

P. 8. no 81.213



KIRIN-AMGEN, INC.

"PROCESSO PARA A PRODUÇÃO DE ERITROPOIETINA"

Fundamento da Invenção

A presente invenção diz respeito, de um modo geral, à manipulação de materiais genéticos e, mais especificadamente, a técnicas recombinantes que tornam possível a produção de polipeptidos que possuem parte ou toda a conformação estrutural primária e/ou uma ou mais das propriedades biológicas da eritropoietina existente na natureza.

A. Manipulação de materiais genéticos

Os materiais genéticos podem ser definidos, de uma maneira geral, como as substâncias químicas que programam e orientam a produção de constituintes de células e vírus e dirigem as respostas de células e vírus. Uma substância polimérica de cadeia comprida, conhecida como ácido deoxiribonucleico (ADN) compreende o material genético de todas as células vivas e de vírus com exceção de determinados vírus que são programados por ácido ribonucleico (ARN). As unidades repetidas em polímeros de ADN são quatro nucleotídeos diferentes, cada um dos quais é constituído ou por uma purina (adenina ou guanina) ou uma pirimidina (timina ou citosina) ligada a um açúcar deoxiribose a que está unido um grupo fosfato. A união de nucleotídeos em forma polimérica linear é feita mediante fusão do fosfato na posição 5' de um nucleotídeo com o grupo hidróxi na posição 3' de um outro. O ADN funcional ocorre na forma de associações estáveis de cadeias duplas com cadeias simples de nucleotídeos (conhecidos como deoxioligonucleotídeos),



associações que se produzem através da ligação de hidrogénio entre as bases purina e pirimidina [isto é, associações "complementares" existentes ou entre a adenina (A) e a timina (T) ou entre a guanina (G) e a citosina (C)]. Por convenção, os nucleotidos são designados pelos nomes das suas bases constituintes purina ou pirimidina e as associações complementares de nucleotidos no ADN de cadeia dupla (isto é, A-T e G-C) são designados por "pares de bases". O ácido ribonucleico é um polinucleotido que contém adenina, guanina, citosina e uracilo (U) em vez de timina, ligados à ribose e a um grupo fosfato.

De uma maneira resumida, a função de programação de ADN é geralmente efectuada através de um processo em que são "transcritas" sequências nucleotídicas de ADN específicas (genes) em polímeros de ARN mensageiro (mARN) relativamente instáveis. O mARN, por sua vez, serve como um template para a formação de proteínas estruturais, reguladoras e catalíticas a partir de aminoácidos. Este processo de "transferência" de mARN compreende as operações de pequenas cadeias de ARN (tARN) que transportam e alinham aminoácidos individuais ao longo da cadeia de mARN para permitir a formação de polipeptidos em sequências de aminoácidos apropriadas. A "mensagem" do mARN, derivada de ADN e proporcionando a base para o fornecimento de tARN e a orientação de qualquer um dos vinte aminoácidos para "expressão" do polipeptido, está sob a forma de "códeços" de triplete-grupos sucessivos de três bases nucleotídicas. Em certo sentido, a formação de uma proteína é a forma derradeira de "expressão" da mensagem genética programada fornecida pela sequência nucleotídica de um gene.

As sequências de ADN "promotoras" "precedem" habitual



mente um gene em um polímero de ADN e fornecem um sítio para início da transcrição no mARN. As sequências de ADN "reguladoras", também habitualmente "a montante" de (isto é, precedendo) um gene em um determinado polímero de ADN, ligam proteínas que determinam a frequência (ou velocidade) de início de transcrição. Colectivamente designadas por "promotor/regulador" ou sequência de ADN "de controlo", estas sequências que precedem um gene escolhido (ou série de genes) em um polímero de ADN funcional cooperam para determinar se deverá ocorrer a transcrição (e expressão eventual) de um gene. As sequências de ADN que "seguem" um gene em um polímero de ADN e fornecem um sinal para termo da transcrição para mARN são designadas por sequências "terminadoras" de transcrição.

Um centro de interesse do tratamento microbiológico nesta última década tem sido a tentativa de preparar, industrialmente e com fins farmacêuticos, substâncias importantes que utilizam organismos que ou não têm inicialmente informação codificada geneticamente relacionada com o produto desejado incluído no seu ADN, ou (no caso de células de mamíferos em cultura) não exprimem ordinariamente um gene cromosómico a níveis apreciáveis. Mais simplesmente, um gene que especifique a estrutura de um produto polipeptídico desejado ou é isolado a partir de um organismo "dador" ou é sintetizado quimicamente e introduzido depois de maneira estável em um outro organismo que seja, de preferência, um organismo unicelular de auto-replicação como bactérias, levedura ou células de mamíferos em cultura. Uma vez feito isto, a maquinaria existente para a expressão do gene nas células de hospedeiros microbianos "transformadas" ou "modificadas" funciona de modo a construir o pro



duto desejado, utilizando o ADN exógeno como um templato (molécula) para a transcrição de mRNA que é depois transferido para uma sequência contínua de restos de aminoácidos.

Existem muitas patentes e publicações relativas a metodologias de "ADN recombinante" para o isolamento, a síntese, a purificação e a amplificação de materiais genéticos para utilização na transformação de organismos-hospedeiros escolhidos. A patente de invenção norte-americana Nº 4.237.224, concedida a Cohen e colab., por exemplo, descreve a transformação de organismos-hospedeiros unicelulares com ADN de plasmídeo viral ou circular "híbrido" que compreende sequências de ADN exógeno. As técnicas da patente de Cohen e colab. envolvem primeiro a preparação de um vector de transformação mediante cisão enzimática de ADN de plasmídeo viral ou circular para formar cadeias de ADN lineares. As cadeias de ADN estranho ("exógeno" ou "heterólogo") escolhido, incluindo habitualmente sequências que codificam o produto desejado, são preparadas em forma linear através da utilização de enzimas similares. O ADN linear viral ou de plasmídeo é incubado com o ADN estranho na presença de enzimas de ligação capazes de efectuarem um processo de restauração e formam-se vectores "híbridos" que contêm o segmento de ADN exógeno escolhido "entrançado" no plasmídeo de ADN viral ou circular.

A transformação de organismos-hospedeiros unicelulares compatíveis com o vector híbrido resulta na formação de cópias múltiplas do ADN exógeno na população de células hospedeiras. Em certos casos, o resultado desejado é simplesmente a amplificação do ADN estranho e o "produto" colhido é o ADN. Mais frequentemente, o objectivo da transformação é a expres-



são, pelas células-hospedeiras, do ADN exógeno sob a forma de síntese, em grande escala, de quantidades isoláveis comercialmente significativas de proteína ou de fragmentos de polipeptido codificados pelo ADN estranho. Ver também, por exemplo, patentes de invenção norte-americanas Nºs 4.264.731 (concedida a Shine), 4.273.875 (concedida a Manis), 4.293.652 (concedida a Cohen) e o pedido de patente de invenção europeia Nº 093.619, publicado em 9 de Novembro de 1983.

O desenvolvimento de sequências de ADN específicas para entrar nos vectores de ADN é efectuado por uma grande variedade de técnicas, dependendo muito do grau de "estranheza" ("foreignness") do "dador" em relação ao hospedeiro projectado e do tamanho do polipeptido a ser expresso no hospedeiro. Correndo o risco de uma simplificação excessiva, pode estabelecer-se que podem ser utilizados três métodos principais de alternativa: (1) o "isolamento" da sequência de ADN de cadeia dupla a partir do ADN genómico do dador; (2) a preparação química de uma sequência de ADN que proporciona um código para um polipeptido de interesse; e (3) a síntese "in vitro" de uma sequência de ADN de cadeia dupla mediante "transcrição inversa" enzimática de mARN isolado a partir de células do dador. Os métodos mencionados antes que compreendem a formação de um "complemento" de ADN de mARN são geralmente designados por métodos de "cADN".

A preparação de sequências de ADN é frequentemente o método de escolha quando se conhece a sequência completa de restos de aminoácidos do produto polipeptídico desejado. As técnicas de preparação de ADN do pedido de patente de invenção norte-americana Nº de Série 483.451, de Alton e colab. (apre-



sentada em 15 de Abril de 1983 e correspondente a PCT US83/
/00605, publicada em 24 de Novembro de 1983 como W083/04053),
por exemplo, proporcionam um meio excelente para conseguir
esses resultados extremamente desejáveis como: proporcionar a
presença de códiços de alternativa geralmente encontrados em
genes que são consideravelmente expressos no organismo-hospe-
deiro escolhido para expressão (por ex. fornecendo códiços "de
preferência" de levedura ou de "E. coli"); evitar a presença
de sequências "intron" não transferidas (habitualmente presen-
tes em sequências de ADN genómico de mamíferos e suas transcri-
ções de mRNA) que não são processadas facilmente por células-
hospedeiras procarióticas; evitar a expressão de sequências
polipeptídicas "condutoras" indesejáveis geralmente codifica-
das por sequências genómicas de ADN e cADN mas frequentemente
não cindidas facilmente do polipeptido de interesse por célu-
las-hospedeiras de bactérias ou de levedura; proporcionar uma
pronta inserção do ADN em vectores de expressão convenientes
em associação com as desejadas sequências de promotor/regula-
dor e de terminador; e proporcionar a construção rápida de ge-
nes que codificam fragmentos e análogos de polipeptidos dos
polipeptidos desejados.

Quando não é conhecida a sequência completa dos res-
tos de aminoácidos do polipeptido desejado, não é possível a
preparação directa de sequências de ADN e o isolamento das se-
quências de ADN que codificam o polipeptido por um método de
cADN torna-se o método de escolha apesar dos inconvenientes
potenciais na facilidade de montagem de vectores de expressão
capazes de proporcionar níveis elevados de expressão microbia-
na referidos antes. Entre as técnicas normais para o isolamento



de sequências de cADN com interesse está a preparação de "coleções" de cADN derivados de plasmídio, resultantes da transcrição inversa do mARN abundante em células dadoras escolhidas como responsáveis pela expressão de alto nível de genes (por exemplo, coleções de cADN derivado de células da pituitária que exprimem quantidades relativamente grandes de produtos da hormona de crescimento). Quando são conhecidas porções substanciais da sequência de aminoácidos do polipeptido, pode utilizar-se sequências-sonda de ADN de cadeia simples, marcadas com um isótopo radioactivo, que duplicam uma sequência putativamente presente no cADN "objectivo" ("alvo"), em técnicas de hibridação de ADN/ADN efectuadas sobre cópias clonadas do cADN que foram desnaturadas até à forma de cadeia simples. [Ver, de um modo geral, a descrição e as discussões concernentes fornecidas pela patente de invenção norte-americana Nº. 4.394.443, concedida a Weissman e colab., e as demonstrações recentes da utilização de sondas da hibridação de oligonucleotidos compridos referidas por Wallace e colab., "Nuc. Acids Res.", 6, 3543-3557 (1979), e Reyes e colab., "P.N.A.S. (U.S.A.)", 79, 3270-3274 (1982), e Jaye e colab., "Nuc. Acids Res.", 11, 2325-2335 (1983). Ver também, patente de invenção norte-americana Nº. 4.358.535, concedida a Falkow e colab., relativa a técnicas de hibridação de ADN/ADN para efectuar diagnósticos; os pedidos de patentes europeias publicados Nºs 0070685 e 0070687 referentes a marcas que emitem luz nas sondas de polinucleotidos de cadeia simples; Davis e colab., "A Manual for Genetic Engineering, Advanced Bacterial Genetics", Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y. (1980) nas pág. 55-58 e 174-176, relativas a técnicas de hibridação em colónia e em placa; e brochu



ras de New England Nuclear (Boston, Mass.) para "Gene Screen" "Hybridization Transfer Membrane materials Providing instruction manuals for the transfer and hybridization of DNA e RNA", Catálogo Nº. NEF-9727.

Entre os progressos recentes mais significativos nas técnicas de hibridação para a selecção de clónios recombinantes está a utilização de sondas de oligonucleotidos sintéticos mistos marcados, cada um das quais é potencialmente o complemento completo de uma sequência de ADN específica na amostra de hibridação que inclui uma mistura heterogénea de ADNs ou ARNs de cadeia simples. Sabe-se que estas técnicas são especialmente úteis na detecção de colónios de cADN provenientes de fontes que fornecem quantidades extremamente pequenas de sequências de mARN do polipeptido com interesse. De uma maneira abreviada, a utilização de condições de hibridação rigorosas, orientadas para evitar a ligação não específica, pode permitir, por ex., a visualização auto-radiográfica de um clónio de cADN específico no caso da hibridação do ADN-objectivo para aquela sonda simples dentro da mistura que é o seu complemento completo. Ver, de um modo geral, Wallace e colab., "Nuc. Acids Res.", 9, 879-897 (1981); Suggs e colab., "P.N.A.S. (U.S.A.)", 78, 6613-6617 (1981); Choo e colab., "Nature", 299, 178-180 (1982); Kurachi e colab., "P.N.A.S. (U.S.A.)", 79, 6461-6464 (1982); OhKubo e colab., "P.N.A.S. (U.S.A.)", 80, 2196-2200 (1983); e Kornblihtt e colab., "P.N.A.S. (U.S.A.)", 80, 3218-3222 (1983). Em geral, as técnicas de sonda mista de Wallace e colab. (1981), Supra, foram desenvolvidas por diversos investigadores até ao ponto de se terem obtido resultados seguros em um isolamento de clónio de cADN que utilizou uma "série" mista de 32 membros de sondas de oligonucleotidos com um compri-



mento de 16 bases (16-mer) de sequências de ADN que variam uniformemente juntamente com um 11-mer simples para efectuar uma confirmação "positiva" de dois sítios da presença de cADN com interesse. Ver, Singer-Sam e colab., "P.N.A.S. (U.S.A.)", 80, 802-806 (1983).

A utilização de isolados de ADN genómico é o menos comum dos três métodos anteriormente citados para o desenvolvimento de sequências de ADN específicas para utilização em técnicas recombinantes. Isto é especialmente verdadeiro na área de técnicas recombinantes orientadas para assegurar a expressão microbiana de polipeptidos de mamíferos e é devido, principalmente, à complexidade de ADN genómico de mamíferos. Assim, embora existam técnicas seguras para desenvolver colecções, derivadas de fases de ADN genómico de origem humana ou de outras espécies de mamíferos [Ver, por ex., Lawn e colab., "Cell", 15, 1157-1174 (1978), relacionadas com técnicas para a criação de uma colecção genómica humana vulgarmente designada por "Colecção Maniatis"; Karn e colab., "P.N.A.S. (U.S.A.)", 77, 5172-5176 (1980) relativa a uma colecção genómica humana com base na técnica de alternativa de fragmentação da endonuclease de restrição; e Blattner e colab., "Science", 196, 161-169 (1977) que descrevem a construção de uma colecção genómica bovina], tem havido relativamente poucas tentativas bem sucedidas na utilização de técnicas de hibridação no isolamento de ADN genómico na ausência de uma previsão extensiva da sequência de aminoácidos ou do ADN. Como exemplo, Fiddes e colab., "J.Mol. and App. Genetics", 1, 3-18 (1981) referem o isolamento bem sucedido de um gene que codifica a sub-unidade α das hormonas glicoproteicas da pituitária humana a partir da Colecção Maniatis através da



utilização de uma sonda de "comprimento total" que compreende um fragmento completo de 621 pares de base de uma sequência de cADN previamente isolada da sub-unidade alfa. Como outro exemplo, Das e colab. "P.N.A.S. (U.S.A.)", 80, 1531-1535 (1983) referem o isolamento de clónios genómicos humanos de HLA-DR humano, utilizando um oligonucleotido sintético de 175 pares de bases. Por último, Anderson e colab., "P.N.A.S. (U.S.A.)", 80, 6838-6842 (1983) referem o isolamento de clone genómico para o inibidor da tripsina pancreática bovina (BPTI), utilizando uma sonda simples de 86 pares de bases em comprimento, e construída de acordo com a sequência de aminoácidos conhecida de BPTI. Os autores registam uma determinação de perspectivas fracas para o isolamento de mARN apropriado para a síntese de uma colecção de cADN devidas aos evidentes níveis fracos de mARN na glândula parótida inicialmente marcada e em tecido de pulmão e dirigem as expectativas de sucesso na sondagem de uma colecção genómica, utilizando uma mistura de sondas marcadas, relatando que: "Mais generalizadamente, têm sido utilizadas sondas de oligodeoxinucleotido de sequência mista para isolar genes de proteína de sequência desconhecida a partir de colecções de cADN. Essas sondas são tipicamente misturas de 8-32 oligonucleotidos, 14-17 nucleótidos em comprimento, representando todas as associações de códiços possíveis para um pequeno alongamento (5-6 restos) da sequência de aminoácidos. Sob condições de hibridação rigorosas que fazem a distinção de sondas de bases incorrectamente emparelhadas, estas misturas são capazes de localizar sequências de genes específicas em colecções de clónios de complexidade fraca-a-moderada. No entanto, devido ao seu comprimento reduzido e à sua heterogeneidade, as sondas mistas carecem muitas vezes



da especificidade necessária para a sondagem de sequências tão complexas como um genoma de mamífero. Isto torna este método impraticável para o isolamento de genes de proteína de mamíferos quando não se dispõe dos mARNs correspondentes". (Citações omitidas).

Deste modo, continua a existir uma necessidade, neste campo, de métodos melhorados para efectuar o isolamento rápido e eficiente de clónios de cADN nos casos em que pouco se conhece da sequência de aminoácidos do polipeptido para tal codificado e em que não se pode dispor facilmente de fontes de tecido "enriquecido" de mARN para utilização na construção de colecções de cADN. Tais métodos melhorados deveriam ser especialmente úteis se pudessem ser aplicados ao isolamento de clónios genómicos de mamíferos em que se dispõe de informação escassa relativa a sequências de aminoácidos do polipeptido codificado pelo gene procurado.

B. Eritropoietina como polipeptido com interesse

A eritropoiese, produção de células de sangue vermelhas (eritrocitos), ocorre continuamente durante toda a vida humana para compensar a destruição das células. A eritropoiese é um mecanismo fisiológico controlado com muita exactidão que permite dispor de um número suficiente de células de sangue vermelhas no sangue para a oxigenação apropriada dos tecidos, mas não de demasiadas células que impeçam a circulação. A formação de células vermelhas do sangue ocorre na medula óssea e está sob o controlo da hormona, eritropoietina.

A eritropoietina, uma glucoproteína ácida com um peso molecular de aproximadamente 34.000 daltons, pode ocorrer em



três formas: α , β e asialo. As formas α e β diferem ligeiramente nos componentes de hidratos de carbono mas têm a mesma potência, a mesma actividade biológica e o mesmo P.M. A forma asialo é uma forma α ou β com o hidrato de carbono terminal removido (ácido siálico). A eritropoietina está presente em concentrações muito reduzidas no plasma quando o corpo se mantém em um estado saudável em que os tecidos recebem uma oxigenação suficiente do número existente de eritrocitos. Esta fraca concentração normal é suficiente para estimular a substituição das células vermelhas do sangue que se vão perdendo normalmente por envelhecimento das mesmas.

A quantidade de eritropoietina na circulação é aumentada sob condições de hipoxia quando é reduzido o transporte de oxigénio pelas células sanguíneas na circulação. A hipoxia pode ser causada pela perda de grandes quantidades de sangue através de uma hemorragia, destruição de eritrocitos mediante uma exposição excessiva a radiações, redução na absorção de oxigénio devida a grandes altitudes ou a inconsciência prolongada ou a diversas formas de anemia. Em resposta ao esforço hipóxico sofrido pelos tecidos, a eritropoietina irá aumentar a produção de células sanguíneas vermelhas estimulando a conversão de células precursoras primitivas na medula óssea em pró-eritroblastos que, em seguida, amadurecem, sintetizam hemoglobina e são libertados na circulação como células vermelhas do sangue. Quando o número de células sanguíneas vermelhas em circulação é superior ao necessário para as exigências de oxigénio do tecido normal, diminui a eritropoietina em circulação.

Ver, de um modo geral, Testa e colab., "Exp. Hematol". 8 (Supl. 8), 144-152 (1980); Tong e colab., "J. Biol. Chem.",



256 (24), 12666-12672 (1981); Goldwasser, "J. Cell. Physiol.", 110 (Supl. 1), 133-135 (1982); Finch, "Blood", 60(6), 1241-1246 (1982); Sytowski e colab., "Exptl. Hematol.", 8 (Supl. 8), 52-64 (1980); Naughton, "Ann. Clin. Lab. Sci.", 13 (5), 432-438 (1983); Weiss e colab., "Am. J. Vet. Res.", 44 (10), 1832-1835 (1983); Lappin e colab., "Exp. Hematol.", 11 (7), 661-666 (1983); Baciú e colab., "Ann. N.Y. Acad. Sci.", 414, 66-72 (1983); Murphy e colab., "Acta Haematologica Japonica", 46 (7), 1380-1396 (1983); Dessypris e colab., "Brit. J. Haematol.", 56, 295-306 (1984); e Emmanouel e colab., "Am. J. Physiol.", 247 (1 Pt 2), F168-176 (1984).

Como a eritropoietina é essencial no processo de formação das células sanguíneas vermelhas, a hormona tem aplicação utilizável potencial tanto na diagnose como no tratamento de perturbações do sangue caracterizadas pela produção fraca ou anormal de células vermelhas do sangue. Ver, de uma maneira geral, Pennathur-Das e colab., "Blood", 63 (5), 1168-1171 (1984) e Haddy, "Am. Jour. Ped. Hematol./Oncol.", 4, 191-196 (1982) relativos à eritropoietina em terapias possíveis na doença das células falciformes, e Eschbach e colab., "J. Clin. Invest.", 74 (2), 434-441 (1984), que descrevem um regime terapêutico para carneiros urémicos com base na resposta "in vivo" a infusões de plasma rico em eritropoietina e propõem uma dosagem de 10/U/EPO/Kg por dia, durante 15-40 dias, como corrector de anemia do tipo associado à incapacidade renal crónica. Ver também, Krane, "Henry Ford Hosp. Med. J.", 31 (3), 177-181 (1983).

Calculou-se recentemente que a disponibilidade de eritropoietina em quantidade deveria permitir o tratamento, em cada ano, de anemias de 1.600.000 pessoas sómente nos Estados



Unidos. Ver, por ex., Morrison, "Bioprocessing in Space -- an Overview", pág. 557-571, em "The World Biotech Report", Volume 2: USA, (Online Publications, New York, N.Y. 1984). Estudos recentes forneceram uma base para projecção da eficácia da terapia com eritropoietina em uma grande variedade de estados de doença, perturbações e estados de irregularidade hematológica: Vedovato e colab., "Acta Haematol.", 71, 211-213 (1984) (beta-talassemia); Vichinsky e colab., "J. Pediatr.", 105 (1), 15-21 (1984) (fibrose cística); Cotes e colab., "Brit. J. Obstet. Gynecol.", 90 (4), 304-311 (1983) (gravidez, perturbações menstruais); Haga e colab., "Acta Pediatr. Scand.", 72, 827-831 (1983) (anemia precoce dos prematuros); Claus-Walker e colab., "Arch. Phys. Med. Rehabil.", 65, 370-374 (1984) (lesão do cordão espinal); Dunn e colab., "Eur. J. Appl. Physiol.", 52, 178-182 (1982) (voo no espaço); Miller e colab., "Brit. J. Haematol.", 52, 545-590 (1982) (perda de sangue aguda); Udupa e colab., "J. Lab. Clin. Med.", 103 (4), 574-580 e 581-588 (1984); e Lipschitz e colab., "Blood", 63 (3), 502-509 (1983) (envelhecimento); e Dainiak e colab., "Cancer", 51 (6), 1101-1106 (1983) e Schwartz e colab., "Otolaryngol", 109, 269-272 (1983) (diversos estados de doenças neoplásticas acompanhados de eritropoiese anormal).

As tentativas anteriormente feitas para se obter eritropoietina com um rendimento bom a partir de plasma ou de urina mostraram ser relativamente mal sucedidas. São necessárias técnicas laboratoriais complicadas e sofisticadas e têm geralmente, como resultado, a obtenção de quantidades muito pequenas de extractos impuros e instáveis que contêm eritropoietina.

A patente de invenção norte-americana Nº 3.033.753



descreve um método de purificação parcial de eritropoietina a partir de plasma de sangue de carneiro que proporciona rendimentos fracos de um extracto sólido bruto que contém eritropoietina.

As tentativas iniciais para isolar eritropoietina a partir de urina proporcionaram preparações instáveis, biologicamente inactivas da hormona. A patente de invenção norte-americana Nº. 3.865.801 descreve um método de estabilização da actividade biológica de uma substância bruta que contém eritropoietina isolada a partir de urina. A preparação bruta resultante que contém eritropoietina retém significativamente 90% de actividade da eritropoietina e é estável.

Um outro método para a purificação de eritropoietina humana de urina de pacientes com anemia aplástica é descrito por Miyake e colab., em "J. Biol. Chem.", Vol. 252, Nº. 15 (10 de Agosto de 1977), pág. 5558-5564. Esta técnica de sete fases compreende uma cromatografia de permuta iónica, uma precipitação com etanol, uma filtração sobre gel e uma cromatografia de adsorção e fornece uma preparação de eritropoietina pura com uma potência de 70.400 unidades/mg de proteína com um rendimento de 21%.

A patente de invenção norte-americana Nº 4.397.840, concedida a Takezawa e colab., descreve métodos para a preparação de "um produto de eritropoietina", a partir de amostras de urina de seres humanos saudáveis, com permutadores iónicos fracamente alcalinos e propõe que os produtos de P.M. baixo "não têm quaisquer efeitos inibidores sobre a eritropoietina".

O pedido de patente de invenção britânica Nº. 2.085.887, apresentado por Sugimoto e colab. e publicado em 6 de Maio de 1977,



1982, descreve um processo para a produção de células linfoblastoides humanas híbridas, referindo níveis de produção compreendidos entre 3 e 420 Unidades de eritropoietina por ml de suspensão de células (distribuídas nas culturas após propagação do hospedeiro-mamífero) que contém até 10^7 células por ml. Nos níveis de produção mais elevados que afirmam ter obtido, pôde calcular-se que a velocidade (taxa) de produção da eritropoietina estava compreendida entre 40 e cerca de 4.000 unidades/ 10^6 células/48 horas na cultura "in vitro" após transferência de células a partir de sistemas de propagação "in vivo". (Ver também a patente de invenção norte-americana Nº. 4.377.513 equivalente). Têm sido feitas muitas sugestões para o isolamento de eritropoietina a partir de fontes de tecidos, incluindo células neoplásticas, mas os rendimentos têm sido muito baixos. Ver, por ex., Jelkman e colab., "Exptl. Hematol.", 11 (7), 581-588 (1983); Tambaurin e colab., "P.N.A.S. (U.S.A.)", 80, 6269-6273 (1983); Katsuoka e colab., "Gann.", 74, 534-541 (1983); Hagiwara e colab., "Blood", 63 (4), 828-835 (1984); e Choppin e colab., "Blood", 64 (2), 341-347 (1984).

Outras técnicas de isolamento utilizadas para se obter eritropoietina pura compreendem técnicas imunológicas. Produz-se um anticorpo policlonal, derivado do soro, dirigido contra a eritropoietina, injectando um animal de preferência um rato ou um coelho, com eritropoietina humana. A eritropoietina humana injectada é considerada como uma substância antigénica estranha pelo sistema imune do animal e induz a produção de anticorpos contra o antigénio. Células diferentes que respondem à estimulação pela substância antigénica produzem e libertam, na circulação, anticorpos ligeiramente diferentes dos produzidos



por outras células correspondentes. A actividade de anticorpo permanece no soro do animal quando o seu sangue é extraído. Embora o soro não purificado ou as preparações de anticorpo purificadas sob a forma de uma fracção de imunoglobulina G do soro possam depois ser utilizados em ensaios para detectar e complexar com eritropoietina humana, os materiais sofrem de uma desvantagem principal. Este anticorpo do soro, constituído por todos os anticorpos diferentes produzidos por células individuais, é de natureza policlonal e irá complexar-se com componentes em extractos brutos diferentes da eritropoietina sózinha.

São de interesse para o fundamento da presente invenção os recentes progressos no campo de produção de culturas contínuas de células capazes de produzirem uma única espécie de anticorpo o qual é específico e imunologicamente reactivo com uma determinante antigénica simples de um antigénio escolhido. Ver na generalidade, Chisholm, "High Technology", Vol. 3, Nº 1, 57-63 (1983). Têm sido feitas tentativas para utilizar técnicas de fusão de células e de hibridação para produzir anticorpos "monoclonais" para a eritropoietina e utilizar estes anticorpos no isolamento e na detecção quantitativa da eritropoietina humana. Como exemplo, apareceu, sob a forma de resumo, em Lee-Huang, Abstract Nº 1463 de "Fed. Proc.", 41, 520 (1982), um relatório do desenvolvimento bem sucedido de séries de células de hibridoma murganho-murganho que segregam anticorpos monoclonais para a eritropoietina humana. Como outro exemplo, aparece uma descrição pormenorizada da preparação e da utilização de um anticorpo monoclonal, anti-eritropoietina, em Weiss e colab., "P.N.A.S. (U.S.A.)", 79, 5465-5469 (1982). Ver tam-



bém, Sasaki, "Biomed. Biochim. Acta", 42(11/12), 5202-5206 (1983); Yanagawa e colab., "Blood", 64 (2), 357-364 (1984); Yanagawa e colab., "J. Biol. Chem.", 259 (5), 2707-2710 (1984); e patente de invenção norte-americana Nº. 4.465.624.

São também de interesse para o fundamento da invenção os relatórios da actividade imunológica de peptidossintéticos que duplicam substancialmente a sequência de aminoácidos existente em proteínas, glucoproteínas e nucleoproteínas naturais. Mais especificadamente, poleptidos de P.M. relativamente baixo têm mostrado participar em reacções imunes que são semelhantes em duração e em grandeza às reacções imunes de proteínas fisiologicamente importantes como antigénios virais, hormonas polipeptídicas, e similares. Está incluída, entre as reacções imunes desses polipeptidos, a indução da formação de anticorpos específicos em animais imunologicamente activos. Ver, por ex. Lerner e colab., "Cell", 23, 309-310 (1981); Ross e colab., "Nature", 294, 654-656 (1981); Walter e colab., "P.N.A.S. (U.S.A.)", 77, 5197-5200 (1980); Lerner e colab., "P.N.A.S. (U.S.A.)", 78, 3403-3407 (1981); Walter e colab., "P.N.A.S. (U.S.A.)", 78, 4882-4886 (1981); Wong e colab., "P.N.A.S. (U.S.A.)", 78 7412-7416 (1981); Green e colab., "Cell", 28, 477-487 (1982); Nigg e colab., "P.N.A.S. (U.S.A.)", 79, 5322-5326 (1982); Baron e colab., "Cell", 28, 395-404 (1982); Dreesman e colab., "Nature", 295, 158-160 (1982); e Lerner, "Scientific American", 248, Nº. 2, 66-74 (1983). Ver também, Kaiser e colab., "Science", 223, 249-255 (1984) que se referem às actividades biológica e imunológica de peptidos sintéticos que partilham aproximadamente estruturas secundárias de hormonas peptídicas mas não partilham a sua conformação estrutural primária. Os estudos



anteriores referem-se evidentemente, a sequências de aminoácidos de proteínas diferentes da eritropoietina, uma substância para a qual não foi publicada qualquer informação substancial sobre a sequência de aminoácidos. No pedido de patente de invenção norte-americana Nº. de Série 463,724, apresentado em 4 de Fevereiro de 1983, por J. Egrie, publicado em 22 de Agosto de 1984 como pedido de patente europeia Nº. 0 116 446, descreve-se uma série de células de hibridoma murganho-murganho (A.T.C.C. Nº. HB 8209) que produz um anticorpo monoclonal, anti-eritropoietina, extremamente específico, que é também especificamente imuno-reactivo com um polipeptido que contém a seguinte sequência de aminoácidos:

NH₂-Ala-Pro-Pro - Arg-Leu - Ile - Cys - Asp - Ser - Arg - Val - Leu - Glu - Arg - Tyr - Leu - Leu - Glu - Ala - Lys - COOH.

A sequência polipeptídica é a atribuída aos primeiros vinte restos de aminoácidos de eritropoietina humana completa, isolada de acordo com o método de Miyake e colab., "J. Biol. Chem.", 252, 5558-5564 (1977), após o que se efectua a análise de aminoácidos pelo sequenciador de fase gasosa (Applied Biosystems, Inc.) de acordo com a técnica de Hewick, M. e colab., "J. Biol. Chem.", 256, 7990-7997 (1981). Ver, também, Sue e colab., "Proc. Nat. Acad. Sci (USA)", 80, 3651-3655 (1981), relativo ao desenvolvimento de anticorpos policlonais contra um 26-mer sintético com base em uma sequência de aminoácidos diferente, e Sytowski e colab., "J. Immunol. Methods", 69, 181-186 (1984).

Embora os anticorpos policlonais e monoclonais descritos antes proporcionem materiais extremamente úteis para utilização em imuno-ensaios para a detecção e a quantificação de eritropoietina e possam ser utilizados na purificação de



afinidade da eritropoietina, parece pouco provável que estes materiais possam fornecer facilmente o isolamento, em grande escala, de quantidades de eritropoietina, a partir de fontes de mamíferos, suficientes para posteriores análise, ensaio clínico e utilização terapêutica potencial, entre limites afastados, da substância no tratamento de, por ex., doença renal crónica em que os tecidos doentes não mantêm a produção de eritropoietina. Por conseguinte, projectou-se para o efeito que os melhores planos para caracterizar completamente a eritropoietina de mamíferos e fornecer grandes quantidades da mesma para diagnóstico potencial e utilização clínica compreendem a aplicação bem sucedida de técnicas recombinantes para efectuar a síntese microbiana do composto em grande escala.

Embora tenham sido feitos esforços substanciais na tentativa de isolamento de sequências de ADN que codificam eritropoietina humana e de outras espécies de mamíferos, nenhum parece ter sido bem sucedido. Isto é devido principalmente à escassez de fontes de tecidos, especialmente de fontes de tecidos humanos, enriquecidos em mRNA que permitam a construção de uma colecção de cADN a partir da qual possa ser isolada, mediante técnicas convencionais, uma sequência de ADN que codifique eritropoietina. Além disso, sabe-se tão pouco da sequência contínua de restos de aminoácidos da eritropoietina que não é possível construir, por ex., sondas longas de polinucleótidos facilmente capazes de utilização segura na avaliação da hibridação ADN/ADN de cADN e, em especial, de colecções de ADN genómico. De maneira ilustrativa, a sequência de vinte aminoácidos utilizada para gerar o anteriormente denominado anticorpo monoclonal produzido por A.T.C.C. Nº. HB 8209 não permitem a



construção de uma sonda inequívoca de oligonucleotido de 60 bases pela maneira descrita por Anderson e colab., Supra. Calcula-se que o gene humano da eritropoietina pode aparecer como um "gene de cópia simples" dentro do genoma humano e, em qualquer caso, é provável que o material genético que codifica a eritropoietina humana constitua menos de 0,00005% de ADN genómico humano total que deveria estar presente em uma colecção genómica.

Até agora, a mais bem sucedida das tentativas referidas conhecidas em métodos recombinantes para proporcionar sequências de ADN apropriadas para utilização na expressão microbiana de quantidades isoláveis de eritropoietina não atingiu o objectivo. Como exemplo, Farber e colab., "Exp. Hematol.", 11, Supl. 14, Abstract 101 (1983) referem a extracção de mRNA a partir de tecidos de rim de babuínos tratados com fenil-hidrazina e a injeção do mRNA em oocitos de "Xenopus laevis" com o resultado bastante transitório de produção "in vitro" de uma mistura de "produtos de transferência" que apresenta as propriedades biológicas da eritropoietina. Mais recentemente, Farber e colab., "Blood", 62, Nº 5, Supl. Nº. 1, Resumo 392, na página 122a (1983) referiram a transferência "in vitro" de mRNA de rim humano por oocitos de rã. Calculou-se que a mistura de produtos de transferência resultante contém aproximadamente 220 mU de um produto de transferência com a actividade da eritropoietina por cada micrograma de mRNA injectado. Embora se saiba que esses níveis de transferência "in vitro" de mRNA exógeno codificando eritropoietina são muito baixos (mesmo em comparação com os níveis anteriormente referidos de transferência de mRNA de babuíno no produto investigado) considera-se que os



resultados obtidos confirmam o rim humano como o sítio de expressão da eritropoietina, permitindo a construção de uma colecção de cADN de rim humano enriquecido a partir da qual pode ser isolado o gene desejado [Ver, também, Farber, "Clin. Res.", 31 (4), 769A (1983)].

Desde a apresentação dos pedidos de patentes norte-americanas N.ºs. de Série 561.024 e 582.185, apareceu apenas uma referência à clonagem e expressão, em "E. coli", do que se afirma ter sido cADN de eritropoietina humana. De uma maneira resumida, inseriu-se um certo número de clónios de cADN em plasmídeos de "E. coli" e observou-se que os produtos de fusão de β -lactamase são imuno-reactivos com um anticorpo monoclonal para um "epitope" não especificado de eritropoietina humana. Ver, Lee-Huang, "Proc. Nat. Acad. Sci. (USA)", 81, 2708-2712 (1984).

Sumário abreviado

A presente invenção proporciona, pela primeira vez, novos produtos polipeptídicos puros e isolados que têm parte ou toda a conformação estrutural primária (isto é, uma sequência contínua de restos de aminoácidos) e uma ou mais das propriedades biológicas (a saber, propriedades imunológicas e actividade biológica "in vivo" e "in vitro") de eritropoietina natural, incluindo as suas variantes alélicas. Estes polipeptídeos são também, de modo único, caracterizados por serem o produto da expressão por um hospedeiro procariótico ou eucariótico (por ex. por bactérias, leveduras e células de mamíferos em cultura) de sequências de ADN exógeno obtidas mediante clonagem genómica ou de cADN ou mediante síntese de genes. Os produtos de expressão microbiana em células de vertebrados (por ex. mamíferos



e aves) podem ainda ser caracterizados pela liberdade de associação com proteínas humanas ou outros contaminantes que podem ser associados com eritropoietina no seu ambiente natural de células de mamíferos ou em fluidos extracelulares como plasma ou urina. Os produtos de levedura típica (por ex. "Saccaromyces cerevisiae") ou de células-hospedeiras procarióticas (por ex. "E. coli") estão livres de associação (não se associam) com quaisquer proteínas de mamíferos. Consoante o hospedeiro utilizado, os polipeptidos de acordo com a invenção podem ser glucosilados com hidratos de carbono de mamíferos ou de outras células eucarióticas ou podem não ser glucosilados. Os polipeptidos de acordo com a invenção podem também conter um resto do aminoácido metionina (na posição 1).

Os novos produtos glucoproteicos de acordo com a presente invenção compreendem os que têm uma conformação estrutural primária suficientemente duplicativa da de uma eritropoietina natural (por ex. humana) para permitir a fruição de uma ou mais das suas propriedades biológicas e têm uma composição média de hidratos de carbono que é diferente da da eritropoietina natural (por ex. humana).

As células de vertebrados (por ex., COS-1 e CHO) fornecidas pela presente invenção compreendem as primeiras células de que alguma vez se pôde dispor que podem ser propagadas continuamente "in vitro" e que, depois de desenvolvidas em cultura, são capazes de produzir, no meio em que se desenvolveram, mais 100 U (de preferência mais de 500 U e, com maior vantagem, mais de 1000 a 5.000 U) de eritropoietina por 10^6 células em 48 horas como se determina mediante ensaio radio-imunológico.

São também fornecidos, pela presente invenção, poli-



peptidos sintéticos total ou parcialmente duplicativos de sequências contínuas de restos de aminoácidos de eritropoietina que aqui são pela primeira vez esclarecidos. Estas sequências, em virtude de partilharem características de conformação e de estrutura primária, secundária ou terciária com a eritropoietina natural, podem possuir uma actividade biológica e/ou propriedades imunológicas em comum com o produto natural de maneira que podem ser utilizados como substitutos biologicamente activos ou imunológicos da eritropoietina em processos terapêuticos e imunológicos. São correspondentemente fornecidos anticorpos monoclonais ou policlonais gerados mediante meios típicos, os quais são imuno-reactivos com esses polipeptidos e, de preferência, são também imuno-reactivos com eritropoietina natural.

São ilustrativas da presente invenção as sequências de ADN clonadas de macaco e da espécie humana e as sequências polipeptídicas apropriadamente delas provenientes que representam, respectivamente, a conformação estrutural primária de eritropoietinas de macaco e da espécie humana.

São também proporcionados pela presente invenção novos vectores de ADN de plasmídeo circular ou viral funcional biologicamente activo que incorporam as sequências de ADN de acordo com a invenção e os organismos-hospedeiros microbianos (por ex. bactérias, levedura e células de mamíferos) estavelmente transformados ou modificados com esses vectores. São correspondentemente proporcionados pela presente invenção novos processos para a preparação de polipeptidos utilizáveis que compreendem o desenvolvimento em cultura desses hospedeiros microbianos transformados ou modificados sob condições que facilitam a



expressão, em grande escala, das sequências de ADN exógeno, provenientes de vectores e o isolamento dos polipeptidos desejados a partir do meio de cultura, de lisados celulares ou de fracções de membranas celulares.

O isolamento e a purificação de polipeptidos expressos microbiologicamente, proporcionados pela presente invenção, podem ser efectuados por processos convencionais que compreendem, por ex. separações cromatográficas preparativas e separações imunológicas que envolvem preparações de anticorpo monoclonal e/ou policlonal.

Tendo sido aqui esclarecida a sequência de restos de aminoácidos de eritropoietina, a presente invenção proporciona a preparação total e/ou parcial de sequências de ADN que codificam a eritropoietina e incluem características vantajosas tais como a incorporação de códiões "preferidos" para expressão por hospedeiros não-mamíferos escolhidos, a provisão de sítios para cisão por enzimas de endonuclease de restrição e a provisão de sequências de ADN iniciais, terminais ou intermédias adicionais que facilitam a construção de vectores prontamente expressos. Correspondentemente, a presente invenção proporciona a preparação (e o desenvolvimento mediante mutagénesse específica do sítio de cADN e de ADN genómico) de sequências de ADN que codificam a expressão microbiana de análogos de polipeptidos ou de derivádos de eritropoietina que diferem das formas existentes na natureza em termos da identidade ou da localização de um ou mais restos de aminoácidos (isto é análogos de supressão que contêm menos de todos os restos especificados para EPO e/ou análogos de substituição em que um ou mais restos especificados são substituídos por outros restos e/ou análogos de adição



em que são adicionados um ou mais restos de aminoácidos a uma porção terminal ou média do polipeptido); e que partilham algumas ou todas as propriedades das formas naturais.

As novas sequências de ADN de acordo com a invenção compreendem todas as sequências utilizáveis para assegurar a expressão em células-hospedeiras procarióticas ou eucarióticas de produtos polipeptídicos que têm pelo menos uma parte da conformação estrutural primária e uma ou mais das propriedades biológicas da eritropietina que são abrangidas: (a) pelas sequências de ADN indicadas nos quadros 5 e 6 mais adiante ou as suas cadeias complementares; (b) pelas sequências de ADN que hibridam (sob condições de hibridação como as ilustradas antes ou sob condições mais rigorosas) para as sequências de ADN definidas em (a) ou seus fragmentos; e (c) pelas sequências de ADN que, a não ser pela degenerescência do código genético, deveriam hibridar para as sequências de ADN definidas antes em (a) e (b). Estão especialmente compreendidas na parte (b) as sequências de ADN genómicas que codificam formas de variantes alélicas de eritropoietina de macaco e humana e/ou codificam outras eritropoietinas de espécies de mamíferos. Estão especificamente compreendidas pela parte (c) as sequências de ADN preparadas que codificam EPO, fragmentos de EPO e análogos de EPO, as quais podem incorporar códiões que facilitam a transferência de ARN mensageiro em hospedeiros não vertebrados.

É abrangida pela presente invenção a classe de polipeptidos codificados por porções do complemento de ADN para a cadeia superior da sequência de ADN genómico humano do quadro 6 adiante, isto é "proteínas invertidas complementares" como descrevem Tramontano e colab., Nucleic Acids Research", 12,



5049-5059 (1984).

São também abrangidas pela invenção as composições farmacêuticas que contêm quantidades eficazes de produtos polipeptídicos de acordo com a presente invenção juntamente com diluentes, adjuvantes e/ou excipientes apropriados que permitam proporcionar uma terapia com eritropoietina, em especial no tratamento de estados de doenças anémicas e muito especialmente de estados anémicos quando se trata da incapacidade renal crónica.

Os produtos polipeptídicos de acordo com a presente invenção podem ser "marcados" mediante associação covalente com uma substância marcadora detectável (por ex. marcados com o isótopo radioactivo ^{125}I) para constituírem reagentes utilizáveis na detecção e na quantificação da eritropoietina em amostras de tecidos sólidos ou de fluídos como sangue ou urina. Os produtos de ADN de acordo com a invenção podem também ser marcados com marcadores detectáveis (como marcas de isótopos radioactivos e marcas não isotópicas como a biotina) e utilizados em processos de hibridação de ADN para localizar a posição do gene da eritropoietina e/ou a posição de qualquer família de genes afins no quadro cromossómico da espécie humana, do macaco e de outros mamíferos. Podem também ser utilizados para identificar as perturbações no gene da eritropoietina ao nível do ADN e utilizados como marcadores de genes para identificar genes vizinhos e suas alterações.

Como se descreverá adiante em pormenor, a presente invenção proporciona ainda melhorias significativas em métodos para a detecção de um polinucleotido específico de cadeia simples de sequência desconhecida em uma amostra celular ou viral



heterogénea que contém polinucleotidos múltiplos de cadeia simples em que

(a) se prepara uma mistura de sondas de polinucleotidos de cadeia simples, marcadas, tendo sequências de bases que variam uniformemente, sendo cada uma das sondas citadas potencial e especificamente complementar em relação a uma sequência de bases que é putativamente única para o polinucleotido a ser detectado,

(b) se fixa a amostra em um substrato sólido,

(c) se trata o substrato que tem a amostra nele fixada para diminuir a ligação ulterior de polinucleotidos ao mesmo, excepto no que se refere à hibridação para polinucleotidos na amostra citada,

(d) se põe transitoriamente o substrato tratado, que nele tem fixada a amostra, em contacto com a citada mistura de sondas marcadas sob condições que facilitam a hibridação apenas entre polinucleotidos totalmente complementares, e

(e) se detecta o polinucleotido específico mediante verificação cuidadosa da presença de uma reacção de hibridação entre ele e uma sonda totalmente complementar dentro da citada mistura de sondas marcadas, como se põe em evidência pela presença de uma densidade mais elevada do material marcado sobre o substrato no local do polinucleotido específico em comparação com uma densidade de base do material marcado resultante da ligação não-específica de sondas marcadas ao substrato.

As técnicas são especialmente eficazes em situações que impõem a utilização de 64, 128, 256, 512, 1024 ou mais sondas de polinucleotidos misturadas com um comprimento de 17 a 20 bases em hibridações ADN/ADN ou ARN/ARN ou ADN/ARN.



Como se descreve adiante, as técnicas melhoradas anteriormente citadas têm permitido ilustrativamente a identificação de clónios de cADN que codificam a eritropoietina de macaco dentro de uma colecção preparada a partir de mRNA de células de rim de macaco anémico. Mais especificadamente, utiliza-se uma mistura de 128 sondas 20-mer que variam uniformemente com base na informação da sequência de aminoácidos, proveniente de fracções de sequenciação da eritropoietina humana, em técnicas de hibridação de colónias para identificar sete clónios cADN de eritropoietina "positivos" dentro de um total de 200.000 colónias. É mesmo muito mais extraordinário que a realização das técnicas melhoradas de acordo com a invenção tenha permitido o isolamento rápido de três clónios positivos a partir de uma avaliação de 1.500.000 placas de fagos que constituem uma colecção genómica humana. Consegue-se isto através da utilização da mistura anteriormente citada de 128 sondas 20-mer juntamente com um segundo conjunto de 128 sondas 17-mer com base na análise dos aminoácidos de uma sequência contínua diferente de eritropoietina humana.

As técnicas ilustrativas anteriormente citadas constituem o primeiro caso conhecido da utilização de sondas misturadas de oligonucleotidos múltiplos em processos de hibridação ADN/ADN orientadas para o isolamento de clónios genómicos de mamíferos e o primeiro caso conhecido da utilização de uma mistura de mais de 32 sondas de oligonucleotidos no isolamento de clónios cADN.

Tornar-se-ão evidentes numerosas modalidades e vantagens da invenção para os entendidos na matéria após consideração da seguinte descrição pormenorizada que proporciona es-



clarecimentos da prática da invenção nos seus modos de realização presentemente preferidos.

Descrição pormenorizada

De acordo com a presente invenção, foram isoladas e caracterizadas sequências de ADN que codificam parte ou toda a sequência polipeptídica da eritropoietina humana e da espécie de macaco (de aqui em diante, por vezes designada por "EPO"). Além disso, o ADN de origem humana ou do macaco foi submetido a expressão eucariótica e procariótica que fornece quantidades isoláveis de polipeptidos que apresentam as propriedades biológicas (por ex. imunológicas) da EPO natural assim como as actividades biológicas da EPO "in vivo" e "in vitro".

O ADN originário do macaco foi isolado a partir de uma colecção de cADN construída com mARN proveniente de tecido do rim de um macaco em estado anémico induzido quimicamente e cujo soro foi avaliado imunologicamente para incluir níveis elevados de EPO em comparação com o soro do macaco normal. Conseguiu-se o isolamento dos clónios cADN desejados que contêm ADN codificador de EPO através da utilização da hibridação de colónia de ADN/ADN, utilizando um lote de 128 sondas de oligonucleotido 20-mer, misturadas e marcadas com um isótopo radioactivo e implicou a avaliação (escolha) rápida de 200.000 colónias. A concepção das sondas de oligonucleotido apoiou-se na informação sobre a sequência de aminoácidos fornecida pela fragmentação enzimática e a sequenciação de uma pequena amostra de EPO humana.

O ADN da espécie humana foi isolado a partir de uma colecção de ADN genómico humano. Conseguiu-se o isolamento de clónios contendo ADN codificador de EPO através da hibridação



em placa de ADN/ADN, utilizando o lote citado antes de 128 sondas de oligonucleotido 20-mer misturadas e um segundo lote de 128 sondas 17-mer marcadas com isótopos radioativos cujas sequências se baseiam na informação sobre sequências de aminoácidos obtidas a partir de um fragmento enzimático e diferente de EPO humana.

As colónias e placas positivas foram verificadas mediante sequenciação dideoxi de ADN clonal, utilizando um subconjunto de 16 sequências dentro do lote de sondas 20-mer e submeteu-se os clónios escolhidos a análise da sequência de nucleotidos de que resulta a dedução da conformação estrutural primária dos polipeptidos da EPO assim codificados. As sequências de polipeptidos deduzidas apresentavam uma homologia de elevado grau de umas para as outras e para uma sequência parcial gerada pela análise de aminoácidos de fragmentos de EPO humana.

Foram inseridos um clónio cADN positivo escolhido de macaco e um clónio genómico positivo escolhido humano em um vector de ADN "shuttle" que foi amplificado em "E. coli" e utilizado para modificar células de mamíferos em cultura. Da cultura de células-hospedeiras modificadas resultaram preparações sobrenadantes do meio de cultura que se avaliou conterem umas 3 000 mU de EPO por ml de fluido de cultura.

Os exemplos seguintes são apresentados a título ilustrativo da invenção e são especificamente dirigidos às técnicas efectuadas antes para identificação de clónios cADN de macaco e clónios genómicos humanos codificadores de EPO, às técnicas de que resulta essa identificação e à sequenciação, desenvolvimento de sistemas de expressão e verificação imunológica da



expressão de EPO nesses sistemas.

Mais especificadamente, o exemplo 1 é dirigido para a sequenciação de aminoácidos de fragmentos de EPO humana e para a construção de misturas de sondas marcadas com isótopos radioactivos com base nos resultados desta sequenciação. O exemplo 2 está orientado, de um modo geral, para técnicas envolvidas na identificação de clónios cADN positivos de macaco e fornece, deste modo, informação relativa ao tratamento do animal e à análise do ensaio rádio-imunológico preliminar de soros animais. O exemplo 3 está orientado para a preparação da colecção de cADN, avaliação da hibridação em colónia e verificação de clónios positivos, sequenciação de ADN de um clónio cADN positivo e a criação de informação sobre a conformação estrutural primária (sequência de aminoácidos) do polipeptido de EPO de macaco. O exemplo 4 está orientado para técnicas envolvidas na identificação de clónios genómicos humanos positivos e fornece, deste modo, informação relativa à fonte da colecção genómica, a técnicas de hibridação em placa e à verificação de clónios positivos. O exemplo 5 dirige-se à sequenciação de ADN de um clónio genómico positivo e à produção de informação sobre a sequência de aminoácidos do polipeptido de EPO humana incluindo uma comparação desta com a informação sobre a sequência de aminoácidos da EPO de macaco. O exemplo 6 está orientado para técnicas de construção de um vector que incorpora ADN codificador de EPO derivado de um clónio cADN positivo de macaco, para a utilização do vector para a modificação de células COS-1 e desenvolvimento, em cultura, das células modificadas. O exemplo 7 está orientado para a construção de um vector que incorpora ADN codificador de EPO, derivado de um clónio genómico hu



mano positivo, para a utilização do vector na transformação de células COS-1 e para o desenvolvimento, em cultura, das células transformadas. O exemplo 8 está dirigido para técnicas de imuno-ensaio efectuadas sobre meios sobrenadantes obtidos a partir da cultura de células transformadas de acordo com os exemplos 6 e 7. O exemplo 9 está orientado para a actividade biológica "in vitro" e "in vivo" de EPO expressa microbiologicamente dos exemplos 6 e 7.

O exemplo 10 refere-se à produção de sistemas de expressão do hospedeiro mamífero para o cADN da EPO de macaco e para o ADN genómico humano que envolve células ("CHO") do ovário do criceto chinês e às actividades imunológica e biológica dos produtos destes sistemas de expressão, assim como à caracterização destes produtos. O exemplo 11 está orientado para a preparação de genes preparados que codificam a EPO humana e análogos de EPO, genes estes que contêm um grande número de códiços preferenciais para expressão em células-hospedeiras de "E.coli" e de levedura e para os sistemas de expressão neles baseados. O exemplo 12 refere-se aos perfis de actividade imunológica e biológica de produtos de expressão dos sistemas do exemplo 11.

Exemplo 1

A. Determinação da sequência de aminoácidos de um fragmento de EPO humana

Isolou-se EPO humana a partir de urina e submeteu-se aquela a digestão triptica de que resultou o desenvolvimento e o isolamento de 17 fragmentos separados em quantidades de aproximadamente 100-150 picomoles.

Os fragmentos, a que se atribuiu números arbitrariamente, foram analisados para determinação da sequência de ami-



noácidos mediante análise de micro-sequência, utilizando um sequenciador de fase gasosa (Applied Biosystems) para fornecer a informação sobre a sequência indicada no Quadro 1 a seguir, em que se utiliza códigos de letra simples e em que o símbolo "X" designa um resto que não foi determinado inequivocamente.



Quadro 1

<u>Fragmento nº.</u>	<u>Resultado da análise da sequência</u>
T4a	A-P-P-R
T4b	G-K-L-K
T9	A-I-G-A-Q-K
T13	V-I-E-R
T16	A-V-S-G-L-R
T18	L-F-R
T21	K-I-F-R
T25	Y-I-L-E-A-K
T26a	L-I-C-D-S-R
T26b	L-Y-T-G-E-A-C-R
T27	T-I-T-A-D-T-F-R
T28	E-A-I-S-P-P-D-A-A-M-A-A-P-L-R
T30	E-A-E-X-I-T-T-G-X-A-E-H-X-S-L- N-E-X-I-T-V-P
T31	V-Y-S-N-F-L-R
T33	S-I-T-T-L-L-R
T35	V-N-F-Y-A-W-K
T38	G-Q-A-L-L-V-X-S-S-Q-P-W- E-P-L-Q-L-H-V-D-K



B. Concepção e construção de misturas de sondas de oligonucleotídeos

As sequências de aminoácidos indicadas no Quadro 1 foram passadas em revista no contexto da degenerescência do código genético com o fim de determinar se era possível aplicar as técnicas de sondas misturadas às técnicas de hibridação de ADN/ADN sobre colecções de cADN e/ou de ADN genómico. Esta análise revelou que, dentro do Fragmento Nº. T35, existia uma série de 7 restos de aminoácidos (Val-Asn-Phe-Tyr-Ala-Trp-Lys) que apenas podia ser caracterizada quando codificada por umas das 128 sequências de ADN possíveis que abrangiam 20 pares de bases. Sintetizou por isso um primeiro conjunto de 128 oligonucleotídeos 20-mer por métodos de fosfoamidito normal [Ver, por ex. Beaucage e colab. "Tetrahedron Letters", 22, 1859-1862 (1981)] sobre um suporte sólido de acordo com a sequência indicada no Quadro 2 a seguir.

Quadro 2

<u>Resto</u>	-	<u>Val</u>	-	<u>Asn</u>	<u>Phe</u>	<u>Tyr</u>	<u>Ala</u>	<u>Trp</u>	<u>Lys</u>	
3'		CAA		TTG	AAG	ATG	CGA	ACC	TT	- 5'
		T		A	A	A	T			
		G					G			
		C					C			

Uma outra análise revelou que, dentro do fragmento Nº. T38, existia uma série de 6 restos de aminoácidos (Gln-Pro-Trp-Glu-Pro-Leu) com base nos quais se podia preparar um lote de 128 sondas 17-mer de oligonucleotídeos misturados como se indica no Quadro 3 a seguir.



Quadro 3

<u>Resto</u>	-	<u>Gln</u>	<u>Pro</u>	<u>Trp</u>	<u>Glu</u>	<u>Pro</u>	<u>Leu</u>	
3'		GTT	GGA	ACC	CTT	GGA	GA	- 5'
		C	T		C	T	A	
			G			G		
			C			C		

As sondas de oligonucleotidos foram marcadas na extremidade 5' com gama - ^{32}P - ATP, 7500-8000 Ci/milimole (ICN) utilizando quinase de polinucleotido T_4 (NEN).

Exemplo 2

A. Técnicas de tratamento do macaco e análise de RIA

Tratou-se macacos *Cynomolgus* fêmeas ("Macaca fascicularias"), com 2,5-3 Kg de peso e 1,5-2 anos de idade, por via subcutânea, com uma solução de cloridrato de fenil-hidrazina de pH 7,0 a um nível de dosagem de 12,5 mg/Kg nos dias 1, 3 e 5. Monitorizou-se o hematocrito antes de cada injeção. No dia 7, ou sempre que o nível do hematocrito caiu abaixo de 25% do nível inicial, colheu-se o soro e os rins após administração de doses de 25 mg/Kg de cloridrato de cetamina. Os materiais colhidos foram imediatamente congelados em azoto líquido a -70°C .

B. RIA da EPO (Ensaio radioimunológico da eritropoietina)

As técnicas de ensaio radioimunológico aplicadas para a detecção quantitativa de EPO em amostras foram efectuadas de acordo com o seguinte modo de realização:

Incubou-se um padrão de eritropoietina ou uma amostra desconhecida juntamente com anti-soro, durante 2 horas, a 37°C . Após incubação de duas horas, arrefeceu-se os tubos com a



amostra sobre gelo, adicionou-se eritropoietina marcada com ^{125}I e incubou-se os tubos a 0°C durante pelo menos 15 horas mais. Cada tubo de ensaio continha $500\ \mu\text{l}$ de mistura de incubação, constituída por $50\ \mu\text{l}$ de soro imune diluído, 10.000 cpm de ^{125}I -eritropoietina, $5\ \mu\text{l}$ de trasilol e $0-250\ \mu\text{l}$ ou do padrão de EPO ou da amostra desconhecida, perfazendo o restante volume com PBS contendo 0,1% de BSA. O anti-soro utilizado foi a segunda sangria experimental de um coelho imunizado com uma preparação com um grau de pureza de 1% de eritropoietina urinária humana. Ajustou-se, no ensaio, a diluição final do anti-soro de maneira que a ^{125}I -EPO ligada ao anticorpo não excedesse 10 a 20% das contagens totais de "input". De um modo geral, isto correspondia a uma diluição final do soro compreendida entre 1:50.000 e 1:100.000.

Precipitou-se a ^{125}I -eritropoietina ligada ao anticorpo mediante adição de $150\ \mu\text{l}$ de Staph A. Após uma incubação de 40 minutos, centrifugou-se as amostras e lavou-se os sedimentos duas vezes com $0,75\ \text{ml}$ de 10mM Tris-HCl, de pH 8, 2, contendo $0,15\ \text{M}$ de cloreto de sódio, 2mM de ácido etilenodiaminotetraacético e 0,05% de Triton X-100. Fez-se a contagem dos sedimentos lavados em um contador gama para determinar a percentagem de ^{125}I -eritropoietina ligada. As contagens ligadas pelos soros pré-ímmes foram subtraídas de todos os valores finais para corrigir a precipitação não específica. Determinou-se o teor de eritropoietina das amostras desconhecidas por comparação com a curva padrão.

A técnica anterior foi aplicada ao soro de macaco obtido na Parte A anterior assim como ao soro de macaco não tratado. Ensauiou-se os níveis do soro normal de maneira a con-



ter aproximadamente 36 mU/ml enquanto o soro de macaco tratado continha desde 1000 a 1700 mU/ml.

Exemplo 3

A. Construção da colecção de cADN de macaco

Isolou-se ARN mensageiro, a partir de rins de macacos normal e anémico, de acordo com a técnica do tiocianato de guanidínio de Chirgwin e colab., "Biochemistry", 18, 5294 (1974) e purificou-se o poli (A)⁺ mARN mediante duas cromatografias de coluna sobre oligo (dT-celulose) como se descreve a pág.197-198 em Maniatis e colab., "Molecular Cloning, A Laboratory Manual" (Cold Springs Harbor Laboratory, Cold Springs, Harbor, N.Y., 1982). A colecção de cADN foi construída de acordo com uma modificação das técnicas gerais de Okayama e colab., "Mol. and Cell. Biol.", 2, 161-170 (1982). As características essenciais das técnicas presentemente preferidas foram as seguintes: (1) utilizou-se pUC 8 como vector único; cortou-se este com PstI e completou-se acrescentando depois com oligo dT 60 a 80 bases em comprimento; (2) utilizou-se a digestão com Hinc II para remover a extremidade oligo dT de uma extremidade do vector; (3) efectuou-se a primeira síntese da cadeia e o acrescentamento do oligo dG de acordo com a técnica publicada; (4) utilizou-se a digestão com Bam HI para remover a extremidade oligo dG de uma ponta do vector; e (5) a substituição da cadeia de ARN por ADN foi feita na presença de dois ligantes (GATCTAAAGACCG TCCCCCCCCC e ACGGTCTTTA) em um excesso três vezes molar em relação ao vector com a extremidade oligo dG.



B. Técnicas de hibridação em colónia para a selecção da colecção de cADN de macaco

Espalhou-se "E. coli" transformado em uma densidade de 9.000 colónias, por placa de 10x10 cm, sobre placas nutritivas que continham 50 microgramas/ml de Ampicilina. Humedeceu-se previamente filtros GeneScreen (New England Nuclear Catalog Nº.NEF-972) sobre uma placa BHI-CAM (37g/l de infusão de cérebro-coração Bacto, 2g/l de casaminoácidos e 15g/l de agar, contendo 500 microgramas/ml de cloranfenicol) que foram utilizados para destacar as colónias da placa. As colónias foram cultivadas no mesmo meio de cultura durante 12 horas ou mais para amplificar os números de cópias de plasmídeo. As colónias aumentadas (face superior da colónia) foram tratadas pondo os filtros em série sobre 2 pedaços de papel Whatman 3 MM saturados com cada uma das seguintes soluções:

- (1) 50 mM de glucose-25 mM de Tris-HCl (pH 8,0)-10mM de ácido etilenodiaminotetraacético (pH 8,0) durante 5 minutos;
- (2) 0,5 M de hidróxido de sódio durante 10 minutos; e
- (3) 1,0 M de Tris-HCl (pH 7,5) durante 3 minutos.

Os filtros foram depois secos ao ar em uma estufa a 80°C, durante 2 horas.

Os filtros foram depois submetidos a digestão com Proteinase K através do tratamento com uma solução que continha 50 microgramas/ml do enzima protease em solução-tampão K $\angle$ 0,1 M de Tris-HCl (pH 8,0)-0,15 M de cloreto de sódio-10 mM de ácido etilenodiaminotetraacético (pH 8,2)-0,2% de SDS $\angle$. Mais especificamente, adicionou-se 5 ml da solução a cada filtro e deixou-se



que a digestão se produzisse a 55°C durante 30 minutos, após o que se removeu a solução.

Tratou-se depois os filtros com 4 ml de uma solução-tampão de pré-hibridação (5×SSPE-0,5% de SDS-100 microgramas/ml de ADN de "E.coli" SS-5×BFP). O tratamento de pré-hibridação foi efectuado a 55°C, geralmente durante 4 horas ou mais, após o que se remove a solução-tampão de pré-hibridação.

O processo de hibridação foi efectuado pela maneira seguinte. Adicionou-se, a cada filtro, 3 ml de solução-tampão de hibridação (5×SSPE-0,5% de SDS-100 microgramas/ml de ARN de levedura) que continha 0,025 picomole de cada uma das 128 sequências de sonda do Quadro 2 (sendo a mistura total designada por mistura EPV) e manteve-se os filtros a 48°C durante 20 horas. Esta temperatura é 2°C inferior à mais baixa das temperaturas de dissociação calculadas (T_d) determinadas para qualquer das sondas.

Depois da hibridação, lavou-se os filtros, três vezes, durante 10 minutos, sobre um agitador, com 6×SSC-0,1% de SDS à temperatura ambiente e lavou-se duas a três vezes com 6×SSC-1% de SDS à temperatura da hibridação (48°C).

A auto-radiografia dos filtros revelou sete clónios positivos entre as 200.000 colónias avaliadas.

A análise da sequência inicial de um dos clónios cADN putativos de macaco (designado clónio 83) foi efectuada com fins de verificação por uma modificação da técnica de Wallace e colab., "Gene", 16, 21-26 (1981). Resumindo, o ADN de plasmídeo do clónio cADN 83 de macaco foi linearizado mediante digestão com Eco RI e desnaturado mediante aquecimento em um banho-Maria fervente. A sequência de nucleotídeos foi determinada



pelo método dideoxi de Sanger e colab., "P.N.A.S. (U.S.A.)", 74, 5463-5467 (1977). Utilizou-se, como iniciador para as reações de sequenciação, um sub-conjunto da mistura EPV de sondas constituído por 16 sequências.

C. Sequenciação de cADN de EPO de macaco

A análise da sequência de nucleotídeos do clónio 83 pelas técnicas de Messing, "Methods in Enzymology", 101, 20-78 (1983). Indica-se, no Quadro 4, uma análise do mapa de restrição preliminar do fragmento clonado com Eco RI/Hind III do clónio 83 com aproximadamente 1600 pares de bases. As localizações aproximadas dos sítios de reconhecimento do enzima endonuclease de restrição são indicadas em termos de números de bases 3' para o sítio Eco RI na extremidade 5' do fragmento. A sequência dos nucleotídeos foi efectuada mediante sequenciação dos fragmentos de restrição individuais com a intenção de emparelhar fragmentos sobrepostos. Por exemplo, uma sobreposição da informação da sequência mediante análise de nucleotídeos em um fragmento de restrição designado C113 (Sau 3 A a ~111/Sma I a ~324) e a sequenciação de ordem inversa de um fragmento designado C73 (Alu I a ~ 424/Bst E II a ~ 203).

Quadro 4

Sítio de reconhecimento
do enzima de restrição

Localização (ões) aproximada(s)

<u>EcoRI</u>	1
<u>Sau3A</u>	111
<u>SmaI</u>	180
<u>BstEII</u>	203
<u>SmaI</u>	324
<u>KpnI</u>	371
<u>RsaI</u>	372
<u>AluI</u>	424
<u>PstI</u>	426
<u>AluI</u>	430
<u>HpaI</u>	466
<u>AluI</u>	546
<u>PstI</u>	601
<u>PvuII</u>	604
<u>AluI</u>	605
<u>AluI</u>	782
<u>AluI</u>	788
<u>RsaI</u>	792
<u>PstI</u>	807
<u>AluI</u>	841
<u>AluI</u>	927
<u>NcoI</u>	946
<u>Sau3A</u>	1014
<u>AluI</u>	1072
<u>AluI</u>	1115
<u>AluI</u>	1223

.../...

Quadro 4

Sítio de reconhecimento

do enzima de restriçãoLocalização (ões) aproximada(s)

<u>PstI</u>	1301
<u>RsaI</u>	1343
<u>AluI</u>	1384
<u>HindIII</u>	1449
<u>AluI</u>	1450
<u>HindIII</u>	1585



A sequenciação de aproximadamente 1342 pares de bases (dentro da região que vai do sítio Sau 3A de 3' até ao sítio EcoRI e o sítio HindIII) e a análise de todos os quadros de leitura possíveis teve em conta o desenvolvimento da informação sobre o ADN e a sequência de aminoácidos indicada no Quadro 5. No quadro, o resto de aminoácido inicial putativo do grupo amino terminal da EPO completa (verificada mediante correlação com a análise da sequência mencionada antes de vinte restos com grupo amino terminal) é designado pelo numeral + 1. A presença de um códão ATG específico da metionina (designado -27) "a montante" do resto alanina inicial com um grupo amino terminal como o primeiro resto designado para a sequência de aminoácidos da proteína completa é indicativa da probabilidade de a EPO ser inicialmente expressa no citoplasma em uma forma precursora que compreende uma região "condutora" de 27 aminoácidos que é cortada antes da entrada da EPO completa em circulação. Os sítios de glucosilação potencial dentro do polipeptido são designados por asteriscos. O P.M. avaliado da região transferida foi determinado como sendo 21.117 daltons e o P.M. dos 165 restos do polipeptido que constitui a EPO de macaco completa foi determinado como sendo 18.236 daltons.



QUADRO 5

Transferência do ADN de EPO do macaco

Sau3A
GATCCGGGCCCCCTGGACAGCCGCCCTCTCTCCAGGCCCGTGGGGCTGGCCCCTGCCCC
CGCTGAACCTCCCGGGATGAGGACTCCCGGTGGTACCCGGCGCCCTAGGTGGCTGAG

Met Gly Val His Glu Cys Pro Ala Trp
-27
GGACCCCGGCCAGCGCGGAGATG GGG GTG CAC GAA TGT CCT GCC TGG
-20

-10

Leu	Trp	Leu	Leu	Ser	Leu	Val	Ser	Leu	Pro	Gly	Leu	Pro
CTG	TGG	CIT	CIC	CIG	TCT	CIC	TGG	CIC	CCT	CTG	CTC	CCA

	-1	+1	10										
Val	Pro	Gly	Ala	Pro	Pro	Arg	Ile	Cys	Asp	Ser	Arg	Val	Leu
GTC	CCG	GGC	GCC	CCA	CCA	CGC	CTC	TGT	GAC	AGC	CGA	GTC	CTG

20
Glu Arg Tyr Leu Leu Glu Ala Lys Glu Ala Glu Asn Val Thr Met
GAG AGG TAC CTC TTG GAG GCC AAG GAG GCC GAG AAT GTC ACG ATG

[illegible]

40

Quadro V (continuação)

Asp	Thr	Lys	Val	Asn	Phe	Tyr	Ala	Trp	Lys	Arg	Met	Glu	Val	Gly
GAC	ACC	AAA	GTT	AAC	TTC	TAT	GCC	TGG	AAG	AGG	ATG	GAG	GTC	GGG
50														
Gln	Gln	Ala	Val	Glu	Val	Trp	Gln	Gly	Leu	Ala	Leu	Leu	Ser	Glu
CAG	CAG	GCT	GTA	GAA	GTC	TGG	CAG	GGC	CTG	GCC	CTG	CTC	TCA	GAA
60														
Ala	Val	Leu	Arg	Gly	Gln	Ala	Val	Leu	Ala	Asn	Ser	Ser	Gln	Pro
GCT	GTC	CTG	CGG	GGC	CAG	GCC	GTC	TTG	GCC	AAC	TCT	TCC	CAG	CCT
80														
*														
Phe	Glu	Pro	Leu	Gln	Leu	His	Met	Asp	Lys	Ala	Ile	Ser	Gly	Leu
TTC	GAG	CCC	CTG	CAG	CTG	CAC	ATG	GAT	AAA	GCC	ATC	AGT	GGC	CTT
90														
Arg	Ser	Ile	Thr	Thr	Leu	Leu	Arg	Ala	Leu	Gly	Ala	Gln	Glu	Ala
CGC	AGC	ATC	ACC	ACT	CTG	CTT	CGG	GCG	CTG	GGA	GCC	CAG	GAA	GCC
110														
Ile	Ser	Leu	Pro	Asp	Ala	Ala	Ser	Ala	Ala	Pro	Leu	Arg	Thr	Ile
ATC	TCC	CTC	CCA	GAT	GCG	GCC	TGC	GCT	GCT	CCA	CTC	CGA	ACC	ATC
120														
Thr	Ala	Asp	Thr	Phe	Cys	Lys	Leu	Phe	Arg	Val	Tyr	Ser	Asn	Phe
ACT	GCT	GAC	ACT	TTC	TGC	AAA	CTC	TTC	CGA	GTC	TAC	TCC	AAT	TTC
140														



Quadro 5 (Continuação)

150 160
 Leu Arg Gly Lys Leu Lys Leu Tyr Thr Gly Glu Ala Cys Arg Arg
 CTC CGG GGA AAG CTG AAG CTG TAC ACG GGG GAG GCC TGC AGG AGA
 165
 Gly Asp Arg OP
 GGG GAC AGA TGA CCAGGTGGTCCAGCTGGGCACATCCACCACCTCCCTCACCAACA
 CTGCCTGTGCCACACCCCTCCCTCACCACTCCCGAACCCCATCGAGGGGCTCTCAGCTAAG
 CGCCAGCCTGTCCCATGGACACTCCAGTGCCAGCAATGACATCTCAGGGGCCAGAGGAAC
 TGTCCAGAGCACAACTCTGAGATCTAAGGATGTCCAGGGCCAACTTGAGGGCCAGAGC
 AGGAAGCATTCAGAGAGCAGCTTTAAACTCAGGAGCAGAGACAATGCAGGGAAACACCT
 GAGCTCACTCGGGCCACCTGC AAAATTTGATGCAGGACACGCTTTGGAGGCAATTTACCTG
 TTTTTCACCTACCATCAGGGACAGGATGACTGGAGAACTTAGGTGGCAAGCTGTGACTT
 CTC AAGGCCCTCAGGGCACCTCCCTTGGTGGCAAGAGCCCCCTTGACACTGAGAGAAATTT
 TTGCAATCTGCAGCAGGAAAAATACGGACAGGTTTTGGAGGTGGAGGGTACTTGACAG
 GTGTGTGGGAAGCAGGGGGTAGGGGTGGAGCTGGGATGCGAGTGAGAACCGTGAAGAC
 AGGATGGGGGCTGGCCTCTGTTCTCGTGGGGTCC AAGCTT
 HindiII





A sequência polipeptídica do quadro 5 pode ser prontamente submetida a análise para avaliação de regiões extremamente hidrófilas e/ou características conformacionais secundárias indicativas de regiões potencialmente muito imunogénicas, por ex. pelos métodos de Hopp e colab., "P.N.A.S. (U.S.A.)", 78, 3824-3828 (1981) e Kyte e colab., "J. Mol. Biol.", 157, 105-132 (1982) e/ou Chou e colab., "Biochem.", 13, 222-245 (1974) e "Advances in Enzimolgy" , 47, 45-47 (1978). Dispõe-se da análise assistida por computador de acordo com o método de Hopp e colab., utilizando um programa designado "PEP Reference Section 6,7" fornecido por Intelligenetics, Inc., 124 University Avenue, Palo Alto, California.

Exemplo 4

A. Colecção genómica humana

Obtém-se uma colecção genómica de fígado de feto humano produzida por fago Ch4 A de acordo com as técnicas de Lawn e colab., "Cell", 18, 533-543 (1979) que é mantida para utilização em um ensaio de hibridação em placa.

B. Técnicas de hibridação em placa para avaliação da colecção genómica humana

Desintegrou-se partículas de fagos e fixou-se os ADN sobre filtros (50.000 placas por filtro) de acordo com as técnicas de Woo, "Methods in Enzimology", 68, 389-395 (1979), excepto no que se refere à utilização de filtros GeneScreen Plus (New England Nuclear Catalog Nº. NEF-976) e a placas NZYAM (5g de NaCl; 2 g de $MgCl_2 \cdot 6H_2O$; 10 g de NZ-Amina A; 5 g de extracto de levedura; 2 g de casaminoácidos; 2 g de maltose; e 15 g de agar por litro).



Os filtros secos ao ar foram cozidos a 80°C durante uma hora e digeridos depois com Proteinase K como se descreve no Exemplo 3, Parte B. A pré-hibridação foi efectuada com uma solução-tampão de 1M NaCl - 1% de SDS, a 55°C durante 4 horas ou mais, após o que se remove a solução-tampão. As lavagens de hibridação e post-hibridação foram efectuadas pela maneira descrita no Exemplo 3, Parte B. Foram utilizadas tanto a mistura de 128 sondas 20-mer designadas EPV como a mistura de 128 sondas 17-mer do Quadro 3 (designada mistura EPQ). A hibridação foi efectuada a 48°C, utilizando a mistura de sondas EPV. A hibridação com a mistura de sondas EPQ foi efectuada a 46°C, 4° abaixo da T_d calculada mais baixa para membros da mistura. A remoção da sonda de hibridação é feita mediante fervura com 1xSSC-0,1% de SDS durante 2 minutos. A auto-radiografia dos filtros revelou três clónios positivos (reactivos com ambas as misturas de sondas) entre as 1.500.000 placas de fagos apreciadas. A verificação de que os clónios positivos codificavam a hemato_poietina (EPO) foi obtida através da sequenciação de ADN e a visualização micrográfica electrónica de formação heteroduplex com o cADN de macaco do Exemplo 3. Esta técnica pôs também em evidência os intrões múltiplos na sequência de ADN genómico.

Exemplo 5

Efectuou-se a análise da sequência de nucleotídeos de um dos clónios positivos (designados hEl) e os resultados obtidos estão indicados no Quadro 6.

Quadro 6

AAGCTTCTGGGCTTCCAGACCCAGCTACTTTGCGGAACCTCAGCAACCCAGGCATCTCTGAGTCTCCGCCCA
AGACCGGGATGCCCCCCAGGGGAGGTGTCCGGGAGCCCAAGCCTTTCCAGATAGCACGCTCCGCCAGTCCC
AAGGGTGGCAACCGGCTGCACTCCCCCTCCCCGGAÇCCAGGGCCÇGGGAGCAGCCCCCATGACCCACACGC
ACGTCITCAGCAGCCCCGCTCACGCCCCGGCAGCCTCAACCCAGGCGTCTTGCCCCCTGCTCTGACCCCGG
GTGGCCCCTACCCCTGGCGACCCCTCACGCACACAGCCTCTCCCCCACCCCCACCCGCGCACACATG
CAGATAACAGCCCCGACCCCGGCCAGAGCCGXAGAGTCCCTGGGCCACCCCGGCCGCTGCCCTGCCGCTG
CGCCGCACCGCGCTGTCTCCCGGAGCCGGAÇCGGGGCCACCGGCCCXGCTCTGCTCCGACACCGCGCCC
CTTGGAÇAGCCGCCCTCTCTCTAGGCCCGTGGGGCTGGCCCTGCACCGCCGAGCTTCCCGGATGAGGXX
-27 -24
Met Gly Val His
ATG GGG GTG CAÇ G
CCCCGTGACCGCGCGCCCCAAGTCGCTGAGGGACCCCGGCCAAGCGCGGAG
GTGAGTACTCGCGGGCTGGGCGCTCCCGCGGCGCGGGTTCCTGTTGAGCGGGGATTTAGCGCCCCGGCT



Quadro 6 (Continuação)

ATTGCCAAGAGGTGGCTGGGTTC AAGACCGCGGACTTGTCAAGACCCCGAAGGGGAGGGGGTGGG
GCAGCCTCCACGTGCCGCGGGGACTTGGGGGAGTTCTTGGGGATGCAAAACCTGGCCTGTGTGAGGGGCA
CAGTTTGGGGTGGGGAGGAGTTTGGGGTCTCTGCTGTCAGTTGTGTCGTGTGTCAGTGTCTCG[I.S.]
TTGCACACGCACAGATCAATAAGCCAGAGGCACCTGAGTGTTCATGGTTGGGACAGGAAGGACGAG
CTGGGGCAGAGACGTGGGGATGAAGGAAGCTGTCTTCCACAGCCACCCCTTCTCCCCCCTGACTCT
CAGCCTGGCTATCTGTCTAG -23 -20
Glu Cys Pro Ala Trp Leu Trp Leu Leu Leu Ser Leu
AA TGT CCT GCC TGG CTG TGG CTT CTC CTG TCC CTG
-10 -1 +1
Leu Ser Leu Pro Leu Gly Leu Val Leu Gly Ala Pro Pro Arg Leu Ile Cys
CTG TCG CTC CCT CTG GGC CTC CCA GTC CTG GGC GCC CCA CCA CGC CTC ATC TGT
Asp Ser Arg Val Leu Glu Arg Tyr Leu Leu Glu Ala Lys Glu Ala Glu Asn Ile
GAC AGC CGA GTC CTG GAG AGG TAC CTC TTG GAG GCC AAG GAG GCC GAG AAT ATC
26
Thr
ACG GTGAGACCCCTTCCCCAGCACATTCACAGAACTCACGCTCAGGGCTTCAGGGAACCTCCTCCAGAT
CCAGGAACCTGGCACTTGGTTGGGGTGGAGTTGGGAAGCTAGACACTGCCCCCTACATAAGAATAAGTC.



Quadro 6 (Continuação)

TGGTGGCCCCAAACCATACCTGAAACTAGGCAAGGAGCAAGCCAGCAGATCCTACGCCCTGTGGGCCAGGG
 CCAGAGCCCTTCAGGGACCCCTTGACTCCCCGGGCTGTGTGCATTTCAG
 His Cys Ser Leu Asn Glu Asn Ile Thr Val Pro Asp Thr Lys Val Asn Phe Tyr
 CAC TGC AGC TTG AAT GAG AAT ATC ACT GTC CCA GAÇ ACC AAA GTT AAT TTC TAT
 * 40
 50 55
 Ala Trp Lys Arg Met Glu
 GCC TGG AAG AGG ATG GAÇ GIGAGTTCCTTTTTTTTTTTCCTTTCTTTTGGAGAAATCTCAT
 TGGGAGCCGTGATTTGGATGAAAGGGAGAATGATCGGGGGAAGGTAAAATGGAGCAGCAGAGATGAGGCT
 GCCTGGGCGCAGAGGCTCACGTCTATAATCCCAGGCTGAGATGGCCGAGATGGGAGAAATGCTTGAGCCCT
 GGAGTTTCAGACCAACCTAGGCAGCATAGTGAGATCCCCCATCTCTACAAACATTTAAAAAAATTAGTCAG
 GTGAAGTGGTGCATGGTGGTAGTCCCAGATATTGGGAAGGCTGAGGCGGGAGGATCGCTTGAGCCCGAGAA
 TTTGAGGCTGCAGTGAGCTGTGATCACACCACCTGCACCTCCAGCCTCAGTGACAGAGTGAGGCCCTGICTCA



Quadro 6 (Continuação)

AAAAAGAAAAAGAAAAATAATGAGGGCTGTATGGAATACATTATTCATTCACTCACTCACT
 CACTCATTCACTTCACTTCAACAAGTCTTATTGCATACCTTCIGTTTGCTCAGCTTGGTGGCTTGG
 GGCTGCTGAGGGGAGGGAGAGGGTGACATGGGTGACGCTCGACTCCAGAGTCCACTCCCTGTAG
 56 60 70
 Val Gly Gln Gln Ala Val Glu Val Trp Gln Gly Leu Ala Leu Leu Ser Glu Ala
 GTC GGG CAG CAG GCC GTA GAA GTC TGG CAG GGC CTG GCC CTG CTG TCG GAA GCT
 Val Leu Arg Gly Gln Ala Leu Leu Val Asn Ser Ser Gln Pro Trp Glu Pro Leu
 GTC CTG CGG GGC CAG GCC CAG GGC CTG TTG GTC AAC TCT TCC CAG CCG TGG GAG CCC CTG
 80 90
 *
 Gln Leu His Val Asp Lys Ala Val Ser Gly Leu Arg Ser Leu Thr Thr Leu Leu
 CAG CTG CAT GTG GAT AAA GCC GTC AGT GGC CTT CGC AGC CTC ACC ACT CTG CTT
 100
 110 115
 Arg Ala Leu Gly Ala Gln
 CGG GCT CTG GGA GCC CAG GTGAGTAGGAGCGGACACTTCGTCTTGCCCTTTCTGTAAAGAGGGGA
 GAAGGGTCTTGCTAAGGAGTACAGGAACGTGCGGTATTCCTTCCCTTTCTGTGGCACTGCAGCGACCTCCT
 116 120
 Lys Glu Ala Ile Ser Pro Pro Asp Ala Ala Ser Ala Ala
 AAG GAA GCC ATC TCC CCT CCA GAT GCG GCC TCA GCT GCT
 GTTTCCTCCTTGGCAG



130 Pro Leu Arg Thr Ile Thr Ala Asp Thr Phe Arg Lys Leu Phe Arg Val Tyr Ser
 CCA CTC CGA ACA ATC ACT GCT GAC ACT TTC CGC AAA CTC TTC CGA GTC TAC TCC
 140
 150 Asn Phe Leu Arg Gly Lys Leu Lys Leu Tyr Thr Gly Glu Ala Cys Arg Thr Gly
 AAT TTC CTC CGG GGA AAG CTG AAG CTG TAC ACA GGG GAG GCC TGC AGG ACA GGG
 160
 166 Asp Arg OP
 GAC AGA TGA CCAGGTGTGTCCACCTGGGCATATCCACCACCCTCCCTCACCACCATTTGCTIGGCCACA
 CCTCCCCCGCCACTCCTGAACCCCGTCGAGGGGCTCTCAGCTCAGCGCCAGCCTGTCCCATGGACACTCC
 AGTGCCAGCAATGACATCTCAGGGGCCAGAGGAACTGTCCAGAGAGCAACTCTGAGATCTAAGGATGTCAC
 AGGGCCAACTTGAAGGGCCAGAGCAGGAAGCATTCAGAGAGCAGCTTAAACTCAGGGACAGAGCCCATGC
 TGGGAAGACGCCTGAGCTCACTCGGCACCCCTGC AAAATTTGATGCCAGGACACGCTTTGGAGGCGATTAC
 CTGTTTTTCGCACCTACCATCAGGGACAGGATGACCTGGAGAACTTAGGTGGCAAGCTGTGACTTCTCCAGG
 TCTCACGGGCATGGGCACCTCCCTTGGTGGCAAGAGCCCCCTTGACACCGGGGTGGTGGGAACCATGAAGAC
 AXGATXGGGGCTGGCCTCTGGCTCTCATGGGGTCCCAAGTTTGTGTATTTCTCAACCTATTGACAGACTGAA
 ACACAATATGAC





No quadro 6, a sequência de ADN contínua inicial designa uma cadeia superior de 620 bases em que está aparentemente uma sequência não transferida que precede imediatamente uma porção transferida do gene de EPO humana. Mais especificadamente, a sequência parece conter a extremidade 5' do gene que conduz a uma região de ADN transferida que codifica os primeiros quatro aminoácidos (-27 a -24) de uma sequência condutora ("pré-sequência"). Não foram ainda determinados inequivocamente quatro pares de bases na sequência antes da codificação do início do condutor, que são, por isso, designados por um "X". Segue-se depois um intrão de cerca de 639 pares de bases (439 pares de bases dos quais foram postos em sequência sendo os restantes 200 pares de bases designados "I.S.") que precede imediatamente um códão para a glutamina que foi designado por resíduo -23 do polipeptido transferido. Verifica-se que a sequência exon imediatamente a seguir codifica os restos de aminoácidos através de um resto de alanina (designado por resto +1 da sequência de aminoácidos da EPO humana completa) para o códão que especifica treonina na posição +26, depois da qual se segue um segundo intrão constituído por 256 pares de bases como se designou especificadamente. A seguir a este intrão está uma sequência exon dos restos de aminoácido 27 a 55 e, começa depois, um terceiro intrão constituído por 612 de pares de bases. O exon seguinte codifica os restos 56 a 115 da hematopoietina humana e principia depois um quarto intrão de 134 bases como se especifica. A seguir ao quarto intrão está um exon que codifica os restos N^{os}. 116 a 166 e um códão de "termo" (TGA). Por último, o quadro 6 identifica uma sequência de 568 pares de bases que parece ser uma região 3' não transferida do gene da EPO humana, de que não foram ainda sequenciados inequivocamente dois pares de bases ("X").



O quadro 6 serve assim para identificar a conformação estrutural primária (sequência de aminoácidos) da EPO humana completa como incluindo 166 restos de aminoácidos especificados (P.M. avaliado: 18.399). É também revelada no quadro a sequência de ADN que codifica uma sequência condutora com 27 restos juntamente com as sequências de ADN 5' e 3' que podem ser importantes para as funções de promotor/operador do operão do gene humano. Os sítios para a glucosilação potencial do políptido da hematopoietina (EPO) humana completa são designados, no quadro, por asteriscos. É de notar que a sequência de aminoácidos específica do Quadro 6 constitui provavelmente a de uma forma alélica natural de eritropoietina humana. O fundamento desta posição encontra-se nos resultados dos esforços contínuos efectuados para a sequenciação de isolados urinários de eritropoietina humana que conduziram à descoberta de um número significativo de moléculas de eritropoietina têm uma metionina no resto 126 oposto à serina como se mostra no quadro.

O quadro VII, adiante, ilustra o grau de homologia de sequência polipeptídica entre a eritropoietina humana e de macaco. Na linha contínua superior do quadro, são utilizadas designações de letras simples para representar as sequências polipeptídicas transferidas deduzidas da EPO humana que começam com o resto -27 e a linha contínua inferior mostra a sequência polipeptídica deduzida da EPO de macaco que começa no resto a que se atribuiu o número -27. São utilizados asteriscos para indicar as homologias das sequências. Deverá notar-se que as sequências de EPO humana e de macaco deduzidas revelam um resto lisina (K) "adicional" na posição 116 (humana). A nota remissiva ao quadro 6 indica que este resto está na margem



de uma junção entrançada putativa de mARN na sequência genômica. A presença do resto de lisina na sequência polipeptídica humana foi ainda verificada mediante sequenciação de um clónio de sequência humana cADN preparado a partir do mARN isolado de células COS-1 transformadas com o ADN genómico humano do exemplo 7, mais adiante.

Quadro 7

Comparação dos polipeptidos da EPO humana e de macaco

	-20	-10	+1	10	20	30	40
Humana	MGVHECPAWLWLLSLPLGLPVLGAPPRLICDSRVLERVLLLEAKEAENITTTGCAEHCSLNENITVPDTK						
Macaco	MGVHECPAWLWLLSLVSLPLGLPVPGAPPRLICDSRVLERVLLLEAKEAENVTMGCSLSESCSLNENITVPDTK						
Humana	50	60	70	80	90	100	110
Macaco	VNFYAWKRMEVGQQAQAVEVWQGLALLSEAVLRGQALLVNSSQPWEPLQLHVDKAVSGLRSLTTLLRALGAQKE						
	***	***	***	***	***	***	***
	VNFYAWKRMEVGQQAQAVEVWQGLALLSEAVLRGQAVLANSSQPFEPQLHMDKAISGLRSITTTLLRALGAQ-E						
Humana	120	130	140	150	160		
Macaco	AISPPDAASAAPLRTITADTFRKLFRVYSNFLRGKLLYTGEACRTGDR						
	***	***	***	***	***		
	AISLPDAASAAPLRTITADTFCCLFRVYSNFLRGKLLYTGEACRRGDR						





Exemplo 6

O sistema de expressão escolhido para as tentativas iniciais na síntese microbiana de quantidades isoláveis de material polipeptídico de EPO, codificado pelo cADN de macaco preparado de acordo com as técnicas descritas no exemplo 3, é um sistema que compreende células de hospedeiro humano (isto é, células COS-1, A.T.C.C. Nº. CRL-1650). As células foram modificadas com um vector "shuttle" (vaivém) capaz de se reproduzir (replicar) de maneira autónoma no hospedeiro "E. coli" (em virtude da presença de ADN derivado de pBR322) e nos hospedeiros mamíferos (em virtude da presença de ADN derivado do vírus SV40).

Mais especificadamente, construiu-se um vector de expressão de acordo com as técnicas seguintes. O clónio de plasmídeo 83, preparado no exemplo 3, foi amplificado em "E. coli" e o ADN que codifica EPO de macaco de aproximadamente 1,4 Kb foi isolado mediante digestão com EcoRI e Hind III. Isolou-se separadamente um fragmento Hind III/Sal I de 4,0 Kb a partir de pBR322. Obteve-se um fragmento "ligante" Eco RI/Sal I de aproximadamente 30 pares de bases a partir de ADN de M13mp10 RF (Laboratórios P e L). Este ligante continha, em série, uma extremidade Eco RI colante, seguida de sítios de reconhecimento Sst I, Sma I, Bam HI e Xba I e uma extremidade colante Sal I. Os três fragmentos citados estavam ligados para proporcionarem um plasmídeo intermédio ("pERS") de aproximadamente 5,4 Kb em que o ADN de EPO estava flanqueado de um lado por um "banco" de sítios de reconhecimento utilizáveis de endonuclease de restrição. O plasmídeo pERS foi depois digerido com Hind III e Sal I para se obter o ADN de EPO e o ligante Eco RI a Sal I (M13mp10).



Ligou-se o fragmento de 1,4Kb a um Bam HI/Sal I de pBR 322 de aproximadamente 4,0 Kbe com um outro fragmento ligante RF M13mp 10 Hind III / Bam HI também com aproximadamente 30 pares de bases. O fragmento ligante M 13 era caracterizado por uma extremidade colante Hind III, seguida de sítios de reconhecimento Pst I, Sal I e Xba I e de uma extremidade colante Bam HI. O produto de ligação foi, novamente, um plasmídeo intermédio utilizável ("pBR-EPO") que continha o ADN de EPO flanqueado em ambos os lados por bancos de sítios de restrição.

O vector escolhido para a expressão do ADN de EPO em células COS-1 ("pDSVL1") tinha sido construído anteriormente para permitir a escolha e a reprodução autónoma em "E. coli". Estas características provêm da origem da reprodução e das sequências de ADN do gene de resistência à Ampicilina na região que abrange os nucleotídeos 2448 a 4362 de pBR322. Esta sequência foi modificada estruturalmente pela adição de um ligante que fornece um sítio de reconhecimento Hind III imediatamente adjacente ao nucleotídeo 2448 antes da incorporação no vector. Entre as outras propriedades utilizáveis do vector escolhido estava a capacidade para se reproduzir de maneira autónoma em células COS-1 e na presença de uma sequência de promotor viral funcional em células de mamíferos. Estas características resultam da origem da sequência do ADN de reprodução e da sequência do ADN do promotor viral "gene derradeiro" presente na sequência de 342 pares de bases que abrange os números dos nucleotídeos 5171 a 270 do genoma SV40. Foi fornecido um único sítio de restrição (Bam HI) imediatamente adjacente à sequência do promotor viral através da utilização de uma sequência ligante existente no comércio (Collaborative Research). Incorporou-se tam-



bém no vector uma sequência de 237 pares de bases (derivados como nucleotidos números 2553 a 2770 de SV40) que continham o sinal de poliadenilação de mRNA viral "gene derradeiro" (geralmente designado terminador de transformação). Este fragmento estava colocado no vector na orientação apropriada em relação ao promotor viral "gene derradeiro" pela via do único sítio Bam HI. Estava também presente no vector um outro gene de mamífero em uma localização não essencial para a transformação potencial de um gene inserido no único sítio Bam HI, entre as sequências do promotor viral e do terminador. [O gene de mamífero compreendia um minigene de redutase dihidrofolato de murganho (DHFR) com 2.500 pares de bases isolado a partir do plasmídeo pMG-1 como descrevem Gasser e colab. em "P.N.A.S. (U.S.A.)" 79, 6522-6526 (1982)]. Novamente, os principais componentes activos do plasmídeo pDSVL 1 compreendem os nucleotidos 2448 a 4362 de pBR 322 juntamente com os nucleotidos 5171 a 270 (342 pares de bases) e 2553 a 2770 (237 pares de bases) de ADN de SV40.

Utilizando as técnicas descritas, por ex. em Mariatis e colab., supra, isolou-se o ADN que codifica EPO a partir do plasmídeo pBR-EPO sob a forma de um fragmento de Bam HI e ligou-se ao plasmídeo pDSVL 1 cortado com Bam HI. Utilizou-se a análise do enzima de restrição para confirmar a inserção do gene de EPO na orientação correcta em dois dos vectores clonados resultantes (vectores duplicados H e L). Ver a Figura 2, que ilustra o plasmídeo pDSVL-MKE. Guardou-se os vectores com genes de EPO na orientação errada para serem utilizados como controlos negativos em experiências de transformação destinadas a determinar níveis de expressão de EPO em hospedeiros transforma-



dos com vectores que têm ADN de EPO na orientação correcta.

Juntou-se os vectores H, L, F, X e G com o ADN do veículo (ADN de fígado e de baço de murganho) e utilizou-se estes para transformar placas de 60 mm em duplicado pelos métodos de microprecipitação com fosfato de cálcio. Também foram transformadas placas de 60 mm em duplicado com ADN veículo como controlo negativo de transformação "simulado" ("mock"). Decorridos cinco dias, ensaiou-se todos os meios de cultura para avaliação da presença de polipeptidos com as propriedades imunológicas da EPO natural.

Exemplo 7

A. Sistema de expressão inicial de EPO envolvendo células COS-1

O sistema escolhido para as tentativas iniciais na síntese microbiana de quantidades isoláveis de polipeptidos de EPO humana codificados pelo clónio de ADN de EPO genómico humano também envolveu a expressão em células-hospedeiras de mamíferos (isto é, células COS-1, A.T.C.C. Nº. CRL-1650). O gene de EPO humana foi primeiro sub-clonado em um vector "shuttle" que é capaz de se reproduzir de uma maneira autónoma tanto em hospedeiros "E. coli" (em virtude da presença de ADN derivado de pBR322) como na série de células COS-1 de mamíferos (devido à presença de ADN derivado de vírus SV40). O vector "shuttle", que contém o gene de EPO, foi depois transformado em células COS-1. O material polipeptídico de EPO foi produzido nas células transformadas e segregado no meio de cultura das células.

Mais especificadamente, construiu-se um vector de expressão de acordo com a técnica seguinte. O ADN isolado a partir do clónio lambda hE1, que contém o gene de EPO genómica



humana, foi digerido com as endonucleases de restrição Bam HI e Hind III e isolado um fragmento de ADN de 5,6 Kb que se sabe conter o gene de EPO total. Misturou-se este fragmento e ligou-se ao plasmídeo bacteriano pUC8 (Bethesda Research Laboratories, Inc.) que tinha sido digerido do mesmo modo, criando o plasmídeo intermédio "pUC8-HuE" que constitui uma fonte conveniente deste fragmento de restrição.

O vector escolhido para a expressão do ADN de EPO em células COS-1 (pSV4 SET) tinha sido construído previamente. O plasmídeo pSV4SET continha sequências de ADN que permitiam a selecção e a reprodução autónoma em "E. coli". Estas características resultam da origem da reprodução e das sequências de ADN do gene da resistência à Ampicilina presentes na região que abrange os nucleotidos 2448 a 4362 do plasmídeo bacteriano pBR322. Esta sequência foi modificada estruturalmente mediante adição de um ligante que fornece um sítio de reconhecimento Hind III imediatamente adjacente ao nucleotido 2448. O plasmídeo pSV4SET foi também capaz de se reproduzir autonomamente em células COS-1. Esta característica é fornecida pela origem de reprodução do vírus SV40 que contém um fragmento com 342 pares de bases (nucleotidos números 5171 a 270). Este fragmento tinha sido modificado pela adição de um ligante que fornece um sítio de reconhecimento de Eco RI adjacente ao nucleotido 270 e de um ligante que fornece um sítio de reconhecimento de Sal I adjacente ao nucleotido 5171. Estava também presente um fragmento SV40 de 1061 pares de bases neste vector (nucleotidos números 1711 a 2772 mais um ligante que fornece um sítio de reconhecimento de Sal I junto do nucleotido número 2772). Dentro deste fragmento estava uma única sequência de reconhecimento de Bam HI.



Resumindo, o plasmídeo pSV4 SET continha sítios de reconhecimento únicos de Bam HI e Hind III que permitiam a inserção do gene de EPO humana, sequências que permitiam a reprodução e a selecção em "E. coli" e sequências que permitiam a reprodução em células COS-1.

Com o fim de inserir o gene de EPO em pSV4SET, digeriu-se o plasmídeo pUC8-HuE com as endonucleases de restrição Bam H I e Hind III e isolou-se o fragmento de ADN que codifica EPO de 5,6 Kb. Também se digeriu o pSV4SET com Bam H I e Hind III e isolou-se o fragmento principal de 2513 pares de bases (conservando todas as funções necessárias). Estes fragmentos foram misturados e ligados, criando o vector final "pSVgHuEPO" (Ver, Figura 3.) Este vector foi propagado em "E. coli" e isolou-se o vector de ADN. Utilizou-se a análise do enzima de restrição para confirmar a inserção do gene de EPO.

Utilizou-se ADN do plasmídeo pSVgHuEPO para exprimir o material polipeptídico de EPO em células COS-1. Mais especificadamente, o ADN de pSVgHuEPO foi associado com o ADN do veículo e transformado em placas de 60 mm, em triplicado, de células COS-1. Como controlo, foi também transformado ADN do veículo sózinho em células COS-1. Foram colhidas amostras dos meios de cultura das células, sete dias mais tarde, que foram ensaiadas para avaliação da presença de polipeptidos que possuam as propriedades imunológicas da EPO humana natural.

B. Segundo sistema de expressão de EPO envolvendo células COS-1

Foi concebido ainda um outro sistema para proporcionar a produção melhorada de material polipeptídico de EPO humana codificado pelo clónio ADN de EPO genómico humano em células



COS-1 (A.T.C.C. Nº. CRI-1650).

No sistema imediatamente precedente, a EPO foi expressa em células COS-1, utilizando o seu próprio promotor que está dentro do fragmento de restrição de Bam HI a Hind III de 5,6 Kb. Na construção seguinte, o gene de EPO foi alterado de maneira que fosse expresso utilizando o último promotor SV 40.

Mais pormenorizadamente, o fragmento genómico de restrição de EPO humana de Bam HI a Hind III de 5,6 Kb, clonado, foi modificado de acordo com as técnicas seguintes. O plasmídeo pUC8-HuE, como se descreveu antes, foi cindido com as endonucleases de restrição Bam HI e BstE II. BstE II efectua a cisão dentro do gene de EPO de 5,6 Kb em uma posição que é a a posição 5' de aproximadamente 44 pares de bases em relação a ATG iniciador que codifica o pré-peptido e a posição 3' de aproximadamente 680 pares de bases em relação ao sítio de restrição Hind III. Isolou-se o fragmento de aproximadamente 4 900 pares de bases. Foi sintetizado e purificado um fragmento de ADN ligante sintético, que continha extremidades colantes Sal I e BstE II e um sítio de reconhecimento Bam HI interno. Os dois fragmentos foram misturados e ligados com o plasmídeo pBR322 que tinha sido cortado com Sal I e BAM HI para produzir o plasmídeo intermédio pBRgHE. O gene genómico de EPO humana pode ser isolado a partir daquele sob a forma de um fragmento da digestão de Bam HI com 4 900 pares de bases que comporta o gene estrutural completo com um único ATG com 44 pares de bases em 3' em relação ao sítio de Bam HI adjacente à região de codificação do grupo amino terminal.

Este fragmento foi isolado e inserido com um fragmento de Bam HI no plasmídeo pDSVL1, vector de expressão cindido



com Bam HI (descrito no exemplo 6). O plasmídeo resultante, pSVLgHuEPO, como se ilustra na Figura 4, foi utilizado para exprimir material polipeptídico de EPO a partir de células COS-1, como se descreveu nos exemplos 6 e 7 A.

Exemplo 8

Analisou-se meios de cultura provenientes do desenvolvimento das seis culturas de células COS-1 transformadas do exemplo 6 mediante ensaio rádio-imunológico, de acordo com as técnicas indicadas no exemplo 2, Parte B. Ensaiou-se cada amostra em níveis de pequenas porções de 250, 125, 50 e 25 microlitros. Os líquidos sobrenadantes de células simuladamente transformadas ou transformadas com vectores que têm uma orientação incorrecta do gene de EPO foram inequivocamente negativos em relação à imuno-reactividade de EPO. Para cada amostra dos dois líquidos sobrenadantes provenientes da cultura de células COS-1 transformadas com vectores (H e L) tendo o ADN de EPO na orientação correcta, a % de inibição da ligação de ^{125}I -EPO ao anticorpo estava compreendida entre 72 e 88%, o que coloca todos os valores na parte superior da curva-padrão. Não foi possível avaliar com segurança a concentração exacta de EPO na cultura sobrenadante. No entanto, fez-se uma estimativa bastante cautelosa de 300 mU/ml a partir do cálculo do valor da porção alíquota maior (250 microlitros).

Ensaiou-se no RIA (ensaio radio-imunológico) um líquido de cultura representativo de acordo com o exemplo 6 e líquidos de cultura de 5 e 7 dias obtidos de acordo com as técnicas descritas no exemplo 7 A com o fim de comparar a actividade de materiais recombinantes de EPO de macaco e humana em relação a



um padrão de EPO humana natural e os resultados estão indicados em forma gráfica na Figura 1. Resumindo, os resultados revelaram, como se esperava, que a EPO de macaco recombinante competia significativamente com o anticorpo de EPO anti-humano embora não fosse capaz de inibir completamente a ligação nas condições do ensaio. Porém, os valores máximos da percentagem de inibição para a EPO humana recombinante, aproximam-se bastante dos do padrão de EPO humana. A natureza paralela das curvas de dose/resposta sugere a identidade imunológica das sequências (epitopes) em comum. Foram novamente avaliadas, nestes níveis de diluição mais elevados as estimativas anteriores de EPO de macaco em fluidos de cultura e verificou-se que estavam compreendidas entre 2,91 e 3,12 U/ml. Os níveis de produção de EPO humana avaliada foram correspondentemente estabelecidos a 392 mU/ml para a amostra de cultura de 5 dias e 567 mU/ml para a amostra de cultura de 7 dias. Os níveis de produção de EPO de macaco avaliados no sistema de expressão do exemplo 7B foram da mesma ordem ou melhores ainda.

Exemplo 9

Submeteu-se fluídos de cultura, preparados de acordo com as técnicas descritas nos exemplos 6 e 7, a um ensaio "in vitro" para avaliação da actividade de EPO, utilizando o método descrito por Goldwasser e colab., "Endocrinology", 97, 2, 315-323 (1975). Os valores de EPO de macaco avaliados para líquidos de cultura estavam compreendidos entre 3,2 e 4,3 U/ml. Os fluídos de cultura de EPO humana foram também activos neste ensaio "in vitro" e, além disso, esta actividade pôde ser neutralizada por anticorpo anti-EPO. Os líquidos de cultura de EPO de macaco recombinante de acordo com o exemplo 6 foram tam-



bém submetidos a um ensaio para avaliação da actividade biológica "in vivo" de acordo com as técnicas gerais de Cotes e colab., "Nature", 191, 1065-1067 (1961) e Hammond e Colab., "Ann. N.Y. Acad. Sci.", 149, 516-527 (1968) e os níveis de actividade estavam compreendidos entre 0,94 e 1,24 U/ml.

Exemplo 10

Nos exemplos anteriores, produziu-se material de EPO humana ou de macaco recombinante a partir de vectores utilizados para transformar células COS-1. Estes vectores reproduzem-se em células COS-1 devido à presença de antigénio SV40 T dentro da célula e a uma origem de reprodução de SV40 sobre os vectores. Embora estes vectores produzam quantidades utilizáveis de EPO em células COS-1, a expressão é só transitória (7 a 14 dias) devido à perda eventual do vector. Adicionalmente, apenas uma pequena percentagem de COS-1 foi transformada produtivamente com os vectores. O presente exemplo descreve sistemas de expressão que utilizam células (CHO) DHFR⁻ de ovário de criceto chinês e o marcador seleccionável DHFR. Para a discussão de sistemas de expressão afins, ver Patente de invenção norte-americana Nº. 4.399.216 e patentes de invenção europeias Nºs. 117058, 117059 e 117060, todas elas publicadas em 29 de Agosto de 1984 7.

As células CHO DHFR⁻ (Dux-B11), células CHO K1, Urlaub e colab., "Proc. Nat. Acad. Sci. (U.S.A.)", Vol. 77, 4461 (1980) não possuem o enzima dihidrofolato-redutase (DHFR) devido a mutações nos genes estruturais e, por isso, exigem a presença de glicina, hipoxantina e timidina nos meios de cultura. Os plasmídios pDSVI-MKE (Exemplo 6) ou pDSVI-gHuEPO (Exemplo 7 B) foram transformados juntamente com o ADN do veículo em células



CHO DHFR⁻ que se desenvolvem em meios que contêm hipoxantina, timidina e glicina em placas de cultura de 60 mm. Misturou-se o plasmídeo pSVgHuEPO (Exemplo 7 A) com o plasmídeo pMG 2 que contém um gene de dihidrofolato-redutase de murganho clonado no vector de plasmídeo bacteriano pBR322 (por Gasser, e colab., supra). A mistura de plasmídeo e ADN do veículo foi transformada em células CHO DHFR⁻. (As células que adquirem um plasmídeo poderão também, de um modo geral, adquirir um segundo plasmídeo). Três dias depois, as células foram dispersas mediante tripsinização em diversas placas de cultura de 100 mm em meios que não têm hipoxantina nem timidina. Apenas as células que tenham sido transformadas de maneira estável com gene DHFR, e por isso o gene de EPO, sobrevivem neste meio de cultura. Decorridos 7-21 dias, as colónias de células sobreviventes tornam-se visíveis. Estas colónias transformantes, após dispersão mediante tripsinização, podem ser propagadas continuamente em meios que carecem de hipoxantina e timidina, criando novas estirpes de células (por ex. CHO pDSVL-MKEPO, CHO pSVgHuEPO, CHO-pDSVL-gHuEPO).

Os líquidos de cultura das estirpes de células mencionadas antes foram ensaiados no ensaio radio-imunológico (RIA) para avaliação da presença de EPO de macaco ou humana recombinante. Os meios de cultura da estirpe CHO pDSVL-MKEPO continham EPO com propriedades imunológicas idênticas às obtidas a partir de células COS-1 transformadas com plasmídeo pDSVL-MKEPO. Um líquido de cultura de 65 horas representativo continha EPO de macaco a 0,60 U/ml.

Os líquidos de cultura a partir de CHO pSVgHuEPO e CHO pDSVL-gHuEPO continham EPO humana recombinante com proprie



dades imunológicas idênticas às obtidas com células COS-1 transformadas com plasmídeo pSVgHuEPO ou pDSVL-gHuEPO. Um líquido de cultura de 3 dias representativo de CHO pSVgHuEPO continha 2,99 U/ml de EPO humana e uma amostra de 5 dias e meio de CHO pDSVL-gHuEPO tinha 18,2 U/ml de EPO humana como se determinou mediante ensaio rádio-imunológico (RIA).

A quantidade de EPO produzida pelas estirpes de células descritas antes pode ser aumentada mediante amplificação do gene obtendo-se novas estirpes de células de grande produtividade. O enzima dihidrofolato-redutase (DHFR), que é o produto codificado pelo gene DHFR, pode ser inibido pela droga metotrexato (MTX). Mais pormenorizadamente, as células propagadas em meios de cultura que não têm hipoxantina nem timidina são inibidas ou mortas por MTX. Sob condições apropriadas (por ex. concentrações mínimas de MTX) podem ser obtidas células resistentes a e capazes de se desenvolver em MTX. Verificou-se que estas células são resistentes a MTX devido a um aumento do número dos seus genes DHFR, o que tem, como resultado, uma produção aumentada do enzima DHFR. As células sobreviventes podem, por sua vez, ser tratadas com concentrações crescentes de MTX, de que resultam estirpes de células que contêm um número maior de genes DHFR. Verifica-se também que os "genes passageiros" (por ex. EPO) transportados sobre o vector de expressão juntamente com o gene DHFR ou transformados com o gene DHFR também sofrem frequentemente um aumento no seu número de cópias de gene.

Como exemplos de realização deste sistema de amplificação, submeteu-se estirpes de células CHO pDSVL-MkE a concentrações crescentes de MTX (0 nM, 30nM e 100nM). Ensaaiou-se



amostras de meios de cultura de 65 horas representativas de cada fase de amplificação mediante ensaio rádio-imunológico (RIA) e determinou-se que continha 0, 60, 2,45 e 6,10 U/ml, respectivamente. A estirpe de células CHO pDSVL-gHuEPO foi submetida a uma série de concentrações crescentes de MTX de 30nM, 50nM, 100nM, 200nM, 1 μ M e 5 μ M de MTX. Uma amostra representativa do meio de cultura de 3 dias, obtida a partir da fase de 100nM de MTX, continha EPO humana em uma quantidade de 3089 ± 129 u/ml como se avaliou mediante ensaio rádio-imunológico. Amostras representativas de meio de cultura de 48 horas, provenientes das fases de 100nM e 1 μ M de MTX, continham, respectivamente, EPO humana em uma quantidade de 466 e 1352 U/ml como se avaliou mediante ensaio rádio-imunológico (média de ensaios em triplicado). Nestas técnicas, colocou-se uma camada de 1×10^6 células em 5 ml de meio de cultura em caixas de cultura de 60 mm. Vinte e quatro horas depois, os meios foram removidos e substituídos por 5 ml de meio de cultura isento de soro (DMEM com elevado teor de glucose suplementado com 0,1mM de aminoácidos não essenciais e L-glutamina. Deixou-se acumular EPO, durante 48 horas, nos meios isentos de soro. Os meios de cultura foram recolhidos para ensaio rádio-imunológico (RIA) e as células foram tripsinizadas e contadas. Os valores médios de RIA de 467 U/ml e 1352 U/ml para as células cultivadas com 100 nM e 1 μ M de MTX, respectivamente, forneceram rendimentos reais de 2335 U/placa e 6750 U/placa. Os números de células médios por placa foram $1,94 \times 10^6$ e $3,12 \times 10^6$ células, respectivamente. As taxas de produção efectiva para estas condições de cultura foram, assim, 1264 e 2167 U/ 10^6 células /48 horas.

As células nas culturas acabadas de descrever são



uma população geneticamente heterogén^{ea}. Estão a ser utilizadas técnicas normais de apreciação (selecção) em uma tentativa para isolar clónios geneticamente homogéneos com a mais elevada capacidade de produção. Ver, Secção A, Parte 2, de "Points to Consider in the Characterization of Cell Lines Used to Produce Biologics" 1 de Junho de 1984, Office of Biologics, Research Review, Center for Drugs and Biologics, U.S. Food and Drug Administration.

A produtividade das séries de células CHO produtoras de EPO descritas antes pode ser melhorada mediante técnicas apropriadas de cultura de células . A propagação de células de mamíferos em cultura requer geralmente a presença de soro nos meios de cultura. Um método para a produção de eritropoietina a partir de células CHO em meios que não contenham soro facilita consideravelmente a purificação da eritropoietina a partir do meio de cultura. O método descrito a seguir é capaz de produzir economicamente eritropoietina em meios de cultura isentos de soro em quantidades grandes suficientes para a produção.

São utilizadas células da estirpe CHO pDSVL-gHu EPO, cultivadas em condições normais de cultura de células, para semear balões rotativos de cultura de células. As células são propagadas, sob a forma de uma suspensão de uma série de células, no balão rotativo de cultura de células em meios de cultura constituídos por uma mistura a 50:50 de DMEM com elevado teor de glicose e de F12 de Ham suplementado com 5% de soro de feto de vitela, L-glutamina, Penicilina e Estreptomicina, 0,05 mM de aminoácidos não essenciais e a concentração apropriada de metotrexato. A cultura de células em suspensão permite que as células



CHO produtoras de EPO se expandam facilmente até volumes grandes. Utiliza-se células CHO, cultivadas em suspensão, para semear frascos rotativos com uma densidade inicial de sementeira de $1,5 \times 10^7$ células viáveis por frasco rotativo de 850 cm^2 em 200 ml de meio de cultura. Deixa-se desenvolver as células, até confluência como uma linha de células aderentes, durante um intervalo de tempo de 3 dias. O meio de cultura utilizado nesta fase do desenvolvimento é o mesmo utilizado para o desenvolvimento em suspensão. No fim do período de desenvolvimento de 3 dias, o meio que contém o soro é removido e substituído por 100 ml de meio isento de soro; mistura a 50:50 de DMEM com elevado teor de glicose e meio F 12 de Ham suplementado com 0,05 mM de aminoácidos não essenciais e glutamina L. Os frascos rotativos voltaram para o incubador de frascos rotativos durante 1 a 3 horas e o meio de cultura é novamente removido e substituído por 100 ml de meio novo isento de soro. A incubação de 1 - 3 horas do meio isento de soro reduz a concentração de proteínas de soro contaminantes. Os frascos rotativos voltaram para o incubador por mais 7 dias durante os quais se acumula eritropoietina no meio de cultura isento de soro. No fim da fase de produção de sete dias, o meio condicionado é removido e substituído por meio novo isento de soro para um segundo ciclo de produção. Como exemplo da realização prática deste sistema de produção, uma amostra representativa do meio isento de soro de 7 dias continha eritropoietina humana em uma quantidade de $3892 \pm 409 \text{ U/ml}$, como se avaliou mediante ensaio rádio-imunológico. Com base na densidade de células avaliada de 0,9 a $1,8 \times 10^5$ células/ cm^2 , cada frasco rotativo de 850 cm^2 continha desde 0,75 a $1,5 \times 10^8$ célu-



las e, deste modo, a taxa de produção de EPO, no dia 7, em 100 ml de cultura era de 750 a 1470 U/10⁶ células/48 horas.

Submeteu-se líquidos de cultura da estirpe de células CHO pDSVI-MkEPO, efectuada em 10nM de MTX, a ensaios da actividade de EPO "in vitro" e "in vivo" mediante ensaio rádio-imunológico (RIA). A amostra do meio condicionado continha 41,2 ± 1,4 U/ml de MkEPO avaliado pelo ensaio rádio-imunológico, 41,2 ± 0,064 U/ml avaliado pelo ensaio da acitvidade biológica "in vitro" e 42,5 ± 5 U/ml avaliado pelo ensaio da actividade biológica "in vivo". A sequenciação de aminoácidos de produtos polipeptídicos revelou a presença de produtos de EPO, uma espécie de origem que tem 3 restos da sequência "condutora" adjacente ao grupo amino terminal putativo da alamina. Desconhece-se presentemente se isto é o resultado do tratamento incorrecto da membrana do polipeptido em células CHO ou se reflecte uma diferença na estrutura do grupo amino terminal da EPO de macaco em relação à EPO humana.

Os líquidos de cultura da estirpe celular CHO pDSVI-gHuEPO foram submetidos a três ensaios. Uma amostra de cinco dias e meio continha EPO recombinante no meio de cultura em um nível de 18,2 U/ml mediante ensaio radio-imunológico, 15,8 ± 4,6 U/ml mediante ensaio "in vitro" e 16,8 ± 3,0 U/ml mediante ensaio "in vivo".

Os líquidos de cultura de células CHO pSDVI-gHu EPO preparadas e amplificadas gradualmente por 100 nM de MTX foram submetidos aos três ensaios. Uma amostra de 3,0 dias continha EPO humana recombinante em um nível de 3089 ± 129 U/ml mediante ensaio rádio-imunológico, 2589 ± 71,5 U/ml pelo ensaio "in vitro" e 2040 ± 160 U/ml pelo ensaio "in vivo". A sequenciação de aminoácidos deste produto revela a presença de um grupo amino ter-



minal que corresponde ao indicado no quadro 6.

Reuniu-se os meios de cultura condicionados de células CHO transformadas com plasmídeo pDSVL-MkE em 10nM de MTX e dialisou-se o MTX durante vários dias, obtendo-se meios de cultura com uma actividade de EPO de $221 \pm 5,1$ U/ml (EPO-CCM). Para determinar o efeito "in vivo" da EPO-CCM sobre níveis de hematocrito em murganhos Balb/C normais, efectuou-se a seguinte experiencia. Ajustou-se com PBS, meios de cultura condicionados de células de células CHO (CCM) e EPO-CCM não transformadas. Utilizou-se CCM no grupo de controlo (3 murganhos) e dois níveis de dose de EPO-CCM - 4 unidades por injeção e 44 unidades por injeção - nos grupos experimentais (2 murganhos/grupo). No decurso de 5 semanas, os sete murganhos foram injectados por via intraperitoneal, 3 vezes por semana. Depois da oitava injeção, determinou-se que os valores médios de hematocrito para o grupo de controlo eram 50,4%; para o grupo 4 U, 55,1%; e, para o grupo 44 U, 67,9%.

Os produtos de expressão de células de mamíferos podem ser isolados facilmente em forma substancialmente pura a partir de meios de cultura, utilizando a cromatografia líquida a pressão elevada (C_4) e um gradiente de etanol, de preferência a pH 7 .

Foi feita uma tentativa preliminar para caracterizar produtos glucoproteicos recombinantes a partir do meio de cultura condicionado de expressão de células COS-1 e CHO do gene de EPO humana em comparação com isolados de EPO urinária humana utilizando a análise de "Western blot" e a de SDS-PAGE. Estes estudos indicaram que o material de EPO produzido por CHO tinha um P.M. um tanto mais elevado que o produto de expressão COS-1, o qual, por sua vez, era ligeiramente maior que os ex-



tractos urinários de origem humana reunidos. Todos os produtos eram um tanto ou quanto heterogêneos. Do tratamento com o enzima neuraminidase para eliminar o ácido siálico, resultaram produtos recombinantes COS-1 e CHO com P.M. aproximadamente iguais, que foram no entanto maiores que o do extracto urinário humano asialo resultante. Do tratamento do produto de CHO recombinante e do produto do extracto urinário com o enzima endoglucosidade F (EC 3, 2, 1) (para remover totalmente os hidratos de carbono de ambos), resultaram produtos substancialmente homogêneos que têm características de P.M. essencialmente idênticas.

A EPO urinária humana purificada e uma EPO recombinante, produzida por células CHO, de acordo com a presente invenção, foram submetidas a análise de hidratos de carbono de acordo com a técnica de Ledeen e colab., "Methods in Enzimology", 83 (Parte D), 139-191 (1982) modificada através da utilização das técnicas de hidrólise de Nesser e colab., "Anal. Biochem"., 142, 58-67 (1984). Os valores da constituição de hidratos de carbono determinada experimentalmente (expressa em relações molares de hidratos de carbono no produto) para o isolado urinário foram os seguintes: Hexoses, 1,73; N-acetilglucosamina, 1; ácido N-acetilneuramínico, 0,93; Fucose, 0; e N-acetilgalactosamina, 0. Valores correspondentes para o produto recombinante (derivado de um meio de cultura CHO pDSVL - gHuEPO com três dias a 100 nM MTX) foram os seguintes: Hexoses, 15,09; N-acetilglicosamida, 1; ácido acetilneuramínico, 0,998; Fucose 0; e N-acetilgalactosamina, 0. Estas conclusões estão de acordo com a análise de "Western blot" e de SDS-PAGE descritas antes.

Os produtos glicoproteicos de acordo com a presente invenção são, deste modo, esclarecedores de produtos que têm



uma conformação estrutural primária suficientemente dupla da da eritropoietina natural para permitir que possuam uma ou mais das suas propriedades biológicas e que tenham uma composição média de hidratos de carbono diferente da da eritropoietina natural.

Exemplo 11

O presente exemplo refere-se à preparação total mediante reunião de bases nucleotídicas de dois genes estruturais que codificam a sequência da eritropoietina da espécie humana do quadro 6 e incorporam, respectivamente, códãos "preferidos" para expressão em células de "E. coli" e de levedura ("S. cerevisiae"). Descreve-se também a construção de genes que codificam análogos de EPO humana. De um modo abreviado, o protocolo utilizado foi, na generalidade, o indicado na descrição anteriormente citada de Alton e colab., (WO 83/04053). Os genes foram projectados para reunião inicial de componentes oligonucleotídicos em duplexes múltiplos os quais, por sua vez, foram reunidos em três secções separadas. Estas secções foram concebidas para amplificação fácil e, após eliminação do sistema de amplificação, puderam ser reunidas sucessivamente ou através de uma ligação múltipla de fragmento em um vector de expressão apropriado.

Os quadros 8 a 14 a seguir ilustram a concepção e a reunião de um gene preparado que codifica um produto de transferência de eritropoietina humana a que falta qualquer condutor ou pré-sequência mas que inclui um resto metionina inicial na posição 1. Além disso, o gene incorporado em parte substancial dos códãos de preferência de "E. coli" e da construção, foi por isso designado gene "ECEPO".



Quadro 8

OLIGONUCLEOTIDOS DA SECÇÃO 1 DE ECEPO

1. AATTCTAGAAACCATGAGGGTAATAAAATA
2. CCATTATTTTATTACCCATCGGTTCTAG
3. ATGGCTCCGCCGCGTCTGATCTGCGAC
4. CTCGAGTCGCAGATCAGACGCGGCGGAG
5. TCGAGAGTTCTGGAACGTTACCTGCTG
6. CTTCCAGCAGGTAACGTTCCAGAACT
7. GAAGCTAAAGAAGCTGAAAACATC
8. GTGGTGATGTTTTTCAGCTTCTTTAG
9. ACCACTGGTTGTGCTGAACACTGTTC
10. CAAAGAACAGTGTTTCAGCACAACCA
11. TTTGAACGAAAACATTACGGTACCG
12. GATCCGGTACCGTAATGTTTTCGTT



Quadro 9

Secção 1 de ECEPO

EcoRI XbaI
 AATTCTAG AAACCATGAG¹ GGTAATAAAA TAATGGCTCC³ GCCGCGTCTG
 GATC TTTGGTACTC² CCATTATTTT ATTACGAGG⁴ CGGCGCAGAC

 ATCTGCGACT⁵ CGAGAGTTCT GGAACGTTAC CTGCTGGAAG⁶ CTAAGAAGC
 TAGACGCTGA GCTCTCAAGA CCTTGCAATG GACGACCTTG⁷ GATTTCTTCG

 TGAAAACATC⁸ ACCACTGGTT⁹ GTGCTGAACA CTGTTCTTTG¹⁰ AACGAAAACA¹¹
 ACTTTTGTAG TGGTGACCAA CACGACTTGT GACAAGAAAC¹² TTGCTTTTGT

KpnI BamHI
 TTACGGTACC G
 AATGCCATGG CCTAG
12

Quadro 10Oligonucleotidos da secção 2 de ECEPO

1. AATTCGGTACCAGACACCAAGGT
2. GTTAACCTTGGTGTCTGGTACCG
3. TAACTTCTACGCTTGGAACGTAT
4. TTCCATACGTTTCCAAGCGTAGAA
5. GGAAGTTGGTCAACAAGCAGTTGAAGT
6. CCAAACCTTCAACTGCTTGTTGACCAAC
7. TTGGCAGGGTCTGGCACTGCTGAGCG
8. GCCTCGCTCAGCAGTGCCAGACCCTG
9. AGGCTGTACTGCGTGGCCAGGCA
10. GCAGTGCCTGGCCACGCAGTACA
11. CTGCTGGTAAACTCCTCTCAGCCGT
12. TTCCCACGGCTGAGAGGAGTTTACCA
13. GGGAAACCGCTGCAGCTGCATGTTGAC
14. GCTTTGTCAACATGCAGCTGCAGCGG
15. AAAGCAGTATCTGGCCTGAGATCTG
16. GATCCAGATCTCAGGCCAGATACT

Quadro 11

Secção 2 de ECEPO

EcoRI KpnI ¹
 A ATTGGGTACC AGACACCAAG GTTAACTTCT ACGCTTGGAA ACGTATGGAA
 GCCATGG TCTGTGGTTC CAATTGAAGA TCGGAACCTT TGCATACCTT ³
²
 GTTGGTCAAC AAGCAGTTGA AGTTTGGCAG GGCTTGGCAC TGCTGAGCGA ⁷
 CAACCAGTTG TTCGTCAACT TCAAACCGTC CCAGACCCGTC ACGACTCGCT ⁸
⁶
 GGCTGTACTG CGTGGCCAGG CACTGCTGGT AAACCTCCTCT CAGCCGTGGG ¹¹
 CCGACATGAC GCACCCGGTCC GTGACGACCA TTTGAGGAGA GTCGGCACCC ¹²
¹⁰
 AACCGCTGCA GCTGCATGTT GACAAAGCAG TATCTGGCCT GAGATCTG ¹⁵ BglII BamHI
 TTGGCGACGT CGACGTACAA CTGTTTCGTC ATAGACCCGA CTCTAGACCTAC ¹⁶
¹⁴



Quadro 12

Secção 3 de ECEPO

1. GATCCAGATCTCTGACTACTCTGC
2. ACGCAGCAGAGTAGTCAGAGATCTG
3. TGCGTGCTCTGGGTGCACAGAAAGAGG
4. GATAGCCTCTTTCTGTGCACCCAGAGC
5. CTATCTCTCCGCCGGATGCTGCATCT
6. CAGCAGATGCAGCATCCGGCGGAGA
7. GCTGCACCGCTGCGTACCATCACTG
8. ATCAGCAGTGATGGTACGCAGCGGTG
9. CTGATACCTTCCGCAAACGTGTTTCG
10. ATACACGAAACAGTTTGCGGAAGGT
11. TGTATACTCTAACTTCCTGCGTGGTA
12. CAGTTTACCACGCAGGAAGTTAGAGT
13. AACTGAAACTGTATACTGGCGAAGC
14. GGCATGCTTCGCCAGTATACAGTTT
15. ATGCCGTA CTGGTGACCGCTAATAG
16. TCGACTATTAGCGGTCACCAGTAC



Quadro 1.3

Secção 3 de ECEPO

BamHI BglII
GA TCCAGATCTCTG
GTCTAGAGAC

ACTACTCTGC TGC GTGCTCT GGGTGCACAG AAAGAGGCTA TCTCTCCGCC
TGATGAGACG ACGC ACGAGA CCCACGTGTC TTTCTCCGAT ACAGAGGCGG

GGATGCTGCA TCTGCTGCAC CGCTGCGTAC CATCACTGCT GATACCTTCC
CCTACGACGT AGACGACGTG GCGACGCATG GTAGTGACGA CTATGGAAGG

GCAAAC TGT TCGTGTATAC TCTAACTTCC TGC GTGGTA ACTGAACTG
 CGTTTGACAA AGCACAATG AGATTGAAGG ACGCACCATT TGAC TTTGAC
 10 11 12 13

TATACTGGCG AAGCATGCCG TACTGGTGAC CGCTAATAG SalI
ATATGACCGC TTCGTACGCG ATGACCACTG GCGATTATC AGCT
14 15 16



Quadro 14

Gene ECEPO

	<u>XbaI</u>			-1	1		<u>MetAla</u>	
	CTAG	AAACCATGAG	GGTAATAAAA	TAATGGCTCC	GCCGCGTCTG			
		TTTGGTACTC	CCATTATTTT	ATTACCGAGG	CGGCGCAGAC			
ATCTGCGACT	CGAGAGTTCT	GGAACGTTAC	CTGCTGGAAG	CTAAAGAAGC				
TAGACGCTGA	GCTCTCAAGA	CCTTGCAATG	GACGACCTTC	GATTTCTTCG				
TGAAAACATC	ACCACTGGTT	GTGCTGAACA	CTGTTCTTTG	AACGAAAACA				
ACTTTTGTAG	TGGTGACCAA	CACGACTTGT	GACAAGAAAC	TTGCTTTTGT				
TTACGGTACC	AGACACCAAG	GTAACTTCT	ACGCTTGGAA	ACGTATGGAA				
AATGCCATGG	TCTGTGGTTC	CAATTGAAGA	TGCGAACCTT	TGCATACCTT				
GTTGGTCAAC	AAGCAGTTGA	AGTTTGGCAG	GGTCTGGCAC	TGCTGAGCGA				
CAACCAGTTG	TTCGTCAACT	TCAAACCGTC	CCAGACCGTG	ACGACTCGCT				
GGCTGTACTG	CGTGGCCAGG	CACTGCTGGT	AAACTCCTCT	CAGCCGTGGG				
CCGACATGAC	GCACCGGTCC	GTGACGACCA	TTTGAGGAGA	GTGGGCACCC				
AACCGCTGCA	GCTGCATGTT	GACAAAGCAG	TATCTGGCCT	GAGATCTCTG				
TTGGCGACGT	CGACGTACAA	CTGTTTCGTC	ATAGACCGGA	CTCTAGAGAC				
ACTACTCTGC	TGCGTGCTCT	GGGTGCACAG	AAAGAGGCTA	TCTCTCCGCC				
TGATGAGACG	ACGCACGAGA	CCCACGTGTC	TTTCTCCGAT	AGAGAGGCGG				
GGATGCTGCA	TCTGCTGCAC	CGCTGCGTAC	CATCACTGCT	GATACCTTCC				
CCTACGACGT	AGACGACGTG	GCGACGCATG	GTAGTGACGA	CTATGGAAGG				
GCAAACTGTT	TCGTGTATAC	TCTAACTTCC	TGCGTGGTAA	ACTGAAACTG				
CGTTTGACAA	AGCACATATG	AGATTGAAGG	ACGCACCATT	TGACTTTGAC				
				<u>SalI</u>				
TATACTGGCG	AAGCATGCCG	TACTGGTGAC	CGCTAATAG					
ATATGACCGC	TTCGTACGGC	ATGACCACTG	GCGATTATCA	GCT				



Mais pormenorizadamente, o quadro 8 ilustra oligonucleotidos utilizados para formar a Secção 1 do gene ECEPO que codifica os restos com grupo amino terminal do polipeptido da espécie humana. Os oligonucleotidos foram reunidos em duplexes (1 e 2, 3 e 4, etc.) e os duplexes foram depois ligados para proporcionarem a Secção 1 de ECEPO como se indica no Quadro 9. É de notar que a secção reunida contém as extremidades colantes EcoRI e BamHI terminais respectivas, sendo a extremidade colante EcoRI "a jusante" um sítio de reconhecimento do enzima de restrição Xba I e a extremidade colante Bam HI "a montante" um sítio de reconhecimento do enzima Kpn I. Foi possível amplificar facilmente a secção 1, utilizando o vector de fago M13 utilizado para a verificação da sequência da secção. Foram encontradas algumas dificuldades no isolamento da secção sob a forma de um fragmento Xba I / Kpn I de ADN de RF gerado em "E. coli", provavelmente devido à metilação das bases do sítio de reconhecimento de Kpn I dentro do hospedeiro. O ADN de fago de cadeia simples foi, por isso, isolado e levado à forma de cadeia dupla "in vitro" mediante extensão inicial e o fragmento de cadeia dupla desejado foi depois isolado facilmente.

As Secções 2 e 3 do gene ECEPO (Quadros 11 e 13) foram construídas por uma maneira análoga a partir dos oligonucleotidos dos quadros 10 e 12, respectivamente. Amplificou-se cada Secção no vector M13 utilizado para verificação da sequência e isolou-se aquela a partir do ADN de fago. Como se torna evidente a partir do Quadro 11, a Secção 2 de ECEPO foi construída com extremidades colantes Eco RI e Bam HI e pôde ser isolada sob a forma de um fragmento Kpn I / Bgl II. Analogamente, a Secção 3 de ECEPO foi preparada com extremidades colantes Bam HI e Sal I e



pôde ser isolada a partir do ADN de RF de fago sob a forma de um fragmento Bgl II / Sal I. As três secções assim preparadas podem ser reunidas facilmente em uma sequência contínua de ADN (Quadro 14) que codifica o polipeptido de eritropoietina da espécie humana completo com um códon de metionina com grupo amino terminal (ATG) para iniciação da transferência de "E. coli". É também de notar que "a montante" do ATG inicial está uma série de pares de bases que duplicam substancialmente a sequência do sítio de ligação do ribosoma do gene OMP-f extremamente expresso de "E. coli".

Pode utilizar-se qualquer vector de expressão apropriado para transportar o ECEPO. O vector especialmente escolhido para expressão do gene de ECEPO é o plasmídeo pCFM536 "sensível à temperatura" - um derivado de plasmídeo pCFM414 (A.T.C.C. 40076)- descrito no pedido de patente de invenção norte-americana Nº. de Série 636.727, depositado em 6 de Agosto de 1984 por Charles F. Morris. Mais especificadamente, o pCFM536 foi digerido com Xba I e Hind III; o fragmento grande foi isolado e utilizado em uma ligação de duas partes com o gene de ECEPO. As Secções 1 (Xba I / Kpn I), 2 (Kpn I / Bgl II) e 3 (Bgl II / Sal I) tinham sido reunidas previamente pela ordem correcta em M13 e o gene de EPO foi isolado, a partir daquele, sob a forma de um fragmento Xba I / Hind III simples. Este fragmento contém uma porção do poli-ligante do fago M13 mp9 que compreende os sítios Sal I a Hind III. O controlo de expressão no plasmídeo de expressão resultante, p536, foi feito por intermédio de um promotor lambda P_L, que pode, por si só, estar sob controlo do gene repressor C_{I857} (tal fornecido na estirpe K 12 Δ Htrp de "E. coli.")).



O gene de ECEPO preparado anterior pode ser modificado de várias maneiras para codificar análogos de eritropoietina tais como $\angle \text{Asn}^2, \text{des-Pro}^2 \text{ através de Ile}^6 \angle \text{hEPO}$ e $\angle \text{His}^7 \angle \text{hEPO}$, como se descreve adiante.

A. $\angle \text{Asn}^2, \text{des-Pro}^2 \text{ a Ile}^6 \angle \text{hEPO}$

O plasmídeo 536, que transporta o gene preparado de ECEPO do quadro 14 sob a forma de uma inserção Xba I a Hind III, foi digerido com Hind III e Xho I. Esta última endonuclease corta o gene de ECEPO em um sítio de reconhecimento único de 6 pares de bases que abrange desde a última base do códão que codifica Asp^8 até à segunda base do códão Arg^{10} . Preparou-se uma sequência "ligante" Xba I / Xho I com a seguinte sequência.

<u>Xba</u> I		+1	2	7	8	9	
	Met	Ala	Asn	Cys	Asp		<u>Xho</u> I
5'	-CTAG	ATG	GCT	AAT	TGC	GAC	- 3'
	3'	-TAC	CGA	TTA	ACG	CTG	AGCT- 5'

O ligante Xba I / Xho I e o fragmento da sequência do gene Xho I / Hind III ECEPO foram inseridos no fragmento grande resultante da digestão, por Xba I e Hind III, do plasmídeo pCFM 526-um derivado de plasmídeo pCFM 414 (A.T.C.C. 40076)-descrito no pedido de patente de invenção norte-americana Nº. de Série 636.727, depositado em 6 de Agosto de 1984 por Charles F. Morris, para formar uma sequência ADN originada em plasmídeo que codifica a expressão de "E. coli" da forma Met^{-1} do análogo desejado.

B. $\angle \text{His}^7 \angle \text{hEPO}$

O plasmídeo 536 foi digerido com Hind III e Xho I



como se descreveu antes na parte A. Preparou-se um ligante

Xba I / Xho I que têm a seguinte sequência:

<u>Xba</u> I		+1	2	3	4	5	6	7	8	9	<u>Xho</u> I
	Met	Ala	Pro	Pro	Arg	Leu	Ile	His	Asp		
5'-CTAG	ATG	GCT	CCG	CCA	CGT	CTG	ATC	CAT	GAC-3'		
3'-TAC	CGA	GGC	GGT	GCA	GAC	TAG	GTA	CTG	AGCT-5'		

O ligante e o fragmento da sequência de ECEPO tratado com Xho I / Hind III foram depois inseridos em pCFM 526 para formar uma sequência de ADN originário do plasmídeo, que codifica a expressão de "E. coli" da forma Met⁻¹ do análogo desejado.

A construção de um gene preparado ("SCEPO") que incorpora códdãos, de preferência de levedura, é a descrita nos quadros 15 a 21 a seguir. Como acontecia com o gene de ECEPO, a construção completa envolveu a formação de três conjuntos de oligonucleotidos (Quadros 15, 17 e 19) que foram formados em duplexes e reunidos em secções (Quadros 16, 18 e 20). É de notar que a síntese foi facilitada em parte pela utilização de alguns códdãos sub-óptimos em ambas as construções de SCEPO e ECEPO, isto é, os oligonucleotidos 7-12 da Secção 1 de ambos os genes eram idênticos como o eram os oligonucleotidos 1-6 da Secção 2 em cada gene.



Quadro 15

Oligonucleotidos da secção 1 de SCEPO

1. AATTCAAGCTTGGATAAAAGAGCT
2. GTGGAGCTCTTTTATCCAAGCTTG
3. CCACCAAGATTGATCTGTGACTC
4. TCTCGAGTCACAGATCAATCTTG
5. GAGAGTTTTGGAAAGATACTTGTTG
6. CTTCCAACAAGTATCTTTCCAAAAC
7. GAAGCTAAAGAAGCTGAAAACATC
8. GTGGTGATGTTTTTCAGCTTCTTTAG
9. ACCACTGGTTGTGCTGAACACTGTTT
10. CAAAGAACAGTGTTTCAGCACAACCA
11. TTTGAACGAAAACATTACGGTACCG
12. GATCCGGTACCGTAATGTTTTCGTT

Quadro 16

Secção 1 de SCEPO

EcoRI HindIII 1
AATTCA AGCTTGGATA
GT TCGAACCTAT
2

3
AAAGAGCTCC ACCAAGATTG ATCTGTGACT CAGAGTTTT
TTTCTCGAGG TGTTCCTAAC TAGACACTGA GCTCTCAAAA
4

5 7
GGAAAGATAC TTGTTGGAAG CTAAAGAAGC TGAAAACATC ACCACTGGTT
CCTTTCTATG AACAACCTTC GATTTCTTCG ACTTTTGTAG TGGTGACCA
6 8

9 11 KpnI BamHI
GTGCTGAACA CTGTTCTTTG AACGAAAACA TTACGGTACC G
CACGACTTGT GACAAGAAAC TTGCTTTTGT AATGCCATGG CCTAG
12



Quadro 17

Oligonucleotidos da secção 2 de SCEPO

1. AATTCGGTACCAGACACCAAGGT
2. GTTAACCTTGGTGTCTGGTACCG
3. TAACTTCTACGCTTGGAACGTAT
4. TTCCATACGTTTCCAAGCGTAGAA
5. GGAAGTTGGTCAACAAGCAGTTGAAGT
6. CCAAACCTTCAACTGCTTGTTGACCAAC
7. TTGGCAAGGTTTGGCCTTGTTATCTG
8. GCTTCAGATAACAAGGCCAAACCTTG
9. AAGCTGTTTTGAGAGGTCAAGCCT
10. AACAAGGCTTGACCTCTCAAACA
11. TGTTGGTTAACTCTTCTCAACCATGGG
12. TGGTTCCCATGGTTGAGAAGAGTTAACC
13. AACCATTGCAATTGCACGTCGAT
14. CTTTATCGACGTGCAATTGCAA
15. AAAGCCGTCTCTGGTTTGAGATCTG
16. GATCCAGATCTCAAACCAGAGACGG



Quadro 18
Secção 2 de SCEPO

EcoRI KpnI
1
A ATTCGGTACC AGACACCAAG
GCCATGG TCTGTGGTTC
2

3 5
GTAACTTCT ACGCTTGGAA ACGTATGGAA GTTGGTCAAC AAGCTGTTGA
CAATTGAAGA TCGGAACCTT TGCATACCTT CAACCAGTTG TTCGACAAC
4 6

7 9
AGTTGGCAA GGTTTGGCCT TGTTATCTGA AGCTGTTTGG AGAGGTCAAG
TCAAACCGTT CCAAACCGGA ACAATAGACT TCGACAAAAC TCTCCAGTTC
8 10

11 13
CCTTGTGGT TAACTCTTCT CAACCATGGG AACCATTGCA ATTGCACGTC
GGAACAACCA ATTGAGAAGA GTTGGTACCC TTGGTAACGT TAACGTGCAG
12 14

15 BglII BamHI
GATAAAGCCG TCTCTGGTTT GAGATCTG
CTATTTCCGC AGAGACCAAA CTCTAGACCTA G
16



Quadro 19

Oligonucleotidos da secção 3 de SCEPO

1. GATCCAGATCTTTGACTACTTTGTT
2. TCTCAACAAAGTAGTCAAAGATCTG
3. GAGAGCTTTGGGTGCTCAAAAGGAAG
4. ATGGCTTCCTTTTGAGCACCCAAAGC
5. CCATTTCCCCACCAGACGCTGCTT
6. GCAGAAGCAGCGTCTGGTGGGGAA
7. CTGCCGCTCCATTGAGAACCATC
8. CAGTGATGGTTCTCAATGGAGCG
9. ACTGCTGATACCTTCAGAAAGTT
10. GAATAACTTTCTGAAGGTATCAG
11. ATTCAGAGTTTACTCCAATTCT
12. CTCAAGAAGTTGGAGTAACTCT
13. TGAGAGGTAAATTGAAGTTGTACAC
14. ACCGGTGTACAACCTCAATTTACCT
15. CGGTGAAGCCTGTAGAACTGGT
16. CTGTCACCAGTTCTACAGGCTTC
17. GACAGATAAGCCCGACTGATAA
18. GTTGTTATCAGTCGGGCTTAT
19. CAACAGTGTAGATGTAACAAAG
20. TCGACTTTGTTACATCTACACT



Quadro 20

Secção 3 de SCEPO

BamHI BglII 1
 GATC CAGATCTTTG ACTACITTTGT TGAGAGCTTT
 GTCTAGAAAC TGATGAAACA ACTCTCGAAA
2

3 GGGTGCTCAA AAGGAAGCCA TTTCCCACC AGACGCTGCT TCTGCCGCTC
 CCCACGAGTT TTCCTTCGGT AAAGGGGTGG TCTGCGACGA AGACGGCGAG
4 5 6

7 CATTGAGAAC CATCACTGCT GATACCTTCA GAAAGTTATT CAGAGTTTAC
 GTAACCTTTG GTAGTGACGA CTATGGAAGT CTTTCAATAA GTCTCAAATG
8 9 10 11 12

TCCAACCTTCT TGAGAGGTAA ATTGAAGTTG TACACCGGTG AAGCCTGTAG
 AGGTTGAAGA ACTCTCCATT TAACTTCAAC ATGTGGCCAC TTCGGACATC
13 14 15 16

AACTGGTAC AGATAAGCCC GACTGATAAC AACAGTGTAG
 TTGACCACTG TCTATTTCGGG CTGACTATTG TTGTCACATC
17 18 19

SalI
 ATGTAACAAA G
 TACATTGTTT CAGCT
20



Quadro 21

Gene SCEP0

-1 +1

HindIII	ArgAla				
AGCTTGGATA	AAAGAGCTCC	ACCAAGATTG	ATCTGTGACT	CGAGAGTTTT	
ACCTAT	TTTCTCGAGG	TGGTTCTAAC	TAGACACTGA	GCTCTCAAAA	

GGAAAGATAC	TTGTTGGAAG	CTAAAGAAGC	TGAAAACATC	ACCACTGGTT	
CCTTTCTATG	AACAACCTTC	GATTTCTTCG	ACTTTTGTAG	TGGTGACCAA	

GTGCTGAACA	CTGTTCTTTG	AACGAAAACA	TTACGGTACC	AGACACCAAG	
CACGACTTGT	GACAAGAAAC	TTGCTTTTGT	AATGCCATGG	TCTGTGGTTC	

GTAACTTCT	ACGCTTGGAA	ACGTATGGAA	GTTGGTCAAC	AAGCTGTTGA	
CAATTGAAGA	TGCGAACCTT	TGCATACCTT	CAACCAGTTG	TTCGACAACT	

AGTTTGGCAA	GGTTTGGCCT	TGTTATCTGA	AGCTGTTTTG	AGAGGTCAAG	
TCAAACCGTT	CCAAACCGGA	ACAATAGACT	TCGACAAAAC	TCTCCAGTTC	

CCTTGTTGGT	TAActCTTCT	CAACCATGGG	AACCATTGCA	ATTGCACGT	
GGAACAACCA	ATTGAGAAGA	GTTGGTACCC	TTGGTAACGT	TAACGTGCAG	

GATAAAGCCG	TCTCTGGTTT	GAGATCTTTG	ACTACTTTGT	TGAGAGCTTT	
CTATTTCCGC	AGAGACCAAA	CTCTAGAAAC	TGATGAAACA	ACTCTCGAAA	

GGGTGCTCAA	AAGGAAGCCA	TTTCCCCACC	AGACGCTGCT	TCTGCCGCTC	
CCCACGAGTT	TTCTTCCGGT	AAAGGGGTGG	TCTGCGACGA	AGACGGCGAG	

CATTGAGAAC	CATCACTGCT	GATACCTTCA	GAAAGTTATT	CAGAGTTTAC	
GTAActCTTG	GTAGTGACGA	CTATGGAAGT	CTTTCAATAA	GTCTCAAATG	

TCCAActTCT	TGAGAGGTAA	ATTGAAGTTG	TACACCGGTG	AAGCCTGTAG	
AGGTTGAAGA	ACTCTCCATT	TAActTCAAC	ATGTGGCCAC	TTCGGACATC	

AACTGGTGAC	AGATAAGCCC	GAActGATAAC	AACAGTGTAG		
TTGACCACTG	TCTATTCGGG	CTGACTATTG	TTGTCAACATC		

	SalI				
ATGTAACAAA	G				
TACATTGTTT	CAGCT				



As secções de SCEPO reunidas foram postas em sequência em M13 e as Secções 1, 2 e 3 puderam ser isoladas do fago sob a forma de fragmentos Hind III / Kpn I, Bgl II e Bgl II / Sal I.

O sistema de expressão presentemente preferido para os produtos do gene SCEPO é um sistema de secreção com base na secreção do factor α de "S. cerevisiae", como se descreve no pedido de patente de invenção norte-americana Nº de Série 487.753, depositado em 22 de Abril de 1983 por Grant A. Bitter, publicado em 31 de Outubro de 1984 como patente de invenção europeia O 123.294. De uma maneira resumida, o sistema implica construções em que o ADN que codifica a sequência condutora do produto genético factor α de levedura está situado imediatamente na posição 5' em relação à região de codificação do gene exógeno a ser expresso. Como resultado, o produto genético transferido contém uma sequência condutora ou de sinal que é "processado" por um enzima de levedura endógeno no decurso da secreção do remanescente do produto. Como a construção utiliza o códão de início de transferência do factor α (ATG), não é necessário fornecer tal códão à posição 1 do gene de SCEPO. Como pode observar-se a partir do quadro 21, a sequência que codifica a alanina (+1) é precedida por uma sequência ligante que permite a inserção directa em um plasmídeo que contém o ADN para os primeiros 80 restos do condutor do factor α que segue o promotor do factor α . A construção preferida específica para expressão do gene de SCEPO envolveu uma ligação em quatro partes que compreende os fragmentos de secção de SCEPO mencionados antes e o fragmento grande da digestão, por Hind III / Sal I, do plasmídeo pC3. A partir do plasmídeo pC3/SCEPO resultante, isolou-se o promotor do factor α e a sequência condutora e o



gene de SCEPO mediante digestão com Bam HI e ligou-se ao plasmídeo pYE digerido por Bam HI para formar o plasmídeo de expressão pYE/SCEPO.

Exemplo 12

O presente exemplo refere-se à expressão de produtos recombinantes dos genes de ECEPO e SCEPO preparados dentro dos sistemas de expressão do exemplo 11.

Na utilização do sistema de expressão concebido para ser utilizado com células-hospedeiras de "E. coli", o plasmídeo p536 do exemplo 11 foi transformado em células de "E. coli" AM7 previamente transformadas com um plasmídeo apropriado, pMW1, que aloja um gene C_{I857} . As culturas de células em caldo LB (Am picilina, 50 $\mu\text{g/ml}$ e canamicina 5 $\mu\text{g/ml}$, de preferência com 10 mM de sulfato de magnésio) foram mantidas a 28° C e, após desenvolvimento das células em cultura até D.O.₆₀₀ = 0,1, induziu-se a expressão de EPO mediante elevação da temperatura da cultura até 42° C. As células cultivadas até cerca de D.O. 40 proporcionaram uma produção de eritropoietina (EPO) (avaliada em gel) de cerca de 5 mg/D0 litro.

As células foram colhidas, desintegradas, quebradas com uma Prensa Francesa (704 Kg/cm²; 10.000 psi) e tratadas com lisozima e detergente NP-40. O sedimento, resultante da centrifugação a 24.000xg, foi solubilizado com cloridrato de guanidina e submetido a purificação ulterior em uma só fase, utilizando uma cromatografia líquida a pressão elevada, de fase inversa C₄ (Vydac) (etanol, 0-80%, 50mM de acetato de amónio, pH 4,5). A sequenciação da proteína revelou que o produto tinha um grau de pureza superior a 95% e os produtos obtidos revelaram dois grupos amino terminais diferentes, A-P-P-R... e



P-P-R... em uma razão quantitativa relativa de cerca de 3 para 1. Esta última observação de produtos h EPO e $\left[\text{des Ala}^1 \right]$ h EPO indica que o "processamento" do grupo amino terminal dentro das células-hospedeiras serve para remover a metionina terminal e, em certos casos, a alanina inicial. A actividade, no ensaio rádio-imunológico, para os isolados era de 150.000 a 160.000 U/mg; a actividade, no ensaio "in vitro", era de 30.000 a 62.000 U/mg ; e a actividade, no ensaio "in vivo", estava compreendida entre cerca de 120 e 720 U/mg. (Cf., padrão de isolado urinário humano de 70.000 U/mg em cada ensaio). A curva de dose resposta para o produto recombinante no ensaio "in vivo" diferia acentuadamente do padrão de eritropoietina urinária humana.

Os plasmídios análogos de eritropoietina formados nas partes A e B do exemplo 11 foram, cada um deles, transformados em células AMT de "E. coli" transformadas por pMW1 e as células foram cultivadas como se indicou antes. Os isolados purificados foram submetidos a ensaio rádio-imunológico e a ensaios "in vitro". Os valores destes ensaios para os produtos de expressão $\left[\text{Asn}^2, \text{des-Pro}^2 \text{ a Ile}^6 \right]$ hEPO foram aproximadamente 11.000 U/mg e 6000 U/mg de proteína, respectivamente, enquanto os valores de ensaio para $\left[\text{His}^7 \right]$ hEPO foram cerca de 41.000 U/mg e 14.000 U/mg de proteína, respectivamente, indicando que os produtos análogos eram desde um quarto a um décimo tão "activos" como o produto de expressão "original" nos ensaios.

No sistema de expressão destinado a ser utilizado por células-hospedeiras de "S. cerevisiae", o plasmídio pYE/SCEPO foi transformado em duas estirpes diferentes, YSDP4 (genotipo $\alpha \alpha$ pep 4-3 trp1) e RK81 (genotipo $\alpha \alpha$ pep4-3 trp1). Os hos



pedeiros YSDP4 transformados foram cultivados em meio SD ["Methods in Yeast Genetics", Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y., pág. 62 (1983)] suplementado com casaminoácidos a 0,5%, de pH 6,5, a 30°C. Os meios colhidos quando as células tinham sido cultivadas até D.O. 36 continham produtos de eritropoietina em concentrações de cerca de 244 U/ml (97 µg/D.O. litro no ensaio rádio-imunológico). As células RK 81 transformadas, cultivadas a D.O. de 6,5 ou a D.O. de 60, forneceram meios com concentrações de eritropoietina de cerca de 80-90 U/ml (34 µg/D.O. litro no ensaio rádio-imunológico). Análises preliminares revelam uma heterogeneidade significativa em produtos produzidos pelo sistema de expressão, provavelmente devido a variações na glucosilação de proteínas expressas, e um teor de manose relativamente elevado do hidrato de carbono associado.

Os plasmídios PαC3 e pYE em células de "E. coli" HB 101 foram depositados, de acordo com as Regras de Prática da Repartição de patentes norte-americana, em 27 de Setembro de 1984, com a "American Type Culture Collection", 12301 Parklawn Drive, Rockville, Maryland, sob os números de depósito A.T.C.C. 39881 e A.T.C.C. 39882, respectivamente. Os plasmídios pCFM 526 em células AM7, pCFM536 em células JM103 e pMW1 em células JM103 foram igualmente depositados em 21 de Novembro de 1984 como A.T.C.C. 33932, 33934 e 33933, respectivamente. Foram depositadas estirpes YSPD4 e RK81 de "Saccharomyces cerevisiae" em 21 de Novembro de 1984 como A.T.C.C. 20734 e 20733, respectivamente.

Deverá tornar-se perfeitamente óbvio depois da consideração dos exemplos ilustrativos anteriores que são proporcio-



nados pela presente invenção nas suas múltiplas modalidades, muitos produtos e processos excepcionalmente valiosos.

Os polipeptidos de acordo com a presente invenção são produtos manifestamente úteis, quer sejam produtos expressos microbiologicamente quer sejam produtos sintéticos, cuja conformação estrutural primária, secundária ou terciária foi pela primeira vez tornada conhecida pela presente invenção.

Como se indicou antes, os produtos obtidos por métodos recombinantes e os produtos sintéticos de acordo com a presente invenção partilham, em graus variáveis, a actividade biológica "in vitro" de isolados de eritropoietina de fontes naturais e, por conseguinte, são concebidos para terem utilidade como substitutos de isolados de eritropoietina em meios de cultura utilizados para desenvolvimento de células eritropoieticas em cultura. Analogamente, pelo facto de os produtos polipeptídicos de acordo com a presente invenção partilham a actividade "in vivo" de isolados de eritropoietina natural, são manifestamente apropriados para utilização em técnicas de terapia com eritropoietina praticadas em mamíferos, incluindo o homem, para desenvolver qualquer efeito ou todos os efeitos até agora atribuídos "in vivo" à eritropoietina, por ex. estimulação da resposta do reticulocito, desenvolvimento de efeitos ferrocinéticos (como os efeitos de "turnover" do ferro do plasma e os efeitos do tempo de passagem pela medula), alterações da massa dos eritrocitos, estimulação da síntese de hemoglobina C (ver, Eschbach e colab., supra) e, como se indicou no exemplo 10, níveis de hematocrito aumentados em mamíferos. Estão compreendidos na classe de seres humanos que podem ser tratados com produtos de acordo com a presente invenção os pacientes que



geralmente necessitam de transfusões de sangue que incluem vítimas de trauma, pacientes cirúrgicos, pacientes de doenças re-
nais como os pacientes de diálise e pacientes com perturbações
que afectam uma variedade de composições de sangue, como hemo-
filia, doença das células em foice, anemias fisiológicas e si-
milares. Pode esperar-se que, da minimização da necessidade de
terapia de transfusão através da utilização de terapia de eri-
tropoietina, resulte uma redução na transmissão de agentes in-
fecciosos. Espera-se que os produtos de acordo com a invenção,
em virtude da sua produção por métodos recombinantes, estejam
isentos de pirogénios, de substâncias inibidoras naturais, e
similares e proporcionem assim uma eficácia global aumentada
em processos terapêuticos relativamente aos produtos derivados
naturalmente. É também de esperar que a terapia da eritropoie-
tina da presente invenção seja utilizada no aumento da capaci-
dade de transporte do oxigénio de indivíduos que se encontram
em condições de ambiente hipóxico e, possivelmente, para pro-
porcionar efeitos cardiovasculares benéficos.

Um método preferido para a administração de produtos
polipeptídicos de acordo com a invenção é o que utiliza a via
parenteral (a saber, intravenosa, intramuscular, subcutânea ou
intraperitoneal) e as composições administradas deverão habi-
tualmente conter quantidades terapêuticamente eficazes de pro-
duto associado a diluentes, excipientes e/ou adjuvantes aceitá-
veis. Estudos farmacocinéticos preliminares indicam uma semi-
-vida mais prolongada para os produtos de eritropoietina de ma
caco quando administrados por via intramuscular em vez da via
intravenosa. É de esperar que as dosagens eficazes variem su-
bstancialmente consoante a situação tratada, mas as doses tera



pêuticas devem estar compreendidas entre 0,1 (~ 7 U) e 100 (~ 7000 U) $\mu\text{g}/\text{Kg}$ de peso de corpo, do produto activo. Considere-se, para as composições farmacêuticas de acordo com a invenção, diluentes convencionais como a albumina de soro humano e excipientes convencionais como o soro fisiológico.

Os adjuvantes apropriados para utilização nas composições de acordo com a presente invenção incluem compostos independentemente indicados para a obtenção de efeitos estimuladores eritropoiéticos, como as testosteronas, os estimuladores de células progenitoras, factor de crescimento análogo à insulina, prostaglandinas, serotonina, ácido 5'-adenílico (AMP), prolactina e triiodotironina, assim como os agentes geralmente utilizados no tratamento de anemia aplástica, como metenoleno, estanozolol e nandrolona [ver, por ex., Resegotti e colab., "Panminerva Medica", 23, 243-248 (1981); McGonigle e colab., "Kidney Int", 25 (2), 437-444 (1984); Pavlovic-Kanterá e colab., "Exptl. Hematol.", 8 (Supl. 8), 283-291 (1980); e Kurtz, "FEBS Letters", 14 a (1), 105-108 (1982)]. São também consideradas como adjuvantes as substâncias que são referidas como aumentando os efeitos da, ou sinergizarem a, eritropoietina ou asialo-eritropoietina, como os agonistas adrenérgicos, as hormonas da tiroide, os androgénios e BPA [ver, Dunn, "Current Concepts in Erythropoiesis", John Wiley and Sons (Chichester, England, 1983); Weiland e colab., "Blut", 44(3), 173-175 (1982); Kalmanti, "Kidney Int.", 22, 383-391 (1982); Shahidi, "New Eng. J. Med.", 289, 72-80 (1973); Fisher e colab., "Steroids", 30(6), 833-845 (1977); Urabe e colab., "J. Exp. Med.", 149, 1314-1325 (1979); e Billat e colab., "Exptl. Hematol." 10(1), 133-140 (1982)] assim como as classes de compostos designados "factores



eritropoiéticos hepáticos" [ver, Naughton e colab., "Acta Haemat.", 69, 171-179 (1983)] e "eritrotropinas" [descritas por Congote e colab. em "Abstract" (Resumo) 364, "Proceedings 7th International Congress of Endocrinology" (Quebec City, Quebec, 1-7 de Julho de 1984); Congote, "Biochem. Biophys. Res. Comm.", 115(2), 447-483 (1983) e Congote, "Anal. Biochem.", 140, 428-433 (1984)] e "eritrogeninas" [descritas por Rothman e colab., "J. Surg. Oncol.", 20, 105-108 (1982)]. Avaliações preliminares destinadas a medir as respostas eritropoiéticas de murganhos policitémicos ex-hipóxicos previamente tratados com 5- α -dihidrotestosterona ou com mandrolona e a que se administrou eritropoietina da presente invenção conduziram a resultados duvidosos.

As utilizações, com fins de diagnóstico, dos polipeptidos de acordo com a presente invenção são analogamente extensivas e compreendem a utilização em formas marcadas ou não marcadas em uma grande variedade de técnicas de imuno-ensaio que incluem os de RIA, de Elisa e similares, assim como uma variedade de ensaios de actividade "in vitro" e "in vivo". Ver por ex., Dunn e colab., "Exptl. Hematol.", 11(7), 590-600 (1983); Gibson e colab., "Pathology", 16, 155-156 (1984); Krystal, "Exptl. Hematol.", 11(7), 649-660 (1983); Saito e colab., "Jap. J. Med.", 23(1), 16-21 (1984); Nathan e colab., "New Eng. J. Med.", 308 (9), 520-522 (1983); e diversas referências relativas a ensaios aqui mencionados. Os polipeptidos de acordo com a presente invenção, que compreendem peptidos sintéticos que contêm sequências de restos de eritropoietina aqui revelados pela primeira vez, também proporcionam materiais puros extremamente activos para gerar anticorpos policlonais e "bancos" de an-



ticorpos monoclonais específicos para distinguir epítopes contínuos ou descontínuos de eritropoietina. Como exemplo, a análise preliminar das sequências de aminoácidos do quadro 6 no contexto de hidropaticidade de acordo com Hopp e colab., "P.N. A.S. (U.S.A.)", 78, 3824-3828 (1981) e de estruturas secundárias de acordo com Chou e colab., "Ann. Rev. Biochem.", 47, 251 (1978) revelou que os peptidos sintéticos duplicativos de sequências contínuas de restos que abrangem as posições 41 a 57 inclusive, 116 a 118 inclusive e 144 a 166 inclusive são susceptíveis de produzir uma resposta altamente antigénica e de gerar anticorpos monoclonais e policlonais utilizáveis imuno-reactivos com o peptido sintético e a proteína completa. Prevê-se que estes anticorpos possam ser utilizados na detecção e na purificação de afinidade de eritropoietina e de produtos relacionados com a eritropoietina.

Ilustrativamente, foram preparados os três peptidos sintéticos seguintes:

- (1) hEPO 41-57, V-P-D-T-K-V-N-F-Y-A-W-K-
R-M-E-V-G;
- (2) hEPO 116-128, K-E-A-I-S-P-P-D-A-A-S-A-A-;
- (3) hEPO 144-166, V-Y-S-N-F-L-R-G-K-L-K-L-Y-
T-G-E-A-C-R-T-G-D-R.

Estudos de imunização preliminares que utilizam os polipeptidos mencionados antes revelaram uma resposta positiva relativamente fraca a hEPO 41-57, nenhuma resposta apreciável a hEPO 116-128 e uma resposta positiva forte a hEPO 144-166, medidas pela capacidade de os anticorpos de soro de coelho imu



no-precipitarem isolados de EPO urinário humano marcado com ¹²⁵I. Estudos de actividade preliminares "in vivo" efectuados sobre os três peptidos não revelaram qualquer actividade significativa quer sozinhos quer associados.

Embora as sequências deduzidas de restos de aminoácidos de eritropoietina de mamíferos fornecidas pelos exemplos ilustrativos definam essencialmente a conformação estrutural primária da eritropoietina completa, deverá entender-se que a sequência específica de 165 restos de aminoácidos da eritropoietina de macaco no quadro 5 e os 166 restos da eritropoietina humana no quadro 6 não limitam o âmbito dos polipeptidos utilizáveis de acordo com a presente invenção. Estão compreendidos na presente invenção as diversas formas alélicas naturais de eritropoietina que as investigações já efectuadas em polipeptidos de mamífero biologicamente activos como o interferon γ humano indicam que poderão existir. [Comparar, por ex. a espécie de interferon imune humano a que se atribui um resto arginina na posição Nº. 140 na eritropoietina, como refere o pedido de patente de invenção europeia publicado Nº. 0077670, e a espécie a que se atribui um resto de glutamina na posição Nº. 140, como referem Gray e colab. em "Nature", 295, 503-508 (1982). Ambas as espécies são caracterizadas por constituírem sequências de interferon γ humano "completo"]. As formas alélicas dos polipeptidos de eritropoietina completa podem diferir umas das outras e das sequências dos quadros 5 e 6 em termos de comprimento da sequência e/ou em termos de supressões, substituições, inserções ou adições de aminoácidos na sequência, com variantes potenciais consequentes na capacidade de glicosilação. Como se referiu antes, admite-se que uma forma alélica putativa da eri



tropoietina da espécie humana contém um resto metionina na posição 126. Como se esperava, podem também ocorrer formas alélicas naturais de EPO codificando ADN genómico e sequências de cADN, as quais codificam os tipos anteriormente citados de polipeptidos alélicos ou utilizam simplesmente códigos diferentes para designação dos mesmos polipeptidos especificados.

Além das formas alélicas naturais de eritropoietina completa, a presente invenção compreende também outros "produtos de eritropoietina" como análogos de eritropoietina e fragmentos de eritropoietina "completa". Utilizando as técnicas descritas no pedido de patente publicado anteriormente citado de Alton e colab. (WO/83/04053) pode conceber-se e preparar-se facilmente genes que codificam a expressão microbiana de polipeptidos com conformações primárias que diferem da especificada aqui para a eritropoietina completa em termos da identidade de localização de um ou mais restos (por ex. substituições, adições terminais e intermédias e supressões). Em alternativa, pode conseguir-se facilmente modificações de cADN e de genes de eritropoietina genómica mediante técnicas de mutagénesse orientadas para o sítio, bem conhecidas, e utilizadas para formar análogos e derivados de eritropoietina. Estes produtos de eritropoietina deveriam partilhar pelo menos uma das propriedades biológicas da eritropoietina mas podem diferir nas outras. Como exemplos, os produtos de eritropoietina projectados de acordo com a presente invenção compreendem os que são encurtados, por exemplo, por supressões $\left[\text{Asn}^2, \text{des-Pro}^2 \text{ a } \text{Ile}^6 \right] \text{hEPO}$, $\left[\text{des-Thr}^{163} \text{ a } \text{Arg}^{166} \right] \text{hEPO}$ e " $\triangle 27-55\text{hEPO}$ ", tendo a última os restos codificados por um exon completo suprimido; ou os que são mais estáveis à hidrólise (e, por isso, podem ter efeitos dura-



douros mais pronunciados ou prolongados que a eritropoietina natural); ou os que têm sido alterados para suprimir um ou mais sítios potenciais para glicosilação (que podem resultar em actividades superiores para produtos derivados da levedura);

os que têm um ou mais restos de cisteína suprimidos ou substituídos por por ex., restos de histidina ou de serina (como o análogo $\text{[His}^7\text{]hEPO}$) e, potencialmente mais fáceis de isolar em forma activa a partir dos sistemas microbianos; ou que têm um ou mais restos de tirosina substituídos por fenilalanina (como os análogos $\text{[Phe}^{15}\text{]hEPO}$, $\text{[Phe}^{49}\text{]hEPO}$ e $\text{[Phe}^{145}\text{]hEPO}$) e podem ligar-se mais ou menos prontamente a receptores de eritropoietina como células de alvo. Estão também compreendidos fragmentos de polipeptido que duplicam apenas uma parte da sequência de aminoácidos contínua ou as conformações secundárias dentro de EPO completa, fragmentos estes que podem possuir uma actividade de EPO (por ex. ligação de receptor) e não outras (por ex. actividade eritropoiética). São especialmente significativos a este respeito os fragmentos potenciais de EPO que são esclarecidos após consideração da sequência de ADN genómico humano do quadro 6, isto é, "fragmentos" da sequência de EPO contínua total que são delineados por sequências intrão e que podem constituir "domínios" distintos de actividade biológica. É de notar que a ausência de actividade "in vivo" para qualquer um ou mais dos "produtos de EPO" de acordo com a invenção não é totalmente impeditiva de utilidade terapêutica (ver Weiland e colab., supra) ou de utilidade em outros contextos, como em ensaios de EPO ou antagonismo de EPO. Os antagonistas da eritropoietina podem ser muito úteis no tratamento de policitémia ou em casos de produção excessiva de eritropoietina [ver, por



ex., Adamson, "Hosp. Practice", 18(12), 49-57 (1983) e Hellmann e colab., "Clin. Lab. Haemat.", 5, 335-342 (1983) 7.

De acordo com uma outra modalidade da presente invenção, as sequências de ADN clonadas aqui descritas que codificam polipeptidos de eritropoietina humana ou de macaco são manifestamente valiosas para a informação que proporcionam relativamente à sequência de aminoácidos de eritropoietina de mamíferos de que não se dispunha anteriormente apesar de décadas de processamento analítico de isolados de produtos naturais. As sequências de ADN são também manifestamente valiosas como produtos úteis para efectuar a síntese microbiana, em grande escala, da eritropoietina mediante uma grande variedade de técnicas recombinantes. Ou, por outras palavras, as sequências de ADN de acordo com a presente invenção podem ser utilizadas para gerar novos vectores de ADN de plasmídeo viral ou circular utilizáveis, novas células procarióticas ou eucarióticas microbianas transformadas ou modificadas utilizáveis (incluindo células de bactérias e de levedura e células de mamíferos desenvolvidas em cultura) e novos métodos que podem ser utilizados no desenvolvimento, mediante cultura, dessas células-hospedeiras microbianas capazes da expressão de eritropoietina e de produtos de eritropoietina. As sequências de ADN da invenção são também materiais manifestamente apropriados para serem utilizados como sondas marcadas no isolamento de EPO e de proteína com aquela relacionada que codificam sequências de cADN e de ADN genómico de mamíferos além das do homem e do macaco aqui especificadas. Não pode ainda calcular-se até que ponto as sequências de ADN da invenção poderão ser utilizadas em diversos processos de alternativa da síntese de proteínas



(por ex. em células de insectos) ou em terapia genética no homem e em outros mamíferos. É de esperar que as sequências de ADN de acordo com a presente invenção possam ser utilizadas no desenvolvimento de espécies transgénicas de mamíferos que podem servir como "hospedeiros" eucarióticos para a produção de eritropoietina e de produtos de eritropoietina em quantidade. Ver, na generalidade, Palmiter e colab., "Science", 222 (4625), 809-814 (1983).

Por isso, consideradas deste modo, as descrições específicas dos exemplos ilustrativos não pretendem nitidamente ser limitativas do âmbito da presente invenção sendo previsíveis numerosas modificações e variações para os entendidos na matéria. Como exemplo, embora as sequências de ADN fornecidas pelos exemplos ilustrativos compreendam sequências de cADN e de ADN genómico, porque este pedido de patente de invenção fornece informação sobre a sequência de aminoácidos essencial à preparação da sequência de ADN, a invenção compreende também as sequências de ADN preparadas que podem ser construídas com base no conhecimento das sequências de aminoácidos de eritropoietina. Estas podem codificar EPO (como no exemplo 12) assim como fragmentos e análogos de polipeptidos de EPO (isto é, "Produtos de EPO") que podem partilhar uma ou mais propriedades biológicas da EPO natural mas não partilham outras (ou possuem outras em graus diferentes).

Observa-se assim que as sequências de ADN de acordo com a presente invenção compreendem todas as sequências de ADN apropriadas para utilização com o fim de assegurar a expressão em uma célula-hospedeira procariótica ou eucariótica de um produto polipeptídico que tenha pelo menos uma parte da confor-



mação estrutural primária e uma ou mais das propriedades biológicas da eritropoietina, e escolhidas entre: (a) as sequências de ADN indicadas nos quadros 5 e 6; (b) as sequências de ADN que hibridam para as sequências de ADN definidas em (a) ou fragmentos destas; e (c) as sequências de ADN que, se não fosse a degeneração do código genético, deveriam hibridar para as sequências de ADN definidas em (a) e (b). Deverá notar-se a este respeito, por exemplo, que existem sequências de gene de EPO de macaco e humana alélicas e é de esperar que haja outras sequências de genes de mamíferos que hibridem para as sequências dos quadros 5 e 6 ou para fragmentos destas. Além disso, se não fosse a degenerescência do código genético, os genes de SCEPO e de ECEPO e as sequências de ADN genómico ou de cADN mutagenizado ou preparado que codificam diversos fragmentos e análogos de EPO deveriam também hibridar para as sequências de ADN mencionadas antes. Essas hibridações podiam ser efectuadas facilmente sob as condições de hibridação aqui descritas em relação ao isolamento inicial do ADN que codifica EPO de macaco ou humana ou sob condições mais rigorosas, para reduzir eventualmente a hibridação de base.

De maneira análoga, embora os exemplos anteriores ilustrem a invenção da expressão microbiana de produtos de EPO no contexto da expressão, em células de mamíferos, de ADN inserido em um vector híbrido de um plasmídeo bacteriano ou de origem genómica viral, está considerada, na presente invenção, uma grande variedade de sistemas de expressão. Estão obviamente considerados, no âmbito da invenção, sistemas de expressão que envolvem vectores de origens homogéneas aplicados a uma variedade de bactérias, de leveduras ou de células de mamíferos em



cultura, assim como sistemas de expressão que não envolvem vectores (como modificação de células com fosfato de cálcio). A este respeito, deverá compreender-se que a expressão de, por ex. ADN de macaco em células hospedeiras de macaco em cultura e células hospedeiras humanas em cultura, constitui realmente casos de expressão de ADN "exógeno" uma vez que o ADN de eritropoietina, cujo nível elevado de expressão se pretende, não deverá ter a sua origem no genoma do hospedeiro. Os sistemas de expressão da invenção abrangem ainda estas práticas do que resulta a formação citoplásmica de produtos de eritropoietina e a acumulação de produtos de EPO glicosilados e não-glicosilados no citoplasma ou na membrana da célula-hospedeira (por ex. acumulação em espaços periplásmicos bacterianos) ou em líquidos sobrenadantes de meios de cultura como se ilustrou antes, ou em sistemas muito pouco comuns, como os sistemas de expressão de "P. aeruginosa" [descritos por Gray e colab. em "Biotechnology", 2, 161-165 (1984)].

As metodologias de hibridação melhoradas da invenção, conquanto que aplicadas ilustrativamente à avaliação da hibridação de ADN/ADN, são igualmente aplicáveis à avaliação de ARN/ARN e ARN/ADN. As técnicas de sondas misturadas, como se ilustrou antes, constituem um grande número de melhorias em processos de hibridação que permitem isolamentos mais rápidos e precisos de polinucleotidos. Todos estes aperfeiçoamentos no processamento individual compreendem: transferência de colónias e técnicas de manutenção melhoradas; utilização de filtros de nylon como GeneScreen e GeneScreen Plus para permitir uma nova sondagem com os mesmos filtros e a utilização repetida do filtro, aplicação de novos tratamentos com protease [comparar, por



ex. com Taub e colab. em "Anal. Biochem.", 126, 222-230 (1982) 7;
utilização de concentrações individuais muito fracas (da ordem
de 0,025 picomole) de um grande número de sondas mistas (por
ex. um excesso de mais de 32); e fases de realização da hibri
dação e da post-hibridação sob temperaturas rigorosas que se
aproximam estreitamente (isto é dentro de 4°C e, de preferên
cia, dentro de 2°C a partir de) da mais baixa temperatura de
dissociação calculada de qualquer das sondas mistas utilizadas.
Estas melhorias reunidas proporcionam resultados que não podiam
ser esperados da sua utilização. Isto é amplamente ilustrado
pelo facto de que as técnicas de sondas mistas que envolvem
quatro vezes o número de sondas alguma vez referido antes como
tendo sido utilizado com êxito em avaliações de cADN sobre espé
cies de ARN mensageiro relativamente pouco abundantes foram apli
cadas com sucesso ao isolamento de um único gene de sequência
em uma avaliação de colecção genómica de 1.500.000 placas de
fagos. Este feito foi conseguido essencialmente ao mesmo tempo
da publicação da opinião considerada de Anderson e colab., supra,
de que os métodos de avaliação com sondas mistas eram "...pou
co práticos para o isolamento de genes proteicos de mamíferos
quando não se dispõe dos ARN correspondentes."



P. 2. 4081.213

R e i v i n d i c a ç õ e s

"processo para a produção de eritropoietina"

1.- Processo para a preparação de um polipeptido purificado e isolado tendo parte ou toda a conformação estrutural primária e uma ou mais das propriedades biológicas da eritropoietina existente na natureza, caracterizado pelo facto de se promover a expressão procariótica ou eucariótica de uma sequência de ADN exógena.

2.- Processo de acordo com a reivindicação 1, caracterizado ainda pelo facto de se eliminar do polipeptido a associação com qualquer proteína de um mamífero.

3.- Processo de acordo com a reivindicação 1, caracterizado pelo facto de a sequência de ADN exógena ser uma sequência de cADN.

4.- Processo de acordo com a reivindicação 1, caracterizado pelo facto de a sequência de ADN exógena ser uma sequência de ADN manufacturada.

5.- Processo de acordo com a reivindicação 1, caracterizado pelo facto de a sequência de ADN exógena ser uma sequência de ADN genómica.

6.- Processo de acordo com a reivindicação 1, caracterizado pelo facto de se efectuar a sequência de ADN exógena sobre um plasmídeo de ADN circular que se reproduz exactamente de maneira autónoma ou sobre um vector viral.

7.- Processo de acordo com a reivindicação 1, caracterizado



pelo facto de o polipeptido obtido possuir parte ou toda a conformação estrutural primária de eritropoietina humana indicada a seguir:

```

AAGCTTC TGGGCTTCCAGACCCAGCTACTTTGCGGAAC TCAGCAACCAGGCATCTCTGAGTCTCCGCCCA
AGACCGGGATGCCCCCAGGGGAGGTGTCCGGGAGCCCCAGCCTTTCCCAGATAGCACGCTCCGCCAGTCCC
AAGGTGCGCAACCGGCTGCACTCCCC TCCCGGACCCAGGGCCCCGGAGCAGCCCCCATGACCCACACGC
ACGTC TGCAGAGCCCCGCTCAGGCCCCGGCGAGCCTCAACCCAGGCGTCTTGCCCCCTGCTCTGACCCCGG
GTGGCCCC TACCCCTGGCGACCCCTCAGGCACACAGCCTCTCCCCCACCACCACCCGCGCACACATG
CAGATAACAGCCCCGACCCCGCCAGAGCCGXAGAGTCCCTGGGCCACCCCGCGCTCGCCTGCCGCTG
CGCCGCACCGCGCTGTCTCCCGGAGCCGACCGGGGCCACCCGCCCCXGCTCTGCTCCGACACCGCGCCC
CTTGGACAGCCGCCCTCTCTCTAGGCCCGTGGGGCTGGCCCTGCACCGCCGAGCTTCCCGGATGAGGXX
                                -27      -24
                                Met Gly Val His
                                ATG GGG GTG CAC G

GTGAGTACTCGCGGGCTGGGCGCTCCCGCGCGCGCGGTTCCTGTTGAGCGGGGATTTAGCGCCCCCGCT

```

::



AT TGGCC AAGAGGTGGCTGGGTTCAAGGACCGCGGACTTGTCAAGGACCCCGGAAGGGGGAGGGGGTGGG
GCAGCCTCCACGTGCCGCGGGGACTTGGGGGAGTTCTTGGGGATGGCAAAACCTGGCCTGTTGAGGGGCA
CAGTTTGGGGTTGGGGAGGAGGTTTGGGGTCTGCTGTGCAGTTGTGTCGTTGTCAGTGTCTCG[I.S.]
TTGCACACGCACAGATCAATAAGCCAGAGGCAGCACCTGAGTGTGCTTGCATGGTTGGGACAGGAAGGACGAG

CTGGGGCAGAGACGTGGGGATGAAGGAAGCTGTCCTCCACAGCCACCCTTCTCCCCCCCCGGCTGACTCT

CAGCCTGGCTATCTGTTCTAG

-10	Leu	Ser	Leu	Pro	Leu	Gly	Leu	Val	Pro	Val	Leu	Gly	Ala	Pro	Pro	Arg	Leu	Ile	Cys
	CTG	TGG	CTC	CCT	CTG	CTG	CTG	CTG	CCA	GTC	CTG	GGC	GGC	CCA	CCA	CGC	CTC	ATC	TGT

10 20 *

Asp Ser Arg Val Leu Glu Arg Tyr Leu Leu Glu Ala Lys Glu Ala Glu Asn Ile
GAC AGC CGA GTC CTG GAG AGG TAC CTC TTG GAG GCC AAG GAG GAG AAT ATC

26 Thr GTGAGACCCCTTCCCAGGCACATTCCACAGAACTCAGGCTCAGGGCTTCAGGGAACTCCTCCCAGAT
ACG

CCAGGAACCTGGCACTTGGTTTGGGGTGGAGTTGGGAAGCTAGACACTGCCCCCTACATAAGAAATAAGTC



TGGTGGCCCCAAACCATACCTGAAACTAGGCAAGGAGCAAGCCAGCAGATCCTACGCCCTGTGGCCAGGG

27 Thr Gly Cys Ala Glu
ACG GGC TGT GCT GAA

CCAGAGCCCTTCAGGGACCCCTTGACTCCCGGGCTGTGTGCAATTCAG

40
* Val Pro Asp Thr Lys Val Asn Phe Tyr
His Cys Ser Leu Asn Glu Asn Ile Thr Val Pro Asp Thr Lys Val Asn Phe Tyr
CAC TGC AGC TTG AAT GAG AAT ATC ACT GTC CCA GAC ACC AAA GTT AAT TTC TAT

50
Ala Trp Lys Arg Met Glu
GCC TGG AAG AGG ATG GAG GTGAGTTCCTTTTTTTTTTCCITTTTGGAGAATCTCAT

TGGGAGCCCTGATTTTGGATGAAGGGAGAAATGATCGGGGAAAGGTAAATGGAGCAGCAGAGATGAGGCT

GCCTGGGGCAGAGGCTCACGTCCTATAATCCCAGGCTGAGATGGCCGAGATGGGAGAATTGCTTGAGCCCT

GGAGTTCAGACCAACCTAGGCAGCATAGTGAGATCCCCCATCTCTACAAACATTTAAAAAATTAGTCAG

GTGAAGTGGTGCATGGTGGTAGTCCCAGATATTGGGAAGGCTGAGCGGGAGGATCGCTTGAGCCCCAGGAA

TTTGAGGCTGCAGTGAGCTGTGATCACACCACCTGCACCTCCAGCCTCAGTGACAGAGTGAGGCCCTGTCTCA



AAAAGAAAAGAAAAAATAATGAGGGCTGTATGGAATACATTTCATTTCCTCACTCACTCACT

CACTCATTTCATTTCATTTCATCAACAAGTCTTATTGCATACCTTCTGTTTGCTCAGCTTGGTGCTTGG

GGCTGCTGAGGGGCAGGAGGGGTGACATGGGTGAGCTCGACTCCAGAGTCCACTCCCTGTAG

56 60 70
Val Gly Gln Gln Ala Val Glu Val Trp Gln Gly Leu Ala Leu Leu Ser Glu Ala
GTC GGG CAG CAG GCC GTA GAA GTC TGG TGG CAG GGC CTG GCC CTG CTG TCG GAA GCT

Val Leu Arg Gly Gln Ala Leu Leu Val Asn Ser Ser Gln Pro Trp Glu Pro Leu
GTC CTG CCG GGC CAG GCC CTG TTG GTC AAC TCT TCC CAG CCG TGG GAG CCC CTG

Gln Leu His Val Asp Lys Ala Val Ser Gly Leu Arg Ser Leu Thr Thr Leu Leu
CAG CTG CAT GTG GAT AAA GCC GTC AGT GGC CTT CGC AGC CTC ACC ACT CTG CTT

110 115
Arg Ala Leu Gly Ala Gln
CGG GCT CTG GGA GCC CAG GTGAGTAGGAGCGGACACTTCTGCTTGCCCTTCTGTAAAGGGGA

GAAGGGCTTGTCTAAGGAGTACAGGAACGTCCGTATTCTCTCCCTTCTGTGGCACTGCAGGACCTCCT

116 120
Lys Glu Ala Ile Ser Pro Pro Asp Ala Ala Ser Ala Ala
AAG GAA GCC ATC TCC CCT CCA GAT GCG GCC TCA GCT GCT
GTTTCTCCTTGGCAG



130 Pro Leu Arg Thr Ile Thr Ala Asp Thr Phe Arg Lys Leu Phe Arg Val Tyr Ser
 CCA CTC CGA ACA ATC ACT GCT GAC ACT TTC CGC AAA CTC TTC CGA GTC TAC TCC
 140
 Asn Phe Leu Arg Gly Lys Leu Lys Leu Tyr Thr Gly Glu Ala Cys Arg Thr Gly
 AAT TTC CTC CGG GGA AAG CTG AAG CTG TAC ACA GGG GAG GCC TGC AGG ACA GGG
 150
 166 ASP Arg OP
 GAC AGA TGA CCAGGTGTGTCCACCTGGGCATATCCACCACCTCCCTCACCACAACTTGCTTGGCCACA
 CCCTCCCCCGCCACTCCTGAACCCCGTCGAGGGGCTCTCAGCTCAGCGCCAGCCTGTCCCATGGACACTCC
 AGTGCCAGCAATGACATCTCAGGGGCCAGAGGAACCTGTCCAGAGAGCAACTCTGAGATCTAAGGATGTCAC
 AGGGCCAACTTGAAGGGCCCAGAGCAGGAAGCATTCAGAGAGCAGCTTTAAACTCAGGGACAGAGCCATGC
 TGGGAAGACGCCTGAGCTCACTCGGCACCCTGCAAAATTTGATGCCAGGACACGCTTTGGAGGCGATTAC
 CTGTTTTCGCACCTACCATCAGGGACAGGATGACCTGGAGAACTTAGGTGGCAAGCTGTGACTTCTCCAGG
 TCTCAGGGGCATGGGCACTCCCCTTGGTGCCAAGAGCCCCCTTGACACCCGGGTGGTGGGAACCATGAAGAC
 AXGATXGGGGCTGGCCTCTGGCTCTCATGGGGTCCAAGTTTTGTGTATTCACAACCTATTGACAGACTGAA
 ACACAATATGAC



ou uma sua variante alélica qualquer existente na natureza.

8.- Processo de acordo com a reivindicação 1, caracteriza-
do pelo facto de o polipeptido obtido possuir parte ou toda a con-
formação estrutural primária de eritropoietina de macaco indicada
a seguir:

Sau3A
GATCCGCGCCCCCTGGACAGCCGCCCTCTCCATCCAGGCCCGTGGGGCTGGCCCCTGCCC
CGCTGAACCTCCCGGGATGAGGACTCCCGGTGGTTCACCGCGCGCCTAGGTGCTGAG

-27	-20
Met Gly Val His Glu Cys Pro Ala Trp	
GGACCCCGCGCAGCGCGGAGATG GGG GTG CAC GAA TGT CCT GCC TGG	

-10	
Leu Trp Leu Leu Ser Leu Val Ser Leu Pro Leu Gly Leu Pro	
CTG TGG CTT CTC CTG TCT CTC CTC GTG TCG CTC CCT CTG GGC CTC CCA	

-1	+1	10
Val Pro Gly Ala Pro Pro Arg Leu Ile Cys Asp Ser Arg Val Leu		
GTC CCG GGC GCC CCA CCA CGC CTC ATC TGT GAC AGC CGA GTC CTG		

20	*
Glu Arg Tyr Leu Leu Glu Ala Lys Glu Ala Glu Asn Val Thr Met	
GAG AGG TAC CTC TTG GAG GCC AAG GAG GCC GAG AAT GTC ACG ATG	

30	*	40
Gly Cys Ser Glu Ser Cys Ser Leu Asn Glu Asn Ile Thr Val Pro		
GGC TGT TCC GAA AGC TGC AGC TTG AAT GAG AAT ATC ACC GTC CCA		

...



50
 Asp Thr Lys Val Asn Phe Tyr Ala Trp Lys Arg Met Glu Val Gly
 GAC ACC AAA GTT AAC TTC TAT GCC TGG AGG AGG ATG GAG GTC GGG

 60
 Gln Gln Ala Val Glu Val Trp Gln Gln Gly Leu Ala Leu Leu Ser Glu
 CAG CAG GCT GTA GAA GTC TTC TGG CAG GCC CTG GCC CTG CTC TCA GAA

 70
 Ala Val Leu Arg Gly Gln Ala Val Leu Ala Asn Ser Ser Gln Pro
 GCT GTC CTG CTG CGG GGC CAG GCC GTG TTG GCC AAC TCT TCC CAG CCT

 80
 *
 Phe Glu Pro Leu Gln Leu His Met Asp Lys Ala Ile Ser Gly Leu
 TTC GAG CCC CTG CAG CTG CAC CAC ATG GAT AAA GCC ATC AGT GGC CTT

 90
 Arg Ser Ile Thr Thr Thr Leu Leu Arg Ala Leu Gly Ala Gln Glu Ala
 CGC AGC ATC ACC ACT ACT CTG CTT CGG GCG GCG CTG GGA GCC CAG GAA GCC

 100
 110
 Ile Ser Leu Pro Asp Ala Ala Ser Ala Ala Pro Leu Arg Thr Ile
 ATC TCC CTC CCA GAT GCG GCC TCG GCT GCT CCA CTC CGA ACC ATC

 120
 130
 Thr Ala Asp Thr Phe Cys Lys Leu Phe Arg Val Tyr Ser Asn Phe
 ACT GCT GAC ACT TTC TGC AAA CTC TTC CGA GTC TAC TCC AAT TTC

 140

150 160
 Leu Arg Gly Lys Leu Lys Leu Tyr Thr Gly Glu Ala Cys Arg Arg
 CTC CGG GGA AAG CTG AAG CTG TAC ACG GGG GAG GCC TGC AGG AGA

 165
 Gly Asp Arg OP
 GGG GAC AGA TGA CCAGGTGGTCCAGCTGGGCACATCCACCACCTCCCTCACCACA
 CTGCCTGTGCCACACCCCTCCCTCACCCTCCCGAACCCCATCGAGGGGCTCTCAGCTAAG

 CGCCAGCCTGTCCCATGGACACTCCAGTGCCAGCAATGACATCTCAGGGGCCAGAGGAAC
 TGTCCAGAGCACAACTCTGAGATCTAAGGATGTCCGAGGGCCAACTTGAGGGCCAGAGC
 AGGAAGCATTGAGAGAGCAGCTTTAACTCAGGAGCAGAGACAATGCAGGGAAACACCT
 GAGCTCACTCGGCCACCTGCAAAATTGTATGCAGGACACGCTTTGGAGGCAATTTACCTG
 TTTTTCACCTACCATCAGGGACAGGATGACTGGAGAACTTAGGTGGCAAGCTGTGACTT
 CTCAGGCCCTACGGGCACTCCCTTGGTGGCAAGAGCCCCCTTGACACTGAGAGAATATT
 TTGCAATCTGCAGCAGGAAAAATTACGGACAGGTTTTGGAGGTGGAGGGTACTTGACAG
 GTGTGTGGGAAGCAGGGCGGTAGGGGTGGAGCTGGGATCGGAGTGAGAACCGTGAAGAC
 AGGATGGGGGCTGGCCCTCTGGTTCCTGTTGGGTCCAAGCTT

HindIII





ou uma sua variante alélica qualquer existente na natureza.

9.- Processo de acordo com a reivindicação 1, caracterizado pelo facto de o polipeptido obtido possuir as propriedades imunológicas da eritropoietina existente na natureza.

10.- Processo de acordo com a reivindicação 1, caracterizado pelo facto de o polipeptido obtido possuir a actividade biológica "in vivo" da eritropoietina existente na natureza.

11.- Processo de acordo com a reivindicação 1, caracterizado pelo facto de o polipeptido obtido possuir a actividade biológica "in vitro" da eritropoietina existente na natureza.

12.- Processo de acordo com a reivindicação 1, caracterizado ainda pelo facto de se associar o polipeptido covalentemente a uma substância de marca detectável.

13.- Processo de acordo com a reivindicação 12, caracterizado pelo facto de a citada marca detectável ser uma marca radioactiva.

14.- Processo para a consecução da expressão em uma célula hospedeira procariótica ou eucariótica de um produto polipeptídico que tem pelo menos uma parte da conformação estrutural primária e uma ou mais das propriedades biológicas da eritropoietina existente na natureza, caracterizado pelo facto de se utilizar uma sequência de ADN escolhida entre:

(a) as sequências de ADN indicadas a seguir:



Transferência do CADN de EPO do macaco

Sau3A
GATCCGGGGCCCCCTGGACAGCCGCCCTCTCTCTCCAGGCCCGTGGGGCTGGCCCCTGCCCC
CGCTGAACATCCCGGGATGAGGACTCCCGGTGTGCTCACCGCGCGCCCTAGGTCGCTGAG

-27	Met	Gly	Val	His	Glu	Cys	Pro	Ala	Trp
	ATG	GGG	GTC	CAC	GAA	TGT	CCT	GCC	TGG

Leu Trp Leu Leu Ser Leu Val Ser Leu Pro Leu Gly Leu Pro
CTG TGG CTT CTC CTG TCT CTC GTG TCG CTC CCT CTG GGC CTC CCA

val	Pro	Gly	GCC	GCC	Ala	Pro	Pro	Arg	Leu	Ile	Cys	Asp	Ser	Arg	Val	Leu
GTC	CCG	GGC	GGC	GGC	GGC	CCA	CCA	CGC	CTC	ATC	TGT	GAC	AGC	CGA	GTC	CTG

Glu	Arg	Tyr	Leu	Leu	Glu	Ala	Lys	Glu	Ala	Glu	Asn	Val	Thr	Met
GAG	AGG	TAC	CTC	TTG	GAG	GCC	AAG	GAG	GCC	GAG	AAT	GTC	ACG	ATG

[illegible]



Asp Thr Lys Val Asn Phe Tyr Ala Trp Lys Arg Met Glu Val Gly
 GAC ACC AAA GTT AAC TTC TAT GCC TGG TGG AAG AGG ATG GAG GAG GTC GGG
 50
 Gln Gln Ala Val Gln Val Trp Gln Gln Gly Leu Ala Leu Leu Ser Glu
 CAG CAG GCT GCT GTA GAA GTC GTC TGG CAG GGC CTG GCC CTC TCA GAA
 60
 Ala Val Leu Arg Gly Gln Ala Val Leu Ala Asn Ser Ser Gln Pro
 GCT GTC CTG CTG CGG GGC CAG GCC GTG TGG TGG GCC AAC TCT TCC CAG CCT
 80
 Phe Glu Pro Leu Gln Leu His Met Asp Lys Ala Ile Ser Gly Leu
 TTC GAG CCC CTG CAG CTG CAG CTG CAC ATG GAT AAA GCC ATC AGT GGC CTT
 90
 Arg Ser Ile Thr Thr Thr Leu Leu Arg Ala Leu Gly Ala Gln Gln Ala
 CGC AGC ATC ACC ACT ACT CTG CTG CTT CGG GCG GCG CTG GGA GCC CAG GAA GCC
 110
 Ile Ser Leu Pro Asp Ala Ala Ser Ala Ala Pro Leu Arg Thr Ile
 ATC TCC CTC CTC CCA GAT GCG GCC TCG TCG GCT GCT CCA CTC CGA ACC ATC
 120
 Thr Ala Asp Thr Phe Cys Lys Leu Phe Arg Val Tyr Ser Asn Phe
 ACT GCT GAC ACT TTC TGC AAA CTC TTC TTC CGA GTC TAC TCC AAT TTC
 140

AAGCTTC TGGGCTTCCAGACCCAGCTACTTTGCGGAAC TCAGCAACCCAGGCATCTCTGAGTCTCCGCCCA
AGACCGGGATGCCCCCAGGGGAGGTGTCCGGAGCCCCAGCCTTTCCCAGATAGCACGGCTCCGCCAGTCCC
AAGGGTGGCAACCGGCTGCAC TCCCC TCCCGGACCCAGGGCCCCGGGAGCAGCCCCCATGACCCACACGC
ACGTCTGCAGAGCCCCGCTACGCCCCCGGAGCCTCAACCCAGGCGTCTGCCCTGCTCTGACCCCCGG
GTGGCCCCTACCCCTGGCGACCCCTCACGCACACAGCCTCTCCCCCACCCCAACCGGCGACGACACATG
CAGATAACAGCCCGACCCCGGCCAGAGCCGXAGAGTCCCTGGGCCACCCCGCGCTCGCCTGCCGCTG
CGCCGCACCGCGCTGTCTCTCCCGAGCCCGGACCGGGGCCACCGCGCCXGCTCTGCTCCGACACCGCGCCC
CTTGGACAGCCGCCCTCTCTCTAGGCCCCGTGGGGCTGGCCCTGCACCGCCGAGCTTCCCGGATGAGGXX
CCCGGTGACCGCGGCCCAAGTCTGTAGGGACCCCGGCCAAGCGCGGAG
GTGAGTACTCGCGGGCTGGGGCTCCCGCGGCCGGGTTCCTGT TTGAGCGGGGATTAGCGCCCCGGCT

-27
Met Gly Val His
ATG GGG GTG CAC G





TGGTGGCCCCAAACCATACCTGAACCTAGGCAAGGAGCAAGCCAGCAGATCCTACGCCTGTGGGCCAGGG

27 30
Thr Gly Cys Ala Glu
ACG GGC TGT GCT GAA

CCAGAGCCTTCAGGGACCCCTTGACTCCCCGGGCTGTGTGCATTTCAG

* 40
His Cys Ser Leu Asn Glu Asn Ile Thr Val Pro Asp Thr Lys Val Asn Phe Tyr
CAC TGC AGC TTG AAT GAG AAT ATC ACT GTC CCA GAC ACC AAA GTT AAT TTC TAT

50 55
Ala Trp Lys Arg Met Glu
GCC TGG AAG AGG ATG GAG GTGAGTTCCTTTTTTTTTTTTCCCTTTTGGAGAATCTCAT

TGCGAGCCTGATTTTGGATGAAAGGGAGAATGATCGGGGAAAGGTAAATGGAGCAGCAGAGATGAGGCT

GCCTGGGGCGCAGAGGCTACGTCTATAATCCCAGGCTGAGATGGCCGAGATGGGAGAAATTGCTTGAGCCCT

GGAGTTTCAGACCAACCTAGGCAGCATAGTGAGATCCCCCATCTCTACAAACATTTAAAAAATTAGTCAG

GTGAAGTGGTGATGGTAGTCCCAGATATTTGGAAGGCTGAGGCGGGAGGATCGCTTGAGCCCCAGGAA

TTTGAGGCTGCAGTGAGCTGTGATCACACCACCTGCACCTCCAGCCTCAGTGACAGAGTGAGGCCCTGTCTCA



130
 Pro Leu Arg Thr Ile Thr Ala Asp Thr Phe Arg Lys Leu Phe Arg Val Tyr Ser
 CCA CTC CGA ACA ATC ACT GCT GAC ACT TTC CGC AAA CTC TTC CGA GTC TAC TCC
 140
 150
 Asn Phe Leu Arg Gly Lys Leu Lys Leu Tyr Thr Gly Glu Ala Cys Arg Thr Gly
 AAT TTC CTC CGG GGA AAG CTG AAG CTG TAC ACA GGG GAG GCC TGC AGG ACA GGG
 160
 Asp Arg OP
 GAC AGA TGA CCAGGTGTGTCCACCTGGGCATATCCACCACCTCCCTCACCAACATTGCTTGTGCCACA
 CCCTCCCCCGCCACTCCTGAACCCCGTCGAGGGGCTCTCAGCTCAGCGCCAGCCTGTCCCATGGACACTCC
 AGTGCCAGCAATGACATCTCAGGGGCCAGAGGAAGTGTCCAGAGAGCAACTCTGAGATCTAAGGATGTAC
 AGGGCCAAC TTGAAGGGGCCAGAGCAGGAAGCATTTCAGAGAGCAGCTTTAAACTCAGGGACAGAGCCATGC
 TGGGAAGACGCCTGAGCTCACTCGGCACCCTGCAAAATTTGATGCCAGGACACGCTTTGGAGGCGATTTAC
 CTGTTTTCGCACCTACCATCAGGGACAGGATGACCTGGAGAACTTAGGTGGCAAGCTGTGACTTCTCCAGG
 TCTCACGGGCATGGGCACTCCCTTGGTGGCAAGAGCCCCCTTGACACCGGGGTGGTGGGAACCATGAAGAC
 AXGATXGGGGCTGGCCTCTGGCTCTCATGGGGTCCAAGTTTTGTGTATTCTCAACCTATTGACAGACTGAA
 ACACAATATGAC

ou as suas cadeias complementares;

(b) sequências de ADN que hibridam para as sequências de ADN definidas em (a) ou seus fragmentos;

e (c) sequências de ADN que, se não fosse a degeneração do código genético, hibridariam para as sequências de ADN definidas em (a) e (b).

15.- Processo para a transformação ou modificação de uma célula de hospedeiro procariótica ou eucariótica, caracterizado pelo facto de se utilizar uma sequência de ADN de acordo com a reivindicação 14 por uma maneira tal que permita, à célula hospedeira,



deira, exprimir o citado produto polipeptídico.

16.- Processo para a preparação de um produto polipeptídico, caracterizado pelo facto de se exprimir uma sequência de ADN de acordo com a reivindicação 14 em um hospedeiro procariótico ou eucariótico.

17.- Processo para a preparação de uma sequência de ADN, caracterizado pelo facto de se purificar e isolar codificando a expressão, em um hospedeiro procariótico, ou eucariótico, um polipeptido que tem parte ou toda a conformação estrutural primária e uma ou mais das propriedades biológicas da eritropoietina.

18.- Processo de acordo com a reivindicação 17, caracterizado pelo facto de se obter uma sequência de cADN.

19.- Processo para a preparação de eritropoietina da espécie de macaco, caracterizado pelo facto de codificar a sequência de ADN de acordo com a reivindicação 18.

20.- Processo de acordo com a reivindicação 19, caracterizado pelo facto de a sequência de ADN utilizada incluir a região de codificação de proteína indicada a seguir:

Sau3A
GATCCGGCCCCCTGGACAGCCGCCCTCTCTCCAGGCCCGTGGGGCTGGCCCTGCCCC

CGCTGAACCTCCCGGGAATGAGGACTCCCGGTGTGGTCACCGCGCCTAGGTCGCTGAG

GGACCCCGGCCAGCGCGGAGATG GGG GTG CAC GAA TGT CCT GCC TGG

Leu Trp Leu Leu Leu Ser Leu Val Ser Leu Pro Leu Gly Leu Pro
CTG TGG CTT CTC CTC CTC TCT CTC CTC GTG TCG TCG CTC CCT CTG GGC CTC CCA

Val Pro Gly Ala Pro Pro Arg Leu Ile Cys Asp Ser Arg Val Leu
GTC CCG GGC GCC GCC CCA CCA CCG CTC CTC ATC ATC TGT GAC AGC CGA GTC CTG

Glu Arg Tyr Leu Leu Glu Ala Lys Glu Ala Glu Asn Val Thr Met
GAG AGG TAC CTC CTC TTT GAG GAG GCC AAG GAG GAG GCC GAG AAT GTC ACG ATG

Gly Cys Ser Glu Ser Cys Ser Leu Asn Glu Asn Ile Thr Val Pro
GGC TGT TCC GCC GAA AGC TGC TGC AGC TTG AAT GAG AAT ATC ACC GTC CCA



50
 Asp Thr Lys Val Asn Phe Tyr Ala Trp Lys Arg Met Glu Val Gly
 GAC ACC AAA GTT AAC TTC TAT GCC TGG AGG AGG ATG GAG GTC GGC

 60
 Gln Gln Ala Val Glu Val Trp Gln Gln Gly Leu Ala Leu Leu Ser Glu
 CAG CAG GCT GTA GAA GTC GTC TGG CAG CAG GGC CTG GCC CTG CTC TCA GAA

 70
 Ala Val Leu Arg Gly Gln Ala Val Leu Ala Asn Ser Ser Gln Pro
 GCT GTC CTG CGG GGC GGC CAG GGC GTG TTG GCC AAC TCT TCC CAG CCT

 80
 Phe Glu Pro Leu Gln Leu His Met Asp Lys Ala Ile Ser Gly Leu
 TTC GAG CCC CTG CAG CTG CAG ATG ATG GAT AAA GCC ATC AGT GGC CTT

 90
 Arg Ser Ile Thr Thr Thr Leu Leu Arg Ala Leu Gly Ala Gln Glu Ala
 CGC AGC ATC ACC ACT ACT CTG CTT CGG GCG GCG CTG GGA GCC CAG GAA GCC

 100
 Ile Ser Leu Pro Asp Ala Ser Ala Ala Pro Leu Arg Arg Thr Ile
 ATC TCC CTC CCA GAT GAT GCG GCC TCG GCT GCT CCA CTC CGA ACC ATC

 110
 Thr Ala Asp Thr Phe Cys Lys Leu Phe Arg Val Tyr Ser Asn Phe
 ACT GCT GAC ACT TTC TGC AAA CTC TTC TTC CGA GTC TAC TCC AAT TTC

 120
 Thr Ala Asp Thr Phe Cys Lys Leu Phe Arg Val Tyr Ser Asn Phe
 ACT GCT GAC ACT TTC TGC AAA CTC TTC TTC CGA GTC TAC TCC AAT TTC

 130
 Thr Ala Asp Thr Phe Cys Lys Leu Phe Arg Val Tyr Ser Asn Phe
 ACT GCT GAC ACT TTC TGC AAA CTC TTC TTC CGA GTC TAC TCC AAT TTC

 140
 Thr Ala Asp Thr Phe Cys Lys Leu Phe Arg Val Tyr Ser Asn Phe
 ACT GCT GAC ACT TTC TGC AAA CTC TTC TTC CGA GTC TAC TCC AAT TTC



150
 Leu Arg Gly Lys Leu Lys Leu Tyr Thr Gly Glu Ala Cys Arg Arg
 CTC CGG GGA AAG CTG AAG CTG TAC ACG GGG GAG GCC TGC AGG AGA

160
 Gly Asp Arg OP
 GGG GAC AGA TGA CCAGGTGCGTGCCAGCTGGGCACATCCACCACCTCCCTCACCAACA

165
 CTGCCTGTGCCACACCCTCCCTCACCACTCCCGAACCCCATCGAGGGGCTCTCAGCTAAG

CGCCAGCCTGTCCCATGGACACTCCAGTGCCAGCAATGACATCTCAGGGGCCAGAGGAAC
 TGTCAGAGCACAACCTCTGAGATCTAAGGATGTCGCAGGGCCAACCTTGAGGGCCCAGAGC
 AGGAAGCATTGAGAGAGCAGCTTTAACTCAGGAGCAGAGACAATGCAGGGAAAACACCT
 GAGCTCACTCGGCCACCTGCAAAATTTGATGCAGGACACGCTTTGGAGGCAATTTACCTG
 TTTTGCACCTACCATCAGGGACAGGATGACTGGAGAACTTAGGTGGCAAGCTGTGACTT
 CTCAAGGCCTCAGGGGCACTCCCTTGGTGGCAAGAGCCCCCTTGACACTGAGAGAATATT
 TTGCAATCTGCAGCAGGAAAAATTACGGACAGGTTTTGGAGGTTGGAGGGTACTTGACAG
 GTGTGTGGGGAAGCAGGGCGGTAGGGGTGGAGCTGGGATGCGAGTGAGAACCGTGAAGAC
 AGGATGGGGGCTGGCCTCTGGTTCTCGTGGGGTCCAAGCTT
HindIII

21.- Processo de acordo com a reivindicação 17, caracterizado pelo facto de se obter uma sequência de ADN genómica.

22.- Processo para a preparação de eritropoietina da espécie humana, caracterizado pelo facto de codificar a sequência de ADN obtida de acordo com a reivindicação 21.

23.- Processo de acordo com a reivindicação 22, caracterizado pelo facto de a sequência de ADN utilizada incluir a região de codificação de proteína indicada a seguir:



AAGCTTC TGGGCTTCCAGACCCAGCTACTTIGCGGAAC TCAGCAACCCAGGCATCTCTGAGTCTCCGCCCA
 AGACCGGGATGCCCCCAGGGGAGGTGTCCGGGAGCCCAGCCTTTCCAGATAGCACGCTCCGCCAGTCCC
 AAGGGTGCACAACCGGCTGCACTCCCCCTCCCGGACCCAGGGCCCGGAGCAGCCCCCATGACCCACACGC
 ACGTCTGCAGCAGCCCCGCTCACGCCCCCGGAGCCTCAACCCAGGGGTCTTGCCCCCTGCTCTGACCCCGG
 GTGGCCCCTACCCCTGGCGACCCCTCACGCACACAGCCTCTCCCCCACCCCGCGCACACACATG
 CAGATAACAGCCCCGACCCCGGCAGAGCCGXAGAGTCCC TGGGCCACCCCGGCCGCTCGCCTGCGCGTG
 CGCCGACCGGCTGTCTCCCGGAGCCGACCGGGGCCACCGCGCCXGCTCTGCTCCGACACCGCGCCC
 CTTGGACAGCGCCCTCTCCCTC TAGGCCCGTGGGGCTGGCCCTGCACCGCCGAGCTTCCCGGATGAGGX
 CCGGGTGACCGGGCGCCCCAAGTCGCTGAGGGACCCCGCCAAAGCGGGAG
 -27 Met Gly Val His
 -24 ATG GGG GTG CAC G
 GTGAGTACTCGCGGGCTGGGGGCTCCCGGGCGCGGGTTCCTGTTGAGCGGGGATTTAGCGCCCCCGGCT



ATTGCCAAGAGGTGGCTGGTTCAAGGACCGCGACTTGTC AAGACCCGGAAGGGGGGGGTGGG
 GCAGCCTCCACGTGCCGCGGGGACTTGGGGAGTTC TTGGGATGGCAAAACCTGGCCTGTTGAGGGGCA
 CAGTTTGGGGTTGGGAGGAGGTTTGGGGTTC TGCTGTGCAGTTGTGTCGTGTCAGTGTCTCG[I.S.]
 TTGCACACGCACAGATCAATAAGCCAGAGGCGACCTGAGTGTGTCATGGTTGGGACAGGAAGGACGAG
 CTGGGGCAGAGACGTGGGGATGAAGGAAGCTGTCTTCCACAGCCACCCCTTCTCCCCCGCCTGACTCT
 CAGCCTGGCTATCIGTTCTAG -23 -20
 Glu Cys Pro Ala Trp Leu Trp Leu Leu Ser Leu
 AA TGT CCT GCC TGG CTG TGG CTT CTC CTG TCC CTG
 -10 -1 +1
 Leu Ser Leu Pro Leu Gly Leu Pro Val Leu Leu Gly Ala Pro Pro Arg Leu Ile Cys
 CTG TCG CTC CCT CTG GGC CTC CCA GTC CTG GGC GCC CCA CCA CGC CTC ATC TGT
 10 20
 Asp Ser Arg Val Leu Glu Arg Tyr Leu Leu Glu Ala Lys Glu Ala Glu Asn Ile
 GAC AGC CGA GTC CTC GAG AGG TAC CTC TTG GAG GCC AAG GAG GCC GAG AAT ATC
 26
 Thr
 ACG GTGAGACCCCTTCCCCAGCACATTCCACAGAACTCAGGCTCAGGGCTTCAGGGAACCTCTCCAGAT
 CCAGGAACCTGGCACCTGGTTTGGGGTGGAGTTGGGAAGCTAGACACTGCCCCCTACATAAGAATAAGTC



TGGTGGCCCCAAACCATACCTGAACTAGGCAAGGAGCAAGCCAGCATCTACGCCGTGTGGCCCCAGGG

27	Thr	Gly	Cys	Ala	30	Glu
	ACG	GGC	TGT	GCT		GAA

CCAGAGCCCTCAGGGACCCCTTGACTCCCCGGCTGTGTGCATTTTCAG

His	Cys	Ser	Leu	Asn	Glu	Asn	Ile	Thr	Val	Pro	Asp	Thr	Lys	Val	Asn	Phe	Tyr
CAC	TGC	AGC	TTG	AAT	GAG	AAT	ATC	ACT	GTC	CCA	GAC	ACC	AAA	GTT	AAT	TTC	TAT

50 Ala Trp Lys Arg Met Glu
55 GCC TGG AAG AGG ATG GAG GTGAGTTCCTTTTCTTTTGGAGAACTCTCATTT

TGGAGCCTGATTTGGATGAAAGGGAGAAATGATCGGGGAAAGGTAAATGGAGCAGCAGAGATGAGGCT

GCCTGGGCGCAGAGGCTCAGCTCTATAATCCAGGCTGAGATGCCGAGATGGGAGAAATTGCTTGAGCCCT

GGAGTTCAGACCAACCTAGGCAGCATAGTGAGATCCCCCATCTCTACAAACATTTAAAAAATTAGTCAG

GTGAAGTGGTGCATGGTGGTAGTCCCAGATATTGGAAAGCTGAGCGGGAGGATCGCTTGAGCCCCAGGAA

TTTGAGGCTGCAGTGAGCTGTGATCACACCCTGCACCTCAGTGACAGAGTGAGGCCCTGTCTCA



AAAAAGAAAGAAAAAATAATGAGGGCTGTATGGAATACATTTCATTTCACATCCTCCT

CACTCATTCATTTCATTTCATCAACAAGTCTTATTGCATACCTTCCTGCTCAGCTTGGTGGTGG

GGCTGCTGAGGGCAGGAGGGGACATGGGTGAGCTCGACTCCAGAGTCCACTCCCTGTAG

56 60 70
Val Gly Gln Gln Ala Val Glu Val Trp Gln Gly Leu Ala Leu Leu Ser Glu Ala
GTC GGG CAG CAG GCC GTA GAA GTC TGG CAG GGC CTG GCC CTG CTG TCG GAA GCT

Val Leu Arg Gly Gln Ala Ala Leu Leu Val Asn Ser Ser Gln Pro Trp Glu Pro Leu
GTC CTG CGG GGC CAG GCC CTG TTG GTC AAC TCT TCC CAG CCG TGG GAG CCC CTG

Gln Leu His Val Asp Lys Ala Val Ser Gly Leu Arg Ser Leu Thr Thr Leu Leu
CAG CTG CAT GTG GAT AAA GCC GTC AGT GGC CTT CGC AGC CTC ACC ACT CTG CTT

110 115
Arg Ala Leu Gly Ala Gln
CGG GCT CTG GGA GCC CAG GTGAGTAGGAGCGGACACTTCTGCTTGCCCTTTCTGTAAAGAGGGGA

GAAGGGTCTTGCTAAGGAGTACAGGAACGTCCGTATTCCTTCCCTTTCTGTGGCACTGCAGCGACCTCCT

116 120
Lys Glu Ala Ile Ser Pro Pro Asp Ala Ala Ser Ala Ala
AAG GAA GCC ATC TCC CCT CCA GAT GCC GCC TCA GCT GCT
GTTTCTCCTTGGCAG



27.- Processo de acordo com a reivindicação 26, caracterizado pelo facto de a sequência de ADN manufacturada obtida incluir a região de codificação da proteína indicada a seguir:

GENE ECEPO

```

                                -1  1
                                MetAla
XbaI
CTAG AAACCATGAG GGTAATAAAA TAATGGCTCC GCCGCGTCTG
      TTTGGTACTC CCATTATTTT ATTACCGAGG CGGCGCAGAC

ATCTGCGACT CGAGAGTTCT GGAACGTTAC CTGCTGGAAG CTAAAGAAGC
TAGACGCTGA GCTCTCAAGA CCTTGCAATG GACGACCTTC GATTTCTTCG

TGAAAACATC ACCACTGGTT GTGCTGAACA CTGTTCTTTG AACGAAAACA
ACTTTTGTAG TGGTGACCAA CACGACTTGT GACAAGAAAC TTGCTTTTGT

TTACGGTACC AGACACCAAG GTTAACTTCT ACGCTTGGAAC ACGTATGGAA
AATGCCATGG TCTGTGGTTC CAATTGAAGA TGCGAACCTT TGCATACCTT

GTTGGTCAAC AAGCAGTTGA AGTTTGGCAG GGTCTGGCAC TGCTGAGCGA
CAACCAGTTG TTCGTCAACT TCAAACCGTC CCAGACCGTG ACGACTCGCT

GGCTGTACTG CGTGGCCAGG CACTGCTGGT AAACCTCCTCT CAGCCGTGGG
CCGACATGAC GCACCGGTCC GTGACGACCA TTTGAGGAGA GTCGGCACCC

AACCGCTGCA GCTGCATGTT GACAAAGCAG TATCTGGCCT GAGATCTCTG
TTGGCGACGT CGACGTAÇAA CTGTTTCGTC ATAGACCGGA CTCTAGAGAC

ACTACTCTGC TCGTGCTCTT GGGTGCACAG AAAGAGGCTA TCTCTCCGCC
TGATGAGACG ACGCACGAGA CCCACGTGTC TTTCTCCGAT AGAGAGGCGG

GGATGCTGCA TCTGCTGCAC CGCTGCGTAC CATCACTGCT GATACCTTCC
CCTACGACGT AGACGACGTG GCGACGCATG GTAGTGACGA CTATGGAAGG

GCAAACTGTT TCGTGTATAC TCTAACTTCC TCGTGGTAA ACTGAAACTG
CGTTTGACAA AGCACATATG AGATTGAAGG ACGCACCATT TGACTTTGAC ...

TATACTGGCG AAGCATGCCG TACTGGTGAC CGCTAATAG
ATATGACCGC TTCGTACGGC ATGACCACTG GCGATTATCA GCT

```

SalI



28.- Processo de acordo com a reivindicação 24, caracterizado pelo facto de a sequência de ADN manufacturada obtida incluir um ou mais códãos preferidos para expressão em células de levedura.

29.- Processo de acordo com a reivindicação 28, caracterizado pelo facto de a sequência de ADN manufacturada obtida codificar para expressão de eritropoietina da espécie humana.

30.- Processo de acordo com a reivindicação 29, caracterizado pelo facto de a sequência de ADN manufacturada obtida incluir a região de codificação de proteína indicada a seguir:



GENE ECEPO

-1 +1

HindIII ArgAla

AGCTTGGATA AAAGAGCTCC ACCAAGATTG ATCTGTGACT CGAGAGTTTT
 ACCTAT TTTCTCGAGG TGGTTCTAAC TAGACACTGA GCTCTCAAAA

GGAAAGATAC TTGTTGGAAG CTAAGAAGC TGAAAACATC ACCACTGGTT
 CCTTTCTATG AACAACCTTC GATTTCTTCG ACTTTTGTAG TGGTGACCAA

GTGCTGAACA CTGTTCTTTG AACGAAAACA TTACGGTACC AGACACCAAG
 CACGACTTGT GACAAGAAAC TTGCTTTTGT AATGCCATGG TCTGTGGTTC

GTAACTTCT ACGCTTGGAA ACGTATGGAA GTTGGTCAAC AAGCTGTTGA
 CAATTGAAGA TGCGAACCTT TGCATACCTT CAACCAGTTG TTCGACAACT

AGTTTGGCAA GGTTCGGCCT TGTTATCTGA AGCTGTTTTG AGAGGTCAAG
 TCAAACCGTT CCAAACCGGA ACAATAGACT TCGACAAAAC TCTCCAGTTC

CCTTGTTGGT TAACTCTTCT CAACCATGGG AACCATTGCA ATTGCACGTC
 GGAACAACCA ATTGAGAAGA GTTGGTACCC TTGGTAACGT TAACGTGCAG

GATAAAGCCG TCTCTGGTTT GAGATCTTTG ACTACTTTGT TGAGAGCTTT
 CTATTTCCGC AGAGACCAAA CTCTAGAAAC TGATGAAACA ACTCTCGAAA

GGGTGCTCAA AAGGAAGCCA TTTCCCCACC AGACGCTGCT TCTGCCGCTC
 CCCACGAGTT TTCCTTCGGT AAAGGGGTGG TCTGCGACGA AGACGGCGAG

CATTGAGAAC CATCACTGCT GATACCTTCA GAAAGTTATT CAGAGTTTAC
 GTAACCTTGT GTAGTGACGA CTATGGAAGT CTTTCAATAA GTCTCAAATG

TCCAACCTTCT TGAGAGGTAA ATTGAAGTTG TACACCGGTG AAGCCTGTAG
 AGGTTGAAGA ACTCTCCATT TAACTTCAAC ATGTGGCCAC TTCGGACATC

AACTGGTGAC AGATAAGCCC GACTGATAAC AACAGTGTAG
 TTGACCACTG TCTATTCGGG CTGACTATTG TTGTCACATC

SalI

ATGTAACAAA G
 TACATTGTTT CAGCT

...



31.- Processo de acordo com a reivindicação 17, caracterizado pelo facto de se associar a sequência de ADN obtida covalentemente com uma substância marcadora detectável.

32.- Processo de acordo com a reivindicação 31, caracterizado pelo facto de a marca detectável ser uma marca radioactiva.

33.- Processo de acordo com a reivindicação 31, caracterizado pelo facto de se obter uma sequência de ADN de cadeia simples.

34.- Processo de acordo com as reivindicações anteriores caracterizado pelo facto de a sequência de ADN obtida codificar para um fragmento de polipeptido ou um análogo de polipeptido de eritropoietina existente na natureza.

35.- Processo de acordo com as reivindicações anteriores caracterizado pelo facto de a sequência de ADN obtida codificar para $\lceil \text{Phe}^{15} \rceil$ hEPO, $\lceil \text{Phe}^{49} \rceil$ hEPO, $\lceil \text{Phe}^{145} \rceil$ hEPO, $\lceil \text{His}^7 \rceil$ hEPO, $\lceil \text{Asn}^2 \text{ des-Pro}^2 \text{ através de Ile}^6 \rceil$ hEPO, $\lceil \text{des-Thr}^{163} \text{ através de Arg}^{166} \rceil$ hEPO, ou $\lceil \triangle 27-55 \rceil$ hEPO.

36.- Processo de acordo com a reivindicação 34, caracterizado pelo facto de a sequência de ADN ser uma sequência manufacturada.

37.- Processo para a preparação de um plasmídeo circular biologicamente funcional ou vector de ADN viral, caracterizado pelo facto de se incluir uma sequência de ADN obtida de acordo com uma qualquer das reivindicações 14, 17, 34 ou 35.

38.- Processo para a preparação de uma célula hospedeira procariótica ou eucariótica, caracterizado pelo facto de se transformar ou modificar estavelmente uma célula com um vector de ADN



obtido de acordo com a reivindicação 37.

39.- Processo para a preparação de um produto polipeptídico da expressão em uma célula hospedeira procariótica ou eucariótica, caracterizado pelo facto de se utilizar uma sequência de ADN obtida de acordo com a reivindicação 17 ou 34.

40.- Processo para a produção de um polipeptido tendo parte ou toda a conformação estrutural primária e uma ou mais das propriedades biológicas da eritropoietina existente na natureza caracterizado pelo facto:

de se cultivar, sob condições nutritivas apropriadas, células hospedeiras procarióticas ou eucarióticas, transformadas ou modificadas com um vector de ADN obtido de acordo com a reivindicação 37, e de se isolar os produtos polipeptídicos desejados da expressão de sequências de ADN no vector citado.

41.- Processo de acordo com a reivindicação 40, caracterizado pelo facto de se obter um polipeptido sintético que possui parte ou toda a sequência de aminoácidos indicada antes na reivindicação 8 e uma ou mais das actividades biológicas "in vivo" ou "in vitro" da eritropoietina de macaco existente na natureza.

42.- Processo de acordo com a reivindicação 40, caracterizado pelo facto de se obter um polipeptido sintético que possui parte ou toda a sequência de aminoácidos indicada antes na reivindicação 7, diferente de uma sequência de restos inteiramente dentro da sequência numerados de 1 a 20, e uma propriedade biológica da eritropoietina humana existente na natureza.

43.- Processo de acordo com a reivindicação 40, caracte-



rizado pelo facto de se obter um polipeptido sintético que possui parte ou toda a conformação secundária de parte ou toda a sequência de aminoácidos indicada antes na reivindicação 7, diferente de uma sequência de restos inteiramente dentro da sequência numerados de 1 a 20, e uma propriedade biológica da eritropoietina humana existente na natureza.

44.- Processo de acordo com a reivindicação 40, caracterizado pelo facto de se obter um produto glicoproteico que possui uma conformação estrutural primária suficientemente duplicativa da de uma eritropoietina existente na natureza para permitir a fruição de uma ou mais das suas propriedades biológicas e com uma composição média de hidratos de carbono que difere da da eritropoietina existente na natureza.

45.- Processo de acordo com a reivindicação 40, caracterizado pelo facto de se obter um produto glicoproteico que possui uma conformação estrutural primária suficientemente duplicativa da de uma eritropoietina humana existente na natureza para permitir a fruição de uma ou mais das suas propriedades biológicas e com uma composição média de hidratos de carbono que difere da da eritropoietina humana existente na natureza.

46.- Células de vertebrados, caracterizadas pelo facto de poderem ser propagadas continuamente "in vitro" e que, após desenvolvimento em cultura são capazes de produzir, no meio do seu desenvolvimento, mais de 100 U de eritropoietina por 10^6 células em 48 horas, como se determina mediante ensaio rádio-imunológico.

47.- Células de vertebrados de acordo com a reivindicação 46, caracterizadas pelo facto de serem capazes de produzir mais de 500 U de eritropoietina por 10^6 células em 48 horas.

48.- Células de vertebrados de acordo com a reivindicação 46, caracterizadas pelo facto de serem capazes de produzir mais de 1.000 U de eritropoietina por 10^6 células em 48 horas.

49.- Células de vertebrados de acordo com a reivindicação 46, caracterizadas pelo facto de serem células de mamíferos ou de aves.

50.- Células de vertebrados de acordo com a reivindicação 49, caracterizadas pelo facto de serem células COS-1 ou células CHO.

51.- Anticorpo com características de substância antibiótica, caracterizado por apresentar imunoreactividade com eritropoietina e com um polipeptido sintético que tem uma conformação estrutural primária substancialmente duplicativa de uma sequência contínua de restos de aminoácidos existente na eritropoietina existente na natureza exceptuando qualquer polipeptido que contenha uma sequência de restos de aminoácido inteiramente compreendidos dentro da sequência A-P-P-R-L-I-C-D-S-R-V-L-E-R-Y-L-L-E-A-K.

52.- Anticorpo de acordo com a reivindicação 51, caracterizado pelo facto de ser um anticorpo monoclonal.

53.- Anticorpo de acordo com a reivindicação 51, caracterizado pelo facto de ser um anticorpo policlonal.

54.- Anticorpo de acordo com a reivindicação 51, caracterizado pelo facto de ser imuno-reactivo com eritropoietina e um polipeptido sintético que tem a sequência escolhida a partir das sequências:



V-P-D-T-K-V-N-F-Y-A-W-K-R-M-E-V-G,
 K-E-A-L-S-P-P-D-A-A-S-A-A, e
 V-Y-S-N-F-L-R-G-K-L-K-L-Y-T-G-E-A-C-R-T-G-D-R.

55.- Processo para a preparação de uma composição farmacêutica, caracterizado pelo facto de se misturar uma quantidade eficaz de um polipeptido de acordo com a reivindicação 1,16,39,40 ou 41 com um diluente, adjuvante ou excipiente aceitável sob o ponto de vista farmacêutico.

56.- Processo de acordo com a reivindicação 14, caracterizado pelo facto de a sequência de ADN purificada e isolada corresponder aos valores indicados em a) ou um seu fragmento ou a cadeia complementar dessa sequência ou desse fragmento.

57.- Processo para a preparação de um produto polipeptídico, caracterizado pelo facto de se promover a expressão de uma sequência de ADN obtida de acordo com a reivindicação 56 em uma célula hospedeira procariótica ou eucariótica.

58.- Método aperfeiçoado para a detecção de um polinucleotídeo específico, de cadeia simples e de sequência desconhecida em uma amostra celular ou viral heterogênea que contém polinucleotídeos múltiplos de cadeia simples, caracterizado pelo facto:

(a) de se preparar uma mistura de sondas de polinucleotídeo de cadeia simples, marcadas com isótopos radioactivos, que têm sequências uniformemente variadas de bases, sendo cada uma das sondas citadas potencial e especificamente complementares em relação a uma sequência de bases que é putativamente única para o polinucleotídeo a ser detectado;

(b) de se fixar a amostra em um substrato sólido;



(c) de se fixar o substrato que tem a amostra nele fixado para diminuir depois a ligação de polinucleotidos ao mesmo, excepto na hibridação para polinucleotidos na amostra citada.

(d) de se pôr transitoriamente o substrato tratado tendo a amostra nele fixada em contacto com a mistura citada de sondas marcadas com isótopos radioactivos sob condições que facilitem a hibridação apenas entre polinucleotidos totalmente complementares, e

(e) de se detectar o polinucleotido específico mediante verificação cuidadosa da presença de uma reacção de hibridação entre ele e uma sonda totalmente complementar dentro da citada mistura de sondas marcadas, como se mostra pela presença de uma maior densidade de material marcado sobre o substrato no local do polinucleotido específico em comparação com uma densidade-base de material marcado resultante da ligação não-específica de sondas marcadas ao substrato;

consistindo o aperfeiçoamento em utilizar mais 32 sondas mistas e em realizar uma ou mais das seguintes operações:

(1) utilização de um papel com nylon como substrato sólido citado;

(2) tratamento com uma protease na fase (c);

(3) utilização de concentrações e sonda marcada individual de aproximadamente 0,025 picomole; e

(4) utilização, como uma das condições de hibridação na fase (d), de temperaturas rigorosas próximas de 4°C e muito longe da Td calculada mais baixa de qualquer das sondas utilizadas.

Lisboa, 2 de Dezembro de 1985

O Agente Oficial da Propriedade Industrial



R E S U M O

"Processo para a produção de eritropoietina"

Descreve-se novos polipeptidos que possuem parte ou toda a conformação estrutural primária e uma ou mais das propriedades biológicas da eritropoietina de mamíferos ("EPO") e que são caracterizados em formas preferidas por serem o produto de expressão de hospedeiro procariótico ou eucariótico de uma sequência de ADN exógena.

Com fins ilustrativos, são incorporadas sequências de ADN genómico, cADN e ADN manufacturado, que codificam parte ou toda a sequência de restos de aminoácido de EPO ou os seus análogos, em um plasmídeo que se reproduz de maneira autónoma ou em vectores virais utilizados para transformar ou modificar células de hospedeiro procariótico ou eucariótico apropriadas como bactérias, levedura ou células de vertebrados em cultura. Após isolamento a partir dos meios de cultura ou de lisados celulares ou fragmentos, os produtos de expressão das sequências de ADN apresentam, a saber, as propriedades imunológicas e as actividades biológicas "in vitro" e "in vivo" de EPO humana ou de macaco. Descreve-se também polipeptidos sintetizados quimicamente que partilham as propriedades bioquímicas e imunológicas de EPO. São também descritos métodos aperfeiçoados para a detecção de polinucleotidos específicos de cadeia simples em uma amostra celular ou viral heterogênea preparada a partir, por ex., do ADN presente em um plasmídeo ou cADN de origem viral ou "coleção" de ADN genómico.

Lisboa, 27 de Setembro de 1985

○ Agente Oficial da Propriedade Industrial

FIG.1 COMPARAÇÃO DE EPO RECOMBINANTE HUMANA E DE MACACO
EM ENSAIO RÁDIO IMUNOLÓGICO

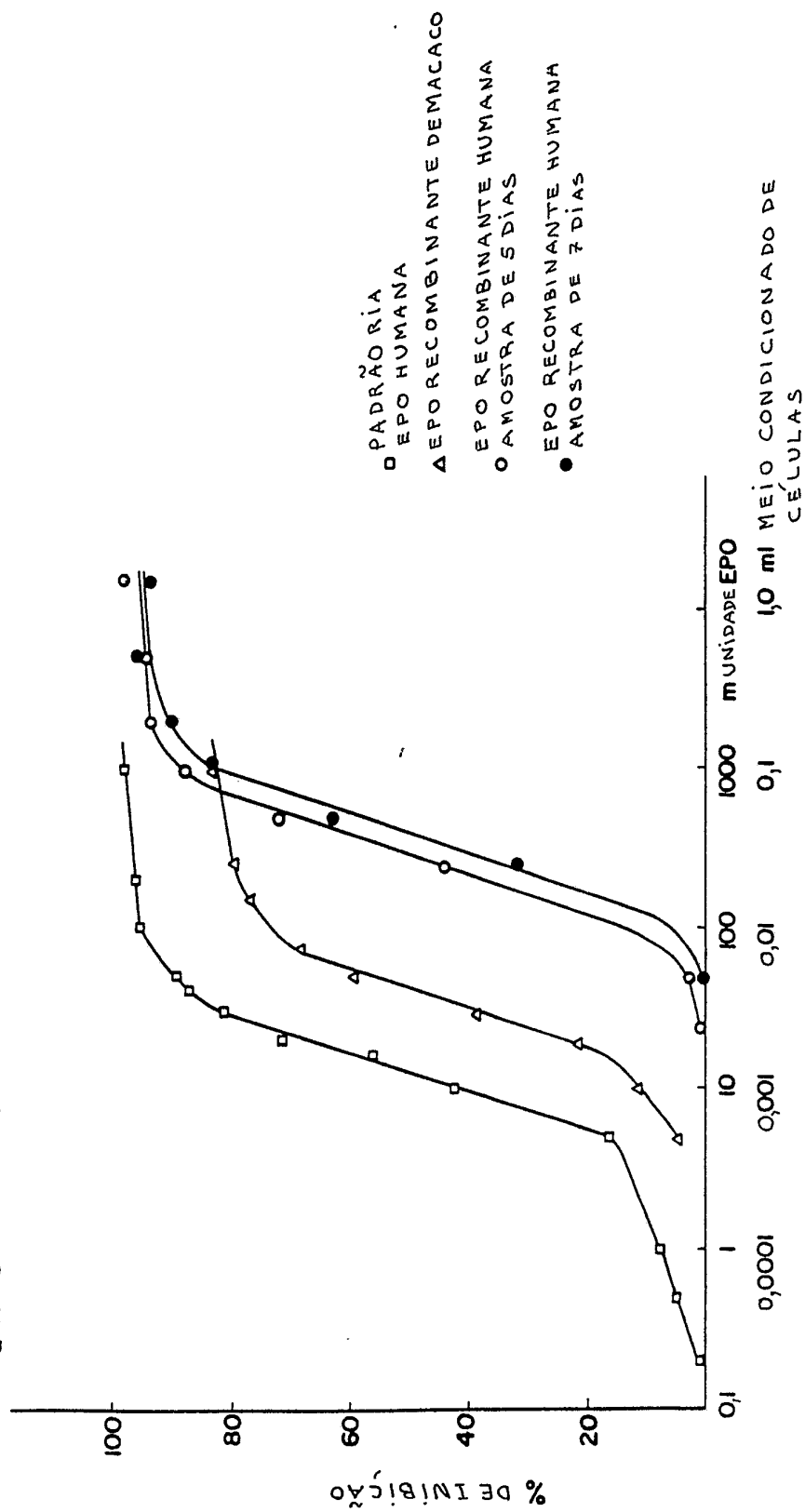


FIG.2

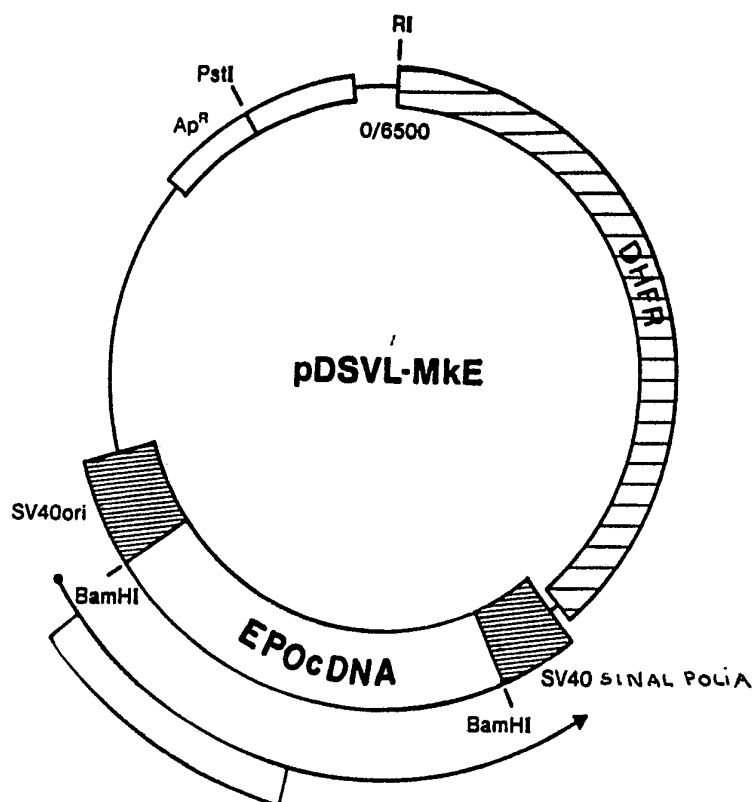


FIG. 3

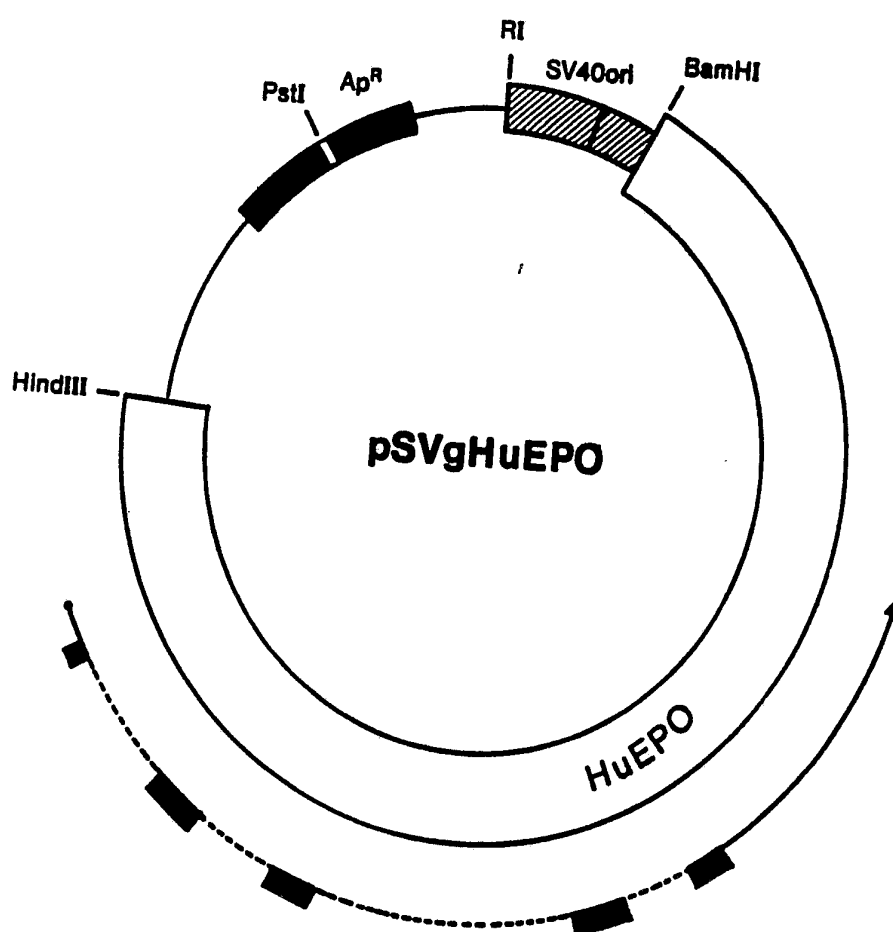


FIG. 4

