

(11) Número de Publicação: **PT 1510580 E**

(51) Classificação Internacional:

C12N 15/51 (2006.01) **A61K 39/29** (2006.01)
C07K 14/005 (2006.01) **C07K 14/18** (2006.01)
C12N 15/86 (2006.01) **A61K 39/00** (2006.01)

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

(22) Data de pedido: **1995.07.26**

(30) Prioridade(s): **1994.07.29 US 282959**
1995.07.25 US UNKNOWN

(43) Data de publicação do pedido: **2005.03.02**

(45) Data e BPI da concessão: **2008.12.10**
034/2009

(73) Titular(es):

NOVARTIS VACCINES AND DIAGNOSTICS, INC.
4560 HORTON STREET EMERYVILLE, CA 94608
US

(72) Inventor(es):

MICHAEL HOUGHTON **US**
MARK SELBY **US**

(74) Mandatário:

FRANCISCO NOVAES CUNHA BRITO SOTO MAIOR DE
ATAYDE
AV. DUQUE D'AVILA, N.º 32, 1º ANDAR 1000-141 LISBOA PT

(54) Epígrafe: **POLIPÉPTIDO E2 TRUNCADO DA HEPATITE C E MÉTODOS PARA A SUA**
OBTENÇÃO

(57) Resumo:

DESCRIÇÃO

| | |
|------------------|---|
| <u>EPÍGRAFE:</u> | <u>"POLIPÉPTIDO E2 TRUNCADO DA HEPATITE C E MÉTODOS PARA A SUA OBTENÇÃO"</u> |
|------------------|---|

Antecedentes da Invenção

Campo de Aplicação

A presente invenção refere-se de uma forma geral a proteínas virais. Em particular, a invenção refere-se a formas truncadas e secretadas das proteínas E1 e E2 do vírus da hepatite C e ao isolamento e produção recombinante das mesmas.

Antecedentes da Invenção

O Vírus da Hepatite C (VHC) é a principal causa da hepatite "não-A, não-B" transmitida por via parentérica, que é transmitida predominantemente por transfusões sanguíneas e contacto sexual. O vírus encontra-se presente em 0,4 a 2,0% dos doadores de sangue. A hepatite crónica desenvolve-se em aproximadamente 50% dos indivíduos infectados e, destes, aproximadamente 20% desenvolvem cirrose hepática que por vezes resulta em carcinoma hepatocelular. Por isso, o estudo e controlo da doença são importantes do ponto de vista médico.

A sequência do genoma viral do VHC é conhecida, uma vez que existem métodos para se obter a sequência. Veja-se, p. ex., as patentes internacionais WO 89/04669; WO 90/11089; e WO 90/14436. Em particular, o VHC possui um genoma de ARN de cadeia simples de sentido positivo com 9,5 kb e é um vírus da família Flaviridae. Actualmente, existem 6 genótipos distintos, mas relacionados, do VHC com base em análises

filogenéticas (Simmonds et al., J. Gen. Virol. (1993) 74:2391-2399). O vírus codifica uma única poliproteína com mais de 3000 resíduos de aminoácidos (Choo et al., Science (1989) 244:359-362; Choo et al., Proc. Natl. Acad. Sci. USA (1991) 88:2451-2455; Han et al., Proc. Natl. Acad. Sci. USA (1991) 88:1711-1715). A poliproteína é processada ao nível da co- e pós-tradução originando proteínas estruturais e não estruturais (NS).

Em particular, existem três proteínas estruturais putativas, que consistem na proteína N-terminal do nucleocapsídeo (denominado "core") e duas glicoproteínas do invólucro, "E1" (também conhecida como E) e "E2" (também conhecida como E2/NS1). (Veja-se, Houghton et al., Hepatology (1991) 14:381-388, para uma discussão sobre as proteínas do VHC, incluindo a E1 e a E2.) A E1 é detectada como uma espécie proteica de 32 a 35 kDa e é convertida numa banda única sensível à endo H de aproximadamente 18 kDa. Em contraste, a E2 apresenta um padrão complexo após imunoprecipitação, consistente com a produção de várias espécies (Grakoui et al., J. Virol. (1993) 67: 1385-1395; Tomei et al., J. Virol. (1993) 67:4017-4026.). As glicoproteínas do invólucro do VHC, as E1 e E2, formam um complexo estável que pode ser identificado por co-imunoprecipitação (Grakoui et al., J. Virol. (1993) 67:1385-1395; Lanford et al., Virology (1993) 197:225-235; Ralston et al., J. Virol. (1993) 67: 6753-6761). As glicoproteínas E1 e E2 do VHC são de grande interesse pois estudos com primatas revelaram que podem ter um efeito protector. (Choo et al., Proc. Natl. Acad. Sci. USA (1994) 91:1294-1298).

O invólucro do virião do VHC ainda não foi caracterizado. Assim, estudos de expressão recorrendo a moldes de cADN recombinante são os únicos métodos actualmente disponíveis para se estudar a biossíntese do invólucro. A E1 e E2 são

retidas no interior das células e não possuem hidratos de carbono complexos quando expressas de forma estável ou num sistema de expressão transiente que utilize o vírus da vacina (Spaete et al., *Virology* (1992) 188:819-830; Ralston et al., *J. Virol.* (1993) 67:6753-6761). Uma vez que as proteínas E1 e E2 se encontram normalmente ligadas às membranas nesses sistemas de expressão, seria desejável a produção de formas que fossem secretadas de forma a possibilitar a purificação das proteínas para posterior utilização.

Observou-se que a remoção do domínio transmembranar das glicoproteínas da superfície celular do vírus da gripe (Sveda, et al., *Cell* (1982) 30:649-656; Gething e Sambrook, *Nature* (1982) 300:598-603) e do vírus da estomatite vesicular (Rose e Bergmann, *Cell* (1982) 30:753-762) resulta na secreção da glicoproteína truncada pelas células mamíferas hospedeiras. Veja-se também a patente do IEP N°. 139 417. Da mesma forma, a gH (glicoproteína) truncada do citomegalovírus é secretada quando expressa em células de baculovírus (Publicação Internacional N° WO 92/02628, publicada a 20 Fevereiro, 1992). Foi descrita uma molécula da E2 truncada na zona C-terminal, que é secretada a partir de células mamíferas. Spaete et al., *Virology* (1992) 188: 819-830 Uma molécula da E2 truncada na zona C-terminal não possuindo uma porção do resíduo 717 foi revelada por Langford et al., *Virology* (1993) 197: 225-235. Esta proteína não é secretada a partir de células de insecto Sf-9. Contudo, ainda se desconhece a região de ancoragem transmembranar da E1 e portanto a produção de formas truncadas da E1 do VHC que sejam secretadas não foi até ao momento revelada. Além disso, não foram anteriormente descritos complexos de polipéptidos E1 e E2 truncados e secretados.

Sumário da Invenção

A presente invenção tem como base a elucidação da sequência da E2 que é importante para a ancoragem das proteínas ao retículo endoplasmático (RE) e para a co-precipitação da E2 com a E1. Assim, a eliminação destas sequências serve para produzir formas truncadas das glicoproteínas que são secretadas, em vez de serem retidas na membrana do RE. Além disso, a produção de formas truncadas possibilita a purificação da proteína E2 sem a proteína E1 associada, e vice-versa. A proteína E2 truncada, quando expressa juntamente com a E1 ou quando é com esta combinada após a expressão, é capaz de formar um complexo.

Assim, é aqui descrito para fins de referência um polipéptido E1 do VHC, sem a totalidade ou sem uma porção do seu domínio transmembranar para que o polipéptido possa ser secretado para o meio de cultura quando expresso de forma recombinante numa célula hospedeira. O polipéptido E1 do VHC não possui pelo menos uma porção do sua região C-terminal, começando no ou aproximadamente próximo do resíduo de aminoácido 370 ou do resíduo de aminoácido 360, numerados com referência à sequência de aminoácidos da E1 do VHC1.

Numa forma de realização, a invenção tem como objecto um polipéptido E2 do VHC que não possui pelo menos uma porção do seu domínio transmembranar para que o polipéptido possa ser secretado no meio de cultura quando expresso de forma recombinante numa célula hospedeira, em que o polipéptido não possui pelo menos uma porção da sua região C-terminal, começando no resíduo de aminoácido 715, numerado com referência à sequência de aminoácidos da E2 do VHC1. Também se descreve aqui um polipéptido E2 do VHC que não possui pelo menos uma porção da sua região C-terminal, começando aproximadamente no resíduo de aminoácido 725, numerado com referência à sequência de aminoácido da E2 do VHC1.

Também são aqui descritos os polinucleótidos que codificam os polipéptidos anteriormente referidos, os vectores que contêm esses polinucleótidos, as células hospedeiras transformadas com os vectores e os métodos para produzir de forma recombinante os polipéptidos.

Também se descreve aqui um complexo de E1 secretado/ E2 secretado compreendendo:

(a) um polipéptido E1 do VHC, sem a totalidade ou sem uma porção do seu domínio transmembranar para que esse polipéptido da E1 possa ser secretado para o meio de cultura quando expresso de forma recombinante numa célula hospedeira; e

(b) um polipéptido E2 do VHC, sem a totalidade ou sem uma porção do seu domínio transmembranar para que o referido polipéptido E2 possa ser secretado para o meio de cultura quando expresso de forma recombinante numa célula hospedeira.

O complexo E1 secretado/ E2 secretado inclui um polipéptido E1 que não possui pelo menos uma porção da sua região C-terminal, começando aproximadamente no aminoácido 360 ou 370, numerado com referência à sequência de aminoácidos da E1 do VHC1, e o referido polipéptido E2 não possui pelo menos uma porção da sua região C-terminal, começando aproximadamente no aminoácido 725 ou 730, numerado com referência à sequência de aminoácidos da E2 do VHC1.

Também são aqui descritas composições de vacinas que contêm os polipéptidos truncados E1 e/ou E2 do VHC e/ou complexos dos polipéptidos E1 e E2.

Breve Descrição das Figuras

A Figura 1 é uma representação esquemática da região E1 de VHC1.

A Figura 2 mostra a sequência nucleotídica e a sequência de aminoácidos correspondente para a E1 do VHC1, incluindo a sequência de sinal N-terminal e o domínio de ancoragem à membrana C-terminal.

A Figura 3 é uma representação esquemática da região E2 do VHC1.

As Figuras 4A a 4C mostram a sequência nucleotídica e a sequência de aminoácidos correspondente para a região E2/NS2 do VHC1, incluindo a sequência de sinal N-terminal para a E2 e o domínio de ancoragem à membrana C-terminal para a E2.

A Figura 5 representa os moldes de cADN da E1 do VHC utilizados para a transfecção descrita nos Exemplos. A região do core até NS2 é mostrada na parte superior e está representada à escala; a região distal NS3 até NS5 não está representada à escala. A região E1 foi aumentada para uma melhor representação dos moldes utilizados. Os números à direita referem-se ao aminoácido de corte utilizado em cada molde.

A Figura 6 representa moldes de cADN da E2 do VHC utilizados para a transfecção descrita nos Exemplos. A região do core até NS2 é mostrada na parte superior e está representada à escala; a região distal NS3 até NS5 não está representada à escala. A região E2/NS2 foi aumentada para uma melhor representação dos moldes utilizados. A coluna à esquerda refere-se ao aminoácido de corte utilizado em cada molde.

A Figura 7 representa o plasmídeo pMHE2-715, que contém um gene que codifica a proteína E2 truncada contendo os aminoácidos 383 a 715 do VHC1. O vector também possui a origem de replicação de SV40, o promotor "immediate early" de citomegalovírus de ratinho (MCMV ie), uma sequência líder de tpa (activador tecidual do plasminogénio), o sinal de poliadenilação de SV40 e cADN da DHFR como marcador de selecção.

Descrição Detalhada da Invenção

A abordagem prática seguida na presente invenção utilizará, excepto quando indicado em contrário, métodos convencionais de virologia, microbiologia, biologia molecular e técnicas de ADN recombinante conhecidas dos especialistas na área. Essas técnicas são explicadas detalhadamente na literatura. Veja-se, p. ex., Sambrook, et al. *Molecular Cloning: A Laboratory Manual* (2^o Edição, 1989); *DNA Cloning: A Practical Approach*, vol. I e II (D. Glover, ed.); *Oligonucleotide Synthesis* (N. Gait, ed., 1984); *Nucleic Acid Hybridization* (B. Hames e S. Higgins, eds., 1985); *Transcription and Translation* (B. Hames e S. Higgins, eds., 1984); *Animal Cell Culture* (R. Freshney, ed., 1986); *Perbal, A Practical Guide to Molecular Cloning* (1984); *Fundamental Virology*, 2^a Edição, vol. I e II (B.N. Fields e D.M. Knipe, eds.)

Conforme usado nesta especificação e nas reivindicações apenas, as formas singulares "um"/"uma" e "o"/"a" incluem referências plurais salvo se o conteúdo especificar claramente em contrário.

I. Definições

Ao descrever a presente invenção, serão utilizados os seguintes termos, que se querem definidos como abaixo indicado.

Um "polipéptido E1" significa uma molécula derivada de uma região da E1 do VHC. Essa molécula pode ser derivada fisicamente da região ou ser produzida de forma recombinante ou sintética, com base na sequência conhecida. A região da E1 madura do VHC1 começa aproximadamente no aminoácido 192 da poliproteína e continua até aproximadamente ao aminoácido 383 (veja-se as Figuras 1 e 2). Os aminoácidos desde a posição cerca de 173 até aproximadamente à 191 funcionam como uma sequência de sinal para a E1. Assim, um "polipéptido E1"

significa uma proteína E1 precursora, incluindo a sequência de sinal, ou um polipéptido E1 maduro que não possui esta sequência, ou mesmo um polipéptido E1 com uma sequência de sinal heteróloga. Além disso, como é aqui explicado, o polipéptido E1 inclui uma sequência C-terminal de ancoragem à membrana que ocorre aproximadamente na posição dos aminoácidos 360 a 383.

Um "polipéptido E2" significa uma molécula derivada de uma região da E2 do VHC. Essa molécula pode ser derivada fisicamente da região ou ser produzida de forma recombinante ou sintética, com base na sequência conhecida. Pensa-se que a região da E2 madura do VHC1 começa aproximadamente no aminoácido 384 a 385 (veja-se as Figuras 3 e 4A a 4C). Um péptido de sinal começa aproximadamente no aminoácido 364 da poliproteína. Assim, um "polipéptido E2" significa uma proteína E2 precursora, incluindo a sequência de sinal, ou um polipéptido E2 maduro que não possui essa sequência, ou mesmo um polipéptido E1 com uma sequência de sinal heteróloga. Além disso, como é aqui explicado, o polipéptido E2 inclui uma sequência C-terminal de ancoragem à membrana que ocorre aproximadamente na posição dos aminoácidos 715 a 730 e que pode prolongar-se até aproximadamente ao resíduo de aminoácido 746 (veja-se, Lin et al., J. Virol. (1994) 68:5063-5073).

As Figuras 2 e 4A a 4C representam regiões representativas da E1 e E2 do VHC1, respectivamente. No âmbito da presente invenção, as regiões E1 e E2 são definidas em relação ao número de aminoácido da poliproteína codificada pelo genoma do VHC1, sendo a metionina iniciadora designada como a posição 1. Contudo, é importante salientar que os termos um "polipéptido E1" ou um "polipéptido E2", conforme são aqui utilizados, não se limitam à sequência do VHC1. Em relação a isso, as regiões da E1 ou E2 correspondentes noutro isolado do VHC podem ser facilmente determinadas por meio do alinhamento das sequências

dos dois isolados de uma forma que produza um grau de alinhamento máximo entre as sequências. Isto pode ser realizado com um de vários pacotes de software informático, tal como o ALIGN 1.0, disponível na University of Virginia, Department of Biochemistry (À atenção do: Dr. William R. Pearson). Veja-se, Pearson et al., Proc. Natl. Acad. Sci. USA (1988) 85: 2444-2448.

Além disso, um " polipéptido E1" ou um "polipéptido E2" conforme aqui definido não está limitado a um polipéptido que tenha exactamente a mesma sequência que se encontra representada nas Figuras. Na realidade, o genoma do VHC encontra-se num estado de fluxo constante e contém vários domínios variáveis que exibem graus elevados de variabilidade entre isolados. Como se tornará aqui claro, o que realmente importa é que a região que serve para ancorar o polipéptido ao retículo endoplasmático seja identificada para que o polipéptido possa ser modificado no sentido de se remover a totalidade ou parte dessa sequência para que ocorra a secreção. É facilmente perceptível que os termos abrangem os polipéptidos E1 e E2 de qualquer um dos vários isolados do VHC, incluindo isolados de qualquer um dos 6 génotipos do VHC descritos em Simmonds et al., J. Gen. Virol. (1993) 74:2391-2399). Além disso, o termo abrange qualquer das proteínas E1 e E2, independentemente do método de produção, incluindo as proteínas produzidas de forma recombinante ou sintética.

Adicionalmente, os termos "polipéptido E1" e "polipéptido E2" englobam proteínas que contêm modificações adicionais à sua sequência nativa, tal como deleções, adições e substituições (geralmente de natureza conservadora) adicionais internas. Estas modificações podem ser deliberadas, como por meio de mutagénese dirigida, ou podem ser acidentais, tais como as que resultam de eventos mutacionais que ocorrem naturalmente. Todas essas modificações são abrangidas pela

presente invenção desde que os polipéptidos E1 e E2 modificados cumpram a sua finalidade. Assim, por exemplo, se os polipéptidos E1 e/ou E2 se destinam a ser utilizados em composições de vacinas, as modificações devem ser de modo a que não se perca a actividade imunológica (i.e., a capacidade para desencadear uma resposta de anticorpos contra o polipéptido). Do mesmo modo, se os polipéptidos se destinam a ser utilizados para fins de diagnóstico, deverão reter essa capacidade.

Um polipéptido E1 ou E2 "sem a totalidade ou sem uma porção do seu domínio transmembranar" é um polipéptido E1 ou E2, respectivamente, conforme acima definido, que foi truncado na região C-terminal para eliminar a totalidade ou uma parte da sequência de ancoragem à membrana cuja função é permitir a ligação do polipéptido ao retículo endoplasmático. Um polipéptido desse tipo é capaz de ser secretado para o meio de cultura em que um organismo que expressa a proteína é cultivado. A extensão da sequência de ancoragem à membrana que o polipéptido truncado não possui tem de ser apenas a suficiente para que se produza a secreção. É possível determinar facilmente a ocorrência de secreção para o meio de cultura usando-se várias técnicas de detecção incluindo, p. ex., electroforese em gel de poliácridamida e afins e técnicas imunológicas tal como ensaios de imunoprecipitação, conforme descrito nos exemplos. No caso de E1, de um modo geral, os polipéptidos que terminem aproximadamente na posição de aminoácido 370 e acima (com base na numeração da E1 do VHC1) serão retidos pelo RE e portanto não serão secretados para o meio de cultura. No caso de E2, os polipéptidos que terminem aproximadamente na posição de aminoácido 731 e acima (também com base na numeração da sequência da E2 do VHC1) serão retidos pelo RE e não serão secretados. É importante realçar

que estas posições de aminoácido não são absolutas e podem variar até certo ponto.

Embora não se tenha apresentado exemplos de todas as trunicações C-terminais possíveis, deverá ficar claro que as trunicações intermédias, tais como p. ex., os polipéptidos E1 que terminem nos aminoácidos 351, 352, 353 e assim por diante ou os polipéptidos E2 que terminem nos aminoácidos 716, 717, 718 e assim por diante, são também abrangidas pela presente invenção. Assim, tenciona-se englobar aqui todos os polipéptidos E1, que terminam aproximadamente nos aminoácido 369 e posições inferiores, e todos os polipéptidos E2, que terminam aproximadamente nos aminoácidos 730 e posições inferiores, que podem ser secretados para o meio de cultura quando expressos de forma recombinante.

Além disso, a trunicação C-terminal pode ser ampliada para além do domínio transmembranar na direcção da região N-terminal. Assim, por exemplo, trunicações do E1 que ocorrem nas posições inferiores, p. ex., a 360 e as trunicações do E2 que ocorrem nas posições inferiores, p. ex., a 715, são também abrangidas pela presente invenção. O necessário é que os polipéptidos E1 e E2 truncados sejam secretados e permaneçam funcionais para o fim a que se destinam. Contudo, prefere-se particularmente construções de E2 com trunicações C-terminais que não ultrapassem a posição de aminoácido 699.

Um "complexo de E1 secretada / E2 secretada" refere-se a um complexo das proteínas E1 e E2, sendo que cada uma delas não possui a totalidade ou uma porção do domínio transmembranar, conforme acima descrito. O modo de associação entre as proteínas E1 e E2 nesse complexo é irrelevante. Na realidade, esse complexo pode formar-se espontaneamente, bastando apenas misturar as proteínas E1 e E2 secretadas que tenham sido produzidas separadamente. Do mesmo modo, quando co-expressas, as proteínas E1 secretada e E2 secretada podem formar um

complexo espontaneamente no meio de cultura. A formação de um "complexo de E1 secretada / E2 secretada" é facilmente determinada usando-se técnicas correntes de detecção de proteínas tal como electroforese em gel de poliacrilamida e técnicas imunológicas tal como imunoprecipitação.

Duas moléculas de polinucleótidos ou proteínas são "substancialmente homólogas" quando pelo menos cerca de 40 a 50%, de preferência pelo menos cerca de 70 a 80%, e mais preferencialmente pelo menos cerca de 85 a 95%, dos nucleótidos ou aminoácidos das moléculas são iguais ao longo de uma extensão definida da molécula. Conforme é aqui utilizado, substancialmente homólogo também se refere a moléculas que tenham sequências que apresentem identidade com a molécula de ácido nucleico ou de proteína especificada. As moléculas de ácido nucleico que são substancialmente homólogas podem ser identificadas numa experiência de hibridação de Southern sob, por exemplo, condições restritivas, conforme definido para esse sistema em particular. A definição das condições apropriadas de hibridação está no âmbito das competências dos especialistas na área. Veja-se, p. ex., Sambrook et al., supracitado; DNA Cloning, vols I e II, supracitado; Nucleic Acid Hybridization, supracitado.

Uma proteína ou um polipéptido "isolada/isolado" é uma proteína que se encontra numa forma discreta e separada de um ser vivo completo com o qual a proteína se encontra normalmente associada na natureza. Assim, uma proteína contida num extracto sem células seria considerada uma proteína "isolada", como o seriam proteínas produzidas de forma sintética ou recombinante. De modo semelhante, um polinucleótido "isolado" é uma molécula de ácido nucleico numa forma discreta e separada de um ser vivo completo no qual a sequência se encontra na natureza; ou uma sequência que não possui, na totalidade ou em parte, sequências com as quais se

encontra normalmente associada na natureza; ou uma sequência, tal como existe na natureza, mas encontrando-se associada a sequências heterólogas (conforme abaixo definido).

Uma "sequência codificante" ou uma sequência que "codifica" uma proteína escolhida, é uma sequência de ácidos nucleicos que é transcrita (no caso do ADN) e traduzida (no caso do mRNA) num polipéptido *in vitro* ou *in vivo* quando colocada sob o controlo de sequências reguladoras apropriadas. Os limites da sequência codificante são determinados pelo códon de iniciação no (amino-)terminal 5' e por um códon de terminação da tradução no (carboxi-)terminal 3'. Uma sequência codificante pode incluir, mas sem que isso constitua uma limitação, cADN de sequências nucleotídicas virais assim como sequências de ADN sintético. Uma sequência de terminação de transcrição pode encontrar-se localizada na região 3' em relação à sequência codificante.

Um "polinucleótido" pode incluir, mas sem que isso constitua uma limitação, sequências virais, sequências procariotas, ARN viral, mRNA eucariota, cADN a partir de mRNA eucariota, sequências de ADN genómico de ADN eucariota (p. ex., mamífero), e mesmo sequências de ADN sintético. O termo também engloba sequências que incluem qualquer um dos análogos de bases de ADN ou ARN conhecidos, tais como, mas sem que isso constitua uma limitação, 4-acetilcitosina, 8-hidroxi-N6-metiladenosina, aziridinilcitosina, pseudoiso-citosina, 5-(carboxihidroxilmetil) uracilo, 5-fluorouracilo, 5-bromouracilo, 5-carboximetilaminometil-2-tiouracilo, 5-carboximetilaminometiluracilo, dihidro-uracilo, inosina, N6-isopenteniladenina, 1-metiladenina, 1-metilpseudouracilo, 1-metilguanina, 1-metilinosina, 2,2-dimetilguanina, 2-metiladenina, 2-metilguanina, 3-metilcitosina, 5-metilcitosina, N6-metiladenina, 7-metilguanina, 5-metilaminometiluracilo, 5-metoxi-aminometil-2-tiouracilo,

beta-D-manosilqueosina, 5'-metoxicarbonilmetiluracilo, 5-metoxiuracilo, 2-metiltio-N6-isopenteniladenina, éster metílico do ácido uracil-5-oxiacético, ácido uracil-5-oxiacético, oxibutoxosina, pseudouracilo, queosina, 2-tiocitosina, 5-metil-2-tiouracil, 2-tiouracilo, 4-tiouracilo, 5-metiluracilo, éster metílico do ácido N-uracil-5-oxiacético, ácido uracil-5-oxiacético, pseudouracilo, queosina, 2-tiocitosina, e 2,6-diaminopurina.

"Elementos de controlo" refere-se colectivamente a sequências promotoras, locais de ligação ao ribossoma, sinais de poliadenilação, sequências de terminação da transcrição, domínios reguladores a montante, estimuladores ("enhancers"), e afins, que colectivamente contribuem para a transcrição e tradução de uma sequência codificante numa célula hospedeira. Nem sempre é necessário que todos estes elementos de controlo estejam presentes num vector recombinante, desde que ocorra a transcrição e tradução do gene desejado.

Um elemento de controlo "dirigir a transcrição" de uma sequência codificante numa célula quando a ARN polimerase se liga à sequência promotora e transcreve a sequência codificante em mRNA, que é então traduzida no polipéptido codificado pela sequência codificante.

"Ligado funcionalmente" refere-se a uma disposição dos elementos em que os componentes descritos são configurados de forma a executar a sua função normal. Assim, os elementos de controlo ligados funcionalmente a uma sequência codificante têm a capacidade de causar a expressão da sequência codificante na presença da ARN polimerase. Não é necessário que os elementos de controlo sejam contíguos à sequência codificante, desde que funcionem de forma a dirigir a expressão da sequência. Assim, por exemplo, podem estar presentes sequências intermédias que não são traduzidas, mas que são transcritas, entre, p. ex., uma sequência promotora e

a sequência codificante e a sequência promotora continuar a ser considerada como estando "ligada funcionalmente" à sequência codificante.

"Recombinante", tal como é aqui utilizado para descrever uma molécula de ácido nucleico, significa um polinucleótido de origem genómica, em cADN, semi-sintética ou sintética que em resultado da sua origem ou manipulação: (1) não se encontra associado com a totalidade ou uma porção de um polinucleótido com o qual se encontra associado na natureza; e/ou (2) encontra-se ligado a um polinucleótido diferente daquele a que se encontra ligado na natureza. O termo "recombinante" quando utilizado em relação a uma proteína ou polipéptido significa um polipéptido produzido por meio da expressão de um polinucleótido recombinante. "Células hospedeiras recombinantes," "células hospedeiras," "células," "linhas de células," "culturas de células," e outros termos semelhantes que se referem a microrganismos procariotas ou linhas de células eucariotas cultivadas como entidades unicelulares, são utilizados de forma equivalente, e referem-se a células que podem ser, ou foram, usadas como células receptoras de vectores recombinantes ou outro ADN de transferência, e incluem a progénie da célula original que foi transfectada. Entende-se que a progénie de uma única célula parental pode não ser necessariamente na sua morfologia ou no seu ADN genómico ou na totalidade do complemento de ADN completamente igual à célula parental original, devido à possibilidade de mutações deliberadas ou acidentais. As progénies da célula parental que sejam suficientemente semelhantes à célula parental, a ser caracterizada de acordo com a propriedade relevante, tal como a presença de uma sequência nucleotídica que codifica uma proteína desejada, são incluídas na progénie a que se destina esta definição, e são abrangidas pelos termos acima referidos.

Um "vector" é um replicação ao qual se encontra ligado um segmento polinucleotídico heterólogo, para que ocorra a replicação e/ou expressão do segmento a ele ligado, tal como um plasmídeo, transposição, fago, etc.

Por "indivíduo vertebrado" entende-se qualquer membro do subfilo Chordata, incluindo, sem limitações, humanos e outros primatas, incluindo primatas não humanos tais como chimpanzés e outros símios e espécies de macacos; animais de criação tal como gado bovino, ovelhas, porcos, cabras e cavalos; mamíferos domésticos tais como cães e gatos; animais de laboratório incluindo roedores como ratinhos, ratos e cobaias; aves, incluindo aves selvagens e aves de caça e domésticas tais como galinhas, perus e outras aves galináceas, patos, gansos, e afins. O termo não se refere a uma idade em particular. Assim, tenciona-se englobar tanto os indivíduos adultos como os recém-nascidos. O sistema acima descrito destina-se a ser utilizado em qualquer uma das espécies de vertebrados anteriormente mencionadas, uma vez que os sistemas imunitários de todos estes vertebrados funcionam de forma semelhante.

II. Formas de Realizar a Invenção - Descritas apenas para fins de referência

A presente invenção tem como base a descoberta de novos polipéptidos E2 truncados na região C-terminal de tal forma que possuem a capacidade de ser secretados para o meio de crescimento quando produzidos de forma recombinante numa célula hospedeira. Surpreendentemente, os polipéptidos secretados conseguem formar complexos com polipéptidos E1. Esta interacção é inesperada uma vez que, como aqui se descreve, a capacidade para os polipéptidos E1 e E2 co-precipitarem desaparece após a eliminação do domínio transmembranar.

Em particular, a partir da análise de transfecções transientes de séries de moldes de diferentes extensões abrangendo a região E2/NS2 obteve-se evidência da existência de três espécies de E2 com diferentes regiões C-terminais. Uma forma revelada por esses estudos foi o polipéptido E2 que termina no aminoácido 729, enquanto as duas espécies maiores eram fusões com as proteínas NS2A e NS2A/NS2B, localizadas a jusante, que terminavam nos aminoácidos 809 e 1026, respectivamente. Usando os mesmos moldes de E2, identificou-se uma região do polipéptido E2 importante para a coimunoprecipitação de E1 e que também impossibilita a secreção de E2. Do mesmo modo, conseguiu-se identificar um domínio transmembranar de E1.

Mais especificamente, identificou-se aproximadamente nas posições de aminoácido 360 a 383 um domínio transmembranar que presumivelmente serve para ancorar o E1 à membrana do RE. Construiu-se uma série de polipéptidos E1 truncados na região C-terminal, que não possuem porções deste domínio transmembranar. (Veja-se a Figura 5). Descobriu-se que os polipéptidos E1 que terminam nas posições de aminoácido 370 e superior não são secretados para o meio de cultura quando expressos de forma recombinante.

Do mesmo modo, uma série de moléculas de polipéptidos E2 que incluem as truncações C-terminais, conforme representado na Figura 6, foi expressa de forma recombinante e testou-se o meio de cultura das células transformadas para a presença de polipéptidos E2 com o objectivo de se determinar que construções truncadas apresentavam a capacidade de secreção. Descobriu-se que as moléculas que terminam nas posições de aminoácido superiores a 730, não são secretadas para o meio de cultura e são presumivelmente retidas na membrana do RE. (Deve ser realçado que se observa uma pequena quantidade de secreção nas construções que terminam na posição de aminoácido 730).

Também foi observada uma relação inversa entre a secreção de E2 e a sua associação com E1. Portanto, os polipéptidos E2 que podem ser secretados não co-precipitam com E1 de lisados, enquanto os que não são secretados, co-precipitam.

Porém, surpreendentemente, quando as formas secretadas dos polipéptidos E1 e E2 são co-expressas, os polipéptidos secretados conseguem formar um complexo, detectável usando-se anticorpos contra quer E1 ou E2. A formação desses complexos é importante uma vez que demonstra que as regiões de E1 e E2 que são importantes para a sua interação são retidas, apesar da eliminação das regiões C-terminais responsáveis pela ancoragem membranar.

Estas descobertas proporcionam métodos eficientes para a purificação de E1, E2 e de complexos E1/E2 secretados, para utilização futura. Em particular, as proteínas secretadas são mais facilmente purificadas do que as proteínas expressas intracelularmente. Do mesmo modo, uma vez que, como acima descrito, se sabe que as proteínas E1 e E2 nativas formam um complexo, a invenção aqui descrita apresenta um método para se obter quer a proteína E1 ou E2 sem a outra proteína. Adicionalmente, caso se queira obter um complexo E1/E2, as proteínas secretadas podem ser co-expressas ou misturadas entre si (em meio de cultura ou na forma purificada ou semi-purificada), para se obter a formação espontânea do complexo.

Os polipéptidos E1 e E2 truncados podem ser produzidos usando-se uma variedade de técnicas. Por exemplo, os polipéptidos podem ser produzidos usando-se técnicas recombinantes, bem conhecidas na área. A este respeito, podem preparar-se sondas oligonucleotídicas com base nas sequências conhecidas do genoma do VHC que podem ser utilizadas para se sondar bibliotecas genômicas ou de cADN para a identificação dos genes de E1 e E2. Os genes podem ser posteriormente isolados usando-se técnicas correntes e pode, p. ex.,

utilizar-se enzimas de restrição para truncar o gene nas partes desejadas da sequência completa. Do mesmo modo, os genes de E1 e E2 podem ser isolados directamente a partir de células e tecidos que contenham esses genes, usando-se técnicas conhecidas, tal como extracção com fenol e a sequência pode ser depois manipulada para produzir as truncações desejadas. Veja-se, p. ex., Sambrook et al., supracitado, para uma descrição de técnicas utilizadas para se obter e isolar ADN. Por fim, os genes que codificam os polipéptidos E1 e E2 truncados podem ser produzidos sinteticamente, com base nas sequências conhecidas. A sequência nucleotídica pode ser preparada com os codões apropriados para a sequência específica de aminoácidos pretendida. De um modo geral, escolhem-se os codões que são mais adequados para o hospedeiro pretendido, no qual a sequência será expressa. A sequência completa é geralmente construída a partir dos oligonucleótidos sobrepostos preparados segundo métodos correntes e agrupados numa sequência codificante completa. Veja-se, p. ex., Edge (1981) *Nature* 292:756; Nambair et al. (1984) *Science* 223:1299; Jay et al. (1984) *J. Biol. Chem.* 259:6311.

Uma vez isoladas ou sintetizadas, as sequências codificantes das proteínas desejadas podem ser clonadas num qualquer vector ou replicação adequado para serem expressas. Os especialistas na área conhecem inúmeros vectores de clonagem, e a selecção de um vector de clonagem apropriado é uma questão de escolha. Exemplos de vectores de ADN recombinante para a clonagem e de células hospedeiras que os vectores podem transformar incluem o bacteriófago λ (*E. coli*), pBR322 (*E. coli*), pACYC177 (*E. coli*), pKT230 (bactérias gram-negativas), pGV1106 (bactérias gram-negativas), pLAFR1 (bactérias gram-negativas), pME290 (bactérias gram-negativas que não a *E. coli*), pHV14 (*E. coli* e *Bacillus subtilis*), pBD9 (*Bacillus*),

pU61 (*Streptomyces*), pUC6 (*Streptomyces*), YIpS (*Saccharomyces*), YCp19 (*Saccharomyces*) e vírus do papiloma bovino (células mamíferas). Veja-se, para referência geral, DNA Cloning: Vols. I e II, supracitado; Sambrook et al., supracitado; B. Perbal, supracitado.

Também se podem utilizar sistemas de expressão que utilizam células de insectos, tal como os sistemas com baculovírus, que são conhecidos dos especialistas na área e se encontram descritos, p. ex., em Summers e Smith, Texas Agricultural Experiment Station Bulletin No. 1555 (1987). Os materiais e métodos para sistemas de expressão que utilizam células de insecto / baculovírus encontram-se comercialmente disponíveis sob a forma de kit sendo comercializados, entre outros, pela Invitrogen, São Diego CA (kit "MaxBac").

Os sistemas virais, tal como o sistema de transfecção / infecção com base no vírus da vacina, conforme descrito em Tomei et al., J. Virol. (1993) 67:4017-4026 e Selby et al., J. Gen. Virol. (1993) 74:1103-1113, também poderá ser utilizados na presente invenção. Neste sistema, as células são primeiro transfectadas *in vitro* com um vírus da vacina recombinante que codifica a ARN polimerase do bacteriófago T7. Esta polimerase exibe uma impressionante especificidade, pois só transcreve moldes contendo promotores T7. Após a infecção, as células são transfectadas com o ADN de interesse, regulado por um promotor T7. A polimerase do vírus da vacina recombinante expressa no citoplasma transcreve o ADN transfectado em ARN que é então traduzido na proteína pelos elementos do sistema de tradução do hospedeiro. O método permite um elevado nível de produção transiente citoplasmática de grandes quantidades de ARN e do(s) seu(s) produto(s) de tradução.

O gene pode ser colocado sob o controlo de um promotor, de um local de ligação ao ribossoma (para expressão bacteriana) e, opcionalmente, de um operador (aqui referidos

colectivamente como elementos de "controlo"), para que a sequência de ADN que codifica o polipéptido E1 ou E2 pretendido seja transcrita para ARN na célula hospedeira transformada por um vector contendo esta construção de expressão. A sequência codificante pode conter ou não um péptido de sinal ou sequência líder. Na presente invenção pode utilizar-se tanto os péptidos de sinal que ocorrem naturalmente ou sequências heterólogas. As sequências líder podem ser removidas pelo hospedeiro por meio de processamento pós-traducional. Veja-se, p. ex., as Patentes US 4 431 739; US 4 425 437; US 4 338 397.

Poderá também ser desejável utilizar outras sequências reguladoras para permitir a regulação da expressão das sequências que codificam a proteína em relação ao crescimento da célula hospedeira. Essas sequências reguladoras são conhecidas dos especialistas na área, e exemplos incluem as sequências reguladoras que fazem com que a expressão do gene seja activada ou inactivada em resposta a um estímulo físico ou químico, incluindo a presença de um composto regulador. Também podem estar presentes no vector outros tipos de elementos reguladores, por exemplo, sequências estimuladoras.

As sequências de controlo e outras sequências reguladoras podem ser ligadas à sequência codificante antes da inserção num vector, tal como os vectores de clonagem acima descritos. Em alternativa, as sequências codificantes podem ser clonadas directamente num vector de expressão que já contém sequências de controlo e um local de restrição apropriado.

Em alguns casos pode ser necessário modificar a sequência codificante para que possa ser ligada às sequências de controlo na orientação apropriada, i.e., para manter a grelha de leitura correcta. Poderá também ser desejável produzir mutantes ou análogos da proteína E1 ou E2. Os mutantes ou análogos podem ser produzidos por meio da deleção de uma

porção da sequência que codifica a proteína, por meio da inserção de uma sequência e / ou por meio da substituição de um ou mais nucleótidos na sequência. As técnicas para a modificação de sequências nucleotídicas, tais como mutagênese dirigida, são bem conhecidas dos especialistas na área. Veja-se, p. ex., Sambrook et al., supracitado; DNA Cloning, Vols. I e II, supracitado; Nucleic Acid Hybridization, supracitado.

O vector de expressão é então utilizado para transformar uma célula hospedeira apropriada. São conhecidas na área várias linhas de células mamíferas disponíveis na American Type Culture Collection (ATCC), tal como, mas sem que isso constitua uma limitação, células de ovário de hamster chinês (CHO), células HeLa, células de rim de hamster lactente (BHK), células de rim de macaco (COS), células do carcinoma hepatocelular humanas (p. ex., Hep G2), células de Madin-Darby do rim bovino ("MDBK"), assim com outras. Do mesmo modo, poderão utilizar-se hospedeiros bacterianos, tais como *E. coli*, *Bacillus subtilis*, e *Streptococcus* spp., nas construções da presente invenção. As leveduras que podem ser úteis como hospedeiros na presente invenção incluem, entre outros, *Saccharomyces cerevisiae*, *Candida albicans*, *Candida maltosa*, *Hansenula polymorpha*, *Kluyveromyces fragilis*, *Kluyveromyces lactis*, *Pichia guilliermondii*, *Pichia pastoris*, *Schizosaccharomyces pombe* e *Yarrowia lipolytica*. As células de insecto que podem ser utilizadas com os vectores de expressão de baculovírus incluem, entre outras, *Aedes aegypti*, *Autographa californica*, *Bombyx mori*, *Drosophila melanogaster*, *Spodoptera frugiperda*, e *Trichoplusia ni*. Dependendo do sistema de expressão e do hospedeiro escolhido, as proteínas da presente invenção são produzidas crescendo-se as células transformadas por um vector de expressão acima descrito sob condições em que a proteína de interesse é expressa. A proteína é então isolada a partir das células hospedeiras e

purificada. Uma vez que a presente invenção proporciona a secreção dos polipéptidos E1 e E2, as proteínas podem ser purificadas directamente a partir do meio de cultura. A escolha das condições de cultura apropriadas e os métodos para a recuperação dos polipéptidos são conhecidos dos especialistas na área.

Os polipéptidos E1 e E2 da presente invenção podem ser também produzidos usando-se métodos convencionais de síntese de proteínas, com base nas sequências de aminoácidos conhecidas. Em geral, estes métodos utilizam a adição sequencial de um ou mais aminoácidos a uma cadeia peptídica em crescimento. Normalmente, o grupo amino ou o grupo carboxilo do primeiro aminoácido é protegido por um grupo protector adequado. O aminoácido protegido ou derivatizado pode então ser ligado a um suporte sólido inerte ou utilizado em solução, adicionando-se o aminoácido seguinte da sequência que possui o grupo complementar (amino ou carboxilo) adequadamente protegido, sob condições que permitam a formação de uma ligação amida. O grupo protector é então removido do resíduo de aminoácido recém-adicionado e adiciona-se o aminoácido seguinte (adequadamente protegido), e assim por diante. Após a ligação dos aminoácidos pretendidos na sequência correcta, quaisquer grupos protectores que permaneçam (e qualquer suporte sólido, se forem utilizadas técnicas de síntese em fase sólida) são removidos sequencialmente ou concomitantemente, obtendo-se assim o polipéptido final. Por uma modificação simples deste procedimento geral, é possível adicionar mais do que um aminoácido de cada vez a uma cadeia em crescimento, por exemplo, acoplado-se (sob condições que não provocam a racemização dos centros quirais) um tripéptido protegido a um dipéptido adequadamente protegido, formando-se, após a desprotecção, um pentapéptido. Veja-se, p. ex., J. M. Stewart e J. D. Young, *Solid Phase Peptide Synthesis*, 2^a Ed.,

Pierce Chemical Co., Rockford, IL (1984) e G. Barany e R. B. Merrifield, *The Peptides: Analysis, Synthesis, Biology*, editores E. Gross e J. Meienhofer, Vol. 2, Academic Press, Nova Iorque, (1980), pp. 3-254, em relação às técnicas para síntese de proteínas em fase sólida; e M. Bodansky, *Principles of Peptide Synthesis*, Springer-Verlag, Berlim (1984) e E. Gross e J. Meienhofer, Eds., *The Peptides: Analysis, Synthesis, Biology*, supracitado, Vol. 1, em relação à síntese clássica em solução.

Conforme se explicou anteriormente, a presente invenção também revela um método para a produção de complexos de E1 secretada / E2 secretada. Esses complexos são facilmente produzidos, p. ex., realizando-se a co-transfecção das células hospedeiras com construções que codificam as proteínas E1 e E2 truncadas. A co-transfecção pode ser realizada em *trans* ou em *cis*, i.e., por meio da utilização de vectores separados ou utilizando um vector único que contenha ambos os genes da E1 e E2. Se for realizada usando-se um único vector, ambos os genes podem ser regulados por um único conjunto de elementos de controlo ou, em alternativa, os genes podem estar presentes no vector em cassetes individuais de expressão, reguladas por elementos de controlo individuais. Após a expressão, as proteínas E1 e E2 secretadas irão associar-se espontaneamente entre si. Em alternativa, os complexos podem ser formados misturando-se as proteínas individuais que foram produzidas separadamente, quer na forma purificada ou semi-purificada, ou mesmo misturando-se os meios de cultura nos quais se cultivaram as células hospedeiras que expressam as proteínas.

Os novos polipéptidos E1 e E2 secretados da presente invenção, os seus complexos, ou os polinucleótidos que os codificam, podem ser utilizados para vários fins terapêuticos ou de diagnóstico. Por exemplo, as proteínas e os polinucleótidos podem ser usados numa variedade de ensaios,

para se determinar a presença de proteínas E1 e E2 numa amostra biológica para auxiliar no diagnóstico da doença do VHC. Os polipéptidos E1 e E2 e os polinucleótidos que codificam os polipéptidos podem também ser utilizados em composições de vacinas, individualmente ou em combinação, p. ex., em vacinas profiláticas (i.e., para evitar a infecção) ou terapêuticas (para tratar o VHC após infecção). De facto, essas glicoproteínas do invólucro secretadas são particularmente úteis na imunização com ácidos nucleicos em que as moléculas que podem ser secretadas podem ser mais eficazes do que as correspondentes proteínas intracelulares na produção de uma resposta imunitária. As vacinas podem conter misturas de uma ou mais das proteínas E1 e E2 (ou sequências nucleotídicas que codificam as proteínas), tal como as proteínas E1 e E2 derivadas de mais do que um isolado viral. A vacina também pode ser administrada em conjunção com outros antigénios ou agentes imunoreguladores, por exemplo, imunoglobulinas, citoquinas, linfoquinas, e quemoquinas, incluindo, mas sem que isso constitua uma limitação, IL-2, IL-2 modificada (cys125→ser125), GM-CSF, IL-12, γ -interferão, IP-10, MIP1 β e RANTES.

As vacinas incluirão geralmente um ou mais "veículos ou excipientes farmacologicamente aceitáveis" tal como água, solução salina, glicerol, etanol, etc. Adicionalmente, substâncias auxiliares, tais como agentes emulsivos ou molhantes, substâncias tamponantes de pH, e afins, podem estar presentes nesses veículos.

Um veículo está opcionalmente presente que é uma molécula que não induz ela própria a produção de anticorpos nocivos para o indivíduo a quem é administrada a composição. Os veículos adequados são tipicamente macromoléculas grandes lentamente metabolizadas, tais como proteínas, polissacáridos, ácidos polilácticos, ácidos poliglicólicos, polímeros de

aminoácidos, copolímeros de aminoácidos, agregados lipídicos (tal como gotículas de óleo ou lipossomas), e partículas virais inativas. Esses veículos são bem conhecidos dos especialistas na área. Além disso, o polipéptido do VHC pode ser conjugado com um toxóide bacteriano, tal como o toxóide da difteria, tétano, cólera, etc.

Os adjuvantes também podem ser utilizados para aumentar a eficácia das vacinas. Esses adjuvantes incluem, mas sem que isso constitua uma limitação: (1) sais de alumínio (alum), tal como hidróxido de alumínio, fosfato de alumínio, sulfato de alumínio, etc.; (2) formulações de emulsões de óleo-em-água (com ou sem outros agentes específicos imuno-estimuladores, tais como os muramil péptidos (veja-se abaixo) ou componentes da parede celular bacteriana), tal como por exemplo (a) MF59 (Patente WO 90/14837), contendo Esqualeno 5%, Tween 80 0,5% e Span 85 (trioleato de sorbitano) 0,5% (contendo opcionalmente diferentes quantidades de MTP-PE (veja-se abaixo), embora isso não seja obrigatório) formulados em partículas submicrônicas usando-se um microfluidificador tal como o microfluidificador Modelo 110Y (Microfluidics, Newton, MA), (b) SAF (formulação de adjuvante Syntex), contendo esqualeno 10%, Tween 80 0,4%, polímero em bloco pluronic L121 5% e thr-MDP (veja-se abaixo) quer microfluidificado num emulsão submicrônica ou agitado em vortex para se produzir uma emulsão com partículas de maior dimensão, e (c) sistema adjuvante RibitTM (RAS), (Ribi Immunochem, Hamilton, MT) contendo Esqualeno 2%, Tween 80 0,2%, e um ou mais componentes da parede celular bacteriana escolhidos a partir do grupo que consiste em monofosforil lípido A (MPL), dimicolato de trealose (TDM), e esqueleto da parede celular (CWS), de preferência MPL + CWS (DetoxTM); (3) podem utilizar-se saponinas como adjuvantes, tal como StimulonTM (Cambridge Bioscience, Worcester, MA) ou utilizar partículas geradas a partir delas como ISCOMs (complexos

imuno-estimuladores); (4) Adjuvante Completo de Freund (CFA) e Adjuvante Incompleto de Freund (IFA); (5) citocinas, tal como as interleuquinas (IL-1, IL-2, etc.), factor estimulante de colónias de macrófagos (M-CSF), factor de necrose tumoral (TNF), etc.; e (6) outras substâncias que actuam como agentes imuno-estimuladores para aumentar a eficácia da composição. Prefere-se o alum e MF59.

Os muramil péptidos incluem, mas sem que isso constitua uma limitação, N-acetil-muramil-L-treonil-D-isoglutamina (thr-MDP), N-acetil-normuramil-L-alanil-D-isoglutamato (nor-MDP), N-acetilmuramil-L-alanil-D-isoglutaminil-L-alanina-2-(1'-2'-dipalmitoil-sn-glicero-3-hidroxifosforil-oxi)-etilamina (MTP-PE), etc.

Tipicamente, as composições das vacinas são preparadas sob a forma injectável, como soluções líquidas ou suspensões; também se podem preparar formas sólidas apropriadas para preparação de uma solução ou suspensão em veículos líquidos antes da injeção. A preparação também pode ser emulsionada ou encapsulada em lipossomas para um maior efeito adjuvante, conforme acima referido.

As vacinas deverão conter uma quantidade terapêuticamente eficaz das proteínas E1 e/ou E2 truncadas, ou complexos das proteínas, ou sequências nucleotídicas que as codificam, e qualquer um dos componentes acima mencionados, conforme necessário. Por "quantidade terapêuticamente eficaz" entende-se uma quantidade de uma proteína E1 e/ou E2 truncadas que irá induzir uma resposta imunológica no indivíduo a que é administrada. Essa resposta irá geralmente resultar no desenvolvimento no indivíduo de uma resposta imunitária celular secretória e / ou uma resposta imunitária mediada por anticorpos após a administração da vacina. Normalmente, essa resposta inclui, mas não se limita, a um ou mais dos seguintes efeitos: a produção de anticorpos de qualquer uma das classes

imunológicas, tais como as imunoglobulinas A, D, E, G ou M; a proliferação de linfócitos B e T; o fornecimento de sinais de activação, crescimento e diferenciação para as células imunológicas; a expansão de células T auxiliares, e células T supressoras e/ou células T citotóxicas e/ou populações de células T $\gamma\delta$.

De preferência, a quantidade eficaz é suficiente para resultar no tratamento ou prevenção dos sintomas da doença. A quantidade exacta necessária irá variar dependendo no indivíduo a ser tratado; da idade e do estado geral do indivíduo a ser tratado; da capacidade do sistema imunitário do indivíduo para sintetizar anticorpos; do grau de protecção pretendido; do grau de severidade da condição patológica a ser tratada; do polipéptido específico do VHC que foi escolhido e do seu modo de administração, entre outros factores. Uma quantidade eficaz apropriada pode ser facilmente determinante por um especialista na área. Uma "quantidade terapeuticamente eficaz" situar-se-á num intervalo relativamente amplo que pode ser determinado por meio de ensaios de rotina.

Uma vez formuladas, as vacinas são convencionalmente administradas por via parenteral, p. ex., por injeção, subcutânea ou intramuscular. As formulações adicionais adequadas para outros modos de administração incluem formulações orais e pulmonares, supositórios e aplicações transdérmicas. A posologia pode consistir num regime de uma dose única ou num regime de doses múltiplas. A vacina pode ser administrada em conjunção com outros agentes imunoreguladores.

Como foi explicado anteriormente, as vacinas que contêm polinucleótidos que codificam as proteínas truncadas secretadas pode ser usadas para imunização com ácidos nucleicos. Esse método compreende a introdução numa célula hospedeira de um polinucleótido que codifica um ou mais dos polipéptidos E1 e/ou E2, para se obter a expressão *in vivo* das

proteínas. O polinucleótido pode ser introduzido directamente no indivíduo receptor, por exemplo, por injeção, inalação ou métodos afins, ou pode ser introduzido *ex vivo*, em células retiradas do hospedeiro. Neste último caso, as células transformadas são reintroduzidas no indivíduo onde se pode desencadear uma resposta imunitária contra a proteína codificada pelo polinucleótido. Os métodos de imunização com ácidos nucleicos são conhecidos na área e são revelados p. ex., na Patente WO 93/14778 (publicada a 5 Agosto de 1993); Patente WO 90/11092 (publicada a 4 Outubro de 1990); Wang et al. Proc. Natl. Acad. Sci. USA (1993) 90:4156-4160; Tang et al. Nature (1992) 356: 152-154; e Ulmer et al. Science (1993) 259: 1745-1749. Geralmente, o polinucleótido é administrado sob a forma de um vector que foi encapsulado num lipossoma e formulado numa composição de vacina conforme acima descrito.

III. Procedimento Experimental

Seguem-se exemplos de formas de realização específicas da presente invenção. Os exemplos são apresentados apenas para fins exemplificativos e não se destinam a limitar de nenhuma forma o escopo da presente invenção.

Embora se tenha tentado garantir a precisão em relação aos valores usados (p. ex., quantidades, temperaturas, etc.), *dever-se-á*, naturalmente, ter em consideração que poderá existir algum erro e desvio experimental.

Materiais e Métodos

Células e expressão transiente

Utilizaram-se células BSC40 e fibroblastos de chimpanzé F503 (Perot et al., J. Gen. Virol. (1992) 73:3281-3284) em experiências de infecção/transfecção conforme descrito anteriormente (Selby et al., J. Gen. Virol. (1993) 74: 1103-1113). Em resumo, camadas de células subconfluentes em placas

de 60 mm foram infectadas com VV_{T7} (moi = 10) em DME sem soro. Após 30 a 60 minutos, o inóculo foi removido e substituído com os moldes apropriados de cADN clonados no vector pTM1 (Elroy-Stein e Moss, Proc. Natl. Acad. Sci. USA (1990) 87:6743-6747), conforme abaixo descrito. O ADN vector foi associado a Lipofectina (BRL) ou Lipofectamina (BRL). Após transfecção durante 2,5 a 3 horas, o ADN foi removido e as células foram submetidas a restrição de nutrientes durante 30 minutos em DME desprovido de met/cys. Adicionou-se aproximadamente 100 a 200 µCi de marcador ³⁵S-Express (NEN) às células durante 3 a 4 horas. As células foram lisadas em tampão de lise 1X (NaCl 100 mM, Tris-HCL 20 mM pH 7,5, EDTA 1 mM, NP40 0,5%, deoxicolato 0,5% e PMSF 100 mM, leupeptina 0,5 µg/ml e 2 mg/ml de aprotinina), armazenado em gelo durante 10 minutos e clarificado por centrifugação (15 000 x g durante 5 minutos). Os lisados foram imunoprecipitados com o anticorpo escolhido imobilizado em Proteína A-Sefarose (BioRad).

Moldes do VHC

Todos os moldes de E1 e E2 foram produzidos por PCR e confirmados por sequenciação. Usou-se o *primer* 5' contendo um resíduo de metionina e um local reconhecido pela *NcoI* juntamente com *primers* 3' que tinham um codão de terminação após o limite escolhido do invólucro e por fim, para o E1, um local reconhecido pela *BamHI*. Ambos os oligos possuíam sequências não específicas nas extremidades para possibilitar digestões mais eficientes com as enzimas *NcoI* e *BamHI*. Os fragmentos de PCR digeridos foram ligados ao vector pTM1 digerido com *NcoI/BamHI* (Elroy-Stein e Moss, Proc. Natl. Acad. Sci. USA (1990) 87:6743-6747). O vector pTM1 contém o promotor T7 e a sequência líder EMC proximal ao local de clonagem da *NcoI* que corresponde ao primeiro resíduo de met codificado pelo ADN escolhido. Os moldes de E2 foram digeridos com *NcoI* e

AscI e clonados no vector pTM1-CE2 digerido com *NcoI* (parcialmente)/*AscI* (Selby et al., J. Gen. Virol. (1993) 74:1103-1113) para produzir clones H em que as traduções começam no aminoácido 1 e codificam o core, a E1 e as regiões escolhidas da E2. No caso dos polipéptidos E1 truncados, os moldes codificantes começam por um resíduo de metionina, seguindo-se uma isoleucina e, então, o aminoácido 172. No caso das construções de polipéptidos E2 truncados, usou-se metionina na posição 364 como a extremidade N-terminal das construções. Após a identificação de clones possíveis e a amplificação do ADN, todos os clones foram sequenciados. A sequenciação revelou que todos os clones de E1 estavam correctos. A maioria dos clones de E2 apresentava a sequência correcta com a excepção da perda de um único resíduo de leucina no polipéptido E2. Este facto não influenciou as conclusões, uma vez que este resíduo não se encontrava na proximidade da extremidade C-terminal.

Imunoprecipitações

As imunoprecipitações foram realizadas nos meios de cultura de células transfectadas conforme descrito em Selby et al., J. Gen. Virol. (1993) 74:1103-1113, excepto que os precipitados foram lavados pelo menos uma vez com tampão de lise contendo NaCl 500 mM em vez de 100 mM. Os precipitados foram ressuspensos em tampão Laemmli, fervidos e analisados em geles de acrilamida a 12,5% ou 15%. O sinal radioactivo dos geles foi aumentado (Amplify, Amersham), e os geles foram secos e revelados após exposição a uma película a -80°C. Os antisoros usados foram já descritos por Selby et al., J. Gen. Virol. (1993) 74:1103-1113. Trataram-se os imunoprecipitados com Endo H de acordo com as instruções do fabricante (Oxford Glycosystems).

Para a imunoprecipitação das proteínas E1 e E2 secretadas, o meio de cultura obtido a partir de células transfectadas foi microcentrifugado e imunoprecipitado em tubos contendo anticorpo monoclonal anti-E1 ou anticorpo de coelho policlonal anti-E2 imobilizado em proteína G - sefarose (Pharmacia) ou proteína A - sefarose (Sigma), respectivamente. Após incubação durante a noite a 4°C, os complexos sefarose Ab-Ag foram lavados com tampão de lise duas vezes, uma vez com tampão de lise contendo NaCl 500 mM e por fim com Tris 120 mM, pH 8,0. Após a aspiração de todo o líquido, adicionou-se aproximadamente 30 µl de tampão de amostra Laemmli, as amostras foram fervidas e a seguir introduzidas num gel de acrilamida a 12,5%. Após a electroforese, os geles foram fixados, amplificados e secos com papel 3MM. Os geles foram revelados por exposição a uma película com ecrã intensificador. Os controlos da transfecção foram efectuados com um molde do cADN da β-galactosidase (pTM1-β-gal).

Sequenciação da NS2B

As células BSC40 foram transfectadas com pE2₁₀₀₆, marcado com ³⁵S-met(NEN) e foram lisadas com tampão de lise contendo inibidores da protease. Os lisados clarificados foram precipitados com anticorpo de coelho anti-E2 durante a noite, lavados e analisados num gel de acrilamida a 15%. O gel foi então transferido para PVDF em metanol a 20%/tampão de corrida 1x durante 2 horas a 50 volts. A região que contém a NS2B foi identificada por autoradiografia e foi cortada. A NS2B foi sequenciada utilizando-se a degradação sequencial de Edman num sequenciador de fase gasosa Applied Biosystems 470A (Speicher, (1989). Microsequencing with PVDF membranes: Efficient electroblotting, direct protein adsorption and sequencer program modifications. in Techniques in Protein Chemistry. Ed. T.E. Hugli. Academic Press, San Diego, CA. pp 24-35.). As

fracções que continham cloreto de butilo foram evaporadas e realizou-se a contagem com líquido de cintilação.

Exemplo 1 - Apenas para fins de referência

Secreção/Retenção de E1

Conforme se explicou anteriormente, foi produzida uma série de moldes de E1 no vector pTM1, usando-se PCR (Figura 5). Em particular, foram produzidos moldes codificantes que começavam por um resíduo de metionina, seguido de isoleucina e, então, do aminoácido 172 da poliproteína do VHC e continuando até ao aminoácido 330, e foram produzidos clones que possuíam incrementos de 10 aminoácidos até ao aminoácido 338. Os aminoácidos 173 até 191 correspondem à região C-terminal do core que aparentemente tem uma função como sequência de sinal. Pensa-se que a E1 madura começa no aminoácido 192 da poliproteína, após a clivagem da sequência de sinal.

Recuperou-se material imunoreactivo a anti-E1 no meio de cultura de células transfectadas com moldes que continham uma sequência até ao aminoácido 360. Não foi detectado E1 no meio de cultura de células transfectadas com clones que terminavam nos aminoácidos 370 e 380.

Exemplo 2

Secreção/Retenção de E2

Foi também produzida uma série de moldes de E2 no vector pTM1, usando-se PCR (Figura 6). Em particular, o primeiro aminoácido codificado de E2 corresponde à metionina na posição 364 que foi usado como a extremidade N-terminal nas construções. O aminoácido 364 corresponde ao começo aproximado do péptido de sinal de E2 (Hijikata et al., Proc. Natl. Acad. Sci. USA (1991) 88:5547-5551; Ralston et al., J. Virol. (1993) 67:6753-6761). Pensa-se que a E2 madura começa no aminoácido 385. As regiões C-terminais desfasadas variavam desde o

aminoácido 661 até ao 1006. Em particular, os clones terminavam nos aminoácidos 661, 699, 710, 715, 720, 725, 730, 760, 780, 807, 837, 906 e 1006.

Recuperou-se material imunoreactivo a anti-E2 no meio de cultura de células transfectadas com os moldes que continham uma sequência até ao aminoácido 725. Detectaram-se pequenas quantidades de E2 no meio de cultura de clones que terminavam no aminoácido 730. Detectou-se pouco ou nenhum E2 no meio de cultura das células transfectadas com clones que terminavam nos aminoácidos localizados depois da posição 730. A sequência precisamente antes da posição 730 é bastante hidrófoba, remanescente de uma sequência de ancoragem à membrana. Portanto, parece que as sequências entre 715 e 730 e que se prolongam até cerca do resíduo de aminoácido 746 (veja-se, Lin et al., J. Virol. (1994) 68:5063-5073) servem para ancorar E2 ao RE, impedindo a secreção.

Foi produzida uma molécula adicional de E2 secretada que incluía os aminoácidos 383_{ala} até 715_{lys}. A molécula de E2 foi expressa usando-se um sistema de expressão de células de ovário de hamster chinês / dihidrofolato reductase (CHO/DHFR) conforme descrito a seguir. Um fragmento de ADN da E2 do VHC desde o aminoácido 383 até ao aminoácido 715 do VHC1 foi produzido por PCR e a seguir ligado ao vector pMH que contém o promotor forte "immediate early" do citomegalovírus de ratinho (MCMVie) e um marcador de selecção DHFR, para originar o plasmídeo pMHE2-715 (veja-se a Figura 7). Este plasmídeo foi então transfectado de forma estável na linha de células CHO, Dg44, conforme se descreve a seguir. Juntaram-se 100 µg de ADN com lipofectina (Gibco-BRL) para transfecção em 1×10^7 células em placas de 10 cm. As células foram transfectadas e incubadas durante 24 horas em meio de cultura não selectivo (mistura de nutrientes F-12 de Ham, JRH Biosciences n° de série 51, prolina, glutamina e piruvato de sódio) e deixou-se a

recuperar durante 24 horas em meio de cultura não selectivo suplementado com soro bovino fetal. O meio de cultura não selectivo foi então substituído por meio de cultura selectivo (mistura de nutrientes F-12 de Ham, JRH Biosciences n° de série 52, prolina, glutamina, piruvato de sódio) suplementado com soro bovino fetal dializado. Mudou-se o meio de cultura em intervalos de 3 a 4 dias até se começarem a formar colónias.

A linha celular DG44 não possui actividade endógena da DHFR. Assim, apenas as células transfectadas com o plasmídeo pMHE2-715 que possui o gene da DHFR podem crescer e formar colónias no meio de cultura selectivo que não possui hipoxantina, aminopterinina e timidina. Cerca de 2200 colónias foram repicadas e cultivadas em placas com 96 poços. Quando os clones atingiram a confluência, 5 dias depois, o meio de cultura foi testado para o E2 secretado recorrendo-se a um ensaio ELISA em que se utilizou um anticorpo monoclonal 3E5-1, produzido contra um determinante linear de E2, como anticorpo de captura e E3 expresso em células CHO e purificado por meio de uma coluna de afinidade com um anticorpo monoclonal (5E5H7, reactivo contra um determinante conformacional de E2) como o padrão. Os 83 clones com níveis mais elevados de expressão foram expandidos em placas de 24 poços. Uma vez atingida a confluência nas placas de 24 poços, o meio de cultura foi analisado e os 41 clones com níveis mais elevados de expressão foram expandidos em placas de 6 poços. Quando os clones atingiram a confluência, o meio de cultura foi analisado e os 21 clones com níveis mais elevados de expressão foram expandidos em frascos de 75 cm². Nesta altura, a expressão foi confirmada por meio do ensaio ELISA e por imunocoloração com fluorescência utilizando-se o anticorpo monoclonal 3E5-1. Em 5 dos clones, 100% das células apresentaram fluorescência. Tendo como base ambos os ensaios, os 3 clones com expressão mais

elevada foram escolhidos para ser purificados por diluição limite.

Juntaram-se os 7 clones com níveis mais elevados de expressão e plaqueou-se em placas de 10 cm para avaliar o crescimento numa gama de concentrações de metotrexato. O MTX é um inibidor da DHFR. Numa população de células cultivadas, as variantes dos clones que sobre-expressam DHFR são resistentes aos efeitos tóxicos do MTX. Foi observado que os outros genes num plasmídeo transfectado em células CHO podem ser co-amplificados com o gene da DHFR. As células que expressam o E2 do VHC foram amplificadas em meio de cultura selectivo com concentrações finais de 10, 20, 50, 100, 200 nM de MTX. As colónias cresceram apenas no meio de cultura com 10 e 20 nM de MTX. Estes 384 clones foram repicados, expandidos e verificados para a expressão, tal como fora feito anteriormente na selecção inicial. Em culturas aderentes, os clones amplificados com níveis mais elevados de expressão apresentavam um nível de expressão 31% superior ao nível mais elevado de expressão de clones não amplificados e apresentavam 100% de fluorescência.

Exemplo 3 - Apenas para fins de referência

Co-immunoprecipitação de E2:E1

Observa-se a imunoprecipitação de E1 com E2 usando-se antisoro monoreactivo contra E1 ou E2 (Grakoui et al., J. Virol. (1993) 67:1385-1395; Lanford et al., Virology (1993) 197: 225-235; Ralston et al., J. Virol. (1993) 67:6753-6761). Esta interacção é muito forte pois é resistente a disrupção com SDS a 0,5% (Grakoui et al., J. Virol. (1993) 7:1385-1395) ou por detergente com concentração salina elevada / não iónico (Ralston et al., J. Virol. (1993) 67:6753-6761); Glazer e Selby, observações não publicadas). Para definir a região de E2 que é importante para esta interacção, células BSC40 com

moldes H (do core até E2) foram transfectadas e os lisados marcados radioativamente foram imunoprecipitados com antisoro policlonal de coelho anti-E2. As espécies de E2 codificadas por moldes que terminavam nos aminoácidos 730, 760 e 780 associaram-se a E1. Não se observou nenhuma diferença quantitativa na co-precipitação de E1 com todos os moldes entre 730 e 1006. Em contraste, espécies de E2 codificadas pelos moldes 661, 699, 710, 715, 720 e 725 não resultaram numa co-precipitação significativa de E1. Como controlo, procedeu-se à imunoprecipitação dos mesmos lisados com antisoro de pacientes LL. As proteínas E1 de todos os moldes exibiram uma precipitação nítida com LL; as proteínas E2 precipitaram eficientemente com a excepção do molde 661H. Tem-se observado consistentemente uma detecção relativamente fraca de E2₆₆₁, possivelmente devido a uma estrutura diferente desta proteína em comparação com as variantes mais longas.

Portanto, as sequências entre os aminoácidos 715 e 730 são importantes para uma associação eficiente de E1. A remoção desta região de ancoragem C-terminal parece impedir a associação que normalmente ocorre entre E1 e E2. Contudo, quando ambos os invólucros são secretados, parece que algum grau de associação ocorre. Este resultado é o oposto dos dados sobre a secreção relativos a E2 acima apresentados, estabelecendo-se assim uma relação inversa entre a secreção de E2 e a interacção de E2 com E1.

Exemplo 4 - Apenas para fins de referência Formação de Complexos de E1 e E2 Truncadas

Para testar se as formas truncadas de E1 e E2, produzidas conforme anteriormente descrito, conseguiam associar-se uma à outra, a construção que codifica um polipéptido E1, truncado na região C-terminal ao nível do aminoácido 360, foi co-transfectada com construções que codificam um polipéptido E2,

truncado na região C- terminal ao nível do aminoácido 715 ou um polipéptido E2, truncado na região C-terminal ao nível do aminoácido 725. As células foram transformadas conforme acima descrito e colheram-se os meios de cultura. Os complexos E1/E2 foram observados claramente com um anticorpo monoclonal anti-E2 em que E1 foi co-precipitado. A precipitação inversa com anticorpo monoclonal anti-E1 resultou em muito menos E1 em comparação com a precipitação anterior. Uma observação semelhante foi realizada por Ralston et al., J. Virol. (1993) 67:6753-6761, em relação a construções não truncadas. Essa formação do complexo é importante porque demonstra que as regiões de E1 e E2 relevantes para a sua interação foram mantidas, apesar da eliminação da totalidade ou de uma parte dos domínios transmembranares.

Exemplo 5

Várias espécies de E2: Estudos com Endo H

O tratamento de E2 com Endo H produziu três bandas de E2, o que sugere a existência de mais do que uma espécie de E2, embora os seus limites permaneçam indefinidos (Grakoui et al., J. Virol. (1993) 67:1385-1395). Além disso, uma banda de elevado peso molecular reactiva tanto com antisoro contra E2 e NS2 levou à especulação da existência de uma possível espécie E2-NS2 (Grakoui et al., J. Virol. (1993) 67:1385-1395; Tomei et al., J. Virol. (1993) 67:4017-4026). Para definir estas três glicoproteínas E2 utilizou-se o tratamento com endoglicosidase H. A Endo H é uma glicosidase adequada para estas experiências uma vez que nos sistemas de expressão transiente apenas se observam glicoproteínas imaturas com um teor elevado de manose (Spaete et al., Virology (1992) 188:819-830; Ralston et al., J. Virol. (1993) 67: 6753-6761). O tratamento com Endo H de proteínas E2 codificadas pelos moldes E2₆₆₁ até E2₇₃₀ produziu bandas individuais que

aumentaram de tamanho com os moldes mais longos. As proteínas E2 codificadas pelos moldes E2₇₆₀, E2₇₈₀ e E2₈₀₇ quando foram tratadas com Endo H produziram uma banda adicional cujo aumento de tamanho parou no aminoácido 807, em comparação com moldes mais longos. Estes resultados sugerem que uma forma de E2 termina próximo do aminoácido 730. Resultados recentes sugerem que a E2 pode terminar aproximadamente no aminoácido 746 (veja-se, Lin et al., J. Virol. (1994) 68:5063-5073). Uma segunda forma de E2 parece terminar próximo da posição de aminoácido 807. Nos moldes E2₉₀₆ e E2₁₀₀₆ observou-se uma terceira banda que migrou para uma zona mais distante do que o par anterior. Os tamanhos das bandas destes dois moldes eram consistentes com moléculas de E2 codificadas pelos moldes da sequência completa. A mobilidade da fusão E2-NS2 no pTM1-VHC apresentou uma ligeira redução em relação a E2₁₀₀₆. Estas observações são consistentes com a existência de uma terceira forma de E2 que termina no mesmo local da junção NS2/NS3 situada no aminoácido 1026 (Grakoui et al., Proc. Natl. Acad. Sci. USA (1993) 90:10583-10587).

A NS2 foi identificada como estando situada numa posição anterior à NS3 (Grakoui et al., J. Virol. (1993) 67:1385-1395; Hijikata et al., Proc. Natl. Acad. Sci. USA (1993) 90:10773-10777; Selby et al., J. Gen. Virol. (1993) 74:1103-1113). Os resultados aqui apresentados e outras publicações recentes (Lin et al., J. Virol. (1994) 68:5063-5073) apontam para a ocorrência de um processamento diferencial de precursores através do qual extensões de E2 para além dos aminoácidos 730 a 746 representam fusões de E2 com NS2. Sugerimos, de forma não conclusiva, que a região que codifica a pequena proteína localizada aproximadamente entre os aminoácidos 730 a 746 e 807 prevista pelas experiências com a Endo H corresponde à NS2A, enquanto a proteína a jusante é a NS2B. Esta nomenclatura tem como base a organização semelhante de

proteínas observadas entre os membros da família Flaviviridae. É possível que a NS2A não seja clivada de forma independente da poliproteína como acontece com a NS2B, e ainda está por determinar se a NS2A é de facto uma proteína não estrutural.

Exemplo 6 - Apenas para Fins de Referência

Co-imunoprecipitação de NS2B:E2 e a região N-terminal de NS2B

Por meio da utilização de soro de pacientes LL detectaram-se as glicoproteínas E2 e duas outras espécies de 14 kDa e 21 kDa obtidas por transfecção com os moldes E2₉₀₆ e E2₁₀₀₆. Estas últimas espécies correspondem à NS2B porque a diferença no tamanho está perfeitamente correlacionada com a diferença no comprimento dos moldes e detectou-se uma NS2B de 23 kDa a partir do molde completo de VHC, o pTM1-VHC (Selby et al., J. Gen. Virol. (1993) 74:1103-1113). Quando se utilizou um anticorpo anti-E2 para imunoprecipitar os mesmos lisados, a NS2B co-precipitou com E2. Sob estas condições também se observou a co-imunoprecipitação de E1. Além disso, também se observaram espécies com uma reactividade mais elevada a anti-E2 em precipitações com 906, 1006 e pTM1-VHC, que correspondem a fusões de E2 com a NS2A e a diferentes comprimentos da NS2B.

Usando-se os vários moldes de E2, foi possível definir uma região de E2 que é importante para a co-precipitação da NS2B. Moldes de E2 que terminavam nos aminoácidos 699, 730, 760, 780 e 807 foram co-transfectados com pTM1-NS25, que codifica os aminoácidos 730 até 3011 (Selby et al., J. Gen. Virol. (1993) 74:1103-1113). Usou-se o pTMI-NS25 pois já fora previamente determinado que este molde codifica a NS2 e era necessário utilizar um molde independente para a expressão da NS2B, de forma a definir as sequências importantes para a interacção entre E2:NS2B. O anticorpo anti-E2 causou a imunoprecipitação de proteínas E2 e também co-precipitou NS2B (23 kDa) e

provavelmente NS4B (27 kDa) de todos os lisados com excepção dos transfectados com o molde E2₆₉₉. Como controlo positivo, a NS2B truncada do molde E2₁₀₀₆ foi submetida a co-precipitação com E2. Estes resultados demonstram que os aminoácidos 699 até 730 são importantes para a associação da NS2B e NS4B com E2.

O anticorpo anti-E2 que co-precipitou NS2B foi produzido contra uma fusão SOD-E2 que terminava no aminoácido 662. Isto exclui a possibilidade de reactividade cruzada com NS2B como uma causa para a detecção de NS2B. A banda de NS2B não apresentava glicosilação, como revela a ausência de resposta ao tratamento com a endo H. Como a co-precipitação de NS2B foi eficiente, apresentando pouco ruído de fundo, esta propriedade de associação foi utilizada para sequenciar com isótopos radioactivos a região N-terminal da NS2B. O imunoprecipitado do lisado marcado com ³⁵S-metionina obtido a partir de uma transfecção com o molde E2₁₀₀₆ foi analisado num gel de acrilamida a 15% e transferido para PVDF para sequenciação. Usando-se como referência um resíduo de metionina situado a uma distância de 18 aminoácidos do resíduo 810, identificou-se o terminal amino da NS2B como estando localizado no aminoácido 810. Esta identificação está de acordo com o C-terminal previsto para a E2 com base no padrão obtido após tratamento com a endo H de moldes de E2 com deleções e vem confirmar os locais de clivagem recentemente descritos por Grakoui et al., Proc. Natl. Acad. Sci. USA (1993) 90:10583-10587 e Mizushima et al., J. Virol. (1994) 68:2731-2734/1994. Prevê-se que a provável região terminal da NS2A se localize nos aminoácidos 730 e 809, caso exista sob a forma de uma proteína independente.

Lisboa, 11 de Fevereiro de 2009

Reivindicações

1. Um método para a produção de um polipéptido do VHC secretado, caracterizado por o referido método compreender:

(a) proporcionar uma população de células hospedeiras de mamíferos ou levedura, em que as referidas células hospedeiras são transformadas com um vector recombinante que compreende:

i) um polinucleótido que compreende uma sequência codificante para o polipéptido E2 do vírus da hepatite C (VHC), em que o referido polipéptido não possui uma porção da sua região C-terminal que começa no resíduo de aminoácido 715, numerado com referência à sequência de aminoácidos de E2 do VHC1 apresentada na Fig. 4; e

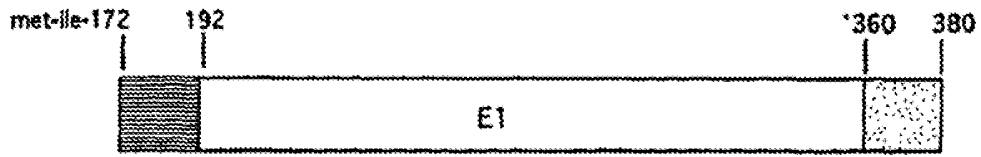
ii) elementos de controlo que se encontram funcionalmente ligados à referida sequência codificante por meio dos quais a referida sequência codificante pode ser transcrita e traduzida numa célula hospedeira; e

(b) cultivar a referida população de células em condições nas quais o referido polipéptido do VHC é expresso e secretado.

2. O método, de acordo com a reivindicação N.º.1, caracterizado por o polinucleótido que compreende um sequência codificante para o polipéptido E2 do vírus da hepatite C (VHC), codificar um fragmento de E2 do VHC a partir do resíduo de aminoácido 383 até ao resíduo de aminoácido 715, numerado com referência à sequência de aminoácidos de E2 do VHC apresentada na Fig. 4.

3. O polipéptido do VHC, caracterizado por ser produzido de acordo com o método das reivindicações N.º.1 ou N.º.2.

Lisboa, 11 de Fevereiro de 2009



= presumível sequência de sinal para o E1, derivada a partir da zona C-terminal da região que codifica o core



= região da poliproteína do VHC que contribui para a retenção de E1 no interior das células

FIG. 1

170 MetIleCysSerPheSerIlePheLeuLeuAlaLeuLeuSerCysLeuThrValProAla
 ATGATTGCTCTTTCTCTATCTTCCTTCIGGCCCTGCTCTCTGCTTGACTGTGCCCGCT
 TACTAAACGAGAAAGAGATAGAAGGAAGACCCGGGACGAGAGAACGAACACTGACACGGGCGA

EI Madura

190 SerAlaTyrGlnValArgAsnSerThrGlyLeuTyrHisValThrAsnAspCysProAsn
 TCGGCTACCAAGTGGCAACTCCACGGGGCTTACCACGTACCAATGATTGCCCTAAC
 AGCCGGATGGTTCACGCGTTGAGGTGCCCGAGATGGTGAGTGGTTACTAACGGGATG

210 SerSerIleValTyrGluAlaAlaAspAlaIleLeuHisThrProGlyCysValProCys
 TCGAGTATTGTGTACGAGGCGGCCGATGCCATCCTGCACACTCCGGGGTGGCTCCCTGCG
 AGCTCATAACACATGCTCCGCCGGCTACGCTAGGACGTGTGAGGCCCCACGCAGGGAACG

230 ValArgGluGlyAsnAlaSerArgCysTrpValAlaMetThrProThrValAlaThrArg
 GTTCGTGAGGGCAACGCCCTCGAGGTGTGGGTGGCGATGACCCCTACGGTGGCCACCAGG
 CAAGCACTCCCGTTGCGGAGCTCCACAACCCACCGCTACTGGGGATGCCACCGTGGTCC

250 AspGlyLysLeuProAlaThrGlnLeuArgArgHisIleAspLeuLeuValGlySerAla
 GATGGCAAACCTCCCGCGACGCAGCTTCGACGTCACATCGATCTGCTTGTGCGGGAGCGCC
 CTACCGTTTGAGGGGCGCTGCGTCCGAGCTGCAGTGTAGCTAGACGAACAGCCCTCGCGG

270 ThrLeuCysSerAlaLeuTyrValGlyAspLeuCysGlySerValPheLeuValGlyGln
 ACCCTCTGTTCCGGCCCTCTACGTGGGGGACCTCTGCGGGTCTGTCTTTCTTGTGCGGCCAA
 TGGGAGACAAGCCGGGAGATGCACCCCTGGAGACGCCCAGACAGAAAGAACAGCCGGTT

290 LeuPheThrPheSerProArgArgHisTrpThrThrGlnGlyCysAsnCysSerIleTyr
 CTGTTTACCTTCTCTCCCAGGCGCCACTGGACGACGCAAGGTTGCAATTGCTCTATCTAT
 GACAAATGGAAGAGAGGGTCCGCGGTGACCTGCTGCGTTCCAACGTTAACGAGATAGATA

310 ProGlyHisIleThrGlyHisArgMetAlaTrpAspMetMetMetAsnTrpSerProThr
 CCCGGCCATATAACGGGTACCCGATGGCATGGGATATGATGATGAACCTGGTCCCTTACG
 GGGCCGGTATATTGCCAGTGGCGTACCGTACCTATACTACTACTTACCAGGGGATGC

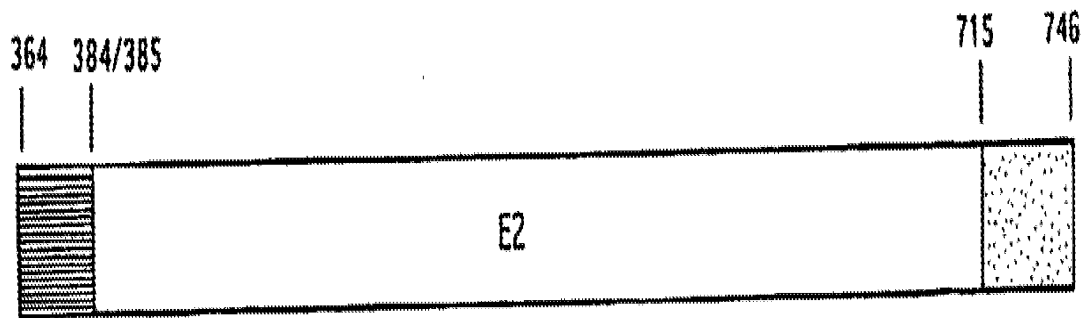
330 ThrAlaLeuValMetAlaGlnLeuLeuArgIleProGlnAlaIleLeuAspMetIleAla
 ACGCGTTPGGTAATGGCTCAGCTGCTCCGGATCCCACAAGCCATCTTGGACATGATCGCT
 TGCCGCAACCATTACCGAGTCGACGAGGCCTAGGGTGTTCGGTAGAACCTGTACTAGCGA

Ancoragem C-terminal

350 GlyAlaHisTrpGlyValLeuAlaGlyIleAlaTyrPheSerMetValGlyAsnTrpAla
 CGTCTCACTGGGGAGTCCCTGGCGGGCATAGCGTATTTCTCCATGGTGGGGAACCTGGGCG
 CCACGAGTGACCCCTCAGGACCGCCCGTATCGCATAAAGAGGTACCACCCCTTGACCCGC

370 LysValLeuValValLeuLeuLeuPheAlaGlyOP
 AAGGTCTGGTAGTGCTGCTGCTATTTGCCGGCTGA
 TTCCAGGACCATCACGACGACGATAAACGGCCGACT

FIG.2



= presumível seqüência de sinal para o E2, derivada a partir da zona C-terminal da região que codifica o E1



= região da poliproteína do VHC que contribui para a retenção de E2 no interior das células

FIG.3

364 MetValGlyAsnTrpAlaLysValLeuValValLeuLeuLeuPheAlaGlyValAspAla
 ATGGTGGGAACTGGGCGAAGGTCTGGTAGTGTCTGCTATTGCGCGGCTCGACGG
 TACCACCCCTGACCCGCTTCCAGGACCATCAGACACGATAAACGGCCGAGCTGCGC

E2 Madura

384 GluThrHisValThrGlyGlySerAlaGlyHisThrValSerGlyPheValSerLeuLeu
 GAAACCCACGTACCGGGGAAGTGCCGGCCACACTGTGTCTGGATTGTTAGCCTCCTC
 CTTTGGGTGACAGTGGCCCCCTTACGGCCGGTGTGACACAGACCTAAACAATCGGAGGAG

404 AlaProGlyAlaLysGlnAsnValGlnLeuIleAsnThrAsnGlySerTrpHisLeuAsn
 GCACCAGCGCCAAGCAGAACGTCCAGCTGATCAACACCAACGGCAGTTGGCACCTCAAT
 CGTGGTCCGCGGTTCTGCTTGCAGGTGACTAGTTGTGGTTGCCGTCAACCGTGGAGTTA

424 SerThrAlaLeuAsnCysAsnAspSerLeuAsnThrGlyTrpLeuAlaGlyLeuPheTyr
 AGCACGGCCCTGAAGTCAATGATAGCCTCAACACCGGCTGGTTGGCAGGGCTTTTCTAT
 TCGTGCCGGGACTTGACGTTACTATCGGAGTGTGGCCGACCAACCGTCCCGAAAAGATA

444 HisHisLysPheAsnSerSerGlyCysProGluArgLeuAlaSerCysArgProLeuThr
 CACCACAAGTTCAACTCTTCAGGCTGTCTGAGAGGCTAGCCAGCTGCCGACCCCTTACC
 GTGGTGTCAAGTTGAGAAGTCCGACAGGACTCTCCGATCGGTGACGGCTGGGGAATGG

464 AspPheAspGlnGlyTrpGlyProIleSerTyrAlaAsnGlySerGlyProAspGlnArg
 GATTTTGACAGGGCTGGGGCCCTATCAGTTATGCCAACGGGAAGCGGCCCGACAGCGC
 CTAACACTGGTCCCGACCCCGGATAGTCAATACGGTTGCCCTTCCCGGGGCTGGTCCGG

484 ProTyrCysTrpHisTyrProProLysProCysGlyIleValProAlaLysSerValCys
 CCCACTGCTGGCACTACCCCCAAAACCTTGCGGTATTGTGCCCGGAAGAGTGTGTGT
 GGGATGACGACCGTGATGGGGGTTTTGGAACGCCATAACACGGGCGCTTCACACACA

504 GlyProValTyrCysPheThrProSerProValValValGlyThrThrAspArgSerGly
 GGTCGGTATATTTGCTTCACTCCCAGCCCGTGGTGGTGGGAACGACCCAGGTCGGGC
 CCAGGCCATATAACGAAGTGAGGGTCCGGGCACCACCACCCCTTGTGGCTGCCAGCCCG

524 AlaProThrTyrSerTrpGlyGluAsnAspThrAspValPheValLeuAsnAsnThrArg
 GCGCCACCTACAGCTGGGGTGAATAATGATACGGAGCTCTTCCCTTAAACAATACCAGG
 CCGGGTGGATGTCGACCCCACTTTTACTATGCTGAGAGCAGGAATTGTTATGGTCC

544 ProProLeuGlyAsnTrpPheGlyCysThrTrpMetAsnSerThrGlyPheThrLysVal
 CCACCGCTGGCAATTTGTTCTGGTGTACCTGGATGAACCTCACTGGATTACCAAAGTG
 GGTGGCGACCCGTTAACAAGCCAACATGGACCTACTTGTAGTTGACCTAAGTGGTTTAC

564 CysGlyAlaProProCysValIleGlyGlyAlaGlyAsnAsnThrLeuHisCysProThr
 TCGGAGCGCTCCTTGTGTCATCGGAGGGGCGGGCAACAACACCCCTGCACTGCCCCACT
 ACGCCTCGCGGAGGAACACAGTAGCTCCCGCCCGTGTGTGGGACGTGACGGGGTGA

584 AspCysPheArgLysHisProAspAlaThrTyrSerArgCysGlySerGlyProTrpIle
 GATGCTTCGCAAGCATCCGGACGCCACATACTCTCGGTGCGGCTCCGGTCCCTGGATC
 CTAACGAAGCGTTCGTAGGCCGCGGTGTATGAGAGCCACGCCGAGGCCAGGGACCTAG

604 ThrProArgCysLeuValAspTyrProTyrArgLeuTrpHisTyrProCysThrIleAsn
 ACACCCAGGTGCCTGGTCGACTACCCGTATAGGCTTTGGCATTATCCTGTACCATCAAC
 TGTGGGTCCACGGACAGCTGATGGGCATATCCGAAACCGTAATAGGAACATGGTAGTTG

624 TyrThrIlePheLysIleArgMetTyrValGlyGlyValGluHisArgLeuGluAlaAla
 TACACCATATTTAAATCAGGATGTACCTGGGAGGGGTCGAACACAGGCTGGAAGCTGCC
 ATGTGGTATAAATTTAGTCTACATGCACCCCTCCCAGCTGTGTGCCACCTTCGACGG

644 CysAsnTrpThrArgGlyGluArgCysAspLeuGluAspArgAspArgSerGluLeuSer
 TGCAACTGGACCGGGGCGAACGTTGCGATCTGGAAGATAGGGACAGGTCGAGCTCAGC
 ACGTTGACCTGCCCGCCGCTTGCACCGCTAGACCTTCTATCCCTGTCCAGGCTCGAGTGC

FIG.4A

664 ProLeuLeuLeuThrThrThrGlnTrpGlnValLeuProCysSerPheThrThrLeuPro
 CCGTTACTGCTGACCACTACACAGTGGCAGGTCCCTCCCGTGTTCCTCACAACCTGCCA
 GGCAATGACGACTGGTGATGTGTACCCGTCCAGGAGGGCACAAGGAAGTGTFTGGGACGGT

684 AlaLeuSerThrGlyLeuIleHisLeuHisGlnAsnIleValAspValGlnTyrLeuTyr
 GCCTTGTCCACCGCCTCATCCACCTCCACCAGAACATTGTGGACGTGACGTACTTGTAC
 CGGAACAGGTGGCCGGAGTAGGTGGAGGTGGTCTTGTAAACACCTGCACGTGATGAACATG

704 GlyValGlySerSerIleAlaSerTrpAlaIleLysTrpGluTyrValValLeuLeuPhe
 GGGGTGGGGTCAAGCATCGCGTCTCTGGGCCATTAAGTGGGAGTACGTGCTCCTCCTGTTC
 CCCCACCCAGTTCGTAGCGCAGGACCCGGTAATTCACCCTCATGCAGCAGGAGACAAG

ancoragem C-terminal

724 LeuLeuLeuAlaAspAlaArgValCysSerCysLeuTrpMetMetLeuLeuIleSerGln
 CTTCTGCTTGCAGACGCGCGTCTGCTCCTGCTTGTGGATGATGCTACTCATATCCCAA
 GAAGACGAACGCTCTGCCGCCGACAGCAGGACGAACACCTACTACGATGAGTATAGGGTT

NS2A

744 AlaGluAlaAlaLeuGluAsnLeuValIleLeuAsnAlaAlaSerLeuAlaGlyThrHis
 GCGGAAGCGCTTGGGAGAACCTCGTAATACTTAATGCAGCATCCCTGGCCGGGACGCAC
 CGCCTTCGCCGAAACCTCTTGGAGCATTATGAATTACGTGCTAGGGACCCGCCCTGCGTG

764 GlyLeuValSerPheLeuValPhePheCysPheAlaTrpTyrLeuLysGlyLysTrpVal
 GGTCTTSTATCCTCCTCGTGTCTTCTGCTTTGCAFGTATCTGAAGGGTAAGTGGGTG
 CCAGAACATAGGAAGGAGCACAAGAAGACGAAACGTACCATAGACTTCCCATTCACCCAC

784 ProGlyAlaValTyrThrPheTyrGlyMetTrpProLeuLeuLeuLeuLeuAlaLeu
 CCGGAGCGGTCTACACCTTCTACGGGATGTGGCCTCCTCCTGCTCCTGTTGGCGTTG
 GGGCCTCGCCAGATGTGGAAGATGCCCTACACCGGAGAGGAGGACGAGGACAACCGCAAC

NS2B

804 ProGlnArgAlaTyrAlaLeuAspThrGluValAlaAlaSerCysGlyGlyValValLeu
 CCCCAGCGGGCTACGCGCTGGACACGGAGGTGGCCGCGTGTGGCGGTGTTGTTCTC
 GGGTGCGCCGATGCCGACCTGTGCCTCCACCGGCGCAGCACACCGCCACAACAAGAG

824 ValGlyLeuMetAlaLeuThrLeuSerProTyrTyrLysArgTyrIleSerTrpCysLeu
 GTCGGGTTGATGGCCGCTAACTCTGTCAACATATTACAAGCGCTATATCAGCTGGTGCTG
 CAGCCCAACTACCGGATTGAGACAGTGGTATAATGTTCCGGATATAGTCGACCACGAAC

844 TrpTrpLeuGlnTyrPheLeuThrArgValGluAlaGlnLeuHisValTrpIleProPro
 TGGTGGCTTCAGTATTTTCTGACCAGAGTGGAAAGCGCAACTGCACGTGTGGATTCCCCC
 ACCACCGAAGTCATAAAAGACTGGTCTCACCTTCGCGTTGACGTGCACACCTAAGGGGGG

864 LeuAsnValArgGlyGlyArgAspAlaValIleLeuLeuMetCysAlaValHisProThr
 CTCACGTCGGAGGGGGCGCGACGCCGTCATCTACTCATGTGTGCTGTACACCCGACT
 GAGTTCAGGCTCCCCCGCGCTGCGGCAGTAGAATGAGTACACACGACATGTGGGCTGA

884 LeuValPheAspIleThrLysLeuLeuLeuAlaValPheGlyProLeuTrpIleLeuGln
 CTGGTATTTGACATCACCAAATTGCTGCTGGCCGCTCTTCGGACCCCTTTGGATTCTTCAA
 GACCATAAACTGTAGTGGTTTAAACGACACCGGCAGAAGCCTGGGGAAACCTAAGAAGTT

904 AlaSerLeuLeuLysValProTyrPheValArgValGlnGlyLeuLeuArgPheCysAla
 GCCAGTTCCTTAAAGTACCCTACTTTGTGCGCGTCCAAGGCCTTCTCCGGTCTGCGCG
 CGSTCAAACGAATTTTCATGGGATGAACACCGCCAGGTTCCGGAAGAGGCCAAGACCGCG

924 LeuAlaArgLysMetIleGlyGlyHisTyrValGlnMetValIleIleLysLeuGlyAla
 TTAGCGCGGAAGATGATCGGAGGCCATTACGTGCAAATGGTCATCATTAAGTTAGGGCG
 AATCGCGCCTTCTACTAGCCTCCGGTAATGCACGTTTACCAGTAGTAATTCAATCCCCCG

944 LeuThrGlyThrTyrValTyrAsnHisLeuThrProLeuArgAspTrpAlaHisAsnGly
 CTTACTGGCACCTATGTTTATAACCATCTCACTCCTCTTCGGGACTGGGCGCACACCGCG
 GAATGACCGTGGATACAAATATTGGTAGAGTGAGGAGAAGCCCTGACCCGCGTGTGGCG

FIG.4B

964 LeuArgAspLeuAlaValAlaValGluProValValPheSerGlnMetGluThrLysLeu
TTGCGAGATCTGGCCGTGGCTGTAGAGCCAGTCGTCTTCTCCCAAATGGAGACCAAGCTC
AACGCTCTAGACCGGCACCGACATCTCGGTCAGCAGAAGAGGGTTTACCTCTGGTTTCGAG

984 IleThrTrpGlyAlaAspThrAlaAlaCysGlyAspIleIleAsnGlyLeuProValSer
ATCACGTGGGGGGCAGATACCGCCCGTGGGGTGACATCATCAACGGCTTGCCCTGTTCC
TAGTGCACCCCCGTCTATGSCGGCCACGCCACTGTAGTAGTTGCCGAACGGACAAAGG

1004 AlaArgArgGlyArgGluIleLeuLeuGlyProAlaAspGlyMetValSerLysGlyTrp
GCCCCAGGGGGCCGGGAGATACTGCTCGGGCCAGCCGATGGAATGGTCTCCAAGGGTTGG
CGGGCGTCCCCGGCCCTCTATGACGAGCCCCGGTGGCTACCTTACCAGAGGTTCCCAACC

1024 ArgLeuLeu
AGGTTGCTG
TCCAACGAC

FIG.4C

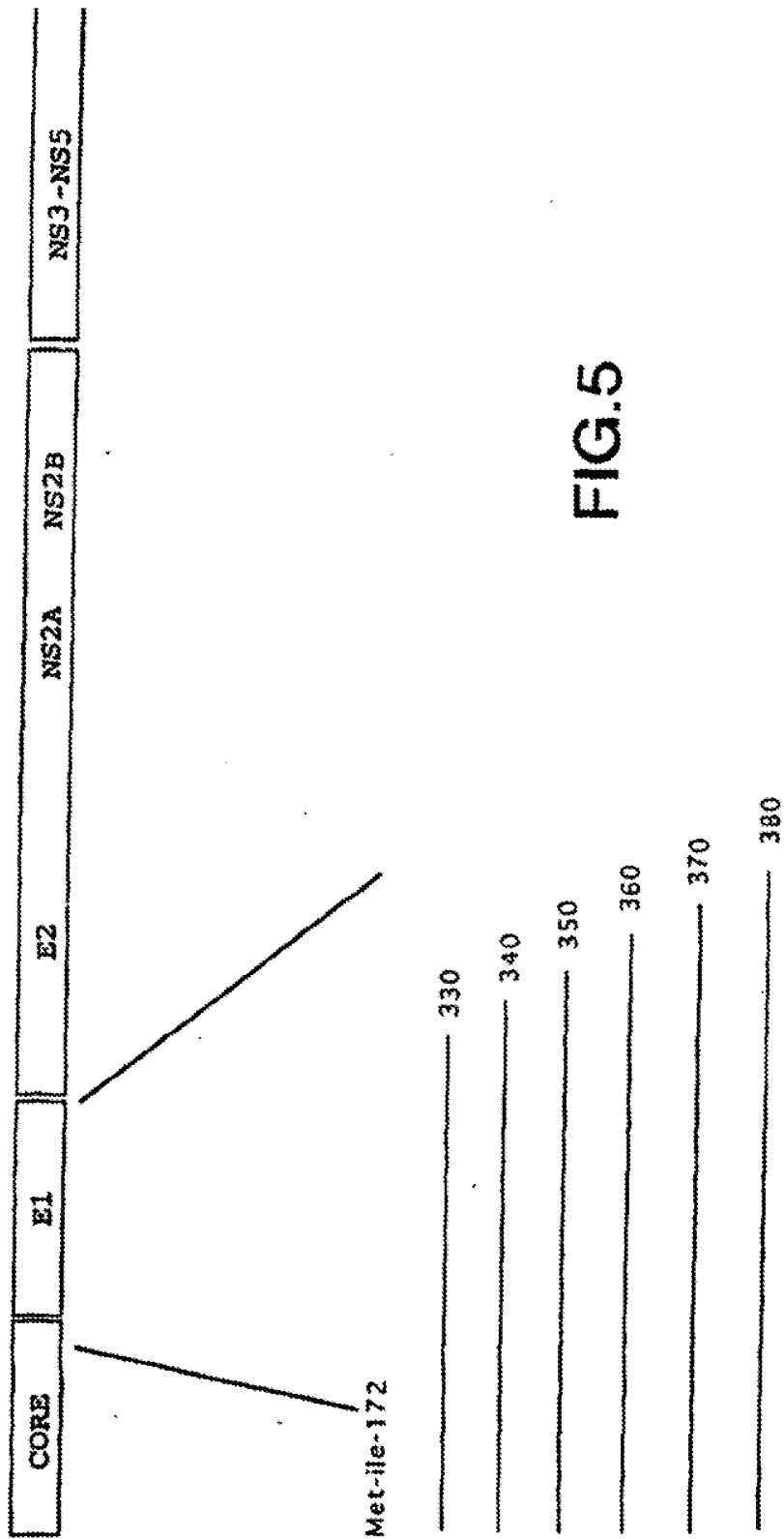


FIG.5

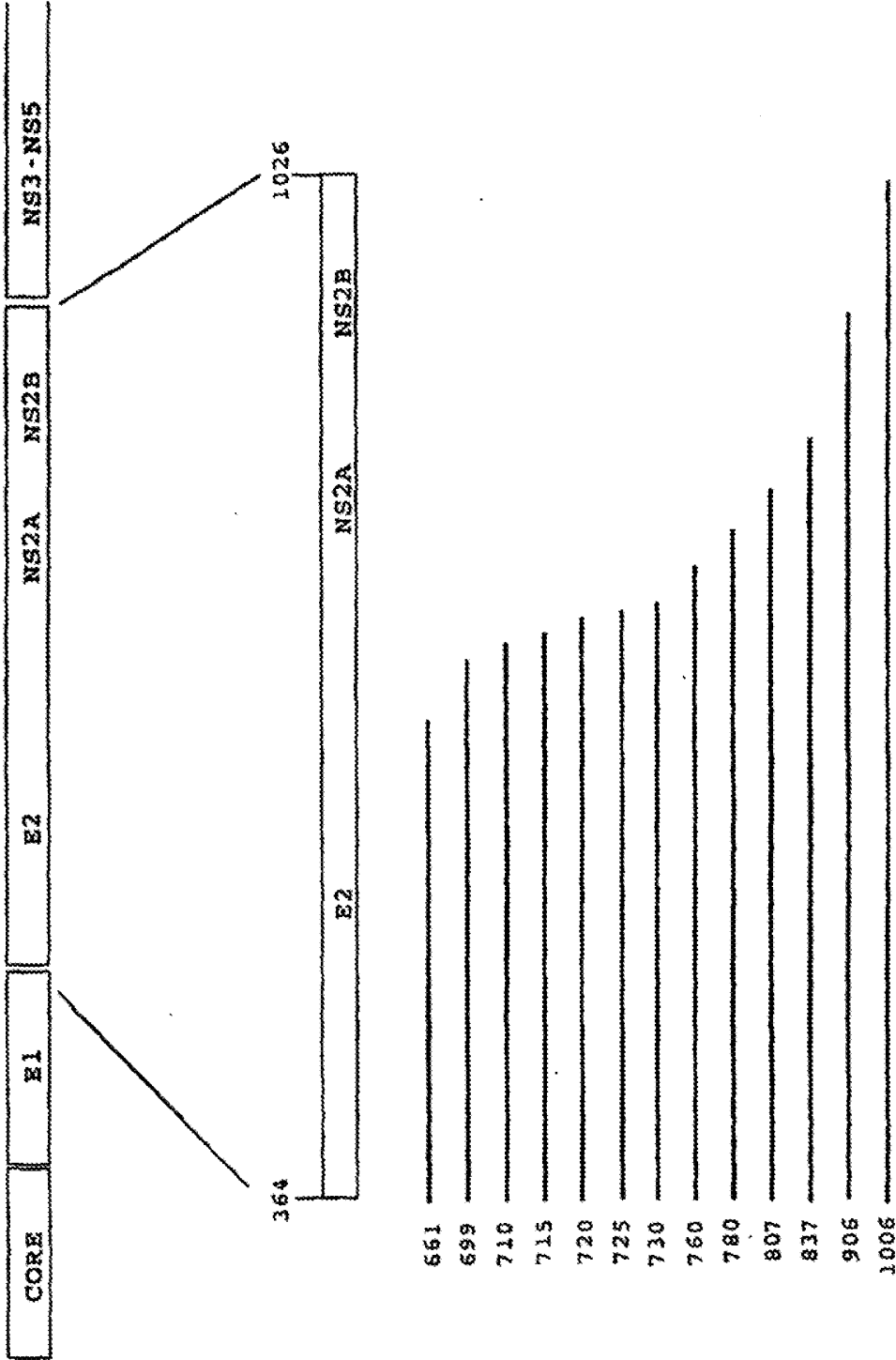


FIG.6

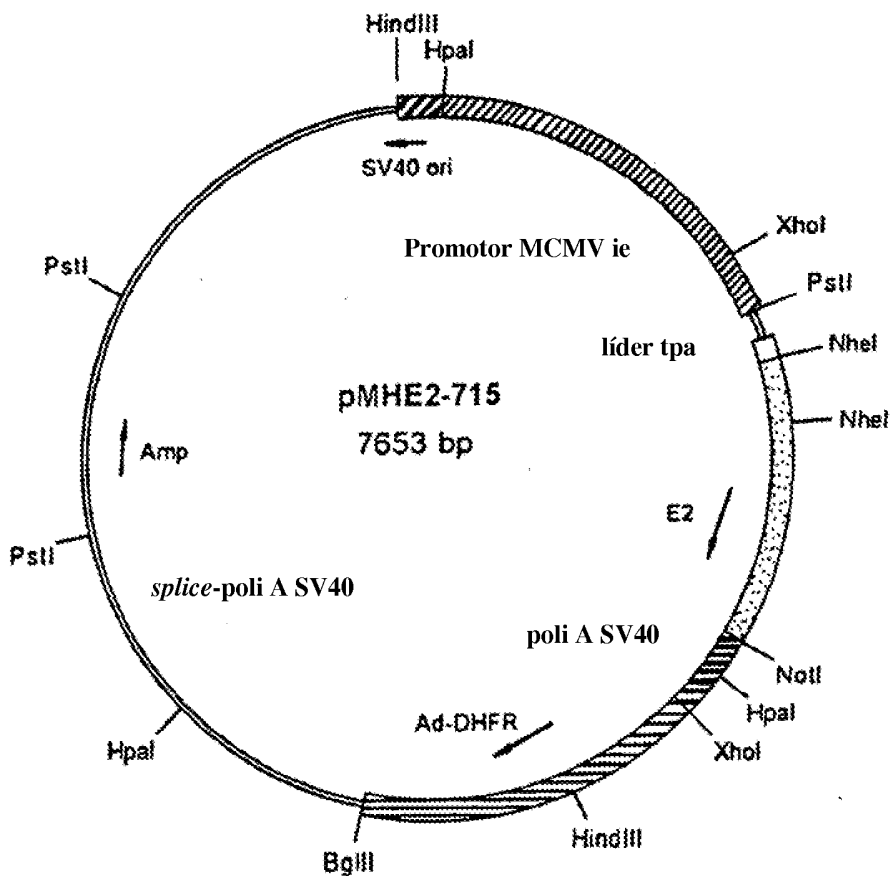


FIG.7