

(11) *Número de Publicação:* **PT 89768 B**

(51) *Classificação Internacional:* (Ed. 5)

C12N015/00 A	C12N001/20 B
C12N001/18 B	C12P021/02 B
C12N001/20 B	C12R001/19 B

(12) *FASCÍCULO DE PATENTE DE INVENÇÃO*

(22) *Data de depósito:* 1989.02.20

(30) *Prioridade:* 1988.02.24 US 160463

(43) *Data de publicação do pedido:*
1989.10.04

(45) *Data e BPI da concessão:*
09/93 1993.09.13

(73) *Titular(es):*

SMITHKLINE BEECHAM CORPORATION
ONE FRANKLIN PLAZA, P.O. BOX 7929 -
PHILADELPHIA PENNSYLVANIA 19101, EUA US

(72) *Inventor(es):*

MARY ELLEN BRAWNER	US
JAMES ALLAN FORNWALD	US
JAMES ARTHOS	US
JESSICA ANGELL GORMAN	US
RAYMOND W. SWEET	US

(74) *Mandatário(s):*

ANTÓNIO JOÃO COIMBRA DA CUNHA FERREIRA
RUA DAS FLORES 74 4/AND. 1294 LISBOA PT

(54) *Epígrafe:* PROCESSO DE PRODUÇÃO DE PROTEÍNAS DE LIGAÇÃO A HIV

(57) *Resumo:*

[Fig.]

68 798

SKB CASE 14414

PATENTE Nº. 89 768



"Processo de produção de proteínas de
ligação a HIV"

para que

SMITHKLINE BECKMAN CORPORATION, pre-
tende obter privilégio de invenção
em Portugal.

R E S U M O

O presente invento refere-se ao processo de produção de proteínas de ligação a HIV, o qual compreende cultivar células de E. coli, Streptomyces ou de levedura que foram transformadas com um vector compreendendo uma sequência de codificação para a proteína de ligação a HIV operativamente ligada a um elemento regulador e recolher daí a proteína de ligação a HIV.

-2-

MEMÓRIA DESCRITIVA

Campo do Invento

Este invento refere-se à produção de proteínas de ligação a HIV, derivadas de proteínas CD-4 solúveis (sCD-4). Mais particularmente, ele refere-se à construção de vectores para a expressão de proteínas de ligação a HIV em E. coli, Streptomyces e levedura.

Antecedentes do Invento

A glicoproteína da célula T humana, CD-4 (anteriormente, algumas vezes, referida por "T4") é uma de várias proteínas não polimórficas do receptor de superfície do linfócito T que têm sido implicadas na mediação de interações eficientes célula T-célula alvo.

A sequência de ADN do receptor humano CD-4 (T-4) é conhecida (Maddon, et al., Cell 42:93 (1985)). A sequência completa da pré-proteína de CD-4 tem 458 aminoácidos de comprimento compreendendo os domínios de comando/^{putativo}secretório com 23 aminoácidos, de superfície (V_1 - V_4) com 372 aminoácidos, de membrana com 23 aminoácidos e citoplásmico com 40 aminoácidos. O domínio de superfície apresenta quatro regiões de homologia limitada, 20-30%, às regiões de variável (V) e união (J) da imunoglobulina. Quatro das cinco fronteiras intrão-exão no domínio de superfície ocorrem próximo das junções destas regiões V-J.

O vírus da imunodeficiência humana (HIV), o agente etiológico do síndrome da imunodeficiência adquirida (SIDA), apresenta uma afinidade marcada com os linfócitos CD-4. A ligação do vírus a CD-4 é mediada pela glicoproteína gp120 do envelope viral. A proteína CD-4 pode ser co-precipitada a partir de células infectadas com HIV, lisadas, com anticorpo anti-env e reciprocamente, a gp120 pode ser co-precipitada com anticorpo monoclonal anti-CD-4, demonstrando assim que a ligação do vírus a CD-4 é mediada pela glicoproteína gp120 do envelope viral.

Recentemente, vários investigadores científicos revelaram as suas observações relativas à interacção entre CD-4 e a infec-

-3-

ção da SIDA usando um derivado de CD-4 expresso em sistemas de células de mamífero.

Deen, et al., Nature 331:82 (1988) revelam o isolamento e expressão de sCD-4 em vários meios celulares. A sCD-4 é um derivado solúvel de CD-4 ao qual falta os domínios da transmembrana e citoplasmático de CD-4. Usando a linha de células de Ovario de Criceto Chinês (CHO) deficiente de hr eles mostraram que a sCD-4 expressa retém as propriedades estruturais e biológicas de CD-4 na superfície da célula, liga-se à proteína do envelope de HIV gp120 e pode inibir a ligação de linfócitos CD-4⁺.

Trautnecker, et al., Nature 331:84 (1988) revelam a utilização de sCD-4 para inibir a junção do vírus à célula alvo. Os dados apresentados mostram a inibição da infecção por HIV, in vitro, usando uma forma recombinante de sCD-4.

Fisher, et al., Nature 331:76 (1988) revelam a purificação de sCD-4 a partir de (CHO) e o seu uso como inibidor potente tanto da replicação do vírus como da fusão da célula induzida por vírus.

Hussey, et al., Nature 331:78 (1988) revelam a utilização do sistema de expressão de baculovírus para produzir sCD-4 que se verificou inibir a fusão da célula induzida por vírus (formação de sincícios) e a infecção por HIV por ligação a células que expressam gp120. Foi também mostrado que a proteína sCD-4 não parece inibir as interações da célula específica da classe II.

Smith, et al., Science 238:1704 (1987) também revelam a utilização do sistema de células de mamífero CHO para produzir sCD-4. Eles revelam a produção de uma proteína sCD-4 em células de CHO que se pode mostrar que se liga à proteína de envelope de HIV gp120 e neutraliza a infecciosidade de HIV-1 in vitro. Os dados indicam que sCD-4 pode ser usada como base para o tratamento terapêutico da infecção da SIDA.

No entanto há ainda necessidade de se conseguirem sistemas bacterianos e eucarióticos inferiores para expressão de proteínas sCD-4.

Sumário do Invento

Este invento reside na verificação de que a função de ligação a HIV da proteína CD-4 pode ser expressa em E. coli e na descoberta de que esta proteína de ligação a HIV pode ser segregada no meio de fermentação. Em particular, portanto, este invento é um vector de expressão em E. coli compreendendo uma proteína de ligação a HIV operativamente ligada a um elemento regulador.

Este invento reside também na verificação de que a função de ligação a HIV da proteína CD-4 pode ser expressa em Streptomyces e na descoberta de que essa proteína de ligação a HIV pode ser segregada no meio de fermentação. Em particular, portanto, este invento é um vector de expressão em Streptomyces compreendendo uma proteína de ligação a HIV operativamente ligada a um elemento regulador.

Este invento também reside na verificação de que a função de ligação a HIV da proteína CD-4 pode ser expressa em levedura. Em particular, portanto, este invento é um vector de expressão em levedura compreendendo uma proteína de ligação a HIV operativamente ligada a um elemento regulador.

Este invento também reside na verificação de que a sequência de comando OMPA pode ser usada para exportar proteínas heterólogas de E. coli. Em particular, portanto este invento é um vector de expressão em E. coli para expressar proteínas exportadas em E. coli compreendendo uma sequência de codificação de ADN operativamente ligada a um elemento regulador, em que a sequência de codificação de ADN codifica para a proteína X-Y onde X é uma sequência de comando OMPA e Y é uma proteína heteróloga.

Descrição Detalhada do Invento

De acordo com o presente invento as proteínas de ligação a HIV derivadas do receptor CD-4 são expressas em quantidades úteis para uso como agentes terapêuticos e/ou profiláticos para inibir a infecção inicial por HIV e para inibir a transmissão célula a célula de HIV a seguir à infecção do hospedeiro; para uso em rastreadores de inibidores da ligação HIV-linfócito; para uso como inibidores das funções de linfócito CD-4⁺; ou para outros usos como descrito abaixo. As proteínas têm todas função de ligação a HIV de CD-4. Exemplos preferidos dessas proteínas são as

-5-

proteínas sCD-4 e seus derivados. As proteínas sCD-4 são derivados de CD-4 às quais faltam os domínios de transmembrana e citoplasmático de CD-4. Exemplos de proteínas sCD-4 são descritas em documentos distintos por Deen et al., Trauneker et al., Fisher et al., e Hussey et al., em Nature 331:76-88 (1988) e por Smith et al., Science 238:1704 (1987).

A proteína sCD-4 aqui exemplificada, designada por SKBsCD-4 é referida por Deen et al. citado acima, em 83. É apresentada na sequência de aminoácidos e nucleótidos seguinte. A sequência mostra uma região 5' não traduzida e uma sequência de comando derivadas do ADNc de CD-4 referido por Maddon et al., Cell 42:93 (1985). O primeiro aminoácido de SKBsCD-4 madura, expresso e segregado por células CHO, é a lisina codificada pelos pares de bases 151-153. O domínio de transmembrana foi deleccionado, restringindo ADNc de CD-4 com HpaII no par de bases 1252 e ligando aí um ligando que restaurou os pares de bases 1253-1257 e que colocou um codão de paragem de tradução, TAA, após a valina do C-terminal. Como referido por Maddon et al., os primeiros três aminoácidos da CD-4 madura são glutamina (apresentado aqui como o aminoácido (-)2) - glicina (apresentado aqui como o aminoácido (-)1) - asparagina (apresentado aqui como lisina, aminoácido (+)1). Por outro lado, a sequência apresentada a seguir é a mesma da revelada por Maddon et al..

-6-

10 30 50
CAAGCCCAGAGCCCTGCCATTTCTGTGGGCTCAGGTCCCTACTGCTCAGCCCTTCCTCC

70 90 110
CTCGGAAAGGCCACAATGAACCGGGGAGTCCCTTTTAGGCACTTGCTTCTGGTGCTGCAA
MetAsnArgGlyValProPheArgHisLeuLeuValLeuGln

130 150 170
CTGGCGCTCCTCCCAGCAGCCACTCAGGGAAGAAAGTGGTGCTGGGCAAAAAGGGGAT
LeuAlaLeuLeuProAlaAlaThrGlnGlyLysLysValValLeuGlyLysLysGlyAsp
-1 +1

190 210 230
ACAGTGGAACCTGACCTGTACAGCTTCCCAGAAGAAGAGCATACAATTCCACTGGAAAAAC
ThrValGluLeuThrCysThrAlaSerGlnLysLysSerIleGlnPheHisTrpLysAsn

250 270 290
TCCAACCAGATAAAGATTCTGGGAAATCAGGGCTCCTTCTTAACTAAAGGTCCATCCAAG
SerAsnGlnIleLysIleLeuGlyAsnGlnGlySerPheLeuThrLysGlyProSerLys

310 330 350
CTGAATGATCGCGCTGACTCAAGAAGAAGCCTTTGGGACCAAGGAAACTTCCCCCTGATG
LeuAsnAspArgAlaAspSerArgArgSerLeuTrpAspGlnGlyAsnPheProLeuIle

370 390 410
ATCAAGAATCTTAAGATAGAAGACTCAGATACTTACATCTGTGAAGTGGAGGACCAGAAG
IleLysAsnLeuLysIleGluAspSerAspThrTyrIleCysGluValGluAspGlnLys

430 450 470
GAGGAGGTGCAATTGCTAGTGTTCCGATTGACTGCCAACTCTGACACCCACCTGCTTCAG
GluGluValGlnLeuLeuValPheGlyLeuThrAlaAsnSerAspThrHisLeuLeuGln
104

490 510 530
GGGCAGAGCCTGACCCTGACCTTGAGAGCCCCCTGGTAGTAGCCCTCAGTGCAATGT
GlyGlnSerLeuThrLeuThrLeuGluSerProProGlySerSerProSerValGlnCys

550 570 590
AGGAGTCCAAGGGCTAAAAACATACAGGGGGGGAAGACCCTCTCCGTGTCTCAGCTGGAG
ArgSerProArgGlyLysAsnIleGlnGlyGlyLysThrLeuSerValSerGlnLeuGlu

610 630 650
CTCCAGGATAGTGGCACCTGGACATGCACTGTCTTGCAGAACCAGAAGAAGGTGGAGTTC
LeuGlnAspSerGlyThrTrpThrCysThrValLeuGlnAsnGlnLysLysValGluPhe
151

670 690 710
AAAaTAGACATCGTGGTGCTAGCTTTCCAGAAGGCCTCCAGCATAGTCTATAAGAAAGAG
LysIleAspIleValValLeuAlaPheGlnLysAlaSerSerIleValTyrLysLysGlu
183

-7-

730 750 770
GGGGAACAGGTGGAGTTCTCCTTCCCACTCGCCTTTACAGTTGAAAAGCTGACGGGCAGT
GlyGluGlnValGluPheSerPheProLeuAlaPheThrValGluLysLeuThrGlySer

790 810 830
GGCGAGCTGTGGTGGCAGGCGGAGAGGGCTTCCTCCTCCAAGTCTTGGATCACCTTTGAC
GlyGluLeuTrpTrpGlnAlaGluArgAlaSerSerSerLysSerTrpIleThrPheAsp

850 870 890
CTGAAGAACAAGGAAGTGTCTGTAAAACGGGTTACCCAGGACCCTAAGCTCCAGATGGGC
LeuLysAsnLysGluValSerValLysArgValThrGlnAspProLysLeuGlnMetGly

910 930 950
AAGAAGCTCCCGCTCCACCTCACCTGCCCCAGGCCTTGCCCTCAGTATGCTGGCTCTGGA
LysLysLeuProLeuHisLeuThrLeuProGlnAlaLeuProGlnTyrAlaGlySerGly

970 990 1010
AACCTCACCTGGCCCTTGAAGCGAAAACAGGAAAGTTGCATCAGGAAGTGAAcCTGGTG
AsnLeuThrLeuAlaLeuGluAlaLysThrGlyLysLeuHisGlnGluValAsnLeuVal

1030 1050 1070
GTGATGAGAGCCACTCAGCTCCAGAAAAATTTGACCTGTGAGGTGTGGGGACCCACCTCC
ValMetArgAlaThrGlnLeuGlnLysAsnLeuThrCysGluValTrpGlyProThrSer

1090 1110 1130
CCTAAGCTGATGCTGAGCTTGAAACTGGAGAACAAGGAGGCAAAGGTCTCGAAGCGGGAG
ProLysLeuMetLeuSerLeuLysLeuGluAsnLysGluAlaLysValSerLysArgGlu

1150 1170 1190
AAGGCGGTGTGGGTGCTGAACCCTGAGGCGGGGATGTGGCAGTGTCTGCTGA_gTGACTCG
LysAlaValTrpValLeuAsnProGluAlaGlyMetTrpGlnCysLeuLeuSerAspSer

1210 1230 1250
GGACAGGTCCTGCTGGAATCCAACATCAAGGTTCTGCCCACATGGTCCACCCCGGtgtaa
GlyGlnValLeuLeuGluSerAsnIleLysValLeuProThrTrpSerThrProValEnd
351 369

1270
tggcgcctctaga

-8-

Porções truncadas de sCD-4 e seus derivados, que retenham substancialmente a afinidade de ligação a HIV de sCD-4 podem também ser expressos de acordo com este invento. De entre estas porções truncadas são preferidas proteínas ou polipéptidos compreendendo a região de ligação a HIV das regiões V1 e V2 que está compreendida aproximadamente nos aminoácidos 1-182. Em relação à sequência apresentada anteriormente, um exemplo particular é a proteína aqui referida por VIJ4 que tem os aminoácidos (-)2-104 (N-glutamina-glicina-lisina- a -alanina-asparagina-serina-C) fundidos aos aminoácidos 351-369 (-glicina-glutamina-valina a -treonina-prolina-valina-C). Um segundo exemplo é a proteína aqui referida por V1V2J4 que tem aminoácidos (-)2-183 (N-glutamina-glicina-lisina- a -lisina-alanina-serina-C) fundidos aos aminoácidos 351-369. Um terceiro exemplo é a proteína aqui referida por YV1, que compreende aminoácidos (-)9-151 (N-alanina-leucina-leucina- a -leucina-ácido glutâmico-leucina-C). Um quarto exemplo é a proteína aqui referida por SKBYsCD-4 que tem aminoácidos (-)9-369. Cada uma destas proteínas de ligação a HIV, baseada em, p. ex. ligação a OKT4A, tem a propriedade de ligação a HIV das proteínas sCD-4, especificamente da SKBsCD-4, e, mais particularmente, tem o domínio de ligação a HIV das regiões V1 e V2 das proteínas sCD-4, especificamente de SKBsCD-4.

Em adição a estas proteínas aqui especificamente exemplificadas, um perito na arte pode construir facilmente sequências de codificação de ADN para mais derivados de sCD-4. Estes derivados retêm substancialmente a função de ligação a HIV das proteínas sCD-4, i.e. a ligação da proteína à proteína env de HIV, gp120, não é significativamente afectado adversamente. Por exemplo, pode construir-se um derivado possuindo apenas a região V1, p.e. aproximadamente os aminoácidos 1-94, ou um derivado possuindo apenas uma porção da região V1 de CD-4. Podem também construir-se derivados das proteínas aqui especificamente exemplificadas, tais como derivados diferindo numa ou em poucas deleções, adições ou substituições de aminoácidos.

Alternativamente, pode construir-se uma proteína híbrida, i.e., uma fusão de tradução, entre uma proteína sCD-4 e um veícu-

-9-

lo de proteína, outro antígeno ou outras moléculas de sCD-4 para preparar uma molécula poli-sCD-4. Ainda noutra alternativa, o mutante de deleção de sCD-4 pode ser conjugado sinteticamente com uma molécula de veículo.

A afinidade das moléculas sCD-4 com HIV pode ser demonstrada por um ensaio de ligação competitiva usando uma molécula de sCD-4 possuindo uma afinidade conhecida ou usando anticorpos que reconhecem o receptor de CD-4, tal como OKT-4 e OKT4A. As proteínas sCD-4 úteis do invento são precipitadas selectivamente de meio condicionado por OKT4A como se mostra nos Exemplos, abaixo. A CD-4, fragmentos e derivados podem ser preparados quimicamente, após expressão ou geneticamente, antes da expressão, por manipulação da sequência de codificação para o domínio de comando e/ou extracelular.

A sequência de codificação de ADN para a proteína sCD-4 ou de ligação a HIV daí derivada pode ser obtida, por exemplo, sintetizando o gene usando a sequência de ADN conhecida, por técnicas de clonação standard baseadas na sequência e por re-isolamento por detecção de proteína i.e. transfecção de clones de ADNc de linhas de células que expressam CD-4 e identificação por anticorpos dirigidos contra a proteína (ver Maddon, et al., supra).

Os clones de ADNc que portam a sequência de codificação da proteína sCD-4 particular podem ser identificados pelo uso de sondas de hibridação de oligonucleótido. As sondas podem ser concebidas com base na sequência particular da proteína. Tendo-se identificado um clone transportando a sequência de codificação de interesse, a sequência de codificação da proteína pode ser excisada pelo uso de endonucleases de restrição e inserida nos vectores de clonação e/ou expressão. Num vector de expressão a sequência de codificação da proteína sCD-4 está operativamente ligada a funções reguladoras necessárias ou desejáveis para a transcrição, tradução e processamento da sequência de codificação.

Na realização do presente invento em E. coli, uma sequência de codificação de ADN que codifica a proteína de ligação a HIV, p.e., sCD-4 ou seus derivados, é ligada operativamente a um

-10-

elemento regulador num vector de ADN para transformação de E. coli. Estão disponíveis numerosos vectores de expressão bacteriana gram-negativos compreendendo esses elementos reguladores. O elemento regulador compreende um promotor que efectua a ligação e transcrição da ARN polimerase. Preferem-se promotores reguláveis, i.e., indutíveis ou desrepressíveis. Está disponível para expressão de polipéptidos heterólogos em E. coli, uma variedade de promotores úteis. Estes incluem o promotor trp e o promotor lambda PL. Ver, p.e., Patente dos E.U. 4 578 355; Courtney et al., Nature 313:149 (1985). Muito preferivelmente, o elemento regulador compreende uma sequência de comando para transporte da proteína de ligação a HIV para a membrana da célula ou através dela, visto que se verificou, de acordo com este invento, que a proteína de ligação a HIV pode ser segregada. Um comando que se mostrou que transportava uma proteína para a membrana exterior de E. coli é o do gene de penicilinase de Bacillus licheniformis. Ver, Sarvas, et al., J. Bacteriol. 155:657 (1983).

É aqui particular o exemplo do comando do polissacárido A da membrana exterior designado aqui por comando ou sinal OMPA. Ver, Movva, et al., J. Molec. Biol. 143:317 (1980). A sequência de codificação de OMPA compreende uma sequência de comando que é útil para segregação da proteína de ligação a HIV no meio de fermentação. Verificou-se agora que o comando OMPA pode ser usado para segregar completamente uma série de proteínas de baixo peso molecular ou polipéptidos nos meios de cultura. Por exemplo, o sinal OMPA tem sido usado para obter exportações de grandes quantidades (p. ex. mais do que 1 a cerca de 10 mg/l) de factor alfa de crescimento transformante humano activo (TGF-alfa), factor de necrose de tumor (TNF) e interleucina-1-beta (IL-1-beta), em adição a SKBsCD-4, V1J4 e V1V2J4.

O sinal OMPA compreende essencialmente os seguintes 21 aminoácidos: N-met-lys-lys-thr-ala-ile-ala-ile-ala-val-ala-leu-ala-gly-phe-ala-thr-val-ala-gln-ala-C. Uma sequência de codificação de ADN para este comando pode ser sintetizada, preferivelmente de modo que um ou mais locais de restrição sejam colocados adjacentes ao aminoácido do C-terminal para inserção de uma sequência de

-11-

codificação de ADN desejada e de modo que um terminador seja colocado a jusante da sequência de codificação desejada. Nos Exemplos, abaixo, apresenta-se uma sequência de codificação, sintética, ilustrativa.

Em adição ao minigene da proteína de ligação a HIV, i.e. a sequência de codificação ligada operativamente a um elemento regulador, o vector usado para a transformação contém preferivelmente pelo menos um marcador de transformação fenotípica, i.e., um marcador de selecção genética, tal como um gene de resistência a antibiótico, um agente cromogénico ou um gene que complemente uma mutação auxotrófica. O vector também compreende regiões necessárias à replicação autónoma.

Construções ilustrativas de vectores de expressão úteis para expressão de proteínas sCD-4 em E. coli são descritos em detalhe nos Exemplos. As proteínas foram expressas no meio e foram aparentemente dobradas adequadamente.

Na realização do presente invento em Streptomyces, uma sequência de codificação de ADN que codifica a proteína de ligação a HIV, p.e., sCD-4 ou seus derivados, é ligada operativamente a um elemento regulador num vector de ADN para transformação de Streptomyces. O elemento regulador compreende um promotor que faz a ligação e transcrição da ARN polimerase. Preferem-se promotores reguláveis, i.e., indutíveis ou desrepressíveis. Está disponível para expressão de polipéptidos heterólogos em Streptomyces uma variedade de promotores úteis. Exemplos incluem o promotor indutível por galactose do operão galactose ^{de} Streptomyces (Fornwald, et al., Proc. Natl. Acad. Sci. USA 84:2130 (1987)), o promotor constitutivo do gene de β -galactosidase de S. lividans (Eckhardt, et al. J. Bacteriol. 169:4249 (1987); Brawner, et al., Patente E.U. 4 717 666) e do gene inibidor da tripsina de S. longisporus (Pedido de Patente Europeia Nº. 87 307 260.7), ou um promotor regulado temporalmente tal como o referido em M. echinospora (Baum, et al., J. Bacteriol. 170:71 (1988)). As regiões para a terminação de transcrição em Streptomyces são derivadas da extremidade 3' de vários genes de Streptomyces por exemplo o sinal de terminação da extremidade do operão de galactose de Streptomyces ou o encontrado na extremidade do gene da neomicina fosfotransferase de S.

-12-

fradiae (Thompson e Gray, Proc. Natl. Acad. Sci USA 80:5190 (1983)). As sequências para a exportação de proteína em Streptomyces incluem as isoladas do gene de β -galactosidase de S. lividans, do gene LEP-10 de S. lividans (Pedido de Patente Europeia Nº. 87 307 260.7) e do gene inibidor da tripsina de S. longisporus.

O gene de sCD-4 é incorporado numa molécula de ADN maior que compreende um sistema marcador de selecção genética. O sistema marcador de selecção pode ser qualquer um de vários sistemas marcadores conhecidos em que o gene marcador confere um novo fenotipo seleccionável à célula transformada. Os exemplos incluem os genes de resistência a droga de Streptomyces tal como metilase ribossómica resistente a tioestrepton (Thompson, et al., Gene 20:51 (1982)), fosfotransferase neomicina (Thompson, et al., supra) e metilase ribossómica resistente a ertromicina (Thompson, et al., supra). A molécula de ADN pode também conter uma sequência para replicação autónoma em Streptomyces, tal como os derivados de pIJ101 (Keiser, et al., Mol. Gen. Genet. 185:223 (1982)) ou um vector derivado de SLPl (Bibb, et al., Mol. Gen. Genet. 184:230 (1981)). A molécula de ADN também pode conter um marcador que permite amplificação do gene. Estes marcadores que servem para amplificar o número de cópias de gene em Streptomyces incluem o gene para resistência a espectinomicina (Hornemann, et al., J. Bacteriol. 169:2360 (1987)) e arginina auxotrófica (Altenbuchner, et al., Mol. Gen. Genet. 195:134 (1984)).

Para os estudos iniciais de expressão, a sequência de codificação de CD-4, mais completamente descrita nos Exemplos abaixo, foi fundida à sequência de sinal β gal numa posição localizada 8 aminoácidos a jusante do sítio de clivagem do péptido de sinal β gal. As funções de replicação em Streptomyces para este vector de expressão foram proporcionadas pelo plasmídeo pIJ702 (Katz, et al., J. Gen. Microbiol. 129:2703 (1983)), que é um derivado de pIJ101 (Keiser, et al., Mol. Gen. Genet. 185:223 (1982)). A estirpe hospedeira usada foi S. lividans 1326 do tipo selvagem (Bibb, et al., Mol. Gen. Genet. 184:230 (1981)). A proteína de fusão β gal-sCD-4 foi detectada tanto no sobrenadante da cultura como em células intactas. Outra série de experiências mostrou

-13-

que o promotor LTI podia ser usado para obter uma expressão mais eficiente das proteínas sCD-4.

Os vectores de expressão LTI/secreção são análogos ao sistema β gal por a expressão/^{do}gene ser dirigida pelo promotor LTI e a exportação da proteína ser facilitada pelo péptido de sinal LTI. As sequências de codificação para várias proteínas de ligação a HIV têm sido clonadas no vector LTI incluindo $V_1V_2J_4$ e V_1J_4 . As sequências de codificação para todas estas proteínas foram inseridas numa posição localizada 8 aminoácidos a jusante do sítio de clivagem do péptido de sinal LTI. Tal como as fusões β gal, as proteínas sCD-4 podem ser transportadas no sobrenadante da cultura via péptido de sinal LTI. As funções de replicação em Streptomyces foram de novo proporcionadas pelos derivados de pIJ101: pIJ702 e pIJ351 (Keiser, et al., supra).

A expressão de SKBsCD-4, V_1J_4 e $V_1V_2J_4$ a partir do sistema de expressão/^{LTI}foi inicialmente examinada em S. lividans 1326. Tanto SKBsCD-4 como $V_1V_2J_4$ estavam igualmente distribuídos entre o sobrenadante da cultura e as células intactas. Embora estas proteínas sejam incompletamente exportadas, as formas intra- e extracelular parecem ter o mesmo peso molecular quando avaliado pelas suas mobilidades em geles SDS-poliacrilamida. Não está claro se o péptido de sinal é removido proteoliticamente da forma intracelular. Os níveis máximos de SKBsCD-4 foram estimados em 30 mg/l no sobrenadante da cultura e nas células intactas. Os níveis de $V_1V_2J_4$ foram estimados em pelo menos 30 mg/l (sobrenadante da cultura e células intactas).

Um problema encontrado para a expressão de sCD-4 em S. lividans é a sua susceptibilidade a actividades da protease. Este problema pode ser eliminado ou pelo menos significativamente diminuído por expressão de sCD-4 em outras espécies de Streptomyces. A proteína sCD-4 pode ser expressa, por exemplo, em S. coelicolor, S. longisporus e S. albus. É esperado que a estabilidade da molécula melhorará usando estes hospedeiros alternativos. A expressão de proteínas sT4 em S. longisporus pode ser vantajosa uma vez que este hospedeiro segrega comparativamente menos proteínas no sobrenadante da cultura o que auxilia no processo de purificação.

-14-

Na realização do presente invento em levedura, uma sequência de codificação de ADN que codifica a proteína de ligação a HIV, p. ex., sCD-4 ou seus derivados, é operativamente ligada a um elemento regulador num vector de ADN para transformação da levedura. Pode-se usar qualquer hospedeiro levedura para o qual estejam disponíveis sistemas de transformação, clonagem e expressão. Exemplos particulares incluem leveduras dos géneros Hansenula, Pichia, Kluveromyces, Schizosaccharomyces, Candida e Saccharomyces. O hospedeiro levedura preferido é Saccharomyces cerevisiae.

O elemento regulador compreende um promotor que efectua a ligação e transcrição da ARN polimerase. Preferem-se promotores reguláveis, i.e., indutíveis ou desrepressores. Está disponível para expressão de polipéptidos heterólogos em levedura, uma variedade de promotores úteis. Estes incluem o promotor do gene da metalotioneína indutível por cobre/^(CUP 1) e o promotor constitutivo dos genes glicolíticos gliceraldeído-3 fosfato desidrogenase (TDH3) e álcool desidrogenase (ADH). As regiões para terminação transcricional em levedura são derivadas da extremidade 3' de qualquer dos vários genes de levedura por exemplo o gene para iso-1-citocromo C (CYC1).

O gene de sCD-4 é incorporado numa molécula de ADN maior que compreende um sistema marcador de selecção genética. O sistema marcador de selecção pode ser qualquer um de vários sistemas marcadores conhecidos, em que o gene marcador confere um novo fenotipo seleccionável à célula transformada. Os exemplos incluem genes de levedura para enzimas biossintéticas tais como fosforribosil antranilato isomerase (TRP1) ou orotidina-5'-fosfato descarboxilase (URA3) ou genes de resistência a drogas heterólogos tal como resistência a G418 ou resistência a benomil (BEN1). A molécula de ADN também pode conter uma sequência para replicação autónoma em levedura, tal como a região de levedura 2-micron-circulo ori ou uma região de replicação autónoma cromossómica (ARS) tal como ARS1 e um centromero de levedura (CEN), tal como CEN3, para permitir a replicação autónoma do plasmídeo.

No sistema de expressão preferido, a sequência de codifi-

-15-

cação da proteína de ligação a HIV é fundida a uma sequência de codificação que estabiliza a expressão visto que se descobriu que a SKBsCD-4 sózinha expressa muito pouco; esta expressão fraca po_{de} ser devida a transcrição e tradução ineficaz ou à rápida degra_{dação} da proteína. Em qualquer caso, foi demonstrado que o produ_{to} de uma fusão de gene, na qual a sequência de codificação de uma ubiquitina é fundida 5' à sequência de codificação de SKBsCD-4 é expressa a níveis relativamente elevados e é clivada para produzir ubiquitina e o produto de gene sCD-4.

Uma sequência de ADN de ubiquitina pode ser isolada a par_{tir} de levedura ou de outras células eucarióticas, incluindo célu_{las} de mamífero, por técnicas de clonação de gene standard. Alternativamente, pode ser sintetizada por técnicas conhecidas standard. A sequência de nucleótidos do gene de ubiquitina modular usada para ilustrar o presente invento é a descrita por Ecker et al., J. Biol. Chem. 262:3524 (1987). A sequência é essencialmente como se segue:

```
AATTCATTATGCAGATCTTCGTCAAGACGTTAACCGGTAAAACCATAACTCTAGA
AGTTGAATCTTCCGATACCATCGACAACGTTAAGTCGAAAATTCAAGACAAGGAA
GGCATTCCACCTGATCAACAAAGATTGATCTTTGCCGGTAAGCAGCTCGAGGACC
GTAGAACGCTGTCTGATTACAACATTTCAGAAGGAGTCGACCTTACATCTTGTCTT
AAGACTAAGAGGTGGTTGAGGTAC
```

Inserindo uma sequência de ADN para um péptido de sinal na extremidade 5' da sequência da proteína de ligação a HIV pode conseguir-se a secreção da proteína no meio de cultura. As se_{quências} de sinal úteis para este fim são, por exemplo, as do fac_{tor} de emparelhamento de levedura, factor alfa (Kurjan e Herskowitz, Cell 30:933 (1982)). Em adição, pode obter-se a expressão por in_{tegração} do módulo de expressão em levedura de sCD-4 no cromosso_{ma} da levedura, seguida por amplificação usando marcadores tais como o gene DHFR de levedura para seleccionar o número de cópia aumentado.

Em geral, as proteínas sCD-4 do invento podem ser purifi_{cadas} a partir dos meios de cultura esgotados usando várias técni_{cas} de purificação de proteína, por exemplo cromatografia de afi-

-16-

nidade; cromatografia de permuta iônica, cromatografia de exclusão por tamanhos, cromatografia hidrofóbica ou cromatografia de fase inversa.

A sCD-4, os seus fragmentos e derivados podem ser purificados por cromatografia de afinidade usando adsorventes gerais e específicos de grupo, por exemplo ligandos de afinidade por corantes ou de ligação de hidratos de carbono; ou usando ligandos que se ligam especificamente a sCD-4, por exemplo, anticorpo monoclonal ou proteína gp120 de HIV ou suas porções.

A purificação compreende, após crescimento das células num meio de crescimento selectivo a uma temperatura de cerca de 28 - cerca de 45°C: a clarificação do meio de fermentação e depois a separação da proteína de ligação a HIV de outras proteínas presentes no meio. No processo preferido as proteínas sCD-4 são purificadas a partir do meio de cultura usando uma série de passos cromatográficos que são baseados nas propriedades físicas da molécula de sCD-4.

Um processo alternativo de purificação das proteínas sCD-4 envolve o uso de anticorpos monoclonais dirigidos contra sCD-4. A proteína pode ser purificada num passo por passagens de meios de cultura clarificados através de um suporte de gel de afinidade ao qual fica ligado o anticorpo monoclonal dirigido contra a proteína. A proteína de interesse ligar-se-á à coluna no local de ligação do anticorpo enquanto que as proteínas contaminantes passam através da coluna. A proteína é então eluída da coluna sob condições que evitem a inactivação.

A caracterização das propriedades biológicas e imunológicas in vitro das proteínas sCD-4 indica que as proteínas têm valor na prevenção e tratamento da SIDA. Os estudos indicaram que as proteínas sCD-4 actuam como inibidor da disseminação extracelular e célula a célula do vírus. Uma vez que a sCD-4, seus fragmentos e derivados descritos neste invento mostraram bloquear a ligação do vírus a células alvo CD-4⁺ em cultura, crê-se que a administração destas proteínas a pessoas infectadas actuará de modo a inibir a disseminação extracelular do vírus. Portanto, as pro-



-17-

teínas sCD-4 têm valor como agente tanto profilático como terapêutico para o tratamento da SIDA ou do complexo relacionado com a SIDA (ARC).

Como agente profilático, as proteínas sCD-4 são administradas a indivíduos de alto risco relativamente à doença ou a indivíduos que apresentam exposição a HIV pela presença de anticorpos para o vírus. A administração de uma quantidade eficaz da proteína numa etapa precoce da doença ou antes do seu início actua de modo a inibir a infecção dos linfócitos CD-4⁺ por HIV. Como agente terapêutico a administração de sCD-4 a pessoas infectadas com HIV actua de modo a inibir a disseminação extracelular do vírus.

A fusão entre células infectadas por HIV e outros linfócitos CD-4⁺ também parece ser uma via da disseminação do vírus. Além disso, a fusão pode ser responsável, em parte, pelo enfraquecimento da função do linfócito CD-4⁺ e em última análise pela depleção de linfócitos CD-4⁺ em indivíduos infectados. A fusão das células é dependente tanto dos produtos de gene de envelope viral como do receptor de CD-4 e pode ser inibida pelo OKT-4A ou por anticorpos monoclonais similares (Mabs) (Sattentau, et al., Science 234:1120 (1986)). As proteínas sCD-4 aqui descritas interferem com a fusão das células e portanto diminuem a disseminação célula a célula do vírus e a perda da função de linfócito CD-4⁺.

O receptor de CD-4 é monomórfico e assim crê-se que as proteínas sCD-4 são um inibidor universal do vírus que reconhece o domínio de superfície do receptor de CD-4, incluindo todo o HIV.

As proteínas sCD-4 podem ser usadas em combinação com outros agentes, por exemplo, em associação com agentes dirigidos contra outras proteínas de HIV, tal como a transcriptase inversa, protease ou tat. Um agente terapêutico eficaz contra HIV deverá prevenir a transmissão da infecção, mediada por vírus e célula a célula. As proteínas sCD-4 também podem ser usadas em combinação com outros agentes anti-virais, por exemplo, azidotimidina (AZT).

As proteínas sCD-4 deste invento também têm utilidade como inibidores da função da célula CD-4⁺ por se ligarem a moléculas alvo extracelulares que normalmente interactivam com o domínio

-18-

de superfície do receptor de CD-4. Numerosos estudos sugerem um papel crucial para o receptor de CD-4 na imunotolerância, particularmente na patogénese e progressão de doenças auto-imunes e na rejeição de enxertos específica de hospedeiro.

Como agente profilático ou terapêutico, as proteínas sCD-4 são administradas parentericamente, de preferência intravenosamente. O agente pode ser administrado por infusão ou por injeção, p. ex., diariamente, semanalmente ou mensalmente. A quantidade e taxa de administração é seleccionada de modo a manter uma quantidade eficaz de proteína sCD-4 em circulação. Um modo alternativo de administração seria extracorporal, empregando proteínas sCD-4 como agentes de diálise.

A proteína sCD-4 deste invento também pode ser usada como reagente para identificar moléculas naturais, sintéticas ou recombinantes que actuam como agentes terapêuticos ou inibidores de interações de células CD-4⁺.

Por exemplo, as proteínas sCD-4 podem ser usadas em ensaios de rastreio, tal como ensaios para interacção de proteína medidos por metodologias baseadas em ELISA, para proporcionar um reagente solúvel em água, bioquimicamente puro que pode ser usado em combinação com outros reagentes para determinar competidores das interacções do domínio de superfície do receptor de CD-4. Por exemplo, uma vez que as proteínas sCD-4 se ligam à proteína env de HIV ou a misturas contendo proteínas env de HIV, as proteínas podem ser usadas para rastrear inibidores da ligação de vírus.

Com base em resultados in vitro que mostram que as proteínas sCD-4 se ligam a células que expressam proteínas env de HIV, estas proteínas também podem servir como moléculas de alvejamento selectivo para células infectadas por HIV in vivo. As proteínas sCD-4 como proteína veículo, específica, alvo, podem servir, por exemplo, como a proteína veículo de distribuição de agentes citotóxicos às células infectadas.

Em adição, com base em resultados que mostram que o receptor de CD-4 se associa especificamente a antígenos de MHC da classe II em células que apresentam antígenos como sugerido pe-

-19-

la restrição de classe de células CD-4⁺, as proteínas sCD-4 podem ser usadas em combinação com antigénios da classe II para testar relativamente a inibidores das interacções linfócito CD-4 - célula alvo. Em adição a estes exemplos, que são baseados em ensaios de ligação directa entre a proteína e as suas moléculas alvo, podem conceber-se ensaios mais complexos que assentam nas respostas bioquímicas do reconhecimento da proteína. Os mutantes de deleção podem ser usados em ensaios de diagnóstico para a detecção de proteínas CD-4 ou das moléculas com que eles interactivam. Por exemplo, a quantificação das células CD-4 e CD-4⁺ e dos anticorpos para CD-4 teria valor para diagnóstico de SIDA.

Em adição, as proteínas sCD-4 podem ser usadas para gerar novos reagentes de diagnóstico, por exemplo, Mabs ou outros tipos de moléculas para uso em ensaios imunológicos standard, i.e., ELISA, imunoensaios de captura, radio-imunoensaios. Uma vez que a sCD-4 apresenta o epitopo OKT-4, o OKT-4A e a maioria, se não a totalidade dos outros epitopos de superfície do receptor de CD-4, as proteínas são especialmente úteis em ensaios de imunodiagnóstico para a quantificação absoluta dos níveis de CD-4 num sistema.

O receptor de CD-4 reside em três ambientes químicos diversos: o oxidante, superfície hidrofílica da célula; a membrana hidrofóbica; e o citoplasma hidrofílico redutor. Estes ambientes diversos evitariam muito provavelmente o isolamento do receptor no seu estado completamente nativo. As proteínas sCD-4 consistem em apenas segmentos seleccionados do domínio extracelular, são segregadas como proteínas solúveis no sobrenadante da célula e a sua conformação parece imitar numa extensão significativa a superfície do domínio de superfície do receptor. Assim, as proteínas sCD-4 são adequadas para análise estrutural detalhada, em particular para cristalografia de raios X. A determinação da estrutura tridimensional das proteínas sCD-4 sozinhas ou numa forma complexada com outras moléculas interactivas poderia proporcionar a base para a concepção racional de antagonista e agonistas selectivos para sCD-4.

-20-

Exemplos

Os exemplos que se seguem são ilustrativos e não limitantes do invento. As enzimas usadas nas manipulações genéticas foram obtidas de fontes comerciais e foram usadas substancialmente de acordo com as instruções do vendedor. Excepto quando for indicado em contrário, os procedimentos foram realizados substancialmente como descrito por Maniatis, et al., Molecular Cloning, Cold Spring Harbor Laboratory, 1982.

Exemplo 1: Construção de vectores de expressão em E. coli

Usando manipulações de ADN recombinante, a sequência de codificação para a SKBsCD-4 de comprimento total e os mutantes de deleção V1V2J4 ou V1J4, seguidos por um código de terminação TAA em estrutura (ver sequência acima) foram colocados a jusante da sequência de sinal OMPA num vector de expressão em E. coli para criar plasmídeos OMPAST48BV1, OMPAV1V2 e OMPAV1. A construção destes vectores é como se segue:

Construção do plasmídeo JRT4: para criar o plasmídeo JRT4 o plasmídeo DSP1 (Pfarr, et al., DNA 4:461 (1985)) foi cortado com XhoI, a região anterior polyA de SV40 foi deleccionada e os locais XhoI foram preenchidos usando fragmento Klenow de ADN polimerase. A região de poliadenilação da hormona de crescimento bovina (BGH) (Pfarr, et al., DNA 5:115 (1986)) foi cortada com PvuII e KpnI e o local KpnI foi tornado rombo por tratamento com ADN polimerase de T4. Este fragmento de 230 pb foi ligado a DSP1 para criar DSP1BGH.

O DSP1BGH foi cortado com SmaI e SalI e a cassette galk (consistindo no promotor anterior de SV40, na região de codificação de galk e na região polyA de BGH) foi ligada em pUC19 (Yanisch-Peron, et al., Gene 33:103 (1985)) no local SalI usando um ligante sintético consistindo de uma extremidade SalI, um local BglII e uma extremidade SmaI. Esta ligação de três partes resultou no plasmídeo DSP1BZBGH.JT.

O DSP1BZBGH.JT, que foi cortado com StuI e BclI para deleccionar a região de codificação de galk, foi ligado a um fragmento de EcoRI (preenchido)-BamHI de 1,7 kb, contendo o ADNc de CD-4

-21-

do plasmídeo pT4B (maddon, et al., Cell 42:93 (1985) para criar o plasmídeo JRT4.

Construção do plasmídeo pUCsT4: para criar o plasmídeo pUCsT4, um fragmento HaeII e HpaII (1125 pb) do ADNc de CD-4 do plasmídeo pT4b foi ligado, através do uso de ligantes sintéticos, ao vector pUC18 que tinha sido cortado com KpnI e XbaI. A extremidade HaeII do ADNc de CD-4 foi ligada ao local KpnI de pUC18 usando um ligante sintético com uma extremidade KpnI e uma extremidade HaeII. A extremidade HaeII do ADNc de CD-4 foi ligada ao local XbaI de pUC18 usando um ligante sintético com uma extremidade HpaII e uma extremidade XbaI. Este ligante também inseriu um código de parada TAA após o nucleótido 1257 da região de codificação de CD-4. O plasmídeo resultante era o pUCsT4.

Construção de pST4sal: para criar o plasmídeo pST4 sal, o plasmídeo JRT4 foi cortado com BglII e SacI e foi isolado um fragmento de 959 pb (consistindo no promotor anterior de SV40 e nos primeiros 602 nucleótidos do ADNc de CD-4). O plasmídeo pUCsT4 foi cortado com SacI e XbaI e isolou-se um fragmento de 660 pb (consistindo no ADNc de CD-4 desde o nucleótido 603 a 1257 seguido pelo código TAA do ligante sintético). Estes dois fragmentos foram ligados em DSP18ZBGHJT que tinha sido cortado com BglII e XbaI para deleccionar o promotor anterior de SV40 e a região de codificação de CD-4 a todo o comprimento. O plasmídeo resultante era o pST4sal.

Construção de pST4DHFR: para criar o plasmídeo pST4DHFR, um fragmento BglII-BamHI contendo uma cassette de expressão β -globina DHFR foi ligada no local BamHI de pST4sal. A cassette de expressão de β -globina DHFR consiste em: o promotor de β -globina do rato (fragmento HincII de 550 pb do plasmídeo pPK288 (Berg, et al., Mol. Cell. Biol. 3:1246 (1983)) modificado na sua extremidade 5' com um ligante sintético para conter um local BglII; a região de codificação de DHFR de rato (fragmento HindIII (preenchido) de 735 pb do plasmídeo pSV2-DHFR (Subramani, et al., Mol. Cell. Biol. 1:854 (1981)); a região anterior polyA de SV40 NheI (preenchido)-BamHI (preenchido) de DSP1 (Pfarr, et al., DNA 4:461 (1985)); e a região terminadora de DHFR de rato (fragmento HindIII (pre-

-22-

enchido) de 907 pb de plasmídeo mDH9 (Frayne, et al., Mol. Cell. Biol. 4:2921 (1984)), modificada na sua extremidade 3' com um ligante sintético para criar um local BamHI.

Construção de psT4BBV1DHFR: para criar o plasmídeo psT4BBV1DHFR, o plasmídeo pST4DHFR foi cortado com EcoRI e XbaI e o fragmento mais pequeno contendo a região de codificação de sCD-4 foi deleccionado. O plasmídeo sT4sal foi cortado com XbaI e BbvI e isolou-se um fragmento de codificação de 1120 pares de bases contendo a sequência para o CD-4 solúvel menos a região de comando. Este fragmento foi ligado ao pST4DHFR cortado com EcoRI/XbaI, usando um ligante sintético com uma extremidade EcoRI, um local KpnI e uma extremidade BbvI. Este fragmento é compatível com a saliência BbvI no fragmento sCD-4 isolado acima. O plasmídeo resultante foi denominado pST4BBV1DHFR.

Construção de OMPAST4: para criar o plasmídeo OMPAST4, o plasmídeo OMPA.GS foi digerido com NcoI e SalI e o fragmento mais pequeno contendo a região ligante foi deleccionada. O OMPA.GS é um derivado de pAS1 (Rosenberg et al., Meth. Enzymol. 101:123 (1983); Patente dos E.U. 4 578 355) possuindo uma sequência sintética inserida no local NdeI na extremidade 3' da sequência de ligação do ribossoma cII. A sequência sintética compreende o comando OMPA seguido por uma sequência multiligante. A sequência sintética foi substancialmente a que se segue:

5'-T ATG AAA AAG ACA GCT ATC GCG ATT GCA GTG GCA CTG GCT
GGT TTC GCT ACC GTA GCG CAG GCC GGC TCT AGA GTC GAC CTA
GTT AAC TAG-3'

O plasmídeo pUCsT4 foi cortado com NcoI e SalI e isolou-se o fragmento de 1149 pares de bases contendo a sequência de sCD-4 (nucleótidos de CD-4 124-1257). Este fragmento foi ligado ao OMPA.GS cortado com NcoI/SalI para criar OMPAST4.

Construção de OMPAST4BbvI: para criar o plasmídeo OMPAST4BbvI, o plasmídeo OMPAST4 foi cortado com NaeI e XbaI. O fragmento mais pequeno resultante deste corte, contendo a região de codificação de sCD-4, foi deleccionado. O plasmídeo ST4BBV1DHFR foi cortado com KpnI e a saliência 3' resultante foi terminada de modo rombo

-23-

com ADN polimerase de T4. O ADN terminado de modo rombo foi então cortado com XbaI e o fragmento de 1124 pares de bases contendo os nucleótidos 145-1257 do ADNc de CD4 foi isolado. O fragmento isolado foi ligado ao plasmídeo OMPAST4 cortado com NaeI/XbaI para criar o plasmídeo OMPAST4BbVI.

Construção de pucST4184: para criar o plasmídeo pucST4184, um fragmento EcoRI-NheI do ADNc de CD-4 [um fragmento de 682 pb codificando os aminoácidos (-25) a (+176)] foi ligado ao ligante sintético sk727/725 (extremidades NheI e AvaI) no local NheI. O sk727/725 codifica para os aminoácidos de CD-4 177-183). A extremidade AvaI de sk727/725 foi ligada a um fragmento AvaI-XbaI de pucST4 (compreendendo os pb 1198-1257 do ADNc humano e codificando os aminoácidos 351-369 do receptor de CD-4 seguidos por um códon de terminação TAA). Esta sequência que é flanqueada por extremidades EcoRI e XbaI, foi inserida em pUC19 nas extremidades EcoRI e XbaI no poliligante pUC19 (sk727/725). O sk727/725 é substancialmente como se segue:

5' ctagctttccagaaggcc 3'
3' gaaaggtcttccggagcc 5'

Construção de pucST4106: para criar o plasmídeo pucST4106, um fragmento EcoRI-AvaII (compreendendo 1-413 pb, e codificando os aminoácidos (-)25 a 87) do ADNc de CD-4 foi ligado ao ligante sintético sk791/792 (extremidades AvaII-AvaI) no local AvaII. O sk791-792 codifica para os aminoácidos de CD-4 88-104. A extremidade AvaI de sk791/792 foi ligada a um fragmento AvaI-XbaI de pucST4 (compreendendo 1198-1257 pb do ADNc humano e codificando os aminoácidos 351-369 do receptor de CD-4 seguido pelo códon de terminação TAA). Esta sequência que é flanqueada por extremidades EcoRI e XbaI foi inserida em pUC19 nas extremidades EcoRI e XbaI no poliligante pUC19. O sk791/792 é substancialmente como se segue:

5' gaccagaaggaggagggtgcaattgctagtggttcggattgactgccaac
gtcttcctcctccacgttaacgatcacaagcctaactgacgggttgagc 5'

Construção de sT4184.DHFR: para criar o plasmídeo sT4184.DHFR, um fragmento EcoRI-XbaI de pucST4184 (codificando os aminoácidos

-24-

-25 a 183 fundidos a 351 a 369) foi ligado às extremidades EcoRI e XbaI de psT4.DHFR em lugar do fragmento EcoRI-XbaI codificando SKBsCD-4.

Construção de sT4106.DHFR: para criar o plasmídeo sT4106.DHFR um fragmento EcoRI-XbaI de pucST4106 (codificando os aminoácidos -25 a 104 fundidos a 351 a 369) foi ligado às extremidades EcoRI e XbaI de psT4.DHFR em lugar do fragmento EcoRI-XbaI codificando SKBsCD-4.

Construção de OMPAV1: para criar OMPAV1, o plasmídeo OMPAST4BbvI foi cortado com AflII e XbaI e o fragmento mais pequeno contendo os nucleótidos 371-1257 da região de codificação do ADNc de CD-4 foi deleccionado. O plasmídeo pst4106.DHFR foi cortado com AflII e XbaI e isolou-se um fragmento de 160 pares de bases contendo os nucleótidos 371-459, 1198-1257 da sequência de ADNc de CD-4. Os dois fragmentos foram ligados para criar o plasmídeo OMPAV1.

Construção de OMPAV1V2: para criar OMPAV1V2, o plasmídeo OMPAST4-BbvI foi cortado com SacI e XbaI e o fragmento mais pequeno contendo os nucleótidos 603-1257 da sequência de ADNc de CD-4 foi deleccionado. O plasmídeo ST4184.DHFR foi cortado com SacI e XbaI e isolou-se um fragmento de 164 pares de bases contendo os nucleótidos 603-696, 1198-1257 da sequência de ADNc de CD-4. Os dois fragmentos foram ligados para criar o plasmídeo OMPAV1V2.

Exemplo 2: Expressão de proteínas sCD-4 em E. coli

As construções OMPAV1 e OMPAV1V2 foram transformadas em lisogénios lambda de E. coli, AR58 (cI857) e AR120 usando procedimentos standard. As induções por calor (Rosenberg, et al., Meth. Enzymol. 101:123 (1983)) ou induções por ácido naladixico (Mott, et al., Proc. Natl. Acad. Sci. USA 82:88 (1985)) foram realizadas como descrito abaixo.

Indução por calor: para a linha de células AR58, 200 ml de LB (contendo 50 µg/ml de ampicilina) foram semeados com 7 ml de uma cultura nocturna (N.O.) de 10 ml de células AR58 contendo o plasmídeo OMPAST4BbvI, OMPAV1 ou OMPAV1V2 e cultivados numa incubadora com agitação a 32°C. A uma densidade óptica a 650 nm (O.D.₆₅₀) de 0,8 unidades de absorvância, 200 ml de LB, aquecidos a 55°C,

-25-

foi adicionado à cultura e o balão foi transferido para uma incubadora com agitação a 42°C. Tomaram-se amostras de um ml da cultura 30, 60 e 90 minutos após a alteração da temperatura para análise. Aos 90 minutos, as células foram arrefecidas durante 15 minutos em água gelada, e em seguida pelotizadas a 3500 rpm durante 15 minutos a 4°C numa centrífuga J-6 Beckman (Beckman Instruments, Fullerton, California). As pelotas de células foram armazenadas congeladas a -20°C até ao processamento. Os meios de condicionamento foram armazenados congelados a -20°C até ao processamento.

Indução com Ácido Nalidíxico: para a linha de células AR120, 400 ml de LB (contendo 50 µg/ml de ampicilina) foram semeados com 13 ml de uma cultura N.O. de 20 ml de células AR120 contendo o plasmídeo OMPAST4Bbv1, OMPAV1 ou OMPAV1V2 e cultivados numa incubadora com agitação a 37°C. A um O.D.₆₅₀ de 0,4, adicionaram-se à cultura 400 µl de ácido nalidíxico (50 mg/ml). A cultura foi mantida a 37°C numa incubadora com agitação e tomaram-se amostras de 1 ml, 1, 3 e 5 horas após a adição do ácido nalidíxico. À 5ª. hora, as células foram arrefecidas durante 15 minutos em água gelada e em seguida pelotizadas a 3500 rpm durante 15 minutos a 4°C numa centrífuga J-6 Beckman. As pelotas de células foram armazenadas congeladas a -20°C até ao processamento. Os meios condicionados foram armazenados congelados a -20°C até ao processamento.

A seguir à indução, as células foram pelotizadas a partir das amostras de 1 ml e analisadas relativamente à expressão das proteínas sT4 de comprimento total e mutante (SKBsCD-4, V1J4 e V1V2J4) por "Western blot", usando um anticorpo policlonal de coelho para uma forma desnaturada de T4 solúvel produzida em bactérias. Detectaram-se proteínas com os tamanhos apropriados, em lisados de células tanto em amostras induzidas por calor como por ácido nalidíxico. Os níveis detectados representam 1-5% da proteína total.

Processamento de Meios Condicionados: os meios condicionados congelados foram descongelados à temperatura ambiente e tomaram-se alíquotas de 30 ml e centrifugaram-se a 4°C durante 1 hora a 24 000 rpm num rotor de cuba oscilante Beckman SW28. A seguir à centrifugação, os sobrenadantes foram reunidos e concentrados 10-

-26-

-20 vezes a 4°C por filtração por pressão em membrana.

Os meios concentrados foram ensaiados em relação às proteínas SKBsCD-4, V1J4 ou V1V2J4 de comprimento total por análise "Western blot" (como acima). ELISA de competição e imunoprecipitação.

Em cada amostra, detectou-se proteína de tamanho apropriado nos meios condicionados concentrados por "Western blot" a níveis representando 1-5% (0,1-0,5 ng/ml de meios condicionados não concentrados) os detectados no interior da célula. Todas as 3 proteínas ligam OKT4A no ensaio ELISA de competição. As proteínas V1J4 e V1V2J4 foram ainda ensaiadas por imunoprecipitação com um painel de anticorpos monoclonais de CD-4 (Ortho, Raritan, New Jersey). A proteína expressa por OMPAV1, V1J4, foi imunoprecipitada por OKT4A e por OKT4D mas não por OKT4, OKT4B, OKT4C, OKT4E ou OKT4F. A proteína expressa por OMPAV1V2, V1V2J4, foi imunoprecipitada por todos os anticorpos monoclonais excepto por OKT4. Em adição, a afinidade desta proteína por OKT4C foi mais baixa do que para os outros anticorpos monoclonais.

Processamento de Pelotas de Células: as pelotas de células do ponto dos 90 minutos de indução por calor OMPAST4Bbv1/AR58, foi descongelada em gelo e ressuspensa em tampão (Tris-HCl 100 mM, pH 8, EDTA 0,5 mM e sacarose 0,5 M) a 4°C numa razão de pelota para tampão de 100 O.D.₆₅₀/ml. A lisozima (10 mg/ml em Tris-HCl 0,05 M, pH 8) foi adicionada até uma concentração final de 0,2 mg/ml e a suspensão foi incubada em gelo durante 15 minutos. Um volume igual de água gelada foi adicionado e a suspensão foi incubada em gelo durante 15 minutos. Adicionou-se PMSF (50 mM) até uma concentração final de 1 mM, seguido pela adição de MgCl₂ (1 M) até uma concentração final de 2 mM. A suspensão de células foi passada através de uma agulha de calibre 18, 3 vezes para reduzir a viscosidade e depois centrifugada a 20 000 x g durante 10 minutos a 4°C.

A seguir à centrifugação a pelota foi ressuspensa em Tris-HCl 20 mM, pH 8, EDTA 5 mM, β ME 5 mM, num volume equivalente ao anterior à centrifugação. A suspensão foi submetida a tratamento

-27-

ultrassônico ("sonicated") durante 60 segundos (rebentamentos de 10 segundos com intervalos de 30 segundos para arrefecimento) e depois centrifugada a 20 000 x g durante 15 minutos a 4°C. O ensaio "Western blot" da pelota e sobrenadante resultantes indicou que 70% da proteína SKBsCD-4 estava na fracção sobrenadante. A fracção sobrenadante foi centrifugada a 100 000 x g durante 60 minutos. A análise "Western blot" da pelota e sobrenadante resultantes indicou que <5% da proteína SKBsCD-4 tinha permanecido na solução. A fracção sobrenadante foi diluída com um volume igual de MES 40 mM pH 6,0 e carregada numa S Sepharose de 1 ml (Pharmacia, Piscataway, New Jersey) equilibrada com MES 20 mM pH 6,0. A coluna foi lavada com 10 ml de MES 20 mM pH 6,0. A SKBsCD-4 ligada foi eluída com MES 20 mM pH 6,0, NaCl 0,5 M, recolhendo-se fracções de 1 ml. A análise "Western blot" das fracções de passagem, lavagem e eluição indicou que <5% da SKBsCD-4 se ligou à coluna e se eluiu nas fracções 2 e 3.

As fracções de passagem, lavagem e eluição foram ensaiadas relativamente à SKBsCD-4 funcional por ELISA de competição. Tanto a fracção de passagem como a de eluição 2 continham SKBsCD-4 que se ligou a OKT4A.

Imunoprecipitação: os meios concentrados (100 µl) foram diluídos com um volume igual de tampão de precipitação (fosfato de sódio 10 mM, pH 7,5, NaCl 100 mM, 0,1% de NP-40, 0,5% de leite magro seco) e pré-clarificados com 3 µg de IgG de coelho durante 15 minutos a 4°C, seguindo-se a adição de 30 µl (volume de enchimento) de pérolas de proteína A-Sepharose (Pharmacia, Piscataway, New Jersey) durante 30 minutos a 4°C. Os sobrenadantes pré-clarificados foram incubados com 5 µg de anticorpos monoclonais OKT4, 4A, 4B, 4C, 4D, 4E, 4F, OKT8 (Ortho Pharmaceuticals), IgG de ratinho (Cooper Biomedical, Malvern, Pennsylvania) ou IgG de coelho anti-ratinho (Cooper Biomedical) durante 30 minutos a 4°C. A OKT4, 4A, 4C, OKT8, IgG de ratinho e a IgG de coelho anti-ratinho foram precipitadas por incubação com 20 µl (volume de enchimento) de pérolas de proteína A-Sepharose durante 30 minutos a 4°C. A OKT4B, 4D, 4E e 4F foram incubadas com 10 µg de IgG de coelho anti-ratinho durante 30 minutos a 4°C antes da adição das pérolas de proteína A

-28-

Sepharose. A seguir à precipitação, as pérolas foram lavadas duas vezes com 200 μ l de tampão de precipitação e uma vez com 200 μ l de tampão de precipitação sem NP-40 e leite magro seco. As pérolas lavadas foram levadas à ebulição durante 5 minutos em 20 μ l de tampão de amostra (Tris-HCl 125 mM, pH 6,8, 20% de glicerol, β ME (beta-mercaptoetanol) 1,4M). A seguir à ebulição, os sobrenadantes foram submetidos a electroforese num gel de SDS-poliacrilamida a 12,5% e analisados por "Western blot" como descrito acima.

Ambas as proteínas V1J4 e V1V2J4 foram imunoprecipitadas por OKT4A. A proteína expressa por OMPAV1 foi imunoprecipitada por OKT4A e OKT4D, mas não por OKT4, OKT4B, OKT4C, OKT4E ou OKT4F. A proteína expressa por OMPAV1V2 foi imunoprecipitada por todos os anticorpos monoclonais excepto por OKT4. Em adição, a afinidade desta proteína por OKT4C foi menor do que para os outros anticorpos monoclonais.

ELISA de Competição: revestiram-se placas de microtitulação (Flow Laboratories, McLean, Virgínia) durante a noite à temperatura ambiente com 100 μ l/cavidade de NaHCO_3 0,1M/ Na_2CO_3 (pH 9,4) contendo 150 ng de sCD-4 purificada.

Incubaram-se alíquotas (2,5-80 μ l) de meios concentrados ou extracto de células com 5 ng de OKT4A na presença de 1 mg/ml de albumina de soro bovino (BSA) num volume final de 100 μ l, durante a noite à temperatura ambiente. As amostras com menos de 100 μ l foram diluídas até 100 μ l com solução salina tamponada com fosfato (PBS). Após incubação, as placas foram decantadas, enxaguadas 3 vezes com PBS e secas. A cada cavidade adicionaram-se 380 μ l de PBS/0,5% de gelatina. As placas foram incubadas a 37°C durante uma hora.

A seguir à incubação a 37°C, as placas foram decantar, enxaguadas 3 vezes com PBS e secas por batida (patted). As incubações de amostras de 100 μ l contendo o anticorpo monoclonal foram adicionadas às cavidades individuais da placa preparada e incubadas durante 5 horas a 37°C.

A seguir à incubação, as placas foram decantadas, enxaguadas 5 vezes com PBS e secas por batida. Adicionaram-se a cada cavidade 100 μ l de uma solução contendo IgG anti-ratinho/cavalo

-29-

biotinilada. Esta preparação de anticorpo foi preparada adicionando uma gota de IgG anti-ratinho/cavalo (Vector Labs, Burlingame, California) a 10 ml de PBS contendo 1% de soro de cavalo. As placas foram incubadas durante 30 minutos a 37°C.

A seguir à incubação, as placas foram decantadas, enxaguadas 5 vezes com PBS e secas por batida. Adicionaram-se 100 µl de complexo ABC (Vector Labs, Burlingame, California) a cada cavidade. As placas foram então incubadas à temperatura ambiente durante 30 minutos.

As placas foram de novo decantadas, enxaguadas 5 vezes com PBS e secas por batida. Adicionaram-se a cada cavidade 100 µl de p-nitrofenilfosfato misto (2 mg/ml de p-nitrofenilfosfato (Sigma Chemical Co., St. Louis, Missouri) em bicarbonato de sódio 0,1M, pH 9,4, MgCl₂ 10 mM). As placas foram incubadas a 37°C no escuro durante 30 minutos. A seguir à incubação, as placas foram lidas a uma absorvância de 405 nm. As concentrações de sCD-4 foram determinadas contra uma curva standard gerada incubando quantidades conhecidas de sCD-4, purificada, com o anticorpo OKT4A. Verificou-se que ambas as proteínas (V1J4 e V1V2J4) se ligavam ao anticorpo monoclonal OKT4A neste ELISA de competição.

Exemplo 3: Construção de vectores de expressão em levedura

Os plasmídeos usados para construir vectores de expressão em levedura para expressão de proteínas sCD-4 são baseados no vector vaivém de levedura, básico de cópias elevadas pYSK102 (Rosenberg, et al., Genetic Engineering 8:151 (1986)). Todos estes contêm um fragmento EcoRI-PstI, de 823 pares de bases do gene TRP1 de levedura (Tsumper e Carbon, Gene 10:157 (1980)), que proporciona um marcador seleccionável; um fragmento EcoRI-PstI de 2,0 kb de levedura do círculo de 2 micra (Hartley e Donaldson, Nature 286:860 (1980)) que proporciona uma origem de replicação em levedura; uma porção de pBR322 (delecionada de BamHI a PvuII) (em Cloning Vectors, Pouwels, Enger, Valk, Eds., 1985), que proporciona a replicação e selecção em E. coli e um módulo promotor de levedura-terminador inserido entre os locais HindIII e BamHI de pBR322. O promotor é quer um fragmento BamHI-RsaI de 431 pares de bases de CUP1 derivado como um fragmento BamHI-EcoRI, de

-30-

pYSK12 (Butt, et al., Proc. Natl. Acad. Sci. USA 81:3332 (1984); Rosenberg et al., Genetic Engineering 8:151 (1986); Gorman et al., Gene 48:13 (1986)) quer uma sequência do promotor TDH3 de 1,1 kb (Holland, et al., J. Biol. Chem. 255:2596 (1980)) derivada como um fragmento BamHI-EcoRI do plasmídeo pYSK18 (Rosenberg, et al., supra; Gorman, et al., supra). O terminador é um fragmento KpnI-HindIII de 361 pares de bases do gene CYC1 (Smith, et al., Cell 16:753 (1979)). Na região entre o promotor e o fragmento de CYC1 podem ser inseridas várias sequências sintéticas como descrito abaixo.

Construção de pYSK147 e pYSK148: os plasmídeos pYSK137 (promotor de CUP1) e pYSK138 (promotor de TDH3) contêm uma sequência sintética de oligonucleótidos entre o local EcoRI na extremidade 3' da sequência do promotor e um local KpnI na extremidade 5' da sequência do terminador. Esta sequência sintética contém locais especiais de restrição XhoI, NcoI e SalI a montante dos códons de terminação de tradução em todas as três estruturas de leitura. O ligante sintético tem a sequência:

GAATTCTCGAGCCATGGCCAGTCGACGTAGTTAAGTAGGTACC

Um fragmento NcoI-SalI do plasmídeo pUCsT4, contendo as bases 124 a 1257 da sequência de sCD-4 foi inserido em pYSK137 ou pYSK138 cortado com NcoI-KpnI para produzir pYSK147 e pYSK148, respectivamente.

Construção de pYSK154 e pYSK155: os vectores para a expressão do gene de fusão de ubiquitina contêm um gene de ubiquitina de levedura, sintético, adaptado a cassete (Ecker, et al., J. Biol. Chem. 262:3524 (1987)) entre o promotor e o terminador de CYC1. A sequência de ubiquitina é idêntica à publicada por Ecker, et al., excepto no facto do fragmento AflII-KpnI terminal ter sido substituído por uma sequência de oligonucleótidos contendo um local de restrição NcoI na extremidade 3'. O fragmento NcoI-KpnI de pYSK147, contendo a sequência de sCD-4 desejada foi inserido entre este local NcoI e o local KpnI na extremidade 5' do fragmento CYC1, para dar pYSK154. Um derivado de pYSK154, no qual a sequência para os primeiros sete aminoácidos do terminal amino da sequência de sCD-4 em pYSK154 (as sequências de péptido de comando) foi deleccio-

-31-

nada, foi construído por clivagem de pYSK154 com KpnI, seguido por preenchimento com o fragmento Klenow de ADN polimerase I de E. coli (polIK) e inserção de uma sequência XpnI-XbaI de sCD-4 (bases 145 a 1257 de sCD-4) de psT4BBV1DHFR que tinha sido tratado com polIK para preencher as extremidades pegajosas. Este plasmídeo foi designado por pYSK155.

Construção de pYSK-159 e pYSK161: os vectores contendo uma sequência de sCD-4 truncada (bases 124-603) foram construídos inserindo uma sequência sintética de oligonucleótidos contendo um codão de terminação no local SacI da sequência sCD-4 (base 602). O ligante sintético (dupla cadeia) tem a sequência:

CTAAGCGGC

TCGAGATTCGCCGGC

Esta sequência sintética foi inserida em pYSK148 e pYSK154 cortados com SacI e NarI para dar pYSK159 e pYSK161, respectivamente. Assim, o ADN de CD-4 em pYSK159 tem os codões para os aminoácidos -9 (alanina) a 151 (leucina) (YV1).

Exemplo 4: Expressão de sT4 em culturas de leveduras

O ADN de plasmídeo (2-10 μ g/ml) foi usado para transformar levedura (Saccharomyces cerevisiae) usando técnicas standard (Sherman et al., Methods in Yeast Genetics, Cold Spring Harbor, 1986), seleccionando para prototrofia de triptofano. As estirpes empregues eram estirpes F762 (Mata, trp1 Δ 1, ura3-52) ou PEC1-88 (Mata, proA Δ ::URA3, ura3-52, trp1-289, his3). As culturas de transformantes seleccionados foram cultivadas em (Yeast Nitrogen Base (YNB, Difco, Detroit, Michigan) contendo suplementos (2% de glucose, 20 μ g/ml de uracil, 20 μ g/ml de histidina). Para plasmídeos utilizando o promotor CUP1, incluiu-se sulfato de cobre a 100 micromolar. As culturas foram colhidas em fase log tardia por centrifugação e foram lavadas e quebradas. O lisado de células foi analisado quanto a sCD-4 por ensaio "immunoblot" standard de proteína.

Em S. cerevisiae contendo pYSK147 e pYSK148, a proteína sCD-4, SKBYsCD-4, (aminoácidos -9 a 369) foi expressa a níveis

-32-

baixos (inferiores a cerca de 0,5 ng/l). Verificou-se que a actividade estava associada à fracção particulada, mas poderia ser facilmente solubilizada por extracção com detergentes não iónicos. A quantidade de SKBYsCD-4 (aminoácidos -9 a 369) em culturas de S. cerevisiae contendo pYSK154 estava entre cerca de 2 e cerca de 6 mg/litro (a 2×10^8 células/ml). Isto foi um melhoramento de 5 vezes a quantidade da proteína CD-4 expressa com pYSK147 e pYSK148, i.e. sem a presença da sequência estabilizante de ubiquitina. O tamanho estimado da proteína sCD-4 sintetizada foi idêntico em todos os casos, indicando que a porção ubiquitina da proteína tinha sido clivada de SKBsCD-4. Isto foi confirmado por procedimentos de "immunoblotting" usando um anticorpo para ubiquitina de levedura.

Os níveis de expressão das proteínas sCD-4 truncados (YV1) em pYSK159 e pYSK161 foram semelhantes às da proteína sCD-4 expressa como fusão de ubiquitina. A quantidade de sCD-4 detectada foi semelhante em cada uma das estirpes hospedeiras examinadas.

A SKBsYCD-4 da estirpe F762 contendo pYSK154 foi parcialmente purificada por passagem através de uma coluna de S-Sepharose. Um extracto em bruto, preparado em tampão Tris 10 mM, pH 6,0, 1% de Tween-20, foi centrifugado a 30 000 x g. O sobrenadante foi ajustado a pH 4,0 e foi aplicado a uma coluna S-Sepharose equilibrada em tampão de acetato 50 mM. A amostra foi eluída em lotes com tampão MES 50 mM [ácido 2-(N-morfolino)-etanossulfónico], pH 6,0) contendo NaCl 0,5 M. A presença de SKBYsCD-4 no eluído foi confirmada por ensaio de "immunoblot". O material era activo como foi determinado por um ensaio baseado num ELISA de competição utilizando SKBsCD-4 purificado de células CHO e anticorpo OKT4A, sugerindo que a conformação da SKBsYCD-4 derivada de levedura imita a do material derivado de células de mamífero (CHO).

Exemplo 5: Construção de Vectores de Streptomyces

Usando manipulações de ADN recombinante, segmentos da sequência de ADNc de CD-4 humana foram fundidos à sequência de sinal do gene do inibidor da tripsina de S. longisporus (LTI) e do gene da β -galactosidase (β gal) de S. lividans. Os minigenes de

-33-

sCD-4 foram unidos a plasmídeos de Streptomyces pIJ702 (Katz, et al., supra) e pIJ351 (Kieser, et al., supra) para criar plasmídeos pLTI:sT4/1, pLTI:sT4/7, pLTI:V1V2 e pBgal:sT4/7. A construção dos vectores foi a que se segue:

Construção de pLTI:sT4/1: O plasmídeo pLTI:sT4/1 foi construído a partir de três outros plasmídeos: pUCs/4, pIJ351 e pLTI450. A construção do plasmídeo pLTI450 é a que se segue:

Construção de pLTI450: um fragmento SacI-KpnI de 0,92 kb contendo o gene LTI foi inserido em pUC18 que tinha sido digerido com SacI e KpnI. Este plasmídeo, pLTI520, foi parcialmente digerido com EagI e totalmente digerido com SalI e depois ligado a um ligante sintético (dupla cadeia) que tinha extremidades EagI e SalI:

5'-GGCCGCCGCCCCCGCG

CGGOGGGGGCGCAGCT-5'

Os plasmídeos obtidos desta ligação foram rastreados para inserção do ligante sintético no local EagI localizado aproximadamente 0,5 kb do local SacI. Este local EagI está localizado no par de bases 86 em relação à extremidade 5' do gene LTI. O plasmídeo resultante contém o promotor LTI e a sequência de codificação para o péptido de sinal e o local de clivagem do péptido de sinal. Os locais EagI e SalI podem ser ambos usados para fundir minigenes de sCD-4 à sequência de sinal LTI.

Construção de pLTI:sT4/1: para criar o plasmídeo pLTI:sT4/1, a sequência de codificação de sCD-4 foi isolada num fragmento BbvI-PstI de aproximadamente 1,1 kb de pUCsT4. Este fragmento contém os pares de bases 149-1257 da sequência de ADNc de CD-4 humana. O plasmídeo pUCsT4 foi digerido com BbvI seguido por tratamento com transcriptase inversa para preencher as extremidades protuberantes 5'. O ADN foi então digerido com PstI. Este fragmento BbvI-(RT)-PstI de 1,1 kb, que codifica SKBsCD-4, foi ligado a pLTI450 que tinha sido tratado com SalI e transcriptase inversa e em seguida digestão com PstI. Esta ligação resultou no plasmídeo pLTI450sT4/1. O passo final na construção de pLTI:sT4/1 requereu a inserção de funções de replicação de Streptomyces em pLTI450sT4/1. Isto foi conseguido digerindo pIJ351 (Kieser, et al., supra) com PstI e ligando-o a pLTI450sT4/1 digerido com PstI.

Construção de pLTI:sT4/7: o plasmídeo pLTI:sT4/7 foi construído a partir de três outros plasmídeos: pUCsT4, pIJ702 e pLTI450. Para criar o plasmídeo pLTI:sT4/7, a sequência de codificação de sCD-4 foi isolada num fragmento NcoI-PstI de 1,1 kb de pUCsT4. Este fragmento contém os pares de bases 124-1257 da sequência de ADNc de CD-4 humana. Estes pares de bases foram colocados entre um códão de iniciação ATG e um códão de terminação TAA como anteriormente descrito no Exemplo 1. O códão de iniciação forma parte da sequência de reconhecimento de NcoI. A sequência de codificação de sCD-4 contida neste fragmento contém 9 aminoácidos da sequência de comando de CD-4. O pUCsT4 foi tratado com NcoI e transcriptase inversa seguida por digestão com PstI. Este fragmento foi então ligado com pLTI450 que tinha sido tratado com HincII e PstI. Esta ligação resultou no plasmídeo pLTI450sT4/7. O replicação de Streptomyces foi proporcionado pelo plasmídeo pIJ702 (Katz, et al., supra). Foi inserido em pLTI450sT4/7 via os locais PstI únicos em ambos os plasmídeos para criar o plasmídeo pLTI:sT4/7.

Construção do plasmídeo pLTI:V1V2: O plasmídeo pLTI:V1V2 foi construído a partir dos plasmídeos psT4.184, pLTI450 e pIJ351. Isolou-se um fragmento XbaI-EcoRI de 0,54 kb a partir de psT4.184. Este fragmento foi tratado com BbvI e transcriptase inversa. A extremidade BbvI corresponde ao par de bases 149 da sequência de ADNc de CD-4 humana. Este fragmento terminado de forma romba foi então ligado a pLTI450 que tinha sido tratado com SalI e transcriptase inversa para criar o plasmídeo pLTI450V1V2. As funções de replicação de Streptomyces foram proporcionadas por pIJ351. Este plasmídeo foi clonado em pLTI450V1V2 via o local PstI único em ambos os plasmídeos pIJ351 e pLTI450V1V2. O plasmídeo resultante foi o pLTI:V1V2.

Construção de p β galsT4/7: O plasmídeo p β galsT4/7 foi construído a partir dos plasmídeos pUCsT4, pIJ702 e p3SSXMCP. p3SSXMCP é um derivado de pUC9 (Vieira and Messing, Gene 19:259 (1982) que contém o promotor de β -galactosidase e a sequência de sinal. Vinte seis pares de bases a jusante da sequência de codificação para o local de clivagem de péptido de sinal existe um local XmnI (Eckhardt, et al., J. Bacteriol. 169:4249 (1987)). Foi inserido neste local um ligante sintético que contém uma sequência de re-

conhecimento de BamHI (New England Biolabs, Beverly, Massachusetts) para gerar o plasmídeo p3SSX10.¹ Este vector foi tratado com BamHI e transcriptase inversa seguindo-se a digestão com XhoI. Foi ligado a este vector um fragmento que tinha uma extremidade NcoI tratado com transcriptase inversa para gerar uma extremidade romba e uma extremidade SalI. A ligação do local BamHI preenchido com o local NcoI preenchido recriou ambos os locais BamHI e NcoI. O plasmídeo resultante era p3SSXMCP. O pUCsT4 foi tratado com XbaI e transcriptase inversa seguindo-se digestão com NcoI. Este fragmento NcoI-XbaI(RT) de 1,1 kb foi então inserido em p3SSXMCP que tinha sido tratado com SacI e ADN polimerase I de T4 seguindo-se digestão com NcoI. As funções de replicação de Streptomyces foram proporcionadas por pIJ702 que foi inserido em p3SSXsT4 via o local BglII único em ambos os plasmídeos. O plasmídeo resultante foi o p β galsT4/7.

Exemplo 6: Expressão de minigenes de sCD-4 em Streptomyces: Os plasmídeos pLTI:sT4/1, pLTI:sT4/7, pLTI:V1V2 e p β gal:sT4/7 foram transformados em S. lividans 1326 (Bibb, et al., supra). Em adição os pLTI:sT4/1, pLTI:sT4/7, pLTI:V1V2 foram introduzidos em S. coelicolor M124 (Hopwood et al., Genetic Manipulation of Streptomyces - A Laboratory Manual, F. Crowe & Sons, Ltd., Norwich, England (1985)), S. albus J1074 (Chater and Wilde, J. Gen. Microbiol. 116:323 (1980)) e num derivado branco de S. longisporus (ATCC accession number 23931). Todos os plasmídeos foram introduzidos nos vários hospedeiros usando o procedimento anteriormente descrito por Thompson, et al., J. Bacteriol. 151:668 (1982). Os transformantes foram seleccionados por deposição nas placas de transformação 3 ml de agar a 0,4% contendo 100 μ g/ml de tioestreptona. Os clones que expressaram minigenes de CD-4 foram identificados propagando os transformantes resistentes a tioestreptona em dez mililitros de caldo de tripticase de soja mais 5 μ g/ml de tioestreptona durante 48-72 horas. Os sobrenadantes da cultura e os lisados de células foram separados num gel de poliacrilamida a 12% (30:0,8 acrilamida:bis) - dodecilsulfato de sódio a 1% e em seguida transferidos para nitrocelulose (Towbin, et al., Proc. Natl. Acad. Sci. USA 76:4350 (1979)). O filtro de nitrocelulose

-36-

foi sondado usando sCD-4-anti-soro de coelho como descrito por Deen, et al., supra.

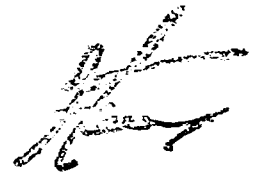
Os transformantes pLTI:sT4/1, pLTI:sT4/7 e p β gal:sT4 de S. lividans expressaram todos uma proteína de aproximadamente 45kD que é o tamanho previsto para a proteína sCD-4 codificada por esses plasmídeos. Similarmente, os transformantes pLTI:V1V2 de S. lividans expressaram uma proteína de aproximadamente 20 kD que é o tamanho previsto da proteína V1V2. A expressão de SKBsCD-4 e V1V2J4 foi também examinada em S. coelicolor, S. albus e S. longisporus. Tal como com S. lividans, as proteínas de 45 kD (os transformantes pLTI:sT4/1 e pLTI:sT4/7) e de 20 kD (os transformantes pLTI:V1V2) foram expressas por outras espécies de Streptomyces.

A quantificação subsequente de SKBsCD-4 presente nos sobrenadantes de cultura e lisados de células de S. lividans mostrou níveis de 1-5 mg/l. A V1V2J4 estava presente a 2-10 mg/l nos sobrenadantes da cultura e lisados de células. Menos do que 1 mg/ml foi expresso pelos transformantes p β gal:sT4 de S. lividans.

Usando uma linha de células pLTI:sT4/7 de S. lividans, a expressão de SKBsCD-4 foi vista num fermentador de 12 litros. De modo semelhante ao dos estudos de expressão em pequena escala, as células foram cultivadas em caldo de tripticase de soja mais 5 μ g/ml de tioestreptona. Os sobrenadantes da cultura e os lisados de células foram analisados quanto à expressão de SKBsCD-4 7,5, 24, 31,5, 54 e 72 horas após a inoculação. O nível de expressão volumétrico máximo foi de 60 mg/ml e ocorreu 24 horas após inoculação. O SKBsCD-4 foi igualmente distribuído entre o sobrenadante de cultura e células intactas.

Mostrou-se por ELISA de competição que os sobrenadantes parcialmente purificados contendo V1V2J4 e SKBsCD-4 se ligam a OKT4A indicando que estas proteínas derivadas de Streptomyces são proteínas de ligação a HIV.

A proteína de ligação a HIV, V1J4 de modo semelhante foi ligada ao promotor LTI e mostrou-se por ensaio western que é expressa em S. lividans.



-37-

Os Exemplos anteriores demonstram que as proteínas de ligação a HIV derivadas de CD-4, p.e., SKBsCD-4, SKBYsCD-4, V1J4, V1V2J4 e YV1, que foram expressas e segregadas pelos organismos recombinantes do invento são proteínas de ligação a HIV possuindo a função de ligação a HIV de sCD-4. Mais particularmente, elas têm o domínio de ligação a HIV da região V1 de SKBsCD-4.

A descrição e Exemplos anteriores revelam completamente o invento incluindo as suas concretizações preferidas. As modificações e melhoramentos às concretizações especificamente reveladas estão dentro do âmbito das reivindicações que se seguem.

R E I V I N D I C A Ç Õ E S

1 - Processo para a produção de uma proteína de ligação a HIV, caracterizado por compreender cultivar células de E. coli que foram transformadas com um vector compreendendo uma sequência de codificação para a proteína de ligação a HIV operativamente ligada a um elemento regulador e recolher daí a proteína de ligação a HIV.

2 - Processo para a produção de uma proteína de ligação a HIV caracterizado por compreender cultivar células de Streptomyces que foram transformadas com um vector compreendendo uma sequência de codificação para a proteína de ligação a HIV operativamente ligada a um elemento regulador e recolher daí a proteína de ligação a HIV.

3 - Processo para a produção de uma proteína de ligação a HIV, caracterizado por compreender cultivar células de levedura que foram transformadas com um vector compreendendo uma sequência de codificação para a proteína de ligação a HIV operativamente ligada a um elemento regulador, e recolher daí a proteína de ligação a HIV.

Lisboa, 20. FEV. 1989

Por SMITHKLINE BECKMAN CORPORATION

Rel - O AGENTE OFICIAL -

(S)
António Vissani