheim, Indianapolis, IN, E.U.A.) até um volume final de 200 µl de tampão, contendo os componentes especificados pelo fabricante. Esta solução foi mantida a 37°C durante 1,5 horas e a solução foi depois extraída com uma mistura de fenol e de clorofórmio. O ADN foi precipitado em etanol e re-suspendido em TE.

Ligaram-se as bibliotecas de expressão de V_H e de V_L II, digeridas com endonuclease de restrição. A mistura reaccional de ligação consistia de 1 µg de V_H e de 1 µg de ADN de biblioteca de fago de V_L II, num meio reaccional de 10 µl, usando os reagentes fornecidos num conjunto experimental de ligação adquirido na Stratagene Cloning Systems (La Jolla, Califórnia, E.U.A.). Após ligação durante 16 h a 4°C, empacotou-se 1 µl do ADN de fago, ligado, com extracto de empacotamento Gigapack Gold II e colocou-se em placa em células XL 1 Azuis preparadas de acordo com as instruções do fabricante. Usou-se uma porção dos 3×10^6 clones obtidos, para determinar a eficácia da combinação. Na Figura 11 mostra-se o vector de expressão de V_H e V_L resultante.

Os clones contendo ambos os vectores de V_H e V_L foram excisados do fago para o pBluescript usando o protocolo de excisão in vitro descrito por Short et al., Nucleic Acid Research, 16:7583-7600 (1988). Os clones escolhidos para excisão exprimiram o decapeptido tag e não clivaram o X-gal na presença de 2 mM de IPTG, permanecendo assim brancos. Os clones com estas características representavam 30% da biblioteca. Tal como se determinou por análise de restrição, 50% dos clones escolhidos para excisão continham um V_H e um V_L . Uma vez que aproximadamente 30% dos clones na biblioteca de V_H exprimiram o decapeptido tag e

50% dos clones na biblioteca de V_L II continham uma sequência V_L , concluiu-se que não mais do que 15% dos clones na biblioteca combinada conteriam ambos os clones V_H e V_L . O número real obtido foi de 15% da biblioteca, indicando que o processo de combinação era muito eficiente.

15. Segregação dos homólogos de ADN de uma proteína de ligação a um antígeno V_H

Para segregar os clones individuais contendo os homólogos de ADN, que codificam uma proteína de ligação a um antígeno V_H , determinou-se o título da biblioteca de expressão de V_H , preparada de acordo com o Exemplo 11. Esta titulação da biblioteca foi realizada usando métodos bem conhecidos dos peritos na arte. Resumidamente, realizaram-se séries de diluições da biblioteca para um tampão contendo 100 mM de NaCl, 50 mM de Tris-HCl, a pH 7,5, e 10 mM de $MgSO_4$. Adicionaram-se 10 μ l de cada diluição a 200 μ l de células de E. coli em crescimento exponencial e mantiveram-se a 37°C durante 15 minutos, para permitir a absorção do fago pelas células bacterianas, com 3 ml de ágar de topo consistindo de 5 g/l de NaCl, 2 g/l de $MgSO_4$, 5 g/l de extracto de levedura, 10 g/l de NZ amina (hidrolisado de caseína) e 0,7% de agarose fundida a 50°C. O fago, as bactérias e o ágar de topo, foram misturados e depois distribuídos equitativamente sobre a superfície de um prato de ágar bacteriano pré-aquecido (5 g/l de NaCl, 2 g/l de $MgSO_4$, 5 g/l de extracto de levedura, 10 g/l de NZ amina (hidrolisado de caseína) e 15 g/l de ágar Difco). Os pratos foram mantidos a 37°C, durante 12 a 24 horas, período de tempo durante o qual as placas lambda se desenvolveram na camada bacteriana. As placas lambda foram contadas para determinar o número total de unidades formando placas por ml, na biblioteca original.

A biblioteca de expressão titulada foi depois colocada em prato de modo a prepararem-se filtros réplicas da biblioteca. Os filtros réplicas serão depois usados para segregar os clones individuais na biblioteca que exprimem as proteínas de ligação a antígenos de interesse. Resumidamente, adicionou-se um volume da



biblioteca titulada, de modo a terem-se 20 000 placas por prato de 150 mililitros, a 600 µl de células de E. coli em crescimento exponencial, e manteve-se a 37°C durante 15 minutos, para permitir a absorção do fago nas células bacterianas. Depois, misturaram-se 7,5 ml de ágar de topo com a solução contendo as células bacterianas e distribui-se, equitativamente, o fago absorvido e toda a mistura, sobre a superfície de um prato de ágar bacteriano pré-aquecido. Este processo foi repetido, um número de vezes suficiente, para colocar em prato um número total de placas pelo menos igual à dimensão da biblioteca. Estes pratos foram depois mantidos a 37°C durante 5 horas. Estes pratos foram depois sobrepostos com filtros de nitrocelulose, que tinham sido pré-tratados com uma solução contendo 10 mM de isopropil-beta-D-tiogalactopiranosido (IPTG) e mantiveram-se a 37°C durante 4 horas. A orientação dos filtros de nitrocelulose em relação ao prato foi marcada fazendo um buraco com uma agulha, mergulhada em tinta à prova de água, através do filtro e nos pratos de bactérias em vários pontos. Os filtros de nitrocelulose foram removidos com fórceps e lavados uma vez com uma solução de TBST contendo 20 mM de Tris-HCl, a pH 7,5, 150 mM de NaCl e 0,05% de monolaurato de polioxietileno-sorbitano (Tween 20). Um segundo filtro, que tinha também sido embebido numa solução contendo 10 mM de IPTG, foi reaplicado nos pratos de bactérias para produzir filtros duplicados. Os filtros foram, adicionalmente, lavados com uma solução fresca de TBST, durante 15 minutos. Os filtros foram depois colocados numa solução bloqueante, consistindo de 20 mM de Tris-HCl a pH 7,5, 150 mM de NaCl e 1% de BSA e agitaram-se, durante 1 hora, à temperatura ambiente. Os filtros de nitrocelulose foram transferidos para uma solução bloqueante fresca, contendo uma diluição 1 para 500 do anticorpo primário, e agitou-se suavemente, durante pelo menos 1 hora, à temperatura ambiente. Depois, os filtros foram agitados na solução contendo o anticorpo primário, os filtros foram lavados 3 a 5 vezes em TBST, durante 5 minutos de cada vez, para remover o anticorpo primário residual não ligado. Os filtros foram transferidos para uma solução contendo solução bloqueante fresca e uma diluição de 1 para 500 a 1 para 1 000 de anticorpo secundário conjugado com

fosfatase alcalina. Os filtros foram agitados suavemente, na solução, durante pelo menos 1 hora, à temperatura ambiente. Os filtros foram lavados 3 a 4 vezes numa solução de TBST, durante pelo menos 5 minutos de cada vez, para remover qualquer anticorpo secundário residual não ligado. Os filtros foram lavados uma vez numa solução contendo 20 mM de Tris-HCl, a pH 7,5, e 150 mM de NaCl. Os filtros foram removidos desta solução e o excesso de humidade foi removido com papel de filtro. A cor foi desenvolvida colocando o filtro numa solução contendo 100 mM de Tris-HCl, a pH 9,5, 100 mM de NaCl, 5 mM de MgCl₂, 0,3 mg/ml de nitro Azul de Tetrazólio (NBT) e 0,15 mg/ml fosfato de 5-bromo-4-cloro-3-indolilo (BCIP) durante, pelo menos, 30 minutos à temperatura ambiente. A solução de desenvolvimento de cor residual, foi removida por lavagem do filtro com uma solução contendo 20 mM de Tris-HCl, a pH 7,5, e 150 mM de NaCl. O filtro foi depois colocado numa solução de paragem consistindo de 20 mM de Tris-HCl, a pH 2,9, e 1 mM de EDTA. O desenvolvimento de uma cor púrpura intensa indica um resultado positivo. Os filtros são usados para localizar a placa de fago que produziu a proteína desejada. Essa placa de fago é segragada e depois cultivada para análise adicional.

Usaram-se várias combinações, diferentes, de anticorpos primários e de anticorpos secundários. A primeira combinação com um anticorpo primário imuno específico de um decapeptido, que apenas será expresso se a proteína de ligação ao antígeno V_H for expressa na cadeia de leitura apropriada, para permitir a leitura através da tradução, para incluir o epítipo de decapeptido ligado covalentemente à proteína de ligação ao antígeno V_H. Este epítipo de decapeptido e um anticorpo imuno específico deste epítipo de decapeptido foram descritos por Green et al., *Cell* 28:477 (1982) e por Niman et al., *Proc. Nat. Acad. Sci. U.S.A.* 80:4949 (1983). Na Figura 11 mostra-se a sequência do decapeptido reconhecido. Um equivalente funcional do anticorpo monoclonal, que é imuno específico do decapeptido, pode ser preparado de acordo com os métodos de Green et al., e de Niemann et al.. O anticorpo secundário usado com este anticorpo primário era uma

IgG de cabra anti-ratinho (Fisher Scientific). Este anticorpo é imuno-específico da região constante da IgG de ratinho e não reconhece qualquer porção da região variável de cadeia pesada. Esta combinação particular de anticorpos primários e secundários, quando usada de acordo com o protocolo anterior, determinou que entre 25% e 30% dos clones exprimiam o decapeptídeo e, deste modo, assumiu-se que estes clones exprimiam também uma proteína de ligação ao antígeno V_H .

Noutra combinação usou-se o anticorpo monoclonal de ratinho anti-decapeptídeo, como anticorpo primário e, como anticorpo secundário usou-se uma Ig de cabra anti-ratinho, purificada por afinidade, disponível comercialmente como parte de um conjunto de imuno-ensaio picoBlue da Stratagene Cloning Systems, La Jolla, CA, E.U.A.. Esta combinação resultou num grande número de clones falsos positivos porque o anticorpo secundário imunorreagiu também com o V_H da cadeia pesada. Deste modo, este anticorpo reagiu com todos os clones, exprimindo qualquer proteína V_H , e esta combinação de anticorpos primários e secundários não detectou especificamente clones com o polinucleótido de V_H na cadeia de leitura apropriada, permitindo assim a expressão do decapeptídeo.

Usam-se várias combinações de anticorpos primários e secundários, onde o anticorpo primário está conjugado com isotiocianato de fluoresceína (FITS) e assim a imuno-especificidade do anticorpo não é importante, porque o anticorpo é conjugado com o antígeno pré-seleccionado (FITC) e é esse antígeno que se deverá ligar às proteínas de ligação ao antígeno V_H produzidas pelos clones na biblioteca de expressão. Após este anticorpo primário ter sido ligado em virtude de ser o anticorpo monoclonal de ratinho p2 5764 (ATCC # HB-9505) conjugado com FITC. O anticorpo secundário usado com este anticorpo primário é uma IgG de cabra anti-ratinho (Fisher Scientific, Pittsburg, PA, E.U.A.) conjugado com fosfatase alcalina, usando o método descrito em Antibodies A Laboratory Manual, Harlow and Love, eds., Cold Spring Harbor, New York, E.U.A. (1988). Se um clone particular, na biblioteca de expressão de V_H , exprime uma

proteína de ligação de V_H que se liga ao FITC acoplado, de um modo covalente ao anticorpo primário, o anticorpo secundário liga-se especificamente e, quando desenvolvida, a fosfatase alcalina provoca a formação de uma cor púrpura distinta.

A segunda combinação de anticorpos do tipo usa um anticorpo primário que é IgG de coelho anti-humana (Fisher Scientific Pittsburg, PA) conjugada com FITC. O anticorpo secundário, usado com este anticorpo primário é uma IgG de cabra anti-coelho conjugada, com fosfatase alcalina, usando os métodos descritos em Antibodies A Laboratory Manual, Harlow and Lane, eds., Cold Spring Harbor, New York, E.U.A. (1988). Se um clone particular na biblioteca de expressão de V_H , exprime uma proteína de ligação de V_H que se liga ao FITC conjugado com o anticorpo primário, o anticorpo secundário liga-se, especificamente e, quando desenvolvida, a fosfatase alcalina provoca a formação de uma cor púrpura distinta.

Outro anticorpo primário foi o anticorpo monoclonal de ratinho p2 5764 (ATCC # HB-9505) conjugado com FITC e ^{125}I . O anticorpo será ligado, por quaisquer das proteínas de ligação ao antígeno V_H , expressas. Depois, uma vez que o anticorpo está também marcado com ^{125}I , faz-se um autorradiograma do filtro, em vez de usar um anticorpo secundário que está conjugado com fosfatase alcalina. Esta produção directa de um autorradiograma permite a segregação dos clones na biblioteca que exprimem uma proteína, de ligação a um antígeno V_H , de interesse.

16. Segregação de homólogos de ADN para um V_H e para um V_L que formam F_V de ligação a antígeno

Para segregar os clones individuais contendo os homólogos de ADN que codificam um V_H e um V_L que formam um F_V que se liga a um antígeno, titulou-se a biblioteca de expressão de V_H e de V_L , de acordo com o Exemplo 15. A biblioteca de expressão, titulada, foi depois pesquisada quanto à presença do decapeptido tag expresso com o V_H , usando os métodos descritos no Exemplo 15. O ADN foi então preparado, a partir dos clones, para exprimir o decapeptido tag. Este ADN foi digerido com a endonuclease de restrição Pvu



II, para determinar se estes clones continham, também, um homólogo de ADN de V_L . A migração mais lenta de um fragmento de endonuclease de restrição Pvu II indicou que o clone particular continha ambos os homólogos de ADN de V_H e de V_L .

Os clones contendo ambos os homólogos de ADN de V_H e de V_L , foram analisados para determinar se estes clones produziam uma molécula de proteína F_v reunida, a partir dos homólogos de ADN de V_H e de V_L .

O fragmento de proteína F_v , produzido em clones contendo V_H e V_L , foi visualizado por imunoprecipitação de proteína radiomarcada expressa nos clones. Inoculou-se uma cultura de 50 ml de caldo LB (5 g/l de extracto de levedura, 10 g/l de triptona e 10 g/l de NaCl, a pH 7,0), contendo 100 $\mu\text{g}/\mu\text{l}$ de ampicilina, com E. coli compreendendo um plasmídeo contendo um V_H e um V_L . A cultura foi mantida a 37°C, com agitação, até a densidade óptica, medida a 550 nm ser de 0,5. A cultura foi centrifugada a 3 000 x g, durante 10 minutos, e depois ressuspendeu-se em 50 ml de meio M9 (6 g/l de Na_2HPO_4 , 3 g/l de KH_2PO_4 , 0,5 g/l de NaCl, 1 g/l de NH_4Cl , 2 g/l de glucose, 2 mM de MgSO_4 , 0,1 mM de CaCl_2) suplementado com aminoácido sem metionina ou cisteína. Esta solução foi mantida a 37°C durante 5 minutos e depois adicionaram-se 0,5 mCi de ^{35}S , na forma de HSO_4^- (New England Nuclear, Boston, MA, E.U.A.) e a solução foi ainda mantida a 37°C durante 2 horas adicionais. A solução foi depois centrifugada a 3 000 x g e rejeitou-se o sobrenadante. O sedimento de células bacterianas resultante foi congelado e descongelado e, depois, ressuspendeu-se numa solução contendo 40 mM de Tris, pH 8,0, 100 mM de sacarose 1 mM de EDTA. A solução foi centrifugada a 10 000 x g durante 10 minutos e rejeitou-se o sedimento resultante. O sobrenadante foi misturado com 10 μl de anticorpo monoclonal anti-decapeptido e mantido durante 30-90 minutos em gelo. Misturaram-se 40 μl de proteína G, acoplada a pérolas de Sepharose (Pharmacia, Piscataway, NJ, E.U.A.) com a solução, e manteve-se esta solução durante 30 minutos em gel para permitir a formação de um imunoprecipitado. A solução foi centrifugada a 10 000 x g durante 10 minutos e ressuspendeu-se o sedimento

resultante em 1 ml de uma solução contendo 100 mM de Tris-HCl, a pH 7,5, e centrifugou-se a 10 000 x g durante 10 minutos. Repetiu-se este procedimento duas vezes. O sedimento imunoprecipitado resultante foi carregado num gel, PhastGel Homogenous 20 (Pharmacia, Piscataway, NY, E.U.A.), de acordo com as instruções do fabricante. O gel foi seco e usado para a exposição de filme de raios X.

Na Figura 12, mostra-se o autorradiograma resultante. Pode-se observar a presença de moléculas de F_y reunidas pela presença de V_L que foi imunoprecipitado pois estava ligado ao V_H-decapeptido tag reconhecido pelo anticorpo precipitante.

17. Geração de uma grande biblioteca combinatória do reportório de imunoglobulina em fago

Nas Figuras 7 e 9 apresentam-se em diagrama, vectores adequados para expressão de sequências de V_H, V_L, Fv e Fab. Como previamente se discutiu, os vectores foram construídos, por modificações do Lambda Zap, por inserção de oligonucleótidos sintéticos no sítio de clonação múltiplo. Os vectores foram concebidos para serem anti-simétricos em relação aos sítios de restrição Not I e EcoR I que flamqueiam as sequências de clonação e de expressão. Como a seguir se descreve, esta assimetria na colocação de sítios de restrição num vector linear, tal como o bacteriófago, é a característica essencial que permite a combinação de uma biblioteca, exprimindo cadeias leves, com uma exprimindo cadeias pesadas, para construir bibliotecas de expressão de Fab combinatórias. O Lambda Zap II V_L II (Figura 9) é concebido para servir como vector de clonação para fragmentos de cadeia leve e o Lambda Zap II V_H (figura 7) é concebido para servir como vector de clonação, para sequências de cadeia pesada no passo inicial de construção da biblioteca. Estes vectores são construídos para clonarem, eficientemente, os produtos de amplificação por PCR com sítios de restrição específicos incorporados em cada extremidade.

A. Amplificação por PCR de fragmentos de anticorpo

A amplificação por PCR de ARNm isolado de células de baço

com oligonucleótidos, que incorporam sítios de restrição nas extremidades do produto amplificado, pode ser usada para clonar e exprimir sequências de cadeia pesada, incluindo sequências de cadeia F_d e kappa. Os iniciadores de oligonucleótido usados para estas amplificações são apresentados nas Tabelas 1 e 2. Os iniciadores são iguais aos que foram usados com sucesso, no Exemplo 5, para amplificação de sequências V_H . O conjunto de iniciadores 5' para a amplificação da cadeia pesada foi idêntico ao previamente usado para amplificar V_H e os usados para amplificação da cadeia leve foram escolhidos com base em princípios similares, Sastry et al., Proc. Natl. Acad. Sci. USA, 86: 5728 (1989) e Orland et al., Proc. Natl. Acad. Sci. USA, 86:3833 (1989). Os iniciadores 3' únicos das sequências de cadeia pesada (IgG1) e leve (k) foram escolhidos para incluir as cisteínas envolvidas na formação da ligação dissulfureto cadeia pesada - cadeia leve. Nesta fase não se construiu qualquer iniciador para amplificar as cadeias leves lambda, uma vez que constituem apenas uma pequena fracção de anticorpos murinos. Adicionalmente, os fragmentos F_v foram construídos usando um iniciador 3', que é complementar com o ARNm da região J (de reunião) (aminoácido 128), e um conjunto de iniciadores 5' únicos, que são complementares com o ADNc da primeira cadeia na região do terminal N conservada da proteína processada. As sequências de reconhecimento de endonuclease de restrição são incorporadas nos iniciadores, para permitir a clonação do fragmento amplificado num vector de fago lambda, numa cadeia de leitura predeterminada, para expressão.

B. Construção da biblioteca

A construção de uma biblioteca combinatória foi realizada em dois passos. No primeiro passo, as bibliotecas da cadeia pesada e leve foram construídas em Lambda Zap II V_H e Lambda Zap II V_L II, respectivamente. No segundo passo, estas duas bibliotecas foram combinadas nos sítios EcoR I anti-simétricos presentes em cada vector. Este procedimento resultou numa biblioteca de clones cada um dos quais co-exprime potencialmente uma cadeia pesada e uma cadeia leve. As combinações reais são aleatórias e não reflectem,

necessariamente, as combinações presentes na população de células B no animal original. Usou-se o vector de expressão Lambda Zap II V_H para criar uma biblioteca de sequências de cadeia pesada a partir do ADN obtido por amplificação por PCR de ARNm isolado a partir do baço de um ratinho 129 G_{ix} + , previamente imunizado com fosforamido de p-nitrofenilo (NPN), antigénio 1, de acordo com a fórmula I (Figura 13) conjugado com hemocianina de lapa Fissurella (KLH). O conjugado NPN-KLH foi preparado por mistura de 250 μ l de uma solução contendo 2,5 mg de NPN de acordo com a fórmula 1 (Figura 13) em dimetilformamida, com 750 μ l de uma solução contendo 2 mg de KLH em tampão de fosfato de sódio 0,01 M (pH 7,2). As duas soluções foram misturadas por adição lenta da solução de NPN à solução de KLH, mantendo a solução de KLH agitada com uma barra de agitação rotativa. Depois, a mistura foi mantida a 40°C durante 1 hora com a mesma agitação para permitir o prosseguimento da conjugação. O conjugado NPN-KLH foi isolado do NPN e da KLH não conjugados por filtração em gel através de Sephadex G-25. O conjugado de NPN-KLH isolado foi usado em imunizações de ratinho como se descreve no Exemplo 2.

O ARNm de baço, resultante das imunizações anteriores, foi isolado e usado para criar uma biblioteca principal de sequências de gene V_H usando o vector de expressão Lambda Zap II V_H . A biblioteca principal contém $1,3 \times 10^6$ ufp e foi pesquisada quanto à expressão do decapeptido tag, para determinar a percentagem de clones exprimindo sequências Fd. A sequência deste péptido apenas fica em cadeia para expressão após a clonagem de um fragmento Fd (ou V_H) no vector. Pelo menos 80% dos clones na biblioteca exprimem fragmentos Fd, com base na imunodeteção do decapeptido tag.

A biblioteca da cadeia leve foi construída da mesma forma que a da cadeia pesada e verificou-se que contém 2×10^6 membros. Um ensaio das placas, usando um anticorpo anti-cadeia kappa indicou que 60% da biblioteca continha inseridos de cadeia leve expressos. Esta percentagem relativamente pequena de inseridos resultou, provavelmente, da desfosforilação incompleta do vector após clivagem com Sac I e Xba I.

Uma vez obtidas, as duas bibliotecas foram usadas para construir uma biblioteca combinatória, pelo seu cruzamento no sítio EcoR I. Para realizar o cruzamento, em primeiro lugar, purificou-se o ADN de cada biblioteca. A biblioteca de cadeia leve foi clivada com endonuclease de restrição Mlu I, as extremidades 5' resultantes foram desfosforiladas e o produto foi digerido com EcoR I. Este processo clivou o braço esquerdo do vector em vários bocados, mas o braço direito, contendo as sequências de cadeia leve, permaneceu intacto. De um modo paralelo, o ADN da biblioteca de cadeia leve foi clivado com Hind III, desfosforilado e clivado com EcoR I, destruindo o braço direito mas deixando o braço esquerdo, contendo as sequências de cadeia pesada, intacto. Os ADN assim preparados foram então combinados e ligados. Após a ligação, apenas os clones que resultaram da combinação de um braço direito de clones, contendo a cadeia leve, com um braço esquerdo de clones, contendo a cadeia pesada, reconstituíram com fago viável. Após ligação e empacotamento, obtiveram-se $2,5 \times 10^7$ clones. Estes constituem a biblioteca de expressão de Fab combinatória que foi pesquisada para identificar clones possuindo afinidade por NPN. Para determinar a frequência dos clones de fago que co-exprimem os fragmentos de cadeia leve e pesada, testaram-se amostras em duplicado das bibliotecas de cadeia leve, de cadeia pesada e combinatória, como anteriormente, para determinar a expressão da cadeia leve e pesada. Neste estudo de, aproximadamente, 500 fagos recombinantes, aproximadamente 60% co-exprimiram proteínas de cadeia leve e pesada.

C. Ligação de antigénio

Analísaram-se todas as três bibliotecas, de cadeia leve, de cadeia pesada e Fab, para determinar se continham fago recombinante que exprimia fragmentos de anticorpo que se ligavam ao NPN. Num procedimento típico colocaram-se 30 000 fagos em prato e ensaiaram-se amostras em duplicado, com nitrocelulose, para determinar a ligação a NPN acoplado a BSA marcado com ^{125}I (Figura 15). Ensaio em duplicado de 80 000 fagos recombinantes, da biblioteca de cadeia leve, e um número similar da biblioteca

de cadeia pesada, não identificaram quaisquer clones que se ligavam ao antígeno. Em contraste, o ensaio de um número similar de clones da biblioteca de expressão de Fab identificou muitas placas de fago que se ligavam ao NPN (Figura 15). Esta observação indica que, sob condições nas quais muitas cadeias pesadas, em combinação com cadeias leves, se ligam ao antígeno, as mesmas cadeias pesada e leve, sózinhas, não o fazem. Deste modo, no caso do NPN, crê-se que existem muitas cadeias pesadas e leves que apenas se ligam ao antígeno quando estão combinadas, respectivamente, com cadeias leves e pesadas específicas.

Para determinar a capacidade para pesquisar grandes números de clones e para obter uma estimativa mais quantitativa da frequência de clones de ligação a antígeno, na biblioteca combinatória, pesquisaram-se um milhão de placas de fago e identificaram-se aproximadamente 100 clones que se ligavam ao antígeno. Para seis clones que se pensava que se ligavam ao NPN, uma região do prato contendo os positivos e aproximadamente 20 placas de bacteriófago circundantes foi nucleada, colocada de novo em prato e pesquisada com amostras em duplicado (Figura 15). Como se esperava, aproximadamente um em vinte fagos ligava-se, especificamente, ao antígeno. Os "núcleos" das regiões do fago em prato, que se pensava serem negativos, não deram resultados positivos quando foram colocados de novo em placa.

Para determinar a especificidade da interação antígeno anticorpo, a ligação de antígeno competiu com antígeno livre não marcado, como se mostra na Figura 16. Os estudos de competição mostraram que os clones individuais podiam ser distinguidos com base na afinidade pelo antígeno. A concentração de haptenos livres requerida para a inibição completa da ligação variou entre $10-100 \times 10^9$ M sugerindo que os fragmentos Fab expressos tinham constantes de ligação na gama dos nanomolares.

D. Composição dos clones e dos seus produtos expressos

Na preparação para caracterização dos produtos de proteína capazes de se ligarem ao NPN, como se descreve no Exemplo 18C, um plasmideo contendo os genes da cadeia pesada e leve foi excisado



a partir da placa de bacteriofago nucleada apropriada usando fago auxiliar M13mp8. Determinação do mapa do plasmídeo excisado demonstrou um padrão de restrição consistente, com a incorporação de sequências de cadeia pesada e leve. Os produtos de proteína de um dos clones foram analisados por ELISA e por manchas Western, para estabelecer a composição da proteína de ligação ao NPN. Concentrou-se um sobrenadante bacteriano, após a indução por IPTG, e submeteu-se a filtração em gel. Reuniram-se as fracções com peso molecular na gama de 40-60 kD, concentraram-se e submeteram-se a uma separação, por filtração em gel, adicional. Como se ilustra na Figura 17, e análise ELISA das fracções eluentes demonstrou que a ligação de NPN estava associada a uma proteína de peso molecular de cerca de 50 kD, cuja detecção imunológica mostrou que continha ambas as cadeias leve e pesada. Uma mancha Western (que não se mostra) de uma preparação de sobrenadante bacteriano concentrado, sob condições não redutoras, foi desenvolvida com anticorpo anti-decapeptido. Esta revelou uma banda de proteína com um peso molecular de 50 kD. Tomados em conjunto, estes resultados são consistentes com o facto da ligação de NPN ser uma função de fragmentos Fab nos quais as cadeias pesada e leve estão ligadas de um modo covalente.

E. Comparação das propriedades do reportório in vivo versus biblioteca combinatória de fago

Neste exemplo preparou-se uma biblioteca relativamente restrita, porque apenas se usou um número limitado de iniciadores para a amplificação por PCR de sequências Fd. Espera-se que a biblioteca contenha apenas clones exprimindo sequências kappa/gama 1. Contudo, esta não é uma limitação inerente do método, uma vez que se podem adicionar iniciadores adicionais para amplificar qualquer classe ou subclasse de anticorpo. Apesar desta restrição, foi possível isolar um grande número de clones de ligação a antigénios.

Uma questão central que resulta deste trabalho é a de como uma biblioteca de fago, preparada como aqui se descreve, se



compara com o reportório de anticorpos in vivo em termos de tamanho, característicos de diversidade e facilidade de acesso.

O tamanho de reportório de anticorpos de mamífero é difícil de avaliar mas é muitas vezes proposto um número de ordem de 10^6 - 10^8 especificidades de antígeno diferentes. Com algumas das reservas que a seguir se discutem, pode-se construir facilmente uma biblioteca de fago deste tamanho, ou maior, por uma modificação do corrente método. De facto, uma vez construída uma biblioteca combinatória inicial, as cadeias pesadas e leves podem ser rearranjadas, para obter bibliotecas com números excepcionalmente grandes.

Em princípio, espera-se que as características de diversidade do reportório nativo (não imunizado), in vivo e da biblioteca de fago correspondente sejam similares, pelo facto de envolverem ambas uma combinação aleatória de cadeias leves e pesadas. Contudo, existem diferentes factores que actuarão para restringir a diversidade expressa por um reportório in vivo e pela biblioteca de fago. Por exemplo, uma modificação fisiológica tal como a tolerância, restringirá a expressão de certas especificidades antigénicas do reportório in vivo, mas estas especificidades podem ainda aparecer na biblioteca de fago. Por outro lado, uma tendência no processo de clonação pode introduzir restrições na diversidade da biblioteca de fago. Por exemplo, pode-se esperar que a representação de ARN para as sequências expressas por células B estimuladas predomine sobre as de células não estimuladas devido a níveis de expressão mais elevados. Tecidos de diferentes fontes (p.e., sangue periférico, medula óssea ou nódulos linfáticos regionais) e diferentes iniciadores de PCR (p.e., os que se espera que amplifiquem diferentes classes de anticorpos) podem resultar em bibliotecas com características de diversidade diferentes.

Outra diferença entre o reportório in vivo e a biblioteca de fago consiste no facto de os anticorpos, isolados a partir do primeiro, poderem ter sido beneficiados por maturação por afinidade, devido a mutações somáticas após a combinação das



cadeias pesadas e leves, enquanto que o segundo combina, aleatoriamente as cadeias pesadas e leves maturadas. Numa dada biblioteca de fago suficientemente grande, derivada de um reportório in vivo particular, as cadeias leves e pesadas, maturadas, originais, recombinar-se-ão. Contudo, uma vez que um dos potenciais benefícios desta nova tecnologia é o de obviar a necessidade de imunização pela geração de uma única biblioteca de fago "genérica", altamente diversa, seria útil ter métodos para otimizar as sequências para compensar a ausência de mutação somática e de selecção clonal. Pelos processos do presente invento ficam facilmente disponíveis três procedimentos. Em primeiro lugar, a mutagénese de saturação pode ser realizada nas CDR e os Fab resultantes podem ser testados para determinar a função aumentada. Em segundo lugar, podem-se recombinar uma cadeia pesada ou uma cadeia leve dum clone, que se liga a um antigénio, com as bibliotecas, inteiras, de cadeia leve ou pesada, respectivamente, num procedimento idêntico ao usado para construir a biblioteca combinatória. Em terceiro lugar, podem-se realizar ciclos iterativos dos dois procedimentos anteriores para otimizar adicionalmente a afinidade ou as propriedades catalíticas da imunoglobulina. Será de notar que os dois últimos procedimentos não são permitidos na selecção clonal de células B, o que sugere que os métodos aqui descritos podem realmente aumentar a capacidade de identificar sequências óptimas.

O acesso é a terceira área onde tem interesse comparar o reportório de anticorpos in vivo e a biblioteca de fago. Em termos práticos a biblioteca de fago tem um acesso muito mais fácil. Os métodos de ensaio permitem pesquisar, pelo menos, 50 000 clones por prato de modo que, por dia, se podem examinar facilmente 10^6 anticorpos. Este factor, por si só, deverá encorajar a substituição da tecnologia do hibridoma pelos métodos que aqui se descrevem. Os métodos de ensaio mais poderosos utilizam a selecção que pode ser realizada incorporando marcadores seleccionáveis no antigénio, tais como grupos rejeitáveis necessários para a replicação de estirpes bacterianas auxotróficas ou substituintes tóxicos susceptíveis de



desactivação catalítica. Existem também outras vantagens relacionadas com o facto de apenas se poder ter acesso ao reportório de anticorpos in vivo, através de uma imunização que é uma selecção com base na afinidade de ligação. A biblioteca de fago não está restringida de um modo similar. Por exemplo, o único método geral para identificar anticorpos com propriedades catalíticas tem sido por pré-selecção, com base na afinidade do anticorpo por um análogo do estado de transição. Estas restrições não se aplicam à biblioteca in vivo onde a catálise pode, em princípio, ser avaliada directamente. A capacidade para testar directamente grandes números de anticorpos, quanto à sua função, pode permitir a selecção de catalisadores em reacções onde não está bem definido um mecanismo ou onde a síntese do análogo do estado de transição é difícil. A determinação de catálise, directamente, elimina a inclinação do procedimento de pesquisa para certos mecanismos de reacção pejorativos para um análogo sintético e, deste modo, é possível para uma dada transformação química, a exploração simultânea de múltiplos percursos reaccionais.

Os processos aqui apresentados descrevem a geração de fragmentos Fab que são claramente diferentes, num certo número de aspectos, dos anticorpos intactos (inteiros). Existe indubitavelmente uma perda ou afinidade quando se têm ligantes, de antigénio, de Fab, monovalentes, mas isto pode ser compensado por selecção de ligantes adequadamente fortes. Para um certo número de aplicações tais como agentes de diagnóstico e biosensores, pode ser preferível ter fragmentos monovalentes. Para aplicações requerendo funções de efector Fc, já existe a tecnologia para a extensão do gene da cadeia pesada e expressão do anticorpo inteiro glicosilado em células de mamífero.

As ideias aqui apresentadas dirigem-se às questões fulcrais na identificação e avaliação de anticorpos. É agora possível construir e pesquisar, pelo menos, três ordens de grandeza, mais clones com mono-especificidades do que era anteriormente possível. As aplicações potenciais do processo deverão englobar a investigação básica e as ciências aplicadas.

71 027

SCRO286P

-88-

A handwritten signature in black ink, consisting of stylized, cursive letters, likely representing the name of the inventor or a legal representative.

Pretende-se que a descrição anterior seja ilustrativa e não limitativa da presente invenção. Podem-se efectuar numerosas variações e modificações sem sair do verdadeiro espírito e âmbito do presente invento.



R E I V I N D I C A Ç Õ E S

1 - Processo para a produção de um vector de expressão de ADN de cadeia dupla, linear, contendo um promotor 5' terminal controlando a expressão de um primeiro e de um segundo polipéptidos, a partir de um primeiro e de um segundo genes, respectivos, sendo os referidos primeiro e segundo polipéptidos capazes de formar um receptor heterodimérico de especificidade predeterminada, caracterizado por compreender:

(a) o isolamento de uma primeira população diversa, de primeiros genes contendo um gene codificando o referido primeiro polipéptido, sendo a referida primeira população constituída por moléculas de ADN de cadeia dupla, linear, definindo um sítio de restrição reconhecido por uma endonuclease, estando o referido sítio de restrição localizado 5' terminal em relação aos referidos genes mas 3' terminal em relação a um promotor controlando a expressão dos referidos genes;

(b) o isolamento de uma segunda população diversa, de segundos genes contendo um gene codificando o referido segundo polipéptido, sendo a referida segunda população constituída por moléculas de ADN de cadeia dupla, linear, definindo o referido sítio de restrição, estando o referido sítio de restrição localizado 3' terminal em relação aos referidos segundos genes, estando os referidos segundos genes localizados 3' terminais em relação a um promotor controlando a expressão dos referidos segundos genes;

(c) a clivagem das referidas moléculas de ADN de cadeia dupla, quando presentes nas referidas populações, com a referida endonuclease para produzir fragmentos de restrição possuindo um terminal coesivo e um dos referidos primeiros genes ou um dos referidos segundos genes; e

(d) a ligação aleatória dos referidos fragmentos de restrição uns aos outros, através do referido terminal coesivo de cada um, para produzir uma população diversa de vectores de expressão de ADN, linear, de cadeia dupla, possuindo um dos referidos primeiros genes e um dos referidos segundos genes, em



tandem, na mesma cadeia de ADN e sob o controlo de um único promotor;

(e) o isolamento de um vector de expressão capaz de exprimir o referido receptor heterodimérico de especificidade predeterminada, a partir da referida população diversa de vectores de expressão de ADN de cadeia dupla.

2 - Processo de acordo com a reivindicação 1, caracterizado por os referidos primeiros genes serem homólogos de ADN codificando o primeiro polipéptido e por o referido isolamento compreender:

(a) a separação das cadeias de um reportório de genes codificando o primeiro polipéptido, compreendendo o referido reportório ácidos nucleicos de cadeia dupla, contendo, cada um, uma cadeia codificando o primeiro polipéptido, temperada (annealed) com uma cadeia complementar;

(b) o tratamento das referidas cadeias separadas sob condições adequadas para amplificação, por reacção em cadeia com polimerase, com primeiros e segundos iniciadores de síntese de polinucleótidos, possuindo cada um dos referidos primeiros iniciadores uma sequência de nucleótidos capaz de se hibridar com uma sequência conservada entre as referida cadeias codificando o primeiro polipéptido, e possuindo cada um dos referidos segundos iniciadores uma sequência de nucleótidos capaz de se hibridar com uma sequência conservada entre as referidas cadeias complementares, sendo os referidos primeiros e segundos iniciadores capazes de iniciar a amplificação de uma pluralidade de homólogos de ADN diferentes, codificando o primeiro polipéptido, de entre o referido reportório de genes codificando o primeiro polipéptido, produzindo o referido tratamento uma pluralidade de homólogos de ADN diferentes codificando o primeiro polipéptido; e

(c) a ligação operativa a um vector de expressão para a expressão, sob o controlo de um promotor, de cada um, de uma pluralidade dos referidos homólogos de ADN diferentes, codificando o primeiro polipéptido.



3 - Processo de acordo com a reivindicação 1, caracterizado por os referidos segundos genes serem homólogos de ADN codificando o segundo polipeptido e por o referido isolamento compreender:

(a) a separação das cadeias de um reportório de genes codificando o segundo polipeptido, compreendendo o referido reportório ácidos nucleicos de cadeia dupla, contendo, cada um, uma cadeia codificando o segundo polipeptido, temperada com uma cadeia complementar;

(b) o tratamento das referidas cadeias separadas sob condições adequadas para amplificação, por reacção em cadeia com polimerase, com primeiros e segundos iniciadores de síntese de polinucleótidos, possuindo cada um dos referidos primeiros iniciadores uma sequência de nucleótidos capaz de se hibridar com uma sequência conservada entre as referidas cadeias codificando o segundo polipeptido, e possuindo cada um dos referidos segundos iniciadores uma sequência de nucleótidos capaz de se hibridar com uma sequência conservada entre as referidas cadeias complementares, sendo os referidos primeiros e segundos iniciadores capazes de iniciar a amplificação de uma pluralidade de homólogos de ADN diferentes, codificando o segundo polipeptido, de entre o referido reportório de genes codificando o segundo polipeptido, produzindo o referido tratamento uma pluralidade de homólogos de ADN diferentes codificando o segundo polipeptido; e

(c) a ligação operativa a um vector de expressão para a expressão, sob o controlo de um promotor, de cada um, de uma pluralidade dos referidos homólogos de ADN diferentes codificando o segundo polipeptido.

4 - Processo de acordo com a reivindicação 1, caracterizado por o referido primeiro polipeptido ser um V_H , por os referidos primeiros genes serem homólogos de ADN codificando o V_H e por o referido isolamento compreender:

(a) a separação das cadeias de um reportório de genes codificando o V_H , compreendendo o referido reportório ácidos nucleicos de cadeia dupla, contendo, cada um, uma cadeia



codificando o V_H temperada com uma cadeia complementar; e

(b) o tratamento das referidas cadeias separadas sob condições adequadas para amplificação, por reacção em cadeia com polimerase, com primeiros e segundos iniciadores de síntese de polinucleótidos, possuindo cada um dos referidos primeiros iniciadores uma sequência de nucleótidos capaz de se hibridar com uma sequência conservada entre as referidas cadeias codificando o V_H , e possuindo cada um dos referidos segundos iniciadores uma sequência de nucleótidos capaz de se hibridar com uma sequência conservada entre as referidas cadeias complementares, sendo os referidos primeiros e segundos iniciadores capazes de iniciar a amplificação de uma pluralidade de homólogos de ADN diferentes, codificando o V_H , de entre o referido reportório de genes codificando V_H , produzindo o referido tratamento uma pluralidade de homólogos de ADN diferentes codificando o V_H .

5 - Processo de acordo com a reivindicação 1, caracterizado por o referido segundo polipéptido ser um V_L , por os referidos segundos genes serem homólogos do ADN codificando o V_L e por o referido isolamento compreender:

(a) a separação das cadeias de um reportório de genes codificando o V_L , compreendendo o referido reportório ácidos nucleicos de cadeia dupla, contendo, cada um, uma cadeia codificando o V_L temperada com uma cadeia complementar; e

(b) o tratamento das referidas cadeias separadas sob condições adequadas para amplificação, por reacção em cadeia com polimerase, com primeiros e segundos iniciadores de síntese de polinucleótidos, possuindo cada um dos referidos primeiros iniciadores uma sequência de nucleótidos capaz de se hibridar com uma sequência conservada entre as referidas cadeias codificando o V_L , e possuindo cada um dos referidos segundos iniciadores uma sequência de nucleótidos capaz de se hibridar com uma sequência conservada entre as referidas cadeias complementares, sendo os referidos primeiros e segundos iniciadores capazes de iniciar a amplificação de uma pluralidade de homólogos de ADN diferentes, codificando o V_L , de entre o referido reportório de genes codifi-



cando o V_L , o referido tratamento produzindo uma pluralidade de homólogos de ADN diferentes codificando o primeiro polipeptido.

6 - Processo de acordo com a reivindicação 1, caracterizado por o referido primeiro polipeptido ser um V_L , por os referidos primeiros genes serem homólogos de ADN codificando o V_L e por o referido isolamento compreender:

(a) a separação das cadeias de um reportório de genes codificando o V_L , compreendendo o referido reportório ácidos nucleicos de cadeia dupla, contendo, cada um, uma cadeia codificando o V_L temperada com uma cadeia complementar; e

(b) o tratamento das referidas cadeias separadas sob condições adequadas para amplificação por reacção em cadeia com polimerase, com primeiros e segundos iniciadores de síntese de polinucleótidos, possuindo cada um dos referidos primeiros iniciadores uma sequência de nucleótidos capaz de se hibridar com uma sequência conservada entre as referidas cadeias codificando o V_L , e possuindo cada um dos referidos segundo iniciadores uma sequência de nucleótidos capaz de se hibridar com uma sequência conservada entre as referidas cadeias complementares, sendo os referidos primeiros e segundos iniciadores capazes de iniciar a amplificação de uma pluralidade de homólogos de ADN diferentes, codificando o V_L , de entre o referido reportório de genes codificando o V_L , o referido tratamento produzindo uma pluralidade de homólogos de ADN diferentes codificando o primeiro polipeptido.

7 - Processo de acordo com a reivindicação 1, caracterizado por o referido segundo polipeptido ser um V_H , por os referidos segundos genes serem homólogos de ADN codificando o V_H e por o referido isolamento compreender:

(a) a separação das cadeias de um reportório de genes codificando o V_H , compreendendo o referido reportório ácidos nucleicos de cadeia dupla, contendo, cada um, uma cadeia codificando o V_H temperada com uma cadeia complementar; e

(b) o tratamento das referidas cadeias separadas sob condições adequadas para amplificação, por reacção em cadeia com



polimerase, com primeiros e segundos iniciadores de síntese de polinucleótidos, possuindo cada um dos referidos primeiros iniciadores uma sequência de nucleótidos capaz de se hibridar com uma sequência conservada entre as referidas cadeias codificando o V_H , e possuindo cada um dos referidos segundos iniciadores uma sequência de nucleótidos capaz de se hibridar com uma sequência conservada entre as referidas cadeias complementares, sendo os referidos primeiros e segundos iniciadores capazes de iniciar a amplificação de uma pluralidade de homólogos de ADN diferentes, codificando o V_H , de entre o referido reportório de genes codificando V_H , o referido tratamento produzindo uma pluralidade de homólogos de ADN diferentes codificando o V_H .

8 - Processo de acordo com a reivindicação 4, caracterizado por o referido segundo polipeptido ser um V_L , por os referidos segundos genes serem homólogos do ADN codificando o V_L e por o referido isolamento compreender:

(a) a separação das cadeias de um reportório de genes codificando o V_L , compreendendo o referido reportório ácidos nucleicos de cadeia dupla, contendo, cada um, uma cadeia codificando o V_L temperada com uma cadeia complementar; e

(b) o tratamento das referidas cadeias separadas sob condições adequadas para amplificação, por reacção em cadeia com polimerase, com primeiros e segundos iniciadores de síntese de polinucleótidos, possuindo cada um dos referidos primeiros iniciadores uma sequência de nucleótidos capaz de se hibridar com uma sequência conservada entre as referidas cadeias codificando o V_L , e possuindo cada um dos referidos segundos iniciadores uma sequência de nucleótidos capaz de se hibridar com uma sequência conservada entre as referidas cadeias complementares, sendo os referidos primeiros e segundos iniciadores capazes de iniciar a amplificação de uma pluralidade de homólogos de ADN diferentes, codificando o V_L , de entre o referido reportório de genes codificando o V_L , o referido tratamento produzindo uma pluralidade de homólogos de ADN diferentes codificando o primeiro polipeptido.

9 - Processo de acordo com a reivindicação 6, caracterizado por o referido segundo polipéptido ser um V_H , por os referidos segundos genes serem homólogos de ADN codificando o V_H e por o referido isolamento compreender:

(a) a separação das cadeias de um reportório de genes codificando o V_H , compreendendo o referido reportório ácidos nucleicos de cadeia dupla, contendo, cada um, uma cadeia codificando o V_H temperada com uma cadeia complementar; e

(b) o tratamento das referidas cadeias separadas sob condições adequadas para amplificação, por reacção em cadeia com polimerase, com primeiros e segundos iniciadores de síntese de polinucleótidos, possuindo cada um dos referidos primeiros iniciadores uma sequência de nucleótidos capaz de se hibridar com uma sequência conservada entre as referidas cadeias codificando o V_H , e possuindo cada um dos referidos segundos iniciadores uma sequência de nucleótidos capaz de se hibridar com uma sequência conservada entre as referidas cadeias complementares, sendo os referidos primeiros e segundos iniciadores capazes de iniciar a amplificação de uma pluralidade de homólogos de ADN diferentes, codificando o V_H , de entre o referido reportório de genes codificando V_H , o referido tratamento produzindo uma pluralidade de homólogos de ADN diferentes codificando o V_H .

10 - Processo de produção de um sistema de clonação compreendendo os primeiros e segundos vectores de ADN de cadeia dupla, linear, caracterizado por se misturarem num meio aquoso, conjuntamente com uma endonuclease, definindo, cada um dos referidos vectores, um promotor, um sítio de restrição, clivável pela referida endonuclease, e um poli-ligante, estando o referido sítio de restrição localizado, entre o referido promotor e o referido poli-ligante, no referido primeiro vector, estando o referido poli-ligante localizado, entre o referido promotor e o referido sítio de restrição, no referido segundo vector, sendo o referido sítio de restrição clivável para produzir, a partir de cada um dos referidos vectores, primeiro e segundo fragmentos de

71 027
SCR0286P

-96-

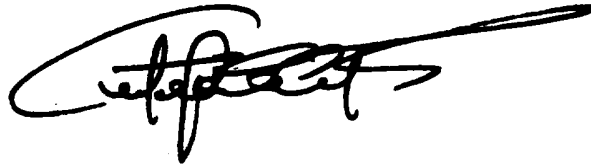
restrição respectivos contendo poli-ligante, possuindo terminais complementares um do outro.

Lisboa,

16. MAI 1990

Por SCRIPPS CLINIC AND RESEARCH FOUNDATION

- O AGENTE OFICIAL -

A handwritten signature in black ink, appearing to be 'Alfred' or similar, with a long horizontal stroke extending to the right.

[Handwritten signature]

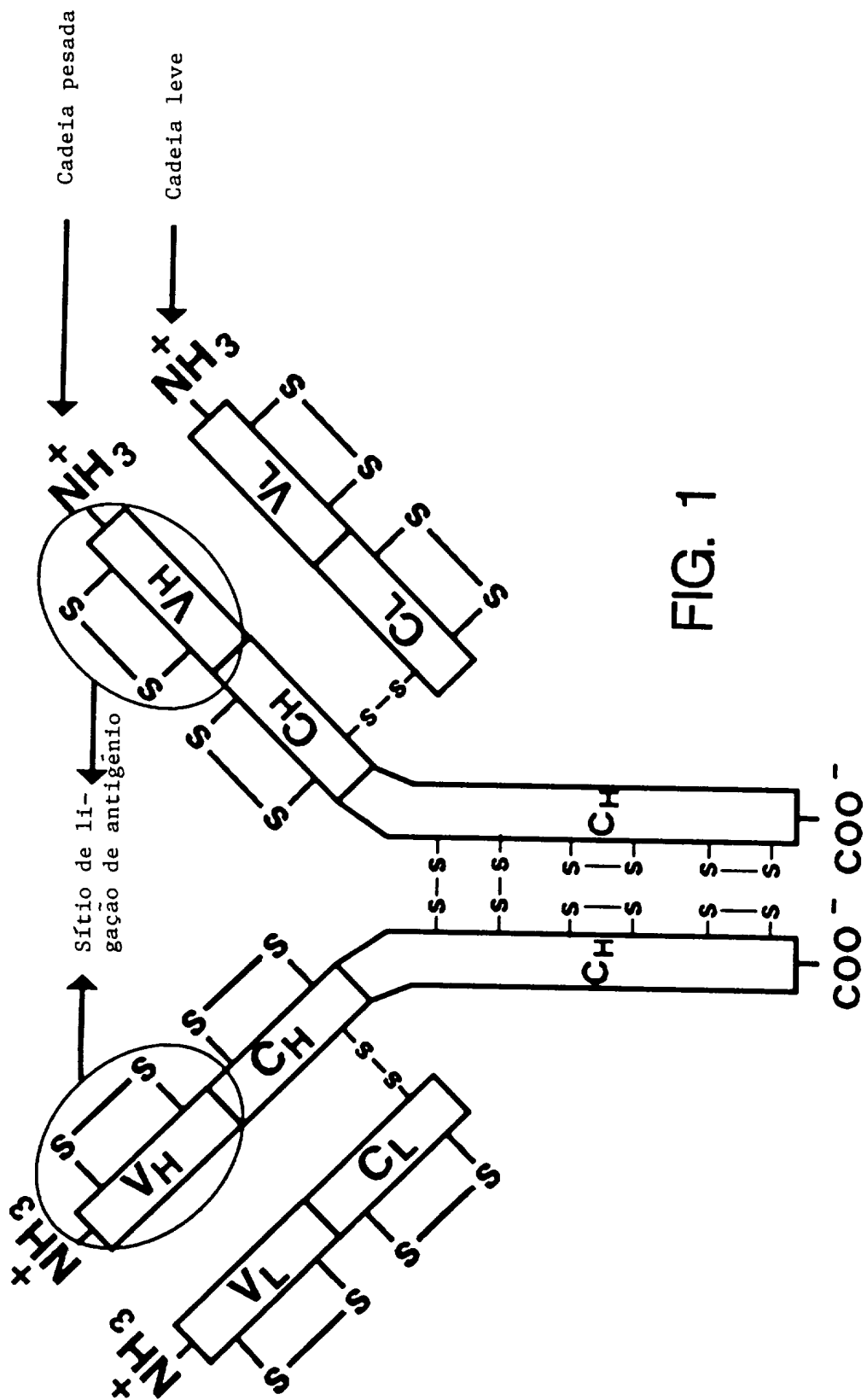


FIG. 1

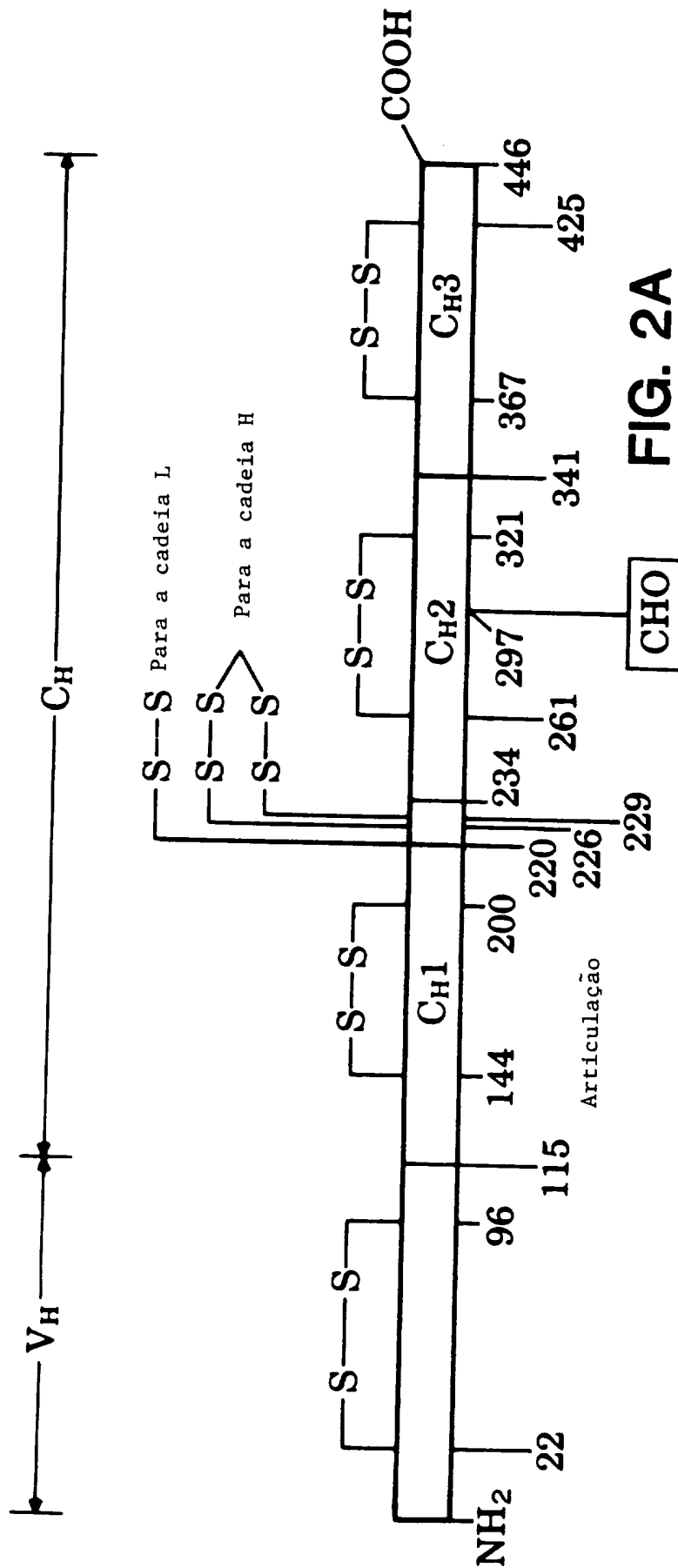


FIG. 2A

FIG. 2B-1

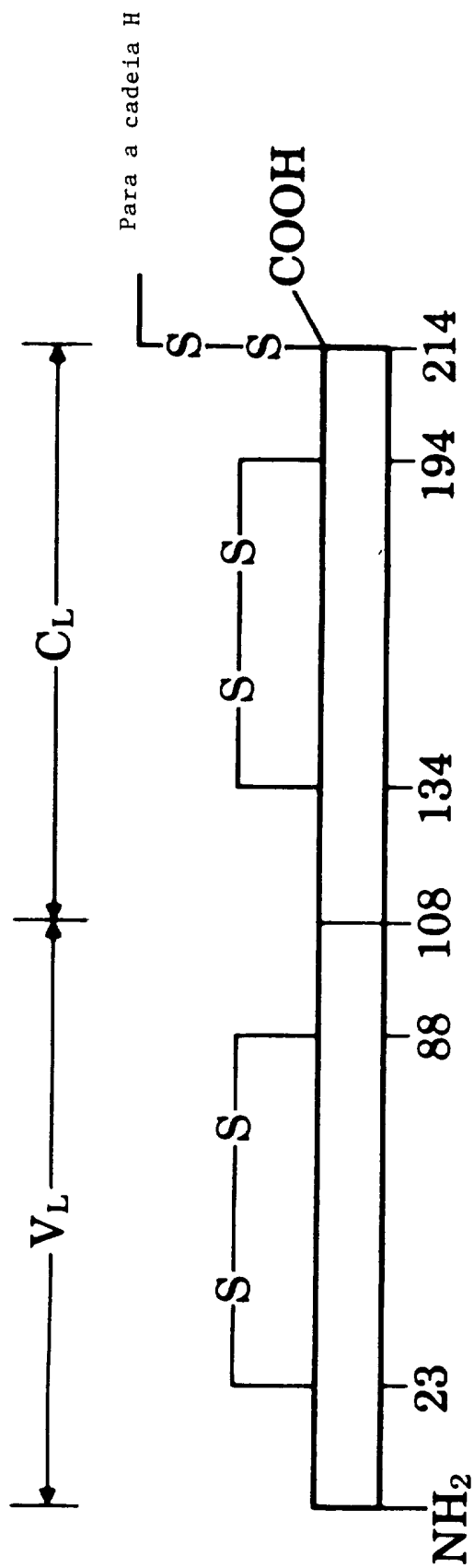


FIG. 2B-2

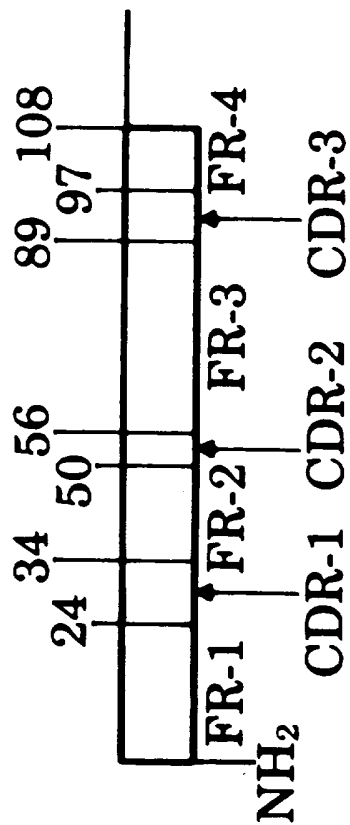


FIG. 3-1

		V	
Classe	Prototipo T15	1	10
		EVKLVESGGGLVQPGGS T15	
IgM	HPCM2		+
	HPCM3		-
	HPCM1		+
	HPCM6		-
	HPCM4		-
IgG	HPCG8		+
	HPCG13		-
	HPCG14		-
	HPCG11		+
	HPCG12		+
IgA	T15		+
	S63		+
	Y5236		+
	S107		+
	H8		+
	M603		+
	W3207		-
	M511		-
	M167	V	-



FIG. 3-2

		V									
Classe	Prototipo T15	20	30	40	50						
IgM	HPCM2	LRLSCATSGTTFSDFYMEWVRQPPGKRLEWIAASRNKAN									+
	HPCM3										-
	HPCM1										+
	HPCM6										-
	HPCM4										-
IgG	HPCG8										+
	HPCG13		L				T				-
	HPCG14			A							-
	HPCG11		I								+
	HPCG12		I								+
IgA	T15										+
	S63										+
	Y5236										+
	S107										+
	H8										+
	M603										+
	W3207										+
	M511						S				-
	M167						T				-
			HV1						S	H	-

V

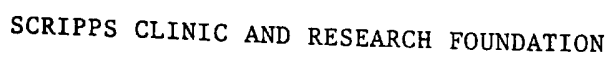


FIG. 3-4

Classe	Prototipo T15	V	D	J	Idiotipo	T15	
		100	110	120			
		YYCARD	YVGSS	YWYFDVWGAGTTVTVSS			
IgM	HPCM2						+
	HPCM3						-
	HPCM1						+
	HPCM6		DYP H				-
	HPCM4		F_RYD G				-
IgG	HPCG8		R				+
	HPCG13		A				-
	HPCG14		V_YD				-
	HPCG11						+
	HPCG12						+
IgA	T15						+
	S63						+
	Y5236						+
	S107						+
	H8		N				+
	M603		N_T				+
	W3207		N_KYD L_V				-
	M511		GD				-
	M167	T	AD_N_YFG				-
			HV3				-

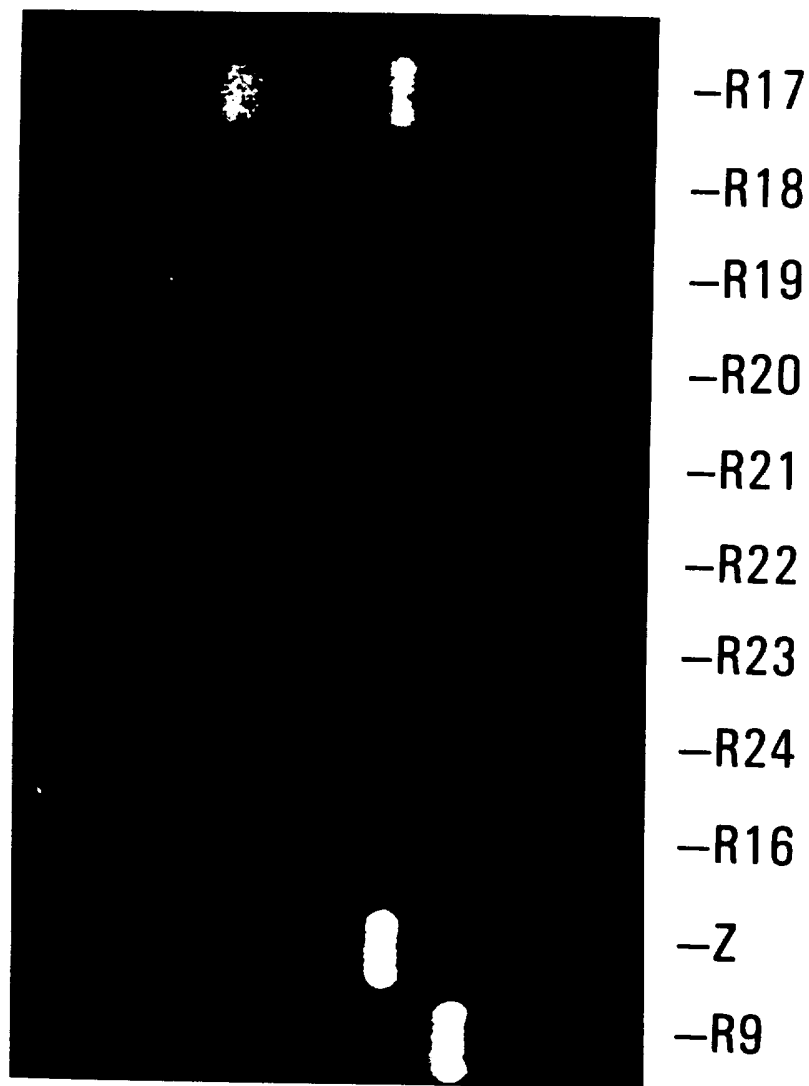


FIG. 4 360 bp 575 bp



FIG. 5-1

Subclasse I (A)

#L39

CTCGAGTCAGGACCTGGCCTCGTGAAACCTTCTCAGTCTCTGTCTCTC
 ACCTGCTCTGTCACTGGCTACTCCATCACCAGTGCTTATTACTGGAAC
 TGGATCCGGCAGTT

Subclasse II (A)

#L11

CTCGAGTCTGGGCCTnAACTGGCAAACCTGGGGCCTCAGTGAAGATG
 TCCTGCAAGGCTTCTGGCCACACCTTGACTAGTTACTGGATACACTGG
 GTAAAAAnAGAGGCC

#L03

CTCGAGTCTGGACCTnAGCTGGTAAAGCCTGGGGTTCAGTGAAGATGT
 CCTGCAAGGCTTCTGGATACACATTACnAGCTATGTTATACACTGGG
 TGAAGCAGAAGCCT

FIG. 5-2

#L32

CTCGAGTCTGGACCTGAACTGGTAAAGCCTGGGACTTCAGTGAAGATG
 TCCTGCAAGGCTTCTGGATACACATTACCAGCTATGTTATGCGCTGG
 GTGAAGCAGAAGCC

Subclasse II (B)

#L37

CTCGAGTCAGGGGCTGAACTGGTGAAGCCTGGGGTTTCAGTGAAGTTG
 TCCTGCAAGGCTTCTGGCTACACCTTCACnAGCTACTATATGTACTGG
 GTGAAGCAGAGGCC

#L06

CTCGAGTCTGGGGCTAAGCTGGTAAAGCCTGGAGCTTnAGTnAAGCTG
 TCCTGnAGGGCTTCTGGCTACTCCTTCACnAGCTACTGGATGAACTGG
 GTGAAGCAGAGGCC



FIG. 5-3

Subclasse II (C)

#L33

CTCGAGTCTGGGGCTGAGCTGGTGGAGCCTGGAGCTTCAGTnAAGCTG
TCCTGCAAGGCCTCTCGTACTCCTTCACCAGCTCCTGATAACTGGGTG
AAGCAGAGGCCTGG

Subclasse III (B)

#L36

CTCGAGTCAGGAGGTGGCCTGGTGCAGCCTGGAGGATCCCTGAAACTC
TCCTGTGCAGCCTCAGGATTCGATTTnAGnAGATACTGGATGAATTGG
GTCCGGCAGCTCCA

#L02

CTCGAGTCTGGAGGTGGCCTGGTGCAGCCTGGAGGATCCCTGAATCTC
CCCTGTGCAGCCTCAGGATTCGATTTnAGnAGATAATGGATGAGTTGG
GTTCCGGCAGGCTCC

FIG. 5-4

#L31

CTCGAGTCTGGAGGTGGCCTGGTGCAGCCTGGAGGATCCCTGAAAGTC
TCCTGTGCAGCCTCAGGATTCGATTTnAGnAGATACTGGATGAGTTGG
GTCCGGCAGCTCCA

#L34

CTCGAGTCTGGAGGTGGCCTGGTGCAGCCTGGAGGATCCCTCAAACCTC
TCCTGTGCAGCCTCAGGATTCGATTTnAGnAGATACTGGATGAGTTGG
GTCCGGCAGCTCCA

#L50

CTCGAGTCAGGAGGTGGCCTGGTGCAGCCTGGAGGAGCCCTGAAACTC
TCCTGTGCAGCCTCAGGATTCGATTTnAGnAGATACTGGATGAGTTGG
GTCCGCAGCTCCAG

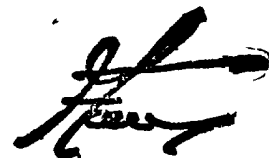


FIG. 5-5

Subclasse III (C)

#L10

CTCGAGTCTGGGGGAGGCTTAGTnCAGCCTGGAGGGTCCCGGAACTC
TCCTGTGCAGCCTCTGGATTCACTTTnAGnAGTTTGGGAATGCACTGG
ATTCGTCAGGCTCC

#L08

CTCGAGTCTGGGGGAGGCTTAGTnnAGCCTGGAGGGTCCCGGAACTC
TCCTGTGCAGCCTCTGGATTCACTTTnAGnAGCTTTGGGAATGCACTGG
GTTACGTCAGGCTC

Subclasse V (A)

#L38

CTCGAGTCAGGGGCTGAACTGGTGAGGCCTGGGCGTTCAGTnAAGATG
TCCTGCAAGGCTTCAGGCTATTCCTTCACCAGCTACTGGATGCACTGG
GTGAAACAGAGGCC

FIG. 5-6

Miscelânea

#47

CTCGAGTCAGGGGCTGAACTGGCAAAACCTGGGGCCTCAGTAAAGATG
TCCTGCAAGGCTTCTGGCTACACCTCTTCTTCCTTCTGGCTGCACTGG
ATAAAAGAAGGCCT

#L35

CTCGAGTCTGGACCTnAGCTGGTGAAGCCTGGGGTTCAGTTAAAATAT
CCTGCAAGGCTTCTGGTTACTCATTTTCTnTCTACTTTGTGAACTGGG
TGATGCAGAGCCAT

#L48

CTCGAGTCAGGGGCTGAACTGGTGAAGCCTGGGGTTCAGTAAGTTGTC
CTGAAGGCTTCTGGCTACACCTTCACCGGCTACTATATGTACTGGGTG
AAGCAGAGGCCTGG

FIG. 6A-1

Vector de expressão V_H

Shine-Dalgarno Met

GGCCGCAAATTCTATTTCAAGGAGACAGTCATAATG
CGTTTAAGATAAAGTTCCTCTGTCAGTATTAC

Sequência Lider

AAATACCTATTGCCTACGGCAGCCGCT
TTTATGGATAACGGATGCCGTCGGCGA

Sequência Lider

GGATTGTTATTACTCGCTGCCCAACCAG
CCTAACAATAATGAGCGACGGGTTGGTC

FIG. 6A-2

Ligante Ligante

NCOI	Esqueleto V _H	XHOI	SPEI
------	--------------------------	------	------

CCATGGCCCAGGTGAACTGCTCGAGATTCTAGACTAGT
GGTACCGGGTCCACTTTGACGAGCTCTAAAGATCTGATCA

Paragem Ligante

TyrProTyrAspValProAspTyrAlaSer
TACCCGTACGACGTTCCGGACTACGGTTCCTAATAGAATTCTG
ATGGGCATGCTGCAAGGCCTGATGCCAAGAATTATCTTAAGCAGCT

FIG. 6B-1

Vector de expressão V_L

Shine-Dalgarno

GGCCGCAAATTCTATTTCAAGGAGACAGTCATA
CGTTTAAGATAAAGTTCCTCTGTCAGTAT

Met Sequência Líder

ATGAAATACCTATTGCCTACGGCAGCCGCTGGA
TACTTTATGGATAACGGATGCCGTCGGCGACCT

Sequência Líder

TTGTTATTACTCGCTGCCCCAACCAG
ACAATAATGAGCGACGGGTTGGTC

FIG. 6B-2

Ligante

NCOI

Esqueleto V_H

CCATGGCCCAGGTGAACTG
GGTACCGGGTCCACTTTGAC

Ligante

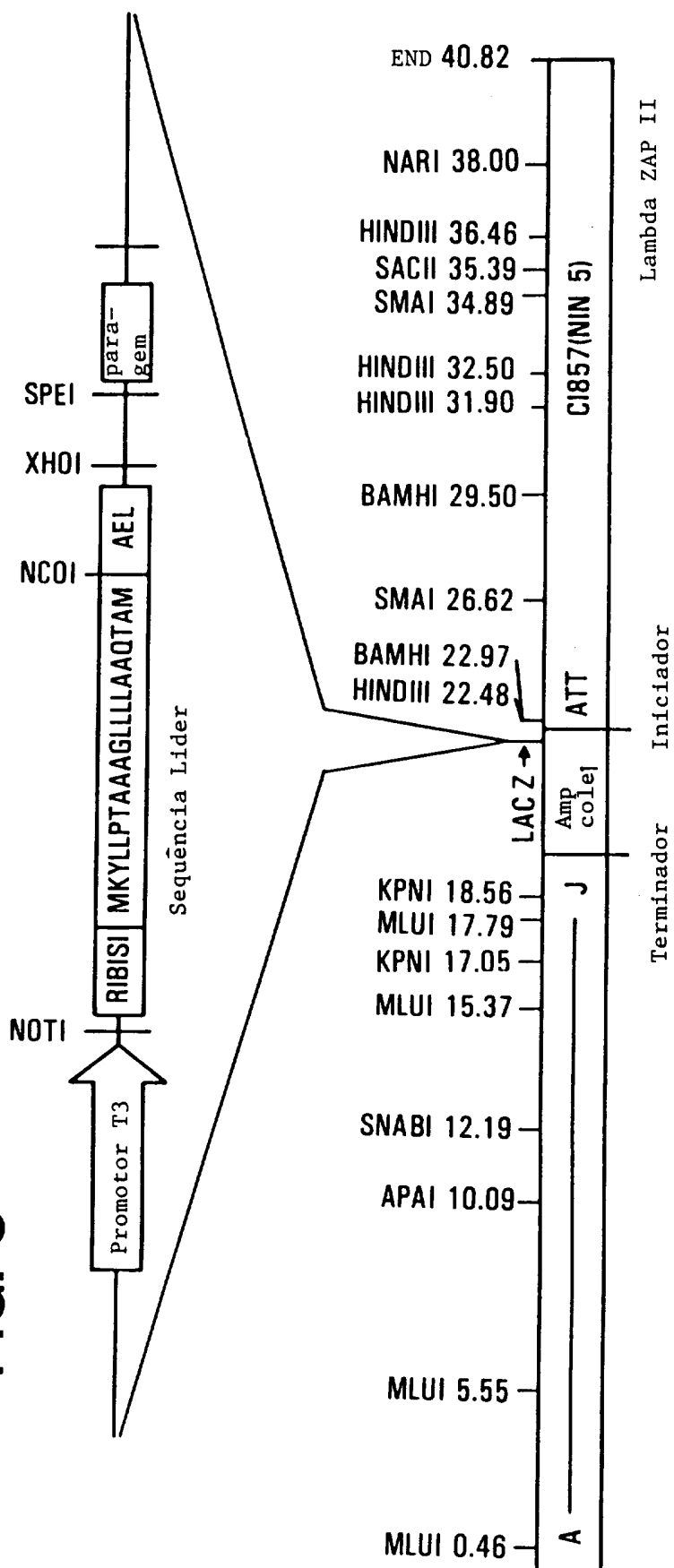
XHOI

SPEI

Paragem

CTCGAGAATTCTAGACTAGTTAATAG
GAGCTCTTAAGATCTGATCAATTATCAGCT

FIG. 8



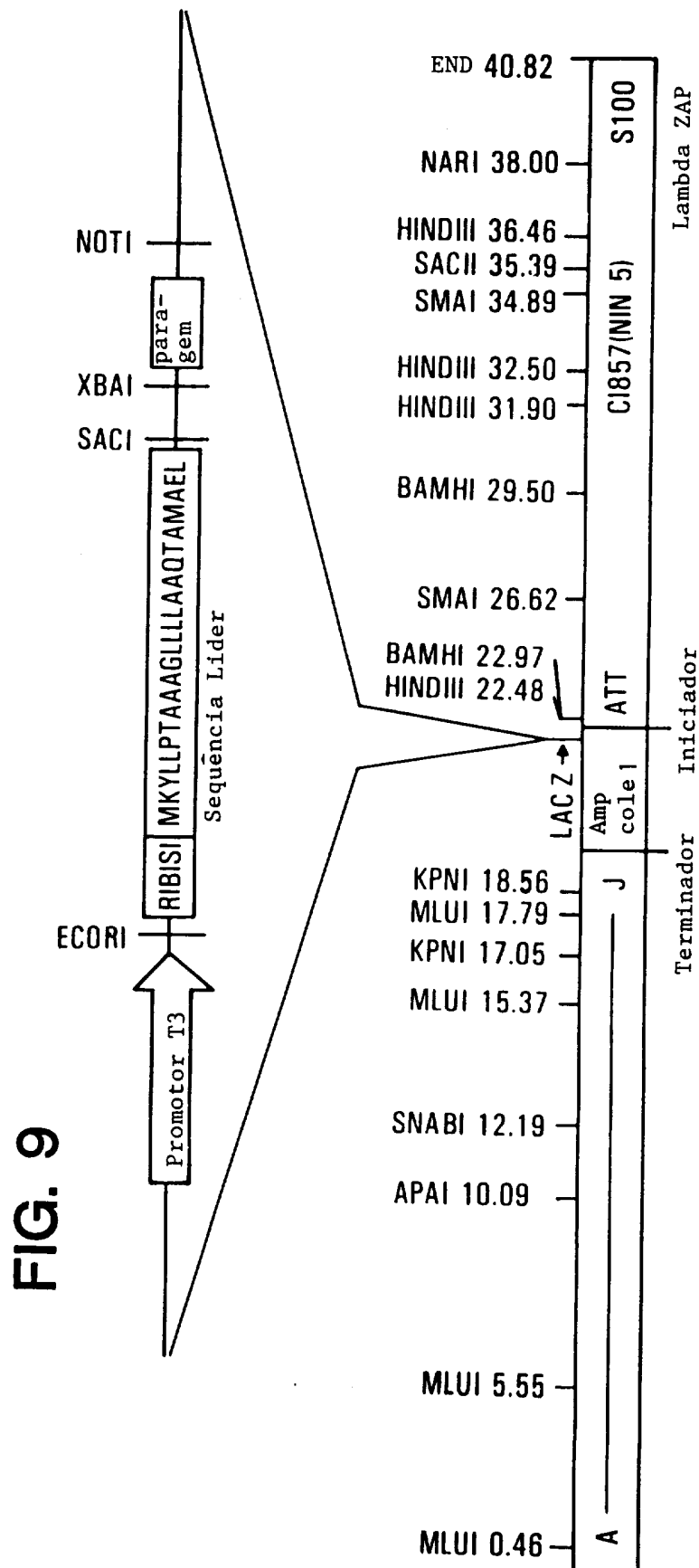


FIG. 10

ECOR I

Shine-Dalgarno
Met
TGAATTCTAAACTAGTCGCCCAAGGAGACAGTCATAATGAAAT
 TCGAACTTAAGATTGATCAGCGGTTCCCTCTGTCAGTATTACTTTA

Sequência Líder

ACCTATTGCCCTACGGCAGCCGCTGGATTGTTATTACTCGCTGCCCAACCAG
TGGATAACGGATGCCGTCGGCGCACCTAACATAATGAGCGACGGTTGGTC

NCO I	SAC I	XBA I	Not I
<u>CCATGGCCGAGCTCGTCAGTTCTAGAGTTAAGCGGCCG</u> GGTACCGGCTCGAGCAGTCAAGATCTCAATTCCGCCGGCAGCT			

FIG. 11-1

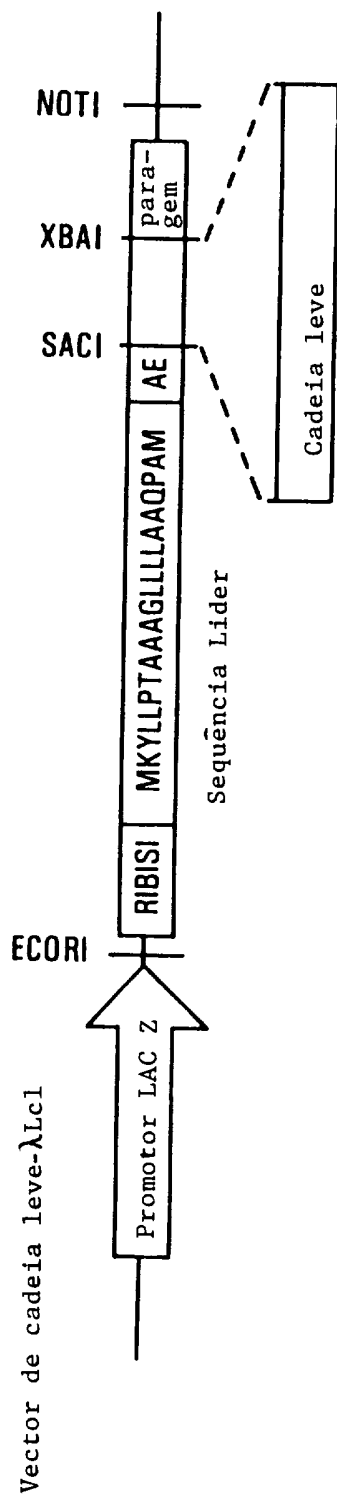
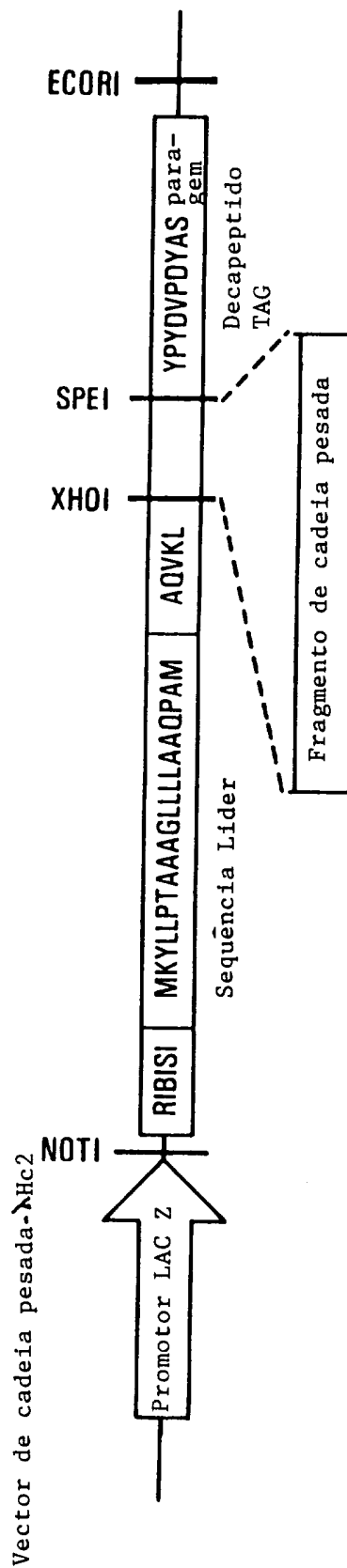
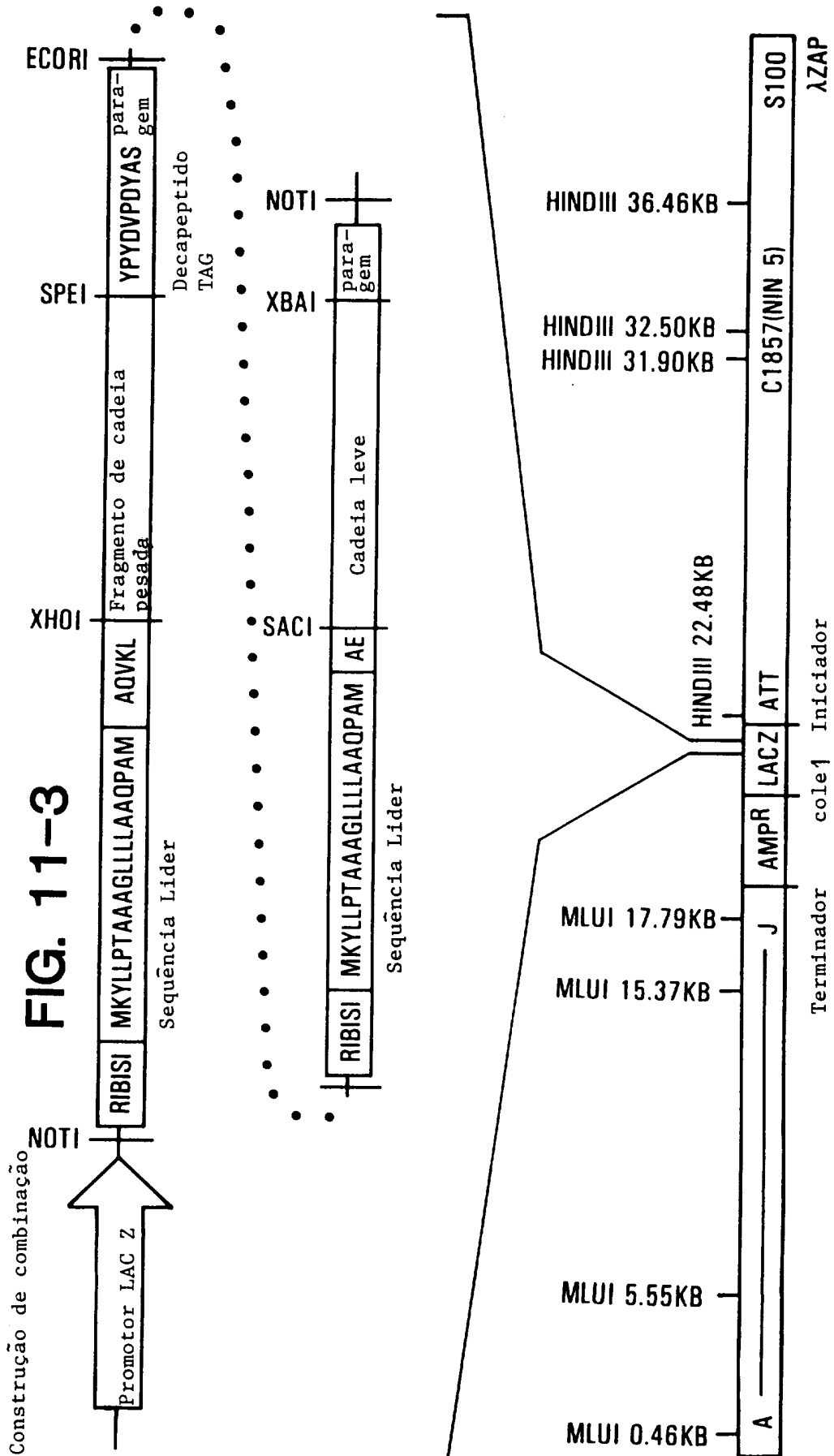


FIG. 11-2





[Handwritten signature]

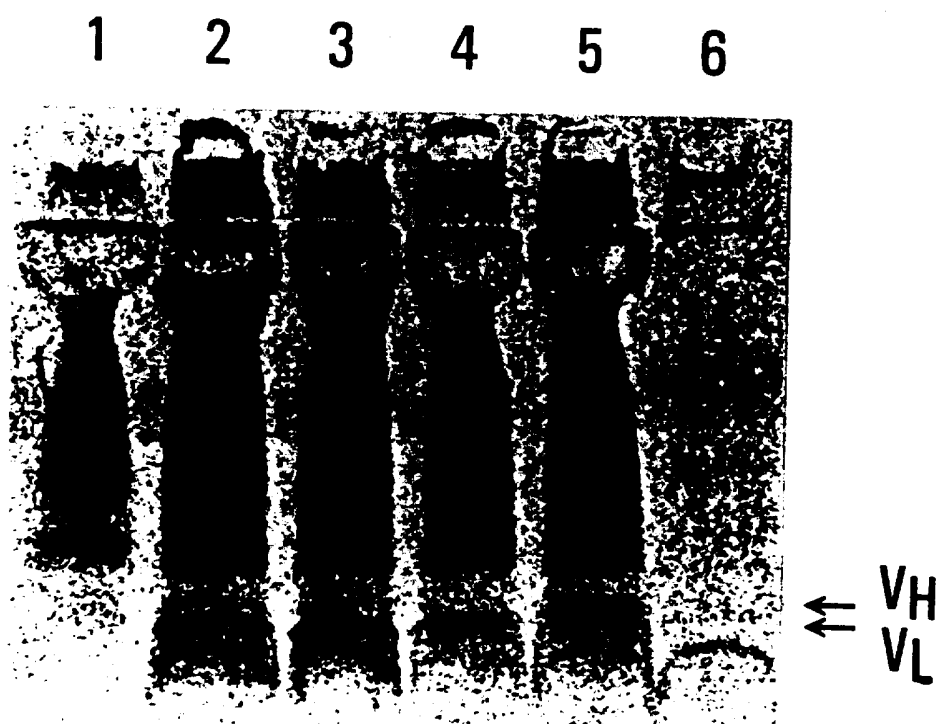
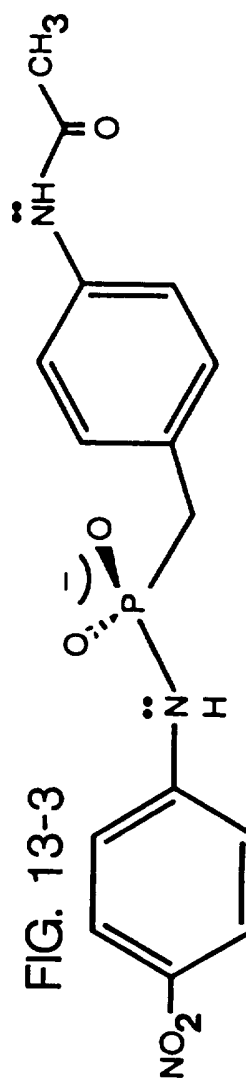
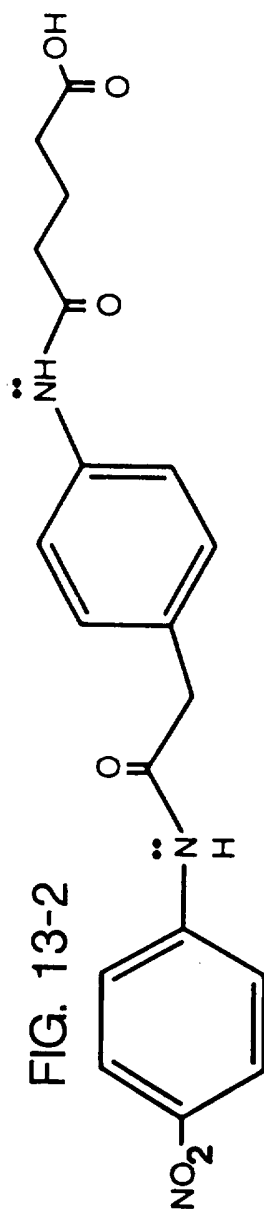
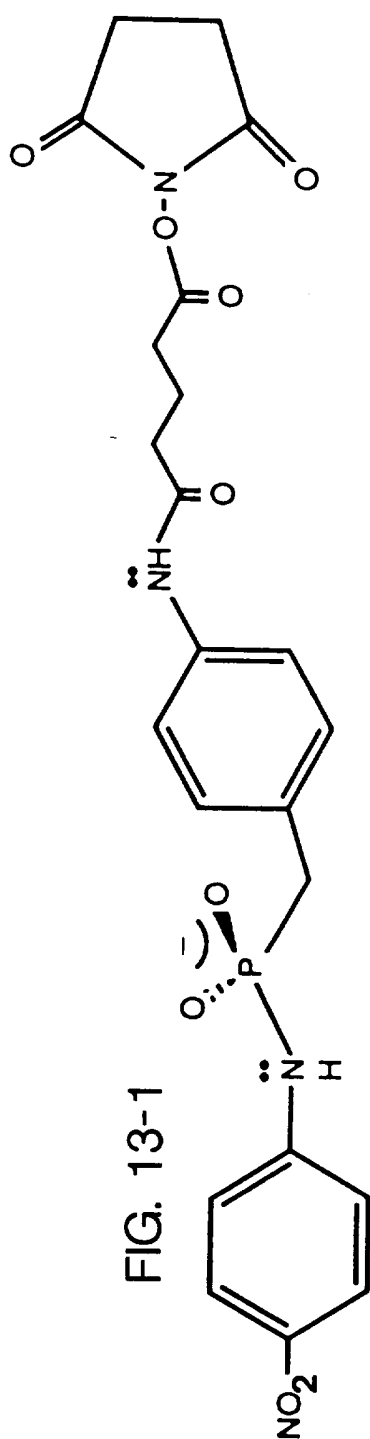
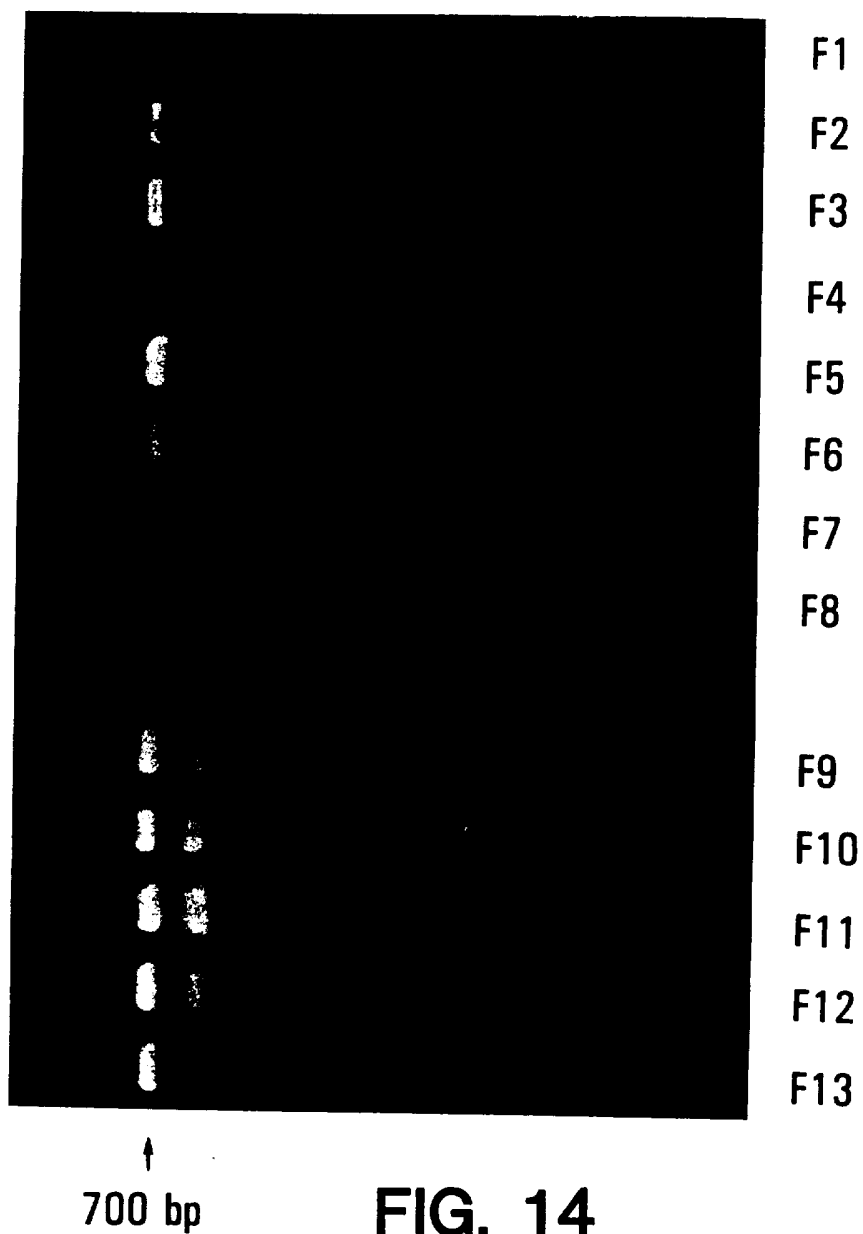


FIG. 12





Handwritten signature

FIG. 15A

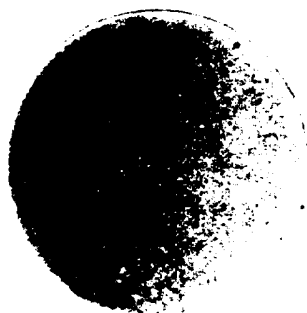


FIG. 15B

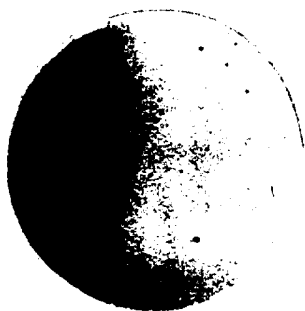


FIG. 15C

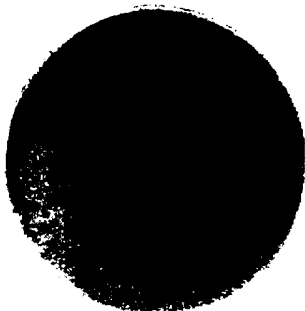


FIG. 15D

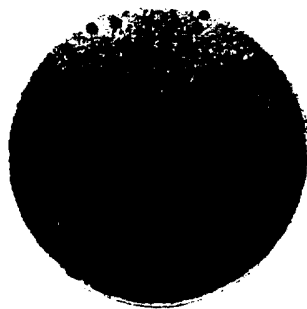


FIG. 15E



FIG. 15F

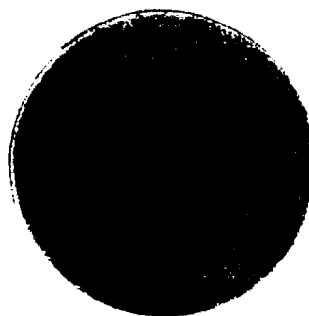


FIG. 15G

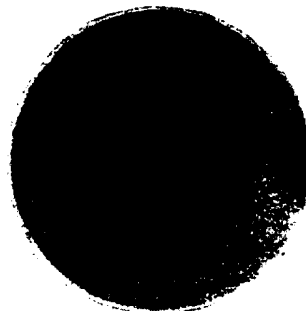


FIG. 15H

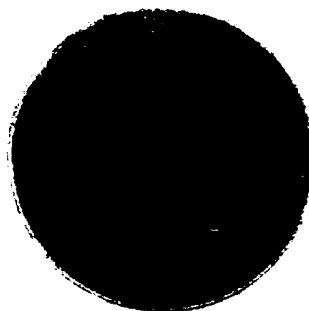


**FIG. 16**