

República Federativa do Brasil
Ministério do Desenvolvimento, Indústria
e do Comércio Exterior
Instituto Nacional da Propriedade Industrial.

(21) **PI0611701-5 A2**

(22) Data de Depósito: 30/06/2006
(43) Data da Publicação: 28/09/2010
(RPI 2073)



(51) *Int.Cl.:*
C12Q 1/68

(54) Título: **MATERIAIS E MÉTODOS PARA A GERAÇÃO DE TRANSCRITOS DE ÁCIDOS NUCLEÍCOS INTEIRAMENTE MODIFICADOS EM 2**

(57) Resumo: MATERIAIS E MÉTODOS PARA A GERAÇÃO DE TRANSCRITOS DE ÁCIDOS NUCLEÍCOS INTEIRAMENTE MODIFICADOS EM 2'. Proporcionam-se materiais e métodos para produzir substâncias terapêuticas de aptâmeros tendo trifosfatos de nucleotídeos inteiramente modificados incorporados em sua seqUência.

(30) Prioridade Unionista: 30/06/2005 US 60/696.292

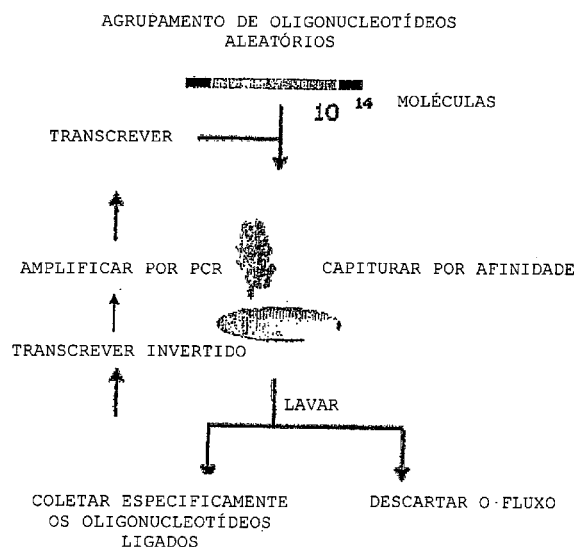
(73) Titular(es): ARCHEMIX CORP.

(72) Inventor(es): ANTHONY DOMINIC KEEFE, CHUNHUA WANG, JOHN L. DIENER, KRISTIN THOMPSON, SHARON T. CLOAD

(74) Procurador(es): RICARDO PINHO

(86) Pedido Internacional: PCT US2006025653 de 30/06/2006

(87) Publicação Internacional: WO 2007/005645 de 11/01/2007



"MATERIAIS E MÉTODOS PARA A GERAÇÃO DE TRANSCRITOS
DE ÁCIDOS NUCLÉICOS INTEIRAMENTE MODIFICADOS EM 2'"

REFERÊNCIA A PEDIDOS RELACIONADOS

Este pedido de patente não provisório reivindica
5 prioridade sob 35 U.S.C. § 119(e) para o seguinte pedido
provisório: Pedido de Patente Provisório U.S. Número de
Série 60/696.292, depositado em 30 de junho de 2005, o qual
é, nesse documento, incorporado por referência em sua
totalidade. A invenção refere-se, em geral, ao campo de
10 ácidos nucleicos e, mais particularmente, aos aptâmeros.

CAMPO DA INVENÇÃO

A invenção refere-se aos materiais e métodos para
transcrever ácidos nucleicos, particularmente as enzimas
modificadas, e materiais e métodos para usar as enzimas
15 modificadas na polimerização direcionada para molde, para
aumentar a incorporação de nucleotídeos modificados nos
ácidos nucleicos, particularmente os aptâmeros.
Adicionalmente, a invenção refere-se aos métodos e materiais
para selecionar seqüências de componentes de moldes de
20 transcrição e ao uso de tais seqüências de componentes no
aumento da produção de transcrito, particularmente no
aumento da produção de transcrito durante o método SELEX®.

ANTECEDENTES DA INVENÇÃO

Um aptâmero, por definição, é uma molécula de
25 ácido nucleico isolada, a qual se liga com alta
especificidade e afinidade a algum alvo, tal como uma
proteína, através de interações diferentes do emparelhamento
de bases de Watson-Crick. Embora os aptâmeros sejam

moléculas à base de ácidos nucleicos, há uma diferença fundamental entre os aptâmeros e as outras moléculas de ácidos nucleicos, tais como os genes e o mRNA. Nessas, a estrutura do ácido nucleico codifica informação através de sua sequência de bases linear e, assim, essa sequência é de importância para a função de armazenagem de informação. Em completo contraste, a função do aptâmero, que é baseada na ligação específica de uma molécula-alvo, não é dependente de uma sequência de bases linear conservada, porém, mais exatamente, de uma estrutura secundária/terciária particular. Ou seja, os aptâmeros são sequências de não codificação. Qualquer potencial de codificação que um aptâmero possa possuir é inteiramente casual e não desempenha nenhuma função, qualquer que seja, na ligação de um aptâmero ao seu alvo cognato. Assim, embora possa ser que aptâmeros que se ligam ao mesmo alvo, e mesmo ao mesmo sítio sobre aquele alvo, compartilhem uma sequência de bases linear similar, a maioria não o faz.

Os aptâmeros devem também ser diferenciados das sequências de ácidos nucleicos de ocorrência natural que se ligam a certas proteínas. Essas sequências são sequências de ocorrência natural incorporadas dentro do genoma do organismo que se ligam a um subgrupo especializado de proteínas que estão envolvidas na transcrição, tradução e transporte de ácidos nucleicos de ocorrência natural, i.e., proteínas de ligação a ácidos nucleicos. Os aptâmeros, por outro lado, são moléculas de ácidos nucleicos de ocorrência não natural, isoladas, curtas. Embora possam ser

identificados aptâmeros que ligam proteínas de ligação de ácidos nucleicos, na maioria dos casos tais aptâmeros têm pouca ou nenhuma identidade de sequência com as sequências reconhecidas pelas proteínas de ligação a ácidos nucleicos na natureza. Mais importantemente, os aptâmeros podem ligar virtualmente qualquer proteína (não somente as proteínas de ligação a ácidos nucleicos), bem como quase qualquer alvo de interesse, incluindo as moléculas pequenas, os carboidratos, os peptídeos, etc. Para a maioria dos alvos, mesmo as proteínas, não existe uma sequência de ácidos nucleicos de ocorrência natural à qual ele se liga. Para esses alvos que de fato têm tal sequência, i.e., as proteínas de ligação a ácidos nucleicos, tais sequências diferirão dos aptâmeros como um resultado da afinidade de ligação relativamente baixa usada na natureza, em comparação com os aptâmeros estreitamente de ligação.

Os aptâmeros, como os peptídeos gerados por exposição em fago ou anticorpos, são capazes de se ligar especificamente aos alvos selecionados e modular a atividade do alvo ou as interações de ligação, p.ex., através de ligação, os aptâmeros podem bloquear a capacidade de atuar do seu alvo. Como com os anticorpos, essa propriedade funcional de ligação específica a um alvo é uma propriedade inerente. Também como com os anticorpos, embora a pessoa versada possa não saber que características estruturais precisas um aptâmero para um alvo terá, a pessoa versada sabe como identificar, preparar e usar tal molécula na ausência de uma definição estrutural precisa.

Os aptâmeros também são análogos a substâncias terapêuticas de moléculas pequenas pelo fato de que uma única alteração estrutural, entretanto aparentemente secundária, pode dramaticamente afetar (por diversas ordens de magnitude) a ligação e/ou outra atividade (ou atividades) do aptâmero. Por outro lado, algumas alterações estruturais terão pouco ou nenhum efeito qualquer que seja. Isto resulta da importância da estrutura secundária/terciária dos aptâmeros. Em outras palavras, um aptâmero é uma estrutura tridimensional mantida em uma conformação fixa que proporciona contatos químicos para ligar especificamente o seu dado alvo. Consequentemente: (1) algumas áreas ou seqüências particulares são essenciais como (a) pontos específicos de contato com o alvo, e/ou como (b) seqüências que posicionam as moléculas em contato com o alvo; (2) algumas áreas ou seqüências particulares têm uma faixa de variabilidade, p.ex., o nucleotídeo X deve ser uma pirimidina, ou o nucleotídeo Y deve ser uma purina, ou os nucleotídeos X e Y devem ser complementares; e (3) algumas áreas ou seqüências particulares podem ser qualquer coisa, i.e., elas são essencialmente elementos espaçadores, p.ex., elas poderiam ser qualquer seqüência de nucleotídeos de um dado comprimento ou mesmo um espaçador que não seja nucleotídeo, tal como uma molécula de PEG.

Descobertos por um processo de seleção *in vitro* a partir de agrupamentos de oligonucleotídeos de seqüências aleatórias, os aptâmeros têm sido gerados por mais de 130 proteínas, incluindo os fatores do crescimento, os fatores

de transcrição, as enzimas, as imunoglobulinas, e os receptores. Um aptâmero típico é 10-15 kDa de tamanho (20-45 nucleotídeos), liga o seu alvo com afinidade nanomolar a subnanomolar, e diferencia contra alvos intimamente relacionados (p.ex., os aptâmeros tipicamente não ligarão outras proteínas da mesma família de genes). Uma série de estudos estruturais mostrou que os aptâmeros são capazes de usar os mesmos tipos de interações de ligação (p.ex., ligação de hidrogênio, complementaridades eletrostáticas, contatos hidrofóbicos, exclusão estérica) que orientam a afinidade e a especificidade em complexos de anticorpo-antígeno.

Os aptâmeros têm diversas características desejáveis para uso como substâncias terapêuticas e substâncias diagnósticas, incluindo alta especificidade e afinidade, eficácia biológica, e excelentes propriedades farmacocinéticas. Além disso, eles oferecem vantagens competitivas específicas sobre os anticorpos e outras substâncias biológicas protéicas, por exemplo:

20 1) Velocidade e controle. Os aptâmeros são produzidos por um processo inteiramente *in vitro*, permitindo a rápida geração de protótipos iniciais, incluindo protótipos terapêuticos. A seleção *in vitro* permite que a especificidade e a afinidade do aptâmero sejam rigidamente controladas e permite a geração de protótipos, incluindo protótipos contra alvos tanto tóxicos quanto não imunogênicos.

25 2) Toxicidade e Imunogenicidade. Os aptâmeros,

como uma classe, têm demonstrado toxicidade terapeuticamente aceitável ou ausência de imunogenicidade. Enquanto que a eficácia de muitos anticorpos monoclonais pode ser severamente limitada por resposta imune aos anticorpos propriamente ditos, é extremamente difícil produzir anticorpos para aptâmeros, mais provavelmente porque os aptâmeros não podem ser apresentados por células T via o MHC e a resposta imune é geralmente direcionada para não reconhecer fragmentos de ácidos nucleicos.

3) Administração. Enquanto que a maior parte das substâncias terapêuticas de anticorpos atualmente aprovadas é administrada por infusão intravenosa (tipicamente mais de 2-4 horas), os aptâmeros podem ser administrados por injeção subcutânea (a biodisponibilidade do aptâmero via administração subcutânea é >80% em estudos com macacos (Tucker e col., J. Chromatography B. 732: 203-212, 1999)). Essa diferença é principalmente devida à solubilidade comparativamente baixa e, assim, aos volumes grandes necessários para a maioria dos mAbs terapêuticos. Com boa solubilidade (>150 mg/mL) e peso molecular comparativamente baixo (aptâmero: 10-50 kDa; anticorpo: 150 kDa), uma dose semanal de aptâmero pode ser distribuída por injeção em um volume de menos do que 0,5 mL. Além disso, o tamanho pequeno dos aptâmeros permite que eles penetrem em áreas de constrições conformacionais que não permitem que anticorpos ou fragmentos de anticorpos penetrem, apresentando ainda uma outra vantagem de substâncias terapêuticas ou profilaxia à base de aptâmeros.

4) Escalabilidade e Custo.

Os aptâmeros terapêuticos são quimicamente sintetizados e consequentemente podem ser prontamente escalonados conforme necessário para atender a demanda de produção. Enquanto que as dificuldades no escalonamento da produção estão atualmente limitando a disponibilidade de algumas substâncias biológicas e o custo de capital de uma planta de produção de proteína em grande escala é enorme, um único sintetizador de oligonucleotídeos de grande escala pode produzir mais de 100 kg/ano e requer um investimento inicial relativamente modesto. O custo atual de produtos para a síntese de aptâmeros na escala de quilogramas é estimado em \$500/g, comparável àquele para anticorpos altamente otimizados. Os aperfeiçoamentos sucessivos no desenvolvimento de processos são esperados diminuir o custo de produtos para <\$100/g em cinco anos.

5) Estabilidade.

Os aptâmeros terapêuticos são quimicamente fortes. Eles estão adaptados intrinsecamente para recuperar a atividade após exposição a fatores tais como calor e desnaturantes e podem ser armazenados por períodos prolongados (>1 ano) na temperatura ambiente como pós liofilizados. Em contraste, os anticorpos devem ser armazenados refrigerados.

Além da estabilidade intrínseca dos aptâmeros, nucleotídeos modificados (p.ex., os nucleotídeos modificados em 2'), os quais são baratos, não tóxicos, e que podem aumentar a resistência à degradação enzimática, química, térmica, e física, podem ser incorporados durante o método

SELEX[®], conforme descrito no pedido de patente U.S. Nº de Série 10/729.851, depositado em 3 de dezembro de 2002, e no pedido de patente U.S. Nº de Série 10/873.856, depositado em 21 de junho de 2004. Embora a incorporação de nucleotídeos modificados durante o processo SELEX[®] seja frequentemente preferível à modificação pós-SELEX[®] devido à perda potencial de afinidade e atividade de ligação que pode ocorrer após a seleção SELEX[®], a incorporação de nucleotídeos modificados, p.ex., 2'-O-metil nucleotídeos ("2-OMe"), durante o processo SELEX[®] tem sido historicamente difícil por causa dos baixos rendimentos de transcrição. As condições da solução e as misturas de transcrição são descritas no pedido de patente U.S. Nº de Série 10/729.851, depositado em 3 de dezembro de 2002, e no pedido de patente U.S. Nº de Série 10/873.856, depositado em 21 de junho de 2004, que dão rendimentos de transcrição aperfeiçoados para aptâmeros incorporando 2'-OMe nucleotídeos. Entretanto, os rendimentos de transcrição para aptâmeros inteiramente 2'-O-metilados permanecem problemáticos.

Além das vantagens dos aptâmeros como agente terapêutico, dada a natureza barata, baixa toxicidade, e resistência à nuclease aumentada conferidas pela incorporação de 2'-OMe nucleotídeos nos aptâmeros, seria benéfico ter materiais e métodos para aumentar as produções de transcritos de aptâmeros inteiramente 2'-O-metilados para, p.ex., prolongar ou aumentar a estabilidade das substâncias terapêuticas de aptâmeros *in vivo*. A presente invenção proporciona materiais e métodos aperfeiçoados para atender

essas e outras necessidades.

RESUMO DA INVENÇÃO

O presente refere-se às T7 RNA polimerases, que podem ser purificadas, isoladas e/ou recombinantes. Como
5 usado nesse documento, o termo isoladas inclui as polimerases da invenção quando expressas de modo recombinante em uma célula ou tecido. Como usado aqui, o termo isoladas inclui as seqüências de ácidos nucleicos da invenção quando engenheiradas em uma célula ou tecido. Em
10 uma modalidade, proporciona-se uma T7 RNA polimerase compreendendo um aminoácido alterado na posição 639 e na posição 784, onde o aminoácido alterado na posição 639 não é uma fenilalanina quando o aminoácido alterado na posição 784 for uma alanina. Em uma outra modalidade, proporciona-se a
15 T7 RNA polimerase acima descrita adicionalmente compreendendo um aminoácido alterado na posição 378. Em uma outra modalidade, proporciona-se a T7 RNA polimerase acima descrita adicionalmente compreendendo um aminoácido alterado na posição 266. Em uma modalidade particular, o aminoácido
20 alterado na posição 639 é uma leucina e o aminoácido alterado na posição 784 é uma alanina. Em uma modalidade adicional, o aminoácido alterado na posição 266 é uma leucina. Em uma modalidade adicional, o aminoácido alterado na posição 378 é uma arginina.
25 Nas modalidades preferidas, os aminoácidos alterados aumentam a produção transcricional de ácidos nucleicos que compreendem modificações com 2'-OMe pela polimerase em uma reação de transcrição compreendendo

somente trifosfatos de 2'-OMe nucleotídeos. Em uma modalidade particular, o aumento na produção de transcrição é relativo a uma T7 RNA polimerase não tendo os aminoácidos alterados quando a transcrição for efetuada tanto para a T7 RNA polimerase com aminoácidos alterados quanto para a T7 RNA polimerase não tendo os aminoácidos alterados, sob condições de transcrição idênticas. Em uma outra modalidade, os aminoácidos alterados diminuem a diferenciação contra os trifosfatos de 2'-OMe nucleotídeos. Em uma modalidade particular, a diferenciação diminuída contra os trifosfatos de 2'-OMe nucleotídeos é em relação a uma T7 RNA polimerase não tendo os aminoácidos alterados quando ambas as polimerases forem usadas sob condições de transcrição idênticas. Nas modalidades particulares desse aspecto, a T7 RNA polimerase não tendo os aminoácidos alterados é a T7 RNA polimerase do tipo selvagem compreendendo um aminoácido na posição 639 alterado para uma fenilalanina e um aminoácido na posição 784 alterado para alanina ou uma polimerase mutante tendo a seqüência de aminoácidos do tipo selvagem, exceto que uma fenilalanina substituiu a tirosina na posição 639, e uma alanina substituiu a histidina na posição 784 e um resíduo de arginina substituiu o resíduo de lisina na posição 378 (Y639F/H784A/K378R).

Em uma modalidade particular, proporciona-se um polipeptídeo isolado compreendendo um aminoácido selecionado a partir do grupo que consiste em: SEQ ID NO 1, SEQ ID NO 2, SEQ ID NO 102 e SEQ ID NO 103. Em uma modalidade particular, proporciona-se um kit compreendendo um recipiente contendo

uma T7 RNA polimerase da invenção.

Em algumas modalidades, proporciona-se um método de transcrever um ácido nucléico de fita simples, compreendendo a incubação de uma T7 RNA polimerase mutante
5 com um ácido nucléico de molde, sob condições de reação suficientes para resultar na transcrição.

Em uma outra modalidade, proporciona-se um ácido nucléico isolado que codifica um polipeptídeo da invenção. Em uma modalidade particular, proporciona-se uma seqüência
10 de ácidos nucléicos, selecionada a partir do grupo que consiste em: SEQ ID NO 122, SEQ ID NO 123, SEQ ID NO 124 e SEQ ID NO 125. Em algumas modalidades, proporciona-se um vetor compreendendo uma seqüência de ácidos nucléicos isolada da invenção. em uma modalidade particular,
15 proporciona-se um vetor de expressão compreendendo um ácido nucléico da invenção operavelmente ligado a um promotor. Em uma outra modalidade da invenção, proporciona-se uma célula compreendendo o vetor de expressão da invenção. Em uma modalidade particular, proporciona-se uma célula, onde a T7
20 RNA polimerase mutante da invenção é expressa pela célula. Em algumas modalidades, proporciona-se um kit compreendendo um recipiente contendo um ácido nucléico codificando uma T7 RNA polimerase da invenção.

Em uma outra modalidade, proporciona-se um método
25 de transcrever um ácido nucléico inteiramente com 2'-OMe, compreendendo as etapas de a) incubar um ácido nucléico de molde em uma mistura de reação sob condições compreendendo uma RNA polimerase mutante, um molde de transcrição de ácido

nucléico e trifosfatos de nucleosídeos, onde os trifosfatos de nucleosídeos são 2'OMe, e b) transcrever a mistura de reação de transcrição para resultar no ácido nucléico de fita simples, onde todos os nucleotídeos dos ácidos
5 nucléicos de fitas simples são modificados com 2'-OMe, exceto que o primeiro nucleotídeo dos transcritos pode ser não modificado em 2'. Em algumas modalidades, o primeiro nucleosídeo do transcrito pode ser a 2'-OH guanosina. Em algumas modalidades do método, a RNA polimerase mutante é
10 uma T7 RNA polimerase mutante compreendendo um aminoácido alterado na posição 639 e na posição 784, particularmente uma T7 RNA polimerase compreendendo um aminoácido alterado na posição 639 e na posição 784, onde o aminoácido alterado na posição 639 não é uma fenilalanina quando o aminoácido
15 alterado na posição 784 for uma alanina, particularmente uma T7 RNA polimerase adicionalmente compreendendo um aminoácido alterado na posição 378 e/ou um aminoácido alterado na posição 266. Em uma modalidade particular, o aminoácido alterado na posição 639 é uma leucina e o aminoácido
20 alterado na posição 784 é uma alanina, na polimerase para uso nos métodos da invenção. Em uma modalidade adicional, o aminoácido alterado na posição 266 é uma leucina da polimerase para uso nos métodos da invenção. Em uma modalidade adicional, o aminoácido alterado na posição 378 é
25 uma arginina na polimerase para uso nos métodos da invenção. Em uma modalidade particular, proporciona-se um polipeptídeo isolado compreendendo um aminoácido selecionado a partir do grupo que consiste em: SEQ ID NO 1, SEQ ID NO 2, SEQ ID NO

102 e SEQ ID NO 103.

Em algumas modalidades do método da invenção, a reação de transcrição adicionalmente compreende íons magnésio. Em uma outra modalidade, a reação de transcrição
5 adicionalmente compreende íons manganês. Em uma outra modalidade, os íons magnésio estão presentes na reação de transcrição em uma concentração que é entre 3,0 a 3,5 vezes maior do que os íons manganês. Em uma outra modalidade, onde cada trifosfato de nucleotídeo está presente na reação
10 de transcrição em uma concentração de 1,0 mM, a concentração de íons magnésio é 6,5 mM, e a concentração de íons manganês é 2,0 mM. Em uma outra modalidade, onde cada trifosfato de nucleotídeo está presente na reação de transcrição em uma concentração de 1,5 mM, a concentração de íons magnésio é 8
15 mM, e a concentração de íons manganês é 2,5 mM. Em uma outra modalidade, onde cada trifosfato de nucleotídeo está presente na reação de transcrição em uma concentração de 2,0 mM, a concentração de íons magnésio é 9,5 mM e a concentração de íons manganês é 3,0 mM.

20 Em uma outra modalidade, a reação de transcrição adicionalmente compreende um resíduo que não é de trifosfato que não é de 2'-OMe guanosina, particularmente onde o resíduo que não é de trifosfato que não é de 2'-OMe guanosina é selecionado a partir do grupo que consiste em:
25 monofosfato de guanosina, difosfato de guanosina, monofosfato de 2'-flúor guanosina, difosfato de 2'-flúor guanosina, monofosfato de 2'-amino guanosina, difosfato de 2'-amino guanosina, monofosfato de 2'-desóxi guanosina, e

difosfato de 2'-desóxi guanosina. Em uma outra modalidade, o molde de transcrição compreende um promotor de T7 RNA polimerase. Em uma outra modalidade, a reação de transcrição adicionalmente compreende o polietileno glicol.

5 Em uma outra modalidade, a reação de transcrição compreende a pirofosfatase inorgânica.

Em um outro aspecto da invenção, proporciona-se um método para identificar aptâmeros. Em uma modalidade, proporciona-se um método para identificar um aptâmero, 10 compreendendo: a) a preparação de uma mistura de reação de transcrição compreendendo uma polimerase mutante da invenção, e um ou mais moldes de transcrição de ácidos nucleicos; b) a transcrição da mistura de reação de transcrição para resultar em uma mistura candidata de ácidos nucleicos de 15 fitas simples, onde todos menos opcionalmente um dos nucleotídeos dos ácidos nucleicos de fitas simples são modificados em 2', c) o contato da mistura candidata com a molécula-alvo, d) a separação dos ácidos nucleicos que têm uma afinidade aumentada pela molécula-alvo, em relação a uma 20 afinidade da mistura candidata, da mistura candidata, e e) a amplificação dos ácidos nucleicos com afinidade aumentada para produzir uma mistura enriquecida em aptâmeros, pelo que são identificados aptâmeros para a molécula-alvo compreendendo nucleotídeo todo modificado em 2', exceto que 25 o primeiro nucleotídeo dos aptâmeros pode ser não modificado em 2'. Em algumas modalidades, a etapa de amplificação f) compreende (i) dissociar opcionalmente os ácidos nucleicos com afinidade aumentada do alvo, ii) transcrever invertido

os ácidos nucléicos com afinidade aumentada, dissociados dos complexos de ácidos nucléicos-alvo, iii) amplificar os ácidos nucléicos com afinidade aumentada transcritos invertidos; e (ii) preparar uma mistura de reação de transcrição compreendendo os ácidos nucléicos com afinidade aumentada transcritos invertidos amplificados como o molde de transcrição e transcrever a mistura de transcrição.

Em algumas modalidades do método de identificação de aptâmeros da invenção, a RNA polimerase mutante é uma T7 RNA polimerase mutante compreendendo um aminoácido alterado na posição 639 e na posição 784, particularmente uma T7 RNA polimerase compreendendo um aminoácido alterado na posição 639 e na posição 784, onde o aminoácido alterado na posição 639 não é uma fenilalanina quando o aminoácido alterado na posição 784 for uma alanina, particularmente, uma T7 RNA polimerase adicionalmente compreendendo um aminoácido alterado na posição 378 e/ou um aminoácido alterado na posição 266. Em uma modalidade particular, o aminoácido alterado na posição 639 é uma leucina e o aminoácido alterado na posição 784 é uma alanina, na polimerase para uso nos métodos da invenção. Em uma modalidade adicional, o aminoácido alterado na posição 266 é uma leucina da polimerase para uso nos métodos da invenção. Em uma modalidade adicional, o aminoácido alterado na posição 378 é uma arginina na polimerase para uso nos métodos da invenção. Em uma modalidade particular, um polipeptídeo isolado compreendendo um aminoácido selecionado a partir do grupo que consiste em: SEQ ID NO 1, SEQ ID NO 2, SEQ ID NO 102 e

SEQ ID NO 103 é usado no método de identificação de aptâmeros da invenção.

Em algumas modalidades, os todos os trifosfatos de nucleotídeos na reação de transcrição são modificados com 2'-OMe. Em uma modalidade, os um ou mais moldes de transcrição de ácidos nucleicos compreendem um promotor de T7 RNA polimerase e uma sequência líder imediatamente 3' ao promotor de T7 RNA polimerase. Em algumas modalidades desse aspecto, o método compreende repetir as etapas a) até e) iterativamente.

Em algumas modalidades do método de identificar aptâmeros da invenção, a reação de transcrição adicionalmente compreende íons magnésio. Em uma outra modalidade, a reação de transcrição adicionalmente compreende íons manganês. Em uma outra modalidade, os íons magnésio estão presentes na reação de transcrição em uma concentração que é entre 3,0 a 3,5 vezes maior do que os íons manganês. Em uma outra modalidade, onde cada trifosfato de nucleotídeo está presente na reação de transcrição em uma concentração de 1,0 mM, a concentração de íons magnésio é 6,5 mM, e a concentração de íons manganês é 2,0 mM. Em uma outra modalidade, onde cada trifosfato de nucleotídeo está presente na reação de transcrição em uma concentração de 1,5 mM, a concentração de íons magnésio é 8 mM, e a concentração de íons manganês é 2,5 mM. Em uma outra modalidade, onde cada trifosfato de nucleotídeo está presente na reação de transcrição em uma concentração de 2,0 mM, a concentração de íons magnésio é 9,5 mM e a

concentração de íons manganês é 3,0 mM.

Em uma outra modalidade desse aspecto, a reação de transcrição para uso no método de identificação de aptâmeros da invenção adicionalmente compreende um resíduo que não é de trifosfato que não é de 2'-OMe guanosina, particularmente onde o resíduo que não é de trifosfato que não é de 2'-OMe guanosina é selecionado a partir do grupo que consiste em: monofosfato de guanosina, difosfato de guanosina, monofosfato de 2' flúor guanosina, difosfato de 2' flúor guanosina, monofosfato de 2'-amino guanosina, difosfato de 2'-amino guanosina, monofosfato de 2'-desóxi guanosina, e difosfato de 2'-desóxi guanosina. Em uma outra modalidade, o molde de transcrição compreende um promotor de T7 RNA polimerase. Em uma outra modalidade, a reação de transcrição adicionalmente compreende o polietileno glicol. Em uma outra modalidade, a reação de transcrição compreende a pirofosfatase inorgânica.

A presente invenção também se refere a um método de selecionar seqüências de componentes de moldes de ácidos nucleicos para orientar a transcrição. Em uma modalidade, a seqüência de componente aumenta a produção de transcrito da transcrição orientada pelo molde. Em uma modalidade particular, a invenção refere-se aos métodos de selecionar seqüências líderes para aumentar a produção de transcrito e às seqüências líderes, aos moldes de ácidos nucleicos compreendendo as seqüências líderes e aos métodos de usar as seqüências líderes e os moldes de ácidos nucleicos da invenção. A presente invenção também se refere às novas

polimerases mutantes e ao seu uso na transcrição, particularmente ao seu uso para aumentar a produção de transcrito onde nucleotídeos modificados em 2' estão sendo incorporados, mais particularmente onde todos os nucleotídeos que estão sendo incorporados são modificados em 2', p.ex., são 2'-OMe. A presente invenção também se refere às condições de reação de transcrição modificadas, para aumentar a produção de transcrito. A presente invenção particularmente se refere às combinações em pares e triplas dos aspectos acima descritos, particularmente para aperfeiçoar a produção de transcrito, onde em todos, exceto o nucleotídeo de partida, os transcritos, são modificados em 2', particularmente modificados com 2'-OMe (transcritos inteiramente com 2'-OMe" ou "mRmY" ou "MNA").

Em uma modalidade do primeiro aspecto da invenção, proporciona-se um método de identificar uma sequência de componente de molde de ácido nucléico para aumentar a transcrição, compreendendo: a) preparar uma biblioteca de candidatos de moldes de transcrição, onde os moldes compreendem um promotor, uma primeira região fixa imediatamente 3' ao promotor, uma região degenerada imediatamente 3' à primeira região fixa e uma segunda região fixa 3' à região degenerada; b) transcrever a biblioteca de candidatos de moldes de transcrição em uma reação de transcrição para dar uma biblioteca de transcritos; c) transcrever invertida a mistura de transcrição para obter uma mistura candidata de cDNA, onde os moldes de cDNA compreendem uma extremidade de 5' e 3'; d) ligar uma

seqüência de DNA codificando o promotor à extremidade de 3' dos moldes de cDNA em uma reação de ligação; e) amplificar os moldes de cDNA para resultar em uma biblioteca de candidatos de moldes de transcrição; e f) identificar um
5 componente de seqüência de ácidos nucléicos para aumentar a transcrição a partir da biblioteca de candidatos de moldes de transcrição, onde o componente de seqüência de ácidos nucléicos compreende uma seqüência derivada de pelo menos uma porção da região degenerada. Em uma modalidade desse
10 método da invenção, a etapa f) compreende i) clonar a biblioteca de candidatos de moldes de transcrição em moldes de transcrição individuais; ii) transcrever os moldes de transcrição individuais em uma reação de transcrição para resultar em uma produção de transcritos; iii) avaliar a
15 produção de transcritos dos moldes de transcrição individuais; e iv) identificar o componente de seqüência de ácidos nucléicos em um molde de transcrição que resulte em uma produção de transcritos predeterminada. Em uma modalidade particular desse método da invenção, a produção
20 de transcritos predeterminada é uma produção maior do que a produção de transcritos obtida na etapa b) por transcrição da mistura de candidatos de moldes de transcrição.

Em uma outra modalidade desse método da invenção, a etapa f) compreende analisar a composição de base da
25 região degenerada da biblioteca de candidatos de moldes de transcrição e identificar o componente da seqüência de ácidos nucléicos com base na composição média de base da biblioteca de candidatos de moldes de transcrição.

Em algumas modalidades desse aspecto da invenção, a etapa b) do método adicionalmente compreende tratar a mistura de transcrição transcrita com DNase. Nas modalidades adicionais desse método da invenção, a etapa b) adicionalmente compreende purificar a mistura de transcrição transcrita por separação dos moldes de transcrição transcritos dos outros componentes da reação de transcrição. Em uma modalidade particular desse método da invenção, a etapa de purificação compreende substituir o tampão de reação de transcrição correndo a reação de transcrição através de uma coluna de dessalinização.

Em uma outra modalidade desse aspecto da invenção, a etapa d) do método é feita antes da etapa c). Em uma outra modalidade desse aspecto da invenção, o método compreende repetir as etapas b) até e) mais do que uma vez, antes de efetuar a etapa f).

Em uma modalidade adicional desse aspecto da invenção, a reação de ligação é uma reação de ligação com esplinte e a reação de ligação compreende uma imobilização de ácido nucléico e um oligonucleotídeo 5'-monofosforilado codificando o promotor.

Em uma modalidade particular desse aspecto da invenção, a reação de transcrição usada no método compreende um ou mais trifosfatos de nucleotídeos modificados e uma polimerase mutada. Em algumas modalidades, o trifosfato de nucleotídeo modificado é um trifosfato de nucleotídeo modificado em 2', particularmente um trifosfato de nucleotídeo modificado com 2'-OMe. Em algumas modalidades,

a polimerase mutada é uma T7 RNA polimerase mutada. Em algumas modalidades, a reação de transcrição usada no método da invenção compreende íons magnésio e manganês (Mn^{2+}) e a T7 RNA polimerase mutada é selecionada a partir do grupo que
5 consiste em: SEQ ID NO 1, SEQ ID NO 2, SEQ ID NO 100 e SEQ ID NO 101.

Em algumas modalidades desse aspecto da invenção, os íons magnésio estão presentes na reação de transcrição em uma concentração que é entre 3,0 a 3,5 vezes maior do que os
10 íons manganês (Mn^{2+}). Nas modalidades adicionais desse aspecto da invenção, cada trifosfato de nucleotídeo está presente na reação de transcrição em uma concentração de 1,0 mM, a concentração de íons magnésio é 6,5 mM, e a concentração de íons manganês é 2,0 mM. Nas modalidades
15 adicionais desse aspecto da invenção, cada trifosfato de nucleotídeo está presente na reação de transcrição em uma concentração de 1,5 mM, a concentração de íons magnésio é 8 mM, e a concentração de íons manganês é 2,5 mM. Nas modalidades ainda adicionais desse aspecto da invenção, cada
20 trifosfato de nucleotídeo está presente na reação de transcrição em uma concentração de 2,0 mM, a concentração de íons magnésio é 9,5 mM e a concentração de íons manganês é 3,0 mM. Em algumas modalidades do método, a reação de transcrição adicionalmente compreende um polialquileno
25 glicol, particularmente o polietileno glicol. Em algumas modalidades do métodos, particularmente nas modalidades em que forem desejados transcritos inteiramente com 2'-OMe, a reação de transcrição adicionalmente compreende um resíduo

de guanosina selecionado a partir do grupo que consiste em: monofosfato de guanosina, difosfato de guanosina, monofosfato ou difosfato de 2' flúor guanosina, monofosfato ou difosfato de 2'-amino guanosina, monofosfato ou difosfato de 2'-desóxi guanosina, ou outros nucleotídeos modificados. Nas modalidades adicionais, a reação de transcrição do método da invenção compreende a pirofosfatase inorgânica. Nas modalidades adicionais, a reação de transcrição do método da invenção opcionalmente compreende a combinação a partir do grupo que consiste em: tampão, detergente (p.ex., Triton X-100), poliamina (p.ex., espermina ou espermidina), e agente redutor (p.ex., DTT ou β ME). Ainda nas modalidades adicionais, a reação de transcrição do método da invenção compreende trifosfatos de nucleotídeos, íons magnésio, íons manganês (p.ex., Mn^{2+}), polietileno glicol, monofosfato de guanosina, pirofosfatase inorgânica, tampão, detergente, poliamina, e DTT, e um ou mais moldes de transcrição de oligonucleotídeos, e uma T7 RNA polimerase, p.ex., uma T7 RNA polimerase mutante, p.ex., uma T7 RNA polimerase mutante selecionada a partir do grupo que consiste em: SEQ ID NO 1, 2, 100 e 101.

Em algumas modalidades do método de identificação da invenção, a primeira região fixa da biblioteca de candidatos de moldes de transcrição consiste em 2, 3, 4 ou 5 resíduos de guanosina. Em algumas modalidades da invenção, a região degenerada da biblioteca de candidatos de moldes de transcrição compreende pelo menos 4, 10, 20 ou 30 nucleotídeos.

Em algumas modalidades do método, a sequência de componente de molde de ácido nucléico a ser identificada é uma sequência líder. Em algumas modalidades, a sequência líder compreende a primeira região fixa e uma sequência derivada de pelo menos uma parte da região degenerada da biblioteca de candidatos de moldes de transcrição. Em algumas modalidades, o método da invenção adicionalmente compreende incorporar a sequência líder identificada em um molde de transcrição de oligonucleotídeo.

A invenção também proporciona a sequência líder identificada pelo método de identificação da invenção. Em algumas modalidades, a sequência líder da invenção compreende a sequência de ácidos nucléicos a partir do nucleotídeo 22 até o nucleotídeo 32 em qualquer uma das sequências selecionadas a partir do grupo que consiste em: SEQ ID NOs 10 a 99. Em algumas modalidades, a sequência líder da invenção compreende a sequência de ácidos nucléicos a partir do nucleotídeo 18 até o nucleotídeo 32 em qualquer uma das sequências selecionadas a partir do grupo que consiste em: SEQ ID NOs 10 a 99. A invenção também proporciona um molde de transcrição de oligonucleotídeo compreendendo uma sequência líder da invenção. Nas modalidades particulares, a invenção proporciona um molde de transcrição de oligonucleotídeo selecionado a partir do grupo que consiste em: SEQ ID NO 3 a 6 e SEQ ID NO 106.

Em um outro aspecto da invenção, proporciona-se um método para aumentar a produção de transcritos de um ácido nucléico, onde a transcrição é orientada por um molde de

transcrição de oligonucleotídeo. Em algumas modalidades desse aspecto da invenção, o método de aumentar a produção de transcritos compreende orientar a transcrição com um molde de transcrição de oligonucleotídeo que compreende uma
5 sequência líder, onde a sequência líder tenha sido identificada pelo método de identificação da invenção usando a composição de nucleotídeos e/ou a polimerase e/ou as condições iguais como usadas na reação de transcrição para a qual o aumento da produção de transcritos é desejado.

10 BREVE DESCRIÇÃO DOS DESENHOS

A Figura 1 é uma representação esquemática do processo de seleção de aptâmeros *in vitro* (SELEX®) a partir de agrupamentos de oligonucleotídeos de sequências aleatórias.

15 A Figura 2 mostra um fluxograma de um método SELEX® de Região Terminal (TR-SELEX®).

A Figura 3 mostra uma análise gráfica da composição de nucleotídeos média combinada de regiões selecionadas a partir de vinte posições degeneradas de uma
20 biblioteca de candidatos de moldes de transcrição antes (R0) e após (R3) a seleção por TR-SELEX®.

A Figura 4 mostra a produção relativa de transcritos, quantificada a partir de imageamento com UV da análise por PAGE-gel para ARC2118, ARC2119, ARC2120, e
25 ARC2121 usando as T7 RNA polimerases mutantes Y639F/H784A/K378R ("FAR") e Y639L/H784A/K378R ("LAR") com uma adição de 2'-OH GTP na mistura de transcrição. * indica que as produções dadas são em relação à ARC2118 transcrita

com a polimerase mutante LAR, que deu a mais alta produção quantificada por imageamento com UV.

A Figura 5A mostra o ácido nucléico (SEQ ID NO 120) e a Figura 5B dá a seqüência de aminoácidos da T7 RNA polimerase do tipo selvagem (SEQ ID NO 121).

A Figura 6A mostra a seqüência de ácidos nucléicos (SEQ ID NO 122) da T7 RNA polimerase mutante Y639L/H784A. A Figura 6B mostra a seqüência de ácidos nucléicos (SEQ ID NO 123) da T7 polimerase mutante Y639L/H784A/K378R. A Figura 6C mostra a seqüência de ácidos nucléicos (SEQ ID NO 124) da T7 polimerase mutante P266L/Y639L/H784A. A Figura 6D mostra a seqüência de ácidos nucléicos (SEQ ID NO 125) da T7 polimerase mutante P266L/Y639L/H784A/K378R.

A Figura 7 mostra a produção relativa de transcritos, quantificada a partir de imageamento com UV da análise por PAGE-gel para ARC2118 e ARC2119, usando a T7 RNA polimerase mutante Y639L/H784A/K378R com uma titulação de rGTP (2'-OH GTP) na mistura de transcrição. * indica que as produções dadas são em relação à ARC2118 transcrita com rGTP a 20 μ M, que deu a mais alta produção quantificada por imageamento com UV.

A Figura 8 mostra a produção relativa de transcritos, quantificada a partir de imageamento com UV da análise por PAGE-gel para ARC2119, usando a T7 RNA polimerase mutante Y639L/H784A/K378R com concentrações variadas de 2'-OMe NTPs (A, U, C, e G), $MgCl_2$ e $MnCl_2$ e nenhum rGTP (2'-OH GTP) na mistura de transcrição. As produções dadas são em relação à condição de transcrição de

1 mM cada de 2'-OMe NTP, $MgCl_2$ a, 6,5 mM, e $MnCl_2$ a 2 mM.

A Figura 9 é uma tabela que mostra uma análise das inserções, deleções e substituições de nucleotídeos da transcrição inteiramente com 2'-OMe (100% de 2'-OMe A, U, C, G) com a T7 RNA polimerase mutante Y639L/H784A/K378R, em comparação com a fidelidade da transcrição com totalmente RNA ou 2'-OMe usando a T7 RNA polimerase mutante Y639F/K378R. Na tabela, (1) indica dados do "Direct in Vitro Selection of a 2'-O-Methyl Aptamer to VEGF", Burmeister e col., (2005) Chemistry and Biology, 12: 25-33, onde as transcrições foram feitas com a T7 polimerase mutante FAR e (2) indica que a transcrição foi feita com a T7 polimerase mutante LAR.

A Figura 10 é uma tabela que mostra uma análise da porcentagem de composição de nucleotídeos dos transcritos inteiramente com 2'-OMe (100% de 2'-OMe A, T, C, G) antes e depois de uma rodada de transcrição inteiramente com 2'-OMe usando a T7 RNA polimerase mutante Y639L/H784A/K378R, seguida por tratamento com DNase, transcrição reversa, ligação com esplinte, e amplificação por PCR.

A Figura 11 é uma representação esquemática de um aptâmero de MNA anti-IgE minimizado, mostrado na direção 5' para 3', tendo um quepi sobre a sua extremidade de 3' (bola de cor preta).

A Figura 12 é uma representação esquemática do aptâmero de MNA anti-IgE minimizado, o aptâmero de MNA anti-IgE minimizado tendo duas substituições de desóxi e o aptâmero de MNA anti-IgE minimizado tendo uma substituição de desóxi e substituições de fosforotioato, cada um mostrado

na direção 5' para 3' e cada um tendo um quepi sobre a sua extremidade de 3' (bola de cor preta).

A Figura 13 é uma ilustração representando diversas estratégias de PEGuilação representando a mono-
5 PEGuilação padrão, a PEGuilação múltipla, e a dimerização via PEGuilação.

DESCRIÇÃO DETALHADA DA INVENÇÃO

Os detalhes de uma ou mais modalidades da invenção são apresentados na descrição abaixo que acompanha. Embora
10 quaisquer métodos e materiais similares ou equivalentes àqueles descritos nesse documento possam ser usados na prática ou no teste da presente invenção, os métodos e os materiais preferidos são agora descritos. Outras características, objetivos, e vantagens da invenção serão
15 aparentes a partir da descrição. No relatório descritivo, as formas singulares também incluem o plural, a não ser que o contexto claramente dite de outro modo. A não ser que de outro modo definido, todos os termos técnicos e científicos usados nesse documento têm o mesmo significado como
20 comumente entendidos por alguém de habilidade comum na técnica à qual pertence essa invenção. No caso de conflito, o presente Relatório Descritivo dominará.

O MÉTODO SELEXTM

O método preferido para gerar um aptâmero é com o
25 processo intitulado "Systematic Evolution of Ligands by Exponential Enrichment" ("SELEX[®]"), representado de um modo geral na Figura 1 e também referido como seleção *in vitro*. O processo SELEX[®] é um método para o desenvolvimento *in*

vitro de moléculas de ácidos nucleicos com ligação altamente específica a moléculas-alvo e é descrito, p.ex., no pedido de patente U.S. Nº de Sér. 07/536.428, depositado em 11 de jun. de 1990, agora abandonado, na Pat. U.S. Nº 5.475.096, intitulada "Nucleic Acid Ligands", e na Pat. U.S. Nº 5.270.163 (ver também o WO 91/19813), intitulada "Nucleic Acid Ligands". Por realização de ciclos iterativos de seleção e amplificação, o SELEX® pode ser usado para obter aptâmeros, também referidos nesse documento como "ligantes de ácidos nucleicos", com qualquer nível desejado de afinidade de ligação ao alvo.

O processo SELEX® é baseado na constatação única que os ácidos nucleicos têm capacidade suficiente para formar uma variedade de estruturas bi e tridimensionais e versatilidade química suficiente disponível dentro de seus monômeros para atuar como ligantes (i.e., formam pares de ligação específicos) com virtualmente qualquer composto químico, quer seja monomérico, quer seja polimérico. As moléculas de qualquer tamanho ou composição podem servir como alvos.

O processo SELEX® é baseado na capacidade de ligar um alvo. Os aptâmeros obtidos através do procedimento SELEX® assim terão a propriedade de ligação ao alvo. A mera ligação ao alvo, entretanto, não proporciona nenhuma informação sobre o efeito funcional, se algum, que pode ser exercido sobre o alvo pela ação da ligação do aptâmero.

A alteração de uma propriedade da molécula-alvo requer que o aptâmero se ligue em certa posição sobre o alvo,

para efetuar uma alteração em uma propriedade do alvo. Na teoria, o método SELEX[®] pode resultar na identificação de um grande número de aptâmeros, onde cada aptâmero liga em um sítio diferente sobre o alvo. Na prática, as interações de ligação entre aptâmero-alvo frequentemente ocorrem em um ou em um número relativamente pequeno de sítios de ligação preferidos sobre o alvo, que proporcionam interfaces estruturais estáveis e acessíveis para a interação. Além disso, quando o método SELEX[®] for efetuado sobre uma molécula-alvo fisiológica, a pessoa versada geralmente não é capaz de controlar a posição do aptâmero ao alvo. Desse modo, a posição do sítio de ligação do aptâmero sobre o alvo pode ou não ser em, ou próxima a, um de potencialmente diversos sítios de ligação que poderiam resultar no efeito desejado, ou pode não ter nenhum efeito sobre a molécula-alvo.

Mesmo onde um aptâmero, em virtude de sua capacidade de ligar o alvo, for verificado ter um efeito, não há nenhum modo de prever a existência desse efeito ou de saber de antemão qual será o efeito. Na realização de um experimento SELEX[®], a pessoa versada pode somente saber com alguma certeza que aptâmeros, na medida em que seja possível obter um aptâmero contra um alvo, terão a propriedade de ligação ao alvo. Pode-se efetuar um experimento SELEX[®] na expectativa que alguns dos aptâmeros identificados também terão um efeito sobre o alvo além da ligação a ele, porém isso é incerto.

O processo SELEX[®] vale-se como um ponto de partida

de uma grande biblioteca ou agrupamento de oligonucleotídeos de fita simples compreendendo seqüências ao acaso. Os oligonucleotídeos podem ser DNA, RNA, ou híbridos de DNA/RNA modificados ou não modificados. Em alguns exemplos, o agrupamento compreende 100% de oligonucleotídeos degenerados ou parcialmente degenerados. Em outros exemplos, o agrupamento compreende oligonucleotídeos degenerados ou parcialmente degenerados contendo pelo menos uma seqüência fixa e/ou seqüência conservada incorporada na seqüência ao acaso. Em outros exemplos, o agrupamento compreende oligonucleotídeos degenerados ou parcialmente degenerados contendo pelo menos uma seqüência fixa e/ou seqüência conservada em sua extremidade de 5' e/ou 3', que pode compreender uma seqüência compartilhada por todas as moléculas do agrupamento de oligonucleotídeos. As seqüências fixas são seqüências comuns aos oligonucleotídeos no agrupamento que são incorporadas para um propósito pré-selecionado, tal como os motivos de CpG descritos adicionalmente abaixo, os sítios de hibridização para os primers da PCR, seqüências promotoras para as RNA polimerases (p.ex., T3, T4, T7, e SP6), sítios de restrição, ou seqüências homopoliméricas, tais como os sistemas de poli A ou poli T, os núcleos catalíticos, os sítios para a ligação seletiva a colunas de afinidade, as seqüências líderes que promovem a transcrição, e outras seqüências para facilitar a clonagem e/ou o seqüenciamento de um oligonucleotídeo de interesse. As seqüências conservadas são seqüências, diferentes das seqüências fixas

anteriormente descritas, compartilhadas por diversos aptâmeros que se ligam ao mesmo alvo.

Os oligonucleotídeos do agrupamento preferivelmente incluem uma porção de sequência degenerada, bem como sequências fixas necessárias para a amplificação eficiente. Tipicamente, os oligonucleotídeos do agrupamento de partida contêm sequências 5' e 3' terminais, fixas, que flanqueiam uma região interna de 30-40 nucleotídeos aleatórios. Os nucleotídeos degenerados podem ser produzidos em diversos modos, incluindo a síntese química e a seleção de tamanho a partir de ácidos nucleicos celulares aleatoriamente clivados. A variação de sequência nos ácidos nucleicos de teste pode também ser introduzida ou aumentada por mutagênese antes ou durante as iterações de seleção/amplificação.

A porção de sequência degenerada do oligonucleotídeo pode ser de qualquer comprimento e pode compreender ribonucleotídeos e/ou desoxirribonucleotídeos e pode incluir nucleotídeos ou análogos de nucleotídeos modificados ou não naturais. Ver, p.ex., a Patente U.S. Nº 5.958.691; a Patente U.S. Nº 5.660.985; a Patente U.S. Nº 5.958.691; a Patente U.S. Nº 5.698.687; a Patente U.S. Nº 5.817.635; a Patente U.S. Nº 5.672.695, e a Publicação PCT WO 92/07065. Os oligonucleotídeos degenerados podem ser sintetizados a partir de nucleotídeos ligados a um fosfodiéster, usando as práticas de síntese de oligonucleotídeos de fase sólida, bastante conhecidas na técnica. Ver, p.ex., Froehler e col, Nucl. Acid Res.

14:5399-5467 (1986) e Froehler e col, Tet. Lett. 27:5575-5578 (1986). Os oligonucleotídeos aleatórios podem também ser sintetizados usando métodos em fase de solução, tais como os métodos de síntese de triésteres. Ver, p.ex., Sood
5 e col, Nucl. Acid Res. 4:2557 (1977) e Hirose e col, Tet. Lett., 28:2449 (1978). As sínteses típicas realizadas em equipamento de síntese de DNA automatizado produzem 10^{16} - 10^{17} moléculas individuais, um número suficiente para a maioria dos experimentos SELEX®. As regiões suficientemente grandes
10 da sequência degenerada no projeto da sequência aumentam a probabilidade que cada molécula sintetizada seja provável de representar uma única sequência.

A biblioteca de partida dos oligonucleotídeos pode ser gerada por síntese química automatizada em um
15 sintetizador de DNA. Para sintetizar as sequências degeneradas, misturas de todos os quatro nucleotídeos são adicionadas em cada etapa de adição de nucleotídeos, durante o processo de síntese, permitindo a incorporação estocástica de nucleotídeos. Conforme acima estabelecido, em uma
20 modalidade, os oligonucleotídeos aleatórios compreendem sequências inteiramente degeneradas; entretanto, em outras modalidades, os oligonucleotídeos degenerados podem compreender extensões de sequências não aleatórias ou parcialmente aleatórias. As sequências parcialmente
25 aleatórias podem ser criadas por adição dos quatro nucleotídeos em diferentes razões molares, em cada etapa de adição.

Nessas situações onde uma biblioteca de RNA é para

ser usada como a biblioteca de partida, ela é tipicamente gerada sintetizando uma biblioteca de DNA, opcionalmente amplificando por PCR, então transcrevendo a biblioteca de DNA *in vitro* usando a T7 RNA polimerase ou uma T7 RNA polimerase modificada, e purificando a biblioteca transcrita. A biblioteca de RNA ou DNA é então misturada com o alvo, sob condições favoráveis para a ligação, e submetida a iterações em etapas de ligação, separação e amplificação, usando o mesmo esquema de seleção geral, para obter virtualmente qualquer critério desejado de afinidade de ligação e seletividade. Mais especificamente, começando com uma mistura contendo o agrupamento inicial de ácidos nucleicos, o método SELEX[®] inclui as etapas de: (a) contatar a mistura com o alvo, sob condições favoráveis para a ligação; (b) separar os ácidos nucleicos não ligados daqueles ácidos nucleicos que tenham se ligado especificamente às moléculas-alvo; (c) opcionalmente dissociar os complexos de ácidos nucleicos-alvo; (d) amplificar os ácidos nucleicos dissociados dos complexos de ácidos nucleicos-alvo, para produzir uma mistura de ácidos nucleicos enriquecida em ligantes; e (e) repetir as etapas de ligar, separar, dissociar e amplificar através de tantos ciclos quanto desejados, para produzir ligantes de ácidos nucleicos altamente específicos, de alta afinidade, para a molécula-alvo. Nas situações onde os aptâmeros de RNA estiverem sendo selecionados, o método SELEX[®] adicionalmente compreende as etapas de: (i) transcrição reversa dos ácidos nucleicos dissociados dos complexos de ácidos nucleicos-alvo,

antes da amplificação na etapa (d); e (ii) transcrição dos ácidos nucléicos amplificados da etapa (d) antes de reiniciar o processo.

Dentro de uma mistura de ácidos nucléicos contendo
5 um grande número de seqüências e estruturas possíveis, há uma ampla faixa de afinidades de ligação por um dado alvo. Uma mistura de ácidos nucléicos compreendendo, por exemplo, um segmento ao acaso de 20 nucleotídeos, pode ter 4^{20} possibilidades de candidatos. Aqueles que tiverem a maior
10 afinidade (menores constantes de dissociação) pelo alvo são mais prováveis de ligarem-se ao alvo. Após a separação, a dissociação e a amplificação, uma segunda mistura de ácidos nucléicos é gerada, enriquecida pelos candidatos de maior afinidade de ligação. As rodadas adicionais de seleção
15 progressivamente favorecem os melhores ligantes, até que a mistura de ácidos nucléicos resultante seja predominantemente composta de somente uma ou poucas seqüências. Essas podem então ser clonadas, seqüenciadas e individualmente testadas como ligantes ou aptâmeros quanto à
20 1) afinidade de ligação pelo alvo; e/ou 2) capacidade de efetuar a função de alvo.

Os ciclos de seleção e amplificação são repetidos até que um objetivo desejado seja atingido. No caso mais geral, a seleção/amplificação é continuada até que nenhum
25 aperfeiçoamento significativo na intensidade de ligação seja atingido na repetição do ciclo. O método é tipicamente usado para amostrar aproximadamente 10^{14} espécies diferentes de ácidos nucléicos, porém pode ser usado para amostrar

tanto quanto cerca de 10^{18} espécies diferentes de ácidos nucleicos. Geralmente, as moléculas de aptâmeros de ácidos nucleicos são selecionadas em um procedimento de 5 a 20 ciclos. Em uma modalidade, a heterogeneidade é introduzida
5 somente nos estágios iniciais de seleção e não ocorre por todo o processo de replicação.

Em uma modalidade do método SELEX[®], o processo de seleção é tão eficiente em isolar aqueles ligantes de ácidos nucleicos que se ligam mais fortemente ao alvo selecionado,
10 que somente um ciclo de seleção e amplificação é requerido. Tal seleção eficiente pode ocorrer, por exemplo, em um processo do tipo cromatográfico, onde a capacidade dos ácidos nucleicos de associarem-se com alvos ligados sobre uma coluna opera em tal maneira que a coluna é
15 suficientemente capaz de permitir a separação e o isolamento dos ligantes de ácidos nucleicos de mais altas afinidades.

Em muitos casos, não é necessariamente desejável efetuar as etapas iterativas do processo SELEX[®] até um único ligante de ácido nucleico ser identificado. A solução de
20 ligante de ácido nucleico específico para o alvo pode incluir uma família de estruturas ou motivos de ácidos nucleicos que tem diversas seqüências conservadas e diversas seqüências que podem ser substituídas ou adicionadas sem afetar significativamente a afinidade dos ligantes de ácidos
25 nucleicos pelo alvo. Pela terminação do processo SELEX[®] antes do término, é possível determinar a seqüência de diversos membros da família da solução de ligante de ácido nucleico.

Sabe-se que existe uma variedade de estruturas primárias, secundárias e terciárias de ácidos nucléicos. As estruturas ou motivos que tenham sido mostrados, de modo mais comum, estarem envolvidos nas interações que não sejam do tipo Watson-Crick são referidos como agrupamentos do tipo grampo de cabelo ("hairpin"), saliências simétricas e assimétricas, pseudo-nós e um grande número de combinações destes. Quase todos os casos conhecidos de tais motivos sugerem que eles podem ser formados em uma seqüência de ácidos nucléicos de não mais do que 30 nucleotídeos. Por essa razão, é frequentemente preferido que os procedimentos SELEX® com segmentos ao acaso contíguos sejam iniciados com seqüências de ácidos nucléicos contendo um segmento ao acaso de entre cerca de 20 a cerca de 50 nucleotídeos e, em algumas modalidades, cerca de 30 a cerca de 40 nucleotídeos. Em um exemplo, a seqüência fixa em 5':aleatória:fixa em 3' compreende uma seqüência aleatória de cerca de 30 a cerca de 40 nucleotídeos.

O método SELEX® básico tem sido modificado para atingir diversos objetivos específicos. Por exemplo, a Patente U.S. Nº 5.707.796 descreve o uso do processo SELEX® em conjunção com a eletroforese em gel, para selecionar moléculas de ácidos nucléicos com características estruturais específicas, tais como DNA dobrado. A Patente U.S. Nº 5.763.177 descreve métodos baseados no SELEX® para selecionar ligantes de ácidos nucléicos contendo grupos fotorreativos, capazes de ligar e/ou foto-reticular com e/ou fotoinativar uma molécula-alvo. A Patente U.S. Nº 5.567.588

e a Patente U.S. Nº 5.861.254 descrevem métodos baseados no SELEX® que atingem uma separação altamente eficiente entre os oligonucleotídeos tendo alta e baixa afinidade por uma molécula-alvo. A Patente U.S. Nº 5.496.938 descreve métodos
5 para obter ligantes de ácidos nucleicos aperfeiçoados após o processo SELEX® ter sido efetuado. A Patente U.S. Nº 5.705.337 descreve métodos para ligar covalentemente um ligante ao seu alvo.

O método SELEX® pode também ser usado para obter
10 ligantes de ácidos nucleicos que se ligam a mais do que um sítio sobre a molécula-alvo, e para obter ligantes de ácidos nucleicos que incluam espécies que não sejam ácidos nucleicos que se ligam a sítios específicos sobre o alvo. O método SELEX® proporciona meios para isolar e identificar
15 ligantes de ácidos nucleicos que se ligam a qualquer alvo imaginável, incluindo as biomoléculas grandes e pequenas, tais como as proteínas de ligação a ácidos nucleicos e as proteínas não sabidas ligarem ácidos nucleicos como parte de sua função biológica, bem como os cofatores e outras
20 pequenas moléculas. Por exemplo, a Patente U.S. Nº 5.580.737 divulga seqüências de ácidos nucleicos, identificadas através do método SELEX®, que são capazes de se ligarem com alta afinidade à cafeína e ao análogo intimamente relacionado, teofilina.

25 O processo Counter-SELEX® é um método para aperfeiçoar a especificidade dos ligantes de ácidos nucleicos para uma molécula-alvo, por eliminação das seqüências de ligantes de ácidos nucleicos com reatividade

cruzada com uma ou mais moléculas que não sejam alvos. O processo Counter-SELEX® é compreendido das etapas de: (a) preparar uma mistura candidata de ácidos nucleicos; (b) contatar a mistura candidata com o alvo, onde os ácidos nucleicos tendo uma afinidade aumentada pelo alvo em relação à mistura candidata podem ser separados do restante da mistura candidata; (c) separar os ácidos nucleicos de afinidade aumentada do restante da mistura candidata; (d) opcionalmente dissociar os ácidos nucleicos de afinidade aumentada do alvo; (e) contatar os ácidos nucleicos de afinidade aumentada com uma ou mais moléculas que não sejam alvos, de modo tal que os ligantes de ácidos nucleicos com afinidade específica pela(s) molécula(s) que não é(são) alvo(s) sejam removidos; e (f) amplificar os ácidos nucleicos com afinidade específica somente pela molécula-alvo, para produzir uma mistura de ácidos nucleicos enriquecida por seqüências de ácidos nucleicos com afinidade relativamente maior e especificidade para ligação à molécula-alvo. Conforme descrito acima para o método SELEX®, os ciclos de seleção e amplificação são repetidos, conforme necessários, até que um objetivo desejado seja atingido.

Um problema potencial encontrado no uso dos ácidos nucleicos como substâncias terapêuticas e vacinas é que os oligonucleotídeos em sua forma de fosfodiéster podem ser rapidamente degradados nos fluidos corpóreos por enzimas intracelulares e extracelulares, tais como as endonucleases e as exonucleases, antes do efeito desejado ser mostrado. O método SELEX®, assim, inclui a identificação de ligantes de

ácidos nucléicos de altas afinidades contendo nucleotídeos modificados que conferem características aperfeiçoadas ao ligante, tais como estabilidade *in vivo* aperfeiçoada ou características de distribuição aperfeiçoadas. Os exemplos

5 de tais modificações incluem as substituições químicas nas posições de açúcar e/ou fosfato e/ou base. Os ligantes de ácidos nucléicos identificados pelo SELEX[®] contendo nucleotídeos modificados são descritos, p.ex., na Patente U.S. Nº 5.660.985, que descreve oligonucleotídeos contendo

10 derivados de nucleotídeos quimicamente modificados na posição 2' da ribose, posição 5 da pirimidinas, e posição 8 das purinas, na Patente U.S. Nº 5.756.703, que descreve oligonucleotídeos contendo diversas pirimidinas modificadas em 2', e na Patente U.S. Nº 5.580.737, que descreve ligantes

15 de ácidos nucléicos altamente específicos contendo um ou mais nucleotídeos modificados com substituintes de 2'-amino (2'-NH₂), 2'-flúor (2'-F), e/ou 2'-O-metila (2'-OMe).

As modificações dos ligantes de ácidos nucléicos contemplados nessa invenção incluem, porém não estão

20 limitadas, aquelas que proporcionam outros grupos químicos que incorporam carga, polarizabilidade, hidrofobicidade, ligação de hidrogênio, interação eletrostática, e variabilidade adicionais às bases do ligante de ácido nucléico ou ao ligante de ácido nucléico como um todo. As

25 modificações para gerar populações de oligonucleotídeos que sejam resistentes às nucleases podem também incluir uma ou mais ligações entre nucleotídeos substitutas, açúcares alterados, bases alteradas, ou suas combinações. Tais

modificações incluem, porém não estão limitadas às, modificações do açúcar na posição 2', modificações da pirimidina na posição 5, modificações da purina na posição 8, modificações em aminas exocíclicas, substituição da 4-tiouridina, substituição da 5-bromo ou 5-iodo-uracil; modificações na cadeia principal, modificações de fosforotioato ou fosfato de alquila, metilações, e combinações incomuns de emparelhamento de bases, tais como as isobases isocitidina e isoguanosina. As modificações podem também incluir as modificações em 3' e 5', tais como o quepi. As modificações podem também incluir as modificações em 3' e 5', tais como o quepi, p.ex., a adição de um quepi de 3'-3'-dT para aumentar a resistência à exonucelase (ver, p.ex., as Patentes U.S. 5.674.685; 5.668.264; 6.207.816; e 6.229.002, cada uma das quais é incorporada por referência nesse documento em sua totalidade).

Em uma modalidade, proporcionam-se oligonucleotídeos nos quais o grupo $P(O)O$ é substituído por $P(O)S$ ("tioato"), $P(S)S$ ("ditioato"), $P(O)NR_2$ ("amidato"), $P(O)R$, $P(O)OR'$, CO ou CH_2 ("formacetal") ou 3'-amina ($-NH-CH_2-CH_2-$), onde cada R ou R' é independentemente H ou alquila substituída ou não substituída. Os grupos de ligação podem ser unidos a nucleotídeos adjacentes através de uma ligação $-O-$, $-N-$, ou $-S-$. Nem todas as ligações no oligonucleotídeo são requeridas serem idênticas.

Nas modalidades adicionais, os oligonucleotídeos compreendem grupos de açúcares modificados, por exemplo, um ou mais dos grupos hidroxila são substituídos com halogênio,

grupos alifáticos, ou funcionalizados como éteres ou aminas. Em uma modalidade, a posição 2' do resíduo de furanose é substituída por qualquer de um grupo O-metila, O-alquila, O-alila, S-alquila, S-alila, ou halo. Os métodos de síntese de açúcares modificados em 2' são descritos, p.ex., em Sproat, e col., Nucl. Acid Res. 19:733-738 (1991); Cotten, e col. Nucl. Acid Res. 19:2629-2635 (1991); e Hobbs, e col., Biochemistry 12:5138-5145 (1973). As outras modificações são conhecidas para alguém de habilidade comum na técnica. Tais modificações podem ser as modificações pré-processo SELEX® ou as modificações pós-processo SELEX® (modificação de ligantes não modificados previamente identificados) ou podem ser feitas por incorporação no processo SELEX®.

As modificações pré-processo SELEX® ou aquelas feitas por incorporação no processo SELEX® produzem ligantes de ácidos nucleicos tanto com especificidade por seu alvo SELEX®, quanto com estabilidade aperfeiçoada, p.ex., estabilidade *in vivo*. As modificações pós-processo SELEX® (p.ex., truncamento, deleção, substituição ou modificações adicionais de nucleotídeos de ligantes previamente identificados, tendo nucleotídeos incorporados por modificação pré-processo SELEX®) aos ligantes de ácidos nucleicos podem resultar em estabilidade aperfeiçoada, p.ex., estabilidade *in vivo*, sem afetar adversamente a capacidade de ligação do ligante de ácido nucleico. Opcionalmente, os aptâmeros em que os nucleotídeos modificados tenham sido incorporados por modificação pré-processo SELEX® podem ser adicionalmente modificados por modificação pós-processo

SELEX[®] (i.e., um pós-processo de modificação SELEX[®] após o SELEX).

O método SELEX[®] inclui combinar oligonucleotídeos selecionados com outros oligonucleotídeos selecionados e
5 unidades funcionais que não sejam oligonucleotídeos, como descrito na Patente U.S. Nº 5.637.459 e na Patente U.S. Nº 5.683.867. O método SELEX[®] adicionalmente inclui combinar ligantes de ácidos nucléicos selecionados com compostos de alto peso molecular, lipofílicos ou não imunogênicos, em um
10 complexo diagnóstico ou terapêutico, como descrito, p.ex., na Patente U.S. Nº 6.011.020, na Patente U.S. Nº 6.051.698, e na Publicação PCT Nº WO 98/18480. Essas patentes e pedidos ensinam a combinação de um amplo arranjo de formatos e outras propriedades, com as propriedades eficientes de
15 amplificação e replicação de oligonucleotídeos, e com as propriedades desejáveis de outras moléculas.

A identificação de ligantes de ácidos nucléicos para peptídeos flexíveis, pequenos, via o método SELEX[®], também tem sido explorada. Os peptídeos pequenos têm
20 estruturas flexíveis e normalmente existem em solução em um equilíbrio de confôrmeros múltiplos e, desse modo, acreditava-se inicialmente que as afinidades de ligação poderiam ser limitadas pela entropia conformacional perdida com a ligação de um peptídeo flexível. Entretanto, a
25 viabilidade de identificar ligantes de ácidos nucléicos para peptídeos pequenos em solução foi demonstrada na Patente U.S. Nº 5.648.214. Nessa patente, foram identificados ligantes de ácidos nucléicos de RNA de alta afinidade para a

substância P, um peptídeo de 11 aminoácidos.

Como parte do processo SELEX®, as seqüências selecionadas para ligarem-se ao alvo são então opcionalmente minimizadas para determinar a seqüência mínima tendo a
5 afinidade de ligação desejada. As seqüências selecionadas e/ou as seqüências minimizadas são opcionalmente modificadas efetuando-se a mutagênese aleatória ou direcionada da seqüência para, p.ex., aumentar a afinidade de ligação ou
alternativamente determinar quais posições na seqüência são
10 essenciais para a atividade de ligação.

O MÉTODO SELEX® MODIFICADO EM 2'

Para que um aptâmero seja adequado para uso como uma substância terapêutica e/ou para tipos particulares de substâncias diagnósticas, ele é preferivelmente barato de
15 sintetizar, seguro e estável *in vivo*. Os aptâmeros de RNA e DNA do tipo selvagem tipicamente não são estáveis *in vivo* por causa de sua suscetibilidade à degradação por nucleases. A resistência à degradação por nucleases pode ser bastante aumentada pela incorporação de grupos modificadores na
20 posição 2'.

Os grupos 2'-flúor e 2'-amino têm sido incorporados com êxito nos agrupamentos de oligonucleotídeos a partir dos quais os aptâmeros têm sido subsequentemente selecionados. Entretanto, essas modificações aumentam muito
25 o custo de síntese do aptâmero resultante, e podem introduzir preocupações de segurança em alguns casos por causa da possibilidade que os nucleotídeos modificados poderiam ser reciclados para o DNA hospedeiro por degradação

dos oligonucleotídeos modificados e uso subsequente dos nucleotídeos como substratos para a síntese de DNA.

Os aptâmeros que contêm 2'-O-metil ("2'-OMe") nucleotídeos, como proporcionados nesse documento, superam
5 muitas dessas desvantagens. Os oligonucleotídeos contendo 2'-OMe nucleotídeos são resistentes às nucleases e baratos de sintetizar. Embora os 2'-OMe nucleotídeos sejam ubíquos nos sistemas biológicos, as polimerases naturais não aceitam os 2'-OMe NTPs como substratos sob condições fisiológicas,
10 assim não há nenhuma preocupação de segurança sobre o reciclo de 2'-OMe nucleotídeos para o DNA hospedeiro. Os métodos SELEX[®] usados para gerar aptâmeros modificados em 2' são descritos, p.ex., no Pedido de Patente Provisório U.S. Nº de Série 60/430.761, depositado em 3 de dezembro de 2002,
15 no Pedido de Patente Provisório U.S. Nº de Série 60/487.474, depositado em 15 de julho de 2003, no Pedido de Patente Provisório U.S. Nº de Série 60/517.039, depositado em 4 de novembro de 2003, no Pedido de Patente U.S. Nº 10/729.581, depositado em 3 de dezembro de 2003, e no Pedido de Patente
20 U.S. Nº 10/873.856, depositado em 21 de junho de 2004, intitulado "Method for *in vitro* Selection of 2'-O-methyl Substituted Nucleic Acids", cada um dos quais é, nesse documento, incorporado por referência em sua totalidade.

A presente invenção inclui os aptâmeros que se
25 ligam ao, e modulam a função do, alvo do aptâmero e que contêm nucleotídeos modificados (p.ex., nucleotídeos que têm uma modificação na posição 2') para tornar o oligonucleotídeo mais estável do que o oligonucleotídeo não

modificado à degradação enzimática e química, bem como à degradação térmica e física. Embora existam diversos exemplos de aptâmeros contendo 2'-OMe na literatura (ver, p.ex., Ruckman e col., J.Biol.Chem, 1998 273, 20556-20567-
5 695), esses foram gerados pela seleção *in vitro* de bibliotecas de transcritos modificados em que os resíduos C e U foram substituídos por 2'-flúor (2'-F) e os resíduos A e G foram por 2'-OH. Assim que as seqüências funcionais foram identificadas, então cada resíduo A e G foi testado quanto à
10 tolerância à substituição com 2'-OMe, e o aptâmero foi sintetizado novamente tendo todos os resíduos A e G, os quais toleravam a substituição com 2'-OMe como resíduos com 2'-OMe. A maior parte dos resíduos A e G de aptâmeros gerados nesse modo de duas etapas tolera a substituição por
15 resíduos com 2'-OMe, embora, em média, aproximadamente 20% não. Consequentemente, os aptâmeros gerados usando esse método tendem a conter de dois a quatro resíduos como 2'-OH, e a estabilidade e o custo de síntese estão comprometidos como um resultado. Pela incorporação de nucleotídeos
20 modificados na reação de transcrição, que geram oligonucleotídeos estabilizados, usados nos agrupamentos de oligonucleotídeos a partir dos quais os aptâmeros são selecionados e enriquecidos pelo método SELEX® (e/ou quaisquer de suas variações e aperfeiçoamentos, incluindo
25 aqueles descritos nesse documento), os métodos da presente invenção eliminam a necessidade por estabilização dos oligonucleotídeos para aptâmeros selecionados por síntese novamente dos oligonucleotídeos de aptâmeros com os

nucleotídeos modificados com 2'-OMe.

Em uma modalidade, a presente invenção proporciona aptâmeros compreendendo combinações de modificações com 2'-OH, 2'-F, 2'-desóxi, e 2'-OMe dos nucleotídeos de ATP, GTP, CTP, TTP, e UTP. Em uma outra modalidade, a presente invenção proporciona aptâmeros compreendendo combinações de modificações com 2-OH, 2'-F, 2'-desóxi, 2'-OMe, 2'-NH₂, e 2'-metoxietila dos nucleotídeos de ATP, GTP, CTP, TTP, e UTP. Em uma modalidade preferida, a presente invenção proporciona aptâmeros compreendendo nucleotídeos de ATP, GTP, CTP, TTP, e/ou UTP totalmente ou substancial e totalmente modificados com 2'-OMe.

Polimerases Modificadas

Os aptâmeros modificados em 2' da invenção são criados usando polimerases modificadas, p.ex., uma T7 polimerase modificada, tendo uma taxa de incorporação de nucleotídeos modificados que têm substituintes grandes na posição 2' da furanose que é mais elevada do que das polimerases do tipo selvagem. Por exemplo, uma T7 polimerase mutante em que o resíduo de tirosina na posição 639 tenha sido alterado para fenilalanina (Y639F) prontamente utiliza trifosfatos de 2'-desóxi, 2'-amino-, e 2'-flúor-nucleotídeos (NTPs) como substratos e tem sido amplamente usada para sintetizar RNAs modificados para uma variedade de aplicações. Entretanto, essa T7 polimerase mutante, ao que consta, não pode prontamente utilizar (i.e., incorporar) os NTPs com substituintes grandes em 2', tais como os substituintes de 2'-OMe ou 2'-azido (2'-N₃). Para a

incorporação de substituintes grandes em 2', uma T7 polimerase tendo a histidina na posição 784 alterada para um resíduo de alanina, além da mutação de Y639F, foi descrita (Y639F/H784A) e foi usada em circunstâncias limitadas para
5 incorporar NTPs com pirimidina modificada. Ver Padilla, R, e Sousa, R., *Nucleic Acids Res.*, 2002, 30(24): 138. Uma T7 RNA polimerase mutante, na qual o resíduo de tirosina na posição 639 tenha sido alterado para fenilalanina, o resíduo de histidina na posição 784 foi alterado para uma alanina, e
10 o resíduo de lisina na posição 378 foi alterado para arginina (Y639F/H784A/K378R), tem sido usada em circunstâncias limitadas para incorporar os NTPs com purina e piridimina modificadas, p.ex, os 2'-OMe NTPs, porém inclui uma adição de 2'-OH GTP para a transcrição. Ver Burmeister
15 e col, (2005) *Chemistry and Biology*, 12: 25-33. A inclusão de uma adição de 2'-OH GTP para a transcrição pode resultar em aptâmeros que não sejam inteiramente com 2'-OMe, porém preferivelmente pode depender da presença de 2-OH GTPs.

Uma T7 polimerase mutante tendo a histidina na
20 posição 784 alterada para um resíduo de alanina (H784A) também foi descrita. Padilla e col., *Nucleic Acids Research*, 2002, 30: 138. Em ambas as T7 polimerases Y639F/H784A mutante e H784A mutante, a alteração para um resíduo de aminoácido menor, tal como a alanina, permite a incorporação
25 de substratos de nucleotídeos maiores, p.ex., os nucleotídeos substituídos com 2'-OMe. Ver Chelliserry, K. e Ellington, A.D., (2004) *Nature Biotech*, 9:1155-60. T7 RNA polimerases adicionais foram descritas com mutações no sítio

ativo da T7 RNA polimerase, que incorporam mais prontamente substratos modificados em 2' maiores, p.ex., uma T7 RNA polimerase mutante tendo o resíduo de tirosina na posição 639 alterado para uma leucina (Y639L). Entretanto, a
5 atividade é frequentemente prejudicada pela especificidade do substrato aumentada conferida por tais mutações, resultando em baixas produções de transcritos. Ver Padilla R e Sousa, R., (1999) *Nucleic Acids Res.*, 27(6): 1561. A mutante de T7 RNA polimerase P266L foi descrita para
10 facilitar a depuração do promotor (Guillerez e col. (2005) *Proc. Nat. Acad. Sci. USA*, 102 (17) 5958). A polimerase faz uma transição da conformação inicial, em que está ligada ao promotor, para a conformação de alongamento, em que não está. Nenhuma das polimerases mutantes acima mencionadas foi
15 descrita resultar em transcritos inteiramente com 2'-OMe.

A presente invenção proporciona materiais e métodos para aumentar o rendimento de transcrição de oligonucleotídeos. Em uma modalidade, a presente invenção proporciona métodos e condições para utilizar as T7 RNA
20 polimerases modificadas para incorporar enzimaticamente nucleotídeos modificados em oligonucleotídeos. Em uma modalidade preferida, a T7 RNA polimerase modificada, usada com os métodos de transcrição da invenção, não requer a presença de 2'-OH GTP. Em uma modalidade preferida, a
25 polimerase modificada é uma T7 RNA polimerase mutante tendo o resíduo de tirosina na posição 639 alterado para um resíduo de leucina e o resíduo de histidina na posição 784 alterado para um resíduo de alanina (Y639L/H784A). Em uma

outra modalidade preferida, a polimerase modificada é uma T7 RNA polimerase mutante tendo o resíduo de tirosina na posição 639 alterado para um resíduo de leucina, o resíduo de histidina na posição 784 alterado para um resíduo de alanina, e o resíduo de lisina na posição 378 alterado para um resíduo de arginina (Y639L/H784A/K378R). Em uma outra modalidade, a polimerase modificada para uso nos métodos da invenção é uma T7 RNA polimerase mutante tendo o resíduo de tirosina na posição 639 alterado para uma leucina (Y639L), enquanto que em mais uma outra modalidade, a T7 RNA polimerase mutante tem o resíduo de tirosina na posição 639 alterado para um resíduo de leucina e o resíduo de lisina na posição 378 alterado para um resíduo de arginina (Y639L/K378R). Embora não desejando estar ligado por qualquer teoria, a mutação de K378R não está próxima ao sítio ativo da polimerase e, desse modo, acredita-se ser uma mutação silenciosa. Em uma outra modalidade, a polimerase modificada para uso nos métodos da invenção é uma T7 RNA polimerase mutante tendo o resíduo de prolina na posição 266 alterado para uma leucina, o resíduo de tirosina na posição 639 alterado para uma leucina e o resíduo de histidina na posição 784 alterado para um resíduo de alanina (P266L/Y639L/H784A), enquanto que em mais uma outra modalidade, a T7 RNA polimerase mutante tem o resíduo de prolina na posição 266 alterado para uma leucina, o resíduo de tirosina na posição 639 alterado para um resíduo de leucina, o resíduo de histidina na posição 784 alterado para um resíduo de alanina e o resíduo de lisina na posição 378

alterado para um resíduo de arginina
(P266L/Y639L/H784A/K378R).

As seqüências de aminoácidos das T7 RNA
polimerases mutantes são mostradas abaixo:

5 Y639L/H784A (SEQ ID NO 1):
MNTINIAKNDFSDIELAAIPFNTLADHYGERLAREQLALEHESYEMGEARFRKMFERQLK
AGEVADNA
AAKPLITTLLPKMIARINDWFEEVKAKRGKRPTAFQFLQEIKPEAVAYITIKTTLACLTS
ADNTTVQAVAS
10 AIGRAIEDEARFGRIRDLEAKHFKNVEEQLNKRVGHVYKKAQVVEADMLSKGLLGGE
AWSSWHK
EDSIHVGVRCEMLIESTGMVSLHRQNAGVVGQDSETIELAPEYAEAIATRAGALAGISP
MFQPCVVPP
KPWTGITGGGYWANGRRPLALVRTHSKKALMRYEDVYMPEVYKAINIAQNTAWKINKKVL
15 AVANVITK
WKHCPVEDIPAIEREELPMKPEDIDMNPEALTAWKRAAAVYRKDKARKSRRISLEFMLE
QANKFANH
KAIWFPYNMDWRGRVYAVSMFNPQGNDMTKGLLTLAKGKPIGKEGYWLKIHGANCAGVD
KVPFPE
20 RIKFIEENHENIMACAKSPLENTWWAEQDSPFCFLAFCFEYAGVQHHGLSYNCSLPLAFD
GSCSGIQH
FSAMLRDEVGGRVNNLLPSETVQDIYGIVAKKVNEILQADAINGTDNEVVTVTDENTGEI
SEKVKLGTK
ALAGQWLAYGVTRSVTKRSVMTLALGSKEFGFRQQVLEDTIQPAIDSGKGLMFTQPNQAA
25 GYMAKLI
WESVSVTVVAAVEAMNWLKSAKLLAAEVKDKKTGEILRKRCVHVWTPDGFPVWQEYKK
PIQTRLN
LMFLGQFRLQPTINTNKDSEIDAHKQESGIAPNFVASQDGSHLRKTVVWAHEKYGIESFA

LIHDSFGTI

PADAANLFKAVRETMVDTYESCDVLADFYDQFADQLHESQLDKMPALPAKGNNLRDILE

SDFAFA

5 Y639L/H784A/K378R (SEQ ID NO 2):

MNTINIAKNDFSDIELAAIPFNTLADHYGERLAREQLALEHESYEMGEARFRKMFERQLK

AGEVADNA

AAKPLITTLTPKMIARNDFEVEVKAKRGKRPTAFQFLQEIKEAVAYITIKTTLACLTSA

DNTTVQAVAS

10 AIGRAIEDEARFGRIRDLEAKHFKNVEEQLNKRVGHVYKKAQVVEADMLSKGLLGGE

AWSSWHK

EDSIHVGVRCEMLIESTGMVSLHRQAGVVGQDSEITELAPEYAEAIATRAGALAGISP

MFQPCVVPP

KPWTGITGGGWANGRRPLALVRTHSKKALMRYEDVYMPEVYKAINIAQNTAWKINKKVL

15 VANVITK

WKHCPVEDIPAIEREELPMKPEDIDMNPEALTAWRRAAAAYRKDKARKSRRISLEFMLE

QANKFANH

KAIWFPYNMDWRGRVYAVSMFNPQGNMTKGLLTLAKGKPIGKEGYWLKIHGANCAGVD

KVPFPE

20 RIKFIEENHENIMACAKSPLENTWWAEQDSPFCFLAFCFEYAGVQHHGLSYNCSLPLAFD

GSCSGQH

FSAMLRDEVGGRAVNLLPSETVQDIYGIVAKKVNEILQADAINGTDNEVVTVTDENTGEI

SEKVKLGTK

ALAGQWLAYGVTRSVTKRSVMTLALGSKEFGFRQQVLEDTIQPAIDSGKGLMFTQPNQAA

25 GYMAKLI

WESVSVTVVAAVEAMNWLKSAKLLAAEVKDKKTGEILRKRCVHVWVTPDGFPVWQYK

PIQTRLN

LMFLGQFRLQPTINTNKDSEIDAHKQESGIAPNFVASQDGSRLRKTVVWAHEKYGIESFA

LIHDSFGTI

PADAANLFKAVRETMVDTYESCDVLADFYDQFADQLHESQLDKMPALPAKGNLNLRDILE

SDFAFA

5 Y639L (SEQ ID NO 100):

MNTINIAKNDFSDIELAAIPFNTLADHYGERLAREQLALEHESYEMGEARFRKMFERQLK

AGEVADNA

AAKPLITTLTPKMIARINDWFEEVKAKRGKRPTAFQFLQEIKPEAVAYITIKTTLACLTS

ADNTTVQAVAS

10 AIGRAIEDEARFGRIRDLEAKHFKNVEEQLNKRVGHVYKKAQVVEADMLSKGLLGGE

AWSSWHK

EDSHVGVRCIEMLIESTGMVSLHRQAGVVGQDSETIELAPEYAEAIATRAGALAGISPM

FQPCVPP

KPWTGITGGGYWANGRRPLALVRTHSKKALMRYEDVYMPEVYKAINIAQNTAWKINKVL

15 AVANVTK

WKHCPVEDIPAIEREELPMKPEDIDMNPEALTAWKRAAAVYRKDKARKSRRISLFMLEQ

ANKFANH

KAIWFPYNMDWRGRVYAVSMFNPQGNDMTKGLLTAKGKPIGKEGYWLKIHGANCAGVD

KVPFPE

20 RIKFIEENHENIMACAKSPLENTWWAEQDSPFCFLAFCFEYAGVQHHGLSYNCSLPLAFD

GSCSGIQH

FSAMLRDEVGGRAVNLLPSETVQDIYGIVAKVNEILQADAINGTDNEVVTVTDENTGEI

SEKVKLGTK

ALAGQWLAYGVTRSVTKRSVMTLALGSKEFGFRQQVLEDTIQPAIDSGKGLMFTQPNQAA

25 GYMAKLI

WESVSVTVVAAVEAMNWLKSAAKLLAAEVKDKKTGEILRKRCVHVWTPDGFVPWQEYKK

PIQTRLN

LMFLGQFRLQOPTINTNKDSEIDAHKQESGIAPNFVHSQDGSHLRKTVVWAHEKYGIESFA

LIHDSFGTI

PADAANLFKAVRETMVDTYESCDVLADFYDQFADQLHESQLDKMPALPAKGNLNLRDILE

SDFafa

5 Y639L/K378R (SEQ ID NO 101):

MNTINIAKNDFSDIELAAIPFNTLADHYGERLAREQLALEHESYEMGEARFRKMFERQLK

AGEVADNA

AAKPLITTLPKMIARINDWFEEVKAKRGKRPTAFQFLQEIKPEAVAYITIKTTLACLTS

ADNTTVQAVAS

10 AGRAIEDEARFGRIRDLEAKHFKNVVEQLNKRVGHVYKKAQFMQWEADMLSKGLLGGEAW

SSWHK

EDSIHVGVRCEIEMLIESTGMVSLHRQNAGVVGQDSETIELAPEYAEAPATRAGALAGISP

MFQPCVVPP

KPWTGITGGGYWANGRRPLALVRTHSKKALMRYEDVYMPEVYKAINIAQNTAWKINKVL

15 AVANVTTK

WKHCPVEDIPAIEREELPMKPEDIDMNPEALTAWRRRAWVYRKDKARKSRRISLEFMLEQA

NKFANH

KAIWFPYNMDWRGRVYAVSMFNPQGNMTKGLLTLAKGKPIGKEGYWLKIHGANCAGVD

KVPFPE

20 RIKFIEENHENIMACAKSPLENTWWAEQDSPFCFLAFCFEYAGVQHHGLSYNCSLPLAFD

GSCSGQH

FSAMLRDEVGGRAVNLLPSETVQDIYGVAKKVNEILQADAINGTDNEWTVDENTGEISE

KVKLGTK

ALAGQWLAYGVTRSVTKRSVMTLALGSKEFGFRQQVLEDTIQPAIDSGKGLMFTQPNQAA

25 GYMAKLI

WESVSVTVVAAVEAMNWLKSAAKLLAAEVKDKKTGEILRKRCVHWVTPDGFPVWQEYKK

PIQTRLN

LMFLGQFRLQPTINTNKDSEIDAHKQESGIAPNFVHSQDGSHLRKTVVWAHEKYGIESFA

LIHDSFGTI

PADAANLFKAVRETMVDTYESCDVLADFYDQFADQLHESQLDKMPALPAKGNLNLRDLES

DFAFA

5 P266L/Y639L/H784A (SEQ ID NO 102)

MNTINIAKNDFSDIELAAIPFNTLADHYGERLAREQLALEHESYEMGEARFRKMFERQLK

AGEVADNA

AAKPLITLLPKMTARINDWFEEVKAKRGKRPTAFQFLQEIKPEAVAYITIKTTLACLTS

ADNTTVQAVAS

10 AIGRAIEDEARFGRIRDLEAKHFKNVEEQLNKRVGHVYKKAQVVEADMLSKGLLGGE

AWSSWHK

EDSIHVGVRCEIEMUESTGMVSLHRQNAGVVGQDSEITELAPEYAEAIATRAGALAGISLM

FQPCVVPP

KPWTGITGGGYWANGRRPLALVRTHSKKALMRYEDVYMPEVYKAINIAQNTAWKINKKVL

15 AVANVITK

WKHCPVEDIPAIEREELPMKPEDIDMNPEALTAWKRAAIAVYRKDKARKSRRISLEFMLE

QANKFANH

KAIWFPYNMDWRGRVYAVSMFNPQGNMTKGLLTLAKGKPIGKEGYWLKIHGANCAGVD

KVPFPE

20 RIKFIEENHENIMACAKSPLENTWWAEQDSPFCFLAFCFEYAGVQHHGLSYNCSLPLAFD

GSCSGIQH

FSAMLRDEVGGRAVNLLPSETVQDIYGIVAKKVNEILQADAINGTDNEVVTVTDENTGEI

SEKVKLGTK

ALAGQWLAYGVTRSVTKRSVMTLALGSKEFGFRQQVLEDTIQPAIDSGKGLMFTQPNQAA

25 GYMAKLI

WESVSVTVVAAVEAMNWLKSAAKLLAEVKDKKTGEILRKRCVHVWTPDGFPVWQEYKKP

IQTRLN

LMFLGQFRLQPTINTNKDSEIDAHKQESGIAPNFVASQDGSHLRKTVVWAHEKYGIESFA

LIHDSFGTI

PADAANLFKAVRETMVDTYESCDVLADFYDQFADQLHESQLDKMPALPAKGNLNLRDILE

SDFafa

5 P266L/Y639L/H784A/ K378R (SEQ ID NO 103)

MNTINIAKNDFSDIELAAIPFNTLADHYGERLAREQLALEHESYEMGEARFRKMFERQLK

AGEVADNA

AAKPLITTLTPKMIARINDWFEEVKAKRGKRPTAFQFLQEIKPEAVAYITIKTTLACLTS

ADNTTVQAVAS

10 AIGRAIEDEARFGRIRDLEAKHFKNVEEQLNKRVGHVYKKAQVVEADMLSKGLLGGE

AWSSWHK

EDSIHVGVRCEIEMLIESTGMVSLHRQAGVVGQDSEITELAPEYAEAIATRAGALAGISL

MFQPCVVPP

KPWTGITGGGYWANGRRPLALVRTHSKKALMRYEDVYMPEVYKAINIAQNTAWKINKKVL

15 AVANVITK

WKHCPVEDIPAIEREELPMKPEDIDMNPEALTAWRRAAAAYRKDKARKSRRISLEFMLE

QANKFANH

KAIWFPYNMDWRGRVYAVSMFNPQGNMTKGLLTLAKGKPIGKEGYWLKIHGANCAGVD

KVPFPE

20 RIKFIEENHENIMACAKSPLENTWWAEQDSPFCFLAFCFEYAGVQHHGLSYNCSLPLAFD

GSCSGIQH

FSAMLRDEVGGRAVNLLPSETVQDIYGIVAKKVNEILQADAINGTDNEWTVDENTGEIS

EKVKLGTK

ALAGQWLAYGVTRSVTKRSVMTLALGSKEFGFRQQVLEDTIQPAIDSGKGLMFTQPNQAA

25 GYMAKLI

WESVSVTVVAVEAMNWLKSAAKLLAAEVKDKKTGEILRKRCVHWVTPDGFVPWQYKPKP

IQTRLN

LMFLGQFRLQPTINTNKDSEIDAHKQESGIAPNFVASQDGSHLRKTVVWAHEKYGIESFA

LIHDSFGTI

PADAANLFKAVRETMVDTYESCDVLADFYDQFADQLHESQLDKMPALPAKGNLNLRDILE

SDFAFA

Para gerar agrupamentos de transcritos de RNA

5 modificados em 2' (p.ex., 2'-OMe), sob condições em que uma polimerase aceita os NTPs modificados em 2', podem ser usadas as T7 RNA polimerases mutantes Y639F, Y639F/ K378R, Y639F/H784A, Y639F/H784A/K378R, Y639L/H784A, Y639L/H784A/K378R, Y639L, Y639L/K378R, P266L/Y639L/H784A ou

10 P266L/Y639L/H784A/ K378R. Uma polimerase preferida é a T7 RNA polimerase mutante Y639L/H784A. Uma outra polimerase preferida é a T7 RNA polimerase mutante Y639L/H784A/K378R. Uma outra polimerase preferida da invenção é a T7 RNA polimerase mutante P266L/Y639L/H784A ou P266L/Y639L/H784A/

15 K378R. Outras T7 RNA polimerases, particularmente aquelas que exibem uma alta tolerância por substituintes grandes em 2', podem também ser usadas nos métodos da presente invenção. Quando usada em uma polimerização molde-direcionada, usando as condições divulgadas nesse documento, a T7 RNA polimerase

20 mutante Y639L/H784A, a Y639L/H784A/K378R, P266L/Y639L/H784A ou P266L/Y639L/H784A/ K378R pode ser usada para a incorporação de NTPs totalmente com 2'-OMe, incluindo o 2'-OMe GTP, com produções de transcritos maiores do que atingidas por utilização das T7 RNA polimerases mutantes

25 Y639F, Y639F/K378R, Y639F/H784A, Y639F/H784A/K378R, Y639L, ou as Y639L/K378R. As T7 RNA polimerases mutantes Y639L/H784A, Y639L/H784A/K378R, P266L/Y639L/H784A ou P266L/Y639L/H784A/ K378R podem ser usadas com, porém não

requerem, o 2'-OH GTP para atingir altas produções de oligonucleotídeos modificados em 2', p.ex., contendo 2'-OMe.

Em uma modalidade preferida, as T7 RNA polimerases mutantes Y639L/H784A ou as Y639L/H784A/K378R da invenção são usadas com uma mistura de transcrição de MNA para promover maiores produções de transcritos inteiramente com 2'-OMe. Em algumas modalidades, as T7 RNA polimerases mutantes Y639L/H784A ou as Y639L/H784A/K378R podem ser usadas com uma mistura de transcrição de rRmY, dRmY, rGmH, fGmH, dGmH, dAmB, rRdY, dRdY ou rN.

Conforme usada aqui, uma mistura de transcrição contendo somente trifosfatos de 2'-OMe A, G, C, e U é referida como uma mistura de MNA, e os aptâmeros selecionados a partir dela são referidos como aptâmeros de MNA e contém somente 2'-O-metil nucleotídeos. Uma mistura de transcrição contendo 2'-OMe C e U e 2'-OH A e G é referida como uma mistura de "rRmY" e os aptâmeros selecionados a partir dela são referidos como aptâmeros de "rRmY". Uma mistura de transcrição contendo desóxi A e G e 2'-OMe U e C é referida como uma mistura de "dRmY" e os aptâmeros selecionados a partir dela são referidos como aptâmeros de "dRmY". Uma mistura de transcrição contendo 2'-OMe A, C, e U, e 2'-OH G é referida como uma mistura de "rGmH" e os aptâmeros selecionados a partir dela são referidos como aptâmeros de "rGmH". Uma mistura de transcrição contendo alternativamente 2'-OMe A, C, U e G e 2'-OMe A, U e C e 2'-F G é referida como uma "mistura alternativa" e os aptâmeros selecionados a partir dela são

referidos como aptâmeros da "mistura alternativa". Uma
mistura de transcrição contendo 2'-OMe A, U, e C, e 2'-F G é
referida como uma mistura de "fGmH" e os aptâmeros
selecionados a partir dela são referidos como aptâmeros de
5 "fGmH". Uma mistura de transcrição contendo 2'-OMe A, U, e
C, e desóxi G é referida como uma mistura de "dGmH" e os
aptâmeros selecionados a partir dela são referidos como
aptâmeros de "dGmH". Uma mistura de transcrição contendo
desóxi A, e 2'-OMe C, G e U é referida como uma mistura de
10 "dAmB" e os aptâmeros selecionados a partir dela são
referidos como aptâmeros de "dAmB". Uma mistura de
transcrição contendo 2'-OH A e 2'-OMe C, G e U é referida
como uma mistura de "rAmB" e os aptâmeros selecionados a
partir dela são referidos como aptâmeros de "rAmB". Uma
15 mistura de transcrição contendo trifosfato de 2'-OH
adenosina e trifosfato de guanosina e trifosfato de desóxi
citidina e trifosfato de timidina é referida como uma
mistura de rRdY e os aptâmeros selecionados a partir dela
são referidos como aptâmeros de "rRdY". Uma mistura de
20 transcrição contendo nucleotídeos totalmente com 2'-OH é
referida como uma mistura de "rN" e os aptâmeros
selecionados a partir dela são referidos como aptâmeros de
"rN", "rRrY" ou RNA, e uma mistura de transcrição contendo
nucleotídeos totalmente com desóxi é referida como uma
25 mistura de "dN" e os aptâmeros selecionados a partir dela
são referidos como aptâmeros de "dN" ou "dRdY" ou DNA.

Os oligonucleotídeos modificados em 2' podem ser
sintetizados inteiramente de nucleotídeos modificados, ou

com um subgrupo de nucleotídeos modificados. Todos os nucleotídeos podem ser modificados, e todos podem conter a mesma modificação. Todos os nucleotídeos podem ser modificados, porém conter diferentes modificações, p.ex., todos os nucleotídeos contendo a mesma base podem ter um tipo de modificação, enquanto os nucleotídeos contendo outras bases podem ter diferentes tipos de modificação. Todos os nucleotídeos de purina podem ter um tipo de modificação (ou são não modificados), enquanto todos os nucleotídeos de pirimidina têm um outro tipo diferente de modificação (ou são não modificados). Nesse modo, os transcritos, ou os agrupamentos de transcritos, são gerados usando qualquer combinação de modificações, incluindo, por exemplo, os ribonucleotídeos (2'-OH), os desoxirribonucleotídeos (2'-desóxi), os 2'-F, e 2'-OMe nucleotídeos. Os oligonucleotídeos modificados adicionalmente podem conter nucleotídeos contendo mais do que uma modificação simultaneamente, tal como uma modificação na ligação entre nucleotídeos (p.ex., fosforotioato) e no açúcar (p.ex., 2'-OMe) e na base (p.ex., inosina).

Condições de Transcrição

Diversos fatores têm sido determinados serem importantes para as condições de transcrição do método SELEX[®] modificado em 2', que podem também se aplicar aos métodos SELEX de Região Terminal descritos abaixo. Por exemplo, podem ser observados aumentos nas produções de transcrito modificado sob algumas condições quando uma

combinação particular de sequência líder/polimerase mutante for usada. Uma sequência líder é uma sequência que pode ser incorporada na extremidade de 3' de uma sequência fixa na extremidade de 5' do molde de transcrição de DNA. A
 5 sequência líder é tipicamente 6-15 nucleotídeos de comprimento, e pode ser composta de uma composição de nucleotídeos predeterminada, por exemplo, ela pode ser totalmente de purinas, ou uma mistura particular de nucleotídeos de purina e pirimidina.

10 Os exemplos de moldes que podem ser usados com as polimerases mutantes e as condições de transcrição da invenção, particularmente em combinação com a Y639L/H784A, a Y639L/H784A/K378R, a P266L/Y639L/H784A ou a P266L/Y639L/H784A/ K378R, são a ARC2118 (SEQ ID NO 3), a
 15 ARC2119 (SEQ ID NO 4), e a ARC3428

GGGAGACAAGAATAAAGCGAGTTNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN
 NNNAAGAGTCGATGATGCTTAGCTAG (SEQ ID NO 137).

Além disso, a presença de 2'-OH GTP tem sido historicamente um fator importante na obtenção de
 20 transcritos incorporando nucleotídeos modificados. A transcrição pode ser dividida em duas fases: a primeira fase é a iniciação, durante a qual um NTP é adicionado à extremidade de 3' do GTP (ou uma outra guanosina substituída) para produzir um dinucleotídeo, o qual é então
 25 estendido por aproximadamente 10-12 nucleotídeos; a segunda fase é o alongamento, durante a qual a transcrição avança para além da adição dos primeiros aproximadamente 10-12 nucleotídeos. Foi anteriormente verificado que pequenas

quantidades de 2'-OH GTP adicionadas a uma mistura de transcrição contendo a T7 RNA polimerase mutante Y639F/K378R ou a mutante Y639F/H784A/K378R e um excesso de 2'-OMe GTP eram suficientes para capacitar a polimerase de iniciar a transcrição usando o 2'-OH GTP (e deu uma produção maior de transcrito contendo 2'-OMe do que sem o 2'-OH GTP), porém, assim que a transcrição entra na fase de alongamento, a diferenciação reduzida entre 2'-OMe e 2'-OH GTP, e o excesso de 2'-OMe GTP sobre 2'-OH GTP, permite a incorporação de principalmente o 2'-OMe GTP.

A presente invenção proporciona T7 RNA polimerases mutantes, p.ex., Y639L/H784A, Y639L/H784A/K378R, P266L/Y639L/H784A ou P266L/Y639L/H784A/K378R, as quais não requerem o 2'-OH GTP na mistura de transcrição para uma alta produção de transcrição com 2'-OMe. Em uma modalidade, a alta produção significa em média pelo menos um transcrito por molde de transcrição de entrada.

Um outro fator na incorporação de nucleotídeos substituídos com 2'-OMe nos transcritos é o uso de magnésio e manganês (Mn^{2+}) ambos divalentes na mistura de transcrição. As combinações diferentes de concentrações de cloreto de magnésio e cloreto de manganês foram verificadas afetarem as produções de transcritos 2'-O-metilados, a concentração ótima do cloreto de magnésio e manganês sendo dependente da concentração na mistura de reação de transcrição dos NTPs que formam complexos de íons de metais divalentes. Para obter as maiores produções de transcritos totalmente 2'-O-metilados (i.e., nucleotídeos de A, C, U e G totalmente com

2'-OMe), as concentrações de cloreto de magnésio de aproximadamente 5 mM e cloreto de manganês a 1,5 mM são preferidas, quando cada NTP estiver presente em uma concentração de 0,5 mM. Quando a concentração de cada NTP for 1,0 mM, as concentrações de aproximadamente 6,5 mM de cloreto de magnésio e 2,0 mM de cloreto de manganês são preferidas. Quando cada NTP estiver presente em uma concentração de 1,5 mM, as concentrações de aproximadamente 8 mM de cloreto de magnésio e 2,5 mM de cloreto de manganês são preferidas. Quando a concentração de cada NTP for 2,0 mM, as concentrações de aproximadamente 9,5 mM de cloreto de magnésio e 3,0 mM de cloreto de manganês são preferidas. Em qualquer caso, os afastamentos dessas concentrações de até duas vezes ainda dão quantidades significativas de transcritos modificados.

Iniciar a transcrição com 2'-OH GMP, guanosina, ou outras 2'-OH guanosinas substituídas em uma posição diferente da posição de 2'-OH do açúcar é também importante para as misturas de transcrição que não contêm o 2'-OH GTP. Esse efeito resulta da especificidade da polimerase pelo nucleotídeo de iniciação. Como um resultado, o nucleotídeo 5' terminal de qualquer transcrito gerado nesse modo é provável de ser a 2'-OH G. Uma concentração preferida de GMP (ou guanosina) é 0,5 mM e ainda mais preferivelmente 1 mM. Foi também verificado que a inclusão do PEG, preferivelmente o PEG-8000, na reação de transcrição é útil para maximizar a incorporação de nucleotídeos modificados.

Para a incorporação máxima de 2'-OMe ATP (100%),

2'-OMe UTP (100%), 2'-OMe CTP (100%) e 2'-OMe GTP (100%) ("MNA") nos transcritos, as seguintes condições podem ser usadas: tampão HEPES a 200 mM, DTT a 40 mM, espermidina a 2 mM, PEG-8000 a 10% (p/v), Triton X-100 a 0,01% (p/v), MgCl₂ a 8 mM, MnCl₂ a 2,5 mM, 2'-OMe NTP (cada) a 1,5 mM, 2'-OH GMP a 1 mM, pH 7,5, T7 RNA Polimerase mutante Y639L/H784A/K378R a 200 nM, 5 unidades/ml de pirofosfatase inorgânica, e um molde de DNA. Em algumas modalidades, o molde de DNA pode estar presente em uma concentração de preferivelmente cerca de 5 a 500 nM. Opcionalmente, o molde de DNA usado com as condições de transcrição acima descritas compreende uma sequência líder totalmente de purinas, que aumenta a produção de transcrição em relação a um molde que não compreende tal sequência líder, quando ambos os moldes forem transcritos sob condições idênticas. Em uma outra modalidade, a sequência líder é uma mistura de purinas e pirimidinas, que aumenta a produção de transcrição em relação a um molde que não compreende tal sequência líder, quando ambos forem transcritos sob condições idênticas. Conforme usada aqui, uma unidade de pirofosfatase inorgânica é definida como a quantidade de enzima que liberará 1,0 mol de ortofosfato inorgânico por minuto, em pH 7,2 e 25 °C. A reação pode ser realizada a partir de 1 até 24 horas.

Em cada caso, os produtos de transcrição podem então ser usados para a entrada no processo SELEX®, para identificar os aptâmeros e/ou para determinar uma sequência conservada que tenha especificidade de ligação para um dado alvo. As sequências resultantes já estão estabilizadas,

eliminando essa etapa do pós-processo de modificação SELEX® e dando um aptâmero mais altamente estabilizado como um resultado.

Conforme descrito abaixo, as produções úteis de transcritos incorporando inteiramente nucleotídeos substituídos em 2' podem ser obtidas sob condições diferentes das condições descritas acima. Por exemplo, as variações nas condições de transcrição acima descritas incluem:

A concentração do tampão HEPES pode variar de 0 a 1 M. A presente invenção também contempla o uso de outros agentes de tamponamento tendo um pKa entre 5 e 10, incluindo, por exemplo, o Tris-hidroximetil-aminometano.

A concentração de DTT pode variar de 0 a 400 mM. Os métodos da presente invenção também proporcionam o uso de outros agentes redutores, incluindo, por exemplo, o mercaptoetanol.

As concentrações de espermidina e/ou espermina podem variar de 0 a 20 mM.

A concentração de PEG-8000 pode variar de 0 a 50 % (p/v). Os métodos da presente invenção também proporcionam o uso de outro polímero hidrofílico, incluindo, por exemplo, um PEG de outro peso molecular ou outros polialquileno glicóis.

A concentração de Triton X-100 pode variar de 0 a 0,1% (p/v). Os métodos da presente invenção também proporcionam o uso de outros detergentes não-iônicos, incluindo, por exemplo, outros detergentes, incluindo outros

detergentes de Triton-X.

A concentração de MgCl_2 pode variar de 0,5 mM a 50 mM. A concentração de MnCl_2 pode variar de 0,15 mM a 15 mM. Tanto o MgCl_2 quanto o MnCl_2 devem estar presentes dentro das faixas descritas e, em uma modalidade preferida, estão presentes em torno de uma razão de 10 para cerca de 3 de $\text{MgCl}_2:\text{MnCl}_2$, preferivelmente, a razão é cerca de 3-5:1, mais preferivelmente, a razão é cerca de 3-4:1.

A concentração de 2'-OMe NTP (cada NTP) pode variar de 5 μM a 5 mM.

A concentração de 2'-OH GTP pode variar de 0 μM a 300 μM . Em uma modalidade preferida, a transcrição ocorre na ausência de 2'-OH GTP (0 μM).

A concentração de 2-OH GMP, guanosina ou outra 2'-OH G substituída em uma posição diferente da posição 2' do açúcar pode variar de 0 a 5 mM. Onde o 2'-OH GTP não estiver incluído na reação, o 2'-OH GMP é requerido e pode variar de 5 μM a 5 mM.

A concentração de molde de DNA pode variar de 5 nM a 5 μM .

A concentração de polimerase mutante pode variar de 2 nM a 20 μM .

A pirofosfatase inorgânica pode variar de 0 a 100 unidades/ml.

O pH pode variar de pH 6 a pH 9. Os métodos da presente invenção podem ser praticados dentro da faixa de pH de atividade da maioria das polimerases que incorporam nucleotídeos modificados.

A reação de transcrição pode ser deixada ocorrer a partir de cerca de uma hora até semanas, preferivelmente de cerca de 1 a cerca de 24 horas.

Além disso, os métodos da presente invenção
5 proporcionam o uso ótimo de agentes quelantes na condição de reação de transcrição, incluindo, por exemplo, EDTA, EGTA, e DTT.

MÉTODO SELEX® DE REGIÃO TERMINAL

Um método para a descoberta de seqüências de
10 moldes de transcrição de ácidos nucleicos que, em algumas modalidades, são usadas para programar uma polimerização de trifosfato de nucleotídeo molde-direcionada, aumentará a produção de transcritos, é uma variante do método SELEX®, conhecido como o método SELEX® de Região Terminal (método
15 TR-SELEX®). A presente invenção proporciona um método para identificar seqüências de componentes de moldes de transcrição de ácidos nucleicos, p.ex., seqüências líderes, cujo uso aumenta a produção de transcritos, particularmente a produção de transcritos contendo nucleotídeos modificados
20 em 2' (p.ex., os 2'-OMe nucleotídeos), quando usadas para programar uma polimerização molde-direcionada, usando o método TR-SELEX®.

Para selecionar as seqüências líderes que promovem uma produção aumentada de transcritos contendo nucleotídeos
25 modificados em 2', uma biblioteca candidata de moldes de transcrição de oligonucleotídeos é gerada, a qual contém uma seqüência promotora que permite a transcrição em um modo dependente do molde, uma primeira região fixa compreendendo

mais do que um nucleotídeo fixo imediatamente 3' ao promotor, para permitir que ocorram ligações com esplinte, desse modo permitindo a amplificação pela extensão de primers ligados aos sítios de ligação a primers sobre o molde ligado; uma
5 região degenerada a partir da qual a sequência líder será selecionada; e uma sequência fixa na extremidade de 3' para permitir a amplificação. Em uma modalidade preferida, a região de degeneração do molde da biblioteca está próxima à extremidade de 5', desse modo reduzindo o comprimento da
10 sequência fixa em 5'.

Essa biblioteca de moldes de transcrição é opcionalmente amplificada por PCR, e então usada para programar a transcrição usando uma mistura de reação de transcrição compreendendo uma polimerase (incluindo, sem
15 limitação, uma T7 RNA polimerase mutada), trifosfatos de nucleotídeos (NTPs) (incluindo, sem limitação, um ou mais NTPs modificados em 2'), e íons magnésio, sob as condições divulgadas nesse documento. A mistura de transcritos resultante é transcrita invertida para obter uma mistura
20 candidata de sequências de cDNA, que são então ligadas a uma sequência de DNA codificando o promotor de T7. Opcionalmente, a mistura de transcritos resultante primeiramente sofre ligação, e é então transcrita invertida. O cDNA que codifica os transcritos é então amplificado por
25 PCR, e os clones são testados quanto à produção de transcrição usando análise por gel. Os moldes de transcrição amplificados nesse modo podem opcionalmente ser usados para efetuar rodadas adicionais do processo TR-SELEX[®],

se necessário, para obter uma maior produção de transcritos (Ver a Figura 2).

As seqüências de clones dos transcritos amplificados podem ser analisadas para identificar o elemento da seqüência líder em 5' que permite a transcrição (incluindo, sem limitação, a transcrição incorporando um ou mais nucleotídeos modificados em 2'). Esses elementos da seqüência líder em 5' são úteis para projetar bibliotecas candidatas de moldes de transcrição de oligonucleotídeos que podem ser usados no SELEX® para promover uma produção aumentada de transcritos de ácidos nucléicos, os quais contêm nucleotídeos modificados em 2'. Os exemplos de bibliotecas preferidas de moldes de transcrição de DNA, que incorporam elementos da seqüência líder em 5' identificados pelo método TR-SELEX® (mostrado sublinhado) e promovem maiores produções de transcritos contendo nucleotídeos modificados em 2', p.ex., 2'-OMe nucleotídeos, usando as condições divulgadas nesse documento, são descritos abaixo.

Para cada uma das seqüências das bibliotecas de
20 moldes de transcrição de DNA listadas abaixo, o elemento da
seqüência líder em 5' é mostrado sublinhado, e todas as
seqüências estão na direção 5'-3'.

ARC 2118 (SEQ ID NO 3)

TAATACGACTCACTATAGGGGAGTACAATAACGTTCTCGNNNNNNNNNNNNNNNNNNNNNNNN

25 NNNNNNNNNGGATCGTTACGACTAGCATCGATG

ARC2119 (SEQ ID NO 4)

TAATACGACTCACTATAGGGGGTGATATTGACGTTCTCGNNNNNNNNNNNNNNNNNNNNNN

NNNNNNNNNGGATCGTTACGACTAGCATCGATG

ARC2120 (SEQ ID NO 5)

TAATACGACTCACTATAGGGGGTGC GCGGTTACGTTCTCGNNNNNNNNNNNNNNNNNNNNNN
NNNNNNNNNGGATCGTTACGACTAGCATCGATG

ARC2121 (SEQ ID NO 6)

5 TAATACGACTCACTATAGGGGGGAGGGGGTGCCGTTCTCGNNNNNNNNNNNNNNNNNNNNNN
NNNNNNNNNGGATCGTTACGACTAGCATCGATG

Para gerar misturas de transcritos de transcritos de RNA modificados em 2' (p.ex., 2'-OMe), sob condições nas quais uma polimerase aceita os NTPs modificados em 2', a T7 RNA polimerase mutante Y639F, Y639F/K378R, Y639F/H784A, Y639F/H784A/K378R, Y639L/H784A, Y639L/H784A/K378R, P266L/Y639L/H784A) ou P266L/Y639L/H784A/K378R) pode ser usada com as seqüências líderes em 5' identificadas pelos métodos proporcionados pela presente invenção. Uma polimerase preferida a ser usada com as seqüências líderes em 5' da presente invenção, dando a maior produção de transcritos de ácidos nucleicos contendo nucleotídeos modificados em 2', é a RNA polimerase mutante Y639L/H784A anteriormente descrita. Uma outra polimerase preferida a ser usada com as seqüências líderes em 5' da invenção é a T7 RNA polimerase mutante Y639L/H784A/K378R. Outras T7 RNA polimerases, particularmente aquelas que exibem uma alta tolerância por substituintes grandes em 2', podem também ser usadas na presente invenção.

25 Além de incorporar seqüências líderes em bibliotecas candidatas e polimerases mutantes que promovem a produção aumentada de transcritos de ácidos nucleicos contendo nucleotídeos modificados em 2' (p.ex., as T7 RNA

polimerases mutantes Y639L/H784A e Y639L/H784A/K378R), os diversos fatores descritos acima, os quais foram determinados serem importantes para as condições de transcrição, podem ser usados para aumentar adicionalmente a produção de transcritos contendo nucleotídeos modificados em 2'.

As seqüências líderes identificadas e as T7 RNA polimerases mutantes Y639F/H784A, Y639F/H784A/K378R, Y639L, Y639L/K378R, Y639L/H874A, Y639L/H874A/K378R, P266L/Y639L/H784A ou P266L/Y639L/H784A/K378R podem ser usadas no SELEX®, com as condições descritas nesse documento, para gerar aptâmeros compreendendo qualquer combinação de nucleotídeos modificados em 2', p.ex., as modificações com 2'-OH, 2'-F, 2'-desóxi, 2'-OMe, e 2'-NH₂ dos nucleotídeos de ATP, GTP, CTP, TTP, e UTP. Os nucleotídeos modificados em 2' incorporados são preferivelmente 2'-O-metil nucleotídeos. Uma composição de aptâmero compreendendo um ou mais 2'-O-metil nucleotídeos é preferida. Uma composição de aptâmero compreendendo 100% de 2'-O-metil purinas e pirimidinas, exceto pelo nucleotídeo de partida, é mais preferida. Em uma modalidade preferida, uma das seqüências líderes identificadas e as T7 RNA polimerases mutantes Y639L/H874A, Y639L/H784A/K378R, P266L/Y639L/H784A ou P266L/Y639L/H784A/K378R são usadas no método SELEX®, com as condições descritas nesse documento, para gerar produções de transcritos maiores de aptâmeros compreendendo nucleotídeos inteiramente com 2'-OMe.

Para a incorporação máxima de 2'-OMe ATP (100%),

UTP (100%), CTP (100%) e GTP (100%) (MNA") nos transcritos, as seguintes condições são preferidas: tampão de HEPES a 200 mM, DTT a 40 mM, espermidina 2 mM, PEG-8000 a 10% (p/v), Triton X-100 a 0,01% (p/v), $MgCl_2$ a 8 mM, $MnCl_2$ a 2,5 mM, 2'-OMe NTP (cada) a 1,5 mM, 2'-OH GMP a 1 mM, pH 7,5, T7 RNA Polimerase Y639L/H784A/K378R a 200 nM, 5 unidades/ml de pirofosfatase inorgânica, e uma sequência líder que aumenta a produção de transcrição sob as condições de transcrição derivadas. Em uma modalidade, a sequência líder é uma sequência líder totalmente de purinas. Em uma outra modalidade, a sequência líder é uma mistura de purinas e pirimidinas. Como usada aqui, uma unidade de pirofosfatase inorgânica é definida como a quantidade de enzima que liberará 1,0 mol de ortofosfato inorgânico por minuto, em pH 7,2 e 25 °C.

Química Medicinal dos Aptâmeros

Assim que forem identificados os aptâmeros que se ligam a um alvo desejado, diversas técnicas podem ser opcionalmente efetuadas para aumentar adicionalmente as características de ligação e/ou funcionais das sequências de aptâmeros identificadas. Os aptâmeros, p.ex., os aptâmeros de MNA, que se ligam a um alvo desejado, identificados através do processo SELEX® (p.ex., o método SELEX® Modificado em 2'), podem ser opcionalmente truncados para obter a sequência de aptâmeros mínima (também referida aqui como "construção minimizada") tendo as características desejadas de ligação e/ou funcionais. Um método de efetuar isso é por utilização de programas de duplicação e análise

da sequência (p.ex., alinhamento de sequências de clones resultantes a partir de uma seleção para procurar motivos conservados e/ou co-variação), para informar o projeto das construções minimizadas. Os experimentos bioquímicos de sondagem podem também ser efetuados para determinar os limites em 5' e 3' de uma sequência de aptâmeros, para informar o projeto das construções minimizadas. As construções minimizadas podem então ser quimicamente sintetizadas e testadas quanto às características de ligação e funcionais, em comparação com a sequência não minimizada a partir da qual elas foram derivadas. As variantes de uma sequência de aptâmeros contendo uma série de deleções em 5', 3' e/ou internas podem também ser direta e quimicamente sintetizadas e testadas quanto às características de ligação e/ou funcionais, em comparação com a sequência de aptâmeros não minimizada a partir da qual elas foram derivadas.

Adicionalmente, as re-seleções dopadas ("doped reselections") podem ser usadas para explorar as exigências das sequências dentro de uma única sequência de aptâmeros ativa, tal como um aptâmero de MNA (i.e., um aptâmero que se liga a um alvo desejado identificado através do processo SELEX® (incluindo o processo SELEX® Modificado em 2'), ou uma única sequência de aptâmeros minimizada. As re-seleções dopadas são realizadas usando um agrupamento degenerado, sintético, que tenha sido projetado com base na sequência única de interesse. O nível de degradação normalmente varia de 70% a 85% da partir do nucleotídeo do tipo selvagem, i.e., a sequência única de interesse. Em geral, as sequências com

mutações neutras são identificadas através do processo de re-seleção dopada, porém, em alguns casos, as alterações na sequência podem resultar em aperfeiçoamentos na afinidade. A informação de sequência compósita a partir de clones
5 identificados usando as re-seleções dopadas pode então ser usada para identificar o motivo de ligação mínimo e auxiliar nos esforços da Química Medicinal.

As sequências de aptâmeros identificadas usando o processo SELEX®, tal como os aptâmeros de MNA (incluindo o
10 processo SELEX Modificado em 2' e as re-seleções dopadas), e/ou as sequências de aptâmeros minimizadas podem também ser opcionalmente modificadas após seleção SELEX® usando a Química Medicinal dos Aptâmeros, para efetuar a mutagênese aleatória ou direcionada da sequência, para aumentar a
15 afinidade de ligação e/ou as características funcionais, ou alternativamente para determinar quais posições na sequência são essenciais para a atividade de ligação e/ou as características funcionais.

A Química Medicinal dos Aptâmeros é uma técnica de
20 aperfeiçoamento dos aptâmeros, na qual grupos de aptâmeros variantes são quimicamente sintetizados. Esses grupos de variantes tipicamente diferem do aptâmero de origem pela introdução de um único substituinte, e diferem um do outro pela posição desse substituinte. Essas variantes são então
25 comparadas uma com a outra e com a origem. Os aperfeiçoamentos nas características podem ser profundos o suficiente que a inclusão de um único substituinte pode ser tudo que é necessário para obter um critério terapêutico

particular.

Alternativamente, a informação colhida a partir do grupo de variantes individuais pode ser usada para projetar grupos adicionais de variantes, em que mais do que um substituinte é introduzido simultaneamente. Em uma estratégia de projeto, todas as variantes de um único substituinte são classificadas, as 4 primeiras são escolhidas e todas as possíveis combinações duplas (6), triplas (4) e quádruplas (1) dessas 4 variantes de um único substituinte são sintetizadas e testadas. Em uma segunda estratégia de projeto, a melhor variante de um único substituinte é considerada ser a nova origem e todas as variantes com duplos substituintes possíveis que incluem essa variante de um único substituinte de maior classificação são sintetizadas e testadas. Podem ser usadas outras estratégias, e essas estratégias podem ser aplicadas repetidamente de modo tal que o número de substituintes seja gradualmente aumentado, ao mesmo tempo continuando a identificar variantes adicionais aperfeiçoadas.

A Química Medicinal dos Aptâmeros pode ser usada particularmente como um método para explorar a introdução local, em vez de global, de substituintes. Porque os aptâmeros são descobertos dentro de bibliotecas que são geradas por transcrição, quaisquer substituintes que sejam introduzidos durante o processo SELEX® devem ser introduzidos globalmente. Por exemplo, se for desejado introduzir ligações de fosforotioato entre os nucleotídeos, então elas podem somente ser introduzidas em cada A (ou cada

G, C, T, U etc.) (globalmente substituídos). Os aptâmeros que requeiram fosforotioatos em algumas As (ou algumas G, C, T, U etc.) (localmente substituídos), porém não podem tolerá-los em outras A, não podem ser prontamente
5 descobertos por esse processo.

Os tipos de substituintes que podem ser utilizados pelo processo de Química Medicinal dos Aptâmeros estão limitados somente pela capacidade de gerá-los como reagentes de síntese de fase sólida e introduzi-los em um esquema de
10 síntese de oligômeros. O processo certamente não está limitado aos nucleotídeos sozinhos. Os esquemas da Química Medicinal dos Aptâmeros podem incluir substituintes que introduzam volume estérico, hidrofobicidade, hidrofili-
15 cidade, lipofobicidade, carga positiva, carga negativa, carga neutra, zwitteríons, polarizabilidade, resistência a nucleases, rigidez conformacional, flexibilidade conformacional, características de ligação a proteínas, massa etc. Os esquemas da Química Medicinal dos Aptâmeros podem incluir modificações das bases, modificações
20 dos açúcares ou modificações da ligação de fosfodiéster.

Quando considerando os tipos de substituintes que são prováveis de serem benéficos dentro do contexto de um aptâmero terapêutico, pode ser desejável introduzir substituições que incidam em uma ou mais das seguintes
25 categorias:

(1) Substituintes já presentes no corpo, p.ex., 2'-desóxi, 2'-ribo, 2'-O-metil purinas ou pirimidinas ou 5-metil citosina.

(2) Substituintes já parte de uma substância terapêutica aprovada, p.ex., os oligonucleotídeos ligados ao fosforotioato.

(3) Substituintes que hidrolisam ou degradam-se até uma das duas categorias acima descritas, p.ex., os oligonucleotídeos ligados ao fosfonato de metila.

(4) Os aptâmeros da presente invenção incluem os aptâmeros desenvolvidos através da química medicinal do aptâmero, como descrito nesse documento.

10 A afinidade de ligação ao alvo dos aptâmeros da presente invenção pode ser avaliada através de uma série de reações de ligação entre o aptâmero e o alvo (p.ex., uma proteína), em que o aptâmero marcado com traço de ^{32}P é incubado com uma série de diluições do alvo em um meio
15 tamponado, então analisado por filtração em nitrocelulose usando um tubo coletor de filtração a vácuo. Referido nesse documento como o ensaio de ligação por "dot blot", esse método utiliza um meio de filtração de três camadas, consistindo (do topo até o fundo) em nitrocelulose, filtro
20 de náilon, e papel blot em gel. O RNA que é ligado ao alvo é capturado sobre o filtro de nitrocelulose, enquanto que o RNA não ligado ao alvo é capturado sobre o filtro de náilon. O papel blot em gel é incluído como um meio de suporte para os outros filtros. Após a filtração, as camadas do filtro
25 são separadas, secadas e expostas sobre uma tela fosforescente e quantificadas usando um sistema de fosforimageamento a partir dela. Os resultados quantificados podem ser usados para gerar curvas de ligação

dos aptâmeros a partir das quais as constantes de dissociação (K_D) pode ser calculadas. Em uma modalidade preferida, o meio tamponado usado para efetuar as reações de ligação é 1X PBS da Dulbecco (com Ca^{++} e Mg^{++}) mais 0,1 mg/mL de BSA.

Geralmente, a capacidade de um aptâmero de modular a atividade funcional de um alvo, i.e., a atividade funcional do aptâmero, pode ser avaliada usando modelos *in vitro* e *in vivo*, que variarão dependendo da função biológica do alvo. Em algumas modalidades, os aptâmeros da presente invenção podem inibir uma função biológica conhecida do alvo, enquanto que em outras modalidades, os aptâmeros da invenção podem estimular uma função biológica conhecida do alvo. A atividade funcional dos aptâmeros da presente invenção pode ser avaliada usando modelos *in vitro* e *in vivo*, projetados para medir uma função conhecida do alvo do aptâmero.

Os aptâmeros da presente invenção podem ser rotineiramente adaptados para propósitos diagnósticos de acordo com qualquer número de práticas empregadas por aqueles versados na técnica. A utilização diagnóstica pode incluir as aplicações diagnósticas tanto *in vivo* quanto *in vitro*. Os agentes diagnósticos necessitam somente de serem capazes de permitir que o usuário identifique a presença de um dado alvo em um local ou uma concentração particular. Simplesmente a capacidade de formar pares de ligação com o alvo pode ser suficiente para desencadear um sinal positivo para propósitos diagnósticos. Aqueles versados na técnica também seriam capazes de adaptar qualquer aptâmero por

procedimentos conhecidos na técnica, para incorporar um marcador de marcação para rastrear a presença de tal ligante. Tal marcador poderia ser usado em diversos procedimentos diagnósticos.

5 APTÂMEROS TENDO MOTIVOS IMUNOESTIMULADORES

O reconhecimento do DNA bacteriano pelo sistema imunológico do vertebrado é baseado no reconhecimento de dinucleotídeos de CG não metilados em contextos particulares da sequência ("motivos de CpG"). Um receptor que reconhece
10 tal motivo é o receptor similar ao Toll 9 ("TLR 9"), um membro de uma família de receptores similares ao Toll (~10 membros) que participam na resposta imune inata por reconhecimento de componentes microbianos distintos. O TLR 9 liga sequências de CpG de oligodesoxinucleotídeos ("ODN")
15 não metiladas em um modo específico para a sequência. O reconhecimento de motivos de CpG desencadeia mecanismos de defesa, resultando em respostas imunes inatas e, no final, adquiridas. Por exemplo, a ativação do TLR 9 em camundongos induz a ativação de células apresentadoras de antígenos, a
20 supra-regulação das moléculas do MHC das classes I e II e a expressão de moléculas e citocinas co-estimuladoras importantes, incluindo a IL-12 e a IL-23. Essa ativação aumenta tanto direta quanto indiretamente as respostas das células B e T, incluindo a supra-regulação forte do IFN-gama
25 de citocina de TH1. Coletivamente, a resposta às sequências de CpG resulta em: proteção contra doenças infecciosas, resposta imune melhorada para as vacinas, uma resposta efetiva contra asma, e citotoxicidade melhorada mediada por

células dependentes de anticorpos. Assim, os ODNs de CpG podem proporcionar proteção contra doenças infecciosas, função como imuno-adjuvantes ou substâncias terapêuticas para cânceres (monoterapia ou em combinação com um mAb ou
5 outras terapias), e podem diminuir a asma e a resposta alérgica.

Os aptâmeros da presente invenção, p.ex., os aptâmeros de MNA, podem compreender um ou mais CpG ou outra sequência imunoestimuladora. Tais aptâmeros podem ser
10 identificados ou gerados por uma variedade de estratégias usando, p.ex., o processo SELEX[®] descrito nesse documento. Em geral, as estratégias podem ser divididas em dois grupos. No grupo um, as estratégias são dirigidas para identificar ou gerar aptâmeros compreendendo tanto um motivo de CpG
15 quanto outra sequência imunoestimuladora, bem como um sítio de ligação para um alvo, onde o alvo (nas partes que seguem, "alvo que não é de CpG") é um alvo diferente de um sabido reconhecer os motivos de CpG ou outras sequências imunoestimuladoras e sabido estimular uma resposta
20 imunológica na ligação a um motivo de CpG. A primeira estratégia desse grupo compreende efetuar o SELEX[®] para obter um aptâmero para um alvo que não seja de CpG, específico, usando um agrupamento de oligonucleotídeos, onde um motivo de CpG tenha sido incorporado em cada membro do
25 agrupamento como, ou como parte de, uma região fixa, p.ex., em algumas modalidades, a região ao acaso dos membros do agrupamento compreende uma região fixa tendo um motivo de CpG incorporado nela, e identificar um aptâmero

compreendendo um motivo de CpG. A segunda estratégia desse grupo compreende efetuar o SELEX[®] para obter um aptâmero para um alvo que não seja de CpG, específico, preferivelmente um alvo, e após a seleção, fixar um motivo de CpG à extremidade de 5' e/ou 3' ou engenheirar um motivo de CpG em uma região, preferivelmente uma região não essencial, do aptâmero. A terceira estratégia desse grupo compreende efetuar o SELEX[®] para obter um aptâmero para um alvo que não seja de CpG, específico, onde durante a síntese do agrupamento, a razão molar dos diversos nucleotídeos é influenciada em uma ou mais etapas de adição de nucleotídeos, de modo que a região ao acaso de cada membro do agrupamento esteja enriquecida em motivos de CpG, e identificar um aptâmero compreendendo um motivo de CpG. A quarta estratégia desse grupo compreende efetuar o SELEX[®] para obter um aptâmero para um alvo que não seja de CpG, específico, e identificar um aptâmero compreendendo um motivo de CpG. A quinta estratégia desse grupo compreende efetuar o SELEX[®] para obter um aptâmero para um alvo que não é de CpG, específico, e identificar um aptâmero que, na ligação, estimule uma resposta imune, porém que não compreenda um motivo de CpG.

No grupo dois, as estratégias são direcionadas para identificar ou gerar aptâmeros compreendendo um motivo de CpG e/ou outras seqüências que são ligadas pelos receptores para os motivos de CpG (p.ex., o TLR9 ou outros receptores semelhantes ao toll) e, com a ligação, estimulam uma resposta imune. A primeira estratégia desse grupo

compreende efetuar o SELEX[®] para obter um aptâmero para um alvo, sabido ligar-se aos motivos de CpG ou outras seqüências imunoestimuladoras e, com a ligação, estimulam uma resposta imune usando um agrupamento de oligonucleotídeos, onde um motivo de CpG tenha sido incorporado em cada membro do agrupamento como, ou como parte de, uma região fixa, p.ex., em algumas modalidades, a região ao acaso dos membros do agrupamento compreende uma região fixa tendo um motivo de CpG incorporado nela, e identificar um aptâmero compreendendo um motivo de CpG. A segunda estratégia desse grupo compreende efetuar o SELEX[®] para obter um aptâmero para um alvo sabido ligar-se aos motivos de CpG ou outras seqüências imunoestimuladoras e, com a ligação, estimular uma resposta imune, e então fixar um motivo de CpG à extremidade de 5' e/ou 3' ou engenheirar um motivo de CpG em uma região, preferivelmente uma região não essencial, do aptâmero. A terceira estratégia desse grupo compreende efetuar o SELEX[®] para obter um aptâmero para um alvo, sabido ligar-se aos motivos de CpG ou outras seqüências imunoestimuladoras e, com a ligação, estimular uma resposta imune, onde durante a síntese do agrupamento, a razão molar dos diversos nucleotídeos é influenciada em uma ou mais etapas de adição de nucleotídeos, de modo que a região ao acaso de cada membro do agrupamento esteja enriquecida em motivos de CpG, e identificar um aptâmero compreendendo um motivo de CpG. A quarta estratégia desse grupo compreende efetuar o SELEX[®] para obter um aptâmero para um alvo, sabido ligar-se aos motivos de CpG ou outras

seqüências imunoestimuladoras e, com a ligação, estimular uma resposta imune, e identificar um aptâmero compreendendo um motivo de CpG. A quinta estratégia desse grupo compreende efetuar o SELEX® para obter um aptâmero para um
5 alvo, sabido ligar-se aos motivos de CpG ou outras seqüências imunoestimuladoras, e identificar um aptâmero que, com a ligação, estimula uma resposta imune, porém que não compreendem um motivo de CpG.

Uma variedade de diferentes classes de motivos de
10 CpG foi identificada, cada uma resultando em reconhecimento em uma cascata diferente de eventos, liberação de citocinas e outras moléculas, e ativação de certos tipos de células. Ver, p.ex., CpG Motifs in Bacterial DNA and Their Immune Effects, Annu. Rev. Immunol. 2002, 20:709-760, incorporado
15 aqui por referência. Os motivos imunoestimuladores adicionais são divulgados nas seguintes Patentes U.S., cada uma das quais é incorporada nesse documento por referência: Patente U.S. Nº 6.207.646; Patente U.S. Nº 6.239.116; Patente U.S. Nº 6.429.199; Patente U.S. Nº 6.214.806;
20 Patente U.S. Nº 6.653.292; Patente U.S. Nº 6.426.434; Patente U.S. Nº 6.514.948 e Patente U.S. Nº 6.498.148. Quaisquer desses motivos de CpG ou outros motivos imunoestimuladores podem ser incorporados em um aptâmero. A escolha dos aptâmeros é dependente da doença ou do distúrbio
25 a ser tratado. Os motivos imunoestimuladores preferidos são como a seguir (mostrados 5' até 3', da esquerda para direita), onde "r" designa uma purina, "y" designa uma pirimidina, e "X" designa qualquer nucleotídeo: AACGTTTCGAG

(SEQ ID NO: 136); AACGTT; ACGT, rCGy; rrCGyy, XCGX, XXCGXX, e $X_1X_2CGY_1Y_2$, onde X_1 é G ou A, X_2 não é C, Y_1 não é G e Y_2 é preferivelmente T.

Naqueles casos onde um motivo de CpG for
 5 incorporado em um aptâmero que se liga a um alvo específico, diferente de um alvo sabido ligar-se aos motivos de CpG, e, com a ligação, estimula uma resposta imune (um "alvo que não seja de CpG"), o CpG é preferivelmente posicionado em uma região não essencial do aptâmero. As regiões não essenciais
 10 dos aptâmeros podem ser identificadas por mutagênese sítio-direcionada, análises de deleção e/ou análises de substituição. Entretanto, pode ser usada qualquer posição que não interfira significativamente com a capacidade do aptâmero de ligar-se ao alvo que não é de CpG. Além de ser
 15 incorporado na sequência de aptâmeros, o motivo de CpG pode ser fixado a uma ou a outra, ou a ambas as extremidades de 5' e 3' ou de outro modo unido ao aptâmero. Qualquer posição ou meio de ligação pode ser usado, desde que a capacidade do aptâmero de ligar-se ao alvo que não seja de
 20 CpG não seja significativamente interferida.

Conforme usado aqui, a "estimulação de uma resposta imune" pode significar (1) a indução de uma resposta específica (p.ex., a indução de uma resposta de Th1) ou da produção de certas moléculas ou (2) a inibição ou
 25 a supressão de uma resposta específica (p.ex., a inibição ou a supressão da resposta de Th2) ou de certas moléculas.

MODULAÇÃO DA FARMACOCINÉTICA E DA BIODISTRIBUIÇÃO DAS SUBSTÂNCIAS TERAPÊUTICAS DE APTÂMEROS

É importante que as propriedades farmacocinéticas para todas as substâncias terapêuticas à base de oligonucleotídeos, incluindo os aptâmeros, sejam reguladas para serem compatíveis com a aplicação farmacêutica desejada.

5 Embora os aptâmeros direcionados contra alvos extracelulares não sofram de dificuldades associadas com a distribuição intracelular (conforme é o caso com as substâncias terapêuticas à base de RNA sem sentido e RNAi), tais aptâmeros devem ainda ser capazes de ser distribuídos para
10 órgãos e tecidos-alvo, e permanecer no corpo (não modificados) por um período de tempo consistente com o regime de dosagem desejado.

Assim, a presente invenção proporciona materiais e métodos para afetar a farmacocinética das composições de
15 aptâmeros, e, em particular, a capacidade de regular a farmacocinética dos aptâmeros. A capacidade de regulação (i.e., a capacidade de modular) da farmacocinética dos aptâmeros é atingida através da conjugação de porções modificadoras (p.ex., polímeros de PEG) ao aptâmero e/ou da
20 incorporação de nucleotídeos modificados (p.ex., 2'-flúor ou 2'-O-metil) para alterar a composição química do ácido nucléico. A capacidade de regular a farmacocinética dos aptâmeros é usada no aperfeiçoamento de aplicações terapêuticas existentes, ou alternativamente, no
25 desenvolvimento de novas aplicações terapêuticas. Por exemplo, em algumas aplicações terapêuticas, p.ex., nas indicações antineoplásticas ou de cuidado agudo, onde a rápida depuração ou desativação da droga possa ser desejada,

é desejável diminuir os tempos de residência dos aptâmeros na circulação. Alternativamente, em outras aplicações terapêuticas, p.ex., as terapias de manutenção, onde for desejada a circulação sistêmica de uma substância terapêutica, pode ser desejável aumentar os tempos de residência dos aptâmeros em circulação.

Além disso, a capacidade de regulação da farmacocinética dos aptâmeros é usada para modificar a biodistribuição de uma substância terapêutica em um paciente.

Por exemplo, em algumas aplicações terapêuticas, pode ser desejável alterar a biodistribuição de uma substância terapêutica de aptâmero em um esforço para alvejar um tipo particular de tecido ou um órgão específico (ou conjunto de órgãos). Nessas aplicações, a substância terapêutica de aptâmero preferencialmente acumula-se em um tecido ou órgão(s) específico(s). Em outras aplicações terapêuticas, pode ser desejável alvejar tecidos que mostrem um marcador celular ou um sintoma associado com uma dada doença, lesão celular ou outra patologia anormal, de modo tal que a substância terapêutica de aptâmero preferencialmente acumule-se no tecido afetado. Por exemplo, conforme descrito no pedido provisório dos Estados Unidos Nº de Série 60/550790, depositado em 5 de março de 2004, e intitulado "Controlled Modulation of the Pharmacokinetics and Biodistribution of Aptamer Therapeutics", e no pedido não provisório dos Estados Unidos Nº de Série 10/---.---, depositado em 7 de março de 2005, e intitulado "Controlled Modulation of the Pharmacokinetics and Biodistribution of

Aptamer Therapeutics", a PEGuilação de uma substância terapêutica de aptâmero (p.ex., a PEGuilação com um polímero de PEG de 20 kDa) é usada para alvejar tecidos inflamados, de modo tal que a substância terapêutica de aptâmero
5 PEGuilada preferencialmente acumule-se no tecido inflamado.

Para determinar os perfis de farmacocinética e biodistribuição das substâncias terapêuticas de aptâmeros (p.ex., os conjugados de aptâmeros ou os aptâmeros tendo químicas alteradas, tais como os nucleotídeos modificados),
10 uma variedade de parâmetros é monitorada. Tais parâmetros incluem, por exemplo, a meia-vida ($t_{1/2}$), a depuração plasmática (Cl), o volume de distribuição (Vss), a área sob a curva de concentração-tempo (AUC), a concentração máxima observada em soro ou plasma ($C_{\text{máx}}$), e o tempo de residência
15 médio (TRM) de uma composição de aptâmero. Conforme usado nesse documento, o termo "AUC" refere-se à área sob o gráfico da concentração plasmática de uma substância terapêutica de aptâmero versus o tempo após a administração do aptâmero. O valor de AUC é usado para estimar a
20 biodisponibilidade (i.e., a porcentagem de substância terapêutica de aptâmero administrada na circulação após a administração do aptâmero) e/ou a depuração total (Cl) (i.e., a taxa na qual a substância terapêutica de aptâmero é removida da circulação) de uma dada substância terapêutica
25 de aptâmero. O volume de distribuição relaciona a concentração plasmática de uma substância terapêutica de aptâmero com a quantidade de aptâmero presente no corpo. Quanto maior o Vss, mais de um aptâmero é encontrado fora do

plasma (i.e., maior o extravasamento).

A presente invenção proporciona materiais e métodos para modular, em um modo controlado, a farmacocinética e a biodistribuição de composições de aptâmeros estabilizadas, p.ex., os aptâmeros de MNA, *in vivo* por conjugação de um aptâmero, p.ex., um aptâmero de MNA, a uma porção moduladora, tal como uma molécula pequena, peptídeo, ou grupo terminal de polímero, ou por incorporação de nucleotídeos modificados em um aptâmero. Conforme descrito nesse documento, a conjugação de uma porção modificadora e/ou a modificação da composição química de nucleotídeo(s) altera aspectos fundamentais de tempo de residência do aptâmero em circulação e distribuição para os tecidos.

Além da depuração pelas nucleases, as substâncias terapêuticas de oligonucleotídeos estão sujeitas à eliminação via filtração renal. Como tal, um oligonucleotídeo resistente a nucleases administrado de modo intravenoso tipicamente exibe uma meia-vida *in vivo* de <10 min, a não ser que a filtração possa ser bloqueada. Isso pode ser efetuado facilitando a rápida distribuição da corrente sanguínea para os tecidos ou aumentando o peso molecular aparente do oligonucleotídeo acima do corte de tamanho efetivo para o glomérulo. A conjugação de substâncias terapêuticas pequenas com um polímero de PEG (PEGuilação), descrita abaixo, pode dramaticamente prolongar os tempos de residência dos aptâmeros em circulação, desse modo diminuindo a frequência de dosagem e aumentando a

eficácia contra os alvos vasculares.

Os aptâmeros podem ser conjugados a uma variedade de porções modificadoras, tais como os polímeros de altos pesos moleculares, p.ex., o PEG; os peptídeos, p.ex., Tat (um fragmento de 13 aminoácidos da proteína Tat de HIV (Vives, e col. (1997), J. Biol. Chem. 272(25): 16010-7)), Ant (uma sequência de 16 aminoácidos derivada da terceira hélice da proteína homeótica de *Drosophila antennapedia* (Pietersz, e col. (2001), Vaccine 19(11-12): 1397-405)) e Arg₇ (peptídeos de permeação de células positivamente carregados, curtos, compostos de poliarginina (Arg₇) (Rothbard, e col. (2000), Nat. Med. 6(11): 1253-7; Rothbard, J e col. (2002), J. Med. Chem. 45(17): 3612-8)); e as moléculas pequenas, p.ex., os compostos lipofílicos, tais como o colesterol. Entre os diversos conjugados descritos nesse documento, as propriedades *in vitro* dos aptâmeros são alteradas mais profundamente por formação de complexo com grupos de PEG. Por exemplo, a formação de complexo de uma substância terapêutica de aptâmero modificada por 2'F e 2'-OMe mista com um polímero de PEG de 20 kDa impede a filtração renal e promove a distribuição de aptâmeros para tecidos tanto saudáveis quanto inflamados. Além disso, o conjugado de polímero de PEG de 20 kDa-aptâmero demonstra ser quase tão efetivo quanto um polímero de PEG de 40 kDa na prevenção da filtração renal de aptâmeros. Embora um efeito da PEGuilação seja sobre a depuração de aptâmeros, a exposição sistêmica prolongada proporcionada pela presença da porção de 20 kDa também facilita a distribuição do

aptâmero para os tecidos, particularmente aqueles de órgãos altamente perfundidos e aqueles no local da inflamação. O conjugado de aptâmero-polímero de PEG de 20 kDa orienta a distribuição do aptâmero para o local da inflamação, de modo tal que o aptâmero PEGuilado preferencialmente acumula-se no tecido inflamado. Em algumas situações, o conjugado de aptâmero PEGuilado de 20 kDa é capaz de acessar o interior das células, tais como, por exemplo, as células dos rins.

Os nucleotídeos modificados podem também ser usados para modular a depuração plasmática de aptâmeros. Por exemplo, um aptâmero não conjugado, o qual incorpora químicas de estabilização tanto de 2'-F quanto de 2'-OMe, que é típico dos aptâmeros de geração atuais, visto que ele exibe um alto grau de estabilidade às nucleases *in vitro* e *in vivo*, mostra perda rápida do plasma (i.e., rápida depuração plasmática) e uma rápida distribuição para os tecidos, principalmente para o rim, quando comparado ao aptâmero não modificado.

ÁCIDOS NUCLÉICOS DERIVADOS COM PEG

Conforme descrito acima, a derivação de ácidos nucléicos com polímeros não imunogênicos de altos pesos moleculares tem o potencial de alterar as propriedades farmacocinéticas e farmacodinâmicas dos ácidos nucléicos, tornando-os agentes terapêuticos mais efetivos. As alterações favoráveis na atividade podem incluir a resistência aumentada à degradação por nucleases, a filtração diminuída através dos rins, a exposição diminuída ao sistema imunológico, e a distribuição alterada da

substância terapêutica através do corpo.

As composições de aptâmeros da invenção podem ser derivadas com porções de polialquileno glicol ("PAG"). Os exemplos de ácidos nucleicos derivados com PAG são encontrados no Pedido de Patente dos Estados Unidos Nº de Série 10/718.833, depositado em 21 de novembro de 2003, que é aqui incorporado por referência em sua totalidade. Os polímeros típicos usados na invenção incluem o polietileno glicol ("PEG"), também conhecido como poli(óxido de etileno) ("PEO"), e o polipropileno glicol (incluindo o poliisopropileno glicol). Adicionalmente, os copolímeros aleatórios ou em blocos de diferentes óxidos de alquileno (p.ex., o óxido de etileno e o óxido de propileno) podem ser usados em muitas aplicações. Na sua forma mais comum, um polialquileno glicol, tal como o PEG, é um polímero linear terminado em cada extremidade com grupos hidroxila: $\text{HO-CH}_2\text{CH}_2\text{O}-(\text{CH}_2\text{CH}_2\text{O})_n-\text{CH}_2\text{CH}_2-\text{OH}$. Esse polímero, alfa-, ômega-diidroxilpolietileno glicol, pode também ser representado como HO-PEG-OH , onde se entende que o símbolo —PEG— representa a seguinte unidade estrutural: $-\text{CH}_2\text{CH}_2\text{O}-(\text{CH}_2\text{CH}_2\text{O})_n-\text{CH}_2\text{CH}_2-$, onde n tipicamente varia de cerca de 4 a cerca de 10.000.

Conforme mostrado, a molécula de PEG é difuncional e é, algumas vezes, referida como "PEG diol". As porções terminais da molécula de PEG são porções de hidroxila relativamente não reativas, os grupos $-\text{OH}$, que podem ser ativadas, ou convertidas em porções funcionais, para a ligação do PEG a outros compostos em locais reativos

sobre o composto. Tais PEG dióis ativados são referidos nesse documento como PEGs bi-ativados. Por exemplo, as porções terminais do PEG diol têm sido funcionalizadas como éster de carbonato ativo, para a reação seletiva com as porções de amino, por substituição das porções de hidroxila relativamente não reativas, $-OH$, com as porções de éster de succinimidila ativas a partir da N-hidróxi succinimida.

Em muitas aplicações, é desejável recobrir a molécula de PEG sobre uma extremidade com uma porção essencialmente não reativa, de modo que a molécula de PEG seja mono-funcional (ou mono-ativada). No caso das substâncias terapêuticas protéicas, as quais geralmente mostram múltiplos locais de reação para os PEGs ativados, os PEGs ativados bi-funcionais resultam em reticulação ampla, produzindo agregados insatisfatoriamente funcionais. Para gerar PEGs mono-ativados, uma porção de hidroxila sobre a extremidade da molécula de PEG diol tipicamente é substituída com a porção de metóxi não reativa na extremidade, $-OCH_3$. A outra extremidade não recoberta da molécula de PEG tipicamente é convertida em uma porção reativa na extremidade que pode ser ativada para a ligação em um local reativo sobre uma superfície ou uma molécula, tal como uma proteína.

Os PAGs são polímeros que tipicamente têm as propriedades de solubilidade em água e em muitos solventes orgânicos, ausência de toxicidade, e ausência de imunogenicidade. Um uso dos PAGs é para unir covalentemente o polímero às moléculas insolúveis para tornar solúvel o

"conjugado" resultante de PAG-moléculas. Por exemplo, foi mostrado que a droga paclitaxel insolúvel em água, quando acoplada ao PEG, torna-se solúvel em água. Greenwald, e col., *J. Org. Chem.*, 60:331-336 (1995). Os conjugados de
5 PAG são frequentemente usados não somente para aumentar a solubilidade e a estabilidade, como também para prolongar a meia-vida de circulação no sangue das moléculas.

Os compostos polialquilados da invenção são tipicamente entre 5 e 80 kDa de tamanho, entretanto,
10 qualquer tamanho pode ser usado, a escolha dependente do aptâmero e da aplicação. Os outros compostos de PAG da invenção são entre 10 e 80 kDa de tamanho. Ainda outros compostos de PAG da invenção são entre 10 e 60 kDa de tamanho. Por exemplo, um polímero de PAG pode ser pelo
15 menos 10, 20, 30, 40, 50, 60, ou 80 kDa de tamanho. Tais polímeros podem ser lineares ou ramificados. Em algumas modalidades, os polímeros são o PEG.

Em contraste com as substâncias terapêuticas protéicas biologicamente expressas, as substâncias
20 terapêuticas de ácidos nucleicos são típica e quimicamente sintetizadas a partir de nucleotídeos de monômeros ativados. Os conjugados de PEG-ácido nucleico podem ser preparados por incorporação do PEG usando a mesma síntese de monômero iterativa. Por exemplo, os PEGs ativados por conversão em
25 uma forma de fosforamidita podem ser incorporados na síntese de oligonucleotídeos em fase sólida. Alternativamente, a síntese de oligonucleotídeos pode ser completada com a incorporação específica para o sítio de um sítio de união ao

PEG reativo. Mais comumente, isso tem sido efetuado por adição de uma amina primária livre na extremidade de 5' (incorporada usando uma fosforamidita modificadora na última etapa de combinação da síntese em fase sólida). Usando essa
5 abordagem, um PEG reativo (p.ex., um que seja ativado de modo que ele reagirá e formará uma ligação com uma amina) é combinado com o oligonucleotídeo purificado e a reação de combinação é efetuada em solução.

A capacidade da conjugação do PEG de alterar a
10 biodistribuição de uma substância terapêutica está relacionada com diversos fatores, incluindo o tamanho aparente (p.ex., conforme medido em termos de raio hidrodinâmico) do conjugado. Os conjugados maiores (>10 kDa) são sabidos bloquear mais efetivamente a filtração via
15 o rim e, conseqüentemente, aumentar a meia-vida em soro das pequenas macromoléculas (p.ex., os peptídeos, os oligonucleotídeos sem sentido). A capacidade dos conjugados de PEG de bloquear a filtração foi mostrada aumentar com o tamanho do PEG até aproximadamente 50 kDa (os aumentos
20 adicionais têm efeito benéfico mínimo, visto que a meia-vida torna-se definida pelo metabolismo mediado por macrófagos, em vez de por eliminação via os rins).

A produção de PEGs de altos pesos moleculares (>10 kDa) pode ser difícil, ineficiente, e cara. Como uma rota
25 para a síntese de conjugados de PEG de alto peso molecular-ácido nucléico, um trabalho anterior concentrou-se para a geração de PEGs ativados de pesos moleculares mais elevados. Um método para gerar tais moléculas envolve a formação de um

PEG ativado ramificado, no qual dois ou mais PEGs estão unidos a um núcleo central que carrega o grupo ativado. As porções terminais dessas moléculas de PEG de pesos moleculares mais elevados, i.e., as porções de hidroxila (-OH) relativamente não reativas, podem ser ativadas, ou convertidas em porções funcionais, para a união de um ou mais dos PEGs a outros compostos, em locais reativos sobre o composto. Os PEGs ativados ramificados terão mais do que duas extremidades, e nos casos onde duas ou mais extremidades tiverem sido ativadas, tais moléculas de PEG de pesos moleculares mais elevados, ativadas, são referidas nesse documento como PEGs multi-ativados. Em alguns casos, nem todas as extremidades em uma molécula de PEG ramificada estão ativadas. Nos casos onde quaisquer duas extremidades de uma molécula de PEG ramificada estiverem ativadas, tais moléculas de PEG são referidas como PEGs bi-ativados. Em alguns casos onde somente uma extremidade em uma molécula de PEG ramificada estiver ativada, tais moléculas de PEG são referidas como mono-ativadas. Como um exemplo dessa abordagem, foi descrito um PEG ativado preparado pela união de dois monometóxi PEGs a um núcleo de lisina, que é subsequentemente ativado para a reação (Harris e col., Nature, vol. 2: 214-221, 2003).

A presente invenção proporciona uma outra rota de custo efetivo para a síntese de conjugados de PEG de alto peso molecular-ácido nucléico (preferivelmente, o aptâmero), incluindo multiplicar os ácidos nucléicos PEGuilados. A presente invenção também inclui oligonucleotídeos

multiméricos ligados ao PEG p.ex., os aptâmeros dimerizados. A presente invenção também diz respeito às composições de altos pesos moleculares onde uma porção estabilizante de PEG é um ligador que separa diferentes porções de um aptâmero, p.ex., o PEG conjugado dentro de uma única sequência de aptâmeros, de modo tal que o arranjo linear da composição de aptâmero de alto peso molecular seja, p.ex., ácido nucleico - PEG - ácido nucleico $(- \text{PEG} - \text{ácido nucleico})_n$, onde n é maior do que, ou igual a, 1.

As composições de altos pesos moleculares da invenção incluem aquelas tendo um peso molecular de pelo menos 10 kDa. As composições tipicamente têm um peso molecular entre 10 e 80 kDa de tamanho. As composições de altos pesos moleculares da invenção são pelo menos 10, 20, 30, 40, 50, 60, ou 80 kDa de tamanho.

Uma porção estabilizante é uma molécula, ou porção de uma molécula, que aperfeiçoa as propriedades farmacocinéticas e farmacodinâmicas das composições de aptâmeros de altos pesos moleculares da invenção. Em alguns casos, uma porção estabilizante é uma molécula ou porção de uma molécula que coloca dois ou mais aptâmeros, ou domínios de aptâmeros, em proximidade, ou proporciona liberdade rotacional global diminuída das composições de aptâmeros de altos pesos moleculares da invenção. Uma porção estabilizante pode ser um polialquilenoglicol, tal como um polietileno glicol, que pode ser linear ou ramificado, um homopolímero ou um heteropolímero. As outras porções estabilizantes incluem polímeros tais como os ácidos

nucléicos de peptídeos (PNA). Os oligonucleotídeos podem também ser porções estabilizantes; tais oligonucleotídeos podem incluir nucleotídeos modificados, e/ou ligações modificadas, tais como os fosforotioatos. Uma porção
5 estabilizante pode ser uma parte integral de uma composição de aptâmero, i.e., ela está covalentemente ligada ao aptâmero.

As composições da invenção incluem as composições de aptâmeros de altos pesos moleculares nas quais duas ou
10 mais porções de ácidos nucléicos estão covalentemente conjugadas a pelo menos uma porção de polialquileno glicol. As porções de polialquileno glicol servem como porções estabilizantes. Nas composições onde uma porção de polialquileno glicol estiver covalentemente ligada, em
15 qualquer extremidade, a um aptâmero, de modo tal que o polialquileno glicol una as porções de ácidos nucléicos conjuntamente em uma molécula, o polialquileno glicol é dito ser uma porção ligadora. Em tais composições, a estrutura primária da molécula covalente inclui o arranjo linear ácido
20 nucléico-PAG-ácido nucléico. Um exemplo é uma composição tendo a estrutura primária ácido nucléico-PEG-ácido nucléico. Um outro exemplo é um arranjo linear de: ácido nucléico — PEG — ácido nucléico — PEG — ácido nucléico.

Para produzir o conjugado de ácido
25 nucléico-PEG-ácido nucléico, o ácido nucléico é originalmente sintetizado de modo tal que ele contenha um único local reativo (p.ex., ele é mono-ativado). Em uma modalidade preferida, esse local reativo é um grupo amino

introduzido na extremidade de 5' por adição de uma fosforamidita modificadora como a última etapa na síntese em fase sólida do oligonucleotídeo. Após a desproteção e a purificação do oligonucleotídeo modificado, ele é reconstituído em alta concentração em uma solução que minimiza a hidrólise espontânea do PEG ativado. Em uma modalidade preferida, a concentração de oligonucleotídeo é 1 mM e a solução reconstituída contém tampão de NaHCO_3 a 200 mM, pH 8,3. A síntese do conjugado é iniciada por adição em etapas, lenta, de PEG bi-funcional altamente purificado. Em uma modalidade preferida, o PEG diol é ativado em ambas as extremidades (bi-ativado) por derivação com propionato de succinimidila. Seguindo a reação, o conjugado de PEG-ácido nucléico é purificado por eletroforese em gel ou cromatografia líquida para separar as espécies inteiramente conjugadas, parcialmente conjugadas, e não conjugadas. As múltiplas moléculas de PAG concatenadas (p.ex., como copolímeros aleatórios ou em blocos) ou as cadeias de PAG menores podem ser ligadas para obter diversos comprimentos (ou pesos moleculares). Ligadores que não sejam de PAG podem ser usados entre as cadeias de PAG de comprimentos variados.

As modificações de nucleotídeos com 2'-O-metil, 2'-flúor e outras modificações de nucleotídeos modificados estabilizam o aptâmero contra as nucleases e aumentam a sua meia-vida *in vivo*. O quepi de 3'-3'-dT também aumenta a resistência à exonuclease. Ver, p.ex., as Patentes U.S. 5.674.685; 5.668.264; 6.207.816; e 6.229.002, cada uma das quais é incorporada por referência nesse documento em sua

totalidade.

DERIVAÇÃO COM PAG DE UM ÁCIDO NUCLÉICO REATIVO

Os conjugados de PAG de alto peso molecular-ácido
nucléico-PAG podem ser preparados por reação de um PEG
5 ativado mono-funcional com um ácido nucléico contendo mais
do que um local reativo. Em uma modalidade, o ácido
nucléico é bi-reativo, ou bi-ativado, e contém dois locais
reativos: um grupo 5'-amino e um grupo 3'-amino introduzidos
no oligonucleotídeo através de síntese convencional de
10 fosforamidita, por exemplo: a 3'-5'-di-PEGuilação como
ilustrada na Figura 13. Nas modalidades alternativas, os
locais reativos podem ser introduzidos em posições internas,
usando, por exemplo, a posição 5 das pirimidinas, a posição
8 das purinas, ou a posição 2' da ribose como locais para a
15 união de aminas primárias. Em tais modalidades, o ácido
nucléico pode ter diversos locais ativados ou reativos e é
dito ser multiplamente ativado. Seguindo a síntese e a
purificação, o oligonucleotídeo modificado é combinado com o
PEG mono-ativado, sob condições que promovam a reação
20 seletiva com os locais reativos do oligonucleotídeo, ao
mesmo tempo minimizando a hidrólise espontânea. Na
modalidade preferida, o monometóxi-PEG é ativado com o
propionato de succinimidila e a reação combinada é efetuada
em pH 8,3. Para orientar a síntese do PEG bi-substituído,
25 proporciona-se PEG em excesso estequiométrico em relação ao
oligonucleotídeo. Após a reação, o conjugado de PEG-ácido
nucléico é purificado por eletroforese em gel ou
cromatografia líquida para separar as espécies inteiramente

conjugadas, parcialmente conjugadas, e não conjugadas.

Os domínios de ligação podem também ter uma ou mais porções de polialquileno glicol unidas a eles. Tais PAGs podem ser de comprimentos variados e podem ser usados em combinações apropriadas para obter o peso molecular desejado da composição.

O efeito de um ligador particular pode ser influenciado tanto por sua composição química quanto pelo comprimento. Um ligador que seja muito longo, muito curto, ou forme interações estéricas e/ou iônicas desfavoráveis com o alvo impedirá a formação de complexo entre o aptâmero e o alvo. Um ligador, o qual é mais longo do que o necessário para transpor a distância entre os ácidos nucleicos, pode reduzir a estabilidade de ligação por diminuição da concentração efetiva do ligante. Assim, é frequentemente necessário otimizar as composições e os comprimentos dos ligadores para maximizar a afinidade de um aptâmero a um alvo.

Todas as publicações e documentos de patentes citados nesse documento são incorporados aqui por referência como se cada tal publicação ou documento fosse específica e individualmente indicado para ser incorporado nesse documento por referência. A citação de publicações e documentos de patentes não é pretendida como uma aceitação que algum seja anterioridade pertinente, nem constitui qualquer aceitação quanto aos conteúdos ou data deles. A invenção tendo agora sido descrita como descrição escrita, aqueles de habilidade na técnica reconhecerão que a invenção

pode ser praticada em uma variedade de modalidades e que a descrição precedente e os exemplos abaixo são para propósitos de ilustração e não limitação das reivindicações que se seguem.

5 Todas as publicações e documentos de patentes citados nesse documento são incorporados aqui por referência como se cada tal publicação ou documento fosse específica e individualmente indicado para ser incorporado nesse documento por referência. A citação de publicações e
10 documentos de patentes não é pretendida como uma aceitação que algum seja anterioridade pertinente, nem constitui qualquer aceitação quanto aos conteúdos ou data deles. A invenção tendo agora sido descrita como descrição escrita, aqueles de habilidade na técnica reconhecerão que a invenção
15 pode ser praticada em uma variedade de modalidades e que a descrição precedente e os exemplos abaixo são para propósitos de ilustração e não limitação das reivindicações que se seguem.

EXEMPLOS

20 EXEMPLO 1: IDENTIFICAÇÃO DE SEQUÊNCIAS LÍDERES EM 5' USANDO O MÉTODO TR-SELEX®

Uma biblioteca de DNA degenerado, com o seguinte projeto (mostrado na direção 5' para 3'):

Promotor de T7 / G₄ / 20 nucleotídeos degenerados
25 / Sequência fixa em 3' foi sintetizada com a seguinte sequência:

5'TAATACGACTCACTATAGGGGNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNACGTAAC
CGGTAAACCCGGGTCGATGCAGTAAGCTAGCT3' (ARC1140, SEQ ID NO 7)).

Essa biblioteca foi amplificada usando o primer em 3' AGCTAGCTTACTGCATCGAC (SEQ ID NO 104) e o primer em 5' TAATACGACTCACTATAG (SEQ ID NO 105). A biblioteca de fita dupla foi então transcrita usando 1X Tampão de Transcrição (HEPES a 200 mM, DTT a 40 mM, espermidina a 2 mM, Triton X-100 a 0,01%), a 37°C, durante a noite, sob as seguintes condições: 2'-OMe ATP CTP, UTP, GTP a 1 mM cada, 2'-OH GTP a 30 µM, MgCl₂ a 6,5 mM, MnCl₂ a 2,0 mM, 10% p/v de PEG-8000, GMP a 1 mM, 0,5 unidade de pirofosfatase inorgânica por 100 µL de reação, e T7 RNA polimerase Y639F/H784A/K378R a 200 nM.

A mistura resultante foi então precipitada (isopropanol, cloreto de sódio, EDTA), purificada em gel (10% de PAGE), excisada e extraída do gel, tratada com DNase (RQ1, Promega, Madison WI), transcrita invertida a 65°C (Thermoscript, Invitrogen, Carlsbad, CA) usando o primer em 3' usado para a PCR, e diluída diretamente para uma reação de ligação com esplinte com os seguintes nucleotídeos.

Oligonucleotídeo 5' fosforilado codificando um promotor de T7 (onde p significa a fosforilação em 5'):

20 pTATAGTGAGTCGTATTA 3' (SEQ ID NO 8)

Esplinte para a ligação:

5'TAATACGACTCACTATAGGGG 3' (SEQ ID NO 9)

Essa mistura foi desnaturada por calor, anelada, e então a T4 DNA ligase (NEB, Beverley MA) foi adicionada, seguida por incubação a 16°C durante a noite. Subseqüente à etapa de ligação, a reação foi diretamente diluída para uma PCR, com os primers já descritos, para amplificar as seqüências transcritas para a entrada na rodada seguinte do

método SELEX[®]. Esse esquema é apresentado na Figura 2.

Após três rodadas de seleção por TR-SELEX[®], a biblioteca foi clonada usando um kit de clonagem TOPO TA conforme instruções do fabricante (Invitrogen, Carlsbad, CA),
5 seqüenciada, e a estatística da ocorrência de nucleotídeo na região degenerada foi analisada. Os clones individuais foram avaliados através de análise por PAGE-gel quanto à sua capacidade de moldar a transcrição de grandes concentrações de transcrito, e as seqüências daqueles que produziram as
10 maiores produções de transcrito foram então utilizadas no projeto de bibliotecas, que foram, por sua vez, testadas através de análise por gel quanto à sua capacidade de moldar a transcrição de altas produções de transcrito. A Figura 3 mostra a porcentagem média da composição de nucleotídeo de
15 regiões das 20 posições degeneradas antes e depois de 3 rodadas da seleção por TR-SELEX[®]. Conforme indicado pela Figura 3, uma forte preferência por G das posições 5 a 13 no transcrito (1 a 9 na região degenerada) foi transcrita, depois nenhum nucleotídeo é preferencialmente transcrito.

20 Os clones descobertos por seqüenciamento após 3 Rodadas da seleção por TR-SELEX[®] foram classificados através de análise por PAGE-gel quanto à sua capacidade de transcrever os 2'-OMe nucleotídeos usando molde a ~200 nM, 1X Tampão de Transcrição (HEPES a 200 mM, DTT a 40 mM,
25 espermidina a 2 mM, Triton X-100 a 0,01%), 2'-OMe ATP CTP, UTP, GTP a 1 mM cada, 2'-OH GTP a 30 uM, MgCl₂ a 6,5 mM, MnCl₂ a 2,0 mM, 10% p/v de PEG-8000, GMP a 1 mM, 0,5 unidade de pirofosfatase inorgânica por 100 µL de reação, e T7 RNA

polimerase mutante Y639F/H784A/K378R a 200 nM, a 37°C, durante a noite. Um exemplo de um clone da Rodada 3, o clone AMX411.D6, deu significativamente mais transcrito de MNA, conforme visualizado por PAGE-gel, quando comparado aos clones da Rodada 0. As seqüências de DNA dos clones gerados a partir da Rodada 3 são listadas abaixo (todas as seqüências listadas estão na direção 5'-3'):

SEQ ID NO 10 >AMX(411)_A1 ARC 1140 Rd 3_411-A1

TAATACGACTCACTATAGGGGGTGGGGCCAATGGCGGGATATACGTAACCGGTTATACCC

10 GGGTCGATGCAGTAAGCTAGCT

SEQ ID NO 11 >AMX(411)_B1 ARC 1140 Rd 3_411-B1

TAATACGACTCACTATAGGGGATGTACATATGTATTTCGTGACGTGACCGGTTAAACCCGG

GTCGATGCAGTAAGCTAGCT

SEQ ID NO 12 >AMX(411)_C1 ARC 1140 Rd 3_411-C1

15 TAATACGACTCACTATAGGGGGAGCGGGGAGACGTAGTCATCACGTAGCCGGTTAAACCC

GGGTCGATGCAGTAAGCTAGCT

SEQ ID NO 13 >AMX(411)_D1 ARC 1140 Rd 3_411-D1

TAATACNACTCACTATAGGGGGTGGGGGTGGTGGTGATAACGTAACCGGTTAAACCCGGG

TCGATGCAGTAAGCTAGCT

20 SEQ ID NO 14 >AMX(411)_E1 ARC 1140 Rd 3_411-E1

TAATACGACTCACTATAGGGGGTGTACCAGATATGCCTTGAACGTAACCCGTTAAACC

CGGGTCGATGCAGTAAGCTAGCT

SEQ ID NO 15 >AMX(411)_F1 ARC 1140 Rd 3_411-F1

TAATACGACTCACTATAGGGGGTAGGGGGCACGCACTAACCAACGTAACCGGTTAAACCC

25 GGGTCGATGCAGTAAGCTAGCT

SEQ ID NO 16 >AMX(411)_G1 ARC 1140 Rd 3_411-G1

TAATACGACTCACTATAGGGGGAGGGGGTGCTGACCNCAAACA

SEQ ID NO 17 >AMX(411)_H1 ARC 1140 Rd 3_411-H1

TAATACGACTCACTATAGGGGTGGGGCTCGGATGAGACAATACGTAACCGGTTAAACCCG
GGTCGATGCAGTAAGCTAGCT

SEQ ID NO 18 >AMX(411)_A2 ARC 1140 Rd 3_411-A2

TAATACGACTCACTATAGGGGGGGGTGGGTAGGCGAGCACTCCACGTAACCGTTAAACC

5 CGGGTCGATGCAGTAAGCTAGCT

SEQ ID NO 19 >AMX(411)_B2 ARC 1140 Rd 3_411-B2

TAATACGACTCACTATAGGGGGGAAGGACGAGCAGACGAGCAACGTAACCTGTAAACCC

GGGTCGATGCAGTAAGCTAGCT

SEQ ID NO 20 >AMX(411)_C2 ARC 1140 Rd 3_411-C2

10 TAATACGACTCACTATAGGGGGGGCGGTTAGAGTGTAAGTACCGACGTAACCGGTTAAA

CCCGGGTCGATGCAGTAAGCTAGCT

SEQ ID NO 21 >AMX(411)_D2 ARC 1140 Rd 3_411-D2

TAATACGACTCACTATAGGGGGGTTGCTGTAGTAACGCCACGTAACCGGTTAACTTGG

TCGATGCAGTAAGCTAGCT

15 SEQ ID NO 22 >AMX(411)_E2 ARC 1140 Rd 3_411-E2

TAATACGACTCACTATAGGGGGGCGGGAGAATGTTATATAGTTACGGTAACCGGTTAAAC

CCGGGTCGATGCAGTAAGCTAGCT

SEQ ID NO 23 >AMX(411)_F2 ARC 1140 Rd 3_411-F2

TAATACGACTCACTATAGGGGAAAGGGGCGGTATGGTACACACGTAACAGGTTAAACCCG

20 GGTCGATGCAGTAAGCTAGCT

SEQ ID NO 24 >AMX(411)_G2 ARC 1140 Rd 3_411-G2

TAATACGACTCACTATAGGGGGGACGTGTAGCATTCCAGAATTCGTAACCTAAACCCGG

GTCGATGCAGTAAGCTAGCT

SEQ ID NO 25 >AMX(411)_H2 ARC 1140 Rd 3_411-H2

25 TAATACGACTCACTATAGGGGGCGTGGGAGATAGGTTCAAGGACGTACCGGTTATACCCG

GGTCGATGCAGTAAGCTAGCT

SEQ ID NO 26 >AMX(411)_A3 ARC 1140 Rd 3_411-A3

TAATACGACTCACTATAGGGGGGCTCCGTGCTATCGTCGGATAACGTAACCCGTTAAACC

CGGGTCGATGCAGTAAGCTAGCT

SEQ ID NO 27 >AMX(411)_B3 ARC 1140 Rd 3_411-B3

TAATACGACTCACTATAGGGGGGAGAAGGTCTTAAGGTCGCCAACGTAACGTGTTAAACC

CGGGTCGATGCAGTAAGCTAGCT

5 SEQ ID NO 28 >AMX(411)_C3 ARC 1140 Rd 3_411-C3

TAATACGACTCACTATAGGGGGGGCATAACGAGTTTAGGTGGAGACGTAACCGGTTAAACC

CGGGTCGATGCAGTAAGCTAGCT

SEQ ID NO 29 >AMX(411)_D3 ARC 1140 Rd 3_411-D3

TAATACGACTCACTATAGGGGGATGATGACTTCCGCGTTAATACGTTACCGGTTAAACCC

10 GGGTCGATGCAGTAAGCTAGCT

SEQ ID NO 30 >AMX(411)_E3 ARC 1140 Rd 3_411-E3

TAATACGACTCACTATAGGGGTGGGACGCCGTCTGAGTATAACGTACCCGGTCGGGTGCA

TGCAGTAAGCTAGCT

SEQ ID NO 31 >AMX(411)_G3 ARC 1140 Rd 3_411-G3

15 TAATACGACTCACTATAGGGGGGGGGGACGTAATCGGCTATCGTTACCGTAACCGGTT

AAACCCGGGTCGATGCAGTAAAGGGCGA

SEQ ID NO 32 >AMX(411)_H3 ARC 1140 Rd 3_411-H3

TAATACGACTCACTATAGGGGTGGGACGGGCAGCGTGGATGTAGGACGTAACCGGTTAAAC

GCGGGTCGATGCAGTAAGCTAGCT

20 SEQ ID NO 33 >AMX(411)_A4 ARC 1140 Rd 3_411-A4

TAATACGACTCACTATAGGGGGGTTTGTCTGAAGTGAAGCAGAACGTAACCGGTTAATCC

CGGGTCGATGCAGTAAGCTAGCT

SEQ ID NO 34 >AMX(411)_B4 ARC 1140 Rd 3_411-B4

TAATACGACTCACTATAGGGGGGGAGGGCACATCATCGTATCAAACGTAACCGTTAATC

25 CCGGGTCGATGCAGTAAGCTAGCT

SEQ ID NO 35 >AMX(411)_C4 ARC 1140 Rd 3_411-C4

TAATACGACTCACTATAGGGGAGGCTAGAGGACGCGACAGAACGTAACCGGTTAAACCCG

GGTCGATGCAGTAAGCTAGCT

SEQ ID NO 36 >AMX(411)_D4 ARC 1140 Rd 3_411-D4

TAATACGACTCACTATAGGGGGCGATCGCGAAGGGATTTC AACGTAACCGGTTAAACCCG
GGTCGATGCAGTAAGCTAGCT

SEQ ID NO 37 >AMX(411)_E4 ARC 1140 Rd 3_411-E4

5 TAATACGACTCACTATAGGGGGGTAGGGAAAGATTACGGGGCTACGTAACCGGTTATACC
TGGGTCGATGCAGTAAGCTAGCT

SEQ ID NO 38 >AMX(411)_F4 ARC 1140 Rd 3_411-F4

TAATACGACTCACTATAGGGGTGGCTATGGCTAACACGTAACCGGTTATACCCGGGTCGA
TGCAGTAAGCTAGCT

10 SEQ ID NO 39 >AMX(411)_G4 ARC 1140 Rd 3_411-G4

TAATACGACTCACTATAGGGGGGGGGCGGTGGCTGTGCAAGCGGAAACGTAACCGGTTAA
ACCCGGGTCGATGCAGTAAGCTAGCT

SEQ ID NO 40 >AMX(411)_H4 ARC 1140 Rd 3_411-H4

TAATACGACTCACTATAGGGGGGTGGGGGCACGGTACTGAGTTACGTTACCGGTTAAACC
15 CGGGTCGATGCAGTAAGCTAGCT

SEQ ID NO 41 >AMX(411)_A5 ARC 1140 Rd 3_411-A5

TAATACGACTCACTATAGGGGGGAGTGGGGACAATTAGAAGATGACGTAACCGTCCGGGT
CGATGCAGTAAGCTAGCT

SEQ ID NO 42 >AMX(411)_B5 ARC 1140 Rd 3_411-B5

20 TAATACNACTCACTATAGGGGTGCAGTGAGGAGCGACNAGTACGTTACCGGTTAAATCCG
AGTCGATGCAGTAAGCTAGCT

SEQ ID NO 43 >AMX(411)_C5 ARC 1140 Rd 3_411-C5

TAATACNACTCACTATAGGGGGACGGGCACTGTGGATGATTTAACGTTACCGGTTAAACC
CGAGTCGATGCAGTAAGCTAGCT

25 SEQ ID NO 44 >AMX(411)_D5 ARC 1140 Rd 3_411-D5

TAATACNACTCACTATAGGGGTCGATGCAGTAAGCTAGCT

SEQ ID NO 45 >AMX(411)_E5 ARC 1140 Rd 3_411-E5

TAATACNACTCACTATAGGGGGTGATATTGACCTCTAACAGCACGTAACCGGTTAAACCC

GGTCGATGCAGTAAGCTAGCT

SEQ ID NO 46 >AMX(411)_F5 ARC 1140 Rd 3_411-F5

TAATACGACTCACTATAGGGGGGGGGTGCAGAGGATGCATCCAAGCTCGTAATCGGTGG

TCGATGCAGTAAGCTAGCT

5 SEQ ID NO 47 >AMX(411)_G5 ARC 1140 Rd 3_411-G5

TAATACGACTCACTATAGGGGGGGGCGGGTGCTUGTGCCTAATCACGTAACCGGTTAAAC

CCGGGTCGATGCAGTAAGCTAGCT

SEQ ID NO 48 >AMX(411)_H5 ARC 1140 Rd 3_411-H5

TAATACGACTCACTATAGGGGTTTGTAATCGAACGTGGAACGCAACCGGTTTAACCGGG

10 TCGATGCAGTAAGCTAGCT

SEQ ID NO 49 >AMX(411)_A6 ARC 1140 Rd 3_411-A6

TAATACGACTCACTATAGGGGGGATGGAAGAGGCTTGATATCACGTAACCGGTTAAACCT

GGGTCGATGCAGTAAGCTAGCT

SEQ ID NO 50 >AMX(411)_B6 ARC 1140 Rd 3_411-B6

15 TAATACGACTCACTATAGGGGGTTATACTAACTCTGTACACAACGTAACCGGCCGGGTGCG

ATGCAGTAAGCTAGCT

SEQ ID NO 51 >AMX(411)_C6 ARC 1140 Rd 3_411-C6

TAATACGACTCACTATAGGGGTATAGGGGGGGTATCGGTGTACGTAACCGGTTAAACCCG

GGTCGATGCAGTAAGCTAGCT

20 SEQ ID NO 52 >AMX(411)_D6 ARC 1140 Rd 3_411-D6

TAATACGACTCACTATAGGGGAGTACAATAAGGTTCCGAGAACGCGACCGGTTAAACCCG

GGTCGATGCAGTAAGCTAGCT

SEQ ID NO 53 >AMX(411)_E6 ARC 1140 Rd 3_411-E6

TAATACGACTCACTATAGGGGTGCGCGGTTACAAGGCAACATACGTAACCGGTTAAACCC

25 GGGTCGATGCAGTAAGCTAGCT

SEQ ID NO 54 >AMX(411)_F6 ARC 1140 Rd 3_411-F6

TAATACNACTCACTATAGGGGGGACGGGTGACAAAGTGTCAACGTAACCGGTTAAACC

CGGGTCGATGCAGTAAGCTAGCT

SEQ ID NO 55>AMX(411)_G6 ARC 1140 Rd 3_411-G6

TAATACGACTCACTATAGGGGAGACGGCGGTACAAGTCCATATGTAACCGGTTAAACCCG
GGTCGATGCAGTAAGCTAGCT

SEQ ID NO 56 >AMX(411)_H6 ARC 1140 Rd 3_411-H6

5 TAATACGACTCACTATAGGGGAGTGGGGGCTTCTCGTTGCCACGTAACCGCTTAAACCCG
GGTCGATGCAGTAAGCTAGCT

SEQ ID NO 57 >AMX(423)_A7 ARC 1140 R3_423-A7

TAATACGACTCACTATAGGGGGGCTGAGCGTGTTTGAGGGACCACGTTACCGGTTAAACC
CGGGTCGATGCAGTAAGCTAGCT

10 SEQ ID NO 58 >AMX(423)_B7 ARC 1140 R3_423-B7

TAATACGACTCACTATAGGGGGGTGGGCGCAATGAAAAGTTGGGCGTAACCGGTTCAACC
CGGGTCGATGCAGTAAGCTAGCT

SEQ LD NO 59 >AMX(423)_C7 ARC 1140 R3_423-C7

TAATACGACTCACTATAGGGGGGTAGTGAAGTAAGGCAGTGTTACGTAACCGGTGAACCCG
15 GGTCGATGCAGTAAGCTAGCT

SEQ ID NO 60 >AMX(423)_D7 ARC 1140 R3_423-D7

TAATACGACTCACTATAGGGGGGAGGGTGGGCTAGAACACACAACGTAACCGGTTAAACC
CGGGTCGATGCAGTAAGCTAGCT

SEQ ID NO 61 >AMX(423)_E7 ARC 1140 R3_423-E7

20 TAATACGACTCACTATAGGGGGGAGAGAGGCGGTTACGTAGGGACGTTACCGATTGAAC
TCAGGTCGATGCAGTAAGCTAGCT

SEQ ID NO 62 >AMX(423)_F7 ARC 1140 R3_423-F7

TAATACGACTCACTATAGGGGGGGGGGCGAATAGGTAGGGCGACGAACGTTACCGGTTA
AACCCGGGTCGATGCAGTAAGCTAGCT

25 SEQ ID NO 63 >AMX(423)_G7 ARC 1140 R3_423-G7

TAATACGACTCACTATAGGGGGAGAGGAGGTCCGGCTAGACAACGTAACCGGTTAAACCC
GGGTCGATGCAGTAAGCTAGCT

SEQ ID NO 64 >AMX(423)_H7 ARC 1140 R3_423-H7

TAATACGACTCACTATAGGGGGGAGGACGGGTCGTACTGTAAACCTGGGTCGATGCAGT
AAGCTAGCT

SEQ ID NO 65 >AMX(423)_B8 ARC 1140 R3_423-B8

TAATACGACTCACTATAGGGGGCGCAACAACGGGAAGTATACGTAACCGGTTTAAACCCG

5 GGTTCGATGCAGTAAGCTAGCT

SEQ ID NO 66 >AMX(423)_C8 ARC 1140 R3_423-C8

TAATACGACTCACTATAGGGGGAAGGAACACGCACATGCATAACGTAACCTGGTTGACCCC

GGGTCGATGCAGTAAGCTAGCT

SEQ ID NO 67 >AMX(423)_D8 ARC 1140 R3_423-D8

10 TAATACGACTCACTATAGGGGAGTGGGGAGTACTGTGGACAACGTGACCGGTTAAACCCG

GGTTCGATGCAGTAAGCTAGCT

SEQ ID NO 68 >AMX(423)_E8 ARC 1140 R3_423-E8

TAATACGACTCACTATAGGGGTCGATGCAGTAAGCTAGCT

SEQ ID NO 69 >AMX(423)_F8 ARC 1140 R3_423-F8

15 TAATACGACTCACTATAGGGGGGGGGGCTAGGGCGGTTCGGATCGGACGTAACCAAGTTAAA

CCCCGGTTCGATGCAGTAAGCTAGCT

SEQ ID NO 70 >AMX(423)_G8 ARC 1140 R3_423-G8

TAATACGACTCACTATAGGGGGGGTGGGGGTTGCTACATGCCCTCGTAACCGGTTAAGCC

CAGGTCGATGCAGTAAGCTAGCT

20 SEQ ID NO 71 >AMX(423)_H8 ARC 1140 R3_423-H8

TAATACGACTCACTATAGGGGGGTGGCGACGATGGAGAGAATAACGTAATCGGTTAAACC

CGGGTCGATGCAGTAAGCTAGCT

SEQ ID NO 72 >AMX(423)_A9 ARC 1140 R3_423-A9

TAATACGACTCACTATAGGGGGTAGGCGGGCCTCATCAACAACGCAACCGGTTAAACCCG

25 GGTTCGATGCAGTAAGCTAGCT

SEQ ID NO 73 >AMX(423)_B9 ARC 1140 R3_423-B9

TAATACGACTCACTATAGGGGGTGGCTGGTAAGGACACAAACACGTAACCTCGTTAAACCC

GGGTCGATGCAGTAAGCTAGCT

SEQ ID NO 74 >AMX(423)_C9 ARC 1140 R3_423-C9

TAATACGACTCACTATAGGGGGGCGGGCAGCGCTTATAGATCCACGTAACCGGTTAAACC
CGGGTCGATGCAGTAAGCTAGCT

SEQ ID NO 75 >AMX(423)_D9 ARC 1140 R3_423-D9

5 TAATACGACTCACTATAGGGGGGGGGTATCTGCGGTTAGGCTATCGACGTACCCAGTTAA
ACCCGGGTCGATGCAGTAAGCTAGCT

SEQ ID NO 76 >AMX(423)_F9 ARC 1140 R3_423-F9

TAATACGACTCACTATAGGGGGGGTAGGGGACATCATAGGTATACGTAACCGGTTAACCC
GGGTCGATGCAGTAAGCTAGCT

10 SEQ ID NO 77 >AMX(423)_H9 ARC 1140 R3_423-H9

TAATACGACTCACTATAGGGGGCGCGTGCGTGTATCCATTAAACGTGACTGGTTAAACCCG
GGTCGATGCAGTAAGCTAGCT

SEQ ID NO 78 >AMX(423)_A10 ALRC 1140 R3_423-A10

TAATACGACTCACTATAGGGGGGGGAGCGTGGATCTTGAGTGTATACGTAACCGGTTAAA
15 CCCGGTCGATGCAGTAAGCTAGCT

SEQ ID NO 79 >AMX(423)_B10 ARC 1140 R3_423-B10

TAATACGACTCACTATAGGGGATGGAGAGGAGTGTACGCATATACAACCGGTTAAACCCG
GGTCGATGCAGTAAGCTAGCT

SEQ ID NO 80 >AMX(423)_C10 ARC 1140 R3_423-C10

20 TAATACGACTCACTATAGGGGCGGGTGGTCGCGATGGTTAACGTAAGTGGTTAAACCCGG
GTCGATGCAGTAAGCTAGCT

SEQ ID NO 81 >AMX(423)_D10 ARC 1140 R3_423-D10

TAATACGACTCACTATAGGGGGGGGGGGGACGTTAGCTTCTCTGTATTTACGTAACCGG
TTAAGCCCGGGTCGATGCAGTAAGCTAGCT

25 SEQ ID NO 82 >AMX(423)_E10 ARC 1140 R3_423-E10

TAATACGACTCACTATAGGGGTCGATGCAGTAAGCTAGCT

SEQ ID NO 83 >AMX(423)_F10 ARC 1140 R3_423-F10

TAATACGACTCACTATAGGGGGGATGGAGTGGGTGCAAATAANACGTAAGTGGTTAAACC

CGGGTCGATGCAGTAAGCTAGCT

SEQ ID NO 84 >AMX(423)_G10 ARC 1140 R3_423-G10

TAATACGACTCACTATAGGGGAGNGTGAQGGGTGAATANTAANGTAANCNGTTAAACCTG

GGTCGATGNNNTANNCTNGNT

5 SEQ ID NO 85 >AMX(423)_H10 ARC 1140 R3_423-H10

NAATNNGACTCACAANAGGGGTCGATGCAGTAAGCTAGCT

SEQ ID NO 86 >AMX(423)_A11 ARC 1140 R3_423-A11

TAATACGACTCACTATAGGGGGGGGTGACGTACGGATCTAAGTAACGTAACCGGTTAAAC

CCGGGTCGATGCAGTAAGCTAGCT

10 SEQ ID NO 87 >AMX(423)_B11 ARC 1140 R3_423-B11

TAATACGACTCACTATAGGGGAGGGACAGACACTTTGTAGACGTAACCAGTTAAACCCGG

GTCGATGCAGTAAGCTAGCT

SEQ ID NO 88 >AMX(423)_C11 ARC 1140 R3_423-C11

TAATACGACTCACTATAGGGGGGGGACTTGGCACTACGTAACAACGTAACCGCTTAAACC

15 CGGGTCGATGCAGTAAGCTAGCT

SEQ ID NO 89 >AMX(423)_D11 ARC 1140 R3_423-D11

TAATACGACTCACTATAGGGGGGGGGCCTCTCGACCAAAGCCCAACGTAACCGGTTAA

ACCCGGGTCGATGCAGTAAGCTAGCT

SEQ ID NO 90 >AMX(423)_E11 ARC 1140 R3_423-E11

20 TAATACNACTCACTATAGGGGGGGGGGATAGTCATGACTGATAAACGTAACGTGTTGAG

CCCGGGTCGATGCAGTAAGCTAGCT

SEQ ID NO 91 >AMX(423)_F11 ARC 1140 R3_423-F11

TAATACGACTCACTATAGGGGACAGTGCTAGTGGAATAGCAACGTAACCAGTTAAACCCG

GGTCGATGCAGTAAGCTAGCT

25 SEQ ID NO 92 >AMX(423)_G11 ARC 1140 R3_423-G11

TAATACGACTCACTATAGGGGACGACCACTATACTCCGAGAACGTAACCGGTTAAACCCG

GGTCGATGCAGTAAGCTAGCT

SEQ ID NO 93 >AMX(423)_H11 ARC 1140 R3_423-H11

TAATACGACTCACTATAGGGGGATGGAGGCGTAGTGTAGTCAACGTTACCGGTTAAACCC
GGGTCGATGCAGTAAGCTAGCT

SEQ ID NO 94 >AMX(423)_A12 ARC 1140 R3_423-A12

TAATACGACTCACTATAGGGGGGAGGTATAGATGGAATGGTTATGTAACCTGTAAACCC

5 GGGTCGATGCAGTAAGCTAGCT

SEQ ID NO 95 >AMX(423)_B12 ARC 1140 R3_423-B12

TAATACGACTCACTATAGGGGTGGGGAGGACCACTTAGATAACGTCACCGGTTAAACCCG

GGTCGATGCAGTAAGCTAGCT

SEQ ID NO 96 >AMX(423)_C12 ARC 1140 R3_423-C12

10 TAATACGACTCACTATAGGGGGGATAGGGGCGAGAGAGTCACAACGTAACCGGTTAATCC

CGGGTCGATGCAGTAAGCTAGCT

SEQ ID NO 97 >AMX(423)_E12 ARC 1140 R3_423-E12

TAATACGACTCACTATAGGGGGGGGATGGCCGAATCATAAAATAACGTAACCGTTAGACC

CGGGTCGATGCAGTAAGCTAGCT

15 SEQ ID NO 98 >AMX(423)_F12 ARC 1140 R3_423-F12

TAATACGACTCACTATAGGGGGCGATTGCTGAGTCAGTtCGTAATCGGTTAAACCCGGGT

QGATGCAGTAAGCTAGCT

SEQ ID NO 99 >AMX(423)_G12 ARC 1140 R3_423-G12

TAATACGACTCACTATAGGGGGGGGAGGATCCGAAACACGGGATCCGTAACCGGTTAAAG

20 CCGGGTCGATGCAGTAAGCTAGCT

EXEMPLO 2: BIBLIOTECAS QUE INCORPORAM SEQUÊNCIAS LÍDERES IDENTIFICADAS PELO MÉTODO TR-SELEX®

Os elementos da sequência líder em 5'

25 identificados (os primeiros 10 nucleotídeos da região
degenerada) a partir de clones maiores produtores de
transcritos modificados em 2', identificados usando a
seleção por TR-SELEX®, como descrito no Exemplo 1, foram

utilizados para projetar bibliotecas que incorporam os elementos da sequência líder na região fixa em 5', com o objetivo de promover um aumento na produção de transcrito contendo nucleotídeos modificados em 2'. Em uma modalidade, a estratégia de projeto incorpora os primeiros 14 nucleotídeos dos clones identificados (as 4 guanossinas compreendendo a região fixa em 5' mais os primeiros 10 nucleotídeos da região degenerada) como a sequência líder em 5', imediatamente seguidos por uns 6-8 nucleotídeos fixos adicionais para facilitar a amplificação por PCR subsequente, imediatamente seguidos por 30-40 nucleotídeos de comprimento de uma região degenerada, imediatamente seguidos por uma região fixa em 3' para também facilitar a amplificação por PCR subsequente.

Os exemplos das sequências de DNA das bibliotecas projetadas, as quais incorporam os elementos da sequência líder identificados, são listados abaixo.

Para cada uma das sequências das bibliotecas de moldes de transcrição de DNA listadas abaixo, o elemento da sequência líder em 5' é mostrado sublinhado, e todas as sequências estão na direção 5'-3'.

ARC2118 (SEQ ID NO 3)

TAATACGACTCACTATAGGGGAGTACAATAACGTTCTCGNNNNNNNNNNNNNNNNNNNNNN
NNNNNNNNNGGATCGTTACGACTAGCATCGATG

ARC2119 (SEQ ID NO 4)

TAATACGACTCACTATAGGGGGTGATATTGACGTTCTCGNNNNNNNNNNNNNNNNNNNNNN
NNNNNNNNNGGATCGTTACGACTAGCATCGATG

ARC2120 (SEQ ID NO 5)

TAATACGACTCACTATAGGGGTGCGCGGTTACGTTCTCGNNNNNNNNNNNNNNNNNNNNNN
 NNNNNNNNNNGGATCGTTACGACTAGCATCGATG

ARC2121 (SEQ ID NO 6)

TAATACGACTCACTATAGGGGGAGGGGGTGCCGTTCTCGNNNNNNNNNNNNNNNNNNNNNN
 5 NNNNNNNNNNGGATCGTTACGACTAGCATCGATG

Um molde de transcrição de DNA de controle, sem uma sequência líder, é listado abaixo, na direção 5'-3'.

TAATACGACTCACTATAGGGGAGAGGAGAGAACGTTCTCGNNNNNNNNNNNN
 NNNNNNNNNNNNNNNNNNNNNNGGATCGTTACGACTAGCATCGATG (ARC2117, SEQ
 10 ID NO 106).

Para testar se as bibliotecas recentemente projetadas promovem uma produção aumentada de transcritos contendo 2'-O-metil nucleotídeos, as bibliotecas foram transcritas usando duas T7 RNA polimerases modificadas
 15 diferentes para comparação, a T7 RNA polimerase mutante Y639F/H784A/K378R, e a polimerase mutante Y639L/H784A/K378R (a mistura de reação de transcrição sem polimerase foi usada como um controle negativo), em uma mistura de transcrição contendo molde a ~200 nM, 1X tampão de transcrição (HEPES a
 20 200 mM, DTT a 40 mM, espermidina a 2 mM, Triton X-100 a 0,01%), 2'-OMe ATP, CTP, UTP, e GTP a 1 mM cada, 2'-OH GTP a 30 uM, MgCl₂ a 6,5 mM, MnCl₂ a 2,0 mM, PEG-8000 a 10% p/v, GMP a 1 mM, T7 RNA polimerase mutante Y639F/H784A/K378R ou
 25 T7 RNA polimerase mutante Y639L/H784A/K378R a 200 nM, 5 unidades/mL de pirofosfatase inorgânica, a 37°C, durante a noite. A produção de transcrito para cada condição foi testada através de análise por PAGE-gel, usando 200 uL da mistura de reação, e a produção de transcrito para cada

condição foi quantificada a partir de imageamento com UV da análise por PAGE-gel usando o software ImageQuant versão 5.2 (Molecular Dynamics).

A Figura 4 resume os resultados quantificados da análise por PAGE-gel, mostrando aumento multiplicado da produção com ambas as T7 RNA polimerases mutantes Y639F/H784A/K378R ("FAR") e Y639L/H784A/K378R ("LAR") em relação ao controle negativo sem polimerase. Conforme pode ser visto na Figura 4, foi observado um aperfeiçoamento significativo na produção de transcritos inteiramente contendo 2'-OMe quando a T7 RNA polimerase mutante Y639L/H784A/K378R foi usada para transcrever as novas bibliotecas que incorporam os novos elementos da sequência líder, em comparação com a polimerase mutante Y639F/H784A/K378R. Notavelmente, a ARC2118, a ARC2119 e a ARC2120 deram produções significativamente maiores de transcritos contendo 2'-OMe quando combinadas com a T7 RNA polimerase mutante Y639L/H784A/K378R, em comparação com a T7 RNA polimerase mutante Y639F/H784A/K378R. Um aumento na produção de transcrito por utilização da T7 RNA polimerase mutante Y639L/H784A/K378R foi também visto com a ARC2117, uma biblioteca primeiramente projetada que não tem os elementos da sequência líder recentemente identificados, sabida transcrever sob as condições dadas com a polimerase mutante Y639F/H794A/K378R, que foi usada como um controle. Esses resultados indicam que as produções de transcrito contendo 2'-OMe podem ser aumentadas por utilização da T7 RNA polimerase mutante Y639L/H784A/K378R em comparação com a

T7 RNA polimerase mutante Y639F/H784A/K378R. Além disso, diversas das bibliotecas (ARC2118 e ARC2119) que incorporam os elementos da sequência líder identificados através do método TR-SELEX[®] também deram produções maiores de
5 transcritos contendo 2'-OMe do que a biblioteca de controle, ARC2117, quando usando a T7 RNA polimerase mutante Y639L/H784A/K378R, indicando que pode ser obtido um aperfeiçoamento na produção de transcrito contendo 2'-OMe por utilização da mutante Y639L/H784A/K378R em combinação
10 com as sequências líderes recentemente identificadas particulares da presente invenção.

Exemplo 3: Expressão e Purificação da Polimerase

A T7 RNA polimerase mutante, para uso nos métodos da invenção, pode ser preparada como a seguir. A T7 RNA
15 polimerase (sequência de ácidos nucleicos e de aminoácidos mostrada nas Figuras 5A e 5B, respectivamente, e descrita em Bull, J.J e col., *J. Mol. Evol.*, 57 (3), 241-248 (2003)) pode ser mutada para resultar na mutante LA (Y639L/H784A), na mutante LAR (Y639L/H784A/K378R), na mutante LLA
20 (P266L/Y639L/H784A) ou na mutante LLAR (P266L/Y639L/H784A/K378R). A T7 RNA polimerase pode estar compreendida em um vetor de expressão (um exemplo de um vetor de expressão de T7 RNA polimerase é descrito na Patente U.S. Nº de Série 5.869.320, aqui incorporada por
25 referência em sua totalidade) ou pode ser inserida em um vetor de expressão seguindo a mutagênese. A T7 RNA polimerase mutada pode ser engenheirada para opcionalmente compreender um marcador de His para a facilidade durante a

purificação da proteína.

As seqüências de oligonucleotídeos complementares que contêm a mutação de Leucina para a posição 639 (agtcacgacgctggctCTGgggtccaaagagttcg (SEQ ID NO 107) e
 5 gaactctttggacccCAGagccagcgtcatgact (SEQ ID NO 108)) podem ser sintetizadas. As seqüências de oligonucleotídeos complementares (ggctggcatctctcTgatgttccaaccttgc (SEQ ID NO 109) e gcaaggttggaacatcAgagagatgccagcc (SEQ ID NO 110)) para a mutação P266L podem ser sintetizadas. As seqüências de
 10 oligonucleotídeos complementares (cgctcctaactttgtaGCcagccaagacggtagc (SEQ ID NO 111) e gctaccgtcttggctgGCtacaaagttaggagcg (SEQ ID NO 112)) para a mutação H784A podem ser sintetizadas. As seqüências de oligonucleotídeos complementares
 15 (gctctcaccgcggtggaGacgtgctgccgctgct (SEQ ID NO 113) e agcagcggcagcacgtCtccacgcggtgagagc (SEQ ID NO 114)) para a mutação K378R podem ser sintetizadas. A mutagênese sítio-direcionada pode ser efetuada usando o Kit de Mutagênese Sítio-Direcionada QuikChange® (Stratagene, La Jolla, CA), de
 20 acordo com as instruções do fabricante, para resultar em seqüências de ácidos nucléicos (Figura 6) que codificam polimerases mutantes tendo a combinação de mutações acima indicada. A seqüência de ácidos nucléicos resultante codificando uma polimerase mutante da invenção pode ser
 25 inserida no vetor de expressão desejado, usando técnicas padrões para a expressão e a purificação.

Expressão e Purificação

O vetor de expressão compreendendo a seqüência de

ácidos nucleicos da T7 polimerase mutante é transformado nas células competentes BL21 (DE3) (Stratagene, CA) e incubado sobre gelo por 20 min. O choque térmico é efetuado colocando-se o tubo em 42°C por 2 min. Após colocar o tubo
5 sobre gelo por 1 minuto, 1 ml de caldo L ("LB") é adicionado e incubado em agitador a 37°C por 45 min. 100 µl de líquido da cultura são preparados sobre placas de LB+Amp ágar e incubados a 37°C durante a noite.

Uma única colônia da placa cultivada durante a
10 noite é inoculada em 100 ml de LB-Amp⁺ (150 µg/ml), a 37°C, durante a noite. No segundo dia, dois frascos de 4 litros contendo 2 litros de LB+Amp pré-aquecido são inoculados com 50 ml da cultura noturna e desenvolvidos a 37°C até que a DO600 atinja entre 0,6-0,8. 200 µl de IPTG a 1M são
15 adicionados a cada 2 L de cultura de células com concentração final de 100 µM e desenvolve-se por outras 3 h a 37°C. As células são precipitadas girando a 5000 rpm por 10 min. As células são suspensas novamente em 200 ml de tampão de lise (Tampão de lise: Tris-Cl a 50 mM, pH 8,0,
20 NaCl a 100 mM, 5% de Glicerol, imidazol a 1 mM, betamercaptoetanol ("BME") a 5 mM) e divididas em 6 tubos cônicos de 50 ml. As células são sonicadas em nível de potência 8, 3x30" para cada tubo, e então o fragmento bacteriano é girado a 11.000 rpm por 60 min e o sobrenadante
25 filtrado através de um filtro de 0,22 µM. O imidazol é adicionado ao filtrado até uma concentração final de 10 mM.

O filtrado é carregado para uma coluna de Ni-NTA de 5 ml (GE Healthcare Bio-Sciences, NJ) com uma bomba de

amostra. A coluna é lavada com 10 volumes de coluna (VC) de tampão A (Tampão A: Tris-Cl a 50 mM, pH 8,0, NaCl a 100 mM, 5% de Glicerol, imidazol a 10 mM, BME a 10 mM) contendo imidazol a 20 mM. A coluna é então lavada com 10 VC de
 5 tampão com um gradiente linear de concentração de imidazol a partir de 40 mM até 70 mM em tampão A. A proteína é eluída com 6 VC de Tampão B (Tampão B: Tris-Cl a 50 mM, pH 8,0, NaCl a 100 mM, 5% de Glicerol, imidazol a 250 mM, BME a 10 mM). Após checar as frações de coleta com 5 µl de amostra
 10 em 4-12% de SDS-PAGE, todas as frações de interesse são combinadas e dialisadas (tubulação de diálise: membrana porosa Spectrum Spectra/por Molecular (VWR) MWCO: 12-14000) em 1 L de tampão de diálise (Tampão de diálise: Tris-Cl a 50 mM, pH 7,9, NaCl a 100 mM, 50% de Glicerol, EDTA a 0,1 mM,
 15 0,1% de Triton X-100, BME a 20 mM), durante a noite. O tampão de diálise é trocado após 12 horas e a diálise é efetuada por umas 4 horas adicionais. A concentração de T7 RNA polimerase é medida usando o ensaio de Bradford, como descrito em Bradford, M. M. (1976) *Anal. Biochem.* 72, 248.

20 Exemplo 4: Transcrição que incorpora 100% de 2'-O-metil nucleotídeos

Exemplo 4A: Transcrição com 2'-O-metil sem 2'-OH GTP

Um experimento foi efetuado para testar a
 25 sensibilidade da polimerase mutante Y639L/H784A/K378R à concentração de 2'-OH GTP, por utilização de uma titulação de 2'-OH GTP.

A ARC2118 e a ARC2119, duas bibliotecas que

incorporaram os novos elementos da sequência líder identificados através da seleção por TR-SELEX® (descrita no Exemplo 1), que mostraram altas produções de transcritos quando usadas com a T7 RNA polimerase mutante Y639L/H784A/K378R (ver o Exemplo 2), foram usadas para testar a sensibilidade de transcrição da T7 RNA polimerase mutante Y639L/H784A/K378R à concentração de 2'-OH GTP. As transcrições foram efetuadas usando uma titulação de 2'-OH GTP (0-160 uM) com 1X tampão de transcrição (HEPES a 200 mM, DTT a 40 mM, espermidina a 2 mM, Triton X-100 a 0,01%), molde a ~200 nM, 2'-OMe ATP, CTP, UTP, e GTP a 1 mM cada, MgCl₂ a 6,5 mM, MnCl₂ a 2,0 mM, PEG-8000 p/v a 10%, GMP a 1 mM, 5 unidades/mL de pirofosfatase inorgânica, T7 RNA polimerase mutante Y639L/H784A/K378R a 200 nM, a 37° C, durante a noite.

A produção de transcritos sob cada condição foi testada através de análise por PAGE-gel, usando 200 uL de mistura de reação, e a produção de transcritos para cada condição foi quantificada a partir de imageamento com UV da análise por PAGE-gel, usando o software ImageQuant versão 5.2 (Molecular Dynamics). A Figura 7 resume os resultados quantificados da análise por PAGE-gel, mostrando o aumento multiplicado da produção de transcritos com cada condição em relação à base. Conforme pode ser visto na Figura 7, a ARC2118 e a ARC2119 foram transcritas com a Y639L/H784A/K378R sob todas as condições, incluindo sem 2'-OH GTP, e a produção na ausência de 2'-OH GTP era comparável à produção de transcrição onde a 2'-OH GTP foi incluída na

mistura de reação. Esses resultados indicam que a T7 RNA polimerase mutante Y639L/H784A/K378R não requer a presença de 2'-OH GTP para a produção de transcrito aumentada, em oposição à T7 RNA polimerase mutante Y639F/H784A/K378R, que
5 requer o 2'-OH GTP para a transcrição (dados não mostrados).

Um experimento foi subsequentemente efetuado para determinar as condições de transcrição ótimas a serem usadas com a T7 RNA polimerase mutante Y639L/H784A/K378R, quando combinada com as seqüências líderes identificadas através de
10 seleção por TR-SELEX® (descrita no Exemplo 1). A ARC2119, uma biblioteca que incorpora os novos elementos da seqüência líder, identificados através da seleção por TR-SELEX®, que mostraram produção de transcritos significativamente maior quando usados com a T7 RNA polimerase mutante
15 Y639L/H784A/K378R (ver o Exemplo 2), foi usada para testar o efeito da variação das concentrações de 2'-OMe NTP, magnésio e manganês sobre a produção de transcritos.

As transcrições foram efetuadas usando 1X tampão de transcrição (HEPES a 200 mM, DTT a 40 mM, espermidina a 2
20 mM, Triton X-100 a 0,01%), molde a ~200 nM, 2'-OMe ATP, CTP, UTP, e GTP (0,5 mM, 1 mM, 1,5 mM, e 2 mM cada), MgCl₂ (5 mM, 6,5 mM, 8 mM, e 9,5 mM), MnCl₂ (1,5 mM, 2 mM, 2,5 mM, 3 mM), PEG-8000 p/v a 10%, GMP a 1 mM, 5 unidades/mL de pirofosfatase inorgânica, T7 RNA polimerase mutante
25 Y639L/H784A/K378R a 200 nM, a 37° C, durante a noite.

A produção de transcritos sob cada condição foi testada através de análise por PAGE-gel, usando 200 uL de mistura de reação, e a produção de transcritos para cada

condição foi quantificada a partir de imageamento com UV da análise por PAGE-gel, usando o software ImageQuant versão 5.2 (Molecular Dynamics). A Figura 8 resume os resultados quantificados da análise por PAGE-gel, mostrando o aumento multiplicado da produção de transcritos com cada condição em relação à base. Com base no custo dos 2'-OMe NTPs, e nos resultados desse experimento, 1,5 mM cada de 2'-OMe NTP (e 8 mM de MgCl₂, 2,5 mM de MnCl₂) foi adotado como as condições preferidas para uso com as seqüências líderes e a T7 RNA polimerase mutante Y639L/H784A/K378R da presente invenção.

Exemplo 4B: Fidelidade e Desvio de Previsão da Transcrição de MNA usando a T7 RNA polimerase mutante Y639L/H784A/K378R

Foram efetuados experimentos adicionais para avaliar a fidelidade e o desvio de previsão da transcrição de MNA usando a T7 RNA polimerase mutante Y639L/H784A/K378R e nenhum 2'-OH GTP. Para testar a fidelidade, uma única seqüência clonada, identificada através de seleção por TR-SELEX[®] (descrita no Exemplo 1), foi amplificada por PCR, usada para programar uma transcrição de MNA usando a polimerase Y639L/H784A/K378R e nenhum 2'-OH GTP, purificada por PAGE, o molde de DNA restante foi digerido usando a DNase RQ1 (a ausência de molde de DNA foi então testada por PCR) e o material transcrito foi transcrito invertido (Thermoscript, Invitrogen, Carlsbad, CA) e então amplificado por PCR. Esse produto da PCR foi seqüenciado e a estatística de deleções, inserções e substituições foi então calculada. Das 1300 bases seqüenciadas nesse experimento, não foi

observada nenhuma deleção e inserção, e três substituições foram observadas (ver a Figura 9). Esses números sugerem que a informação de sequência codificada dentro de uma região degenerada de 30 nucleotídeos teria uma chance de 93% de ser transmitida fielmente para a próxima rodada de SELEX®, esse número é tão alto que excede aquele medido para o RNA do tipo selvagem.

Para testar o desvio de previsão de nucleotídeos, a biblioteca ARC2118 foi transcrita sob as seguintes condições: HEPES a 200 mM, DTT a 40 mM, espermidina a 2mM, Triton X-100 a 0,01%, molde a 200 nM, 2'-OMe ATP, CTP, UTP, e GTP a 1 mM cada (nenhum 2'-OH GTP), $MgCl_2$ (6,5 mM), $MnCl_2$ (2 mM), PEG-8000 p/v a 10%, GMP a 1 mM, 5 unidades/mL de pirofosfatase inorgânica, T7 RNA polimerase mutante Y639L/H784A/K378R a 200 nM, a 37° C, durante a noite, purificada por PAGE, o molde de DNA restante foi digerido usando a DNase RQ1 (a ausência de molde de DNA foi então testada por PCR) e o material transcrito foi transcrito invertido e amplificado usando PCR, antes da clonagem e do seqüenciamento. 48 clones da biblioteca amplificada e 48 clones da biblioteca de partida foram seqüenciados. A estatística da ocorrência de nucleotídeos na região degenerada foi examinada para ver se o desvio de previsão ocorreu. Conforme indicado pela Figura 10, a porcentagem de composição de nucleotídeos após a transcrição era muito similar à porcentagem de composição de nucleotídeos da biblioteca de partida, na qual a porcentagem de cada nucleotídeo (A, T, C e G) era aproximadamente igual,

indicando que nenhum desvio de previsão de nucleotídeos ocorre com a T7 RNA polimerase mutante Y639L/H784A/K378R usada para a transcrição.

Exemplo 4C: Comparação da Produção Transcricional

5 com Diversas Sequências Líderes

Moldes 1 a 4:

Para comparar as produções transcricionais usando a T7 RNA polimerase mutante Y639L/H784A/K378R com múltiplas seqüências líderes diferentes, 4 moldes compreendendo razões variadas de purinas para pirimidinas na seqüência líder (posições 1 a 14 nas SEQ ID NOs 126 a 129 abaixo) foram sintetizados com diferentes regiões constantes. Os moldes de DNA foram sintetizados usando um sintetizador de DNA ABI EXPEDITE® (Applied Biosystems, Foster City, CA), e desprotegidos por métodos padrões. As seqüências (mostradas na direção 5' para 3') são como a seguir:

Molde 1

GGAGAATTCCGACCAGAAAGCTTNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN
NNNNNNNNNNNNNCATATGTGCGTCTACATGGATCCTCA (SEQ ID NO 126)

20 Molde 2

GGGAGAGCGGAAGCCGTGCTGGGGCCNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN
NNNNNNNNNNNNNNNNNNCATAACCCAGAGGTCGATGGATC (SEQ ID NO 127)

Molde 3

GGGAGAGACAAGCTTGGGTCNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN
25 NNNNNNNNNNAGAAGAGAAAGAGAAGTTAATTAAGGATCCTCAG (SEQ ID NO 128)

Molde 4

GGGAGAATCCGACCACAAGNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN
NNNNNNNNNNCATATGTGCGTCTACATGGATCCTCA (SEQ ID NO 129).

Os moldes foram amplificados com os seus respectivos primers, conforme indicado abaixo:

Molde 1:

Primer em 5'

5 TAATACGACTCACTATAGGGAGAATTCCGACCAGAAGCTT (SEQ ID NO 130)

Primer em 3' TGAGGATCCATGTAGACGCACATATG (SEQ ID NO 131)

Molde 2:

Primer em 5'

10 TAATACGACTCACTATAGGGAGAGCGGAAGCCGTGCTGGGGCC (SEQ ID NO 149)

Primer em 3' GATCCATCGACCTCTGGGTTATG (SEQ ID NO 132)

Molde 3:

Primer em 5' TAATACGACTCACTATAGGGAGAGACAAGCTTGGGTC

15 (SEQ ID NO 133)

Primer em 3' CTGAGGATCCTTAATTAAGTTCTCTTTCTCTTCT (SEQ ID NO 134)

Molde 4:

Primer em 5' TAATACGACTCACTATAGGGAGAATTCCGACCACAAG

20 (SEQ ID NO 135)

Primer em 3' TGAGGATCCATGTAGACGCACATATG (SEQ ID NO 148)

Os moldes foram usados em uma reação de transcrição in vitro de 15 mL com a T7 RNA polimerase (Y639L/ H784A/ K378R). As transcrições foram feitas usando Hepes a 200 mM, DTT a 40 mM, espermidina a 2 mM, 0,01 % de TritonX-100, 10% de PEG-8000, MgCl₂ a 8 mM, MnCl₂ a 2,5 mM, mCTP a 1,5 mM, mUTP a 1,5 mM, mGTP a 1,5 mM, mATP a 1,5 mM,

GMP a 1 mM, 0,01 unidade/ μ L de pirofosfatase inorgânica, e ~9 μ g/ml de T7 polimerase (Y639L/ H784A/ K378R) e molde de DNA a 0,2 μ M. O RNA foi precipitado e purificado sobre PAGE de desnaturação a 10%. O RNA foi eluído do gel em NaOAc a 300 mM, EDTA a 20 mM, durante a noite, precipitado e quantificado com um Espec. de UV. As produções não diferiram muito entre as quatro seqüências líderes testadas e são mostradas na Tabela 1A abaixo.

Tabela 1A

Agrupamento	Produção (nmols)
Agrupamento 1	18,7
Agrupamento 2	17,4
Agrupamento 3	20,2
Agrupamento 4	27,8

10. Moldes 5 a 8

Como para os Moldes 1 a 4 acima descritos, as produções transcricionais para múltiplas seqüências líderes foram avaliadas. Quatro moldes foram sintetizados com diferentes regiões constantes. Os moldes de DNA foram sintetizados usando um sintetizador de DNA ABI EXPEDITE® (Applied Biosystems, Foster City, CA), e desprotegidos por métodos padrões. As seqüências (mostradas na direção 5' para 3') são como a seguir:

Molde 5:

20 GGCCTTGTAGCGTGCAATTCTTGNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN
NNNCTAACATACTCCGAATCTGTCGAA (SEQ ID NO 138)

Molde 6:

GGAGCCTTCCTCCGGANNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN

NNNNNTCCGGTTCCCGAGCTT (SEQ ID NO 139)

Molde 7:

[illegible]

NNNNNNNNNNNNNTTCGACAGGAGGCTCACAACAGGC (SEQ ID NO 140)

5 Molde 8:

[illegible]

NNGGATCGTTACGACTAGCATCGATG (SEQ ID NO 150)

Os moldes foram amplificados com os seus respectivos primers:

10 Molde 5:

Primer em 5'

TAATACGACTCACTATAGGGCCTTGTAGCGTGCATTCTTG (SEQ ID

Primer em 3'

15 TTCGACAGATTCGGAGTATGTTAG (SEQ ID NO 141)

Molde 6:

Primer em 5'

TAATACGACTCACTATAGGAGCCTTCCTCCGGA (SEQ ID NO 142)

Primer em 3'

20 AAGCTCGGGAAACCGGA (SEQ ID NO 143)

Molde 7:

Primer em 5'

TAATACGACTCACTATAGGGAGACAAGAATAAACGCTCAA (SEQ ID

25 Primer em 3'

GCCTGTTGTGAGCCTCCTGTCGAA (SEQ ID NO 145)

Molde 8:

Primer em 5'

TAATACGACTCACTATAGGGGAGTACAATAACCAGACAT (SEQ ID NO 146)

Primer em 3'

CATCGATGCTAGTCGTAAACGATCC (SEQ ID NO 147)

5 Os moldes foram então usados para uma transcrição in vitro de 0,5 mL com a T7 RNA polimerase (Y639L/ H784A/ K378R). As transcrições foram feitas usando Hepes a 200 mM, DTT a 40 mM, espermidina a 2 mM, 0,01 % de Triton X-100, 10% de PEG-8000, MgCl₂ a 8 mM, MnCl₂ a 2,5 mM, mCTP a 1,5 mM, 10 mUTP a 1,5 mM, mGTP a 1,5 mM, mATP a 1,5 mM, GMP a 1 mM, 0,01 unidade/ μ L de pirofosfatase inorgânica, e ~9 μ g/ml de T7 polimerase (Y639L/ H784A/ K378R) e molde de DNA a 0,2 μ M. O RNA foi precipitado e carregado sobre PAGE de desnaturação a 10%. O RNA foi visualizado por absorbância de UV sobre 15 gel. As produções não diferiram muito entre as quatro seqüências líderes testadas e são mostradas na Tabela 1B abaixo (Produções de transcrição relativas dadas):

Tabela 1B

Agrupamento	Produção (em relação ao Agrupamento 4)
Agrupamento 1	80%
Agrupamento 2	88%
Agrupamento 3	130%
Agrupamento 4	100%

20 Nas modalidades particulares, os moldes identificados acima descritos podem ser usados nos métodos de transcrição e/ou nos métodos de seleção de aptâmeros da invenção.

Exemplo 4D: Transcrição de MNA usando a T7 RNA polimerase mutante P266L/Y639L/H784A/K378R

O molde de DNA e os primers a seguir foram usados para programar uma reação em cadeia por polimerase, para gerar um molde de transcrição de fita dupla. N indica uma posição degenerada com uma probabilidade aproximadamente igual de ser cada uma de ATGC, todas as seqüências são listadas na direção 5' para 3':

Molde de PCR (ARC2118)

10 TAATACGACTCACTATAGGGGAGTACAATAACGTTCTCGNNNNNNNNNNNN
NNNNNNNNNNNNNNNNNNNNNGGATCGTTACGACTAGCATCGATG (SEQ ID NO 3)

primer em 5'

AAAAAAAAAAAAAAAAAAAAAAAAATAATACGACTCACTATAGGGGAG
TACAATAACGTTCTCG (SEQ ID NO 115)

15 primer em 3' CATCGATGCTAGTCGTAACG (SEQ TD NO 116)

O molde de transcrição de fita dupla resultante foi então usado para programar misturas de transcrição de 200 uL para cada amostra, como a seguir: HEPES (200 mM), DTT (40 mM), Espermidina (2 mM), Triton X-100 (0,01%), MgCl₂ (8 mM), MnCl₂ (2,5 mM), PEG-8000 (10% p/v), 1,5 mM cada de 2'-OMe NTP, GMP a 1 mM, molde de transcrição a 100-200 nM, Pirofosfatase Inorgânica (1 unidade), pH 7,5, a T7 polimerase mutante P266L/Y639L/H784A/K378R foi diluída conforme indicado abaixo. A mistura de transcrição foi incubada a 37°C, durante a noite (16 h).

Após incubação, as misturas foram precipitadas com isopropanol, o precipitado resultante foi dissolvido e quantificado usando PAGE de desnaturação (12,5% de

acrilamida) por 60 min, a 25 W. As amostras foram visualizadas e quantificadas através de imagem por UV a 260 nm.

Tabela 2: Produção Transcricional

Enzima	Concentração de Enzima	Produção de Transcrito de MNA Normalizada
K378R/Y639L/H784A	2,1 µg/ml	100
P266L/K378R/Y639L/H784A	11 µg/ml	130
P266L/K378R/Y639L/H784A	2,6 µg/ml	65
P266L/K378R/Y639L/H784A	0,66 µg/ml	13
P266L/K378R/Y639L/H784A	0,16 µg/ml	8,3

5 EXEMPLO 5: Seleção de Aptâmeros usando a T7 RNA polimerase mutante Y639L/H784A/K378R

—Uma seleção foi efetuada para identificar os aptâmeros para Ang2 humana (daqui por diante, "h-Ang2") usando um agrupamento consistindo em 2'-OMe nucleotídeos de purina e pirimidina (daqui por diante, "MNA"). A estratégia de seleção produziu aptâmeros de altas afinidades, específicos para h-Ang2.

A Ang2 humana foi adquirida de R&D Systems, Inc. (Minneapolis, MN). A T7 RNA polimerase (Y639L/H784A/K378R) foi expressa e purificada conforme descrito no Exemplo 3 acima. Os 2'-OMe nucleotídeos de purina e pirimidina foram adquiridos da TriLink BioTechnologies (San Diego, CA).

Seleção do aptâmero para Ang2

Preparação do Agrupamento

20 Um molde de DNA com a sequência 5'-TAATACGACTCACTA

TAGGGGAGTACAATAACGTTCTCGNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNGGATCG
 TTACGACTAGCATCGATG ARC2118 (SEQ ID NO 3) foi sintetizado
 usando um sintetizador de DNA ABI EXPEDITE® (Applied
 Biosystems, Foster City, CA), e desprotegido por métodos
 5 padrões. Os moldes foram amplificados com os primers (5'-
 (GATCGATCGATCGATCGATCTAATACGACTCACTATAGGGGAGTACAAT
 AACGTTCTCG-3') (SEQ ID NO 118) e (5'-
 CATCGATGCTAGTCGTAACGATCC-3') (SEQ ID NO 119) e então usados
 como um molde para a transcrição in vitro com a T7 RNA
 10 polimerase (Y639L/ H784A/ K378R). As transcrições foram
 feitas usando Hepes a 200 mM, DTT a 40 mM, espermidina a 2
 mM, 0,01 % de TritonX-100, 10% de PEG-8000, MgCl₂ a 8 mM,
 MnCl₂ a 2,5 mM, mCTP a 1,5 mM, mUTP a 1,5 mM, mGTP a 1,5 mM,
 mATP a 1,5 mM, GMP a 1 mM, 0,01 unidade/ μ L de pirofosfatase
 15 inorgânica, e ~ 9 μ g/mL de T7 polimerase (Y639L/ H784A/
 K378R) e molde de DNA a 0,5 μ M, para gerar o agrupamento de
 ARC2118 mRmY.

Seleção

A seleção foi iniciada por incubação de 330 pmols
 20 (2×10^{14} moléculas) do agrupamento de MNA ARC 2118 com 100
 pmols de proteína, em um volume final de tampão de seleção
 de 100 μ L (1X PBS da Dulbecco (DPBS)), por 1 h, na
 temperatura ambiente. Os complexos de RNA-proteína e as
 moléculas de RNA não ligadas foram separados usando uma
 25 coluna de rotação de nitrocelulose de 0,45 μ m (Schleicher e Schuell, Keene, NH). A coluna foi pré-tratada
 com KOH (Embeber o filtro da coluna em 1 mL de KOH a 0,5 M,
 TR de 15 min; girar completamente. Embeber o filtro em 1 mL

de dH₂O, TR de 5 min.; girar completamente), lavada 2 x com 1 mL de 1X PBS, e então a solução contendo os complexos de agrupamento:Ang2 foi adicionada à coluna e centrifugada a 1500 x g por 2 minutos. O filtro foi lavado duas vezes com 5 500 µL de DPBS para remover os ligantes não específicos. O RNA foi eluído por adição de 2 x 100 µL de tampão de eluição (uréia a 7 M, acetato de sódio a 100 mM, EDTA a 3 mM, pré-aquecido para 95°C) e então precipitado com etanol. O RNA foi transcrito invertido com o sistema ThermoScript RT-PCR® 10 (Invitrogen, Carlsbad, CA), de acordo com as instruções do fabricante, usando o primer SEQ ID NO 119. O cDNA foi amplificado por PCR com Taq polimerase (New England Biolabs, Beverly, MA), de acordo com as instruções do fabricante, usando a SEQ ID NO 118 e a SEQ ID NO 119. Os moldes foram 15 transcritos conforme descrito acima para a preparação do agrupamento e purificados sobre um gel de poliacrilamida de desnaturação.

A rodada 2 foi efetuada com o mesmo método que a rodada 1. As rodadas 3-12 foram realizadas com h-Ang2 20 imobilizada sobre placas hidrofóbicas. Cada rodada de seleção foi iniciada por imobilização de 20 pmols de h-Ang2 na superfície de uma placa hidrofóbica Maxisorp da Nunc, por 1 hora, na temperatura ambiente, em 100 µL de 1X DPBS. A placa foi lavada 5x com 120 µL de DPBS, então incubada com 25 tampão de bloqueio (1X DPBS, e 0,1 mg/mL de BSA), por 1 hora. O sobrenadante foi então removido e os poços foram lavados 5 vezes com 120 µL de 1X DPBS. O RNA do agrupamento foi incubado por 1 hora na temperatura ambiente nos poços vazios,

então por 1 hora em um poço que tinha sido previamente bloqueado com 100 μ L de tampão de bloqueio. A partir da rodada 3 em diante, os poços com o alvo imobilizado foram bloqueados por 1 hora na temperatura ambiente, em 100 μ L de tampão de bloqueio (1X PBS, 0,1 mg/mL de tRNA, 0,1 mg/mL de ssDNA e 0,1 mg/mL de BSA), antes da etapa de seleção positiva. Em todos os casos, o RNA do agrupamento ligado à h-Ang2 imobilizada foi transcrito invertido diretamente na placa de seleção, pela adição da mistura de transcrição reversa ("TR") (primer em 3', SEQ ID NO 119, e Thermoscript RT, Invitrogen, Carlsbad, CA), seguida por incubação a 65°C por 1 hora. O cDNA resultante foi usado como um molde para a PCR (Taq polimerase, New England Biolabs, Beverly, MA) e a transcrição, conforme descrito para a rodada 1. As condições para cada rodada estão na Tabela 3.

Tabela 3. Resumo das Rodadas

Roda-da	Agrupa-mento (nM)	Platafo-rma	Negati-vo	Tampão	Compe-tidor	Alvo (nM)
1	3300	Filtro com KOH	nenhuma	1X DPBS	nenhuma	1000
2	1000	Filtro com KOH	Filtro com KOH	1X DPBS	nenhuma	1000
3	1000	Placa	Placa; Placa Com BSA	1X DPBS	nenhuma	200
4	1000	Placa	Placa;	1X	nenhuma	200

			Placa com BSA	DPBS		
5	1000	Placa	Placa; Placa com BSA	1X DPBS	nenhuma	200
6	1000	Placa	Placa; Placa com BSA	1X DPBS	nenhuma	200
7	1000	Placa	Placa; Placa com BSA	1X DPBS	nenhuma	200
8	1000	Placa	Placa; Placa Com BSA	1X DPBS	nenhuma	200
9	1000	Placa	Placa; Placa com BSA	1X DPBS	0,1 mg/mL tRNA	200
10	1000	Placa	Placa; Placa Com BSA	1X DPBS	0,1 mg/mL tRNA	200
11	1000	Placa	Placa; Placa Com BSA	1X DPBS	0,1 mg/mL tRNA	200
12	1000	Placa	Placa; Placa Com BSA	1X DPBS	0,1 mg/mL tRNA	200

Análise da Ligação do Aptâmero de MNA

Os ensaios de ligação por dot blot foram efetuados

por todas as seleções, para monitorar a afinidade de ligação dos agrupamentos pelas proteínas. O RNA do agrupamento marcado na extremidade com traços de ^{32}P foi combinado com a h-Ang2 e incubado na temperatura ambiente, por 30 minutos, em tampão de DPBS, em um volume final de 30 μL . A mistura foi aplicada a uma aparelhagem de dot blot (Minifold-1 Dot Blot, Acrílico, Schleicher e Schuell, Keene, NH), montada (do topo até o fundo) com nitrocelulose, náilon, e membranas blot em gel. O RNA que está ligado à proteína é capturado sobre o filtro de nitrocelulose; enquanto que o RNA não ligado à proteína é capturado sobre o filtro de náilon. O enriquecimento para a ligação da Ang2 foi visto começando na rodada 9. Os moldes do agrupamento das rodadas 9, 10 e 12 foram clonados usando o kit de clonagem TOPO TA (Invitrogen, Carlsbad, CA), de acordo com as instruções do fabricante, e 26 clones únicos foram escolhidos para a síntese química, e as constantes de dissociação (K_D) foram determinadas. Resumidamente, os RNAs sintéticos foram marcados na extremidade de 5' com γ - ^{32}P ATP e os valores de K_D foram determinados usando o ensaio por dot blot e as condições de tampão de 1X DPBS (c/ Ca^{2+} e Mg^{2+}) (Gibco, Catálogo nº 14040, Invitrogen, Carlsbad, CA). As K_D s foram estimadas adaptando-se os dados à equação: fração de RNA ligado = amplitude * (((AptConc+[h-Ang2]+ K_D) - SQRT((AptConc+[h-Ang2] + K_D)² - 4(AptConc*[h-Ang2])))/(2*AptConc)) + base. Os resultados são descritos na Tabela 4 abaixo.

Dentro das 26 seqüências únicas, 8 compartilharam um motivo similar e tinham atividades de ligação e

inibidoras similares. Essas seqüências são identificadas como Família I. A Família II compreende 2 seqüências com um motivo compartilhado que tinham atividades de ligação e inibidoras similares.

5 Análise da Função do Aptâmero de MNA

Ensaio Elisa

 Alguns dos aptâmeros foram testados em um ensaio ELISA que foi configurado para medir a sua capacidade de interferir com a ligação da Ang2 ao receptor de Tie2. Para

10 capturar o receptor de Tie2, 150 ng de Tie2-Fc (R&D systems 313-TI-100-CF, Minneapolis, MN) em 100 µL de PBS (pH 7,4) foram colocados sobre uma placa Maxisorb de 96 poços (NUNC nº 446612, Rochester, NY) e incubados durante a noite, a 4 °C. Durante a captura, 50 µL de diversas concentrações de

15 RNA sintético foram misturados com 50 µL de Ang2 a 3,6 nM (200 ng/mL) (R&D systems, 623-AN-025/CF, Minneapolis, MN) (em PBS com 0,2% de BSA), com concentração final de Ang2 a 1,8 nM (100 ng/mL) em PBS com 0,1% de BSA, e incubados na temperatura ambiente por 1 hora. A solução de captura foi

20 removida após uma incubação noturna e a placa foi lavada com 200 µL de TBST (Tris-HCl a 25 mM, pH 7,5, NaCl a 150 mM e 0,01% de Tween 20), três vezes. A placa foi então bloqueada com 200 µL de TBST contendo 5% de leite seco sem gordura, por 30 minutos, na temperatura ambiente. Após o bloqueio, a

25 placa foi lavada com 200 µL de TBST novamente três vezes, na temperatura ambiente, e a mistura de RNA sintético:Ang2 foi adicionada à placa e incubada na temperatura ambiente por 1 hora. A placa foi então lavada com 200 µL de TBST três

vezes, e 100 μ L de anticorpo anti-Ang2 em cabra biotinilado (1:1000, R&D Systems BAF623, Minneapolis, MN) foram adicionados e incubados por 1 hora, na temperatura ambiente. Após três lavagens com 200 μ L de TBST, 100 μ L de Estreptavidina ligada à HRP (1:200; R&D systems nº DY998, Minneapolis, MN) foram adicionados e incubados na temperatura ambiente, por 0,5 hora. Então, a placa foi lavada novamente com 200 μ L de TBST três vezes, e 100 μ L de solução de TMP (Pierce, nº 34028) foram adicionados e incubados no escuro, na temp. ambiente, por 5 minutos. Uma solução de 100 μ L contendo H_2SO_4 a 2 N foi adicionada para interromper a reação e a placa foi lida por SpectroMax a 450 nm. Os resultados são dados na coluna final da Tabela 4 abaixo.

15 Ensaio de FACS

A célula endotelial da veia umbilical humana ("HUVEC") (ATCC) e a célula K293, uma linhagem celular superexpressando o receptor de Tie2 humano, foram usadas para determinar a IC_{50} dos aptâmeros de MNA para Ang2 específicos que inibem a ligação da Ang2 ao receptor de Tie2 sobre a membrana celular. Em resumo, o vetor de expressão de mamífero recombinante pCDNA3.1-Tie2 foi transfectado em 293 células (ATCC, Manassas, VA) e os clones estáveis foram então obtidos após seleção com G418 (Invitrogen, Carlsbad, CA). A citometria de fluxo demonstrou a expressão da proteína Tie2 sobre ambas as células HUVEC e K293. Um ensaio de titulação de Ang2 adicionalmente determinou as quantidades de Ang2 (R&D Systems, Minneapolis, MN) para o

ensaio de inibição do aptâmero nas células HUVEC e K293, que eram 1 e 0,1 µg/mL, respectivamente.

No ensaio de ligação por citometria de fluxo, as células HUVEC e K293 (2×10^5 células/poço) foram precipitadas na placa com 96 poços de fundo em V e foram subsequentemente suspensas novamente e incubadas em soluções de aptâmero de MNA/Ang2 por 2 horas. As soluções de aptâmero/Ang2 foram preparadas por pré-incubação de dosagens diferentes de aptâmeros (100 nM, 33,3 nM, 11,1 nM, 3,7 nM, 1,2 nM, 0,411 nM, 0,137 nM, e 0,0456 nM) com Ang2 em tampão de FACs (1% de BSA, 0,2% de azida sódica em PBS), por 30 min, sobre gelo. Após três lavagens com tampão de FACs, as células foram incubadas, 30 minutos, com anticorpo Anti-Ang2 humana biotinilado (5 µg/mL; R&D Systems, Minneapolis, MN), seguidos por uma outra incubação de 30 minutos com Estreptavidina PE (1:10; BD Biosciences, San Jose, CA). A análise por FACS foi completada usando o FACScan (BD Biosciences, San Jose, CA). Os resultados são descritos na Tabela 4 abaixo.

Tabela 4: Resumo dos resultados de ligação e funcionais para os aptâmeros de MNA anti-Ang2

Aptâmero de MNA	Rodada de Seleção	Família	K _D (nM)	IC ₅₀ (293-Tie2 FACs) (nM)	IC ₅₀ ELISA (nM)
1	10 e 12	I	0,7	0,7	Não testado
2	10 e 12	I	Não	0,5	Não

			testado		testado
3	12	I	0,2	0,5	Não testado
4	10 e 12	I	20,0	1,0	Não testado
5	12	I	34,0	0,7	Não testado
6	10 e 12	I	9,0	0,5	1,0
7	12	I	17,0	1,6	0,3
8	10	II	19,0	Não testado	1,5
9	10	I	120,0	Não testado	Não testado
10	12	I	70,0	Não testado	Não testado
11	12	I	Nenhuma Ligação	Não testado	Não testado
12	12	I	170,0	Não testado	Não testado
13	12	I	82,0	Não testado	Não testado
14	12	I	Nenhuma Ligação	Não testado	Não testado
15	12	I	Nenhuma Ligação	Não testado	Não testado
16	12	I	Nenhuma Ligação	Nenhuma Inibição	Não testado
17	12	I	20,0	Nenhuma	Não

				Inibição	testado
18	12	I	90,0	Não testado	Não testado
19	12	II	25,0	1,1	2,4
20	12	I	Nenhuma Ligação	Nenhuma Inibição	Não testado
21	12	I	2,9	0,5	Não testado
22	12	I	17,0	0,6	Não testado
23	12	I	Nenhuma Ligação	Não testado	Não testado
24	12	I	Nenhuma Ligação	Não testado	Não testado
25	12	I	Nenhuma Ligação	Nenhuma Inibição	Não testado
26	12	I	Nenhuma Ligação	Não testado	Não testado

EXEMPLO 6: Seleção de Aptâmeros usando a T7 RNA polimerase mutante Y639L/H784A/K378R

Uma seleção foi efetuada para identificar os aptâmeros para a IgE humana (daqui por diante, "h-IgE")
5 usando um agrupamento consistindo em 2'-OMe nucleotídeos de purina e pirimidina (daqui por diante, "mRmY"). A estratégia de seleção produziu aptâmeros de altas afinidades, específicos para a h-IgE.

A IgE humana foi adquirida de Athens Research &
10 Technology (Cat. nº 16-16-090705 Athens, GA)). A T7 RNA

polimerase (Y639L/H784A/K378R) foi expressa e purificada conforme descrito no Exemplo 3 acima. Os 2'-OMe nucleotídeos de purina e pirimidina foram adquiridos da TriLink BioTechnologies (San Diego, CA).

5 Seleção do aptâmero para IgE

Preparação do Agrupamento

Um molde de DNA com a sequência 5'-TAATACGACTCACTA TAGGGGAGTACAATAACGTTCTCGNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNGGATCG TTACGACTAGCATCGATG ARC2118 (SEQ ID NO 3) foi sintetizado
10 usando um sintetizador de DNA ABI EXPEDITE® (Applied Biosystems, Foster City, CA), e desprotegido por métodos padrões. Os moldes foram amplificados com os primers (5'-
(GATCGATCGATCGATCGATCTAATACGACTCACTATAGGGGAGTACAAT AACGTTCTCG-3') (SEQ ID NO 118) e (5'-
15 CATCGATGCTAGTCGTAACGATCC-3') (SEQ ID NO 119) e então usados como um molde para a transcrição in vitro com a T7 RNA polimerase (Y639L/ H784A/ K378R). As transcrições foram feitas usando Hepes a 50 mM, DTT a 10 mM, espermidina a 0,5 mM, 0,0025 % de TritonX-100, 10% de PEG-8000, MgCl₂ a 8 mM,
20 MnCl₂ a 2,5 mM, mCTP a 1,5 mM, mUTP a 1,5 mM, mGTP a 1,5 mM, mATP a 1,5 mM, GMP a 1 mM, 0,01 unidade/μL de pirofosfatase inorgânica, e ~ 9 μg/mL de T7 polimerase mutante (Y639L/ H784A/ K378R) e molde de DNA a 0,3 μM, para gerar o agrupamento de ARC2118 MNA.

25 Seleção

A seleção foi iniciada por incubação de 330 pmols (2x10¹⁴ moléculas) do agrupamento de MNA ARC 2118 com 24 pmols de proteína ligada a uma placa hidrofóbica bloqueada

com BSA (placa Maxisorp, Nunc, Rochester, NY), em um volume final de tampão de seleção de 100 μ L (1X PBS da Dulbecco (DPBS)), por 1 h, na temperatura ambiente. O poço foi lavado quatro vezes com 120 μ L de DPBS para remover os ligantes não específicos. O RNA foi eluído e transcrito invertido com o sistema ThermoScript RT-PCR® (Invitrogen, Carlsbad, CA), de acordo com as instruções do fabricante, usando o primer SEQ ID NO 119. O cDNA foi amplificado por PCR com Taq polimerase (New England Biolabs, Beverly, MA), de acordo com as instruções do fabricante, usando a SEQ ID NO 118 e a SEQ ID NO 119. Os moldes foram transcritos conforme descrito acima para a preparação do agrupamento e purificados sobre um gel de poliacrilamida de desnaturação.

Todas as rodadas foram realizadas com h-IgE imobilizada sobre placas hidrofóbicas. Cada rodada de seleção foi iniciada por imobilização de 24 pmols de h-IgE na superfície de uma placa hidrofóbica Maxisorp da Nunc, por 1 hora, na temperatura ambiente, em 100 μ L de 1X DPBS. A placa foi lavada quatro vezes com 120 μ L de DPBS, então incubada com tampão de bloqueio (1X DPBS, e 0,1 mg/mL de BSA), por 1 hora. O sobrenadante foi então removido e os poços foram lavados, quatro vezes, com 120 μ L de 1X DPBS. Começando na Rodada 2, o RNA do agrupamento foi incubado por 1 hora na temperatura ambiente nos poços vazios, então por 1 hora em um poço que tinha sido previamente bloqueado com 100 μ L de tampão de bloqueio. A partir da Rodada 2 em diante, um competidor não específico foi adicionado à etapa de seleção positiva (0,1 mg/mL de tRNA, e 0,1 mg/mL de ssDNA).

Em todos os casos, o RNA do agrupamento ligado à h-IgE imobilizada foi transcrito invertido diretamente na placa de seleção, pela adição da mistura de transcrição reversa ("TR") (primer em 3', SEQ ID NO 119, e Thermoscript RT, Invitrogen, Carlsbad, CA), seguida por incubação a 65°C por 1 hora. O cDNA resultante foi usado como um molde para a PCR (Taq polimerase, New England Biolabs, Beverly, MA) e a transcrição, conforme descrito para a rodada 1. As condições para cada rodada estão na Tabela 5.

10

Tabela 5. Resumo das Rodadas

Roda-da	Agrupamento	Plata-forma	Negativo	Tampão	Compe-tidor	Lava-gens	Alvo
1	3300	Placa	Nenhum	1X DPBS	Nenhum	4 x 120 μL	24 pmols
2	500	Placa	Placa; placa com BSA	1X DPBS	0,1 mg/mL tRNA; 0,1 mg/mL ssDNA	4 x 120 μL	24 pmols
3	1000	Placa	Placa; placa com BSA	1X DPBS	0,1 mg/mL tRNA; 0,1 mg/mL ssDNA	4 x 120 μL	24 pmols
4	1000	Placa	Placa; placa	1X DPBS	0,1 mg/mL	4 x 120 μL	24 pmols

			com BSA		tRNA; 0,1 mg/mL ssDNA		
5	1000	Placa	Placa; placa com BSA	1X DPBS	0,1 mg/mL tRNA; 0,1 mg/mL ssDNA	4 x 120 μ L	24 pmols
6	1000	Placa	Placa; placa com BSA	1X DPBS	0,1 mg/mL tRNA; 0,1 mg/mL ssDNA	4 x 120 μ L	24 pmols
7	1000	Placa	Placa; placa com BSA	1X DPBS	0,1 mg/mL tRNA; 0,1 mg/mL ssDNA	4 x 120 μ L	24 pmols
8	1000	Placa	Placa; placa com BSA	1X DPBS	0,1 mg/mL tRNA; 0,1 mg/mL ssDNA	4 x 120 μ L	24 pmols

9	1000	Placa	Placa; placa com BSA	1X DPBS	0,1 mg/mL tRNA; 0,1 mg/mL ssDNA	4 x 120 μ L	24 pmols
10	1000	Placa	Placa; placa com BSA	1X DPBS	0,1 mg/mL tRNA; 0,1 mg/mL ssDNA	8 x 120 μ L (última lavagem = 15 min.)	24 pmols
11	500	Placa	Placa; placa com BSA	1X DPBS	0,1 mg/mL tRNA; 0,1 mg/mL ssDNA	8 x 120 μ L (última lavagem = 15 min.)	24 pmols
12	500	Placa	Placa; placa com BSA	1X DPBS	0,1 mg/mL tRNA; 0,1 mg/mL ssDNA	8 x 120 μ L (última lavagem = 15 min.)	24 pmols

Análise da Ligação do Aptâmero de MNA

Os ensaios de ligação por dot blot foram efetuados por todas as seleções, para monitorar a afinidade de ligação dos agrupamentos pelas proteínas. O RNA do agrupamento

marcado na extremidade com traços de ^{32}P foi combinado com a h-IgE e incubado na temperatura ambiente, por 30 minutos, em tampão de DPBS, em um volume final de 30 μL . A mistura foi aplicada a uma aparelhagem de dot blot (Minifold-1 Dot Blot, Acrílico, Schleicher e Schuell, Keene, NH), montada (do topo até o fundo) com nitrocelulose, náilon, e membranas blot em gel. O RNA que está ligado à proteína é capturado sobre o filtro de nitrocelulose; enquanto que o RNA não ligado à proteína é capturado sobre o filtro de náilon. O

5 enriquecimento para a ligação da h-IgE foi visto começando na Rodada 8. Os moldes do agrupamento das Rodadas 5, 8 e 12 foram clonados usando o kit de clonagem TOPO TA (Invitrogen, Carlsbad, CA), de acordo com as instruções do fabricante. Os dados de seqüenciamento revelaram que o agrupamento da

10 Rodada 8 tinha se convergido em um único clone principal que compreendia 59% das seqüências totais. Esse clone principal e três mínimos possíveis foram escolhidos para a síntese química, e as constantes de dissociação (K_D) foram determinadas. Resumidamente, os RNAs sintéticos foram

15 marcados na extremidade de 5' com γ - ^{32}P ATP e os valores de K_D foram determinados usando o ensaio por dot blot e as condições de tampão de 1X DPBS (c/ Ca^{2+} e Mg^{2+}) (Gibco, Catálogo nº 14040, Invitrogen, Carlsbad, CA). As K_D s foram estimadas adaptando-se os dados à equação: fração de RNA

20 ligado = amplitude * (((AptConc+[h-IgE]+ K_D) - $\text{SQRT}((\text{AptConc} + [\text{h-IgE}] + K_D)^2 - 4(\text{AptConc} * [\text{h-IgE}])))/(2 * \text{AptConc})) + \text{base}$. O clone principal tinha uma K_D de cerca de 800 pM. O melhor

25 mínimo de ligação foi também testado quanto à ligação à

IgE de macaco (m-IgE), porém não demonstrou ligação reativa cruzada com a proteína IgE de macaco. Essa ausência de reatividade cruzada foi também confirmada por ELISA. Os minimeros com um dT invertido sobre a extremidade de 3' foram usados como a molécula de origem para o processo de química medicinal.

Química Medicinal

A composição química de um dos minimeros de MNA específicos para IgE (Figura 11) foi alterada para aperfeiçoar a afinidade e a potência, ao mesmo tempo mantendo a estabilidade plasmática do composto. O processo incluiu o projeto, a síntese e a avaliação de uma série de derivados do aptâmero para IgE minimizado, onde cada derivado da série compreendia uma única modificação em cada ocorrência de um nucleotídeo predeterminado, para determinar quais resíduos toleraram a substituição. O primeiro grupo de modificações foi a substituição de cada 2'-OMe nucleotídeo único por um desóxi nucleotídeo. Em uma rodada separada de modificação, uma série de derivados foi sintetizada, em que cada derivado compreendia uma única modificação com fosforotioato em uma posição de ligação entre nucleotídeos diferente. Os dados gerados nessas fases iniciais de modificação foram usados para estabelecer uma relação de atividade da estrutura (SAR) para o aptâmero minimizado. Em uma fase subsequente de modificação, os aptâmeros foram sintetizados e testados com grupos compósitos de substituições que foram projetados com base nos dados iniciais da SAR. A partir do painel de

substituições compósitas, identificou-se um aptâmero de 39 nucleotídeos de comprimento com duas substituições de 2'-OMe por 2'-desóxi introduzidas em sua composição. Além disso, identificou-se um aptâmero minimizado modificado resultante, 39 nucleotídeos de comprimento, com uma substituição de 2'-OMe por 2'-desóxi e quatro substituições de fosfato por fosforotioato incorporadas em sua composição. Conforme mostrado na Figura 12, esse aptâmero modificado com desóxi/fosforotioato demonstra afinidade de ligação aumentada, comparado tanto ao aptâmero de origem minimizado, porém não modificado, quanto ao aptâmero de origem minimizado tendo duas substituições de 2'-OMe por desóxi.

Estabilidade em Soro

O aptâmero de origem não modificado e o modificado com desóxi/fosforotioato, minimizados, foram testados para determinar a sua estabilidade em soros humanos, de ratos e de macacos. Cada aptâmero foi adicionado a 1 ml de soro reunido, até uma concentração final de 5 μ M em 90% de soro. Os aptâmeros foram incubados a 37 °C, com agitação, e os pontos de tempo foram coletados a 0, 0,5, 1, 4, 24, 48, 72, e 98 horas. Em cada ponto de tempo, 90 μ l de estoque a partir das amostras incubadas foram adicionados a 10 μ l de EDTA a 0,5 M e congelados a -20 °C, para análise posterior da estabilidade, usando um sistema BIACORE 2000.

Todas as medições de ligação com biossensor foram efetuadas a 25°C, usando um BIACORE 2000 equipado com um chip de biossensor CM5 de grau de pesquisa (BIACORE Inc., Piscataway, NJ). A IgE humana recombinante purificada

(Athens Research & Technology, Athens, GA) foi imobilizada à superfície do biossensor usando a química de amino-acoplamento. Para obter isso, as superfícies de duas células em fluxo foram primeiramente ativadas por 7 min com uma mistura a 1 : 1 de NHS (*N*-hidroxissuccinimida) a 0,1 M e EDC (3-(*N,N*-dimetilamina) propil-*N*-etilcarbodiimida) a 0,4 M, em uma vazão de 5 µl/min. Após a ativação das superfícies, uma célula em fluxo foi injetada com 50 µg/ml de IgE, a 10 µl/min, por 20 min, para permitir o estabelecimento de ligações covalentes à superfície ativada. A seguir, o cloridrato de etanolamina a 1 M, pH 8,5, foi injetado por 7 min, a 5 µl/min, para inativar os ésteres residuais. Para a célula em fluxo usada como branco, o cloridrato de etanolamina a 1 M, pH 8,5, foi injetado por 7 min, para inativar os ésteres residuais, sem injeção de proteína.

Um grupo de padrões de aptâmeros foi corrido através do chip preparado, para gerar uma curva padrão antes de todos os pontos de tempo serem analisados. Para estabelecer uma curva padrão, os aptâmeros foram diluídos em série (a partir de 200 nM até 12,5 nM) em tampão de HBS-P (HEPES a 10 mM, pH 7,4, NaCl a 150 mM, 0,005% de Tensoativo 20) suplementado com 4% de soro humano e EDTA a 50 mM. Todas as amostras diluídas foram injetadas no Biacore 2000 para a ligação, a 20 µl/min, por 5 min, e espera por 3 minutos. Para regenerar o chip, o NaCl a 1N foi injetado por 60 segundos, a 30 µl/min. A resposta de pico RU no final da fase de ligação foi plotada contra a concentração de aptâmero e uma curva padrão foi gerada usando uma função

logística de Quatro Parâmetros. Para medir a concentração de aptâmero ativo nos soros humanos, de ratos, e de macacos, as amostras nos pontos de tempos foram diluídas 22,5 vezes em HBS-P, para tornar a concentração final em soro a 4%,
5 imediatamente antes da injeção no Biacore 2000. As concentrações de aptâmeros funcionais, em cada período de incubação em soro, foram calculadas por conversão da unidade de resposta RU em concentração, usando a curva padrão gerada acima. Como uma medição de controle de qualidade adicional,
10 dois padrões de aptâmeros foram independentemente testados no final do experimento, para certificar-se que as concentrações medidas no BIACORE estavam menos do que 20% desviadas dos padrões. O aptâmero de origem não modificado e o modificado com desóxi/fosforotioato, minimizados, foram
15 ambos determinados serem mais do que 90% ativos, em 98 horas, nos soros humanos, de ratos, e de macacos.

A invenção, tendo agora sido descrita como forma de descrição escrita e exemplo, aqueles de habilidade na técnica reconhecerão que a invenção pode ser praticada em
20 uma variedade de modalidades e que a descrição e os exemplos acima são para propósitos de ilustração, e não limitação, das reivindicações a seguir.

LISTAGEM DE SEQUÊNCIAS

<110> Archemix Corp., et col.

<120> "MATERIAIS E MÉTODOS PARA A GERAÇÃO DE TRANSCRITOS DE ÁCIDOS
NUCLÉICOS INTEIRAMENTE MODIFICADOS EM 2'"

<130> 23239-587-061

<150> 60/696.292

<151> 30-06-2005

<160> 151

<170> PatentIn versão 3.2

<210> 1

<211> 883

<212> PRT

<213> Sequência Artificial

<220>

<223> T7 polimerase modificada, quimicamente sintetizada Y639/H784A

<400> 1

Met Asn Thr Ile Asn Ile Ala Lys Asn Asp Phe Ser Asp Ile Glu Leu

1

5

10

15

Ala Ala Ile Pro Phe Asn Thr Leu Ala Asp His Tyr Gly Glu Arg Leu

20

25

30

Ala Arg Glu Gln Leu Ala Leu Glu His Glu Ser Tyr Glu Met Gly Glu
 35 40 45
 Ala Arg Phe Arg Lys Met Phe Glu Arg Gln Leu Lys Ala Gly Glu Val
 50 55 60
 Ala Asp Asn Ala Ala Ala Lys Pro Leu Ile Thr Thr Leu Leu Pro Lys
 65 70 75 80
 Met Ile Ala Arg Ile Asn Asp Trp Phe Glu Glu Val Lys Ala Lys Arg
 85 90 95
 Gly Lys Arg Pro Thr Ala Phe Gln Phe Leu Gln Glu Ile Lys Pro Glu
 100 105 110
 Ala Val Ala Tyr Ile Thr Ile Lys Thr Thr Leu Ala Cys Leu Thr Ser
 115 120 125
 Ala Asp Asn Thr Thr Val Gln Ala Val Ala Ser Ala Ile Gly Arg Ala
 130 135 140
 Ile Glu Asp Glu Ala Arg Phe Gly Arg Ile Arg Asp Leu Glu Ala Lys
 145 150 155 160
 His Phe Lys Lys Asn Val Glu Glu Gln Leu Asn Lys Arg Val Gly His
 165 170 175
 Val Tyr Lys Lys Ala Phe Met Gln Val Val Glu Ala Asp Met Leu Ser
 180 185 190
 Lys Gly Leu Leu Gly Gly Glu Ala Trp Ser Ser Trp His Lys Glu Asp
 195 200 205
 Ser Ile His Val Gly Val Arg Cys Ile Glu Met Leu Ile Glu Ser Thr
 210 215 220
 Gly Met Val Ser Leu His Arg Gln Asn Ala Gly Val Val Gly Gln Asp
 225 230 235 240
 Ser Glu Thr Ile Glu Leu Ala Pro Glu Tyr Ala Glu Ala Ile Ala Thr
 245 250 255

Arg	Ala	Gly	Ala	Leu	Ala	Gly	Ile	Ser	Pro	Met	Phe	Gln	Pro	Cys	Val
260				265				270							
Val	Pro	Pro	Lys	Pro	Trp	Thr	Gly	Ile	Thr	Gly	Gly	Gly	Tyr	Trp	Ala
275				280				285							
Asn	Gly	Arg	Arg	Pro	Leu	Ala	Leu	Val	Arg	Thr	His	Ser	Lys	Lys	Ala
290				295				300							
Leu	Met	Arg	Tyr	Glu	Asp	Val	Tyr	Met	Pro	Glu	Val	Tyr	Lys	Ala	Ile
305				310				315				320			
Asn	Ile	Ala	Gln	Asn	Thr	Ala	Trp	Lys	Ile	Asn	Lys	Lys	Val	Leu	Ala
				325				330				335			
Val	Ala	Asn	Val	Ile	Thr	Lys	Trp	Lys	His	Cys	Pro	Val	Glu	Asp	Ile
340				345				350							
Pro	Ala	Ile	Glu	Arg	Glu	Glu	Leu	Pro	Met	Lys	Pro	Glu	Asp	Ile	Asp
355				360				365							
Met	Asn	Pro	Glu	Ala	Leu	Thr	Ala	Trp	Lys	Arg	Ala	Ala	Ala	Ala	Val
370				375				380							
Tyr	Arg	Lys	Asp	Lys	Ala	Arg	Lys	Ser	Arg	Arg	Ile	Ser	Leu	Glu	Phe
385				390				395				400			
Met	Leu	Glu	Gln	Ala	Asn	Lys	Phe	Ala	Asn	His	Lys	Ala	Ile	Trp	Phe
				405				410				415			
Pro	Tyr	Asn	Met	Asp	Trp	Arg	Gly	Arg	Val	Tyr	Ala	Val	Ser	Met	Phe
420				425				430							
Asn	Pro	Gln	Gly	Asn	Asp	Met	Thr	Lys	Gly	Leu	Leu	Thr	Leu	Ala	Lys
435				440				445							
Gly	Lys	Pro	Ile	Gly	Lys	Glu	Gly	Tyr	Tyr	Trp	Leu	Lys	Ile	His	Gly
450				455				460							
Ala	Asn	Cys	Ala	Gly	Val	Asp	Lys	Val	Pro	Phe	Pro	Glu	Arg	Ile	Lys
465				470				475				480			

Phe Ile Glu Glu Asn His Glu Asn Ile Met Ala Cys Ala Lys Ser Pro			
	485	490	495
Leu Glu Asn Thr Trp Trp Ala Glu Gln Asp Ser Pro Phe Cys Phe Leu			
	500	505	510
Ala Phe Cys Phe Glu Tyr Ala Gly Val Gln His His Gly Leu Ser Tyr			
	515	520	525
Asn Cys Ser Leu Pro Leu Ala Phe Asp Gly Ser Cys Ser Gly Ile Gln			
	530	535	540
His Phe Ser Ala Met Leu Arg Asp Glu Val Gly Gly Arg Ala Val Asn			
	545	550	555
Leu Leu Pro Ser Glu Thr Val Gln Asp Ile Tyr Gly Ile Val Ala Lys			
	565	570	575
Lys Val Asn Glu Ile Leu Gln Ala Asp Ala Ile Asn Gly Thr Asp Asn			
	580	585	590
Glu Val Val Thr Val Thr Asp Glu Asn Thr Gly Glu Ile Ser Glu Lys			
	595	600	605
Val Lys Leu Gly Thr Lys Ala Leu Ala Gly Gln Trp Leu Ala Tyr Gly			
	610	615	620
Val Thr Arg Ser Val Thr Lys Arg Ser Val Met Thr Leu Ala Leu Gly			
	625	630	635
Ser Lys Glu Phe Gly Phe Arg Gln Gln Val Leu Glu Asp Thr Ile Gln			
	645	650	655
Pro Ala Ile Asp Ser Gly Lys Gly Leu Met Phe Thr Gln Pro Asn Gln			
	660	665	670
Ala Ala Gly Tyr Met Ala Lys Leu Ile Trp Glu Ser Val Ser Val Thr			
	675	680	685
Val Val Ala Ala Val Glu Ala Met Asn Trp Leu Lys Ser Ala Ala Lys			
	690	695	700

Leu Leu Ala Ala Glu Val Lys Asp Lys Lys Thr Gly Glu Ile Leu Arg
 705 710 715 720
 Lys Arg Cys Ala Val His Trp Val Thr Pro Asp Gly Phe Pro Val Trp
 725 730 735
 Gln Glu Tyr Lys Lys Pro Ile Gln Thr Arg Leu Asn Leu Met Phe Leu
 740 745 750
 Gly Gln Phe Arg Leu Gln Pro Thr Ile Asn Thr Asn Lys Asp Ser Glu
 755 760 765
 Ile Asp Ala His Lys Gln Glu Ser Gly Ile Ala Pro Asn Phe Val Ala
 770 775 780
 Ser Gln Asp Gly Ser His Leu Arg Lys Thr Val Val Trp Ala His Glu
 785 790 795 800
 Lys Tyr Gly Ile Glu Ser Phe Ala Leu Ile His Asp Ser Phe Gly Thr
 805 810 815
 Ile Pro Ala Asp Ala Ala Asn Leu Phe Lys Ala Val Arg Glu Thr Met
 820 825 830
 Val Asp Thr Tyr Glu Ser Cys Asp Val Leu Ala Asp Phe Tyr Asp Gln
 835 840 845
 Phe Ala Asp Gln Leu His Glu Ser Gln Leu Asp Lys Met Pro Ala Leu
 850 855 860
 Pro Ala Lys Gly Asn Leu Asn Leu Arg Asp Ile Leu Glu Ser Asp Phe
 865 870 875 880
 Ala Phe Ala

<210> 2

<211> 883

<212> PRT

<213> Sequência Artificial

<220>

<223> T7 polimerase modificada, quimicamente sintetizada

Y639L/H784A/K378R

<400> 2

Met Asn Thr Ile Asn Ile Ala Lys Asn Asp Phe Ser Asp Ile Glu Leu

1 5 10 15

Ala Ala Ile Pro Phe Asn Thr Leu Ala Asp His Tyr Gly Glu Arg Leu

20 25 30

Ala Arg Glu Gln Leu Ala Leu Glu His Glu Ser Tyr Glu Met Gly Glu

35 40 45

Ala Arg Phe Arg Lys Met Phe Glu Arg Gln Leu Lys Ala Gly Glu Val

50 55 60

Ala Asp Asn Ala Ala Ala Lys Pro Leu Ile Thr Thr Leu Leu Pro Lys

65 70 75 80

Met Ile Ala Arg Ile Asn Asp Trp Phe Glu Glu Val Lys Ala Lys Arg

85 90 95

Gly Lys Arg Pro Thr Ala Phe Gln Phe Leu Gln Glu Ile Lys Pro Glu

100 105 110

Ala Val Ala Tyr Ile Thr Ile Lys Thr Thr Leu Ala Cys Leu Thr Ser

115 120 125

Ala Asp Asn Thr Thr Val Gln Ala Val Ala Ser Ala Ile Gly Arg Ala

130 135 140

Ile Glu Asp Glu Ala Arg Phe Gly Arg Ile Arg Asp Leu Glu Ala Lys

145 150 155 160

His Phe Lys Lys Asn Val Glu Glu Gln Leu Asn Lys Arg Val Gly His

165 170 175

Val Tyr Lys Lys Ala Phe Met Gln Val Val Glu Ala Asp Met Leu Ser

180

185

190

Lys Gly Leu Leu Gly Gly Glu Ala Trp Ser Ser Trp His Lys Glu Asp

195

200

205

Ser Ile His Val Gly Val Arg Cys Ile Glu Met Leu Ile Glu Ser Thr

210

215

220

Gly Met Val Ser Leu His Arg Gln Asn Ala Gly Val Val Gly Gln Asp

225

230

235

240

Ser Glu Thr Ile Glu Leu Ala Pro Glu Tyr Ala Glu Ala Ile Ala Thr

245

250

255

Arg Ala Gly Ala Leu Ala Gly Ile Ser Pro Met Phe Gln Pro Cys Val

260

265

270

Val Pro Pro Lys Pro Trp Thr Gly Ile Thr Gly Gly Gly Tyr Trp Ala

275

280

285

Asn Gly Arg Arg Pro Leu Ala Leu Val Arg Thr His Ser Lys Lys Ala

290

295

300

Leu Met Arg Tyr Glu Asp Val Tyr Met Pro Glu Val Tyr Lys Ala Ile

305

310

315

320

Asn Ile Ala Gln Asn Thr Ala Trp Lys Ile Asn Lys Lys Val Leu Ala

325

330

335

Val Ala Asn Val Ile Thr Lys Trp Lys His Cys Pro Val Glu Asp Ile

340

345

350

Pro Ala Ile Glu Arg Glu Glu Leu Pro Met Lys Pro Glu Asp Ile Asp

355

360

365

Met Asn Pro Glu Ala Leu Thr Ala Trp Arg Arg Ala Ala Ala Val

370

375

380

Tyr Arg Lys Asp Lys Ala Arg Lys Ser Arg Arg Ile Ser Leu Glu Phe

385		390		395		400									
Met	Leu	Glu	Gln	Ala	Asn	Lys	Phe	Ala	Asn	His	Lys	Ala	Ile	Trp	Phe
		405				410								415	
Pro	Tyr	Asn	Met	Asp	Trp	Arg	Gly	Arg	Val	Tyr	Ala	Val	Ser	Met	Phe
		420				425								430	
Asn	Pro	Gln	Gly	Asn	Asp	Met	Thr	Lys	Gly	Leu	Leu	Thr	Leu	Ala	Lys
		435						440						445	
Gly	Lys	Pro	Ile	Gly	Lys	Glu	Gly	Tyr	Tyr	Trp	Leu	Lys	Ile	His	Gly
		450				455								460	
Ala	Asn	Cys	Ala	Gly	Val	Asp	Lys	Val	Pro	Phe	Pro	Glu	Arg	Ile	Lys
465					470					475				480	
Phe	Ile	Glu	Glu	Asn	His	Glu	Asn	Ile	Met	Ala	Cys	Ala	Lys	Ser	Pro
				485				490						495	
Leu	Glu	Asn	Thr	Trp	Trp	Ala	Glu	Gln	Asp	Ser	Pro	Phe	Cys	Phe	Leu
		500						505						510	
Ala	Phe	Cys	Phe	Glu	Tyr	Ala	Gly	Val	Gln	His	His	Gly	Leu	Ser	Tyr
		515						520						525	
Asn	Cys	Ser	Leu	Pro	Leu	Ala	Phe	Asp	Gly	Ser	Cys	Ser	Gly	Ile	Gln
		530						535						540	
His	Phe	Ser	Ala	Met	Leu	Arg	Asp	Glu	Val	Gly	Gly	Arg	Ala	Val	Asn
545					550					555				560	
Leu	Leu	Pro	Ser	Glu	Thr	Val	Gln	Asp	Ile	Tyr	Gly	Ile	Val	Ala	Lys
				565						570				575	
Lys	Val	Asn	Glu	Ile	Leu	Gln	Ala	Asp	Ala	Ile	Asn	Gly	Thr	Asp	Asn
		580						585						590	
Glu	Val	Val	Thr	Val	Thr	Asp	Glu	Asn	Thr	Gly	Glu	Ile	Ser	Glu	Lys
		595						600						605	
Val	Lys	Leu	Gly	Thr	Lys	Ala	Leu	Ala	Gly	Gln	Trp	Leu	Ala	Tyr	Gly

610	615	620	
Val Thr Arg Ser Val Thr Lys Arg Ser Val Met Thr Leu Ala Leu Gly			
625	630	635	640
Ser Lys Glu Phe Gly Phe Arg Gln Gln Val Leu Glu Asp Thr Ile Gln			
	645	650	655
Pro Ala Ile Asp Ser Gly Lys Gly Leu Met Phe Thr Gln Pro Asn Gln			
	660	665	670
Ala Ala Gly Tyr Met Ala Lys Leu Ile Trp Glu Ser Val Ser Val Thr			
	675	680	685
Val Val Ala Ala Val Glu Ala Met Asn Trp Leu Lys Ser Ala Ala Lys			
	690	695	700
Leu Leu Ala Ala Glu Val Lys Asp Lys Lys Thr Gly Glu Ile Leu Arg			
705	710	715	720
Lys Arg Cys Ala Val His Trp Val Thr Pro Asp Gly Phe Pro Val Trp			
	725	730	735
Gln Glu Tyr Lys Lys Pro Ile Gln Thr Arg Leu Asn Leu Met Phe Leu			
	740	745	750
Gly Gln Phe Arg Leu Gln Pro Thr Ile Asn Thr Asn Lys Asp Ser Glu			
	755	760	765
Ile Asp Ala His Lys Gln Glu Ser Gly Ile Ala Pro Asn Phe Val Ala			
	770	775	780
Ser Gln Asp Gly Ser His Leu Arg Lys Thr Val Val Trp Ala His Glu			
785	790	795	800
Lys Tyr Gly Ile Glu Ser Phe Ala Leu Ile His Asp Ser Phe Gly Thr			
	805	810	815
Ile Pro Ala Asp Ala Ala Asn Leu Phe Lys Ala Val Arg Glu Thr Met			
	820	825	830
Val Asp Thr Tyr Glu Ser Cys Asp Val Leu Ala Asp Phe Tyr Asp Gln			

835 840 845
 Phe Ala Asp Gln Leu His Glu Ser Gln Leu Asp Lys Met Pro Ala Leu
 850 855 860
 Pro Ala Lys Gly Asn Leu Asn Leu Arg Asp Ile Leu Glu Ser Asp Phe
 865 870 875 880
 Ala Phe Ala

<210> 3

<211> 93

<212> DNA

<213> Sequência Artificial

<220>

<223> molde de transcrição de DNA, quimicamente sintetizado, ARC2118

<220>

<221> característica diversa

<222> (40)..(69)

<223> n é a, t, c ou g

<400> 3

taatacgact cactataggg gagtacaata acgttctcgn nnnnnnnnnn nnnnnnnnnn 60
 nnnnnnnnnng gatcggttacg actagcatcg atg 93

<210> 4

<211> 93

<212> DNA

<213> Sequência Artificial

<220>

<223> molde de transcrição de DNA, quimicamente sintetizado, ARC2119

<220>

<221> característica diversa

<222> (40)..(69)

<223> n é a, t, c ou g

<400> 4

taatacgact cactataggg ggtgatattg acgttctcgn nnnnnnnnnn nnnnnnnnnn 60

nnnnnnnnng gatcggttacg actagcatcg atg 93

<210> 5

<211> 93

<212> DNA

<213> Sequência Artificial

<220>

<223> molde de transcrição de DNA, quimicamente sintetizado, ARC2120

<220>

<221> característica diversa

<222> (40)..(69)

<223> n é a, t, c ou g

<400> 5

taatacgact cactataggg gtgcgcggtt acgttctcgn nnnnnnnnnn nnnnnnnnnn 60

nnnnnnnnng gatcggttacg actagcatcg atg

93

<210> 6

<211> 93

<212> DNA

<213> Sequência Artificial

<220>

<223> molde de transcrição de DNA, quimicamente sintetizado, ARC2121

<220>

<221> característica diversa

<222> (40)..(69)

<223> n é a, t, c ou g

<400> 6

taatacgact cactataggg ggagggggtg cggttctcgn nnnnnnnnnn nnnnnnnnnn 60

nnnnnnnnng gatcggttacg actagcatcg atg 93

<210> 7

<211> 81

<212> DNA

<213> Sequência Artificial

<220>

<223> molde de transcrição de DNA, quimicamente sintetizado, ARC1140

<220>

<221> característica diversa

<222> (22)..(41)

<223> n é a, t, c ou g

<400> 7

taatacgact cactataggg gnnnnnnnnn nnnnnnnnnn nacgtaaccg gttaaaccgc 60

ggtcgatgca gtaagctagc t 81

<210> 8

<211> 17

<212> DNA

<213> Sequência Artificial

<220>

<223> oligonucleotídeo de promotor de T7 fosforilado em 5', quimicamente
sintetizado

<220>

<221> base modificada

<222> (1)..(1)

<223> onde a timidina na posição 1 está fosforilada em 5'

<400> 8

tatagtgagt cgtatta 17

<210> 9

<211> 21

<212> DNA

<213> Sequência Artificial

<220>

<223> esplinte de ligação quimicamente sintetizado

<400> 9

taatacgact cactataggg g

21

<210> 10

<211> 82

<212> DNA

<213> Sequência Artificial

<220>

<223> sequência quimicamente sintetizada

<400> 10

taatacgact cactataggg ggtggggcca atggcgggat atacgtaacc ggttataccc 60

gggtcgatgc agtaagctag ct 82

<210> 11

<211> 80

<212> DNA

<213> Sequência Artificial

<220>

<223> sequência quimicamente sintetizada

<400> 11

taatacgact cactataggg gatgtacata tgtattcgtg acgtgaccgg ttaaaccogg 60

gtc gat gcag taagctagct

80

<210> 12

<211> 82

<212> DNA

<213> Sequência Artificial

<220>

<223> sequência quimicamente sintetizada

<400> 12

taatac gact cactataggg ggagcgggga gacgtagtca tcacgtagcc ggttaaacc 60

gggtc gatgc agtaagctag ct 82

<210> 13

<211> 79

<212> DNA

<213> Sequência Artificial

<220>

<223> sequência quimicamente sintetizada

<220>

<221> característica diversa

<222> (7)..(7)

<223> n é a, c, g, ou t

<400> 13

taatacnact cactataggg ggtgggggtg gtggtgataa cgtaaccggt taaacccggg 60
tcgatgcagt aagctagct 79

<210> 14

<211> 83

<212> DNA

<213> Sequência Artificial

<220>

<223> sequência quimicamente sintetizada

<400> 14

taatacgact cactataggg ggtgtgcacc agatatgcct tgaacgtaac ccgttaaacc 60
cgggtcgaatg cagtaagcta gct 83

<210> 15

<211> 82

<212> DNA

<213> Sequência Artificial

<220>

<223> sequência quimicamente sintetizada

<400> 15

taatacgact cactataggg ggtagggggc acgcactaac caacgtaacc ggttaaacc 60
gggtcgaatg agtaagctag ct 82

<210> 16

<211> 43

<212> DNA

<213> Sequência Artificial

<220>

<223> sequência quimicamente sintetizada

<220>

<221> característica diversa

<222> (37)..(37)

<223> n é a, c, g, ou t

<400> 16

taatacact cactataggg ggagggggtg ctgaccncaa aca

43

<210> 17

<211> 81

<212> DNA

<213> Sequência Artificial

<220>

<223> sequência quimicamente sintetizada

<400> 17

taatacgact cactataggg gtggggctcg gatgagacaa tacgtaaccg gttaaaccg 60

ggtcgatgca gtaagctagc t 81

<210> 18

<211> 83

<212> DNA

<213> Sequência Artificial

<220>

<223> sequência quimicamente sintetizada

<400> 18

taatacgact cactataggg gggggtgggt aggcgagcac tccacgtaac cagttaaacc 60

cgggtcgatg cagtaagcta gct 83

<210> 19

<211> 82

<212> DNA

<213> Sequência Artificial

<220>

<223> sequência quimicamente sintetizada

<400> 19

taatacgact cactataggg gggaaggacg agcagacgag caacgtaacc tgttaaacc 60

gggtcgatgc agtaagctag ct 82

<210> 20

<211> 85

<212> DNA

<213> Sequência Artificial

<220>

<223> sequência quimicamente sintetizada

<400> 20

taatacgact cactataggg gggggcggtt agagtgtgtaag taccgacgta accgggttaa 60

cccgggtcga tgcagtaagc tagct 85

<210> 21

<211> 79

<212> DNA

<213> Sequência Artificial

<220>

<223> sequência quimicamente sintetizada

<400> 21

taatacgact cactataggg gggttgctgt tagtaacgcc acgtaaccgg ttaaacttgg 60

tcgatgcagt aagctagct 79

<210> 22

<211> 84

<212> DNA

<213> Sequência Artificial

<220>

<223> sequência quimicamente sintetizada

<400> 22

taatacgact cactataggg gggcgggaga atgttatata gttacggtaa ccggttaa 60

ccgggtcgat gcagtaagct agct

84

<210> 23

<211> 81

<212> DNA

<213> Sequência Artificial

<220>

<223> sequência quimicamente sintetizada

<400> 23

taatacgact cactataggg gaaaggggcg gtatgttaca cacgtaacag gttaaaccgg 60

ggtcgatgca gtaagctagc t 81

<210> 24

<211> 80

<212> DNA

<213> Sequência Artificial

<220>

<223> sequência quimicamente sintetizada

<400> 24

taatacgact cactataggg gggacgtgtt agcattccag aattcgtaac ctaaaccgg 60

gtcgacgcag taagctagct 80

<210> 25

<211> 81

<212> DNA

<213> Sequência Artificial

<220>

<223> sequência quimicamente sintetizada

<400> 25

taatacgact cactataggg ggcgtgggag ataggttcaa ggacgtaccg gttatacccg 60

ggtcgatgca gtaagctagc t 81

<210> 26

<211> 83

<212> DNA

<213> Sequência Artificial

<220>

<223> sequência quimicamente sintetizada

<400> 26

taatacgact cactataggg gggctccgtg ctatcgtcgg ataacgtaac ccgttaaacc 60

cgggctgatg cagtaagcta gct 83

<210> 27

<211> 83

<212> DNA

<213> Sequência Artificial

<220>

<223> sequência quimicamente sintetizada

<400> 27

taatacgact cactataggg ggggagaagg tcttaaggtc gccaacgtaa ctgttaaacc 60

cgggtcgatg cagtaagcta gct 83

<210> 28

<211> 83

<212> DNA

<213> Sequência Artificial

<220>

<223> sequência quimicamente sintetizada

<400> 28

taatacgact cactataggg ggggcatacg agtttaggtg gagacgtaac cggttaaacc 60

cgggtcgatg cagtaagcta gct 83

<210> 29

<211> 82

<212> DNA

<213> Sequência Artificial

<220>

<223> sequência quimicamente sintetizada

<400> 29

taatacgact cactataggg ggatgatgac ttccgcgta atacgttacc ggttaaacc 60

gggtcgatgc agtaagctag ct

82

<210> 30

<211> 75

<212> DNA

<213> Sequência Artificial

<220>

<223> sequência quimicamente sintetizada

<400> 30

taatacgact cactataggg gtgggacgcc gtctgagtat aacgtacccg gtcgggtcga 60

tgcaagtaagc tagct 75

<210> 31

<211> 87

<212> DNA

<213> Sequência Artificial

<220>

<223> sequência quimicamente sintetizada

<400> 31

taatacgact cactataggg ggggggggac gtaatcggct atcggttcacg taaccgggta 60

aacccggggtc gatgcagtaa agggcga 87

<210> 32

<211> 84

<212> DNA

<213> Sequência Artificial

<220>

<223> sequência quimicamente sintetizada

<400> 32

taatacgact cactataggg tgggacgggc agcgtggatg taggacgtaa ccggttaaac 60

gcgggtcgat gcagtaagct agct 84

<210> 33

<211> 83

<212> DNA

<213> Sequência Artificial

<220>

<223> sequência quimicamente sintetizada

<400> 33

taatacgact cactataggg gggtttgtct gaagtgaagc agaacgtaac ccggttaatcc 60

cgggtcgatg cagtaagcta gct 83

<210> 34

<211> 84

<212> DNA

<213> Sequência Artificial

<220>

<223> sequência quimicamente sintetizada

<400> 34

taatacgact cactataggg ggggagggca catcatcgta tcaaacgtaa ccagttaatc 60
ccgggtcgat cagtaagct agct 84

<210> 35

<211> 81

<212> DNA

<213> Sequência Artificial

<220>

<223> sequência quimicamente sintetizada

<400> 35

taatacgact cactataggg gaggctagag gacgcgacag aacgtaaccg gttaaaccg 60
ggtcgatgca gtaagctagc t 81

<210> 36

<211> 81

<212> DNA

<213> Sequência Artificial

<220>

<223> sequência quimicamente sintetizada

<400> 36

taatacgact cactataggg ggcgatcgcg aagggtttc aacgtaaccg gttaaaccg 60
ggtcgatgca gtaagctagc t 81

<210> 37

<211> 83

<212> DNA

<213> Sequência Artificial

<220>

<223> sequência quimicamente sintetizada

<400> 37

taatacgact cactatagg gggtagggaa agattacggg gctacgtaac cggttatacc 60

tgggtcgatg cagtaagcta gct 83

<210> 38

<211> 75

<212> DNA

<213> Sequência Artificial

<220>

<223> sequência quimicamente sintetizada

<400> 38

taatacgact cactatagg gtggctatgg ctaacacgta accggttata cccgggtcga 60

tgcaagtaagc tagct 75

<210> 39

<211> 86

<212> DNA

<213> Sequência Artificial

<220>

<223> sequência quimicamente sintetizada

<400> 39

taatacgact cactataggg ggggggcggt ggctgtgcaa gcggaaacgt aaccggttaa 60

acccgggtcg atgcagtaag ctagct 86

<210> 40

<211> 83

<212> DNA

<213> Sequência Artificial

<220>

<223> sequência quimicamente sintetizada

<400> 40

taatacgact cactataggg ggggtgggggc acggtactga gttacgttac cggttaaacc 60

cgggtcgatg cagtaagcta gct 83

<210> 41

<211> 78

<212> DNA

<213> Sequência Artificial

<220>

<223> sequência quimicamente sintetizada

<400> 41

taatacgact cactataggg gggagtgggg acaattagaa gatgacgtaa ccgtccgggt 60

cgatgcagta agctagct 78

<210> 42

<211> 81

<212> DNA

<213> Sequência Artificial

<220>

<223> sequência quimicamente sintetizada

<220>

<221> característica diversa

<222> (7)..(7)

<223> n é a, c, g, ou t

<220>

<221> característica diversa

<222> (38)..(38)

<223> n é a, c, g, ou t

<400> 42

taatacnact cactataggg gtgcagtgag gagcgacnag tacgttaccg gttaaattccg 60

agtcgatgca gtaagctagc t 81

<210> 43

<211> 83

<212> DNA

<213> Sequência Artificial

<220>

<223> sequência quimicamente sintetizada

<220>

<221> característica diversa

<222> (7)..(7)

<223> n é a, c, g, ou t

<400> 43

taatacnact cactataggg ggacggcac tgtggatgat ttaacggttac cggttaaacc 60

cgagtcgatg cagtaagcta gct 83

<210> 44

<211> 40

<212> DNA

<213> Sequência Artificial

<220>

<223> sequência quimicamente sintetizada

<220>

<221> característica diversa

<222> (7)..(7)

<223> n é a, c, g, ou t

<400> 44

taatacnact cactataggg gtcgatgcag taagctagct

40

<210> 45

<211> 81

<212> DNA

<213> Sequência Artificial

<220>

<223> sequência quimicamente sintetizada

<220>

<221> característica diversa

<222> (7)..(7)

<223> n é a, c, g, ou t

<400> 45

taatacnact cactataggg ggtgatattg acctctaaca gcacgtaacc ggttaaacc 60

ggtcgatgca gtaagctagc t 81

<210> 46

<211> 79

<212> DNA

<213> Sequência Artificial

<220>

<223> sequência quimicamente sintetizada

<400> 46

taatacgact cactataggg gggggggtgc agaggatgca tccaagctcg taatcggtgg 60
tcgatgcagt aagctagct 79

<210> 47

<211> 84

<212> DNA

<213> Sequência Artificial

<220>

<223> sequência quimicamente sintetizada

<400> 47

taatacgact cactataggg gggggcggtt gcttgtgcct aatcacgtaa ccggttaaac 60
ccgggtcgat gcagtaagct agct 84

<210> 48

<211> 79

<212> DNA

<213> Sequência Artificial

<220>

<223> sequência quimicamente sintetizada

<400> 48

taatacgact cactataggg gtttggtaat cgaacgtgga acgcaaccgg ttaaccggg 60
tcgatgcagt aagctagct 79

<210> 49

<211> 82

<212> DNA

<213> Sequência Artificial

<220>

<223> sequência quimicamente sintetizada

<400> 49

taatacgact cactataggg gggatggaag aggcttgata tcacgtaacc ggttaaacct 60

gggtcgatgc agtaagctag ct 82

<210> 50

<211> 76

<212> DNA

<213> Sequência Artificial

<220>

<223> sequência quimicamente sintetizada

<400> 50

taatacgact cactataggg gggtatacta actctgtaca caacgtaacc ggccgggtcg 60

atgcagtaag ctagct 76

<210> 51

<211> 81

<212> DNA

<213> Sequência Artificial

<220>

<223> sequência quimicamente sintetizada

<400> 51

taatacgact cactataggg gtataggggg ggtatcgggtg tacgtaaccg gttaaaccgg 60
ggtcgaatgca gtaagctagc t 81

<210> 52

<211> 81

<212> DNA

<213> Sequência Artificial

<220>

<223> sequência quimicamente sintetizada

<400> 52

taatacgact cactataggg gactacaata aggttccgag aacgcgaccg gttaaaccgg 60
ggtcgaatgca gtaagctagc t 81

<210> 53

<211> 82

<212> DNA

<213> Sequência Artificial

<220>

<223> sequência quimicamente sintetizada

<400> 53

taatacgact cactataggg gtgcgcgggtt acaaggcaac atacgtaacc ggttaaacc 60
 gggtcgatgc agtaagctag ct 82

<210> 54

<211> 83

<212> DNA

<213> Sequência Artificial

<220>

<223> sequência quimicamente sintetizada

<220>

<221> característica diversa

<222> (7)..(7)

<223> n é a, c, g, ou t

<220>

<221> característica diversa

<222> (42)..(42)

<223> n é a, c, g, ou t

<400> 54

taatacnact cactataggg gggacggggt gacaaagtgt cnaacgtaac cggtaaacc 60
 cgggtcgatg cagtaagcta gct 83

<210> 55

<211> 81

<212> DNA

<213> Sequência Artificial

<220>

<223> sequência quimicamente sintetizada

<400> 55

taatacgact cactataggg gagacggcgg tacaagtcca tatgtaaccg gttaaaccg 60

ggtcgatgca gtaagctagc t 81

<210> 56

<211> 81

<212> DNA

<213> Sequência Artificial

<220>

<223> sequência quimicamente sintetizada

<400> 56

taatacgact cactataggg gaggggggc ttctcgttgc cacgtaaccg cttaaaccg 60

ggtcgatgca gtaagctagc t 81

<210> 57

<211> 83

<212> DNA

<213> Sequência Artificial

<220>

<223> sequência quimicamente sintetizada

<400> 57

taatacgact cactataggg gggctgagcg tgtttgaggg accacgttac cggttaaacc 60
cgggtcgatg cagtaagcta gct 83

<210> 58

<211> 83

<212> DNA

<213> Sequência Artificial

<220>

<223> sequência quimicamente sintetizada

<400> 58

taatacgact cactataggg ggggtgggcgc aatgaaaagt tgggcgtaac cggttcaacc 60
cgggtcgatg cagtaagcta gct 83

<210> 59

<211> 81

<212> DNA

<213> Sequência Artificial

<220>

<223> sequência quimicamente sintetizada

<400> 59

taatacgact cactataggg ggtagtgaag taaggcagtg ttacgtaacc ggtgaacccg 60
ggtcgatgca gtaagctagc t 81

<210> 60

<211> 83

<212> DNA

<213> Sequência Artificial

<220>

<223> sequência quimicamente sintetizada

<400> 60

taatacgact cactataggg gggaggggtgg gctagaacac acaacgtaac cggttaaacc 60

cgggtcgatg cagtaagcta gct 83

<210> 61

<211> 84

<212> DNA

<213> Sequência Artificial

<220>

<223> sequência quimicamente sintetizada

<400> 61

taatacgact cactataggg ggggagagag gcggttacgt agggacgtta ccgattgaac 60

tcaggtcgat gcagtaagct agct 84

<210> 62

<211> 87

<212> DNA

<213> Sequência Artificial

<220>

<223> sequência quimicamente sintetizada

<400> 62

taatacgact cactataggg ggggggggcg aataggtagg gcgacgaacg ttaccggtta 60
aaccggggtc gatgcagtaa gctagct 87

<210> 63

<211> 82

<212> DNA

<213> Sequência Artificial

<220>

<223> sequência quimicamente sintetizada

<400> 63

taatacgact cactataggg ggagaggagg tccggctaga caacgtaacc ggttaaacc 60
gggtcgatgc agtaagctag ct 82

<210> 64

<211> 69

<212> DNA

<213> Sequência Artificial

<220>

<223> sequência quimicamente sintetizada

<400> 64

taatacgact cactataggg gggaggacgg gtcgtactgt taaacctggg tcgatgcagt 60
aagctagct 69

<210> 65

<211> 81

<212> DNA

<213> Sequência Artificial

<220>

<223> sequência quimicamente sintetizada

<400> 65

taatacgact cactataggg ggcgcaaca cgggaagtat acgtaaccgg tttaaaccgg 60
ggtcgatgca gtaagctagc t 81

<210> 66

<211> 82

<212> DNA

<213> Sequência Artificial

<220>

<223> sequência quimicamente sintetizada

<400> 66

taatacgact cactataggg ggaaggaaca cgcacatgca taacgtaact gggtgacccc 60
gggtcgatgc agtaactag ct 82

<210> 67

<211> 81

<212> DNA

<213> Sequência Artificial

<220>

<223> sequência quimicamente sintetizada

<400> 67

taatacgact cactataggg gagtggggag tactgtggac aacgtgaccg gttaaaccg 60

ggtcgatgca gtaagctagc t 81

<210> 68

<211> 40

<212> DNA

<213> Sequência Artificial

<220>

<223> sequência quimicamente sintetizada

<400> 68

taatacgact cactataggg gtcgatgcag taagctagct 40

<210> 69

<211> 85

<212> DNA

<213> Sequência Artificial

<220>

<223> sequência quimicamente sintetizada

<400> 69

taatacgact cactataggg gggggggcta gggcggtcgg atcggacgta accagttaaa 60

cccgggtcga tgcagtaagc tagct 85

<210> 70

<211> 83

<212> DNA

<213> Sequência Artificial

<220>

<223> sequência quimicamente sintetizada

<400> 70

taatacgact cactataggg ggggtggggg ttgctacatg ccctcgtaac cggttaagcc 60

caggtcgatg cagtaagcta gct 83

<210> 71

<211> 83

<212> DNA

<213> Sequência Artificial

<220>

<223> sequência quimicamente sintetizada

<400> 71

taatacgact cactataggg ggggtggcgac gatggaaga ataacgtaat cggttaaacc 60

cggtcgcgatg cagtaagcta gct

83

<210> 72

<211> 81

<212> DNA

<213> Sequência Artificial

<220>

<223> sequência quimicamente sintetizada

<400> 72

taatacgact cactataggg ggtaggcggg cctcatcaac aacgcaaccg gttaaaccg 60

ggtcgatgca gtaagctagc t 81

<210> 73

<211> 82

<212> DNA

<213> Sequência Artificial

<220>

<223> sequência quimicamente sintetizada

<400> 73

taatacgact cactataggg ggtggctggt aagacacaa acacgtaact cgttaaacc 60

gggtcgcgatgc agtaagctag ct 82

<210> 74

<211> 83

<212> DNA

<213> Sequência Artificial

<220>

<223> sequência quimicamente sintetizada

<400> 74

taatacgact cactataggg gggcgggcag cgcttataga tccacgtaac cggttaaacc 60

cgggtcgatg cagtaagcta gct 83

<210> 75

<211> 86

<212> DNA

<213> Sequência Artificial

<220>

<223> sequência quimicamente sintetizada

<400> 75

taatacgact cactataggg ggggggtatc tgcggttagg ctatcgacgt acccagttaa 60

acccggtcgc atgcagtaag ctagct 86

<210> 76

<211> 82

<212> DNA

<213> Sequência Artificial

<220>

<223> sequência quimicamente sintetizada

<400> 76

taatacgact cactataggg ggggtagggg acatcatagg tatacgtaac cggttaaccc 60
gggtcgatgc agtaagctag ct 82

<210> 77

<211> 81

<212> DNA

<213> Sequência Artificial

<220>

<223> sequência quimicamente sintetizada

<400> 77

taatacgact cactataggg gcgcgtgcgt gtatccatta aacgtgactg gttaaaccgc 60
ggtcgatgca gtaagctagc t 81

<210> 78

<211> 84

<212> DNA

<213> Sequência Artificial

<220>

<223> sequência quimicamente sintetizada

<400> 78

taatacgact cactataggg gggggagcgt ggatcttgag tgtatacgta accggttaaa 60
cccggtcgat gcagtaagct agct 84

<210> 79

<211> 81

<212> DNA

<213> Sequência Artificial

<220>

<223> sequência quimicamente sintetizada

<400> 79

taatacgact cactataggg gatggagag agtgtacgca tataacaaccg gttaaaccgg 60

ggtcgatgca gtaagctagc t 81

<210> 80

<211> 80

<212> DNA

<213> Sequência Artificial

<220>

<223> sequência quimicamente sintetizada

<400> 80

taatacgact cactataggg gcgggtggc gcgatggta acgtaactgg ttaaaccgg 60

gtcgatgcag taagctagct 80

<210> 81

<211> 90

<212> DNA

<213> Sequência Artificial

<220>

<223> sequência quimicamente sintetizada

<400> 81

taatacgact cactataggg gggggggggg acgtagctt ctctgtattt acgtaaccgg 60

ttaagcccgg gtcgatgcag taagctagct 90

<210> 82

<211> 40

<212> DNA

<213> Sequência Artificial

<220>

<223> sequência quimicamente sintetizada

<400> 82

taatacgact cactataggg gtcgatgcag taagctagct 40

<210> 83

<211> 83

<212> DNA

<213> Sequência Artificial

<220>

<223> sequência quimicamente sintetizada

<220>

<221> característica diversa

<222> (43)..(43)

<223> n é a, c, g, ou t

<400> 83

taatacgact cactataggg gggatggagt gggcgcaaataa anacgtaac tggttaaacc 60

cgggtcgatg cagtaagcta gct 83

<210> 84

<211> 81

<212> DNA

<213> Sequência Artificial

<220>

<223> sequência quimicamente sintetizada

<220>

<221> característica diversa

<222> (24)..(24)

<223> n é a, c, g, ou t

<220>

<221> característica diversa

<222> (39)..(39)

<223> n é a, c, g, ou t

<220>

<221> característica diversa

<222> (43)..(43)

<223> n é a, c, g, ou t

<220>

<221> característica diversa

<222> (48)..(48)

<223> n é a, c, g, ou t

<220>

<221> característica diversa

<222> (50)..(50)

<223> n é a, c, g, ou t

<220>

<221> característica diversa

<222> (69)..(71)

<223> n é a, c, g, ou t

<220>

<221> característica diversa

<222> (74)..(75)

<223> n é a, c, g, ou t

<220>

<221> característica diversa

<222> (78)..(78)

<223> n é a, c, g, ou t

<220>

<221> característica diversa

<222> (80)..(80)

<223> n é a, c, g, ou t

<400> 84

taatacgact cactataggg gagngtgagg ggtgaatant aangtaancn gttaaacctg 60

ggtcgatggn ntannctngn t 81

<210> 85

<211> 40

<212> DNA

<213> Sequência Artificial

<220>

<223> sequência quimicamente sintetizada

<220>

<221> característica diversa

<222> (1)..(1)

<223> n é a, c, g, ou t

<220>

<221> característica diversa

<222> (5)..(6)

<223> n é a, c, g, ou t

<220>

<221> característica diversa

<222> (16)..(16)

<223> n é a, c, g, ou t

<400> 85

naatnngact cacaanaggg gtcgatgcag taagctagct

40

<210> 86

<211> 84

<212> DNA

<213> Sequência Artificial

<220>

<223> sequência quimicamente sintetizada

<400> 86

taatacgact cactataggg gggggtgacg tacggatcta agtaacgtaa ccggttaaac 60

ccgggtcgat gcagtaagct agct 84

<210> 87

<211> 80

<212> DNA

<213> Sequência Artificial

<220>

<223> sequência quimicamente sintetizada

<400> 87

taatacgact cactataggg gagggacaga cactttgtag acgtaaccag ttaaaccgg 60

gtcgatgcag taagctagct 80

<210> 88

<211> 83

<212> DNA

<213> Sequência Artificial

<220>

<223> sequência quimicamente sintetizada

<400> 88

taatacgact cactataggg gggggacttg gcactacgta acaacgtaac cgcttaaacc 60

cgggtcgatg cagtaagcta gct 83

<210> 89

<211> 86

<212> DNA

<213> Sequência Artificial

<220>

<223> sequência quimicamente sintetizada

<400> 89

taatacgact cactataggg gggggggcct ctcgaccaa agcccaacgt aaccggttaa 60

acccgggtcg atgcagtaag ctagct 86

<210> 90

<211> 85

<212> DNA

<213> Sequência Artificial

<220>

<223> sequência quimicamente sintetizada

<220>

<221> característica diversa

<222> (7)..(7)

<223> n é a, c, g, ou t

<400> 90

taatacnact cactataggg ggggggggat agtcatgact gataaaacgt aactgttgag 60

cccggtcga tgcagtaagc tagct 85

<210> 91

<211> 81

<212> DNA

<213> Sequência Artificial

<220>

<223> sequência quimicamente sintetizada

<400> 91

taatacgact cactataggg gacagtgcta gtggaatagc aacgtaacca gttaaaccgg 60

ggtcgatgca gtaagctagc t 81

<210> 92

<211> 81

<212> DNA

<213> Sequência Artificial

<220>

<223> sequência quimicamente sintetizada

<400> 92

taatacgact cactataggg gacgaccact atactccgag aacgtaaccg gttaaaccg 60

ggtcgatgca gtaagctagc t 81

<210> 93

<211> 82

<212> DNA

<213> Sequência Artificial

<220>

<223> sequência quimicamente sintetizada

<400> 93

taatacgact cactataggg ggatggaggc gtagtgtagt caacgttacc ggtaaacc 60

gggtcgatgc agtaagctag ct 82

<210> 94

<211> 82

<212> DNA

<213> Sequência Artificial

<220>

<223> sequência quimicamente sintetizada

<400> 94

taatacgact cactataggg gggaggtata gatggaatgg ttatgtaacc tgttaaacc 60

gggtcgatgc agtaagctag ct 82

<210> 95

<211> 81

<212> DNA

<213> Sequência Artificial

<220>

<223> sequência quimicamente sintetizada

<400> 95

taatacgact cactataggg gtggggagga ccacttagat aacgtcaccg gttaaaccg 60

ggtcgatgca gtaagctagc t 81

<210> 96

<211> 83

<212> DNA

<213> Sequência Artificial

<220>

<223> sequência quimicamente sintetizada

<400> 96

taatacgact cactataggg gggatagggg cgagagagtc acaacgtaac cggttaatcc 60
cgggtcgatg cagtaagcta gct 83

<210> 97

<211> 83

<212> DNA

<213> Sequência Artificial

<220>

<223> sequência quimicamente sintetizada

<400> 97

taatacgact cactataggg ggggatggc cgaatcataa aataacgtaa cggtagacc 60
cgggtcgatg cagtaagcta gct 83

<210> 98

<211> 78

<212> DNA

<213> Sequência Artificial

<220>

<223> sequência quimicamente sintetizada

<400> 98

taatacgact cactataggg ggcgattgct gagtcagttc gtaatcggtt aaaccgggt 60
cgatgcagta agctagct 78

<210> 99

<211> 85

<212> DNA

<213> Sequência Artificial

<220>

<223> sequência quimicamente sintetizada

<400> 99

taatacgact cactataggg gggggaggat ccgaaacaca gggatccgta accggttaaa 60

gccgggtcga tgcagtaagc tagct 85

<210> 100

<211> 883

<212> PRT

<213> Sequência Artificial

<220>

<223> T7 polimerase modificada, quimicamente sintetizada Y639L

<400> 100

Met Asn Thr Ile Asn Ile Ala Lys Asn Asp Phe Ser Asp Ile Glu Leu

1 5 10 15

Ala Ala Ile Pro Phe Asn Thr Leu Ala Asp His Tyr Gly Glu Arg Leu

20 25 30

Ala Arg Glu Gln Leu Ala Leu Glu His Glu Ser Tyr Glu Met Gly Glu

35 40 45

Ala Arg Phe Arg Lys Met Phe Glu Arg Gln Leu Lys Ala Gly Glu Val
 50 55 60
 Ala Asp Asn Ala Ala Ala Lys Pro Leu Ile Thr Thr Leu Leu Pro Lys
 65 70 75 80
 Met Ile Ala Arg Ile Asn Asp Trp Phe Glu Glu Val Lys Ala Lys Arg
 85 90 95
 Gly Lys Arg Pro Thr Ala Phe Gln Phe Leu Gln Glu Ile Lys Pro Glu
 100 105 110
 Ala Val Ala Tyr Ile Thr Ile Lys Thr Thr Leu Ala Cys Leu Thr Ser
 115 120 125
 Ala Asp Asn Thr Thr Val Gln Ala Val Ala Ser Ala Ile Gly Arg Ala
 130 135 140
 Ile Glu Asp Glu Ala Arg Phe Gly Arg Ile Arg Asp Leu Glu Ala Lys
 145 150 155 160
 His Phe Lys Lys Asn Val Glu Glu Gln Leu Asn Lys Arg Val Gly His
 165 170 175
 Val Tyr Lys Lys Ala Phe Met Gln Val Val Glu Ala Asp Met Leu Ser
 180 185 190
 Lys Gly Leu Leu Gly Gly Glu Ala Trp Ser Ser Trp His Lys Glu Asp
 195 200 205
 Ser Ile His Val Gly Val Arg Cys Ile Glu Met Leu Ile Glu Ser Thr
 210 215 220
 Gly Met Val Ser Leu His Arg Gln Asn Ala Gly Val Val Gly Gln Asp
 225 230 235 240
 Ser Glu Thr Ile Glu Leu Ala Pro Glu Tyr Ala Glu Ala Ile Ala Thr
 245 250 255
 Arg Ala Gly Ala Leu Ala Gly Ile Ser Pro Met Phe Gln Pro Cys Val
 260 265 270

Val Pro Pro Lys Pro Trp Thr Gly Ile Thr Gly Gly Gly Tyr Trp Ala
 275 280 285
 Asn Gly Arg Arg Pro Leu Ala Leu Val Arg Thr His Ser Lys Lys Ala
 290 295 300
 Leu Met Arg Tyr Glu Asp Val Tyr Met Pro Glu Val Tyr Lys Ala Ile
 305 310 315 320
 Asn Ile Ala Gln Asn Thr Ala Trp Lys Ile Asn Lys Lys Val Leu Ala
 325 330 335
 Val Ala Asn Val Ile Thr Lys Trp Lys His Cys Pro Val Glu Asp Ile
 340 345 350
 Pro Ala Ile Glu Arg Glu Glu Leu Pro Met Lys Pro Glu Asp Ile Asp
 355 360 365
 Met Asn Pro Glu Ala Leu Thr Ala Trp Lys Arg Ala Ala Ala Val
 370 375 380
 Tyr Arg Lys Asp Lys Ala Arg Lys Ser Arg Arg Ile Ser Leu Glu Phe
 385 390 395 400
 Met Leu Glu Gln Ala Asn Lys Phe Ala Asn His Lys Ala Ile Trp Phe
 405 410 415
 Pro Tyr Asn Met Asp Trp Arg Gly Arg Val Tyr Ala Val Ser Met Phe
 420 425 430
 Asn Pro Gln Gly Asn Asp Met Thr Lys Gly Leu Leu Thr Leu Ala Lys
 435 440 445
 Gly Lys Pro Ile Gly Lys Glu Gly Tyr Tyr Trp Leu Lys Ile His Gly
 450 455 460
 Ala Asn Cys Ala Gly Val Asp Lys Val Pro Phe Pro Glu Arg Ile Lys
 465 470 475 480
 Phe Ile Glu Glu Asn His Glu Asn Ile Met Ala Cys Ala Lys Ser Pro
 485 490 495

Leu Glu Asn Thr Trp Trp Ala Glu Gln Asp Ser Pro Phe Cys Phe Leu
 500 505 510
 Ala Phe Cys Phe Glu Tyr Ala Gly Val Gln His His Gly Leu Ser Tyr
 515 520 525
 Asn Cys Ser Leu Pro Leu Ala Ile Asp Gly Ser Cys Ser Gly Ile Gln
 530 535 540
 His Phe Ser Ala Met Leu Arg Asp Glu Val Gly Gly Arg Ala Val Asn
 545 550 555 560
 Leu Leu Pro Ser Glu Thr Val Gln Asp Ile Tyr Gly Ile Val Ala Lys
 565 570 575
 Lys Val Asn Glu Ile Leu Gln Ala Asp Ala Ile Asn Gly Thr Asp Asn
 580 585 590
 Glu Val Val Thr Val Thr Asp Glu Asn Thr Gly Glu Ile Ser Glu Lys
 595 600 605
 Val Lys Leu Gly Thr Lys Ala Leu Ala Gly Gln Trp Leu Ala Tyr Gly
 610 615 620
 Val Thr Arg Ser Val Thr Lys Arg Ser Val Met Thr Leu Ala Leu Gly
 625 630 635 640
 Ser Lys Glu Phe Gly Phe Arg Gln Gln Val Leu Glu Asp Thr Ile Gln
 645 650 655
 Pro Ala Ile Asp Ser Gly Lys Gly Leu Met Phe Thr Gln Pro Asn Gln
 660 665 670
 Ala Ala Gly Tyr Met Ala Lys Leu Ile Trp Glu Ser Val Ser Val Thr
 675 680 685
 Val Val Ala Ala Val Glu Ala Met Asn Trp Leu Lys Ser Ala Ala Lys
 690 695 700
 Leu Leu Ala Ala Glu Val Lys Asp Lys Lys Thr Gly Glu Ile Leu Arg
 705 710 715 720

Lys Arg Cys Ala Val His Trp Val Thr Pro Asp Gly Phe Pro Val Trp
 725 730 735
 Gln Glu Tyr Lys Lys Pro Ile Gln Thr Arg Leu Asn Leu Met Phe Leu
 740 745 750
 Gly Gln Phe Arg Leu Gln Pro Thr Ile Asn Thr Asn Lys Asp Ser Glu
 755 760 765
 Ile Asp Ala His Lys Gln Glu Ser Gly Ile Ala Pro Asn Phe Val His
 770 775 780
 Ser Gln Asp Gly Ser His Leu Arg Lys Thr Val Val Trp Ala His Glu
 785 790 795 800
 Lys Tyr Gly Ile Glu Ser Phe Ala Leu Ile His Asp Ser Phe Gly Thr
 805 810 815
 Ile Pro Ala Asp Ala Ala Asn Leu Phe Lys Ala Val Arg Glu Thr Met
 820 825 830
 Val Asp Thr Tyr Glu Ser Cys Asp Val Leu Ala Asp Phe Tyr Asp Gln
 835 840 845
 Phe Ala Asp Gln Leu His Glu Ser Gln Leu Asp Lys Met Pro Ala Leu
 850 855 860
 Pro Ala Lys Gly Asn Leu Asn Leu Arg Asp Ile Leu Glu Ser Asp Phe
 865 870 875 880
 Ala Phe Ala

<210> 101

<211> 883

<212> PRT

<213> Sequência Artificial

<220>

<223> T7 polimerase modificada, quimicamente sintetizada, Y639L/K378R

<400> 101

Met Asn Thr Ile Asn Ile Ala Lys Asn Asp Phe Ser Asp Ile Glu Leu

1 5 10 15

Ala Ala Ile Pro Phe Asn Thr Leu Ala Asp His Tyr Gly Glu Arg Leu

20 25 30

Ala Arg Glu Gln Leu Ala Leu Glu His Glu Ser Tyr Glu Met Gly Glu

35 40 45

Ala Arg Phe Arg Lys Met Phe Glu Arg Gln Leu Lys Ala Gly Glu Val

50 55 60

Ala Asp Asn Ala Ala Ala Lys Pro Leu Ile Thr Thr Leu Leu Pro Lys

65 70 75 80

Met Ile Ala Arg Ile Asn Asp Trp Phe Glu Glu Val Lys Ala Lys Arg

85 90 95

Gly Lys Arg Pro Thr Ala Phe Gln Phe Leu Gln Glu Ile Lys Pro Glu

100 105 110

Ala Val Ala Tyr Ile Thr Ile Lys Thr Thr Leu Ala Cys Leu Thr Ser

115 120 125

Ala Asp Asn Thr Thr Val Gln Ala Val Ala Ser Ala Ile Gly Arg Ala

130 135 140

Ile Glu Asp Glu Ala Arg Phe Gly Arg Ile Arg Asp Leu Glu Ala Lys

145 150 155 160

His Phe Lys Lys Asn Val Glu Glu Gln Leu Asn Lys Arg Val Gly His

165 170 175

Val Tyr Lys Lys Ala Phe Met Gln Val Val Glu Ala Asp Met Leu Ser

180 185 190

Lys Gly Leu Leu Gly Gly Glu Ala Trp Ser Ser Trp His Lys Glu Asp
 195 200 205
 Ser Ile His Val Gly Val Arg Cys Ile Glu Met Leu Ile Glu Ser Thr
 210 215 220
 Gly Met Val Ser Leu His Arg Gln Asn Ala Gly Val Val Gly Gln Asp
 225 230 235 240
 Ser Glu Thr Ile Glu Leu Ala Pro Glu Tyr Ala Glu Ala Ile Ala Thr
 245 250 255
 Arg Ala Gly Ala Leu Ala Gly Ile Ser Pro Met Phe Gln Pro Cys Val
 260 265 270
 Val Pro Pro Lys Pro Trp Thr Gly Ile Thr Gly Gly Gly Tyr Trp Ala
 275 280 285
 Asn Gly Arg Arg Pro Leu Ala Leu Val Arg Thr His Ser Lys Lys Ala
 290 295 300
 Leu Met Arg Tyr Glu Asp Val Tyr Met Pro Glu Val Tyr Lys Ala Ile
 305 310 315 320
 Asn Ile Ala Gln Asn Thr Ala Trp Lys Ile Asn Lys Lys Val Leu Ala
 325 330 335
 Val Ala Asn Val Ile Thr Lys Trp Lys His Cys Pro Val Glu Asp Ile
 340 345 360
 Pro Ala Ile Glu Arg Glu Glu Leu Pro Met Lys Pro Glu Asp Ile Asp
 355 360 365
 Met Asn Pro Glu Ala Leu Thr Ala Trp Arg Arg Ala Ala Ala Val
 370 375 380
 Tyr Arg Lys Asp Lys Ala Arg Lys Ser Arg Arg Ile Ser Leu Glu Phe
 385 390 395 400
 Met Leu Glu Gln Ala Asn Lys Phe Ala Asn His Lys Ala Ile Trp Phe
 405 410 415

Pro Tyr Asn Met Asp Trp Arg Gly Arg Val Tyr Ala Val Ser Met Phe
 420 425 430
 Asn Pro Gln Gly Asn Asp Met Thr Lys Gly Leu Leu Thr Leu Ala Lys
 435 440 445
 Gly Lys Pro Ile Gly Lys Glu Gly Tyr Tyr Trp Leu Lys Ile His Gly
 450 455 460
 Ala Asn Cys Ala Gly Val Asp Lys Val Pro Phe Pro Glu Arg Ile Lys
 465 470 475 480
 Phe Ile Glu Glu Asn His Glu Asn Ile Met Ala Cys Ala Lys Ser Pro
 485 490 495
 Leu Glu Asn Thr Trp Trp Ala Glu Gln Asp Ser Pro Phe Cys Phe Leu
 500 505 510
 Ala Phe Cys Phe Glu Tyr Ala Gly Val Gln His His Gly Leu Ser Tyr
 515 520 525
 Asn Cys Ser Leu Pro Leu Ala Phe Asp Gly Ser Cys Ser Gly Ile Gln
 530 535 540
 His Phe Ser Ala Met Leu Arg Asp Glu Val Gly Gly Arg Ala Val Asn
 545 550 555 560
 Leu Leu Pro Ser Glu Thr Val Gln Asp Ile Tyr Gly Ile Val Ala Lys
 565 570 575
 Lys Val Asn Glu Ile Leu Gln Ala Asp Ala Ile Asn Gly Thr Asp Asn
 580 585 590
 Glu Val Val Thr Val Thr Asp Glu Asn Thr Gly Glu Ile Ser Glu Lys
 595 600 605
 Val Lys Leu Gly Thr Lys Ala Leu Ala Gly Gln Trp Leu Ala Tyr Gly
 610 615 620
 Val Thr Arg Ser Val Thr Lys Arg Ser Val Met Thr Leu Ala Leu Gly
 625 630 635 640

Ser Lys Glu Phe Gly Phe Arg Gln Gln Val Leu Glu Asp Thr Ile Gln
 645 650 655
 Pro Ala Ile Asp Ser Gly Lys Gly Leu Met Phe Thr Gln Pro Asn Gln
 660 665 670
 Ala Ala Gly Tyr Met Ala Lys Leu Ile Trp Glu Ser Val Ser Val Thr
 675 680 685
 Val Val Ala Ala Val Glu Ala Met Asn Trp Leu Lys Ser Ala Ala Lys
 690 695 700
 Leu Leu Ala Ala Glu Val Lys Asp Lys Lys Thr Gly Glu Ile Leu Arg
 705 710 715 720
 Lys Arg Cys Ala Val His Trp Val Thr Pro Asp Gly Phe Pro Val Trp
 725 730 735
 Gln Glu Tyr Lys Lys Pro Ile Gln Thr Arg Leu Asn Leu Met Phe Leu
 740 745 750
 Gly Gln Phe Arg Leu Gln Pro Thr Ile Asn Thr Asn Lys Asp Ser Glu
 755 760 765
 Ile Asp Ala His Lys Gln Glu Ser Gly Ile Ala Pro Asn Phe Val His
 770 775 780
 Ser Gln Asp Gly Ser His Leu Arg Lys Thr Val Val Trp Ala His Glu
 785 790 795 800
 Lys Tyr Gly Ile Glu Ser Phe Ala Leu Ile His Asp Ser Phe Gly Thr
 805 810 815
 Ile Pro Ala Asp Ala Ala Asn Leu Phe Lys Ala Val Arg Glu Thr Met
 820 825 830
 Val Asp Thr Tyr Glu Ser Cys Asp Val Leu Ala Asp Phe Tyr Asp Gln
 835 840 845
 Phe Ala Asp Gln Leu His Glu Ser Gln Leu Asp Lys Met Pro Ala Leu
 850 855 860

Pro Ala Lys Gly Asn Leu Asn Leu Arg Asp Ile Leu Glu Ser Asp Phe

865 870 875 880

Ala Phe Ala

<210> 102

<211> 883

<212> PRT

<213> Sequência Artificial

<220>

<223> T7 polimerase modificada, quimicamente sintetizada,

P266L/Y639L/H784A

<400> 102

Met Asn Thr Ile Asn Ile Ala Lys Asn Asp Phe Ser Asp Ile Glu Leu

1 5 10 15

Ala Ala Ile Pro Phe Asn Thr Leu Ala Asp His Tyr Gly Glu Arg Leu

20 25 30

Ala Arg Glu Gln Leu Ala Leu Glu His Glu Ser Tyr Glu Met Gly Glu

35 40 45

Ala Arg Phe Arg Lys Met Phe Glu Arg Gln Leu Lys Ala Gly Glu Val

50 55 60

Ala Asp Asn Ala Ala Ala Lys Pro Leu Ile Thr Thr Leu Leu Pro Lys

65 70 75 80

Met Ile Ala Arg Ile Asn Asp Trp Phe Glu Glu Val Lys Ala Lys Arg

85 90 95

Gly Lys Arg Pro Thr Ala Phe Gln Phe Leu Gln Glu Ile Lys Pro Glu

100 105 110

Ala Val Ala Tyr Ile Thr Ile Lys Thr Thr Leu Ala Cys Leu Thr Ser
 115 120 125
 Ala Asp Asn Thr Thr Val Gln Ala Val Ala Ser Ala Ile Gly Arg Ala
 130 135 140
 Ile Glu Asp Glu Ala Arg Phe Gly Arg Ile Arg Asp Leu Glu Ala Lys
 145 150 155 160
 His Phe Lys Lys Asn Val Glu Glu Gln Leu Asn Lys Arg Val Gly His
 165 170 175
 Val Tyr Lys Lys Ala Phe Met Gln Val Val Glu Ala Asp Met Leu Ser
 180 185 190
 Lys Gly Leu Leu Gly Gly Glu Ala Trp Ser Ser Trp His Lys Glu Asp
 195 200 205
 Ser Ile His Val Gly Val Arg Cys Ile Glu Met Leu Ile Glu Ser Thr
 210 215 220
 Gly Met Val Ser Leu His Arg Gln Asn Ala Gly Val Val Gly Gln Asp
 225 230 235 240
 Ser Glu Thr Ile Glu Leu Ala Pro Glu Tyr Ala Glu Ala Ile Ala Thr
 245 250 255
 Arg Ala Gly Ala Leu Ala Gly Ile Ser Leu Met Phe Gln Pro Cys Val
 260 265 270
 Val Pro Pro Lys Pro Trp Thr Gly Ile Thr Gly Gly Gly Tyr Trp Ala
 275 280 285
 Asn Gly Arg Arg Pro Leu Ala Leu Val Arg Thr His Ser Lys Lys Ala
 290 295 300
 Leu Met Arg Tyr Glu Asp Val Tyr Met Pro Glu Val Tyr Lys Ala Ile
 305 310 315 320
 Asn Ile Ala Gln Asn Thr Ala Trp Lys Ile Asn Lys Lys Val Leu Ala
 325 330 335

Val Ala Asn Val Ile Thr Lys Trp Lys His Cys Pro Val Glu Asp Ile
 340 345 350
 Pro Ala Ile Glu Arg Glu Glu Leu Pro Met Lys Pro Glu Asp Ile Asp
 355 360 365
 Met Asn Pro Glu Ala Leu Thr Ala Trp Lys Arg Ala Ala Ala Ala Val
 370 375 380
 Tyr Arg Lys Asp Lys Ala Arg Lys Ser Arg Arg Ile Ser Leu Glu Phe
 385 390 395 400
 Met Leu Glu Gln Ala Asn Lys Phe Ala Asn His Lys Ala Ile Trp Phe
 405 410 415
 Pro Tyr Asn Met Asp Trp Arg Gly Arg Val Tyr Ala Val Ser Met Phe
 420 425 430
 Asn Pro Gln Gly Asn Asp Met Thr Lys Gly Leu Leu Thr Leu Ala Lys
 435 440 445
 Gly Lys Pro Ile Gly Lys Glu Gly Tyr Tyr Trp Leu Lys Ile His Gly
 430 455 460
 Ala Asn Cys Ala Gly Val Asp Lys Val Pro Phe Pro Glu Arg Ile Lys
 465 470 475 480
 Phe Ile Glu Glu Asn His Glu Asn Ile Met Ala Cys Ala Lys Ser Pro
 485 490 495
 Leu Glu Asn Thr Trp Trp Ala Glu Gln Asp Ser Pro Phe Cys Phe Leu
 500 505 510
 Ala Phe Cys Phe Glu Tyr Ala Gly Val Gln His His Gly Leu Ser Tyr
 515 520 525
 Asn Cys Ser Leu Pro Leu Ala Phe Asp Gly Ser Cys Ser Gly Ile Gln
 530 535 540
 His Phe Ser Ala Met Leu Arg Asp Glu Val Gly Gly Arg Ala Val Asn
 545 550 555 560

Leu Leu Pro Ser Glu Thr Val Gln Asp Ile Tyr Gly Ile Val Ala Lys
 565 570 575
 Lys Val Asn Glu Ile Leu Gln Ala Asp Ala Ile Asn Gly Thr Asp Asn
 580 585 590
 Glu Val Val Thr Val Thr Asp Glu Asn Thr Gly Glu Ile Ser Glu Lys
 595 600 605
 Val Lys Leu Gly Thr Lys Ala Leu Ala Gly Gln Trp Leu Ala Tyr Gly
 610 615 620
 Val Thr Arg Ser Val Thr Lys Arg Ser Val Met Thr Leu Ala Leu Gly
 625 630 635 640
 Ser Lys Glu Phe Gly Phe Arg Gln Gln Val Leu Glu Asp Thr Ile Gln
 645 650 655
 Pro Ala Ile Asp Ser Gly Lys Gly Leu Met Phe Thr Gln Pro Asn Gln
 660 665 670
 Ala Ala Gly Tyr Met Ala Lys Leu Ile Trp Glu Ser Val Ser Val Thr
 675 680 685
 Val Val Ala Ala Val Glu Ala Met Asn Trp Leu Lys Ser Ala Ala Lys
 690 695 700
 Leu Leu Ala Ala Glu Val Lys Asp Lys Lys Thr Gly Glu Ile Leu Arg
 705 710 715 720
 Lys Arg Cys Ala Val His Trp Val Thr Pro Asp Gly Phe Pro Val Trp
 725 730 735
 Gln Glu Tyr Lys Lys Pro Ile Gln Thr Arg Leu Asn Leu Met Phe Leu
 740 745 750
 Gly Gln Phe Arg Leu Gln Pro Thr Ile Asn Thr Asn Lys Asp Ser Glu
 755 760 765
 Ile Asp Ala His Lys Gln Glu Ser Gly Ile Ala Pro Asn Phe Val Ala
 770 775 780

[illegible]

<210> 103

<211> 883

<212> PRT

<213> Sequência Artificial

<220>

<223> T7 polimerase modificada, quimicamente sintetizada,
P2EEL/Y639L/H784A/K378R

<400> 103

Met Asn Thr Ile Asn Ile Ala Lys Asn Asp Phe Ser Asp Ile Glu Leu
1 5 10 15
Ala Ala Ile Pro Phe Asn Thr Leu Ala Asp His Tyr Gly Glu Arg Leu
20 25 30

Ala Arg Glu Gln Leu Ala Leu Glu His Glu Ser Tyr Glu Met Gly Glu			
35	40	45	
Ala Arg Phe Arg Lys Met Phe Glu Arg Gln Leu Lys Ala Gly Glu Val			
50	55	60	
Ala Asp Asn Ala Ala Ala Lys Pro Leu Ile Thr Thr Leu Leu Pro Lys			
65	70	75	80
Met Ile Ala Arg Ile Asn Asp Trp Phe Glu Glu Val Lys Ala Lys Arg			
	85	90	95
Gly Lys Arg Pro Thr Ala Phe Gln Phe Leu Gln Glu Ile Lys Pro Glu			
	100	105	110
Ala Val Ala Tyr Ile Thr Ile Lys Thr Thr Leu Ala Cys Leu Thr Ser			
	115	120	125
Ala Asp Asn Thr Thr Val Gln Ala Val Ala Ser Ala Ile Gly Arg Ala			
	130	135	140
Ile Glu Asp Glu Ala Arg Phe Gly Arg Ile Arg Asp Leu Glu Ala Lys			
145	150	155	160
His Phe Lys Lys Asn Val Glu Glu Gln Leu Asn Lys Arg Val Gly His			
	165	170	175
Val Tyr Lys Lys Ala Phe Met Gln Val Val Glu Ala Asp Met Leu Ser			
	180	185	190
Lys Gly Leu Leu Gly Gly Glu Ala Trp Ser Ser Trp His Lys Glu Asp			
	195	200	205
Ser Ile His Val Gly Val Arg Cys Ile Glu Met Leu Ile Glu Ser Thr			
	210	215	220
Gly Met Val Ser Leu His Arg Gln Asn Ala Gly Val Val Gly Gln Asp			
225	230	235	240
Ser Glu Thr Ile Glu Leu Ala Pro Glu Tyr Ala Glu Ala Ile Ala Thr			
	245	250	255

Arg Ala Gly Ala Leu Ala Gly Ile Ser Leu Met Phe Gln Pro Cys Val
 260 265 270
 Val Pro Pro Lys Pro Trp Thr Gly Ile Thr Gly Gly Gly Tyr Trp Ala
 275 280 285
 Asn Gly Arg Arg Pro Leu Ala Leu Val Arg Thr His Ser Lys Lys Ala
 290 295 300
 Leu Met Arg Tyr Glu Asp Val Tyr Met Pro Glu Val Tyr Lys Ala Ile
 305 310 315 320
 Asn Ile Ala Gln Asn Thr Ala Trp Lys Ile Asn Lys Lys Val Leu Ala
 325 330 335
 Val Ala Asn Val Ile Thr Lys Trp Lys His Cys Pro Val Glu Asp Ile
 340 345 350
 Pro Ala Ile Glu Arg Glu Glu Leu Pro Met Lys Pro Glu Asp Ile Asp
 355 360 365
 Met Asn Pro Glu Ala Leu Thr Ala Trp Arg Arg Ala Ala Ala Ala Val
 370 375 380
 Tyr Arg Lys Asp Lys Ala Arg Lys Ser Arg Arg Ile Ser Leu Glu Phe
 385 390 395 400
 Met Leu Glu Gln Ala Asn Lys Phe Ala Asn His Lys Ala Ile Trp Phe
 405 410 415
 Pro Tyr Asn Met Asp Trp Arg Gly Arg Val Tyr Ala Val Ser Met Phe
 420 425 430
 Asn Pro Gln Gly Asn Asp Met Thr Lys Gly Leu Leu Thr Leu Ala Lys
 435 440 445
 Gly Lys Pro Ile Gly Lys Glu Gly Tyr Tyr Trp Leu Lys Ile His Gly
 450 455 460
 Ala Asn Cys Ala Gly Val Asp Lys Val Pro Phe Pro Glu Arg Ile Lys
 465 470 475 480

Phe Ile Glu Glu Asn His Glu Asn Ile Met Ala Cys Ala Lys Ser Pro			
485	490	495	
Leu Glu Asn Thr Trp Trp Ala Glu Gln Asp Ser Pro Phe Cys Phe Leu			
500	505	510	
Ala Phe Cys Phe Glu Tyr Ala Gly Val Gln His His Gly Leu Ser Tyr			
515	520	525	
Asn Cys Ser Leu Pro Leu Ala Phe Asp Gly Ser Cys Ser Gly Ile Gln			
530	535	540	
His Phe Ser Ala Met Leu Arg Asp Glu Val Gly Gly Arg Ala Val Asn			
545	550	555	560
Leu Leu Pro Ser Glu Thr Val Gln Asp Ile Tyr Gly Ile Val Ala Lys			
565	570	575	
Lys Val Asn Glu Ile Leu Gln Ala Asp Ala Ile Asn Gly Thr Asp Asn			
580	585	590	
Glu Val Val Thr Val Thr Asp Glu Asn Thr Gly Glu Ile Ser Glu Lys			
595	600	605	
Val Lys Leu Gly Thr Lys Ala Leu Ala Gly Gln Trp Leu Ala Tyr Gly			
610	615	620	
Val Thr Arg Ser Val Thr Lys Arg Ser Val Met Thr Leu Ala Leu Gly			
625	630	635	640
Ser Lys Glu Phe Gly Phe Arg Gln Gln Val Leu Glu Asp Thr Ile Gln			
645	650	655	
Pro Ala Ile Asp Ser Gly Lys Gly Leu Met Phe Thr Gln Pro Asn Gln			
660	665	670	
Ala Ala Gly Tyr Met Ala Lys Leu Ile Trp Glu Ser Val Ser Val Thr			
675	680	685	
Val Val Ala Ala Val Glu Ala Met Asn Trp Leu Lys Ser Ala Ala Lys			
690	695	700	

Leu Leu Ala Ala Glu Val Lys Asp Lys Lys Thr Gly Glu Ile Leu Arg
 705 710 715 720
 Lys Arg Cys Ala Val His Trp Val Thr Pro Asp Gly Phe Pro Val Trp
 725 730 735
 Gln Glu Tyr Lys Lys Pro Ile Gln Thr Arg Leu Asn Leu Met Phe Leu
 740 745 750
 Gly Gln Phe Arg Leu Gln Pro Thr Ile Asn Thr Asn Lys Asp Ser Glu
 755 760 765
 Ile Asp Ala His Lys Gln Glu Ser Gly Ile Ala Pro Asn Phe Val Ala
 770 775 780
 Ser Gln Asp Gly Ser His Leu Arg Lys Thr Val Val Trp Ala His Glu
 785 790 795 800
 Lys Tyr Gly Ile Glu Ser Phe Ala Leu Ile His Asp Ser Phe Gly Thr
 805 810 815
 Ile Pro Ala Asp Ala Ala Asn Leu Phe Lys Ala Val Arg Glu Thr Met
 820 825 830
 Val Asp Thr Tyr Glu Ser Cys Asp Val Leu Ala Asp Phe Tyr Asp Gln
 835 840 845
 Phe Ala Asp Gln Leu His Glu Ser Gln Leu Asp Lys Met Pro Ala Leu
 850 855 860
 Pro Ala Lys Gly Asn Leu Asn Leu Arg Asp Ile Leu Glu Ser Asp Phe
 865 870 875 880
 Ala Phe Ala

<210> 104

<211> 20

<212> DNA

<213> Sequência Artificial

<220>

<223> primer em 5', quimicamente sintetizado

<400> 104

agctagctta ctgcatcgac

20

<210> 105

<211> 18

<212> DNA

<213> Sequência Artificial

<220>

<223> primer em 3', quimicamente sintetizado

<400> 105

taatacgact cactatag

18

<210> 106

<211> 94

<212> DNA

<213> Sequência Artificial

<220>

<223> molde de transcrição de DNA de controle, quimicamente sintetizado,
sem sequência líder, ARC2117

<220>

<221> característica diversa

<222> (41)..(70)

<223> n é a, t, c ou g

<400> 106

taatacgact cactataggg gagaggagag aacgttctcg nnnnnnnnnn nnnnnnnnnn 60

nnnnnnnnnn ggatcggtac gactagcatc gatg 94

<210> 107

<211> 35

<212> DNA

<213> Sequência Artificial

<220>

<223> sequência quimicamente sintetizada

<400> 107

agtcattgac ctggctctgg ggtccaaaga gttcg 35

<210> 108

<211> 34

<212> DNA

<213> Sequência Artificial

<220>

<223> sequência quimicamente sintetizada

<400> 108

gaactctttg gaccccagag ccagcgatcat gact

34

<210> 109

<211> 31

<212> DNA

<213> Sequência Artificial

<220>

<223> sequência quimicamente sintetizada

<400> 109

ggctggcatc tctctgatgt tccaaccttg c

31

<210> 110

<211> 31

<212> DNA

<213> Sequência Artificial

<220>

<223> sequência quimicamente sintetizada

<400> 110

gcaagggttg aacatcagag agatgccagc c

31

<210> 111

<211> 34

<212> DNA

<213> Sequência Artificial

<220>

<223> sequência quimicamente sintetizada

<400> 111

cgctcctaac ttgtagcca gccaaagacgg tagc

34

<210> 112

<211> 34

<212> DNA

<213> Sequência Artificial

<220>

<223> sequência quimicamente sintetizada

<400> 112

gctaccgtct tggctggcta caaagttagg agcg

34

<210> 113

<211> 33

<212> DNA

<213> Sequência Artificial

<220>

<223> sequência quimicamente sintetizada

<400> 113

gctctcaccg cgtggagacg tgctgccgct gct

33

<210> 114

<211> 33

<212> DNA

<213> Sequência Artificial

<220>

<223> sequência quimicamente sintetizada

<400> 114

agcagcggca gcacgtctcc acgcggtgag agc

33

<210> 115

<211> 66

<212> DNA

<213> Sequência Artificial

<220>

<223> primer em 5', quimicamente sintetizado

<400> 115

aaaaaaaaaa aaaaaaaaaa aaaaaataa tacgactcac tataggggag tacaataacg 60

ttctcg 66

<210> 116

<211> 20

<212> DNA

<213> Sequência Artificial

<220>

<223> primer em 3', quimicamente sintetizada

<400> 116

catcgatgct agtcgtaacg

20

<210> 117

<211> 24

<212> DNA

<213> Sequência Artificial

<220>

<223> molde de transcrição de DNA, quimicamente sintetizada

<220>

<221> característica diversa

<222> (5)..(20)

<223> n é a, t, c ou g

<220>

<221> característica diversa

<222> (21)..(24)

<223> n é a, c , ou t

<400> 117

ggggnnnnnnn nnnnnnnnnn nnnn

24

<210> 118

<211> 59

<212> DNA

<213> Sequência Artificial

<220>

<223> primer quimicamente sintetizado

<400> 118

gatcgatcga tcgatcgatc taatacgact cactataggg gagtacaata acgttctcg 59

<210> 119

<211> 24

<212> DNA

<213> Sequência Artificial

<220>

<223> primer quimicamente sintetizado

<400> 119

catcgatgct agtcgtaacg atcc 24

<210> 120

<211> 2652

<212> DNA

<213> Sequência Artificial

<220>

<223> sequência de T7 RNA polimerase do tipo selvagem, quimicamente sintetizada

<400> 120

atgaacacga ttaacatcgc taagaacgac ttctctgaca tcgaactggc tgctatcccg	60
ttcaaacactc tggctgacca ttacggtgag cgtttagctc gcgaacagtt ggcccttgag	120
catgagtctt acgagatggg tgaagcacgc ttccgcaaga tgtttgagcg tcaacttaaa	180
gctggtgagg ttgcggataa cgctgccgcc aagcctctca tcactaccct actccctaag	240
atgattgcac gcatcaacga ctggtttgag gaagtgaaag ctaagcgcg caagcgcccg	300
acagccttcc agttcctgca agaaatcaag ccggaagccg tagcgtacat caccattaag	360
accactctgg cttgcctaac cagtgtgac aatacaaccg ttcaggctgt agcaagcgca	420
atcggtcggg ccattgagga cgaggctcgc ttcggtcgta tccgtgacct tgaagctaag	480
cacttcaaga aaaacgttga ggaacaactc aacaagcgcg tagggcacgt ctacaagaaa	540
gcatttatgc aagttgtcga ggctgacatg ctctctaagg gtctactcgg tggcgaggcg	600
tggtcttcgt ggcataagga agactctatt catgtaggag tacgctgcat cgagatgctc	660
attgagtcaa ccggaatggt tagcttacac cgccaaaatg ctggcgtagt aggtcaagac	720
tctgagacta tcgaactcgc acctgaatac gctgaggcta tcgcaaccg tgcaggtgcg	780
ctggctggca tctctccgat gttccaacct tgcgtagttc ctctaagcc gtggactggc	840
attactggtg gtggctattg ggctaacggt cgtcgtcctc tggcgctggt gcgtactcac	900
agtaagaaag cactgatgcg ctacgaagac gtttacatgc ctgaggtgta caaagcgatt	960
aacattgcgc aaacacccgc atggaaaatc aacaagaaag tcctagcggc cgccaacgta	1020
atcaccaagt ggaagcattg tccggtcgag gacatccctg cgattgagcg tgaagaactc	1080
ccgatgaaac cggaagacat cgacatgaat cctgaggctc tcaccgcgtg gaaacgtgct	1140
gccgctgctg tgtaccgcaa ggacaaggct cgcaagtctc gccgtatcag ccttgagtgc	1200
atgcttgagc aagccaataa gtttgctaac cataaggcca tctggttccc ttacaacatg	1260
gactggcgcg gtcgtgttta cgctgtgtca atgttcaacc cgcaaggtaa cgatatgacc	1320
aaaggactgc ttacgctggc gaaaggtaaa ccaatcggtg aggaaggtta ctactggctg	1380
aaaatccacg gtgcaaactg cgcggtgtc gataagggtc cgttccctga gcgcatcaag	1440
ttcattgagg aaaccacga gaacatcatg gcttgcgcta agtctccact ggagaacact	1500
tggtgggctg agcaagattc tccgttctgc ttccttgctg tctgctttga gtacgctggg	1560

```

gtacagcacc acggcctgag ctataactgc tcccttccgc tggcgtttga cgggtcttgc 1620
tctggcatcc agcacttctc cgcgatgctc cgagatgagg taggtggtcg cgcggttaac 1680
ttgcttccta gtgaaaccgt tcaggacatc tacgggattg ttgctaagaa agtcaacgag 1740
attctacaag cagacgcaat caatgggacc gataacgaag tagttaccgt gaccgatgag 1800
aacactgggtg aaatctctga gaaagtcaag ctgggcacta aggcactggc tggtaaatgg 1660
ctggcttacg gtgttactcg cagtgtgact aagcgttcag tcatgacgct ggcttacggg 1920
tccaaagagt tcggcttccg tcaacaagtg ctggaagata ccattcagcc agctattgat 1980
tccggcaagg gtctgatgtt cactcagccg aatcaggctg ctggatacat ggctaagctg 2040
atttggaat ctgtgagcgt gacggtggtg gctgcggttg aagcaatgaa ctggcttaag 2100
tctgctgcta agctgctggc tgctgaggtc aaagataaga agactggaga gattcttcgc 2160
aagcgttgcg ctgtgcattg ggtaactcct gatggtttcc ctgtgtggca ggaatacaag 2220
aagcctattc agacgcgctt gaacctgatg ttcctcggtc agttccgctt acagcctacc 2280
attaacacca acaaagatag cgagattgat gcacacaaac aggagtctgg tatcgctcct 2340
aactttgtac acagccaaga cggtagccac cttcgtaaga ctgtagtgtg ggcacacga 2400
aagtacggaa tcgaatcttt tgcaactgatt cagcactcct tcggtaccat tccggctgac 2460
gctgcgaacc tgttcaaagc agtgcgcgaa actatggttg acacatatga gtcttgtgat 2520
gtactggctg atttctacga ccagttcgct gaccagttgc acgagtctca attggacaaa 2580
atgccagcac ttccggctaa aggtaacttg aacctccgtg acatcttaga gtcggacttc 2640
gcgttgcggt aa 2652

```

<210> 121

<211> 883

<212> PRT

<213> Sequência Artificial

<220>

<223> sequência de T7 RNA polimerase do tipo selvagem, quimicamente sintetizada

<400> 221

Met Asn Thr Ile Asn Ile Ala Lys Asn Asp Phe Ser Asp Ile Glu Leu

1 5 10 15

Ala Ala Ile Pro Phe Asn Thr Leu Ala Asp His Tyr Gly Glu Arg Leu

20 25 30

Ala Arg Glu Gln Leu Ala Leu Glu His Glu Ser Tyr Glu Met Gly Glu

35 40 45

Ala Arg Phe Arg Lys Met Phe Glu Arg Gln Leu Lys Ala Gly Glu Val

50 55 60

Ala Asp Asn Ala Ala Ala Lys Pro Leu Ile Thr Thr Leu Leu Pro Lys

65 70 75 80

Met Ile Ala Arg Ile Asn Asp Trp Phe Glu Glu Val Lys Ala Lys Arg

85 90 95

Gly Lys Arg Pro Thr Ala Phe Gln Phe Leu Gln Glu Ile Lys Pro Glu

100 105 110

Ala Val Ala Tyr Ile Thr Ile Lys Thr Thr Leu Ala Cys Leu Thr Ser

115 120 125

Ala Asp Asn Thr Thr Val Gln Ala Val Ala Ser Ala Ile Gly Arg Ala

130 135 140

Ile Glu Asp Glu Ala Arg Phe Gly Arg Ile Arg Asp Leu Glu Ala Lys

145 150 155 160

His Phe Lys Lys Asn Val Glu Glu Gln Leu Asn Lys Arg Val Gly His

165 170 175

Val Tyr Lys Lys Ala Phe Met Gln Val Val Glu Ala Asp Met Leu Ser

180 185 190

Lys Gly Leu Leu Gly Gly Glu Ala Trp Ser Ser Trp His Lys Glu Asp

195 200 205

Ser Ile His Val Gly Val Arg Cys Ile Glu Met Leu Ile Glu Ser Thr
 210 215 220
 Gly Met Val Ser Leu His Arg Gln Asn Ala Gly Val Val Gly Gln Asp
 225 230 235 240
 Ser Glu Thr Ile Glu Leu Ala Pro Glu Tyr Ala Glu Ala Ile Ala Thr
 245 250 255
 Arg Ala Gly Ala Leu Ala Gly Ile Ser Pro Met Phe Gln Pro Cys Val
 260 265 270
 Val Pro Pro Lys Pro Trp Thr Gly Ile Thr Gly Gly Gly Tyr Trp Ala
 275 280 285
 Asn Gly Arg Arg Pro Leu Ala Leu Val Arg Thr His Ser Lys Lys Ala
 290 295 300
 Leu Met Arg Tyr Glu Asp Val Tyr Met Pro Glu Val Tyr Lys Ala Ile
 305 310 315 320
 Asn Ile Ala Gln Asn Thr Ala Trp Lys Ile Asn Lys Lys Val Leu Ala
 325 330 335
 Val Ala Asn Val Ile Thr Lys Trp Lys His Cys Pro Val Glu Asp Ile
 340 345 350
 Pro Ala Ile Glu Arg Glu Glu Leu Pro Met Lys Pro Glu Asp Ile Asp
 355 360 365
 Met Asn Pro Glu Ala Leu Thr Ala Trp Lys Arg Ala Ala Ala Val
 370 375 380
 Tyr Arg Lys Asp Lys Ala Arg Lys Ser Arg Arg Ile Ser Leu Glu Phe
 385 390 395 400
 Met Leu Glu Gln Ala Asn Lys Phe Ala Asn His Lys Ala Ile Trp Phe
 405 410 415
 Pro Tyr Asn Met Asp Trp Arg Gly Arg Val Tyr Ala Val Ser Met Phe
 420 425 430

Asn Pro Gln Gly Asn Asp Met Thr Lys Gly Leu Leu Thr Leu Ala Lys
 435 440 445
 Gly Lys Pro Ile Gly Lys Glu Gly Tyr Tyr Trp Leu Lys Ile His Gly
 450 455 460
 Ala Asn Cys Ala Gly Val Asp Lys Val Pro Phe Pro Glu Arg Ile Lys
 465 470 475 480
 Phe Ile Glu Glu Asn His Glu Asn Ile Met Ala Cys Ala Lys Ser Pro
 485 490 495
 Leu Glu Asn Thr Trp Trp Ala Glu Gln Asp Ser Pro Phe Cys Phe Leu
 500 505 510
 Ala Phe Cys Phe Glu Tyr Ala Gly Val Gln His His Gly Leu Ser Tyr
 515 520 525
 Asn Cys Ser Leu Pro Leu Ala Phe Asp Gly Ser Cys Ser Gly Ile Gln
 530 535 540
 His Phe Ser Ala Met Leu Arg Asp Glu Val Gly Gly Arg Ala Val Asn
 545 550 555 560
 Leu Leu Pro Ser Glu Thr Val Gln Asp Ile Tyr Gly Ile Val Ala Lys
 565 570 575
 Lys Val Asn Glu Ile Leu Gln Ala Asp Ala Ile Asn Gly Thr Asp Asn
 580 585 590
 Glu Val Val Thr Val Thr Asp Glu Asn Thr Gly Glu Ile Ser Glu Lys
 595 600 605
 Val Lys Leu Gly Thr Lys Ala Leu Ala Gly Gln Trp Leu Ala Tyr Gly
 610 615 620
 Val Thr Arg Ser Val Thr Lys Arg Ser Val Met Thr Leu Ala Tyr Gly
 625 630 635 640
 Ser Lys Glu Phe Gly Phe Arg Gln Gln Val Leu Glu Asp Thr Ile Gln
 645 650 655

Pro Ala Ile Asp Ser Gly Lys Gly Leu Met Phe Thr Gln Pro Asn Gln
 660 665 670
 Ala Ala Gly Tyr Met Ala Lys Leu Ile Trp Glu Ser Val Ser Val Thr
 675 680 685
 Val Val Ala Ala Val Glu Ala Met Asn Trp Leu Lys Ser Ala Ala Lys
 690 695 700
 Leu Leu Ala Ala Glu Val Lys Asp Lys Lys Thr Gly Glu Ile Leu Arg
 705 710 715 720
 Lys Arg Cys Ala Val His Trp Val Thr Pro Asp Gly Phe Pro Val Trp
 725 730 735
 Gln Glu Tyr Lys Lys Pro Ile Gln Thr Arg Leu Asn Leu Met Phe Leu
 740 745 750
 Gly Gln Phe Arg Leu Gln Pro Thr Ile Asn Thr Asn Lys Asp Ser Glu
 755 760 765
 Ile Asp Ala His Lys Gln Glu Ser Gly Ile Ala Pro Asn Phe Val His
 770 775 780
 Ser Gln Asp Gly Ser His Leu Arg Lys Thr Val Val Trp Ala His Glu
 785 790 795 800
 Lys Tyr Gly Ile Glu Ser Phe Ala Leu Ile His Asp Ser Phe Gly Thr
 805 810 815
 Ile Pro Ala Asp Ala Ala Asn Leu Phe Lys Ala Val Arg Glu Thr Met
 820 825 830
 Val Asp Thr Tyr Glu Ser Cys Asp Val Leu Ala Asp Phe Tyr Asp Gln
 835 840 845
 Phe Ala Asp Gln Leu His Glu Ser Gln Leu Asp Lys Met Pro Ala Leu
 850 855 860
 Pro Ala Lys Gly Asn Leu Asn Leu Arg Asp Ile Leu Glu Ser Asp Phe
 865 870 875 880

Ala Phe Ala

<210> 122

<211> 2652

<212> DNA

<213> Sequência Artificial

<220>

<223> sequência de T7 RNA polimerase mutante, quimicamente sintetizada,
Y639L/H784A

<400> 122

```

atgaacacga ttaacatcgc taagaacgac ttctctgaca tcgaactggc tgctatcccg      60
ttcaacactc tggctgacca ttacggtgag cgtttagctc gcgaacagtt ggcccttgag      120
catgagtctt acgagatggg tgaagcacgc ttccgcaaga tgtttgagcg tcaacttaaa      180
gctggtgagg ttgcgataa cgctgccgcc aagcctctca tcactaccct actccctaag      240
atgattgcac gcatcaacga ctggtttgag gaagtgaaag ctaagcgcg caagcgcccc      300
acagccttcc agttcctgca agaaatcaag ccggaagccg tagcgtacat caccattaag      360
accactctgg cttgcctaac cagtgctgac aatacaaccg ttcaggctgt agcaagcgca      420
atcggtcggg ccattgagga cgaggctcgc ttcggtcgta tccgtgacct tgaagctaag      480
cacttcaaga aaaacgttga ggaacaactc aacaagcgcg tagggcacgt ctacaagaaa      540
gcatttatgc aagttgtcga ggctgacatg ctctctaagg gtctactcgg tggcgaggcg      600
tggtcttcgt ggcataagga agactctatt catgtaggag tacgctgcat cgagatgctc      660
attgagtcaa ccggaatggt tagcttacac cgccaaaatg ctggcgtagt aggtcaagac      720
tctgagacta tcgaactcgc acctgaatac gctgaggcta tcgcaacccg tgcagggtgcg      780
ctggctggca tctctccgat gttccaacct tgcgtagttc ctcctaagcc gtggactggc      840
attactggtg gtggctattg ggctaacggt cgtcgtcctc tggcgctggt gcgtactcac      900
agtaagaaag cactgatgcg ctacgaagac gtttacatgc ctgagggtgta caaagcgatt      960
aacattgcgc aaaacaccgc atggaaaatc aacaagaaag tcctagcggt cgccaacgta     1020

```

atcaccaagt ggaagcattg tccggtcgag gacatccctg cgattgagcg tgaagaactc	1080
ccgatgaaac cggaagacat cgacatgaat cctgaggctc tcaccgcgtg gaaacgtgct	1140
gccgctgctg tgtaccgcaa ggacaaggct cgcaagtctc gccgtatcag ccttgagttc	1200
atgcttgagc aagccaataa gtttgctaac cataaggcca tctggttccc ttacaacatg	1260
gactggcgcg gtcgtgttta cgctgtgtca atgttcaacc cgcaaggtaa cgatatgacc	1320
aaaggactgc ttacgctggc gaaaggtaaa ccaatcggtg aggaaggtta ctactggctg	1380
aaaatccacg gtgcaaacgt tgcgggtgtc gataaggttc cgttccctga ggcgatcaag	1440
ttcattgagg aaaaccacga gaacatcatg gcttgcgcta agtctccact ggagaacact	1500
tggtgggctg agcaagattc tccgttctgc ttccttgctg tctgctttga gtacgctggg	1560
gtacagcacc acggcctgag ctataactgc tcccttccgc tggcgtttga cgggtcttgc	1620
tctggcatCc agcacttctc cgcgatgctc cgagatgagg taggtggctg cgcggttaac	1680
ttgcttccta gtgaaaccgt tcaggacatc tacgggattg ttgctaagaa agtcaacgag	1740
attctacaag cagacgcaat caatgggacc gataacgaag tagttaccgt gaccgatgag	1800
aacactggtg aaatctctga gaaagtcaag ctgggcacta aggcactggc tggccaatgg	1860
ctggccttacg gtgttactcg cagtgtgact aagcgttcag tcatgacgct ggctctgggg	1920
tccaaagagt tgggcttccg tcaacaagtg ctggaagata ccattcagcc agctattgat	1980
tccggcaagg gtctgatgtt cactcagccg aatcaggctg ctggatacat ggctaagctg	2040
atttggaat ctgtgagcgt gacggtggta gctgcggttg aagcaatgaa ctggcttaag	2100
tctgctgcta agctgctggc tgctgaggct aaagataaga agactggaga gattcttcgc	2160
aagcgttgcg ctgtgcattg ggtaactcct gatggtttcc ctgtgtggca ggaatacaag	2220
aagcctattc agacgcgctt gaacctgatg ttctcggctc agctccgctt acagcctacc	2280
attaacacca acaaagatag cgagattgat gcacacaaac aggagtctgg tatcgctcct	2340
aactttgtag ccagccaaga cggtagccac cttcgtaaga ctgtagtgtg ggcacacgag	2400
aagtacggaa tcgaatcttt tgcactgatt cacgactcct tcggtaccat tccggctgac	2460
gctgcgaacc tgttcaaagc agtgcgcgaa actatggttg acacatatga gtcttgtgat	2520
gtactggctg atttctacga ccagttcgct gaccagttgc acgagtctca attggacaaa	2580
atgccagcac ttccggctaa aggtaacttg aacctccgtg acatcttaga gtcggacttc	2640
gcgttcgcgt aa	2652

<210> 123

<211> 2652

<212> DNA

<213> Sequência Artificial

<220>

<223> sequência de T7 RNA polimerase mutante, quimicamente sintetizada,
Y639L/I-1784A/K378R

<400> 123

```

atgaacacga ttaacatcgc taagaacgac ttctctgaca tcgaactggc tgctatcccg 60
ttcaaacactc tggctgacca ttacggtgag cgtttagctc gcgaacagtt ggcccttgag 120
catgagtctt acgagatggg tgaagcacgc ttccgcaaga tgtttgagcg tcaacttaaa 180
gctggtgagg ttgcgataa cgctgccgcc aagcctctca tcactaccct actccctaag 240
atgattgcac gcatcaacga ctggtttgag gaagtgaag ctaagcgcg caagcgcccc 300
acagccttcc agttcctgca agaaatcaag ccggaagccg tagcgtacat caccattaag 360
accactctgg cttgcctaac cagtgctgac aatacaaccg ttcaggctgt agcaagcgca 420
atcggtcggg ccattgagga cgaggctcgc ttcggtcgta tccgtgacct tgaagctaag 480
cacttcaaga aaaacgttga ggaacaactc aacaagcgcg tagggcacgt ctacaagaaa 540
gcatttatgc aagttgtcga ggctgacatg ctctctaagg gtctactcgg tggcgaggcg 600
tggtcttcgt ggcataagga agactctatt catgtaggag tacgctgcat cgagatgctc 660
attgagtcaa ccggaatggt tagcttacac cgccaaaatg ctggcgtagt aggtcaagac 720
tctgagacta tcgaactcgc acctgaatac gctgaggcta tcgcaaccgg tgcaggtgcg 780
ctggctggca tctctccgat gttccaacct tgcgtagttc ctccctaagcc gtggactggc 840
attactggtg gtggctattg ggctaacggt cgtcgtcttc tggcgctggt gcgtactcac 900
agtaagaaag cactgatgcg ctacgaagac gtttacatgc ctgaggtgta caaagcgatt 960
aacattgcgc aaaacaccgc atggaaaatc aacaagaaag tcctagcggg cgccaacgta 1020

```

atcaccaagt ggaagcattg tccggctcgag gacatccctg cgattgagcg tgaagaactc	1080
ccgatgaaac cggaagacat cgacatgaat cctgaggctc tcaccgctg gagacgtgct	1140
gccgctgctg tgtaccgcaa ggacaaggct cgcaagtctc gccgtatcag ccttgagttc	1200
atgcttgagc aagccaataa gtttgctaac cataaggcca tctggttccc ttacaacatg	1260
gactggcgcg gtcgtgttta cgctgtgtca atgttcaacc cgcaaggtaa cgatatgacc	1320
aaaggactgc ttacgctggc gaaaggtaaa ccaatcggtg aggaaggtta ctactggctg	1380
aaaatccacg gtgcaaactg Cgcgggtgtc gataaggttc cgttccctga gcgcatacaag	1440
ttcattgagg aaaaccacga gaacatcatg gcttgcgcta agtctccaact ggagaacact	1500
tggtgggctg agcaagattc tccgttctgc ttccttgctg tctgctttga gtacgctggg	1560
gtacagcacc acggcctgag ctataactgc tcccttccgc tggcgtttga cgggtcttgc	1620
tctggcatcc agcacttctc cgcgatgctc cgagatgagg taggtggctg cgcggttaac	1680
ttgcttcccta gtgaaaccgt tcaggacatc tacgggattg ttgctaagaa agtcaacgag	1740
attctacaag cagacgcaat caatgggacc gataacgaag tagttaccgt gaccgatgag	1800
aacactgggtg aaatctctga gaaagtcaag ctgggcacta aggcactggc tggtaaatgg	1860
ctggcttacg gtgttactcg cagtgtgact aagcgttcag tcatgacgct ggctctgggg	1920
tccaaagagt tccgcttccg tcaacaagtg ctggaagata ccattcagcc agctattgat	1980
tccggcaagg gtctgatgtt cactcagccg aatcaggctg ctggatacat ggctaagctg	2040
atttggaat ctgtgagcgt gacggtggta gctgcggttg aagcaatgaa ctggcttaag	2100
tctgtgcta agctgctggc tgctgaggct aaagataaga agactggaga gattcttcgc	2160
aagcgttgctg ctgtgcattg ggtaactcct gatggtttcc ctgtgtggca ggaatacaag	2220
aagcctattc agacgcgctt gaacctgatg ttctcggctc agttccgctt acagcctacc	2280
attaacacca acaaagatag cgagattgat gcacacaaac aggagtctgg tatcgctcct	2340
aactttgtag ccagccaaga cggtagccac cttcgtaaga ctgtagtggtg ggcacacgag	2400
aagtacggaa tcgaatcttt tgcactgatt cacgactcct tcggtaccat tccggctgac	2460
gctgcgaacc tgttcaaagc agtgcgcgaa actatggttg acacatatga gtcttgat	2520
gtactggctg atttctacga ccagttcgt gaccagttgc acgagtctca attggacaaa	2580
atgccagcac ttccggctaa aggtaacttg aacctccgtg acatcttaga gtcggacttc	2640
gcgttcgctg aa	2652

<210> 124

<211> 2652

<212> DNA

<213> Sequência Artificial

<220>

<223> sequência de T7 RNA polimerase mutante, quimicamente sintetizada,
P266L/Y639L/H784A

<400> 124

```

atgaacacga ttaacatcgc taagaacgac ttctctgaca tcgaactggc tgctatcccg      60
ttcaacactc tggctgacca ttacggtgag cgtttagctc gcgaacagtt ggcccttgag      120
catgagtctt acgagatggg tgaagcacgc ttccgcaaga tgtttgagcg tcaacttaaa      180
gctggtgagg ttgcgataa cgctgccgcc aagcctctca tcactaccct actccctaag      240
atgattgcac gcatcaacga ctggtttgag gaagtgaag ctaagcgcg caagcgcccc      300
acagccttcc agttcctgca agaaatcaag ccggaagccg tagcgtacat caccattaag      360
accactctgg cttgcctaac cagtgctgac aatacaaccg ttcaggctgt agcaagcgca      420
atcggtcggg ccattgagga cgaggctcgc ttcggtcgta tccgtgacct tgaagcctaag      480
cacttcaaga aaaacgttga ggaacaactc aacaagcgcg tagggcacgt ctacaagaaa      540
gcatttatgc aagttgtcga ggctgacatg ctctctaagg gtctactcgg tggcgaggcg      600
tggtcttcgt ggcataagga agactctatt catgtaggag tacgctgcat cgagatgctc      660
attgagtcaa ccggaatggt tagcttacac cgccaaaatg ctggcgtagt aggtcaagac      720
tctgagacta tcgaactcgc acctgaatac gctgaggcta tcgcaaccgg tgcaggtgcg      780
ctggctggca tctctctgat gttccaacct tgcgtagttc ctctaagcc gtggactggc      840
attactggtg gtggctattg ggctaacggt cgtcgtcttc tggcgctggt gcgtactcac      900
agtaagaaaag cactgatgcg ctacgaagac gtttacatgc ctgaggtgta caaagcgatt      960
aacattgcgc aaaacaccgc atggaaaatc aacaagaaag tcctagcggg cgccaacgta     1020

```

atcaccaagt ggaagcattg tccggtcgag gacatccctg cgattgagcg tgaagaactc	1080
ccgatgaaac cggaagacat cgacatgaat cctgaggctc tcaccgcgtg gaaacgtgct	1140
gccgctgctg tgtaccgcaa ggacaaggct cgcaagtctc gccgtatcag ccttgagttc	1200
atgcttgagc aagccaataa gtttgctaac cataaggcca tctggttccc ttacaacatg	1260
gactggcgcg gtcgtgttta cgctgtgtca atgttcaacc cgcaaggtaa cgatatgacc	1320
aaaggactgc ttacgctggc gaaaggtaaa ccaatcggta aggaaggtta ctactggctg	1380
aaaatccacg gtgcaaactg tgcgggtgtc gataaggttc cgttcctga gcgcatcaag	1440
ttcattgagg aaaaccacga gaacatcatg gcttgcgcta agtctccact ggagaacact	1500
tggtgggctg agcaagattc tccgttctgc ttccttgcgt tctgctttga gtacgctggg	1560
gtacagcacc acggcctgag ctataactgc tcccttcgcg tggcgtttga cgggtcttgc	1620
tctggcatcc agcacttctc cgcgatgctc cgagatgagg taggtggctg cgcggttaac	1680
ttgcttccta gtgaaaccgt tcaggacatc tacgggattg ttgctaagaa agtcaacgag	1740
attctacaag cagacgcaat caatgggacc gataacgaag tagttaccgt gaccgatgag	1800
aacactggtg aaatctctga gaaagtcaag ctgggcacta aggcactggc tggccaatgg	1860
ctggcttacg gtgttactcg cagtgtgact aagcgttcag tcatgacgct ggctctgggg	1920
tccaaagagt tcggcttccg tcaacaagtg ctggaagata ccattcagcc agctattgat	1980
tccggcaagg gtctgatgtt cactcagccg aatcaggctg ctggatacat ggctaagctg	2040
at ttgggaat ctgtgagcgt gacggtggta gctgcggtt gaaagcaatgaa ctggcttaag	2100
tctgctgcta agctgctggc tgctgaggtc aaagataaga agactggaga gattcttcgc	2160
aagcgttgcg ctgtgcattg ggtaactcct gatggtttcc ctgtgtggca ggaatacaag	2220
aagcctattc agacgcgctt gaacctgatg ttcctcggtc agttccgctt acagcctacc	2280
attaacacca acaaagatag cgagattgat gcacacaaac aggagtctgg tatcgctcct	2340
aactttgtag ccagccaaga cggtagccac cttcgtaaga ctgtagtgtg ggcacacgag	2400
aagtacggaa tcgaatcttt tgcactgatt cacgactcct tcggtaccat tccggctgac	2460
gctgcgaacc tgttcaaagc agtgccgcaa actatggttg acacatatga gtcttgat	2520
gtactggctg atttctacga ccagttcgt gaccagttgc acgagtctca attggacaaa	2580
atgccagcac ttccggctaa aggttaacttg aacctccgtg acatcttaga gtcggacttc	2640
gcgttcgctg aa	2652

<210> 125

<211> 2652

<212> DNA

<213> Sequência Artificial

<220>

<223> sequência de T7 RNA polimerase mutante, quimicamente sintetizada,
P266L/Y639L/H784A/K378R

<400> 125

```

atgaacacga ttaacatcgc taagaacgac ttctctgaca tcgaactggc tgctatcccg      60
ttcaaacactc tggctgacca ttacggtgag cgtttagctc gcgaacagtt ggcccttgag      120
catgagtctt acgagatggg tgaagcacgc ttccgcaaga tgtttgagcg tcaacttaaa      180
gctggtgagg ttgcggataa cgctgccgcc aagcctctca tcactaccct actccctaag      240
atgattgcac gcatcaacga ctggtttgag gaagtgaag ctaagcgcg caagcgccc      300
acagccttcc agttcctgca agaaatcaag ccggaagccg tagcgtagat caccattaag      360
accactctgg cttgcctaac cagtgcctgac aatacaaccg ttcaggctgt agcaagcgca      420
atcggtcggg ccattgagga cgaggctcgc ttcggtcgta tcggtgacct tgaagctaag      480
cacttcaaga aaaacgttga ggaacaactc aacaagcgcg tagggcacgt ctacaagaaa      540
gcatttatgc aagttgtcga ggctgacatg ctctctaagg gtctactcgg tggcgaggcg      600
tggtcttcgt ggcataagga agactctatt catgtaggag tacgctgcat cgagatgctc      660
attgagtcaa ccggaatggt tagcttacac cgccaaaatg ctggcgtagt aggtcaagac      720
tctgagacta tcgaactcgc acctgaatac gctgaggcta tcgcaaccgg tgcaggtgcg      780
ctggctggca tctctctgat gttccaacct tgcgtagttc ctcttaagcc gtggactggc      840
attactggtg gtggctattg ggctaaCggt cgtcgtcctc tggcgctggt gcgtactcac      900
agtaagaaag cactgatgcg ctacgaagac gtttacatgc ctgaggtgta caaagcgatt      960
aacattgcgc aaaacaccgc atggaaaatc aacaagaaag tcctagcggg cgccaacgta     1020

```

atcaccaagt ggaagcattg tccggctcgag gacatccctg cgattgagcg tgaagaactc	1080
ccgatgaaac cggaagacat cgacatgaat cctgaggctc tcaccgcgtg gagacgtgct	1140
gccgctgctg tgtaccgcaa ggacaaggct cgcaagtctc gccgtatcag ccttgagttc	1200
atgcttgagc aagccaataa gtttgctaac cataaggcca tctggttccc ttacaacatg	1260
gactggcgcg gtcgtgttta cgctgtgtca atgttcaacc cgcaaggtaa cgatatgacc	1320
aaaggactgc ttacgctggc gaaaggtaaa ccaatcggta aggaaggtta ctactggctg	1380
aaaatccacg gtgcaaactg tgcgggtgtc gataaggttc cgttccctga gcgcatcaag	1440
ttcattgagg aaaaccacga gaacatcatg gcttgcgcta agtctccact ggagaacact	1500
tggtgggctg agcaagattc tccgttctgc ttccttgctg tctgctttga gtacgctggg	1560
gtacagcacc acggcctgag ctataactgc tcccttccgc tggcgtttga cgggtcttgc	1620
tctggcatcc agcacttctc cgcgatgctc cgagatgagg taggtggctg cgcggttaac	1680
ttgcttcccta gtgaaaCCgt tcaggaCata tacgggattg ttgctaagaa agtcaacgag	1740
attctacaag cagacgcaat caatgggacc gataacgaag tagttaccgt gaccgatgag	1800
aacactgggtg aaatctctga gaaagtcaag ctgggcacta aggactggc tggtaaatgg	1860
ctggcttacg gtgttactcg cagtgtgact aagcgttcag tcatgacgct ggctctgggg	1920
tccaaagagt tcggcttccg tcaacaagtg ctggaagata ccattcagcc agctattgat	1980
tccggcaagg gtctgatgtt cactcagccg aatcaggctg ctggatacat ggctaagctg	2040
atttggaat ctgtgagcgt gacggtggtg gctgcggttg aagcaatgaa ctggcttaag	2100
tctgtgcta agctgctggc tgctgaggtc aaaga ¹ aaga agactggaga gattcttcgc	2160
aagcgttgcg ctgtgcattg ggtaactcct gatggtttcc ctgtgtggca ggaatacaag	2220
aagcctattc agacgcgctt gaacctgatg ttcctcggtc agttccgctt acagcctacc	2280
attaacacca acaaagatag cgagattgat gcacacaaac aggagtctgg tatcgctcct	2340
aactttgtag ccagccaaga cggtagccac cttcgtaaga ctgtagtgtg ggcacacgag	2400
aagtacggaa tcgaatcttt tgcactgatt cacgactcct tcggtaccat tccggtgac	2460
gctgcgaacc tgttcaaagc agtgccgcaa actatgggtg acacatatga gtcttgat	2520
gtactggctg atttctacga ccagttcgct gaccagttgc acgagtctca attggacaaa	2580
atgccagcac ttccggctaa aggttaacttg aacctcgtg acatcttaga gtcggacttc	2640
gcgttcgctg aa	2652

<210> 126

<211> 89

<212> DNA

<213> Sequência Artificial

<220>

<223> molde de transcrição de DNA, quimicamente sintetizado

<220>

<221> característica diversa

<222> (24)..(63)

<223> n é a, t, c ou g

<400> 126

gggagaattc cgaccagaag cttnnnnnnnn nnnnnnnnnnn nnnnnnnnnnn nnnnnnnnnnn 60

nnncatatgt gcgtctacat ggatcctca 89

<210> 127

<211> 89

<212> DNA

<213> Sequência Artificial

<220>

<223> molde de transcrição de DNA, quimicamente sintetizado

<220>

<221> característica diversa

<222> (27)..(66)

<223> n é a, t, c ou g

<400> 127

gggagagcgg aagccgtgct ggggccnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 60

nnnnnncata acccagaggt cgatggatc 89

<210> 128

<211> 94

<212> DNA

<213> Sequência Artificial

<220>

<223> molde de transcrição de DNA, quimicamente sintetizado

<220>

<221> característica diversa

<222> (21)..(60)

<223> n é a, t, c ou g

<400> 128

gggagagaca agcttgggtc nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 60

agaagagaaa gagaagttaa ttaaggatcc tcag 94

<210> 129

<211> 86

<212> DNA

<213> Sequência Artificial

<220>

<223> molde de transcrição de DNA, quimicamente sintetizado

<220>

<221> característica diversa

<222> (21)..(60)

<223> n é a, t, c ou g

<400> 129

gggagaattc cgaccacaag nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 60

catatgtgcg tctacatgga tcctca 86

<210> 130

<211> 40

<212> DNA

<213> Sequência Artificial

<220>

<223> sequência de primer em 5', quimicamente sintetizada

<400> 130

taatacgact cactataggg agaattccga ccagaagctt 40

<210> 131

<211> 26

<212> DNA

<213> Sequência Artificial

<220>

<223> sequência de primer em 3', quimicamente sintetizada

<400> 131

tgaggatcca ttagacgca catatg

26

<210> 132

<211> 23

<212> DNA

<213> Sequência Artificial

<220>

<223> sequência de primer em 3', quimicamente sintetizada

<400> 132

gatccatoga cctctgggtt atg

23

<210> 133

<211> 37

<212> DNA

<213> Sequência Artificial

<220>

<223> sequência de primer 5', quimicamente sintetizada

<400> 133

taatacgact cactataggg agagacaagc ttgggtc

37

<210> 134

<211> 34

<212> DNA

<213> Sequência Artificial

<220>

<223> sequência de primer 3', quimicamente sintetizada

<400> 134

ctgaggatcc ttaattaact tctctttctc ttct

34

<210> 135

<211> 37

<212> DNA

<213> Sequência Artificial

<220>

<223> sequência de primer 5', quimicamente sintetizada

<400> 135

taatacgact cactataggg agaattccga ccacaag

37

<210> 136

<211> 10

<212> DNA

<213> Sequência Artificial

<220>

<223> motivo imunoestimulador quimicamente sintetizado

<400> 136

aacgttcgag

10

<210> 137

<211> 76

<212> DNA

<213> Sequência Artificial

<220>

<223> molde de DNA, quimicamente sintetizado, ARC3428

<220>

<221> característica diversa

<222> (24)..(53)

<223> n é a, t, c, ou g

<400> 137

gggagacaag aataaagcga gttnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnaagagtc 60

gatgatgctt agctag 76

<210> 138

<211> 77

<212> DNA

<213> Sequência Artificial

<220>

<223> molde de transcrição de DNA, quimicamente sintetizado

<220>

<221> característica diversa

<222> (24)..(53)

<223> n é a, c, g, ou t

<400> 138

gggccttgta gcgtgcattc ttgnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnctaacat 60

actccgaatc tgtcgaa 77

<210> 139

<211> 73

<212> DNA

<213> Sequência Artificial

<220>

<223> molde de transcrição de DNA, quimicamente sintetizado

<220>

<221> característica diversa

<222> (17)..(56)

<223> n é a, c, ou t

<400> 139

ggagccttcc tccggannnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnntccg 60

gtttcccagag ctt 73

<210> 140

<211> 87

<212> DNA

<213> Sequência Artificial

<220>

<223> molde de transcrição de DNA, quimicamente sintetizado

<220>

<221> característica diversa

<222> (24)..(63)

<223> n é a, c, g, ou t

<400> 140

gggagacaag aataaacgct caannnnnnnn nnnnnnnnnnn nnnnnnnnnnn nnnnnnnnnnn 60

nnnttcgaca ggaggctcac aacaggc 87

<210> 141

<211> 24

<212> DNA

<213> Sequência Artificial

<220>

<223> sequência de primer em 3', quimicamente sintetizado

<400> 141

ttcgacagat tcggagtatg ttag

<210> 142

<211> 33

<212> DNA

<213> Sequência Artificial

<220>

<223> sequência de primer em 5', quimicamente sintetizado

<400> 142

taatacgact cactatagga gccttcctcc gga

33

<210> 143

<211> 17

<212> DNA

<213> Sequência Artificial

<220>

<223> sequência de primer em 3', quimicamente sintetizado

<400> 143

aagctcggga aaccgga

17

<210> 144

<211> 40

<212> DNA

<213> Sequência Artificial

<220>

<223> sequência de primer em 5', quimicamente sintetizado

<400> 144

taatacgact cactataggg agacaagaat aaacgctcaa

40

<210> 145

<211> 24

<212> DNA

<213> Sequência Artificial

<220>

<223> sequência de primer em 3', quimicamente sintetizado

<400> 145

gcctgttggtg agcctcctgt cgaa

24

<210> 146

<211> 39

<212> DNA

<213> Sequência Artificial

<220>

<223> sequência de primer em 5', quimicamente sintetizado

<400> 146

taatacgact cactataggg gagtacaata accagacat

39

<210> 147

<211> 24

<212> DNA

<213> Sequência Artificial

<220>

<223> sequência de primer em 3', quimicamente sintetizado

<400> 147

catcgatgct agtcgtaacg atcc

24

<210> 148

<211> 26

<212> DNA

<213> Sequência Artificial

<220>

<223> sequência de primer em 3', quimicamente sintetizado

<400> 148

tgaggatcca tgtagacgca catatg

26

<210> 149

<211> 43

<212> DNA

<213> Sequência Artificial

<220>

<223> sequência de primer em 5', quimicamente sintetizado

<400> 149

taatacgact cactataggg agagcggaag ccgtgctggg gcc

43

<210> 150

<211> 76

<212> DNA

<213> Sequência Artificial

<220>

<223> molde de transcrição de DNA, quimicamente sintetizado

<220>

<221> característica diversa

<222> (23)..(52)

<223> n é a, c, g, ou t

<400> 150

ggggagtaca ataaccagac atnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnggatcggt

60

acgactagca tcgatg

76

<210> 151

<211> 40

<212> DNA

<213> Sequência Artificial

<220>

<223> sequência de primer em 5', quimicamente sintetizado

<400> 151

taatacgact cactataggg ccttgtagcg tgcattcttg

40

REIVINDICAÇÕES

1. T7 RNA polimerase isolada, **CARACTERIZADA** por compreender um aminoácido alterado na posição 639 e na posição 784, onde o aminoácido alterado na posição 639 não é uma fenilalanina quando o aminoácido alterado na posição 784 for uma alanina.

2. T7 RNA polimerase isolada, de acordo com a reivindicação 1, **CARACTERIZADA** por adicionalmente compreender um aminoácido alterado na posição 378.

10 3. T7 RNA polimerase isolada, de acordo com quaisquer das reivindicações acima mencionadas, **CARACTERIZADA** por adicionalmente compreender um aminoácido alterado na posição 266.

15 4. T7 RNA polimerase isolada, de acordo com quaisquer das reivindicações acima mencionadas, **CARACTERIZADA** pelo fato de que o aminoácido alterado na posição 639 é uma leucina e o aminoácido alterado na posição 784 é uma alanina.

20 5. T7 RNA polimerase isolada, de acordo com quaisquer das reivindicações acima mencionadas, **CARACTERIZADA** pelo fato de que o aminoácido alterado na posição 266 é uma leucina.

25 6. T7 RNA polimerase isolada, de acordo com quaisquer das reivindicações acima mencionadas, **CARACTERIZADA** pelo fato de que o aminoácido alterado na posição 378 é uma arginina.

7. T7 RNA polimerase isolada, de acordo com quaisquer das reivindicações acima mencionadas, **CARACTERIZADA** pelo fato de que os aminoácidos alterados aumentam a produção

transcricional de ácidos nucleicos que compreendem modificações de 2'-OMe pela polimerase em uma reação de transcrição compreendendo somente trifosfatos de 2'-OMe nucleotídeos.

8. T7 RNA polimerase isolada, de acordo com a reivindicação 8, **CARACTERIZADA** pelo fato de que o aumento na produção de transcrição é relativo a uma T7 RNA polimerase não tendo os aminoácidos alterados quando a transcrição for efetuada sob condições de transcrição idênticas.

9. T7 RNA polimerase isolada, de acordo com quaisquer das reivindicações acima mencionadas, **CARACTERIZADA** pelo fato de que os aminoácidos alterados diminuem a diferenciação contra os trifosfatos de 2'-OMe nucleotídeos.

10. T7 RNA polimerase isolada, de acordo com a reivindicação 9, **CARACTERIZADA** pelo fato de que a diferenciação diminuída contra os trifosfatos de 2'-OMe nucleotídeos é em relação a uma T7 RNA polimerase não tendo os aminoácidos alterados.

11. T7 RNA polimerase isolada, de acordo com a reivindicação 8 ou 10, **CARACTERIZADA** pelo fato de que a T7 RNA polimerase não tendo os aminoácidos alterados é a T7 RNA polimerase do tipo selvagem compreendendo um aminoácido na posição 639 alterado para uma fenilalanina e um aminoácido na posição 784 alterado para alanina.

12. Polipeptídeo isolado, **CARACTERIZADO** por compreender um aminoácido selecionado a partir do grupo que consiste em: SEQ ID NO 1, SEQ ID NO 2, SEQ ID NO 102 e SEQ ID NO 103.

13. Método de transcrever um ácido nucleico de fi-

ta simples, **CARACTERIZADO** por compreender a incubação de uma polimerase de acordo com qualquer uma das reivindicações precedentes com um ácido nucléico de molde, sob condições de reação suficientes para resultar na transcrição.

5 14. Ácido nucléico isolado, **CARACTERIZADO** pelo fato de que codifica um polipeptídeo de acordo com quaisquer das reivindicações precedentes.

 15. Sequência de ácidos nucléicos isolada, **CARACTERIZADA** por ser selecionada a partir do grupo que con-
10 siste em: SEQ ID NO 122, SEQ ID NO 123, SEQ ID NO 124 e SEQ ID NO 125.

 16. Vetor, **CARACTERIZADO** por compreender uma sequência de ácidos nucléicos isolada de acordo com a reivindicação 14 ou 15.

15 17. Vetor de expressão, **CARACTERIZADO** por compreender um ácido nucléico da reivindicação 14 ou 15 operavelmente ligado a um promotor.

 18. Célula, **CARACTERIZADA** por compreender o vetor de expressão da reivindicação 17.

20 19. Célula, de acordo com a reivindicação 18, **CARACTERIZADA** pelo fato de que a T7 RNA polimerase mutante é expressa pela célula.

 20. Kit, **CARACTERIZADO** por compreender um recipiente contendo uma T7 RNA polimerase de acordo com quaisquer
25 das reivindicações precedentes.

 21. Kit, **CARACTERIZADO** por compreender um recipiente contendo um ácido nucléico codificando uma T7 RNA polimerase de acordo com quaisquer das reivindicações preceden-

tes.

22. Método de transcrever um ácido nucléico inteiramente com 2'-OMe, **CARACTERIZADO** por compreender as etapas de a) incubar um ácido nucléico de molde em uma mistura de
5 reação compreendendo uma RNA polimerase mutante, um molde de transcrição de ácido nucléico e trifosfatos de nucleosídeos, onde os trifosfatos de nucleosídeos são 2'-OMe, e

b) transcrever a mistura de reação de transcrição por um tempo suficiente para resultar no ácido nucléico de
10 fita simples, onde todos os nucleotídeos dos ácidos nucléicos de fitas simples são modificados com 2'-OMe, exceto que o primeiro nucleotídeo dos transcritos podem ser não modificados em 2'.

23. Método, de acordo com a reivindicação 22,
15 **CARACTERIZADO** pelo fato de que a RNA polimerase mutante é uma T7 RNA polimerase mutante compreendendo um aminoácido alterado na posição 639 e na posição 784.

24. Método, de acordo com a reivindicação 23,
CARACTERIZADO pelo fato de que a T7 RNA polimerase mutante é
20 uma T7 RNA polimerase mutante de acordo com qualquer uma das reivindicações 1 a 12.

25. Método, de acordo com qualquer uma das reivindicações 22 a 24, **CARACTERIZADO** pelo fato de que a reação de transcrição adicionalmente compreende íons magnésio.

26. Método, de acordo com qualquer uma das reivindicações 22 a 25, **CARACTERIZADO** pelo fato de que a reação de transcrição adicionalmente compreende íons manganês.

27. Método, de acordo com qualquer uma das reivin-

dicações 22 a 26, **CHARACTERIZADO** pelo fato de que a reação de transcrição adicionalmente compreende um resíduo que não é de trifosfato que não é de 2'-OMe guanosina.

28. Método, de acordo com qualquer uma das reivindicações 22 a 27, **CHARACTERIZADO** pelo fato de que o molde de transcrição compreende um promotor de T7 RNA polimerase.

29. Método, de acordo com as reivindicações 26 a 28, **CHARACTERIZADO** pelo fato de que os íons magnésio estão presentes na reação de transcrição em uma concentração que é entre 3,0 a 3,5 vezes maior do que os íons manganês.

30. Método, de acordo com qualquer uma das reivindicações 26 a 29, **CHARACTERIZADO** pelo fato de que cada trifosfato de nucleotídeo está presente na reação de transcrição em uma concentração de 1,0 mM, a concentração de íons magnésio é 6,5 mM, e a concentração de íons manganês é 2,0 mM.

31. Método, de acordo com qualquer uma das reivindicações 26 a 30, **CHARACTERIZADO** pelo fato de que cada trifosfato de nucleotídeo está presente na reação de transcrição em uma concentração de 1,5 mM, a concentração de íons magnésio é 8 mM, e a concentração de íons manganês é 2,5 mM.

32. Método, de acordo com qualquer uma das reivindicações 26 a 31, **CHARACTERIZADO** pelo fato de que cada trifosfato de nucleotídeo está presente na reação de transcrição em uma concentração de 2,0 mM, a concentração de íons magnésio é 9,5 mM e a concentração de íons manganês é 3,0 mM.

33. Método, de acordo com quaisquer das reivindicações 22 a 32, **CHARACTERIZADO** pelo fato de que a reação de

transcrição adicionalmente compreende o polietileno glicol.

34. Método, de acordo com quaisquer das reivindicações 22 a 33, **CARACTERIZADO** pelo fato de que o resíduo que não é de trifosfato que não é de 2'-OMe guanosina é selecionado a partir do grupo que consiste em: monofosfato de guanosina, difosfato de guanosina, monofosfato de 2'-flúor guanosina, difosfato de 2'-flúor guanosina, monofosfato de 2'-amino guanosina, difosfato de 2'-amino guanosina, monofosfato de 2'-desóxi guanosina, e difosfato de 2'-desóxi guanosina.

35. Método, de acordo com qualquer uma das reivindicações 22 a 34, **CARACTERIZADO** pelo fato de que a reação de transcrição compreende a pirofosfatase inorgânica.

36. Método para identificar aptâmeros, **CARACTERIZADO** por compreender:

a) a preparação de uma mistura de reação de transcrição compreendendo uma polimerase mutante de acordo com qualquer uma das reivindicações 1 a 12, e um ou mais moldes de transcrição de ácidos nucleicos;

b) a transcrição da mistura de reação de transcrição para resultar em uma mistura candidata de ácidos nucleicos de fitas simples, onde todos menos opcionalmente um dos nucleotídeos dos ácidos nucleicos de fitas simples são modificados em 2',

c) o contato da mistura candidata com a molécula-alvo,

d) a separação dos ácidos nucleicos que têm uma afinidade aumentada pela molécula-alvo, em relação a uma a-

finidade da mistura candidata, da mistura candidata, e

e) a amplificação dos ácidos nucléicos com afinidade aumentada para produzir uma mistura enriquecida em aptâmeros, pelo que são identificados aptâmeros para a molécula-alvo compreendendo nucleotídeo todo modificado em 2', exceto que o primeiro nucleotídeo dos aptâmeros pode ser não modificado em 2'.

37. Método, de acordo com a reivindicação 36, **CARACTERIZADO** pelo fato de que a etapa de amplificação compreende (i) dissociar opcionalmente os ácidos nucléicos com afinidade aumentada do alvo, ii) transcrever invertido os ácidos nucléicos com afinidade aumentada, dissociados dos complexos de ácidos nucléicos-alvo, iii) amplificar os ácidos nucléicos com afinidade aumentada transcritos invertidos; e (ii) preparar uma mistura de reação de transcrição compreendendo os ácidos nucléicos com afinidade aumentada transcritos invertidos amplificados como o molde de transcrição e transcrever a mistura de transcrição.

38. Método, de acordo com qualquer uma das reivindicações 36 a 37, **CARACTERIZADO** pelo fato de que todos os trifosfatos de nucleotídeos na reação de transcrição são modificados com 2'-OMe.

39. Método, de acordo com qualquer uma das reivindicações 36 a 38, **CARACTERIZADO** pelo fato de que os um ou mais moldes de transcrição de ácidos nucléicos compreendem um promotor de T7 RNA polimerase e uma sequência líder imediatamente 3' ao promotor de T7 RNA polimerase.

40. Método, de acordo com qualquer uma das reivin-

dicações 37 a 39, **CARACTERIZADO** pelo fato de que compreende repetir as etapas a) até e) iterativamente.

41. Método de identificar uma sequência de componente de molde de ácido nucléico para aumentar a transcrição,
5 **CARACTERIZADO** pelo fato de que compreende:

a) preparar uma biblioteca de candidatos de moldes de transcrição, onde os moldes compreendem um promotor, uma primeira região fixa imediatamente 3' ao promotor, uma região degenerada imediatamente 3' à primeira região fixa e uma
10 segunda região fixa 3' à região degenerada.

b) transcrever a biblioteca de candidatos de moldes de transcrição em uma reação de transcrição para resultar em uma mistura de transcrição tendo uma produção de transcrito,

15 c) transcrever invertida a mistura de transcrição para obter uma mistura candidata de moldes de cDNA, onde os moldes de cDNA compreendem uma extremidade de 5' e 3',

d) ligar uma sequência de DNA codificando o promotor à extremidade de 3' dos moldes de cDNA em uma reação de
20 ligação,

e) amplificar os moldes de cDNA para resultar em uma biblioteca de candidatos de moldes de transcrição, e

f) identificar um componente de sequência de ácidos nucléicos para aumentar a transcrição a partir da biblioteca de candidatos de moldes de transcrição, onde o componente de sequência de ácidos nucléicos compreende uma sequência derivada de pelo menos uma porção da região degenerada.
25

42. Método, de acordo com a reivindicação 45, **CARACTERIZADO** pelo fato de que a etapa f) compreende as etapas de:

i) clonar a biblioteca de candidatos de moldes de transcrição em moldes de transcrição individuais,

ii) transcrever os moldes de transcrição individuais em uma reação de transcrição para resultar em uma produção de transcrito,

iii) avaliar a produção de transcrito dos moldes de transcrição individuais, e

iv) identificar o componente de sequência de ácidos nucleicos em um molde de transcrição que resulte em uma produção de transcrito predeterminada.

43. Método, de acordo com a reivindicação 42, **CARACTERIZADO** pelo fato de que a produção de transcrito predeterminada é uma produção maior do que a produção de transcrito obtida na etapa b) por transcrição da mistura de candidatos de moldes de transcrição.

44. Método, de acordo com a reivindicação 41, **CARACTERIZADO** pelo fato de que a etapa f) compreende analisar a composição de base da região degenerada da biblioteca de candidatos de moldes de transcrição e identificar o componente da sequência de ácidos nucleicos com base na composição média de base da biblioteca de candidatos de moldes de transcrição.

45. Método, de acordo com qualquer uma das 41 a 44, **CARACTERIZADO** pelo fato de que a etapa b) adicionalmente compreende tratar a mistura de transcrição transcrita com

DNase.

46. Método, de acordo com qualquer uma das reivindicações 41 a 45, **CHARACTERIZADO** pelo fato de que a etapa adicionalmente compreende purificar a mistura de transcrição
5 transcrita por separação dos moldes de transcrição transcritos dos outros componentes da reação de transcrição.

47. Método, de acordo com a reivindicação 46, **CHARACTERIZADO** pelo fato de que a etapa de purificação compreende substituir o tampão de reação de transcrição corren-
10 do a reação de transcrição através de uma coluna de dessalinização.

48. Método, de acordo com quaisquer das reivindicações 41 a 47, **CHARACTERIZADO** pelo fato de que a etapa d) é feita antes da etapa c).

15 49. Método, de acordo com quaisquer das reivindicações 41 a 48, **CHARACTERIZADO** pelo fato de que a reação de ligação é uma reação de ligação imobilizada e a reação de ligação compreende uma imobilização de ácido nucléico e um oligonucleotídeo 5'-monofosforilado codificando o promotor.

20 50. Método, de acordo com quaisquer das reivindicações 41 a 49, **CHARACTERIZADO** por adicionalmente compreender repetir as etapas b) até e) mais do que uma vez, antes de efetuar a etapa f).

25 51. Método, de acordo com quaisquer das reivindicações 41 a 50, **CHARACTERIZADO** pelo fato de que a reação de transcrição compreende um ou mais trifosfatos de nucleotídeos modificados e uma polimerase mutada.

52. Método, de acordo com a reivindicação 51,

CARACTERIZADO pelo fato de que os um ou mais trifosfatos de nucleotídeos modificados são um trifosfato de nucleotídeo modificado em 2'.

53. Método, de acordo com a reivindicação 51 ou 52,
5 **CARACTERIZADO** pelo fato de que a polimerase mutada é uma T7 RNA polimerase mutada.

54. Método, de acordo com qualquer uma das reivindicações 51 a 53, **CARACTERIZADO** pelo fato de que o trifosfato de nucleotídeo modificado é um trifosfato de 2'-OMe nucleotídeo.
10

55. Método, de acordo com quaisquer das reivindicações 51 a 54, **CARACTERIZADO** pelo fato de que a reação de transcrição compreende íons magnésio e íons manganês.

56. Método, de acordo com a reivindicação 55,
15 **CARACTERIZADO** pelo fato de que a T7 RNA polimerase mutada é selecionada a partir do grupo que consiste em: SEQ ID NO 1, SEQ ID NO 2, SEQ ID NO 100, SEQ ID NO 101, SEQ ID NO 102 e SEQ ID NO 103.

57. Método, de acordo com qualquer uma das reivindicações 55 a 56, **CARACTERIZADO** pelo fato de que os íons magnésio estão presentes na reação de transcrição em uma concentração que é entre 3,0 a 3,5 vezes maior do que a concentração de íons manganês.
20

58. Método, de acordo com qualquer uma das reivindicações 55 a 56, **CARACTERIZADO** pelo fato de que cada trifosfato de nucleotídeo está presente na reação de transcrição em uma concentração de 1,0 mM, a concentração de íons magnésio é 6,5 mM, e a concentração de íons manganês é 2,0
25

mM.

59. Método, de acordo com qualquer uma das reivindicações 55 a 56, **CARACTERIZADO** pelo fato de que cada trifosfato de nucleotídeo está presente na reação de transcri-
5 ção em uma concentração de 1,5 mM, a concentração de íons magnésio é 8 mM, e a concentração de íons manganês é 2,5 mM.

60. Método, de acordo com qualquer uma das reivindicações 55 a 56, **CARACTERIZADO** pelo fato de que cada trifosfato de nucleotídeo está presente na reação de transcri-
10 ção em uma concentração de 2,0 mM, a concentração de íons magnésio é 9,5 mM e a concentração de íons manganês é 3,0 mM.

61. Método, de acordo com quaisquer das reivindicações 41 a 60, **CARACTERIZADO** pelo fato de que a reação de transcrição adicionalmente compreende o polialquileno glicol.

15 62. Método, de acordo com a reivindicação 61, **CARACTERIZADO** pelo fato de que o polialquileno glicol é o polietileno glicol.

63. Método, de acordo com quaisquer das reivindicações 41 a 61, **CARACTERIZADO** pelo fato de que a reação de
20 transcrição adicionalmente compreende um resíduo de guanossina selecionado a partir do grupo que consiste em: monofosfato de guanossina, difosfato de guanossina, monofosfato de 2' flúor guanossina, difosfato de 2' flúor guanossina, monofosfato de 2'-amino guanossina, difosfato de 2'-amino guanossina,
25 monofosfato de 2'-desóxi guanossina, e difosfato de 2'-desóxi guanossina.

64. Método, de acordo com qualquer uma das reivindicações 41 a 63, **CARACTERIZADO** pelo fato de que a reação de

transcrição compreende a pirofosfatase inorgânica.

65. Método, de acordo com qualquer uma das reivindicações 41 a 64, **CARACTERIZADO** pelo fato de que a primeira região fixa consiste em 2, 3, 4 ou 5 resíduos de guanosina.

5 66. Método, de acordo com qualquer uma das reivindicações 41 a 65, **CARACTERIZADO** pelo fato de que a região degenerada compreende pelo menos 4, 10, 20 ou 30 nucleotídeos.

10 67. Método, de acordo com qualquer uma das reivindicações 41 a 66, **CARACTERIZADO** pelo fato de que a sequência de componente de molde de ácido nucléico a ser identificada é uma sequência líder.

15 68. Método, de acordo com a reivindicação 67, **CARACTERIZADO** pelo fato de que a sequência líder compreende a primeira região fixa e uma sequência derivada da região degenerada do molde de transcrição da biblioteca de candidatos.

20 69. Sequência líder, **CARACTERIZADA** por ser identificada por um método de acordo com qualquer uma das reivindicações 41 a 68.

70. Sequência líder, **CARACTERIZADA** por compreender a sequência de ácidos nucléicos a partir do nucleotídeo 22 até o nucleotídeo 32 de qualquer uma das sequências selecionadas a partir do grupo que consiste em: SEQ ID NOs 10 a 99.

25 71. Sequência líder, **CARACTERIZADA** por compreender a sequência de ácidos nucléicos a partir do nucleotídeo 18 até o nucleotídeo 32 em qualquer uma das sequências selecionadas a partir do grupo que consiste em: SEQ ID NOs 10 a 99.

72. Molde de transcrição, **CARACTERIZADO** por compreender uma sequência líder de acordo com qualquer uma das reivindicações 70 a 71.

73. Molde de transcrição de oligonucleotídeo,
5 **CARACTERIZADO** por ser selecionado a partir do grupo que consiste em: SEQ ID NOS 3 a 6 e 106.

FIG. 1

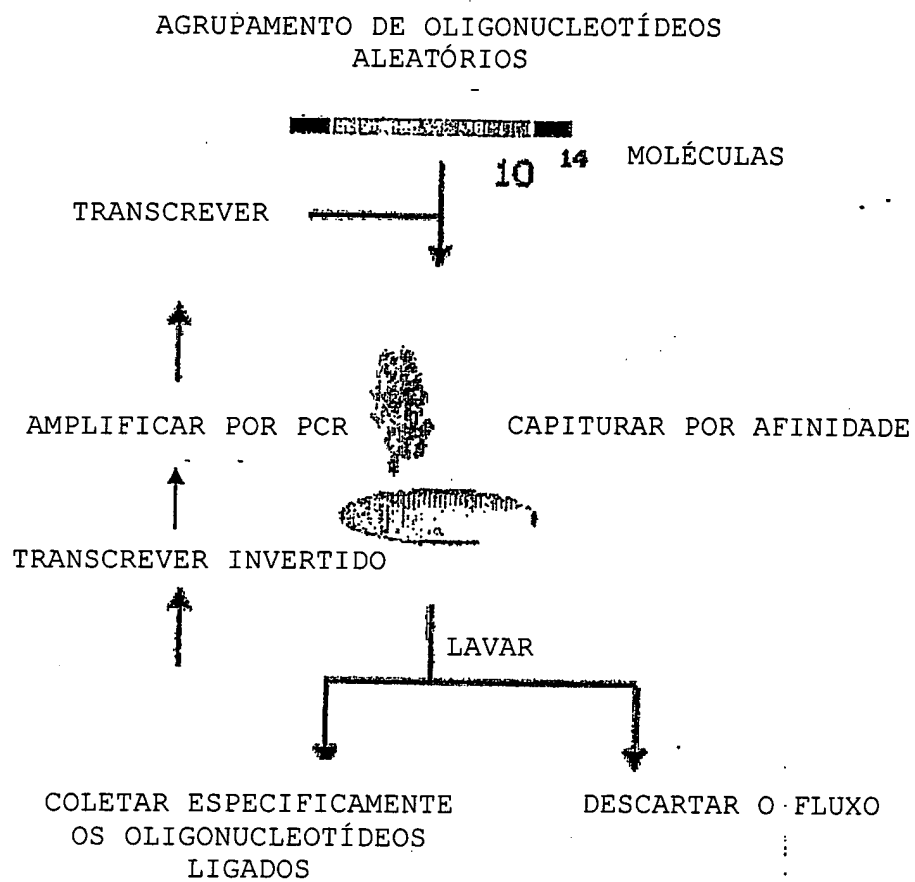


FIG. 3

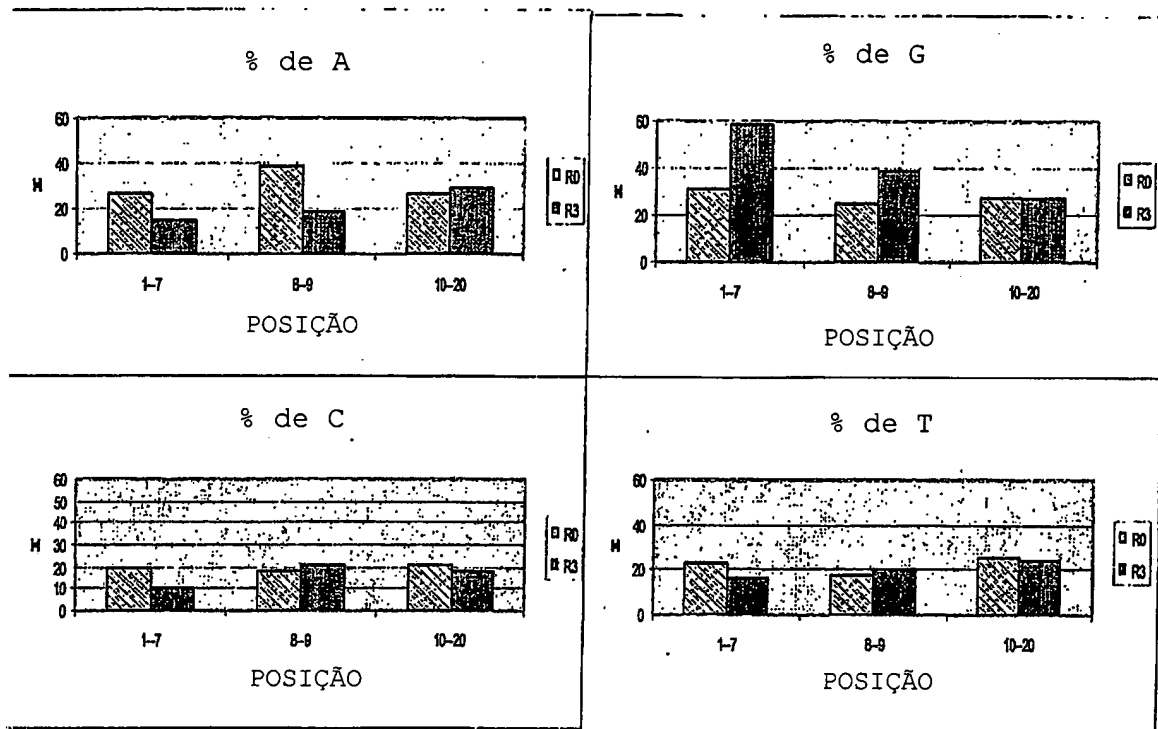


FIG. 4

MOLDE	ARC2117	ARC2118	ARC2119	ARC2120	ARC2121
PRODUÇÃO RELATIVA DE TRANSCRITOS DE mEmY* COM A POLIMERASE LAR	<10	<10	<10	<10	<10
PRODUÇÃO RELATIVA DE TRANSCRITOS DE mEmY* COM A POLIMERASE LAR	30	100	80	20	<10

FIG. 5A

atgaacacgattaacatcgctaagaacgacttctctgacatcgaaactggctgctatcccggtcaacactctggctgaccattacgggtgagcgttagctcgcgaacagttg
gcccttgagcatgagcttacgagatgggtgaagcacgcttcgcaagatgttgagcgtcaacttaagctggtaggggtgggataacgctgcccgaagccttca
tcactacccttactccctaagatgattgcacgcatcaacgactgggttgaaggaagtgaagctaaagcgcggcaagcggccgacagccttccagttctcgaagaatca
agccggaagccgtagcgtacatcaccattagaccactctggcttgcctaacagctgctgacaatacaaccgttcaggctgtagcaagcgcaatcggtcggccattg
aggacgaggctcgttcggctgatacctgacgttaagctaacgactcaagaaacagttgaggaacactcaacgaagcgcgtaggcgacgtctacaagaagca
ttatgcaagttgctgaggctgacatgctcttaagggttactgggtggcaggcgtggcttctggtgcataaggaagactctatcatgtaggagtagcgtcatcga
gatgctcattgagtaaacgggaatggttagcttacaccgccaataatgctggcgttaggtgaagactctgagactatcgaaactcgcaactgaatacgtgaggtatc
gcaaccggtgaggtcgtggctggcatcttccgatgttccaacttgcgttagtcttccctaagccgtggactggcattactgggtggtggtatgggtaacgggtc
tcgtctctggcgtgggtgctgactacagtaagaagacactgatgcgtacggaagacgtttacatgctgaggtgtacaagcgaataacattgcgcaaacaccgca
tggaaaatacaagaagctctagcgggtgcgaacgtaatacgaagtggaagcattgtccggctgaggaacatccctgcgattgagcgtgaagaaactccgatgaa
accggaagacatcgacatgaatctgtaggctctaccgctggaaacgtgctgctgctgtgtaccgcaaggacaaggctcgcaagctcgcgctatcagccttg
agttcatgcttgagcaagcaataagttgtaaccataaggccatctggttcccttacaacatgtagctggcgcggtgctgtttacgctgtgtcaatgttaacccgcaag
gtaacgatatgaccaaaggactgcttacgctggcgaaggttaacccaatcgtaaggaaggttactactggctgaataatccaggtgcaaaactgtcgggtgtcgata
aggttccgttccctgagcgcacatcaagttcattgaggaacaccagagaacatcatggttgcgchagcttccactggagaacacttgggtggctgagcaagattctcc
gttctgtcttctgctgttctgtgtagtgcgtgggtacagcaccagccgctgagctataactgctcccttccgctggcgtttgacgggttctgtctgcatccagcac
ttctccgctgctccgagatgaggtggttggctgcgggttaacttgccttactgtaaacccgttcaggacatctacgggattgttctgaagaagtaacgagattcta
caagcagacgcaatcaatgggaccgataacgaagtagttaccgtgacccgatgagaacactggtgaatctctgagaagtaacgctggcactaaggcactggctg
gtcaatggctgcttacgggttactcgcagtgtagtaagcgttcagtcacgtgctggcttaccgggtccaaagagttcggcttccgtaacaagtgctggaagatacc
attcagcagcgtattgattccggcaagggtctgtatgttactcagccgaatcaggctgctggatatactggctaaagctgatttgggaatctgtgagcgtgacggtgtagc
tgcgggtgaagcaatgaactggcttaagctgctgctgaagcgtgctgctgaggtcaagataagaagactggagagattcttccgaagcgttgcgtgctgattgg
gtactcctgatggttccctgtgtgaggaatacaagaagcctattcagacgcgttgaacctgatttctcctggctagttccgcttacagcctaccattaacaccaac
aaagatagcagattgtagcacacaacaggaggtctgtatcgtccttaacttgtacacgccaagacggtgaccccttctgaagactgtatgttgggcacacgag
aagtagcgaatcgaatcttggcactgattcagcactccttgggtaccattccgctgacgctgcaacctgttcaagcagtgccggaactatggtgacacatatgag
tctgtgtagtactgctgatttctacgaccagttcgtgaccagttgacgagttcgaatggacaataatggcagcacttccggctaaaggttaactgaacctcctgac
atcttagagtcggacttgcgttccgta (SEQ ID NO 120)

FIG. 5B

MNTINIAKNDFSDELAIPFNTLADHYGERLAREQLALEHESYEMGEARFRKMFERQLKAGEVADNAAAKP
LITLLLPKMIARINDWFEEVKAKRGKRPTAFQFLQEKPEAVA YITIKTTLACLT SADNTTVQAVAS AIGRAIED
EARFGRIRDLEAKHFKKNVEBQLNKR VGHVYKKA FMQVVBADMLSKGLLGGEAWSSWHKEDSIHVGVRCI
EMLIESTGMVSLHRQNA GVVGGQDSEITELAPY ABAIATRAGALAGISPMFQPCVVPKPWTGITGGGYWAN
GRRPLALVRTHSKALMRYEDVYMP EVYKAINIAQNTA WKINKKVLAVANVITKWKHC PVEDIPA IEREELP
MKPEDIDMNP EALTA WKRAAAAVYRKDKARKSRRISLEFMLEQANKFANHKA IWFPPYNM DWRGRVYAVS
MFNPQGN DMTKGLLTLAKGKPIGKEGY YWLKIHGANCA GVDKVPFPERIKFIEENHENIMACAKSPLENTW
WABQDSPFCFLAFCEYAGVQHHLGLSYNCSLPLAFD GSCSGIQHFSAMLRDEVGGRAVNLLPSETVQDIYGIV
AKKVNEILQADAINGTDNEVVTVTDENTGBISEKVKLGTKALAGQW LAYGVTRSVTKRSVMTLAYGSKEFG
FRQQVLEDTIQPAIDSGKGLMFTQPNQAAGYMAKLIWESVSVTVVAAVEAMNWLKSAAKLLAAEVKDKKT
GEILRKRCVHVWVTPDGFPVWQ EYKKPIQTRLNLMFLGQFRLQPTINTNKDSEIDAHKQESGIAPNFVHSQDG
SHLRKTVVWAHEKYGIBSFALIHDSFGTTPADAANLFKAVRETMVD TYBSCDVLADFYDQFADQLHESQLDK
MPALPAKG NNLNRDILESDFAF (SEQ ID NO 121)

atgaacacgattaacatcgctaagaacgacttctctgacatcgaaactggctgtatccggttaacactctggctgaccattacggtagcggttagctcgcgaacagttg
 gcccttgagcatgagtttaccgagatgggtgaagcacgcttccgcaagatgtttgagcgtcaacttaagctggtaggttgcggataacgctgccgccaagcctctca
 tcactaccctactccctaagatgattgcacgcatcaacgactgtttgaggaagtgaaagctaagcggcgcaagcggccgacagccttcagttcctgcaagaaatca
 agccgggaagccgtagcgtacatcaccatgaagaccactctggcttgcctaacagtgctgacaatacaaccgttcaggctgtagcaagcgaatcggtcgggccattg
 aggacgaggctcgtctcggctgtatccgtgacctgaagctaagcacttcaagaaaaacgttgaggaaacaactcaacaagcgcgttagggcacgtctacaagaaagca
 tttatgcaagttgtcagggtgacatgtctctaaagggctactcgggtggcggaggcgtggcttctcgtggcataaggaagactctattcatgtaggagtagcgtgcatcga
 gatgtctattgagtaaccgggaatggtagcttacaccgcaaaatgctggcgttagtaggtaagactctgagactatgaactcgcacctgaatacgtgaggctatc
 gcaaccgctgaggtgcgtgctggcgtcctcctgatgttccaaacttgcgtatgtcctcctaagccgtggactggcattactgggtggctattgggctaaccgtcgt
 tegtctctggcgtgggtgcttactcacagtaagaagcactgatgcgtacggaagcgtttacatgcctgagggtgtacaagcgattaacattgcgcaaaacaccgca
 tggaaaatcaacaagaaagtcctagcggctgcgaacgtatcaccaagtggaagcatgtccggctgaggacatccctgcgattgagcgtgaagaactcccgatgaa
 accgggaagacatcgacatgaatcctgaggctctcaccgctggaaacgtgctcggctgtgtgtaccgcaaggacaaggctcgaagttcgccttatcagccttg
 agttcatgcttgagcaagccaataagttgctaaccataagccatctgggttcccttaacaacatggactggcggctgtgtttacgctgtgtcaatgttcaaccgcaag
 gtaacgatattgacaaaggactgtctacgtggcgaaggtaaaccaatcggttaaggaagggttactactgtgtgaaatccacgggtgcaaaactgtgctgggtgtcgtata
 aggttccgttccctgagcgcataagttcattgaggaaccacgagaaacatcatggcttgcgttaagctcactggagaacacttgggtggctgagcaagatttctc
 gttctgtcttctgtcttctgttctgagtagcgtgggtgacagcaccacggcctgagctataactgtcctctccgtgctggttgcgggtctgtcttgcgtacacgac
 tttcgcgtgatgctcggatgaggtagggtggtcgcgggttaactgtcttctagtgaacacgttcaggacatctacgggattgtgtctaagaaggtcaacgagattcta
 caagcagacgcaatcaatgggaccgataacgaagtagtaccgtgaccgatgagaacactggtaaatctctgagaaagtcaagctgggcaactaaggcactggctg
 gtcaatggctggttaccgtgttactcgcagtgtagtactaangcgttcagtcagtcagcgtggctCTGGgggtcacaagagttcggcttccgtcaucuaigtgtggaagata
 ccaitcagccagctattgattcggcgaagggtctgatgttactcagccgaatcaggctgctggatatactgcttaagctgatttgggaatcgtgagcgtgacgggtgta
 gctgcgggtgaagcaatgaactggcttaagctgtctgctaagctgtggtcgtgaggtcaagataagaagactggagagattcttcaagcgttgcgtgtgctt
 gggttaactcctgatgttccctgtgtggcaggaatacaagaagcctattcagacgcgttgaacctgatgttctcgggtcagttccgttacagcctacatcaaccca
 acaagatagcgagatgatgcacacaacaggagctgtggtatcgtccttaactttgtaGCagccaagacggtagccaccttctgaagactgtatgtgtggcacaac
 gagaagtaggaatcgaatcttttgcactgattcacgactccttgggtaccattccggctgacgtgcgaacctgttcaagcagtgccgcaaaactatggttgacacatat
 gagtctgtgtgtacttggctgatttctacgaccagttcgtgaccagttgcagagctcaattggacaaaatgccagcacttccggtaaaaggtaacttgaacctcgt
 gacatcttagagtcggacttgcgttcgctaa

(SEQ ID NO 122)

FIG. 6B

1 atgaacacga ttaacatcgc taagaacgac ttctctgaca tcgaactggc tgctatcccg
 61 ttcaacactc tggctgacca ttacggtgag cgtttagctc gcgaacagtt ggcccttgag
 121 catgagtcct acgagatggg tgaagcacgc ttccgcaaga tgtttgagcg tcaacttaaa
 181 gctggtgagg ttgcggataa cgctgccgcc aagcctctca tcactaccct actccctaag
 241 atgattgcac gcatcaacga ctggtttgag gaagtgaag ctaagcgcgg caagcgccccg
 301 acagccttcc agttcctgca agaaatcaag ccggaagccg tagcgatcat caccattaag
 361 accactctgg ctgacctaac cagtctgac aatacaaccg ttacggctgt agcaagcgca
 421 atcggtcggg ccatgagga cagggtcgc ttcggtcgta tccgtgacct tgaagctaa
 481 cactcaaga aaaacgttga ggaacaactc aacaagcgcg tagggcacgt ctacaagaaa
 541 gcatttatgc aagtgtcga ggctgacatg ctctctaagg gtctactcgg tggcgaggcg
 601 tggcttctgt ggcatagga agactctatt catgtaggag tacgctgcat cgagatgctc
 661 attgagtcac ccggaatggt tagcttacac cgccaaatg ctggcgtagt aggtcaagac
 721 tctgagacta tcgaactcgc acctgaatac gctgaggcta tcgcaacccg tgcagggtcg
 781 ctggctggca tctctccat gtccaacct tgcgtagtc ctcttaagcc gtggactggc
 841 attactggtg gtggctattg ggctaaccgt cgtcgtctc tggcgctggt gcgtactcac
 901 agtaagaaag cactgatcgc ctacgaagac gtttcatgc ctgagggtga caaagcgatt
 961 aacattgcgc aaaacaccgc atggaaaatc aacaagaaag tcttagcggt cgccaacgta
 1021 atcacaagt ggaagcattg tccggtcgag gacatccctg cgattgagcg tgaagaactc
 1081 ccgatgaaac cggaagacat cgacatgaat cctgaggctc tcaccgcgtg gaGacgtgct
 1141 gccgctgctg tgtaccgcaa ggacaaggct cgcaagtctc gccgtatcag ccttgagttc
 1201 atgcttgagc aagccaataa gtttgtaac cataaggcca tctggtccc ttacaacatg
 1261 gactggcgcg gtctgttga cgtgtgtca atgttaacc cgcaaggtaa ogatgtgacc
 1321 aaaggactgc ttacgtggc gaaaggtaaa ccaatcggtg aggaaggta ctactggctg
 1381 aaaatccacg gtgcaaacgt tgcgggtgtc gataagggtc cgttccctga gcgcatcaag
 1441 ttcatgagg aaaaccacga gaacatcatg gcttgcgta agtctccact ggagaacact
 1501 tgggtggctg agcaagattc tccgttctgc ttcttgcgt tctgcttga gtacgtggg
 1561 gtacagcacc acggcctgag ctataactgc tccctccgc tggcgttga cgggtcttgc
 1621 tctggcatcc agcacttctc cgcgatgctc cgagatgagg taggtggtog cgcggttaac
 1681 ttgttctta gtgaaaccgt tcaggacatc tacgggattg ttgctaagaa agtcaacag
 1741 attctacaag cagacgcaat caatgggacc gataacgaag agttaccgt gaccgatgag
 1801 aacactgggtg aaatctctga gaaagtcaag ctgggcacta aggcactggc tggtaaatgg
 1861 ctggcttacg gtgttactcg cagtgtgact aagcgttcag tcatgacgct ggctCTGggg
 1921 tccaaagagt tgggttccg tcaacaagtg ctggaagata ccattcagcc agctattgat
 1981 tccggcaagg gtctgatgti cactcagccg aatcaggctg ctggatacat ggctaagctg
 2041 atttgggaat ctgtgagcgt gacgggtgta gctcgggtg aagcaatgaa ctggcttaag
 2101 tctgtgcta agctgtggc tctgaggtc aaagataaga agactggaga gattcttgc
 2161 aagcgttgcg ctgtgcattg ggtaactcct gatggttcc ctgtgtggca ggaatacaag
 2221 aagcctatc agacgcgtt gaacctgatg ttctcggc agttccgtt acagcctacc
 2281 attaacacca acaagatag cgagattgat gcacacaaac aggagtctgg tatcgtctct
 2341 aactttgtaG Ccagccaaga cgttagccac ctctgtaaga ctgtagtgtg ggcacacgag
 2401 aagtacgga tcgaatctt tgcactgatt cagactcct tcggtaccat tccggctgac
 2461 gctgcgaacc tgtcaaacg agtgcgcgaa actatgggtg acacatatga gtctgtgat
 2521 gtactggctg atttctacga ccagttcgt gaccagttgc acgagtctca attggacaaa
 2581 atgccagcac ttccggctaa aggttaactg aacctccgtg acatcttga gtcggacttc
 2641 gcgttcgct aa (SEQ ID NO123)

FIG. 7

PRODUÇÕES RELATIVAS DE TRANSCRITOS*	ARC2118	ARC2119
[rGTP] 0uM	40	49
[rGTP] 5uM	74	62
[rGTP] 10uM	94	21
[rGTP] 20uM	100	30
[rGTP] 40uM	95	36
[rGTP] 80uM	93	41
[rGTP] 160uM	95	32

FIG. 8

CONCENTRAÇÃO DO COMPONENTE DE TRANSCRIÇÃO	CONDIÇÕES DE TRANSCRIÇÃO 1	CONDIÇÕES DE TRANSCRIÇÃO 2	CONDIÇÕES DE TRANSCRIÇÃO 3	CONDIÇÕES DE TRANSCRIÇÃO 4
[MgCl ₂]	5mM	6.5mM	8mM	9.5mM
[MnCl ₂]	1.5mM	2mM	2.5mM	3mM
[2'-OMe NTP] (cada)	0.5mM	1mM	1.5mM	2mM
PRODUÇÃO RELATIVA	0.8	1.0	1.5	1.6

FIG. 9

	DELEÇÕES	INSERÇÕES	SUBSTITUIÇÕES	N30 PERFEITO	N
2' -OH RNA ¹	0.7%	0.0%	0.3%	74%	1936
2' -OMe RNA ¹	0.6%	0.6%	1.2%	49%	1584
2' -OMe RNA ²	0.0%	0.0%	0.23%	93%	1300

FIG. 10

	Total	A	G	C	T	N
BIBLIOTECA DE PARTIDA	1206	21,9 %	28,3 %	24,1 %	25,0 %	0.66%
BIBLIOTECA APÓS A TRANSCRIÇÃO	1386	25%	25,3 %	21,5 %	28.1 %	0.0 %

FIG. 11

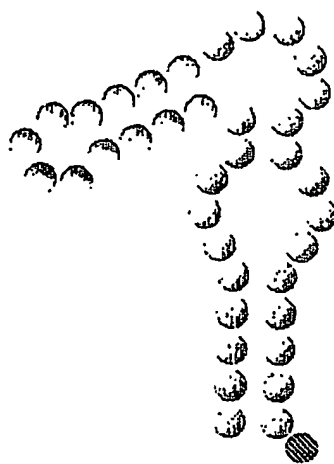


FIG. 12

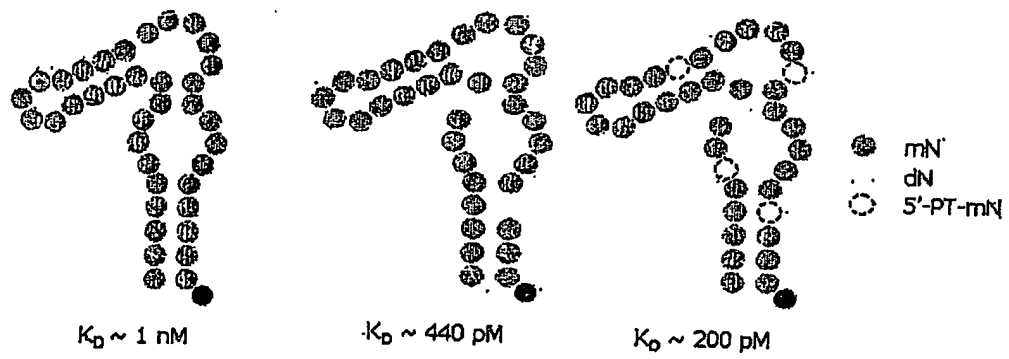
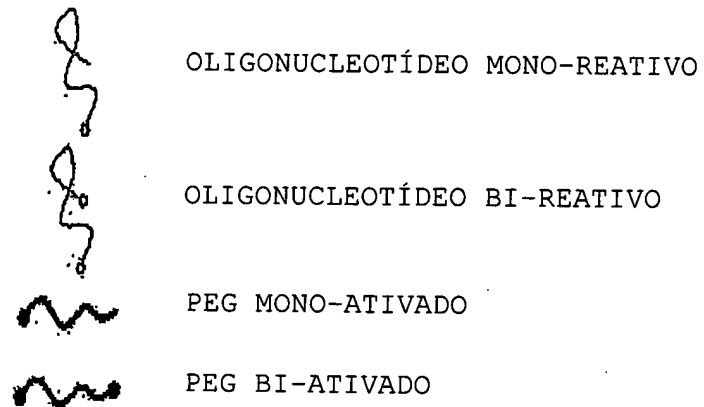
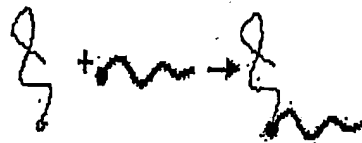


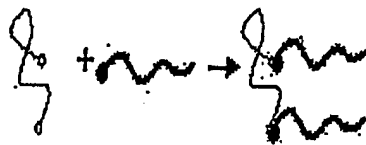
FIG. 13



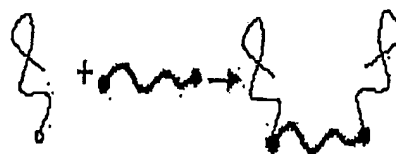
PEGuilação MÚLTIPLA



PEGuilação PADRÃO



DIMERIZAÇÃO VIA PEGuilação



RESUMO

"MATERIAIS E MÉTODOS PARA A GERAÇÃO DE TRANSCRITOS
DE ÁCIDOS NUCLÉICOS INTEIRAMENTE MODIFICADOS EM 2'"

5 Proporcionam-se materiais e métodos para produzir
substâncias terapêuticas de aptâmeros tendo trifosfatos de
nucleotídeos inteiramente modificados incorporados em sua
seqüência.