

(11) *Número de Publicação:* PT 850320 E

(51) *Classificação Internacional:* (Ed. 6)
C12Q001/68 A

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

<p>(22) <i>Data de depósito:</i> 1997.01.23</p> <p>(30) <i>Prioridade:</i> 1996.01.23 US 10436 P 1996.03.21 US 15402 P</p> <p>(43) <i>Data de publicação do pedido:</i> 1998.07.01</p> <p>(45) <i>Data e BPI da concessão:</i> 1999.12.09</p>	<p>(73) <i>Titular(es):</i> QIAGEN GENOMICS, INC 1725-220TH STREET S.E.SUITE 2000, BOTHELL WAHINGTON 98021 US</p> <p>(72) <i>Inventor(es):</i> JEFFREY VAN NESS JOHN C. TABONE JAMES JEFFRY HOWBERT JOHN T. MULLIGAN US US US US</p> <p>(74) <i>Mandatário(s):</i> FRANCISCO NOVAIS CUNHA BRITO MAIOR DE ATAYDE AVENIDA DUQUE D'ÁVILA, 32 1º AND. 1000 LISBOA PT</p>
---	--

(54) *Epígrafe:* MÉTODOS PARA A DETECÇÃO DA LIGAÇÃO NUM PAR DE LIGANTES COM SENSIBILIDADE AUMENTADA

(57) *Resumo:*

DIRECÇÃO DE SERVIÇOS DE PATENTES

CAMPO DAS CEBOLAS, 1100 LISBOA
TEL.: 888 51 51 / 2 / 3 LINHA AZUL 888 1078
TELEFAX: 87 53 08

FOLHA DO RESUMO

PAT. INV.	MOD. UTI.	MOD. IND.	DES. IND.	TOP. SEMIC.	Classificação Internacional (51)
<input type="checkbox"/>					

N.º 850 320 11

Data do pedido: ___ / ___ / ___

(22)

Requerente(s) (71) : (Nome e Morada)

RAPIGENE, INC.
1725, 220th Street S.E., Suite 200
Bothell, Washington 98021, E.U.A.

Código Postal _____

Inventores (72) :

Reivindicação de prioridade(s) (30)

Data do pedido	País de Origem	N.º de pedido
23.JAN.96	E.U.A.	10436 P
21.MAR.96	E.U.A.	15402 P

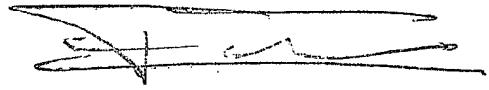
Figura (para interpretação do resumo)

Epígrafe: (54)

“MÉTODOS PARA A DETECÇÃO
DA LIGAÇÃO NUM PAR DE
LIGANTES COM SENSIBILIDADE
AUMENTADA”

Resumo: (máx. 150 palavras) (57)

A presente invenção refere-se a métodos para a detecção da ligação entre um primeiro membro e um segundo membro de um par de ligantes, compreendendo os passos de (a) combinar um conjunto de primeiros membros marcados com uma amostra biológica que poderá conter um ou mais segundos membros, sob condições e durante o tempo suficiente para permitir a ligação de um primeiro membro a um segundo membro, sendo que o dito marcador se correlaciona com um primeiro membro em particular e é detectável por espectrometria não fluorescente ou potenciometria; (b) separar os primeiro e segundo membros ligados dos membros não ligados; (c) separar o marcador do primeiro membro marcado; e (d) detectar o dito marcador por espectrometria não fluorescente ou potenciometria, e deste modo detectar a ligação do primeiro membro ao segundo membro.



DESCRIÇÃO

<u>EPÍGRAFE:</u>	"MÉTODOS PARA A DETECCÃO DA LIGAÇÃO NUM PAR DE LIGANTES COM SENSIBILIDADE AUMENTADA"
------------------	---

Campo técnico

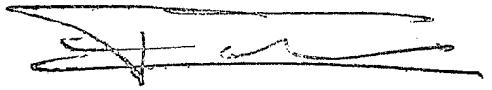
O presente invento relaciona-se de modo geral com métodos e composições para análise de moléculas de ácidos nucleicos, e mais especificamente com o uso de marcadores e linkers especializados que podem ser utilizados com o fim de aumentar a sensibilidade da análise de uma larga variedade de ensaios biológicos.

Antecedentes do invento

As técnicas de detecção e análise de moléculas de ácidos nucleicos são das mais importantes no campo da biologia. Estas técnicas desempenham um papel central na área da biologia molecular e a sua relevância nas restantes áreas da biologia cresce rapidamente.

Em geral, na sequência de essencialmente todas as reacções bioquímicas, a análise contempla sob alguma forma um passo de detecção. Neste âmbito, a detecção das hibridizações de ácidos nucleicos e da ligação antígeno-anticorpo são alvo de especial preocupação. Idealmente, a detecção deverá ser sensível e permitir o processamento de múltiplas amostras. No entanto, as técnicas actuais de detecção encontram-se algo limitadas no que respeita a ambas as características referidas.

A hibridização de moléculas de ácidos nucleicos é geralmente detectada por autoradiografia ou análise imanológica por fosforescência, quando a sonda de hibridização contém um marcador radioactivo, ou por densitometria quando a sonda de hibridização contém um marcador, tal como a biotina ou a digoxina, que é reconhecido por um anticorpo, ou ligando, acoplado a um enzima. Quando é utilizada uma sonda radiomarcada, a detecção por autoradiografia sofre de limitações inerentes à película, tais como a falha de reciprocidade e a não-linearidade. Estas limitações da película podem ser ultrapassadas através da detecção do marcador por análise imanológica por fosforescência. No entanto, os requisitos de segurança ligados ao uso de radiomarcadores obrigam a uma utilização aumentada dos recursos, à disponibilização de equipamento especializado e ao treino de



pessoal. Por estas razões, o uso de marcadores não radioactivos tem conhecido uma popularidade crescente. Nestes sistemas, os nucleótidos contêm um marcador, tal como a biotina ou a digoxina, que pode ser detectado por um anticorpo, ou outra molécula, marcado com um enzima que reage com um substrato cromogénico. Alternativamente, poderão usarse marcadores fluorescentes. Estes sistemas não levantam os problemas de segurança acima descritos, mas utilizam componentes que são frequentemente lábeis e que poderão dar origem a reacções inespecíficas, resultando num elevado efeito de fundo (i.e., baixa razão sinal-ruído).

As reacções de ligação anticorpo-antígeno podem ser detectadas através de diversos procedimentos. Tal como acontece com a hibridização de ácidos nucleicos, um marcador, radioactivo ou não radioactivo, é tipicamente conjugado com o anticorpo. Os tipos de marcadores são semelhantes: um enzima que reage com um substrato cromogénico, um marcador fluorescente, um haptene que é detectado por um ligando ou por outro anticorpo, e outros. Estes métodos de detecção apresentam limitações semelhantes às verificadas na detecção da hibridização de ácidos nucleicos.

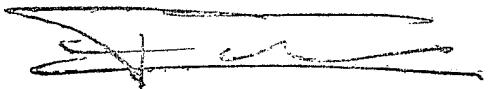
A Patente WO-A-9 apresenta um reagente marcador composto por uma porção molecular de analito, formada por dois resíduos de analito, ligada a uma porção molecular de marcador que possui um grupo repórter. É utilizada uma colecção destes reagentes para a sequenciação de ácidos nucleicos.

O presente invento fornece novas composições que podem ser utilizadas numa larga variedade de procedimentos baseados em ácidos nucleicos ou proteínas (p.ex. anticorpos), e proporciona ainda outras vantagens relacionadas.

Sumário do invento

Abreviadamente, o presente invento fornece composições e métodos que podem ser utilizados para aumentar a sensibilidade e o número de amostras processadas numa grande variedade de ensaios biológicos. Em particular, tomando como base os inventos aqui descritos, muitos ensaios que até ao presente têm requerido longos períodos de tempo para a sua execução podem agora ser realizados num tempo entre dez a mais do que cem vezes mais curto. Os métodos aqui descritos representam, assim, um acentuado e importante melhoramento em relação aos ensaios anteriormente disponíveis.

Por exemplo, num dos aspectos do invento é facultado um método para a detecção da ligação de um primeiro membro a um segundo membro de um par de ligandos, compreendendo os passos de:



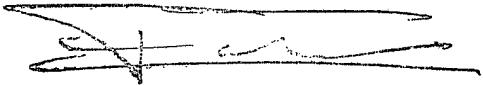
- (a) combinar um conjunto de primeiros membros marcados com uma amostra biológica que poderá conter um ou mais segundos membros, sob condições e durante o tempo suficiente para permitir a ligação de um primeiro membro a um segundo membro, sendo que o dito marcador se correlaciona com um primeiro membro em particular e é detectável por espectrometria ou potenciometria;
- (b) separar os primeiro e segundo membros ligados dos membros não ligados;
- (c) separar o dito marcador do dito primeiro membro marcado; e
- (d) detectar o dito marcador por espectrometria ou potenciometria, e deste modo detectar a ligação do dito primeiro membro ao dito segundo membro.

Poderá utilizar-se uma grande variedade de pares de primeiro e segundo membros no âmbito do presente invento, incluindo, por exemplo, moléculas de ácidos nucleicos (p. ex., DNA, RNA e análogos de ácidos nucleicos, como o PNA, ou qualquer combinação destes), proteínas ou polipeptídos (p.ex., um anticorpo ou fragmento de anticorpo; p.ex., um anticorpo monoclonal, um anticorpo policlonal ou um parceiro de ligação, como um CDR), oligossacáridos, hormonas, moléculas orgânicas e outros substratos (p.ex., xenobióticos como uma molécula de glucuronidase-fármaco), ou qualquer outro par de ligandos. No âmbito das várias formas de realização do invento, o primeiro e segundo membros poderão corresponder ao mesmo tipo de molécula ou serem de tipos diferentes. Por exemplo, os pares representativos de ligandos com primeiro membro e segundo membro incluem: molécula de ácido nucleico/molécula de ácido nucleico; anticorpo/molécula de ácido nucleico; anticorpo/hormona; anticorpo/xenobiótico; e anticorpo/proteína.

De preferência, o primeiro membro reconhecerá especificamente um segundo membro seleccionado (i.e., com exclusão de outras moléculas relacionadas), ou identificará como segundo membro uma classe de moléculas relacionadas (p.ex., uma classe de receptores relacionados). Preferencialmente, o primeiro membro ligar-se-á ao segundo membro com uma afinidade de pelo menos cerca de $10^{-5}/M$, e de preferência de $10^{-6}/M$, $10^{-7}/M$, $10^{-8}/M$, $10^{-9}/M$ ou superior a $10^{-12}/M$. A afinidade de uma primeira molécula para uma segunda molécula pode facilmente ser determinada por um técnico de perícia média (ver Scatchard, *Ann. N.Y. Acad. Sci.* 51:660-672, 1949).

Segundo outros aspectos do invento, é fornecido um método para a análise do padrão de expressão genética a partir de uma amostra biológica seleccionada, compreendendo os passos de:

- (a) expôr os ácidos nucleicos de uma amostra biológica;

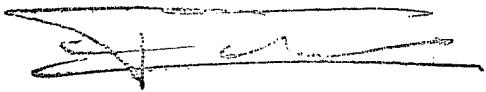


- (b) combinar os ditos ácidos nucleicos expostos com uma ou mais sondas marcadas de ácidos nucleicos seleccionadas, sob condições e durante o tempo suficiente para que as ditas sondas se hibridizem com os ditos ácidos nucleicos, sendo que o dito marcador se correlaciona com uma sonda de ácidos nucleicos em particular e é detectável por espectrometria ou potenciometria;
- (c) separar as sondas hibridizadas das não hibridizadas;
- (d) separar o dito marcador do dito fragmento marcado; e
- (e) detectar o dito marcador por espectrometria ou potenciometria, e desse modo determinar o padrão de expressão genética da dita amostra biológica.

Dentro do contexto do presente invento, deve considerar-se que a expressão "amostras biológicas" abrange não só as amostras provenientes de organismos vivos (i.e., mamíferos, peixes, bactérias, parasitas, vírus, fungos e semelhantes) ou do meio ambiente (i.e., ar, água ou amostras sólidas), como também os materiais biológicos que podem ser produzidos artificialmente ou por síntese (i.e., colecções de fagos, colecções de moléculas orgânicas, "pools" de clones genómicos e semelhantes). Os exemplos representativos de amostras biológicas incluem os fluidos biológicos (p.ex., sangue, sémen, líquido cefalorraquidiano, urina), células biológicas (p.ex., células pluripotentes, células B ou T, células hepáticas, fibroblastos e semelhantes) e tecidos biológicos.

No âmbito das diversas formas de realização dos métodos acima descritos, as sondas e/ou moléculas de ácidos nucleicos do presente invento poderão ser geradas através de, por exemplo, uma reacção de ligação, de corte ou de extensão (p.ex., PCR). Segundo outros aspectos, as sondas ou moléculas de ácidos nucleicos poderão ser marcadas a nível da sua extremidade 5', funcionando as moléculas assim marcadas como primers oligonucleotídicos ou terminadores didesoxinucleotídicos.

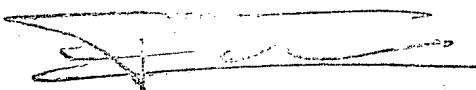
Noutras formas de realização do invento, poderão ser utilizadas em simultâneo numa dada reacção 4, 5, 10, 15, 20, 25, 30, 35, 40, 45, 50, 60, 70, 80, 90, 100, 200, 250, 300, 350, 400, 450 ou um número superior a 500 moléculas diferentes e marcadas com um marcador específico, em que cada marcador é único para um fragmento de ácido nucleico, sonda, primeiro membro ou segundo membro seleccionados e pode ser identificado em separado. Ainda em outras formas de realização do invento, o(s) marcador(es) pode(m) ser detectado(s) por fluorometria, espectrometria de massa, espectrometria de infravermelhos, espectrometria de ultravioleta ou amperometria potenciotáctica (p.ex., utilizando detectores



coulométricos ou amperométricos). Os exemplos representativos de técnicas espectrométricas adequadas incluem a espectrometria de massa de tempo-de-voo, espectrometria de massa quadrupolar, espectrometria de massa de sector magnético e espectrometria de massa de sector eléctrico. As formas específicas de realização destas técnicas incluem a espectrometria de massa por captura de iões, espectrometria de massa com ionização por electrospray, espectrometria de massa com spray de iões, espectrometria de massa com ionização em líquido, espectrometria de massa com ionização a pressão atmosférica, espectrometria de massa com ionização electrónica, espectrometria de massa de ionização por bombardeamento com átomos acelerados, espectrometria de massa MALDI, espectrometria de massa de tempo-de-voo com fotoionização, espectrometria de massa com micro-sonda de laser, espectrometria de massa MALDI-TOF, espectrometria de massa APCI, espectrometria de massa com nanospray, espectrometria de massa com ionização em spray nebulizado, espectrometria de massa com ionização química, espectrometria de massa com ionização por ressonância, espectrometria de massa com ionização secundária e espectrometria de massa com termospray.

No contexto de outras formas de realização do invento, os primeiro e segundo membros ligados ou os ácidos nucleicos expostos, ligados, poderão ser separados dos membros ou moléculas não ligados através de métodos tais como a electroforese em gel, electroforese capilar, electroforese em microcanais, HPLC, cromatografia de exclusão por dimensão, filtração, electroforese em gel de poliacrilamida, cromatografia líquida, cromatografia reversa de exclusão por dimensão, cromatografia de troca iônica, cromatografia líquida de fase reversa, electroforese em campo pulsado, electroforese de inversão de campo, diálise e separação de gotículas líquidas activadas por fluorescência. Alternativamente, o primeiro ou o segundo membro, ou os ácidos nucleicos expostos, poderão ser ligados a um suporte sólido [p.ex., fibras ocas (Amicon Corporation, Danvers, Mass.), esferas (Polysciences, Warrington, Pa.), esferas magnéticas (Robbin Scientific, Mountain View, Calif.), suportes, placas e matrizes (Corning Glass Works, Corning, N.Y.), crivos (Becton Dickinson, Mountain View, Calif.), ecrans e fibras sólidas (ver Edelman et al., Patente dos E.U.A. nº 3.843.324; ver também Kuroda et al., Patente dos E.U.A. nº 4.416.777), membranas (Millipore Corp., Bedford, Mass.) e tiras]. Se o primeiro ou o segundo membros, ou os ácidos nucleicos expostos, forem ligados a um suporte sólido, em certas formas de realização do invento os métodos aqui apresentados poderão ainda compreender o passo de lavagem para retirar o material não ligado do suporte sólido.

Em outras formas de realização, os primeiros membros marcados poderão ser removidos por métodos químicos, de oxidação, redução, labilidade a ácidos, labilidade a



bases, enzimáticos, electroquímicos, térmicos e de fotolabilidade. Ainda em outras formas de realização, os passos de separação, remoção e detecção poderão ser realizados de modo contínuo (como num fluxo contínuo); por exemplo, num só dispositivo que poderá ser automatizado.

Estes e outros aspectos do presente invento tornar-se-ão evidentes por consulta da descrição detalhada que se segue e das figuras em anexo. Adicionalmente, são aqui apresentadas várias referências que descrevem em maior detalhe certos procedimentos ou composições (p.ex., plasmídeos, etc.), e que são, deste modo, incorporadas integralmente para fins de referência.

Breve descrição das figuras

A Figura 1 representa o diagrama de síntese de ésteres de pentafluorofenilo de marcadores para espectrometria de massa removíveis por via química, libertando-se marcadores com terminações de amida carboxílica.

A Figura 2 representa o diagrama de síntese de ésteres de pentafluorofenilo de marcadores para espectrometria de massa removíveis por via química, libertando-se marcadores com terminações de ácido carboxílico.

As Figuras 3-6 e 8 representam os diagramas de síntese de ésteres de tetrafluorofenilo de um conjunto de 36 marcadores para espectrometria de massa removíveis por via fotoquímica.

A Figura 7 representa o diagrama de síntese de um conjunto de 36 marcadores para espectrometria de massa cujas moléculas possuem uma terminação de amina e são removíveis por via fotoquímica.

A Figura 9 representa a síntese de 36 oligonucleótidos marcados com marcadores para espectrometria de massa removíveis por via fotoquímica, produzidos a partir do conjunto correspondente de 36 ésteres de tetrafluorofenilo de marcadores ácidos para espectrometria de massa removíveis por via fotoquímica.

A Figura 10 representa a síntese de 36 oligonucleótidos marcados com marcadores para espectrometria de massa removíveis por via fotoquímica, produzidos a partir do conjunto correspondente de 36 marcadores para espectrometria de massa com uma terminação amina e removíveis por via fotoquímica.

A Figura 11 ilustra a detecção em simultâneo de múltiplos marcadores através de espectrometria de massa.

A Figura 12 apresenta o spectrograma de massa da matriz alfa-ciano isolada.

A Figura 13 representa um fragmento de ácido nucleico marcado construído de forma modular.

Descrição detalhada do invento

Tal como assinalado acima, o presente invento fornece métodos que compreendem o uso de marcadores e linkers que poderão ser utilizados para aumentar a sensibilidade e o número de amostras processadas numa grande variedade de ensaios biológicos. São abaixo descritos em maior detalhe marcadores e linkers representativos que poderão ser utilizados, a grande variedade de métodos a que poderão aplicar-se estes marcadores e os métodos adequados à detecção dos marcadores.

Resumidamente, o presente invento fornece métodos que compreendem o uso de compostos nos quais a molécula de interesse, ou um seu precursor, se encontra ligada por meio de uma ligação lábil (ou ligações lábeis) a um marcador. Assim, poderá definir-se para os ditos compostos a seguinte fórmula geral:

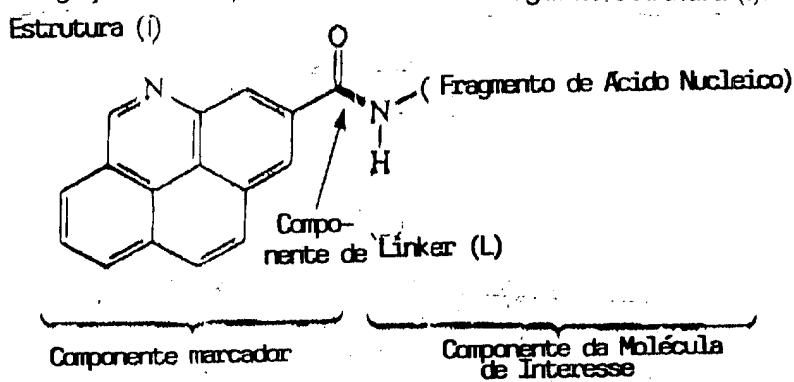


em que o T corresponde ao componente marcador, L é o componente de linker, o qual tanto poderá conter uma ligação lábil como corresponder à própria ligação, e X é o componente de molécula de interesse (MOI) ou um componente de grupo funcional (L_h) através do qual a MOI se pode ligar aos componentes T-L. Desta modo, os ditos compostos poderão ser representados através das fórmulas gerais mais específicas:



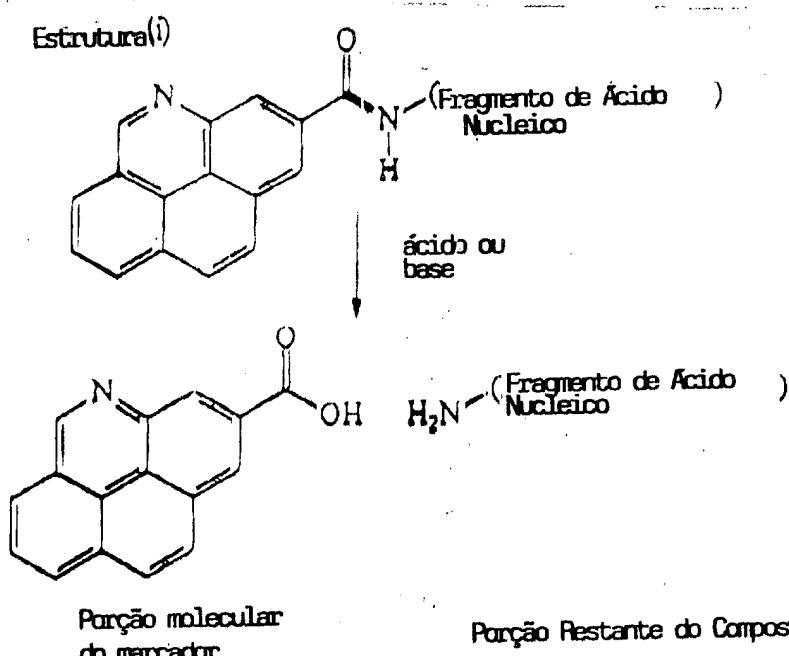
Por razões que são abaixo descritas em detalhe, os conjuntos de compostos T-L-MOI poderão ser propositadamente submetidos a condições que causem a quebra da(s) ligação(ões) lábil(eis), libertando-se deste modo uma porção molecular de marcador a partir da restante porção molecular do composto. A porção molecular de marcador é então caracterizada através de uma ou mais técnicas analíticas de modo a se obterem informações directas sobre a estrutura da porção molecular de marcador, e (o que é mais importante) informações indirectas sobre a identidade da MOI correspondente.

A título de simples exemplo ilustrativo de um composto representativo em que L corresponde a uma ligação directa, é feita referência à seguinte estrutura (i):

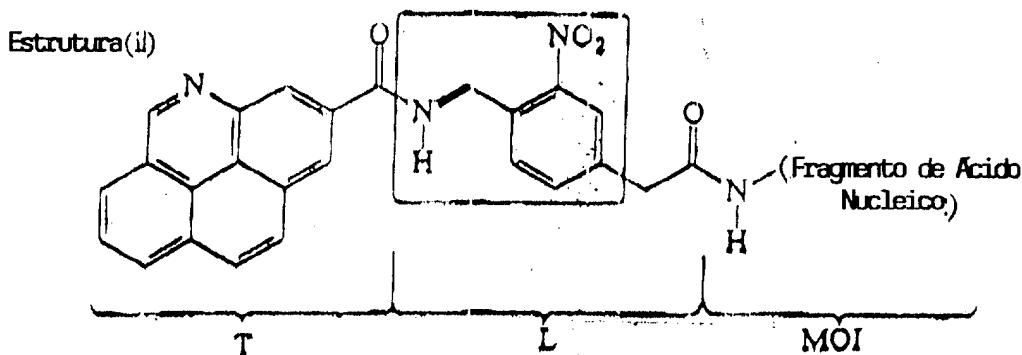


Silva

Na estrutura (i), T é uma molécula arornática policíclica contendo azoto que se encontra ligada a um grupo carbonilo, X é uma MOI (e especificamente um fragmento de ácido nucleico que termina num grupo amina) e L é a ligação que forma um grupo amida. A ligação amida é lábil relativamente às ligações presentes em T, uma vez que, tal como é conhecido da técnica, uma ligação amida pode ser quimicamente quebrada em condições básicas ou acídicas que deixam inalteradas as ligações no interior do componente marcador. Assim, poderá libertar-se uma porção molecular de marcador (i.e., o produto da reacção de quebra de ligações que contém T), tal como abaixo exposto:



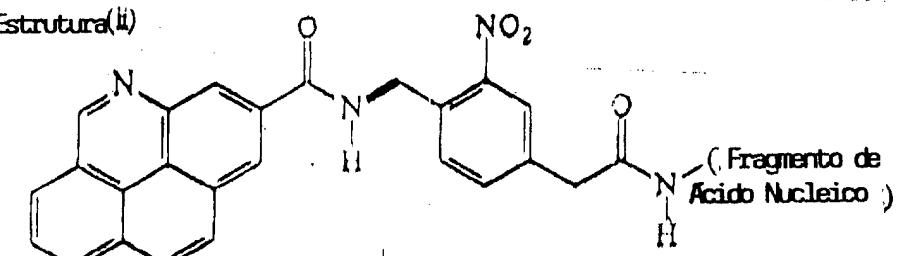
No entanto, o linker L poderá constituir mais do que uma ligação directa, tal como mostra o exemplo seguinte, em que é feita referência a outro composto representativo que possui a estrutura (ii) abaixo apresentada:



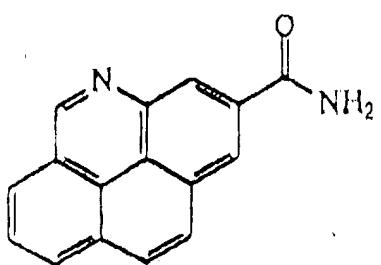
~~SECRETARIA DE ESTADO
DE SAÚDE
BRASIL~~

É bem conhecido que os compostos que possuem um grupo *ortho*-benzilamina (ver átomos contidos no quadrado no interior da estrutura (ii)) são fotoliticamente instáveis, já que a exposição de tais compostos a radiação actínica com um comprimento de onda específico causará o corte selectivo da ligação benzilamina (ver a ligação assinalada com uma linha mais carregada na estrutura (ii)). Assim, a estrutura (ii) possui os mesmos grupos T e MOI que a estrutura (i), mas o grupo de linker contém múltiplos átomos e ligações no interior dos quais existe uma ligação particularmente lábil. A fotólise da estrutura (ii) liberta, deste modo, uma porção molecular de marcador (porção molecular que contém T) da porção restante do composto, tal como mostrado abaixo.

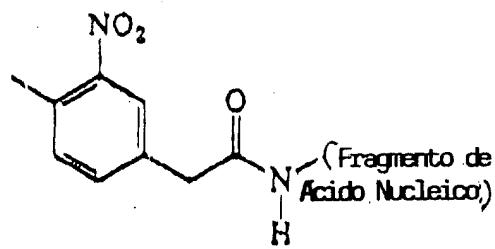
Estrutura (ii)



$\downarrow h\nu$



Porção Molecular do Marcador



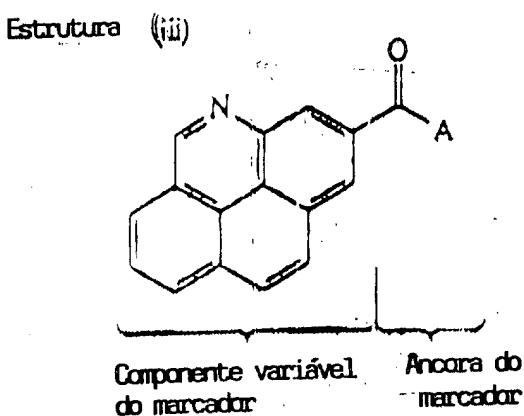
Porção Restante do Composto

Deste modo, o invento proporciona métodos que compreendem o uso de compostos que, por exposição a condições apropriadas à quebra de ligações, sofram reacções de quebra de ligações que libertam uma porção molecular de marcador da porção restante do composto. Os ditos compostos podem ser descritos em termos da porção molecular de marcador, da MOI (ou do seu precursor, L_h) e da(s) ligação(ões) lábil(eis) que unem os dois grupos. Alternativamente, estes compostos podem ser descritos em termos dos componentes a partir dos quais são formados. Assim, os compostos podem ser descritos como o produto de reacção de um reagente de marcador, de um reagente de linker e de um reagente de MOI, tal como se segue.

O reagente de marcador é constituído por uma "âncora" química (Th) e um componente variável (Tvc), de tal modo que é atribuída ao reagente de marcador a seguinte estrutura geral:

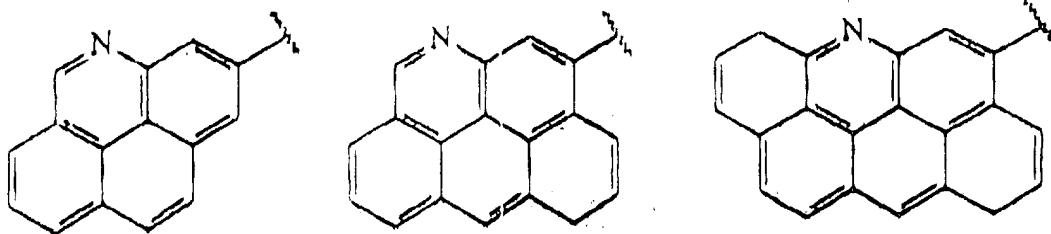
[Signature]
Tvc-Th

Para ilustrar esta nomenclatura, é feita referência à estrutura (iii), que apresenta um reagente de marcador que poderá ser usado para preparar o composto com a estrutura (ii). O reagente de marcador que possui a estrutura (iii) contém um componente variável do marcador e uma âncora do marcador, tal como apresentado abaixo:

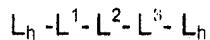


Na estrutura (iii), a âncora do marcador ($-C(=O)-A$) proporciona simplesmente uma via de reacção do reagente de marcador com o reagente de linker, de modo a formar uma porção molecular T-L. O grupo "A" na estrutura (iii) indica que o grupo carboxilo se encontra num estado quimicamente activo, pelo que se encontra pronto para se ligar a outras âncoras. "A" poderá corresponder, por exemplo, a um grupo hidroxilo ou a um grupo pentafluorofenoxy, entre muitas outras possibilidades. Existe um grande número de possibilidades para as âncoras de marcador que poderão ligar-se a um componente variável de marcador, tal como se expõe abaixo em detalhe. O componente variável de marcador constitui, assim, uma parte de "T" na fórmula T-L-X, e fará igualmente parte da porção molecular de marcador que se forma a partir da reacção que quebra L.

Tal como se expõe igualmente em detalhe abaixo, o componente variável de marcador é assim denominado porque, na preparação de conjuntos de compostos que serão usados nos métodos do invento, é desejável que os membros de um conjunto possuam componentes variáveis individualizados, de modo a que os membros individuais possam ser distinguidos entre si através de uma técnica analítica. Como exemplo, o componente variável do marcador da estrutura (iii) poderá corresponder a um membro do conjunto que se segue, sendo que os membros do conjunto podem diferenciar-se através do seu espectro em UV ou do espectro de massa:

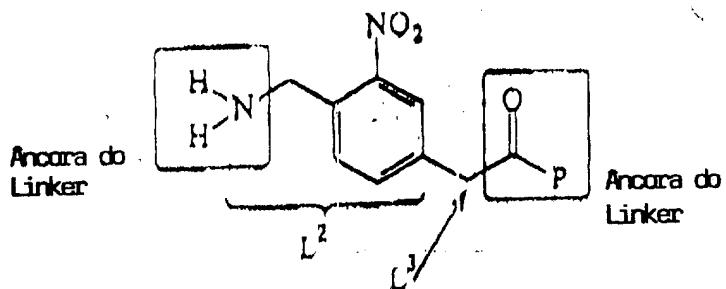


Do mesmo modo, o reagente de linker pode ser descrito em termos das suas âncoras químicas (são necessárias pelo menos duas, cada uma das quais se pode designar por L_h), as quais flanqueiam um componente lável do linker, sendo que o componente lável do linker consiste na porção lável requerida (L^2) e em porções lábeis opcionais (L^1 e L^3), que as porções lábeis opcionais servem para separar L^2 das âncoras L_h , e que a porção lável requerida se destina a proporcionar uma ligação lável no seio do componente lável do linker. Assim, poderá considerar-se que o reagente de linker possui a fórmula geral de:



A nomenclatura usada para descrever o reagente de linker pode ser ilustrada em função da estrutura (iv), a qual se baseia de novo no composto com a estrutura (ii):

Estrutura (iv)



Tal como ilustra a estrutura (iv), os átomos poderão possuir mais do que um papel funcional. Assim, na estrutura (iv), o azoto do grupo benzilo funciona como âncora química ao possibilitar a ligação entre o reagente de linker e o reagente de marcador através de uma reacção em que se forma uma amida, e posteriormente funciona também como uma parte necessária da estrutura da porção lável L^2 , uma vez que a ligação carbono-azoto do grupo benzilo é particularmente susceptível ao corte fotolítico. A estrutura (iv) ilustra igualmente o facto de que um reagente de linker poderá possuir um grupo L^3 (neste caso, um grupo metileno) ainda que não possua um grupo L^1 . Do mesmo modo, os reagentes de linker poderão possuir um grupo L^1 e não um grupo L^3 , ou poderão ter grupos L^1 e L^3 , ou poderão ainda não possuir grupos L^1 nem L^3 . Na estrutura (iv), a presença do grupo "P" junto ao

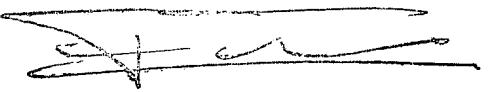


grupo carbonilo indica que o grupo carbonilo se encontra protegido da reacção. Dada esta configuração, o grupo carboxilo activado do reagente de marcador (iii) poderá facilmente reagir com o grupo amina do reagente de linker (iv) para formar uma ligação amida e originar um composto com a fórmula T-L-L_h.

O reagente de MOI corresponde a uma forma adequadamente reactiva de uma molécula de interesse. No caso de a molécula de interesse ser um fragmento de ácido nucleico, um reagente de MOI adequado será um fragmento de ácido nucleico ligado através do seu grupo 5' hidroxilo a um grupo fosfodiéster e em seguida a uma cadeia de alquíleno que termina num grupo amina. Este grupo amina poderá então reagir com o grupo carbonilo da estrutura (iv) (após, como é evidente, se ter procedido à desprotecção do grupo carbonilo, e de preferência após activar o grupo carbonilo para a reacção com o grupo amina), para, deste modo, se reunir a MOI ao linker.

Quando examinada por ordem cronológica, a preparação dos compostos do invento requer um reagente de marcador (possuindo uma âncora química do marcador e um componente variável do marcador), um reagente de linker (possuindo duas âncoras químicas do linker, uma porção lábil requerida e de 0 a 2 porções lábeis opcionais) e um reagente de MOI (possuindo um componente de molécula de interesse e uma âncora química da molécula de interesse) para formar a molécula T-L-MOI. Deste modo, para formar T-L-MOI, tanto se poderá fazer reagir inicialmente o reagente de marcador com o reagente de linker para formar T-L-L_h e de seguida fazer reagir o reagente de MOI com a molécula T-L-L_h para formar T-L-MOI, como (com menor grau de preferência) fazer reagir inicialmente o reagente de linker e o reagente de MOI para originar a molécula L_h-L-MOI e de seguida promover a reacção de L_h-L-MOI com o reagente marcador para formar T-L-MOI. Por uma questão de conveniência, os compostos que possuam a fórmula T-L-MOI serão descritos em termos do reagente de marcador, do reagente de linker e do reagente de MOI que tenham sido utilizados na formação destes compostos. Evidentemente, os mesmos compostos com a fórmula T-L-MOI poderão ser preparados por meio de outros métodos (tipicamente, mais laboriosos).

Em qualquer dos casos, o invento prevê que um composto T-L-MOI seja submetido a condições de quebra de ligações, de modo a que uma porção molecular de marcador se liberte da porção restante do composto. A porção molecular de marcador compreenderá pelo menos o componente variável do marcador, e tipicamente incluirá ainda alguns ou todos os átomos da âncora do marcador, alguns ou todos os átomos da âncora do linker que foi utilizada para reunir o reagente de marcador ao reagente de linker, a porção lábil opcional L¹ se este grupo estiver presente em T-L-MOI, e poderá também conter uma parte



da porção lábil requerida L^2 , dependendo da própria estrutura de L^2 e da natureza da reacção química de quebra de ligações. Por uma questão de conveniência, a porção molecular de marcador poderá ser referida como a porção molecular contendo T, uma vez que T constituirá tipicamente a porção principal (em termos de massa) da porção molecular de marcador.

Após esta introdução aos componentes a usar em associação aos métodos do presente invento, os vários componentes T, L e X serão descritos em detalhe. Esta descrição inicia-se com as definições de certos termos que serão doravante utilizados para descrever T, L e X.

Tal como aqui é usado, o termo "fragmento de ácido nucleico" designa uma molécula que é complementar a uma molécula-alvo de ácido nucleico seleccionada (i.e., complementar a toda a molécula ou a uma porção desta), e que pode derivar da natureza ou ser produzida por via sintética ou recombinante, incluindo moléculas que não ocorrem na natureza, e que poderão encontrar-se sob a forma de cadeia simples ou dupla, segundo apropriado; o termo inclui um oligonucleótido (DNA ou RNA), um primer, uma sonda, um análogo de ácido nucleico (p.ex., PNA), um oligonucleótido que sofre uma extensão na direcção 5' para 3' por acção de uma polimerase, um ácido nucleico que é cortado quimicamente ou enzimaticamente, um ácido nucleico que é terminado por um terminador didesoxi ou recebe na extremidade 3' ou 5' uma cauda de um composto que impede a polimerização a nível da extremidade 5' ou 3', e combinações destes elementos. A complementaridade de um fragmento de ácido nucleico para uma molécula-alvo de ácido nucleico seleccionada significa geralmente que o emparelhamento específico de bases corresponde no mínimo a 70% ao longo de toda a extensão do fragmento. De preferência, o fragmento de ácido nucleico exibe um emparelhamento específico de bases de pelo menos 80%; e mais preferencialmente de pelo menos cerca de 90%. Os ensaios destinados a determinar a percentagem de não emparelhamento (e, deste modo, a percentagem de emparelhamento específico de bases) são bem conhecidos da técnica e baseiam-se na percentagem de não emparelhamento como função do T_m por comparação com o controlo que possui um emparelhamento total de bases.

Tal como aqui é usado, o termo "alquilo", aplicado isoladamente ou em combinação, refere-se a um radical de hidrocarboneto saturado de cadeia linear ou ramificada contendo de entre 1 a 10, de preferência de 1 a 6, e mais preferencialmente de 1 a 4 átomos de carbono. Os exemplos de tais radicais incluem, de modo não limitativo, os radicais metilo, etilo, n-propilo, iso-propilo, n-butilo, iso-butilo, sec-butilo, tert-butilo, pentilo, iso-amilo, hexilo, decilo e semelhantes. O termo "alquileno" refere-se a um di-radical de hidrocarboneto

saturado de cadeia linear ou ramificada contendo de entre 1 a 10, de preferência de 1 a 6, e mais preferencialmente de 1 a 4 átomos de carbono. Os exemplos de tais di-radicais incluem, de modo não limitativo, os di-radicais metileno, etileno ($-\text{CH}_2\text{-CH}_2-$), propileno e semelhantes.

O termo “alquenilo”, aplicado isoladamente ou em combinação, refere-se a um radical de hidrocarboneto de cadeia linear ou ramificada que possui pelo menos uma ligação dupla carbono-carbono num total de 2 a 10, de preferência de 2 a 6, e mais preferencialmente de 2 a 4 átomos de carbono. Os exemplos de tais radicais incluem, de forma não limitativa, os radicais etenilo, E- e Z-propenilo, isopropenilo, E- e Z-butenilo, E- e Z-isobutenilo, E- e Z-pentenilo, decenilo e semelhantes. O termo “alquenileno” refere-se a um di-radical de hidrocarboneto de cadeia linear ou ramificada que possui pelo menos uma ligação dupla carbono-carbono num total de 2 a 10, de preferência de 2 a 6, e mais preferencialmente de 2 a 4 átomos de carbono. Os exemplos de tais di-radicais incluem, de modo não limitativo, os di-radicais metilideno ($=\text{CH}_2$), etilideno ($-\text{CH}=\text{CH}-$), propilideno ($-\text{CH}_2\text{-CH}=\text{CH-}$) e semelhantes.

O termo “alquinilo”, utilizado isoladamente ou em combinação, refere-se a um radical de hidrocarboneto de cadeia linear ou ramificada que possui pelo menos uma ligação tripla carbono-carbono num total de 2 a 10, de preferência de 2 a 6, e mais preferencialmente de 2 a 4 átomos de carbono. Os exemplos de tais radicais incluem, de modo não limitativo, os radicais etinilo (acetilenilo), propinilo (propargilo), butinilo, hexinilo, decinilo e semelhantes. O termo “alquinileno”, usado isoladamente ou em combinação, refere-se a um di-radical de hidrocarboneto de cadeia linear ou ramificada que possui pelo menos uma ligação tripla carbono-carbono num total de 2 a 10, de preferência de 2 a 6, e mais preferencialmente de 2 a 4 átomos de carbono. Os exemplos de tais di-radicais incluem, de modo não limitativo, os di-radicais etinileno ($-\text{C}\equiv\text{C-}$), propinileno ($-\text{CH}_2\text{-C}\equiv\text{C-}$) e semelhantes.

O termo “cicloalquilo”, usado isoladamente ou em combinação, refere-se a um arranjo saturado e cíclico de átomos de carbono compreendendo de entre 3 a 8, e preferencialmente de 3 a 6, átomos de carbono. Os exemplos de tais radicais cicloalquílicos incluem, de forma não limitativa, os radicais ciclopropilo, ciclobutilo, ciclopentilo, ciclohexilo e semelhantes. O termo “cicloalquileno” refere-se a uma forma di-radical de um cicloalquilo.

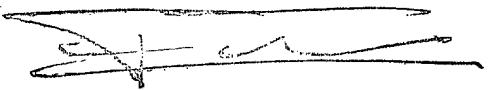
O termo “cicloalquenilo”, aplicado isoladamente ou em combinação, refere-se a um radical cíclico de carbono contendo de entre 4 a 8, e preferencialmente 5 ou 6, átomos de carbono e uma ou mais ligações duplas. Os exemplos de tais radicais cicloalquenílicos incluem, de forma não limitativa, os radicais ciclopentenilo, ciclohexenilo, ciclopentadienilo e

semelhantes. O termo "cicloalquenileno" refere-se a uma forma di-radical de um cicloalquenilo.

O termo "arilo" refere-se a um grupo carbocíclico (consistindo inteiramente em carbono e hidrogénio) aromático, seleccionado a partir do grupo formado pelos radicais fenilo, naftilo, indenilo, indanilo, azulenilo, fluorenilo e antracenilo; ou a um grupo heterocíclico aromático seleccionado a partir do grupo formado pelos radicais furilo, tienilo, piridilo, pirrolilo, oxazolilo, tiazolilo, imidazolilo, pirazolilo, 2-pirazolinilo, pirazolidinilo, isoxazolilo, isotiazolilo, 1,2,3-oxadiazolilo, 1,2,3-triazolilo, 1,3,4-tiodiazolilo, piridazinilo, pirimidinilo, pirazinilo, 1,3,5-triazinilo, 1,3,5-tritianilo, indolizinilo, indolilo, isoindolilo, 3H-indolilo, indolinilo, benzo[b]furanilo, 2,3-dihidrobenzofuranilo, benzo[b]tiofenilo, 1H-indazolilo, benzimidazolilo, benziazolilo, purinilo, 4H-quinolizinilo, quinolinilo, isoquinolinilo, cinolinilo, ftalazinilo, quinazolinilo, quinoxalinilo, 1,8-naftiridinilo, pteridinilo, carbazolilo, acridinilo, fenazinilo, fenotiazinilo e fenoxazinilo.

Os grupos "aryl", tal como definido neste pedido de patente, poderão conter independentemente de um a quatro substituintes que são seleccionados independentemente a partir do grupo formado por hidrogénio, halogénio, hidroxilo, amino, nitro, trifluorometil, trifluorometoxi, alquilo, alquenilo, alquinilo, ciano, carboxi, carboalcoxi, 1,2-dioxietileno, alcoxi, alquenoxi ou alquinoxi, alquilamino, alquenilamino, alquinilamino, acilo aromático ou alifático, alcoxi-carbonilamino, alquilsulfonilamino, morfolinocarbonilamino, tiomorfolinocarbonilamino, N-alquilguanidino, aralquilaminosulfonilo; aralcoxialquilo; N-aralcoxiureia; N-hidroxilureia; N-alquenilureia; N,N-(alquil, hidroxil)ureia; heterociclico; arilo tioariloxi-substituído; N,N-(aril, alquil)hidrazino; sulfonilheterociclico Ar'-substituído; heterociclico aralquil-substituído; heterociclico cicloalquilo-substituído e cicloalquenilo-substituído; arilo cicloalquil-fundido; alquilo ariloxi-substituído; heterociclikamino; acilaminocarbonilo alifático ou aromático; alquenilo alifático ou aromático acil-substituído; aminocarboniloxi Ar'-substituído; arilo Ar', Ar'-disubstituído; acilo alifático ou aromático acil-substituído; cicloalquilcarbonilalquilo; amino cicloalquil-substituído; ariloxicarbonilalquilo; ácido ou éster fosforodiamídílico.

"Ar" é um grupo arilo carboxílico ou heterocíclico tal como definido acima, possuindo de um a três substituintes seleccionados a partir do grupo constituído por hidrogénio, halogénio, hidroxilo, amino, nitro, trifluorometilo, trifluorometoxi, alquilo, alquenilo, alquinilo, 1,2-dioximetileno, 1,2-dioxietileno, alcoxi, alquenoxi, alquinoxi, alquilamino, alquenilamino ou alquinilamino, alquilcarboniloxi, acilo alifático ou aromático, alquilcarbonilamino, alcoxycarbonilamino, alquilsulfonilamino, N-alquil-ureia ou N,N-dialquilureia.



O termo "alcoxi", usado isoladamente ou em combinação, refere-se a um radical de éter de alquilo, tendo o termo "alquilo" sido já acima definido. Os exemplos de radicais de éter de alquilo adequados incluem, de modo não limitativo, os radicais metoxi, etoxi, n-propoxi, isso-propoxi, n-butoxi, isso-butoxi, sec-butoxi, tert-butoxi e semelhantes.

O termo "alquenoxi", utilizado isoladamente ou em combinação, refere-se a um radical com a fórmula alquenil-O, em que o termo "alquenilo" corresponde à definição acima, com a ressalva de que o radical não corresponda a um éter de enol. Os exemplos de radicais alquenoxi adequados incluem, de modo não limitativo, os radicais aliloxi, E- e Z-3-metil-2-propenoxi e semelhantes.

O termo "alquiniloxi", utilizado isoladamente ou em combinação, refere-se a um radical com a fórmula alquinil-O, em que o termo "alquinilo" corresponde à definição acima, com a ressalva de que o radical não corresponda a um éter de inol. Os exemplos de radicais alquinoxí adequados incluem, de modo não limitativo, os radicais propargiloxi, 2-butiniloxi e semelhantes.

O termo "tioalcoxi" refere-se a um radical de tioéter com a fórmula alquil-S-, em que o alquilo corresponde à definição acima apresentada.

O termo "alquilamino", usado isoladamente ou em combinação, refere-se a um radical amino mono- ou di-alquilsustituído (i.e., um radical com a fórmula alquil-NH- ou (alquil)₂-N-), em que o termo "alquilo" corresponde à definição acima. Os exemplos de radicais alquilamino adequados incluem, de modo não limitativo, os radicais metilamino, etilamino, propilamino, isopropilamino, t-butilamino, N,N-dietylamino e semelhantes.

O termo "alquenilamino", aplicado isoladamente ou em combinação, refere-se a um radical com a fórmula alquenil-NH- ou (alquenil)₂N-, em que o termo "alquenilo" corresponde à definição acima, com a ressalva de que o radical não corresponda a uma enamina. Um exemplo de tais radicais alquenilamino é dado pelo radical alilamino.

O termo "alquinilamino", usado isoladamente ou em combinação, refere-se a um radical com a fórmula alquinil-NH- ou (alquinil)₂N-, em que o termo "alquinilo" corresponde à definição acima, com a ressalva de que o radical não corresponda a uma inamina. Um exemplo de tais radicais alquinilamino é dado pelo radical propargilamino.

O termo "amida" refere-se a radicais $-N(R')-C(=O)-$ ou $-C(=O)-N(R')$, em que R' é aqui definido como incluindo hidrogénio, assim como outros grupos. O termo "amida substituída" refere-se à situação em que R' não é o hidrogénio, enquanto que o termo "amida não substituída" se refere à situação em que R' é o hidrogénio.

O termo "ariloxi", usado isoladamente ou em combinação, refere-se a um radical de fórmula aril-O, em que o aril corresponde à definição acima. Os exemplos de radicais ariloxi



adequados incluem, de modo não limitativo, os radicais fenoxi, naftoxi, piridiloxi e semelhantes.

O termo "arilamino", aplicado isoladamente ou em combinação, refere-se a um radical com a fórmula aril-NH-, em que o aril corresponde à definição acima. Os exemplos de radicais arilamino adequados incluem, de modo não limitativo, os radicais fenilamino (anilido), naftilamino, 2-, 3- e 4-piridilamino e semelhantes.

O termo "cicloalquilo aril-fundido", usado isoladamente ou em combinação, refere-se a um radical cicloalquilo que partilha dois átomos adjacentes com um radical arilo, sendo que os termos "cicloalquilo" e "arilo" correspondem às definições acima apresentadas. Um exemplo de um radical cicloalquilo aril-fundido é dado pelo radical ciclobutilo benzofundido.

O termo "alquilcarbonilamino", usado isoladamente ou em combinação, refere-se a um radical com a fórmula alquil-CONH, em que o termo "alquilo" corresponde à definição acima exposta.

O termo "alcoxicarbonilamino", usado isoladamente ou em combinação, refere-se a um radical com a fórmula álquil-OCONH-, em que o termo "alquilo" corresponde à definição acima exposta.

O termo "alquilsulfonilamino", usado isoladamente ou em combinação, refere-se a um radical com a fórmula alquil-SO₂NH-, em que o termo "alquilo" corresponde à definição acima exposta.

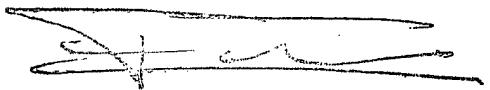
O termo "arilsulfonilamino", usado isoladamente ou em combinação, refere-se a um radical com a fórmula aril-SO₂NH-, em que o termo "arilo" corresponde à definição acima exposta.

O termo "N-alquilureia", usado isoladamente ou em combinação, refere-se a um radical com a fórmula alquil-NH-CO-NH-, em que o termo "alquilo" corresponde à definição acima exposta.

O termo "N-arilureia", usado isoladamente ou em combinação, refere-se a um radical com a fórmula aril-NH-CO-NH-, em que o termo "arilo" corresponde à definição acima exposta.

O termo "halogénio" designa os radicais fluoreto, cloreto, brometo e iodeto.

O termo "radical de hidrocarboneto" refere-se a um arranjo de átomos de carbono e hidrogénio que necessita apenas de um átomo de hidrogénio para se tornar numa molécula estável independente. Assim, um radical de hidrocarboneto possui um local de valência aberta sobre um átomo de carbono, através do qual o radical de hidrocarboneto se pode ligar a outro(s) átomo(s). Os radicais alquilo, alquenilo, cicloalquilo, etc., constituem exemplos de radicais de hidrocarboneto.



O termo “di-radical de hidrocarboneto” refere-se a um arranjo de átomos de carbono e hidrogénio que necessita de dois átomos de hidrogénio para se tornar numa molécula estável independente. Assim, um di-radical de hidrocarboneto possui dois locais de valência aberta sobre um ou dois átomos de carbono, através do qual o di-radical de hidrocarboneto se pode ligar a outro(s) átomo(s). Os radicais alquíleno, alquenileno, alquinileno, cicloalquíleno, etc., constituem exemplos de di-radicais de hidrocarboneto.

O termo “hidrocarbilo” refere-se a qualquer arranjo estável constituído inteiramente por carbono e hidrogénio e com um só local de valência através do qual se liga a outra porção molecular, incluindo, desta forma, radicais conhecidos como alquilo, alquenilo, alquinilo, cicloalquilo, cicloalquenilo, arilo (sem incorporação de heteroátomos no anel de arilo), arilalquilo, alquilarilo e semelhantes. “Radical de hidrocarboneto” constitui outro nome para hidrocarbilo.

O termo “hidrocarbíleno” refere-se a qualquer arranjo estável constituído inteiramente por carbono e hidrogénio e com dois locais de valência através do qual se liga a outras porções moleculares, incluindo, desta forma, os radicais alquíleno, alquenileno, alquinileno, cicloalquíleno, cicloalquenileno, arileno (sem incorporação de heteroátomos no anel de arileno), arilalquíleno, alquilarileno e semelhantes. “Di-radical de hidrocarboneto” constitui outro nome para hidrocarbíleno.

O termo “hidrocarbil-O-hidrocarbíleno” refere-se a um grupo hidrocarbilo ligado a um átomo de oxigénio, em que o átomo de oxigénio se encontra igualmente ligado a um grupo hidrocarbíleno a nível de um dos dois locais de valência através dos quais o grupo hidrocarbíleno se liga a outras porções moleculares. Os termos “hidrocarbil-S-hidrocarbíleno”, “hidrocarbil-NH-hidrocarbíleno” e “hidrocarbil-amida-hidrocarbíleno” têm significados equivalentes, sendo o oxigénio substituído por enxofre, -NH- ou um grupo amida, respectivamente.

O termo N-(hidrocarbil)hidrocarbíleno refere-se a um grupo hidrocarbíleno em que um dos dois locais de valência se encontra ligado a um átomo de azoto, estando este átomo de azoto simultaneamente ligado a um hidrogénio e a um grupo hidrocarbilo. O termo N,N-di(hidrocarbil)hidrocarbíleno refere-se a um grupo de hidrocarbíleno em que um dos dois locais de valência se encontra ligado a um átomo de azoto, estando este átomo de azoto simultaneamente ligado a dois grupos hidrocarbilo.

O termo “hidrocarbilacil-hidrocarbíleno” refere-se a um grupo hidrocarbilo ligado através de um grupo acilo (-C(=O)-) a um de dois locais de valência de um grupo hidrocarbíleno.



Os termos "heterociclilhidrocarbilo" e "heterociclilo" referem-se a um arranjo cílico estável de átomos que inclui átomos de carbono e até quatro átomos (designados por heteroátomos) de oxigénio, azoto, fósforo ou enxofre. O arranjo cílico poderá encontrar-se sob a forma de um anel monocíclico de 3-7 átomos, ou de um anel bicíclico de 8-11 átomos. Os anéis poderão ser saturados ou insaturados (incluindo anéis aromáticos), e opcionalmente poderão ser benzofundidos. Os átomos de azoto e enxofre presentes no anel poderão encontrar-se sob qualquer forma oxidada, incluindo a forma quaternária do azoto. Poderá existir uma ligação entre um heterociclilhidrocarbilo e qualquer carbono, ou heteroátomo, endocíclico que resulte na criação de uma estrutura estável. Os heterociclilhidrocarbilos preferidos incluem heterociclos monocíclicos de 5-7 membros contendo um ou dois heteroátomos de azoto.

O termo "heterociclilhidrocarbilo substituído" refere-se a um heterociclilhidrocarbilo tal como acima definido, em que pelo menos um átomo do anel se encontra ligado a um substituinte indicado que se prolonga para fora do anel.

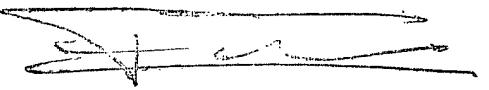
Em referência a grupos hidrocarbilo e hidrocarbileno, o termo "derivados de qualquer das moléculas seguintes em que ocorre a substituição de um ou mais hidrogénios por igual número de átomos de flúor" refere-se a moléculas que contêm átomos de carbono, hidrogénio e flúor mas não contêm nenhum outro átomo.

O termo "éster activado" designa um éster contendo um "grupo de partida" que é facilmente deslocado por um nucleófilo, tal como uma amina, um álcool ou um tiol. Tais grupos de partida são bem conhecidos e incluem, de modo não limitativo, N-hidroxisuccinimida, N-hidroxibenzotriazol, halogénios (haletos), grupos alcoxi, incluindo tetrafluorofenolatos, grupos tioalcoxi e semelhantes. O termo "éster protegido" refere-se a um grupo éster que se encontra oculto ou se mostra de outra forma não reactivo. Ver, p.ex., Greene, "Protecting Groups In Organic Solutions".

Tendo em consideração as definições acima, outros termos químicos que serão usados ao longo deste pedido de patente poderão ser facilmente compreendidos pelos peritos na técnica. Os termos poderão ser usados isoladamente ou em qualquer combinação pretendida. As extensões preferidas e mais preferidas para as cadeias dos radicais aplicam-se a todas estas combinações.

A. CRIAÇÃO DE FRAGMENTOS DE ÁCIDO NUCLEICO MARCADOS

Tal como assinalado acima, um dos aspectos do presente invento fornece um esquema geral para a sequenciação de DNA que permite o uso de mais do que 16 marcadores em cada carreira; com uma detecção contínua, os marcadores poderão ser



detectados e a sequência lida à medida que ocorre a separação por dimensões, tal como acontece com a sequenciação convencional baseada em fluorescência. Este esquema é aplicável a qualquer uma das técnicas de sequenciação de DNA que se baseiam na separação por dimensões das moléculas marcadas. São abaixo expostos em maior detalhe marcadores e linkers apropriados para utilização com o presente invento, assim como métodos para a sequenciação de ácidos nucleicos.

1. Marcadores

O termo "marcador", tal como aqui é usado, refere-se genericamente a uma espécie química que é usada para identificar de modo individual uma "molécula de interesse", e refere-se mais especificamente ao componente variável do marcador, assim como a qualquer parcela do reagente de marcador, do componente de marcador ou da porção molecular do marcador que possa estar estreitamente ligada a este componente variável.

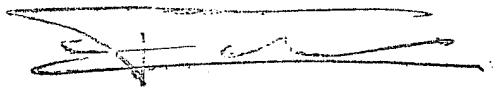
Um marcador útil no âmbito do presente invento possui diversos atributos:

1) Tem a capacidade de se distinguir de todos os outros marcadores. Esta discriminação a partir de outras espécies químicas poderá basear-se no comportamento cromatográfico do marcador (particularmente após a reacção de corte de ligações), nas suas propriedades espectroscópicas ou potenciométricas ou numa combinação destes factores. Os métodos espectroscópicos através dos quais se podem distinguir os marcadores incluem a espectroscopia de massa (MS), de infravermelhos (IR), de ultravioleta (UV) e de fluorescência, sendo o MS, o IR e o UV os métodos espectroscópicos preferidos e o MS o mais preferido. A amperometria potenciométrica constitui um método potenciométrico preferido.

2) O marcador tem a capacidade de ser detectado quando se encontra presente em concentrações de 10^{-22} a 10^{-6} mole.

3) O marcador possui uma âncora química através da qual se pode ligar à MOI que se pretende que seja identificada de modo individual pelo marcador. A ligação poderá ser estabelecida directamente entre o marcador e a MOI, ou indirectamente através de um grupo de "linker".

4) O marcador é quimicamente estável face a todas as manipulações a que é submetido, incluindo as reacções de ligação à MOI e de separação desta, mantendo-se igualmente estável face a quaisquer manipulações da MOI enquanto a esta se mantém ligado.



5) O marcador não interfere de modo significativo com as manipulações da MOI enquanto se mantém ligado a esta. Por exemplo, se o marcador estiver ligado a um oligonucleótido, o marcador não interferirá de modo significativo com quaisquer reacções de hibridização ou enzimáticas (p.ex., reacções de sequenciação por PCR) que sejam realizadas sobre o oligonucleótido. De forma similar, se o marcador se encontrar ligado a um anticorpo, não interferirá de modo significativo com o reconhecimento do antígeno por parte do anticorpo.

Uma porção molecular de marcador que se destine a ser detectada através de um dado método espectrométrico ou potenciométrico deverá possuir propriedades que aumentem a sensibilidade e a especificidade de detecção através desse método. Tipicamente, a porção molecular de marcador possuirá estas propriedades porque estas fizeram parte da concepção do componente variável do marcador, o qual constitui tipicamente a parte principal da porção molecular de marcador. Na exposição que se segue, o uso da palavra "marcador" refere-se tipicamente à porção molecular de marcador (i.e., o produto de quebra de ligações que contém o componente variável do marcador); no entanto, este termo pode também considerar-se como designando o próprio componente variável do marcador, uma vez que esta parte da porção molecular de marcador é tipicamente responsável pelas propriedades de detectabilidade individualizada. Nos compostos com a fórmula T-L-X, a porção "T" contém o componente variável de marcador. Nos casos em que o componente variável de marcador foi concebido para ser caracterizado por, p.ex., espectrometria de massa, a porção "T" da T-L-X poderá ser designado por T^{MS} . Do mesmo modo, o produto da quebra de ligações de T-L-X que contém T poderá ser designado por porção molecular que contém T^{MS} . Poderão utilizar-se os métodos espectroscópicos e potenciométricos que se seguem para caracterizar as porções moleculares que contêm T^{MS} .

a. Características dos Marcadores MS

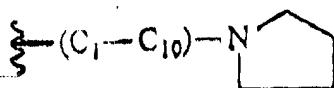
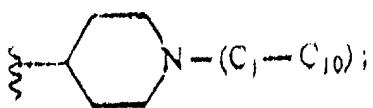
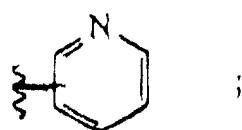
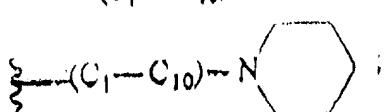
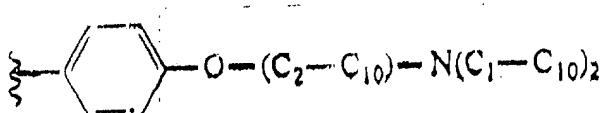
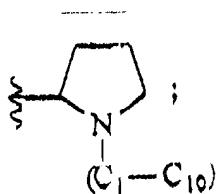
Quando um marcador é analisável por espectrometria de massa (i.e., é um marcador legível por MS, aqui também designado por marcador MS ou "porção molecular contendo T^{MS} "), a característica principal deste marcador é a de que pode ser ionizado. Deste modo, constitui um elemento preferido na concepção de marcadores legíveis por MS a incorporação nestes de uma funcionalidade química que poderá exibir uma carga positiva ou negativa sob condições de ionização no MS. Esta característica proporciona a melhoria da eficiência de formação de iões e uma maior sensibilidade global de detecção, particularmente na ionização por electrospray. A funcionalidade química que suporta uma carga ionizada poderá derivar de T^{MS} , ou de L, ou de ambos. Os factores que poderão

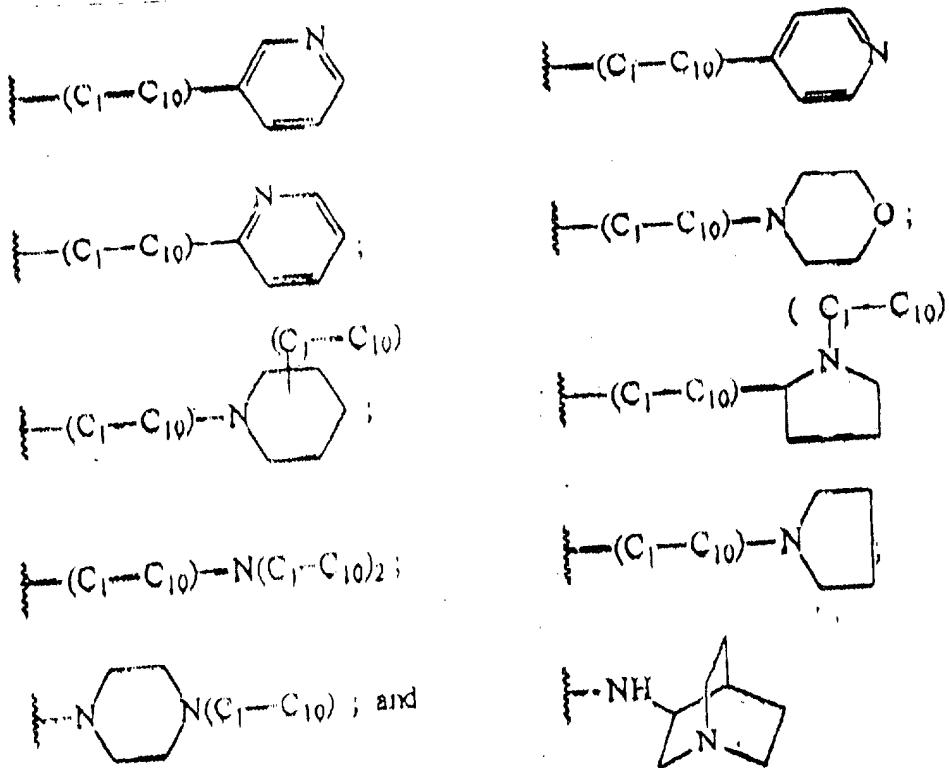
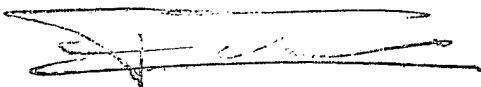
aumentar a sensibilidade relativa de um analito que se destina a detecção por espectrometria de massa são expostos em, p.ex., Sunner, J. et al., *Anal. Chem.* 60:1300-1307 (1988).

Uma funcionalidade preferida para facilitar a produção de uma carga negativa é um ácido orgânico, tal como hidroxifenol, ácido carboxílico, fosfonato, fosfato, tetrazol, sulfonilureia, perfluoroálcool e ácido sulfônico.

As funcionalidades preferidas para facilitar a produção de uma carga positiva são as aminas alifáticas ou aromáticas. Os exemplos de grupos funcionais de aminas que proporcionam uma melhoria da detectabilidade dos marcadores MS incluem aminas quaternárias (i.e., aminas que possuem quatro ligações, cada uma a átomos de carbono; ver Aebersold, Patente dos E.U.A. nº 5.240.859) e aminas terciárias (i.e., aminas que possuem três ligações, cada uma a átomos de carbono, o que inclui grupos C=N-C tais como os presentes na piridina; ver Hess et al., *Anal. Biochem.* 224:373, 1995; Bures et al., *Anal. Biochem.* 224:364, 1995). As aminas terciárias bloqueadas são particularmente preferidas. As aminas terciárias e quaternárias poderão ser constituídas por grupos alquilo ou arilo. Uma porção molecular contendo T^{MS} terá de conter pelo menos uma espécie ionizável, mas poderá possuir mais do que uma espécie ionizável. O estado de carga preferido é o de uma só espécie ionizada por marcador. Do mesmo modo, é preferido que cada porção molecular contendo T^{MS} (e cada componente variável do marcador) contenha apenas um grupo de amina bloqueada ou de ácido orgânico.

Os radicais adequados contendo aminas que poderão fazer parte da porção molecular contendo T^{MS} incluem os seguintes elementos:





A identificação de um marcador por espectrometria de massa baseia-se preferencialmente na sua razão de massa molecular/carga (m/z). O intervalo preferido de massas moleculares dos marcadores MS corresponde a cerca de 100 a 2.000 daltons, e de preferência a porção molecular contendo T^{MS} possui uma massa de pelo menos cerca de 250 daltons, mais preferencialmente pelo menos cerca de 300 daltons, e ainda mais preferencialmente pelo menos cerca de 350 daltons. É geralmente difícil, com os espectrómetros de massa, distinguir entre porções moleculares que possuam iões-mãe com massas abaixo de cerca de 200-250 daltons (dependendo do instrumento usado), e deste modo as porções moleculares contendo T^{MS} deste invento exibem massas superiores às deste intervalo.

Tal como se explica acima, a porção molecular contendo T^{MS} poderá conter átomos diferentes dos presentes no componente variável de marcador, e até do que os presentes no próprio T^{MS} . Assim, a massa do T^{MS} poderá ser inferior a cerca de 250 daltons, desde que a a porção molecular contendo T^{MS} possua uma massa de pelo menos cerca de 250 daltons. Assim, a massa do T^{MS} poderá variar desde 15 (i.e., um radical metilo) até cerca de 10.000 daltons, de preferência de 100 a cerca de 5.000 daltons, e mais preferencialmente de cerca de 200 a cerca de 1.000 daltons.

É relativamente difícil distinguir marcadores por espectrometria de massa quando estes marcadores incorporam átomos que possuem mais do que um isótopo em abundância significativa. Assim, os grupos T preferidos que se destinam a determinação por

espectroscopia de massa (grupos T^{MS}) contém carbono, pelo menos um átomo de hidrogénio e flúor, e átomos opcionais seleccionados a partir do grupo consistindo em oxigénio, azoto, enxofre, fósforo e iodo. Se bem que possam estar presentes outros átomos no T^{MS} , a sua presença poderá dificultar a análise dos dados do espectro de massa. De preferência, os grupos T^{MS} possuirão apenas átomos de carbono, azoto e oxigénio, para além de hidrogénio e/ou flúor.

O flúor constitui um átomo opcional, se bem que preferido, para inclusão num grupo T^{MS} . Em comparação com o hidrogénio, o flúor é, evidentemente, muito mais pesado. Assim, a presença de átomos de flúor em lugar de átomos de hidrogénio leva a grupos T^{MS} de massas superiores, permitindo assim que o grupo T^{MS} atinja e exceda uma massa de 250 daltons, o que está conforme o pretendido, tal como se explicou acima. Para além disto, a substituição de hidrogénio por flúor confere uma maior volatilidade à porção molecular contendo T^{MS} , e uma maior volatilidade do analito aumenta a sensibilidade quando a espectrometria de massa é usada como método de detecção.

A fórmula molecular de T^{MS} encontra-se no intervalo de $C_{1-500}N_{0-100}O_{0-100}S_{0-10}P_{0-10}H_{\alpha}F_{\beta}I_{\delta}$, em que a soma de α , β e δ é suficiente para satisfazer as valências de outro modo não satisfeitas dos átomos de C, N, O, S e P. A designação $C_{1-500}N_{0-100}O_{0-100}S_{0-10}P_{0-10}H_{\alpha}F_{\beta}I_{\delta}$ significa que T^{MS} contém pelo menos um átomo de carbono, e poderá conter qualquer número destes átomos dentro do intervalo de 1 a 500, para além de conter opcionalmente um número de átomos de azoto de até 100 (" N_0- " significa que T^{MS} não necessita de conter quaisquer átomos de azoto), até 100 átomos de O, até 10 átomos de enxofre e até 10 átomos de fósforo. Os símbolos α , β e δ representam o número de átomos de hidrogénio, flúor e iodo presentes em T^{MS} , sendo que qualquer conjunto de dois destes números poderá corresponder a zero e que a soma destes números equivale ao total das valências de outro modo não satisfeitas dos átomos de C, N, O, S e P. De preferência, T^{MS} possui uma fórmula molecular que se encontra no âmbito da fórmula $C_{1-50}N_{0-10}O_{0-10}H_{\alpha}F_{\beta}$, em que a soma de α e β equivale ao número de átomos de hidrogénio e flúor, respectivamente, que se encontram presentes na porção molecular.

b. Características dos Marcadores IR

Existem duas formas primárias de detecção por IR de grupos químicos orgânicos: IR de deslocamento Raman e IR de absorção. Os espectros IR de deslocamento Raman e os espectros IR de absorção constituem métodos espectroscópicos complementares. Em geral, a excitação de Raman depende das alterações da polarizabilidade das ligações, enquanto que a absorção em IR depende das alterações do

S/N 2010/00000000000000000000000000000000

momento dipolar das ligações. As linhas fracas de absorção em IR tornam-se linhas Raman fortes, e vice-versa. O número de onda é a unidade característica dos espectros IR. Existem 3 regiões espectrais para os marcadores IR que possuem aplicações distintas: IR próximo, de 12500 a 4000 cm⁻¹, IR médio, de 4000 a 600 cm⁻¹, e IR afastado, de 600 a 30 cm⁻¹. Para os usos aqui descritos, em que se pretende que um composto sirva como marcador para identificação de uma MOI, sonda ou primer, as regiões médias do espectro são as preferidas. Por exemplo, o alongamento carbonílico (1850 a 1750 cm⁻¹) seria medido para os ácidos carboxílicos, para os ésteres e amidas carboxílicos e para os carbonatos, carbamatos e cetonas alquílicos e arílicos. A deformação N-H (1750 a 160 cm⁻¹) seria usada para identificar as aminas, os iões de amónio e as amidas. No intervalo de 1400 a 1250 cm⁻¹, a deformação R-OH é detectada juntamente com o alongamento C-N nas amidas. Os padrões de substituição aromática são detectados no intervalo de 900 a 690 cm⁻¹ (deformação C-H, deformação N-H para ArNH₂). As ligações C-H saturadas, as olefinas, os anéis aromáticos, ligações duplas e triplas, ésteres, acetais, cetais, sais de amónio, compostos N-O tais como as oximas, compostos nitro, N-óxidos e nitratos, os compostos azo, as hidrazonas, quinonas, ácidos carboxílicos, amidas e lactams possuem dados de correlação vibracional em infravermelhos (ver Pretsch et al., *Spectral Data for Structure Determination of Organic Compounds*, Springer-Verlag, Nova Iorque, 1989). Os compostos preferidos incluiriam um nitrilo aromático que exibe uma vibração de alongamento muito forte no intervalo de 2230 a 2210 cm⁻¹. Outros tipos de compostos úteis são os alquinos aromáticos, os quais apresentam uma forte vibração de alongamento que dá origem a uma forte banda de absorção entre 2140 e 2100 cm⁻¹. Um terceiro tipo de composto preferido corresponde às tiazidas aromáticas, que exibem uma banda intensa de absorção na região de 2160 a 2120 cm⁻¹. Os tiocianatos constituem elementos representativos dos compostos que exibem uma forte absorção na região de 2275 a 2263 cm⁻¹.

c. Características dos Marcadores UV

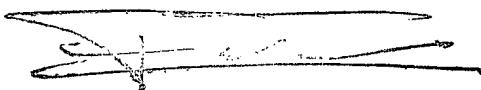
Scott (*Interpretation of the UV Spectra of Natural Products*, Pergamon Press, Nova Iorque, 1962) apresenta uma compilação dos tipos cromóforos orgânicos e das suas respectivas propriedades em UV-visível. Um cromóforo é um átomo, ou grupo de átomos ou electrões, que é responsável pela absorção de luz. Existem regras empíricas para os máximos de π para π^* em sistemas conjugados (ver Pretsch et al., *Spectral Data for Structure Determination of Organic Compounds*, p.B65 e B70, Springer-Verlag, Nova Iorque, 1989). Os compostos preferidos (com sistemas conjugados) possuiriam transições de n para n^* e de π para π^* . Como exemplos de tais compostos encontram-se o Violeta Ácido 7,

[Handwritten signature]

Laranja de Acridina, Amarelo de Acridina G, Azul Brilhante G, Vermelho do Congo, Cristal Violeta, oxalato de Verde Malaquita, Amarelo de Metanilo, Azul de Metileno, Laranja de Metilo, Violeta de Metilo B, Verde de Naftol B, Azul de Óleo N, Vermelho de Óleo O, 4-fenilazofenol, Safranie O, Verde de Solvente 3 e Laranja do Sudão G, sendo que todos estes compostos se encontram disponíveis no mercado (Aldrich, Milwaukee, WI). Outros compostos adequados encontram-se listados em, p.ex., Jane, I., et al., *J. Chrom.* 323:191-225 (1985).

d. Características de um Marcador Fluorescente

As sondas fluorescentes são identificadas e quantificadas directamente através dos seus comprimentos de onda e intensidades de absorção e de emissão de fluorescência. Os espectros de emissão (fluorescência e fosforescência) são muito mais sensíveis e permitem medições mais específicas do que os espectros de absorção. Outras características fotofísicas, tais como o tempo de vida em estado excitado e a anisotropia de fluorescência, são menos utilizadas. Os parâmetros de intensidade de maior utilidade geral são o coeficiente de extinção molar (ϵ) para a absorção e o rendimento quântico (QY) para a fluorescência. O valor de ϵ é especificado para um só comprimento de onda (que corresponde geralmente ao máximo de absorção da sonda), enquanto que o QY constitui uma medida da emissão total de fotões ao longo de todo o perfil do espectro de fluorescência. É geralmente utilizada uma banda óptica estreita (< 20 nm) para a excitação da fluorescência (para a absorção), enquanto que a largura de banda para a detecção da fluorescência é muito mais variável, abrangendo desde a totalidade do espectro, para garantir uma sensibilidade máxima, até uma banda estreita (~20 nm), para obter uma resolução máxima. A intensidade de fluorescência por molécula de sonda é proporcional ao produto de ϵ e QY. O intervalo abrangido por estes parâmetros, entre os fluoróforos de importância prática, é de aproximadamente 10.000 a 100.000 $\text{cm}^{-1}\text{M}^{-1}$ para ϵ e de 0,1 a 1,0 para QY. Como compostos que podem funcionar como marcadores fluorescentes encontram-se os seguintes: fluoresceína, rodamina, azul Lambda 470, verde Lambda, vermelho Lambda 664, vermelho Lambda 665, laranja de acridina e iodeto de propídio, os quais são comercializados pela Lambda Fluorescence Co. (Pleasant Gap, PA). Os compostos fluorescentes vermelho-nilo, vermelho Texas, lissamineTM e BODIPYTM podem ser adquiridos à empresa Molecular Probes (Eugene, OR).

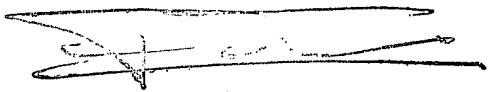


e. Características dos Marcadores Potenciométricos

O princípio da detecção electroquímica (ECD) baseia-se na oxidação ou redução de compostos, sendo que ao aplicar certas voltagens estes compostos se comportam como dadores ou aceitadores de electrões, produzindo-se desta forma uma corrente que pode ser medida. Quando certos compostos são submetidos a uma diferença de potencial, as moléculas sofrem um rearranjo molecular a nível da superfície de trabalho dos eléctrodos, verificando-se uma perda (oxidação) ou ganho (redução) de electrões; tais compostos são designados por electrónicos e sofrem reacções electroquímicas. Os detectores electroquímicos aplicam uma voltagem sobre a superfície de um eléctrodo pela qual passa o fluxo do eluente de uma HPLC. Os compostos electroactivos que são eluídos a partir da coluna poderão doar electrões (oxidando-se) ou adquirir electrões (reduzindo-se), gerando um pico de corrente em tempo real. É importante notar que a quantidade de corrente gerada depende tanto da concentração do analito como da voltagem aplicada, sendo que cada composto possui uma voltagem específica a partir da qual se inicia a sua redução ou oxidação. O detector electroquímico actualmente mais popular é o detector amperométrico, em que o potencial se mantém constante e é medida a corrente produzida a partir da reacção electroquímica. Este tipo de espectrometria é designado por "amperometria potenciotáctica". Estes amperómetros são comercializados pela empresa ESA, Inc., Chelmsford, MA.

Quando a eficiência de detecção corresponde a 100%, os detectores especializados são designados por "coulométricos". Os detectores coulométricos apresentam diversas vantagens práticas a nível de selectividade e sensibilidade, vantagens essas que tornam útil a utilização destes tipos de detectores num conjunto. Nos detectores coulométricos, para uma dada concentração de analito, a corrente de sinal é representada em gráfico como uma função do potencial aplicado ao eléctrodo de trabalho (voltagem). O gráfico sigmoidal resultante é designado por curva de corrente-voltagem, ou voltamagrama hidrodinâmico (HDV). O HDV permite fazer a melhor escolha do potencial a aplicar ao eléctrodo de trabalho de modo a obter a maximização do sinal observado. Uma grande vantagem do HDV consiste na sua sensibilidade inerente, obtendo-se níveis de detecção na gama subfemtomolar.

Existem numerosas substâncias químicas e compostos que apresentam actividade electroquímica, incluindo muitas substâncias bioquímicas, produtos farmacêuticos e pesticidas. Os compostos co-eluídos por cromatografia podem ser resolvidos de modo eficaz mesmo que os seus potenciais de semi-onda (o potencial correspondente a metade do máximo de sinal) difiram em apenas 30-60 mV.

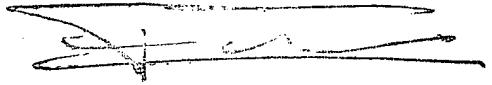


Os sensores coulométricos recentemente desenvolvidos proporcionam selectividade, identificação e resolução de compostos co-eluídos quando são utilizados como detectores em separações por cromatografia líquida. Deste modo, este conjunto de detectores adiciona um outro conjunto de separações que são realizadas pelo próprio detector. Os instrumentos actuais possuem 16 canais que em princípio são limitados apenas pela velocidade a que se dá a aquisição de dados. O número de compostos que podem ser resolvidos no conjunto electroquímico é cromatograficamente limitado (i.e., limitado pelas contagens por placa). No entanto, se um ou mais compostos cromatograficamente co-eluídos possuírem uma diferença de potenciais de semi-onda de 30-60 mV, o conjunto tem a capacidade de distinguir os compostos. A actividade electroquímica exibida por um composto depende da existência de um grupo electroquimicamente activo neste composto (i.e., -OH, -O, -N, -S).

Os compostos que foram detectados com sucesso através do uso de detectores coulométricos incluem a 5-hidroxitriptamina, 3-metoxi-4-hidroxifenilglicol, ácido homogentísico, dopamina, metanefrina, 3-hidroxiquinurenina, acetaminofeno, 3-hidroxitriptofol, ácido 5-hidroxi-indolacético, ácido octanosulfónico, fenol, o-cresol, pirogalhol, 2-nitrofenol, 4-nitrofenol, 2,4-dinitrofenol, 4,6-dinitrocresol, 3-metil-2-nitrofenol, 2,4-diclorofenol, 2,6-diclorofenol, 2,4,5-triclorofenol, 4-cloro-3-metilfenol, 5-metilfenol, 4-metil-2-nitrofenol, 2-hidroxianilina, 4-hidroxianilina, 1,2-fenilenodiamina, benzocatequina, buturon, clortoluron, diuron, isoproturon, linuron, metobromuron, metoxuron, monolinuron, monuron, metionina, triptofano, tirosina, ácido 4-aminobenzóico, ácido 4-hidroxibenzóico, ácido 4-hidroxicumárico, 7-metoxicumarina, apigenina, baicaleína, ácido cafeico, catequina, centaureína, ácido clorogénico, daidzeína, datiscetina, diosmetina, galhato de epicatequina, epigalhocatequina, galhato de epigalhocatequina, eugenol, eupatorina, ácido ferúlico, fisetina, galangina, ácido gállico, gardenina, genisteína, ácido gentísico, hesperidina, irigenina, kaemferol, leucocianidina, luteolina, mangostina, morina, miricetina, naringina, narirutina, pelargondina, peonidina, floretina, pratenseína, ácido protocatecuico, ramnetina, quercetina, sakuranetina, escutelareína, escopoletina, siringaldeído, ácido siríngico, tangeritina, troxerutina, umbelifera, ácido vanílico, 1,3-dimetiltetrahidroisoquinolina, 6-hidroxidopamina, r-salsolinol, N-metil-r-salsolinol, tetrahidroisoquinolina, amitriptilina, apomorfina, capsaicina, clordiazepóxido, clorpromazina, daunorubicina, desipramina, doxepina, fluoxetina, flurazepam, imipramina, isoprotenerol, metoxamina, morfina, morfina-3-glucuronido, nortriptilina, oxazepam, fenilefrina, trimipramina, ácido ascórbico, N-acetylserotoninina, 3,4-dihidroxibenzilamina, ácido 3,4-dihidroximandélico (DOMA), ácido 3,4-dihidroxifenilacético (DOPAC), 3,4-dihidroxifenilalanina (L-DOPA), 3,4-dihidroxifenilglicol

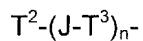
(DHPG), ácido 3-hidroxiantranílico, ácido 2-hidroxifenilacético (2HPAC), ácido 4-hidroxibenzóico (4HBAC), ácido 5-hidroxiindol-3-acético (5-HIAA), 3-hidroxiquinurenina, ácido 3-hidroximandélico, 3-hidroxi-4-metoxifeniletilamina, ácido 4-hidroxifenilacético (4HPAC), ácido 4-hidroxifenil-láctico (4HPLA), 5-hidroxitriptofano (5HTP), 5-hidroxitriptofol (5HTOL), 5-hidroxitriptamina (5HT), sulfato de 5-hidroxitriptamina, 3-metoxi-4-hidroxifenilglicol (MHPG), 5-metoxitriptamina, 5-metoxitriptofano, 5-metoxitriptofol, 3-metoxitiramina (3MT), 3-metoxitirosina (3-OM-DOPA), 5-metilcisteína, 3-metilguanina, bufotenina, dopamina, dopamina-3-glucuronido, dopamina-3-sulfato, dopamina-4-sulfato, epinefrina, epinina, ácido fólico, glutationa (reduzida), guanina, guanosina, ácido homogentísico (HGA), ácido homovanílico (HVA), álcool homovanilílico (HVOL), ácido homoverártico, sulfato de HVA, hipoxantina, indol, ácido indol-3-acético, ácido indol-3-láctico, quinurenina, melatonina, metanefrina, N-metiltriptamina, N-metiltiramina, N,N-dimetiltriptamina, N,N-dimetiltiramina, norepinefrina, normetanefrina, octopamina, piridoxal, fosfato de piridoxal, piridoxamina, sinefrina, triptofol, triptamina, tiramina, ácido úrico, ácido vanilmandélico (VMA), xantina e xantosina. Encontram-se listados outros compostos adequados em, p.ex., Jane, I., et al., *J. Chrom.* 323:191-225 (1985) e Musch, G., et al., *J. Chrom.* 348:97-110 (1985). Estes compostos poderão ser incorporados em compostos com a fórmula T-L-X recorrendo a métodos conhecidos da técnica. Por exemplo, os compostos que contenham um grupo de ácido carboxílico poderão ser colocados em reacção com uma amina, hidroxilo, etc., para formar ligações amida, éster e outras entre T e L.

Para além das propriedades acima expostas, e independentemente do método de detecção pretendido, prefere-se que o marcador possua uma estrutura química modular. Esta estrutura facilita a construção de grandes quantidades de marcadores estruturalmente relacionados por meio de técnicas de química combinatorial. Por exemplo, é desejado que o grupo T^{MS} possua diversas propriedades. Será de desejar que T^{MS} possua um grupo funcional que suporte um único estado ionizado de carga quando a porção molecular contendo T^{MS} for submetida a espectrometria de massa (mais simplesmente, este será um grupo "de melhoria da sensibilidade da espectrometria de massa", ou MSSE). Adicionalmente, é desejado que o grupo T^{MS} constitua um dos membros de uma família de porções moleculares contendo T^{MS} , em que cada membro da família possui uma razão massa/carga diferente mas todos apresentam aproximadamente a mesma sensibilidade no espetrómetro de massa. Assim, é desejado que todos os membros da família tenham o mesmo MSSE. Para possibilitar a criação de famílias de compostos, verificou-se ser conveniente a criação de reagentes de marcador através de um esquema de síntese



modular, de tal modo que os próprios componentes do marcador se possam considerar compostos por módulos.

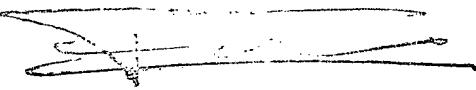
Numa abordagem modular preferida à estrutura do grupo T^{MS} , T^{MS} apresenta a fórmula



em que T^2 é uma porção molecular orgânica formada a partir de carbono e um ou mais átomos de hidrogénio, flúor, iodo, oxigénio, azoto, enxofre e fósforo, possuindo uma massa contida no intervalo de 15 a 500 daltons; T^3 é uma porção molecular orgânica formada a partir de carbono e um ou mais átomos de hidrogénio, flúor, iodo, oxigénio, azoto, enxofre e fósforo, possuindo uma massa contida no intervalo de 50 a 1000 daltons; J é uma ligação directa ou um grupo funcional como um grupo amida, éster, amina, sulfureto, éter, tioéster, dissulfureto, tioéter, ureia, tioureia, carbamato, tiocarbamato, base de Schiff, base de Schiff reduzida, imina, oxima, hidrazona, fosfato, fosfonato, fosforamida, fosfonamida, sulfonato, sulfonamida ou uma ligação carbono-carbono; e n é um índice que varia entre 1 e 50, de tal modo que, quando n é superior a 1, cada um dos elementos T^3 e J é seleccionado independentemente.

A estrutura modular $T^2-(J-T^3)_n-$ proporciona uma entrada conveniente para as famílias de compostos T-L-X, em que cada membro da família possui um grupo T diferente. Por exemplo, quando T corresponde a T^{MS} e todos os membros da família têm o mesmo MSSE, tal como desejado, um dos grupos T^3 poderá fornecer esta estrutura MSSE. De modo a tornar possível a variabilidade entre os membros de uma família no que diz respeito à massa de T^{MS} , o grupo T^2 poderá variar entre os membros da família. Por exemplo, um membro da família poderá ter $T^2 =$ metilo enquanto outro tem $T^2 =$ etilo e outro ainda tem $T^2 =$ propilo, etc.

Para proporcionar saltos grandes, ou "grosseiros" de massa, poderá ser concebido um grupo T^3 que adicione uma quantidade significativa de unidades de massa (p.ex., uma centena ou várias centenas) a T-L-X. Um tal grupo T^3 poderá ser designado como grupo ajustador do peso molecular ("WRA"). Um WRA torna-se bastante útil se se estiver a trabalhar com um só conjunto de grupos T^2 , que possuem massas inseridas num intervalo limitado. Poderá ser usado um só conjunto de grupos T^2 para criar grupos T^{MS} possuam um vasto intervalo de massas incorporando simplesmente um ou mais grupos T^3 WRA no TMS. Assim, usando um exemplo simples, se um conjunto de grupos T^2 proporcionar um intervalo de massas de 250-340 daltons ao T^{MS} , se for feita a adição de um só WRA que tenha, por exemplo, uma massa de 100 daltons, este grupo (que funcionará como T^3) proporcionará o



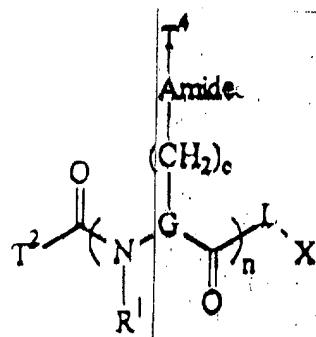
acesso ao intervalo de massas de 350-440 daltons ao ser utilizado o mesmo conjunto de grupos T². Do mesmo modo, a adição de dois grupos WRA com 100 daltons (cada um na qualidade de grupo T³) proporciona o acesso ao intervalo de massas de 450-540 daltons, podendo prosseguir-se esta adição de grupos WRA de modo a proporcionar o acesso a um intervalo bastante vasto de massas para o grupo TMS. Os compostos preferidos com a fórmula T²-(J-T³-)_n-L-X possuem a fórmula R_{VWC}-(R_{WRA})_n-R_{MSSE}-L-X, em que VWC é um grupo "T²" e cada um dos grupos WRA e MSSE são grupos "T³". Esta estrutura é ilustrada na Figura 12 e representa uma abordagem modular à preparação de T^{MS}.

Na fórmula T²-(J-T³-)_n-, T² e T³ são preferencialmente seleccionados a partir do grupo formado por hidrocarbilo, hidrocarbil-O-hidrocarbileno, hidrocarbil-S-hidrocarbileno, hidrocarbil-NH-hidrocarbileno, hidrocarbil-amida-hidrocarbileno, N-(hidrocarbil)hidrocarbileno, N,N-di(hidrocarbil)hidrocarbileno, hidrocarbilacil-hidrocarbileno, heterociclílhidrocarbilo em que o(s) heteroátomo(s) é(são) seleccionado(s) a partir dos átomos de oxigénio, azoto, enxofre e fósforo, heterociclílhidrocarbilo substituído em que o(s) heteroátomo(s) é(são) seleccionado(s) a partir dos átomos de oxigénio, azoto, enxofre e fósforo e os substituintes são seleccionados a partir do grupo constituído por hidrocarbilo, hidrocarbil-O-hidrocarbileno, hidrocarbil-NH-hidrocarbileno, hidrocarbil-S-hidrocarbileno, N-(hidrocarbil)hidrocarbileno, N,N-di(hidrocarbil)hidrocarbileno e hidrocarbilacil-hidrocarbileno. Adicionalmente, T² e/ou T³ poderão constituir derivados de um dos grupos T²/T³ potenciais previamente listados, em que um ou mais hidrogénios são substituídos por átomos de flúor.

Também em relação à fórmula T²-(J-T³-)_n-, um grupo T³ preferido possui a fórmula -G(R2)-, em que G é uma cadeia de alquíleno C1-6 possuindo um só substituinte R2. Assim, se G for o etíleno (-CH₂-CH₂-), um dos dois átomos de etíleno poderão possuir um substituinte R2, e R2 é seleccionado a partir do grupo constituído por alquilo, alquenilo, alquinilo, cicloalquilo, cicloalquilo aril-fundido, cicloalquenilo, arilo, aralquilo, alquenilo aril-substituído ou alquinilo, alquilo cicloalquil-substituído, cicloalquilo cicloalquenil-substituído, biarido, alcoxi, alquenoxy, alquinoxi, aralcoxi, alquenoxy ou alquinoxi aril-substituídos, alquilamino, alquenilamino ou alquinilamino, alquilamino aril-substituído, alquenilamino ou alquinilamino aril-substituído, ariloxi, arilamino, alquilo N-alquilureia-substituído, alquilo N-arilureia-substituído, alquilo alquilcarbonilamino-substituído, alquilo aminocarbonil-substituído, heterociclílo, alquilo heterociclíl-substituído, amino heterociclíl-substituído, aralquilo carboxialquil-substituído, arilo e heterociclílalquilo oxocarbociclíl-fundidos; cicloalquenilo, alquilo e aralquilo aril-substituídos, alquilo hidroxi-substituído, alquilo alcoxi-substituído, alquilo aralcoxi-substituído, alquilo amino-substituído, alquilo(alquilocarbonilamino aril-substituído)-substituído, alquilo tiol-substituído, alquilo

alquilsulfonil-substituído, alquilo(alquitio hidroxi-substituído)-substituído alquilo tioalcoxí-substituído, alquilo hidrocabilacilamino-substituído, alquilo heterocicilacilamino-substituído, alquilo substituído por heterocicilacilamino hidrocabil-substituído, alquilo alquilsulfonilamino-substituído, alquilo arilsulfonilamino-substituído, morfolinoalquilo, tiomorfolinoalquilo, alquilo morfolinocarbonil-substituído, alquilo tiomorfolinocarbonil-substituído, alquilo [N-(alquil, alquenil ou alquinil)- ou N,N-[dialquil, dialquenil, dialquinil ou (alquil, alquenil)-amino]carbonil-substituído, heterocicilaminocarbonilo, heterocicilalquelenoaminocarbonilo, alquilo heterocicilaminocarbonil-substituído, alquilo heterocicilalquelenoaminocarbonil-substituído, N,N-[dialquil]alquelenoaminocarbonilo, alquilo N,N-[dialquil]alquelenoaminocarbonil-substituído, heterocicilcarbonilo alquil-substituído, heterocicilcarbonil-alquilo alquil-substituído, alquilo carboxil-substituído, acilaminoalquilo dialquilamino-substituído e cadeias laterais de aminoácidos seleccionadas a partir do grupo constituído por arginina, asparagina, glutamina, S-metilcisteína, metionina e seus derivados de sulfóxido e sulfona, glicina, leucina, isoleucina, alo-isoleucina, tert-leucina, norleucina, fenilalanina, tirosina, triptofano, prolina, alanina, ornitina, histidina, glutamina, valina, treonina, serina, ácido aspártico, beta-cianoalanina e alotreonina; alinilo e heterocicilcarbonilo, aminocarbonilo, amido, mono- ou dialquilaminocarbonilo, mono- ou diarilaminocarbonilo, alquilarilaminocarbonilo, diarilaminocarbonilo, mono- ou diacilaminocarbonilo, acilo aromático ou alifático, alquilo opcionalmente substituído por substituintes seleccionados a partir do grupo constituído por amino, carboxi, hidroxi, mercapto, mono- ou dialquilamino, mono- ou diarilamino, alquilarilamino, diarilamino, mono- ou diacilamino, alcoxi, alquenoxi, ariloxi, tioalcoxi, tioalquenoxi, tioalquinoxi, tioariloxi e heterociclico.

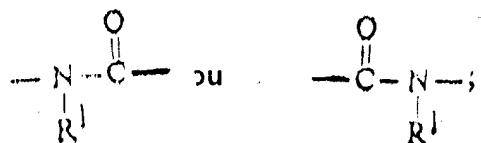
Um composto preferido com a fórmula $T^2-(J-T^3)_n-L-X$ possui a estrutura:



em que G corresponde a $(CH_2)_{1-6}$; de tal modo que um hidrogénio em um e só um dos grupos CH_2 representado por um só "G" é substituído por $-(CH_2)_c\text{-Amida-}T^4$; T^2 e T^4 são porções moleculares orgânicas com a fórmula $C_{1-25}N_{0-9}O_{0-9}H_{\alpha}F\beta$, de tal modo que a

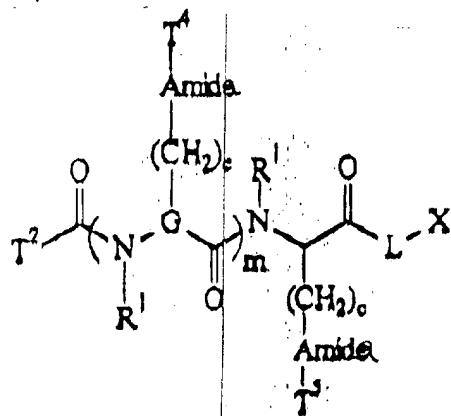
~~S. C. S.~~

soma de α e β é suficiente para satisfazer as valências de outro modo não satisfeitas dos átomos de C, N e O; a amida é



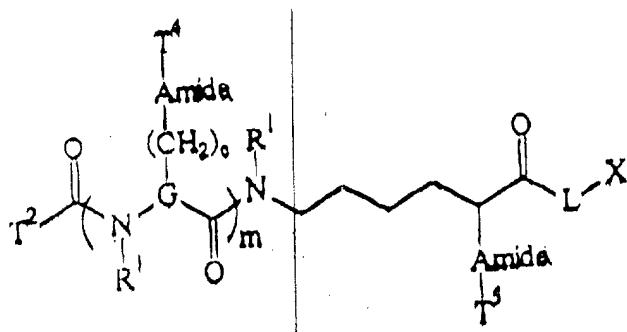
R^1 é um hidrogénio ou um alquilo C_{1-10} ; c é um índice variando de 0 a 4; e n é um índice que varia de 1 a 50, de tal forma que, quando n é superior a 1, G , c , Amida, R^1 e T^4 são seleccionados de modo independente.

Numa outra forma de realização preferida, um composto com a fórmula $T^2-(J-T^3)_n-L-X$ possui a seguinte estrutura:



em que T^5 é uma porção molecular orgânica com a fórmula $C_{1-25}N_{0-9}O_{0-9}H_\alpha F\beta$, de tal modo que a soma de α e β é suficiente para satisfazer as valências de outro modo não satisfeitas dos átomos de C, N e O; T^5 inclui uma amina terciária ou quaternária ou um ácido orgânico; m é um índice que varia de 0-49, e T^2 , T^4 , R^1 , L e X foram já previamente definidos.

Outro composto preferido com a fórmula $T^2-(J-T^3)_n-L-X$ possui a estrutura:

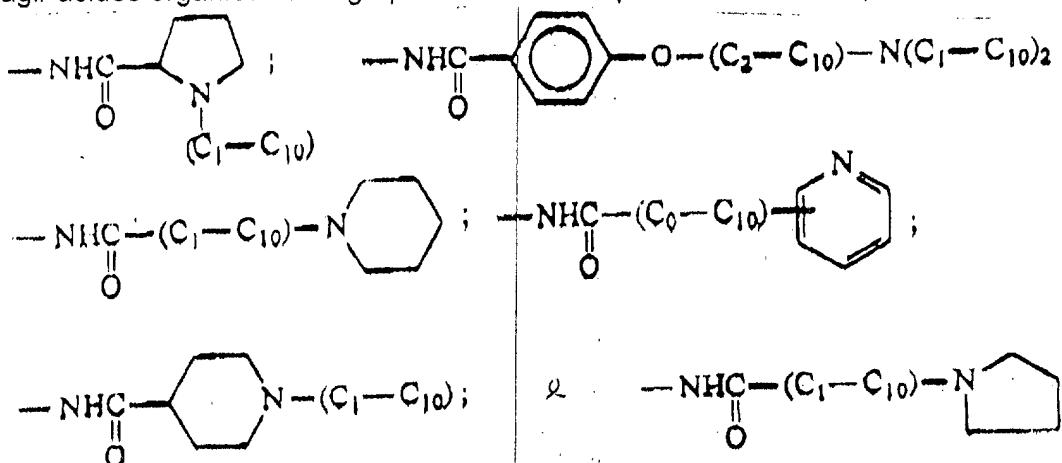


em que T^5 é uma porção molecular orgânica com a fórmula $C_{1-25}N_{0-9}O_{0-9}H_\alpha F\beta$, de tal modo que a soma de α e β é suficiente para satisfazer as valências de outro modo não satisfeitas dos átomos de C, N e O; T^5 inclui uma amina terciária ou quaternária ou um

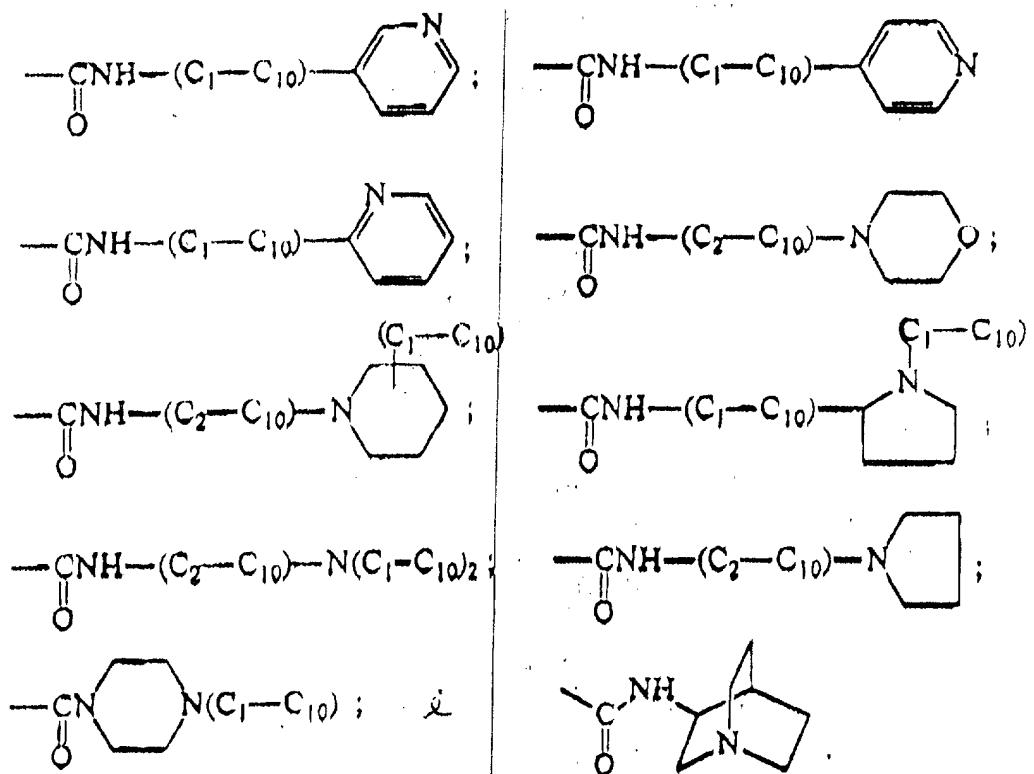
~~SECRET~~

ácido orgânico; m é um índice que varia de 0-49, e T², T⁴, c, R¹, "Amida", L e X foram já previamente definidos.

Nas estruturas acima que possuem um grupo T⁵, o conjunto -Amida-T⁵ traduz-se preferencialmente por um dos seguintes grupos, que são facilmente produzidos fazendo reagir ácidos orgânicos com grupos amino livres que se estendem a partir de "G":

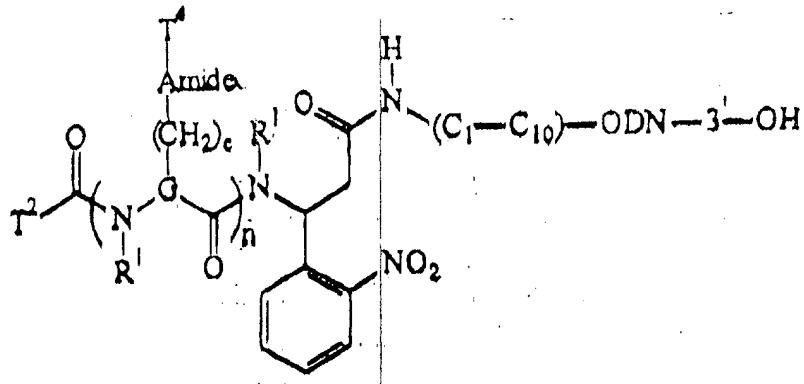


Nos casos em que os compostos acima possuem um grupo T⁵ e o grupo "G" inclui um grupo carboxilo livre (ou um seu equivalente reactivo), são então preferidos os seguintes grupos -Amida-T⁵, que são facilmente produzidos fazendo reagir as aminas orgânicas apropriadas com um grupo carboxilo livre que se estende a partir de "G":

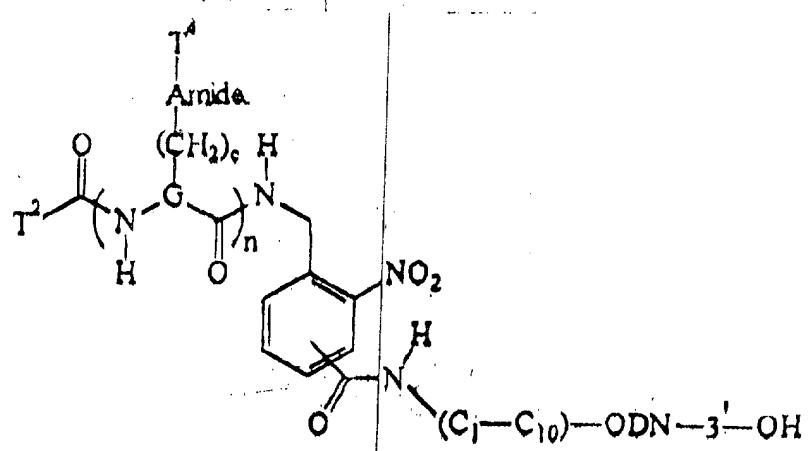


~~SECRET~~

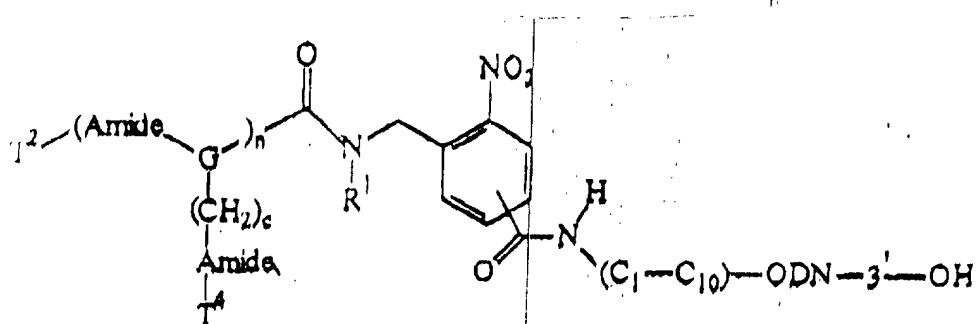
Em três formas de realização preferidas para o invento, T-L-MOI possui a estrutura:



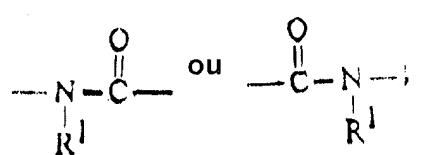
ou a estrutura:



ou a estrutura:



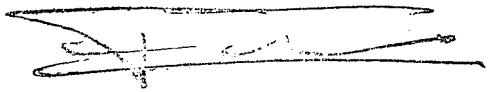
em que T2 e T4 são porções moleculares orgânicas com a fórmula $C_{1-25}N_0\cdot O_{0-9}S_{0-3}P_{0-3}H_\alpha F_\beta I_\delta$, em que a soma de α , β e δ é suficiente para satisfazer as valências de outro modo não satisfeitas dos átomos de C, N, O, S e P; G corresponde a $(CH_2)_{1-6}$, de tal modo que um e só um hidrogénio dos grupos CH_2 representado por cada "G" é substituído por $-(CH_2)_c\text{-Amida-}T^4$; a Amida é



R^1 é um hidrogénio ou um alquilo C_{1-10} ; c é um índice variando entre 0 e 4; “ C_2-C_{10} ” representa um grupo hidrocarbílico possuindo entre 2 a 10 átomos de carbono; “ODN-3'-OH” representa um fragmento de ácido nucleico com um grupo hidroxilo na extremidade 3' (i.e., um fragmento de ácido nucleico ligado a (C_1-C_{10}) por outro local que não a extremidade 3' deste mesmo ácido nucleico); e n é um índice que varia de 1 a 50, de tal modo que, quando n é superior a 1, então G, c, Amida, R^1 e T^4 são seleccionados de modo independente. De preferência, não se verifica a ligação de três heteroátomos a um só átomo de carbono.

Nas estruturas acima expostas que contêm um grupo $T^2-C(=O)-N(R^1)$, este grupo poderá ser formado através da reacção de uma amina de fórmula $HN(R^1)$ - com um ácido seleccionado a partir da lista que se segue, a qual se apresenta como exemplo e não pretende constituir uma lista exaustiva de ácidos orgânicos potenciais: Ácido fórmico, ácido acético, ácido propiólico, ácido propiónico, ácido fluoroacético, ácido 2-butanóico, ácido ciclopropanocarboxílico, ácido butírico, ácido metoxiacético, ácido difluoroacético, ácido 4-pentinóico, ácido ciclobutanocarboxílico, ácido 3,3-dimetilacrílico, ácido valérico, N,N-dimetilglicina, N-formil-Gly-OH, ácido etoxiacético, ácido (metiltio)acético, ácido pirrol-2-carboxílico, ácido 3-furóico, ácido isoxazol-5-carboxílico, ácido trans-3-hexenóico, ácido trifluoroacético, ácido hexanóico, Ac-Gly-OH, ácido 2-hidroxi-2-metilbutírico, ácido benzóico, ácido nicotínico, ácido 2-pirazinacarboxílico, ácido 1-metil-2-pirrolcarboxílico, ácido 2-ciclopenteno-1-acético, ácido ciclopentilacético, ácido (S)-(-)-2-pirrolidona-5-carboxílico, N-metil-L-prolina, ácido heptanóico, Ac-b-Ala-OH, ácido 2-etyl-2-hidroxibutírico, ácido 2-(2-metoxietoxi)acético, ácido p-toluico, ácido 6-metilnicotínico, ácido 5-metil-2-pirazinacarboxílico, ácido 2,5-dimetilpirrol-3-carboxílico, ácido 4-fluorobenzóico, ácido 3,5-dimetilisoxazol-4-carboxílico, ácido 3-ciclopentilpropiónico, ácido octanóico, ácido N,N-dimetsuccinâmico, ácido fenilpropilólico, ácido cinâmico, ácido 4-etylbenzóico, ácido p-anísico, ácido 1,2,5-trimetilpirrol-3-carboxílico, ácido 3-fluoro-4-metilbenzóico, Ac-DL-propargilglicina, ácido 3-(Trifluorometil)butírico, ácido 1-piperidinopropiónico, N-acetilprolina, ácido 3,5-difluorobenzóico, Ac-L-Val-OH, ácido indol-2-carboxílico, ácido 2-benzofuranocarboxílico, ácido benzotriazol-5-carboxílico, ácido 4-n-propilbenzóico, ácido 3-dimetilaminobenzóico, ácido 4-etoxibenzóico, ácido 4-(metiltio)benzóico, N-(2-furoil)glicina, ácido 2-(metiltio)nicotínico, ácido 3-fluoro-4-metoxibenzóico, Tfa-Gly-OH, ácido 2-naftóico, ácido quinalídico, Ac-L-Ile-OH, ácido 3-metilindeno-2-carboxílico, ácido 2-quinoxalinacarboxílico, ácido 1-metilindol-2-carboxílico, ácido 2,3,6-trifluorobenzóico, N-formil-L-Met-OH, ácido 2-[2-(2-metoxietoxi)etoxi]acético, ácido 4-n-butilbenzóico, n-benzoilglicina, ácido 5-fluoroindol-2-carboxílico, ácido 4-n-propoxibenzóico, ácido 4-acetil-

3,5-dimetil-2-pirrolcarboxílico, ácido 3,5-dimetoxibenzóico, ácido 2,6-dimetoxinicótínico, ácido ciclohexanopentanóico, ácido 2-naftilacético, ácido 4-(1H-pirrol-1-il)benzóico, ácido indol-3-propiónico, ácido m-trifluorometilbenzóico, ácido 5-metoxiindol-2-carboxílico, ácido 4-pentilbenzóico, Bz-b-Ala-OH, ácido 4-dietilaminobenzóico, ácido 4-n-butoxibenzóico, ácido 3-metil-5-CF₃-isoxazol-4-carboxílico, ácido (3,4-dimetoxifenil)acético, ácido 4-bifenilcarboxílico, pivaloil-Pro-OH, octanoil-Gly-OH, ácido (2-naftoxi)acético, ácido indol-3-butírico, ácido 4-(trifluorometil)fenilacético, ácido 5-metoxiindol-3-acético, ácido 4-(trifluorometoxi)benzóico, Ac-L-Phe-OH, ácido 4-Pentiloxibenzóico, Z-Gly-OH, 4-carboxi-N-(fur-2-ilmetil)pirrolidin-2-ona, ácido 3,4-dietoxibenzóico, ácido 2,4-dimetil-5-CO₂Et-pirrol-3-carboxílico, ácido N-(2-fluorofenil)succinâmico, ácido 3,4,5-trimetoxibenzóico, ácido N-fenilantranílico, ácido 3-fenoxibenzóico, nonanoil-Gly-OH, ácido 2-fenoxipirimidina-3-carboxílico, ácido 2,5-dimetil-1-fenilpirrol-3-carboxílico, ácido trans-4-(trifluorometil)cinâmico, ácido (5-metil-2-feniloxazol-4-il)acético, ácido 4-(2-ciclohexeniloxi)benzóico, ácido 5-metoxi-2-metilindol-3-acético, ácido trans-4-cotininacarboxílico, ácido Bz-5-aminovalérico, ácido 4-hexiloxibenzóico, ácido n-(3-metoxifenil)succinâmico, Z-Sar-OH, ácido 4-(3,4-dimetoxifenil)butírico, Ac-o-fluoro-DL-Phe-OH, ácido N-(4-fluorofenil)glutarâmico, ácido 4-etyl-4-bifenilcarboxílico, ácido 1,2,3,4-tetrahidroacridinacarboxílico, ácido 3-fenoxifenilacético, ácido N-(2,4-difluorofenil)succinâmico, N-decanoil-Gly-OH, ácido (+)-6-metoxi-a-metil-2-naftalenoacético, ácido 3-(trifluorometoxi)cinâmico, N-formil-DL-Trp-OH, ácido (R)-(+)-a-metoxi-a-(trifluorometil)fenilacético, Bz-DL-Leu-OH, ácido 4-(trifluorometoxi)fenoxiacético, ácido 4-heptiloxibenzóico, ácido 2,3,4-trimetoxicinâmico, 2,6-dimetoxibenzoil-Gly-OH, ácido 3-(3,4,5-trimetoxifenil)propiónico, ácido 2,3,4,5,6-pentafluorofenoxiacético, ácido N-(2,4-difluorofenil)glutarâmico, N-undecanoil-Gly-OH, ácido 2-(4-fluorobenzoil)benzóico, ácido 5-trifluorometoxiindol-2-carboxílico, ácido N-(2,4-difluorofenil)diglicolâmico, Ac-L-Trp-OH, Tfa-L-fenilglicina-OH, ácido 3-iodobenzóico, ácido 3-(4-n-pentilbenzoil)propiónico, ácido 2-fenil-4-quinolinacarboxílico, ácido 4-octiloxibenzóico, Bz-L-Met-OH, ácido 3,4,5-trietoxibenzóico, N-lauroil-Gly-OH, ácido 3,5-bis(trifluorometil)benzóico, Ac-5-metil-DL-Trp-OH, ácido 2-iodofenilacético, ácido 3-ido-4-metilbenzóico, ácido 3-(4-n-hexilbenzoil)propiónico, N-hexanoil-L-Phe-OH, ácido 4-noniloxibenzóico, ácido 4'-(trifluorometil)2-bifenilcarboxílico, Bz-L-Phe-OH, N-tridecanoil-Gly-OH, ácido 3,5-bis(trifluorometil)fenilacético, ácido 3-(4-n-heptilbenzoil)propiónico, N-heptanoil-L-Phe-OH, ácido 4-decilogibenzóico, ácido N-(α,α,α -trifluoro-m-tolil)antranílico, ácido niflúmico, ácido 4-(2-hidroxihexafluoroisopropil)benzóico, N-miristoil-Gly-OH, ácido 3-(4-n-octilbenzoil)propiónico, N-octanoil-L-Phe-OH, ácido 4-undecilogibenzóico, 3-(3,4,5-trimetoxifenil)propionil-Gly-OH, ácido 8-iodonaftóico, N-pentadecanoil-Gly-OH, ácido 4-

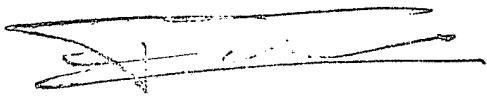


dodecilogibenzóico, N-palmitoil-Gly-OH e N-estearoil-Gly-OH. Estes ácidos orgânicos são comercializados por uma ou várias das seguintes empresas: Advanced ChemTech, Louisville, KY; Bachem Bioscience, Inc., Torrance, CA; Calbiochem-Novabiochem Corp., San Diego, CA; Farchan Laboratories Inc., Gainesville FL; Lancaster Synthesis, Wyndham, NH; e MayBridge Chemical Company (c/o Ryan Scientific), Columbia, SC. Os catálogos destas companhias usam as abreviaturas indicadas acima para identificar os ácidos.

f. A Química Combinatorial como Meio para a Preparação de Marcadores

A química combinatorial constitui um tipo de estratégia de síntese que conduz à produção de grandes colecções químicas (ver, por exemplo, a Publicação do Pedido de Patente PCT Nº WO 94/08051). Estas colecções combinatoriais podem ser usadas como marcadores para a identificação de moléculas de interesse (MOIs). A química combinatorial pode definir-se como a ligação covalente sistemática e repetitiva de um conjunto de diversos "blocos de construção" com estruturas que variam entre si, por forma a originar um vasto conjunto de diversas entidades moleculares. Os blocos de construção podem tomar muitas formas, tanto naturais como sintéticas, tais como nucleófilos, electrófilos, dienos, agentes alquilantes ou acilantes, diaminas, nucleótidos, aminoácidos, açúcares, lípidos, monómeros orgânicos, synthons e combinações de todos estes. Como reacções químicas de ligação dos blocos de construção poderão usar-se a alquilação, acilação, oxidação, redução, hidrólise, substituição, eliminação, adição, ciclização, condensação e outras. Este processo pode produzir colecções de compostos oligoméricos, não-oligoméricos ou resultantes de combinações destes. No caso dos compostos oligoméricos, estes poderão ser ramificados, não ramificados ou cílicos. Os exemplos de estruturas oligoméricas que podem ser preparadas através de métodos combinatoriais incluem oligopeptídos, oligonucleótidos, oligossacáridos, polilípidos, poliésteres, poliamidas, poliuretanos, poliureias, poliéteres, poli(derivados de fósforo), p.ex., fosfatos, fosfonatos, fosforamidas, fosfonamidas, fosfitos, fosfinamidas, etc., e poli(derivados de enxofre), p.ex., sulfonas, sulfonatos, sulfitos, sulfonamidas, sulfenamidas.

Um tipo comum de colecção combinatorial oligomérica corresponde à colecção combinatorial de péptidos. As inovações recentes na química de péptidos e na biologia molecular possibilitaram a preparação e utilização de colecções formadas por dezenas a centenas de milhões de sequências peptídicas diferentes. Tais colecções podem dividir-se em três categorias gerais. Um das categorias de colecções envolve a síntese química de colecções de péptidos solúveis não ligados a suportes (p.ex., Houghten et al.,



Nature 354:84, 1991). Uma segunda categoria envolve a síntese química de colecções de péptidos ligados a suportes, apresentados em suportes sólidos como alfinetes de plástico, esferas de resina ou algodão (Geysen et al., *Mol. Immunol.* 23:709, 1986; Lam et al., *Nature* 354:82, 1991; Eichler e Houghten, *Biochemistry* 32:11035, 1993). Nestas primeiras duas categorias, os blocos de construção correspondem tipicamente a L-aminoácidos, D-aminoácidos, aminoácidos não naturais ou a uma mistura ou combinação destes. Uma terceira categoria utiliza abordagens da biologia molecular para preparar péptidos ou proteínas à superfície de partículas de fagos filamentosos ou plasmídeos (Scott e Craig, *Curr. Opinion Biotech.* 5:40, 1994). As colecções de péptidos solúveis não ligados a suportes parecem ser adequadas a diversas aplicações, incluindo o uso como marcadores. O repertório disponível de diversidades químicas em colecções de péptidos pode ser expandido através de passos como a permetilação (Ostresh et al., *Proc. Natl. Acad. Sci. USA* 91:11138, 1994).

É possível obter numerosas variantes de colecções combinatoriais de péptidos em que a cadeia principal do péptido é modificada e/ou as ligações amina são substituídos por grupos miméticos. Os grupos miméticos de amina que podem ser utilizados incluem ureias, uretanos, e grupos de carbonilmetileno. A reestruturação da cadeia principal de tal modo que as cadeias laterais emanem dos azotos de amida de cada aminoácido, em lugar dos carbonos alfa, fornece colecções de compostos conhecidos por peptóides (Simon et al., *Proc. Natl. Acad. Sci. USA* 89:9367, 1992).

Outro tipo comum de colecção combinatorial oligomérica corresponde à colecção combinatorial de oligonucleótidos, em que os blocos de construção constituem derivados de nucleótidos ou de polissacáridos, de ocorrência natural ou sintéticos, incluindo compostos em que vários grupos orgânicos e inorgânicos poderão substituir a ligação fosfato, e o azoto ou o enxofre poderão substituir o oxigénio numa ligação éter (Schneider et al., *Biochem.* 34:9599, 1995; Freier et al., *J. Med. Chem.* 38:344, 1995; Frank, *J. Biotechnology* 41:259, 1995; Schneider et al., PCT Publicada Nº WO 942052; Ecker et al., *Nucleic Acids Res.* 21:1853, 1993).

Mais recentemente, tem sido descrita a produção combinatorial de colecções de compostos não oligoméricos de moléculas pequenas (DeWitt et al., *Proc. Natl. Acad. Sci. USA* 90:690, 1993; Bunin et al., *Proc. Natl. Acad. Sci. USA* 91:4708, 1994). As estruturas adequadas a elaboração em colecções de moléculas pequenas incluem uma grande variedade de moléculas orgânicas, por exemplo estruturas heterocíclicas, aromáticas, alicíclicas, alifáticas, esteróides, antibióticos, inibidores enzimáticos, ligandos, hormonas,

fármacos, alcalóides, opióides, terpenos, porfirinas, toxinas, catalisadores, assim como combinações destes.

g. Métodos Específicos de Síntese Combinatorial de Marcadores

São abaixo expostos dois métodos de preparação e uso de um conjunto de diversos marcadores MS contendo o grupo amina. Em ambos os métodos é utilizada a síntese em fase sólida de modo a permitir a síntese simultânea e em paralelo de um grande número de linkers marcados, utilizando as técnicas da química combinatorial. No primeiro método, o eventual corte da ligação entre o marcador e o oligonucleotídeo resulta na libertação de uma amida carboxílica. No segundo método, o corte da ligação com o marcador produz um ácido carboxílico. Os componentes químicos e elementos de ligação usados nestes métodos são abreviados tal como se segue:

R	=	resina
FMOC	=	grupo protector fluorenilmetoxicarbonilo
All	=	grupo protector alilo
CO ₂ H	=	grupo de ácido carboxílico
CONH ₂	=	grupo de amida carboxílica
NH ₂	=	grupo amino
OH	=	grupo hidroxilo
CONH	=	ligação amida
COO	=	ligação éster
NH ₂ –Rink– CO ₂ H	=	ácido 4-[(α -amino)-2,4-dimetoxibenzil]-fenoxibutírico (linker de Rink)
OH –1MeO– CO ₂ H	=	ácido (4-hidroximetil)fenoxibutírico
OH –2MeO– CO ₂ H	=	ácido (4-hidroximetil-3-metoxi)fenoxiacético
NH ₂ -A-COOH	=	aminoácido com uma funcionalidade de amina alifática ou aromática na cadeia lateral
X ₁X _n -COOH	=	conjunto de diversos (n) ácidos carboxílicos com pesos moleculares individualizados
oligo1.....oligo(n)	=	conjunto de n oligonucleótidos
HBTU	=	hexafluorofosfato de O-benzotriazol-1-il-N,N,N',N'- Tetrametilurônio

~~SECRET~~

A sequência de passos do Método 1 é a seguinte:

$\text{OH} - 2\text{MeO}-\text{CONH}-\text{R}$

↓ $\text{FMOC}-\text{NH}-\text{Rink}-\text{CO}_2\text{H}$; acoplamento (p.ex., HBTU)

$\text{FMOC}-\text{NH}-\text{Rink}-\text{COO}-2\text{MeO}-\text{CONH}-\text{R}$

↓ piperidina (remoção de FMOC)

$\text{NH}_2-\text{Rink}-\text{COO}-2\text{MeO}-\text{CONH}-\text{R}$

↓ $\text{FMOC}-\text{NH}-\text{A}-\text{COOH}$; acoplamento (p.ex., HBTU)

$\text{FMOC}-\text{NH}-\text{A}-\text{CONH}-\text{Rink}-\text{COO}-2\text{MeO}-\text{CONH}-\text{R}$

↓ piperidina (remoção de FMOC)

$\text{NH}_2-\text{A}-\text{CONH}-\text{Rink}-\text{COO}-2\text{MeO}-\text{CONH}-\text{R}$

↓ dividir em n alíquotas

↓↓↓↓ acoplar a diferentes ácidos $\text{X}_1 \dots \text{X}_n - \text{COOH}$

$\text{X}_1 \dots \text{X}_n - \text{CONH}-\text{A}-\text{CONH}-\text{Rink}-\text{COO}-2\text{MeO}-\text{CONH}-\text{R}$

↓↓↓↓ Separar os linkers marcados da resina com TFA a 1%

$\text{X}_1 \dots \text{X}_n - \text{CONH}-\text{A}-\text{CONH}-\text{Rink}-\text{CO}_2\text{H}$

↓↓↓↓ acoplar a n oligos (oligo1.....oligo(n))

(p.ex., através de ésteres Pfp)

$\text{X}_1 \dots \text{X}_n - \text{CONH}-\text{A}-\text{CONH}-\text{Rink}-\text{CONH}-\text{oligo1....oligo}(n)$

↓ fazer um "pool" dos oligos marcados

↓ realizar a reacção de sequenciação

↓ separar os fragmentos de diferentes extensões resultantes da reacção de sequenciação (p.ex., por HPLC ou CE)

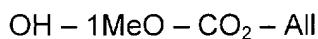
↓ separar os marcadores dos linkers com TFA a 25%-100%

$\text{X}_1 \dots \text{X}_n - \text{CONH}-\text{A}-\text{CONH}$

↓ analisar por espectrometria de massa



A sequência de passos do Método 2 é a seguinte:



↓ FMOC - NH - A - CO₂H; acoplamento (p.ex., HBTU)



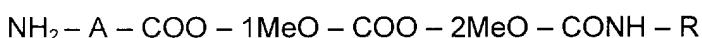
↓ Paládio (remoção do Alilo)



↓ OH - 2MeO - CONH - R; acoplamento (p.ex., HBTU)



↓ piperidina (remoção de FMOC)

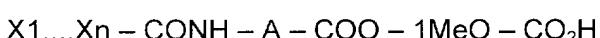


↓ dividir em n alíquotas

↓↓↓↓↓ acoplar a diferentes ácidos X₁....X_n - CO₂H



↓↓↓↓↓ Separar os linkers marcados da resina com TFA a 1%



↓↓↓↓↓ acoplar a n oligos (oligo1.....oligo(n))

(p.ex., através de ésteres Pfp)



↓ fazer um "pool" dos oligos marcados

↓ realizar a reacção de sequenciação

↓ separar os fragmentos de diferentes extensões resultantes da reacção de sequenciação (p.ex., por HPLC ou CE)

↓ separar os marcadores dos linkers com TFA a 25%-100%



↓

analisar por espectrometria de massa

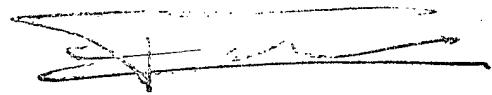
2. Linkers

Um componente de “linker” (ou L) tal como aqui é usado, tanto pode traduzir uma ligação covalente directa como um grupo químico orgânico que é utilizado para ligar um “marcador” (ou T) a uma “molécula de interesse” (MOI) através de ligações químicas covalentes. Adicionalmente, a ligação directa em si, ou uma ou mais ligações no seio do componente de linker, são passíveis de corte sob condições que permitem que T seja libertado (por outras palavras, cortado) a partir da porção remanescente do composto T-L-X (incluindo o componente MOI). O componente variável do marcador que se encontra presente no seio de T deverá ser estável face às condições de corte de ligações. De preferência, a reacção de corte é efectuada rapidamente, no espaço de alguns minutos, e de preferência em cerca de 15 segundos ou menos.

Em geral, um linker é usado para ligar um elemento pertencente a um grande conjunto de marcadores a um elemento de um conjunto similarmente vasto de MOIs. Tipicamente, uma só combinação de marcador-linker é ligada a cada MOI (originando diversos T-L-MOI), mas em alguns casos poderá ser ligada mais do que uma combinação de marcador-linker a cada MOI individual (originando diversos (T-L)_n-MOI). Em outra forma de realização do presente invento, são ligados dois ou mais marcadores a um só linker através de locais múltiplos e independentes no linker, e esta combinação múltipla de marcador-linker é então ligada a uma MOI individual (originando diversos (T)_n-L-MOI).

Após diversas manipulações do conjunto de MOIs marcadas, são usadas condições químicas e/ou físicas especiais para cortar uma ou mais ligações covalentes no seio do linker, resultando na libertação dos marcadores a partir das MOIs. A(s) ligação(ões) passível(eis) de corte poderão corresponder ou não a algumas das ligações que se formaram quando o marcador, o linker e a MOI foram ligados entre si. A concepção do linker determinará em grande medida o conjunto de condições sob as quais será realizada a reacção de corte de ligações. Assim, os linkers poderão também ser identificados segundo as condições de corte a que são particularmente susceptíveis. Quando um linker é fotolábil (i.e., susceptível a corte de ligações por exposição a radiação actínica), o linker pode ser designado por L^{hv}. Do mesmo modo, poderão usar-se as designações L^{ácido}, L^{base}, L^[O], L^[R], L^{enz}, L^{elc}, L^Δ e L^{ss} para designar linkers que são particularmente susceptíveis a corte de ligações por ácido, base, oxidação química, redução química, a actividade catalítica de uma enzima (mais simplesmente “enzima”), oxidação ou redução electroquímica, temperatura elevada (“térmica”) e troca de tióis, respectivamente.

Certos tipos de linker são lábeis a um só tipo de reacção de corte de ligações, enquanto que outros são lábeis face a vários tipos de condições de corte. Adicionalmente,



em linkers que têm a capacidade de se ligar a múltiplos marcadores (originando estruturas do tipo (T)n-L-MOI), cada um dos locais de ligação ao marcador poderá ser lábil face a diferentes condições de corte. Por exemplo, num linker que esteja ligado a dois marcadores, um dos marcadores poderá ser lábil apenas a bases e o outro apenas a fotólise.

Um linker com utilidade no âmbito deste invento possui vários atributos:

1) O linker possui uma âncora química (L_h) através da qual se pode ligar a uma MOI.
2) O linker possui uma segunda âncora química separada (L_h) através da qual o marcador se liga ao linker. Se ocorrer a ligação de diversos marcadores a um só linker (estruturas de tipo (T)n-L-MOI), existirá uma âncora química separada para cada marcador.

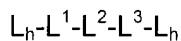
3) O linker é estável face a todas as manipulações a que é submetido, com a excepção das condições que permitem o corte das ligações de modo a que uma porção molecular contendo T se liberte da porção remanescente do composto, incluindo a MOI. Assim, o linker é estável durante os passos de ligação do marcador ao linker, de ligação do linker à MOI e de quaisquer manipulações da MOI enquanto o marcador e o linker (T-L) se encontram ligados a esta.

4) O linker não interfere de modo significativo com as manipulações efectuadas sobre a MOI enquanto o conjunto T-L se mantém ligado a esta. Por exemplo, se T-L estiver ligado a um oligonucleótido, o T-L não poderá interferir de forma significativa com quaisquer reacções de hibridização ou reacções enzimáticas (p.ex., PCR) realizadas sobre o oligonucleótido. Do mesmo modo, se T-L estiver ligado a um anticorpo, não poderá interferir significativamente com o reconhecimento do抗ígeno pelo anticorpo.

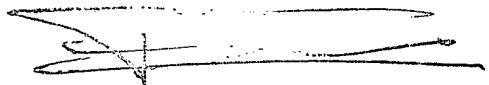
5) O corte do marcador a partir da porção remanescente do composto ocorre de modo altamente controlado, usando processos físicos ou químicos que não afectam adversamente a detectabilidade do marcador.

Para qualquer linker, é preferido que o linker se possa ligar a uma grande variedade de MOIs, e que uma grande variedade de marcadores se possa ligar ao linker. Esta flexibilidade é vantajosa porque permite que uma colecção de conjugados T-L, uma vez preparada, possa ser usada com vários conjuntos diferentes de MOIs.

Tal como acima foi exposto, um linker preferido possui a fórmula



em que cada L_h é uma âncora química que pode ser usada para ligar o linker a um reagente de marcador e a um reagente de molécula de interesse. L^2 forma uma parte essencial do linker, uma vez que L^2 confere labilidade ao linker. L^1 e L^3 são grupos opcionais que servem, na realidade, para separar L^2 das âncoras L_h .

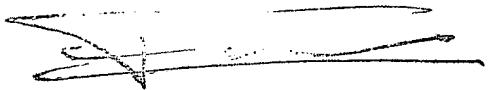


L^1 (que, por definição, se encontra mais próximo de T do que L^3) serve para separar T da porção lábil requerida L^2 . Esta separação poderá ser útil quando a reacção de corte origina espécies particularmente reactivas (p.ex., radicais livres) que poderão causar alterações aleatórias na estrutura da porção molecular contendo T. Uma vez que o local de corte de ligações se encontra mais separado da porção molecular contendo T, existe uma menor probabilidade de que a espécie reactiva formada ao nível do local de corte quebre a estrutura da porção molecular contendo T. Adicionalmente, uma vez que os átomos de L^1 se encontrarão tipicamente presentes na porção molecular contendo T, estes átomos de L^1 poderão conferir uma qualidade desejável à porção molecular contendo T. Por exemplo, nos casos em que a porção molecular contendo T corresponde a uma porção molecular contendo T^{MS} , e uma amina bloqueada se encontra desejavelmente presente como parte da estrutura da porção molecular contendo T^{MS} (para servir, p.ex., como um MSSE), a amina bloqueada poderá encontrar-se presente na porção lábil L^1 .

Em outros casos, L^1 e/ou L^3 poderão encontrar-se presentes num componente de linker simplesmente por o fornecedor comercial do linker optar por comercializar o linker sob uma forma que inclua um grupo L^1 e/ou L^3 . Em tal caso, não existe inconveniente em utilizar linkers que possuam grupos L^1 e/ou L^3 (desde que esse grupo não iniba a reacção de corte de ligações), mesmo que estes grupos possam não contribuir com qualquer vantagem particular de performance para os compostos que os incorporam. Assim, o presente invento permite que estejam presentes no componente de linker grupos L^1 e/ou L^3 .

Os grupos L^1 e/ou L^3 poderão corresponder a uma ligação directa (caso em que o grupo não estará efectivamente presente), a um grupo hidrocarbíleno (p.ex., alquíleno, aríleno, cicloalquíleno, etc.), -O-hidrocarbíleno (p.ex., -O-CH₂-, O-CH₂CH(CH₃)-, etc.) ou hidrocarbíleno-(O-hidrocarbíleno)_w-, em que w é um índice que varia entre 1 e 10 (p.ex., -CH₂-O-Ar-, -CH₂-(O-CH₂CH₂)₄-, etc.).

Com o advento da síntese em fase sólida, foi desenvolvido um grande volume de publicações que se debruçam sobre linkers lábeis a condições de reacção específicas. Numa reacção típica de síntese em fase sólida, um suporte sólido é ligado através de um linker a um local reactivo, e uma molécula a ser sintetizada é gerada a nível do local reactivo. Quando a síntese da molécula se encontra completa, a construção de suporte sólido-linker-molécula é submetida a condições de corte de ligações que libertam a molécula do suporte sólido. Os linkers lábeis que foram desenvolvidos para uso neste contexto (ou que podem ser usados neste contexto) poderão também ser usados como reagente de linker neste invento.



Lloyd-Williams, P., et al., "Convergent Solid-Phase Peptide Synthesis", Relatório Tetrahedron Nº 347, 49(48):11065-11133 (1993) fornece uma extensa lista de linkers lábeis à radiação actínica (i.e., à fotólise), assim como a ácidos, bases e outras condições de quebra de ligações. São bem conhecidas da técnica outras fontes de informação sobre linkers lábeis.

Tal como se descreveu acima, as diferentes concepções do linker farão com que a sua susceptibilidade ao corte ("labilidade") se verifique sob diferentes condições específicas, físicas ou químicas. Os exemplos de condições que proporcionam o corte de ligações em diversos tipos de linker incluem as condições ácidas, básicas, de oxidação, redução, fluoretos, troca de tióis, fotólise e acção enzimática.

Os exemplos de linkers susceptíveis ao corte que satisfazem os critérios gerais para linkers acima apresentados são bem conhecidos dos técnicos e incluem os compostos listados no catálogo de Pierce (Rockford, JL). Os exemplos incluem:

- glicobis(succinimidilsuccinato) de etileno (EGS), um reagente de "cross-linking" (que estabelece ligações cruzadas) reactivo a aminas e que sofre quebra de ligações por acção da hidroxilamina (1 M a 37°C durante 3-6 horas);
- tartarato de disuccinimidilo (DST) e sulfo-DST, que são reagentes de cross-linking reactivos a aminas, susceptíveis ao corte por periodato de sódio a 0,015 M;
- bis[2-(succinimidoxicarboniloxi)etil]sulfona (BSOCOES) e sulfo-BSOCOES, que são reagentes de cross-linking reactivos a aminas, susceptíveis ao corte por base (pH 11,6);
- 1,4-di-[3'-(2'-piridilditio(propionamido))butano (DPDPB), um piridilditiol, reagente de crosslinking susceptível ao corte por troca de tióis ou por redução;
- N-[4-(p-azidosalicilamido)-butil]-3'-(2'-piridilditio)propionamida (APDP), um piridilditiol, reagente de crosslinking susceptível ao corte por troca de tióis ou por redução;
- Bis-[beta-4-(azidosalicilamido)etil]disulfureto, um reagente de crosslinking fotoreactivo que é susceptível ao corte por troca de tióis ou por redução;
- N-succinimidil-(4-azidofenil)-1,3'ditiopropionato (SADP), um reagente de crosslinking fotoreactivo que é susceptível ao corte por troca de tióis ou por redução;
- sulfosuccinimidil-2-(7-azido-4-metilcumarina-3-acetamida)etil-1,3'-ditiopropionato (SAED), um reagente de crosslinking fotoreactivo que é susceptível ao corte por troca de tióis ou por redução;

- sulfosuccinimidil-2-(m-azido-o-nitrobenzamido)-etil-1,3'-ditiopropionato (SAND),
 - um reagente de crosslinking fotoreactivo que é susceptível ao corte por troca de tióis ou por redução.

São em seguida fornecidos outros exemplos de linkers susceptíveis a corte e de condições que podem ser usadas para libertar os marcadores. Um grupo linker de sililo pode ser cortado por fluoretos ou por condições acídicas. O linker 2-nitrobenziloxi 3-, 4-, 5- ou 6-substituído ou o 4-nitrobenziloxi 2-, 3-, 5- ou 6-substituído podem ser cortados por uma fonte de fotões (fotólise). O linker 2-alcoxifenoxi 3-, 4-, 5- ou 6-substituído ou o 4-alcoxifenoxi 2-, 3-, 5- ou 6-substituído podem ser cortados por $\text{Ce}(\text{NH}_4)_3(\text{NO}_3)_6$ (oxidação). Um linker de NCO_2 (uretano) pode ser cortado por hidróxido (base), ácido ou LiAlH_4 (redução). Um grupo linker de 3-pentenilo, 2-butenilo ou 1-butenilo pode ser cortado por O_3 , $\text{O}_5\text{O}_4|\text{IO}_4^-$ ou KMnO_4 (oxidação). Um linker de 2-[furil 3-, 4- ou 5-substituído]oxi pode ser cortado por O_2 , Br_2 , MeOH ou ácido.

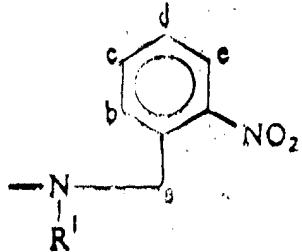
As condições para o corte de outros linkers lábeis incluem: os linkers de t-alquiloxi podem ser cortados por ácido; os linkers de metil(dialquil)metoxi ou 2-alquil-1,3-dioxano-2-iloxi 4-substituído podem ser cortados por H_3O^+ ; os grupos 2-sililetoxi podem ser cortados por fluoretos ou ácido; os linkers 2-(X)-etoxi (em que X = ceto, amida de éster, ciano, NO_2 , sulfureto, sulfóxido, sulfona) podem ser cortados sob condições alcalinas; os linkers benziloxi 2-, 3-, 4-, 5- ou 6-substituídos podem ser cortados por condições acídicas ou de redução; os linkers de buteniloxi podem ser cortados por $(\text{Ph}_3\text{P})_3\text{RhCl}(\text{H})$; os linkers 2-bromofenoxi 3-, 4-, 5- ou 6-substituídos podem ser cortados por Li, Mg ou BuLi; os linkers de metiltiometoxi podem ser cortados por Hg^{2+} ; os linkers de 2-(X)-etiloxi (em que X = um halogénio) podem ser cortados por Zn ou Mg; os linkers de 2-hidroxietiloxi podem ser cortados por oxidação (p.ex., com $\text{Pb}(\text{OAc})_4$).

Os linkers preferidos são os susceptíveis a corte por ácido ou por fotólise. Diversos linkers que foram desenvolvidos para a síntese de péptidos em fase sólida têm utilidade na ligação de marcadores às MOIs. Alguns destes linkers estão descritos numa publicação recente de Lloyd-Williams et al. (*Tetrahedron* 49:11065-11133, 1993). Um tipo útil de linker é baseado em álcoois p-alcoxibenzílicos, dois dos quais, o ácido 4-hidroximetilfenoxiacético e o ácido 4-(4-hidroximetil-3-metoxifenoxy)butírico, são comercializados pela Advanced ChemTech (Louisville, KY). Ambos os linkers se podem ligar a um marcador através de uma ligação éster ao benzilálcool, e a uma MOI que contenha uma amina através de uma ligação amida ao ácido carboxílico. Os marcadores ligados a estas moléculas são libertados da MOI por ação de concentrações variáveis de ácido trifluoroacético. O corte de ligações destes

linkers resulta na libertação de um ácido carboxílico no marcador. O corte por ácido das ligações entre marcadores e linkers como o 2,4-dimetoxi-4'-(carboximetiloxi)-benzidrilamina (fornecido pela Advanced ChemTech sob forma protegida por Fmoc) resulta na libertação de uma amida carboxílica no marcador libertado.

Na sua maioria, os linkers fotolábeis com utilidade no âmbito deste pedido de patente foram também desenvolvidos para a síntese de péptidos em fase sólida (ver a publicação de Lloyd-Williams). Estes linkers têm geralmente por base os 2-nitrobenzilésteres ou as 2-nitrobenzilamidas. Como exemplos de linkers fotolábeis recentemente referidos na literatura encontram-se o ácido 4-(4-(1-Fmoc-amino)etil)-2-metoxi-5-nitrofenoxi)butanóico (Holmes e Jones, *J. Org. Chem.* 60:2318-2319, 1995) e o ácido (3-Fmoc-amino)3-(2-nitrofenil)propiónico (Brown et al., *Molecular Diversity* 1:4-12, 1995). Ambos os linkers se podem ligar por meio do ácido carboxílico a uma amina presente na MOI. A ligação do marcador ao linker é efectuada através da formação de uma amida entre um ácido carboxílico do marcador e a amina do linker. O corte de ligações dos linkers fotolábeis é geralmente efectuado por meio de radiação UV com um comprimento de onda de 350 nm, usando intensidades e períodos de tempo conhecidos da técnica. O corte de ligações dos linkers resulta na libertação de uma amida primária no marcador. Os exemplos de linkers fotolábeis incluem ésteres de nitrofenil-glicina, cloretos e metanosulfonatos de exo- e endo-2-benzonorborneilo e o ácido 3-amino-3(2-nitrofenil)propiónico. Como exemplos de corte enzimático de ligações encontram-se o corte de ligações éster pelas esterases, o corte das ligações fosfodiéster por acção das nucleases, o corte das ligações peptídicas pelas proteases, etc.

Um componente de linker preferido possui uma estrutura orto-nitrobenzílica, tal como se mostra abaixo:

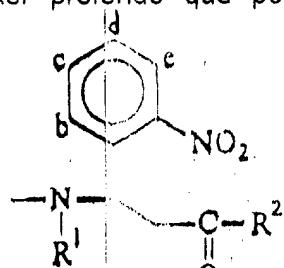


em que um átomo de carbono numa das posições **a**, **b**, **c**, **d** ou **e** é substituído por $\text{L}^3\text{-X}$, e L^1 (que corresponde de preferência a uma ligação directa) está presente à esquerda de $\text{N}(\text{R}^1)$ na estrutura acima. Um tal componente de linker é susceptível ao corte selectivo foto-induzido da ligação entre o carbono marcado como "a" e $\text{N}(\text{R}^1)$. Tipicamente, a identidade de R^1 não é crítica para a reacção de corte da ligação; no entanto, R^1 corresponde preferencialmente a hidrogénio ou hidrocarbilo. O presente invento prevê que,

S. C. S.

na estrutura acima, $-N(R^1)-$ possa ser substituído por $-O-$. Igualmente na estrutura acima, uma ou mais das posições b, c, d ou e poderá opcionalmente ser substituída com um grupo alquilo, alcoxi, fluoreto, cloreto, hidroxilo, carboxilato ou amida, sendo estes substituintes seleccionados de modo independente em cada ocorrência:

Um outro componente de linker preferido que possui uma âncora química L_n apresenta a seguinte estrutura:



em que uma ou mais das posições **b**, **c**, **d** ou **e** é substituída com hidrogénio, alquilo, alcoxi, fluoreto, cloreto, hidroxilo, carboxilato ou amida, R^1 corresponde a hidrogénio ou hidrocarbilo e R^2 corresponde a $-OH$ ou a um grupo que protege ou activa um ácido carboxílico para a ligação com outra porção molecular. Os grupos de fluorocarbono e hidrofluorocarbono constituem grupos preferidos que activam um ácido carboxílico para a ligação a outra porção molecular.

3. Molécula de Interesse (MOI)

Os exemplos de MOIs incluem ácidos nucleicos ou análogos de ácidos nucleicos (p.ex., PNA), fragmentos de ácidos nucleicos, ácidos nucleicos sintéticos ou seus fragmentos, oligonucleótidos (p.ex., DNA ou RNA), proteínas, péptidos, anticorpos ou fragmentos de anticorpos, receptores, ligandos de receptores, membros de um par de ligandos, citoquinas, hormonas, oligossacáridos, moléculas orgânicas sintéticas, fármacos e combinações de todos estes elementos.

As MOIs preferidas correspondem a fragmentos de ácidos nucleicos. Os fragmentos de ácidos nucleicos preferidos constituem sequências de primer que são complementares de sequências presentes em vectores, sendo os vectores usados em sequenciação de bases. De preferência, um fragmento de ácido nucleico é ligado directa ou indirectamente a um marcador num local distinto da extremidade 3' do fragmento; e mais preferencialmente a nível da extremidade 5' do fragmento. Os fragmentos de ácidos nucleicos podem ser adquiridos comercialmente ou preparados a partir de bases de dados genéticas (p.ex., Dib et al., Nature 380:152-154, 1996 e CEPH Genotype Database, <http://www.cephb.fr>) e fornecedores comerciais (p.ex., Promega, Madison, WI).

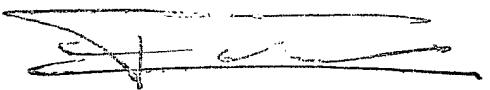


Tal como aqui é usado, o termo MOI inclui derivados de uma MOI que contêm funcionalidades úteis para a ligação da MOI ao composto T-L-L_h. Por exemplo, uma molécula de ácido nucleico que possui uma ligação fosfodiéster a nível da extremidade 5', em que o fosfodiéster se encontra igualmente ligado a uma alquilenamina, constitui uma MOI. Uma tal MOI encontra-se descrita, p.ex., na Patente dos EUA nº 4.762.779, que é aqui incorporada para referência. Um ácido nucleico contendo uma modificação interna constitui igualmente uma MOI. Um exemplo de modificação interna de um fragmento de ácido nucleico ocorre quando uma base (p.ex., adenina, guanina, citosina, timidina, uracilo) é modificada para adicionar um grupo funcional reactivo. Tais fragmentos de ácidos nucleicos com uma modificação interna são comercializados, p.ex., pela empresa Glen Research, Herndon, VA. Outro exemplo de modificação interna de um fragmento de ácido nucleico ocorre quando um fosforamidato abásico é usado para sintetizar um fosfodiéster modificado que é interposto entre um grupo açúcar e um grupo fosfato de um fragmento de ácido nucleico. O fosforamidato abásico contém um grupo reactivo que possibilita a ligação de um fragmento de ácido nucleico contendo esta porção molecular derivada do fosforamidato a uma outra porção molecular, p.ex. um composto T-L-L_h. Tais fosforamidatos abásicos são comercializados, p.ex., pela empresa Clonitech Laboratories, Inc., Palo Alto, CA.

4. Âncoras químicas (L_h)

Uma âncora química constitui um arranjo atómico estável mas reactivo presente como parte de uma primeira molécula, sendo que a âncora pode sofrer uma reacção química com uma âncora química complementar presente como parte de uma segunda molécula de modo a se formar uma ligação covalente entre as duas moléculas. Por exemplo, a âncora química poderá ser um grupo hidroxilo e a âncora química complementar poderá ser um grupo de ácido carboxílico (ou um seu derivado activado, p.ex., um éster de hidrofluoroarilo), sendo que a reacção entre estas duas âncoras forma uma ligação covalente (especificamente, uma ligação éster) que une as duas moléculas.

As âncoras químicas poderão ser usadas num grande número de reacções formadoras de ligações covalentes que são adequadas para ligar marcadores a linkers, e linkers a MOIs. Tais reacções incluem a alquilação (p.ex., para formar éteres, tioéteres), acilação (p.ex., para formar ésteres, amidas, carbamatos, ureias, tioureias), fosforilação (p.ex., para formar fosfatos, fosfonatos, fosforamidas, fosfonamidas), sulfonilação (p.ex. para formar sulfonatos, sulfonamidas), condensação (p.ex., para formar iminas, oximas, hidrazonas), sililação, formação de dissulfuretos e geração de intermediários reactivos, tais como nitrenos e carbenos, por fotólise. Em geral, as âncoras e as reacções de formação de



ligações que são adequadas para a ligação de marcadores a linkers são também apropriadas para a ligação de linkers a MOIs, e vice-versa. Em alguns casos, a MOI poderá sofrer uma prévia modificação ou derivatização para se obter a âncora necessária à ligação com o linker.

Um tipo de ligação especialmente útil para a ligação de linkers a MOIs é a ligação dissulfureto. A sua formação requer a presença de um grupo tiol ("âncora") no linker e de outro grupo tiol na MOI. É então suficiente utilizar condições suaves de oxidação para ligar os dois tióis sob a forma de dissulfureto. A formação da ligação dissulfureto poderá também ser induzida através do uso de um excesso de um reagente apropriado de transferência de dissulfuretos, p.ex., disulfuretos de piridilo. Uma vez que a formação de dissulfuretos é facilmente reversível, o dissulfureto poderá também ser usado como ligação susceptível de corte para libertar o marcador, se desejado. Tal é tipicamente realizado sob condições suaves semelhantes, usando um excesso de um reagente apropriado de transferência de tióis, p.ex., ditiotreitol.

A formação de ligações amida reveste-se de particular interesse na ligação de marcadores (ou marcadores com linkers) a oligonucleótidos. Podem introduzir-se facilmente âncoras de aminas alifáticas primárias em oligonucleótidos sintéticos com fosforamiditos, tal como o 6-monometoxitritilhexilcianoetil-N,N-diisopropil-fosforamidito (disponível através da Glenn Research, Sterling, VA). As aminas encontradas em nucleótidos naturais como a adenosina e a guanosina são virtualmente não reactivas quando comparadas com a amina primária introduzida. Esta diferença de reactividade faz com que seja possível a formação selectiva de amidas e grupos de ligação relacionados (p.ex., ureias, tioureias, sulfonamidas) com a amina primária introduzida e não com as aminas nucleotídicas.

Tal como listado no catálogo de Sondas Moleculares (Eugene, OR), uma enumeração parcial de grupos funcionais reactivos com aminas inclui ésteres carboxílicos activados, isocianatos, isotiocianatos, halatos de sulfônico e diclorotriazenos. Os ésteres activos constituem excelentes reagentes para a modificação de aminas, uma vez que os produtos de amida que se formam são muito estáveis. Adicionalmente, estes reagentes exibem uma boa reactividade com as aminas alifáticas e baixa reactividade com as aminas nucleotídicas dos oligonucleótidos. Os exemplos de ésteres activos incluem os ésteres de N-hidroxisuccinimida, os ésteres de pentafluorofenilo, os ésteres de tetrafluorofenilo e os ésteres de p-nitrofenilo. Os ésteres activos são também úteis por poderem ser produzidos a partir de virtualmente qualquer molécula que contenha um ácido carboxílico. Os métodos de produção de ésteres activos encontram-se listados em Bodansky (*Principles of Peptide Chemistry* (2^a ed.), Springer Verlag, Londres, 1993).

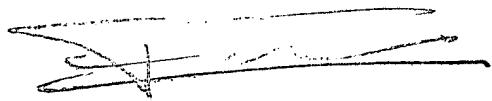
5. Ligação ao Linker

Tipicamente, é usado apenas um tipo de linker para ligar um conjunto ou família particular de marcadores a um conjunto ou família particular de MOIs. Numa forma de realização preferida para o invento, pode seguir-se um procedimento único e uniforme para criar todas as variantes de estruturas T-L-MOI. Tal torna-se especialmente vantajoso quando o conjunto de estruturas T-L-MOI é vasto, uma vez que permite a preparação do conjunto através da utilização dos métodos da química combinatorial ou de outra tecnologia de processamento em paralelo. De forma similar, o uso de um só tipo de linker permite a utilização de um procedimento único e uniforme para o corte de ligações em todas as variantes de estruturas T-L-MOI. De novo, isto torna-se vantajoso quando se tem um conjunto vasto de estruturas T-L-MOI, uma vez que o conjunto pode ser processado de forma paralela, repetitiva e/ou automatizada.

Existe, no entanto, outra forma de realização do presente invento segundo a qual são usados dois ou mais tipos de linker para ligar diferentes subconjuntos de marcadores a subconjuntos correspondentes de MOIs. Neste caso, poderão utilizar-se condições de corte selectivo de ligações para cortar cada um dos linkers de modo independente, sem cortar os linkers presentes em outros subconjuntos de MOIs.

Existem muitas reacções de formação de ligações covalentes que são adequadas para ligar os marcadores aos linkers, e os linkers às MOIs. Estas reacções incluem a alquilação (p.ex., para formar éteres, tioéteres), acilação (p.ex., para formar ésteres, amidas, carbamatos, ureias, tioureias), fosforilação (p.ex., para formar fosfatos, fosfonatos, fosforamidas, fosfonamidas), condensação (p.ex., para formar iminas, oximas, hidrazonas), sililação, formação de dissulfuretos e geração de intermediários reactivos, tais como os nitrenos ou carbenos, por fotólise. Em geral, as âncoras e as reacções de formação de ligações que são adequadas para efectuar a ligação dos marcadores aos linkers são também apropriadas para ligar os linkers às MOI, e vice-versa. Em alguns casos, a MOI poderá sofrer uma prévia modificação ou derivatização para fornecer a âncora necessária à ligação com o linker.

Um tipo de ligação especialmente útil para a ligação de linkers a MOIs é a ligação dissulfureto. A sua formação requer a presença de um grupo tiol ("âncora") no linker e de outro grupo tiol na MOI. É então suficiente utilizar condições suaves de oxidação para ligar os dois tióis sob a forma de dissulfureto. A formação da ligação dissulfureto poderá também ser induzida através do uso de um excesso de um reagente apropriado de transferência de dissulfuretos, p.ex., disulfuretos de piridílio. Uma vez que a formação de dissulfuretos é



facilmente reversível, o dissulfureto poderá também ser usado como ligação susceptível de corte para libertar o marcador, se desejado. Tal é tipicamente realizado sob condições suaves semelhantes, usando um excesso de um reagente apropriado de transferência de tióis, p.ex., ditiotreitol.

A formação de ligações amida reveste-se de particular interesse na ligação de marcadores a oligonucleótidos. Podem introduzir-se facilmente âncoras de aminas alifáticas primárias em oligonucleótidos sintéticos com fosforamiditos, tal como o 6-monometoxitritilhexilcianoetil-N,N-diisopropil-fosforamidito (disponível através da Glenn Research, Sterling, VA). As aminas encontradas em nucleótidos naturais como a adenosina e a guanosina são virtualmente não reactivas quando comparadas com a amina primária introduzida. Esta diferença de reactividade faz com que seja possível a formação selectiva de amidas e grupos de ligação relacionados (p.ex., ureias, tioureias, sulfonamidas) com a amina primária introduzida e não com as aminas nucleotídicas.

Tal como listado no catálogo de Sondas Moleculares (Eugene, OR), uma enumeração parcial de grupos funcionais reactivos com aminas inclui ésteres carboxílicos activados, isocianatos, isotiocianatos, haletos de sulfônico e diclorotriazenos. Os ésteres activos constituem excelentes reagentes para a modificação de aminas, uma vez que os produtos de amida que se formam são muito estáveis. Adicionalmente, estes reagentes exibem uma boa reactividade com as aminas alifáticas e baixa reactividade com as aminas nucleotídicas dos oligonucleótidos. Os exemplos de ésteres activos incluem os ésteres de N-hidroxisuccinimida, os ésteres de pentafluorofenilo, os ésteres de tetrafluorofenilo e os ésteres de p-nitrofenilo. Os ésteres activos são também úteis por poderem ser produzidos a partir de virtualmente qualquer molécula que contenha um ácido carboxílico. Os métodos de produção de ésteres activos encontram-se listados em Bodansky (*Principles of Peptide Chemistry* (2^a ed.), Springer Verlag, Londres, 1993).

Existem numerosos reagentes comerciais de cross-linking que podem usar-se como linkers (p.ex., ver Cross-linkers no Catálogo Pierce, Pierce Chemical Co., Rockford, IL). Entre estes encontram-se reagentes de cross-linking homobifuncionais reactivos com aminas que são exemplificados por imidoésteres homobifuncionais e ésteres de N-hidroxisuccinimidilo (NHS). Existem igualmente reagentes heterobifuncionais de cross-linking possuindo dois ou mais grupos reactivos diferentes que são responsáveis pelas reacções sequenciais. Os imidoésteres reagem rapidamente com aminas primárias ou secundárias. As maleimidas, os haletos de alquilo e arilo, os alfa-haloacilos e os dissulfuretos de piridilo reagem com tióis. As maleimidas são específicas para os grupos sulfidrilo dos tióis a pH 6,5-

7,5, e a pH alcalino podem tornar-se reactivas com aminas. A ligação tioéter é estável sob condições fisiológicas. Os reagentes de cross-linking de tipo alfa-haloacetilo contêm o grupo iodoacetilo e são reactivos face aos sulfidrilos. Os imidazóis têm a capacidade de reagir com a porção molecular de iodoacetilo, mas a reacção é muito lenta. Os dissulfuretos de piridilo reagem com os grupos tiol para formar uma ligação dissulfureto. As carbodiimidas estabelecem a ligação de carboxilos com aminas primárias de hidrazidas, o que dá origem à formação de uma ligação acil-hidrazina. As arilazidas constituem reagentes de fotoafinidade que são quimicamente inertes até serem expostas a radiação UV ou a luz visível. Quando estes compostos são fotolizados a 250-460 nm, forma-se um nitreno de arilo reactivo. O nitreno de arilo reactivo é relativamente não específico. Os glixoxais são reactivos face à porção de guanidinilo da arginina.

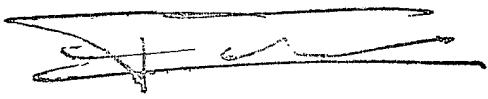
Numa forma de realização típica do presente invento, um marcador é inicialmente ligado a um linker; em seguida, a combinação de marcador e linker é ligada a uma MOI para criar a estrutura T-L-MOI. Alternativamente, a mesma estrutura é criada ligando inicialmente um linker a uma MOI, e ligando então o conjunto de linker e MOI a um marcador. Como exemplo temos a situação em que a MOI é um primer de DNA ou um oligonucleótido. Neste caso, o marcador é tipicamente ligado em primeiro lugar a um linker, sendo então o T-L ligado a um primer de DNA ou a um oligonucleótido, usando-se posteriormente o conjunto, por exemplo, numa reacção de sequenciação.

Um modo útil de ligar reversivelmente um marcador a uma MOI (p.ex., um oligonucleótido ou um primer para sequenciação de DNA) consiste em utilizar um linker quimicamente lável. Uma concepção preferida para o linker permite-lhe ser cortado quando é exposto a um ácido orgânico volátil, por exemplo, ácido trifluoroacético (TFA). O TFA, em particular, é compatível com a maioria dos métodos de ionização em espectrometria de massa, incluindo o electrospray.

O invento comprehende também composições para a análise de mutações. Uma composição útil para a análise de mutações comprehende um par de compostos com a fórmula:

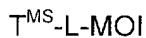


em que T^{MS} é um grupo orgânico detectável por espectrometria de massa, formado por carbono, pelo menos um átomo de hidrogénio e de flúor, e átomos opcionais seleccionados a partir do grupo formado por oxigénio, azoto, enxofre, fósforo e iodo. Na fórmula, L é um grupo orgânico que permite que se dê o corte da ligação entre a porção molecular contendo T^{MS} e a porção remanescente do composto, sendo que a porção molecular contendo T^{MS} comprehende um grupo funcional que suporta um estado ionizado de

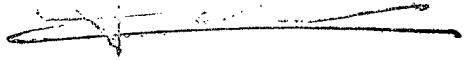


carga única quando o composto é submetido a espectrometria de massa e é seleccionado a partir do grupo de amina terciária, amina quaternária e ácido orgânico. Na fórmula, MOI corresponde a um fragmento de ácido nucleico, sendo que L é conjugado com MOI em outro local que não a extremidade 3' da MOI. A composição comprehende pares de compostos em que os membros de um par possuem grupos T^{MS} não idênticos e têm sequências idênticas excepto em uma posição de base, posição essa em que as bases não são idênticas. Em outra forma de realização da composição do invento, os membros do par de compostos possuem grupos T^{MS} não idênticos e têm sequências idênticas excepto em uma posição de base, na qual as bases não são idênticas. Estas composições são então adicionadas a uma sequência de ácido nucleico ligada a um suporte, sequência essa que é idêntica à sequência de um dos membros de cada par. Assim, o invento fornece uma composição que comprehende uma pluralidade de pares de compostos, tal como descrito acima, e que comprehende ainda uma mesma pluralidade de ácidos nucleicos imobilizados num suporte sólido, em que cada membro da pluralidade de ácidos nucleicos possui uma sequência de bases que é exactamente complementar à de um dos membros de cada par.

O invento proporciona ainda um kit para a análise de mutações que comprehende uma pluralidade de conjuntos de reacção. Cada conjunto de reacção comprehende um par de compostos com a fórmula:



em que T^{MS} é um grupo orgânico detectável por espectrometria de massa, formado por carbono, pelo menos um átomo de hidrogénio e de flúor, e átomos opcionais seleccionados a partir do grupo formado por oxigénio, azoto, enxofre, fósforo e iodo. Na fórmula, L é um grupo orgânico que permite que se dê o corte da ligação entre a porção molecular contendo T^{MS} e a porção remanescente do composto, sendo que a porção molecular contendo T^{MS} comprehende um grupo funcional que suporta um estado ionizado de carga única quando o composto é submetido a espectrometria de massa e é seleccionado a partir do grupo de amina terciária, amina quaternária e ácido orgânico. Na fórmula, MOI corresponde a um fragmento de ácido nucleico, sendo que L é conjugado com MOI em outro local que não a extremidade 3' da MOI. No kit, os compostos de cada par possuem grupos T^{MS} não idênticos e têm sequências idênticas excepto em uma ou duas posições de bases, posições essas em que as bases não são idênticas. Num kit preferido, a pluralidade é de pelo menos 3, e mais preferencialmente de pelo menos 5.

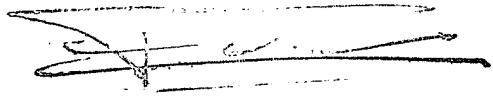


B. ENSAIOS

Tal como salientado acima, o presente invento contempla uma grande variedade de ensaios, nos quais os marcadores e a metodologia de detecção aqui apresentados podem ser utilizados para aumentar grandemente a sensibilidade e o rendimento do ensaio. Num dos aspectos do invento, poderão utilizar-se estes métodos para detectar a ligação de um primeiro membro a um segundo membro de um par de ligandos, compreendendo os passos de (a) combinar um conjunto de primeiros membros marcados com uma amostra biológica que poderá conter um ou mais segundos membros, sob condições e durante o tempo suficientes para permitir a ligação de um primeiro membro a um segundo membro, sendo que o dito marcador se correlaciona com um primeiro membro em particular e é detectável por espectrometria não fluorescente ou por potociometria, (b) separar o primeiro e o segundo membros ligados dos membros não ligados, (c) separar o marcador do primeiro membro marcado (por corte da ligação entre eles); e (d) detectar o dito marcador por espectrometria não fluorescente ou por potociometria, e deste modo detectar a ligação do primeiro membro ao segundo membro.

Poderá utilizar-se uma grande variedade de pares de primeiro e segundo membros dentro do contexto do presente invento, incluindo, por exemplo, moléculas de ácidos nucleicos (p.ex., DNA, RNA, análogos de ácidos nucleicos como o PNA, ou qualquer combinação destes elementos), proteínas ou polipéptidos (p.ex., um anticorpo ou fragmentos de um anticorpo; p.ex., anticorpos monoclonais, anticorpos policlonais ou parceiros de ligação como um CDR), oligossacáridos, hormonas, moléculas orgânicas e outros substratos (p.ex., xenobióticos como uma molécula de glucuronidase-fármaco), ou qualquer outro ligando de um par de ligandos. Dentro de várias formas de realização do invento, o primeiro e segundo membros poderão possuir o mesmo tipo de molécula ou pertencer a grupos de moléculas diferentes. Por exemplo, os pares representativos de ligandos de primeiro membro/segundo membro incluem: molécula de ácido nucleico/molécula de ácido nucleico; anticorpo/molécula de ácido nucleico; anticorpo/hormona; anticorpo/xenobiótico; e anticorpo/proteína.

Com o objectivo de aprofundar a compreensão dos ensaios que podem ser realizados no âmbito da presente exposição, é abaixo fornecida uma descrição de certos ensaios particularmente preferidos.



1. Ensaios de Ácidos Nucleicos

a. *Introdução*

Tal como salientado acima, o presente invento apresenta igualmente uma grande variedade de métodos no âmbito dos quais os marcadores e/ou linkers susceptíveis de corte acima descritos podem ser utilizados em substituição dos marcadores tradicionais (p.ex., marcadores radioactivos, fluorescentes ou enzimáticos), de modo a aumentar, para um dado método, a especificidade, a sensibilidade ou o número de amostras que podem ser analisadas em simultâneo. Os exemplos representativos de métodos que podem ser melhorados incluem as reacções convencionais de hibridização de ácidos nucleicos (ver Sambrook et al., *supra*), as reacções de diagnóstico como a Tecnologia de Ciclagem de Sonda (CPT) (ver Patentes dos EUA nos. 4.876.187 e 5.011.769) ou o Ensaio de Ligação entre Oligonucleótidos (OLA) (Burket et al., *Science* 196:180, 1987). Estas e outras técnicas serão abaixo apresentadas em maior detalhe.

b. *Técnicas de Hibridização*

O êxito da clonagem e sequenciação de um gene permite a investigação da sua estrutura e expressão ao tornar possível a detecção deste gene ou do seu mRNA a partir de um vasto "pool" de moléculas de DNA ou RNA não relacionadas entre si. A quantidade de mRNA que codifica para uma proteína específica presente num tecido é um parâmetro importante para a actividade de um gene e pode relacionar-se significativamente com a actividade dos sistemas funcionais. A sua regulação depende da interacção entre sequências presentes no interior do gene (elementos cis-actuantes) e proteínas de ligação ao DNA específicas de sequência (factores trans-actuantes), que são activadas especificamente por cada tecido ou por sistemas de hormonas e segundos mensageiros.

Encontram-se disponíveis diversas técnicas para a análise de um gene em particular, das suas sequências reguladoras, do seu mRNA específico e da regulação da sua expressão; estas técnicas incluem a análise por Southern Blot ou por Northern Blot, o ensaio de protecção de ribonuclease (RNase) e a hibridização *in situ*.

As variações da composição nucleotídica de um dado gene poderão possuir uma grande relevância fisiopatológica. Quando as variações ocorrem em regiões não codificantes (regiões flanqueadoras 5', 3' e intron), a regulação da expressão genética poderá ser afectada, causando uma activação ou inibição anormais. Quando as variações se localizam nas regiões codificante do gene (exons), estas podem resultar na alteração do funcionamento das proteínas ou na produção de proteínas disfuncionais.

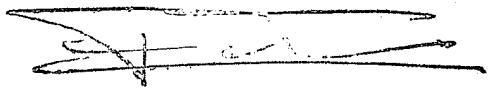
Assim, uma dada sequência de um gene pode correlacionar-se com uma doença específica e poderá ter utilidade como marcador da doença. Um dos objectivos principais da pesquisa no campo da medicina é, deste modo, o de detectar estas variações genéticas para as utilizar como ferramentas de diagnóstico e reunir informações importantes para o conhecimento dos fenómenos fisiopatológicos.

O método básico para análise de uma população relativamente às variações que ocorrem em determinado gene é a análise de DNA através da técnica de Southern Blot. Abreviadamente, é feita a digestão de DNA genómico preparado com uma enzima de restrição (RE), obtendo-se um grande número de fragmentos de DNA com diferentes extensões, determinados pela presença do local específico de reconhecimento da RE no genoma. Os alelos de um dado gene que possuam mutações no interior deste local de restrição serão cortados em fragmentos que são diferentes no que respeita ao seu número e extensão. Este fenómeno é designado por polimorfismo da extensão dos fragmentos de restrição (RFLP) e pode constituir um importante marcador de diagnóstico, possuindo diversas aplicações.

O fragmento a ser analisado é separado do "pool" de fragmentos de DNA e distinguido de outras espécies de DNA através da utilização de uma sonda específica. Assim, o DNA é submetido a fraccionamento electroforético usando um gel de agarose, seguindo-se a sua transferência e fixação a uma membrana de nylon ou de nitrocelulose. O DNA de cadeia simples que foi fixado é hibridizado com um DNA marcado que é complementar ao DNA a detectar. Após a remoção das hibridizações não específicas, o fragmento de DNA de interesse poderá ser visualizado por espectrometria de massa MALDI (MALDI-MS), tal como abaixo se expõe em maior detalhe.

A presença e a quantificação de um produto específico de transcrição de um gene e a sua regulação através de parâmetros fisiológicos pode ser analisada através da técnica de Northern Blot e do ensaio de protecção de Rnase.

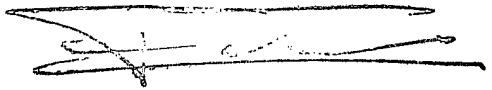
Estes métodos baseiam-se na hibridização de um "pool" de RNA celular total a uma sonda específica. Na técnica de Northern Blot, o RNA total de um tecido é fraccionado por electroforese usando um gel de agarose, sendo depois transferido e imobilizado sobre um RNA anti-sentido marcado (cDNA), complementar ao RNA a detectar. Esta sonda de cRNA é então marcada como aqui se descreve. Através da aplicação de condições de lavagem rigorosas, as moléculas ligadas de modo inespecífico são eliminadas. As moléculas especificamente ligadas podem então ser detectadas por MALDI-MS. Adicionalmente, a especificidade pode ser controlada através da comparação das dimensões do mRNA detectado com as dimensões previstas para o mRNA de interesse.



Mais rápido mas menos específico, o método "dot-blot" é executado à semelhança da técnica de Northern Blot, diferindo no facto de o RNA ser transferido directamente para a membrana sem se fazer um fraccionamento prévio. Assim, no dot-blot, o RNA é imobilizado inespecificamente.

O método mais específico para a detecção de uma espécie de mRNA é o ensaio de protecção de Rnase. Abreviadamente, o RNA total de um tecido ou cultura celular é hibridizado com um cRNA específico marcado de homologia completa. A especificidade é conseguida mediante uma digestão subsequente por RNase. Os fragmentos não hibridizados da cadeia simples de RNA e os fragmentos inespecificamente hibridizados, mesmo com pequenas não-correspondências, serão reconhecidos e cortados, enquanto que o RNA de cadeia dupla com homologia completa não é acessível à enzima e estará protegido. Após a remoção da RNase por digestão com proteinase K e extracção por fenol, o fragmento protegido específico pode ser separado dos produtos de degradação, geralmente num gel desnaturante de poliacrilamida, e a extensão prevista pode ser verificada por HPLC. Todos os ensaios acima descritos podem ser quantificados por espectrometria não fluorescente ou por potenciometria.

A localização precisa de um dado mRNA numa população específica de células que fazem parte de um tecido pode ser determinada através de hibridização in-situ. Este método é análogo à técnica imunocitoquímica e pode, de facto, ser utilizado em simultâneo com a imunocitoquímica na mesma secção para descobrir, por exemplo, se uma certa proteína é realmente sintetizada a nível local ou se é captada a partir de outras fontes. Para além da possibilidade de identificar o tipo celular que expressa um tipo específico de mRNA, a hibridização in situ poderá revelar-se um método ainda mais sensível do que a análise de uma preparação de RNA de tecido total utilizando as técnicas acima descritas. Isto acontece quando o mRNA é expresso em elevadas concentrações por uma região ou tipo celular muito restritos no interior do tecido, que seriam diluídos por homogeneização do tecido total. A análise da expressão genética através da hibridização in-situ reveste-se, assim, de particular importância em tecidos heterogéneos, como o cérebro. Para a hibridização in-situ, os tecidos têm de ser congelados ou fixados por perfusão e seccionados de acordo com o protocolo histoquímico. O protocolo de hibridização para as secções de tecido e para as sondas marcadas é semelhante ao de outros métodos de hibridização acima descritos. É possível realizar uma análise semiquantitativa.



c. cDNAs como Populações Representativas de mRNAs, e seu Uso como Sondas

A maioria dos mRNAs é transcrita a partir de sequências de cópia simples. Outra propriedade dos cDNAs é a de que estes representam uma região mais longa do genoma, devido aos introns que se encontram presentes na versão cromossómica da maioria dos genes. A representação varia de gene para gene mas poderá ser bastante significativa, já que muitos genes abrangem mais do que 100 kb no DNA genómico, extensão representada num só cDNA. Uma utilização possível para a hibridização molecular consiste no uso de sondas de uma espécie para detectar clones provenientes de outra espécie. A divergência de sequências entre os mRNA do rato e do homem permite a reassociação específica cruzada de sequências longas, mas, exceptuando o caso das regiões altamente conservadas, evita a hibridização cruzada dos primers de PCR.

A análise diferencial em amostras biológicas complexas, tal como o sistema nervoso em desenvolvimento, utilizando sondas de cDNA produzidas a partir de células isoladas é agora possível graças ao desenvolvimento das técnicas de amplificação baseadas em PCR e em cRNA. Diversos grupos de trabalho referiram já a criação de colecções de cDNA a partir de pequenas quantidades de poli(A) + RNA (1 ng ou menos) preparadas a partir de 10-50 células (Belyav et al., *Nuc. Acids Res.* 17:2919, 1989). Se bem que estas colecções se tenham mostrado suficientemente representativas da complexidade do mRNA, a dimensão média do fragmento de inserção do cDNA nestas colecções era bastante pequeno (< 2 kb).

Mais recentemente, combinaram-se várias metodologias para gerar sondas baseadas em PCR (Lambolez et al., *Neuron* 9:247, 1992) e baseadas em cRNA (Van Gelder et al., *Proc. Natl. Acad. Sci. USA* 87:1663, 1990) a partir de células isoladas. Após recolha de registos eléctricos, o conteúdo citoplasmático de uma célula isolada foi aspirado através de microeléctrodos de fixação para a realização da síntese e amplificação *in situ* do cDNA. Foi utilizada a técnica de PCR para amplificar o cDNA de mRNAs do receptor selectivo do glutamato de células de Purkinje isoladas, e do mRNA GFAP de células isoladas da glia, provenientes de uma cultura cerebelar organotípica (Lambolez et al., *Neuron* 9:247, 1992). No caso da amplificação do cRNA, foram concebidas sequências promotoras da transcrição sob a forma de primers para síntese de cDNA, gerando-se cRNAs antisentido complexos por transcrição com RNA polimerases de bacteriófagos.

Assim, numa das formas de realização do invento, poderão ser utilizados cRNAs marcados como sondas marcadas para analisar colecções de cDNA aleatoriamente ou em experiências de "definição do perfil de expressão", analisando produtos de Southern Blot que contenham fragmentos de cDNA com interesse (receptores, factores de crescimento,

canais iónicos, etc.). Parece que a falta de linearidade da amplificação, frequentemente encontrada nas abordagens baseadas em PCR, é minimizada com os métodos baseados em cRNAs.

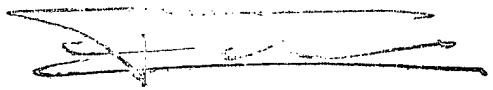
d. Ensaio de Ligação Entre Oligonucleótidos

O ensaio de ligação entre oligonucleótidos constitui uma extensão das análises por PCR que utiliza um ensaio baseado em ELISA (OLA, Nickerson et al., Proc. Natl. Acad. Sci. USA 87:8923, 1990) para detectar os produtos de PCR que contêm as sequências-alvo. Assim, são eliminados os processos de electroforese em gel e de hibridização em colónias. Abreviadamente, o OLA utiliza dois nucleótidos adjacentes: uma sonda "repórter" (marcada a nível da extremidade 5') e uma sonda de "âncora" 5'-fosforilada/3'-biotinilada. Os dois oligonucleótidos, que são complementares a sequências internas dos primers de PCR, são submetidos a annealing com o DNA-alvo e, se existir complementaridade perfeita, as duas sondas são ligadas por acção da DNA ligase de T4. A captura da sonda de âncora biotinilada em estreptavidina imobilizada e a análise para detecção da sonda repórter covalentemente ligada permitem testar a presença ou ausência de sequências-alvo entre os produtos de PCR.

e. Aplicação de Técnicas de Hibridização

i. Área Forense

A identificação de indivíduos a nível da variação de sequências de DNA oferece diversas vantagens práticas sobre os critérios convencionais como as impressões digitais, o tipo sanguíneo ou as características físicas. Em contraste com a maioria dos marcadores fenotípicos, a análise de DNA permite prontamente a dedução de relações entre os indivíduos, tal como requerido no teste de paternidade. A análise genética mostrou-se de grande utilidade em transplantes de medula óssea, em que é necessário distinguir entre as células estreitamente relacionadas do dador e do recipiente. São actualmente utilizados dois tipos de sondas para análise do DNA por técnicas de Blot. As sondas de DNA polimórficas de mini-satélite identificam sequências múltiplas de DNA, cada uma das quais se encontra presente sob formas variáveis em diferentes indivíduos, gerando-se deste modo padrões complexos e altamente variáveis entre indivíduos. As sondas VNTR identificam sequências isoladas no genoma, mas estas sequências poderão estar presentes em até 30 formas diferentes na população humana, distinguindo-se pela dimensão dos fragmentos identificados. A probabilidade de que indivíduos não relacionados possuam padrões idênticos de hibridização com múltiplas VNTR ou sondas mini-satélite é muito baixa. Uma



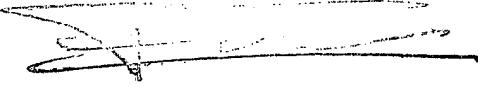
quantidade de tecido muito menor do que a requerida para produzir os blots de DNA, até um cabelo isolado, fornece DNA suficiente para realizar uma análise baseada em PCR dos marcadores genéticos. Do mesmo modo, pode utilizar-se tecido parcialmente degradado para análise, uma vez que são apenas necessários pequenos fragmentos de DNA. As análises forenses de DNA serão eventualmente realizadas sobre sequências polimórficas de DNA que podem ser estudadas através de ensaios automatizados simples, como o OLA. Por exemplo, a análise de 22 sequências de genes distintas, estando cada uma delas presente sob duas formas diferentes na população, poderá gerar 10¹⁰ resultados diferentes, permitindo a identificação individualizada de seres humanos.

ii. Diagnóstico de Tumores

A detecção de oncogenes virais ou celulares constitui outro campo de aplicação importante das técnicas de diagnóstico com ácidos nucleicos. Os oncogenes virais (oncogenes v) são transmitidos por retrovírus, enquanto que os seus correspondentes celulares (oncogenes c) se encontram já presentes em células normais. Os oncogenes celulares podem, no entanto, ser activados por modificações específicas, tais como mutações pontuais (como acontece com o oncogene c-K-ras no carcinoma da bexiga e nos tumores colorectais), a indução de promotores, a amplificação de genes (como acontece com o gene N-myc no neuroblastoma) ou o rearranjo de cromossomas (como na translocação do oncogene c-abl do cromossoma 9 para o cromossoma 22, na leucemia mielóide crónica). Cada um dos processos de activação conduz, juntamente com processos degenerativos adicionais, a um crescimento celular aumentado e incontrolado. Os chamados "oncogenes recessivos" que têm de ser inactivados para que se dê a formação de um tumor (como acontece com o gene Rb do retinoblastoma e no osteosarcoma) podem também ser detectados com a ajuda de sondas de DNA. Utilizando sondas contra os genes de imunoglobulinas e contra os genes dos receptores de células T, é possível a detecção de linfomas a células B e da leucemia linfoblástica.

iii. Análise de Transplantes

A reacção de rejeição de tecidos transplantados é controlada de forma decisiva por uma classe específica de抗énios de histocompatibilidade (HLA). Estes são expressos à superfície das células sanguíneas apresentadoras do抗énio, p.ex., os macrófagos. O complexo formado entre o HLA e o抗énio estranho é reconhecido pelas células T-helper através de receptores correspondentes nas células T que estão presentes na superfície celular. A interacção entre o HLA, o抗énio e o receptor da célula T desencadeia uma



complexa reacção de defesa que leva a uma resposta imune "em cascata" por parte do organismo.

O reconhecimento de diferentes抗énios estranhos é mediado por regiões variáveis específicas de抗énio presentes no receptor da célula T – num processo análogo ao da reacção抗énio-anticorpo. Na rejeição de um transplante, as células T expressando um receptor de células T específico que reconhece o抗énio estranho poderão, desta forma, ser eliminadas do "pool" de células T. Estas análises são possibilitadas pela identificação de sequências variáveis de DNA específicas do抗énio que são amplificadas por PCR e que, desta forma, sofrem um incremento selectivo. A reacção de amplificação específica permite a identificação isolada, específica de célula, de um receptor de célula T específico.

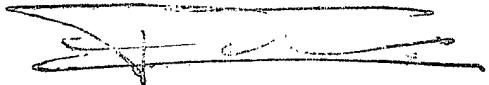
São presentemente efectuadas reacções semelhantes para a identificação de doenças autoimunes como a diabetes juvenil, a arteriosclerose, a esclerose múltipla, a artrite reumatóide ou a encefalomielite.

iv. Diagnóstico do Genoma

Quatro por cento dos recém-nascidos nascem com defeitos genéticos; de entre as 3.500 doenças hereditárias descritas que são causadas pela modificação de um só gene, só em cerca de 400 se conhecem os defeitos moleculares primários que as causam.

As doenças hereditárias são desde há muito diagnosticadas através de análises fenotípicas (anamneses, p.ex., deficiências do sangue: talassémias), análises dos cromossomas (cariotipo, p.ex., mongolismo: trissomia 21) ou análises de produtos dos genes (proteínas modificadas, p.ex., fenilcetonúria: deficiência do enzima fenilalanina hidroxilase que resulta no aumento de concentrações do ácido fenilpirúvico). O uso adicional de métodos de detecção de ácidos nucleicos alarga consideravelmente o intervalo de utilização das técnicas de diagnóstico do genoma.

No caso de certas doenças genéticas, a modificação de apenas um de dois alelos é suficiente para que a doença se manifeste (defeitos monogénicos de transmissão dominante); em muitos casos, é necessária a modificação dos dois alelos (defeitos monogénicos de transmissão recessiva). Num terceiro tipo de defeito genético, o desencadeamento da doença não é determinado apenas pela modificação do gene mas também por factores como os hábitos alimentares (no caso da diabetes ou da arteriosclerose) ou o estilo de vida (no caso do cancro). Muito frequentemente, estas doenças ocorrem em idades avançadas. As doenças como a esquizofrenia, a depressão maníaca ou a epilepsia devem também ser mencionadas neste contexto; encontra-se em



investigação a hipótese de o desencadeamento da doença nestes casos se encontrar dependente de factores ambientais, juntamente com a modificação de diversos genes em diferentes localizações cromossómicas.

Utilizando a análise directa e indirecta de DNA, tornou-se possível efectuar o diagnóstico de uma série de doenças genéticas: anemia de células falciformes, talassémias, deficiência em α -1-antitripsina, síndrome de Lesch-Nyhan, fibrose quística/mucoviscidose, distrofia muscular de Duchenne-Becker, doença de Alzheimer, deficiência mental dependente do cromossoma X, coreia de Huntington.

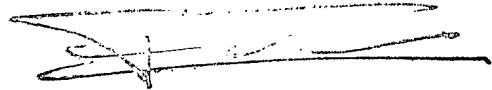
v. Doenças Infecciosas

A aplicação de métodos de DNA recombinante ao diagnóstico de doenças infecciosas tem sido mais exaustivamente explorada para as infecções virais, já que o diagnóstico actual destas infecções se revela ainda trabalhoso e demorado. A hibridização in situ em tecidos ou culturas de células tornou possível o diagnóstico da infecção aguda e crónica por herpes. Foi já referida a utilidade da utilização de tecidos frescos ou fixados por formalina na detecção do papilomavírus no carcinoma invasivo da cérvix uterina e na detecção do HIV, tendo as culturas celulares já sido usadas para a detecção do citomegalovírus e do vírus de Epstein-Barr. A aplicação de métodos de DNA recombinante ao diagnóstico de doenças microbianas possui potencial para substituir os métodos actuais baseados no crescimento microbiano, se for possível satisfazer certos requisitos em termos de razão custo-eficácia, rapidez e precisão. As situações clínicas em que foi já iniciada a aplicação de procedimentos de DNA recombinante incluem a identificação de: *Neisseria gonorrhoeae* resistente à penicilina, através da presença de um transposão; o género *Chlamydia*, muito exigente em termos de condições de crescimento; micróbios em alimentos; e ainda a aplicação de métodos simples para o seguimento da disseminação de uma infecção no seio de uma comunidade. O desafio epidemiológico que as doenças causadas por parasitas como a leishmania e o plasmódio colocam a nível mundial está já a ser alvo de estudo por métodos recombinantes.

2. Ensaios Baseados em Proteínas

a. Introdução

Tal como acima foi referido, uma grande variedade de ensaios baseados em proteínas pode igualmente ser melhorada por utilização dos marcadores aqui descritos (ver, p.ex.: *Antibodies: A Laboratory Manual*, Harlow e Lane (eds.), Cold



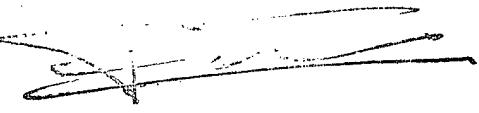
Spring Harbor Press, 1988. Os exemplos representativos incluem ensaios de antigénio-anticorpo tais como: imunolectroforese em contra-corrente (CIEP), ensaios ELISA, ensaios de competição ou inibição e ensaios de sanduíche, imunoensaios em simultâneo e ensaios de imunofiltração. No entanto, poderá ser melhorada uma grande variedade de outros ensaios, incluindo, por exemplo, os ensaios receptor-ligando e semelhantes.

b. Imunoensaios

Desde o desenvolvimento dos métodos de RIA para a análise da insulina e da tiroxina, os antigénios marcados com radioisótopos têm sido largamente utilizados na medição de moléculas de haptenos, tais como hormonas e fármacos. Estes métodos baseiam-se na competição entre um antigénio marcado e um antigénio não marcado para uma quantidade limitada de anticorpo. Estes métodos poderão igualmente ser descritos como "métodos de reagente limitado", devido à limitada quantidade de anticorpo que é utilizada no ensaio.

Se bem que a utilização de anticorpos marcados em métodos de imunofluorescência tenha sido iniciada em 1941, a sua aplicação mais alargada a métodos quantitativos só se iniciou com a introdução de anticorpos marcados com radioisótopos, nos métodos IRMA. Os ensaios IRMA, tal como outros ensaios de duplo anticorpo ou de "sanduíche" baseados em fase sólida (ELISA, IFMA, ensaios de marcação por imunofluorescência), caracterizam-se pela existência de um excesso de anticorpos relativamente aos antigénios; desta forma, poderão designar-se por "métodos de excesso de reagente". Em princípio, o uso de reagentes em excesso reduz o tempo de incubação e aumenta potencialmente a sensibilidade. A fase sólida facilita a separação, e o sinal é directamente proporcional à quantidade de antigénio – o que se opõe à relação inversa característica dos ensaios competitivos.

O uso da tecnologia de avidina-biotina tem assumido uma importância crescente em numerosas áreas da biotecnologia, biologia molecular e medicina, incluindo a detecção de proteínas através de ensaios não radioactivos, a coloração citoquímica, a separação de células e o isolamento de ácidos nucleicos e detecção de sequências específicas de DNA/RNA por hibridização. A técnica deve a sua utilidade à afinidade extremamente elevada da interacção avidina-biotina (constante de associação 1015M⁻¹) e à possibilidade de biotinilação de uma grande variedade de biomoléculas-alvo, tais como anticorpos, ácidos nucleicos e lípidos. O primeiro passo no isolamento de uma molécula-alvo corresponde à


sua biotinilação, ou à biotinilação de uma biomolécula que se ligará subsequentemente à molécula-alvo (p.ex., um anticorpo ou sonda de hibridização que forma um complexo-alvo). A molécula biotinilada do complexo-alvo é então separada de outras moléculas presentes numa mistura heterogénea através da utilização de meios de afinidade que se baseiam nas interacções entre a avidina e a biotina.

Assim, em uma das formas de realização do invento, poderá realizar-se qualquer tipo de imunoensaio convencional utilizando reagentes marcados não radioactivamente, em lugar dos reagentes típicos que incluem marcadores de isótopos. Estes métodos resultam num grande aumento da sensibilidade, assim como na capacidade de analisar um grande número de amostras em simultâneo.

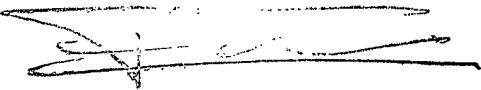
3. Análise da expressão genética

Um dos inventos aqui apresentados consiste num método de alto rendimento para a análise da expressão de numerosos genes (1-2000) numa só medição. O método permite também que esta análise seja feita em paralelo com um número superior a cem amostras por processo. O método é aplicável à monitorização de fármacos, à biologia do desenvolvimento, a estudos de medicina molecular e semelhantes. Assim, num aspecto do invento são fornecidos métodos para análise do padrão de expressão genética a partir de uma amostra biológica seleccionada, compreendendo os passos de (a) expôr os ácidos nucleicos de uma amostra biológica; (b) combinar os ditos ácidos nucleicos expostos com uma ou mais sondas marcadas de ácidos nucleicos seleccionadas, sob condições e durante o tempo suficiente para que as ditas sondas se hibridizem com os ditos ácidos nucleicos, sendo que o dito marcador se correlaciona com uma sonda de ácidos nucleicos em particular e é detectável por espectrometria ou potenciometria; (c) separar as sondas hibridizadas das não hibridizadas; (d) separar o marcador do fragmento marcado; e (e) detectar o dito marcador por espectrometria não fluorescente ou potenciometria, e desse modo determinar o padrão de expressão genética da dita amostra biológica.

Numa forma particularmente preferida de realização do invento, são fornecidos ensaios ou métodos que se descrevem tal como segue: o RNA proveniente de uma fonte-alvo é ligado a um suporte sólido através de um passo de hibridização específico (i.e., captura de mRNA poli(A) através de uma sonda de captura com oligo(dT)). O suporte sólido é então lavado e o cDNA é sintetizado sobre o suporte sólido utilizando métodos convencionais (i.e., transcriptase reversa). A cadeia de RNA é então removida por hidrólise. O resultado é a criação de uma população de DNA que se encontra covalentemente imobilizada sobre o suporte sólido e que reflecte a diversidade, abundância e complexidade

do RNA a partir do qual foi sintetizado o cDNA. O suporte sólido é então interrogado (hbridizado) com sondas que se encontram presentes em número que varia entre 1 e vários milhares e que são complementares a uma sequência genética de interesse. Cada tipo de sonda está marcado com um marcador de espectrometria de massa susceptível de corte ou com outro tipo de marcador suscetível de corte. Após o passo de interrogação, o excesso de sonda ou a sonda não hibridizada é removido por lavagem, o suporte sólido é colocado em (por exemplo) um poço de uma placa de microtitulação e o marcador de espectrometria de massa é removido do suporte sólido por uma reacção de corte de ligações. O suporte sólido é removido do poço da placa que contém a amostra e o conteúdo do poço é medido num espectrómetro de massa. O aparecimento de marcadores específicos de espectrometria de massa indica a presença de RNA na amostra e prova que um gene específico é expresso numa dada amostra biológica. O método poderá igualmente ser quantificável.

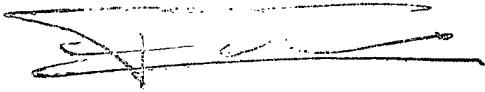
As composições e metodologias desenvolvidas para a medição rápida da expressão genética usando marcadores susceptíveis de corte podem descrever-se em detalhe tal como se segue. Abreviadamente, poderão usar-se como fontes de RNA um tecido (figado, músculo, etc.), linhas celulares primárias ou transformadas, tipos celulares isolados ou purificados ou qualquer outra fonte de material biológico para o qual se torna útil a determinação da expressão genética. No método preferido, o material da fonte biológica é lisado na presença de um agente caotrópico, de modo a suprimir nucleases e proteases e a apoiar a hibridização rigorosa da molécula-alvo de ácido nucleico com o suporte sólido. Os tecidos, células e fontes biológicas poderão ser lisados de modo eficaz utilizando sais caotrópicos em concentrações de 1 a 6 molar (cloridrato de guanidina, tiocianato de guanidina, perclorato de sódio, etc.). Após a lise da amostra biológica de fonte, a solução é misturada com um suporte sólido para efectuar a captura da molécula-alvo de ácido nucleico presente no lisado. Em uma alteração do método, o RNA é capturado usando uma sonda de captura com cauda de oligo(dT). Os suportes sólidos poderão incluir esferas de nylón, microsferas de polistireno, esferas de vidro e superfícies de vidro, ou qualquer outro tipo de suporte sólido ao qual seja possível ligar oligonucleótidos de modo covalente. Os suportes sólidos são preferencialmente revestidos com um polímero de amina, tal como a polietileno(imina), a acrilamida, os amina-dendrímeros, etc. Os grupos amina dos polímeros são utilizados para a imobilização covalente dos oligonucleótidos. Os oligonucleótidos são preferencialmente sintetizados usando uma amina 5' (geralmente uma hexilamina que inclui um braço espaçador com seis carbonos e uma amina distal). Os oligonucleótidos poderão possuir uma extensão de 15 a 50 nucleótidos. Os oligonucleótidos são activados com



reagentes de cross-linking homo-bifuncionais ou hetero-bifuncionais, tal como o cloreto cianúrico. Os oligonucleótidos activados são purificados do excesso de reagente de cross-linking (i.e., cloreto cianúrico) por cromatografia de exclusão. O oligonucleótilo activado é então misturado com os suportes sólidos para a realização da ligação covalente. Após a ligação covalente dos oligonucleótidos, as aminas do suporte sólido que não reagiram são protegidas com uma cauda (i.e., com anidrido succínico) para eliminar a carga positiva do suporte sólido.

Os suportes sólidos podem ser usados em paralelo e são de preferência configurados num formato de 96 poços ou de 384 poços. Os suportes sólidos poderão ser ligados a ganchos, hastes ou varetas numa configuração de 96 poços ou de 384 poços, podendo os suportes sólidos ser destacáveis ou íntegros para uma dada configuração. A configuração particular dos suportes sólidos não assume uma importância crítica para o funcionamento do ensaio, mas afecta a capacidade de adaptação do ensaio a um funcionamento automatizado.

Os suportes sólidos são misturados com o lisado durante um período que varia desde 15 minutos até várias horas, de modo a se efectuar a captura da molécula-alvo de ácido nucleico pelo suporte sólido. Em geral, a "captura" da molécula-alvo de ácido nucleico é feita através do emparelhamento complementar de bases entre o RNA alvo e a sonda de captura que se encontra imobilizada no suporte sólido. Uma variante utiliza a cauda de 3' poli(A) encontrada na maioria dos RNA mensageiros eucarióticos, hibridizando-a com uma cauda de oligo(dT) presente no suporte sólido. Outra variante consiste na utilização de um oligonucleótilo específico ou de sondas extensas (com mais do que 50 bases) para capturar um RNA que possui uma sequência definida. Outra possibilidade é a de empregar primers degenerados (oligonucleótidos) que efectuariam a captura de numerosas sequências relacionadas da população de RNA-alvo. Os tempos de hibridização são ditados pela complexidade das sequências da população de RNA e pelo tipo de sonda de captura utilizada. As temperaturas de hibridização são determinadas pelo tipo de agente caotrópico utilizado e pela concentração final do agente caotrópico (ver Van Ness e Chen, *Nuc. Acids Res.*, para obter orientações gerais). O lisado é de preferência agitado em contínuo juntamente com o suporte sólido, de modo a efectuar a difusão do RNA alvo. Uma vez concluído o passo de captura da molécula-alvo de ácido nucleico, o lisado é removido do suporte sólido por lavagem, e toda a solução de agente caotrópico ou de solução de hibridização é removida. O suporte sólido é preferencialmente lavado com soluções que contêm detergentes iónicos ou não iónicos, tampões e sais. O passo seguinte consiste na síntese de DNA complementar ao RNA capturado. Neste passo, o oligonucleótilo de



captura serve como primer de extensão para a transcriptase reversa. A reacção é geralmente efectuada a uma temperatura entre 25°C e 37°C, sendo preferencialmente feita uma agitação durante a reacção de polimerização. Após a síntese do cDNA, este fica covalentemente ligado ao suporte sólido, uma vez que o oligonucleótido de captura serviu como primer de extensão. O RNA é então hidrolisado a partir do duplex de cDNA/RNA. O passo pode ser efectuado por acção do calor, que desnatura o duplex, ou por acção de base (i.e., NaOH 0,1 N), para hidrolisar quimicamente o RNA. O resultado pretendido para este passo é o de tornar o cDNA disponível para hibridização subsequente com sondas definidas. O suporte sólido ou conjunto de suportes sólidos são então lavados para remover o RNA ou os fragmentos de RNA. Neste ponto, o suporte sólido contém uma população aproximadamente representativa de moléculas de cDNA que representa a população de RNA em termos de abundância de sequências, complexidade e diversidade.

O passo seguinte consiste em hibridizar sondas seleccionadas com o suporte sólido, de modo a identificar a ausência ou presença e abundância relativa de sequências específicas de cDNA. As sondas correspondem preferencialmente a oligonucleótidos com uma extensão de 15 a 50 nucleótidos. A sequência das sondas é ditada pelo objectivo final do ensaio. Por exemplo, se o objectivo correspondesse ao estudo da expressão genética de uma resposta inflamatória num tecido, seleccionar-se-iam sondas complementares a numerosos mRNAs de citoquinas, RNAs que codificam para enzimas que modulam lípidos, RNAs que codificam para factores que regulam as células envolvidas numa resposta inflamatória, etc. Uma vez definido para estudo um conjunto preciso de sequências, cada sequência é construída sob a forma de sonda oligonucleotídica e a cada sonda é atribuído um marcador específico susceptível de corte. O(s) marcador(es) é(são) então ligado(s) ao(s) oligonucleótido(s) correspondente(s). O(s) oligonucleótido(s) é(são) então hibridizado(s) com o cDNA do suporte sólido sob condições de hibridização apropriadas. Após a conclusão do passo de hibridização, o suporte sólido é lavado para remover as moléculas de sonda não hibridizada. O suporte sólido ou conjunto de suportes são, em seguida, colocados em soluções que efectuam o corte de ligações dos marcadores de espectrometria de massa. Os marcadores de espectrometria de massa são então submetidos a análise por espectrometria de massa, a massa de cada marcador presente é identificada e é determinada a presença (e abundância) ou ausência de um mRNA expresso.

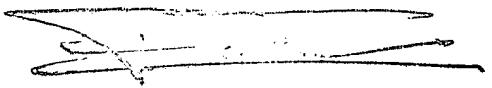
4. Detecção de Microorganismos, de uma Expressão Genética Específica ou de Sequências Específicas num Ácido Nucleico

O uso de sondas de DNA possuindo marcadores susceptíveis de corte poderá ter utilidade na detecção da presença ou ausência de microorganismos em qualquer tipo de amostra ou espécime. Tipicamente, a amostra será submetida a um passo de lise utilizando detergentes iónicos ou agentes caotrópicos, sendo então o ácido nucleico imobilizado especificamente ou inespecificamente sobre um suporte sólido e sondado em seguida por meio de sondas marcadas de DNA. As moléculas de sonda não hibridizadas são removidas num passo de lavagem e os marcadores são separados das suas respectivas sondas por meio de reacções de corte de ligações e, em seguida, medidos.

Os ácidos nucleicos detectáveis podem incluir mRNA, DNA genómico, DNA ou RNA plasmídico, rRNA e DNA ou RNA viral. Para efectuar a detecção da molécula-alvo de ácido nucleico, o alvo requer algum tipo de imobilização, uma vez que os ensaios aqui descritos não são homogéneos. São possíveis dois tipos de imobilização, não específica e específica. No primeiro caso, os ácidos nucleicos são imobilizados num suporte ou substrato sólido que possui um certo grau de afinidade para o ácido nucleico. Os ácidos nucleicos poderão ser purificados ou não purificados antes de se proceder a uma imobilização não específica. Os suportes sólidos poderão incluir membranas de nylon, membranas compostas por nitrocelulose, etc. Os suportes sólidos são então sondados com oligonucleótidos marcados de sequência pré-determinada para identificar a molécula-alvo de ácido nucleico com interesse. As moléculas de sonda não hibridizadas são removidas num passo de lavagem, os marcadores são removidos por corte de ligações das suas respectivas sondas, e então medidos.

Outro método, que resulta numa especificidade mais elevada para a análise de uma população no que respeita à presença de um determinado gene ou sequência de DNA, utiliza a técnica de Southern Blot. O DNA preparado é digerido por uma enzima de restrição (RE), resultando a produção de um grande número de fragmentos de diferentes extensões, determinados pela presença no genoma de um local específico de reconhecimento da enzima de restrição. Os alelos de um dado gene que possuam mutações no interior deste local de restrição serão cortados em fragmentos que são diferentes no que respeita ao seu número e extensão. O resultante polimorfismo da extensão dos fragmentos de restrição (RFLP) pode constituir um importante elemento de diagnóstico de um microorganismo se o fragmento puder ser identificado de forma específica.

O fragmento a ser analisado deve ser separado do "pool" de fragmentos de DNA e distinguido de outras espécies de DNA utilizando sondas específicas. Assim, o DNA é

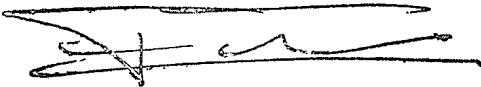


submetido a fraccionamento electroforético utilizando algum tipo de gel ou de cromatografia, seguindo-se a transferência e fixação a uma membrana de nylon ou nitrocelulose. O DNA fixado de cadeia simples é hibridizado com um oligonucleótido marcado, complementar ao DNA a ser detectado. Após a remoção das hibridizações não específicas, o fragmento de DNA com interesse é identificado por corte do(s) marcador(es) a partir da sonda hibridizada. Com a tecnologia aqui descrita é possível produzir em simultâneo um número de sondas superior a 100.

A presença e quantificação de um produto específico de transcrição genética podem ser verificadas através de uma análise por Northern Blot e de um ensaio de protecção de RNase. O princípio em que se fundamentam estes métodos é o da hibridização do "pool" total de RNA celular com uma sonda específica marcada ou com um conjunto de sondas específicas marcadas. Na técnica de Northern Blot, o RNA total de um tecido é fraccionado electroforeticamente utilizando um gel de agarose e é então transferido e immobilizado sobre um suporte sólido (nylon, nitrocelulose, etc.). O RNA é hibridizado com um oligonucleótido marcado, complementar ao RNA a ser detectado. Após a remoção das hibridizações não específicas, o fragmento de DNA com interesse é identificado por corte do(s) marcador(es) a partir da sonda hibridizada. Por aplicação de condições rigorosas de lavagem, as moléculas ligadas de forma não específica são eliminadas devido a o seu grau de hibridização ser mais fraco em comparação com o das moléculas hibridizadas de modo específico. Mais rápido, mas menos específico, o método de Dot-Blot é realizado tal como o método de Northern Blot, diferindo no facto de o RNA ser directamente transferido para a membrana, sem fraccionamento prévio.

O ensaio de protecção de RNase constitui um método específico de detecção de um mRNA. O RNA total proveniente de um tecido ou de uma cultura celular é hibridizado com uma sonda marcada de ribonucleótido ou desoxirribonucleótido. A especificidade é conseguida através de uma digestão subsequente com RNase. Os fragmentos não hibridizados da cadeia simples de RNA e os fragmentos inespecificamente hibridizados, mesmo com pequenas não-correspondências, serão reconhecidos e cortados, enquanto que o RNA de cadeia dupla ou os duplexes DNA/RNA com homologia completa não estão acessíveis à enzima e estarão protegidos. O fragmento protegido específico pode ser separado dos produtos de degradação e o(s) marcador(es) cortados a partir da respectiva sonda e subsequentemente medidos.

A localização precisa de um dado mRNA (ou de qualquer sequência de ácidos nucleicos) numa população específica de células pertencentes a um dado tecido pode ser determinada por hibridização *in situ*. A hibridização *in situ* poderá revelar-se um método



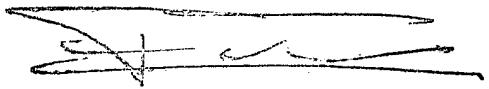
ainda mais sensível do que a análise de uma preparação de RNA de tecido total utilizando as técnicas acima descritas. Isto acontece quando o mRNA é expresso em elevadas concentrações por uma região ou tipo celular muito restritos no interior do tecido, que seriam diluídos por homogeneização do tecido total. Para a hibridização in-situ, os tecidos têm de ser congelados ou fixados por perfusão e seccionados de acordo com o protocolo histoquímico. O protocolo de hibridização para as secções de tecido e as sondas marcadas usadas são semelhantes aos de outros métodos de hibridização acima descritos. É possível realizar uma análise quantitativa.

5. Técnicas de Detecção de Mutações

A detecção de doenças tem um papel cada vez mais importante na sua prevenção e tratamento. Se por um lado é difícil conceber testes genéticos para a detecção de doenças multifactoriais, por outro verifica-se que mais do que 200 doenças humanas conhecidas são causadas por um defeito num único gene, frequentemente uma alteração num só resíduo de aminoácido (Olsen, *Biotechnology: An Industry Comes of Age*, National Academic Press, 1986). Muitas destas mutações resultam num aminoácido alterado que causa um estado de doença.

As técnicas sensíveis de detecção de mutações oferecem extraordinárias possibilidades de pesquisa de mutações. Por exemplo, estas análises poderão até ser realizadas antes de fazer a implantação de um ovo fertilizado (Holding e Monk, *Lancet* 3:532, 1989). Os testes genéticos cada vez mais eficientes poderão também permitir o despiste de mutações oncogénicas em células exfoliadas a partir do tracto respiratório ou da bexiga, testes esses que poderão vir a realizar-se em associação a um checkup médico (Sidransky et al., *Science* 252:706, 1991). Adicionalmente, quando um gene desconhecido causa uma doença genética, os métodos de monitorização de variantes das sequências de DNA são úteis para o estudo da hereditariedade da doença através de análise de associação genética. No entanto, a detecção e o diagnóstico de mutações em genes individuais coloca alguns problemas tecnológicos e económicos. Têm sido desenvolvidas diversas abordagens, mas nenhuma delas é suficientemente económica e eficiente para permitir uma aplicação em larga escala.

As mutações que envolvem um só nucleótido podem ser identificadas numa amostra por meios físicos, químicos ou enzimáticos. Geralmente, os métodos para a detecção de mutações podem dividir-se em técnicas de varrimento, que são adequadas para a identificação de mutações anteriormente desconhecidas, e técnicas concebidas para detectar, distinguir ou quantificar variantes conhecidas de sequências.



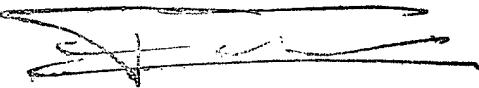
Foram desenvolvidas diversas técnicas de varrimento para a detecção de mutações em heteroduplexes de cadeias de DNA de complementaridade imperfeita, derivadas de sequências de tipo selvagem e mutantes, que exibem um comportamento anormal especialmente quando desnaturadas. Este fenómeno é explorado pelos métodos de electroforese em gel de desnaturação e de electroforese em gel com gradiente de temperatura (DGGE e TGGE, respectivamente). Mesmo os duplexes com um grau de emparelhamento imperfeito mínimo, de até uma posição de nucleótido, se podem desnaturar parcialmente, resultando numa migração retardada, quando submetidos a electroforese em gel com um gradiente de desnaturação crescente (Myers et al., *Nature* 313:495, 1985; Abrams et al., *Genomics* 7:463, 1990; Henco et al., *Nucl. Acids Res.* 18:6733, 1990). Se bem que as mutações possam ser detectadas, não é obtida qualquer informação relativa à localização precisa de uma mutação. As formas mutantes terão de ser isoladas mais completamente e submetidas a análise por sequenciação de DNA.

Alternativamente, um heteroduplex de uma sonda de RNA e de uma cadeia-alvo poderá ser cortado por uma RNase A numa posição em que as cadeias não se encontrem bem emparelhadas. O local de corte poderá então ser determinado por electroforese da sonda desnaturada. No entanto, algumas mutações poderão escapar à detecção, já que nem todos os emparelhamentos imperfeitos são cortados de forma eficiente pela RNase A.

As bases mal emparelhadas num duplex são também susceptíveis à modificação química. Uma tal modificação poderá tornar as cadeias susceptíveis ao corte no local do emparelhamento imperfeito, ou poderá causar a paragem de uma polimerase numa reacção de extensão subsequente. A técnica de corte químico das ligações permite a identificação de uma mutação em sequências-alvo com até 2 kb e fornece informações sobre a localização aproximada do(s) nucleótido(s) mal emparelhado(s) (Cotton et al., *PNAS USA* 85:4397, 1988; Ganguly et al., *Nucl. Acids Res.* 18:3933, 1991). No entanto, esta técnica é trabalhosa e poderá não identificar a localização precisa da mutação.

Uma estratégia alternativa para a detecção de uma mutação numa cadeia de DNA é a de substituir (durante a síntese) um dos nucleótidos normais por um nucleótido modificado, alterando o peso molecular ou outro parâmetro físico do produto. Uma cadeia que possui um número aumentado ou diminuído deste nucleótido modificado relativamente à sequência de tipo selvagem exibe uma mobilidade electroforética alterada (Naylor et al., *Lancet* 337:635, 1991). Esta técnica detecta a presença de uma mutação, mas não fornece a sua localização.

Duas outras estratégias permitem visualizar mutações num segmento de DNA através da alteração da migração em gel. Na técnica de polimorfismo em conformação de



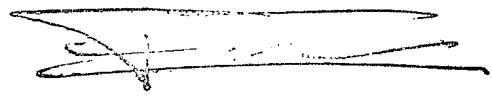
cadeia simples (SSCP), as mutações fazem com que as cadeias desnaturadas adoptem diferentes estruturas secundárias, influenciando desta forma a mobilidade durante a electroforese em gel nativo. As moléculas de DNA num heteroduplex, contendo locais internos de desemparelhamento, podem também ser separadas das moléculas correctamente emparelhadas por meio de electroforese (Orita, *Genomics* 5:874, 1989; Keen, *Trends Genet.* 7:5, 1991). Tal como as técnicas acima expostas, poderá detectar-se a presença de uma mutação mas não a sua localização. Do mesmo modo, muitas destas técnicas não distinguem entre mutações simples ou duplas.

Todas as técnicas acima mencionadas indicam a presença de uma mutação num segmento limitado de DNA, e algumas delas permitem obter uma indicação aproximada da sua localização no interior do segmento. No entanto, a análise da sequência continua a ser necessária para revelar o efeito da mutação sobre o potencial codificante do segmento. A análise de sequência constitui uma ferramenta muito poderosa, permitindo, por exemplo, a pesquisa da mesma mutação em outros indivíduos de uma família afectada, a monitorização da progressão do estado de doença, no caso de uma doença maligna, ou a detecção de células malignas residuais na medula óssea antes de um transplante autólogo. Apesar destas vantagens, os elevados custos envolvidos não tornam provável que o procedimento venha a ser adoptado como método de diagnóstico de rotina.

Foi desenvolvido um grande número de outras técnicas para a análise de variantes de sequências conhecidas. A automatização e a economia constituem considerações muito importantes neste tipo de análises, já que estas poderão vir a ser aplicadas a pesquisas em indivíduos e na população em geral. Nenhuma das técnicas abaixo descritas combina os critérios de economia e automatização com o grau de especificidade requerida.

As mutações poderão ser identificadas segundo os efeitos desestabilizantes que exercem sobre a hibridização de sondas oligonucleotídicas curtas com uma sequência-alvo (ver Wetmur, *Crit. Rev. Biochem. Mol. Biol.*, 26:227, 1991). Esta técnica, a hibridização de oligonucleótidos específica de alelo, envolve geralmente a amplificação de sequências-alvo e a hibridização subsequente com sondas oligonucleotídicas curtas. Neste produto amplificado podem, então, pesquisar-se numerosas variantes de sequência possíveis através da determinação do padrão de hibridização do produto com um conjunto de sondas oligonucleotídicas imobilizadas.

No entanto, esta e todas as outras possibilidades, de entre uma variedade de estratégias que podem usar-se na identificação das diferenças dentro das sequências de nucleótidos, dependem das enzimas que são usadas nessa identificação (Saiki, *PNAS USA* 86:6230, 1989; Zhang, *Nucl. Acids Res.* 19:3929, 1991).



Por exemplo, as enzimas de restrição reconhecem sequências com cerca de 4-8 nucleótidos. Com base no teor médio de G+C, aproximadamente metade das posições nucleotídicas num segmento de DNA podem ser monitorizadas utilizando um painel de 100 enzimas de restrição. Como alternativa, poderão criar-se à volta de uma posição variável sequências artificiais para reconhecimento por enzima de restrição, utilizando primers de PCR com uma incorrecção parcial do emparelhamento. Através desta técnica, tanto a sequência mutante como a sequência de tipo selvagem isolada poderão ser reconhecidas e cortadas por uma enzima de restrição após a amplificação (Chen et al., *Anal. Biochem.* 195:51, 1991; Levi et al., *Cancer Res.* 51:3497, 1991).

Outro método explora o facto de que um primer oligonucleotídico que se encontra incorrectamente emparelhado com uma sequência alvo a nível da penúltima posição da extremidade 3' exibe uma capacidade reduzida para funcionar como primer de PCR. No entanto, alguns emparelhamentos imperfeitos 3', notavelmente o G-T, são menos inibitórios que outros, o que limita a utilidade deste método. Na tentativa de melhorar esta técnica, são incorporados no primer outros locais de emparelhamento imperfeito na terceira posição a contar da extremidade 3'. Isto resulta em duas posições mal emparelhadas nos três nucleótidos 3' do primer que se hibridiza com um dos variantes alélicas, e num emparelhamento imperfeito da terceira posição a contar da extremidade 3' quando o primer se hibridiza com o outro variante alélico (Newton et al., *Nucl. Acids Res.* 17:2503, 1989). Torna-se necessário definir condições de amplificação que favoreçam significativamente a amplificação de um produto de emparelhamento imperfeito com 1 bp.

Têm igualmente sido usadas DNA polimerases para distinguir variantes de sequências alélicas, determinando-se qual o nucleótido que é adicionado a um primer oligonucleotídico imediatamente a montante de uma posição variável da cadeia alvo.

Foi desenvolvido um ensaio de ligação em que duas sondas oligonucleotídicas que se hibridizam com uma cadeia-alvo em justaposição imediata são ligadas por acção de uma DNA ligase. A ligação é inibida se existir um emparelhamento imperfeito na zona em que as duas sondas contactam entre si.

a. Ensaios para Detecção de Mutações

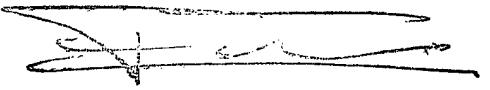
As mutações constituem uma alteração de um só par de bases no DNA genómico. No contexto deste invento, a maioria destas alterações pode ser facilmente detectada por hibridização com oligonucleótidos que são complementares à sequência em questão. No sistema aqui descrito, são utilizados dois oligonucleótidos para detectar uma mutação. Um dos oligonucleótidos possui a sequência de tipo selvagem e o outro oligonucleótido

apresenta a sequência mutante. Quando os dois oligonucleótidos são usados como sondas numa sequência genómica alvo de tipo selvagem, o oligonucleótilde de tipo selvagem formará uma estrutura de bases perfeitamente emparelhadas e a sequência oligonucleotídica mutante formará um duplex com um emparelhamento imperfeito a nível de um só par de bases.

Tal como exposto acima, uma diferença de 6 a 7°C entre a T_m de um duplex de tipo selvagem e o duplex de emparelhamento imperfeito permite a pronta identificação ou discriminação dos dois tipos de duplexes. Para efectuar esta discriminação, é realizada uma hibridização a nível da T_m do duplex de emparelhamento imperfeito no seio da respectiva solução hibotrópica. A extensão da hibridização é então medida para o conjunto de sondas oligonucleotídicas. Quando é medida a razão entre a extensão de hibridização da sonda de tipo selvagem e a da sonda de emparelhamento imperfeito, são obtidos valores de desde 10/1 até valores superiores a 20/1. Estes tipos de resultados permitem o desenvolvimento de ensaios robustos para a detecção de mutações.

A título de exemplo, um formato de ensaio para a detecção de mutações utiliza uma molécula-alvo de ácido nucleico (p.ex., DNA genómico) e sondas oligonucleotídicas que abrangem a área de interesse. As sondas oligonucleotídicas possuem extensões maiores ou iguais a 24 nt (com um máximo de cerca de 36 nt) e estão marcadas com um fluorocromo a nível da extremidade 3' ou 5'. A molécula-alvo de ácido nucleico é obtida através da lise de culturas celulares, de tecidos, de organismos, etc., no seio da respectiva solução de hibridização. A solução lisada é então aquecida a uma temperatura que desnatura a molécula-alvo de ácido nucleico (15-25°C acima da T_m do duplex de ácido nucleico alvo). As sondas de oligonucleótidos são adicionadas à temperatura de desnaturação, e a hibridização é conduzida mantendo a T_m do duplex de emparelhamento imperfeito durante 0,5 a 24 horas. O DNA genómico é então recolhido, passando através de um filtro de fibra de vidro GF/C (GF/B, e semelhantes). O filtro é então lavado com a solução de hibridização respectiva para remover quaisquer sondas oligonucleotídicas não hibridizadas (RNA, oligos curtos e ácido nucleico não se ligam à fibra de vidro sob estas condições). A sonda oligo de hibridização pode então ser termicamente eluída, separando-se do DNA alvo, e medida (por fluorescência, por exemplo). Para os ensaios que requerem níveis de sensibilidade muito elevados, as sondas são concentradas e medidas.

Poderão usar-se outros protocolos de hibridização de elevada sensibilidade. Os métodos do presente invento permitem a análise simples de um ácido nucleico que contém uma mutação suspeita de se encontrar presente em células, amostras, etc., i.e., uma molécula-alvo de ácido nucleico. A "molécula-alvo de ácido nucleico" contém a sequência

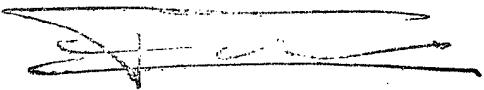


nucleotídica do ácido desoxirribonucleico (DNA) ou ribonucleico (RNA) cuja presença se reveste de interesse, e cuja presença ou ausência se destina a ser detectada por meio do ensaio de hibridização. Os métodos de hibridização do presente invento podem igualmente ser aplicados a uma mistura biológica complexa de ácido nucleico (RNA e/ou DNA). Uma tal mistura biológica complexa inclui uma vasta gama de células eucarióticas e procarióticas, incluindo protoplastos; e/ou outros materiais biológicos que contêm um ácido nucleico polinucleotídico. Deste modo, o método é aplicável a culturas celulares, células animais, tecido animal, células sanguíneas (p.ex., reticulócitos, linfócitos), células vegetais, bactérias, leveduras, vírus, micoplasmas, protozoários, fungos e semelhantes. Através da detecção de uma hibridização específica entre sondas de ácidos nucleicos de origem conhecida, pode ser estabelecida a presença específica de uma molécula-alvo de ácido nucleico.

Um protocolo típico de ensaio de hibridização para a detecção de uma molécula-alvo de ácido nucleico no seio de uma população complexa de ácidos nucleicos é descrito como se segue: As moléculas-alvo de ácido nucleico são separadas de acordo com o seu tamanho numa matriz de gel (electroforese), clonadas e isoladas, subdivididas em pools, ou reunidas na categoria de população complexa. As moléculas-alvo de ácidos nucleicos são transferidas, impressas ou imobilizadas num suporte sólido, tal como uma membrana de nylon ou uma membrana de nitrocelulose. (Esta "imobilização" é também referida como "arraying"). Os ácidos nucleicos imobilizados são então submetidos a um passo de aquecimento ou a radiação UV, condições que imobilizam irreversivelmente o ácido nucleico. As membranas são então imersas em "agentes bloqueantes", que incluem o reagente de Dendhart (Dendhart, *Biochem. Biophys. Res. Comm.* 23:641, 1966), a heparina (Singh e Jones, *Nucleic Acids Res.* 12:5627, 1984), e leite magro evaporado (Jones et al., *Gene Anal. Tech.* 1:3, 1984). Os agentes bloqueantes são geralmente incluídos tanto no passo de pré-hibridização como no passo de hibridização quando é usada a nitrocelulose. As moléculas-alvo de ácido nucleico são então sondadas com sondas oligonucleotídicas marcadas sob as condições acima descritas em soluções hibotrópicas. A enzima não marcada é então removida por lavagem e a membrana é imersa numa solução de substrato. O sinal é então detectado por MALDI-MS essencialmente como abaixo se descreve.

b. Sequenciação por hibridização

A análise da sequência do DNA é convencionalmente realizada através da hibridização de um primer ao DNA-alvo e da realização de extensões da cadeia utilizando uma polimerase. As paragens específicas são controladas pela inclusão de um didesoxinucleótido. A especificidade conferida pelo uso do primer neste tipo de análise pode



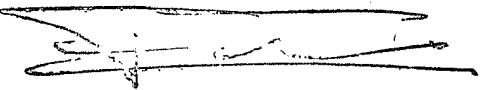
ser aumentada pela inclusão de um agente hibotrópico no tampão de annealing e/ou pela incorporação de um resíduo abásico no primer e a realização da reacção de annealing a uma temperatura discriminativa.

Outros métodos de análise de sequências envolvem a hibridização do alvo com um conjunto de oligonucleótidos curtos escolhidos ao acaso. A sequência é construída por análise da hibridização por sobreposição. Nesta técnica, é essencial obter uma hibridização precisa. A utilização de agentes hibotrópicos ou de resíduos abásicos e a realização da reacção de annealing a uma temperatura discriminativa são factores benéficos na redução ou eliminação do emparelhamento imperfeito na hibridização. O objectivo é o de desenvolver métodos de hibridização automatizada, de modo a sondar grandes conjuntos de sondas oligonucleotídicas ou grandes conjuntos de amostras de ácidos nucleicos. A aplicação de tais tecnologias inclui o mapeamento de genes, a caracterização de clones, a genética médica e a descoberta de genes, a análise de sequências de DNA por hibridização, e, finalmente, a verificação da sequenciação.

Para a automatização ou o controle da formação de multiplexes de sondas oligonucleotídicas é necessário o controlo de muitos parâmetros. A estabilidade das sondas respectivas deve ser semelhante; do mesmo modo, o grau de não correspondência com a molécula-alvo de ácido nucleico, a temperatura, a força iônica, o conteúdo em A+T da sonda (ou do alvo), assim como outros parâmetros quando a sonda é curta, devem ser semelhantes. Usualmente, as condições da experiência e a sequência da sonda são ajustadas até que seja termodinamicamente favorecida a formação da sonda com um emparelhamento perfeito de bases, em prejuízo da formação de qualquer duplex que contivesse um emparelhamento imperfeito. As aplicações em larga escala das sondas, tais como a sequenciação por hibridização (SBH) ou o teste de loci altamente polimórficos, como o locus da proteína transmembranar da fibrose quística, requer um controlo mais rigoroso dos multiplexes de sondas.

6. Conjuntos para análise

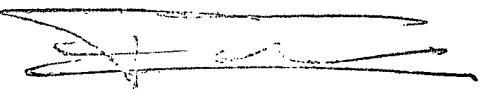
A hibridização de ácidos nucleicos com conjuntos de amostras de DNA tem vindo a ser usada há já algum tempo numa grande variedade de aplicações de pesquisa básica no campo da biologia, e começa agora a ser utilizada também nas áreas do diagnóstico médico, da medicina forense e da agricultura. Tal como abaixo se descreve em maior detalhe, as moléculas de ácido nucleico ou proteínas poderão encontrar-se ligadas a um suporte sólido de modo a formar um conjunto para análise e ser testadas por recurso a moléculas marcadas preparadas de acordo com o presente invento.



Por exemplo, segundo uma forma de realização deste invento, as amostras de DNA dispostas num conjunto para análise podem ser utilizadas na identificação de clones individuais. Abreviadamente, são marcadas moléculas conhecidas de DNA de modo a formar sondas marcadas, sendo estas sondas então testadas por hibridização com um conjunto de clones desconhecidos. Os clones que exibem uma hibridização específica com a sonda podem então ser isolados. Estes ensaios podem ser realizados utilizando conjuntos não ordenados de clones (Sambrook et.al., "Molecular Cloning: a Laboratory Manual", Cold Spring Harbor, NY, 1989). Alternativamente, poderão adquirir-se membranas que possuam conjuntos de clones com um espaçamento regular entre conjuntos, possuindo os clones uma identidade individual conhecida (embora tipicamente uma sequência desconhecida) (p.ex., conjuntos de clones BAC da Research Genetics, Huntsville, AL).

Dentro de outras formas de realização do invento, poderão utilizar-se conjuntos para análise com a finalidade de medir os níveis de transcrição de um grande número de genes em simultâneo (ver, genericamente, Gess et al., *Mammalian Genome* 3:609-619, 1992). Abreviadamente, são marcados pools de cDNA, os quais são utilizados como sondas sob a forma de grandes conjuntos de clones de cDNA para identificar os genes expressos em abundância em tecidos específicos. Poderão usar-se microconjuntos formados a partir de clones individuais de cDNA para medir quantitativamente a expressão relativa de cada gene no conjunto em duas amostras de RNA diferentes (Schena et al., *Science* 270:467-470, 1995). Mais especificamente, poderão utilizar-se robôs para produzir microconjuntos de produtos para PCR a partir de clones individuais: cada elemento do conjunto para análise corresponde a um só clone de cDNA. As sondas para os conjuntos são preparadas através da marcação da primeira cadeia de cDNA de cada tecido com um marcador. Para comparar a expressão de genes em duas amostras de tecido, o cDNA de cada uma é marcado com um marcador diferente. As duas amostras são reunidas num pool e são ambas hibridizadas com o conjunto para análise. Após a hibridização das sondas com o conjunto para análise, os marcadores poderão ser cortados e analisados, tal como descrito neste pedido de patente para cada marcador hibridizado com cada amostra no conjunto para análise. Para um dado gene, a taxa de hibridização com cada amostra de complexo marcado de cDNA constitui uma medida da expressão genética relativa nas duas amostras de tecido. O uso de controlos internos e de dois (e potencialmente até quatro) marcadores distintos é crucial para esta aplicação.

Muitas das aplicações a seguir descritas constituem variações desta experiência básica, usando-se diferentes fontes de DNA no conjunto para análise e diferentes fontes de DNA de sonda, mas cada uma das aplicações é limitada pelo uso de métodos

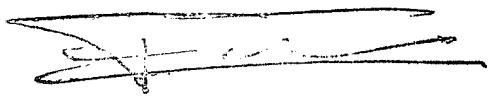


convencionais de detecção a menos do que 4-6 sondas distinguíveis entre a mistura de hibridização.

Outra das aplicações da hibridização com DNA em conjuntos para análise, que foi demonstrada em princípio e que possui potencial para uma aplicação bastante vasta, é a sequenciação por hibridização (SBH). O conceito de sequenciação por hibridização (DBH) recorre ao uso de um conjunto de todos os oligómeros de n-nucleótidos (n-meros) possíveis para identificar n-meros presentes numa sequência desconhecida de DNA. Podem então usar-se abordagens computadorizadas para montar a sequência completa (ver, genericamente, Drmanac et al., *Science* 260:1649-1652, 1993). As aplicações da SBH incluem o mapeamento físico (ordenamento) de clones de DNA com sequências de extremidades sobreponíveis, verificação de sequências, comparações detalhadas entre genes normais e genes causadores de doenças, e a identificação de fragmentos de DNA com motivos particulares de sequências presentes em colecções genómicas e em colecções de DNAs complementares.

Os conjuntos de DNA para análise possuem também uma vasta gama de aplicações na detecção de variações genéticas e de polimorfismos. As alterações de um só par de bases, as deleções e inserções, mutações e polimorfismos podem detectar-se imobilizando variantes de sequências conhecidas e sondando então estas variantes com produtos de PCR marcados obtidos a partir de pacientes ou de patogénios (ver, p.ex., Guo et al., *Nucleic Acids Res.* 22: 5456-5465, 1994). Do mesmo modo, podem usar-se conjuntos de oligonucleótidos para medir a variação genética, incluindo a detecção de variantes resistentes aos fármacos e variantes sensíveis aos fármacos do vírus HIV (ver, p.ex., Lipshutz et al., *Biotechniques* 19:442-447, 1995).

Os conjuntos de DNA para análise podem ser produzidos usando pelo menos duas técnicas diferentes: síntese *in situ* e deposição de amostras produzidas separadamente ("spotting"). Uma das técnicas mais proeminentes para a produção das amostras de DNA *in situ* é a síntese de oligonucleótidos dirigida pela luz, descrita em Pease et al., *PNAS USA* 91: 5022-6, 1994. Abreviadamente, são produzidos conjuntos de sequências de DNA definidas recorrendo a grupos bloqueadores fotolábeis para dirigir a síntese de oligonucleótidos num conjunto, utilizando métodos fotolitográficos modernos. São preparadas máscaras para aplicar a cada elemento do conjunto que necessite de uma base particular no passo de síntese seguinte, expondo-se então o conjunto à luz. É adicionado um só resíduo nucleotídico a cada cadeia que foi exposta pela máscara, termina o ciclo de síntese, o ciclo seguinte é iniciado pela aplicação de outra máscara e é adicionado outro resíduo de oligonucleótido. A aplicação sequencial deste protocolo pode ser usada para



construir rapidamente conjuntos muito vastos de oligonucleótidos. É descrita uma versão de deposição por robô em Schena et al. (1995) ligada ao uso de conjuntos para análise por transcrição.

Numa forma de realização do invento, os segundos membros são reunidos num conjunto sobre um suporte sólido como a sílica, quartzo ou vidro. O conjunto poderá então ser tratado para bloquear a hibridização não específica, seguindo-se a incubação das sondas de primeiro membro marcadas sobre o suporte sólido. Em certas formas de realização preferidas, o conjunto é então lavado com uma solução (a um grau definido de intensidade de lavagem) de forma a remover os ácidos nucleicos hibridizados inespecificamente, coberto por uma solução que inclui um material de matriz apropriado para espectrometria ou potenciometria (p.ex., para espectrometria de massa assistida por matriz com desadsorção e ionização por laser), secado de modo a formar uma matriz apropriada e exposto à luz para promover o corte dos marcadores, separando-os das sondas de ácidos nucleicos. Os marcadores cortados podem então ser analisados por meio de técnicas espectrométricas ou potenciométricas (p.ex., MALDI-MS).

Em certas formas de realização, o corte dos marcadores e a desadsorção por laser ocorrem num só passo. Em outras variações, a desadsorção por laser e a ionização são efectuados sem uma matriz. Em algumas experiências, os oligonucleótidos ligados a marcadores de referência, ou outros compostos marcados, são adicionados à solução de matriz para o controlo de variações da eficiência do corte fotolítico, da desadsorção por laser e da eficiência da detecção por MS. Por medição da razão de abundâncias entre um marcador de teste e uma série de marcadores de referência, são extraídas informações quantitativas a partir dos dados de MALDI-MS.

Em outras formas de realização, o conjunto é composto por oligonucleótidos de extensão inferior a 50 bp. Este conjunto pode ser usado na detecção de polimorfismos (p.ex., alterações em um só par de bases) para mapeamento genético, ou na detecção da presença ou ausência de um DNA particular numa amostra, para análise ou separação de clones, em testes de paternidade, na área forense e no mapeamento genético. Os conjuntos para análise poderão também ser compostos por proteínas.

C. SEPARAÇÃO DOS FRAGMENTOS DE ÁCIDOS NUCLEICOS

Uma amostra para análise é frequentemente constituída por uma mistura de muitos componentes numa matriz complexa. Para amostras que contêm compostos desconhecidos, os componentes terão de ser separados entre si para que cada componente individual possa ser identificado por outros métodos analíticos. As propriedades de



separação dos componentes numa mistura mantêm-se constantes sob condições constantes, e assim, uma vez determinadas, poderão ser usadas para identificar e quantificar cada um dos componentes. Tais procedimentos são típicos nas separações analíticas cromatográficas e electroforéticas.

1. Cromatografia Líquida de Alto Rendimento (HPLC)

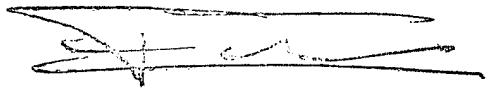
A cromatografia líquida de alto rendimento (HPLC) constitui uma técnica de separação cromatográfica de compostos que se encontram dissolvidos em solução. Os instrumentos de HPLC são constituídos por um reservatório da fase móvel, uma bomba, um injector, uma coluna de separação e um detector. Os compostos são separados através da injecção de uma alíquota da mistura de amostra na coluna. Os diferentes componentes da mistura passam através da coluna a taxas de fluxo diferentes, devido às diferenças no seu comportamento de partição entre a fase móvel líquida e a fase estacionária.

Recentemente, as IP-RP-HPLC em partículas de PS/DVB não porosas ligadas por via química a cadeias alquilo têm-se evidenciado como alternativas rápidas à electroforese capilar na análise de ácidos nucleicos de cadeias simples e duplas, proporcionando níveis semelhantes de resolução (Huber et al., Anal. Biochem., 212, p351; Huber et al., 1993, Nucl. Acids Res., 21, p1061; Huber et al., 1993, Biotechniques, 16, p898). Em contraste com a cromatografia de troca iônica, que nem sempre retém o DNA de cadeia dupla em função da extensão das cadeias (uma vez que os pares de bases AT interagem com a fase estacionária carregada positivamente, mais fortemente do que os pares de bases G-C), a IP-RP-HPLC proporciona uma separação estritamente dependente da extensão das cadeias.

Foi desenvolvido um método que utiliza acetato de trietilamónio a 100 mM como reagente de emparelhamento de iões, em que os fosfodiéster-oligonucleótidos são separados de modo eficaz por utilização de partículas alquiladas não porosas de poli(estireno-divinilbenzeno) a 2,3 μ M numa cromatografia líquida de alto rendimento (Oefner et al., Biochem., 223, p39). A técnica descrita permitiu a separação de produtos de PCR com uma extensão diferindo em apenas 4 a 8 pares de bases dentro de um intervalo de dimensões de 50 a 200 nucleótidos.

2. Electroforese

A electroforese constitui uma técnica separativa que se baseia na mobilidade de iões (ou de DNA, como no caso aqui descrito) num campo eléctrico. As moléculas de DNA carregadas negativamente migram em direcção a um eléctrodo positivo e os iões



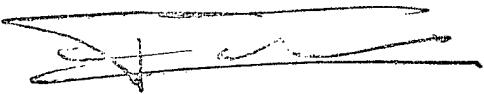
carregados positivamente migram em direcção a um eléctrodo negativo. Por razões de segurança, um dos eléctrodos está geralmente em terra e o outro está carregado positiva ou negativamente. As espécies carregadas possuem velocidades diferentes de migração, dependendo da sua carga total, dimensão e forma, e podem deste modo ser separadas. Um aparelho de electroforese é constituído por uma fonte de alimentação de alta voltagem, eléctrodos, tampão e um suporte para o tampão, tal como o gel de poliacrilamida ou um tubo capilar. Frequentemente, os tubos capilares são usados para muitos tipos de amostras e os outros suportes de gel são utilizados para amostras biológicas, tais como misturas de proteínas ou fragmentos de DNA.

3. Electroforese Capilar (CE)

A electroforese capilar (CE), nas suas variadas manifestações (solução livre, isotacoforese, focalização isoeléctrica, gel de poliacrilamida, "cromatografia" electrocinética micelar) encontra-se em desenvolvimento como método para separações rápidas de alta resolução para amostras de misturas complexas com volumes muito pequenos. Em combinação com a sensibilidade e selectividade inerentes à MS, a CE-MS constitui uma técnica potencialmente poderosa no campo da bioanálise. Na nova aplicação aqui apresentada, a ligação entre estes dois métodos conduzirá à obtenção de métodos superiores de sequenciação de DNA que ultrapassarão em diversas ordens de magnitude os actuais métodos de sequenciação.

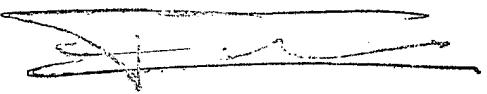
A correspondência entre as taxas de fluxo obtidas em CE e em ionização por electrospray (ESI), e o facto de que ambas as técnicas são facilitadas pela presença de (e primariamente usadas para) espécies iónicas em solução, fornece a base para uma combinação extremamente apelativa. Foi já descrita a utilização de uma combinação de electroforese de zona capilar (CZE) e de isotacoforese capilar em associação ao uso de espectrómetros de massa quadrupolares baseados em ESI (Olivares et al., Anal. Chem. 59:1230, 1987; Smith et al., Anal. Chem. 60:436, 1988; Loo et al., Anal. Chem. 179:404, 1989; Edmonds et al., J. Chromatog. 474:21, 1989; Loo et al., J. Microcolumn Sep. 1:223, 1989; Lee et al., J. Chromatog. 458:313, 1988; Smith et al., J. Chromatog. 480:211, 1989; Grese et al., J. Am. Chem. Soc. 111:2835, 1989). Os péptidos pequenos são facilmente adaptáveis à análise por CZE, obtendo-se uma boa sensibilidade (da ordem das femtomoles).

O método mais poderoso para a separação de fragmentos de DNA é a electroforese em gel de poliacrilamida (PAGE), geralmente num formato de placa de gel. No entanto, a principal limitação da tecnologia actual consiste no relativamente longo período de tempo que é necessário dispenser na electroforese em gel de fragmentos de DNA produzidos nas



reacções de sequenciação. Pode conseguir-se um aumento da velocidade (numa ordem de grandeza de 10 vezes) por meio do uso da electroforese capilar, que utiliza geles ultrafinos. Em solução livre, e numa primeira aproximação, todos os DNA migram com a mesma mobilidade, já que a adição de uma base resulta na compensação da massa com a carga. Nos geles de poliacrilamida, os fragmentos de DNA migram em função do tamanho, e esta abordagem foi agora aplicada à CE. Foi agora atingido um número notável de placas por metro através do uso de poliacrilamida com ligações cruzadas (10^7 placas por metro, Cohen et al., Proc. Natl. Acad. Sci. USA 85:9660, 1988). Tais colunas de CE, tal como descrito, podem ser utilizadas para a sequenciação de DNA. O método de CE é, em princípio, 25 vezes mais rápida do que a electroforese em placa de gel realizada num sequenciador convencional. Por exemplo, é possível ler cerca de 300 bases por hora. A velocidade de separação na electroforese em placa de gel é limitada pela intensidade do campo eléctrico que pode ser aplicado ao gel sem que ocorra uma produção excessiva de calor. Assim, a maior velocidade da CE é conseguida através do uso de forças de campo superiores (300 V/cm na CE versus 10 V/cm na electroforese em placa de gel). O formato capilar reduz a amperagem, e, desse modo, a potência e a geração de calor resultante.

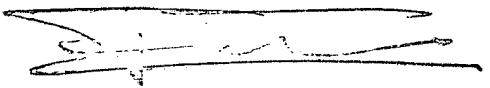
Smith e outros (Smith et al., *Nuc. Acids Res.* 18:4417, 1990) sugeriram o uso de múltiplos capilares em paralelo para obter um aumento do rendimento. Do mesmo modo, Mathies e Huang (Mathies e Huang, *Nature* 359:167, 1992) introduziram a electroforese capilar em que as separações são realizadas em conjuntos de capilares em paralelo e demonstraram a ocorrência de uma sequenciação de alto rendimento (Huang et al., *Anal. Chem.* 64:967, 1992, Huang et al., *Anal. Chem.* 64:2149, 1992). A principal desvantagem da electroforese capilar consiste na limitada quantidade de amostra com que é possível carregar o capilar. Ao concentrar uma grande quantidade de amostra no início do capilar, antes da separação, a capacidade de carga é aumentada e os níveis de detecção podem baixar em diversas ordens de grandeza. O método mais popular de pré-concentração na CE consiste no "empilhamento" de amostras. O empilhamento de amostras foi recentemente analisado numa publicação (Chien e Burgi, *Anal. Chem.* 64:489A, 1992). O empilhamento de amostras depende da diferença de matrizes (pH, força iônica) entre o tampão da amostra e o tampão do capilar, sendo o campo eléctrico ao longo da zona da amostra superior ao verificado na zona capilar. No empilhamento de amostras, um grande volume de amostra presente num tampão de baixa concentração é introduzido para pré-concentração na cabeça da coluna capilar. O capilar está cheio de um tampão com a mesma composição mas de concentração superior. Quando os íões da amostra atingem o tampão do capilar e o



campo eléctrico mais baixo, eles empilham-se numa zona concentrada. O empilhamento de amostras possibilita um aumento de detectabilidade de 1-3 ordens de grandeza.

Um outro método de pré-concentração consiste na aplicação de isotacoforese (ITP) antes da separação dos analitos em CE de zona livre. A ITP é uma técnica electroforética que permite o carregamento do capilar com volumes de amostra da ordem do microlitro, em contraste com os baixos volumes de injecção (nL) tipicamente associados à CE. A técnica baseia-se na inserção da amostra entre dois tampões (electrólitos de vanguarda e de rectaguarda), de mobilidades respectivamente mais alta e mais baixa do que a do analito. A técnica constitui inherentemente uma técnica de concentração, em que os analitos se concentram em zonas puras que migram à mesma velocidade. A técnica é actualmente menos popular do que os métodos de empilhamento acima descritos devido à necessidade de fazer diversas escolhas de electrólitos de vanguarda e rectaguarda, e ao facto de ser apenas possível separar espécies catiónicas ou aniónicas durante um processo de separação.

A característica central do processo de sequenciação de DNA reside na separação notavelmente selectiva de DNA ou fragmentos oligonucleotídicos. Esta separação é notável porque cada um dos fragmentos é resolvido e pode diferir dos outros em apenas um nucleótido. Têm já sido obtidas separações de até 1000 fragmentos (1000 bp). Uma vantagem adicional da sequenciação por meio de marcadores susceptíveis de corte é a seguinte: quando são usados marcadores susceptíveis de corte, não existe a necessidade de usar um formato de placa de gel na separação de fragmentos de DNA por gel de poliacrilamida. Uma vez que são combinadas numerosas amostras (4 a 2000), não há a necessidade de processar as amostras em paralelo, tal como acontece com os métodos actuais de corante-primer ou corante-terminador (i.e., sequenciador ABI373). Uma vez que não é necessário processar carreiras em paralelo no gel, não existe nenhuma razão para usar uma placa de gel. Deste modo, é possível utilizar um formato de gel em tubo no método de separação electroforética. Grossman e outros (Grossman et al., *Genet. Anal. Tech. Appl.* 9:9, 1992) demonstraram que se obtém vantagens significativas por utilização de um formato de gel em tubo no lugar de um formato de gel em placa. Tal deve-se à maior facilidade em dissipar a energia calorífica de Joule quando se utiliza um formato de tubo, o que resulta na redução dos tempos de corrida (em 50%) e numa resolução muito mais elevada dos fragmentos de DNA de elevado peso molecular (superior a 100 nt). As leituras extensas são críticas para a sequenciação genómica. Assim, o uso de marcadores susceptíveis de corte na sequenciação possui a vantagem adicional de facultar ao utilizador



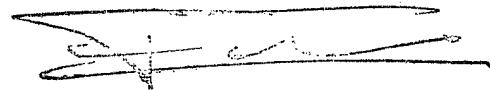
a utilização do método de separação de DNA que apresenta maior eficiência e sensibilidade e que possui igualmente a mais alta resolução.

4. Dispositivos Microfabricados

A electroforese capilar (CE) constitui um poderoso método para sequenciação, análise forense, análise de produtos de PCR e separação dos fragmentos de restrição por dimensões. A CE é bastante mais rápida do que a PAGE tradicional em placa, uma vez que em geles capilares se pode aplicar um potencial de campo bastante superior. No entanto, a CE apresenta a desvantagem de só permitir o processamento de uma amostra por gel. O método combina a maior rapidez dos tempos de separação da CE com a capacidade de analisar amostras múltiplas em paralelo. O conceito subjacente ao uso de dispositivos microfabricados é o de que a densidade da informação obtida na electroforese é aumentada ao miniaturizar a dimensão das carreiras do gel para cerca de 100 micrómetros. A indústria electrónica usa rotineiramente a microfabricação para produzir circuitos com estruturas de tamanho inferior a um micron. A densidade actual dos conjuntos de capilares é limitada pelo diâmetro exterior do tubo capilar. A microfabricação de canais possibilita a obtenção de uma maior densidade para os conjuntos de capilares. A microfabricação permite igualmente a execução de montagens que não é possível realizar utilizando fibra de vidro, e liga os canais directamente a outros dispositivos num chip. Poucos dispositivos foram já produzidos em microchips para aplicação às tecnologias de separação. Foram fabricados um cromatógrafo de gás (Terry et al., *IEEE Trans. Electron Device*, ED-26:1880, 1979) e um cromatógrafo de líquido (Manz et al., *Sens. Actuators B1:249*, 1990) em chips de silicone, mas estes dispositivos não têm conhecido uma utilização muito alargada. Diversos grupos têm referido a separação de corantes fluorescentes e de aminoácidos em dispositivos microfabricados (Manz et al., *J. Chromatography* 593:253, 1992, Effenhauser et al., *Anal. Chem.* 65:2637, 1993). Recentemente, Woolley e Mathies (Woolley e Mathies, *Proc. Natl. Acad. Sci.* 91:11348, 1994) demonstraram que a fotolitografia e a gravação química por acção de ácido podem ser usados para produzir grandes números de canais de separação em substratos de vidro. Os canais são preenchidos com matrizes de separação de hidroxietilcelulose (HEC). Foi demonstrado que é possível separar fragmentos de restrição de DNA em períodos de tempo tão breves como dois minutos.

D. CORTE DE MARCADORES

Tal como se descreveu acima, diferentes concepções do linker conferirão susceptibilidades ao corte ("labilidades") sob diferentes condições físicas ou químicas. Os



exemplos de condições que servem para cortar diferentes tipos de linker incluem as condições de ácido, base, oxidação, redução, fluoretos, troca de tióis, fotólise e acção enzimática.

Os exemplos de linkers susceptíveis ao corte que satisfazem os critérios gerais para linkers acima expostos são bem conhecidos dos técnicos e incluem os compostos listados no catálogo de Pierce (Rockford, JL). Os exemplos incluem:

- glicobis(succinimidilsuccinato) de etileno (EGS), um reagente de "cross-linking" (que estabelece ligações cruzadas) reactivo a aminas e que sofre quebra de ligações por acção da hidroxilamina (1 M a 37°C durante 3-6 horas);
- tartarato de disuccinimidilo (DST) e sulfo-DST, que são reagentes de cross-linking reactivos a aminas, susceptíveis ao corte por periodato de sódio a 0,015 M;
- bis[2-(succinimidiloxicarboniloxi)etil]sulfona (BSOCOES) e sulfo-BSOCOES, que são reagentes de cross-linking reactivos a aminas, susceptíveis ao corte por base (pH 11,6);
- 1,4-di-[3'-(2'-piridilditio(propionamido))butano (DPDPB), um piridilditiol, reagente de crosslinking susceptível ao corte por troca de tióis ou por redução;
- N-[4-(p-azidosalicilamido)-butil]-3'-(2'-piridilditio)propionamida (APDP), um piridilditiol, reagente de crosslinking susceptível ao corte por troca de tióis ou por redução;
- Bis-[beta-4-(azidosalicilamido)etil]disulfureto, um reagente de crosslinking fotoreactivo que é susceptível ao corte por troca de tióis ou por redução;
- N-succinimidil-(4-azidofenil)-1,3'ditiopropionato (SADP), um reagente de crosslinking fotoreactivo que é susceptível ao corte por troca de tióis ou por redução;
- sulfosuccinimidil-2-(7-azido-4-metilcumarina-3-acetamida)etil-1,3'-ditiopropionato (SAED), um reagente de crosslinking fotoreactivo que é susceptível ao corte por troca de tióis ou por redução;
- sulfosuccinimidil-2-(m-azido-o-nitrobenzamido)-etil-1,3'-ditiopropionato (SAND), um reagente de crosslinking fotoreactivo que é susceptível ao corte por troca de tióis ou por redução.

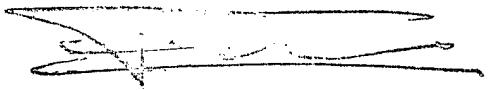
São em seguida fornecidos outros exemplos de linkers susceptíveis a corte e de condições que podem ser usadas para libertar os marcadores. Um grupo linker de sililo pode ser cortado por fluoretos ou por condições acídicas. O linker 2-nitrobenziloxi 3-, 4-, 5- ou 6-

substituído ou o 4-nitrobenziloxi 2-, 3-, 5- ou 6-substituído podem ser cortados por uma fonte de fotões (fotólise). O linker 2-alcoxifenoxy 3-, 4-, 5- ou 6-substituído ou o 4-alcoxifenoxy 2-, 3-, 5- ou 6-substituído podem ser cortados por $\text{Ce}(\text{NH}_4)_3(\text{NO}_3)_6$ (oxidação). Um linker de NCO_2 (uretano) pode ser cortado por hidróxido (base), ácido ou LiAlH_4 (redução). Um grupo linker de 3-pentenilo, 2-butenilo ou 1-butenilo pode ser cortado por O_3 , $\text{O}_5\text{O}_4/\text{IO}_4^-$ ou KMnO_4 (oxidação). Um linker de 2-[furil 3-, 4- ou 5-substituído]oxi pode ser cortado por O_2 , Br_2 , MeOH ou ácido.

As condições para o corte de outros linkers lábeis incluem: os linkers de t-alquiloxi podem ser cortados por ácido; os linkers de metil(dialquil)metoxi ou 2-alquil-1,3-dioxano-2-iloxi 4-substituído podem ser cortados por H_3O^+ ; os grupos 2-sililetoxi podem ser cortados por fluoretos ou ácido; os linkers 2-(X)-etoxi (em que X = ceto, amida de éster, ciano, NO_2 , sulfureto, sulfóxido, sulfona) podem ser cortados sob condições alcalinas; os linkers benziloxi 2-, 3-, 4-, 5- ou 6-substituídos podem ser cortados por condições acídicas ou de redução; os linkers de buteniloxi podem ser cortados por $(\text{Ph}_3\text{P})_3\text{RhCl}(\text{H})$; os linkers 2-bromofenoxy 3-, 4-, 5- ou 6-substituídos podem ser cortados por Li, Mg ou BuLi; os linkers de metiltiometoxi podem ser cortados por Hg^{2+} ; os linkers de 2-(X)-etiloxi (em que X = um halogénio) podem ser cortados por Zn ou Mg; os linkers de 2-hidroxietiloxi podem ser cortados por oxidação (p.ex., com $\text{Pb}(\text{OAc})_4$).

Os linkers preferidos são os susceptíveis a corte por ácido ou por fotólise. Diversos linkers que foram desenvolvidos para a síntese de péptidos em fase sólida têm utilidade na ligação de marcadores às MOIs. Alguns destes linkers estão descritos numa publicação recente de Lloyd-Williams et al. (*Tetrahedron* 49:11065-11133, 1993). Um tipo útil de linker é baseado em álcoois p-alcoxibenzílicos, dois dos quais, o ácido 4-hidroximetilfenoxiacético e o ácido 4-(4-hidroximetil-3-metoxifenoxy)butírico, são comercializados pela Advanced ChemTech (Louisville, KY). Ambos os linkers se podem ligar a um marcador através de uma ligação éster ao benzilálcool, e a uma MOI que contenha uma amina através de uma ligação amida ao ácido carboxílico. Os marcadores ligados a estas moléculas são libertados da MOI por acção de concentrações variáveis de ácido trifluoroacético. O corte de ligações destes linkers resulta na libertação de um ácido carboxílico no marcador. O corte por ácido das ligações entre marcadores e linkers como o 2,4-dimetoxi-4'-(carboximetiloxi)-benzidrilamina (fornecido pela Advanced ChemTech sob forma protegida por Fmoc) resulta na libertação de uma amida carboxílica no marcador libertado.

Na sua maioria, os linkers fotolábeis com utilidade no âmbito deste pedido de patente foram também desenvolvidos para a síntese de péptidos em fase sólida (ver a publicação de Lloyd-Williams). Estes linkers têm geralmente por base os 2-nitrobenzilésteres ou as 2-

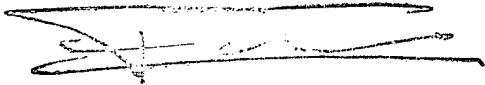


nitrobenzilamidas. Como exemplos de linkers fotolábeis recentemente referidos na literatura encontram-se o ácido 4-(4-(1-Fmoc-amino)etil)-2-metoxi-5-nitrofenoxi)butanóico (Holmes e Jones, *J. Org. Chem.* 60:2318-2319, 1995) e o ácido (3-Fmoc-amino)3-(2-nitrofenil)propiónico (Brown et al., *Molecular Diversity* 1:4-12, 1995). Ambos os linkers se podem ligar por meio do ácido carboxílico a uma amina presente na MOI. A ligação do marcador ao linker é efectuada através da formação de uma amida entre um ácido carboxílico do marcador e a amina do linker. O corte de ligações dos linkers fotolábeis é geralmente efectuado por meio de radiação UV com um comprimento de onda de 350 nm, usando intensidades e períodos de tempo conhecidos da técnica. Como exemplos de fontes comerciais de instrumentos que podem ser utilizados para o corte fotoquímico citam-se a Aura Industries Inc. (Staten Island, NY) e Agrenetics (Wilmington, MA). O corte de ligações dos linkers resulta na libertação de uma amida primária no marcador. Os exemplos de linkers fotolábeis incluem ésteres de nitrofenil-glicina, cloretos e metanosulfonatos de exo- e endo-2-benzonorborneilo e o ácido 3-amino-3(2-nitrofenil)propiónico. Como exemplos de corte enzimático de ligações encontram-se o corte de ligações éster pelas esterases, o corte das ligações fosfodiéster por acção das nucleases, o corte das ligações peptídicas pelas proteases, etc.

E. DETEÇÃO DE MARCADORES

Os métodos de detecção baseiam-se tipicamente na absorção ou emissão em algum tipo de zona do espectro de radiações. Quando os átomos ou as moléculas absorvem luz, a energia recebida excita uma estrutura quântica, fazendo-a passar para um nível mais elevado de energia. O tipo de excitação depende do comprimento de onda da luz. Os electrões são promovidos a orbitais mais elevadas através da luz ultravioleta ou visível, as vibrações moleculares são excitadas pela luz infravermelha e as rotações são excitadas pelas microondas. Um espectro de absorção é dado pela absorção da luz em função do comprimento de onda. O espectro de um átomo ou molécula depende da sua estrutura de níveis de energia. Os espectros de absorção têm utilidade na identificação de compostos. Os métodos espectroscópicos de absorção incluem a espectroscopia de absorção atómica (AA), a espectroscopia de infravermelhos (IR) e a espectroscopia UV-vis (uv-vis).

Os átomos ou moléculas que são excitados até níveis energéticos elevados podem decair até níveis mais baixos através da emissão de radiações. Esta emissão de luz é designada por fluorescência se a transição ocorrer entre estados com o mesmo spin, e fosforescência se a transição ocorrer entre estados com spins diferentes. A intensidade de emissão de um analito é linearmente proporcional à concentração (para concentrações

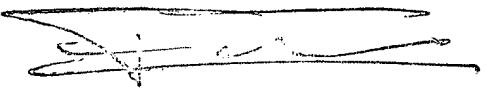


reduzidas), e é útil para a quantificação da espécie emissora. Os métodos espectroscópicos de emissão incluem a espectroscopia de emissão atómica (AES), a espectroscopia de fluorescência atómica (AFS), a fluorescência molecular induzida por laser (LIF) e a fluorescência a raios X (XRF).

Quando a matéria é atravessada por radiação electromagnética, a maior parte da radiação continua na sua direcção original mas uma pequena fracção é difractada em outras direcções. A luz que é difractada ao mesmo comprimento de onda que a luz incidente é designada por difracção de Rayleigh. A luz que é difractada em sólidos transparentes devido a vibrações (fonões) é designada por difracção de Brillouin. A luz difractada de Brillouin encontra-se desfasada de 0,1 a 1 número de onda relativamente à luz incidente. A luz que é difractada devido a vibrações das moléculas ou de fonões ópticos em sólidos opacos é designada por difracção de Raman. A luz difractada de Raman encontra-se desfasada em até 4000 números de onda relativamente à luz incidente. Os métodos espectroscópicos de difracção incluem a espectroscopia de Raman.

A espectroscopia de IR consiste na medição do comprimento de onda e da intensidade da absorção de luz da região média do infravermelho por parte de uma amostra. A luz da região média do infravermelho ($2,5 - 50 \mu\text{m}$, $4000 - 200 \text{ cm}^{-1}$) é suficientemente energética para excitar a transferência das vibrações moleculares para níveis energéticos mais elevados. O comprimento de onda das bandas de absorção em IR são características de tipos específicos de ligações químicas, e a espectroscopia de IR é geralmente muito útil para a identificação de moléculas orgânicas e organometálicas.

A espectroscopia de absorção em infravermelho próximo (NIR) consiste na medição do comprimento de onda e da intensidade da absorção de luz da região do infravermelho próximo por parte de uma amostra. A luz de infravermelho próximo abrange o intervalo de comprimentos de onda de $800 \text{ nm} - 2,5 \mu\text{m}$ ($12.500 - 4000 \text{ cm}^{-1}$), e é suficientemente energética para excitar sobretons e combinações de vibrações moleculares até níveis mais elevados de energia. A espectroscopia de NIR é tipicamente usada para a medição quantitativa de grupos funcionais orgânicos, especialmente O-H, N-H e C=O. Os componentes e a concepção dos instrumentos de NIR são similares aos dos espetrómetros de absorção uv-vis. A fonte de luz é geralmente uma lâmpada de tungsténio e o detector é geralmente um detector de PbS de fase sólida. Os suportes da amostra poderão ser de vidro ou quartzo e os solventes típicos são o CCl_4 e o CS_2 . A instrumentação conveniente da espectroscopia de NIR torna esta técnica adequada à monitorização on-line e ao controlo de processos.



A Espectroscopia de Absorção de Ultravioleta e Visível (uv-vis) constitui a medição do comprimento de onda e da intensidade da absorção de luz da região do ultravioleta próximo e da luz visível por parte de uma amostra. A absorção no UV de vácuo ocorre a 100 – 200 nm ($105\text{-}50.000\text{ cm}^{-1}$); no UV de quartzo a 200-350 nm ($50.000\text{-}28.570\text{ cm}^{-1}$); e no visível a 350-800 nm ($28.570\text{-}12.500\text{ cm}^{-1}$), e é descrita pela lei de Beer-Lambert-Bouquet. A luz ultravioleta e a luz visível são suficientemente energéticas para promover os electrões externos a níveis energéticos mais elevados. A espectroscopia UV-vis pode geralmente ser aplicada a moléculas e a iões ou complexos inorgânicos em solução. Os espectros uv-vis são limitados pelas características alargadas dos espectros. A fonte de luz corresponde geralmente a uma lâmpada de hidrogénio ou deutério para medições UV e a uma lâmpada de tungsténio para medições no visível. Os comprimentos de onda destas fontes de luz contínua são seleccionados através de um separador de comprimentos de onda, tal como um prisma ou um monocromador de grelha. Os espectros são obtidos através do varrimento do separador de comprimentos de onda, e as medições quantitativas podem ser feitas a partir de um espectro ou a um só comprimento de onda.

Os espectrómetros de massa usam a diferença entre as razões massa/carga (m/z) de átomos ou moléculas ionizados para os separar entre si. A espectrometria de massa é, deste modo, útil para a quantificação de átomos ou moléculas e também para a determinação de informações químicas e estruturais sobre moléculas. As moléculas possuem padrões distintos de fragmentação que proporcionam informações estruturais para a identificação de compostos. As operações gerais de um espectrómetro de massa são a seguir descritas. São criados iões em fase gasosa, os iões são separados no espaço ou no tempo com base na sua razão massa/carga, e a quantidade de iões correspondentes a cada razão massa/carga é medida. A potência de separação de iões de um espectrómetro de massa é descrita pela resolução, a qual é definida pela fórmula $R = m / \Delta m$, em que m é a massa do ião e Δm é a diferença em massa entre dois picos resolúveis de um espectro de massa. Por exemplo, um espectrómetro de massa com uma resolução de 1000 pode distinguir um ião com uma m/z de 100,0 de outro ião com uma m/z de 100,1.

Em geral, um espectrómetro de massa (MS) consiste em uma fonte de iões, um analisador selectivo de massas e um detector de iões. Os espectrómetros de sector magnético, quadrupolar e de tempo-de-voo requerem igualmente sistemas ópticos de extracção e aceleração de iões para transferir os iões para a região de fonte do analisador de massas. Os detalhes dos diversos tipos de analisadores de massa (para o MS de sector magnético, o MS quadrupolar ou o MS de tempo-de-voo) são abaixo expostos. Os analisadores de Focagem Simples para a MS de sector magnético utilizam uma trajectória



para o feixe de partículas de 180, 90 ou 60 graus. As várias forças que influenciam as partículas separam os iões com diferentes razões massa/carga. Com os analisadores de focagem dupla, é adicionado um analisador electrostático a este tipo de instrumento para separar partículas com diferentes energias cinéticas.

Um filtro quadrupolar de massa para a MS quadrupolar consiste em quatro cilindros metálicos dispostos em paralelo. As voltagens aplicadas afectam a trajectória dos iões que percorrem a trajectória centrada entre os quatro cilindros. Para dadas voltagens AC e DC, apenas os iões com uma certa razão massa/carga passam através do filtro quadrupolar, e todos os outros iões são desviados da sua trajectória original. É obtido um espectro de massa através da monitorização dos iões que passam através do filtro quadrupolar à medida que se fazem variar as voltagens nos cilindros.

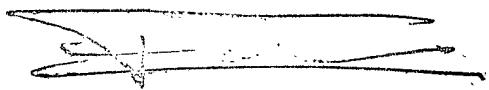
Um espectrómetro de massa de tempo-de-voo usa as diferenças no tempo de trânsito através de uma "região de flutuação" para separar os iões de diferentes massas. O aparelho funciona num modo pulsado, de tal forma que os iões terão de ser produzidos em impulsos e/ou extraídos em impulsos. Um campo eléctrico pulsado acelera todos os iões até que estes penetram numa região de flutuação subtraída à acção do campo com uma energia cinética de qV , em que q é a carga do ião e V é a voltagem aplicada. Uma vez que a energia cinética dos iões corresponde a $0,5 mV^2$, os iões mais leves adquirirão maior velocidade do que os iões mais pesados e atingirão mais depressa o detector localizado no final da região de flutuação. Os dados obtidos no detector de iões são visualizados num osciloscópio em função do tempo, produzindo-se o espectro de massa.

O processo de formação de iões constitui o ponto de partida para as análises por espectrometria de massa. A ionização química é um método que utiliza um reagente formador de iões que reage com as moléculas de analito (marcadores) para formar iões através da transferência de um protão ou hidreto. Os reagentes formadores de iões são produzidos através da introdução de metano em grande excesso (relativamente ao marcador) numa fonte de iões de impacto electrónico (EI). As colisões electrónicas produzem CH_4^+ e CH_3^+ , que prosseguem a reacção com o metano para originar CH_5^+ e C_2H_5^+ . Outro método de ionização de marcadores consiste no uso de plasma e de uma descarga em tubo de electrões. O plasma é um gás quente parcialmente ionizado que excita e ioniza átomos de modo eficaz. Uma descarga em tubo de electrões corresponde a um plasma a baixa pressão mantido entre dois eléctrodos. A ionização por impacto electrónico utiliza um feixe de electrões, usualmente gerado a partir de um filamento de tungsténio, para ionizar átomos ou moléculas em fase gasosa. Um electrão do feixe retira, por colisão, um electrão dos átomos ou moléculas de analito, criando iões. A ionização por electrospray



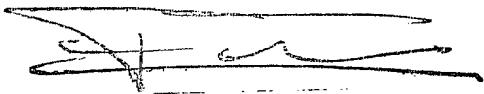
utiliza uma agulha muito fina e uma série de divisores de fluxo. Uma solução de amostra é vaporizada para a câmara de fonte, formando-se gotículas. As gotículas transportam uma carga ao sair do sistema capilar, e, à medida que o solvente se vaporiza, as gotículas desaparecem deixando moléculas de analito altamente carregadas. A ESI é particularmente útil para grandes moléculas biológicas que são difíceis de vaporizar ou ionizar. O bombardeamento por átomos rápidos (FAB) utiliza um feixe de elevada energia de átomos neutros, tipicamente Xe ou Ar, que atingem uma amostra sólida causando desadsorção e ionização. Esta técnica é usada para grandes moléculas biológicas que atingem dificilmente a fase gasosa. A técnica de FAB provoca um pequeno grau de fragmentação e proporciona geralmente um pico de um grande ião molecular, o que a torna útil para a determinação de pesos moleculares. O feixe atómico é produzido através da aceleração de iões provenientes de uma fonte de iões através de uma célula de conversão de cargas. Os iões captam um electrão em resultado de colisões com átomos neutros, formando-se um feixe de átomos de elevada energia. A ionização por laser (LIMS) é um método em que um impulso de laser remove material da superfície de uma amostra e cria um microplasma que ioniza alguns dos constituintes da amostra. A ionização por desadsorção com laser assistida por matriz (MALDI) constitui um método LIMS para vaporização e ionização de grandes moléculas biológicas, tais como proteínas ou fragmentos de DNA. As moléculas biológicas são dispersas numa matriz sólida, como o ácido nicotínico. Um impulso de laser UV remove a matriz, que arrasta algumas das moléculas de grandes dimensões para a fase gasosa sob uma forma ionizada, o que faz com que estas moléculas possam ser extraídas para um espectrómetro de massa. A ionização por desadsorção em plasma (PD) utiliza o decaimento de ^{252}Cf , o qual produz dois fragmentos de fissão que se deslocam em direcções opostas. Um dos fragmentos atinge a amostra, produzindo 1-10 iões de analito. O outro fragmento atinge o detector e desencadeia o processo de aquisição de dados. Este método de ionização é especialmente útil para grandes moléculas biológicas. A ionização por ressonância (RIMS) é um método em que um ou mais feixes de laser são sintonizados em ressonância com transições de um átomo ou molécula em fase gasosa de modo a promovê-lo, num processo em etapas, a um nível acima do seu potencial de ionização, criando-se um ião. A ionização secundária (SIMS) utiliza um feixe de iões, formado por $^3\text{He}^+$, $^{16}\text{O}^+$ ou $^{40}\text{Ar}^+$, que é focado sobre a superfície de uma amostra e lança material para a fase gasosa. A ionização por faísca é um método que ioniza analitos em amostras sólidas através da aplicação de uma corrente eléctrica pulsada entre dois eléctrodos.

Um marcador poderá receber uma carga antes, durante ou após o seu corte a partir da molécula a que se encontra ligado. Os métodos de ionização que se baseiam na



"desadsorção" de iões, a formação ou emissão directa de iões a partir de superfícies sólidas ou líquidas, têm permitido a utilização crescente de compostos não voláteis e termolábeis. Estes métodos eliminam a necessidade de volatilização das moléculas neutras antes da ionização e minimizam de maneira geral a degradação térmica das espécies moleculares. Estes métodos incluem a desadsorção em campo (Becky, *Principles of Field Ionization and Field Desorption Mass Spectrometry*, Pergamon, Oxford, 1977), a desadsorção por plasma (Sundqvist e Macfarlane, *Mass Spectrom. Rev.* 4:421, 1985), desadsorção por laser (Karas e Hillenkamp, *Anal. Chem.* 60:2299, 1988); Karas et al., *Angew. Chem.* 101:805, 1989), bombardeamento por partículas rápidas (p.ex., bombardeamento por átomos rápidos, FAB, e espectrometria de massa de iões secundários, SIMS, Barber et al., *Anal. Chem.* 54:645A, 1982), e ionização por termospray (TS) (Vestal, *Mass Spectrom. Rev.* 2:447, 1983). A técnica de termospray é largamente utilizada em combinação on-line com a cromatografia líquida. Os métodos de FAB em fluxo contínuo (Caprioli et al., *Anal. Chem.* 58:2949, 1986) demonstraram igualmente possuir um potencial significativo. Uma lista mais completa de combinações de ionização/espectrometria de massa inclui a espectrometria de massa por captura de iões, espectrometria de massa com ionização por electrospray, espectrometria de massa com spray de iões, espectrometria de massa com ionização em líquido, espectrometria de massa com ionização a pressão atmosférica, espectrometria de massa com ionização por electrões, espectrometria de massa com ionização por bombardeamento com átomos metaestáveis, espectrometria de massa com ionização por bombardeamento com átomos rápidos, espectrometria de massa MALDI, espectrometria de massa de tempo-de-voo com fotoionização, espectrometria de massa com micro-sonda de laser, espectrometria de massa MALDI-TOF, espectrometria de massa APCI, espectrometria de massa por nanospray, espectrometria de massa com ionização em spray nebulizado, espectrometria de massa com ionização química, espectrometria de massa com ionização por ressonância, espectrometria de massa com ionização secundária, espectrometria de massa por termospray.

Os métodos de ionização adaptáveis a compostos biológicos não voláteis possuem intervalos de aplicabilidade sobreponíveis. As eficiências de ionização são altamente dependentes da composição da matriz e do tipo de composto. Os resultados actualmente disponíveis indicam que a massa molecular máxima analisável por TS é de cerca de 8000 daltons (Jones e Krolik, *Rapid Comm. Mass Spectrom.* 1:67, 1987). Uma vez que a TS é realizada principalmente em associação a espetrómetros quadrupolares de massa, a sensibilidade é tipicamente afectada de modo desproporcionado por elevadas razões de massa/carga (m/z). Os espetrómetros de massa de tempo-de-voo (TOF) estão disponíveis

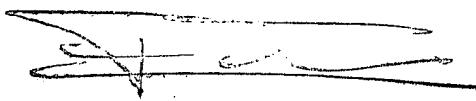


no mercado e possuem a vantagem de que o intervalo de razões m/z é limitado apenas pela eficiência do detector. Foram recentemente introduzidos dois outros métodos de ionização. Estes dois métodos são agora designados por desadsorção por laser assistida por matriz (MALDI, Karas e Hillenkamp, *Anal. Chem.* 60:2299, 1988; Karas et al., *Angew. Chem.* 101:805, 1989) e ionização por electrospray (ESI). Ambas as metodologias exibem uma eficiência de ionização muito elevada (i.e., uma razão de [iões moleculares produzidos]/[moléculas consumidas] muito elevada). A sensibilidade, que define em última análise o potencial da técnica, está dependente da dimensão da amostra, da quantidade de iões, da taxa de fluxo, da eficiência de detecção e da eficiência real de ionização.

A MS por electrospray baseia-se numa ideia inicialmente proposta nos anos 60 (Dole et al., *J. Chem. Phys.* 49:2240, 1968). A ionização por electrospray (ESI) constitui um meio de produção de moléculas carregadas para análise por espectrometria de massa. Abreviadamente, a ionização por electrospray produz gotículas altamente carregadas através da nebulização de líquidos num campo electrostático forte. As gotículas altamente carregadas, geralmente formadas num banho seco de gás à pressão atmosférica, diminuem de tamanho por evaporação do solvente neutro até que a repulsão de cargas se sobrepõe às forças de coesão, conduzindo a uma "explosão de Coulomb". O mecanismo exacto de ionização é controverso, tendo sido sugeridas várias hipóteses por diversos grupos de trabalho (Blades et al., *Anal. Chem.* 63:2109-14, 1991; Kebarle et al., *Anal. Chem.* 65:A972-86, 1993; Fenn, *J. Am. Soc. Mass Spectrom.*:524-35, 1993). Independentemente do processo de formação de iões utilizado, a ESI produz moléculas carregadas sob condições suaves a partir de uma solução.

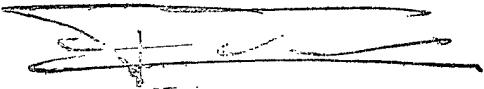
A capacidade de obter dados úteis por espectrometria de massa a partir de pequenas quantidades de uma molécula orgânica depende da produção eficiente de iões. A eficiência de ionização da ESI está relacionada com a extensão do processo de associação de uma carga positiva à molécula. Um outro método de melhoramento da ionização consiste em usar aminas quaternárias quando tal se mostra possível (ver Aebersold et al., *Protein Science* 1:494-503, 1992; Smith et al., *Anal. Chem.* 60:436-41, 1988).

A ionização por electrospray é em seguida descrita em maior detalhe. A produção de iões por electrospray requer dois passos: dispersão de gotículas altamente carregadas a pressão próxima da atmosférica, seguida de condições que induzem a evaporação. É passada uma solução de moléculas de analito através de uma agulha mantida a um potencial eléctrico elevado. Na extremidade da agulha, a solução dispersa-se num spray de gotículas pequenas e altamente carregadas contendo as moléculas de analito. As pequenas gotículas evaporam-se rapidamente e, através de um processo de desadsorção em campo



ou de evaporação residual, as moléculas protonadas de proteínas são libertadas para a fase gasosa. O electrospray é geralmente produzido por aplicação de um campo eléctrico elevado a um pequeno fluxo de líquido (geralmente 1-10 $\mu\text{l}/\text{min}$) num tubo capilar. É tipicamente aplicada uma diferença de potencial de 3-3 kV entre o capilar e um contra-electrodo situado a uma distância de 0,2-2 cm (zona a partir da qual os iões, os agregados moleculares carregados e até as gotículas carregadas, dependendo da extensão da dessolvatação, poderão ser recolhidos pelo MS através de um pequeno orifício). O campo eléctrico resulta numa acumulação de cargas à superfície do líquido na extremidade do capilar; desta forma, a taxa de fluxo do líquido, a resistividade e a tensão superficial são factores importantes na produção de gotículas. O elevado campo eléctrico resulta numa quebra de continuidade da superfície do líquido e na formação de gotículas líquidas altamente carregadas. Podem produzir-se gotículas com carga positiva ou negativa, dependendo da tendência conferida pelo capilar. O modo de iões negativos requer a presença de um captador de electrões, como o oxigénio, para inibir a descarga eléctrica.

Pode utilizar-se uma grande variedade de líquidos para produzir electrostaticamente um spray para um vácuo, sendo também possível utilizar um agente nebulizador. O uso de campos exclusivamente eléctricos para a nebulização debate-se com algumas restrições práticas a nível do intervalo de condutividade do líquido e da constante dieléctrica. É necessário ter uma condutividade da solução menor do que 10^{-5} ohms à temperatura ambiente para que seja possível obter um electrospray estável a taxas de fluxo de líquido úteis correspondendo a uma solução aquosa de electrólito de concentração $< 10^{-4}$ M. No modo considerado mais útil para a ESI-MS, uma taxa de fluxo de líquido apropriada resulta na dispersão do líquido numa névoa fina. A uma curta distância do capilar, o diâmetro das gotículas é frequentemente bastante uniforme e da ordem de 1 μm . Reveste-se de importância particular o facto de a corrente total de iões no electrospray sofrer apenas um aumento ligeiro com taxas de fluxo de líquido superiores. Existem indicações de que o aquecimento é útil para a manipulação do electrospray. Por exemplo, um ligeiro aquecimento possibilita uma produção rápida de electrospray a partir de soluções aquosas, presumivelmente devido à diminuição da viscosidade e da tensão superficial. Tanto o electrospray termicamente assistido como o electrospray assistido por nebulização de gás permitem a utilização de taxas mais elevadas de fluxo de líquido, mas diminuem a intensidade da carga das gotículas. A formação de iões moleculares requer condições que proporcionem a evaporação da população inicial de gotículas. Isto pode ser conseguido a pressões mais elevadas através de um fluxo de gás seco a temperaturas moderadas ($< 60^\circ\text{C}$), através de aquecimento durante o transporte através da interface, ou

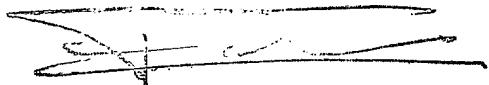


(particularmente no caso dos métodos de captura de iões) por colisões energéticas a pressão relativamente baixa.

Se bem que os processos detalhados subjacentes à ESI permaneçam obscuros, as gotículas muito pequenas produzidas pela ESI parecem possibilitar que praticamente qualquer espécie em solução que transporte uma carga residual seja transferida para a fase gasosa após a evaporação do solvente residual. A detecção por espectrometria de massa requer, então, que os iões possuam uma razão m/z apropriada (< 4000 daltons para os instrumentos quadrupolares) após a dessolvatação, e que a sua produção e transmissão sejam suficientemente eficientes. A larga gama de solutos que se verificou já adequar-se à ESI-MS, e a ausência de uma dependência substancial da eficiência de ionização relativamente ao peso molecular, sugerem que este processo de ionização é altamente não-discriminante e largamente aplicável.

A "fonte" de iões do electrospray funciona a uma pressão próxima da atmosférica. A "fonte" de electrospray consiste tipicamente num capilar de vidro ou metal que incorpora um método para imprimir uma tendência eléctrica à solução líquida relativamente a um contra-electrodo. As soluções, tipicamente misturas de metanol-água que contêm o analito e frequentemente outros aditivos como o ácido acético, fluem para a extremidade do terminal. Foi descrita uma fonte de ESI (Smith et al., *Anal. Chem.* 62:885, 1990) que pode adaptar-se a essencialmente qualquer sistema de solventes. As taxas de fluxo típicas em ESI encontram-se entre 1-10 $\mu\text{l}/\text{min}$. O principal requisito de uma interface ESI-MS é o de que a recolha e transporte de iões a partir da região de alta pressão para o MS se faça da maneira mais eficiente possível.

A eficiência da ESI pode atingir níveis muito elevados, fornecendo uma base para medições extremamente sensíveis, o que se torna útil no contexto do invento aqui descrito. A performance do instrumento pode fornecer uma corrente iónica total a nível do detector de cerca de 10^{-12} A , ou cerca de 10^7 contagens/s para as espécies de carga única. Com base na performance do instrumento, concentrações tão baixas como 10^{-10} M ou cerca de 10^{-18} mol/s de uma espécie de carga única fornecerão uma corrente iónica detectável (de cerca de 10 contagens/s) se o analito se encontrar completamente ionizado. Por exemplo, obtiveram-se limites de detecção da ordem das atomoles com iões de amónio quaternário usando uma interface entre ESI e a electroforese capilar de zona (Smith et al., *Anal. Chem.* 59:1230, 1988). Para um composto com um peso molecular de 1000, o número médio de cargas é igual a 1, o número aproximado de estados de carga é 1, a largura de picos (m/z) é 1 e a intensidade máxima (ião/s) é de 1×10^{12} .



O consumo de amostra necessário para a obtenção de um espectro de ESI é notavelmente baixo (Smith et al., *Anal. Chem.* 60:1948, 1988). Poderão igualmente obter-se ganhos substanciais através do uso de um conjunto de detectores em instrumentos de sector, permitindo a detecção em simultâneo de porções do espectro. Uma vez que actualmente apenas cerca de 10^{-5} de todos os iões formados por ESI são detectados, a atenção dedicada aos factores que limitam a performance do instrumento poderão fornecer a base para uma sensibilidade melhorada. Tornar-se-á evidente aos peritos na técnica que o presente invento contempla e se adapta a melhoramentos das tecnologias de ionização e detecção.

É preferencialmente colocada uma interface entre a instrumentação de separação (p.ex., gel) e o detector (p.ex., espectrómetro de massa). A interface possui preferencialmente as seguintes propriedades: (1) a capacidade para recolher os fragmentos de DNA em curtos intervalos de tempo, (2) concentrar os fragmentos de DNA, (3) remover os fragmentos de DNA a partir dos tampões de electroforese e do meio, (4) cortar as ligações entre o marcador e o fragmento de DNA, (5) separar o marcador do fragmento de DNA, (6) descartar o fragmento de DNA, (7) colocar o marcador numa solução volátil, (8) volatilizar e ionizar o marcador, e (9) colocar ou transportar o marcador para um dispositivo de electrospray que introduza o marcador no espetrómetro de massa.

A interface possui também a capacidade de "recolher" os fragmentos de DNA à medida que estes vão sendo eluídos a partir do fundo de um gel. O suporte de gel poderá ser composto por uma placa de gel, um gel tubular, um capilar, etc. Os fragmentos de DNA poderão ser recolhidos por meio de diversos métodos. O primeiro método corresponde ao uso de um campo eléctrico, sendo os fragmentos de DNA recolhidos a partir do elecetrodo ou na vizinhança deste. Num segundo método, os fragmentos de DNA são recolhidos fazendo passar um fluxo de líquido pelo fundo do gel. Poderão combinar-se aspectos dos dois métodos, sendo o DNA recolhido a partir de um fluxo de líquido que pode posteriormente ser concentrado através do uso de um campo eléctrico. O resultado final é o de que os fragmentos de DNA são removidos do meio ao qual foi aplicado o método de separação. Assim, os fragmentos de DNA poderão ser "arrastados" de um tipo de solução para outro através do uso de um campo eléctrico.

Uma vez colocados os fragmentos de DNA em solução apropriada (compatível com o electrospray ou a espetrometria de massa), o marcador pode ser cortado do fragmento de DNA. O fragmento de DNA (ou seus resíduos) pode então ser separado do marcador através da aplicação de um campo eléctrico (de preferência, o marcador possui uma carga

oposta à do fragmento de DNA). O marcador é em seguida introduzido no dispositivo de electrospray por meio de um campo eléctrico ou de um fluxo de líquido.

Os marcadores fluorescentes podem ser identificados e quantificados mais directamente através dos seus comprimentos de onda e intensidades de absorção e de emissão de fluorescência.

Se bem que um espectrofluorómetro seja um aparelho extremamente flexível, fornecendo intervalos contínuos de comprimentos de onda de excitação e emissão (λ_{EX} , λ_{S1} , λ_{S2}), os instrumentos mais especializados, tais como os citómetros de fluxo e os microscópios de varrimento por laser, requerem sondas que sejam excitáveis a um único comprimento de onda fixo. Nos instrumentos actuais, isto corresponde geralmente à linha de 488 nm do laser de árgon.

A intensidade da fluorescência por molécula de sonda é proporcional ao produto de ϵ e QY. O intervalo destes parâmetros, para os fluoróforos de importância prática, corresponde a cerca de 10.000 a 100.000 $\text{cm}^{-1}\text{M}^{-1}$ para ϵ e de 0,1 a 1,0 para QY. Quando a absorção é levada até à saturação através do uso de iluminação de elevada intensidade, a destruição irreversível do fluoróforo excitado (fotolexiviação) torna-se o factor limitante da detectabilidade da fluorescência. O impacto prático da fotolexiviação depende da técnica de detecção de fluorescência envolvida.

Tornar-se-á evidente a um técnico que é possível colocar um dispositivo (uma interface) entre os passos de separação e de detecção para permitir o funcionamento contínuo da separação por dimensões e da detecção do marcador (em tempo real). Este procedimento reune a metodologia e instrumentação de separação com a metodologia e instrumentação de detecção, formando um só dispositivo. Por exemplo, é interposta uma interface entre uma técnica de separação e a detecção por espectrometria de massa ou amperometria potenciosistática.

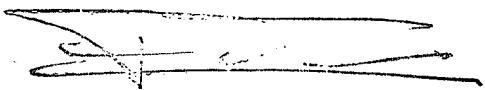
A função da interface é primariamente a libertação do marcador (p.ex., de espectrometria de massa) a partir do analito. Existem diversas implementações representativas da interface. A concepção da interface está dependente da escolha dos linkers susceptíveis ao corte. No caso da luz ou dos linkers susceptíveis ao corte fotolítico, é necessária uma fonte de energia ou de fotões. No caso de um linker ácido-lábil, de um linker lábil a bases ou de um linker de dissulfureto, é necessária a adição do reagente à interface. No caso dos linkers lábeis ao calor, é necessária uma fonte de energia calorífica. É requerida a adição de enzimas para um linker sensível a enzimas, como acontece com uma protease específica e um linker de péptidos, uma nuclease e um linker de DNA ou RNA, uma glicosilase, HRP ou fosfatase e um linker que é instável após o corte (p.ex., similar aos



substratos quimioluminescentes). Outras características da interface incluem um alargamento mínimo das bandas e a separação entre o DNA e os marcadores antes da injecção destes num espectrómetro de massa. As técnicas de separação incluem as que se baseiam em métodos e técnicas electroforéticas, técnicas de afinidade, retenção por dimensões (diálise), filtração e semelhantes.

É igualmente possível concentrar os marcadores (ou a construção de ácido nucleico-linker-marcador), capturá-los electroforeticamente e libertá-los então para uma corrente alternativa de reagente que seja compatível com o tipo particular de método de ionização seleccionado. A interface poderá igualmente possuir a capacidade de capturar os marcadores (ou a construção de ácido nucleico-linker-marcador) em microesferas, lançando a(s) esfera(s) para a câmara e realizando então a desadsorção por laser/vaporização. É também possível fazer uma extracção por fluxo para um tampão alternativo (p.ex., de um tampão de electroforese capilar para um tampão hidrofóbico através de uma membrana permeável). Em alguns casos poderá igualmente ser desejável a libertação intermitente de marcadores para o espectrómetro de massa, o que acrescentaria uma outra função à interface. Outra das funções da interface consiste na libertação de marcadores a partir de múltiplas colunas para um espectrómetro de massa, através de uma fenda de introdução rotativa em função do tempo para cada coluna. Adicionalmente, é possível promover a libertação de marcadores a partir de uma só coluna para detectores MS múltiplos, de uma forma espaçada no tempo, recolher cada conjunto de marcadores durante alguns milisegundos, e libertá-los em seguida pra um espectrómetro de massa.

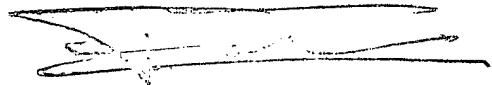
É a seguir apresentada uma lista de fornecedores representativos de tecnologias de separação e detecção que podem ser usadas no âmbito do presente invento. A Hoefner Scientific Instruments (San Francisco, CA) fabrica equipamentos de electroforese (Two StepTM, Poker FaceTM II) para aplicações em reacções de sequenciação. A Pharmacia Biotech (Piscataway, NJ) fabrica equipamentos de electroforese para separações e sequenciação de DNA (PhastSystem para análise por PCR-SSCP, MacroPhor System para a sequenciação de DNA). A Perkin Elmer/Applied Biosystems Division (ABI, Foster City, CA) fabrica sequenciadores semi-automatizados baseados em corantes fluorescentes (ABI373 e ABI377). A Analytical Spectral Devices (Boulder, CO) fabrica espectrómetros de UV. A Hitachi Instruments (Tóquio, Japão) fabrica espectrómetros de Absorção Atómica, espectrómetros de Fluorescência, Espectrómetros de Massa LC e GC, espectrómetros de NMR e Espectrómetros de UV-VIS. A PerSeptive Biosystems (Framingham, MA) produz Espectrómetros de Massa (VoyagerTM Elite). A Bruker Instruments Inc. (Manning Park, MA) fabrica Espectrómetros FTIR (Vector 22), Espectrómetros FT-Raman, Espectrómetros de



Massa de Tempo-de-Voo (Reflex IITM), Espectrómetro de Massa de Captura de Electrões (EsquireTM) e um Espectrómetro de Massa MALDI. A Analytical Technology Inc. (ATI, Boston, MA) produz unidades de Electroforese Capilar em Gel, detectores UV e Detectores de Conjunto de Díodos. A Teledyne Electronic Technologies (Mountain View, CA) fabrica um Espectrómetro de Massa de Captura de Iões (3DQ DiscoveryTM e 3DQ ApogeeTM). A Perkin Elmer/Applied Biosystems Division (Foster City, CA) fabrica um Espectrómetro de Massa Sciex (quadrupolo triplo LC/MS/MS, API 100/300) que é compatível com o método de electrospray. A Hewlett-Packard (Santa Clara, CA) produz Detectores Selectivos de Massa (HP 5972A), Espectrómetros de Massa MALDI-TOF (HP G2025A), Detectores de Conjunto de Díodos, unidades de CE, unidades de HPLC (HP1090) e Espectrómetros de UV. A Finigan Corporation (San Jose, CA) fabrica espectrómetros de massa (de sector magnético (MAT 95 STM), quadrupolares (MAT 95SQTM) e quatro outros tipos de espectrómetros associados). A Rainin (Emeryville, CA) fabrica instrumentos de HPLC.

Os métodos e composições aqui descritos possibilitam que os marcadores cortados sejam usados como mapas de um tipo de amostra em particular e da identidade de nucleótidos. No início de cada método de sequenciação, é atribuído um primer particular (seleccionado) a um marcador particular individualizado. Os marcadores mapeiam um tipo de primer, que por sua vez mapeia um tipo de vector, que por seu lado mapeia uma identidade de amostra. O marcador poderá também mapear um tipo de terminador didesoxi (ddTTP, ddCTP, ddGTP, ddATP), sendo o primer marcado colocado na respectiva reacção didesoxinucleotídica para funcionar como referência. A reacção de sequenciação é então realizada e os fragmentos resultantes são sequencialmente separados por dimensões ao longo do tempo.

Os marcadores são cortados a partir dos fragmentos numa estrutura temporal e são medidos e registados numa estrutura temporal. A sequência é construída através da comparação do mapa do marcador com a estrutura temporal. Isto significa que todas as identidades dos marcadores são registadas ao longo do tempo após o passo de separação por dimensões e se relacionam entre si numa estrutura temporal. O passo de separação por dimensões separa os fragmentos de ácido nucleico por incrementos de um nucleótido, e desta forma as identidades de marcador relacionadas encontram-se separadas entre si por um só nucleótido. Por meio do conhecimento prévio do mapa do terminador didesoxi ou do mapa de nucleótidos e do tipo de amostra, a sequência é prontamente deduzida de modo linear.



Os exemplos que se seguem são fornecidos com fins ilustrativos e não pretendem limitar o âmbito do invento.

Salvo indicação em contrário, os produtos químicos usados nos exemplos poderão ser adquiridos à Aldrich Chemical Company, Milwaukee, WI. São usadas as abreviaturas que se seguem, com os significados indicados:

ANP = ácido 3-(Fmoc-amino)-3-(2-nitrofenil)propiónico

NBA = ácido 4-(Fmoc-aminometil)-3-nitrobenzóico

HATU = Hexafluorofosfato de O-7-azabenzotriazol-1-il-N,N,N',N'-tetrametilurónio

DIEA = diisopropiletilamina

MCT = monoclorotriazina

NMM = 4-metilpirrolidona

ACT357 = sintetizador de péptidos ACT357 da Advanced ChenTech, Inc., Louisville, KY

ACT = Advanced ChenTech, Inc., Louisville, KY

NovaBiochem = CalBiochem-NovaBiochem International, San Diego, CA

TFA = Ácido trifluoroacético

Tfa = Trifluoroacetilo

INIP = ácido N-metilisopecótico

Tfp = Tetrafluorofenilo

DIAEA = 2(Diisopropilamino)etilamina

MCT = monoclorotriazeno

5'-AH-ODN = oligonucleórido marcado com aminohexilo em 5'

EXEMPLOS

EXEMPLO 1

PREPARAÇÃO DE LINKERS ÁCIDO-LÁBEIS PARA USO EM SEQUENCIAÇÃO COM MARCADORES SUSCEPTÍVEIS DE CORTE

- A. Síntese de Ésteres de Pentafluorofenilo de Marcadores de Espectroscopia de Massa Susceptíveis ao Corte por Via Química, para Libertar Marcadores com Extremidades de Amida Carboxílica

A Figura 1 apresenta o esquema de reacção.

Passo A. A resina TentaGel S AC (composto II; comercializada pela ACT; 1 eq.) é suspensa em DMF no vaso de recolha do sintetizador de péptidos ACT357 (ACT). São adicionados o composto I (3 eq.), HATU (3 eq.) e DIEA (7,5 eq.) em DMF, e o vaso de recolha é agitado durante 1 h. O solvente é removido e a resina é lavada com NMP (2X), MeOH (2X) e DMF (2X). O acoplamento de I à resina e os passos de lavagem são repetidos para originar o composto III.

Passo B. A resina (composto III) é misturada com piperidina a 25% em DMF e agitada durante 5 min. A resina é filtrada, sendo então misturada com piperidina a 25% em DMF e agitada durante 10 min. O solvente é removido e a resina é lavada com NMP (2X), MeOH (2X) e DMF (2X), sendo então directamente usada no passo C.

Passo C. A resina desprotegida obtida no passo B é suspensa em DMF e é adicionada à suspensão um aminoácido protegido por Fmoc, contendo uma funcionalidade amina na sua cadeia lateral (composto IV, p.ex. alfa-N-Fmoc-3-(3-piridil)-alanina, comercializado pela Synthetech, Albany, OR; 3 eq.), HATU (3 eq.) e DIEA (7,5 eq.) em DMF. O vaso é agitado durante 1 hora. O solvente é removido e a resina é lavada com NMP (2X), MeOH (2X) e DMF (2X). O acoplamento do composto IV à resina e os passos de lavagem são repetidos para originar o composto V.

Passo D. A resina (composto V) é tratada com piperidina tal como descrito no passo B para remover o grupo Fmoc. A resina desprotegida é então dividida em 16 partes iguais pelo ACT357 a partir do vaso de recolha e passado para 16 vasos de reacção.

Passo E. As 16 alíquotas de resina desprotegida obtida no passo D são suspensas em DMF. A cada vaso de reacção é adicionado o ácido carboxílico apropriado VI₁₋₁₆ ($R_{1-16}CO_2H$; 3 eq.), HATU (3 eq.) e DIEA (7,5 eq.) em DMF. Os vasos de reacção são agitados durante 1 hora. O solvente é removido e as alíquotas de resina são lavadas com NMP (2X), MeOH (2X) e DMF (2X). O acoplamento de VI₁₋₁₆ às alíquotas de resina e os passos de lavagem são repetidos para originar os compostos VII₁₋₁₆.

Passo F. As alíquotas de resina (compostos VII₁₋₁₆) são lavadas com CH₂Cl₂ (3X). É adicionado a cada vaso de reacção TFA a 1% em CH₂Cl₂ e os vasos de reacção são agitados durante 30 min. O solvente é filtrado a partir dos vasos de reacção para tubos individuais. As alíquotas de resina são lavadas com CH₂Cl₂ (2X) e MeOH (2X) e os filtrados

são combinados nos tubos individuais. Os tubos individuais são evaporados em vácuo, fornecendo os compostos VIII₁₋₁₆.

Passo G. Cada um dos ácidos carboxílicos livres VIII₁₋₁₆ é dissolvido em DMF. A cada solução é adicionada piridina (1,05 eq.), seguindo-se a adição de trifluoroacetato de pentafluorofenilo (1,1 eq.). As misturas são agitadas durante 45 min. à temperatura ambiente. As soluções são diluídas com EtOAc, lavadas com ácido cítrico em solução aquosa a 1 M (3X) e com NaHCO₃ em solução aquosa a 5% (3X), secadas em Na₂SO₄, filtradas e evaporadas em vácuo, fornecendo os compostos IX₁₋₁₆.

B. Síntese de Ésteres de Pentafluorofenilo de Marcadores de Espectroscopia de Massa Susceptíveis ao Corte por Via Química, para Libertar Marcadores com Extremidades de Ácido Carboxílico

A Figura 2 apresenta o esquema de reacção.

Passo A. O ácido 4-(hidroximetil)fenoxibutírico (composto I; 1 eq.) é combinado com DIEA (2,1 eq.) e brometo de alilo (2,1 eq.) em CHCl₃ e aquecido até refluxo durante 2 h. A mistura é diluída com EtOAc, lavada com HCl 1 N (2X), tampão de carbonato a pH 9,5 (2X) e solução concentrada de sal (1X), secada em Na₂SO₄ e evaporada em vácuo para fornecer o éster de alilo do composto I.

Passo B. O éster de alilo do composto I obtido no passo A (1,75 eq.) é combinado em CH₂Cl₂ com um aminoácido protegido por Fmoc contendo uma funcionalidade amina na sua cadeia lateral (composto II, p.ex. alfa-N-Fmoc-3-(3-piridil)-alanina, comercializado pela Synthetech, Albany, OR; 1 eq.), N-metilmorfolina (2,5 eq.) e HATU (1,1 eq.), com agitação à temperatura ambiente durante 1 h. A mistura é diluída com CH₂Cl₂, lavada com ácido cítrico em solução aquosa a 1 M (2X), água (1X) e NaHCO₃ em solução aquosa a 5% (2X), secada em Na₂SO₄ e evaporada em vácuo. O composto III é isolado por cromatografia de flash (CH₂Cl₂ → EtOAc).

Passo C. O composto 3 é dissolvido em CH₂Cl₂, adicionando-se então Pd(PPh₃)₄ (0,07 eq.) e N-metilanilina (2 eq.), sendo a mistura agitada à temperatura ambiente durante 4 horas. A mistura é diluída com CH₂Cl₂, lavada com ácido cítrico em solução aquosa a 1 M (2X) e



água (1X), secada em Na_2SO_4 e evaporada em vácuo. O composto IV é isolado por cromatografia de flash ($\text{CH}_2\text{Cl}_2 \rightarrow \text{EtOAc} + \text{HOAc}$).

Passo D. A resina TentaGel S AC (composto V; 1 eq.) é suspensa em DMF no vaso de recolha do sintetizador de péptidos ACT357 (Advanced ChemTech Inc. (ACT), Louisville, KY). São adicionados o composto IV (3 eq.), HATU (3 eq.) e DIEA (7,5 eq.) em DMF, e o vaso de recolha é agitado durante 1 h. O solvente é removido e a resina é lavada com NMP (2X), MeOH (2X) e DMF (2X). O acoplamento de IV à resina e os passos de lavagem são repetidos para originar o composto VI.

Passo E. A resina (composto VI) é misturada com piperidina a 25% em DMF e agitada durante 5 min. A resina é filtrada, sendo então misturada com piperidina a 25% em DMF e agitada durante 10 min. O solvente é removido e a resina é lavada com NMP (2X), MeOH (2X) e DMF (2X). A resina desprotegida é então dividida em 16 partes iguais pelo ACT357 a partir do vaso de recolha e passado para 16 vasos de reacção.

Passo F. As 16 alíquotas de resina desprotegida obtida no passo E são suspensas em DMF. A cada vaso de reacção é adicionado o ácido carboxílico apropriado VII_{1-16} ($\text{R}_{1-16}\text{CO}_2\text{H}$; 3 eq.), HATU (3 eq.) e DIEA (7,5 eq.) em DMF. Os vasos de reacção são agitados durante 1 hora. O solvente é removido e as alíquotas de resina são lavadas com NMP (2X), MeOH (2X) e DMF (2X). O acoplamento de VII_{1-16} às alíquotas de resina e os passos de lavagem são repetidos para originar os compostos VIII_{1-16} .

Passo G. As alíquotas de resina (compostos VIII_{1-16}) são lavadas com CH_2Cl_2 (3X). É adicionado a cada vaso de reacção TFA a 1% em CH_2Cl_2 e os vasos de reacção são agitados durante 30 min. O solvente é filtrado a partir dos vasos de reacção para tubos individuais. As alíquotas de resina são lavadas com CH_2Cl_2 (2X) e MeOH (2X) e os filtrados são combinados nos tubos individuais. Os tubos individuais são evaporados em vácuo, fornecendo os compostos IX_{1-16} .

Passo H. Cada um dos ácidos carboxílicos livres IX_{1-16} é dissolvido em DMF. A cada solução é adicionada piridina (1,05 eq.), seguindo-se a adição de trifluoroacetato de pentafluorofenilo (1,1 eq.). As misturas são agitadas durante 45 min. à temperatura ambiente. As soluções são diluídas com EtOAc, lavadas com ácido cítrico em solução aquosa a 1 M (3X) e com



NaHCO₃ em solução aquosa a 5% (3X), secadas em Na₂SO₄, filtradas e evaporadas em vácuo, fornecendo os compostos X₁₋₁₆.

EXEMPLO 2

DEMONSTRAÇÃO DO CORTE FOTOLÍTICO DE T-L-X

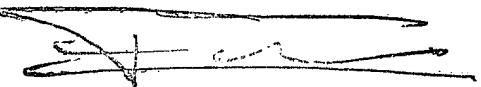
Um composto T-L-X como o preparado no Exemplo 11 foi irradiado com luz de UV próximo durante 7 minutos à temperatura ambiente. É usada uma lâmpada de fluorescência UV Rayonet (Southern New England Ultraviolet Co., Middletown, CT) com um pico de emissão a 350 nm como fonte de luz UV. A lâmpada é colocada a uma distância de 15 cm das caixas de Petri que contêm as amostras. A electroforese em gel de SDS mostra que >85% do conjugado é cortado nestas condições.

EXEMPLO 3

PREPARAÇÃO DE PRIMERS COM MARCADOR FLUORESCENTE E DEMONSTRAÇÃO DO CORTE DO FLUORÓFORO

Síntese e Purificação de Oligonucleótidos

Os oligonucleótidos (ODNs) são preparados em sintetizadores automatizados de DNA usando os reagentes químicos convencionais de fosforamidito comercialmente disponíveis, ou os reagentes químicos de H-fosfonato (Glenn Research Sterling, VA). Os fosforamiditos de dA, dG, dC e T bloqueados encontram-se comercialmente disponíveis nestas formas, e os nucleósidos sintéticos podem facilmente ser convertidos a uma forma apropriada. Os oligonucleótidos são preparados usando o fosforamidito convencional existente no mercado ou o reagente químico de H-fosfonato. Os oligonucleótidos são purificados através de adaptações dos métodos convencionais. Os oligonucleótidos com grupos 5'-trítilo são submetidas a cromatografia de HPLC usando uma coluna de fase reversa de 12 micrómetros e 300 # Rainin (Emeryville, CA) Dynamax C-8 com 4,2x250 mm, usando um gradiente de 15% a 55% de MeCN em Et₃NH⁺OAc⁻, pH 7,0, durante 20 minutos. Após a realização da destilação, os oligonucleótidos são adicionalmente purificados por cromatografia de exclusão em gel. As verificações analíticas relativas à qualidade dos oligonucleótidos são efectuadas numa coluna de PRP (Alltech, Deerfield, IL) a pH alcalino e por PAGE.



Preparação de oligonucleótidos derivatizados com 2,4,6-triclorotriazina: fazem-se reagir 10 a 1000 µg de oligonucleótido ligado a amina pela extremidade 5' com um excesso de cloreto cianúrico recristalizado em n-metil-pirrolidona a 10% em tampão alcalino (pH 8,3 a 8,5, de preferência), a 19 – 25°C, durante 30 a 120 minutos. As condições de final de reacção consistem em borato de sódio 0,15 M, pH 8,3, cloreto cianúrico recristalizado a 2 mg/ml e respectivo oligonucleótido a 500 µg/ml. O cloreto cianúrico que não reagiu é removido por cromatografia de exclusão por dimensões numa coluna de Sephadex G-50 (Pharmacia, Piscataway, NJ).

Em seguida, faz-se reagir o oligonucleótido activado purificado com um excesso molar de 100 vezes de cistamina em borato de sódio 0,15 M, pH 8,3, durante 1 hora à temperatura ambiente. A cistamina que não reagiu é removida por cromatografia de exclusão por dimensões numa coluna de Sephadex G-50. Fazem-se então reagir os ODNs derivatizados com fluorocromos reactivos a aminas. A preparação de ODNs derivatizados é dividida em três porções e cada porção é posta em reacção com (a) um excesso molar de 20 vezes de cloreto de sulfônico de Texas Red (Molecular Probes, Eugene, OR), com (b) um excesso molar de 20 vezes de cloreto de sulfônico de Lissamina (Molecular Probes, Eugene, OR) ou (c) um excesso molar de 20 vezes de isotiocianato de fluoresceína. As condições de final de reacção consistem em borato de sódio 0,15 M, pH 8,3, durante 1 hora à temperatura ambiente. Os fluorocromos que não reagiram são removidos por cromatografia de exclusão por dimensões numa coluna de Sephadex G-50.

Para cortar o fluorocromo a partir do oligonucleótido, os ODNs são ajustados a uma concentração de 1×10^{-5} molar e são então feitas diluições (diluições de 12,3 vezes) em TE (TE corresponde a Tris 0,01 M, pH 7,0, EDTA 5 mM). A volumes de 100 µl de ODNs são adicionados 25 µl de ditiotreitol (DTT) 0,01 M. A um conjunto idêntico de controlos não é adicionado DTT. A mistura é incubada durante 15 minutos à temperatura ambiente. A fluorescência é medida numa placa negra de microtitulação. A solução é removida dos tubos de incubação (150 microlitros) e colocada numa placa negra de microtitulação (Dynatek Laboratories, Chantilly, VA). As placas são então lidas directamente por meio de um fluorómetro Fluoroskan II (Flow Laboratories, McLean, VA) utilizando um comprimento de onda de excitação de 495 nm e monitorizando a emissão a 520 nm para a fluoresceína, utilizando um comprimento de onda de excitação de 591 nm e monitorizando a emissão a 612 nm para o Texas Red, e utilizando um comprimento de onda de excitação de 570 nm e monitorizando a emissão a 590 nm para a lissamina.

Moles de Fluorocromo	RFU não cortado	RFU cortado	RFU Livre
$1,0 \times 10^5$ M	6,4	1200	1345
$3,3 \times 10^6$ M	2,4	451	456
$1,1 \times 10^6$ M	0,9	135	130
$3,7 \times 10^7$ M	0,3	44	48
$1,2 \times 10^7$ M	0,12	15,3	16,0
$4,1 \times 10^7$ M	0,14	4,9	5,1
$1,4 \times 10^8$ M	0,13	2,5	2,8
$4,5 \times 10^9$ M	0,12	0,8	0,9

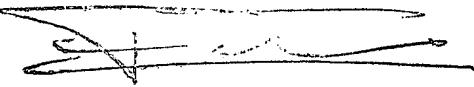
Os dados indicam que se verifica um aumento de cerca de 200 vezes na fluorescência relativa quando o fluorocromo é cortado a partir do ODN.

EXEMPLO 4

PREPARAÇÃO DE PRIMERS DE SEQUÊNCIA M13 MARCADOS E DEMONSTRAÇÃO DO CORTE DOS MARCADORES

Preparação de oligonucleótidos derivatizados com 2,4,6-triclorotriazina: fazem-se reagir 1000 µg de oligonucleótilo ligado a amina pela extremidade 5' (5'-hexilamina-TGTAAACGACGGCCAGT-3') (ID da Seq. Nº 1) com um excesso de cloreto cianúrico recristalizado em n-metil-pirrolidona a 10% em tampão alcalino (pH 8,3 a 8,5, de preferência), a 19 – 25°C, durante 30 a 120 minutos. As condições de final de reacção consistem em borato de sódio 0,15 M, pH 8,3, cloreto cianúrico recristalizado a 2 mg/ml e respectivo oligonucleótilo a 500 µg/ml. O cloreto cianúrico que não reagiu é removido por cromatografia de exclusão por dimensões numa coluna de Sephadex G-50 (Pharmacia, Piscataway, NJ).

Em seguida, faz-se reagir o oligonucleótilo activado purificado com um excesso molar de 100 vezes de cistamina em borato de sódio 0,15 M, pH 8,3, durante 1 hora à temperatura ambiente. A cistamina que não reagiu é removida por cromatografia de exclusão por dimensões numa coluna de Sephadex G-50. Fazem-se então reagir os ODNs derivatizados com diversas aminas. A preparação derivatizada de ODNs é dividida em 12 porções e cada porção é posta em reacção (em excesso molar de 25 vezes) com os ésteres de pentafluorofenilo de: (1) ácido 4-metoxibenzóico, (2) ácido 4-fluorobenzóico, (3) ácido toluico, (4) ácido benzóico, (5) ácido indol-3-acético, (6) ácido 2,6-difluorobenzóico, (7) N-óxido de ácido nicotínico, (8) ácido 2-nitrobenzóico, (9) ácido 5-acetilsalicílico, (10) ácido 4-



etoxibenzóico, (11) ácido cinâmico, (12) ácido 3-aminonicotínico. A reacção é efectuada durante 2 horas a 37°C em borato de Na 0,2 M, pH 8,3. Os ODNs derivatizados são purificados por cromatografia de exclusão em gel numa coluna de Sephadex G-50.

Para cortar o marcador a partir do oligonucleótido, os ODNs são ajustados a uma concentração de 1×10^{-5} molar e são então feitas diluições (diluições de 12,3 vezes) em TE (TE corresponde a Tris 0,01 M, pH 7,0, EDTA 5 mM) com 50% de EtOH (V/V). A volumes de 100 μ l de ODNs são adicionados 25 μ l de ditioreitol (DTT) 0,01 M. A um conjunto idêntico de controlos não é adicionado DTT. A mistura é incubada durante 30 minutos à temperatura ambiente. É adicionado NaCl a 0,1 M e são adicionados 2 volumes de EtOH para precipitar os ODNs. Os ODNs são removidos da solução por centrifugação a 14.000 x G a 4°C durante 15 minutos. Os sobrenadantes são reservados e secos por completo. O pellet é então dissolvido em 25 μ L de MeOH. O pellet é então testado por espectrometria de massa quanto à presença de marcadores.

O espectrómetro de massa usado neste trabalho é um espectrómetro de massa de transformação de Fourier com fonte externa de iões (FTMS). As amostras preparadas para análise MALDI são depositadas na ponta de uma sonda directa e inseridas na fonte de iões. Quando a amostra é irradiada com um impulso de laser, são extraídos iões da fonte que passam para um longo canal quadrupolar de guia de iões, que focaliza e transporta os iões para uma célula de analisador FTMS localizada no interior do orifício de um magneto supercondutor.

Os espectros fornecem a informação que se segue. São encontrados picos de intensidade que varia entre 25 a 100 unidades de intensidade relativa para os seguintes pesos moleculares: (1) 212,1 amu indicando um derivado do ácido 4-metoxibenzóico, (2) 200,1 amu indicando um derivado do ácido 4-fluorobenzóico, (3) 196,1 amu indicando um derivado do ácido toluico, (4) 182,1 amu indicando um derivado do ácido benzóico, (5) 235,2 amu indicando um derivado do ácido indol-3-acético, (6) 218,1 amu indicando um derivado do ácido 2,6-difluorobenzóico, (7) 199,1 amu indicando um derivado N-óxido do ácido nicotínico, (8) 227,1 amu indicando 2-nitrobenzamida, (9) 179,18 amu indicando um derivado do ácido 5-acetilsalicílico, (10) 226,1 amu indicando um derivado do ácido 4- etoxibenzóico, (11) 209,1 amu indicando um derivado do ácido cinâmico, (12) 198,1 amu indicando um derivado do ácido 3-aminonicotínico.

Os resultados indicam que os marcadores são cortados a partir dos primers e são detectáveis por espectrometria de massa.

EXEMPLO 5

PREPARAÇÃO DE UM CONJUNTO DE COMPOSTOS COM A FÓRMULA



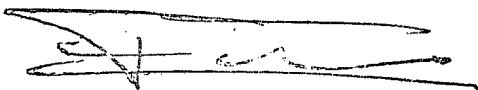
A Figura 3 ilustra a síntese em paralelo de um conjunto de 36 compostos T-L-X (X = L_h), em que L_h é um éster activado (especificamente, o éster de tetrafluorofenilo), L² é um grupo orto-nitrobenzilamina e L³ é um grupo metíleno que liga L_h e L², T possui uma estrutura modular em que o grupo de ácido carboxílico da lisina foi ligado ao átomo de azoto do grupo benzilamina de L² para formar uma ligação amida, e um componente de peso variável R₁₋₃₆ (em que estes grupos R correspondem a T² tal como aqui definido, e podem ser introduzidos por meio de qualquer um dos ácidos carboxílicos aqui listados) se encontra ligado através do grupo α -amino da lisina, existindo também um grupo melhorador da sensibilidade da espectrometria de massa (introduzido através do ácido N-metilisonipecótico) que se encontra ligado através do grupo ε -amino da lisina.

Por referência à Figura 3:

Passo A. A Resina NovaSyn HMP (comercializada pela NovaBiochem; 1 eq.) é suspensa em DMF no vaso de recolha do ACT357. São adicionados o composto I (ANP, comercializado pela ACT; 3 eq.), HATU (3 eq.) e NMM (7,5 eq.) em DMF, e o vaso de recolha é agitado durante 1 h. O solvente é removido e a resina é lavada com NMP (2X), MeOH (2X) e DMF (2X). O acoplamento de I à resina e os passos de lavagem são repetidos para originar o composto II.

Passo B. A resina (composto II) é misturada com piperidina a 25% em DMF e agitada durante 5 min. A resina é filtrada, sendo então misturada com piperidina a 25% em DMF e agitada durante 10 min. O solvente é removido e a resina é lavada com NMP (2X), MeOH (2X) e DMF (2X), sendo então directamente usada no passo C.

Passo C. A resina desprotegida obtida no passo B é suspensa em DMF e é adicionada à suspensão um aminoácido protegido por Fmoc, contendo uma funcionalidade amina na sua cadeia lateral (Fmoc-Lys-(Aloc)-OH, comercializado pela PerSeptive Biosystems; 3 eq.), HATU (3 eq.) e NMM (7,5 eq.) em DMF. O vaso é agitado durante 1 hora. O solvente é removido e a resina é lavada com NMP (2X), MeOH (2X) e DMF (2X). O acoplamento de Fmoc-Lys-(Aloc)-OH à resina e os passos de lavagem são repetidos para originar o composto IV.



Passo D. A resina (composto IV) é lavada com CH_2Cl_2 (2X), sendo então suspensa numa solução de $(\text{PPh}_3)_4\text{Pd}$ (0) (0,3 eq.) e PhSiH_3 (10 eq.) em CH_2Cl_2 . A mistura é agitada durante 1 hora. O solvente é removido e a resina é lavada com CH_2Cl_2 (2X). O passo de paládio é repetido. O solvente é removido e a resina é lavada com CH_2Cl_2 (2X), dietilditiocarbamato de N,N-diisopropiletilamónio em DMF (2X) e DMF (2X) para originar o composto V.

Passo E. A resina desprotegida do passo D é acoplada ao ácido N-metilisonipecótico tal como descrito no passo C para originar o composto VI.

Passo F. A resina VI protegida por Fmoc é dividida em 36 partes iguais pelo ACT357 a partir do vaso de recolha para 36 vasos de reacção, originando os compostos VI_{1-36} .

Passo G. A resina (compostos VI_{1-36}) é tratada com piperidina como descrito no passo B, para remover o grupo FMOC.

Passo H. As 36 alíquotas de resina desprotegida obtida no passo G são suspensas em DMF. A cada vaso de reacção é adicionado o ácido carboxílico apropriado ($\text{R}_{1-36}\text{CO}_2\text{H}$; 3 eq.), HATU (3 eq.) e NMM (7,5 eq.) em DMF. Os vasos de reacção são agitados durante 1 hora. O solvente é removido e as alíquotas de resina são lavadas com NMP (2X), MeOH (2X) e DMF (2X). O acoplamento de $\text{R}_{1-16}\text{CO}_2\text{H}$ às alíquotas de resina e os passos de lavagem são repetidos para originar os compostos VIII_{1-36} .

Passo I. As alíquotas de resina (compostos VIII_{1-36}) são lavadas com CH_2Cl_2 (3X). É adicionado a cada vaso de reacção TFA:H₂O:CH₂Cl₂ a 90:5:5 e os vasos de reacção são agitados durante 120 min. O solvente é filtrado a partir dos vasos de reacção para tubos individuais. As alíquotas de resina são lavadas com CH_2Cl_2 (2X) e MeOH (2X) e os filtrados são combinados nos tubos individuais. Os tubos individuais são evaporados em vácuo, fornecendo os compostos IX_{1-36} .

Passo J. Cada um dos ácidos carboxílicos livres IX_{1-36} é dissolvido em DMF. A cada solução é adicionada piridina (1,05 eq.), seguindo-se a adição de trifluoroacetato de tetrafluorofenilo (1,1 eq.). As misturas são agitadas durante 45 min. à temperatura ambiente. As soluções são diluídas com EtOAc, lavadas com NaHCO_3 em solução aquosa a 5% (3X), secadas em Na_2SO_4 , filtradas e evaporadas em vácuo, fornecendo os compostos X_{1-36} .

EXEMPLO 6

PREPARAÇÃO DE UM CONJUNTO DE COMPOSTOS COM A FÓRMULA



A Figura 4 ilustra a síntese em paralelo de um conjunto de 36 compostos T-L-X ($X = L_h$), em que L_h é um éster activado (especificamente, o éster de tetrafluorofenilo), L^2 é um grupo orto-nitrobenzilamina e L^3 é uma ligação directa entre L_h e L^2 , em que L_h se encontra directamente ligado ao anel aromático do grupo L^2 , T possui uma estrutura modular em que o grupo de ácido carboxílico da lisina foi ligado ao átomo de azoto do grupo benzilamina de L^2 para formar uma ligação amida, e um componente de peso variável R_{1-36} (em que estes grupos R correspondem a T^2 tal como aqui definido, e podem ser introduzidos por meio de qualquer um dos ácidos carboxílicos aqui listados) se encontra ligado através do grupo α -amino da lisina, existindo também um grupo melhorador da sensibilidade da espectrometria de massa (introduzido através do ácido N-metilisonipeptólico) que se encontra ligado através do grupo ε -amino da lisina.

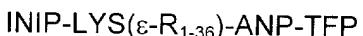
Por referência à Figura 4:

Passo A. A Resina NovaSyn HMP é acoplada ao composto I (NBA preparado de acordo com o procedimento de Brown et al., Molecular Diversity, 1, 4 (1995)), de acordo com o procedimento descrito no passo A do Exemplo 5, para originar o composto II.

Passos B-J. A resina (composto II) é tratada tal como descrito nos passos B-J do Exemplo 5 para originar os compostos X_{1-36} .

EXEMPLO 7

PREPARAÇÃO DE UM CONJUNTO DE COMPOSTOS COM A FÓRMULA



A Figura 5 ilustra a síntese em paralelo de um conjunto de 36 compostos T-L-X ($X = L_h$), em que L_h é um éster activado (especificamente, o éster de tetrafluorofenilo), L^2 é um grupo orto-nitrobenzilamina e L^3 é um grupo metileno que liga L_h a L^2 , T possui uma estrutura modular em que o grupo de ácido carboxílico da lisina foi ligado ao átomo de azoto do grupo benzilamina de L^2 para formar uma ligação amida, e um componente de peso variável R_{1-36} (em que estes grupos R correspondem a T^2 tal como aqui definido, e podem ser introduzidos por meio de qualquer um dos ácidos carboxílicos aqui listados) se encontra



ligado através do grupo ϵ -amino da lisina, existindo também um grupo melhorador da sensibilidade da espectrometria de massa (introduzido através do ácido N-metilisonipeptônico) que se encontra ligado através do grupo α -amino da lisina.

Por referência à Figura 5:

Passos A-C. Tal como no Exemplo 5.

Passo D. A resina (composto IV) é tratada com piperidina, como descrito no passo B do Exemplo 5, para remover o grupo FMOC.

Passo E. A α -amina desprotegida na resina do passo D é acoplada com o ácido N-metilisonipeptônico, tal como descrito no passo C do Exemplo 5, para originar o composto V.

Passo F. Tal como no Exemplo 5.

Passo G. A resina (compostos VI₁₋₃₆) é tratada com paládio, tal como descrito no passo D do Exemplo 5, para remover o grupo Alloc.

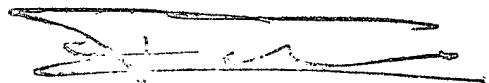
Passos H-J. Os compostos X₁₋₃₆ são preparados do mesmo modo que no Exemplo 5.

EXEMPLO 8

PREPARAÇÃO DE UM CONJUNTO DE COMPOSTOS COM A FÓRMULA



A Figura 6 ilustra a síntese em paralelo de um conjunto de 36 compostos T-L-X (X = L_h), em que L_h é um éster activado (especificamente, o éster de tetrafluorofenilo), L² é um grupo orto-nitrobenzilamina e L³ é um grupo metíleno que liga L_h a L². T possui uma estrutura modular em que o grupo de ácido α -carboxílico do ácido glutâmico foi ligado ao átomo de azoto do grupo benzilamina de L² para formar uma ligação amida, e um componente de peso variável R₁₋₃₆ (em que estes grupos R correspondem a T² tal como aqui definido, e podem ser introduzidos por meio de qualquer um dos ácidos carboxílicos aqui listados) se encontra ligado através do grupo α -amino do ácido glutâmico, existindo também um grupo melhorador da sensibilidade da espectrometria de massa (introduzido através da 2-(diisopropilamino)etilamina) que se encontra ligado através do grupo de ácido γ -carboxílico do ácido glutâmico.



Por referência à Figura 6:

Passos A-B. Tal como no Exemplo 5.

Passo C. A resina desprotegida (composto III) é acoplada a Fmoc-Glu-(OAI)-OH , usando o método de acoplamento descrito no passo C do Exemplo 5, para originar o composto IV.

Passo D. O éster de alilo da resina (composto IV) é lavada com CH₂Cl₂ (2X), sendo então suspensa numa solução de (PPh₃)₄Pd (0) (0,3 eq.) e N-metilanilina (3 eq.) em CH₂Cl₂. A mistura é agitada durante 1 hora. O solvente é removido e a resina é lavada com CH₂Cl₂ (2X). O passo de paládio é repetido. O solvente é removido e a resina é lavada com CH₂Cl₂ (2X), dietilditiocarbamato de N,N-diisopropiletilamónio em DMF (2X) e DMF (2X) para originar o composto V.

Passo E. A resina desprotegida do passo D é suspensa em DMF e activada através da mistura com HATU (3 eq.) e NMM (7,5 eq.). Os vasos são agitados durante 15 minutos. O solvente é removido e a resina é lavada com NMP (1X). A resina é misturada com 2-(diisopropilamino)etilamina (3 eq.) e NMM (7,5 eq.). Os vasos são agitados durante 1 hora. O acoplamento entre a 2-(diisopropilamino)etilamina e a resina e os passos de lavagem são repetidos, para originar o composto VI.

Passos F-J. Tal como no Exemplo 5.

EXEMPLO 9

PREPARAÇÃO DE UM CONJUNTO DE COMPOSTOS COM A FÓRMULA



A Figura 7 ilustra a síntese em paralelo de um conjunto de 36 compostos T-L-X (X = L_h), em que L_h é uma amina (especificamente, o grupo ε-amino de uma porção molecular derivada da lisina), L² é um grupo orto-nitrobenzilamina e L³ é um grupo alquienoaminoacilalquileno carboxamido-substituído que liga L_h a L², T possui uma estrutura modular em que o grupo de ácido carboxílico da lisina foi ligado ao átomo de azoto do grupo benzilamina de L² para formar uma ligação amida, e um componente de peso variável R₁₋₃₆ (em que estes grupos R correspondem a T² tal como aqui definido, e podem ser introduzidos por meio de qualquer um dos ácidos carboxílicos aqui listados) se encontra ligado através do grupo α-amino da lisina, existindo também um grupo melhorador da sensibilidade da

[Handwritten signature]

espectrometria de massa (introduzido através do ácido N-metilisonipeptílico) que se encontra ligado através do grupo ϵ -amino da lisina.

Por referência à Figura 7:

Passo A. A Resina Fmoc-Lys(Boc)-SRAM (comercializada pela ACT; composto I) é misturada com piperidina a 25% em DMF, agitando-se durante 5 minutos. A resina é filtrada e em seguida misturada com piperidina a 25% em DMF, agitando-se durante 10 min. O solvente é removido, a resina é lavada com NMP (2X), MeOH (2X) e DMF (2X), e é usada directamente no passo B.

Passo B. São adicionados a resina (composto II), ANP (comercializado pela ACT; 3 eq.), HATU (3 eq.) e NMM (7,5 eq.) em DMF, e o vaso de recolha é agitado durante 1 h. O solvente é removido e a resina é lavada com NMP (2X), MeOH (2X) e DMF (2X). O acoplamento de I à resina e os passos de lavagem são repetidos para originar o composto III.

Passos C-J. A resina (composto III) é tratada tal como nos passos B-I no Exemplo 5 para formar os compostos X_{1-36} .

EXEMPLO 10

PREPARAÇÃO DE UM CONJUNTO DE COMPOSTOS COM A FÓRMULA



A Figura 8 ilustra a síntese em paralelo de um conjunto de 36 compostos T-L-X ($X = L_h$), em que L_h é um éster activado (especificamente, o éster de tetrafluorofenilo), L^2 é um grupo orto-nitrobenzilamina e L^3 é um grupo metíleno que liga L_h a L^2 , T possui uma estrutura modular em que o grupo de ácido carboxílico de um primeiro resíduo de lisina foi ligado ao átomo de azoto do grupo benzilamina de L^2 para formar uma ligação amida, um grupo melhorador da sensibilidade da espectrometria de massa (introduzido através do ácido N-metilisonipeptílico) está ligado através do grupo ϵ -amino da primeira lisina, uma segunda molécula de lisina está ligada à primeira lisina através do grupo α -amino da primeira lisina, um grupo ajustador do peso molecular (contendo uma estrutura de trifluoroacetilo) está ligado ao grupo ϵ -amino da segunda lisina, e um componente de peso variável R_{1-36} (em que estes grupos R correspondem a T^2 tal como aqui definido, e podem ser introduzidos por meio de qualquer um dos ácidos carboxílicos aqui listados) se encontra ligado através do grupo α -amino da segunda lisina.

Por referência à Figura 8:

Passos A-E. Estes passos são idênticos aos passos A-E do Exemplo 5.

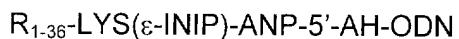
Passo F. A resina (composto VI) é tratada com piperidina, tal como descrito no passo B do Exemplo 5, para remover o grupo FMOC.

Passo G. A resina desprotegida (composto VII) é acoplada a Fmoc-Lys-(Tfa)-OH , usando o método de acoplamento descrito no passo C do Exemplo 5, para originar o composto VIII.

Passos H-K. A resina (composto VIII) é tratada tal como nos passos F-J do Exemplo 5 para originar os compostos XII₁₋₃₆.

EXEMPLO 11

PREPARAÇÃO DE UM CONJUNTO DE COMPOSTOS COM A FÓRMULA



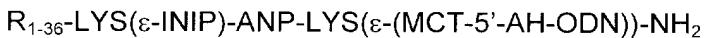
A Figura 9 ilustra a síntese em paralelo de um conjunto de 36 compostos T-L-X (X = MOI, em que MOI é um fragmento de ácido nucleico, ODN) derivados dos ésteres do exemplo 5 (poderá recorrer-se ao mesmo procedimento utilizado com outros compostos T-L-X em que X é um éster activado). A MOI é conjugada com T-L através da extremidade 5' da MOI por meio de um grupo fosfodiéster-alquienoamina.

Por referência à figura 9:

Passo A. São preparados compostos XII₁₋₃₆ de acordo com um procedimento de biotinilação modificado descrito em Van Ness et al., Nucleic Acids Res., 19, 3345 (1991). À solução de um dos oligonucleótidos de 5'-aminohexilo (compostos XI₁₋₃₆, 1 mg) em borato de sódio 200 mM (pH 8,3 , 250 mL) é adicionado um dos ésteres de tetrafluorofenilo (compostos X₁₋₃₆ do exemplo 5, excesso molar de 100 vezes em 250 mL de NMP). A reacção é incubada durante doze horas à temperatura ambiente. Os ésteres de tetrafluorofenilo hidrolisados e os que não reagiram são removidos dos compostos XII₁₋₃₆ por cromatografia com coluna de Sephadex G-50.

EXEMPLO 12

PREPARAÇÃO DE UM CONJUNTO DE COMPOSTOS COM A FÓRMULA



A Figura 10 ilustra a síntese em paralelo de um conjunto de 36 compostos T-L-X (X = MOI, em que MOI é um fragmento de ácido nucleico, ODN) derivados dos ésteres do Exemplo 9 (poderá recorrer-se ao mesmo procedimento utilizado com outros compostos T-L-X em que X é uma amina). A MOI é conjugada com T-L através da extremidade 5' da MOI por meio de um grupo fosfodiéster-alquienoamina.

Por referência à figura 10:

Passo A. Os oligonucleótidos 5'-[6-(4,6-dicloro-1,3,5-triazin-2-ilamino)hexil] XII₁₋₃₆ são preparados como descrito em Van Ness et al., Nucleic Acids Res., 19, 3345 (1991).

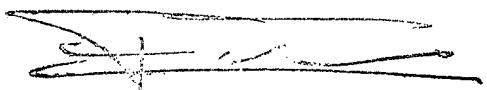
Passo B. À solução de um dos oligonucleótidos 5'-[6-(4,6-dicloro-1,3,5-triazin-2-ilamino)hexil] (compostos XII₁₋₃₆) a uma concentração de 1 mg/mL em borato de sódio 100 mM (pH 8,3) foi adicionado um excesso molar de 100 vezes de uma amina primária seleccionada a partir dos compostos R₁₋₃₆-Lys(ε-iNIP)-ANP-Lys(ε-NH₂)-NH₂ (compostos X₁₋₃₆ do Exemplo 11). A solução é misturada durante doze horas à temperatura ambiente. A amina que não reagiu é removida por ultrafiltração através de uma membrana com cutoff de PM 3000 (Amicon, Beverly, MA) usando H₂O como solução de lavagem (3X). Os compostos XIII₁₋₃₆ são isolados por redução do volume até 100 mL.

EXEMPLO 13

DEMONSTRAÇÃO DA DETECÇÃO SIMULTÂNEA DE MARCADORES MÚLTIPLOS POR ESPECTROMETRIA DE MASSA

Este Exemplo proporciona uma descrição da capacidade de detectar em simultâneo múltiplos compostos (marcadores) por espectrometria de massa. Neste Exemplo em particular, 31 compostos são misturados com uma matriz, depositados e secos sobre um suporte sólido e então desadsorvidos com um laser. Os iões resultantes são então introduzidos num espectrómetro de massa.

Os compostos que se seguem (comercializados pela Aldrich, Milwaukee, WI) são misturados numa base equimolar a uma concentração final de 0,002 M (por composto): benzamida (121,14), nicotinamida (122,13), pirazinamida (123,12), ácido 3-amino-4-pirazolcarboxílico (127,10), 2-tiofenocarboxamida (127,17), 4-aminobenzamida (135,15), toluimida (135,17), 6-metilnicotinamida (136,15), 3-aminonicotinamida (137,14), N-óxido de

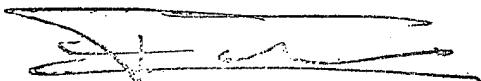


nicotinamida (138,12), 3-hidropicolinamida (138,13), 4-fluorobenzamida (139,13), cinamamida (147,18), 4-metoxibenzamida (151,17), 2,6-difluorobenzamida (157,12), 4-amino-5-imidazol-carboxiamida (162,58), 3,4-piridina-dicarboxiamida (165,16), 4-etoxibenzamida (165,19), 2,3-pirazinadicarboxamida (166,14), 2-nitrobenzamida (166,14), ácido 3-fluoro-4-metoxibenzóico (170,4), indol-3-acetamida (174,2), 5-acetilsalicilamida (179,18), 3,5-dimetoxibenzamida (181,19), 1-naftalenoacetamida (185,23), 8-cloro-3,5-diamino-2-pirazinacarboxiamida (187,59), 4-trifluorometilo-benzamida (189,00), 5-amino-5-fenil-4-pirazol-carboxamida (202,22), 1-metil-2-benzilmalonamato (207,33), 4-amino-2,3,5,6-tetrafluorobenzamida (208,11), ácido 2,3-naftalenodicarboxílico (212,22). Os compostos são colocados em DMSO à concentração acima descrita. Um μ L do material é então misturado com uma matriz alfa-ciano de ácido 4-hidroxicinâmico (após uma diluição de 1:10.000) e depositados sobre um suporte sólido de aço inoxidável.

O material é então desdsorvido com um laser utilizando um Espectrómetro de Massa TOF para Proteínas (Brucker, Manning Park, MA) e os iões resultantes são medidos tanto no modo de funcionamento linear como no modo de reflectrão. São observados os seguintes valores para a razão m/z (Figura 11):

- 121,1 ----> benzamida (121,14)
- 122,1 ----> nicotinamida (122,13)
- 123,1 ----> pirazinamida (123,12)
- 124,1
- 125,2
- 127,3 ----> ácido 3-amino-4-pirazolcarboxílico (127,10)
- 127,2 ----> 2-tiofenocarboxamida (127,17)
- 135,1 ----> 4-aminobenzamida (135,15)
- 135,1 ----> toluimida (135,17)
- 136,2 ----> 6-metilnicotinamida (136,15)
- 137,1 ----> 3-aminonicotinamida (137,14)
- 138,2 ----> N-óxido de nicotinamida (138,12)
- 138,2 ----> 3-hidropicolinamida (138,13)
- 139,2 ----> 4-fluorobenzamida (139,13)
- 140,2
- 147,3 ----> cinamamida (147,18)
- 148,2
- 149,2
- 4-metoxibenzamida (151,17)

- ~~S. L. C. 2010~~
- 152,2
2,6-difluorobenzamida (157,12)
- 158,3
4-amino-5-imidazol-carboxiamida (162,58)
- 163,3
- 165,2 ----> 3,4-piridina-dicarboxiamida (165,16)
- 165,2 ----> 4-etoxibenzamida (165,19)
- 166,2 ----> 2,3-pirazinodicarboxamida (166,14)
- 166,2 ----> 2-nitrobenzamida (166,14)
ácido 3-fluoro-4-metoxibenzóico (170,4)
- 171,1
- 172,2
- 173,4
indol-3-acetamida (174,2)
- 178,3
- 179,3 ----> 5-acetilsalicilamida (179,18)
- 181,2 ----> 3,5-dimetoxibenzamida (181,19)
- 182,2 ---->
1-naftalenoacetamida (185,23)
- 186,2
8-cloro-3,5-diamino-2-pirazinocarboxiamida (187,59)
- 188,2
- 189,2 ----> 4-trifluorometil-benzamida (189,00)
- 190,2
- 191,2
- 192,3
5-amino-5-fenil-4-pirazol-carboxamida (202,22)
- 203,2
- 203,4
1-metil-2-benzil-malonamato (207,33)
4-amino-2,3,5,6-tetrafluorobenzamida (208,11)
- 212,2 ----> ácido 2,3-naftalenodicarboxílico (212,22)
- 219,3
- 221,2
- 228,2



234,2
237,4
241,4

Os dados indicam que 22 dos 31 compostos apareceram no espectro com a massa prevista, enquanto 9 dos 31 compostos apareceram no espectro com uma massa de $n + H$ (1 unidade de massa atómica, amu) acima do valor de massa previsto. Este fenómeno é provavelmente devido à protonação de uma amina existente nos compostos. Assim sendo, 31 de 31 compostos são detectados por espectroscopia de massa MALDI. Mais importante que isto, o exemplo demonstra que múltiplos marcadores podem ser detectados em simultâneo por um método espectroscópico.

A matriz alfa-ciano por si só (Figura 11) exibe picos a 146,2, 164,1, 172,1, 173,1, 189,1, 190,1, 191,1, 192,1, 212,1, 224,1, 228,0 e 234,3. Outras massas identificadas no espectro são devidas a contaminantes dos compostos comercializados, já que nenhum esforço foi feito para purificar mais completamente os compostos.

EXEMPLO 14

ENSAIO DA EXPRESSÃO GENÉTICA UTILIZANDO SONDAS MÚLTIPLAS

Foram extemporaneamente preparados tampões de borato de sódio (SBB) a partir de ácido bórico e de hidróxido de sódio. O tampão APB é constituído por NaCl 0,18 M, Tris 0,05 M, pH 7,6, EDTA 5 mM e Tween 20R a 0,5 %. O tampão TMNZ é constituído por Tris 0,05 M, pH 9,5, MgCl₂ 1 mM, ZnCl₂ 0,5 mM. FW (lavado do filtro) é formado por NaCl 0,09 M, Tris 50 mM, pH 7,6, EDTA 25 mM. SDS/FW corresponde a FW com dodecilsulfato de sódio (SDS) a 0,1 %. A solução de lise e hibridação é constituída por tiocianato de guanidínio 3 M, N-lauroilsarcosina (sarcosil) a 2 %, Tris 50 mM, pH 7,6, e EDTA 25 mM. O tampão CAP é formado por citrato de sódio 0,1 M e fosfato de sódio 0,2 M, pH 6,5. A solução de substrato de HRP (peroxidase de rábano) corresponde a citrato de sódio 0,1 M, pH 6,5, fosfato de sódio 0,2 M, 4-metoxi-1-naftol 2,87 mM, cloridrato de hidrazona 3-metil-2-benzotiazolinona 0,093 mM e peróxido de hidrogénio 4 mM. A solução de substrato de AP (fosfatase alcalina) é constituída por 5-bromo-4-clorindoil-3-fosfato 1 mM, nitroazul de tetrazólio 1 mM e Tween 20 a 0,01% em TMNZ. O substrato fluorescente para a fosfatase alcalina é constituído por fosfato de 4-metil-umbelifera 0,5 mM, tris 0,05 M, pH 9,5, MgCl₂ 1 mM, ZnCl₂ 0,5 mM. A poli(etileneimina) foi adquirida à empresa Polysciences (Warrington,

PA). As esferas de nylon polidas ou não polidas foram adquiridas à empresa The Hoover Group (Sault St. Marie, MI). O tetrafluoroborato de trietiloxónio, o anidrido succínico e a 1-metil-2-pirrolidinona foram adquiridas à Aldrich Chemical (Milwaukee, WI). O Tween 20R e a NHS-LC-Biotina foram adquiridas à empresa Pierce (Rockford, IL). O tiocianato de guanidina (GuSCN) foi adquirido à Kodak (Rochester, NY). O cloreto cianúrico foi adquirido à Aldrich Chemical Co. (Milwaukee, WI) e foi recristalizado a partir do tolueno.

A. Síntese de ODNs

Foram sintetizados ODNs complementares (5'CCTTAGGACAGTCTTCTCACGC) de regiões conservadas ou hipervariáveis do RNA ribossomal 16S (rRNA) da *Porphyromonas gingivalis* (Pg), por meio de um sintetizador de DNA automatizado ABI380B ou Milligen 7500 utilizando a química convencional do cianoetyl-N,N-diisopropilamino-fosforamidito (CED-fosforamidito). Foram incorporadas caudas de amina na extremidade 5' usando o N-monometoxitritilaminoihex-6-iloxi-CED-fosforamidito. Os ODNs com grupos 5'-monometoxitritilo foram submetidos a cromatografia por HPLC utilizando uma coluna de fase reversa (7,0 x 305 mm) Hamilton PRP-1 aplicando um gradiente de 5% a 45% de CH₃CN em Et₃NH⁺OAc⁻ 0,1 M, pH 7,5, durante 20 minutos. Após detritilação com ácido acético a 80%, os ODNs foram precipitados por adição de acetato de sódio 3 M e 1-butanol. As verificações analíticas da qualidade dos ODNs foram efectuadas por HPLC de troca iônica usando uma coluna Toso-Haas DEAE-NPR e por electroforese em gel desnaturante de poliacrilamida (PAGE).

B. Preparação da esfera de nylon revestida a polímero

Foram agitadas durante 5 min. e à temperatura ambiente esferas de nylon não polidas (25.000, 3/32 polegadas de diâmetro) em 1-metil-2-pirrolidinona anidra (1800 mL). Adicionou-se tetrafluoroborato de trietiloxónio (200 mL, 1 M em diclorometano), agitando-se durante 30 min. à temperatura ambiente). O líquido foi decantado e as esferas foram rapidamente lavadas com 1-metil-2-pirrolidinona (4 X 500 mL). As esferas foram então agitadas durante 12-24 h em 1 L de uma solução de poli(etileneimina) de PM 70.000 a 3% (p/v) em 1-metil-2-pirrolidinona (preparada a partir de uma solução aquosa de poli(etileneimina) a 30%). Após a decantação da solução de poli(etileneimina), as esferas foram lavadas com 1-metil-2-pirrolidinona (2 x 1 L, SDS/FW (2 x 1 L), H₂O (10 x 2 L), e finalmente com etanol a 95% (1 x 500 mL). As esferas foram secas em vácuo elevado durante 4 a 5 h. O conteúdo em aminas das esferas foi determinado através de reacção com o ácido picrilsulfónico.

C. Preparação de ODNs de 5'-[6-(4,6-dicloro-1,3,5-triazin-2-ilamino)-hexilo]

A uma solução de ODNs de 5'-aminohexilo (1 mL, 10 mg/mL) em SBB 0,1 M preparado extemporaneamente (pH 8,3, 3,2 mL) e H₂O (1,8 mL) foi adicionada uma solução de cloreto cianúrico recristalizado em acetonitrilo (1 mL, 50 mg/mL). A solução foi misturada durante 30-120 minutos à temperatura ambiente. O cloreto cianúrico que não reagiu foi removido por ultrafiltração através duma membrana com um cutoff de PM 3.000 (Amicon, Beverly, MA) usando SBB de preparação extemporânea a 0,1 M (pH 9,3, 4x10 mL) como solução de lavagem. Após a lavagem final, o volume foi reduzido para 1 mL. Os ODNs de 5'-[6-(4,6-dicloro-1,3,5-triazin-2-ilamino)-hexilo] são estáveis durante uma semana a 4°C em SSB 0,1 M (pH 8,3) sem que ocorra uma decomposição detectável.

D. Ligação dos ODNs às esferas de nylon

As esferas de nylon revestidas por PEI (500 esferas), acima descritas, foram colocadas num igual volume de SBB preparado extemporaneamente a 0,1 M (pH 9,3) e agitadas vigorosamente durante 30 minutos para re-hidratar as esferas. A solução de borato foi decantada e as esferas foram lavadas uma vez com SBB 0,1 M (pH 8,3), e então tratadas com igual volume de SBB 0,1 M fresco. A solução de ODNs 5'-[6-(4,6-dicloro-1,3,5-triazin-2-ilamino)-hexilo] em borato (1 mL, 500 mg/mL) foi então adicionada às esferas. A mistura foi vigorosamente agitada à temperatura ambiente durante 60 minutos. A solução foi decantada e as esferas foram então lavadas com SSB 0,1 M (pH 8,3, 2x500 mL). As esferas foram tratadas com anidrido succínico (10 mg/mL) em 9:1 de 1-metil-2-pirrolidinona: SBB 1,0 M (pH 8,3), utilizando-se um volume equivalente a três vezes o volume das esferas. A mistura de reacção foi agitada durante 1 h à temperatura ambiente. As esferas foram então lavadas com um 1-metil-2-pirrolidinona (3x250 mL), dH₂O (2x1 L), SDS/FW (5x250 mL) e então com dH₂O (4x1 L). As esferas foram guardadas em EDTA 25 mM.

E. Concepção e marcação das sondas

Nesta parte do Exemplo 5 são concebidas sondas que permitirão a expressão diferencial de mRNA de células estimuladas versus células não estimuladas de linfoma humano de células T Jurkat (JRT 3,5).

Fazem-se reagir 100 µg de cada um dos oligonucleótidos ligados a amina pela extremidade 5' (acima descritos) com um excesso de cloreto cianúrico recristalizado em n-metil-pirrolidona a 10 % em tampão alcalino (de preferência a pH 8,3 - 8,5) a 19°C – 25°C durante 30 a 120 minutos. As condições de final de reacção consistem em borato de sódio 0,15 M a pH 8,3, cloreto cianúrico recristalizado a 2 mg/ml e respectivo oligonucleótilo a 500

μg/ml. O cloreto cianúrixo que não reagiu por cromatografia de exclusão por dimensões numa coluna de Sephadex G-50. O oligonucleótido activado purificado é então posto em reacção com um excesso molar de 100 vezes de cistamina em borato de sódio 0,15 M a pH 8,3, durante 1 hora, à temperatura ambiente. A cistamina que não reagiu é removida por cromatografia de exclusão por dimensões numa coluna de Sephadex G-50. Os ODNs derivatizados são então postos em reacção com fluorocromos reactivos a aminas. A preparação derivatizada de ODNs é então dividida em 3 porções e cada porção éposta em reacção com (a) um excesso molar de 20 vezes de cloreto de sulfônico de Texas Red (Molecular Probes, Eugene, OR), com (b) um excesso molar de 20 vezes de cloreto de sulfônico de Lissamina (Molecular Probes, Eugene, OR) ou (c) um excesso molar de 20 vezes de isotiocianato de fluoresceína. As condições de final de reacção consistem em borato de sódio 0,15 M, pH 8,3, durante 1 hora à temperatura ambiente. Os fluorocromos que não reagiram são removidos por cromatografia de exclusão por dimensões numa coluna de Sephadex G-50. Foram marcadas moléculas de IL-2, IFN-γ, e GM-CSF com Texas Red. As moléculas de IL-4 c-fos e de PKC-γ foram marcadas com lissamina e as moléculas de CTLA4/CD28 e de GMP quinase foram marcadas com fluoresceína. As sondas de IL-2, c-fos e CTLA4 foram reunidas num pool. As sondas de IFN-γ, IL-4 e GMP quinase foram reunidas num pool, e as sondas de GM-CSF e PKC-γ foram igualmente reunidas num pool.

F. Síntese de cDNA em suporte sólido para ensaio da expressão genética

Oligo DMO 596 5'-ACTACTGATCAGGCGCGCTTTTTTTTTTTTTTTT -3'

Espaçador Asc I (poli dT)20

G. Estimulação e preparação de RNA

A linha celular Jurkat JRT 3.5 é estimulada durante 6 horas, a uma densidade celular de 1×10^6 células/ml, em meio RPMI isento de soro (Life Technologies, Gaithersburg, MD) na presença de forbol-12-miristato-13-acetato a 10 ng/ml (Calbiochem, San Diego, CA) e ionomicina a 100 ng/ml (Calbiochem). É preparado um pellet de células que é lavado em 1x PBS (Life Technologies), fazendo-se então nova preparação de um pellet que é em seguida lisado em tampão, a 0,5 ml por 10^{-6} células, contendo isotiocianato de guanidina 4 M/N-laurilsarcosina a 1%/citrato de sódio 25 mM, pH 7,1 (Fisher Scientific, Pittsburg, PA). É adicionado um décimo de volume de acetato de sódio 2 M a pH 4,2 (Fisher Scientific), seguido da adição de um volume de fenol saturado com água (Amresco, Solon, OH). Após a mistura é adicionado um quarto de volume de clorofórmio:álcool isoamílico (29:1) (Fisher Scientific) e a solução é vigorosamente misturada, sendo então incubada em gelo durante

10 minutos. O lisado é então recolhido, a fase aquosa é removida e extraída com igual volume de clorofórmio:álcool isoamílico. É feito então um pool da fase aquosa e o RNA é precipitado com 2 volumes de EtOH (Quantum Chemical Corp., Tuscola, IL). Após centrifugação, o EtOH é decantado e o RNA é brevemente seco ao ar , sendo ressuspenso em seguida em água isenta de RNase para obter uma concentração situada entre 1 e 5 mg/ml.

H. Captura e síntese da primeira cadeia

Uma esfera de nylon ligada por ligação covalente ao oligonucleótido 5' – ACTACTGATCAGGCGCGCTTTTTTTTTTTTTTTTT –3' (GenSet, La Jolla, CA) é adicionada a 10 µg de RNA celular total, diluído numa quantidade de água isenta de RNase suficiente para cobrir a esfera, num tubo esterilizado de microcentrífuga de 1,5 ml (Fisher Scientific). O RNA e a esfera são incubados a 65°C durante 5 minutos. É adicionado um igual volume de tampão de hibridização de 2X RNA consistindo em Tris 50 mM, pH 7,5, NaCl 1 M (Fisher Scientific) e BSA acetilado a 20 µg/ml (New England Biolabs, Beverly, MA) a cada tubo, sendo os tubos agitados suavemente durante 2 horas à temperatura ambiente. O sobrenadante é removido e a esfera é então lavada 3 vezes em tampão de hibridização de 1X mRNA. Após completada a última lavagem, é adicionada a cada tubo uma mistura de transcrição reversa constituída por 1 X tampão de MMLV-transcriptase reversa, uma mistura de dNTPs a 1 mM, DTT 2 mM (Life Technologies), 20 unidades de RNasina (Promega, Madison, WI) e BSA acetilado a 10 µg/ml (New England Biolabs), seguindo-se a adição de 600 unidades de MMLV-transcriptase reversa (Life Technologies). Esta reacção é agitada suavemente a 42°C durante 2 horas. É então adicionada uma unidade de RNase H (Boehringer Manheim, Indianapolis, IN) e a reacção prossegue durante mais 0,5 h. O sobrenadante é novamente removido e cada esfera é lavada três vezes em Tris 10 mM, pH 8,0, EDTA 1 mM, pH 8 (Fisher Scientific). O molde remanescente de RNA é removido ao levar as esferas a ebuição em TE com SDS a 0,01% (Fisher Scientific).

O suporte sólido de nylon foi então hibridizado com 100 nanogramas por ml das seguintes sondas oligonucleotídicas marcadas:

- (5'-GAACTCAAACCTCTGGAGGAAGTG-3', IL-2,
- 5'-CAGTGCAGAGGCTCGCGAGCTATA-3', IFN-gama
- 5'-CTTGACCATGATGGCCAGCCACTA-3', GM-CSF
- 5'-CATTCCCACGGTCACTGCCATCTC-3', c-fos
- 5'-GCGACTGTGCTCCGGCAGTTCTAC-3', IL-4

5'-GTGGTTCATCGACGATGCCACGAA-3', PKC-gama
5'-GAGCTCATGTACCCACCTCCGTAC-3', CTLA4/CD28
5'-ATCTTCGTGCAGCCGCCCTCACTG-3', GMP quinase)

(Todos os oligos correspondem aos homólogos humanos com excepção da GMP quinase, que se baseou na sequência bovina). A hibridização foi efectuada em GuSCN 3 M durante 8 horas a 37°C. A mistura de reacção foi misturada suavemente durante a hibridização para promover a difusão da sonda para o suporte sólido. Após o período de incubação de 1 hora, o suporte sólido foi lavado duas vezes com GuSCN 3 M, 5 vezes com 0,1x SSC e foi então colocado em ditiotreitol 0,01 M para cortar o fluorocromo a partir do oligonucleótido. A mistura é incubada durante 15 minutos à temperatura ambiente. A fluorescência é medida numa placa negra de microtitulação (Dynatek Laboratories, Chantilly, VA). As placas são então lidas directamente por meio de um fluorómetro Fluoroskan II (Flow Laboratories, McLean, VA) utilizando um comprimento de onda de excitação de 495 nm e monitorizando a emissão a 520 nm para a fluoresceína, utilizando um comprimento de onda de excitação de 591 nm e monitorizando a emissão a 612 nm para o Texas Red, e utilizando um comprimento de onda de excitação de 570 nm e monitorizando a emissão a 590 nm para a lissamina. Os resultados da sondagem são os seguintes:

	Não estimulados	Estimulados
IL-2	1,2 rfu	230 rfu
IFN	0,8 rfu	120 rfu
GM-CSF	21 rfu	38 rfu
c-fos	16 rfu	76 rfu
IL-4	33 rfu	12 rfu
PKC	10 rfu	130 rfu
CTLA-4	ND	ND
GMP quinase	450 rfu	420 rfu

EXEMPLO 15

DETECÇÃO DE UM EMPARELHAMENTO INCORRECTO DE UM SÓ PAR DE BASES NUMA FASE SÓLIDA

Este Exemplo descreve a detecção de um emparelhamento incorrecto de um só par de bases numa sonda imobilizada usando oligonucleótidos complementares marcados por

fluorescência. O conjunto de oligonucleótidos de sonda consiste numa sonda que forma um emparelhamento perfeito de pares de bases e num oligonucleótido que apresenta o emparelhamento incorrecto quando é hibridizado. Os dois oligonucleótidos são marcados com diferentes fluorocromos, e após se permitir a ocorrência da hibridização à T_m do emparelhamento incorrecto é determinada a razão de fluorocromos hibridizados.

Um oligonucleótido "alvo" (DMO501: 5'-TTGATTCCCAATTATGCGAAGGAG-3') foi immobilizado num conjunto de suportes sólidos. As esferas de ODN (3/32 polegadas de diâmetro) foram preparadas como anteriormente descrito (Van Ness et al., *Nucl. Acids Res.* 19:3345, 1991). As esferas de ODN contiveram entre 0,01 e 1,2 mg/esfera de ODN immobilizado por via covalente. DMO578 constitui o complemento de DMO501 (complementaridade perfeita). DMO1969 constitui o complemento de DMO501 com uma alteração de G→T a nível da posição 11. DMO1971 constitui o complemento de DMO501 com uma alteração de A→T a nível da posição 12. Cada sonda de oligonucleótido foi marcada com BIODIPY, TAMRA ou Texas Red. As reacções de hibridização foram efectuadas em GuSCN 3 M, Tris 0,01 M, pH 7,6, EDTA 5 mM a 50 ng/ml com cada sonda respectiva. Foram usadas razões molares iguais de cada tipo de sonda em cada hibridização, na presença de 3 suportes sólidos por tubo. As hibridizações são realizadas a 42°C durante 30 minutos com agitação constante. As esferas foram lavadas duas vezes com GuSCN 3 M e em seguida com SDS/FW por cinco vezes.

Para desnaturar o oligonucleótido da sonda, os suportes sólidos são colocados em 200 µl de TE (TE corresponde a Tris 0,01 M, pH 7,0, EDTA 5 mM). A mistura é incubada durante 10 minutos à 100°C. A fluorescência é medida numa placa negra de microtitulação. A solução é removida dos tubos de incubação (200 microlitros) e colocada numa placa negra de microtitulação (Dynatek Laboratories, Chantilly, VA). As placas são então lidas directamente por meio de um fluorómetro Fluoroskan II (Flow Laboratories, McLean, VA) utilizando um comprimento de onda de excitação de 495 nm e monitorizando a emissão a 520 nm para a fluoresceína, utilizando um comprimento de onda de excitação de 591 nm e monitorizando a emissão a 612 nm para o Texas Red, e utilizando um comprimento de onda de excitação de 570 nm e monitorizando a emissão a 590 nm para a lissamina ou o TAMRA.

Os resultados obtidos são os apresentados a seguir:

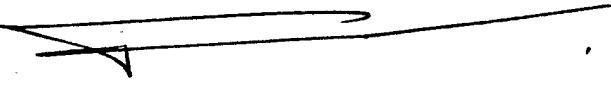
Tabela 10

Mistura de Sondas	Razão entre fluorocromos na mistura de hibridização	Razão entre fluorocromos após a desnaturação
578TR/578BD	1,9/1	1,9/1

578TR/1969BD	2,0/1	25/1
578TR/1971TA	0,025/1	0,58/1
578BD/1971TA	0,014/1	0,48/1

Os resultados indicam que o fluorocromo não exerce qualquer efeito sobre a hibridização; assim, tal como indica a linha 1, o oligonucleótido 578 marcado com Texas Red (TR) e o 578-BD (BIODIPY) competiram de modo uniforme para a hibridização com o alvo imobilizado, uma vez que a razão entre marcadores não se alterou após a hibridização. O enriquecimento médio do GuSCN de cerca de 20 vezes em sondas com emparelhamento perfeito entre bases relativamente às sondas de emparelhamento imperfeito permite a detecção certa dos emparelhamentos incorrectos entre pares de bases.

Lisboa, - 3 MAR, 2000
 O AGENTE OFICIAL DA PROPRIEDADE INDUSTRIAL





REIVINDICAÇÕES

- 1- Um método para a detecção da ligação de um primeiro membro a um segundo membro de um par de ligandos, caracterizado pelo facto de compreender os passos de:
 - (a) combinar um conjunto de primeiros membros marcados com uma amostra biológica que poderá conter um ou mais segundos membros, sob condições e durante o tempo suficiente para permitir a ligação de um primeiro membro a um segundo membro, sendo que o dito marcador se correlaciona com um primeiro membro em particular e é detectável por espectrometria ou potenciometria;
 - (b) separar os primeiro e segundo membros ligados dos membros não ligados;
 - (c) separar o dito marcador do dito primeiro membro marcado; e
 - (d) detectar o dito marcador por espectrometria ou potenciometria, e deste modo detectar a ligação do dito primeiro membro ao dito segundo membro.
- 2- Um método, conforme reivindicado pela reivindicação 1, caracterizado pelo facto de os ditos primeiros membros se encontrarem ligados a um suporte sólido.
- 3- Um método, conforme reivindicado pela reivindicação 2, caracterizado pelo facto de compreender, subsequentemente ao passo de separação dos primeiro e segundo membros ligados, a remoção por lavagem dos membros não ligados a partir do dito suporte sólido.
- 4- Um método, conforme reivindicado nas reivindicações 1-3, caracterizado pelo facto de a detecção do marcador é feita por espectrometria de massa, espectrometria de infravermelhos, espectrometria de ultravioleta ou amperometria potenciotáctica.
- 5- Um método, conforme reivindicado nas reivindicações 1-4, caracterizado pelo facto de ser combinado um número superior a 4 primeiros membros marcados e sendo cada marcador exclusivo de um fragmento de ácido nucleico seleccionado.
- 6- Um método, conforme reivindicado por qualquer das reivindicações 1 e 4-5, caracterizado pelo facto de os ditos primeiro e segundo membros ligados serem separados dos membros não ligados, através de um método seleccionado a partir do grupo que consiste em electroforese em gel, electroforese capilar, electroforese em micro-canais, HPLC, cromatografia de exclusão por dimensões e filtração.
- 7- Um método, conforme reivindicado pelas reivindicações 1-6, caracterizado pelo facto de os ditos primeiros membros marcados são cortados por um método seleccionado a partir do grupo que inclui os métodos de oxidação, redução, labilidade a ácidos, labilidade a bases, enzimático, electroquímico, térmico e de fotolabilidade.
- 8- Um método, conforme reivindicado pelas Reivindicações 1-7, caracterizado pelo facto de o marcador ser detectado por espectrometria de massa de tempo-de-voo,



espectrometria de massa quadrupolar, espectrometria de massa de sector magnético e espectrometria de massa de sector eléctrico.

9- Um método, conforme reivindicado nas reivindicações 1-7, caracterizado pelo facto de o dito marcador ser detectado por amperometria potenciostática utilizando detectores seleccionados a partir do grupo formado por detectores coulométricos e detectores amperométricos.

10- Um método, conforme reivindicado nas reivindicações 1-9, caracterizado pelo facto de os passos b, c e d serem efectuados em modo contínuo.

11- Um método, conforme reivindicado nas reivindicações 1-10, caracterizado pelo facto de os passos b, c e d serem efectuados em modo contínuo num só dispositivo.

12- Um método, conforme reivindicado nas reivindicações 1-11, caracterizado pelo facto de os passos b, c e d serem automatizados.

13- Um método, conforme reivindicado nas reivindicações 1-12, caracterizado pelo facto de o dito primeiro membro ser uma molécula de ácido nucleico.

14- Um método, conforme reivindicado nas reivindicações 1-13, caracterizado pelo facto de o dito segundo membro ser uma molécula de ácido nucleico.

15- Um método, conforme reivindicado nas reivindicações 1-14, caracterizado pelo facto de os ditos primeiro e segundo membros serem moléculas de ácido nucleico e diferindo em, pelo menos, uma não correspondência de um par de bases.

16- Um método, conforme reivindicado nas reivindicações 13-15, caracterizado pelo facto de a dita molécula de ácido nucleico ser gerada por extensão de primer.

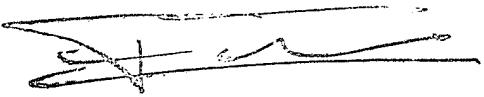
17- Um método, conforme reivindicado nas reivindicações 13-15, caracterizado pelo facto de a dita molécula de ácido nucleico ser gerada a partir de primers de oligonucleótidos não marcados em 3'.

18- Um método, conforme reivindicado nas reivindicações 13-15, caracterizado pelo facto de a dita molécula de ácido nucleico é gerada a partir de terminadores de didesoxinucleótidos marcados.

19- Um método, conforme reivindicado nas reivindicações 1-12, caracterizado pelo facto de o dito primeiro membro é uma proteína, hormona ou uma molécula orgânica.

20- Um método, conforme reivindicado na reivindicação 19, caracterizado pelo facto de a dita proteína ser seleccionada a partir do grupo que consiste em anticorpos e receptores.

21- Um método, conforme reivindicado nas reivindicações 1-20, caracterizado pelo facto de o dito marcador ser detectado por espectrometria não-fluorescente ou potenciometria.



22- Um método para a análise do padrão de expressão genética a partir de uma amostra biológica seleccionada, caracterizado pelo facto de compreender os passos de:

- (a) expôr os ácidos nucleicos de uma amostra biológica;
- (b) combinar os ditos ácidos nucleicos expostos com uma ou mais sondas marcadas de ácidos nucleicos seleccionadas, sob condições e durante o tempo suficiente para que as ditas sondas se hibridizem com os ditos ácidos nucleicos, sendo que o dito marcador se correlaciona com uma sonda de ácidos nucleicos em particular e é detectável por espectrometria ou potenciometria;
- (c) separar as sondas hibridizadas das não hibridizadas;
- (d) separar o dito marcador do dito fragmento marcado; e
- (e) detectar o dito marcador por espectrometria ou potenciometria, e desse modo determinar o padrão de expressão genética da dita amostra biológica.

23- Um método, conforme reivindicado na reivindicação 22, caracterizado pelo facto de a dita amostra biológica ser seleccionada a partir do grupo que consiste em células de mamíferos, bactérias e leveduras.

24- Um método, conforme reivindicado na reivindicação 23, caracterizado pelo facto de as ditas células de mamíferos conterem vírus.

25- Um método, conforme reivindicado nas reivindicações 22-24, caracterizado pelo facto de os ditos ácidos nucleicos expostos se encontrarem ligados a um suporte sólido.

26- Um método, conforme reivindicado na reivindicação 25, caracterizado pelo facto de o dito suporte sólido ser um polímero.

27- Um método, conforme reivindicado na reivindicação 25, caracterizado pelo facto de compreender, subsequentemente ao passo de separação, a lavagem do suporte sólido.

28- Um método, conforme reivindicado nas reivindicações 22-24, caracterizado pelo facto de as ditas sondas hibridizadas serem separadas das sondas não hibridizadas através de um método seleccionado a partir do grupo que consiste em electroforese em gel, electroforese capilar, electroforese em micro-canais, HPLC, filtração e electroforese em gel de poliacrilamida.

29- Um método, conforme reivindicado nas reivindicações 22-28, caracterizado pelo facto de as ditas sondas marcadas serem cortadas por um método seleccionado a partir do grupo que inclui os métodos de oxidação, redução, labilidade a ácidos, labilidade a bases, enzimático, electroquímico, térmico e de fotolabilidade.

- 30- Um método, conforme reivindicado nas reivindicações 22-29, caracterizado pelo facto de o dito marcador ser detectado por um método seleccionado a partir do grupo que consiste em espectrometria de massa de tempo-de-voo, espectrometria de massa quadrupolar, espectrometria de massa de sector magnético e espectrometria de massa de sector eléctrico.
- 31- Um método, conforme reivindicado nas reivindicações 22-28, caracterizado pelo facto de o dito marcador ser detectado por amperometria potenciostática utilizando detectores seleccionados a partir do grupo formado por detectores coulométricos e detectores amperométricos.
- 32- Um método, conforme reivindicado nas reivindicações 22-31, caracterizado pelo facto de os passos c, d e e serem efectuados em modo contínuo.
- 33- Um método, conforme reivindicado nas reivindicações 22-32, caracterizado pelo facto de os passos c, d e e serem efectuados em modo contínuo num só dispositivo.
- 34- Um método, conforme reivindicado na reivindicação 33, caracterizado pelo facto de o dito dispositivo ser automatizado.
- 35- Um método, conforme reivindicado nas reivindicações 22-29 e 32-34, caracterizado pelo facto de o dito marcador ser detectado por espectrometria não-fluorescente ou potenciometria.

Lisboa,

- 3 MAR. 2000

O AGENTE OFICIAL DA PROPRIEDADE INDUSTRIAL

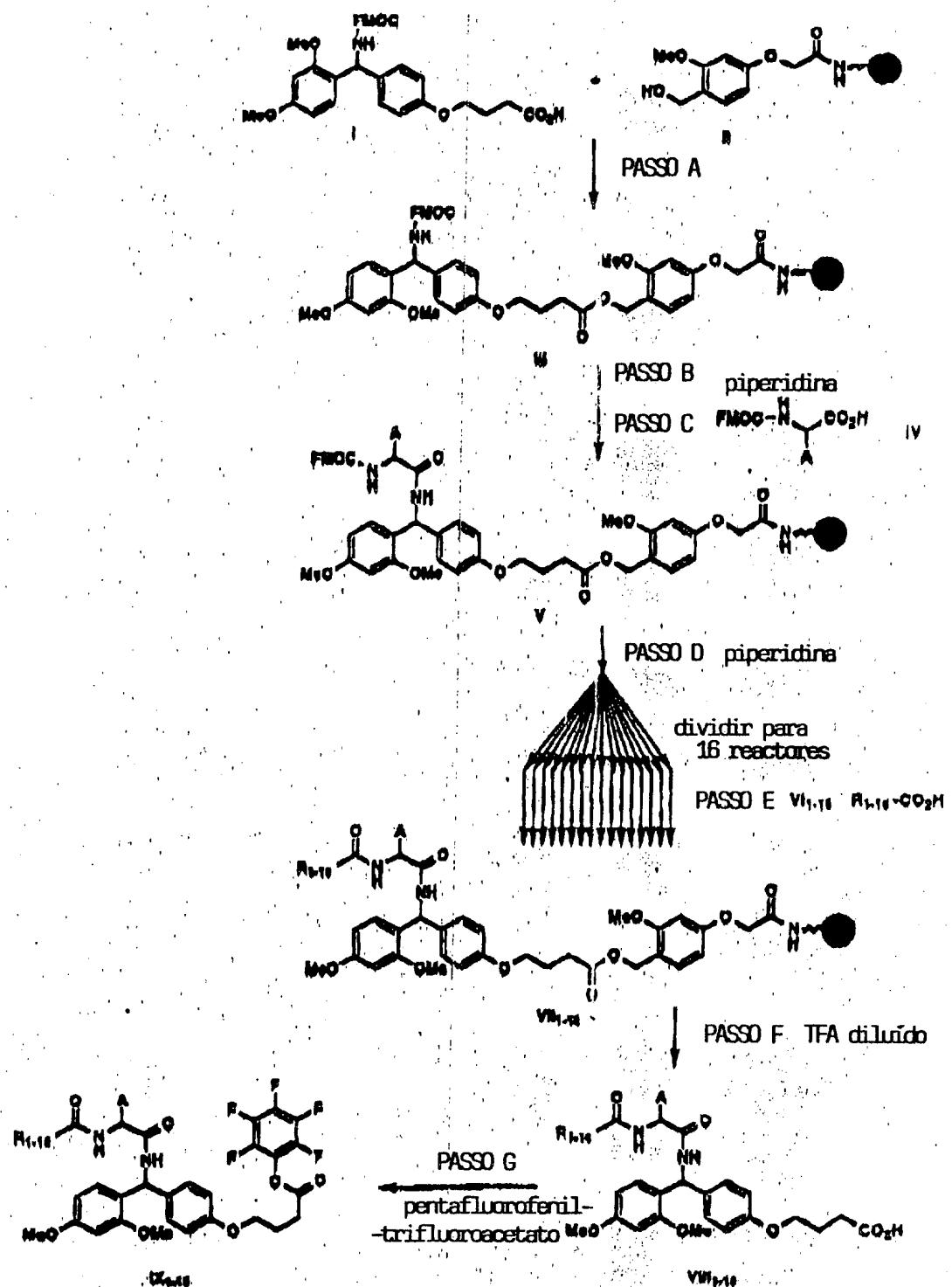


FIGURA 1

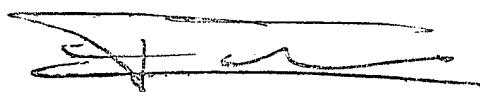
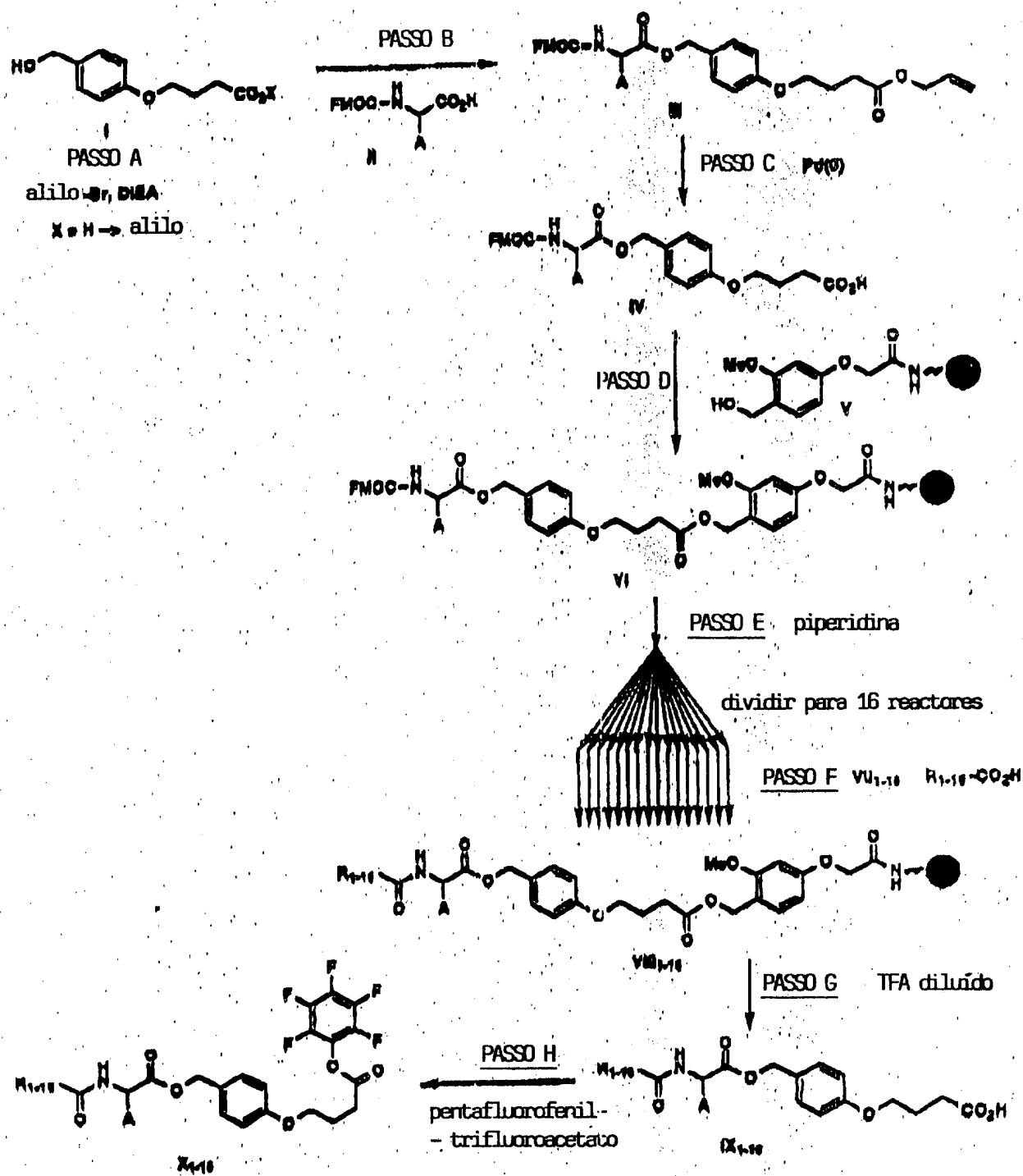



FIGURA 2

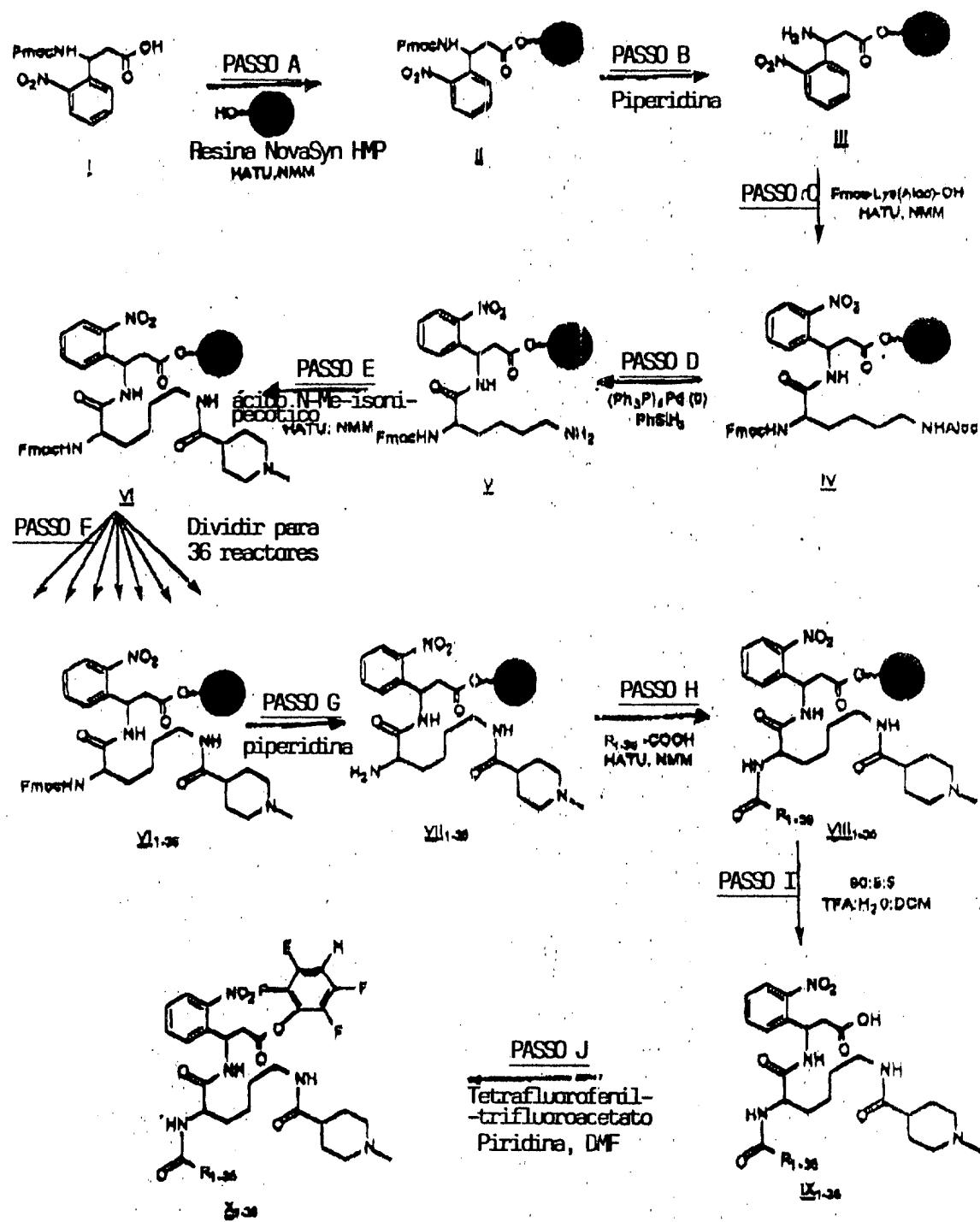


FIGURA 3

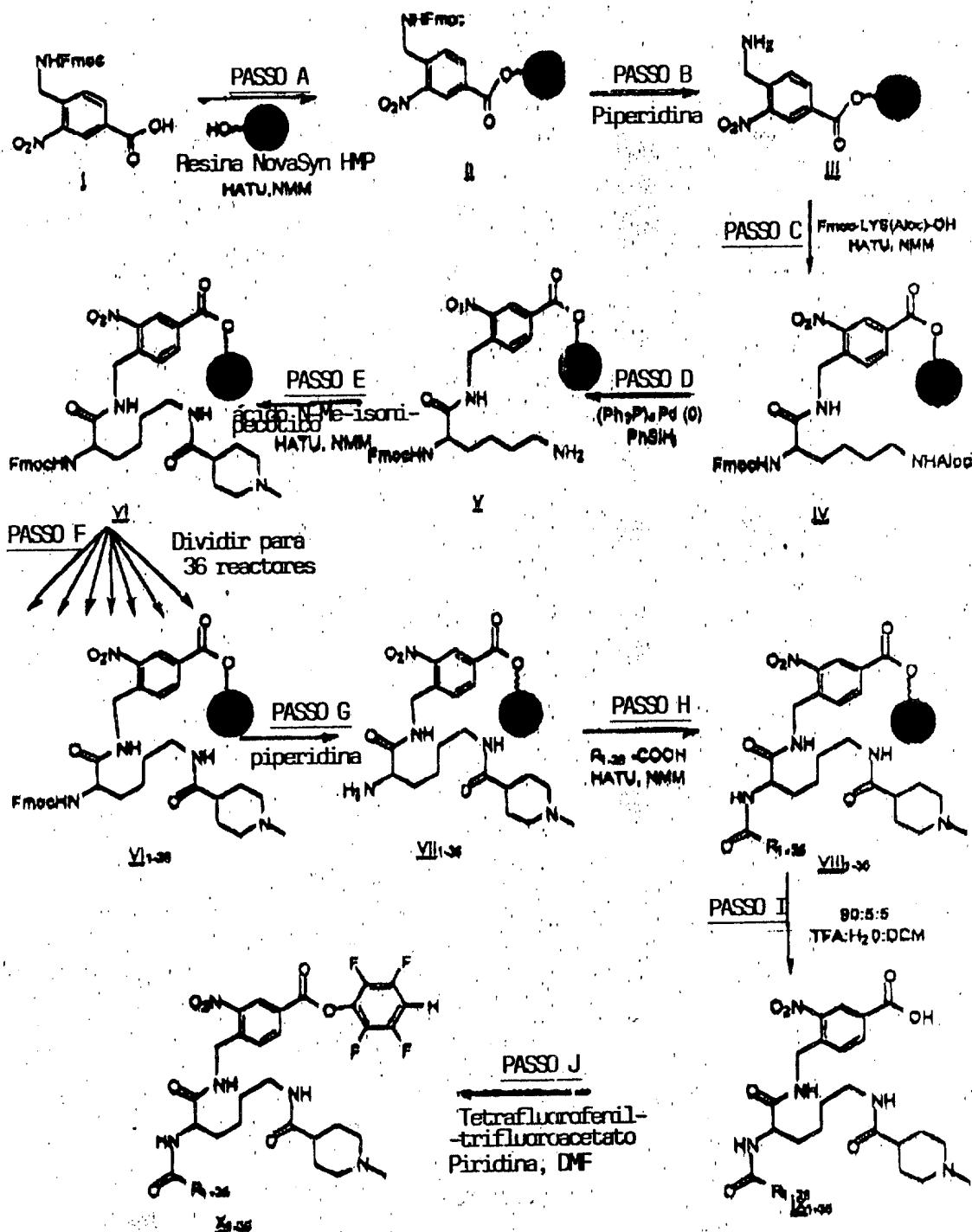


FIGURA 4

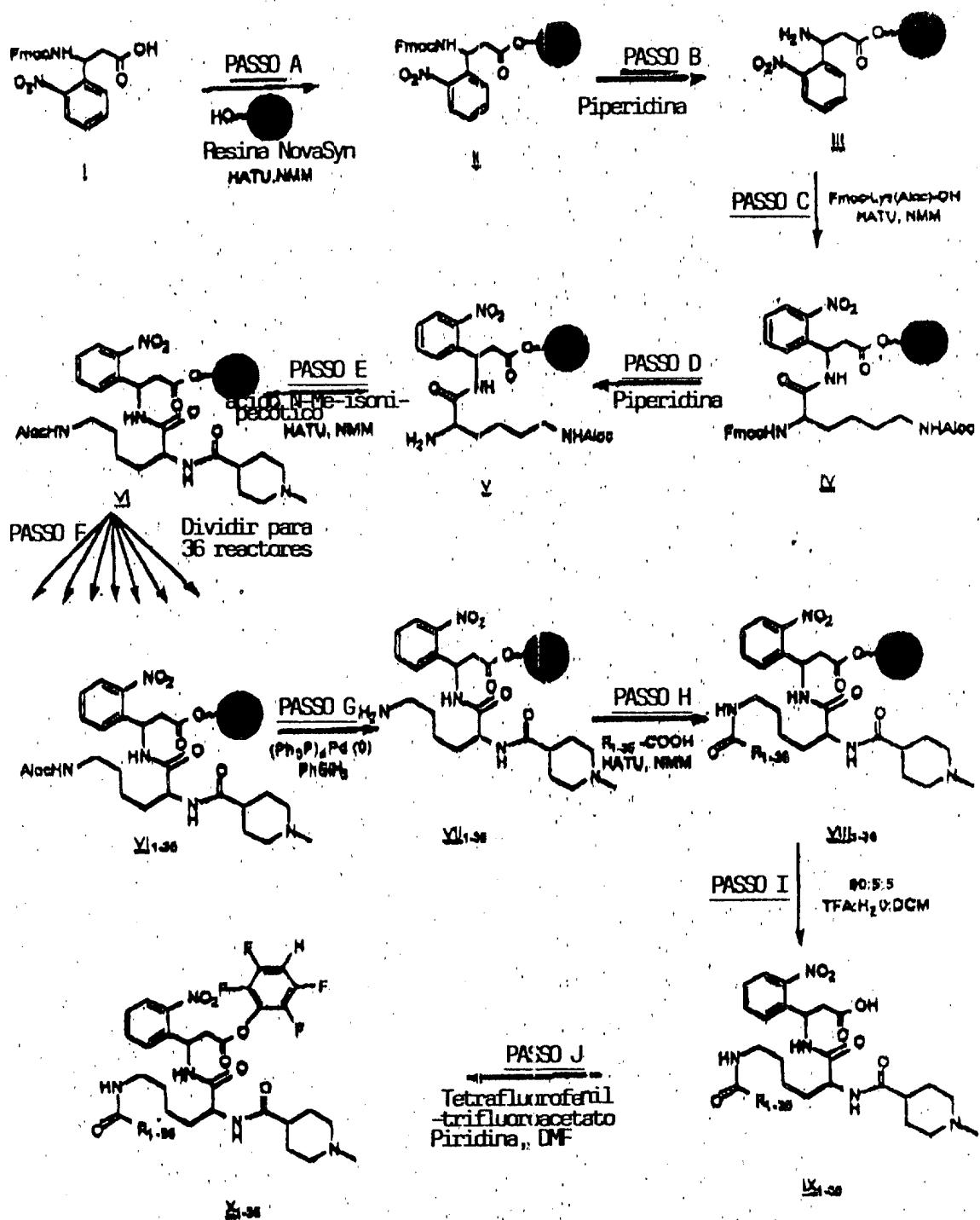


FIGURA 5

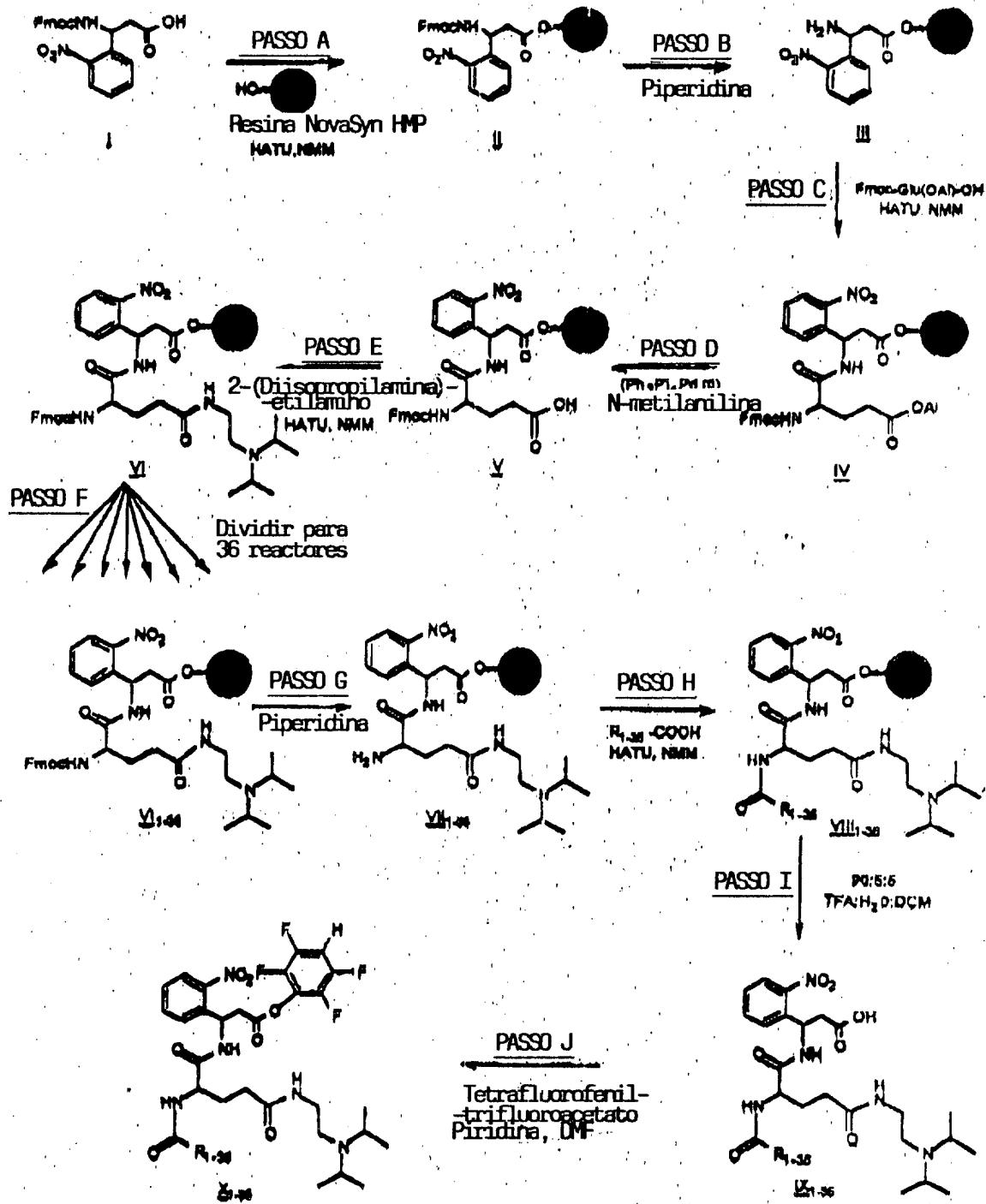


FIGURA 6

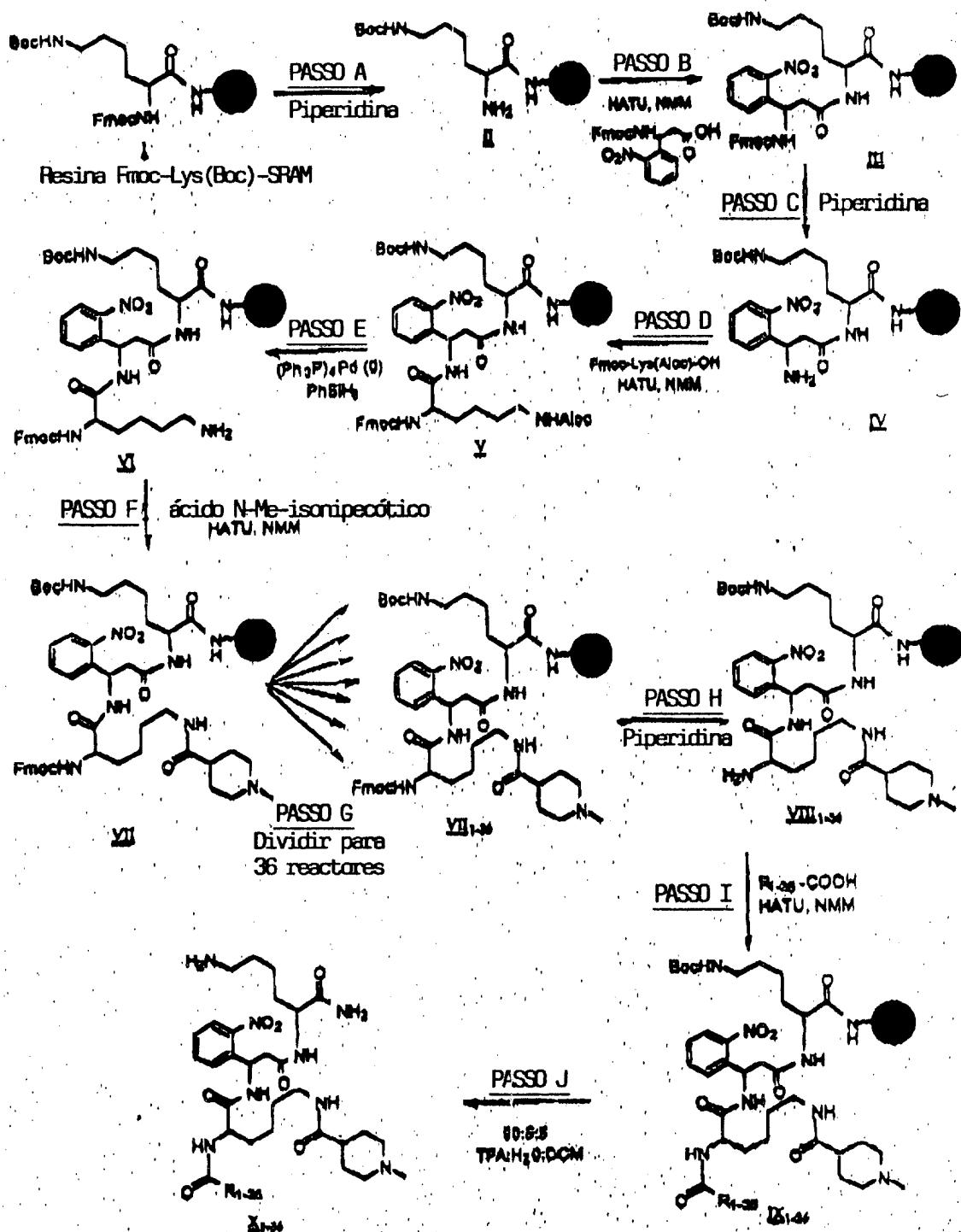


FIGURA 7

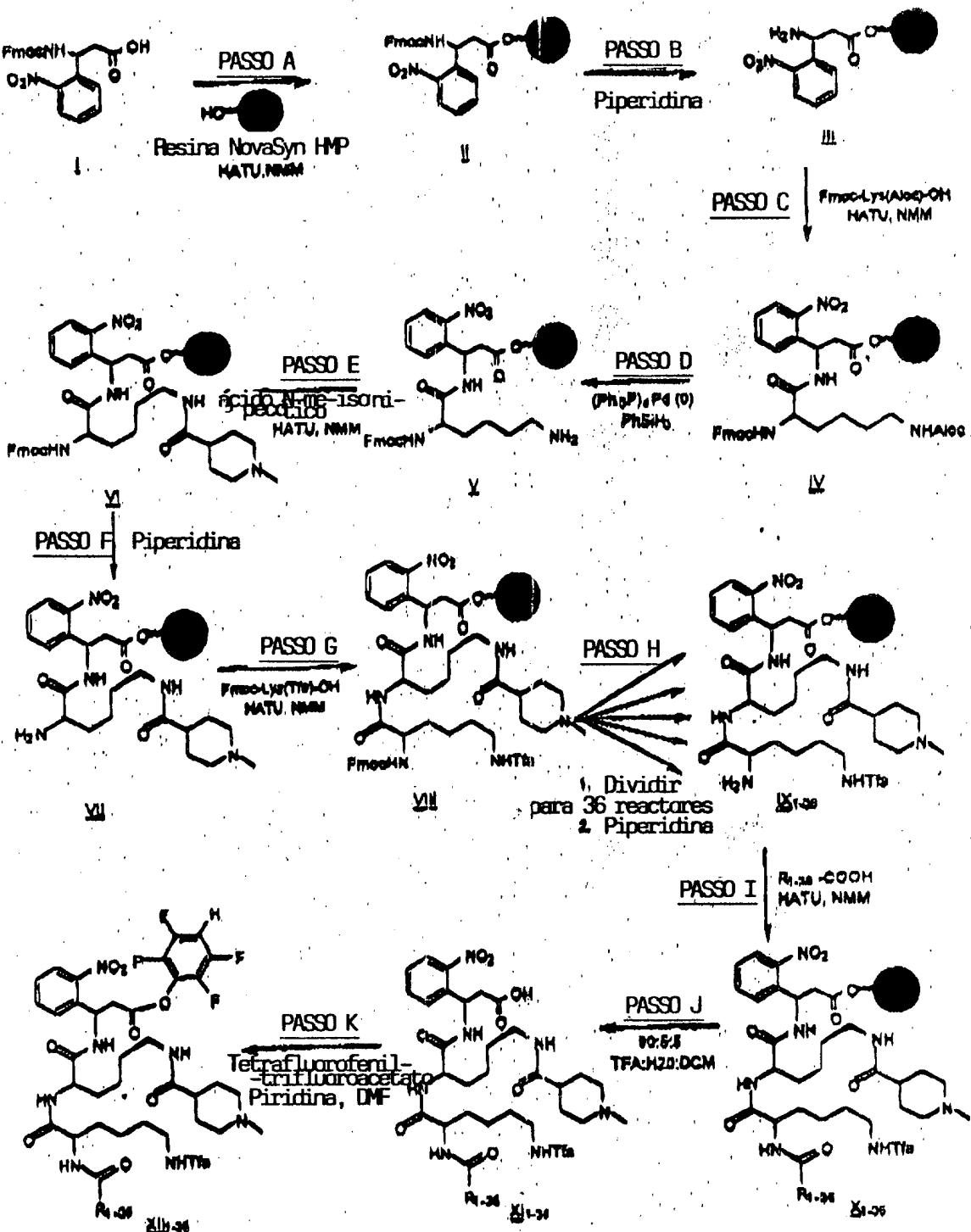


FIGURA 8

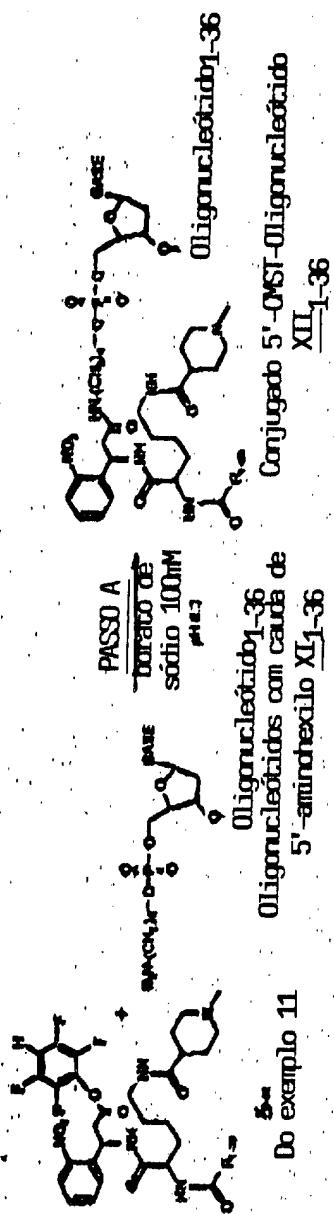


FIGURA 9

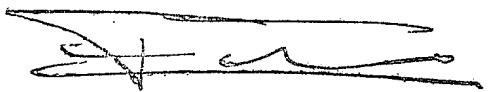
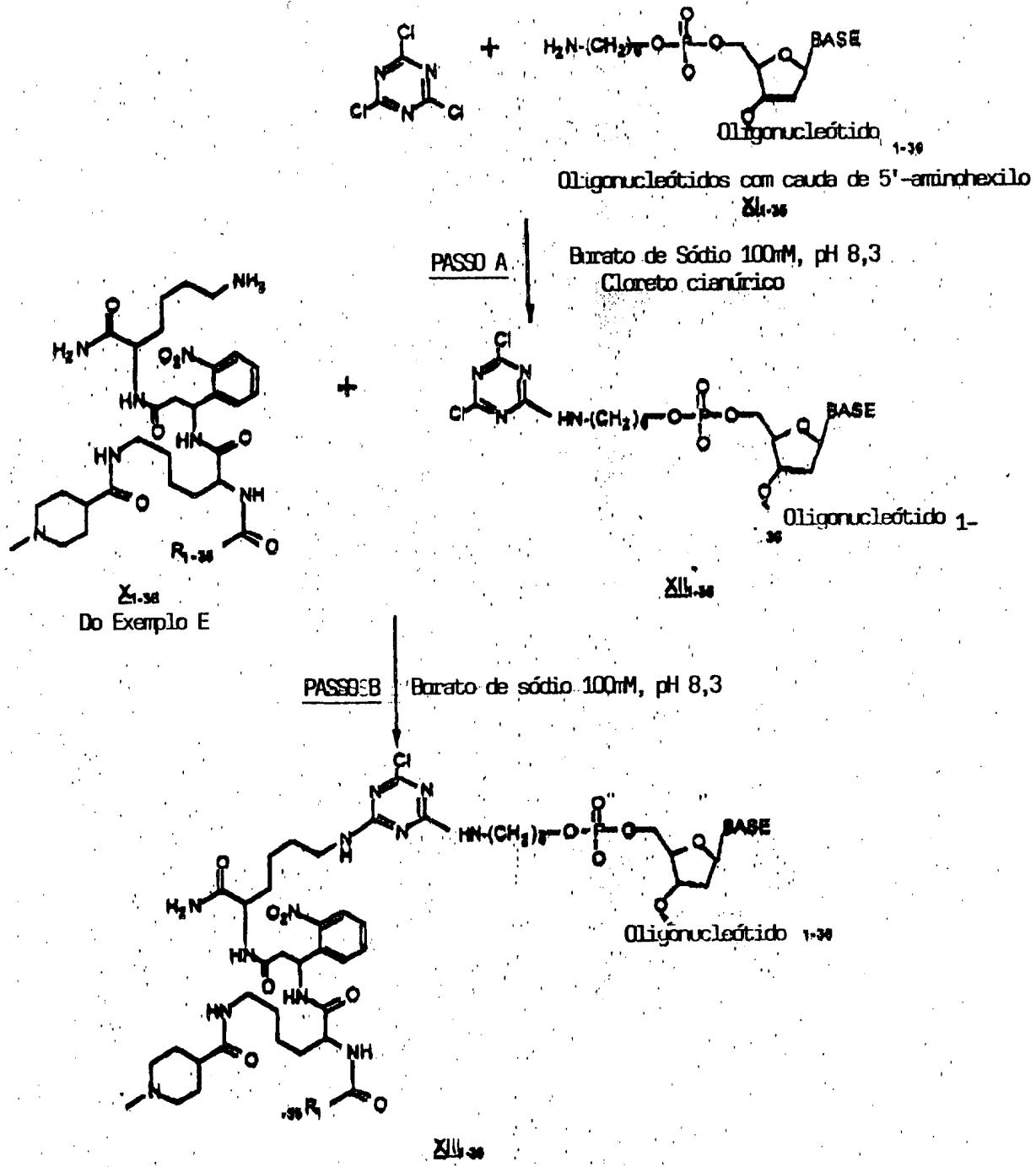



FIGURA 10

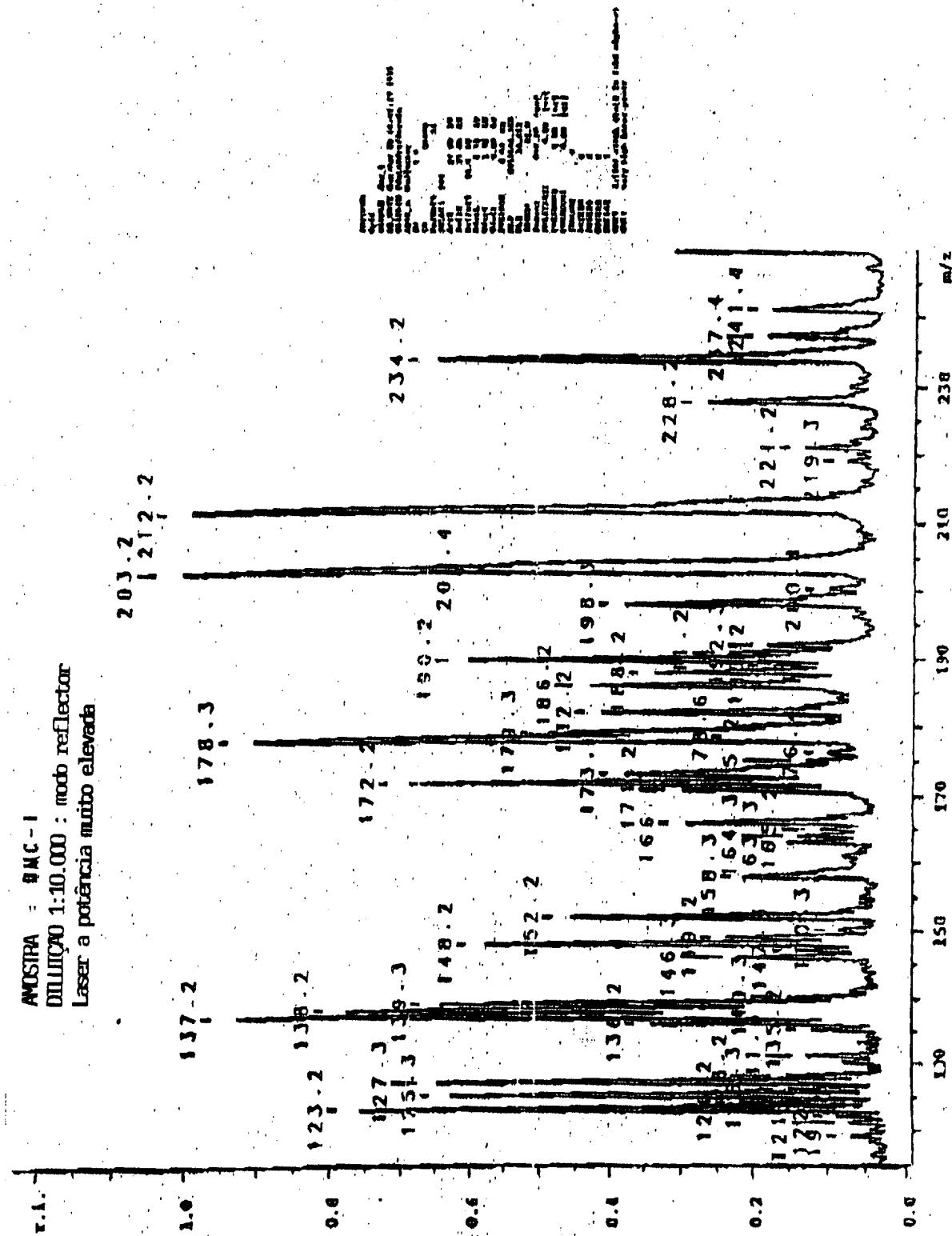


FIGURA 11

M



FIGURA 12

Componente de Massa Variável	Ajustador do Intervalo de Massas	Melhorador da Sensibilidade Da Espectr. Massa	Linker suscetível a corte fotolítico	Oligonucleotido com Cauda de 5'-Aminohexilo
R_{wra}	R_{wra}	R_{msse}	NO_2	$N-(C_2-C_{10})-ODN-3'-OH$

FIGURA 13