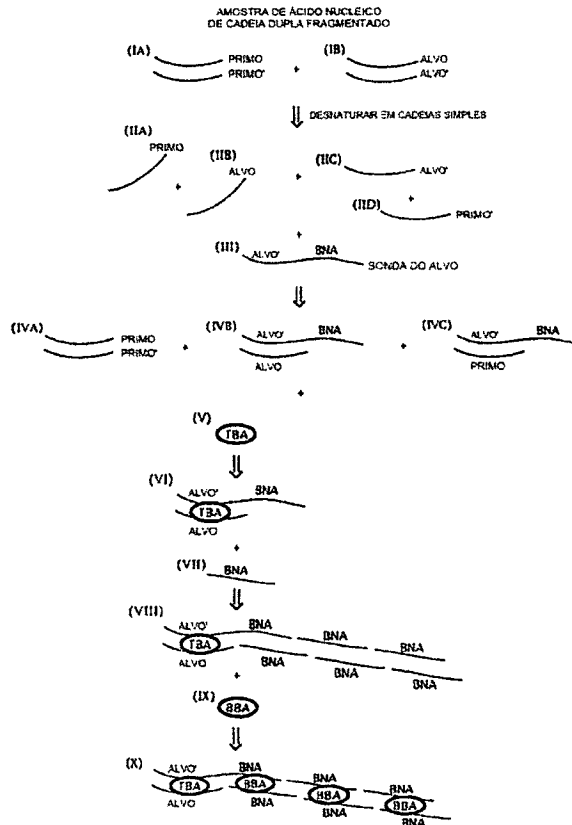


RESUMO

"Método de detecção de ácidos nucleicos com uma composição de sequência específica"

O presente invento consiste num novo método para detecção e localização de sequências de ácido nucleico específicas numa amostra com um elevado grau de sensibilidade e especificidade. O método e as novas composições utilizadas no método envolvem a utilização de Ácidos Nucleicos Sonda, a produção de regiões de ligação a ácidos nucleicos e a utilização de Conjuntos de Ligação ao Alvo de ácido nucleico para detectar e localizar Ácidos Nucleicos Alvo específicos. A detecção e a localização do Ácido Nucleico Alvo são conseguidas mesmo na presença de ácidos nucleicos que possuam sequências semelhantes. O método proporciona um elevado grau de amplificação do sinal produzido por cada acontecimento de ligação específica. Em particular, são apresentados métodos e composições para a detecção de ácido nucleico de HIV e de HPV em amostras. Estes métodos e composições têm utilização em diagnóstico de doença, monitorização genética, análises forenses e análise de misturas de ácidos nucleicos. Algumas das novas composições utilizadas no método de detecção são úteis na prevenção ou tratamento de condições patogénicas.



DESCRIÇÃO

"Método de detecção de ácidos nucleicos com uma composição de sequência específica"

Antecedentes do Invento

1. Campo do Invento

O presente invento proporciona um método e composições para utilizar em ligação, detecção e amplificação da detecção de sequências de Ácido Nucleico Alvo específicas numa amostra com fidelidade e precisão, mesmo na presença de ácidos nucleicos intimamente relacionados mas diferentes. A ligação pode envolver o acompanhamento e montagem de moléculas específicas em Conjuntos de Ligação ao Alvo que se ligam especificamente a Regiões de Ligação ao Alvo formadas através da hibridação de Ácidos Nucleicos Sonda com sequências de Ácido Nucleico Alvo. A amplificação pode envolver o acompanhamento e/ou a montagem de moléculas específicas em Conjuntos de Ligação de Reforço que se ligam especificamente a Regiões de Ligação de Reforço formadas através da hibridação de Ácidos Nucleicos de Reforço com Ácidos Nucleicos Sonda, Ácidos Nucleicos Alvo ou outros Ácidos Nucleicos de Reforço. A detecção envolve proporcionar um ou mais marcadores detectáveis, incluindo moléculas radioactivas, emissoras de luz ou fluorescentes, enzimáticas ou outras moléculas detectáveis ou geradoras de sinais, em associação com o Ácido Nucleico Sonda, o Conjunto de Ligação ao Alvo, o Ácido Nucleico de Reforço ou o Conjunto de Ligação de Reforço.

2. Antecedentes e Descrição da Arte Relacionada

Existe um número crescente de casos nos quais é importante ser capaz de detectar ácidos nucleicos contendo uma sequência específica, daqui em diante designados Ácidos Nucleicos Alvo (TNA), numa amostra. É desejável ser-se capaz de detectar TNA com o menor número de passos de processamento, com os componentes mais simples e até à exclusão de outros ácidos nucleicos semelhantes mas diferentes, daqui em diante designados Ácidos Nucleicos Primos (CNA). É desejável ser-se capaz de detectar TNA específicos até à exclusão de qualquer um e de todos os CNA na amostra de detecção sem a necessidade de amplificação ou de outro processamento pós-deteção.

Existem numerosos métodos que utilizam ácidos nucleicos imobilizados ou marcados como sondas para TNA. No entanto, utilizando métodos conhecidos, é difícil discriminar entre um TNA ligado ao Ácido Nucleico Sonda (PNA) em oposição a um CNA ligado ao PNA. Por exemplo, um ou mais erros de emparelhamento de bases entre o PNA e um CNA podem mesmo assim resultar numa hibridação CNA-PNA que é quase indistinguível de uma hibridação TNA-PNA. Assim, apenas a hibridação não é um indicador óptimo de que um PNA tenha hibridado com um único TNA.

Existem muitas situações nas quais seria utilizado um PNA para tentar determinar se um TNA estava presente numa amostra que pudesse conter CNA. A hibridação do PNA com qualquer CNA nesta situação limitaria o valor de diagnóstico que o PNA pudesse ter para a detecção de um TNA, ausente de verificação adicional. Para além disso, é desejável ser-se capaz de detectar e localizar TNA com baixo número de cópias em amostras que possam conter muitas cópias de CNA, sem a necessidade de criar cópias adicionais do TNA. Seria também desejável ser-se capaz de confirmar a presença de CNA, independentes dos TNA, sem a necessidade de separação dos CNA e TNA na amostra.

Para além disso, seria desejável ser-se capaz de amplificar o sinal de uma hibridação mesmo de baixa frequência de um determinado TNA-PNA. Para este fim, seria desejável um método de polimerização de múltiplas cópias de um marcador, daqui em diante referido como um Ácido Nucleico de Reforço (BNA) sobre o TNA-PNA.

O presente invento proporciona métodos e composições para alcançar os objectivos anteriores desejados. Tal como revelado pela seguinte revisão, as presentes composições e métodos não foram relatados ou sugeridos na arte. Uma revisão geral e abrangente do estado da arte da detecção de ácidos nucleicos é proporcionada em Keller, H., M.M. Manak, "DNA Probes", Stockton Press, 1989.

Foi relatado um método para detecção de erros de emparelhamento de bases através de meios químicos para determinar se um PNA tinha hibridado com um CNA em vez de com

um TNA. Na Patente U.S. 4 794 075 de Ford *et al.*, é discutido um método para distinguir fragmentos de ADN que contenham erros de emparelhamento de uma base em relação aos seus homólogos perfeitamente emparelhados. As regiões de cadeia simples dentro de um fragmento dúplice são modificadas com carbodiimida, que reage com resíduos não emparelhados de guanina (G) e timina (T) no ADN. Moléculas de ADN dúplice lineares não reagem, enquanto que moléculas de ADN com emparelhamentos errados de uma base reagem quantitativamente. Após a reacção com carbodiimida, as moléculas de ADN são fraccionadas em géis de poliacrilamida de elevada percentagem de modo a que os fragmentos modificados e não modificados possam ser distinguidos. Ford *et al.* aplicaram esta técnica para localizar e purificar diferenças na sequência de ADN responsáveis por variação fenotípica e doença hereditária. Embora este método seja útil para seguir variações em material genético, possui um grande número de passos, requer componentes dispendiosos e não oferece um meio directo para determinar se um PNA hibridou com o TNA em exclusivo relativamente aos CNA na amostra.

Houve algumas tentativas para assegurar que pelo menos uma porção da hibridação entre o PNA e outro ácido nucleico fosse complementar. Um método envolve a monitorização dos produtos de transcrição que são produzidos se o PNA hibridar com um ácido nucleico de forma suficiente para ser transcrito a partir de um local promotor contido na sonda. A Patente U.S. 5 215 899 de Dattagupta divulga o modo como sequências de ácido nucleico específicas são amplificadas através da utilização de uma sonda em "gancho de cabelo" que, após hibridação e ligação a uma sequência alvo, é capaz de ser transcrita. A sonda compreende uma sequência auto-complementar de cadeia simples que, sob condições de hibridação, forma uma estrutura em "gancho de cabelo" possuindo uma região promotora funcional e compreende ainda uma sequência sonda de cadeia simples prolongando-se a partir da extremidade 3' da sequência em "gancho de cabelo". Após hibridação com uma sequência alvo complementar à sequência sonda e ligação da extremidade 3' da sequência alvo hibridada com a extremidade 5' da sonda em "gancho de cabelo", a sequência alvo é tornada passível de transcrição na presença de uma ARN-polimerase adequada e dos ribonucleósido-trifosfatos apropriados (rNTP). A amplificação

é conseguida através da hibridação da sequência do TNA desejado com a sonda, ligando o TNA ao PNA, adicionando a ARN-polimerase e os rNTP aos híbridos separados e permitindo que a transcrição prossiga até se acumular uma quantidade desejada do produto da transcrição do ARN. Esse método envolve, geral e especificamente, a utilização de ADN em "gancho de cabelo" formado com uma extremidade de cadeia simples desemparelhada para hibridar com uma sequência alvo. Quando a sequência alvo se liga, é permitida a produção de produtos da transcrição do ARN. Assim, o método envolve a detecção de produtos de transcrição secundários em vez da utilização de um conjunto de ligação de ácido nucleico para imobilizar e/ou localizar directamente uma sequência alvo. Um CNA pode ligar-se facilmente à sonda e a falta de complementaridade não interferirá necessariamente com a formação de um híbrido CNA-PNA que possa então suportar a produção de produtos de transcrição indesejados.

Um CNA ligado ao PNA pode ser detectado se a falta de complementaridade interferir com a susceptibilidade do par híbrido CNA-PNA para ser cortado por uma endonuclease de restrição. Na Patente U.S. 5 118 605 de Urdea e na Patente U.S. 4 775 619 de Urdea, foram proporcionados novos métodos para ensaio de um ácido nucleico a analisar, que empregam polinucleótidos possuindo sequências oligonucleotídicas substancialmente homólogas a uma sequência de interesse na substância a analisar, onde a presença ou ausência de hibridação a um rigor predeterminado proporciona a libertação de um marcador de um suporte. São empregues várias técnicas para a ligação de um marcador a um suporte, após o que a clivagem de uma cadeia simples ou dupla permite a libertação de um marcador de um suporte e a libertação do marcador pode ser detectada como indicadora da presença de uma determinada sequência polinucleotídica numa amostra. No entanto, esta técnica possui a desvantagem de um par CNA-PNA poder ser cortado pela endonuclease de restrição, mesmo se houver um emparelhamento errado, desde que o emparelhamento errado esteja fora da região de reconhecimento da endonuclease. Isto conduziria ao insucesso do ensaio para identificar um híbrido CNA-PNA.

Outro método utiliza uma sonda de ADN ramificada para detectar ácidos nucleicos. A Patente U.S. 5 124 246 de Urdea *et al.* divulga múltimeros oligonucleotídicos lineares ou ramificados úteis como amplificadores em ensaios bioquímicos que compreendem (1) pelo menos uma primeira unidade oligonucleotídica de cadeia simples (PNA) que é complementar a uma sequência oligonucleotídica de cadeia simples de interesse (TNA), e (2) uma multiplicidade de segundas unidades oligonucleotídicas de cadeia simples que são complementares a um oligonucleótido marcado de cadeia simples. Embora sejam descritas hibridações de ácidos nucleicos em sanduíche amplificadas e imunoenaios utilizando os múltimeros, o método tem a limitação de poderem ocorrer hibridações PNA-CNA e resultar na produção de um sinal indesejado.

Para além dos métodos para identificação de TNA, foram divulgados métodos para a amplificação deste ADN. Na Patente U.S. 5 200 314 de Urdea, uma cadeia polinucleotídica a analisar possuindo uma sequência a analisar (TNA) é detectada dentro de uma amostra contendo polinucleótidos através do contacto do polinucleótido a analisar com uma sonda de captura (PNA) sob condições de hibridação, onde a sonda de captura possui um primeiro parceiro de ligação específico para o TNA, e uma segunda sequência de ligação específica para um terceiro parceiro de ligação de fase sólida. O dúplex resultante é então imobilizado através de ligação específica entre os parceiros de ligação, e os polinucleótidos não ligados são separados das espécies ligadas. O polinucleótido a analisar é opcionalmente deslocado da fase sólida, depois amplificado através de PCR. Os iniciadores de PCR possuem, cada um, uma região polinucleotídica capaz de hibridar com uma região do polinucleótido a analisar e pelo menos um dos iniciadores possui ainda um parceiro de ligação adicional capaz de se ligar a um parceiro de ligação de fase sólida. O produto amplificado é então separado da mistura reaccional através da ligação específica entre os parceiros de ligação e o produto amplificado é detectado. Embora seja possível confirmar (através de PCR) que um determinado ácido nucleico hibridou com o PNA, a confirmação é dispendiosa e envolve múltiplos passos.

Quanto aos relatórios que envolvem a interacção de um ácido nucleico de cadeia dupla e uma proteína de ligação a ADN, foi descrito um método através do qual uma sequência de ADN immobilizada que contém locais de ligação a uma única proteína é utilizada para purificar essa proteína. A Patente U.S. 5 122 600 de Kawaguchi et al. divulga uma microesfera com ADN immobilizado compreendendo cadeias de ADN possuindo sequências de bases que se ligam especificamente a uma determinada proteína, e um transportador possuindo um tamanho de partícula de não mais de 50 μm e não menos de 0,01 μm que não absorve qualquer proteína, sendo as cadeias do referido transportador e do referido ADN ligadas uma à outra através de uma ligação química, e um processo para purificação de uma proteína utilizando a referida microesfera. Uma vez que isto é um método de purificação para uma proteína, não divulga um método de detecção de um TNA nem um método através do qual uma proteína se ligue a um ácido nucleico de cadeia dupla para fins de detecção e localização de sequências de TNA específicas.

Em EP-A-0453301 descreve-se um método para detecção de uma sequência alvo polinucleotídica numa amostra, em que as sequências num TNA são detectadas através da hibridação do primeiro e do segundo PNA com o TNA. Cada PNA contém um dúplex pré-formado, ou um dúplex que é formado através de prolongamento da cadeia, capaz de se ligar a uma proteína de ligação específica da sequência nucleotídica.

Em EP-A-0147665 divulga-se também a utilização de proteínas de ligação a ADN dúplex específicas da sequência como meios de detecção num ensaio de hibridação. Novamente, a sonda dúplex é pré-formada.

Em EP-A-0450594 divulga-se a possibilidade de marcação das designadas moléculas reveladoras com compostos específicos da sequência dúplex, p.ex. certos intercaladores. Estes compostos são unidos às moléculas reveladoras antes da hibridação.

Em US-A-4556643 divulga-se a detecção não radioactiva de sequências nucleotídicas específicas numa amostra, envolvendo

a hibridação de uma sonda contendo sequências específicas da proteína de ligação ao ADN.

WO 93/00446 divulga oligonucleótidos de cadeia simples compreendendo uma porção que, quando tornada de cadeia dupla, se liga à proteína UL9 derivada do vírus de Herpes simples, e outra porção que, quando tornada de cadeia dupla, se liga a compostos intercaladores.

Sumário do Invento

O presente invento é definido nas reivindicações anexas. Proporciona métodos através dos quais sequências de Ácido Nucleico Alvo (TNA) específicas são detectadas através da utilização de Ácidos Nucleicos Sonda (PNA) que, após hibridação com os TNA, são capazes de se ligar a Conjuntos de Ligação ao Alvo (TBA). Cada TBA liga-se pelo menos a uma região específica do par híbrido PNA-TNA, a Região de Ligação Alvo (TBR). O TBA é constituído por uma ou mais moléculas, das quais uma ou mais se podem ligar a sequências TBR de um modo específico e dependente da sequência ou da conformação. O TBA pode compreender uma ou mais sequências piloto, designadas "PILOTOS" ou "Sequências de Assimetria", que se montam e restringem os componentes de ligação aos nucleótidos do TBA a geometrias específicas. Os PILOTOS actuam montando unidades de reconhecimento de ácido nucleico específicas ou outros pilotos aos quais as unidades de reconhecimento de ácidos nucleicos específicas se unem nos TBA de um modo predeterminado. O TBA pode também conter uma ou mais moléculas que ancoram ou localizam o TBA.

Os PNA de acordo com o presente invento são definidos na reivindicação 1. As utilizações destes ácidos nucleicos são definidas noutras reivindicações.

Os PNA, para além das sequências específicas de TNA, contêm também uma ou mais sequências, 1/2 BBR, capazes de hibridar com 1/2 BBR complementares em Ácidos Nucleicos de Reforço (BNA). Através da hibridação dos BNA adicionados aos 1/2 BBR iniciais presentes nos PNA, os prolongamentos dos PNA são feitos na forma de híbridos PNA-BNA e depois BNA-BNA. Estes prolongamentos contêm uma ou mais Regiões de Ligação de

Reforço (BBR). Cada BBR é capaz de se ligar a um Conjunto de Ligação de Reforço (BBA). O BBA é constituído por moléculas, uma ou mais das quais pode ligar-se a uma BBR de um modo específico e dependente da sequência ou da conformação. O BBA pode compreender uma ou mais sequências piloto, designadas "PILOTOS" ou "Sequências de Assimetria", que montam e restringem os componentes de ligação dos nucleótidos do TBA a geometrias específicas. Os PILOTOS actuam montando unidades de reconhecimento de ácidos nucleicos específicas ou outros pilotos aos quais unidades de reconhecimento de ácidos nucleicos específicas se unem nos BBA de um modo predeterminado. O BBA pode conter moléculas que ancoram ou localizam o BBA ou que permitem a detecção dos BBA ligados e deste modo dos complexos TBA-TNA-PNA aos quais elas, por sua vez, se ligam. São divulgados métodos e composições para utilização dos 1/2 BBR, BNA, BBR, BBA e PILOTOS BBA, incluindo a sua utilização como componentes de estojos de testes de diagnóstico e forenses.

São divulgados métodos e composições para os procedimentos de teste e a produção de um estojo de teste contendo PNA, TBA, TBR, BNA, BBR e BBA para a detecção, localização e diferenciação de sequências de ácido nucleico específicas, incluindo sequências de ácido nucleico que se verificam em células humanas, no Vírus da Imunodeficiência Humana (HIV), no Vírus do Papiloma Humano (HPV) e noutros sistemas contendo ácidos nucleicos incluindo vírus e bactérias.

Assim, é um objectivo deste invento proporcionar métodos e composições para utilizar na ligação, detecção e amplificação da detecção de sequências de Ácido Nucleico Alvo numa amostra com fidelidade e precisão, mesmo na presença de sequências de ácido nucleico intimamente relacionadas mas diferentes.

Assim, é um objectivo deste invento proporcionar métodos e composições para a criação de Conjuntos de Ligação ao Alvo que se ligam especificamente a Regiões de Ligação Alvo formadas através da hibridação de Ácidos Nucleicos Sonda com sequências do Ácido Nucleico Alvo.

Outro objectivo deste invento é o de proporcionar um método e composições para a criação de Conjuntos de Ligação de Reforço que se liguem especificamente a Regiões de Ligação de Reforço formadas através da hibridação de sequências de Ácido Nucleico de Reforço com Ácidos Nucleicos Sonda, Ácidos Nucleicos de Reforço e Ácidos Nucleico em "gancho de cabelo".

Outro objectivo deste invento é o de proporcionar um método e composições para utilizar na amplificação da detecção de conjuntos de Ligação ao Alvo ligados a Regiões de Ligação Alvo utilizando Conjuntos de Ligação de Reforço e Ácidos Nucleicos de Reforço.

Outro objectivo deste invento é o de proporcionar um método e composições que permitam a utilização de um ou mais marcadores detectáveis, incluindo mas limitado a marcadores radioactivos, emissores de luz, fluorescentes, enzimáticos ou outras moléculas geradores de sinal. Estes marcadores são utilizados em associação com Ácidos Nucleicos Sonda, Conjuntos de Ligação ao Alvo, Conjuntos de Ligação de Reforço, Ácidos Nucleicos de Reforço ou Ácidos Nucleicos em "gancho de cabelo".

Breve Descrição dos Desenhos

As seguintes ilustrações estão contidas na **Figura 1**: a **Figura 1-I** é um PNA contendo um 1/2 TBR, que é uma sequência de cadeia simples que é complementar a um TNA e a uma sequência de 1/2 BBR. A **Figura 1-IIa** é um TNA ao qual são adicionados os componentes da **Figura 1-I** e, sob condições de hibridação, se liga ao PNA para formar os componentes da **Figura 1-IIIa**, um híbrido PNA-TNA contendo pelo menos uma TBR. A **Figura 1-IVa** é um BNA que é adicionado aos componentes da **Figura 1-IIIa** e, sob condições de hibridação, se liga a 1/2 BBR da **Figura 1-IIIa** para formar um híbrido PNA-BNA contendo uma BBR mostrada na **Figura 1-Va**.

A **Figura 1-IIb** é um BNA que é adicionado aos componentes da **Figura 1-I**, e que, sob condições de hibridação, se liga ao PNA para formar os componentes da **Figura 1-IIIb**, um híbrido PNA-TNA contendo uma BBR. A **Figura 1-IVb** é um TNA ao qual se adicionam os componentes da **Figura 1-IIIb** e que, sob condições

de hibridação, se liga a 1/2 TBR da Figura 1-IIIb para formar um híbrido PNA-BNA contendo uma TBR mostrada na **Figura 1-Vb**.

A **Figura 1-IIc** é um HNA que é adicionado aos componentes da Figura 1-I, e que, sob condições de hibridação, se liga ao PNA para formar os componentes da **Figura 1-IIIc**, um híbrido PNA-HNA contendo uma BBR. A **Figura 1-IVc** é um TNA que é adicionado aos componentes da Figura 1-IIIc e que, sob condições de hibridação, se liga a 1/2 TBR da Figura 1-IIIc para formar um híbrido PNA-BNA contendo uma BBR mostrada na **Figura 1-Vb**.

Os híbridos que formam as TBR e BBR são úteis no presente invento. Os PNA e BNA, tal como indicado na Figura 1, podem não conter um suporte e/ou indicador ligado ou conter um suporte ligado ou outro meio de localização, incluindo, mas não se limitando a, ligação a esferas, polímeros e superfícies, e/ou indicadores (OSA).

A **Figura 2a** é um diagrama de estratégias para polimerização de BNA sobre PNA e protecção das extremidades por HNA.

A **Figura 2b** é um diagrama de estratégias adicionais para amplificação de sinais PNA-TNA através da polimerização de BNA e protecção das extremidades por HNA.

A **Figura 3** é um diagrama mostrando a utilização de BNA contendo múltiplas 1/2 BBR por BNA.

A **Figura 4a** é um diagrama mostrando a ligação de TBA e BBA a TBR e BBR, e a capacidade do TBA para discriminar entre TNA e CNA. De acordo com esta concretização, se o TBA for imobilizado, numa esfera, superfície de uma placa de microtítulo ou qualquer outra superfície, apenas complexos tais como o complexo X seriam mantidos e detectados, enquanto que complexos tais como os complexos XI não.

A **Figura 4b** é um diagrama que exemplifica acontecimentos semelhantes aos mostrados na Figura 4a mas numa ordem de ocorrência ligeiramente diferente.

A **Figura 5** é um diagrama exemplificando PNA contendo entre uma 1/2 TBR e nenhuma 1/2 BBR até PNA contendo até cinco 1/2 TBR e uma 1/2 BBR. Os membros (a) e (b) de cada número (I, II, III, IV, V) formam um conjunto que, após hibridação com um TNA, proporciona TBR com (membros (a)) ou sem (membros (b)) uma 1/2 BBR disponível para amplificação através de hibridação com BNA possuindo 1/2 BBR complementares.

A **Figura 6a** é um diagrama exemplificando um determinado TNA possuindo duas 1/2 TBR que, após ligação a um PNA apropriado, forma duas TBR intimamente associadas capazes de se ligar a dois TBA. É também proporcionada uma 1/2 BBR para amplificação.

A **Figura 6b** é um diagrama mostrando os mesmos acontecimentos que a Figura 6a excepto que aqui é utilizado um duplo TBA de modo a que a discriminação entre TBR simples que ocorrem em amostras celulares normais possa ser discriminada de TBR duplos, anormais.

A **Figura 6c** é um diagrama mostrando o mesmo cenário que a Figura 6a excepto que aqui são identificadas cinco TBR no TNA. Cada TBR pode ser ligada a um TBA igual ou diferente e cada TBA pode ser diferenciadamente marcado, permitindo a confirmação de todos os cinco locais presentes no TNA.

A **Figura 6d** é um diagrama dos mesmos acontecimentos que na Figura 6c excepto que aqui é mostrado um duplo TBA, prolongando o que é mostrado na Figura 6b à utilização do duplo TBA. Um exemplo do TNA mostrado no item II das Figuras 6a, 6b, 6c e 6d é o ADN ou ARN de cadeia simples de HIV.

A **Figura 7** mostra a LTR de HIV como um TNA e dois PNA, e uma estratégia para a detecção do TNA utilizando os PNA.

A **Figura 8** é um esquema de uma concretização do presente invento em que é utilizada um conjunto de ligação ao alvo para ligar um híbrido TNA-PNA, e são utilizados conjuntos de ligação de reforço para ligar BNA polimerizados.

A **Figura 9** é um esquema de um TBA modular no qual as sequências do conjunto, as sequências ligantes e as sequências

de assimetria são utilizadas para acompanhar unidades de reconhecimento do ácido nucleico desejadas juntas para formar um TBA.

A **Figura 10** mostra TBA modulares úteis na detecção de sequências específicas de HIV.

A **Figura 11** mostra TBA modulares úteis na detecção de sequências de vírus do papiloma humano. Cada unidade de E2 é de facto um dímero da porção de ligação ao ADN de E2.

A **Figura 12a** é um esquema do fraccionamento do TNA e da mudança de mobilidade devida à ligação de um TBA.

A **Figura 12b** é um esquema do fraccionamento do TNA e do aumento na mudança de mobilidade devida à ligação de BBA para além de TBA.

A **Figura 13** mostra uma estratégia de detecção para sequências de deleção; um exemplo da utilização desta estratégia é para um ensaio de integração de vírus do papiloma humano.

A **Figura 14** mostra a montagem de TBA de ordem superior através da utilização de unidades de reconhecimento de ácidos nucleicos, sequências ligantes, de montagem e de assimetria de modo a serem formados vários Conjuntos de Ligação ao Alvo específicos para locais de ligação na LTR de HIV.

A **Figura 15** mostra a montagem de TBA de ordem superior através da utilização de unidades de reconhecimento de ADN, sequências ligantes, de montagem e de assimetria de modo a serem formados vários Conjuntos de Ligação ao Alvo específicos para locais de ligação no genoma de HPV.

A **Figura 16** mostra a discriminação alcançada através da utilização de um TBA complexo e a capacidade de moléculas de ligação alvo competidoras endógenas para eliminar a ligação do TBA a uma molécula de ácido nucleico prima mas não do TNA que contém a orientação apropriada de mais de um local reconhecido pelo TBA.

A **Figura 17** mostra a capacidade de um TBA para ser especificamente direccionado para se ligar a locais de emparelhamento errado da sequência e a ligar-se de preferência aqueles locais em locais primos que não contenham todos os emparelhamentos errados alvo.

Breve Descrição das Sequências

SEQ ID NO:1 corresponde à Figura 5-Ia-I e mostra o local de ligação a NF-kB de MHC de classe I.

SEQ ID NO:2 corresponde à Figura 5(Ia) e mostra o local de ligação a NF-kB da microglobina B2.

SEQ ID NO:3 corresponde à Figura 5(Ia) e mostra o local de ligação a NF-kB da imunoglobulina capa.

SEQ ID NO:4 corresponde à Figura 5 (Ia) e mostra um dos locais de ligação a NF-kB de HIV.

SEQ ID NO:5 corresponde à Figura 5 (Ia) e mostra um dos locais de ligação a NF-kB de HIV.

SEQ ID NO:6 corresponde à Figura 5 (Ia) e mostra o local de ligação a NF-kB de c-myc.

SEQ ID NO:7 corresponde à Figura 5 (IIa) e mostra um local de ligação duplo a NF-kB de HIV.

SEQ ID NO:8 corresponde à Figura 5 (IIa) e mostra um local de ligação duplo a NF-kB de HIV.

SEQ ID NO:9-16 correspondem à Figura 5 (IIa) e mostram um local de ligação duplo com um local sendo um local de ligação a NF-kB de HIV e o outro local sendo um local de ligação a SP1 de HIV.

SEQ ID NO:17-18 correspondem à Figura 5 (IIa) e mostram um local de ligação duplo a SP1 de HIV.

SEQ ID NO:19-31 correspondem à Figura 5 (IIIa) e mostram um local de ligação duplo a NF-kB de HIV e um local de ligação a SP1 de HIV.

SEQ ID NO:32-33 correspondem à Figura 5 (IVa) e mostram um local de ligação quádruplo onde dois locais são locais de ligação a NF-kB de HIV e dois locais são locais de ligação a SP1 de HIV.

SEQ ID NO:34 corresponde à Figura 5 (VIa) e mostra um local de ligação quádruplo onde dois locais são locais de ligação a NF-kB de HIV e três locais são locais de ligação a SP1 de HIV.

SEQ ID NO:35 é um exemplo de uma 1/2 BBR, neste caso os elementos OL1, OL2 e OL3 do operador esquerdo do bacteriófago lambda, incluindo sequências intervenientes.

SEQ ID NO:36 é um exemplo de uma 1/2 BBR, neste caso os elementos OR3, OR2 e OR1 do operador direito do bacteriófago lambda, incluindo sequências intervenientes.

SEQ ID NO:37 é a LTR de HIV.

SEQ ID NO:38 é um PNA complementar ao PNA da LTR de HIV.

SEQ ID NO:39 é um PNA complementar a um PNA da LTR de HIV diferente do de SEQ ID NO:38.

SEQ ID NO:40 é um PNA complementar a parte da LTR de HIV e contém também 1/2 BBR e uma sequência solta para polimerização dos BNA no PNA.

SEQ ID NO:41 é um BNA complementar à 1/2 BBR de SEQ ID NO:40.

SEQ ID NO:42 é um BNA que polimerizará sobre o BNA de SEQ ID NO:41 e que, com as SEQ ID NO:40 e 41, cria um local de reconhecimento *Pst*I.

SEQ ID NO:43 é um BNA que é complementar ao BNA de SEQ ID NO:42 e que completa um local de reconhecimento *Bam*HI.

SEQ ID NO:44 é um HNA que possui um local de reconhecimento *Bam*HI que hibridará com o local de reconhecimento *Bam*HI criado por SEQ ID NO:42 e 43 no polímero em crescimento.

SEQ ID NO:45 é um segundo PNA que, tal como SEQ ID NO:40, é complementar a parte da LTR de HIV, mas não à mesma sequência que SEQ ID NO:40. SEQ ID NO:45 codifica também uma 1/2 BBR e uma ponta solta que permitirá a polimerização de BNA começando com um local de reconhecimento *Sph*I.

SEQ ID NO:46-62 são PNA específicos do vírus do papiloma humano (HPV) que, após hibridação com sequências de HPV, formam TBR que se ligam a proteínas de ligação ao ADN de HPV.

SEQ ID NO:63-71 são unidades de reconhecimento de ADN de NF-kB para incorporação em TBA.

SEQ ID NO:72 é uma sequência de localização nuclear.

SEQ ID NO:73 é uma unidade de reconhecimento da sequência de SP1.

SEQ ID NO:74 é uma unidade de reconhecimento de proteínas de ligação a TATA.

SEQ ID NO:75-84 são unidades de reconhecimento do ADN de E2 do vírus de papiloma.

SEQ ID NO:85-92 são sequências de assimetria.

SEQ ID NO:93 é uma unidade de reconhecimento de proteínas de ligação a TATA de *Arabidopsis*.

SEQ ID NO:94 é uma unidade de reconhecimento de proteínas de ligação ao ADN de HPV-16-E2-1.

SEQ ID NO:95 é uma unidade de reconhecimento de proteínas de ligação ao ADN de HPV-16-E2-2.

SEQ ID NO:96 é uma unidade de reconhecimento de proteínas de ligação ao ADN de HPV-18-E2.

SEQ ID NO:97 é uma unidade de reconhecimento de proteínas de ligação ao ADN de HPV-33-E2.

SEQ ID NO:98 é uma unidade de reconhecimento de proteínas de ligação ao ADN de E2 do vírus do papiloma bovino.

SEQ ID NO:99-102 são sequências ligantes exemplificativas.

SEQ ID NO:103 é uma sequência sinal de localização nuclear (NLS) exemplificativa.

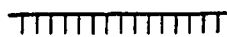
SEQ ID NO:104-108 são sequências de acompanhamento exemplificativas.

SEQ ID NO:109-116 são sequências de TBA montadas exemplificativas.

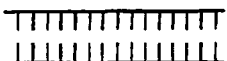
SEQ ID NO:117 é um local de ligação a NF-kB de consenso.

SEQ ID NO:118 é uma sequência de aminoácidos de Tat de HIV.

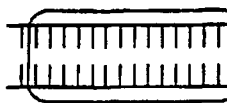
Abreviaturas



ácido nucleico de cadeia simples



ácido nucleico de cadeia dupla



Região de ligação no ácido nucleico



sem suporte nem indicadores, ou com suporte sólido ou outros meios de localização, incluindo, mas não se limitando a esferas, polímeros e superfícies ou indicadores = OSA

BBA conjunto de ligação de reforço (*Booster Binding Assembly*)

BBR região de ligação de reforço (*Booster Binding Region*)

BNA	ácido nucleico de reforço (<i>Booster Nucleic Acid</i>)
CNA	ácido nucleico primo (<i>Cousin Nucleic Acid</i>)
1/2 BBR	região de cadeia simples que, quando hibridada com a sequência complementar de um HNA ou de um BNA, se pode ligar a um BBA
1/2 TBR	região de cadeia simples do PNA que, quando hibridada com a sequência complementar de um TNA, se pode ligar a um TBA
OSA	suporte ou ligação opcional, círculo com caixa (<i>Optional Support ou Attachment</i>)
PNA	ácido nucleico sonda (<i>Probe Nucleic Acid</i>)
TBA	conjunto de ligação ao alvo (<i>Target Binding Assembly</i>)
TBR	região de ligação alvo (<i>Target Binding Region</i>)
TNA	ácido nucleico alvo (<i>Target Nucleic Acid</i>)
HNA	Ácido Nucleico em "gancho de cabelo" (<i>Hairpin Nucleic Acid</i>)

Definições

Deve também entender-se a partir da descrição que se segue que quando se mencionam termos como conjuntos de ligação ao alvo (TBA), conjuntos de ligação de reforço (BBA), proteínas de ligação ao ADN, proteínas de ligação ao ácido nucleico ou proteínas de ligação ao ARN, o que se pretende são composições constituídas por moléculas que se ligam a sequências de ácido nucleico alvo (TNA) de ADN ou ARN independentemente da especificidade da categoria das moléculas de ligação a partir das quais são derivadas. Assim, por exemplo, um TBA adaptado para se ligar a sequências do vírus da imunodeficiência humana pode ser muito semelhante a um factor de transcrição NF-kB que tipicamente se liga a sequências de ADN. No entanto, tal como aqui se utiliza, entender-se-á que o TBA pode ser adaptado para uma utilização óptima para se ligar a sequências de ARN de uma determinada composição ou conformação de sequência.

A fidelidade do método de detecção aqui divulgado depende em grande medida da ligação selectiva dos TBA e BBA a motivos de ácido nucleico particulares. Deve entender-se ao longo desta divulgação que a base da discriminação de TBA e BBA dos TNA de sequências aparentadas (ácidos nucleicos primos ou CNA) pode ser a formação de segmentos híbridos precisos de ácido nucleico sonda (PNA)-ácido nucleico alvo (TNA) (híbridos PNA-TNA). No entanto, a base da discriminação pode também ser a formação de uma determinada conformação e pode não requerer a completa ausência de emparelhamento errado de bases no híbrido TNA-PNA. Assim, a base da operação de TBA ou BBA deve ser sempre entendida como dependendo da discriminação de qualquer propriedade única do híbrido TNA-PNA em oposição a quaisquer propriedades mostradas por quaisquer híbridos PNA-CNA que se possam formar numa amostra de teste colocada em contacto com um dado PNA.

Descrição Detalhada do Invento

O presente invento é definido pelas reivindicações anexas. Proporciona um método para identificar especificamente um ácido nucleico alvo (TNA) numa amostra através da utilização de conjuntos de ligação ao alvo (TBA) que incorporam proteínas de ligação a ácido nucleico específicas. Através da utilização de ácidos nucleicos sonda (PNA) específicos para uma dada sequência de TNA e um TBA que seja específico para a região de ligação alvo (TBR) dúplice formada após a formação de sequências híbridas TNA-PNA, forma-se um complexo TBA-TNA-PNA estável. Proporcionando adicionalmente sequências amplificáveis específicas no PNA, para além de sequências que contribuam especificamente para a formação da TBR reconhecida pelo TBA, a ligação do PNA ao TNA é detectada e a detecção amplificada. Para este fim, pode ser utilizado qualquer um de vários sistemas de amplificação de ácidos nucleicos, incluindo reacção em cadeia da polimerase ou a utilização de ADN ramificado, cada ramificação do qual contém um marcador detectável. Em particular, é aqui descrito um novo método de amplificação onde a porção amplificável do PNA contém sequências sobre as quais podem ser polimerizados ácidos nucleicos de reforço (BNA). Após a formação de cada híbrido BNA-PNA, forma-se uma região de ligação de reforço (BBR) à qual um conjunto de ligação de reforço (BBA) se liga

especificamente. Se marcados de forma detectável, os BBA ou BNA proporcionam uma amplificação essencialmente ilimitada do acontecimento de ligação TNA-PNA original.

De acordo com este invento, entender-se-á que o TNA inclui sequências de ácido nucleico específicas. O TBA entender-se-á ser qualquer conjunto molecular que possa de forma específica e apertada ligar-se a um híbrido TNA-PNA formado. O TBA conterá uma ou mais moléculas cujas sequências são suficientes para se ligarem à TBR. Os domínios de ligação ao ácido nucleico que são conhecidos podem ser utilizados directamente como componentes do TBA ou modificados de acordo com os ensinamentos aqui proporcionados. As moléculas mais facilmente disponíveis com tais sequências são os domínios de ligação ao ADN das proteínas de ligação ao ADN. Especificamente, conhecem-se muitas proteínas de ligação ao ADN ou ao ARN que podem ser utilizadas directamente como a proteína não modificada conhecida, ou o TBA pode ser uma proteína de ligação ao ácido nucleico modificada de acordo com os ensinamentos específicos aqui proporcionados. No último caso, as modificações específicas que são desejáveis incluiriam a optimização de afinidades de ligação, remoção de actividades indesejadas (tais como actividade de nuclease e reorganização do TBA na presença de outras moléculas com uma afinidade para componentes do TBA), optimização da selectividade de uma sequência alvo sobre sequências intimamente aparentadas, e optimização da estabilidade.

Exemplos de proteínas de ligação ao ADN que podiam ser utilizadas de acordo com este invento são as porções de ligação ao ADN do factor de transcrição NF-kB (p50 e p65), NF-IL6, NF-AT, rel, TBP, a proteína E2 do vírus do papiloma humano, sp1, os repressores *cro* e CI do bacteriófago lambda e proteínas semelhantes bem como proteínas conhecidas cuja porção de ligação ao ADN foi isolada, clonada, sequenciada e caracterizada. Adicionalmente, é incluída qualquer outra proteína de ligação ao ADN ou porção de uma proteína que seja necessária e suficiente para se ligar a um híbrido de TBR ou a uma BBR. Isto inclui proteínas ou porções de proteínas de tipo selvagem com actividade de ligação ao ADN alterada bem como proteínas criadas com especificidade de ligação ao ADN alterada, tal como a troca de uma hélice de reconhecimento da

ligação ao ADN de uma proteína para outra. Adicionalmente, proteínas que exibem ligação ao ácido nucleico e outras funções no ácido nucleico, tais como endonucleases de restrição, podem ser utilizadas como função de ligação ao ácido nucleico. Proteínas que se ligam a regiões alvo em híbridos ADN-ARN bem como híbridos ARN-ARN estão incluídas (ver, por exemplo, Shi, 1995; DeStefano, 1993, Zhu, 1995; Gonzales, 1994; Salazar, 1993; Jaishree, 1993; Wang, 1992; Roberts, 1992; Kainz, 1992; Salazar, 1993 (b)). Os conjuntos de ligação podem ser construídos com a utilização de uma molécula que acompanha porções do conjunto de ligação de modo a possam ser alcançadas combinações e geometrias de componentes específicas. Esta molécula é aqui designada como um PILOTO. Os pilotos podem ser constituídos por proteínas ou qualquer combinação de materiais orgânicos e inorgânicos que alcancem a selecção combinatória e/ou induzam geometrias específicas entre membros do TBA ou BBA. Um acompanhante é uma estrutura estável sobre a qual pode ser construído um TBA ou BBA de modo a que seja proporcionada a conformação correcta do TBA ou BBA enquanto que ao mesmo tempo elimina propriedades indesejáveis de uma proteína de ligação a ácido nucleico de ocorrência natural. Como exemplo específico desta concretização, é proporcionada uma versão modificada do factor de transcrição pleiotrópico, NF-kB, utilizando uma proteína cro do bacteriófago lambda modificada como acompanhante. Cada dímero de ligação de NF-kB retém a afinidade de ligação picomolar para o local de ligação de NF-kB enquanto que ao mesmo tempo o conjunto de ligação apresenta várias características vantajosas de produção, estabilidade e especificidade.

Tendo em vista o antecedente, os vários aspectos e concretizações deste invento são descritos abaixo em detalhe.

1. Os Ácidos Nucleicos Sonda (PNA) e sua preparação. Os PNA do presente invento compreendem pelo menos três partes principais unidas umas às outras. Com referência à Figura 1(I) dos desenhos, a primeira parte do PNA é uma ou mais sequências de bases, designada "1/2 TBR". Com referência à Figura 1(I e IIa) dos desenhos, a 1/2 TBR no PNA é complementar a uma sequência de interesse numa amostra, contendo o TNA uma 1/2 TBR. Com referência à Figura 1(IIIa) dos desenhos, o TNA,

quando adicionado ao PNA sob condições de hibridação, forma um híbrido PNA-TNA contendo uma TBR. Com referência à Figura 1(I) dos desenhos, a segunda parte do PNA é uma sequência de bases, designada "1/2 BBR". Com referência à Figura 1(I, IIb, IIc e IVa) dos desenhos, a 1/2 BBR no PNA é complementar a uma 1/2 BBR contida num BNA ou num HNA. Com referência à Figura 1(IIIb, IIIc e Va) dos desenhos, o BNA ou HNA, quando adicionado ao PNA sob condições de hibridação, forma um híbrido PNA-BNA ou um híbrido PNA-HNA, respectivamente, contendo uma BBR. Com referência à Figura 1(I) dos desenhos, a terceira parte do PNA é o OSA, designado por um círculo dentro de uma caixa. O OSA é a ausência de suporte e/ou de um indicador, ou um suporte sólido ou outro meio de localização, incluindo, mas não se limitando a, ligação a esferas, polímeros e superfícies e/ou indicadores que é/são ligados covalentemente ou não covalentemente, mas especificamente, associados ao PNA. O OSA pode ser um átomo ou molécula que ajuda na separação e/ou localização tal como um grupo de ligação a um suporte sólido ou marcador que pode ser detectado através de vários meios físicos incluindo, mas não se limitando a, adsorção ou imagiologia de partículas emitidas ou luz. Os métodos para unir indicadores a oligonucleótidos ou para imobilizar oligonucleótidos a suportes sólidos são bem conhecidos na arte (ver Keller e Manak, *supra*).

O PNA do presente invento pode ser preparado através de qualquer método adequado. Tais métodos, em geral, incluirão síntese de oligonucleótidos e clonagem num vector replicável. Os métodos para a síntese de ácidos nucleicos são bem conhecidos na arte. Quando clonados ou sintetizados, a purificação da cadeia e a separação podem ser necessárias para utilizar o produto como um PNA puro. Os métodos de preparação de sondas de ARN são bem conhecidos (ver por exemplo Blais 1993, Blais 1994, que utiliza transcrição *in vitro* a partir de uma reacção PCR incorporando um promotor da ARN-polimerase de T7).

Será entendido pelos peritos na arte que o comprimento e a sequência específica do PNA dependem do comprimento e da sequência a detectar num TNA, e das restrições para se alcançar uma ligação apertada e específica do TBA particular a utilizar (ver discussão sobre TBA abaixo). Em geral, PNA com

comprimentos de sequência entre cerca de 10 e cerca de 300 nucleótidos de comprimento são adequados, sendo os de cerca de 5-100 nucleótidos de comprimento desejáveis para muitas das concretizações especificamente exemplificadas aqui.

Deve entender-se também que o PNA pode ser construído de modo a conter mais de uma 1/2 TBR e a produzir mais de uma TBR para um ou mais TBA, iguais ou diferentes, bem como TBR complexas reconhecidas por novos TBA dúplice e múltipla (ver descrição abaixo em relação a estes novos TBA) após hibridação dos PNA e TNA. A Figura 5 ilustra PNA específicos que contêm uma ou mais 1/2 TBR. As sequências específicas que correspondem às sequências de 1/2 TBR ilustradas na Figura 5 (Ia, IIa, IIIa, IVa e Va) são SEQ ID NO:1-34 (ver Descrição das Sequências acima).

Tal como mostrado nas Figuras 2a e 2b, o PNA, contendo uma 1/2 TBR, pode ser hibridado com um ou mais BNA (ver descrição abaixo) e a cadeia de BNA polimerizada até qualquer potencial comprimento desejado para amplificação do acontecimento de hibridação TNA-PNA. De preferência, estarão presentes no PNA entre 1 e 10 1/2 BBR.

Tal como mostrado nas Figuras 6a e 6b, o PNA pode conter várias 1/2 TBR, iguais ou diferentes, que podem hibridar com várias 1/2 TBR num TNA. De cada vez que uma 1/2 TBR no PNA coincide com uma 1/2 TBR num TNA, forma-se uma Região de Ligação Alvo, TBR, que se pode ligar a uma TBA. Para além disso, não é essencial que todas as TBR estejam num único PNA contíguo. Assim, numa concretização do presente invento, são utilizados dois PNA diferentes para detectar sequências num determinado TNA. Como ilustração deste aspecto do presente invento, a Figura 7 mostra uma representação da repetição terminal longa (LTR) do vírus da imunodeficiência humana (HIV). Tal como se sabe na arte, a LTR de HIV compreende dois locais de ligação a NF-kB e três locais de ligação a SP1, muito próximos, em que NF-kB e SP1 são proteínas de ligação ao ADN conhecidas. A Figura 7 proporciona dois PNA, PNA1 (SEQ ID NO:38) e PNA2 (SEQ ID NO:39), cada um dos quais é complementar à cadeia oposta mostrada como um TNA (SEQ ID NO:37), que mostra os dois locais de ligação a NF-kB e os três locais de ligação a SP1 da LTR de HIV. De acordo com este aspecto do

presente invento, PNA1 hibrida especificamente com a secção do TNA mostrada na Figura 7 com as bases sublinhadas com um símbolo "+", enquanto que PNA2 hibrida especificamente com a secção do TNA mostrada na Figura 7 com as bases sublinhadas com um símbolo "=". Cada um de PNA1 e PNA2 pode também conter sequências (indicadas pelos símbolos "#" ou "**") que hibridarão com sequências 1/2 BBR de um BNA (ver abaixo). Adicionalmente, cada um de PNA1 e PNA2 pode ser marcado de modo diferencial com um OSA, tal como um fluoróforo como um marcador de fluoresceína ou rodamina, que permitirá a confirmação de que ambas as sondas se ligaram ao TNA. Se for detectada apenas um ou nenhum marcador, conclui-se que o TNA não está presente na amostra a testar.

Noutro aspecto da concretização mostrada na Figura 7, é mostrado um método para alterar a especificidade do presente método de ensaio. Através da alteração do comprimento do intervalo entre PNA1 e PNA2, de modo a que seja alterada a região de TNA que permanece sem hibridar, um executante deste invento é capaz de alterar a discriminação do ensaio.

Para exemplificar mais claramente este aspecto do presente invento, é necessário enfatizar que a TBR pode ter uma estrutura helicoidal. Assim, enquanto PNA1 cria TBR numa "face" da hélice, PNA2 cria uma TBR na mesma face ou numa face diferente da hélice, dependendo da distância entre o meio de cada TBR (sublinhado na Figura 7). Se o meio de cada local de ligação for um produto inteiro de 10,5 bases de distância, as TBR estarão no mesmo lado da hélice, enquanto que produtos não inteiros de 10,5 bases de distância colocarão as TBR em lados opostos da hélice. Deste modo, qualquer cooperatividade na ligação pelo TBA reconhecendo a TBR de PNA1 e pelo TBA reconhecendo a TBR de PNA2 pode ser manipulada (ver Hochschild, A., M. Ptashne, *Cell* 44: 681-687, 1986, mostrando este efeito para a ligação do repressor do bacteriófago lambda a dois locais diferentes do operador situados a diferentes distâncias um do outro numa hélice de ADN). Tal como descrito por Perkins et al. (*EMBO J.* 12: 3551-3558, 1993), é necessária cooperatividade entre os locais de NF-kB e SP1 para se alcançar a activação da LTR de HIV. No entanto, para os fins do presente invento, pode-se tirar vantagem do motivo do local de ligação duplo NF-kB-triplo SP1 na LTR de HIV proporcionando

uma nova proteína de ligação única capaz de se ligar a ambos os locais simultaneamente, mas apenas se o espaço entre os locais for geometricamente praticável. Isto é controlado tanto pela estrutura do TBA seleccionado como pelo TNA utilizado. Assim, na concretização exemplificada na Figura 7, as duas sondas podem ser utilizadas com uma região inter-sondas suficientemente grande de ADN de cadeia simples restante de modo a que, mesmo se os locais de ligação a NF-kB e SP1 estiverem em lados opostos da hélice, a região de cadeia simples entre as sondas proporciona uma "charneira" suficientemente flexível de modo a que o ADN se possa dobrar e torcer para acomodar a geometria do TBA. Alternativamente, pode ser desenhado um ensaio mais rigoroso através do estreitamento da distância inter-sondas de modo a que o ADN se possa apenas dobrar, mas não torcer. Finalmente, as sondas podem ser tão pouco espaçadas ou ser utilizado um único PNA, de modo a que o ADN se possa dobrar mas não torcer. Assim, esta figura exemplifica e permite a produção de sistemas de detecção com qualquer grau desejado de discriminação entre os ácidos nucleicos alvo possuindo sequências semelhantes, mas diferentes justaposições destas sequências.

Em termos de um estojo de diagnóstico ou forense para HIV, os peritos na arte compreenderão que os aspectos acima mencionados deste invento permitem a obtenção de componentes do estojo de diagnóstico ou forense à medida para corresponder ao que se sabe a dado momento sobre as estirpes prevalentes de HIV ou de outro patógeno ou condição de doença. Será apreciado pelos peritos na arte que, embora a detecção de infecção de HIV não seja a única utilidade do presente invento, devido à mutabilidade do genoma de HIV, é provavelmente um dos ambientes de teste mais complexos para um tal diagnóstico. No entanto, é precisamente num tal ambiente mutável que a flexibilidade do presente método, ligada à sua capacidade para discriminar entre sequências muito intimamente aparentadas, podem ser melhor apreciadas. Em ambientes menos mutáveis, alguma da sofisticação à qual este método se sujeita não necessita ser utilizada. Assim, num estojo de diagnóstico para infecção de vírus do papiloma, todas as características de discriminação da interacção TBA-TBR estão disponíveis, juntamente com a capacidade para amplificar o sinal utilizando os BNA e BBA, mas pode ser utilizado um único PNA simples, tal

como qualquer um de SEQ ID NO:44-62, que identifique sequências únicas do vírus do papiloma, que também se saiba que se ligam a um TBA tal como a proteína E2 do vírus do papiloma ou porções de ligação ao ADN truncadas desta (ver Hedge et al., *Nature* 359: 505-512, 1992; Monini et al., *J. Virol.* 65: 2124-2130, 1991).

Na aplicação do presente método à detecção de um determinado TNA para fins de avaliação de se estão presentes certos ácidos nucleicos que estão associados com a progressão de melanoma, hepatoma, cancros da mama, cervical, do pulmão, do cólon, da próstata, pancreático ou dos ovários, o TNA pode ser obtido a partir de materiais de biopsia tomados de órgãos e fluidos suspeitos de conter as células cancerosas. Para a detecção de deficiências genéticas, o TNA pode ser obtido a partir de amostras de pacientes contendo as células afectadas. Para detecção de contaminantes e produtos de fermentação no fabrico de alimentos, produtos químicos ou biotecnológicos ou na bio-resolução de desperdícios, o TNA pode ser obtido a partir de amostras tomadas a vários estádios no processo de fermentação ou tratamento. Para a detecção de patogénios ou contaminantes em alimentos ou fármacos, a amostra de TNA pode ser obtida a partir do alimento ou fármaco, esfregaços de alimentos ou superfícies em contacto com os alimentos, fluidos em contacto com os alimentos, materiais de processamento, fluidos e semelhantes associados ao fabrico ou em contacto com os alimentos, fármacos ou amostras biológicas tomadas daqueles em contacto com os alimentos ou fármacos ou semelhantes.

2. Ácidos Nucleicos de Reforço (BNA), Regiões de Ligação de Reforço (BBR) e sua preparação. Os BNA do presente invento são constituídos por pelo menos uma ou mais 1/2 BBR ligadas a um OSA. As 1/2 BBR podem hibridar com 1/2 BBR complementares contidas no PNA, outros BNA ou num HNA.

Com referência à Figura 1 (I, IIb e IIIb) dos desenhos, o BNA mais simples é constituído por duas partes. Com referência à Figura 1(IIb) dos desenhos, a primeira parte do BNA mais simples é uma sequência de bases que é complementar à sequência no PNA que é designada "1/2 BBR". Com referência à Figura 1(IIb) dos desenhos, a segunda parte do BNA mais simples é o OSA, designado por um círculo dentro de uma caixa.

O OSA é ausência de suporte e/ou de indicador, ou um suporte sólido ou outro meio de localização, incluindo, mas não se limitando a, ligação a esferas, polímeros e superfícies e/ou indicadores que são ligados covalentemente ou não covalentemente, mas especificamente, associados ao BNA.

Com referência à Figura 2a (II e III) dos desenhos, o BNA pode conter mais de uma sequência de 1/2 BBR. O BNA ilustrado na Figura 3(II) contém uma sequência que é complementar ao PNA ilustrado na Figura 3(I) e duas outras sequências de 1/2 BBR. O BNA ilustrado na Figura 3(III) contém duas sequências de 1/2 BBR que são complementares a duas das sequências de 1/2 BBR no BNA ilustrado na Figura 3(II), mais até "n" 1/2 BBR adicionais para polimerização de mais BNA.

Sob condições de hibridação, o BNA ilustrado na Figura 3(II), quando combinado com o PNA ilustrado na Figura 3(I), cria o híbrido PNA-BNA ilustrado na Figura 3(IVa) contendo uma BBR e uma extensão não hibridada com duas sequências de 1/2 BBR adicionais ou sequências de "reforço". As BBR criadas através da referida hibridação podem ser idênticas, semelhantes ou dissemelhantes na sequência. As BBR criadas através da referida hibridação podem ligar-se a BBA idênticas, semelhantes ou dissemelhantes (ver abaixo). Os BNA podem ter sido preparados de forma análoga aos PNA.

Sob condições de hibridação, o híbrido BNA-BNA ilustrado na Figura 3(IVb), quando combinado com o PNA ilustrado na Figura 3(Vb), cria o híbrido PNA-BNA ilustrado na Figura 3(VI) contendo uma BBR, dois híbridos BNA-BNA adicionais contendo BBR e uma extensão não hibridada com uma sequência de 1/2 BBR adicional, uma sequência de "reforço". As BBR criadas pela referida hibridação podem ser idênticas, semelhantes ou dissemelhantes na sequência. As BBR criadas pela referida hibridação podem-se ligar a BBA idênticas, semelhantes ou dissemelhantes (ver abaixo). Os BNA podem ser preparados de um modo análogo ao da preparação dos PNA.

3. Os Ácidos Nucleicos Alvo (TNA) e sua preparação. O primeiro passo na detecção e amplificação de sinais produzidos através da detecção de um determinado TNA de acordo com o presente método é a hibridação de tal alvo com o PNA numa

mistura adequada. Tal hibridação é alcançada sob condições adequadas bem conhecidas na arte.

A amostra suspeita de conter ou que se sabe conter o TNA pretendido pode ser obtida a partir de uma variedade de fontes. Pode ser uma amostra biológica, uma amostra de alimento ou agrícola, uma amostra ambiental e por aí adiante. Na aplicação do presente método para a detecção de um determinado TNA para fins de diagnóstico médico ou forense, o TNA pode ser obtido a partir de uma amostra de biopsia, um fluido corporal ou exsudado tal como urina, sangue, leite, líquido cerebrospinal, expectoração, saliva, fezes, aspirados pulmonares, esfregaços da garganta ou genitais e semelhantes. Adicionalmente, a detecção pode ser *in situ* (ver por exemplo Embretson, 1993, Patterson, 1993; Adams, 1994).

Assim, PNA específicos para vertebrados (incluindo mamíferos e incluindo humanos) ou para qualquer um ou para todos os seguintes microrganismos de interesse podem ser considerados e utilizados de acordo com o presente método:

Corinebactérias

Corynebacterium diphtheria

Bacilos

Bacillus thuringiensis

Pneumococos

Diplococcus pneumoniae

Estreptococos

Streptococcus pyogenes

Streptococcus salivarius

Estafilococos

Staphylococcus aureus

Staphylococcus albus

Pseudomonas

Pseudomonas stutzeri

Neissérias

Neisseria meningitidis

Neisseria gonorrhoea

Enterobacteriáceas*Escherichia coli**Aerobacteria aerogenes**Klebsiella pneumoniae*

As bactérias coliformes

*Salmonella typhosa**Salmonella choleraesuis*As *Salmonellae**Salmonella typhimurium**Shigellae dysenteriae**Shigellae schmitzii**Shigellae arabinotarda**Shigellae flexneri*As *Shigellae**Shigellae boydii**Shigellae sonnei*Outros bacilos entéricos*Proteus vulgaris**Proteus mirabilis*Espécies de *Proteus**Proteus morgani**Pseudomonas aeruginosa**Alcaligenes faecalis**Vibrio cholerae*Grupo Hemófilos-Bordetelas*Hemophilus influenza, H. Ducryi**Hemophilus hemophilus**Hemophilus aegypticus**Hemophilus parainfluenzae**Bordetella pertussis*Pasteurelas*Pasteurella pestis**Pasteurella tularensis*Brucelas*Brucella melitensis**Brucella abortus**Brucella suis*

Bacilos Aeróbicos Formadores de Esporos

Bacillus anthracis

Bacillus subtilis

Bacillus megaterium

Bacillus cereus

Bacilos Anaeróbicos Formadores de Esporos

Clostridium botulinum

Clostridium tetani

Clostridium perfringens

Clostridium novyi

Clostridium septicum

Clostridium histolyticum

Clostridium tertium

Clostridium bifermentans

Clostridium sporogenes

Micobactérias

Mycobacterium tuberculosis hominis

Mycobacterium bovis

Mycobacterium avium

Mycobacterium leprae

Mycobacterium paratuberculosis

Actinomicetos (bactérias tipo fungo)

Actinomyces israeli

Actinomyces bovis

Actinomyces naeslundii

Nocardia asteroides

Nocardia brasiliensis

As Espiroquetas

Treponema pallidum

Treponema pertenue

Treponema carateum

Spirillum minus

Streptobacillus moniliformis

Borrelia recurrentis

Leptospira icterohemorrhagiae

Leptospira canicola

TripanassomasMicoplasmas*Mycoplasma pneumoniae*Outros patogénios*Listeria monocytogenes**Erysipelothrix rhusiopathiae**Streptobacillus moniliformis**Donvania granulomatis**Bartonella bacilliformis*Rickettsias (parasitas tipo bactérias)*Rickettsia prowazekii**Rickettsia mooseri**Rickettsia rickettsii**Rickettsia conori**Rickettsia australis**Rickettsia sibiricus**Rickettsia akari**Rickettsia tsutsugamushi**Rickettsia burnetti**Rickettsia quintana*Clamídias (parasitas bacterianos/virais inclassificáveis)agentes *Chlamydia* (denominação incerta)Fungos*Cryptococcus neoformans**Blastomyces dermatidis**Histoplasma capsulatum**Coccidioides immitis**Paracoccidioides brasiliensis**Candida albicans**Aspergillus fumigatus**Mucor corymbifera* (*Absidia corymbifera*)*Rhizopus oryzae**Rhizopus arrhizus**Rhizopus nigricans**Sporotrichum schenkii**Flonsecaea pedrosoi**Fonsecaea compact*

Phycomycetes

Fonsecaea dermatidis
Cladosporium carrioni
Phialophora verrucosa
Aspergillus nidulans
Madurella mycetomi
Madurella grisea
Allescheria boydii
Phialophora jeanselmei
Microsporum gypsum
Trichophyton mentagrophytes
Keratinomyces ajelloi
Microsporum canis
Trichophyton rubrum
Microsporum adouini

Vírus

Adenovírus

Herpesvírus

Herpes simplex
Varicela (Chicken pox)
Herpes zoster (Shingles)
Vírus B
Citomegalovírus

Poxvírus

Variola (smallpox)
Vaccinia
Poxvirus bovis
Paravaccinia
Molluscum contagiosum

Picornavírus

Poliovírus
Vírus Coxsackie
Ecovírus
Rinovírus

Mixovírus

Influenza (A, B e C)
Parainfluenza (1-4)
Vírus da papeira
Vírus da Doença de Newcastle

Vírus do sarampo
Vírus da peste bovina (*Rinderpest*)
Vírus da esgana canina
Vírus sincicial respiratório
Vírus da rubéola

Arbovírus

Vírus da encefalite equina Oriental
Vírus da encefalite equina Ocidental
Vírus Sindbis
Vírus Chikugunya
Vírus da floresta Semliki
Vírus de Mayora
Vírus da encefalite de St. Louis
Vírus da encefalite da Califórnia
Vírus da febra da carraça do Colorado
Vírus da febre amarela
Vírus do Dengue

Reovírus

Reovírus tipos 1-3

Retrovírus

Vírus da imunodeficiência humana (HIV)
Vírus linfotrófico de células T humanas I & II (HTLV)

Hepatite

Vírus da hepatite A
Vírus da hepatite B
Vírus da hepatite nãoA-nãoB
Hepatite, C, D, E

Vírus Tumorais

Vírus da leucemia de Rauscher
Vírus Gross
Vírus da leucemia de Maloney
Vírus dos papilomas humanos

Será entendido por um perito na arte que é geralmente necessário tratar amostras suspeitas de conter um determinado TNA de um tal modo de forma a produzir fragmentos que possam facilmente hibridar com o PNA. Pode ser necessário tratar a amostra de teste para efectuar a libertação ou para extrair o

TNA para hibridação, tal como expondo sangue ou outras células a um ambiente hipotónico, ou rompendo de outro modo a amostra utilizando meios mais vigorosos. Quando se pensa que o TNA está presente na forma de cadeia dupla, será naturalmente desejável separar as cadeias para tornar o TNA hibridável na forma de cadeia simples através de métodos bem conhecidos na arte, incluindo mas não se limitando a aquecimento ou exposição limitada a condições alcalinas que possam ser neutralizadas após a adição do PNA de cadeia simples para permitir que a hibridação ocorra. Os métodos para preparação de alvos de ARN são bem conhecidos (ver Waterhouse, 1993; Mitchell, 1992).

A fragmentação de amostras de ácido nucleico contendo TNA é habitualmente necessária para diminuir a viscosidade da amostra e para aumentar a acessibilidade dos TNA aos PNA. Tal fragmentação é alcançada através de meios aleatórios ou específicos conhecidos na arte. Assim, por exemplo, nucleases específicas que se saiba cortarem com uma determinada frequência no genoma particular a analisar, podem ser utilizadas para produzir fragmentos de um tamanho molecular médio conhecido. Adicionalmente, outras nucleases, fosfodiesterases, exonucleases e endonucleases, fragmentação física e ultra-sons são todos métodos possíveis para este fim. Estes processos são bem conhecidos na arte. A utilização de enzimas de restrição para fins de fragmentação do ADN é geralmente preferida. No entanto, o ADN pode também ser fragmentado através de uma variedade de meios químicos tais como a utilização dos seguintes tipos de reagentes: EDTA-Fe(II) (de acordo com Stroebel et al., *J. Am. Chem. Soc.* 110: 7927, 1988; Dervan, *Science* 232: 464, 1986); Cu(II)-fenantrolina (de acordo com Chen e Sigman, *Science* 237: 1197, 1987); enzima de restrição de classe IIS (de acordo com Kim et al., *Science* 240: 504, 1988); ADNase híbrida (de acordo com Corey et al., *Biochem.* 28: 8277, 1989); bleomicina (de acordo com Umezawa et al., *J. Antibiot. (Tokyo) Ser. A.* 19: 200, 1986); neocarzinostatina (Goldberg et al., "Second Annual Bristol-Myers Symposium in Cancer Research", Academic Press, New York, pág. 163, 1981); e metidiumpropil-EDTA-Fe(II) (de acordo com Hrtzberg et al., *J. Am. Chem. Soc.* 104: 313, 1982). A remoção de proteínas, como através do tratamento com uma protease, é também geralmente desejável e os métodos para

efectuar a remoção das proteínas de amostras de ácidos nucleicos, sem perda apreciável de ácido nucleico, são bem conhecidos na arte.

Os TNA do presente invento devem ser suficientemente longos para que exista uma quantidade suficiente de híbrido de cadeia dupla a flanquear a TBR de modo a que um TBA se possa ligar sem ser perturbado pelas extremidades dos fragmentos não ligados. Tipicamente, são utilizados fragmentos no intervalo de cerca de 10 nucleótidos a cerca de 100 000 nucleótidos, e de preferência no intervalo de cerca de 20 nucleótidos a cerca de 1000 nucleótidos como tamanho médio para fragmentos de TNA. Exemplos de sequências específicas de TNA que podem ser detectadas são sequências complementares às sequências de PNA aqui descritas para a detecção de sequências de ácido nucleico celulares normais, celulares anormais (como em oncogenes activados, genes estranhos integrados, genes geneticamente deficientes) e específicas de patogénios, para as quais são conhecidas proteínas de ligação a ácido nucleico específicas, ou que podem ser produzidas de acordo com métodos descritos nesta divulgação. Com referência à Figura 7, um TNA relacionado com HIV específico é mostrado como SEQ ID NO:37.

4. Extensões do PNA utilizando BNA, sua preparação e amplificação do sinal. Sob condições de hibridação, podem ser adicionados BNA que hibridam com os PNA, híbridos PNA-BNA, BNA e/ou híbridos BNA-BNA. As adições acima mencionadas podem ser feitas de um modo polimérico não vectorial ou de um modo vectorial, com uma ordem conhecida de BNA.

Com referência à Figura 2a, é apresentado um reforço simples. Um polímero de reforço é produzido através da adição de dois BNA, ilustrados na Figura 2a (Ib e Ic), que quando combinados sob condições de hibridação com o PNA, formam os híbridos PNA-BNA-BNA, constituídos pelo PNA e pelas "extensões de reforço", ilustrados na Figura 2a (IIa, IIb, IIc e IID) deixando pelo menos uma sequência de 1/2 BBR desemparelhada. Cada sequência de 1/2 BBR desemparelhada, ilustrada na Figura 2a (IIa, IIb, IIc e IID), pode hibridar com BNA adicionais para formar extensões de "reforço" adicionais. Cada sequência de 1/2 BBR desemparelhada, ilustrada na Figura 2a (IIa, IIb, IIc e IID) pode hibridar com HNA adicionais,

ilustrados na Figura 2a (IIIa e IIIb). A hibridação dos HNA, que não podem hibridar com BNA adicionais, actua "rematando" da adição dos BNA no PNA, tal como ilustrado na Figura 2a (IVa, IVb, IVc e IVd).

Com referência à Figura 2b, é possível controlar e especificar a ordem e os componentes das extensões do PNA. Se for necessária uma única BBR, um HNA contendo a sequência complementar à 1/2 BBR no PNA é adicionado ao PNA para produzir uma única BBR e para "rematar" evitando quaisquer extensões de "reforço" no PNA. Se houver BBR adicionais a adicionar ao PNA, pode ser alcançada uma extensão controlada do PNA.

Com referência à Figura 2b, é apresentado um reforço simples. A extensão polimérica vectorial é alcançada através da adição de um BNA que seja específico para o PNA, tal como ilustrado na Figura 2b (Ia e IIa), que quando combinado sob condições de hibridação com o PNA, forma híbridos PNA-BNA-BNA, constituídos pelo PNA e extensões de "reforço". Estas extensões, se marcadas com um OSA, proporcionam um método para amplificar grandemente qualquer sinal produzido após a ligação de um PNA a um TNA na amostra. Para além disso, através da ligação de BBA marcados às BBR do polímero, é alcançada uma amplificação adicional.

Qualquer um de vários métodos pode ser utilizado para preparar os BNA, incluindo, p.ex., síntese através de química conhecida ou através de métodos de produção de ADN recombinante. No último método, pode ser produzido um número essencialmente ilimitado de BNA de modo simples e económico, por exemplo, através da produção em procariotas (*E. coli* por exemplo) de um ADN plasmídico possuindo múltiplas repetições das sequências de BNA específicas flanqueadas por locais de restrição possuindo extremidades soltas. Deste modo, por exemplo, os locais operadores esquerdo ou direito do bacteriófago lambda, ou qualquer outro ADN ou outras sequências de ácido nucleico que se saiba ligarem-se especificamente e fortemente a um determinado BBA, tal como uma proteína de ligação a ADN ou ARN, podem ser produzidas num número de cópias essencialmente ilimitado, com cada cópia flanqueada por um *EcoRI*, *PstI*, *BamHI* ou qualquer um de vários

outros locais de nucleases de restrição vulgares. Alternativamente, um polímero em locais repetidos pode ser excisado através de locais de restrição únicos não presentes dentro do polímero. Grandes quantidades de plasmídeo pBR322, pUC ou outro plasmídeo possuindo múltiplas cópias destas sequências são produzidas através de métodos bem conhecidos na arte, o plasmídeo é cortado com a enzima de restrição flanqueando o local polimerizado e as múltiplas cópias libertadas dos operadores são isoladas através de cromatografia ou de quaisquer outros meios convenientes conhecidos na arte. O BNA, antes da utilização, é então separado da cadeia e é então passível de polimerização sobre um PNA codificando uma cópia complementar de cadeia simples do operador como uma 1/2 BBR. Os BNA podem ser polimerizados vectorialmente sobre o PNA através da utilização de diferentes enzimas de restrição para flanquear cada repetição do polímero no plasmídeo utilizado para produzir múltiplas cópias do BNA. Alternativamente, o polímero de BNA pode ser hibridado com o PNA através de extremidades soltas numa ou em ambas as extremidades do polímero de BNA, sem necessidade de separar as cadeias e de hibridar cada segmento de BNA. Exemplos de sequências de BNA específicas são proporcionados acima na secção intitulada Descrição das Sequências, como SEQ ID NO:35-36. Para estabilizar o polímero de BNA, pode ser utilizada ADN-ligase para ligar de modo covalente os BNA hibridados.

5. Os Ácidos Nucleicos em "gancho de cabelo" (HNA) e sua preparação. Os HNA aqui descritos compreendem pelo menos duas partes principais unidas uma à outra: Uma sequência de cadeia simples, que é complementar a uma 1/2 BBR, e uma região de ácido nucleico de cadeia dupla formada, sob condições de hibridação, pela auto-associação de sequências auto-complementares dentro do HNA. Com referência à Figura 1 (IIc) dos desenhos, a 1/2 BBR no HNA pode ser construída de modo a ser complementar à 1/2 BBR no PNA. Com referência à Figura 1 (I, IIc e IIIc) dos desenhos, o HNA acima mencionado, quando adicionado ao PNA sob condições de hibridação, forma um híbrido PNA-HNA contendo uma BBR. Com referência à Figura 1 (IIIc, IVc e Vc) dos desenhos, um híbrido PNA-HNA, sob condições de hibridação, após a adição do TNA, pode formar um híbrido TNA-PNA-HNA contendo uma TBR e uma BBR.

Com referência à Figura 2a e 2b, os HNA podem ser utilizados para "rematar" ou terminar a adição de extensões de BNA ao PNA. Os dois BNA da Figura 2a (Ib e Ic) podem associar-se para formar o híbrido mostrado na Figura 3(IVb) ou podem hibridar directamente e individualmente com o PNA tal como ilustrado na Figura 2a (Ia-c, IIa-d). Os dois HNA (mostrados na Figura 2a (IIIa e IIIb)) podem terminar a hibridação do BNA com outros BNA que se estendem do PNA, tal como ilustrado na Figura 2a (IVa-d). O HNA na Figura 2a (IIIa) pode terminar os híbridos PNA-BNA mostrados na Figura 2a (IIb e IId) e qualquer híbrido PNA-BNA com uma 1/2 BBR de cadeia simples que seja complementar à 1/2 BBR no HNA ilustrado na Figura 2a (IIIa). De modo semelhante, o HNA na Figura 2a(IIIb) pode terminar os híbridos PNA-BNA mostrados na Figura 2a (IIa e IIc) e qualquer híbrido PNA-BNA com duas 1/2 BBR de cadeia simples que sejam complementares às 1/2 BBR no HNA ilustrado na Figura 2a(IIIb).

São construídos HNA que terminarão os híbridos PNA-BNA que são construídos a partir da adição sequencial de BNA ao PNA tal como ilustrado na Figura (2b). As sequências de cadeia simples de 1/2 BBR ilustradas na Figura 2b (Ia, IIIa, Va e VIIa) são especificamente complementares às sequências de 1/2 BBR de cadeia simples ilustradas na Figura 2b (Ib, IIIb, Vb e VIIb) e produzem os híbridos PNA-BNA-HNA rematados únicos ilustrados na Figura 2b (Ic, IIIc, Vc e VIIc).

As sequências auto-complementares no HNA e a sequência de volta que se liga às sequências em "gancho de cabelo" auto-complementares podem ser de qualquer composição e comprimento, desde que não impeçam substancialmente nem inibam a apresentação da 1/2 BBR de cadeia simples que compreende parte do HNA pelo HNA ou se liguem selectivamente ao BBA ou ao TBA. As sequências de volta devem ser seleccionadas de modo a que a formação da volta não impeça a formação do "gancho de cabelo". Um exemplo de um HNA útil neste pedido é proporcionado como SEQ ID NO:44 (ver Descrição das Sequências acima).

6. Os Conjuntos de Ligação ao Alvo (TBA) e sua preparação. Um TBA pode ser qualquer substância que se liga a uma determinada TBR formada através de hibridação de determinados TNA e PNA, desde que o TBA tenha pelo menos os seguintes atributos:

- a) O TBA tem de se ligar à(s) TBR de um modo que seja altamente específico para a(s) TBR de interesse. Isto é, o TBA tem de discriminar entre as TBR presentes no híbrido TNA-PNA e sequências dúplex semelhantes formadas por híbridos PNA-CNA. O TBA tem de se ligar ao híbrido PNA-CNA com uma avides suficientemente baixa para que após lavagem do complexo TBA-TNA-PNA, o híbrido PNA-CNA seja deslocado e o híbrido PNA-TNA não seja deslocado;
- b) O TBA tem de se ligar de forma ávida à(s) TBR criada(s) através da hibridação do TNA com o PNA. Afinidades de ligação no intervalo de 10^{-5} a cerca de 10^{-12} ou mais são geralmente consideradas suficientes. Tal como observado abaixo, nalguns casos pode ser desejável utilizar um determinado TBA que possua uma avides muito baixa para uma determinada TBR, mas que possua uma afinidade muito maior quando é proporcionada uma determinada configuração de múltiplas TBR de modo a que o quadrado da afinidade do TBA para cada TBR se torne a afinidade relevante para esse TBA em particular.

Exemplos de componentes de ligação ao ADN úteis na formação dos TBA incluem, mas não se limitam a NF-kB, proteína E2 do vírus do papiloma, factor de transcrição SP1, enzimas de restrição inactivas, anticorpos, etc. Foi reconhecido na arte que cada uma destas proteínas contém sequências que se ligam a determinadas sequências de ácido nucleico e as afinidades destas interacções são conhecidas. Naturalmente, o método do presente invento não se limita à utilização destas proteínas de ligação ao ADN conhecidas ou a fragmentos destas. A partir da presente divulgação, será aparente para um vulgar perito na arte que o presente método pode facilmente ser aplicado à utilização de novos TBA exibindo pelo menos os atributos necessários observados acima. Assim, por exemplo, em WO 92/20698, foi descrita uma molécula de ligação ao ADN específica da sequência compreendendo um conjugado oligonucleotídico formado pela ligação covalente de um fármaco de ligação ao ADN a um oligonucleótido formando um tríplex. O método dessa divulgação podia ser utilizado para produzir novos TBA para utilizar de acordo com a presente divulgação, desde que os TBA assim formados verificassem os critérios descritos acima. Adicionalmente, os métodos das Patentes U.S. 5 096 815, 5 198 346 e WO 88/06601, podem ser

utilizados para criar novos TBA para utilizar de acordo com o método deste invento. Podem ser utilizados anticorpos específicos ou porções destes (ver por exemplo Blais, 1994).

Quando o TBA é uma proteína, ou um complexo de proteínas, será reconhecido que pode ser utilizado qualquer um de vários métodos de rotina na arte para produzir o TBA. O TBA pode ser isolado a partir do seu ambiente natural ou, se isto for impraticável, produzido através de técnicas padrão de biologia molecular. Assim, utilizando NF- κ B como exemplo, utilizando as porções de ligação ao ADN das subunidades p50 ou p65, este conjunto de ligação pode ser produzido de acordo com métodos recombinantes conhecidos na arte (ver por exemplo Ghosh, *Cell* 62: 1019-1029, 1990, que descreve a clonagem da subunidade de ligação ao ADN p50 de NF- κ B e a homologia dessa proteína com *rel* e *dorsal*).

São conhecidas muitas proteínas de ligação ao ADN e a outros ácidos nucleicos que podem ser utilizadas como ou em TBA de acordo com este invento. Uma vez conhecida a sequência de aminoácidos de qualquer proteína de ligação a ADN, ARN:ADN, ARN ou outro ácido nucleico, uma sequência de ADN apropriada codificando a proteína pode ser preparada através de meios sintéticos ou pode ser utilizada uma cópia de ADNc do ARNm codificando a proteína a partir de uma fonte de tecido apropriada. Para além disso, podem ser obtidas cópias genómicas codificando a proteína e os intrões serem removidos de acordo com métodos conhecidos na arte. Para além disso, os TBA podem ser quimicamente sintetizados.

Uma vez obtida uma sequência de codificação apropriada, pode ser utilizada mutagénese dirigida ao local para alterar a sequência de aminoácidos codificada para produzir proteínas de ligação ao ácido nucleico mutantes exibindo características de ligação mais desejáveis que as da proteína de ligação ao ácido nucleico original. Como exemplo deste processo, a sequência de aminoácidos das porções de ligação ao ADN de NF- κ B podem ser alteradas de forma a produzir uma molécula NF- κ B' que se liga de forma mais forte ao local de ligação de NF- κ B (ver exemplos abaixo HIV-Detect e HIV-Lock).

Para proporcionar uma melhor perspectiva deste aspecto do presente invento, deve-se observar as seguintes considerações. Utilizando NF- κ B como exemplo, pode ser preparado um TBA utilizando a molécula NF- κ B de ocorrência natural. No entanto, como esta molécula está presente em quantidades pequenas e decrescentes nas células e como as subunidades desta proteína de ligação ao ADN foram clonadas, seria mais razoável preparar grandes quantidades do complexo através de meios de ADN recombinante tal como foi já alcançado para esta proteína (ver por exemplo Ghosh, *Cell* 62: 1019-1029, 1990). NF- κ B é um indutor pleiotrópico de genes envolvidos em respostas imunitárias, inflamatórias e reguladoras do crescimento a confrontos patogénicos primários (virais, bacterianos ou de stress) ou a confrontos patogénicos secundários (citoquinas inflamatórias). NF- κ B é uma proteína de ligação ao ADN dimérica compreendendo uma subunidade p50 e uma p65, ambas as quais contactam e ligam-se a sequências de ADN específicas. Num estado inactivado, NF- κ B reside no citoplasma celular, complexado com um inibidor específico, I- κ B, para formar um heterodímero citoplasmático. Após activação, o inibidor é descomplexado e o dímero p50-p65 realocaliza-se através de um sinal de localização nuclear (NLS) específico no núcleo celular onde se pode ligar ao ADN e efectuar o seu papel como activador da transcrição de numerosos genes (ver Grimm e Baeuerle, *Biochem. J.* 290: 297-308, 1993, para uma revisão do estado da arte em relação a NF- κ B).

O dímero p50-p65 liga-se com uma afinidade picomolar a sequências correspondentes à de consenso GGGAMTNYCC (SEQ ID NO:117), com afinidades ligeiramente diferentes dependendo da sequência exacta. Deve notar-se que se observou também a ocorrência de homodímeros de p50 e p65. Estes homodímeros apresentam propriedades bioquímicas diferentes bem como afinidades ligeiramente diferentes das sequências de ligação dentro e semelhantes à de consenso acima. Assim, dependendo das características de ligação desejadas do TBA, pode ser utilizado um heterodímero p50-p65, um homodímero p50-p50 ou um homodímero p65-p65 ou fragmentos dos dímeros acima mencionados.

Um modo pelo qual podem ser produzidos vários novos TBA é mostrado esquematicamente na Figura 9. As unidades de

reconhecimento dos ácidos nucleicos do TBA podem ser montadas e associadas a unidades de reconhecimento de ácidos nucleicos de TBA semelhantes ou diferentes através de um "acompanhante". O acompanhante é uma estrutura na qual são construídos os vários elementos de reconhecimento do TBA e que confere propriedades desejáveis nas unidades de reconhecimento dos ácidos nucleicos. O acompanhante é constituído por qualquer sequência que proporcione sequências de montagem de modo a que as mesmas ou diferentes unidades de reconhecimento dos ácidos nucleicos sejam colocadas em íntima e estável associação umas com as outras. Assim, por exemplo, no caso de um TBA desenhado para ligar fortemente a TBR de NF-kB, um TBA é montado proporcionando sequências *cro* de lambda como sequências de montagem, ligadas às sequências de ligação ao ácido nucleico para p50 ou p65 de NF-kB. As sequências de ligação ao ácido nucleico de p50 ou p65 são ligadas às sequências *cro* no terminal carboxilo ou amino de *cro* e no terminal carboxilo ou amino da unidade de reconhecimento do ácido nucleico da p50 ou p65. Sequências de ligação são opcionalmente proporcionadas para permitir o espaçamento apropriado das unidades de reconhecimento do ácido nucleico para uma óptima ligação à TBR.

As sequências de montagem, exemplificadas acima pelas sequências *cro* e CI (SEQ ID NO:104-108), compreendem quaisquer oligopéptidos estáveis que se liguem naturalmente e fortemente a sequências semelhantes. Assim, no caso de *cro*, é bem conhecido que um dímero de *cro* se liga aos locais operadores do bacteriófago lambda (Anderson et al., *Nature* 290: 754-758, 1981; Harrison e Aggarwal, *Ann. Rev. Biochem.* 59: 933-969). As unidades monoméricas de *cro* associam-se de forma apertada e específica umas às outras. Assim, através da ligação de sequências de unidades de reconhecimento de ADN às sequências *cro*, é alcançada uma associação íntima e apertada.

As sequências ligantes opcionais compreendem qualquer sequência de aminoácidos que não interfira com a montagem do TBA ou a ligação ao ácido nucleico, e que não seja instável de modo a libertar a unidade de reconhecimento do ácido nucleico do TBA completo. É desejável mas não necessário que as sequências ligantes sejam covalentemente ligadas a outros componentes do conjunto de ligação. A associação deve ser

específica de modo a ajudar na montagem e fabrico dos conjuntos de ligação. Exemplos de tais sequências incluem, mas não se limitam a, sequências tão bem conhecidas como as verificadas a ligar vários domínios em proteínas estruturais. Assim, por exemplo, na proteína repressora de lambda, existe uma sequência ligante entre o domínio de ligação ao ADN e o domínio de dimerização que é útil para este fim. São conhecidas muitas outras destas sequências e a sequência exacta destas não é crítica para este invento, desde que seja conduzida experimentação de rotina para assegurar a estabilidade e a não interferência com a ligação ao ácido nucleico alvo. Exemplos de tais sequências são aqui proporcionados como Met Ser e SEQ ID NO:99-102. A inserção de locais de proteólise específicos conhecidos nestes ligantes é também uma parte integral deste invento. A presença de tais locais nas sequências ligantes proporcionará vantagens de fabrico, permitindo que moléculas diferentes sejam montadas na estrutura de acompanhamento.

Para além das unidades de reconhecimento do ácido nucleico, sequências de ligação opcionais e sequências de montagem, os novos TBA deste invento possuem opcionalmente sequências de TNA de assimetria ou PILOTO e uma ou mais unidades OSA. As sequências de assimetria são proporcionadas para favorecer ou prevenir determinadas associações desejáveis ou indesejáveis. Por exemplo, no caso de se desejar que um TBA possua unidades de reconhecimento de ADN de p50 homodiméricas, as sequências de assimetria são proporcionadas para destruir a associação naturalmente mais forte das subunidades p50 e das subunidades p65 de NF-kB, ao mesmo tempo que não destrói as sequências de montagem que mantêm as subunidades p50 juntas. Exemplos de tais sequências são aqui proporcionados como SEQ ID NO:85-92 e SEQ ID NO:105 e 106.

Numa configuração diferente, sequências da subunidade p50 de NF-kB são colocadas em íntima associação com sequências da unidade de reconhecimento de ADN do factor de transcrição SP1. Isto é desejável no caso de um motivo de ligação a NF-kB/SP1 ser significativo, como na LTR de HIV onde se sabe existir um motivo de pelo menos seis locais de reconhecimento da proteína de ligação ao ADN, dois NF-kB, três SP1 e um local TATA. Uma vez que se sabe também que o segundo local NF-kB e o primeiro

SP1 são significativos para a regulação da transcrição do HIV (Perkins et al., *EMBO J.* 12: 3551-3558, 1993), esta configuração particular de TBA é útil não apenas na detecção de HIV, mas como terapêutico ou profilático contra a infecção de HIV (ver abaixo). Do mesmo modo, a região de controlo longa (LCR) do vírus do papiloma humano pode ser utilizada como região de controlo chave para sondar de acordo com este método.

Tendo em vista os diferentes elementos que podem ser associados, em modo de cassete, de acordo com este método de formação do TBA, é produzida uma variedade essencialmente ilimitada de TBA. Na Figura 10, é exemplificada uma série de moléculas diferentes, referida como "HIV-detect I-IV" em que "CHAP" denomina o acompanhante, "nfkb" denomina subunidades de NF-kB, "sp1" denomina a unidade de reconhecimento do ácido nucleico do factor de transcrição SP1 e "TATA" denomina um dímero da unidade de reconhecimento de ADN de uma proteína de ligação ao ADN na sequência TATA (TBP), também conhecida como proteína de ligação a TATA ou TBP. Estas configurações são ainda exemplificadas abaixo e são todas partes integrais do presente invento.

Ainda noutra concretização, a estrutura modular mostrada na Figura 9 é adaptada à detecção e/ou ao tratamento ou profilaxia de um patógeno completamente diferente. Na Figura 11, de um modo semelhante ao das moléculas "HIV-detect I-IV" acima descritas, é produzida uma série de moléculas "HPV-Detect I-IV". Nesta concretização é tirada vantagem das propriedades de ligação ao ADN da proteína E2 do vírus do papiloma humano (HPV). Adicionalmente, é tirado partido dos papéis de SP1 e TBP proporcionando unidades de reconhecimento de ADN específicas adaptadas para se ligarem a estas sequências no genoma de HPV. Na formação de TBA específicos de E2 para utilizar na detecção de infecção de HPV, pode ser desejável utilizar qualquer uma de SEQ ID NO:75-84 ou 93-98 como unidades de reconhecimento do ADN de E2. Um TBA contendo um domínio de ligação ao ADN do dímero de E2 bovino e do dímero de E2 humano pode ser particularmente útil.

As várias sequências descritas acima podem ser quimicamente ligadas utilizando materiais de partida

oligopeptídicos puros ou podem ser ligadas através do fornecimento de ácidos nucleicos recombinantes codificando, através do bem conhecido código genético, os vários sub-elementos. No caso da produção recombinante, a ligação de sequências de codificação *cro* a sequências de unidades de reconhecimento do ácido nucleico para formar TBA é vantajosa porque não só *cro* actua como sequências de montagem no acompanhante como também actua a dirigir a dobragem correcta dos elementos de reconhecimento do ácido nucleico. Sequências exemplificativas para acompanhantes são aqui proporcionadas como SEQ ID NO:104-108. Para além disso, no caso de serem desejadas estruturas de ordem superior compreendendo múltiplos locais de ligação, como num TBA pentamérico NF-kB/NF-kB/SP1/SP1/SP1, o desenho correcto das sequências de assimetria permite que tais estruturas sejam feitas.

Do modo antecedente, são preparados TBA que se ligam aos seus locais de ligação cognatos com elevada afinidade. Por exemplo, espera-se que os componentes de ligação ao ADN de NF-kB dos TBA da Figura 10 se liguem à LTR de HIV com uma afinidade de entre cerca de 10^{-8} e 10^{-12} molar. Sequências úteis como unidades de reconhecimento do ADN são proporcionadas como SEQ ID NO:63-71, 73-84, 93-98 e 104-108 e exemplificadas mais abaixo.

Tendo em vista a descrição antecedente da montagem dirigida das proteínas de ligação ao ácido nucleico utilizando sequências de montagem e assimetria (ou pilotagem), os peritos na arte reconhecerão que é proporcionado por este invento um método geralmente aplicável para montagem de estruturas proteicas. A generalidade deste método é ainda demonstrada pela consideração, como outro exemplo, da utilização de uma interacção anticorpo-epítipo na montagem de estruturas desejadas. Por meio da especificidade, uma estrutura proteica de ligação ao ADN pode ser montada através da ligação de uma subunidade p50 de NF-kB a um antigénio, tal como uma hormona estimuladora dos melanócitos (MSH) circularizada (através de ligações dissulfureto). Esta pró molécula MSH pode então ser ligada através de um anticorpo anti-MSH para proporcionar um novo conjunto de ligação ao ácido nucleico, com o antigénio e o anticorpo a actuar como sequências de montagem.

A estrutura modular proporcionada pela Figura 9 revela que pode ser montada uma grande variedade de TBA utilizando diferentes combinações dos componentes. Assim, são proporcionadas concretizações representativas desta estrutura geral como SEQ ID NO:109-116.

7. Os Conjuntos de Ligação de Reforço (BBA) e sua preparação. Um BBA pode ser qualquer substância que se ligue a uma determinada BBR formada através da hibridação de determinados PNA e BNA, incluindo quando múltiplos BNA (até e incluindo "n" BNA, i.e., BNA_n , em que "n" é teoricamente $0-\infty$, mas na prática está entre 1 e 100) são polimerizados sobre o PNA para amplificação do sinal, desde que o BBA tenha pelo menos os seguinte atributos:

- a) O BBA tem de se ligar às BBR de um modo que seja altamente específico para a BBR de interesse. Isto é, o BBA tem de discriminar entre as BBR presentes no híbrido PNA-BNA e sequências dúplice semelhantes em híbridos BNA-CNA ou outros CNA. Assim, mesmo quando ocorre apenas um único emparelhamento errado de bases ou diferenças conformacionais com ou sem emparelhamentos errados de bases na produção do híbrido PNA- BNA_n ou PNA- BNA_n -HNA, o BBA tem de se ligar ao híbrido com uma avidéz suficientemente baixa para que após lavagem do complexo TBA-TNA-PNA- BNA_n , o BBA seja deslocado das sequências de CNA mas não das sequências de BBR.
- b) O BBA tem de se ligar avidamente à(s) BBR(s). Afinidades de ligação no intervalo de 10^{-5} a cerca de 10^{-9} ou mais são geralmente consideradas suficientes.

Exemplos de BBA incluem, mas não se limitam a *cro* e à proteína repressora do bacteriófago lambda, CI. Adicionalmente, ver Patente U.S. 4 556 643, que sugere outras sequências de ADN e proteínas de ligação específicas tais como repressores, histonas, enzimas modificadoras do ADN e a proteína activadora de genes de catabolitos. Ver também EP 0453301, que sugere uma variedade de proteínas de ligação específicas da sequência nucleotídica (NSSBP) tais como o repressor da tetraciclina, o repressor lac e o repressor do triptofano. Foi reconhecido na arte que cada um destes BBA se liga a determinadas sequências de ácido nucleico conhecidas e as afinidades destas interacções são conhecidas. Naturalmente,

o método do presente invento não se limita à utilização destes BBA conhecidos. A partir da presente divulgação, um perito na arte pode facilmente aplicar a utilização de novos BBA exibindo pelo menos os atributos necessários observados acima no presente método.

Exemplos de novos BBA úteis de acordo com este aspecto do presente invento incluem novas proteínas com base no motivo de uma proteína de ligação a ADN ou ARN ou ADN:ARN conhecida tal como *cro* ou a proteína repressora de λ CI. De preferência, tais modificações são feitas para melhorar o manuseamento destes componentes do presente invento. Assim, pode ser desejável adicionar uma elevada concentração de *cro* a um ensaio. Uma das qualidades negativas de *cro* é que a elevadas concentrações, a ligação de *cro* ao seu alvo de ADN entra em competição com interações *cro-cro*. Assim, por exemplo, pode ser produzida uma *cro* acompanhada ou mutada que não possua este inconveniente. Exemplos de tais acompanhantes alterados são SEQ ID NO:105-106 e 108. Métodos conhecidos na arte, tais como a produção de novas proteínas de ligação utilizando populações variegadas de ácidos nucleicos e selecção de bacteriófagos que se ligam a determinados alvos pré-seleccionados (i.e., a designada tecnologia de apresentação de fagos, ver discussão acima para a produção de novos TBA) podem ser utilizados para produzir tais novos BBA bem como os novos TBA acima mencionados.

Quando o BBA é uma proteína ou um complexo de proteínas, será reconhecido que qualquer um de vários métodos de rotina na arte pode ser utilizado para produzir o BBA. O BBA pode ser isolado a partir do seu ambiente natural ou, se isto for impraticável, produzido através de técnicas padrão de biologia molecular. Assim, por exemplo, a sequência da proteína *cro* é conhecida e qualquer clone molecular do bacteriófago lambda pode ser utilizado para obter os ácidos nucleicos apropriados codificando *cro* para produção recombinante desta. Adicionalmente, os TBA aqui descritos podem ser utilizados como BBA, desde que sejam utilizados TBA diferentes para se ligarem a TBR e BBR.

8. A utilização de BBA e BBR para localizar e amplificar a localização dos complexos PNA-TNA-TBA (ver Figura 8). Numa concretização deste invento, a ligação altamente específica e extremamente forte dos TBA constituídos por componentes de ligação ao ácido nucleico é utilizada para produzir um ensaio

em sanduíche de ácidos nucleicos amplificável. De acordo com um aspecto desta concretização, um suporte sólido é revestido com um primeiro TBA criando um TBA imobilizado. Em solução, um PNA e um TNA são postos em contacto sob condições de hibridação e depois postos em contacto com o TBA imobilizado. Apenas as interações PNA-TNA que formam a TBR específica reconhecida pelo TBA imobilizado são mantidas após a lavagem da superfície sólida que se liga ao complexo TBA-TBR.

A detecção da TBR ligada é alcançada através da ligação de Ácidos Nucleicos de Reforço, BNA, às 1/2 BBR presentes nos PNA sob condições de hibridação. Deste modo, mesmo que apenas se ligue um único complexo TBA-TBR ao TBA imobilizado, pode ser produzido um grande sinal amplificado através da polimerização de múltiplos BNA no TNA imobilizado. Cada BNA que se liga ao TNA forma uma BBR que se pode ligar através de BBA que, tal como os TBA imobilizados na superfície sólida, podem ser escolhidos pela sua ligação muito forte e específica a determinadas estruturas de ácido nucleico. Assim, de acordo com esta concretização, o TBA imobilizado pode conter a porção de ligação ao ADN de NF-kB, que se liga de forma muito específica e forte aos locais de ligação de NF-kB formados após hibridação do TNA e PNA para formar um tal local.

Porque é bem conhecido que existem locais de ligação de NF-kB tanto no genoma humano normal como nas repetições terminais longas do vírus da imunodeficiência humana (HIV), este invento proporciona um método de discriminar entre os locais humanos "normais" e os locais presentes em células devido a infecção de HIV. Portanto, num teste desenhado para determinar a presença ou ausência de ADN de HIV numa amostra de ADN humano, os locais de ligação de NF-kB de HIV podem ser vistos como o TNA e os locais de ligação de NF-kB humanos normais podem ser vistos como CNA. De acordo com o método deste invento, a discriminação entre estes TNA e CNA é alcançada tirando vantagem do facto de na LTR de HIV existirem dois locais de ligação de NF-kB, seguidos de três locais de SP1 (ver, por exemplo, Koken et al., *Virology* 191: 968-972, 1992), enquanto que locais de ligação de NF-kB celulares com as mesmas sequências não se verificam em cadeia.

Em casos onde o TNA contém mais de uma 1/2 TBR e é desejável perseguir as aplicações terapêuticas e profiláticas dos TBA, pode ser desejável utilizar mais de um TBA, cada um com a capacidade de se ligar a um TBR no complexo TNA-PNA. Neste caso, pode ser vantajoso seleccionar, como componentes dos TBA, domínios de ligação ao ADN ou de ligação ao ARN com menor afinidade para a sua TBR que o domínio de ligação ao ADN ou de ligação ao ARN de tipo selvagem. Dado que os TBA que estão envolvidos na ligação a múltiplas TBR se podem montar antes da ligação às suas TBR ou montar após a ligação às suas TBR, os TBA individuais não bloquearão as correspondentes TBR nos outros genomas para além do genoma alvo a menos que as TBR sejam espacialmente capazes de se ligar ao complexo de TBA montado. Uma característica da montagem multimérica de TBA que é especificamente reivindicada aqui como parte deste invento é que se espera que uma tal montagem multimérica tenha uma afinidade muito reduzida para um único local dentro do TNA. No entanto, uma vez que a ligação é dramaticamente aumentada relativamente a qualquer TBA, esperar-se-ia que o complexo de TBA não competisse pela ligação a qualquer TBR única com as proteínas nativas correspondentes *in situ* mas se ligasse fortemente a sequências no híbrido PNA-TNA contendo as TBR para cada um dos componentes de ligação ao ácido nucleico montados no TBA. O complexo de TBA deve ser montado e os ligantes ajustados nos TBA individuais de forma a permitir que as regiões de ligação ao ácido nucleico contidas no complexo de TBA alcancem e se liguem simultaneamente a estes alvos.

Uma vez formados os híbridos TNA-PNA e colocados em contacto com o TBA imobilizado, o ácido nucleico não ligado é lavado da superfície imobilizada e os híbridos imobilizados são detectados. Isto é alcançado através de um de vários modos. Num aspecto deste invento, o PNA é marcado com um OSA, tal como um radionuclido, esferas coloridas ou uma enzima capaz de formar um produto de reacção colorido. Para além disso, para além de possuir uma ou mais 1/2 TBR, o PNA pode também conter pelo menos uma 1/2 BBR. As sequências de 1/2 BBR são escolhidas de forma a serem complementares a sequências de 1/2 BBR únicas nos BNA. Na concretização descrita acima, por exemplo, onde o TBA é NF-kB e a TBR formada após hibridação TNA-PNA é um ou mais locais de ligação de NF-kB, a 1/2 BBR pode proporcionar sequências que podem hibridar (isto é, de

cadeia simples e complementares) dos operadores esquerdo e direito do bacteriófago lambda (ver, por exemplo, Ptashne, *Scientific American* 247: 128-140, 1982, e referências aí citadas para sequências destes operadores). Estas podem ser polimerizadas sobre as 1/2 BBR do PNA de um modo vectorial (ver Figuras 2 e 3) proporcionando até "n" BBR e cada BBR forma um local de ligação *cro*. A *cro* marcada enzimaticamente, radioactivamente ou de outro modo é colocada em contacto com o complexo TBA-TNA-PNA-(BNA)_n. Deste modo é produzido um sinal altamente selectivo e amplificado. O sinal produzido utilizando um PNA possuindo uma única 1/2 TBR indica o sucesso do ensaio em alcançar a ligação TBA-TBR e a polimerização dos BNA para produzir o sinal a partir de locais celulares (i.e., a partir de CNA). A ausência de sinal quando é utilizado um TBA dimerizado indica que no TNA não havia LTR de HIV uma vez que não estavam presentes locais de ligação duplos de NF-kB. Por outro lado, a presença de sinal utilizando o dímero de NF-kB indica infecção de HIV. Como exemplo específico da descrição antecedente desta concretização do presente invento, ver o Exemplo 6 descrevendo um estojo de teste de HIV.

Naturalmente, os peritos na arte reconhecerão que a descrição antecedente está sujeita a várias modificações na escolha dos PNA, TNA, TBA, BNA e BBA. Para além disso, em sistemas que não de HIV, os peritos na arte reconhecerão que o método geral descrito acima pode igualmente ser aplicado. No entanto, estas outras aplicações podem ser mais simples que o método descrito acima uma vez que os TBA utilizados podem não reconhecer quaisquer locais celulares normais e portanto o recorrer à dimerização ou a outros métodos de discriminação entre TNA e CNA pode ser menos crítico. No desenho de sondas e conjuntos de ligação para estes outros sistemas, o perito na arte será guiado pelos seguintes princípios e considerações.

Na concretização acima descrita, o interesse de utilizar as porções de ligação ao ADN da proteína NF-kB como o TBA e os elementos de ligação de reconhecimento de NF-kB como as TBR é que estes elementos formam um "ponto de controlo" importante para a replicação do HIV. Isto é, sabe-se que o HIV necessita de utilizar NF-kB como uma característica crítica no seu ciclo de vida replicativo. Pontos de controlo semelhantes para

outros patogénios são escolhidos e utilizados como base para a detecção de acordo com os métodos aqui descritos.

Da descrição antecedente de características gerais deste invento e do modo da sua operação, um perito na arte reconhecerá que existe uma variedade de modos específicos para a prática deste invento. Como exemplo, o método deste invento é adaptável a um método e dispositivos utilizando estojos de teste cromatográficos descritos nas Patentes U.S. 4 690 691 e 5 310 650 (as patentes '691 e '650). Nessas patentes, foi utilizado um meio poroso para imobilizar um TNA ou uma sonda de captura e foi utilizado um solvente para transportar uma fase móvel contendo um PNA marcado, se o TNA estivesse imobilizado, ou o TNA, se estivesse imobilizada uma sonda de captura, para a "zona de captura". Uma vez o TNA ligado na zona de captura, através da sua imobilização directa ou através de captura, um PNA marcado era sujeito a cromatografia através da zona de captura e era detectado qualquer marcador ligado.

A adaptação do presente invento a um tal sistema proporciona a melhoria da utilização de um Conjunto de Ligação ao Alvo na zona de captura e, deste modo, a captura de apenas sequências de TBR que coincidem perfeitamente ou de outras TBR que representem conformações do ácido nucleico que se ligam especificamente ao TBA dentro dos dúplices TNA-PNA em virtude da discriminação sensível anteriormente descrita pelo TBA entre TNA e CNA.

Uma vez os híbridos TNA-PNA ligados ao TBA imobilizado, o sinal é amplificado através da adição de BNA ou de BNA de cromatografia através da zona de captura. Finalmente, o sinal pode ser ainda amplificado através da adição de BBA ou de BBA marcados cromatograficamente através da zona de captura. Deste modo, a facilidade de realização dos passos de análise descritos nas patentes '691 e '650 é aqui melhorada proporcionando uma capacidade adicional de aumentar a especificidade e, através da amplificação, a sensibilidade do método descrito nessas patentes.

Os peritos na arte reconhecerão também que o método do presente invento é passível de ser realizado em placas de

microtítulo ou de ser automatizado. A utilização de máquinas incorporando o método deste invento recai deste modo naturalmente dentro do âmbito da presente divulgação e das reivindicações anexas a ela. Assim, por exemplo, este invento é adaptável para utilização em instrumentos tais como o analisador de bancada IMx de Abbott Laboratories (Abbott Park, IL). O IMx é actualmente desenhado para correr o imunoensaio de polarização fluorescente (FPZA, ver Kier, *KCLA* 3: 13-15, 1983) e o imunoensaio enzimático de micropartículas (MEZA, ver "Laboratory Medicine", Vol. 20. N° 1, Janeiro de 1989, págs. 47-49). O método MEZA é facilmente transformado num método de detecção de ácidos nucleicos utilizando o presente invento através da utilização de um TBA como molécula de captura revestida numa micropartícula com tamanho submicrométrico (<0,5 µm em média) suspensa em solução. As micropartículas revestidas com TBA são pipetadas para uma célula de reacção. O IMx pipeta então a amostra (PNA-TNA hibridados) para a célula de reacção, formando um complexo com o TBA. Após um período de incubação apropriado, a solução é transferida para uma matriz de fibra de vidro inerte para a qual as partículas têm uma forte afinidade e à qual as micropartículas aderem. Antes ou após a filtração da mistura reaccional através da matriz de fibra de vidro, são adicionados os BNA e BBA ou é utilizado outro meio de amplificação e detecção do sinal que dependa da formação específica de híbridos TNA-PNA. O complexo imobilizado é lavado e o material não ligado flui através da matriz de fibra de vidro.

Os complexos ligados são detectados por meio de BBA marcados com fosfatase alcalina ou BBA marcados de outro modo (radioactivamente, enzimaticamente, por fluorescência). No caso de BBA marcados com fosfatase alcalina, pode ser adicionado o substrato fluorescente 4-metil-umbeliferil-fosfato ou um reagente semelhante. Alternativamente, a enzima pode ser obviada através da marcação directa dos BBA com este ou com um reagente semelhante. Em qualquer caso, o sinal fluorescente ou outro é proporcional à quantidade de híbridos PNA-TNA presentes.

A fluorescência é detectada à superfície da matriz por meio de um fluorómetro de superfície frontal tal como descrito pelo fabricante do IMx. Com ajustes menores que podem ser

feitos através de experimentação de rotina para otimizar um instrumento tal como o IMx para a hibridação de ácidos nucleicos e interacções ácido nucleico-TBA, o presente invento é completamente adaptável a análises automáticas de amostras de TNA.

9. Outras aplicações de diagnóstico deste invento. Embora a descrição antecedente permita a utilização do presente invento de várias maneiras diferentes, muitas utilidades adicionais deste invento são facilmente apreciadas, por exemplo, num sistema de atraso da mobilidade.

Nesta concretização do presente invento, uma melhoria do bem conhecido ensaio de mudança da mobilidade electroforética (EMSA) é conduzida como se segue (ver Figuras 12a e 12b):

Uma amostra de ADN é fragmentada, através de clivagem aleatória ou através de tratamento com endonucleases de restrição específicas. O ADN na amostra é então dividido em duas alíquotas iguais e é adicionado um TNA específico à primeira alíquota mas não à segunda. A primeira e a segunda alíquotas são então sujeitas a electroforese num gel de acrilamida ou agarose e o padrão das bandas de ADN (visualizadas através de ligação a brometo de etídio ou por serem radioactivamente marcadas antes da electroforese) é então comparado para as duas alíquotas. Os fragmentos de ADN possuindo locais de ligação para os quais o TBA é específico são atrasados na sua migração ao longo do meio electroforético. Através da utilização de um TBA apropriado, qualquer número de sequências de ADN ou de outro ácido nucleico pode ser seguido deste modo.

Numa modificação do EMSA descrito acima, o TNA fragmentado é hibridado com um PNA e fraccionado numa primeira dimensão. O ADN fraccionado é então feito reagir com um TBA apropriado e é observada a mudança na mobilidade dos fragmentos de ADN. O aumento do atraso é possível através da adição de BBA tal como descrito acima (ver, por exemplo, Vijg e referências aí citadas para técnicas conhecidas de electroforese de ácidos nucleicos a duas (2) dimensões, às quais o presente método pode ser aplicado).

Os seguintes Exemplos são proporcionados para orientar melhor os peritos na arte nos métodos da prática deste invento. As técnicas de ADN recombinante padrão são divulgadas em Sambrook, Fritsch e Maniatis "Molecular Cloning: A Laboratory Manual", 2ª Ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 1989, e textos mais recentes não são divulgados uma vez que estes estão agora bem dentro dos conhecimentos de uma pessoa competente na matéria.

Exemplo 1 - Preparação dos PNA e Marcação dos PNA

Ácidos nucleicos sonda, PNA, podem ser preparados por meios bem conhecidos na arte. Assim, PNA polinucleotídicos de cadeia simples de sequência definida podem ser preparados através de síntese química de fase sólida de acordo com Merrifield. Os PNA podem ser preparados através de síntese automática utilizando tecnologia comercialmente disponível, tal como resinas e máquinas produzidas ou comercializadas por Applied Biosystems, ABI, ou outros fabricantes. Alternativamente, através de métodos de ADN recombinante conhecidos, determinadas sequências de PNA são sintetizadas *in vivo*, por exemplo, através de clonagem de um PNA dúplice num vector que se possa replicar em *E. coli*, podem ser preparadas grandes quantidades de PNA dúplice. Os multímeros do PNA podem ser clonados no vector de forma a que para cada mole de vector são libertadas várias moles de PNA após digestão do vector com um fragmento de restrição a flanquear a sequência do PNA. Subsequente à síntese ou à produção recombinante, os PNA são purificados através de métodos bem conhecidos na arte tais como através de electroforese em gel ou cromatografia líquida de alta pressão (HPLC). Se o PNA for produzido como um dúplice, antes da utilização num ensaio de hibridação para detecção das sequências de ácido nucleico alvo, as cadeias do PNA são separadas através de aquecimento ou de outros métodos conhecidos na arte.

A sequência específica de bases no PNA é escolhida para reflectir a sequência a detectar num TNA, com a condição de que, de acordo com este invento, o PNA contenha uma sequência de 1/2 BBR, que é uma que após hibridação do PNA e TNA forma uma TBR. Uma vez que existe um número essencialmente ilimitado de tais sequências conhecidas na arte, a escolha da sequência

do PNA é passível de selecção pelo perito na arte para uma dada aplicação. A sequência da LTR de HIV é uma sequência tal que após hibridação das porções de codificação de um PNA da LTR com os TNA codificando a LTR de HIV, formará TBR capazes de se ligar às proteínas de ligação ao ADN NF-kB e SP1.

Para além das sequências que formarão uma TBR após hibridação, o PNA pode também conter uma 1/2 BBR. Esta sequência é uma que, após hibridação com um ácido nucleico de reforço, BNA, forma uma BBR que é capaz de se ligar a uma BBA. A BBA é de preferência uma proteína de ligação ao ADN possuindo elevada afinidade para a sequência da BBR.

Neste exemplo particular, a hibridação entre um PNA possuindo como 1/2 TBR, SEQ ID NO:4, e na extremidade 3' dessa sequência, uma sequência de 1/2 BBR mostrada como SEQ ID NO:35. O PNA codificando estas sequências é utilizado sem marcação ou é marcado com um isótopo radioactivo tal como P³², S³⁵ ou um isótopo semelhante, de acordo com métodos conhecidos na arte. Alternativamente, o PNA é ligado a uma esfera de entre 0,01 a 10 µm, que pode ser colorida para facilidade de detecção visual. Este marcador forma o OSA tal como descrito no fascículo. Esta sonda hibrida com as sequências da LTR de HIV para formar uma TBR que se liga a NF-kB. Adicionalmente, o PNA hibrida com BNA possuindo uma 1/2 BBR complementar para formar um operador esquerdo do bacteriófago lambda que se liga a *cro* ou a proteínas repressoras de lambda.

De um modo semelhante ao descrito acima, são utilizados PNA em que a 1/2 TBR é qualquer uma de SEQ ID NO:5 ou SEQ ID NO:7-34, e uma 1/2 BBR tal como SEQ ID NO:35 ou SEQ ID NO:36 está na extremidade 3' ou 5' da 1/2 TBR.

Exemplo 2 - Preparação e Marcação de BNA

De forma semelhante aos métodos descritos no Exemplo 1 para preparação e marcação de PNA, os BNA são preparados e marcados de acordo com métodos conhecidos na arte. Tal como descrito na Patente U.S. 4 556 643, aqui incorporada por referência (ver particularmente o Exemplo 1), as sequências de ácido nucleico codificando sequências de ligação ao ácido nucleico particulares podem ser produzidas em massa através de

clonagem num vector replicável. Para além disso, de forma semelhante a essa divulgação, as sequências de 1/2 TBR e de 1/2 BBR podem ser co-linearmente produzidas deste modo, no entanto, com a distinção de que, de acordo com o presente invento, a sequência de 1/2 TBR forma ela própria um local de reconhecimento do componente de ligação ao ácido nucleico e a 1/2 BBR, ao mesmo tempo que forma um local de reconhecimento do componente de ligação ao ácido nucleico, proporciona também um meio de amplificar o sinal produzido após a ligação da 1/2 TBR a sequências complementares no TNA proporcionando polimerização de BNA sobre o TNA ligado ao PNA. Para permitir isto, é proporcionada uma sequência tal como SEQ ID NO:35, que codifica o operador esquerdo do bacteriófago lambda, com sequências adicionais de modo a que seja criada uma sequência solta numa ou em ambas as extremidades do BNA após hibridação com o PNA.

Como exemplo específico, a polimerização vectorial dos BNA sobre um TNA é proporcionada através de SEQ ID NO:40-43. Neste exemplo, SEQ ID NO:40 codifica duas 1/2 TBR que hibridarão com duas 1/2 TBR num TNA para formar dois locais de ligação de NF-kB, enquanto que ao mesmo tempo proporciona uma 1/2 BBR do operador esquerdo do bacteriófago lambda, que é adicionalmente terminada na extremidade 3' com o local de reconhecimento para a enzima de restrição *PstI*. A adição do BNA, SEQ ID NO:41, com a 1/2 BBR complementar à 1/2 BBR do PNA, SEQ ID NO:40, completa a BBR enquanto que ao mesmo tempo completa o local de reconhecimento *PstI*, deixando uma extremidade solta de quatro bases para hibridação com BNA adicionais. Concordantemente, é adicionada SEQ ID NO:42 que possui uma sequência de quatro pares de bases na extremidade 3' que é complementar à extremidade solta de quatro bases restante da hibridação de SEQ ID NO:40 e 41. Adicionalmente, é proporcionada SEQ ID NO:42 com uma sequência de cinco bases na sua extremidade 5' que faz parte de um local de reconhecimento *BamHI*. O polímero em crescimento de BNA é ainda prolongado através da adição do BNA de SEQ ID NO:43 que é complementar a SEQ ID NO:42, completando a BBR ao mesmo tempo que completa o local de reconhecimento *BamHI* e deixa uma extremidade solta de quatro bases que pode ainda hibridar com BNA possuindo sequências complementares. Deste modo, os BNA podem ser

hibridados extensamente de modo a amplificar grandemente o sinal de um único acontecimento de hibridação PNA-TNA.

Tal como com os PNA descritos no Exemplo 1, os BNA podem ser utilizados numa forma não marcada ou podem ser marcados de acordo com métodos conhecidos na arte e descritos no Exemplo 1. Será também apreciado que, em vez de se produzir o polímero de BNA através da adição sequencial de BNA ao complexo PNA-TNA, o polímero de BNA pode ser pré-formado e adicionado directamente ao complexo PNA-TNA. Um método simples para pré-formação de um tal polímero de BNA inclui a produção recombinante de um vector no qual são proporcionados múltiplos do BNA com um único local de restrição em cada extremidade do polímero. Este polímero de BNA contendo múltiplas BBR é cortado do vector e hibrida com uma 1/2 BBR de cadeia simples que permanece no PNA após a hibridação do PNA com o TNA. Isto é alcançado proporcionando uma sequência de cadeia simples no PNA complementar a uma extremidade solta produzida no polímero de BNA quando é excisado do vector de produção.

Exemplo 3 - Produção de HNA e Sua Utilização para Protecção de Polímeros de BNA

Os HNA deste invento são produzidos de acordo com métodos conhecidos na arte para produção de polinucleótidos tal como descrito nos Exemplos 1 e 2 para PNA e BNA. No entanto, na produção dos HNA a sequência do HNA é especificamente desenhada para que uma porção substancial do HNA forme um palindroma auto-complementar para formar um "gancho de cabelo", enquanto que ao mesmo tempo, deixa numa forma de cadeia simples suficientes bases para ser capaz de hibridar com sequências de cadeia simples na cadeia crescente de BNA descrita no Exemplo 2.

Neste Exemplo é proporcionado um HNA de SEQ ID NO:44 para rematar a extensão de BNA no PNA no Exemplo 2 após a adição do BNA, SEQ ID NO:43. Isto é alcançado porque SEQ ID NO:44, para além de possuir uma sequência palindrômica que forma um "gancho de cabelo" estável possui também uma sequência na extremidade 5' do HNA que completa a sequência de BamHI formada através da hibridação de SEQ ID NO:42 e SEQ ID NO:43.

Naturalmente, a terminação do polímero após a adição de apenas 3 BNA é para fins de simplicidade na demonstração do presente invento. Tal como descrito acima, esta polimerização pode ser continuada essencialmente de forma indefinida para amplificar o sinal do acontecimento de hibridação PNA-TNA.

Exemplo 4 - Preparação de TBA e BBA, Marcação e Imobilização Destes

As TBA e as BBA que podem ser utilizadas de acordo com o presente invento incluem qualquer substância que se possa ligar de modo específico às TBR e BBR formadas através de hibridação dos PNA, TNA e BNA. A utilização de proteínas de ligação ao ADN forma um exemplo de tais substâncias.

Para este exemplo, a TBA é o dímero da porção de ligação ao ADN de p50, e a BBA é a proteína *cro* de lambda. Estas proteínas podem ser produzidas de acordo com métodos conhecidos na arte. Os genes para ambas estas proteínas foram clonados. Assim, estas proteínas são produzidas de forma recombinante e purificadas de acordo com métodos conhecidos na arte. Para além disso, estas proteínas são marcadas, com um radioisótopo, tal como iodo radioactivo, ou com uma enzima, tal como beta-galactosidase ou peroxidase de rábano, ou com um corante fluorescente tal como fluoresceína ou rodamina, de acordo com métodos bem conhecidos na arte. Adicionalmente, um ou ambos os TBA e BBA podem ser imobilizados numa superfície sólida tal como a superfície de uma placa de microtítulo ou a superfície de uma esfera, tal como uma esfera colorida de diâmetro entre 0,01 e 10 μm . Os marcadores nos TBA e BBA podem ser os mesmos ou diferentes.

Neste exemplo, o TBA contendo o domínio de ligação ao ADN de p50 dimérico é marcado com rodamina, enquanto que o BBA, *cro*, é marcado com fluoresceína. Assim, após a hibridação dos PNA, TNA, BNA e HNA tal como descrito nesta divulgação de patente e os antecedentes e seguintes exemplos, os híbridos de ácidos nucleicos, se formados, são colocados em contacto com um excesso de TBA e *cro* marcados. A fluorescência destes marcadores é medida de acordo com métodos conhecidos e a detecção de ambos os sinais é indicadora da presença de sequências de 1/2 TBR no TNA. O sinal diferencial produzido

pela fluorescência da NF-kB e da *cro* é uma medida do grau ao qual a polimerização dos BNA sobre o híbrido PNA-TBA resultou em amplificação do sinal. Está contemplada uma amplificação de uma a mais de mil vezes de acordo com o método deste invento.

Exemplo 5 - Hibridação de dois PNA com um TNA e Discriminação Entre um TNA e um CNA

Os PNA, PNA1, SEQ ID NO:40 e PNA2, SEQ ID NO:45, são utilizados num excesso molar de cerca de dez vezes em relação à concentração de TNA numa amostra de teste. Para este exemplo, uma LTR de HIV dúplice isolada, em que uma cadeia da qual possui a sequência SEQ ID NO:37, mostrada na Figura 7, e a outra cadeia da qual é complementar à sequência mostrada na Figura 7, é utilizada como o TNA. É também utilizado neste exemplo um CNA dúplice isolado, do qual uma cadeia possui a mesma sequência que SEQ ID NO:37 excepto que, no primeiro local de ligação de NF-kB mostrado na Figura 7, no centro da do local de ligação, posição 1 na Figura 7, em vez de um "T" existe um "A", cuja cadeia complementar portanto emparelha erradamente com o PNA de SEQ ID NO:40 nessa localização.

SEQ ID NO:40 e SEQ ID NO:50 são ambas adicionadas para separar reacções, a primeira contendo o TNA acima descrito e a segunda contendo o CNA acima descrito. As amostras são solubilizadas num tampão de hibridação apropriado, tal como Tris 10 mM (pH 7,5), EDTA 1 mM. As amostras são aquecidas a cerca de 90°C durante cerca de cinco minutos para separar as cadeias dos TNA dúplice e dos CNA nas amostras e depois as amostras são deixadas a arrefecer para permitir que as cadeias dos PNA, TNA e CNA hibridem.

Uma vez completada a hibridação, o que pode ser determinado de acordo com métodos conhecidos tais como através do cálculo de $t_{1/2}$ com base nas composições de bases e na temperatura de hibridação de acordo com métodos conhecidos, o PNA de SEQ ID NO:40 é polimerizado através da adição de BNA tal como no Exemplo 2 e a sonda PNA2 de SEQ ID NO:45 é polimerizada com BNA a começar com a extremidade solta de reconhecimento de *Sph1*. Após a adição dos BNA e de um breve período de hibridação, são adicionadas amostras separadas a esferas revestidas com NF-kB covalentemente imobilizada e a

NF-kB é deixada a ligar-se a quaisquer TBR formadas nas amostras de TNA e CNA. Após cerca de 15 minutos de ligação, as amostras foram lavadas duas vezes com cerca de três volumes de um tampão de lavagem apropriado, tal como Tris 10 mM, pH 7,5, NaCl 100 mM ou outro tampão predeterminado para não interferir com NF-kB ou com a actividade de ligação da proteína repressora CI do bacteriófago lambda. Após cada lavagem, as esferas foram deixadas a assentar sob gravidade ou através de uma breve centrifugação. Isto remove quaisquer ácidos nucleicos que não possuam um local de ligação a NF-kB perfeito formado através de hibridação das sequências de PNA1 e TNA.

Após a lavagem final, a proteína repressora CI do bacteriófago lambda marcada com um isótopo radioactivo, tal como com iodo radioactivo, ou marcada com uma enzima, tal como peroxidase de rábano, com esferas coloridas ou com um marcador fluorescente, é adicionada a cada amostra. As amostras são então lavadas várias vezes (cerca de 3) com vários volumes (cerca de 2) de um tampão de lavagem apropriado tal como Tris 10 mM, pH 7,5, NaCl 100 mM ou outro tampão predeterminado para não interferir com NF-kB ou com a actividade de ligação da proteína repressora CI do bacteriófago lambda. Após cada lavagem, as esferas foram deixadas assentar sob gravidade ou através de uma breve centrifugação. Após a última vez que se deixa assentar ou a última centrifugação, o marcador ligado é quantificado através da detecção da radioactividade ligada, da cor libertada num ensaio enzimático, da cor das esferas ou da detecção de fluorescência. Alternativamente, pode ser adicionado um anticorpo anti-CI e efectuado um imunoensaio enzimático em sanduíche ou um radioimunoensaio padrão para detectar o repressor ligado. Adicionalmente, como controlo negativo (fundo) todas as manipulações antecedentes são realizadas em cadeia com uma amostra na qual as esferas são utilizadas sem possuírem NF-kB imobilizada.

Como resultado do ensaio antecedente, as amostras contendo o controlo e o CNA possuem sinais semelhantemente baixos enquanto que a amostra contendo TNA possui um sinal bem acima do fundo.

Exemplo 6 - Estojo de Teste para a Detecção de HIVA. Conteúdo do estojo:

1. Placa de microtítulo.
2. Solução 1 mg/ml de NF-kB produzida de forma recombinante em solução salina tamponada com tris.
3. Tubo contendo PNA de HIV de cadeia simples (uma mistura de oligonucleótidos pré-misturados codificando dois 1/2 locais de ligação de NF-kB, i.e. uma mistura de SEQ ID NO:7 e 8).
4. Tubo contendo PNA genômico humano de cadeia simples, SEQ ID NO:1.
5. Tubo de nuclease (*Pst*I).
6. Tubo de protease.
7. Tubo contendo BNA pré-polimerizados, 100 unidades de repetição de O_R do bacteriófago lambda, rematados com um HNA mas com 1/2 BBR disponíveis para ligação a híbridos PNA-TNA.
8. Tubo de cro conjugada com peroxidase de rábano (hrp).
9. Tubo de substrato colorido de hrp.
10. Solução salina tamponada com tris, 100 ml.
11. Lanceta.
12. Tubos de reacção A, B, C, contendo cada um 250 µl de água destilada.
13. Conta-gotas.

B. Método de ensaio:

- a) A placa de microtítulo (item 1) é revestida com a solução de NF-kB produzida de forma recombinante (item 2) a uma concentração de 1 mg/ml em solução salina tamponada com tris de um dia para o outro a 4°C com agitação.
- b) Obtêm-se três gotas de sangue do sujeito a testar picando um dedo com a lanceta (reagente 11) e é dispensada uma gota para cada tubo de reacção A, B e C (reagente 12).
- c) Para cada tubo é dispensada uma gota de solução de protease (reagente 6) com o conta-gotas (item 12) e o tubo é agitado e deixado a repousar durante 5 minutos.

- d) Uma gota de nuclease (item 5) é adicionada a cada um dos tubos A-C utilizando o conta-gotas e os tubos são agitados e deixados a repousar durante 10 minutos.
- e) Uma gota do item 3 é adicionada ao tubo A (amostra de teste); uma gota de item 4 é adicionada ao tubo B (controlo positivo); e uma gota de solução salina (item 12) é adicionada ao tubo C como controlo negativo. Os tubos são aquecidos a 50°C em água quente e deixados arrefecer à temperatura ambiente durante uma hora.
- f) Embora se deixe ocorrer a hibridação no passo (d), a proteína em excesso é drenada da superfície e da placa de microtítulo, do passo (a) e a placa é lavada com solução salina tamponada com tris (tubo 10).
- g) Os conteúdos dos tubos A-C do passo (e) são transferidos para três poços da placa de microtítulo e deixados a repousar durante 1 hora com agitação.
- h) Os poços de microtítulo contendo os conteúdos dos tubos A-C são lavados com solução salina tamponada com tris e esvaziados.
- i) É adicionada uma gota de item 7 a cada poço e deixa-se hibridar com quaisquer locais 1/2 BBR ligados à placa, durante uma hora, seguido de três lavagens com solução salina tamponada com tris.
- j) É adicionada uma gota de item 8 a cada poço e deixa-se a cro ligar a quaisquer BNA ligados durante 10 minutos, seguido de cinco lavagens de 1 ml com solução salina tamponada com tris.
- k) É adicionada uma gota de substrato de hrp a cada poço e deixa-se revelar a cor.

C. Resultados:

Se os poços A e B mostrarem ambos revelação de cor e o poço C não, o teste é válido e o sujeito foi infectado com HIV. Se apenas o poço A apresentar revelação de cor ou se o poço C apresentar revelação de cor, o teste foi efectuado incorrectamente e é inválido. Se os poços A e C não apresentarem revelação de cor mas o B sim, o teste é válido e o indivíduo não foi infectado com HIV.

Exemplo 7 - Produção de Vários Novos TBA

Novos TBA para utilizar de acordo com o presente invento são preparados como se segue:

a) NFkB/NF-kB (HIV-Detect I). Um ácido nucleico codificando qualquer uma de SEQ ID NO:63-71 ou uma proteína de ligação ao ADN semelhante a NF-kB, é fundido, enquadrado, com uma sequência nucleotídica codificando uma sequência de montagem, tal como *cro*, de modo a que a sequência de reconhecimento do ADN de NF-kB seja codificada no terminal amino ou carboxilo da sequência de *cro*. Opcionalmente, é proporcionada uma sequência ligante entre a sequência de NF-kB e a sequência de *cro*. No outro terminal de *cro* é opcionalmente proporcionada uma sequência sinal de localização nuclear, tal como SEQ ID NO:72. Mais, são opcionalmente proporcionadas sequências de assimetria no terminal de *cro* não utilizado pela sequência de reconhecimento de NF-kB. Exemplos de TBA completos são mostrados abaixo.

b) NF-kB/SP1 (HIV-Detect II). De um modo semelhante ao descrito em (a) acima, é preparada uma sequência de codificação recombinante codificando um domínio de reconhecimento de NF-kB. Numa construção separada, em vez de SEQ ID NO:63-72, é incluída a sequência de codificação para a porção de reconhecimento do ADN de SP1. Uma tal sequência deve codificar toda ou uma parte funcional de SEQ ID NO:73, que é a porção do factor de transcrição SP1 que exhibe ligação ao ADN (ver Kadonaga et al., Cell 51: 1079-1090, 1987). O vector codificando NF-kB e o vector codificando SP1 são então co-transfectados num sistema de expressão apropriado tal como é bem conhecido na arte. Uma unidade de reconhecimento de NF-kB monomérica é adicionada para completar o dímero de reconhecimento de NF-kB após a montagem das unidades de reconhecimento de SP1 e NF-kB pelo acompanhante. As sequências de assimetria impedem a formação de dímeros de NF-kB ou SP1 e dirigem, por sua vez, a formação de heterodímeros NFkB-SP1 (i.e., HIV-Detect II), que são então isolados do sistema de expressão (células de mamífero ou bacterianas) através de métodos conhecidos.

c) TBA de SP1/SP1 (HIV-Detect III). Tal como descrito em (b) acima, é preparada uma construção de TBA codificando SP1. No entanto, apenas esta construção é transfectada para o sistema de expressão e estão incluídas as sequências de assimetria que permitem a formação de dímeros SP1-SP1.

d) SP1-TATA (HIV-Detect IV). Tal como descrito em (b) acima, é produzido um recombinante de TBA codificando SP1. Adicionalmente, é preparado um recombinante codificando um TBA possuindo a sequência de ligação SEQ ID NO:74 ou uma sequência semelhante codificando uma unidade de reconhecimento TATA com sequências de assimetria complementares às incluídas na construção codificando TBA de SP1. Estas construções são co-transfectadas e os heterodímeros isolados através de métodos padrão, incluindo purificação de afinidade numa coluna de ADN possuindo as regiões de ligação ao alvo SP1-TATA apropriadas.

e) SP1-E2 (HPV-Detect I). Uma construção codificando SP1 é preparada tal como em (b) acima. Uma construção codificando um TBA de E2 é preparada através da utilização de uma sequência codificando qualquer uma de SEQ ID NO:75-84 e 94-98 que são unidades de reconhecimento do ADN de E2 (ver Hedge *et al.*, *Nature* 359: 505-512, 1992) ou unidades de reconhecimento semelhantes, é preparada e co-transformada ou co-transfectada com a construção codificando o TBA de SP1. A unidade de reconhecimento de E2 monomérica é adicionada ao dímero de reconhecimento de E2 completo após a montagem da unidade de reconhecimento de E2-SP1 pelo acompanhante. O heterodímero HPV-Detect I é isolado de acordo com métodos conhecidos.

f) E2-E2 (HPV-Detect II). Tal como descrito acima em (e), é preparada uma construção codificando TBA de E2, excepto que são incluídas sequências de assimetria que permitem a formação de dímeros de E2. Os dímeros expressos são então isolados através de métodos conhecidos incluindo afinidade para um local de ligação de E2 dimérico numa coluna de afinidade de ADN.

g) E2-TATA (HPV-Detect III). Tal como descrito acima em (e) e (d), são preparados TBA de ligação de E2 e TATA (respectivamente), excepto que são incluídas sequências de assimetria que aumentam a formação de heterodímeros em vez de

homodímeros. Estas construções são então co-expressas e os heterodímeros são isolados.

h) TATA-TATA (HPV-Detect IV). Tal como descrito acima em (a) e (d), é preparada uma construção codificando TBA de ligação a TATA utilizando sequências de assimetria que favorecem a formação deste homodímero e o homodímero é isolado.

i) Outros TBA. Tal como descrito acima para os TBA de HIV e HPV, TBA para qualquer dado patogénio ou estado de doença podem ser produzidos através da identificação de proteínas de ligação ao ADN específicas e da formação de uma construção de expressão utilizando sequências ligantes, de montagem e de assimetria apropriadas.

Exemplo 8

De um modo semelhante ao do ensaio descrito no Exemplo 5, é produzido um ensaio mais rigoroso através da utilização da proteína de ligação NF-kB-SP1 dúplex preparada de acordo com o Exemplo 6. Assim, as sondas mostradas na Figura 7 e utilizadas no Exemplo 5 podem ser alongadas para reduzir a distância inter-sondas e deste modo reduzir a flexibilidade do ADN no TNA.

Exemplo 9 - Produção de TBA de "Ordem Superior"

Através da utilização apropriada de sequências de assimetria, são produzidos TBA que são dímeros, trímeros, tetrâmeros, pentâmeros ou hexâmeros de determinadas unidades de reconhecimento de ADN. Deste modo, é produzido um TBA hexamérico através da produção de um primeiro TBA dimérico de p50 de NF-kB utilizando sequências de assimetria que permitem a formação do dímero. Adicionalmente, as sequências de assimetria permitem a tetramerização do dímero de p50 com um dímero SP1-SP1. Finalmente, sequências de assimetria adicionais dirigem a hexamerização com um dímero exibindo sequências de localização nuclear. Isto é alcançado através da incorporação, por exemplo, de sequências de assimetria de insulina, que na natureza forma hexâmeros. Esta formação de hexâmeros é dirigida pelas sequências SEQ ID NO:85 (A) e

86 (B), 87 (A) e 88 (B), 89 (A) e 90 (B), e 91 (A) e 92 (B) (ver Figuras 13 e 14).

Devido à afinidade extremamente elevada para a LTR de HIV que pode ser gerada utilizando um TBA multimérico, os compostos possuindo esta estrutura e que podem ser utilizados para este fim são aqui referidos como "HIV-Lock".

Um HIV-Lock óptimo é definido por "*footprinting*" (de acordo com métodos bem conhecidos na arte) de TBA ligados a TBR na LTR de HIV para confirmar que a afinidade de ligação de cada proteína de ligação ao ADN que contribui para a formação do complexo multimérico de TBA é diminuída relativamente à afinidade de qualquer sequência alvo natural (i.e. CNA) da qual a unidade de reconhecimento de ligação ao ADN do TBA é derivada. Qualquer perda concomitante na afinidade de ligação para as TBR de HIV é mais do que compensada após a formação do multímero tal como descrito abaixo.

Pode haver competição entre a ligação de cada TBA componente pela sua TBR e montagem, através das sequências de assimetria para formar o multímero. Isto é evitado ajustando os ligantes entre as sequências de acompanhamento e de assimetria em cada componente do TBA de modo a que estes acontecimentos de competição estejam desacoplados. A redução resultante na dimensionalidade da difusão (eficaz aumento da concentração) para os componentes de assimetria e de montagem do TBA resulta na formação eficiente do complexo multimérico.

Com base no "*footprinting*", o comprimento e a composição dos ligantes são ajustados para alcançar uma discriminação óptima entre as sequências de HIV alvo e as sequências naturais. Deste modo, embora cada TBA componente terá baixa afinidade para as sequências de CNA e TBR, o complexo multimérico terá uma afinidade extremamente elevada para a TBR agora expandida reconhecida pelo complexo multimérico (o quadrado da afinidade de cada TBR reconhecida por cada TBA componente do TBA multimérico), ao mesmo tempo que possui ainda uma baixa afinidade para os CNA. Do mesmo modo são preparados outros complexos multiméricos de TBA, para além do HIV-Lock.

Os TBA que podem ser formados deste modo incluem as seguintes seqüências, que são montadas através da ligação de subunidades proteicas ou seqüências de ácidos nucleicos codificando estas subunidades, como se segue:

<u>Conjunto</u>	<u>Ligar as Sequências dos Grupos</u>
A	I + II + III
B	IV + V + III
C	IV + III

em que os grupos I-V consistem em seqüências seleccionadas de entre:

<u>Grupo</u>	<u>Seleccionado de entre as Sequências</u>
I	Qualquer uma de SEQ ID NO:85-92
II	Met Ser, ligada a qualquer uma de SEQ ID NO:104-106, cada uma das quais é ligada a SEQ ID NO:99
III	SEQ ID NO:100 ligada a qualquer uma de SEQ ID NO:75-84 ou 94-98; SEQ ID NO:101 ligada a SEQ ID NO:74 ou SEQ ID NO:93; ou SEQ ID NO:102 ligada a SEQ ID NO:74 ou SEQ ID NO:93; ou qualquer uma de SEQ ID NO:72, 103, 73 ou 63-71.
IV	Qualquer uma de SEQ ID NO:104-106.
V	SEQ ID NO:99.

Exemplos específicos de tais TBA são SEQ ID NO:109-106, montadas como se segue:

<u>Conjunto</u>	<u>SEQ ID NO.</u>	<u>Ligar SEQ IDS.</u>
A	109	85 + Met Ser + 104 + 99 + 100 + 94
A	110	85 + Met Ser + 104 + 99 + 72
A	111	86 + Met Ser + 105 + 99 + 102 + 74
A	112	86 + Met Ser + 106 + 99 + 73
A	113	89 + Met Ser + 106 + 99 + 63
C	114	106 + 64
C	115	105 + 64
B	116	106 + 99 + 73

Deste modo, escolhendo entre seqüências de assimetria, seqüências de montagem e unidades de reconhecimento do ADN

apropriadas, podem ser formados muitos TBA diferentes. Para além disso, conjuntos destes, tais como SEQ ID NO:114 e 115, associar-se-ão uns aos outros mas não se formarão dímeros de SEQ ID NO:114 ou 115 devido à repulsão de cargas nas sequências de montagem mutadas (SEQ ID NO:104 é *cro*; SEQ ID NO:105 é uma nova *cro* mutada, carregada negativamente, e SEQ ID NO:106 é uma nova *cro* mutada, carregada positivamente).

Naturalmente, dada a sequência de aminoácidos destes TBA, um vulgar perito na arte pode produzir clones de ácidos nucleicos recombinantes codificando-as e tais clones recombinantes formam naturalmente uma parte integral deste invento.

Exemplo 10 - Teste de HIV Utilizando "HIV-Lock"

Praticamente como no método utilizado no Exemplo 6, o "HIV-Lock" produzido de acordo com o Exemplo 9 é utilizado como TBA, reagente 2, com resultados semelhantes.

Exemplo 11 - Teste de HIV Utilizando "HIV-Lock" Quando se Testa Sangue para Doação

Quando a quantidade de sangue a testar não é limitante, tal como quando se testam amostras de sangue para doação quanto a contaminação por HIV, são realizados testes semelhantes aos do Exemplo 6, mas para cada um dos tubos A-C, são sedimentados cerca de 5 ml de sangue numa centrífuga de bancada. Os outros reagentes são proporcionais conforme o necessário para manusear a maior quantidade de TNA presente na amostra.

Exemplo 12 - "HIV-Lock" como Agente Terapêutico Anti-HIV

O "HIV-Lock" produzido de acordo com o Exemplo 9 é formulado como uma solução 1 mg/ml em lipossomas e injectado intravenosamente num sujeito que foi testado e confirmado estar infectado com HIV. Uma dose de cerca de 0,1 mg a 100 mg de "HIV-Lock"/quilograma de massa corporal é infundida durante um período de vinte e quatro horas e é monitorizada a concentração de p24 de HIV no soro do paciente. O tratamento é

repetido tão frequentemente conforme o necessário tal como quando ocorrem elevações na p24 sérica.

Exemplo 13 - Utilização de uma Construção de TBA de HIV como Terapêutico

Um vector retroviral recombinante ou semelhante é utilizado para distribuir uma construção codificando um TBA de ligação à LTR de HIV a um paciente infectado. O vector codifica um acompanhante, tal como *cro*, e sequências de ADN para porções de ligação de p50. O mesmo vector codifica também um acompanhante no qual se dobra um TBA de SP1. São proporcionadas sequências de assimetria de modo que após a co-expressão do TBA de p50 e do TBA de SP1 numa única célula infectada com HIV *in vivo*, ocorre uma associação imediata entre estes TBA, ao mesmo tempo que se impede qualquer associação entre a porção de ligação ao ADN de p50 e monómeros endógenos de p50 ou p65. São também proporcionadas sequências NLS nos TBA de modo a que, após a formação dos dímeros, o TBA se re-localize imediatamente no núcleo da célula e se ligue especificamente às sequências de HIV integradas, impedindo assim qualquer transcrição a partir desse *locus*.

Para este fim, é desejável seleccionar sequências codificando domínios de ligação ao ADN de modo a que os monómeros expressos sejam montados num TBA que não se ligue a sequências humanas naturais. Assim, é apenas após a ligação dos componentes do TBA às suas sequências alvo que a associação entre todos os componentes do TBA ocorre para formar um complexo que se liga fortemente e especificamente à LTR de HIV.

Exemplo 14 - Estojo de Teste de Diagnóstico para Vírus do Papiloma Humano

Este diagnóstico para o vírus do papiloma humano tira vantagem do diferencial conhecido entre HPV benigno e carcinogénico para proporcionar um teste que indica a susceptibilidade à malignidade num paciente. Os vírus do papiloma são um grupo de pequenos vírus de ADN associados a tumores de células epiteliais escamosas benignos nos vertebrados superiores. Foram verificados pelo menos 27 tipos

humanos distintos de vírus do papiloma (HPV); muitos destes foram associados a lesões clínicas específicas. Quatro destes, HPV-6, HPV-11, HPV-16, HPV-18 e HPV-33 foram associados a lesões do tracto genital humano. Em geral, os ADN de HPV-6 e HPV-11 foram verificados estar associados a lesões benignas do tracto genital. HPV-16, HPV-18 e HPV-33 foram também verificados estar associados a lesões pré-malignas e malignas e são transcritos na maioria das linhas celulares estabelecidas a partir de carcinomas cervicais. É provável que HPV-16, HPV-18 e HPV-33 sejam apenas dois membros de um grande conjunto de ADN de HPV associados a carcinomas cervicais humanos malignos.

Modelos animais mostraram que lesões benignas do vírus do papiloma podem progredir para lesões malignas na presença de um co-carcinogénio. O ADN de HPV foi verificado em metástases de carcinomas cervicais. Em lesões cervicais malignas, o ADN de HPV está habitualmente integrado no genoma humano, mas pode também estar presente ADN de HPV extracromossómico. A integração de HPV para formar o pró-vírus resulta habitualmente na destruição do enquadramento de leitura aberto (ORF) de E2. Apesar da destruição do ORF de E2, o exame de linhas celulares de vários carcinomas cervicais mostrou HPV-16 e HPV-18 transcricionalmente activos e integrados. Quando os genomas de HPV-16 que estão presentes nas linhas celulares de carcinoma cervical humano SiHa e CaSki são examinados, há diferenças verificadas na integração de HPV-16. Na linha SiHa, a integração única do genoma de HPV-16 ocorreu nas bases 3132 e 3384, destruindo os ORF de E1 e E2 com uma deleção de 0,3 kb. Uma deleção adicional de 50 pares de bases do ADN de HPV-16 resultou na fusão dos ORF de E2 e E4. A porção 5' do ADN de HPV-16, que consiste no ORF de E2 destruído, é ligada a sequências flanqueadoras à direita humanas contínuas. Adicionalmente, é detectada uma única guanina adicional no nucleótido 1138 no meio do ORF de E1. Esta adição de um par de bases resulta na fusão dos ORF de E1a e E1b num único ORF de E1.

O genoma completo de HPV-16 está disponível em GenBank com o número de acesso K02718; o genoma completo de HPV-33 está disponível em GenBank como número de acesso M12732; o

genoma completo de HPV-18 está disponível em GenBank como número de acesso X05015.

Como pesquisa preliminar, o facto de uma infecção de HPV ser estabelecida para uma dada amostra de biopsia cervical através de uma simples análise de tipo "sim/não" utilizando, por exemplo, qualquer um ou todos os PNA SEQ ID NO:46-53 e um TBA de E2 tal como descrito acima (i.e., fragmentar o ADN, ligar ao PNA, imobilizar com o TBA e detectar o sinal com BNA e BBA).

Uma vez a biopsia se verifique positiva para HPV, obtém-se informação adicional quanto à potencial malignidade do HPV através da análise do estado de integração do vírus no genoma humano.

1. Fragmentar o ADN na amostra da biopsia cervical e hibridar com uma sonda de bloqueio possuindo a sequência SEQ ID NO:60. Esta sonda ligar-se-á a todos os fragmentos no ADN que não foram removidos do fragmento de 0,3 kb.
2. Expor o ADN da amostra da biopsia a um PNA possuindo a sequência SEQ ID NO:61. Esta sonda ligar-se-á apenas a fragmentos que foram suprimidos do fragmento de 0,3 kb (a sonda de bloqueio impedirá a formação de uma volta dos grandes segmentos de deleção se presentes).
3. Um PNA possuindo SEQ ID NO:62 é hibridado com SEQ ID NO:41 para formar uma BBR que se ligará a *cro* ou ao repressor CI de λ como BBA, deixando uma porção de cadeia simples capaz de hibridar com o local TATA em SEQ ID NO:61. Este foi adicionado para formar uma TBR na extremidade 5' da grande deleção.
4. A TBR é imobilizada por um TBA possuindo uma unidade de reconhecimento do ADN da proteína de ligação a TATA.
5. Os fragmentos ligados são detectados através da adição de BNA e BBA tal como descrito acima.

A detecção do sinal neste ensaio indica que o fragmento grande está suprimido no HPV presente no TNA. Uma vez que esta

deleção se correlaciona com malignidade, este ensaio proporciona uma perspectiva sobre o potencial de malignidade da infecção de HPV. Esta conclusão pode ser confirmada através da realização de um ensaio análogo com base na deleção do fragmento de 52 pares de bases que também se correlaciona com malignidade induzida por HPV.

A unidade de reconhecimento de TBP utilizada no TBA para este ensaio pode ser escolhida, por exemplo, a partir de uma sequência tal como SEQ ID NO:70 ou SEQ ID NO:93.

Exemplo 15 - Produção de HIV-Lock™ Recombinante

Fase Um - Preparação de ADN para Produzir o HIV-Lock™. É efectuada mutagênese *in vitro* das regiões de codificação dos componentes de ocorrência natural e clonados do HIV-Lock™ que necessitam ser modificados com um estojo MutaGene Phagemid. O protocolo modificado inclui a utilização de um plasmídeo Blue-script contendo cada um dos componentes de ligação de HIV-Lock™. Estes são transformados para células competentes e os fagemídeos contendo uracilo são criados. O ADN de cadeia simples é extraído e utilizado como molde para a cadeia mutagénica. Oligonucleótidos contendo as mutações desejadas, incluindo a incorporação de um novo local de restrição, são sintetizados e tratados com polinucleótido-quinase e ATP. Os oligonucleótidos tratados com a quinase são hibridados com o molde de cadeia simples e a cadeia mutagénica é sintetizada e ligada de acordo com o protocolo MutaGene, com a excepção de que Sequenase 2.0 proporciona a polimerase. As bibliotecas são pesquisadas utilizando nucleótidos marcados nas extremidades com g-³²P contendo sequências complementares às mutações introduzidas e através de isolamento do ADN plasmídico e identificação dos mutantes através da presença do local de restrição introduzido. As mutações são também confirmadas através de sequenciação com um estojo Sequenase. O ADN de HIV-Lock™ é clonado no sistema de expressão de baculovírus com um promotor de poliedro.

Fase Dois - Produção de Proteínas HIV-Lock™ Utilizando Baculovírus. Células Sf-9 são cultivadas até uma densidade predeterminada (cerca de 1×10^6 células/ml, fase logarítmica), infectadas com o baculovírus contendo as instruções de HIV-

Lock™ e colhidas para recuperar as proteínas recombinantes compreendendo o HIV-Lock™. No processo de ampliação, as culturas são expandidas de balões para garrafas e subsequentemente para biorreactores. Após infecção as células são colhidas às 12, 24, 36 e 48 horas pela proteína. Os índices de viabilidade são monitorizados ao longo de todo o processo.

Fase Três - Purificação das Proteínas HIV-Lock™. As proteínas colhidas são primeiro separadas dos particulados através de ultracentrifugação de fluxo para facilitar a purificação a jusante. O produto centrifugado é então filtrado para esterilização. Os extractos são então centrifugados a 40 000 rpm a 4°C durante 30 minutos e as alíquotas são imunoprecipitadas com anticorpo policlonal de coelho contra um dos componentes de HIV-Lock™. As proteínas imunoprecipitadas são corridas numa gel de SDS-PAGE a 10%.

Fase Quatro - Teste das Proteínas HIV-Lock™ Contra ADN de HIV. Os ensaios de mudança de mobilidade são realizados utilizando uma sonda oligonucleotídica compreendendo elementos da repetição terminal longa de HIV e fragmentos contendo NFκB que se liga a ADN associado à cadeia leve capa e a regulação de microglobulina. O oligonucleótido é hibridado com a sua cadeia complementar e marcado na extremidade com ATP γ -³²P.

O "footprinting" é alcançado combinando uma pequena quantidade (10^{-15} M) de ADN da LTR de HIV radiomarcado com uma quantidade ligeiramente maior de HIV-Lock™ num tampão à temperatura ambiente durante 10 minutos. O ditiotreitól é adicionado antes da adição da proteína. São adicionados ferro (II), EDTA, peróxido de hidrogénio e ascorbato de sódio e a mistura reaccional é incubada. É adicionado um agente de extinção e os produtos são analisados utilizando electroforese em gel desnaturante. Isto é feito para diferentes concentrações de proteína. O gel resultante é visualizado utilizando um digitalizador "Phosphoimager" e o ficheiro de imagem de elevada resolução resultante é analisado para calcular a afinidade de ligação de HIV-Lock™ para o ADN de HIV relativamente ao ADN celular.

Pode ser utilizado desenho múltiplo e interações de teste para refinar a ligação de HIV-Lock™ e outros TBA para HIV e outros organismos. Este processo torna possível desenhar conjuntos de ligação de modo a que o conjunto de ligação não seja competitivo com as proteínas de tipo selvagem para locais de ligação únicos nas amostras de genoma. O desenvolvimento de TBA para outros organismos e de TNA para sequências dentro destes organismos pode ser feito utilizando o método acima mencionado. Este método é válido quando a produção de conjuntos de ligação para todas as TBR de ácidos nucleicos incluindo híbridos ADN-ADN, ADN-ARN e ARN-ARN e combinações destes híbridos.

Exemplo 16 - Método para identificação de moléculas de ligação a ácidos nucleicos para produção de TBA e BBA do presente invento:

No método deste invento, os conjuntos de ligação ao alvo e os conjuntos de ligação de reforço são montados através da identificação de moléculas de ligação ao ácido nucleico, e da ligação das porções de ligação ao ácido nucleico das moléculas de modo a alcançar TBA que discriminam entre determinadas sequências alvo e até sequências intimamente aparentadas. Um método para identificação das moléculas de ligação ao ácido nucleico envolve os seguintes passos:

1. Obtenção de uma amostra biológica contendo o ácido nucleico alvo. Esta pode ser por exemplo, um organismo ou um extracto de tecido infectado com um patogénio.
2. Fragmentação da amostra de modo a expor os ácidos nucleicos e a reduzir a complexidade do tamanho do ácido nucleico contido na amostra.
3. Colocação de uma primeira alíquota dos ácidos nucleicos fragmentados em contacto com um meio tampão de controlo e colocação de uma segunda alíquota de ácidos nucleicos fragmentados em contacto com o meio tampão de controlo contendo um perfil conhecido de moléculas de ligação ao ácido nucleico.
4. Análise das duas alíquotas para identificar fragmentos que possuam um comportamento alterado na alíquota em contacto com as moléculas de ligação ao alvo em oposição à alíquota de controlo. Isto é alcançado através de

electroforese em gel de uma dimensão, electroforese em gel de duas dimensões, cromatografia líquida de alta resolução, cromatografia em papel ou através de outros meios que revelem um comportamento diferente dos fragmentos de ácido nucleico quando ligados a uma molécula de ligação ao ácido nucleico em oposição a quando o fragmento de ácido nucleico não está ligado.

5. Identificação e isolamento de fragmentos que exibem um comportamento alterado quando colocados em contacto com a molécula de ligação ao ácido nucleico e sequenciação do fragmento de ácido nucleico para determinar se os motivos da molécula de ligação ao ácido nucleico conhecidos estão presentes, ou identificação directa da molécula de ligação ao ácido nucleico ligada ao ácido nucleico. A última pode ser alcançada, por exemplo, através do contacto de uma grelha bidimensional dos ácidos nucleicos sujeitos a electroforese com anticorpos marcados diferencialmente que se ligam às várias moléculas de ligação ao ácido nucleico.

Neste método, de preferência os motivos do ácido nucleico são utilizados para fins diagnósticos ou terapêuticos em que o ácido nucleico alvo possui mais de um alvo molécula de ligação ao ácido nucleico utilizável. Deste modo, pode ser gerado um conjunto de ligação ao alvo complexo que tira vantagem da proximidade de diferentes motivos moleculares de ligação ao ácido nucleico para aumentar a especificidade do TBA montado a partir dos componentes de ligação ao ácido nucleico individuais identificados. As várias porções de ligação ao ácido nucleico das moléculas de ligação ao ácido nucleico são então montadas nos TBA completos tal como descrito acima, por exemplo, para HIV-LockTM.

Exemplo 17 - Método de identificação de sequências de ARN específicas numa amostra

De acordo com os métodos e composições ensinados neste invento, qualquer sequência de ácido nucleico pode ser especificamente identificada. A identificação do ARN de HIV alvo numa amostra é alcançada através da obtenção de uma amostra de sangue do paciente ou de outro fluido biológico ou extracto que possa conter o ARN de HIV, e de testes da

presença de locais de ligação a TAR. Tat é um regulador positivo da replicação de HIV que se liga à região TAR do ARN de HIV. A forma completamente activa mais pequena de ocorrência natural de Tat de HIV tem 72 aminoácidos de comprimento, SEQ ID NO:118 aqui. Tat contém pelo menos dois domínios funcionais e transactiva a expressão génica da repetição terminal longa de HIV (LTR de HIV). Tat liga-se a uma estrutura de caule e volta do ARN formada através da auto-hibridação de sequências em TAR, que ficam imediatamente a 5' da LTR de HIV. O ARN da TAR de HIV forma uma protuberância dinucleotídica e duas estruturas de caule e volta (Rhim et al., *Virology* 202: 202-211, 1994). Tat (SEQ ID NO:118) liga-se a esta estrutura com menor avidéz do que variantes de Tat em que Ala58 é uma treonina ou onde His65 é um resíduo Asp (Derse et al., *Virology* 194: 530-536, 1993). A utilização destes factos no presente método é alcançada através de:

1. Fragmentação de uma amostra biológica para expor os ácidos nucleicos e reduzir a complexidade do tamanho dos ácidos nucleicos.
2. Colocação de um TBA em contacto com a amostra que identifica uma sequência proteica de ligação a TAR híbrida e uma sequência de flanqueamento próxima no genoma de HIV. O TBA utilizado para este fim é montado em cro como acompanhante utilizando Tat como molécula de ligação específica do ARN de HIV. Para proporcionar especificidade tal como a comunicação cruzada entre o local TAR de HIV e locais TAR intimamente aparentados que possam estar presentes devido a outros patogénios tais como citomegalovírus, o TBA tem também um componente de anticorpo que reconhece a região de ligação ao híbrido ADN-ARN alvo formada quando o ácido nucleico sonda se liga ao ARN da LTR de HIV.
3. Eliminação de qualquer "comunicação cruzada" produzida através da ligação de Tat à região TAR do ARN de HIV devido a tais contaminantes (ARN primos) tais como a sequência TAR de CMV através do contacto da reacção com excesso de variante de Tat (as variantes Ala58 para Thr ou His65 para Asp) que se ligam mais avidamente. Deste modo, acontecimentos de ligação únicos devidos à ligação de TBA a ARN primos competem com a amostra de ácido nucleico pela variante de Tat. Por outro lado, através da

selecção apropriada da afinidade da dupla ligação alcançada em resultado do anticorpo e de Tat, o TBA não é deslocado dos verdadeiros alvos. Este processo é ilustrado na Figura 16. Noutro aspecto deste mesmo método, o TBA pode ser um em que, em vez de se utilizar uma variante de Tat, é utilizado um anticorpo que reconhece este segmento de ácido nucleico e o TBA utilizado é um TBA de duplo anticorpo.

Numa versão alternativa deste método, pode ser utilizado uma ácido nucleico sonda que hibride com o ARN da LTR de HIV. Concordantemente, pode ser criado um segmento dúplice dos locais Sp1 da LTR como parte da região de ligação ao alvo. Esta região do ARN de HIV flanqueia a região TAR que fica a 5' da LTR mas muito próximo desta. Um TBA contendo Tat e duas unidades de ligação de Sp1 é acompanhado para proporcionar ligação de Tat a TAR e ligação de Sp1 aos locais de ligação de Sp1. A amplificação e detecção são então realizadas através da adição de BNA, BBA e HNA apropriados. Ainda noutra alternativa, podem ser utilizados PNA possuindo SEQ ID NO:38 e SEQ ID NO:39 (ver Figura 7). É utilizado um TBA que contém uma ou mais unidades de ligação de Sp1 e uma unidade de anticorpo que se liga ao ADN-ARN híbrido produzido a partir do ARN da amostra e do PNA de SEQ ID NO:38. Os BNA, BBA e HNA apropriados são então adicionados para amplificar o sinal.

Naturalmente, os peritos na arte reconhecerão que podem ser utilizadas outras combinações de TBA e TNA para otimizar os métodos aqui exemplificados.

Entender-se-á que as sequências aqui proporcionadas são apenas exemplificativas e que outras sequências sugeridas por estas podem ser utilizadas nos métodos deste invento. Será também entendido que embora qualquer sequência aqui proporcionada possa ser desenhada como linear, pode ser utilizada numa forma circular ou permutada de outro modo e que embora designada como não sendo anti-sentido, pode ser utilizada na forma de codificação ou de não codificação ou para se ligar a sequências complementares codificantes ou não codificantes.

LISTAGEM DE SEQUÊNCIAS

(1) INFORMAÇÃO GERAL:

(i) REQUERENTE:

Nome(s) do(s) Requerente(s): THE GENE POOL, INC.
Rua: 300 Queen Anne Ave. N., Suite 392
Cidade: Seattle
Estado/Província: Washington
País: EUA
Código Postal/Zip: 98109-4599
Número de telefone: (206) 526-8617
Número de Fax:

(ii) TÍTULO DA INVENÇÃO: MÉTODO DE DETECÇÃO DE ÁCIDO NUCLEICOS
COM UMA COMPOSIÇÃO DE SEQUÊNCIA ESPECÍFICA

(iii) NÚMERO DE SEQUÊNCIAS: 118

(iv) ENDEREÇO PARA CORRESPONDÊNCIA:

(A) DESTINATÁRIO: Saliwanchik & Saliwanchik
(B) RUA: 2421 N.W. 41st St., Suite A-1
(C) CIDADE: Gainesville
(D) ESTADO: Florida
(E) PAÍS: EUA
(F) CÓDIGO POSTAL: 32606

(v) FORMATO LEGÍVEL COMPUTADOR:

(A) TIPO DE MEIO: disquete
(B) COMPUTADOR: compatível com PC IBM
(C) SISTEMA OPERATIVO: PC-DOS/MS-DOS
(D) SOFTWARE: PatentIn Release #1.0, Versão #1.25

(vi) DADOS DO PRESENTE PEDIDO:

(A) NÚMERO DE PEDIDO:
(B) DATA DE APRESENTAÇÃO:
(C) CLASSIFICAÇÃO:

(viii) INFORMAÇÃO SOBRE REPRESENTANTE/AGENTE:

(A) NOME: Gerard H Bencen
(B) NÚMERO DE REGISTRO: 35,746
(C) NÚMERO DE REFERÊNCIA/PROCESSO: GP-100.C1

(ix) INFORMAÇÃO PARA TELECOMUNICAÇÕES:

(A) TELEFONE: (904) 375-8100
(B) TELEFAX: (904) 372-5800

(2) INFORMAÇÃO PARA SEQ ID NO:1:

(i) CARACTERÍSTICAS DA SEQUÊNCIA:

(A) COMPRIMENTO: 13 pares de bases
(B) TIPO: ácido nucleico
(C) TIPO DE CADEIA: ambos
(D) TOPOLOGIA: linear

(ii) TIPO DE MOLÉCULA: ADNc

(iii) HIPOTÉTICA: NÃO

(iv) ANTI-SENTIDO: NÃO

(xi) DESCRIÇÃO DA SEQUÊNCIA: SEQ ID NO:1:
TGGGGATTCC CCA 13

(2) INFORMAÇÃO PARA SEQ ID NO:2:

(i) CARACTERÍSTICAS DA SEQUÊNCIA:

(A) COMPRIMENTO: 13 pares de bases

(B) TIPO: ácido nucleico

(C) TIPO DE CADEIA: ambos

(D) TOPOLOGIA: linear

(ii) TIPO DE MOLÉCULA: ADNc

(iii) HIPOTÉTICA: NÃO

(iv) ANTI-SENTIDO: NÃO

(xi) DESCRIÇÃO DA SEQUÊNCIA: SEQ ID NO:2:
AAGGGACTTT CCC 13

(2) INFORMAÇÃO PARA SEQ ID NO:3:

(i) CARACTERÍSTICAS DA SEQUÊNCIA:

(A) COMPRIMENTO: 13 pares de bases

(B) TIPO: ácido nucleico

(C) TIPO DE CADEIA: ambos

(D) TOPOLOGIA: linear

(ii) TIPO DE MOLÉCULA: ADNc

(iii) HIPOTÉTICA: NÃO

(iv) ANTI-SENTIDO: NÃO

(xi) DESCRIÇÃO DA SEQUÊNCIA: SEQ ID NO:3:
AGGGGACTTT CCG 13

(2) INFORMAÇÃO PARA SEQ ID NO:4:

(i) CARACTERÍSTICAS DA SEQUÊNCIA:

(A) COMPRIMENTO: 15 pares de bases

(B) TIPO: ácido nucleico

(C) TIPO DE CADEIA: ambos

(D) TOPOLOGIA: linear

(ii) TIPO DE MOLÉCULA: ADNc

(iii) HIPOTÉTICA: NÃO

(iv) ANTI-SENTIDO: NÃO

(xi) DESCRIÇÃO DA SEQUÊNCIA: SEQ ID NO:4:
GCTGGGGACT TTCCA 15

(2) INFORMAÇÃO PARA SEQ ID NO:5:

(i) CARACTERÍSTICAS DA SEQUÊNCIA:

- (A) COMPRIMENTO: 15 pares de bases
- (B) TIPO: ácido nucleico
- (C) TIPO DE CADEIA: ambos
- (D) TOPOLOGIA: linear

(ii) TIPO DE MOLÉCULA: ADnc

(iii) HIPOTÉTICA: NÃO

(iv) ANTI-SENTIDO: NÃO

(xi) DESCRIÇÃO DA SEQUÊNCIA: SEQ ID NO:5:
ACAAGGGACT TTCCG 15

(2) INFORMAÇÃO PARA SEQ ID NO:6:

(i) CARACTERÍSTICAS DA SEQUÊNCIA:

- (A) COMPRIMENTO: 13 pares de bases
- (B) TIPO: ácido nucleico
- (C) TIPO DE CADEIA: ambos
- (D) TOPOLOGIA: linear

(ii) TIPO DE MOLÉCULA: ADnc

(iii) HIPOTÉTICA: NÃO

(iv) ANTI-SENTIDO: NÃO

(xi) DESCRIÇÃO DA SEQUÊNCIA: SEQ ID NO:6:
CCGGGTTTTTCCC 13

(2) INFORMAÇÃO PARA SEQ ID NO:7:

(i) CARACTERÍSTICAS DA SEQUÊNCIA:

- (A) COMPRIMENTO: 27 pares de bases
- (B) TIPO: ácido nucleico
- (C) TIPO DE CADEIA: ambos
- (D) TOPOLOGIA: linear

(ii) TIPO DE MOLÉCULA: ADnc

(iii) HIPOTÉTICA: NÃO

(iv) ANTI-SENTIDO: NÃO

(xi) DESCRIÇÃO DA SEQUÊNCIA: SEQ ID NO:7:
AAGGGACTTT CCGCTGGGGA CTTTCCA 27

(2) INFORMAÇÃO PARA SEQ ID NO:8:

(i) CARACTERÍSTICAS DA SEQUÊNCIA:

- (A) COMPRIMENTO: 27 pares de bases
- (B) TIPO: ácido nucleico
- (C) TIPO DE CADEIA: ambos
- (D) TOPOLOGIA: linear

(ii) TIPO DE MOLÉCULA: ADnc

(iii) HIPOTÉTICA: NÃO

(iv) ANTI-SENTIDO: NÃO

(xi) DESCRIÇÃO DA SEQUÊNCIA: SEQ ID NO:8:

AAGGGACTTT CCGCTGGGGA CTTCCG 27

(2) INFORMAÇÃO PARA SEQ ID NO:9:

(i) CARACTERÍSTICAS DA SEQUÊNCIA:

- (A) COMPRIMENTO: 26 pares de bases
- (B) TIPO: ácido nucleico
- (C) TIPO DE CADEIA: ambos
- (D) TOPOLOGIA: linear

(ii) TIPO DE MOLÉCULA: ADnc

(iii) HIPOTÉTICA: NÃO

(iv) ANTI-SENTIDO: NÃO

(xi) DESCRIÇÃO DA SEQUÊNCIA: SEQ ID NO:9:

GCTGGGGACT TTCCAGGGAG GCGTGG 26

(2) INFORMAÇÃO PARA SEQ ID NO:10:

(i) CARACTERÍSTICAS DA SEQUÊNCIA:

- (A) COMPRIMENTO: 26 pares de bases
- (B) TIPO: ácido nucleico
- (C) TIPO DE CADEIA: ambos
- (D) TOPOLOGIA: linear

(ii) TIPO DE MOLÉCULA: ADnc

(iii) HIPOTÉTICA: NÃO

(iv) ANTI-SENTIDO: NÃO

(xi) DESCRIÇÃO DA SEQUÊNCIA: SEQ ID NO:10:

GCTGGGGACT TTCCAGGGGA GGTGTG 26

(2) INFORMAÇÃO PARA SEQ ID NO:11:

(i) CARACTERÍSTICAS DA SEQUÊNCIA:

- (A) COMPRIMENTO: 26 pares de bases
- (B) TIPO: ácido nucleico
- (C) TIPO DE CADEIA: ambos
- (D) TOPOLOGIA: linear

(ii) TIPO DE MOLÉCULA: ADNc

(iii) HIPOTÉTICA: NÃO

(iv) ANTI-SENTIDO: NÃO

(xi) DESCRIÇÃO DA SEQUÊNCIA: SEQ ID NO:11:
GCTGGGGACT TTCCGGGGAG CGTGGC 26

(2) INFORMAÇÃO PARA SEQ ID NO:12:

(i) CARACTERÍSTICAS DA SEQUÊNCIA:

(A) COMPRIMENTO: 26 pares de bases

(B) TIPO: ácido nucleico

(C) TIPO DE CADEIA: ambos

(D) TOPOLOGIA: linear

(ii) TIPO DE MOLÉCULA: ADNc

(iii) HIPOTÉTICA: NÃO

(iv) ANTI-SENTIDO: NÃO

(xi) DESCRIÇÃO DA SEQUÊNCIA: SEQ ID NO:12:
GCTGGGGACT TTCCGGGGAG GCGCGG 26

(2) INFORMAÇÃO PARA SEQ ID NO:13:

(i) CARACTERÍSTICAS DA SEQUÊNCIA:

(A) COMPRIMENTO: 26 pares de bases

(B) TIPO: ácido nucleico

(C) TIPO DE CADEIA: ambos

(D) TOPOLOGIA: linear

(ii) TIPO DE MOLÉCULA: ADNc

(iii) HIPOTÉTICA: NÃO

(iv) ANTI-SENTIDO: NÃO

(xi) DESCRIÇÃO DA SEQUÊNCIA: SEQ ID NO:13:
GCTGGGGACT TTCCAGAGAG GCGTGG 26

(2) INFORMAÇÃO PARA SEQ ID NO:14:

(i) CARACTERÍSTICAS DA SEQUÊNCIA:

(A) COMPRIMENTO: 26 pares de bases

(B) TIPO: ácido nucleico

(C) TIPO DE CADEIA: ambos

(D) TOPOLOGIA: linear

(ii) TIPO DE MOLÉCULA: ADNc

(iii) HIPOTÉTICA: NÃO

(iv) ANTI-SENTIDO: NÃO

(xi) DESCRIÇÃO DA SEQUÊNCIA: SEQ ID NO:14:
GCTGGGGACT TTCCAGGGGA GGCGTG 26

(2) INFORMAÇÃO PARA SEQ ID NO:15:

(i) CARACTERÍSTICAS DA SEQUÊNCIA:

- (A) COMPRIMENTO: 26 pares de bases
- (B) TIPO: ácido nucleico
- (C) TIPO DE CADEIA: ambos
- (D) TOPOLOGIA: linear

(ii) TIPO DE MOLÉCULA: ADnc

(iii) HIPOTÉTICA: NÃO

(iv) ANTI-SENTIDO: NÃO

(xi) DESCRIÇÃO DA SEQUÊNCIA: SEQ ID NO:15:
GCTGGGGACT TTCCAGGGAG GCGTGG 26

(2) INFORMAÇÃO PARA SEQ ID NO:16:

(i) CARACTERÍSTICAS DA SEQUÊNCIA:

- (A) COMPRIMENTO: 26 pares de bases
- (B) TIPO: ácido nucleico
- (C) TIPO DE CADEIA: ambos
- (D) TOPOLOGIA: linear

(ii) TIPO DE MOLÉCULA: ADnc

(iii) HIPOTÉTICA: NÃO

(iv) ANTI-SENTIDO: NÃO

(xi) DESCRIÇÃO DA SEQUÊNCIA: SEQ ID NO:16:
GCTGGGGACT TTCCAGGGAG GCTGCC 26

(2) INFORMAÇÃO PARA SEQ ID NO:17:

(i) CARACTERÍSTICAS DA SEQUÊNCIA:

- (A) COMPRIMENTO: 33 pares de bases
- (B) TIPO: ácido nucleico
- (C) TIPO DE CADEIA: ambos
- (D) TOPOLOGIA: linear

(ii) TIPO DE MOLÉCULA: ADnc

(iii) HIPOTÉTICA: NÃO

(iv) ANTI-SENTIDO: NÃO

(xi) DESCRIÇÃO DA SEQUÊNCIA: SEQ ID NO:17:
TTTCCAGGGA GGCGTGGCCT GGGCGGGACT GGG 33

(2) INFORMAÇÃO PARA SEQ ID NO:18:

(i) CARACTERÍSTICAS DA SEQUÊNCIA:

- (A) COMPRIMENTO: 33 pares de bases
- (B) TIPO: ácido nucleico
- (C) TIPO DE CADEIA: ambos
- (D) TOPOLOGIA: linear

(ii) TIPO DE MOLÉCULA: ADNc

(iii) HIPOTÉTICA: NÃO

(iv) ANTI-SENTIDO: NÃO

(xi) DESCRIÇÃO DA SEQUÊNCIA: SEQ ID NO:18:

CGTGGCCTGG GCGGGACTGG GGAGTGGCGT CCC 33

(2) INFORMAÇÃO PARA SEQ ID NO:19:

(i) CARACTERÍSTICAS DA SEQUÊNCIA:

- (A) COMPRIMENTO: 45 pares de bases
- (B) TIPO: ácido nucleico
- (C) TIPO DE CADEIA: ambos
- (D) TOPOLOGIA: linear

(ii) TIPO DE MOLÉCULA: ADNc

(iii) HIPOTÉTICA: NÃO

(iv) ANTI-SENTIDO: NÃO

(xi) DESCRIÇÃO DA SEQUÊNCIA: SEQ ID NO:19:

CTACAAGGGA CTTTCCGCTG GGGACTTTCC AGGGAGGCGT GGCCT 45

(2) INFORMAÇÃO PARA SEQ ID NO:20:

(i) CARACTERÍSTICAS DA SEQUÊNCIA:

- (A) COMPRIMENTO: 46 pares de bases
- (B) TIPO: ácido nucleico
- (C) TIPO DE CADEIA: ambos
- (D) TOPOLOGIA: linear

(ii) TIPO DE MOLÉCULA: ADNc

(iii) HIPOTÉTICA: NÃO

(iv) ANTI-SENTIDO: NÃO

(xi) DESCRIÇÃO DA SEQUÊNCIA: SEQ ID NO:20:

CAGCAAGGGA CTTTCCGCTG GGGACTTTCC AGGGGAGGTG TGGCCT 46

(2) INFORMAÇÃO PARA SEQ ID NO:21:

(i) CARACTERÍSTICAS DA SEQUÊNCIA:

- (A) COMPRIMENTO: 46 pares de bases
- (B) TIPO: ácido nucleico
- (C) TIPO DE CADEIA: ambos
- (D) TOPOLOGIA: linear

(ii) TIPO DE MOLÉCULA: ADNc

(iii) HIPOTÉTICA: NÃO

(iv) ANTI-SENTIDO: NÃO

(xi) DESCRIÇÃO DA SEQUÊNCIA: SEQ ID NO:21:

CATCAAGGGA CTTTCCGCTG GGGACTTTC AGGGGAGGTG TGGCCT 46

(2) INFORMAÇÃO PARA SEQ ID NO:22:

(i) CARACTERÍSTICAS DA SEQUÊNCIA:

(A) COMPRIMENTO: 46 pares de bases

(B) TIPO: ácido nucleico

(C) TIPO DE CADEIA: ambos

(D) TOPOLOGIA: linear

(ii) TIPO DE MOLÉCULA: ADNc

(iii) HIPOTÉTICA: NÃO

(iv) ANTI-SENTIDO: NÃO

(xi) DESCRIÇÃO DA SEQUÊNCIA: SEQ ID NO:22:

CAACAAGGGA CTTTCCGCTG GGGACTTTC AGGGGAGGTG TGGCCT 46

(2) INFORMAÇÃO PARA SEQ ID NO:23:

(i) CARACTERÍSTICAS DA SEQUÊNCIA:

(A) COMPRIMENTO: 45 pares de bases

(B) TIPO: ácido nucleico

(C) TIPO DE CADEIA: ambos

(D) TOPOLOGIA: linear

(ii) TIPO DE MOLÉCULA: ADNc

(iii) HIPOTÉTICA: NÃO

(iv) ANTI-SENTIDO: NÃO

(xi) DESCRIÇÃO DA SEQUÊNCIA: SEQ ID NO:23:

CTACAAGGGA CTTTCCGCTG GGGACTTTC AGGGAGGCGT GGCAT 45

(2) INFORMAÇÃO PARA SEQ ID NO:24:

(i) CARACTERÍSTICAS DA SEQUÊNCIA:

(A) COMPRIMENTO: 44 pares de bases

(B) TIPO: ácido nucleico

(C) TIPO DE CADEIA: ambos

(D) TOPOLOGIA: linear

(ii) TIPO DE MOLÉCULA: ADNc

(iii) HIPOTÉTICA: NÃO

(iv) ANTI-SENTIDO: NÃO

(xi) DESCRIÇÃO DA SEQUÊNCIA: SEQ ID NO:24:
CTACAAGGGA CTTTCCGCTG GGGACTTTCC GGGGAGCGTG GCCT 44

(2) INFORMAÇÃO PARA SEQ ID NO:25:

(i) CARACTERÍSTICAS DA SEQUÊNCIA:

- (A) COMPRIMENTO: 44 pares de bases
- (B) TIPO: ácido nucleico
- (C) TIPO DE CADEIA: ambos
- (D) TOPOLOGIA: linear

(ii) TIPO DE MOLÉCULA: ADnc

(iii) HIPOTÉTICA: NÃO

(iv) ANTI-SENTIDO: NÃO

(xi) DESCRIÇÃO DA SEQUÊNCIA: SEQ ID NO:25:
CTACAAGGGA CTTTCCGCTG GGGACTTTCC GGGGAGGCGC GGCT 44

(2) INFORMAÇÃO PARA SEQ ID NO:26:

(i) CARACTERÍSTICAS DA SEQUÊNCIA:

- (A) COMPRIMENTO: 45 pares de bases
- (B) TIPO: ácido nucleico
- (C) TIPO DE CADEIA: ambos
- (D) TOPOLOGIA: linear

(ii) TIPO DE MOLÉCULA: ADnc

(iii) HIPOTÉTICA: NÃO

(iv) ANTI-SENTIDO: NÃO

(xi) DESCRIÇÃO DA SEQUÊNCIA: SEQ ID NO:26:
CTACAAGGGA CTTTCCGCTG GGGACTTTCC AGAGAGGCGT GGACT 45

(2) INFORMAÇÃO PARA SEQ ID NO:27:

(i) CARACTERÍSTICAS DA SEQUÊNCIA:

- (A) COMPRIMENTO: 46 pares de bases
- (B) TIPO: ácido nucleico
- (C) TIPO DE CADEIA: ambos
- (D) TOPOLOGIA: linear

(ii) TIPO DE MOLÉCULA: ADnc

(iii) HIPOTÉTICA: NÃO

(iv) ANTI-SENTIDO: NÃO

(xi) DESCRIÇÃO DA SEQUÊNCIA: SEQ ID NO:27:
CTACAAGGGA CTTTCCGCTG GGGACTTTCC AGGGGAGGCG TGGACT 46

(2) INFORMAÇÃO PARA SEQ ID NO:28:

(i) CARACTERÍSTICAS DA SEQUÊNCIA:

- (A) COMPRIMENTO: 46 pares de bases
- (B) TIPO: ácido nucleico
- (C) TIPO DE CADEIA: ambos
- (D) TOPOLOGIA: linear

(ii) TIPO DE MOLÉCULA: ADNc

(iii) HIPOTÉTICA: NÃO

(iv) ANTI-SENTIDO: NÃO

(xi) DESCRIÇÃO DA SEQUÊNCIA: SEQ ID NO:28:

CTACAGGGGA CTTTCCGCTG GGGACTTTC AGGGAGGCGT GGGGAG 46

(2) INFORMAÇÃO PARA SEQ ID NO:29:

(i) CARACTERÍSTICAS DA SEQUÊNCIA:

- (A) COMPRIMENTO: 43 pares de bases
- (B) TIPO: ácido nucleico
- (C) TIPO DE CADEIA: ambos
- (D) TOPOLOGIA: linear

(ii) TIPO DE MOLÉCULA: ADNc

(iii) HIPOTÉTICA: NÃO

(iv) ANTI-SENTIDO: NÃO

(xi) DESCRIÇÃO DA SEQUÊNCIA: SEQ ID NO:29:

CTACAGGGGA CTTTCCGCTG GGGACTTTC AGGGAGGCTG CCT 43

(2) INFORMAÇÃO PARA SEQ ID NO:30:

(i) CARACTERÍSTICAS DA SEQUÊNCIA:

- (A) COMPRIMENTO: 48 pares de bases
- (B) TIPO: ácido nucleico
- (C) TIPO DE CADEIA: ambos
- (D) TOPOLOGIA: linear

(ii) TIPO DE MOLÉCULA: ADNc

(iii) HIPOTÉTICA: NÃO

(iv) ANTI-SENTIDO: NÃO

(xi) DESCRIÇÃO DA SEQUÊNCIA: SEQ ID NO:30:

CTTTCCGCTG GGGACTTTC AGGGAGGCGT GGCCTGGGCG GGACTGGG 48

(2) INFORMAÇÃO PARA SEQ ID NO:31:

(i) CARACTERÍSTICAS DA SEQUÊNCIA:

- (A) COMPRIMENTO: 45 pares de bases
- (B) TIPO: ácido nucleico
- (C) TIPO DE CADEIA: ambos
- (D) TOPOLOGIA: linear

(ii) TIPO DE MOLÉCULA: ADnc

(iii) HIPOTÉTICA: NÃO

(iv) ANTI-SENTIDO: NÃO

(xi) DESCRIÇÃO DA SEQUÊNCIA: SEQ ID NO:31:

TTTCCAGGGA GGCGTGGCCT GGGCGGGACT GGGGAGTGGC GTCCC 45

(2) INFORMAÇÃO PARA SEQ ID NO:32:

(i) CARACTERÍSTICAS DA SEQUÊNCIA:

(A) COMPRIMENTO: 59 pares de bases

(B) TIPO: ácido nucleico

(C) TIPO DE CADEIA: ambos

(D) TOPOLOGIA: linear

(ii) TIPO DE MOLÉCULA: ADnc

(iii) HIPOTÉTICA: NÃO

(iv) ANTI-SENTIDO: NÃO

(xi) DESCRIÇÃO DA SEQUÊNCIA: SEQ ID NO:32:

CTACAAGGGA CTTTCCGCTG GGGACTTTCC AGGGAGGCGT GGCCTGGGCG GGACTGGGG 59

(2) INFORMAÇÃO PARA SEQ ID NO:33:

(i) CARACTERÍSTICAS DA SEQUÊNCIA:

(A) COMPRIMENTO: 59 pares de bases

(B) TIPO: ácido nucleico

(C) TIPO DE CADEIA: ambos

(D) TOPOLOGIA: linear

(ii) TIPO DE MOLÉCULA: ADnc

(iii) HIPOTÉTICA: NÃO

(iv) ANTI-SENTIDO: NÃO

(xi) DESCRIÇÃO DA SEQUÊNCIA: SEQ ID NO:33:

TTTCCGCTGG GGACTTTCCA GGGAGGCGTG GCCTGGGCGG GACTGGGGAG TGGCGTCCC 59

(2) INFORMAÇÃO PARA SEQ ID NO:34:

(i) CARACTERÍSTICAS DA SEQUÊNCIA:

(A) COMPRIMENTO: 70 pares de bases

(B) TIPO: ácido nucleico

(C) TIPO DE CADEIA: ambos

(D) TOPOLOGIA: linear

(ii) TIPO DE MOLÉCULA: ADnc

(iii) HIPOTÉTICA: NÃO

(iv) ANTI-SENTIDO: NÃO

(xi) DESCRIÇÃO DA SEQUÊNCIA: SEQ ID NO:34:

CTACAAGGGA CTTTCCGCTG GGGACTTTC AGGGAGGCGT GCCTGGGCG GGACTGGGA 60
GTGGCGTCCC 70

(2) INFORMAÇÃO PARA SEQ ID NO:35:

(i) CARACTERÍSTICAS DA SEQUÊNCIA:

(A) COMPRIMENTO: 61 pares de bases

(B) TIPO: ácido nucleico

(C) TIPO DE CADEIA: ambos

(D) TOPOLOGIA: linear

(ii) TIPO DE MOLÉCULA: ADnc

(iii) HIPOTÉTICA: NÃO

(iv) ANTI-SENTIDO: NÃO

(xi) DESCRIÇÃO DA SEQUÊNCIA: SEQ ID NO:35:

TATCACCGCC AGTGGTATTT ATGTCAACAC CGCCAGAGAT AATTTATCAC CGCAGATGGT 60
T 61

(2) INFORMAÇÃO PARA SEQ ID NO:36:

(i) CARACTERÍSTICAS DA SEQUÊNCIA:

(A) COMPRIMENTO: 64 pares de bases

(B) TIPO: ácido nucleico

(C) TIPO DE CADEIA: ambos

(D) TOPOLOGIA: linear

(ii) TIPO DE MOLÉCULA: ADnc

(iii) HIPOTÉTICA: NÃO

(iv) ANTI-SENTIDO: NÃO

(xi) DESCRIÇÃO DA SEQUÊNCIA: SEQ ID NO:36:

TATCACCGCA AGGATAAAT ATCTAACACC GTGCGTGTG ACTATTTTAC CTCTGGCGGT 60
GATA 64

(2) INFORMAÇÃO PARA SEQ ID NO:37:

(i) CARACTERÍSTICAS DA SEQUÊNCIA:

(A) COMPRIMENTO: 70 pares de bases

(B) TIPO: ácido nucleico

(C) TIPO DE CADEIA: ambos

(D) TOPOLOGIA: linear

(ii) TIPO DE MOLÉCULA: ADnc

(iii) HIPOTÉTICA: NÃO

(iv) ANTI-SENTIDO: NÃO

(xi) DESCRIÇÃO DA SEQUÊNCIA: SEQ ID NO:37:

CTACAAGGGA CTTTCCGCTG GGGACTTTC AGGGAGGCGT GGCCTGGGCG GGACTGGGGA 60

GTGGCGTCCC 70

(2) INFORMAÇÃO PARA SEQ ID NO:38:

(i) CARACTERÍSTICAS DA SEQUÊNCIA:

(A) COMPRIMENTO: 37 pares de bases

(B) TIPO: ácido nucleico

(C) TIPO DE CADEIA: ambos

(D) TOPOLOGIA: linear

(ii) TIPO DE MOLÉCULA: ADNc

(iii) HIPOTÉTICA: NÃO

(iv) ANTI-SENTIDO: NÃO

(xi) DESCRIÇÃO DA SEQUÊNCIA: SEQ ID NO:38:

CTACAAGGGA CTTTCCGCTG GGGACTTTC AGGGAGG 37

(2) INFORMAÇÃO PARA SEQ ID NO:39:

(i) CARACTERÍSTICAS DA SEQUÊNCIA:

(A) COMPRIMENTO: 22 pares de bases

(B) TIPO: ácido nucleico

(C) TIPO DE CADEIA: ambos

(D) TOPOLOGIA: linear

(ii) TIPO DE MOLÉCULA: ADNc

(iii) HIPOTÉTICA: NÃO

(iv) ANTI-SENTIDO: NÃO

(xi) DESCRIÇÃO DA SEQUÊNCIA: SEQ ID NO:39:

CGGGACTGGG GAGTGGCGTC CC 22

(2) INFORMAÇÃO PARA SEQ ID NO:40:

(i) CARACTERÍSTICAS DA SEQUÊNCIA:

(A) COMPRIMENTO: 103 pares de bases

(B) TIPO: ácido nucleico

(C) TIPO DE CADEIA: ambos

(D) TOPOLOGIA: linear

(ii) TIPO DE MOLÉCULA: ADNc

(iii) HIPOTÉTICA: NÃO

(iv) ANTI-SENTIDO: NÃO

(xi) DESCRIÇÃO DA SEQUÊNCIA: SEQ ID NO:40:

CTACAAGGGA CTTTCCGCTG GGGACTTTC AGGGAGGTAT CACCGCCAGT GGTATTTATG 60
TCAACACCCG CAGAGATAAT TTATCACCCG AGATGGTTCT GCA 103

(2) INFORMAÇÃO PARA SEQ ID NO:41:

(i) CARACTERÍSTICAS DA SEQUÊNCIA:

- (A) COMPRIMENTO: 62 pares de bases
- (B) TIPO: ácido nucleico
- (C) TIPO DE CADEIA: ambos
- (D) TOPOLOGIA: linear

(ii) MÅ

(iii) HIPOTÉTICA: NÃO

(iv) ANTI-SENTIDO: NÃO

(xi) DESCRIÇÃO DA SEQUÊNCIA: SEQ ID NO:41;

GAACCATCTG CGGTGATAAA TTATCTCTGG CGGTGTTGAC ATAAATACCA CTGGCGGTGA 60
TA 62

(2) INFORMAÇÃO PARA SEQ ID NO:42:

(i) CARACTERÍSTICAS DA SEQUÊNCIA:

- (A) COMPRIMENTO: 71 pares de bases
- (B) TIPO: ácido nucleico
- (C) TIPO DE CADEIA: ambos
- (D) TOPOLOGIA: linear

(ii) TIPO DE MOLÉCULA: ADNc

(iii) HIPOTÉTICA: NÃO

(iv) ANTI-SENTIDO: NÃO

(xi) DESCRIÇÃO DA SEQUÊNCIA: SEQ ID NO:42:

GATCCAACCA TCTGCCGTGA TAAATTATCT CTGGCGGTGT TGACATAAAT ACCACTGGCG 60
GTGATACTGC A 71

(2) INFORMAÇÃO PARA SEQ ID NO:43:

(i) CARACTERÍSTICAS DA SEQUÊNCIA:

- (A) COMPRIMENTO: 63 pares de bases
- (B) TIPO: ácido nucleico
- (C) TIPO DE CADEIA: ambos
- (D) TOPOLOGIA: linear

(ii) TIPO DE MOLÉCULA: ADNc

(iii) HIPOTÉTICA: NÃO

(iv) ANTI-SENTIDO: NÃO

(xi) DESCRIÇÃO DA SEQUÊNCIA: SEQ ID NO:43:

GTATCACCGC CAGTGGTATT TATGTCAACA CCGCCAGAGA TAATTATCA CCGCAGATGG 60

TTG 63

(2) INFORMAÇÃO PARA SEQ ID NO:44:

(i) CARACTERÍSTICAS DA SEQUÊNCIA:

(A) COMPRIMENTO: 21 pares de bases

(B) TIPO: ácido nucleico

(C) TIPO DE CADEIA: ambos

(D) TOPOLOGIA: linear

(ii) TIPO DE MOLÉCULA: ADNc

(iii) HIPOTÉTICA: NÃO

(iv) ANTI-SENTIDO: NÃO

(xi) DESCRIÇÃO DA SEQUÊNCIA: SEQ ID NO:44:

GATCCGGGGG GATACCCCC G 21

(2) INFORMAÇÃO PARA SEQ ID NO:45:

(i) CARACTERÍSTICAS DA SEQUÊNCIA:

(A) COMPRIMENTO: 91 pares de bases

(B) TIPO: ácido nucleico

(C) TIPO DE CADEIA: ambos

(D) TOPOLOGIA: linear

(ii) TIPO DE MOLÉCULA: ADNc

(iii) HIPOTÉTICA: NÃO

(iv) ANTI-SENTIDO: NÃO

(xi) DESCRIÇÃO DA SEQUÊNCIA: SEQ ID NO:45:

CGGGACTGGG GAGTGGCGTC CCTATCACCG CAAGGGATAA ATATCTAACA CCGTGCCTGT 60

TGACTATTTT ACCTCTGGCG GTGATAGCAT G 91

(2) INFORMAÇÃO PARA SEQ ID NO:46:

(i) CARACTERÍSTICAS DA SEQUÊNCIA:

(A) COMPRIMENTO: 53 pares de bases

(B) TIPO: ácido nucleico

(C) TIPO DE CADEIA: ambos

(D) TOPOLOGIA: linear

(ii) TIPO DE MOLÉCULA: ADNc

(iii) HIPOTÉTICA: NÃO

(iv) ANTI-SENTIDO: NÃO

(xi) DESCRIÇÃO DA SEQUÊNCIA: SEQ ID NO:46:
CTAAGGGCGT AACCGAAATC GGTGAACCG AAACCGTTA GTATAAAAGC AGA 53

(2) INFORMAÇÃO PARA SEQ ID NO:47:

- (i) CARACTERÍSTICAS DA SEQUÊNCIA:
 (A) COMPRIMENTO: 54 pares de bases
 (B) TIPO: ácido nucleico
 (C) TIPO DE CADEIA: ambos
 (D) TOPOLOGIA: linear

(ii) TIPO DE MOLÉCULA: ADnc

(iii) HIPOTÉTICA: NÃO

(iv) ANTI-SENTIDO: NÃO

(xi) DESCRIÇÃO DA SEQUÊNCIA: SEQ ID NO:47:
AAAAGGGAGT AACCGAAAAC GGTCGGGACC GAAAACGGTG TATATAAAAG ATGT 54

(2) INFORMAÇÃO PARA SEQ ID NO:48:

- (i) CARACTERÍSTICAS DA SEQUÊNCIA:
 (A) COMPRIMENTO: 54 pares de bases
 (B) TIPO: ácido nucleico
 (C) TIPO DE CADEIA: ambos
 (D) TOPOLOGIA: linear

(ii) TIPO DE MOLÉCULA: ADnc

(iii) HIPOTÉTICA: NÃO

(iv) ANTI-SENTIDO: NÃO

(xi) DESCRIÇÃO DA SEQUÊNCIA: SEQ ID NO:48:
AGTAGGGTGT AACCGAAAGC GGTCAACCG AAAACGGTGC ATATATAAAG CAAA 54

(2) INFORMAÇÃO PARA SEQ ID NO:49:

- (i) CARACTERÍSTICAS DA SEQUÊNCIA:
 (A) COMPRIMENTO: 24 pares de bases
 (B) TIPO: ácido nucleico
 (C) TIPO DE CADEIA: ambos
 (D) TOPOLOGIA: linear

(ii) TIPO DE MOLÉCULA: ADnc

(iii) HIPOTÉTICA: NÃO

(iv) ANTI-SENTIDO: NÃO

(xi) DESCRIÇÃO DA SEQUÊNCIA: SEQ ID NO:49:
GCTTCAACCG AATTCGGTTG CATG 24

(2) INFORMAÇÃO PARA SEQ ID NO:50:

(i) CARACTERÍSTICAS DA SEQUÊNCIA:

- (A) COMPRIMENTO: 24 pares de bases
- (B) TIPO: ácido nucleico
- (C) TIPO DE CADEIA: ambos
- (D) TOPOLOGIA: linear

(ii) TIPO DE MOLÉCULA: ADNc

(iii) HIPOTÉTICA: NÃO

(iv) ANTI-SENTIDO: NÃO

(xi) DESCRIÇÃO DA SEQUÊNCIA: SEQ ID NO:50:

TGTGCAACCG ATTTGGTTG CCTT 24

(2) INFORMAÇÃO PARA SEQ ID NO:51:

(i) CARACTERÍSTICAS DA SEQUÊNCIA:

- (A) COMPRIMENTO: 24 pares de bases
- (B) TIPO: ácido nucleico
- (C) TIPO DE CADEIA: ambos
- (D) TOPOLOGIA: linear

(ii) TIPO DE MOLÉCULA: ADNc

(iii) HIPOTÉTICA: NÃO

(iv) ANTI-SENTIDO: NÃO

(xi) DESCRIÇÃO DA SEQUÊNCIA: SEQ ID NO:51:

TATGCAACCG AAATAGGTTG GGCA 24

(2) INFORMAÇÃO PARA SEQ ID NO:52:

(i) CARACTERÍSTICAS DA SEQUÊNCIA:

- (A) COMPRIMENTO: 24 pares de bases
- (B) TIPO: ácido nucleico
- (C) TIPO DE CADEIA: ambos
- (D) TOPOLOGIA: linear

(ii) TIPO DE MOLÉCULA: ADNc

(iii) HIPOTÉTICA: NÃO

(iv) ANTI-SENTIDO: NÃO

(xi) DESCRIÇÃO DA SEQUÊNCIA: SEQ ID NO:52:

TGCCTAACCG TTTTCGGTTA CTTG 24

(2) INFORMAÇÃO PARA SEQ ID NO:53:

(i) CARACTERÍSTICAS DA SEQUÊNCIA:

- (A) COMPRIMENTO: 24 pares de bases
- (B) TIPO: ácido nucleico
- (C) TIPO DE CADEIA: ambos
- (D) TOPOLOGIA: linear

(ii) TIPO DE MOLÉCULA: ADNc

(iii) HIPOTÉTICA: NÃO

(iv) ANTI-SENTIDO: NÃO

(xi) DESCRIÇÃO DA SEQUÊNCIA: SEQ ID NO:53:

GGACTAACCG TTTTAGGTCA TATT 24

(2) INFORMAÇÃO PARA SEQ ID NO:54:

(i) CARACTERÍSTICAS DA SEQUÊNCIA:

(A) COMPRIMENTO: 52 pares de bases

(B) TIPO: ácido nucleico

(C) TIPO DE CADEIA: ambos

(D) TOPOLOGIA: linear

(ii) TIPO DE MOLÉCULA: ADNc

(iii) HIPOTÉTICA: NÃO

(iv) ANTI-SENTIDO: NÃO

(xi) DESCRIÇÃO DA SEQUÊNCIA: SEQ ID NO:54:

GACGACTATC CAGCGACCAA GATCAGAGCC AGACACCGGA AACCCCTGCC AC 52

(2) INFORMAÇÃO PARA SEQ ID NO:55:

(i) CARACTERÍSTICAS DA SEQUÊNCIA:

(A) COMPRIMENTO: 53 pares de bases

(B) TIPO: ácido nucleico

(C) TIPO DE CADEIA: ambos

(D) TOPOLOGIA: linear

(ii) TIPO DE MOLÉCULA: ADNc

(iii) HIPOTÉTICA: NÃO

(iv) ANTI-SENTIDO: NÃO

(xi) DESCRIÇÃO DA SEQUÊNCIA: SEQ ID NO:55:

GACGACACGG TATCCGCTAC TCAGCTTGTT AACAGCTAC AGCACACCCC CTC 53

(2) INFORMAÇÃO PARA SEQ ID NO:56:

(i) CARACTERÍSTICAS DA SEQUÊNCIA:

(A) COMPRIMENTO: 60 pares de bases

(B) TIPO: ácido nucleico

(C) TIPO DE CADEIA: ambos

(D) TOPOLOGIA: linear

(ii) TIPO DE MOLÉCULA: ADNc

(iii) HIPOTÉTICA: NÃO

(iv) ANTI-SENTIDO: NÃO

(xi) DESCRIÇÃO DA SEQUÊNCIA: SEQ ID NO:56:

GACGACGACC TGCAGACACC ACAGACACCG CCCAGCCCCT TACAAAGCTG TTCTGTGCAG 60

(2) INFORMAÇÃO PARA SEQ ID NO:57:

(i) CARACTERÍSTICAS DA SEQUÊNCIA:

- (A) COMPRIMENTO: 68 pares de bases
- (B) TIPO: ácido nucleico
- (C) TIPO DE CADEIA: ambos
- (D) TOPOLOGIA: linear

(ii) TIPO DE MOLÉCULA: ADNc

(iii) HIPOTÉTICA: NÃO

(iv) ANTI-SENTIDO: NÃO

(xi) DESCRIÇÃO DA SEQUÊNCIA: SEQ ID NO:57:

CATACCAAAG CCGTGCCTT GGGCACCGAA GAAACACAAC CACTAAGTTG TTGCACAGAG 60
ACTCAGTG 68

(2) INFORMAÇÃO PARA SEQ ID NO:58:

(i) CARACTERÍSTICAS DA SEQUÊNCIA:

- (A) COMPRIMENTO: 77 pares de bases
- (B) TIPO: ácido nucleico
- (C) TIPO DE CADEIA: ambos
- (D) TOPOLOGIA: linear

(ii) TIPO DE MOLÉCULA: ADNc

(iii) HIPOTÉTICA: NÃO

(iv) ANTI-SENTIDO: NÃO

(xi) DESCRIÇÃO DA SEQUÊNCIA: SEQ ID NO:58:

TAATGTAATT GATTGTAATG ACTCTATGTG CAGTACCAGT ACCGTATTCC AGCACCGTGT 60
CCGTGGGCAC CGCAAAG 77

(2) INFORMAÇÃO PARA SEQ ID NO:59:

(i) CARACTERÍSTICAS DA SEQUÊNCIA:

- (A) COMPRIMENTO: 80 pares de bases
- (B) TIPO: ácido nucleico
- (C) TIPO DE CADEIA: ambos
- (D) TOPOLOGIA: linear

(ii) TIPO DE MOLÉCULA: ADNc

(iii) HIPOTÉTICA: NÃO

(iv) ANTI-SENTIDO: NÃO

(xi) DESCRIÇÃO DA SEQUÊNCIA: SEQ ID NO:59:

ACAGACAACG ATAACCGACC ACCACAAGCA GCGGCCAAAC ACCCCGCCTT GGACAATAGA 60
ACAGCACGTA CTGCAACTAA 80

(2) INFORMAÇÃO PARA SEQ ID NO:60:

(i) CARACTERÍSTICAS DA SEQUÊNCIA:

- (A) COMPRIMENTO: 266 pares de bases
- (B) TIPO: ácido nucleico
- (C) TIPO DE CADEIA: ambos
- (D) TOPOLOGIA: linear

(ii) TIPO DE MOLÉCULA: ADNc

(iii) HIPOTÉTICA: NÃO

(iv) ANTI-SENTIDO: NÃO

(xi) DESCRIÇÃO DA SEQUÊNCIA: SEQ ID NO:60:

CATATGCAAT ACAATGCATT ATACAACTG GACACATATA TATATTTGTG AAGAAGCATC 60
AGTAACTGTG GTAGAGGGTC AAGTTGACTA TTATGGTTTA TATTATGTTT ATGAAGGAAT 120
ACGAACATAT TTTGTGCAGT TTAAGATGA TGCAGAAAA TATAGTAAAA ATAAAGTATG 180
GGAAGTTCAT GCGGGTGGTC AGGTAATATT ATGTCCTACA TCTGTGTTTA GCAGCAACGA 240
AGTATCCTCT CCTGAAATTA TTAGGC 266

(2) INFORMAÇÃO PARA SEQ ID NO:61:

(i) CARACTERÍSTICAS DA SEQUÊNCIA:

- (A) COMPRIMENTO: 95 pares de bases
- (B) TIPO: ácido nucleico
- (C) TIPO DE CADEIA: ambos
- (D) TOPOLOGIA: linear

(ii) TIPO DE MOLÉCULA: ADNc

(iii) HIPOTÉTICA: NÃO

(iv) ANTI-SENTIDO: NÃO

(xi) DESCRIÇÃO DA SEQUÊNCIA: SEQ ID NO:61:

AGGATGTATA AAAAAACATG GATATACAGT GGAAGTGCAG TTTGATGGAG ACATATGCTA 60
TTAGGCAGCA CTTGGCCAAC CACCCCGCCG CGACC 95

(2) INFORMAÇÃO PARA SEQ ID NO:62:

(i) CARACTERÍSTICAS DA SEQUÊNCIA:

- (A) COMPRIMENTO: 81 pares de bases
- (B) TIPO: ácido nucleico
- (C) TIPO DE CADEIA: ambos
- (D) TOPOLOGIA: linear

- (ii) TIPO DE MOLÉCULA: ADNC
- (iii) HIPOTÉTICA: NÃO
- (iv) ANTI-SENTIDO: NÃO
- (xi) DESCRIÇÃO DA SEQUÊNCIA: SEQ ID NO:62:

CATGTTTTT TATACATCCA TATCACCGCC AGTGGTATTT ATGTCAACAC CGCCAGAGAT 60

AATTATCAC CGCAGATGGT T 81

(2) INFORMAÇÃO PARA SEQ ID NO:63:

- (i) CARACTERÍSTICAS DA SEQUÊNCIA:
 - (A) COMPRIMENTO: 322 aminoácidos
 - (B) TIPO: aminoácido
 - (D) TOPOLOGIA: linear

(ii) TIPO DE MOLÉCULA: péptido

(iii) HIPOTÉTICA: NÃO

(iv) ANTI-SENTIDO: NÃO

(v) TIPO DE FRAGMENTO: interno

(xi) DESCRIÇÃO DA SEQUÊNCIA: SEQ ID NO:63:

**Met Ala Asp Asp Asp Pro Tyr Gly Thr Gly Gln Met Phe His Leu Asn
1 5 10 15**

**Thr Ala Leu Thr His Ser Ile Phe Asn Ala Glu Leu Tyr Ser Pro Glu
20 25 30**

**Ile Pro Leu Ser Thr Asp Gly Pro Tyr Leu Gln Ile Leu Glu Gln Pro
35 40 45**

**Lys Gln Arg Gly Phe Arg Phe Arg Tyr Val Cys Glu Gly Pro Ser His
50 55 60**

**Gly Gly Leu Pro Gly Ala Ser Ser Glu Lys Asn Lys Lys Ser Tyr Pro
65 70 75 80**

**Gln Val Lys Ile Cys Asn Tyr Val Gly Pro Ala Lys Val Ile Val Gln
85 90 95**

**Leu Val Thr Asn Gly Lys Asn Ile His Leu His Ala His Ser Leu Val
100 105 110**

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

(2) INFORMAÇÃO PARA SEQ ID NO:64:

- (i) CARACTERÍSTICAS DA SEQUÊNCIA:
 - (A) COMPRIMENTO: 325 aminoácidos
 - (B) TIPO: aminoácido
 - (D) TOPOLOGIA: linear
- (ii) TIPO DE MOLÉCULA: péptido
- (iii) HIPOTÉTICA: NÃO
- (iv) ANTI-SENTIDO: NÃO
- (v) TIPO DE FRAGMENTO: interno

(xi) DESCRIÇÃO DA SEQUÊNCIA: SEQ ID NO:64:

Met Ala Glu Asp Asp Pro Tyr Leu Gly Arg Pro Glu Gln Met Phe His
 1 5 10 15
 Leu Asp Pro Ser Leu Thr His Thr Ile Phe Asn Pro Glu Val Phe Gln
 20 25 30
 Pro Gln Met Ala Leu Pro Thr Ala Asp Gly Pro Tyr Leu Gln Ile Leu
 35 40 45
 Glu Gln Pro Lys Gln Arg Gly Phe Arg Phe Arg Tyr Val Cys Glu Gly
 50 55 60
 Pro Ser His Gly Gly Leu Pro Gly Ala Ser Ser Glu Lys Asn Lys Lys
 65 70 75 80
 Ser Tyr Pro Gln Val Lys Ile Cys Asn Tyr Val Gly Pro Ala Lys Val
 85 90 95
 Ile Val Gln Leu Val Thr Asn Gly Lys Asn Ile His Leu His Ala His
 100 105 110
 Ser Leu Val Gly Lys His Cys Glu Asp Gly Ile Cys Thr Val Thr Ala
 115 120 125
 Gly Pro Glu Asp Cys Val His Gly Phe Ala Asn Leu Gly Ile Leu His
 130 135 140
 Val Thr Lys Lys Lys Val Phe Glu Thr Leu Glu Ala Arg Met Thr Glu
 145 150 155 160
 Ala Cys Ile Arg Gly Tyr Asn Pro Gly Leu Leu Val His Pro Asp Leu
 165 170 175
 Ala Tyr Leu Gln Ala Glu Gly Gly Gly Asp Arg Gln Leu Gly Asp Arg
 180 185 190
 Glu Lys Glu Leu Ile Arg Gln Ala Ala Leu Gln Gln Thr Lys Glu Met
 195 200 205
 Asp Leu Ser Val Val Arg Leu Met Phe Thr Ala Phe Leu Pro Asp Ser
 210 215 220
 Thr Gly Ser Phe Thr Arg Arg Leu Glu Pro Val Val Ser Asp Ala Ile
 225 230 235 240
 Tyr Asp Ser Lys Ala Pro Asn Ala Ser Asn Leu Lys Ile Val Arg Met
 245 250 255
 Asp Arg Thr Ala Gly Cys Val Thr Gly Gly Glu Glu Ile Tyr Leu Leu
 260 265 270
 Cys Asp Lys Val Gln Lys Asp Asp Ile Gln Ile Arg Phe Tyr Glu Glu
 275 280 285
 Glu Glu Asn Gly Gly Val Trp Glu Gly Phe Gly Asp Phe Ser Pro Thr
 290 295 300
 Asp Val His Arg Gln Phe Ala Ile Val Phe Lys Thr Pro Lys Tyr Lys
 305 310 315 320
 Asp Ile Asn Ile Thr
 325

(2) INFORMAÇÃO PARA SEQ ID NO:65:

(i) CARACTERÍSTICAS DA SEQUÊNCIA:

(A) COMPRIMENTO: 268 aminoácidos

(B) TIPO: aminoácido

(D) TOPOLOGIA: linear

(ii) TIPO DE MOLÉCULA: péptido

(iii) HIPOTÉTICA: NÃO

(iv) ANTI-SENTIDO: NÃO

(v) TIPO DE FRAGMENTO: interno

(xi) DESCRIÇÃO DA SEQUÊNCIA: SEQ ID NO:65:

Met	Glu	Pro	Ala	Asp	Leu	Leu	Pro	Leu	Tyr	Leu	Gln	Pro	Glu	Trp	Gly	1		5		10				15		
Glu	Gln	Glu	Pro	Gly	Gly	Ala	Thr	Pro	Phe	Val	Glu	Ile	Leu	Glu	Gln		20		25					30		
Pro	Lys	Gln	Arg	Gly	Met	Arg	Phe	Arg	Tyr	Lys	Cys	Glu	Gly	Arg	Ser		35		40				45			
Ala	Gly	Ser	Ile	Pro	Gly	Glu	His	Ser	Thr	Asp	Ser	Ala	Arg	Thr	His		50		55				60			
Pro	Thr	Ile	Arg	Val	Asn	His	Tyr	Arg	Gly	Pro	Gly	Arg	Val	Arg	Val	65		70			75					80
Ser	Leu	Val	Thr	Lys	Asp	Pro	Pro	His	Gly	Pro	His	Pro	His	Glu	Leu		85		90							95
Val	Gly	Arg	His	Cys	Gln	His	Gly	Tyr	Tyr	Glu	Ala	Glu	Leu	Ser	Pro		100		105							110
Asp	Arg	Ser	Ile	His	Ser	Phe	Gln	Asn	Leu	Gly	Ile	Gln	Cys	Val	Lys		115		120							125
Lys	Arg	Glu	Leu	Glu	Ala	Ala	Val	Ala	Glu	Arg	Ile	Arg	Thr	Asn	Asn		130		135							140
Asn	Pro	Phe	Asn	Val	Pro	Met	Glu	Glu	Arg	Gly	Ala	Glu	Tyr	Asp	Leu	145		150			155					160

Ser Ala Val Arg Leu Cys Phe Gln Val Trp Val Asn Gly Pro Gly Gly
 165 170 175
 Leu Cys Pro Leu Pro Pro Val Leu Ser Gln Pro Ile Tyr Asp Asn Arg
 180 185 190
 Ala Pro Ser Thr Ala Glu Leu Arg Ile Leu Pro Gly Asp Arg Asn Ser
 195 200 205
 Gly Ser Cys Gln Gly Gly Asp Glu Ile Phe Leu Leu Cys Asp Lys Val
 210 215 220
 Gln Lys Glu Asp Ile Glu Val Arg Phe Trp Ala Glu Gly Trp Glu Ala
 225 230 235 240
 Lys Gly Ser Phe Ala Ala Ala Asp Val His Arg Gln Val Ala Ile Val
 245 250 255
 Phe Arg Thr Pro Pro Phe Arg Glu Arg Ser Leu Arg
 260 265

(2) INFORMAÇÃO PARA SEQ ID NO:66:

(i) CARACTERÍSTICAS DA SEQUÊNCIA:

(A) COMPRIMENTO: 263 aminoácidos

(B) TIPO: aminoácido

(D) TOPOLOGIA: linear

(ii) TIPO DE MOLÉCULA: péptido

(iii) HIPOTÉTICA: NÃO

(iv) ANTI-SENTIDO: NÃO

(v) TIPO DE FRAGMENTO: interno

(xi) DESCRIÇÃO DA SEQUÊNCIA: SEQ ID NO:66:

Met Asp Asp Leu Phe Pro Leu Ile Phe Pro Ser Glu Pro Ala Gln Ala
 1 5 10 15
 Ser Gly Pro Tyr Val Glu Ile Ile Glu Gln Pro Lys Gln Arg Gly Met
 20 25 30
 Arg Phe Arg Tyr Lys Cys Glu Gly Arg Ser Ala Gly Ser Ile Pro Gly
 35 40 45
 Glu Arg Ser Thr Asp Thr Thr Lys Thr His Pro Thr Ile Lys Ile Asn
 50 55 60
 Gly Tyr Thr Gly Pro Gly Thr Val Arg Ile Ser Leu Val Thr Lys Asp
 65 70 75 80

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

(2) INFORMAÇÃO PARA SEQ ID NO:67:

(i) CARACTERÍSTICAS DA SEQUÊNCIA:

(A) COMPRIMENTO: 263 aminoácidos

(B) TIPO: aminoácido

(D) TOPOLOGIA: linear

(ii) TIPO DE MOLÉCULA: péptido

(iii) HIPOTÉTICA: NÃO

(iv) ANTI-SENTIDO: NÃO

(v) TIPO DE FRAGMENTO: interno

(xi) DESCRIÇÃO DA SEQUÊNCIA: SEQ ID NO:67:

Met Asp Glu Leu Phe Pro Leu Ile Phe Pro Ala Glu Pro Ala Gln Ala
 1 5 10 15

Ser Gly Pro Tyr Val Glu Ile Ile Glu Gln Pro Lys Gln Arg Gly Met
 20 25 30
 Arg Phe Arg Tyr Lys Cys Glu Gly Arg Ser Ala Gly Ser Ile Pro Gly
 35 40 45
 Glu Arg Ser Thr Asp Thr Thr Lys Thr His Pro Thr Ile Lys Ile Asn
 50 55 60
 Gly Tyr Thr Gly Pro Gly Thr Val Arg Ile Ser Leu Val Thr Lys Asp
 65 70 75 80
 Pro Pro His Arg Pro His Pro His Glu Leu Val Gly Lys Asp Cys Arg
 85 90 95
 Asp Gly Phe Tyr Glu Ala Glu Leu Cys Pro Asp Arg Cys Ile His Ser
 100 105 110
 Phe Gln Asn Leu Gly Ile Gln Cys Val Lys Lys Arg Asp Leu Glu Gln
 115 120 125
 Ala Ile Ser Gln Arg Ile Gln Thr Asn Asn Asn Pro Phe Gln Val Pro
 130 135 140
 Ile Glu Glu Gln Arg Gly Asp Tyr Asp Leu Asn Ala Val Arg Leu Cys
 145 150 155 160
 Phe Gln Val Thr Val Arg Asp Pro Ser Gly Arg Pro Leu Arg Leu Pro
 165 170 175
 Pro Val Leu Pro His Pro Ile Phe Asp Asn Arg Ala Pro Asn Thr Ala
 180 185 190
 Glu Leu Lys Ile Cys Arg Val Asn Arg Asn Ser Gly Ser Cys Leu Gly
 195 200 205
 Gly Asp Glu Ile Phe Leu Leu Cys Asp Lys Val Gln Lys Glu Asp Ile
 210 215 220
 Glu Val Tyr Phe Thr Gly Pro Gly Trp Glu Ala Arg Gly Ser Phe Ser
 225 230 235 240
 Gln Ala Asp Val His Arg Gln Val Ala Ile Val Phe Arg Thr Pro Pro
 245 250 255
 Tyr Ala Asp Pro Ser Leu Gln
 260

(2) INFORMAÇÃO PARA SEQ ID NO:68:

- (i) CARACTERÍSTICAS DA SEQUÊNCIA:
 - (A) COMPRIMENTO: 299 aminoácidos
 - (B) TIPO: aminoácido
 - (D) TOPOLOGIA: linear
- (ii) TIPO DE MOLÉCULA: péptido
- (iii) HIPOTÉTICA: NÃO
- (iv) ANTI-SENTIDO: NÃO
- (v) TIPO DE FRAGMENTO: interno

(xi) DESCRIÇÃO DA SEQUÊNCIA: SEQ ID NO:68:

Met Phe Pro Asn Gln Asn Asn Gly Ala Ala Pro Gly Gln Gly Pro Ala
 1 5 10 15
 Val Asp Gly Gln Gln Ser Leu Asn Tyr Asn Gly Leu Pro Ala Gln Gln
 20 25 30
 Gln Gln Gln Leu Ala Gln Ser Thr Lys Asn Val Arg Lys Lys Pro Tyr
 35 40 45
 Val Lys Ile Thr Glu Gln Pro Ala Gly Lys Ala Leu Arg Phe Arg Tyr
 50 55 60
 Glu Cys Glu Gly Arg Ser Ala Gly Ser Ile Pro Gly Val Asn Ser Thr
 65 70 75 80
 Pro Glu Asn Lys Thr Tyr Pro Thr Ile Glu Ile Val Gly Tyr Lys Gly
 85 90 95
 Arg Ala Val Val Val Val Ser Cys Val Thr Lys Asp Thr Pro Tyr Arg
 100 105 110
 Pro His Pro His Asn Leu Val Gly Lys Glu Gly Cys Lys Lys Gly Val
 115 120 125
 Cys Thr Leu Glu Ile Asn Ser Glu Thr Met Arg Ala Val Phe Ser Asn
 130 135 140
 Leu Gly Ile Gln Cys Val Lys Lys Lys Asp Ile Glu Ala Ala Leu Lys
 145 150 155 160
 Ala Arg Glu Glu Ile Arg Val Asp Pro Phe Lys Thr Gly Phe Ser His
 165 170 175
 Arg Phe Gln Pro Ser Ser Ile Asp Leu Asn Ser Val Arg Leu Cys Phe
 180 185 190
 Gln Val Phe Met Glu Ser Glu Gln Lys Gly Arg Phe Thr Ser Pro Leu
 195 200 205
 Pro Pro Val Val Ser Glu Pro Ile Phe Asp Lys Lys Ala Met Ser Asp
 210 215 220
 Leu Val Ile Cys Arg Leu Cys Ser Cys Ser Ala Thr Val Phe Gly Asn
 225 230 235 240
 Thr Gln Ile Ile Leu Leu Cys Glu Lys Val Ala Lys Glu Asp Ile Ser
 245 250 255
 Val Arg Phe Phe Glu Glu Lys Asn Gly Gln Ser Val Trp Glu Ala Phe
 260 265 270
 Gly Asp Phe Gln His Thr Asp Val His Lys Gln Thr Ala Ile Thr Phe
 275 280 285
 Lys Thr Pro Arg Tyr His Thr Leu Asp Ile Thr
 290 295

(2) INFORMAÇÃO PARA SEQ ID NO:69:

(i) CARACTERÍSTICAS DA SEQUÊNCIA:

(A) COMPRIMENTO: 261 aminoácidos

(B) TIPO: aminoácido

(D) TOPOLOGIA: linear

(ii) TIPO DE MOLÉCULA: péptido

(iii) HIPOTÉTICA: NÃO

(iv) ANTI-SENTIDO: NÃO

(v) TIPO DE FRAGMENTO: interno

(xi) DESCRIÇÃO DA SEQUÊNCIA: SEQ ID NO:69:

Met	Asp	Phe	Leu	Thr	Asn	Leu	Arg	Phe	Thr	Glu	Gly	Ile	Ser	Glu	Pro	1	5	10	15
Tyr	Ile	Glu	Ile	Phe	Glu	Gln	Pro	Arg	Gln	Arg	Gly	Thr	Arg	Phe	Arg	20	25	30	
Tyr	Lys	Cys	Glu	Gly	Arg	Ser	Ala	Gly	Ser	Ile	Pro	Gly	Glu	His	Ser	35	40	45	
Thr	Asp	Asn	Asn	Lys	Thr	Phe	Pro	Ser	Ile	Gln	Ile	Leu	Asn	Tyr	Phe	50	55	60	
Gly	Lys	Val	Lys	Ile	Arg	Thr	Thr	Leu	Val	Thr	Lys	Asn	Glu	Pro	Tyr	65	70	75	80
Lys	Pro	His	Pro	His	Asp	Leu	Val	Gly	Lys	Gly	Cys	Arg	Asp	Gly	Tyr	85	90	95	
Tyr	Glu	Ala	Glu	Phe	Gly	Pro	Glu	Arg	Gln	Val	Leu	Ser	Phe	Gln	Asn	100	105	110	
Leu	Gly	Ile	Gln	Cys	Val	Lys	Lys	Lys	Asp	Leu	Lys	Glu	Ser	Ile	Ser	115	120	125	
Leu	Arg	Ile	Ser	Lys	Lys	Asn	Pro	Phe	Asn	Val	Pro	Glu	Glu	Gln	Leu	130	135	140	

His Asn Ile Asp Glu Tyr Asp Leu Asn Val Val Arg Leu Cys Phe Gln
 145 150 155 160
 Ala Phe Leu Pro Asp Glu His Gly Asn Tyr Thr Leu Ala Leu Pro Pro
 165 170 175
 Leu Ile Ser Asn Pro Ile Tyr Asp Asn Arg Ala Pro Asn Thr Ala Glu
 180 185 190
 Leu Arg Ile Cys Arg Val Asn Lys Asn Cys Gly Ser Val Lys Gly Gly
 195 200 205
 Asp Glu Ile Phe Leu Leu Cys Asp Lys Val Gln Lys Asp Asp Ile Glu
 210 215 220
 Val Arg Phe Val Leu Gly Asn Trp Glu Ala Lys Gly Ser Phe Ser Gln
 225 230 235 240
 Ala Asp Val His Arg Gln Val Ala Ile Val Phe Arg Thr Pro Pro Phe
 245 250 255
 Leu Gly Asp Ile Thr
 260

(2) INFORMAÇÃO PARA SEQ ID NO:70:

(i) CARACTERÍSTICAS DA SEQUÊNCIA:

- (A) COMPRIMENTO: 262 aminoácidos
- (B) TIPO: aminoácido
- (D) TOPOLOGIA: linear

(ii) TIPO DE MOLÉCULA: péptido

(iii) HIPOTÉTICA: NÃO

(iv) ANTI-SENTIDO: NÃO

(v) TIPO DE FRAGMENTO: interno

(xi) DESCRIÇÃO DA SEQUÊNCIA: SEQ ID NO:70:

Met Asp Phe Leu Thr Asn Leu Arg Phe Thr Glu Gly Ile Ser Glu Pro
 1 5 10 15
 Tyr Ile Glu Ile Phe Glu Gln Pro Arg Gln Arg Gly Met Arg Phe Arg
 20 25 30
 Tyr Lys Cys Glu Gly Arg Ser Ala Gly Ser Ile Pro Gly Glu His Ser
 35 40 45
 Thr Asp Asn Asn Lys Thr Phe Pro Ser Ile Gln Ile Leu Asn Tyr Phe
 50 55 60
 Gly Lys Val Lys Ile Arg Thr Thr Leu Val Thr Lys Asn Glu Pro Tyr
 65 70 75 80

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

(2) INFORMAÇÃO PARA SEQ ID NO:71:

(i) CARACTERÍSTICAS DA SEQUÊNCIA:

(A) COMPRIMENTO: 314 aminoácidos

(B) TIPO: aminoácido

(D) TOPOLOGIA: linear

(ii) TIPO DE MOLÉCULA: péptido

(iii) HIPOTÉTICA: NÃO

(iv) ANTI-SENTIDO: NÃO

(v) TIPO DE FRAGMENTO: interno

(xi) DESCRIÇÃO DA SEQUÊNCIA: SEQ ID NO:71:

Met Ser Asn Lys Lys Gln Ser Asn Arg Leu Thr Glu Gln His Lys Leu
 1 5 10 15

Ser Gln Gly Val Ile Gly Ile Phe Gly Asp Tyr Ala Lys Ala His Asp
 20 25 30
 Leu Ala Val Gly Glu Val Ser Lys Leu Val Lys Lys Ala Leu Ser Asn
 35 40 45
 Glu Tyr Pro Gln Leu Ser Phe Arg Tyr Arg Asp Ser Ile Lys Lys Thr
 50 55 60
 Glu Ile Asn Glu Ala Leu Lys Lys Ile Asp Pro Asp Leu Gly Gly Thr
 65 70 75 80
 Leu Phe Val Ser Asn Ser Ser Ile Lys Pro Asp Gly Gly Ile Val Glu
 85 90 95
 Val Lys Asp Asp Tyr Gly Glu Trp Arg Val Val Leu Val Ala Glu Ala
 100 105 110
 Lys His Gln Gly Lys Asp Ile Ile Asn Ile Arg Asn Gly Leu Leu Val
 115 120 125
 Gly Lys Arg Gly Asp Gln Asp Leu Met Ala Ala Gly Asn Ala Ile Glu
 130 135 140
 Arg Ser His Asn Ile Ser Glu Ile Ala Asn Phe Met Leu Ser Glu Ser
 145 150 155 160
 His Phe Pro Tyr Val Leu Phe Leu Glu Gly Ser Asn Phe Leu Thr Glu
 165 170 175
 Asn Ile Ser Ile Thr Arg Pro Asp Gly Arg Val Val Asn Leu Glu Tyr
 180 185 190
 Asn Ser Gly Ser Glu Ser His Phe Pro Tyr Val Leu Phe Leu Glu Gly
 195 200 205
 Ser Asn Phe Leu Thr Glu Asn Ile Ser Ile Thr Arg Pro Asp Gly Arg
 210 215 220
 Val Val Asn Leu Glu Tyr Asn Ser Gly Ile Leu Asn Arg Leu Asp Arg
 225 230 235 240
 Leu Thr Ala Ala Asn Tyr Gly Met Pro Ile Asn Ser Asn Leu Cys Ile
 245 250 255
 Asn Lys Phe Val Asn His Lys Asp Lys Ser Ile Met Leu Gln Ala Ala
 260 265 270
 Ser Ile Tyr Thr Gln Gly Asp Gly Arg Glu Trp Asp Ser Lys Ile Met
 275 280 285
 Phe Glu Ile Met Phe Asp Ile Ser Thr Thr Ser Leu Arg Val Leu Gly
 290 295 300
 Arg Asp Leu Phe Glu Gln Leu Thr Ser Lys
 305 310

(2) INFORMAÇÃO PARA SEQ ID NO:72:

(i) CARACTERÍSTICAS DA SEQUÊNCIA:

- (A) COMPRIMENTO: 17 aminoácidos
- (B) TIPO: aminoácido
- (D) TOPOLOGIA: linear

(ii) TIPO DE MOLÉCULA: péptido

- (iii) HIPOTÉTICA: NÃO
 (iv) ANTI-SENTIDO: NÃO
 (v) TIPO DE FRAGMENTO: interno
 (xi) DESCRIÇÃO DA SEQUÊNCIA: SEQ ID NO:72:

Cys Asp Thr Asp Asp Arg His Arg Ile Glu Glu Lys Arg Lys Arg Lys
 1 5 10 15

Thr

(2) INFORMAÇÃO PARA SEQ ID NO:73:

- (i) CARACTERÍSTICAS DA SEQUÊNCIA:
 (A) COMPRIMENTO: 168 aminoácidos
 (B) TIPO: aminoácido
 (D) TOPOLOGIA: linear
 (ii) TIPO DE MOLÉCULA: péptido
 (iii) HIPOTÉTICA: NÃO
 (iv) ANTI-SENTIDO: NÃO
 (v) TIPO DE FRAGMENTO: interno
 (x) DESCRIÇÃO DA SEQUÊNCIA: SEQ ID NO:73:

Gly Asp Pro Gly Lys Lys Lys Gln His Ile Cys His Ile Gln Gly Cys
 1 5 10 15

Gly Lys Val Tyr Gly Lys Thr Ser His Leu Arg Ala His Leu Arg Trp
 20 25 30

His Thr Gly Glu Arg Pro Phe Met Cys Thr Trp Ser Tyr Cys Gly Lys
 35 40 45

Arg Phe Thr Arg Ser Asp Glu Leu Gln Arg His Lys Arg Thr His Thr
 50 55 60

Gly Glu Lys Lys Phe Ala Cys Pro Glu Cys Pro Lys Arg Phe Met Arg
 65 70 75 80

Ser Asp His Leu Ser Lys His Ile Lys Thr His Gln Asn Lys Lys Gly
 85 90 95

Gly Pro Gly Val Ala Leu Ser Val Gly Thr Leu Pro Leu Asp Ser Gly
 100 105 110

Ala Gly Ser Glu Gly Ser Gly Thr Ala Thr Pro Ser Ala Leu Ile Thr
 115 120 125

Thr Asn Met Val Ala Met Glu Ala Ile Cys Pro Glu Gly Ile Ala Arg
 130 135 140

Leu Ala Asn Ser Gly Ile Asn Val Met Gln Val Ala Asp Leu Gln Ser
 145 150 155 160

Ile Asn Ile Ser Gly Asn Gly Phe
 165

(2) INFORMAÇÃO PARA SEQ ID NO:74:

(i) CARACTERÍSTICAS DA SEQUÊNCIA:

- (A) COMPRIMENTO: 181 aminoácidos
- (B) TIPO: aminoácido
- (D) TOPOLOGIA: linear

(ii) TIPO DE MOLÉCULA: péptido

(iii) HIPOTÉTICA: NÃO

(iv) ANTI-SENTIDO: NÃO

(v) TIPO DE FRAGMENTO: interno

(xi) DESCRIÇÃO DA SEQUÊNCIA: SEQ ID NO:74:

```

Ser Gly Ile Val Pro Gln Leu Gln Asn Ile Val Ser Thr Val Asn Leu
1           5           10           15
Gly Cys Lys Leu Asp Leu Lys Thr Ile Ala Leu Arg Ala Arg Asn Ala
                20           25           30
Glu Tyr Asn Pro Lys Arg Phe Ala Ala Val Ile Met Arg Ile Arg Glu
                35           40           45
Pro Arg Thr Thr Ala Leu Ile Phe Ser Ser Gly Lys Met Val Cys Thr
                50           55           60
Gly Ala Lys Ser Glu Glu Gln Ser Arg Leu Ala Ala Arg Lys Tyr Ala
65           70           75           80
Arg Val Val Gln Lys Leu Gly Phe Pro Ala Lys Phe Leu Asp Phe Lys
                85           90           95
Ile Gln Asn Met Val Gly Ser Cys Asp Val Lys Phe Pro Ile Arg Leu
                100          105          110
Glu Gly Leu Val Leu Thr His Gln Gln Phe Ser Ser Tyr Glu Pro Glu
                115          120          125
Leu Phe Pro Gly Leu Ile Tyr Arg Met Ile Lys Pro Arg Ile Val Leu
                130          135          140
Leu Ile Phe Val Ser Gly Lys Val Val Leu Thr Gly Ala Lys Val Arg
145          150          155          160
Ala Glu Ile Tyr Glu Ala Phe Glu Asn Ile Tyr Pro Ile Leu Lys Gly
                165          170          175
Phe Arg Lys Thr Thr
                180

```

(2) INFORMAÇÃO PARA SEQ ID NO:75:

(i) CARACTERÍSTICAS DA SEQUÊNCIA:

- (A) COMPRIMENTO: 85 aminoácidos
- (B) TIPO: aminoácido
- (D) TOPOLOGIA: linear

(ii) TIPO DE MOLÉCULA: péptido

(iii) HIPOTÉTICA: NÃO

(iv) ANTI-SENTIDO: NÃO

(v) TIPO DE FRAGMENTO: interno

(xi) DESCRIÇÃO DA SEQUÊNCIA: SEQ ID NO:75:

Ser Cys Phe Ala Leu Ile Ser Gly Thr Ala Asn Gln Val Lys Cys Tyr
1 5 10 15

Arg Phe Arg Val Lys Lys Asn His Arg His Arg Tyr Glu Asn Cys Thr
20 25 30

Thr Thr Trp Phe Thr Val Ala Asp Asn Gly Ala Glu Arg Gln Gly Gln
35 40 45

Ala Gln Ile Leu Ile Thr Phe Gly Ser Pro Ser Gln Arg Gln Asp Phe
50 55 60

Leu Lys His Val Pro Leu Pro Pro Gly Met Asn Ile Ser Gly Phe Thr
65 70 75 80

Ala Ser Leu Asp Phe
85

(2) INFORMAÇÃO PARA SEQ ID NO:76:

(i) CARACTERÍSTICAS DA SEQUÊNCIA:

(A) COMPRIMENTO: 87 aminoácidos

(B) TIPO: aminoácido

(D) TOPOLOGIA: linear

(ii) TIPO DE MOLÉCULA: péptido

(iii) HIPOTÉTICA: NÃO

(iv) ANTI-SENTIDO: NÃO

(v) TIPO DE FRAGMENTO: interno

(xi) DESCRIÇÃO DA SEQUÊNCIA: SEQ ID NO:76:

Cys Pro Cys Leu Leu Ile Gly Thr Ser Gly Asn Gly Asn Gln Val Lys
1 5 10 15

Cys Tyr Ser Phe Arg Val Lys Arg Trp His Asp Arg Asp Lys Tyr His
20 25 30

His Thr Thr Thr Trp Trp Ala Val Gly Gly Gln Gly Ser Glu Arg Pro
35 40 45

Gly Asp Ala Thr Val Ile Val Thr Phe Lys Asp Gln Ser Gln Arg Ser
50 55 60

His Phe Leu Gln Gln Val Pro Leu Pro Pro Gly Met Ser Ala His Gly
65 70 75 80

Val Thr Met Thr Val Asp Phe
85

(2) INFORMAÇÃO PARA SEQ ID NO:77:

(i) CARACTERÍSTICAS DA SEQUÊNCIA:

(A) COMPRIMENTO: 84 aminoácidos

(B) TIPO: aminoácido

(D) TOPOLOGIA: linear

(ii) TIPO DE MOLÉCULA: péptido

(iii) HIPOTÉTICA: NÃO

(iv) ANTI-SENTIDO: NÃO

(v) TIPO DE FRAGMENTO: interno

(xi) DESCRIÇÃO DA SEQUÊNCIA: SEQ ID NO:77:

Pro Pro Val Ile Cys Leu Lys Gly Gly His Asn Gln Leu Lys Cys Leu
 1 5 10 15
 Arg Tyr Arg Leu Lys Ser Lys His Ser Ser Leu Phe Asp Cys Ile Ser
 20 25 30
 Thr Thr Trp Ser Trp Val Asp Thr Thr Ser Thr Cys Arg Leu Gly Ser
 35 40 45
 Gly Arg Met Leu Ile Lys Phe Ala Asp Ser Glu Gln Arg Asp Lys Phe
 50 55 60
 Leu Ser Arg Val Pro Leu Pro Ser Thr Thr Gln Val Phe Leu Gly Asn
 65 70 75 80
 Phe Tyr Gly Leu

(2) INFORMAÇÃO PARA SEQ ID NO:78:

(i) CARACTERÍSTICAS DA SEQUÊNCIA:

(A) COMPRIMENTO: 84 aminoácidos

(B) TIPO: aminoácido

(D) TOPOLOGIA: linear

(ii) TIPO DE MOLÉCULA: péptido

(iii) HIPOTÉTICA: NÃO

(iv) ANTI-SENTIDO: NÃO

(v) TIPO DE FRAGMENTO: interno

(xi) DESCRIÇÃO DA SEQUÊNCIA: SEQ ID NO:78:

Pro Pro Val Ile Leu Val Arg Gly Gly Ala Asn Thr Leu Lys Cys Phe
 1 5 10 15
 Arg Asn Arg Ala Arg Val Arg Tyr Arg Gly Leu Phe Lys Tyr Phe Ser
 20 25 30
 Thr Thr Trp Ser Trp Val Ala Gly Asp Ser Thr Glu Arg Leu Gly Arg
 35 40 45
 Ser Arg Met Leu Ile Leu Phe Thr Ser Ala Cys Gln Arg Glu Lys Pro
 50 55 60
 Asp Glu Thr Val Lys Tyr Pro Lys Gly Val Asp Thr Ser Tyr Gly Asn
 65 70 75 80
 Leu Asp Ser Leu

(2) INFORMAÇÃO PARA SEQ ID NO:79:

(i) CARACTERÍSTICAS DA SEQUÊNCIA:

(A) COMPRIMENTO: 84 aminoácidos

(B) TIPO: aminoácido

(D) TOPOLOGIA: linear

(ii) TIPO DE MOLÉCULA: péptido

(iii) HIPOTÉTICA: NÃO

(iv) ANTI-SENTIDO: NÃO

(v) TIPO DE FRAGMENTO: interno

(x) DESCRIÇÃO DA SEQUÊNCIA: SEQ ID NO:79:

Pro Pro Val Val Cys Val Lys Gly Gly Ala Asn Gln Leu Lys Cys Leu
1 5 10 15

Arg Tyr Arg Leu Lys Ala Ser Thr Gln Val Asp Phe Asp Ser Ile Ser
20 25 30

Thr Thr Trp His Trp Thr Asp Arg Lys Asn Thr Glu Arg Ile Gly Ser
35 40 45

Ala Arg Met Leu Val Lys Phe Ile Asp Glu Ala Gln Arg Glu Lys Phe
50 55 60

Leu Glu Arg Val Ala Leu Pro Arg Ser Val ser Val Phe Leu Gly Gln
65 70 75 80

Phe Asn Gly Ser

(2) INFORMAÇÃO PARA SEQ ID NO:80:

(i) CARACTERÍSTICAS DA SEQUÊNCIA:

(A) COMPRIMENTO: 84 aminoácidos

(B) TIPO: aminoácido

(D) TOPOLOGIA: linear

(ii) TIPO DE MOLÉCULA: péptido

(iii) HIPOTÉTICA: NÃO

(iv) ANTI-SENTIDO: NÃO

(v) TIPO DE FRAGMENTO: interno

(xi) DESCRIÇÃO DA SEQUÊNCIA: SEQ ID NO:80:

Thr Pro Ile Val Gln Leu Gln Gly Asp Ser Asn Cys Leu Lys Cys Phe
1 5 10 15

Arg Tyr Arg Leu Asn Asp Lys Tyr Lys His Leu Phe Glu Leu Ala Ser
20 25 30

Ser Thr Trp His Trp Ala Ser Pro Glu Ala Pro His Lys Asn Ala Ile
35 40 45

Val Thr Leu Thr Tyr Ser Ser Glu Glu Gln Arg Gln Gln Phe Leu Asn
50 55 60

Ser Val Lys Ile Pro Pro Thr Ile Arg His Lys Val Gly Phe Met Ser
65 70 75 80

Leu His Leu Leu

(2) INFORMAÇÃO PARA SEQ ID NO:81:

(i) CARACTERÍSTICAS DA SEQUÊNCIA:

(A) COMPRIMENTO: 84 aminoácidos

(B) TIPO: aminoácido

(D) TOPOLOGIA: linear

(ii) TIPO DE MOLÉCULA: péptido

(iii) HIPOTÉTICA: NÃO

(iv) ANTI-SENTIDO: NÃO

(v) TIPO DE FRAGMENTO: interno

(xi) DESCRIÇÃO DA SEQUÊNCIA: SEQ ID NO:81:

Thr Pro Ile Val Gln Phe Gln Gly Glu Ser Asn Cys Leu Lys Cys Phe
1 5 10 15

Arg Tyr Arg Leu Asn Arg Asp His Arg His Leu Phe Asp Leu Ile Ser
20 25 30

Ser Thr Trp His Trp Ala Ser Ser Lys Ala Pro His Lys His Ala Ile
35 40 45

Val Thr Val Thr Tyr Asp Ser Glu Glu Gln Arg Gln Gln Phe Leu Asp
50 55 60

Val Val Lys Ile Pro Pro Thr Ile Ser His Lys Leu Gly Phe Met Ser
65 70 75 80

Leu His Leu Leu

(2) INFORMAÇÃO PARA SEQ ID NO:82:

(i) CARACTERÍSTICAS DA SEQUÊNCIA:

(A) COMPRIMENTO: 80 aminoácidos

(B) TIPO: aminoácido

(D) TOPOLOGIA: linear

(ii) TIPO DE MOLÉCULA: péptido

(iii) HIPOTÉTICA: NÃO

(iv) ANTI-SENTIDO: NÃO

(v) TIPO DE FRAGMENTO: interno

(xi) DESCRIÇÃO DA SEQUÊNCIA: SEQ ID NO:82:

Thr Pro Ile Ile His Leu Lys Gly Asp Arg Asn Ser Leu Lys Cys Leu
1 5 10 15
Arg Tyr Arg Leu Arg Lys His Ser Asp His Tyr Arg Asp Ile Ser Ser
20 25 30
Thr Trp His Trp Thr Gly Ala Gly Asn Glu Lys Thr Gly Ile Leu Thr
35 40 45
Val Thr Tyr His Ser Glu Thr Gln Arg Thr Lys Phe Leu Asn Thr Val
50 55 60
Ala Ile Pro Asp Ser Val Gln Ile Leu Val Gly Tyr Asn Thr Met Tyr
65 70 75 80

(2) INFORMAÇÃO PARA SEQ ID NO:83:

(i) CARACTERÍSTICAS DA SEQUÊNCIA:

- (A) COMPRIMENTO: 80 aminoácidos
- (B) TIPO: aminoácido
- (D) TOPOLOGIA: linear

(ii) TIPO DE MOLÉCULA: péptido

(iii) HIPOTÉTICA: NÃO

(iv) ANTI-SENTIDO: NÃO

(v) TIPO DE FRAGMENTO: interno

(xi) DESCRIÇÃO DA SEQUÊNCIA: SEQ ID NO:83:

Thr Pro Ile Val His Leu Lys Gly Asp Ala Asn Thr Leu Lys Cys Leu
1 5 10 15
Arg Tyr Arg Phe Lys Lys His Cys Thr Leu Tyr Thr Ala Val Ser Ser
20 25 30
Thr Trp His Trp Thr Gly His Asn Tyr Lys His Lys Ser Ala Ile Val
35 40 45
Thr Leu Thr Tyr Asp Ser Glu Trp Gln Arg Asp Gln Phe Leu Ser Gln
50 55 60
Val Lys Ile Pro Lys Thr Ile Thr Val Ser Thr Gly Phe Met Ser Ile
65 70 75 80

(2) INFORMAÇÃO PARA SEQ ID NO:84:

(i) CARACTERÍSTICAS DA SEQUÊNCIA:

- (A) COMPRIMENTO: 81 aminoácidos
- (B) TIPO: aminoácido
- (D) TOPOLOGIA: linear

(ii) TIPO DE MOLÉCULA: péptido

(iii) HIPOTÉTICA: NÃO

(iv) ANTI-SENTIDO: NÃO

(v) TIPO DE FRAGMENTO: interno

(xi) DESCRIÇÃO DA SEQUÊNCIA: SEQ ID NO:84:

Ala	Pro	Ile	Val	His	Leu	Lys	Gly	Glu	Ser	Asn	Ser	Leu	Lys	Cys	Leu
1				5					10					15	
Arg	Tyr	Arg	Leu	Lys	Pro	Tyr	Asn	Glu	Leu	Tyr	Ser	Ser	Met	Ser	Ser
			20					25					30		
Thr	Trp	His	Trp	Thr	Ser	Asp	Asn	Lys	Asn	Ser	Lys	Asn	Gly	Ile	Val
		35					40						45		
Thr	Val	Thr	Phe	Val	Thr	Gly	Gln	Gln	Gln	Gln	Met	Phe	Leu	Gly	Thr
	50					55				60					
Val	Lys	Ile	Pro	Pro	Thr	Val	Gln	Ile	Ser	Thr	Gly	Phe	Met	Thr	Leu
65					70					75					80
Val															

(2) INFORMAÇÃO PARA SEQ ID NO:85:

(i) CARACTERÍSTICAS DA SEQUÊNCIA:

(A) COMPRIMENTO: 21 aminoácidos

(B) TIPO: aminoácido

(D) TOPOLOGIA: linear

(ii) TIPO DE MOLÉCULA: péptido

(iii) HIPOTÉTICA: NÃO

(iv) ANTI-SENTIDO: NÃO

(v) TIPO DE FRAGMENTO: interno

(xi) DESCRIÇÃO DA SEQUÊNCIA: SEQ ID NO:85:

Gly	Ile	Val	Glu	Gln	Cys	Cys	Thr	Ser	Ile	Cys	Ser	Leu	Tyr	Gln	Leu
1				5					10					15	
Glu	Asn	Tyr	Cys	Asn											
			20												

(2) INFORMAÇÃO PARA SEQ ID NO:86:

(i) CARACTERÍSTICAS DA SEQUÊNCIA:

(A) COMPRIMENTO: 30 aminoácidos

(B) TIPO: aminoácido

(D) TOPOLOGIA: linear

(ii) TIPO DE MOLÉCULA: péptido

(iii) HIPOTÉTICA: NÃO

(iv) ANTI-SENTIDO: NÃO

(v) TIPO DE FRAGMENTO: interno

(xi) DESCRIÇÃO DA SEQUÊNCIA: SEQ ID NO:86:

Phe	Val	Asn	Gln	His	Leu	Cys	Gly	Ser	His	Leu	Val	Glu	Ala	Leu	Tyr
1				5					10					15	
Leu	Val	Cys	Gly	Glu	Arg	Gly	Phe	Phe	Tyr	Thr	Pro	Lys	Thr		
			20					25					30		

(2) INFORMAÇÃO PARA SEQ ID NO:87:

(i) CARACTERÍSTICAS DA SEQUÊNCIA:

(A) COMPRIMENTO: 21 aminoácidos

(B) TIPO: aminoácido

(D) TOPOLOGIA: linear

(ii) TIPO DE MOLÉCULA: péptido

(iii) HIPOTÉTICA: NÃO

(iv) ANTI-SENTIDO: NÃO

(v) TIPO DE FRAGMENTO: interno

(xi) DESCRIÇÃO DA SEQUÊNCIA: SEQ ID NO:87:

Gly	Ile	Val	Glu	Gln	Cys	Cys	Ala	Ser	Val	Cys	Ser	Leu	Tyr	Gln	Leu
1				5					10					15	
Glu	Asn	Tyr	Cys	Asn											
			20												

(2) INFORMAÇÃO PARA SEQ ID NO:88:

(i) CARACTERÍSTICAS DA SEQUÊNCIA:

(A) COMPRIMENTO: 30 aminoácidos

(B) TIPO: aminoácido

(D) TOPOLOGIA: linear

(ii) TIPO DE MOLÉCULA: péptido

(iii) HIPOTÉTICA: NÃO

(iv) ANTI-SENTIDO: NÃO

(v) TIPO DE FRAGMENTO: interno

(xi) DESCRIÇÃO DA SEQUÊNCIA: SEQ ID NO:88:

Phe	Val	Asn	Gln	His	Leu	Cys	Gly	Ser	His	Leu	Val	Glu	Ala	Leu	Tyr
1				5					10					15	
Leu	Val	Cys	Gly	Glu	Arg	Gly	Phe	Phe	Tyr	Thr	Pro	Lys	Thr		
			20					25					30		

(2) INFORMAÇÃO PARA SEQ ID NO:89:

(i) CARACTERÍSTICAS DA SEQUÊNCIA:

(A) COMPRIMENTO: 24 aminoácidos

(B) TIPO: aminoácido

(D) TOPOLOGIA: linear

- (ii) TIPO DE MOLÉCULA: péptido
- (iii) HIPOTÉTICA: NÃO
- (iv) ANTI-SENTIDO: NÃO
- (v) TIPO DE FRAGMENTO: interno
- (xi) DESCRIÇÃO DA SEQUÊNCIA: SEQ ID NO:89:

Gln Leu Tyr Ser Ala Leu Ala Asn Lys Cys Cys His Val Gly Cys Ile
 1 5 10 15
 Lys Arg Ser Leu Ala Arg Phe Cys
 20

(2) INFORMAÇÃO PARA SEQ ID NO:90:

- (i) CARACTERÍSTICAS DA SEQUÊNCIA:
 - (A) COMPRIMENTO: 33 aminoácidos
 - (B) TIPO: aminoácido
 - (D) TOPOLOGIA: linear
- (ii) TIPO DE MOLÉCULA: péptido
- (iii) HIPOTÉTICA: NÃO
- (iv) ANTI-SENTIDO: NÃO
- (v) TIPO DE FRAGMENTO: interno
- (xi) DESCRIÇÃO DA SEQUÊNCIA: SEQ ID NO:90:

Asp Ser Trp Met Glu Glu Val Ile Lys Ile Cys Gly Arg Glu Leu Val
 1 5 10 15
 Arg Ala Gln Ile Ala Ile Cys Gly Met Ser Thr Trp Ser Lys Arg Ser
 20 25 30
 Leu

(2) INFORMAÇÃO PARA SEQ ID NO:91:

- (i) CARACTERÍSTICAS DA SEQUÊNCIA:
 - (A) COMPRIMENTO: 24 aminoácidos
 - (B) TIPO: aminoácido
 - (D) TOPOLOGIA: linear
- (ii) TIPO DE MOLÉCULA: péptido
- (iii) HIPOTÉTICA: NÃO
- (iv) ANTI-SENTIDO: NÃO
- (v) TIPO DE FRAGMENTO: interno

(xi) DESCRIÇÃO DA SEQUÊNCIA: SEQ ID NO:91:

Glu Glu Lys Met Gly Thr Ala Lys Lys Cys Cys Ala Ile Gly Cys Ser
1 5 10 15

Thr Glu Asp Phe Arg Met Val Cys
20

(2) INFORMAÇÃO PARA SEQ ID NO:92:

(i) CARACTERÍSTICAS DA SEQUÊNCIA:

(A) COMPRIMENTO: 40 aminoácidos

(B) TIPO: aminoácido

(D) TOPOLOGIA: linear

(ii) TIPO DE MOLÉCULA: péptido

(iii) HIPOTÉTICA: NÃO

(iv) ANTI-SENTIDO: NÃO

(v) TIPO DE FRAGMENTO: interno

(xi) DESCRIÇÃO DA SEQUÊNCIA: SEQ ID NO:92:

Arg Pro Asn Trp Glu Glu Arg Ser Arg Leu Cys Gly Arg Asp Leu Ile
1 5 10 15

Arg Ala Phe Ile Tyr Leu Cys Gly Gly Thr Arg Trp Thr Arg Leu Pro
20 25 30

Asn Phe Gly Asn Tyr Pro Ile Met
35 40

(2) INFORMAÇÃO PARA SEQ ID NO:93:

(i) CARACTERÍSTICAS DA SEQUÊNCIA:

(A) COMPRIMENTO: 182 aminoácidos

(B) TIPO: aminoácido

(D) TOPOLOGIA: linear

(ii) TIPO DE MOLÉCULA: péptido

(iii) HIPOTÉTICA: NÃO

(iv) ANTI-SENTIDO: NÃO

(v) TIPO DE FRAGMENTO: interno

(xi) DESCRIÇÃO DA SEQUÊNCIA: SEQ ID NO:93:

Ser Gly Ile Val Pro Thr Leu Gln Asn Ile Val Ser Thr Val Asn Leu
 1 5 10 15
 Asp Cys Lys Leu Asp Leu Lys Ala Ile Ala Leu Gln Ala Arg Asn Ala
 20 25 30
 Glu Tyr Asn Pro Lys Arg Phe Ala Ala Val Ile Met Arg Ile Arg Glu
 35 40 45
 Pro Lys Thr Thr Ala Leu Ile Phe Ala Ser Gly Lys Met Val Cys Thr
 50 55 60
 Gly Ala Lys Ser Glu Asp Phe Ser Lys Met Ala Ala Arg Lys Tyr Ala
 65 70 75 80
 Arg Ile Val Gln Lys Leu Gly Phe Pro Ala Lys Phe Lys Asp Phe Lys
 85 90 95
 Ile Gln Asn Ile Val Gly Ser Cys Asp Val Lys Phe Pro Ile Arg Leu
 100 105 110
 Glu Gly Leu Ala Tyr Ser His Ala Ala Phe Ser Ser Tyr Glu Pro Glu
 115 120 125
 Leu Phe Pro Gly Leu Ile Tyr Arg Met Lys Val Pro Lys Ile Val Leu
 130 135 140
 Leu Ile Phe Val Ser Gly Lys Ile Val Ile Thr Gly Ala Lys Met Arg
 145 150 155 160
 Asp Glu Thr Tyr Lys Ala Phe Glu Asn Ile Tyr Pro Val Leu ser Glu
 165 170 175
 Phe Arg Lys Ile Gln Gln
 180

(2) INFORMAÇÃO PARA SEQ ID NO:94:

(i) CARACTERÍSTICAS DA SEQUÊNCIA:

- (A) COMPRIMENTO: 84 aminoácidos
- (B) TIPO: aminoácido
- (D) TOPOLOGIA: linear

(ii) TIPO DE MOLÉCULA: péptido

(iii) HIPOTÉTICA: NÃO

(iv) ANTI-SENTIDO: NÃO

(v) TIPO DE FRAGMENTO: interno

(xi) DESCRIÇÃO DA SEQUÊNCIA: SEQ ID NO:94:

Asn Ser Asn Ser Thr Pro Ile Val His Leu Lys Gly Asp Ala Asn Thr
1 5 10 15

Leu Lys Cys Leu Arg Tyr Arg Phe Lys Lys His Cys Thr Leu Tyr Thr
20 25 30

Ala Val Ser Ser Thr Trp His Trp Thr Gly His Asn Val Lys His Lys
35 40 45

Ser Ala Ile Val Thr Leu Thr Tyr Asp Ser Glu Trp Gln Arg Asp Gln
50 55 60

Phe Leu Ser Gln Val Lys Ile Pro Lys Thr Ile Thr Val Ser Thr Gly
65 70 75 80

Phe Met Ser Ile

(2) INFORMAÇÃO PARA SEQ ID NO:95:

(i) CARACTERÍSTICAS DA SEQUÊNCIA:

- (A) COMPRIMENTO: 84 aminoácidos
- (B) TIPO: aminoácido
- (D) TOPOLOGIA: linear

(ii) TIPO DE MOLÉCULA: péptido

(iii) HIPOTÉTICA: NÃO

(iv) ANTI-SENTIDO: NÃO

(v) TIPO DE FRAGMENTO: interno

(xi) DESCRIÇÃO DA SEQUÊNCIA: SEQ ID NO:95:

Asn Ser Asn Thr Thr Pro Ile Val His Leu Lys Gly Asp Ala Asn Thr
1 5 10 15

Leu Lys Cys Leu Arg Tyr Arg Phe Lys Lys His Cys Thr Leu Tyr Thr
20 25 30

Ala Val Ser Ser Thr Trp His Trp Thr Gly His Asn Val Lys His Lys
35 40 45

Ser Ala Ile Val Thr Leu Thr Tyr Asp Ser Glu Trp Gln Arg Asp Gln
50 55 60

Phe Leu Ser Gln Val Lys Ile Pro Lys Thr Ile Thr Val Ser Thr Gly
65 70 75 80

Phe Met Ser Ile

(2) INFORMAÇÃO PARA SEQ ID NO:96:

(i) CARACTERÍSTICAS DA SEQUÊNCIA:

- (A) COMPRIMENTO: 83 aminoácidos
- (B) TIPO: aminoácido
- (D) TOPOLOGIA: linear

(ii) TIPO DE MOLÉCULA: péptido

(iii) HIPOTÉTICA: NÃO

(iv) ANTI-SENTIDO: NÃO

(v) TIPO DE FRAGMENTO: interno

(xi) DESCRIÇÃO DA SEQUÊNCIA: SEQ ID NO:96:

Ser Gly Asn Thr Thr Pro Ile Ile His Leu Lys Gly Asp Arg Asn Ser
 1 5 10 15
 Leu Lys Cys Leu Arg Tyr Arg Leu Arg Lys His Ser Asp His Tyr Arg
 20 25 30
 Asp Ile Ser Ser Thr Trp His Trp Thr Gly Ala Gly Asn Glu Lys Thr
 35 40 45
 Gly Ile Leu Thr Val Thr Tyr His Ser Glu Thr Gln Arg Thr Lys Phe
 50 55 60
 Leu Asn Thr Val Ala Ile Pro Asp Ser Val Gln Ile Leu Val Gly Tyr
 65 70 75 80
 Met Thr Met

(2) INFORMAÇÃO PARA SEQ ID NO:97:

(i) CARACTERÍSTICAS DA SEQUÊNCIA:

(A) COMPRIMENTO: 84 aminoácidos

(B) TIPO: aminoácido

(D) TOPOLOGIA: linear

(ii) TIPO DE MOLÉCULA: péptido

(iii) HIPOTÉTICA: NÃO

(iv) ANTI-SENTIDO: NÃO

(v) TIPO DE FRAGMENTO: interno

(xi) DESCRIÇÃO DA SEQUÊNCIA: SEQ ID NO:97:

Ser Gly Asn Thr Ala Pro Ile Val His Leu Lys Gly Glu Ser Asn Ser
 1 5 10 15
 Leu Lys Cys Leu Arg Tyr Arg Leu Lys Pro Tyr Lys Glu Leu Tyr Ser
 20 25 30
 Ser Met Ser Ser Thr Trp His Trp Thr Ser Asp Asn Lys Asn Ser Lys
 35 40 45
 Asn Gly Ile Val Thr Val Thr Phe Val Thr Glu Gln Gln Gln Met
 50 55 60
 Phe Leu Gly Thr Val Lys Ile Pro Pro Thr Val Gln Ile Ser Thr Gly
 65 70 75 80
 Phe Met Thr Leu

(2) INFORMAÇÃO PARA SEQ ID NO:98:

(i) CARACTERÍSTICAS DA SEQUÊNCIA:

(A) COMPRIMENTO: 89 aminoácidos

(B) TIPO: aminoácido

(D) TOPOLOGIA: linear

(ii) TIPO DE MOLÉCULA: péptido

(iii) HIPOTÉTICA: NÃO

(iv) ANTI-SENTIDO: NÃO

(v) TIPO DE FRAGMENTO: interno

(xi) DESCRIÇÃO DA SEQUÊNCIA: SEQ ID NO:98:

Ser Gly Asn Thr Ser Cys Phe Ala Leu Ile Ser Gly Thr Ala Asn Gln
 1 5 10 15

Val Lys Cys Tyr Arg Phe Arg Val Lys Lys Asn His Arg His Arg Tyr
 20 25 30

Glu Asn Cys Thr Thr Thr Trp Phe Thr Val Ala Asp Asn Gly Ala Glu
 35 40 45

Arg Gln Gly Gln Ala Gln Ile Leu Ile Thr Phe Gly Ser Pro Ser Gln
 50 55 60

Arg Gln Asp Phe Leu Lys His Val Pro Leu Pro Pro Gly Met Asn Ile
 65 70 75 80

Ser Gly Phe Thr Ala Ser Leu Asp Phe
 85

(2) INFORMAÇÃO PARA SEQ ID NO:99:

(i) CARACTERÍSTICAS DA SEQUÊNCIA:

(A) COMPRIMENTO: 7 aminoácidos

(B) TIPO: aminoácido

(D) TOPOLOGIA: linear

(ii) TIPO DE MOLÉCULA: péptido

(iii) HIPOTÉTICA: NÃO

(iv) ANTI-SENTIDO: NÃO

(v) TIPO DE FRAGMENTO: C-terminal

(xi) DESCRIÇÃO DA SEQUÊNCIA: SEQ ID NO:99:

Ser Asn Lys Lys Thr Thr Ala
 1 5

(2) INFORMAÇÃO PARA SEQ ID NO:100:

(i) CARACTERÍSTICAS DA SEQUÊNCIA:

(A) COMPRIMENTO: 4 aminoácidos

(B) TIPO: aminoácido

(D) TOPOLOGIA: linear

(ii) TIPO DE MOLÉCULA: péptido

(iii) HIPOTÉTICA: NÃO

(iv) ANTI-SENTIDO: NÃO

(v) TIPO DE FRAGMENTO: interno

(xi) DESCRIÇÃO DA SEQUÊNCIA: SEQ ID NO:100:

Asn Ser Asn Thr**1**

(2) INFORMAÇÃO PARA SEQ ID NO:101:

(i) CARACTERÍSTICAS DA SEQUÊNCIA:

(A) COMPRIMENTO: 4 aminoácidos

(B) TIPO: aminoácido

(D) TOPOLOGIA: linear

(ii) TIPO DE MOLÉCULA: péptido

(iii) HIPOTÉTICA: NÃO

(iv) ANTI-SENTIDO: NÃO

(v) TIPO DE FRAGMENTO: interno

(xi) DESCRIÇÃO DA SEQUÊNCIA: SEQ ID NO:101:

Ser Gly Asn Thr**1**

(2) INFORMAÇÃO PARA SEQ ID NO:102:

(i) CARACTERÍSTICAS DA SEQUÊNCIA:

(A) COMPRIMENTO: 6 aminoácidos

(B) TIPO: aminoácido

(D) TOPOLOGIA: linear

(ii) TIPO DE MOLÉCULA: péptido

(iii) HIPOTÉTICA: NÃO

(iv) ANTI-SENTIDO: NÃO

(v) TIPO DE FRAGMENTO: interno

(xi) DESCRIÇÃO DA SEQUÊNCIA: SEQ ID NO:102:

Ser Ser Gly Ser Ser Gly**1****5**

(iii) HIPOTÉTICA: NÃO

(iv) ANTI-SENTIDO: NÃO

(v) TIPO DE FRAGMENTO: N-terminal

(x) DESCRIÇÃO DA SEQUÊNCIA: SEQ ID NO:105:

Met Glu Gln Glu Ile Thr Leu Lys Asp Tyr Ala Met Arg Phe Gly Gln
1 5 10 15

Thr Lys Thr Ala Lys Asp Leu Gly Val Tyr Gln Ser Ala Ile Asn Lys
20 25 30

Ala Ile His Ala Gly Arg Lys Ile Phe Leu Thr Ile Asn Ala Asp Gly
35 40 45

Ser Val Tyr Ala Glu Glu Val Lys Pro Phe Pro Ser Asn Lys Lys Thr
50 55 60

Thr Ala
65

(2) INFORMAÇÃO PARA SEQ ID NO:106:

(i) CARACTERÍSTICAS DA SEQUÊNCIA:

(A) COMPRIMENTO: 66 aminoácidos

(B) TIPO: aminoácido

(D) TOPOLOGIA: linear

(ii) TIPO DE MOLÉCULA: proteína

(iii) HIPOTÉTICA: NÃO

(iv) ANTI-SENTIDO: NÃO

(v) TIPO DE FRAGMENTO: N-terminal

(xi) DESCRIÇÃO DA SEQUÊNCIA: SEQ ID NO:106:

Met Arg Gln Arg Ile Thr Leu Lys Asp Tyr Ala Met Arg Phe Gly Gln
1 5 10 15

Thr Lys Thr Ala Lys Asp Leu Gly Val Tyr Gln Ser Ala Ile Asn Lys
20 25 30

Ala Ile His Ala Gly Arg Lys Ile Phe Leu Thr Ile Asn Ala Asp Gly
35 40 45

Ser Val Tyr Ala Glu Glu Val Lys Pro Phe Pro Ser Asn Lys Lys Thr
50 55 60

Thr Ala
65

(2) INFORMAÇÃO PARA SEQ ID NO:107:

(i) CARACTERÍSTICAS DA SEQUÊNCIA:

(A) COMPRIMENTO: 96 aminoácidos

(B) TIPO: aminoácido

(D) TOPOLOGIA: linear

(ii) TIPO DE MOLÉCULA: péptido

(iii) HIPOTÉTICA: NÃO

(iv) ANTI-SENTIDO: NÃO

(v) TIPO DE FRAGMENTO: N-terminal

(xi) DESCRIÇÃO DA SEQUÊNCIA: SEQ ID NO:107:

```

Ser Thr Lys Lys Lys Pro Leu Thr Gln Glu Gln Leu Glu Asp Ala Arg
1           5           10           15
Arg Leu Lys Ala Ile Tyr Glu Lys Lys Lys Asn Glu Leu Gly Leu Ser
          20           25           30
Gln Glu Ser Val Ala Asp Lys Met Gly Met Gly Gln Ser Gly Val Gly
          35           40           45
Ala Leu Phe Asn Gly Ile Asn Ala Leu Asn Ala Tyr Asn Ala Ala Leu
          50           55           60
Leu Ala Lys Ile Leu Lys Val Ser Val Glu Glu Phe Ser Pro Ser Ile
65           70           75           80
Ala Arg Glu Ile Tyr Glu Met Tyr Glu Ala Val Ser Met Glu Pro ser
          85           90           95

```

(2) INFORMAÇÃO PARA SEQ ID NO:108:

(i) CARACTERÍSTICAS DA SEQUÊNCIA:

(A) COMPRIMENTO: 96 aminoácidos

(B) TIPO: aminoácido

(D) TOPOLOGIA: linear

(ii) TIPO DE MOLÉCULA: péptido

(iii) HIPOTÉTICA: NÃO

(iv) ANTI-SENTIDO: NÃO

(v) TIPO DE FRAGMENTO: N-terminal

(xi) DESCRIÇÃO DA SEQUÊNCIA: SEQ ID NO:108:

```

Ser Thr Lys Lys Lys Pro Leu Thr Gln Glu Gln Leu Glu Asp Ala Arg
1           5           10           15
Arg Leu Lys Ala Ile Tyr Glu Lys Lys Lys Asn Glu Leu Gly Leu Ser
          20           25           30
Gln Glu Ser Val Ala Asp Lys Met Gly Met Gly Gln Ser Gly Val Gly
          35           40           45
Ala Leu Phe Asn Gly Ile Asn Ala Leu Asn Ala Tyr Asn Ala Ala Leu
          50           55           60
Leu Ala Lys Ile Leu Lys Val Ser Val Glu Glu Phe Ser Pro Ser Ile
65           70           75           80
Ala Arg Glu Ile Tyr Glu Met Cys Glu Ala Val Ser Met Glu Pro Ser
          85           90           95

```


(2) INFORMAÇÃO PARA SEQ ID NO:109:

(i) CARACTERÍSTICAS DA SEQUÊNCIA:

- (A) COMPRIMENTO: 180 aminoácidos
- (B) TIPO: aminoácido
- (D) TOPOLOGIA: linear

(ii) TIPO DE MOLÉCULA: proteína

(iii) HIPOTÉTICA: NÃO

(iv) ANTI-SENTIDO: NÃO

(v) TIPO DE FRAGMENTO: N-terminal

(xi) DESCRIÇÃO DA SEQUÊNCIA: SEQ ID NO:109:

```

Gly Ile Val Glu Gln Cys Cys Thr ser Ile Cys Ser Leu Tyr Gln Leu
1           5           10           15
Glu Asn Tyr Cys Asn Met Ser Met Glu Gln Arg Ile Thr Leu Lys Asp
20          25          30
Tyr Ala Met Arg Phe Gly Gln Thr Lys Thr Ala Lys Asp Leu Gly Val
35          40          45
Tyr Gln Ser Ala Ile Asn Lys Ala Ile His Ala Gly Arg Lys Ile Phe
50          55          60
Leu Thr Ile Asn Ala Asp Gly Ser Val Tyr Ala Glu Glu Val Lys Pro
65          70          75          80
Phe Pro Ser Asn Lys Lys Thr Thr Ala Ser Asn Lys Lys Thr Thr Ala
85          90          95
Asn Ser Asn Thr Thr Pro Ile Val His Leu Lys Gly Asp Ala Asn Thr
100         105         110
Leu Lys Cys Leu Arg Tyr Arg Phe Lys Lys His Cys Thr Leu Tyr Thr
115         120         125
Ala Val Ser Ser Thr Trp His Trp Thr Gly His Asn Val Lys His Lys
130         135         140
Ser Ala Ile Val Thr Leu Thr Tyr Asp Ser Glu Trp Gln Arg Asp Gln
145         150         155         160
Phe Leu Ser Gln Val Lys Ile Pro Lys Thr Ile Thr Val ser Thr Gly
165         170         175
Phe Met Ser Ile
180

```

(2) INFORMAÇÃO PARA SEQ ID NO:110:

(i) CARACTERÍSTICAS DA SEQUÊNCIA:

- (A) COMPRIMENTO: 113 aminoácidos
- (B) TIPO: aminoácido
- (D) TOPOLOGIA: linear

(ii) TIPO DE MOLÉCULA: proteína

(iii) HIPOTÉTICA: NÃO

(iv) ANTI-SENTIDO: NÃO

(v) TIPO DE FRAGMENTO: N-terminal

(xi) DESCRIÇÃO DA SEQUÊNCIA: SEQ ID NO:110:

Gly	Ile	Val	Glu	Gln	Cys	Cys	Thr	Ser	Ile	Cys	Ser	Leu	Tyr	Gln	Leu
1			5						10					15	
Glu	Asn	Tyr	Cys	Asn	Met	Ser	Met	Glu	Gln	Arg	Ile	Thr	Leu	Lys	Asp
	20							25					30		
Tyr	Ala	Met	Arg	Phe	Gly	Gln	Thr	Lys	Thr	Ala	Lys	Asp	Leu	Gly	Val
	35						40					45			
Tyr	Gln	Ser	Ala	Ile	Asn	Lys	Ala	Ile	His	Ala	Gly	Arg	Lys	Ile	Phe
	50					55					60				
Leu	Thr	Ile	Asn	Ala	Asp	Gly	Ser	Val	Tyr	Ala	Glu	Glu	Val	Lys	Pro
65					70					75					80
Phe	Pro	Ser	Asn	Lys	Lys	Thr	Thr	Ala	Ser	Asn	Lys	Lys	Thr	Thr	Ala
				85					90						95
Cys	Asp	Thr	Asp	Asp	Arg	His	Arg	Ile	Glu	Glu	Lys	Arg	Lys	Arg	Lys
			100					105						110	

Thr

(2) INFORMAÇÃO PARA SEQ ID NO:111:

(i) CARACTERÍSTICAS DA SEQUÊNCIA:

(A) COMPRIMENTO: 292 aminoácidos

(B) TIPO: aminoácido

(D) TOPOLOGIA: linear

(ii) TIPO DE MOLÉCULA: proteína

(iii) HIPOTÉTICA: NÃO

(iv) ANTI-SENTIDO: NÃO

(v) TIPO DE FRAGMENTO: N-terminal

(xi) DESCRIÇÃO DA SEQUÊNCIA: SEQ ID NO:111:

Phe Val Asn Gln His Leu Cys Gly Ser His Leu Val Glu Ala Leu Tyr
 1 5 10 15
 Leu Val Cys Gly Glu Arg Gly Phe Phe Tyr Thr Pro Lys Thr Met Ser
 20 25 30
 Met Glu Gln Glu Ile Thr Leu Lys Asp Tyr Ala Met Arg Phe Gly Gln
 35 40 45
 Thr Lys Thr Ala Lys Asp Leu Gly Val Tyr Gln Ser Ala Ile Asn Lys
 50 55 60
 Ala Ile His Ala Gly Arg Lys Ile Phe Leu Thr Ile Asn Ala Asp Gly
 65 70 75 80
 Ser Val Tyr Ala Glu Glu Val Lys Pro Phe Pro Ser Asn Lys Lys Thr
 85 90 95
 Thr Ala Ser Asn Lys Lys Thr Thr Ala Ser Ser Gly Ser Ser Gly Ser
 100 105 110
 Gly Ile Val Pro Gln Leu Gln Asn Ile Val Ser Thr Val Asn Leu Gly
 115 120 125
 Cys Lys Leu Asp Leu Lys Thr Ile Ala Leu Arg Ala Arg Asn Ala Glu
 130 135 140
 Tyr Asn Pro Lys Arg Phe Ala Ala Val Ile Met Arg Ile Arg Glu Pro
 145 150 155 160
 Arg Thr Thr Ala Leu Ile Phe Ser Ser Gly Lys Met Val Cys Thr Gly
 165 170 175
 Ala Lys Ser Glu Glu Gln Ser Arg Leu Ala Ala Arg Lys Tyr Ala Arg
 180 185 190
 Val Val Gln Lys Leu Gly Phe Pro Ala Lys Phe Leu Asp Phe Lys Ile
 195 200 205
 Gln Asn Met Val Gly Ser Cys Asp Val Lys Phe Pro Ile Arg Leu Glu
 210 215 220
 Gly Leu Val Leu Thr His Gln Gln Phe Ser Ser Tyr Glu Pro Glu Leu
 225 230 235 240
 Phe Pro Gly Leu Ile Tyr Arg Met Ile Lys Pro Arg Ile Val Leu Leu
 245 250 255
 Ile Phe Val Ser Gly Lys Val Val Leu Thr Gly Ala Lys Val Arg Ala
 260 265 270
 Glu Ile Tyr Glu Ala Phe Glu Asn Ile Tyr Pro Ile Leu Lys Gly Phe
 275 280 285
 Arg Lys Thr Thr
 290

(2) INFORMAÇÃO PARA SEQ ID NO:112:

(i) CARACTERÍSTICAS DA SEQUÊNCIA:

- (A) COMPRIMENTO: 273 aminoácidos
- (B) TIPO: aminoácido
- (D) TOPOLOGIA: linear

(ii) TIPO DE MOLÉCULA: proteína

(iii) HIPOTÉTICA: NÃO

(iv) ANTI-SENTIDO: NÃO

(v) TIPO DE FRAGMENTO: N-terminal

(xi) DESCRIÇÃO DA SEQUÊNCIA: SEQ ID NO:112:

```

Phe Val Asn Gln His Leu Cys Gly Ser His Leu Val Glu Ala Leu Tyr
1          5          10          15
Leu Val Cys Gly Glu Arg Gly Phe Phe Tyr Thr Pro Lys Thr Met Ser
20          25          30
Met Arg Gln Arg Ile Thr Leu Lys Asp Tyr Ala Met Arg Phe Gly Gln
35          40          45
Thr Lys Thr Ala Lys Asp Leu Gly Val Tyr Gln Ser Ala Ile Asn Lys
50          55          60
Ala Ile His Ala Gly Arg Lys Ile Phe Leu Thr Ile Asn Ala Asp Gly
65          70          75          80
Ser Val Tyr Ala Glu Glu Val Lys Pro Phe Pro Ser Asn Lys Lys Thr
85          90          95
Thr Ala Ser Asn Lys Lys Thr Thr Ala Gly Asp Pro Gly Lys Lys Lys
100         105         110
Gln His Ile Cys His Ile Gln Gly Cys Gly Lys Val Tyr Gly Lys Thr
115         120         125
Ser His Leu Arg Ala His Leu Arg Trp His Thr Gly Glu Arg Pro Phe
130         135         140
Met Cys Thr Trp Ser Tyr Cys Gly Lys Arg Phe Thr Arg Ser Asp Glu
145         150         155         160
Leu Gln Arg His Lys Arg Thr His Thr Gly Glu Lys Lys Phe Ala Cys
165         170         175
Pro Glu Cys Pro Lys Arg Phe Met Arg Ser Asp His Leu Ser Lys His
180         185         190

```

Ile Lys Thr His Gln Asn Lys Lys Gly Gly Pro Gly Val Ala Leu Ser
 195 200 205

Val Gly Thr Leu Pro Leu Asp Ser Gly Ala Gly Ser Glu Gly Ser Gly
 210 215 220

Thr Ala Thr Pro Ser Ala Leu Ile Thr Thr Asn Met Val Ala Met Glu
 225 230 235 240

Ala Ile Cys Pro Glu Gly Ile Ala Arg Leu Ala Asn Ser Gly Ile Asn
 245 250 255

Val Met Gln Val Ala Asp Leu Gln Ser Ile Asn Ile Ser Gly Asn Gly
 260 265 270

Phe

(2) INFORMAÇÃO PARA SEQ ID NO:113:

(i) CARACTERÍSTICAS DA SEQUÊNCIA:

- (A) COMPRIMENTO: 421 aminoácidos
- (B) TIPO: aminoácido
- (D) TOPOLOGIA: linear

(ii) TIPO DE MOLÉCULA: proteína

(iii) HIPOTÉTICA: NÃO

(iv) ANTI-SENTIDO: NÃO

(v) TIPO DE FRAGMENTO: N-terminal

(xi) DESCRIÇÃO DA SEQUÊNCIA: SEQ ID NO:113:

Gln Leu Tyr Ser Ala Leu Ala Asn Lys Cys Cys His Val Gly Cys Ile
 1 5 10 15

Lys Arg Ser Leu Ala Arg Phe Cys Met Ser Met Arg Gln Arg Ile Thr
 20 25 30

Leu Lys Asp Tyr Ala Met Arg Phe Gly Gln Thr Lys Thr Ala Lys Asp
 35 40 45

Leu Gly Val Tyr Gln Ser Ala Ile Asn Lys Ala Ile His Ala Gly Arg
 50 55 60

Lys Ile Phe Leu Thr Ile Asn Ala Asp Gly Ser Val Tyr Ala Glu Glu
 65 70 75 80

Val Lys Pro Phe Pro Ser Asn Lys Lys Thr Thr Ala Ser Asn Lys Lys
 85 90 95

Thr Thr Ala Met Ala Asp Asp Asp Pro Tyr Gly Thr Gly Gln Met Phe
 100 105 110

His Leu Asn Thr Ala Leu Thr His Ser Ile Phe Asn Ala Glu Leu Tyr
 115 120 125
 Ser Pro Glu Ile Pro Leu Ser Thr Asp Gly Pro Tyr Leu Gln Ile Leu
 130 135 140
 Glu Gln Pro Lys Gln Arg Gly Phe Arg Phe Arg Tyr Val Cys Glu Gly
 145 150 155 160
 Pro Ser His Gly Gly Leu Pro Gly Ala Ser Ser Glu Lys Asn Lys Lys
 165 170 175
 Ser Tyr Pro Gln Val Lys Ile Cys Asn Tyr Val Gly Pro Ala Lys Val
 180 185 190
 Ile Val Gln Leu Val Thr Asn Gly Lys Asn Ile His Leu His Ala His
 195 200 205
 Ser Leu Val Gly Lys His Cys Glu Asp Gly Val Cys Thr Val Thr Ala
 210 215 220
 Gly Pro Lys Asp Met Val Val Gly Phe Ala Asn Leu Gly Ile Leu His
 225 230 235 240
 Val Thr Lys Lys Lys Val Phe Glu Thr Leu Glu Ala Arg Met Thr Glu
 245 250 255
 Ala Cys Ile Arg Gly Tyr Asn Pro Gly Leu Leu Val His Ser Asp Leu
 260 265 270
 Ala Tyr Leu Gln Ala Glu Gly Gly Gly Asp Arg Gln Leu Thr Asp Arg
 275 280 285
 Glu Lys Glu Ile Ile Arg Gln Ala Ala Val Gln Gln Thr Lys Glu Met
 290 295 300
 Asp Leu Ser Val Val Arg Leu Met Phe Thr Ala Phe Leu Pro Asp Ser
 305 310 315
 Thr Gly Ser Phe Thr Arg Arg Leu Glu Pro Val Val Ser Asp Ala Ile
 325 330 335
 Tyr Asp Ser Lys Ala Pro Asn Ala Ser Asn Leu Lys Ile Val Arg Met
 340 345 350
 Asp Arg Thr Ala Gly Cys Val Thr Gly Gly Glu Glu Ile Tyr Leu Leu
 355 360 365
 Cys Asp Lys Val Gln Lys Asp Asp Ile Gln Ile Arg Phe Tyr Glu Glu
 370 375 380
 Glu Glu Asn Gly Gly Val Trp Glu Gly Phe Gly Asp Phe Ser Pro Thr
 385 390 395 400
 Asp Val His Arg Gln Phe Ala Ile Val Phe Lys Thr Pro Lys Tyr Lys
 405 410 415
 Asp Val Asn Ile Thr
 420

(2) INFORMAÇÃO PARA SEQ ID NO:114:

(i) CARACTERÍSTICAS DA SEQUÊNCIA:

- (A) COMPRIMENTO: 391 aminoácidos
- (B) TIPO: aminoácido
- (D) TOPOLOGIA: linear

(ii) TIPO DE MOLÉCULA: proteína

(iii) HIPOTÉTICA: NÃO

(iv) ANTI-SENTIDO: NÃO

(v) TIPO DE FRAGMENTO: N-terminal

(xi) DESCRIÇÃO DA SEQUÊNCIA: SEQ ID NO:114:

Met	Arg	Gln	Arg	Ile	Thr	Leu	Lys	Asp	Tyr	Ala	Met	Arg	Phe	Gly	Gln	1	5	10	15
Thr	Lys	Thr	Ala	Lys	Asp	Leu	Gly	Val	Tyr	Gln	Ser	Ala	Ile	Asn	Lys	20	25	30	
Ala	Ile	His	Ala	Gly	Arg	Lys	Ile	Phe	Leu	Thr	Ile	Asn	Ala	Asp	Gly	35	40	45	
Ser	Val	Tyr	Ala	Glu	Glu	Val	Lys	Pro	Phe	Pro	Ser	Asn	Lys	Lys	Thr	50	55	60	
Thr	Ala	Met	Ala	Glu	Asp	Asp	Pro	Tyr	Leu	Gly	Arg	Pro	Glu	Gln	Met	65	70	75	80
Phe	His	Leu	Asp	Pro	Ser	Leu	Thr	His	Thr	Ile	Phe	Asn	Pro	Glu	Val	85	90	95	
Phe	Gln	Pro	Gln	Met	Ala	Leu	Pro	Thr	Ala	Asp	Gly	Pro	Tyr	Leu	Gln	100	105	110	
Ile	Leu	Glu	Gln	Pro	Lys	Gln	Arg	Gly	Phe	Arg	Phe	Arg	Tyr	Val	Cys	115	120	125	
Glu	Gly	Pro	Ser	His	Gly	Gly	Leu	Pro	Gly	Ala	Ser	Ser	Glu	Lys	Asn	130	135	140	
Lys	Lys	Ser	Tyr	Pro	Gln	Val	Lys	Ile	Cys	Asn	Tyr	Val	Gly	Pro	Ala	145	150	155	160
Lys	Val	Ile	Val	Gln	Leu	Val	Thr	Asn	Gly	Lys	Asn	Ile	His	Leu	His	165	170	175	
Ala	His	Ser	Leu	Val	Gly	Lys	His	Cys	Glu	Asp	Gly	Ile	Cys	Thr	Val	180	185	190	

Thr Ala Gly Pro Glu Asp Cys Val His Gly Phe Ala Asn Leu Gly Ile
 195 200 205
 Leu His Val Thr Lys Lys Lys Val Phe Glu Thr Leu Glu Ala Arg Met
 210 215 220
 Thr Glu Ala Cys Ile Arg Gly Tyr Asn Pro Gly Leu Leu Val His Pro
 225 230 235 240
 Asp Leu Ala Tyr Leu Gln Ala Glu Gly Gly Gly Asp Arg Gln Leu Gly
 245 250 255
 Asp Arg Glu Lys Glu Leu Ile Arg Gln Ala Ala Leu Gln Gln Thr Lys
 260 265 270
 Glu Met Asp Leu Ser Val Val Arg Leu Met Phe Thr Ala Phe Leu Pro
 275 280 285
 Asp Ser Thr Gly Ser Phe Thr Arg Arg Leu Glu Pro Val Val Ser Asp
 290 295 300
 Ala Ile Tyr Asp Ser Lys Ala Pro Asn Ala Ser Asn Leu Lys Ile Val
 305 310 315 320
 Arg Met Asp Arg Thr Ala Gly Cys Val Thr Gly Gly Glu Glu Ile Tyr
 325 330 335
 Leu Leu Cys Asp Lys Val Gln Lys Asp Asp Ile Gln Ile Arg Phe Tyr
 340 345 350
 Glu Glu Glu Glu Asn Gly Gly Val Trp Glu Gly Phe Gly Asp Phe Ser
 355 360 365
 Pro Thr Asp Val His Arg Gln Phe Ala Ile Val Phe Lys Thr Pro Lys
 370 375 380
 Tyr Lys Asp Ile Asn Ile Thr
 385 390

(2) INFORMAÇÃO PARA SEQ ID NO:115:

- (i) CARACTERÍSTICAS DA SEQUÊNCIA:
 - (A) COMPRIMENTO: 391 aminoácidos
 - (B) TIPO: aminoácido
 - (D) TOPOLOGIA: linear
- (ii) TIPO DE MOLÉCULA: proteína
- (iii) HIPOTÉTICA: NÃO
- (iv) ANTI-SENTIDO: NÃO
- (v) TIPO DE FRAGMENTO: N-terminal

(xi) DESCRIÇÃO DA SEQUÊNCIA: SEQ ID NO:115:

Met Glu Gln Glu Ile Thr Leu Lys Asp Tyr Ala Met Arg Phe Gly Gln
 1 5 10 15
 Thr Lys Thr Ala Lys Asp Leu Gly Val Tyr Gln Ser Ala Ile Asn Lys
 20 25 30
 Ala Ile His Ala Gly Arg Lys Ile Phe Leu Thr Ile Asn Ala Asp Gly
 35 40 45
 Ser Val Tyr Ala Glu Glu Val Lys Pro Phe Pro Ser Asn Lys Lys Thr
 50 55 60
 Thr Ala Met Ala Glu Asp Asp Pro Tyr Leu Gly Arg Pro Glu Gln Met
 65 70 75 80
 Phe His Leu Asp Pro Ser Leu Thr His Thr Ile Phe Asn Pro Glu Val
 85 90 95
 Phe Gln Pro Gln Met Ala Leu Pro Thr Ala Asp Gly Pro Tyr Leu Gln
 100 105 110
 Ile Leu Glu Gln Pro Lys Gln Arg Gly Phe Arg Phe Arg Tyr Val Cys
 115 120 125
 Glu Gly Pro Ser His Gly Gly Leu Pro Gly Ala Ser Ser Glu Lys Asn
 130 135 140
 Lys Lys Ser Tyr Pro Gln Val Lys Ile Cys Asn Tyr Val Gly Pro Ala
 145 150 155 160
 Lys Val Ile Val Gln Leu Val Thr Asn Gly Lys Asn Ile His Leu His
 165 170 175
 Ala His Ser Leu Val Gly Lys His Cys Glu Asp Gly Ile Cys Thr Val
 180 185 190
 Thr Ala Gly Pro Glu Asp Cys Val His Gly Phe Ala Asn Leu Gly Ile
 195 200 205
 Leu His Val Thr Lys Lys Lys Val Phe Glu Thr Leu Glu Ala Arg Met
 210 215 220
 Thr Glu Ala Cys Ile Arg Gly Tyr Asn Pro Gly Leu Leu Val His Pro
 225 230 235 240
 Asp Leu Ala Tyr Leu Gln Ala Glu Gly Gly Gly Asp Arg Gln Leu Gly
 245 250 255
 Asp Arg Glu Lys Glu Leu Ile Arg Gln Ala Ala Leu Gln Gln Thr Lys
 260 265 270
 Glu Met Asp Leu Ser Val Val Arg Leu Met Phe Thr Ala Phe Leu Pro
 275 280 285

Asp Ser Thr Gly Ser Phe Thr Arg Arg Leu Glu Pro Val Val Ser Asp
 290 295 300
 Ala Ile Tyr Asp Ser Lys Ala Pro Asn Ala Ser Asn Leu Lys Ile Val
 305 310 315 320
 Arg Met Asp Arg Thr Ala Gly Cys Val Thr Gly Gly Glu Glu Ile Tyr
 325 330 335
 Leu Leu Cys Asp Lys Val Gln Lys Asp Asp Ile Gln Ile Arg Phe Tyr
 340 345 350
 Glu Glu Glu Glu Asn Gly Gly Val Trp Glu Gly Phe Gly Asp Phe Ser
 355 360 365
 Pro Thr Asp Val His Arg Gln Phe Ala Ile Val Phe Lys Thr Pro Lys
 370 375 380
 Tyr Lys Asp Ile Asn Ile Thr
 385 390

(2) INFORMAÇÃO PARA SEQ ID NO:116:

(i) CARACTERÍSTICAS DA SEQUÊNCIA:

- (A) COMPRIMENTO: 241 aminoácidos
- (B) TIPO: aminoácido
- (D) TOPOLOGIA: linear

(ii) TIPO DE MOLÉCULA: proteína

(iii) HIPOTÉTICA: NÃO

(iv) ANTI-SENTIDO: NÃO

(v) TIPO DE FRAGMENTO: N-terminal

(xi) DESCRIÇÃO DA SEQUÊNCIA: SEQ ID NO:116:

Met Arg Gln Arg Ile Thr Leu Lys Asp Tyr Ala Met Arg Phe Gly Gln
 1 5 10 15
 Thr Lys Thr Ala Lys Asp Leu Gly Val Tyr Gln Ser Ala Ile Asn Lys
 20 25 30
 Ala Ile His Ala Gly Arg Lys Ile Phe Leu Thr Ile Asn Ala Asp Gly
 35 40 45
 Ser Val Tyr Ala Glu Glu Val Lys Pro Phe Pro Ser Asn Lys Lys Thr
 50 55 60
 Thr Ala Ser Asn Lys Lys Thr Thr Ala Gly Asp Pro Gly Lys Lys Lys
 65 70 75 80
 Gln His Ile Cys His Ile Gln Gly Cys Gly Lys Val Tyr Gly Lys Thr
 85 90 95

Ser His Leu Arg Ala His Leu Arg Trp His Thr Gly Glu Arg Pro Phe
 100 105 110
 Met Cys Thr Trp Ser Tyr Cys Gly Lys Arg Phe Thr Arg Ser Asp Glu
 115 120 125
 Leu Gln Arg His Lys Arg Thr His Thr Gly Glu Lys Lys Phe Ala Cys
 130 135 140
 Pro Glu Cys Pro Lys Arg Phe Met Arg Ser Asp His Leu Ser Lys His
 145 150 155 160
 Ile Lys Thr His Gln Asn Lys Lys Gly Gly Pro Gly Val Ala Leu Ser
 165 170 175
 Val Gly Thr Leu Pro Leu Asp Ser Gly Ala Gly Ser Glu Gly Ser Gly
 180 185 190
 Thr Ala Thr Pro Ser Ala Leu Ile Thr Thr Asn Met Val Ala Met Glu
 195 200 205
 Ala Ile Cys Pro Glu Gly Ile Ala Arg Leu Ala Asn Ser Gly Ile Asn
 210 215 220
 Val Met Gln Val Ala Asp Leu Gln Ser Ile Asn Ile Ser Gly Asn Gly
 225 230 235 240
 Phe

(2) INFORMAÇÃO PARA SEQ ID NO:117:

(i) CARACTERÍSTICAS DA SEQUÊNCIA:

- (A) COMPRIMENTO: 10 pares de bases
- (B) TIPO: ácido nucleico
- (C) TIPO DE CADEIA: ambos
- (D) TOPOLOGIA: linear

(ii) TIPO DE MOLÉCULA: ADNc

(iii) HIPOTÉTICA : NO

(iv) ANTI-SENTIDO: NÃO

(xi) DESCRIÇÃO DA SEQUÊNCIA: SEQ ID NO:117:

GGGAMTNYCC 10

(2) INFORMAÇÃO PARA SEQ ID NO:118:

(i) CARACTERÍSTICAS DA SEQUÊNCIA:

- (A) COMPRIMENTO: 72 aminoácidos
- (B) TIPO: aminoácido
- (D) TOPOLOGIA: linear

(ii) TIPO DE MOLÉCULA: proteína

(iii) HIPOTÉTICA: NÃO

(iv) ANTI-SENTIDO: NÃO

(v) TIPO DE FRAGMENTO: N-terminal

(xi) DESCRIÇÃO DA SEQUÊNCIA: SEQ ID NO:118:

Met	Glu	Pro	Val	Asp	Pro	Arg	Leu	Glu	Pro	Trp	Lys	His	Pro	Gly	Ser
1				5					10					15	
Gln	Pro	Lys	Thr	Ala	Cys	Thr	Asn	Cys	Tyr	Cys	Lys	Lys	Cys	Cys	Phe
			20					25					30		
His	Cys	Gln	Val	Cys	Phe	Ile	Thr	Lys	Ala	Leu	Gly	Ile	Ser	Tyr	Gly
		35					40					45			
Arg	Lys	Lys	Arg	Arg	Gln	Arg	Arg	Arg	Ala	His	Gln	Asn	Ser	Gln	Thr
	50					55					60				
His	Gln	Ala	Ser	Leu	Ser	Lys	Gln								
65					70										

Lisboa,

REIVINDICAÇÕES

1. Ácido nucleico sonda (PNA) compreendendo duas sequências diferentes que são:

- a) uma sequência de cadeia simples (1/2 TBR) que é capaz de formar, sob condições de hibridação, um híbrido (TBR) com um ácido nucleico alvo (TNA); e
- b) uma sequência de cadeia simples (1/2 BBR) que é capaz de formar, sob condições de hibridação, um híbrido (BBR) com uma sequência de cadeia simples presente num ácido nucleico de reforço (BNA);

em que a referida TBR é capaz de se ligar com elevada afinidade a uma substância (TBA) capaz de discriminar entre um híbrido emparelhado (TBR) e um híbrido possuindo nucleótidos desemparelhados, e em que a referida BBR é capaz de se ligar com elevada afinidade a uma substância (BBA) capaz de discriminar entre um híbrido emparelhado (BBR) e um híbrido possuindo nucleótidos desemparelhados.

2. PNA de acordo com a reivindicação 1 em que a TBR é constituída por um ou mais locais de reconhecimento para uma proteína de ligação a um ácido nucleico, uma proteína de ligação ao ADN, uma proteína de ligação a um híbrido ADN-ARN ou uma proteína de ligação a ARN.

3. PNA de acordo com a reivindicação 2 em que a TBR é um local de reconhecimento de uma proteína de ligação a um ácido nucleico presente no genoma de um patógeno ou é um local de ligação associado a uma condição patogénica numa genoma de vertebrado ou é um local de reconhecimento de uma proteína de ligação a um ácido nucleico presente no genoma de um organismo que contamina um processo de fermentação.

4. PNA de acordo com a reivindicação 2 em que a TBR é a LTR de HIV ou uma sua porção.

5. Método para detecção ou localização de uma sequência de TNA específica, compreendendo os passos de:

- a) hibridação do referido TNA com o PNA da reivindicação 1;

- b) hibridação do referido PNA com um BNA contendo uma 1/2 BBR cuja sequência é complementar a uma sequência de 1/2 BBR no PNA;
- c) adição dos produtos dos passos (a) e (b) contendo uma TBR e uma BBR, a uma superfície, a um líquido ou a outro meio contendo um TBA;
- d) adição de BBA à mistura no passo (c) em que os referidos BBA compreendem:
 - i) uma molécula ou uma porção de uma molécula que é capaz de se ligar selectivamente a uma BBR;
 - ii) um indicador detectável; e
- e) detecção do sinal produzido pelo indicador ligado ao BBA.

6. Método de acordo com a reivindicação 5 em que o referido indicador é uma proteína, incluindo enzimas capazes de catalisar reacções que conduzam à produção de produtos de reacção coloridos; um radionuclido; esferas coloridas.

7. Método *in vitro* de amplificação do sinal obtido através da ligação do PNA de acordo com a reivindicação 1 a um TNA, que compreende a ligação dos BNA ao híbrido PNA-TNA e a ligação dos BBA marcados aos BNA.

8. Método para detecção ou localização de sequências de ácido nucleico específicas com um elevado grau de sensibilidade e especificidade que compreende:

- a) adição de PNA tal como definidos na reivindicação 1, contendo uma 1/2 BBR e uma 1/2 TBR, a uma amostra contendo ou suspeita de conter TNA contendo sequências de 1/2 TBR, para formar um complexo possuindo regiões de ligação ao alvo, TBR, formadas pela hibridação de 1/2 TBR complementares presentes nos PNA e TNA, respectivamente;
- b) ligação das TBR formadas no passo (a) a um TBA imobilizado para formar um complexo TBA-TNA-PNA;
- c) adição de Ácidos Nucleicos de Reforço, BNA, contendo regiões de ligação de reforço, 1/2 BBR, ao complexo formado no passo (b) de modo a que as 1/2 BBR nos BNA hibridem com as sequências de 1/2 BBR presentes nos PNA ou com as 1/2 BBR presentes nos BNA já ligados ao PNA,

para formar BBR, de modo que se formem complexos TBA-TNA-PNA-(BNA)_n;

- d) adição de Ácidos Nucleicos em "gancho de cabelo", HNA, contendo sequências de 1/2 BBR, ao complexo formado no passo (c) de modo a que as 1/2 BBR nos HNA hibridem com quaisquer sequências de 1/2 BBR disponíveis presentes nos BNA do complexo do passo (c), rematando deste modo a extensão dos BNA nos complexos TBA-TNA-PNA-(BNA)_n do passo (c) para formar complexos TBA-TNA-PNA-(BNA)_n-HNA;
- e) adição de Conjuntos de Ligação de Reforço, BBA, ligados a porções indicadoras, aos complexos TBA-TNA-PNA-(BNA)_n-HNA formados no passo (d) para formar complexos TBA-TNA-PNA-(BNA-BBA)_n-HNA; e
- f) detecção dos sinais produzidos pelas porções indicadoras ligadas aos TBA, PNA, BNA, BBA ou HNA nos complexos TBA-TNA-PNA-(BNA-BBA)_n-HNA do passo (e);

em que o TNA compreende:

- i) uma ou mais sequências de ácido nucleico de 1/2 TBR específicas, cuja presença ou ausência numa determinada amostra se pretende confirmar;

o BNA compreende:

- i) uma 1/2 BBR, tal como mostrada na Figura 1 (IIb), que possui uma sequência que é complementar a uma sequência de 1/2 BBR num PNA e que é capaz de formar, sob condições de hibridação, um híbrido, BBR, com o PNA;
- ii) um OSA, que é sem suporte ou indicador ligado ou com suporte ligado, ou outro meio de localização, incluindo, mas não se limitando a, ligação a esferas, polímeros e superfícies e/ou indicadores;
- iii) locais de hibridação adicionais, 1/2 BBR, para outros BNA; e
- iv) sequências, 1/2 BBR, que podem hibridar com BNA já hibridados com o PNA;

o BBA compreende:

- i) uma molécula ou uma porção de uma molécula que é capaz de se ligar selectivamente a uma BBR; e

- ii) um OSA, que é sem suporte e/ou indicador ligado ou com suporte ligado, ou outro meio de localização, incluindo, mas não se limitando a, ligação a esferas, polímeros e superfícies e/ou indicadores;

e o TBA compreende:

- i) uma molécula ou uma porção de uma molécula que é capaz de se ligar selectivamente a uma TBR; e
- ii) sem suporte e/ou indicador ligado ou com suporte ligado, ou outro meio de localização, incluindo, mas não se limitando a, ligação a esferas, polímeros e superfícies e/ou indicadores.

9. Método de hibridação de fase sólida para detecção da presença de um polinucleótido alvo através da utilização de um PNA tal como definido na reivindicação 1, envolvendo: a imobilização de um polinucleótido alvo, se presente numa amostra de teste, directamente ou através de uma estrutura de captura intermediária, numa fase sólida num local de captura; antes, durante ou após a referida imobilização, a ligação de um marcador detectável ao referido polinucleótido alvo, se presente; e detecção do referido marcador, se houver algum, no referido local de captura; em que a imobilização compreende a utilização de um conjunto de ligação ao alvo (TBA) que se liga apenas a um único híbrido do ácido nucleico alvo e um ácido nucleico sonda (PNA) compreendendo uma 1/2 BBR capaz de se ligar a um ácido nucleico de reforço (BNA) contendo uma 1/2 BBR complementar de cadeia simples que, após hibridação com a 1/2 BBR no PNA, forma uma BBR capaz de se ligar a conjuntos de ligação de reforço marcados (BBA), em que os termos TBA, BNA, BBR e BBA são tal como definidos na reivindicação 1.

10. Estojo de teste de diagnóstico ou forense para a detecção de uma sequência de ácido nucleico alvo numa amostra de ácido nucleico, que compreende uma primeira e uma segunda sondas de ácido nucleico e uma primeira e uma segunda proteínas de ligação ao ácido nucleico,

em que a primeira sonda possui uma sequência complementar à sequência alvo e uma sequência adicional;

a primeira proteína de ligação é específica para o primeiro dúplex sonda-alvo;

a segunda sonda é complementar à sequência adicional na primeira sonda; e

a segunda proteína de ligação liga-se especificamente ao dúplice primeira sonda-segunda sonda e é marcada com um marcador detectável.

11. Estojo de acordo com a reivindicação 10, em que a primeira sonda é complementar à LTR de HIV e, após hibridação da primeira sonda com a LTR de HIV, forma-se um local de ligação para NF-kB ou uma sua subunidade, SP1, proteína de ligação a TATA, HIV-Detect I, II, III ou IV ou HIV-Lock.

12. Estojo de acordo com a reivindicação 11, em que a primeira proteína de ligação é NF-kB ou uma sua subunidade, SP1, proteína de ligação a TATA, HIV-Detect I, II, III ou IV ou HIV-Lock.

13. Estojo de acordo com qualquer uma das reivindicações 10 a 12, em que a primeira sonda, para além de ser complementar à LTR de HIV, compreende uma sequência que codifica o operador esquerdo ou direito do bacteriófago lambda, e a segunda sonda compreende a sequência complementar à referida sequência do operador esquerdo ou direito do bacteriófago lambda, de modo que, após hibridação da primeira e da segunda sondas, se forma um local de ligação para a proteína repressora CI do bacteriófago lambda, a proteína *cro* do bacteriófago lambda ou um seu derivado ou homólogo.

14. Estojo de acordo com a reivindicação 13, em que a segunda proteína de ligação é a proteína repressora CI do bacteriófago lambda, a proteína *cro* do bacteriófago lambda ou um seu derivado ou homólogo.

Lisboa,

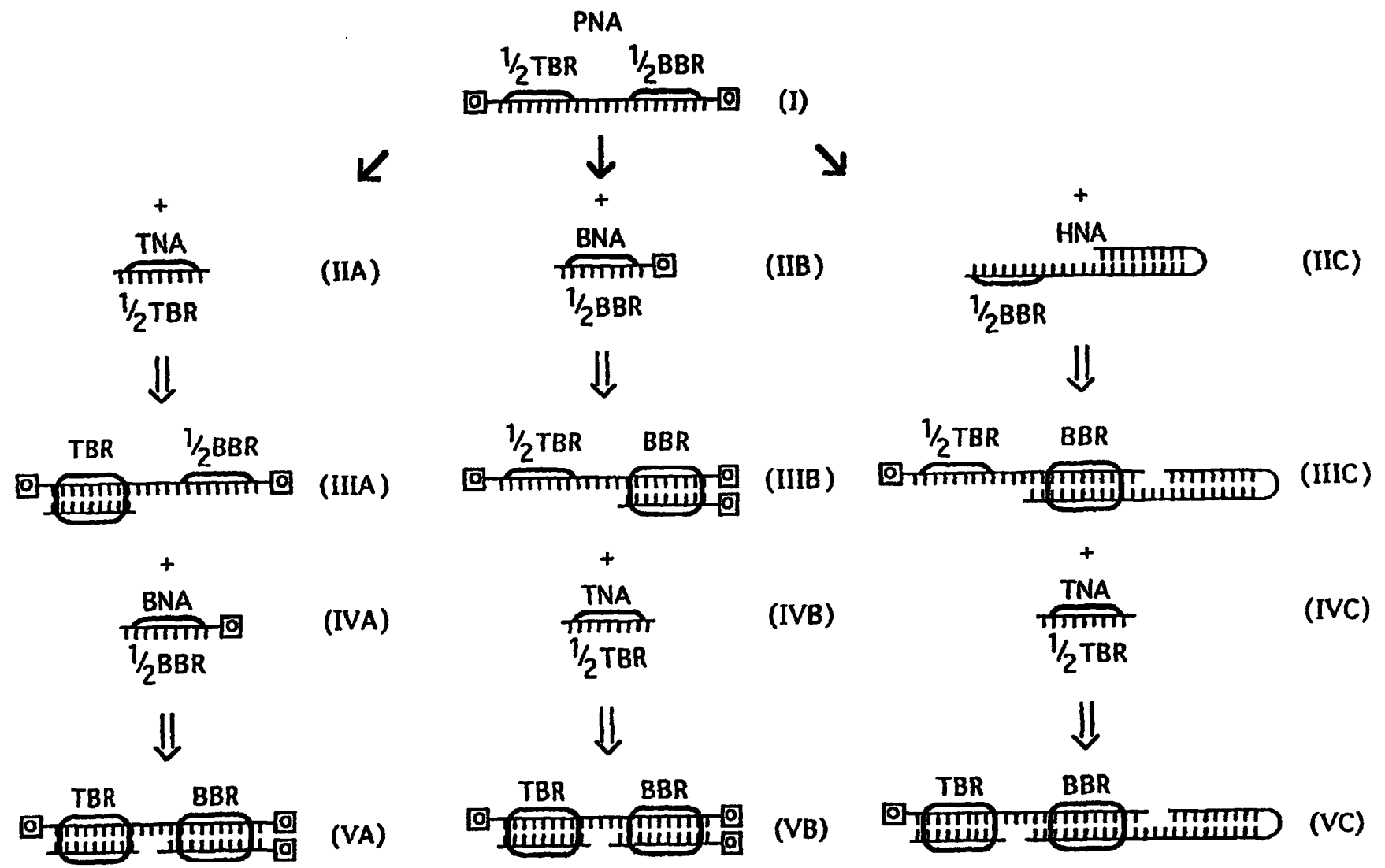


FIGURA 1

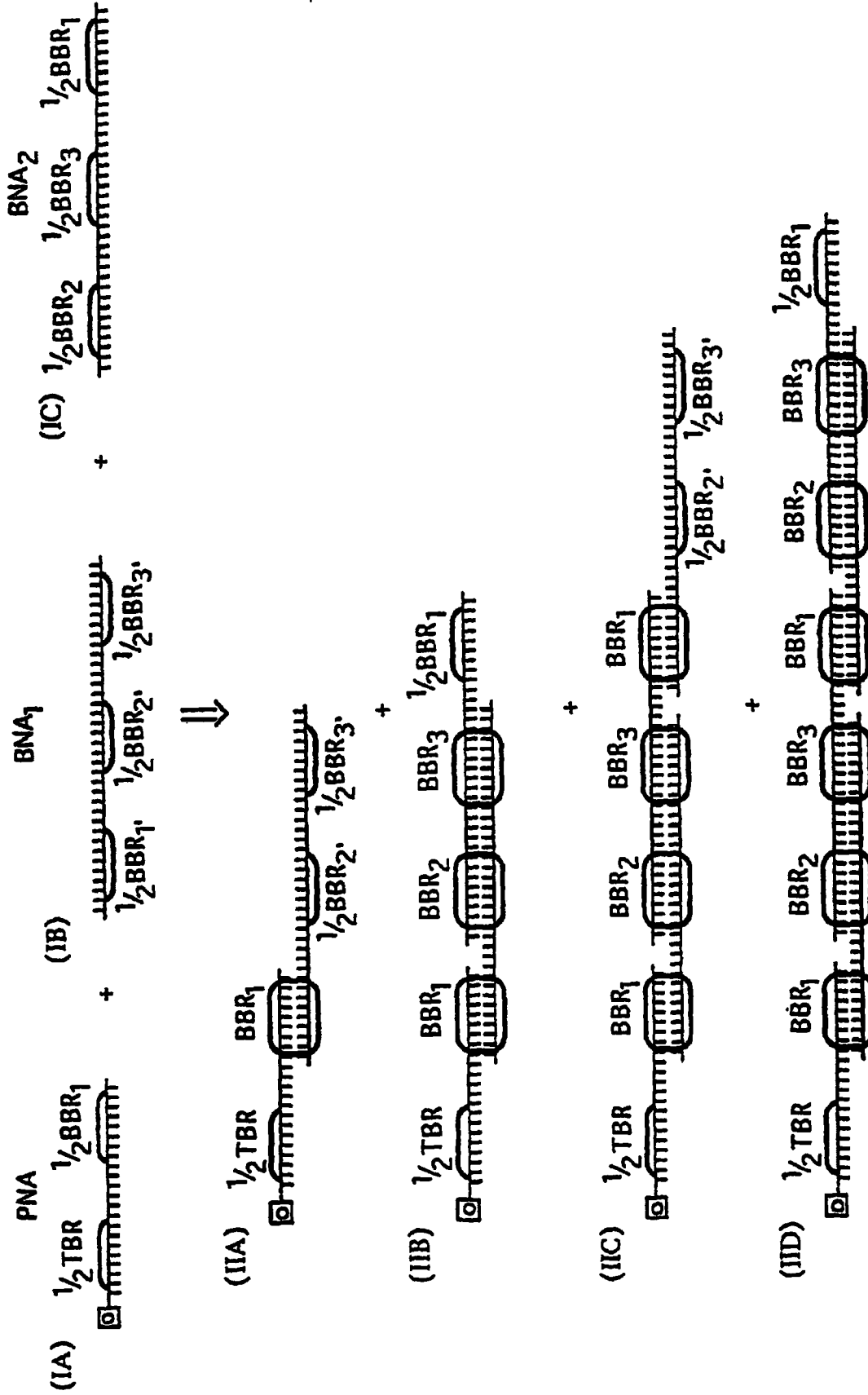


FIGURA 2A

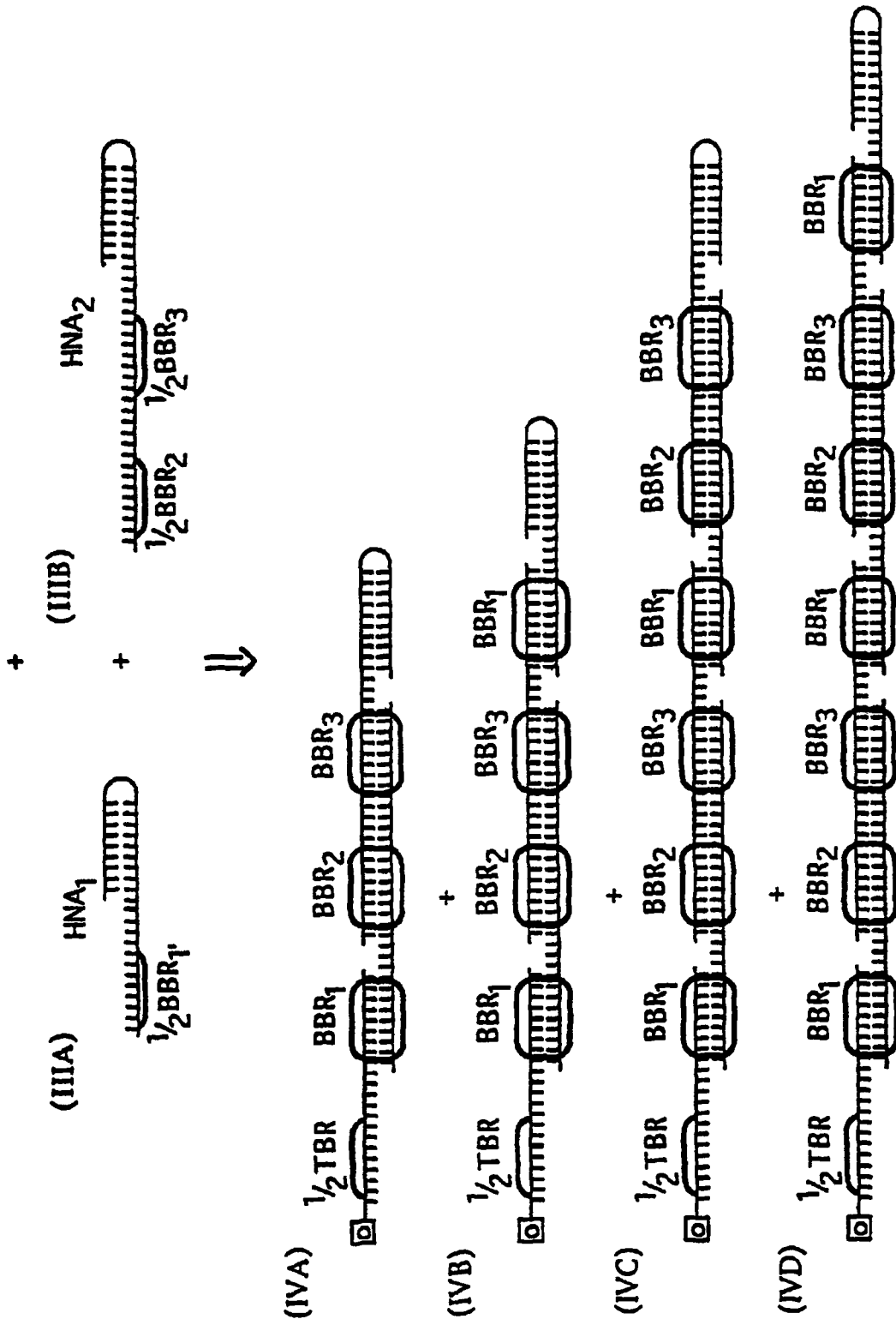


FIGURA 2B

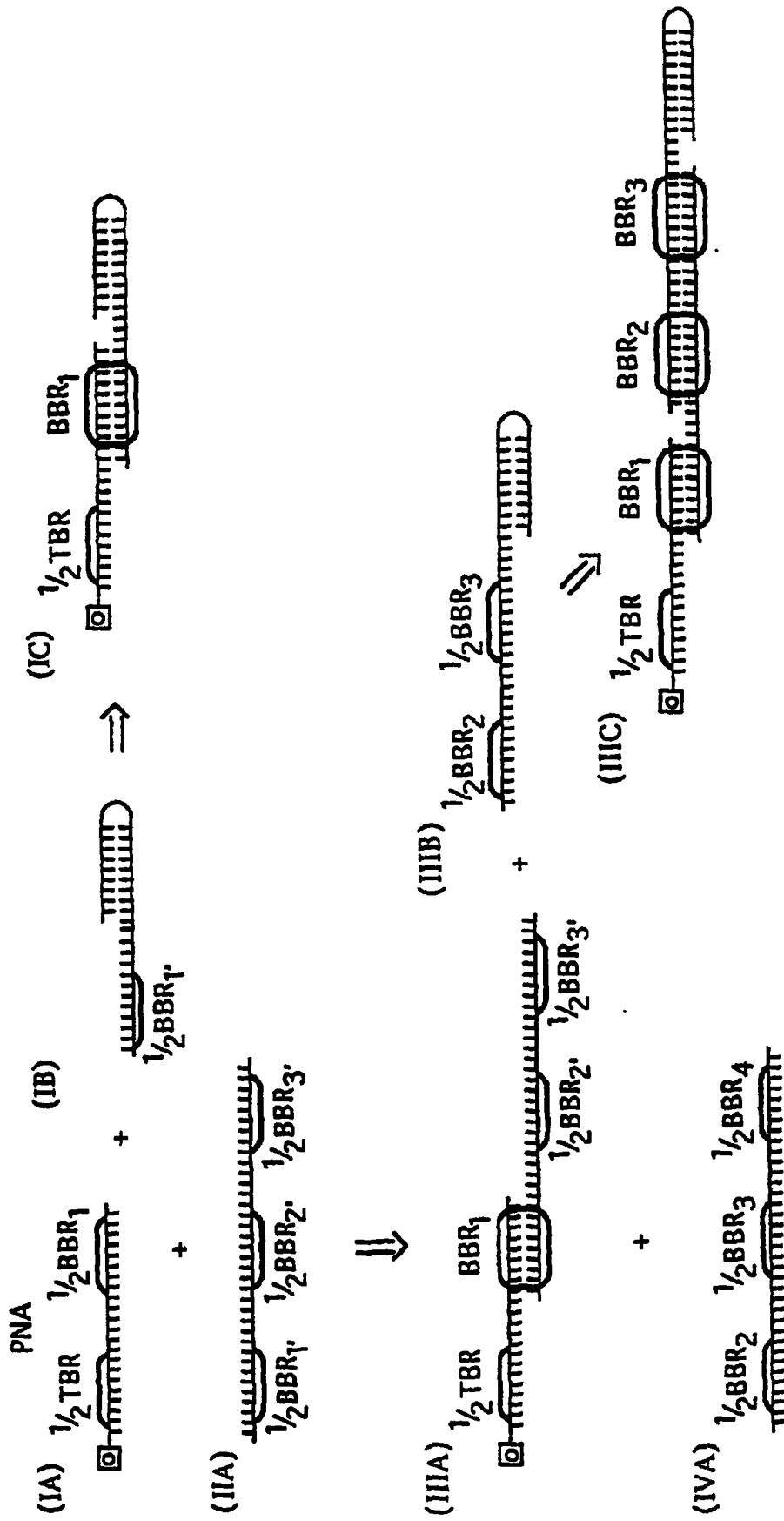


FIGURA 2C

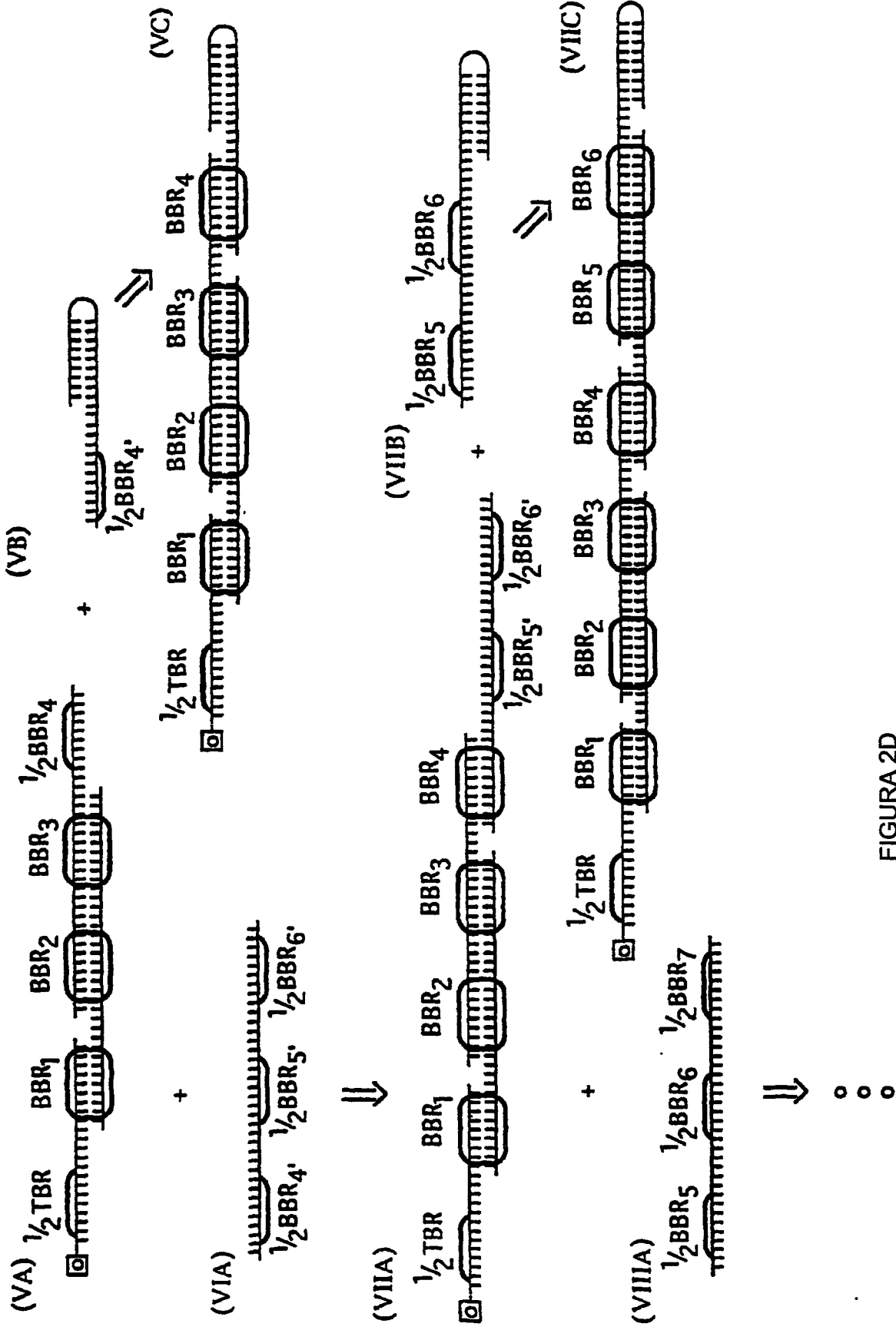


FIGURA 2D

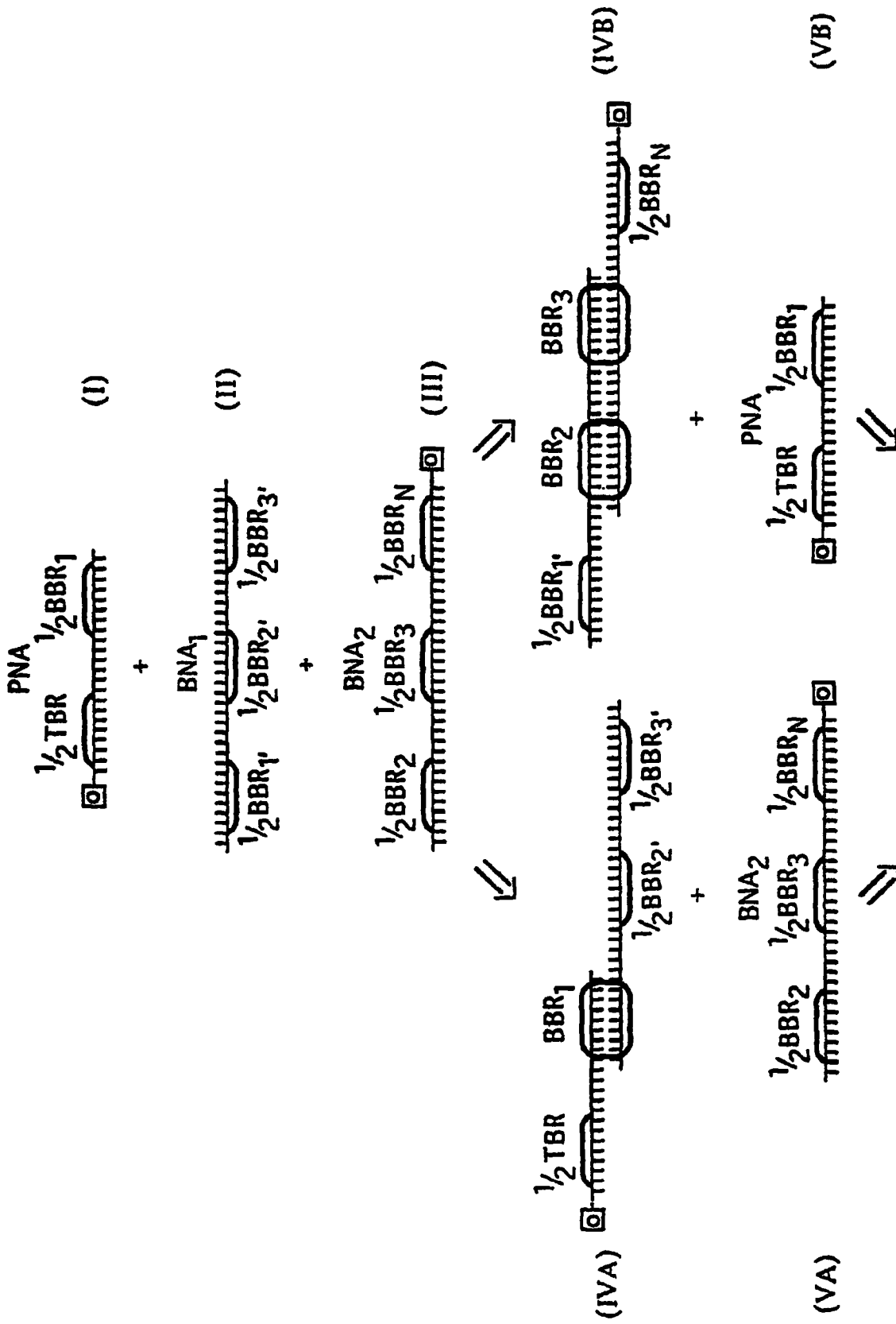


FIGURA 3A

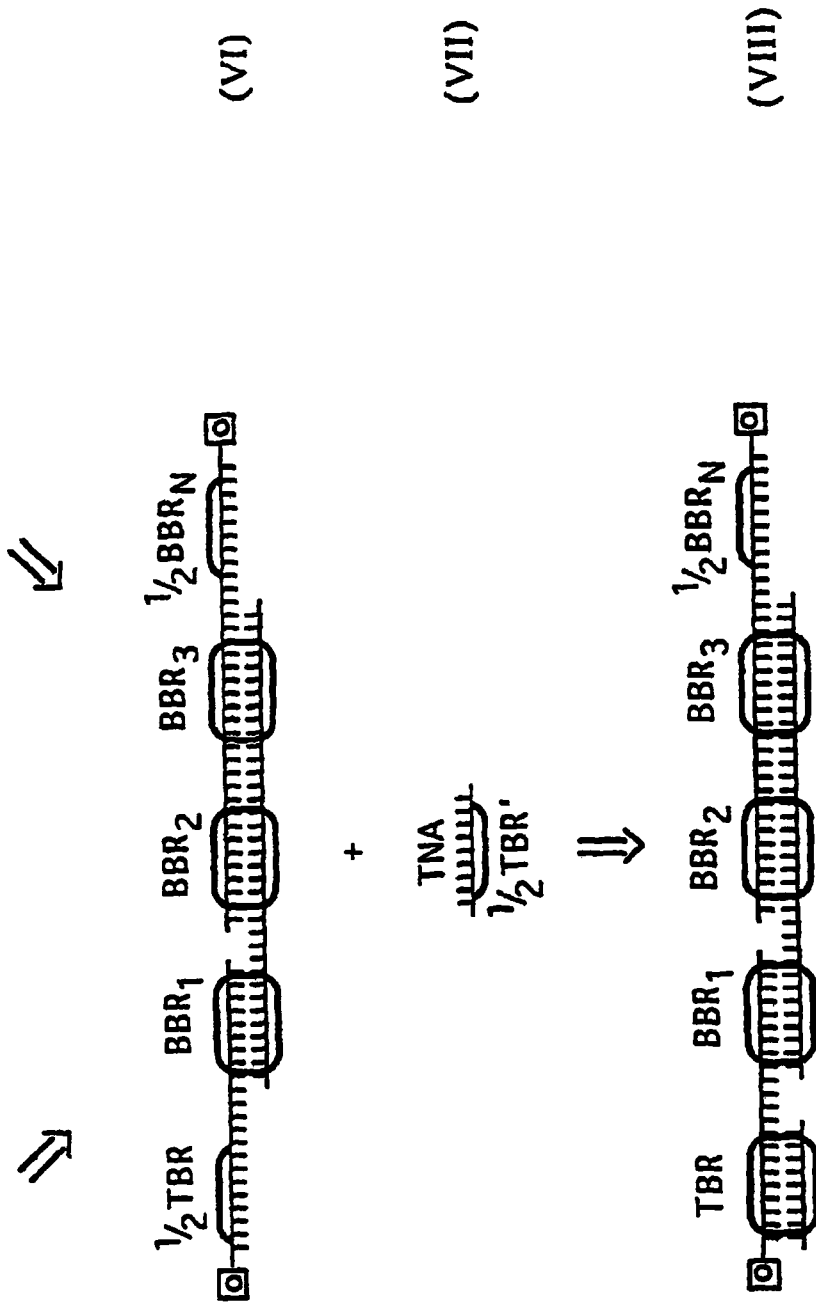


FIGURA 3B

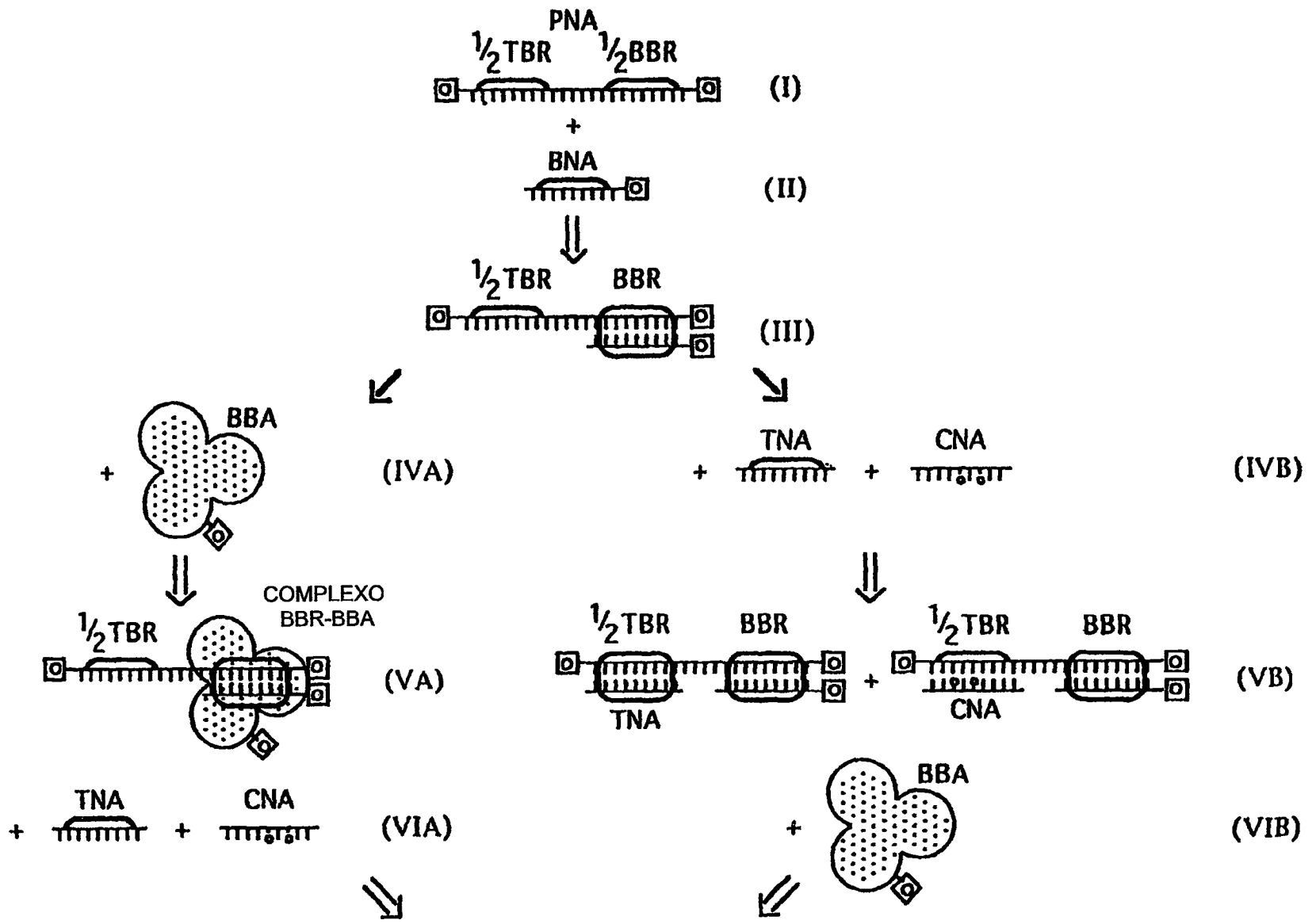


FIGURA 4A

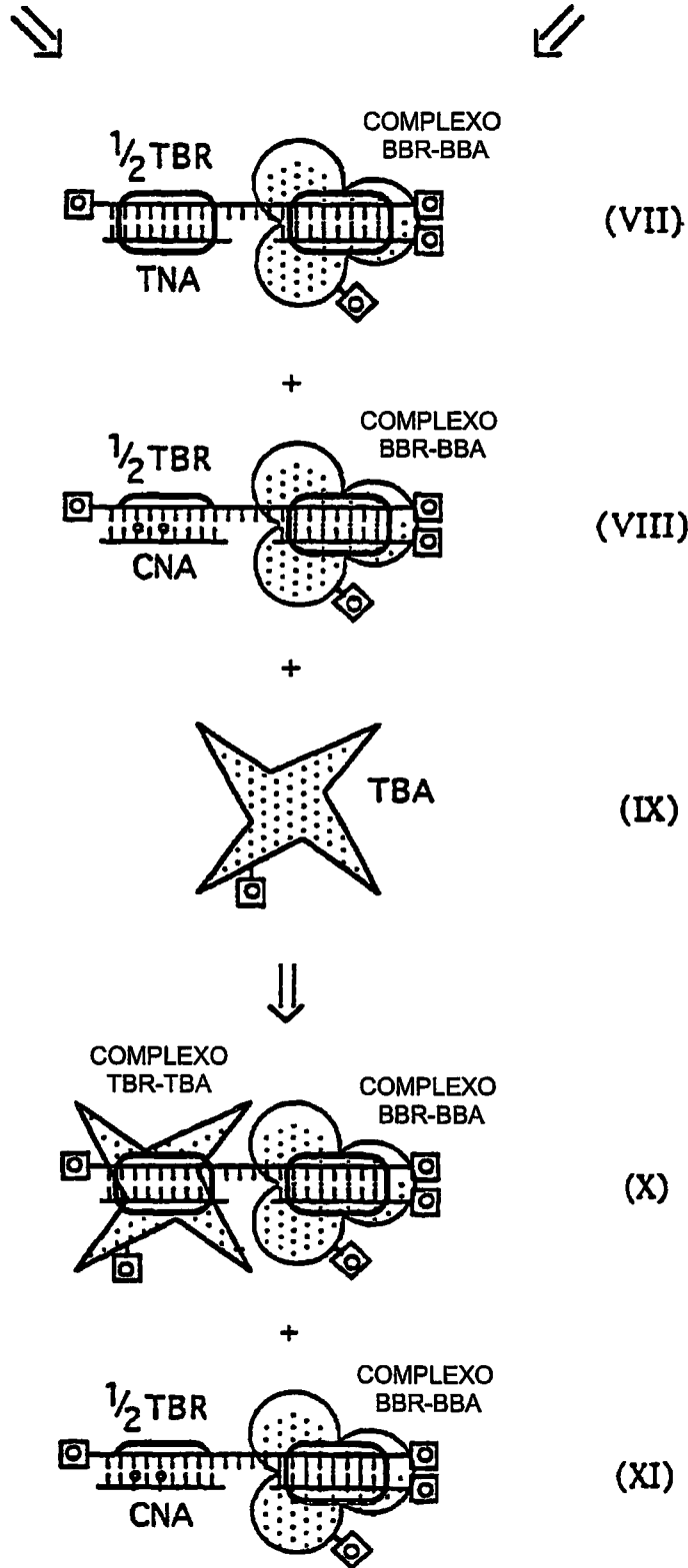


FIGURA 4B

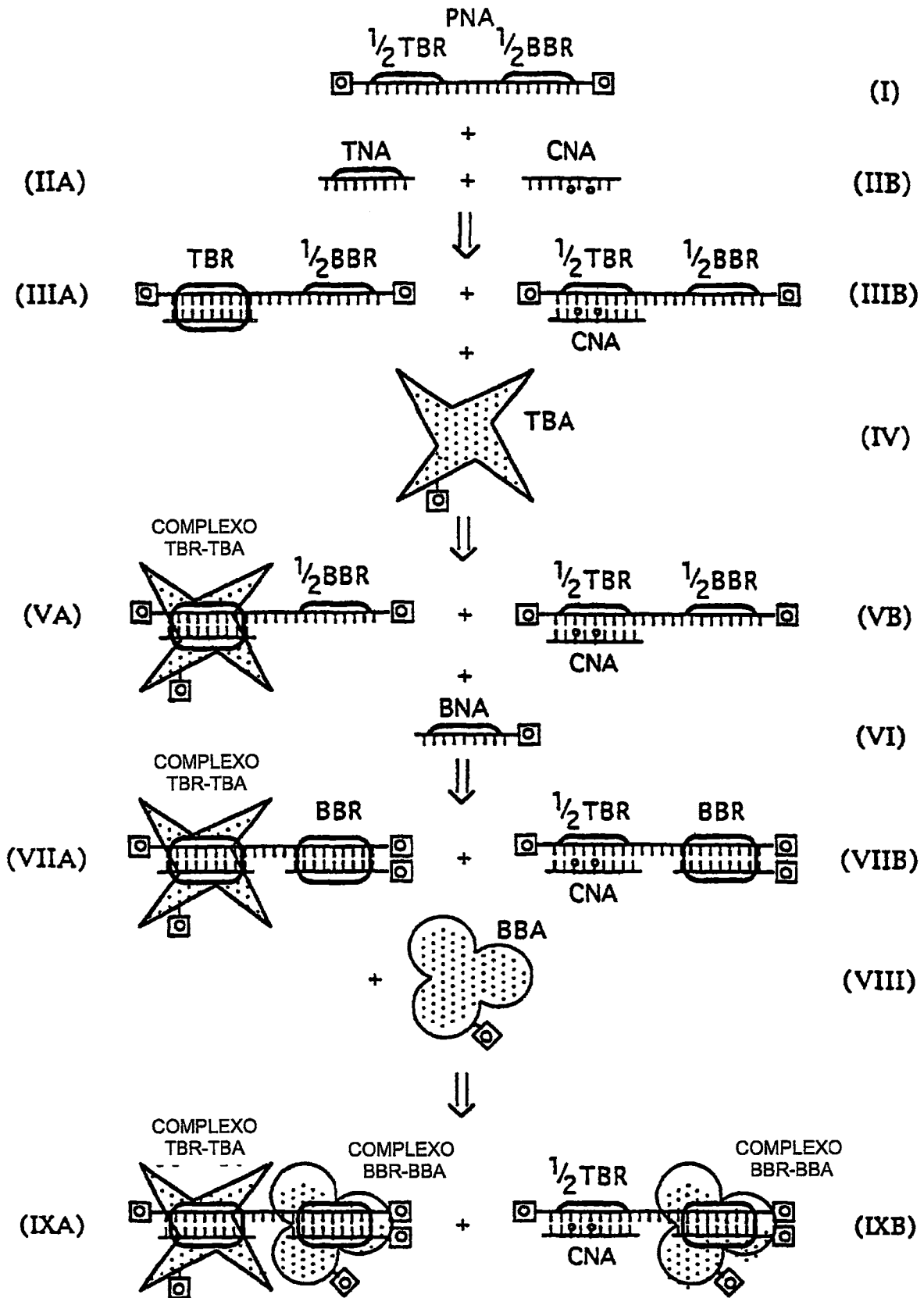


FIGURA 4C

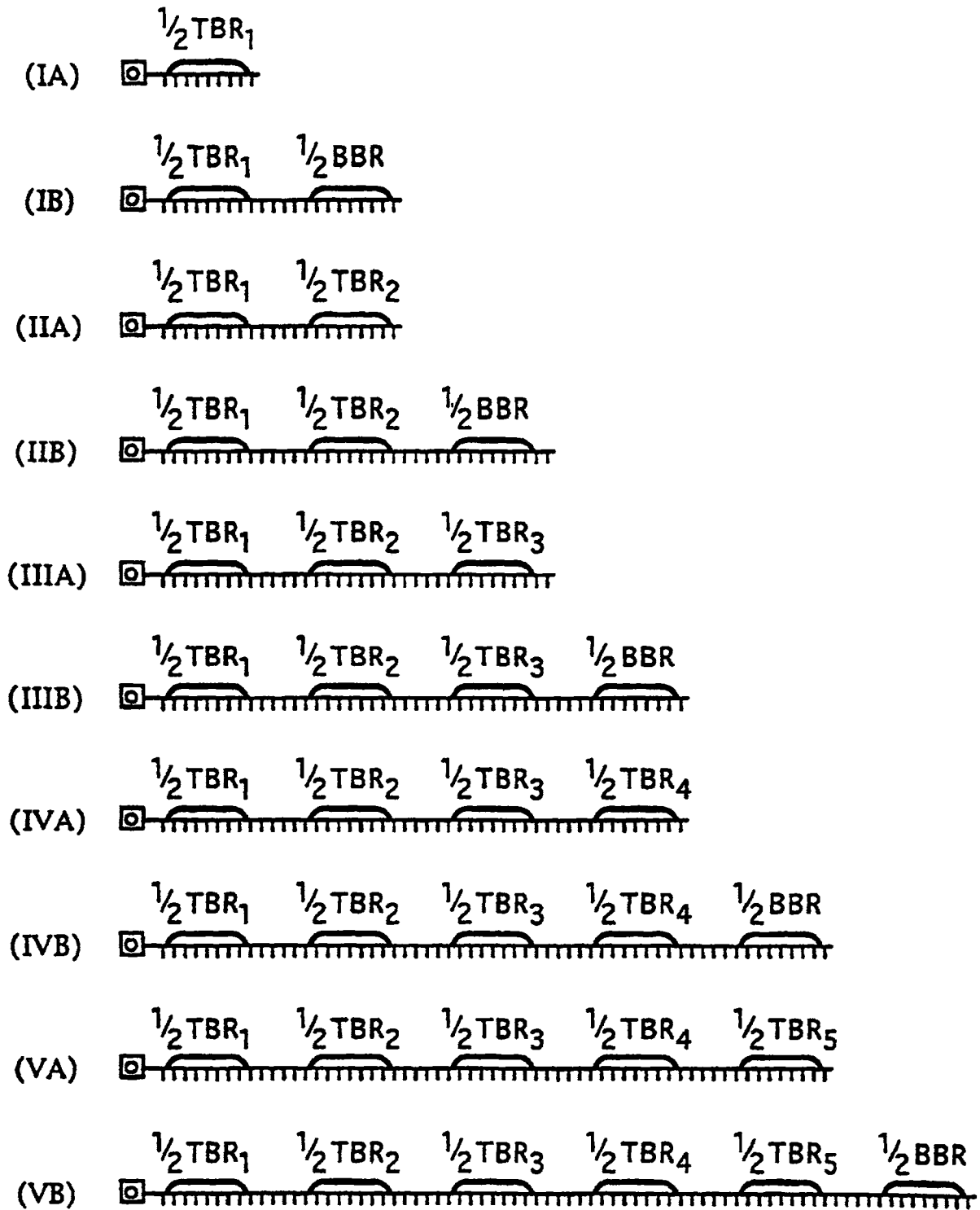


FIGURA 5

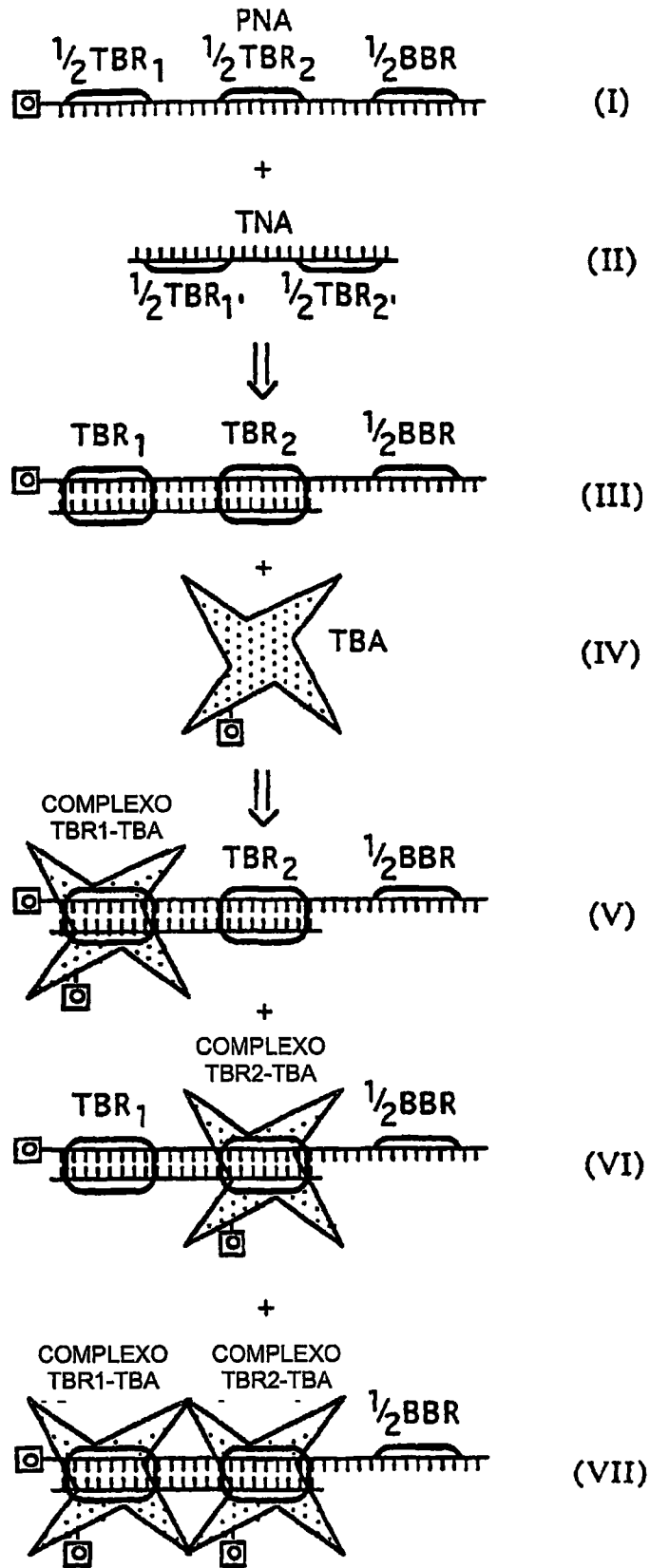
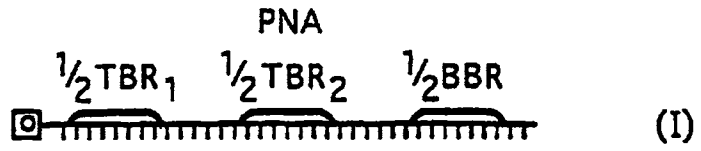


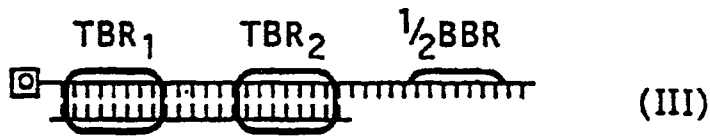
FIGURA 6A



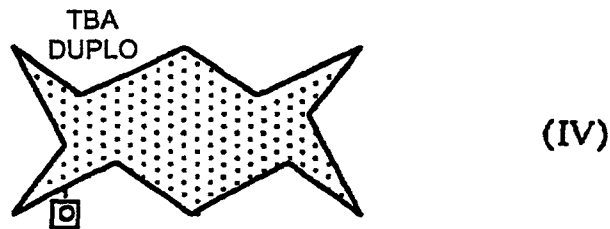
+



⇓



+



⇓

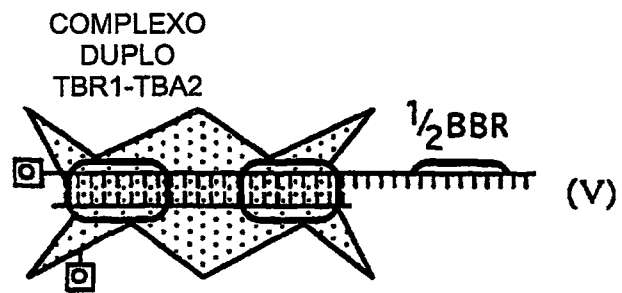


FIGURA 6B

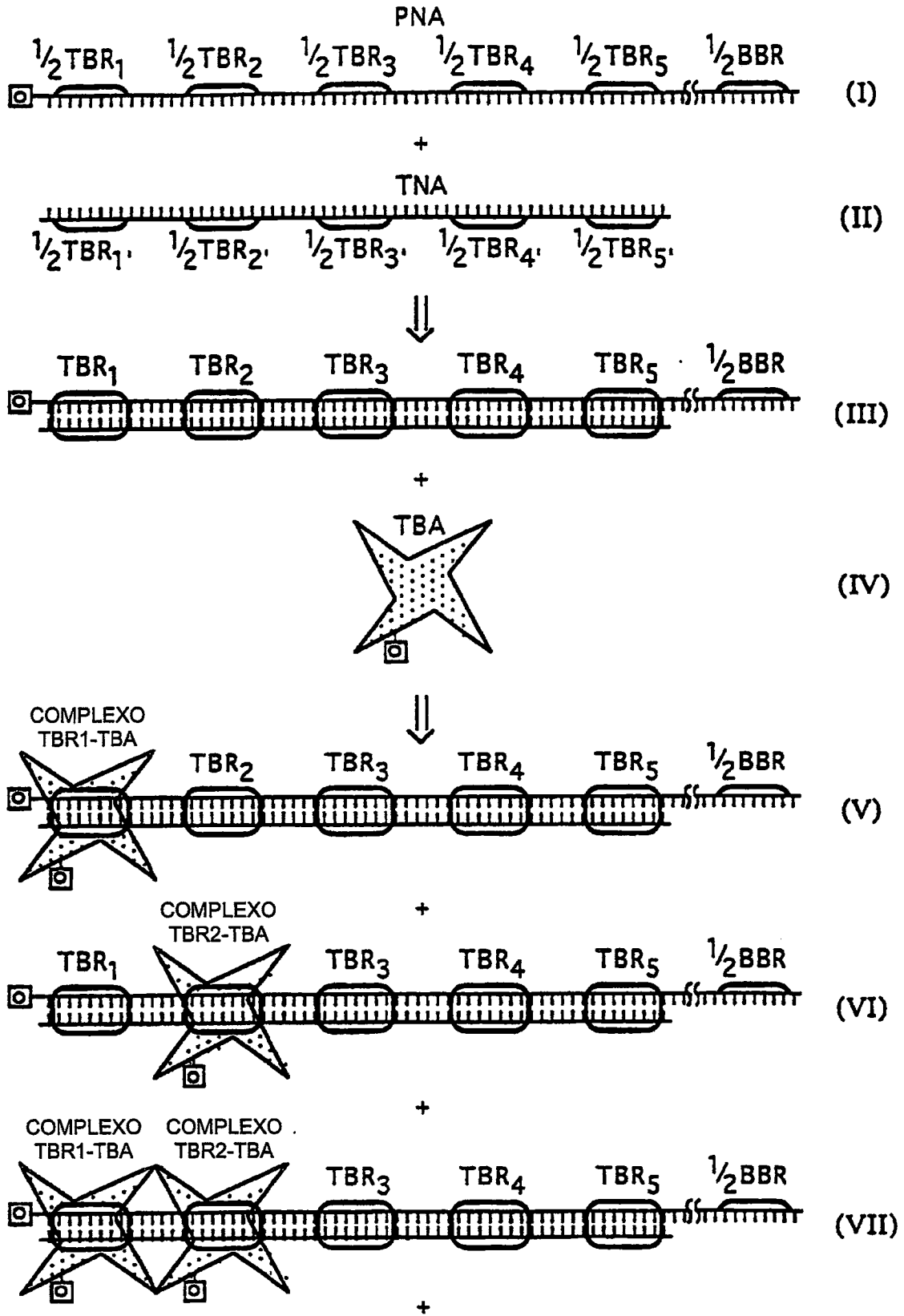


FIGURA 6C

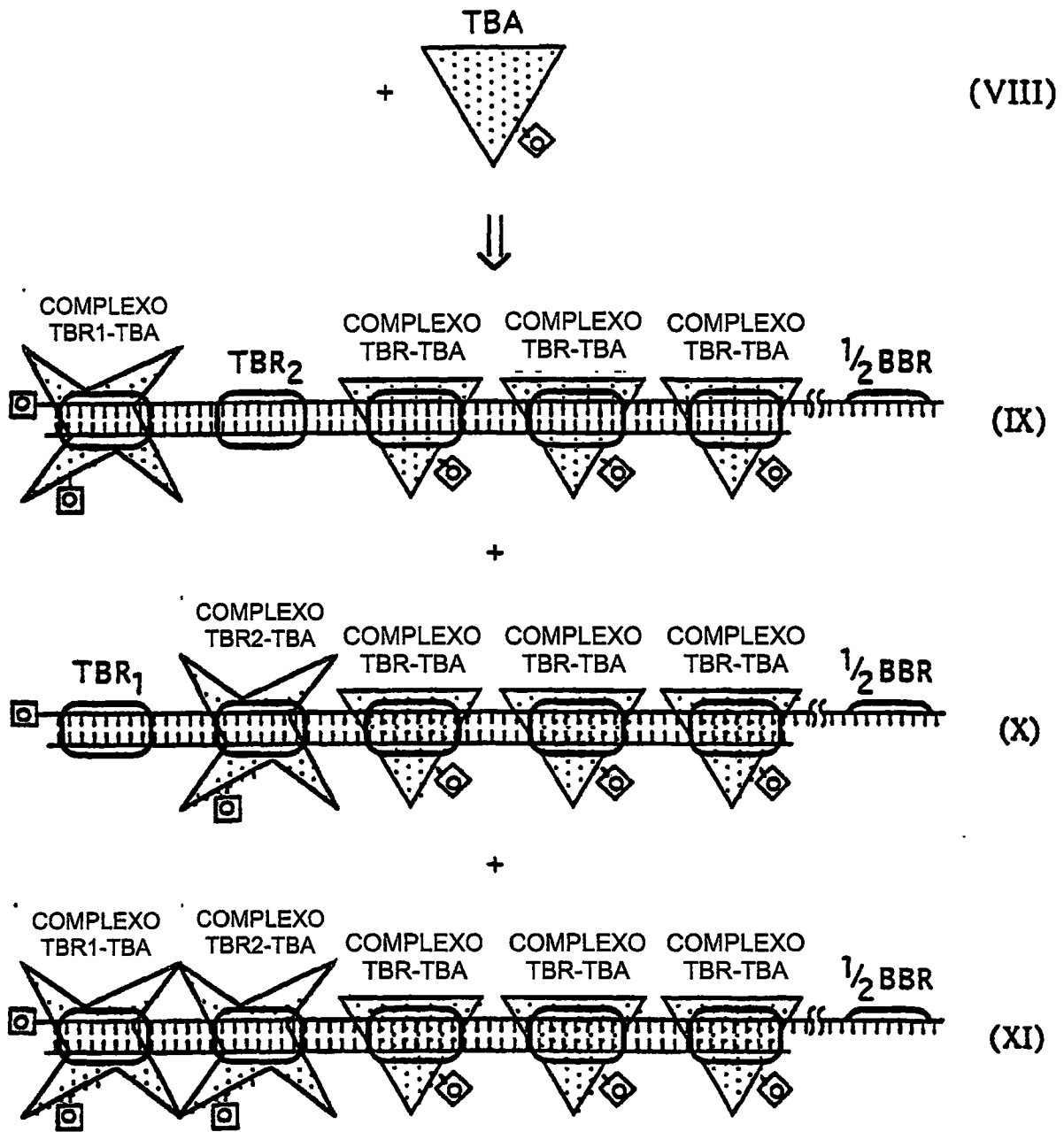


FIGURA 6D

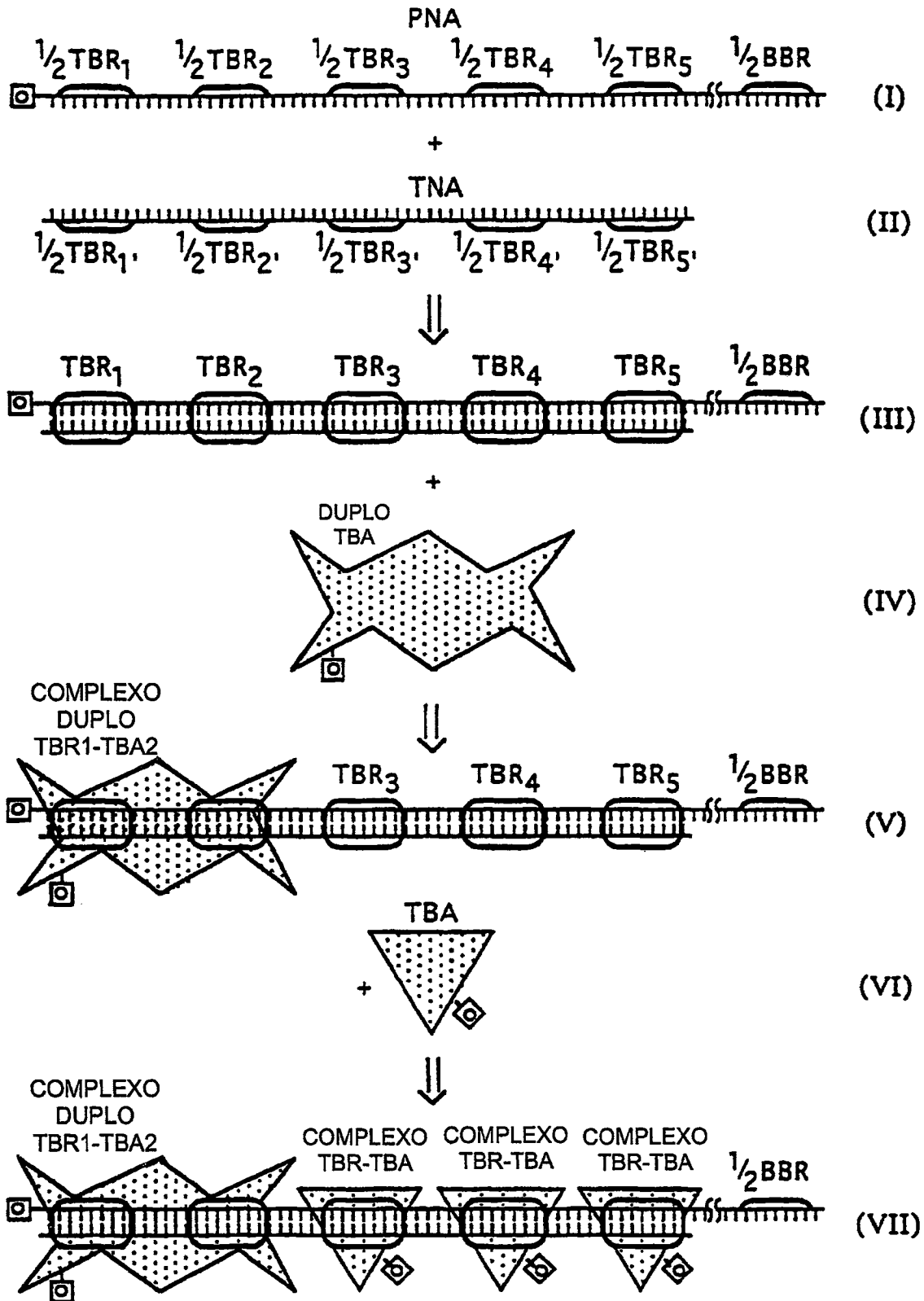


FIGURA 6E

SEQ. ID: 37:

12345678901234567890123456789012345678901234567890

CTACAAGGGACTTTCCGCTGGGGACTTTCCAGGGAGGCGTGGCCTGGGCGGGACTGGGGAGTGGCGTCCC

+++++

NF-kB NF-kB SP1 SP1 SP1

PNA1 do Estoj de Teste de HIV (+++ de cima), SEQ ID NO:38:

CTACAAGGGACTTTCCGCTGGGGACTTTCCAGGGAGG

PNA2 do Estoj de Teste de HIV (==== de cima), SEQ ID NO:39:

###CGGGACTGGGGAGTGGCGTCCC###

A sequêcia de extremidades adesivas em PNA2 é complementar a uma das extremidades do ADN do operador formado a partir de:

###OL1-OL2-OL3

OL1'-OL2'-OL3'***

ou

###OR3-OR2-OR1

OR3'-OR2'-OR1'***

FIGURA 7

AMOSTRA DE ÁCIDO NUCLEICO
DE CADEIA DUPLA FRAGMENTADO

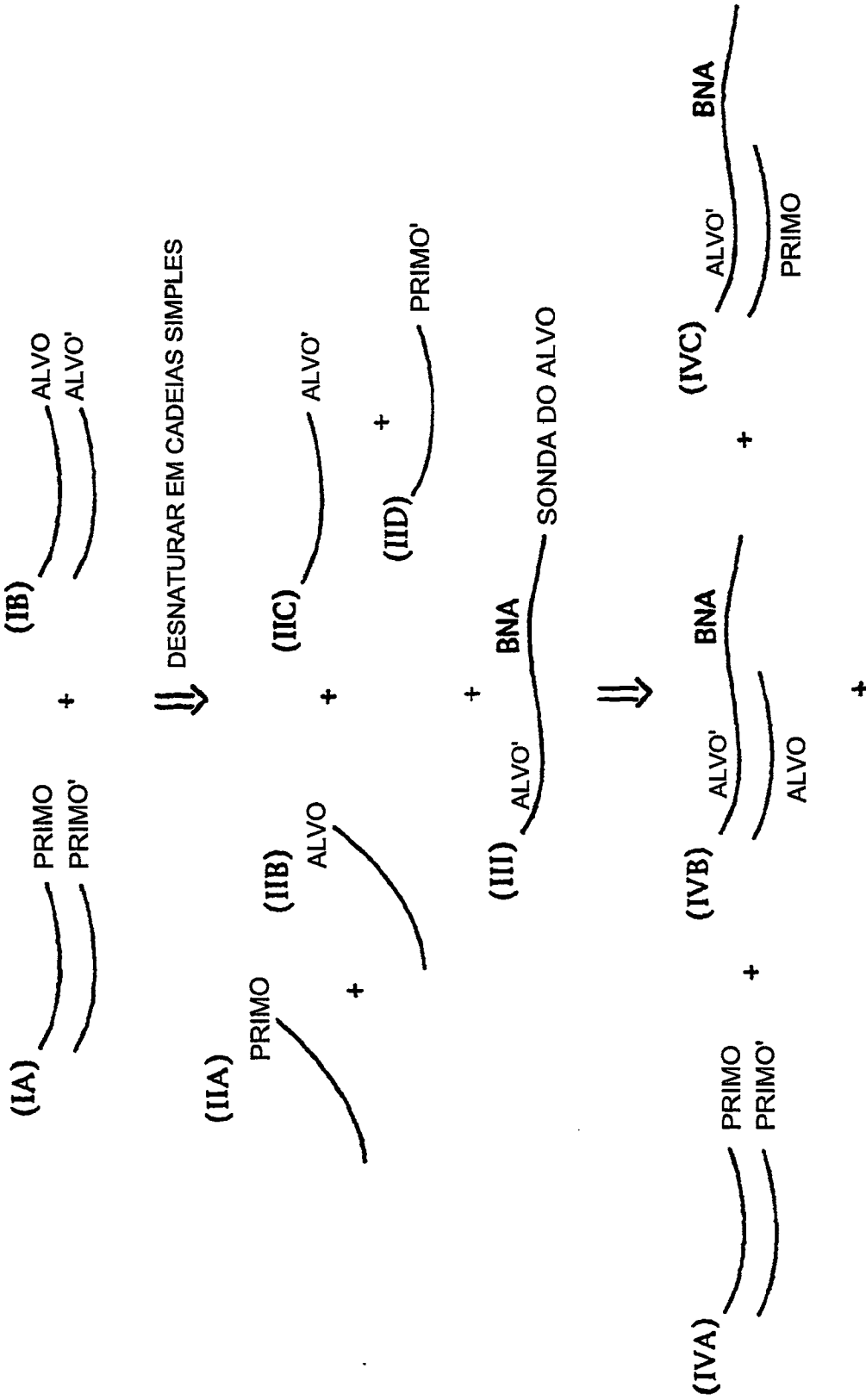


FIGURA 8A

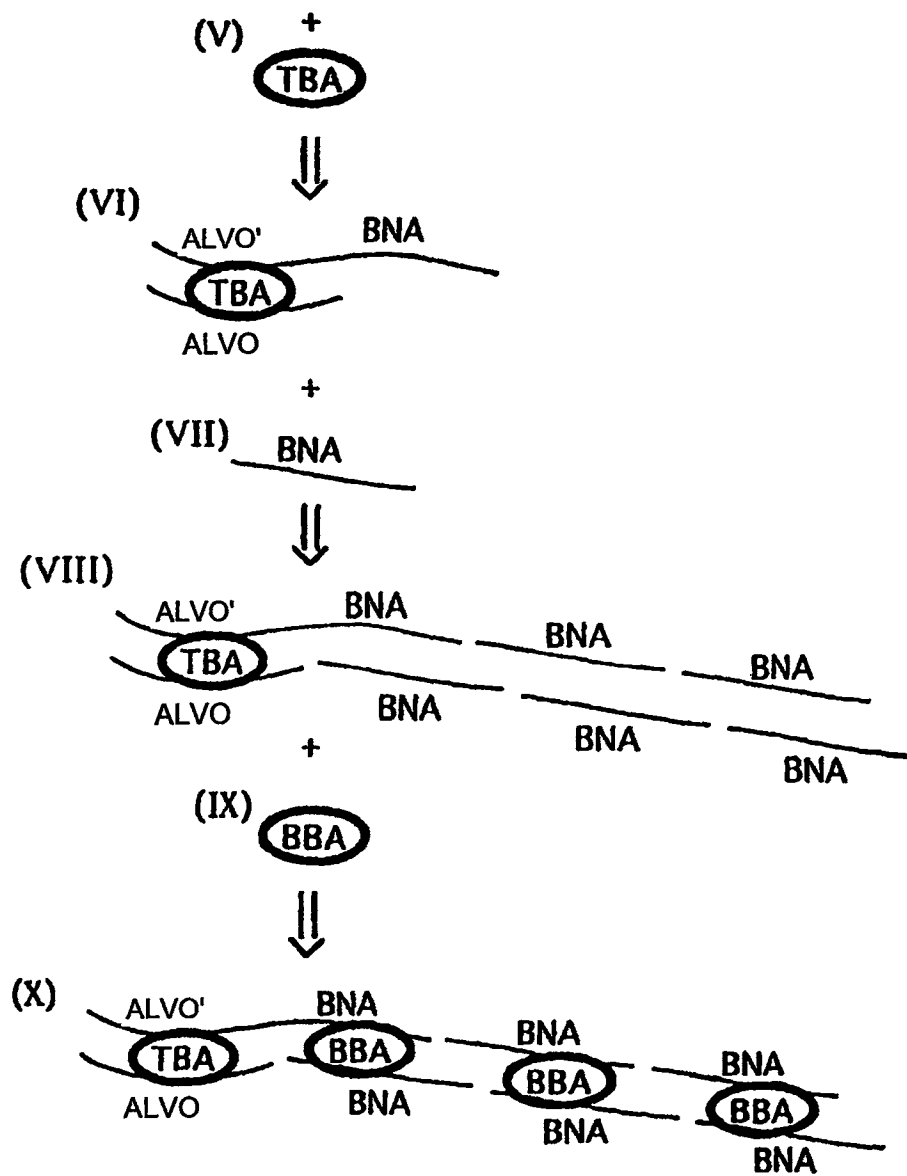


FIGURA 8B

TBA: CONJUNTO DE LIGAÇÃO AO ALVO

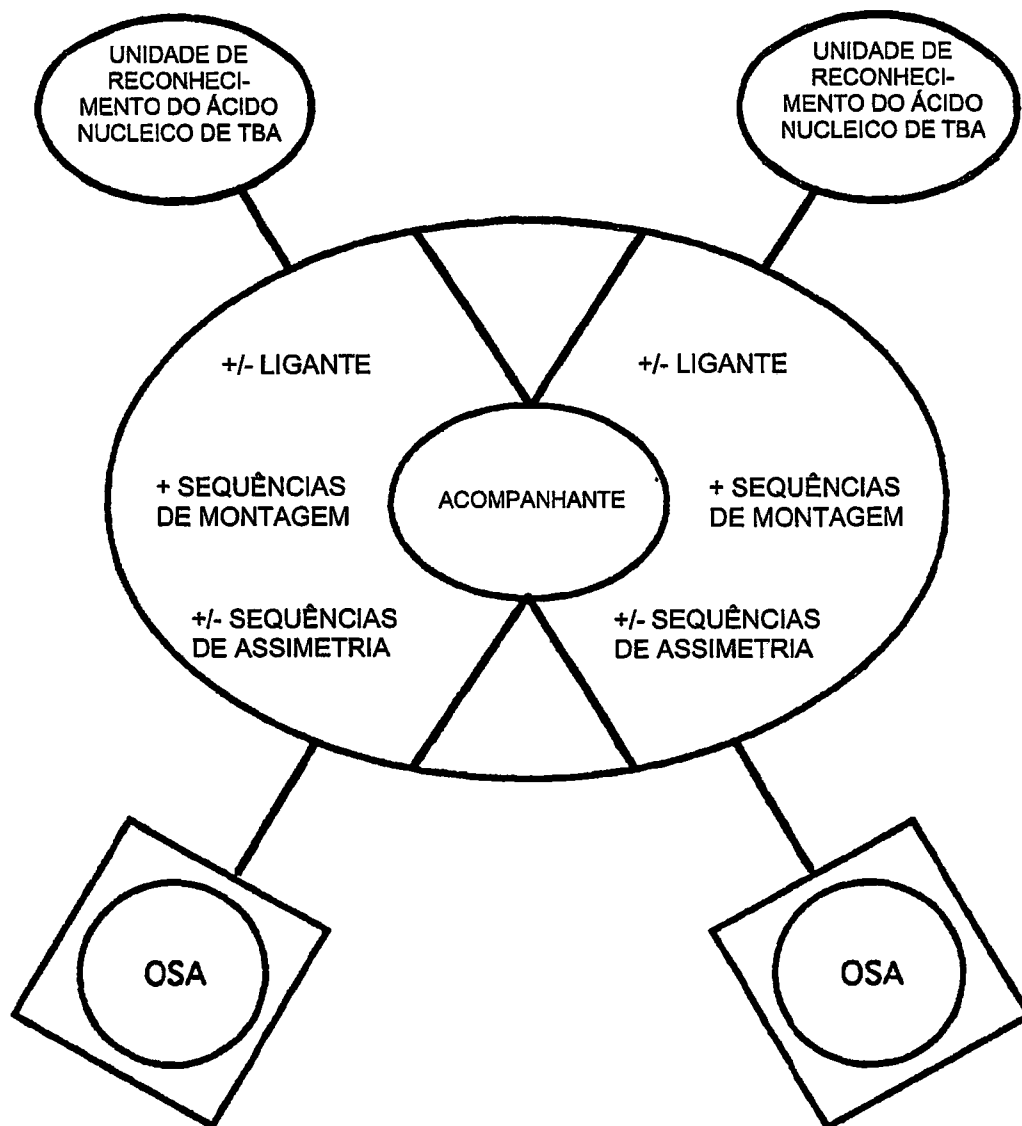
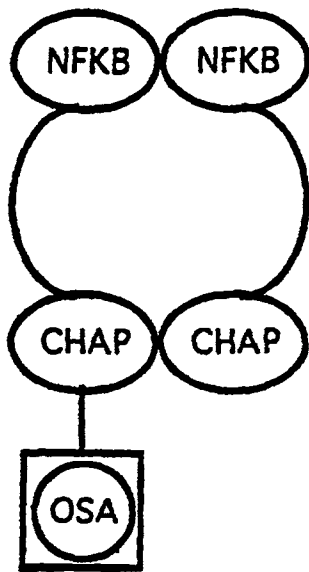
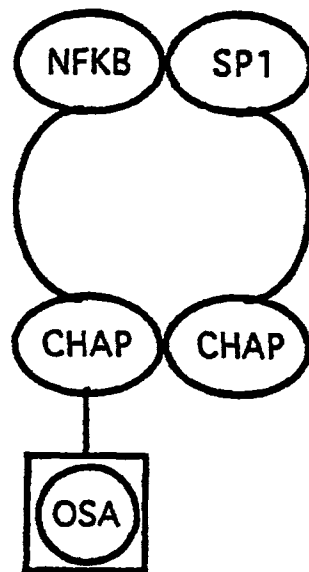


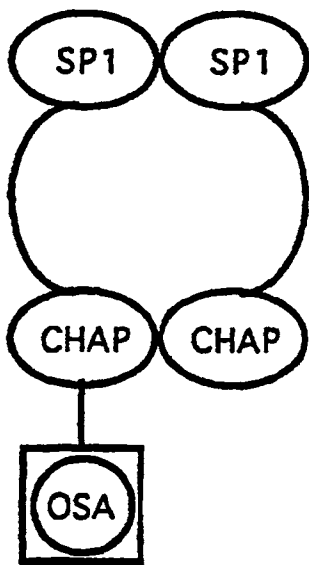
FIGURA 9



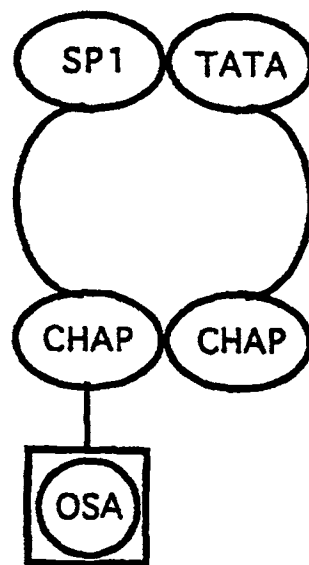
HIV-DETECT I



HIV-DETECT II

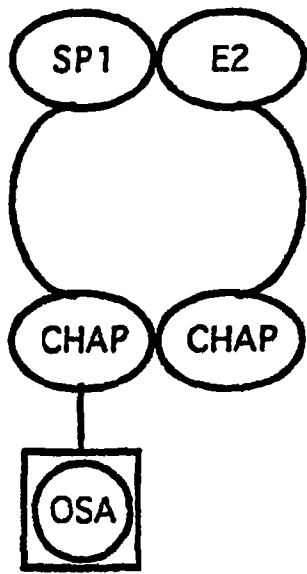


HIV-DETECT III

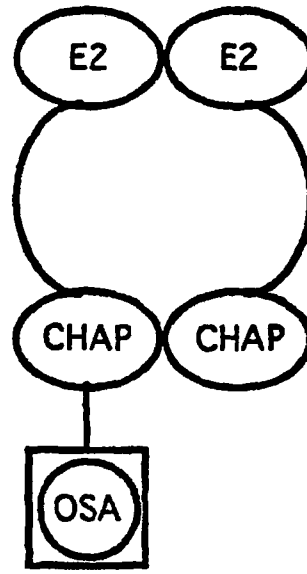


HIV-DETECT IV

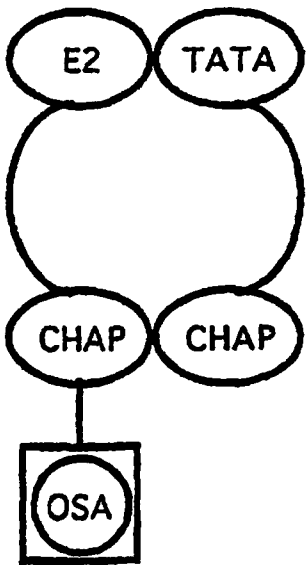
FIGURA 10



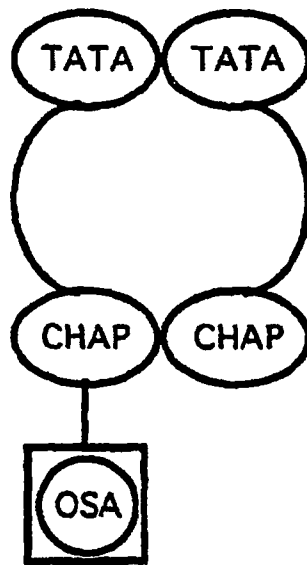
HPV-DETECT I



HPV-DETECT II



HPV-DETECT III



HPV-DETECT IV

FIGURA 11

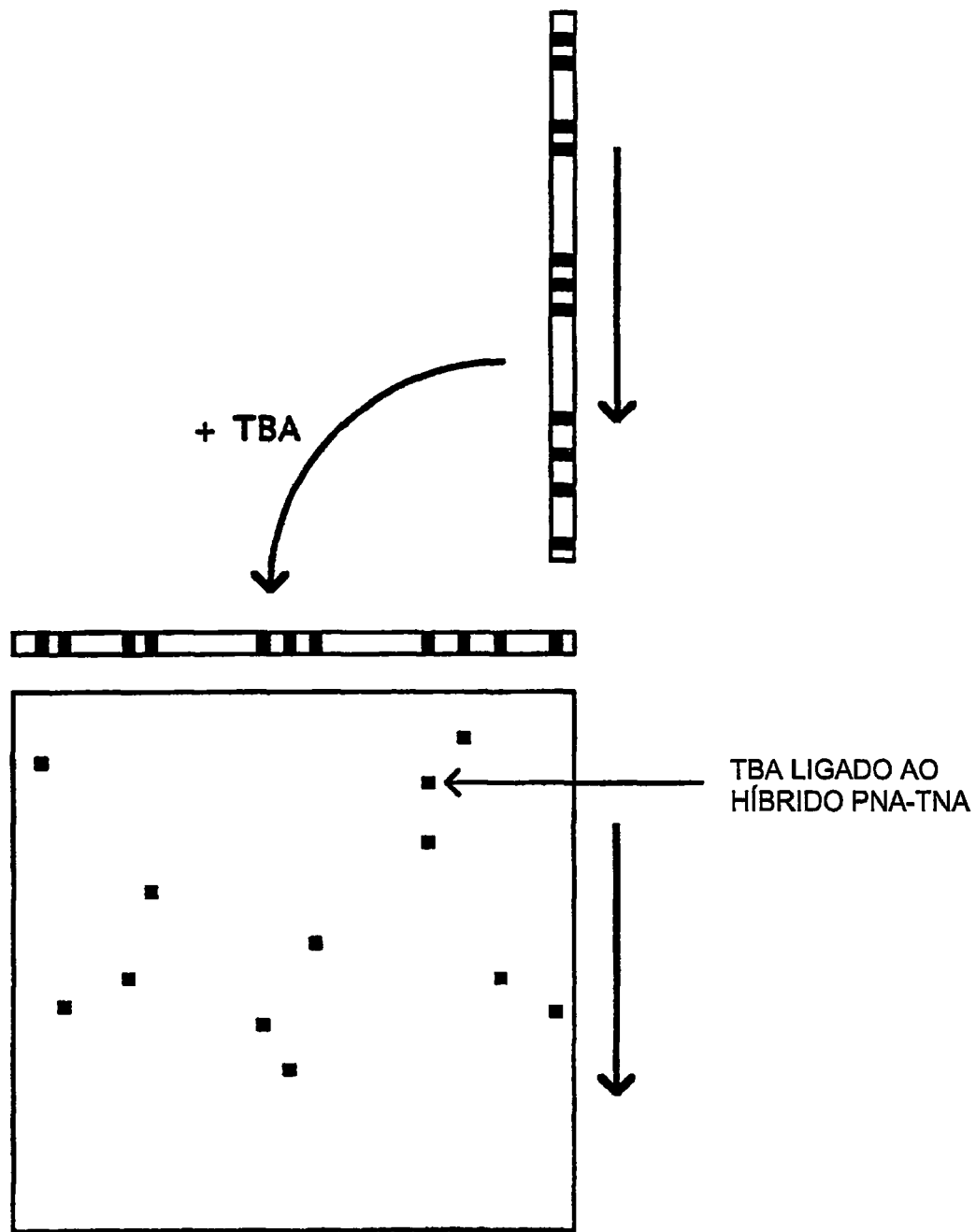


FIGURA 12A

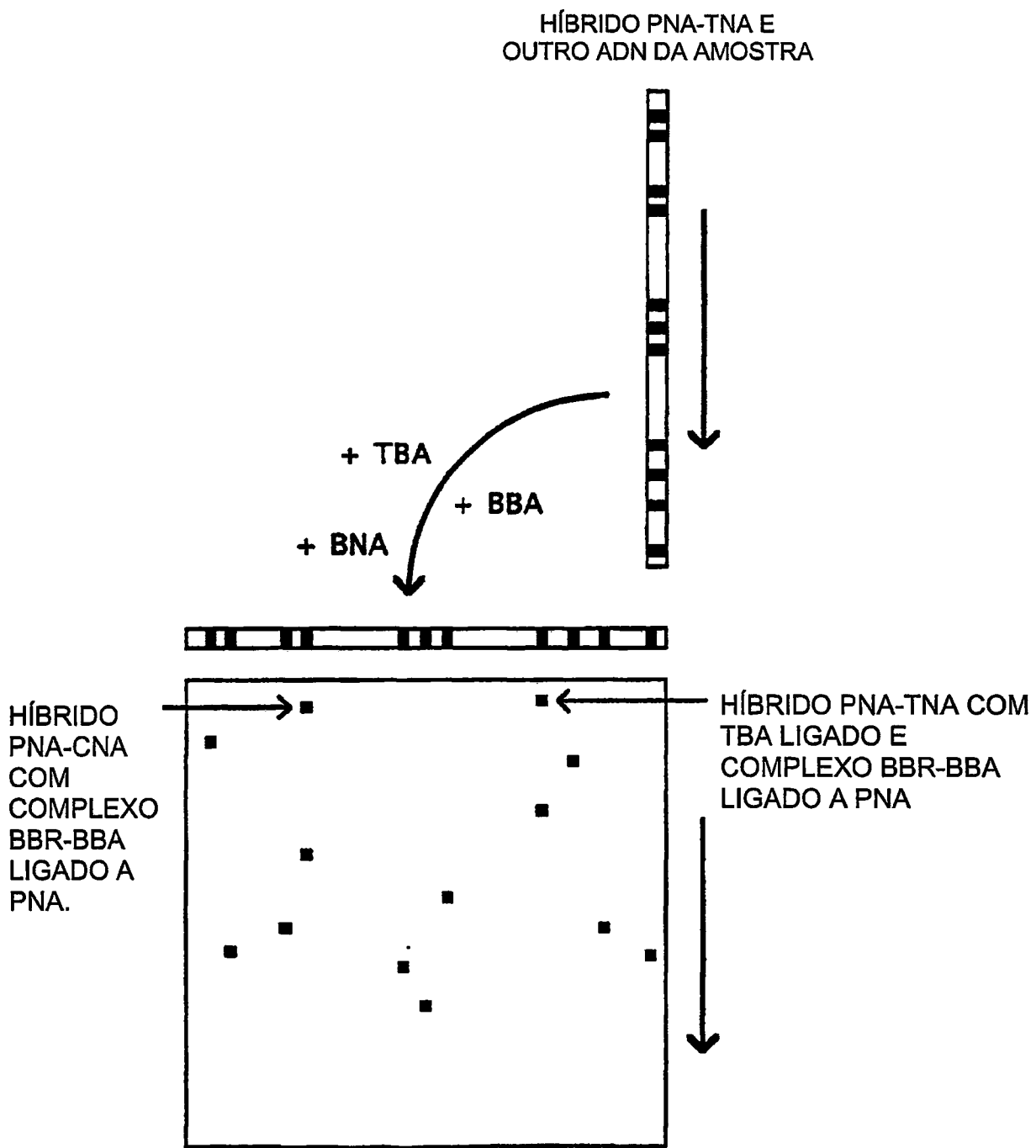


FIGURA 12B

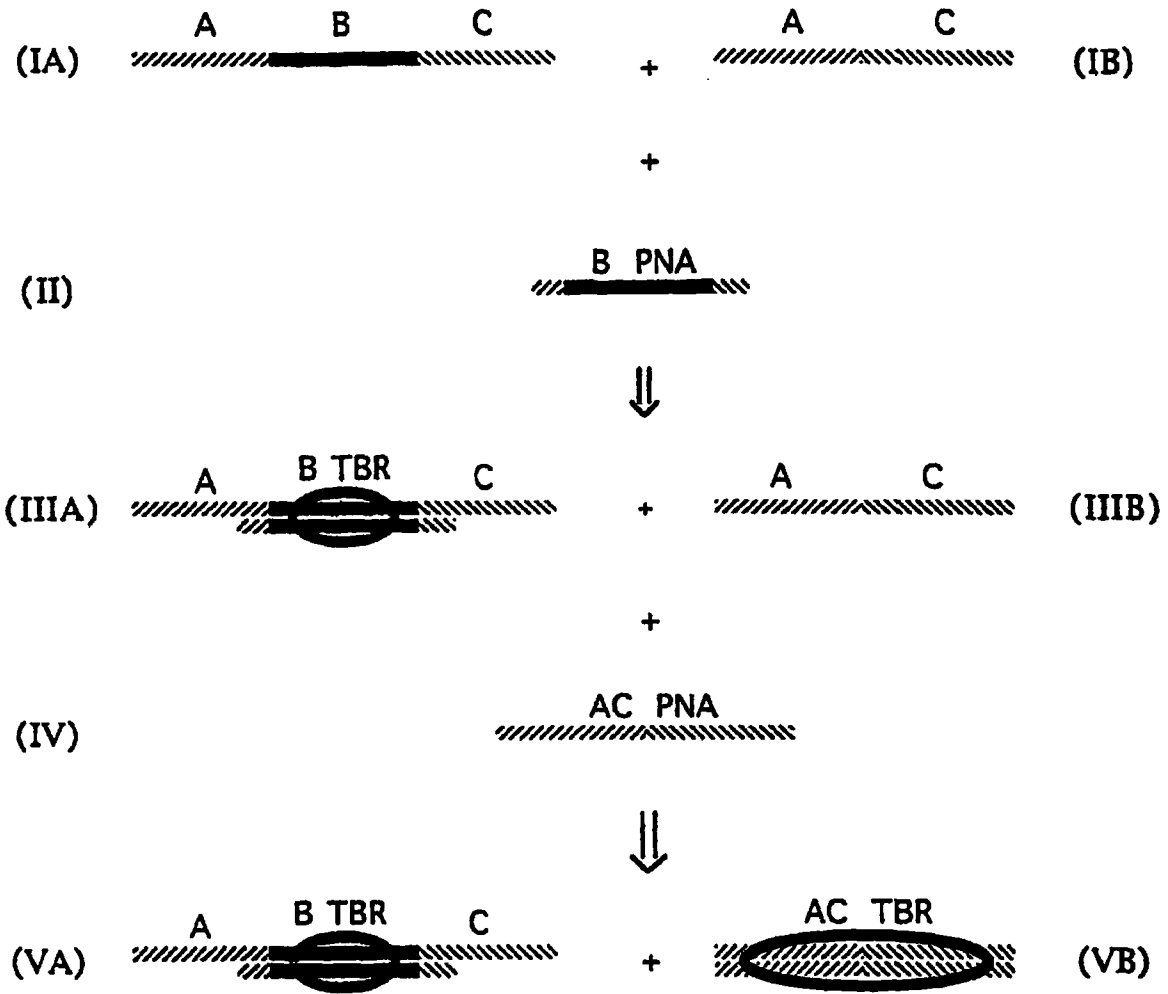


FIGURA 13

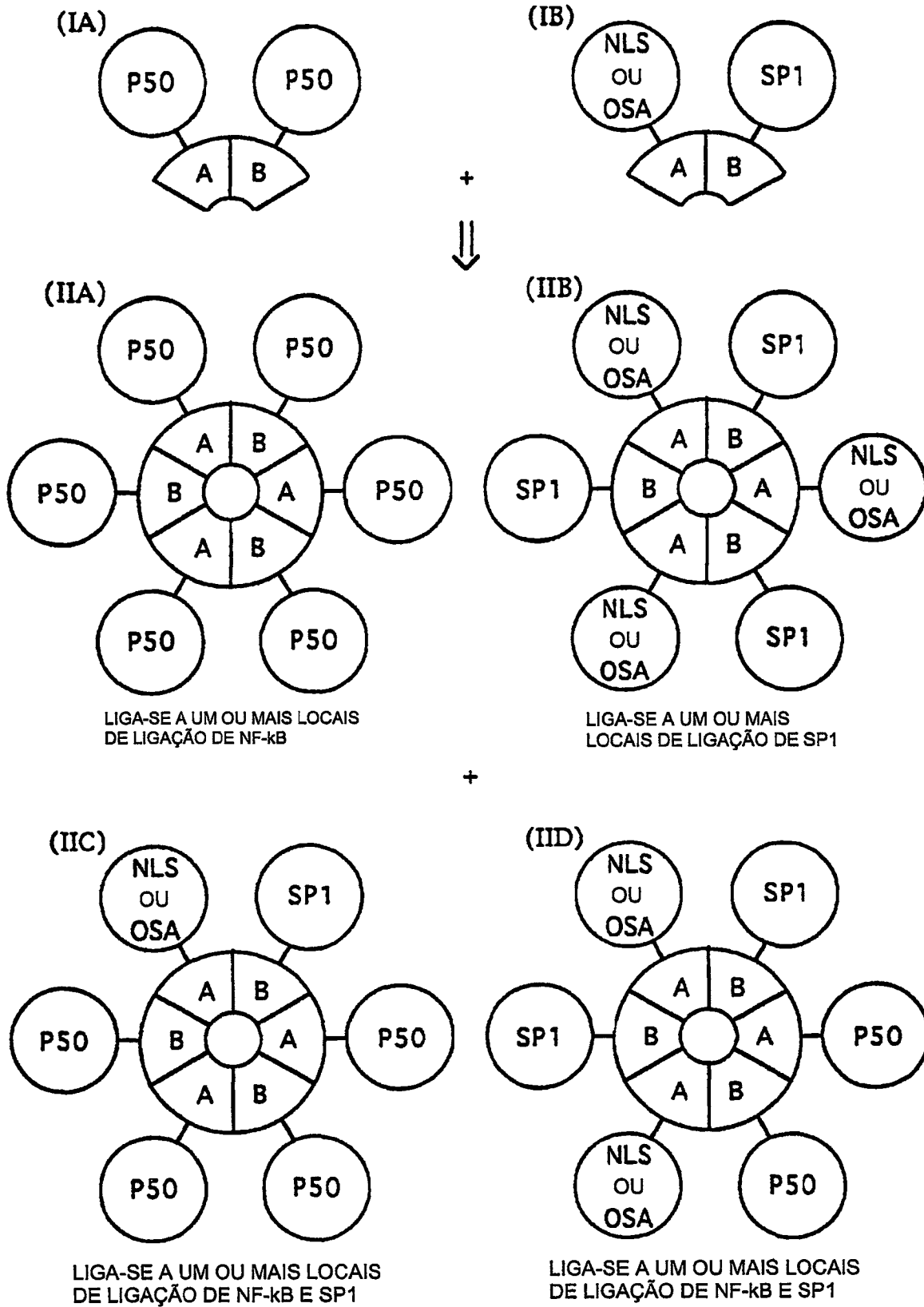


FIGURA 14

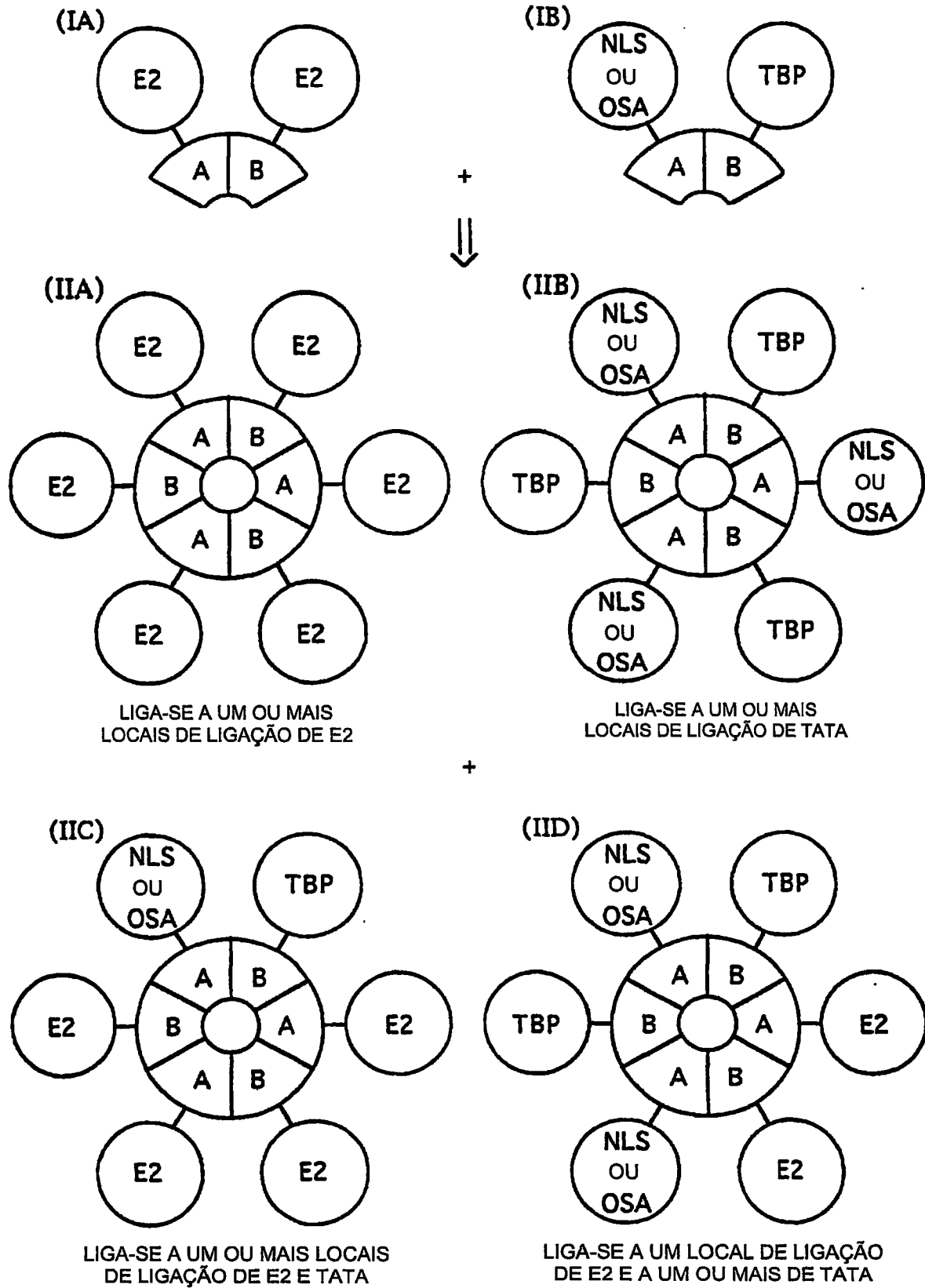


FIGURA 15

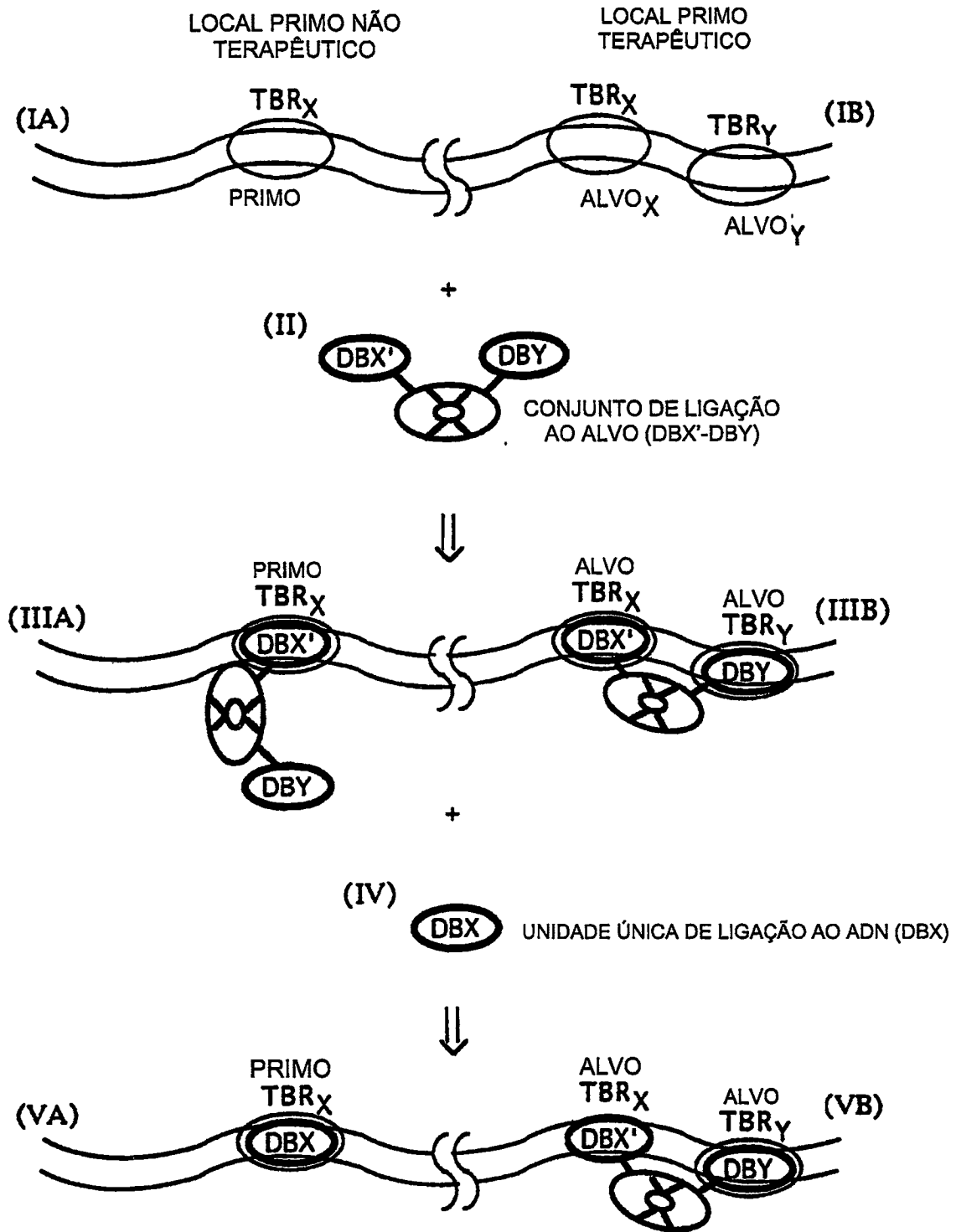


FIGURA 16

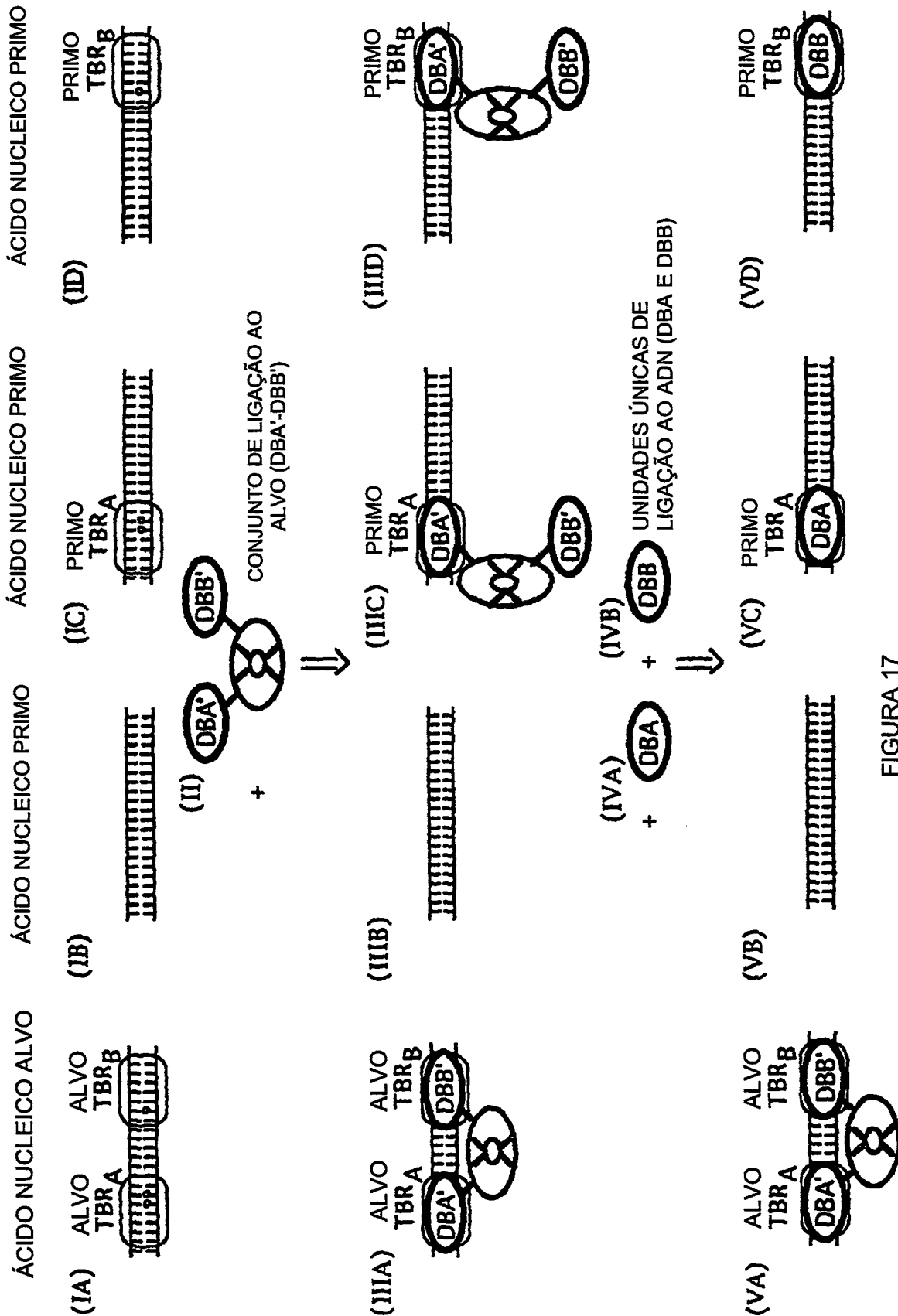


FIGURA 17