



INSTITUTO NACIONAL DA PROPRIEDADE INDUSTRIAL

DIRECÇÃO DE SERVIÇOS DE PATENTES

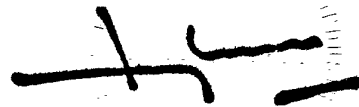
CAMPO DAS CEBOLAS, 1100 LISBOA
TEL.: 888 51 51 / 2 / 3 TELEX: 18356 INPI
TELEFAX: 87 53 08

FOLHA DO RESUMO

Modalidade e n.º (11)	T D	Data do pedido: (22)	Classificação Internacional (51)	
PAT. No. 101.014				
Requerente (71) THE GENERAL HOSPITAL CORPORATION, norte-americana, industrial, com sede em Fruit Street, Boston, Massachusetts 02114, Estados Unidos da América do Norte				
Inventores (72): ROBERT E. KINGSTON e OPHELIA PAPOULAS residentes nos Estados Unidos da América do Norte				
Reivindicação de prioridade(s) (30)		Figura (para interpretação do resumo)		
Data do pedido	Pais de Origem			N.º de pedido
30.10.91	E.U.A.			07/785,567
Epigrafe: (54) "PARCEIROS, MOTIVOS, ENSAIOS DE DESPISTE E UTILIZAÇÕES DA LIGAÇÃO DE C-MYC A DNA"				
Resumo: (máx. 150 palavras) (57)				
O presente invento diz respeito ao desenvolvimento de uma linha celular de mamífero que expressa abundantemente Myc e a purificação de quantidades significativas de c-Myc a partir destas células. São descritos três tipos de oligomerização de proteína (ou formações de complexos) conduzidos por c-Myc: (1) complexos de homo-oligómero (daqui em diante designados complexos C1) formados pela associação de pelo menos dois peptídeos de c-Myc, (2) complexos de hetero-oligómeros (aqui designados complexos C2) formados por heterodimerização de pelo menos dois peptídeos, pelo menos um dos quais não é o peptídeo c-Myc e especificamente hetero-oligomerização entre c-Myc e um factor de 26-29 Kd e (3) complexos hetero-oligoméricos (aqui designado complexo C2') formado na presença de c-Myc, no entanto tais proteínas hetero-oligoméricas não contendo quaisquer peptídeos que sejam c-Myc. O invento é ainda dirigido a um método reprodutível e preciso para a classificação objectiva de compostos, incluindo fármacos humanos, como inibidores da actividade de c-Myc.				

NÃO PREENCHER AS ZONAS SOMBREADAS

101.074



**"PARCEIROS, MOTIVOS, ENSAIOS DE DESPISTE E UTILIZAÇÕES DA LIGAÇÃO
DE C-MYC A DNA"**

=====

MEMÓRIA DESCRITIVA

Resumo

O presente invento diz respeito ao desenvolvimento de uma linha celular de mamífero que expressa abundantemente Myc e a purificação de quantidades significativas de c-Myc a partir destas células. São descritos três tipos de oligomerização de proteína (ou formações de complexos) conduzidos por c-Myc: (1) complexos de homo-oligómero (daqui em diante designados complexos C1) formados pela associação de pelo menos dois peptídeos de c-Myc, (2) complexos de hetero-oligómeros (aqui designados complexos C2) formados por heterodimerização de pelo menos dois peptídeos, pelo menos um dos quais não é o peptídeo c-Myc e especificamente hetero-oligomerização entre c-Myc e um factor de 26-29 Kd e (3) complexos hetero-oligoméricos (aqui designado complexo C2') formado na presença de c-Myc, no entanto tais proteínas hetero-oligoméricas não contendo quaisquer peptídeos que sejam c-Myc. O invento é ainda dirigido a um método reproduzível e preciso para a classificação objectiva de compostos, incluindo fármacos humanos, como inibidores da actividade de c-Myc.

Referência a Pedidos de Patente Relacionados

Este pedido de patente é uma continuação em parte do Pedido de Patente Nº 07/510,253, entregue em 19 de Abril, 1990.

Campo do Invento

Este invento é dirigido a métodos para a purificação da proteína Myc de mamíferos e métodos para a identificação de compostos que inibem a actividade de transcrição de c-Myc.

FUNDAMENTO DO INVENTO

Myc é um oncogene nuclear cuja expressão aberrante está associada a muitos tipos diferentes de cancros humanos em tecidos muito diferentes (Cole, M.D., Ann. Rev. Genet. 20:361-384 (1986)). Se bem que o mecanismo da acção da oncoproteína c-Myc permaneça desconhecido, sabe-se claramente que ela desempenha um papel no controle do crescimento e diferenciação celulares (Luscher e Eisenman, Genes & Dev. 4:2025-2035 (1990); Penn et al., Sem. Cancer Biol. 1:69 (1990)). Um mecanismo plausível para a acção de Myc é como regulador da transcrição numa via que controla directamente a proliferação e diferenciação. Este modelo é consistente com várias observações. Primeiro, sabe-se desde há muito tempo que Myc é uma proteína nuclear com uma afinidade geral para DNA (Abrams et al., Cell 29:427-439 (1982); Alitalo et al. Nature 306:274-277 (1983); Donner et al. Nature 296:262-265 (1982); Persson e Leder, Science 225:718-721 (1984)) e recentemente foi identificado um sítio que é ligado especificamente por variantes de c-Myc expressas em bactérias (Blackwell et al. Science 250:1149-1151 (1990)); Prendergast e Ziff, Science 251:186-189 (1991)). Em segundo lugar, mostrou-se que c-Myc de tamanho completo activa e reprime genes em ensaios de transfecção

transitória (Kaddurah-Daouk et al. Genes & Dev. 1:347-357; Yang et al., Mol. Cell. Biol. 11:2291-2295 (1991)) e estimulam fracamente a transcrição quando fundidos a um domínio heterólogo de ligação a DNA (Lech et al., Cell 52:179-184 (1988); Kato et al. Mol. Cell. Biol. 10:5914-5920 (1990)). E finalmente, semelhanças de sequências descritas abaixo colocam Myc no grupo dos factores de transcrição conhecidos.

Myc contem dois domínios que sugerem que se oligomeriza, talvez como um dímero, e liga-se especificamente a DNA: um domínio tipo fecho de correr de leucina e um domínio básico helix-ansa-helix (B-HLH). O domínio em fecho de correr de leucina é uma estrutura em hélice alfa encontrada em proteínas que se ligam a DNA numa sequência específica como seja Fos e Jun onde medeia a homo ou heterodimerização via uma interacção enrolado-enrolado (Landschulz et al., Science 240: 1759-1764 (1988); O'Shea et al., Science 243:538-542 (1982); e revisto em Busch e Sassoned-Corsi, TIG 6:36-40 (1990)). Esta dimerização é necessária para a ligação ao DNA (Gentz et al., Science 243:1695-1699 (1989); Halazonetis et al., Cell 55:917-924 (1988); Kouzarides e Ziff, Nature 336:646-651 (1988); Turner e Tijan, Science 243:1698-1694 (1989)). A região HLH também parece mediar a oligomerização necessária para a ligação a DNA em várias proteínas importantes no desenvolvimento (Murre et al., Cell 58:537-544 (1989); Murre et al., Cell 56: 777-783 (1989)). As proteínas HLH formam uma família grande e cada vez maior de genes e incluem os produtos dos genes achaete-scute e daughterless responsáveis pelo desenvolvimento neural em Drosophila, a família dos genes R que regulam o padrão de pigmentação no milho, MyoD e várias outras proteínas envolvidas na diferenciação específica dos músculos em vertebrados e uma proteína de ligação ao centrómero, CBF1, de levedura (Braun et al., EMBO J. 8:701-709 (1989); Cai e Davies, Cell 61:437-446 (1990); Caudy et al., Cell 55:1061-1067 (1988);

Cromiller et al. Genes & Dev. 2:1666-1676 (1988); Davies et al., Cell 51:987-1000 (1987); Edmondson e Olson, Genes & Dev. 3:628-640 (1989); Ludwig e Wessler, Cell 62:849-851 (1990); Pinney et al. Cell 53:781-793 (1988); Rhodes e Konieczny, Genes & Dev. 3:2050-2061 (1989); Villares e Cabrera, Cell 50:415-424 (1987); Wright et al., Cell 56:607-617 (1989)). Se bem que muitas proteínas contenham um HLH ou um motivo em fecho de correr de leucina, Myc é uma de um pequeno número de proteínas que contem HLH e simultaneamente um domínio em fecho de correr com leucina. Tanto as proteínas contendo o domínio em fecho de correr com leucina como as proteínas com HLH requerem um segmento de aminoácidos básicos adjacentes ao motivo de dimerização para o contacto com o DNA (revisto em Busch e Sassone-Corsi, TIG 6:36-40 (1990); Jones, N., Cell 61:9-11 (1990)). É interessante que todas as proteínas B-HLH parecem ligar-se a sequências de DNA estreitamente relacionadas conhecidas como as caixas E. Estes são motivos de sequências encontradas na imunoglobulina e noutros potenciadores específicos de tecidos tendo um núcleo NNCANNTGNN [SEQ ID Nº 16] onde diferentes bases centrais são preferidas por diferentes proteínas B-HLH e as sequências delimitantes podem afectar a afinidade da ligação (Blackwell et al., Science 250:1149-1151 (1990); Blackwell e Weintraub, Science 250:1104-1110 (1990)). O núcleo do sítio de ligação descrito para c-Myc, CACGTG, enquadra-se neste padrão e tem a mesma sequência do núcleo do elemento da sequência a montante (USE) do principal promotor tardio dos Adenovírus (Blackwell et al., Science 250:1149-1151 (1990); Prendergast e Ziff, Science 251:186-189 (1991)). Um factor de transcrição celular (USF ou MLTF) que se liga ao USE foi recentemente clonado e também contem um domínio B-HLH adjacente ao domínio em fecho de correr de leucina (Gregor et al. Genes Dev. 4:1730-1740 (1990)).

Observou-se que muitas destas proteínas B-HLH ou com o domínio em fecho de correr de leucina foram homodímeros e heterodímeros com outras proteínas tendo motivos de dimerização semelhantes (revisto em Busch e Sassone-Corsi, TIG 6:36-40 (1990); Jones N., Cell 61:9-11 (1990)). A heterodimerização entre grupos específicos de proteínas com B-HLH ou com domínios em fecho de correr de leucina pode alterar as suas propriedades de ligação ao DNA. Se bem que os homodímeros se devam ligar fracamente, os heterodímeros com o parceiro adequado podem ligar-se com uma maior afinidade e nalguns casos com uma nova especificidade (Jones, N., Cell 9:9-11 (1990)), Blackwell e Weintraub, Science 250: 1104-1110 (1990); Wright et al., Mol. Cell. Biol. 11:4104-4110 (1991)). Myc é capaz de formar um homo-oligómero em concentrações elevadas in vitro (Dang et al., Nature 337:664-666 (1989); Kerhöff e Bister, Oncogene 6:93-102 (1991)), ainda que não seja claro o facto da formação do homo-oligómero in vivo (Dang et al. Mol. Cell Biol. 11:954-962 (1991)). Parece provável que Myc tenha uma interacção directa com outras proteínas celulares para formar hetero-oligómeros e de facto tal proteína parceira designada Max foi recentemente identificada (Blackwood e Eisenmann, Science 251:1211-1217 (1991)). O efeito que tais proteínas parceiras possam ter na especificidade da ligação de Myc a DNA é provável que sejam fundamental para a compreensão da função de Myc.

Muito do trabalho in vitro feito com proteínas B-HLH utilizou proteína transcrita e traduzida in vitro ou usou proteína expressa em bactérias. Myc expressa por estes meios foi usada para determinar a especificidade da ligação e para demonstração de que Myc pode formar heterodímeros com Max (Blackwell et al., Science 250:1149-1151 (1990); Prendergast e Ziff, Science 251: 186-189 (1991); Blackwood e Eisenmann, Science 251: 1211-1217 (1991)). Myc, no entanto, é modificada após a tradução por pelo

menos uma fosforilação em células de memífero (Hann e Eisenmann, Mol. Cell. Biol. 4: 2486-2497 (1984); Ramsay et al., Proc. Natl. Acad. Sci. USA 81:7742-7746 (1984)) e pensa-se que as modificações após tradução regulem a função de muitas proteínas, incluindo os factores de transcrição Myb, Fos, HSF, CREB e SP-1 (Abate et al., Science 249:1157-1161 (1990); Jackson et al., Cell 63:155-165 (1990); Luscher et al., Nature 344:517-522 (1990); Sorger et al., Nature 329:81-84 (1987); Yamamoto et al., Nature 334:494-498 (1988)). Em adição, foi descrito que Myc produzido em células aviárias se liga mais fortemente a DNA-celulose do que Myc produzido em bactérias (Kerhoff e Bister, Oncogene 6:93-102 (1991)).

Várias linhas de evidência argumentam que as funções bioquímicas de Myc serão determinadas em grande parte pela hetero-oligomerização com Max e talvez com outros factores ainda não identificados. Para uma compreensão completa da função de c-Myc será pois necessária a identificação de todas as proteínas parceiras e uma caracterização funcional dos complexos que estas proteínas formam na ausência ou presença de c-Myc. Para descobrir a natureza da função de Myc será necessário determinar não só as propriedades de ligação de todos os complexos importantes como também determinar se diferem no seu modo de acção depois de ligados. A modificação após tradução deve desempenhar um papel na modulação da formação, ligação ou outras actividades destes complexos e a disponibilidade de grandes quantidades de c-Myc modificada como aqui se descreve deverá facilitar uma abordagem bioquímica para este problema. Tais estudos deverão conduzir-nos a uma compreensão dos complexos disponíveis em tempos diferentes em diferentes tipos celulares e as consequências para cada célula em termos de crescimento e diferenciação adequados ou oncogénese.

Ainda, até à data não foram identificados inibidores da acção de c-Myc. A identificação de tais inibidores deve-se à ausência de identificação de uma sequência específica de ligação ao DNA a que se ligue c-Myc e da inexistência de um ensaio simples, barato e reprodutível que possa rapidamente identificar potenciais inibidores e seus derivados activos. Assim, existe também a necessidade de ensaios de despiste rápidos e económicos que identifiquem actividade inibidora específica de c-Myc.

SUMÁRIO DO INVENTO

Reconhecendo a importância potencial dos inibidores da actividade da oncoproteína c-Myc no tratamento de muitas formas de cancro e sabendo da inexistência de um sistema de ensaio simples em que tais inibidores possam ser identificados, os inventores estudaram a ligação ao DNA de c-myc.

Estes esforços conduziram ao desenvolvimento de uma linha celular de mamífero que expressa abundantemente Myc e à purificação de quantidades significativas de c-Myc a partir destas células. Estes esforços culminaram na descoberta de três tipos de oligomerização proteica ou formação de complexos conduzida por c-Myc: (1) complexos de homo-oligómeros (aqui designados complexos C1) formados por associação de pelo menos dois peptídeos de c-Myc, (2) complexos de hetero-oligómeros (aqui designados complexos C2) formados por heterodimerização de pelo menos dois peptídeos, pelo menos um dos quais não é o peptídeo c-Myc e especificamente hetero-oligomerização entre c-Myc e um factor de 26-29 Kd, e (3) complexos hetero-oligoméricos dependentes de c-Myc (aqui designados complexo C2') formado na presença de c-Myc, no entanto, tais proteínas oligoméricas não contendo quaisquer peptídeos que sejam c-Myc.

Assim, o invento é dirigido a um método reprodutível e preciso para a purificação de Myc a partir de uma fonte de mamífero.

O invento é ainda dirigido à utilização de oligómeros contendo o motivo de DNA 5'-CACGTG-3', na sua forma de cadeia dupla, como um método reprodutível e preciso para a detecção da presença de complexos C1 numa amostra.

O invento é ainda dirigido à utilização do motivo de DNA 5'-CACGTG-3', na sua forma de DNA de cadeia dupla, como um método reprodutível e preciso para a detecção de complexos C2 numa amostra.

O invento é ainda dirigido à utilização do motivo de DNA 5'-CAGCTG-3', na sua forma de DNA de cadeia dupla, como um método reprodutível e seguro para a detecção de complexo C2' numa amostra.

O invento é ainda dirigido a uma fracção de proteína de 26-29 Kd purificada a partir de células de ovário de hamster chinês (CHO) ou de baculovírus, tal fracção proteica contendo pelo menos um peptídeo capaz de formar oligómeros do complexo C2 com c-Myc.

O invento é ainda dirigido a uma fracção proteica de 40-50 Kd purificada a partir de células CHO, tal fracção proteica contendo pelo menos um peptídeo capaz de formar oligómeros do complexo C2' na presença de c-Myc.

O invento é ainda dirigido a um método reprodutível e preciso para a classificação objectiva de compostos, incluindo fármacos humanos, como inibidores da actividade de c-Myc e

especialmente como um inibidor da formação do complexo C2 ou da formação do complexo C2'.

O invento é ainda dirigido a um método reprodutível e preciso para a classificação objectiva de compostos, incluindo fármacos humanos, como inibidores da actividade de c-Myc e especialmente como um inibidor da ligação a DNA do complexo C1, da ligação a DNA do complexo C2 ou da ligação a DNA do complexo C2'.

O invento proporciona ainda um método para a identificação e classificação do mecanismo de acção de um composto bioactivo inibidor de c-Myc.

O invento proporciona ainda um ensaio para o controle do isolamento e/ou purificação de um peptídeo capaz de formar um complexo C2 ou C2' ou uma mistura de tais peptídeos a partir de uma preparação bruta.

O invento ainda proporciona um ensaio para o controle do isolamento e/ou purificação de um composto inibidor de c-Myc ou mistura de tais compostos a partir de uma preparação bruta de tais compostos.

BREVE DESCRIÇÃO DOS DESENHOS

Fig. 1 Proteína c-Myc purificada. A) 1 μ g de proteína c-Myc purificada a partir da linha celular CHO expressando abundantemente 5A foi sujeito a electroforese em gel bidimensional. Um gel em tubo para isoelectrofocagem correu com anfolitos pH 5-7 seguido de SDS-PAGE e coloração com prata. As proteínas c-Myc estão entre parêntesis e as setas distinguem as espécies de 60, 62 e 72 KDa. O gel foi arranjado para esta figura; a gama de

pI para as proteínas Myc foi de 5,0-5,6. B) 0,5 μ g da proteína c-Myc purificada derivada das linhas celulares indicadas foi sujeita a electroforese num gel de SDS e visualizada por coloração com prata (faixa da esquerda) ou electrotransferido para nitrocelulose e sujeito a imunotransferência usando o anticorpo policlonal ST-2 (2 faixas da direita).

Fig. 2 Ligação a DNA das Proteínas c-Myc Purificadas. O EMSA foi realizado como descrito em materiais e métodos usando quantidades iguais (aproximadamente 2 ng) das sondas que se seguem e 0,5 μ g de c-Myc purificado produzido por CHO ou c-Myc produzido por baculovírus: (μ E2)₃ faixas 1 e 7, (μ SE)₃ faixas 2 e 8, MLC-A faixas 3 e 9, MLC-B faixas 4 e 10, (USE)₃ faixas 5, 11 e 12 e HES faixas 6 e 13. Nos materiais e métodos são dadas as sequências de tamanho completo da sonda. As faixas 1-6 e as faixas 7-13 são exposições diferentes das faixas do mesmo gel.

Fig. 3 A actividade de ligação de C1 está presente em fatias de géis de SDS contendo Myc. 400 μ g de c-Myc produzido por CHO ou 163 μ g de c-Myc produzido por baculovírus foram separados num gel de SDS-PAGE. As proteínas de fatias de 0,5 cm foram recuperadas, renaturadas como descrito em materiais e métodos e analisadas por EMSA usando a sonda (USE)₃. 0,4 μ g da Myc de CHO aplicada e 5 μ l da proteína da fatia contendo Myc de CHO foram analisados no mesmo gel (painel da esquerda). 0,37 μ g de Myc de baculovírus aplicado e 5 μ l da proteína da fatia Myc de baculovírus foram analisados no mesmo gel (painel da direita). As fatias das outras gamas de peso molecular do mesmo gel não apresentaram qualquer ligação (resultados não apresentados).

Fig. 4. A actividade é formada por c-Myc e um factor de 26-29 KD. Proteínas de fatias de gel foram recuperadas e

analizadas por EMSA como descrito em materiais e métodos usando a sonda (USE)₃. As faixas 1-4 representam proteínas do mesmo gel aplicadas com Myc produzido por baculovírus descrito na Fig. 5. Estas faixas contêm 0,37 µg do material aplicado (faixa 1), 0,75 µg de BSA com 7,5 µl de proteínas de uma fatia de Myc (faixa 2) ou uma fatia de 26-29 KD (faixa 3) ou 7,5 de cada fatia usada para as faixas 1 e 2 mais 0,2 µg de BSA (faixa 4). As faixas 5-8 e 10 contêm proteínas de géis aplicadas com Myc purificada a partir de células CHO. Estas faixas contêm 0,47 do aplicado no gel (faixa 5), 4 µl de material de uma fatia de Myc de um gel a que se aplicou 400 µg de Myc (faixa 6), 7 µl de material de uma fatia de 26-29 KD de um gel semelhante mais 0,8 µg de Proteína A (faixa 7) e 4 µl da fatia de Myc e 7 µg da fatia de 26-29 KD (faixa 8). As faixas 9-12 utilizam as proteínas de fusão Proteína A-Myc expressas por bactérias contendo os domínios B-HLH e em fecho de correr de leucina de Myc (aminoácidos 353-439) ou sem a região básica e contendo os aminoácidos 372-439 de Myc. Estas foram expressas e purificadas como descrito em materiais e métodos. A faixa 9 contem 0,5 µg de Proteína A-Myc(353-439) e a faixa 10 contem o mesmo mais 7 µl da fatia de 26-29 KD. A faixa 11 contem 1 µg da Proteína A-Myc(372-439) e a faixa 12 contem 0,5 µg de Proteína A-Myc(372-439) mais 7 µl da fatia de 26-29-KD.

Fig. 5. A actividade de ligação de C2' requiere um factor de 40-50 KD. A) 101 µg de c-Myc produzido por CHO foi separado num gel de SDS. As proteínas foram recuperadas, ressuspensas em 100 µl e renaturadas e analisadas por EMSA usando a sonda ERP3/4. Esta sonda contem a fracção do potenciador MLC que engloba o sítio µE2. As amostras de EMSA continham 0,3 µg do material aplicado no gel de SDS (faixa 1), 7,5 µl das proteínas da fatia de Myc (faixa 2) ou a fatia de 40-50 KD (faixa 3) ou 7,5 µl de ambas as fatias renaturadas em conjunto (faixa 4). B) As amostras de EMSA continham 0,9 µg de c-Myc produzido por



baculovírus (faixa 5), 3 μ l de proteína derivada da fatia de 40-50 KD de um gel em que se aplicou 400 μ g de c-Myc produzido por CHO (faixa 6) ou ambos renaturadas em conjunto (faixa 7). A sonda foi ERP1/2. C) As amostras de EMSA continham 10 μ l (0,9 μ g) da proteína de fusão de c-Myc produzida por bactérias contendo os aminoácidos 353-439 de Myc (faixa 8), 0,47 μ g de c-Myc produzido por CHO (faixa 9), 5 μ l de proteína derivada da fatia de gel 40-50 KD de um gel em que se aplicou 400 μ g da Myc produzida por CHO mostrada na faixa 9 (faixa 10) ou 5 μ l do mesmo material de 40-50 KD renaturado na presença de 0,9 μ g de Myc produzida por baculovírus mostrada na faixa 5 (faixa 11), 2 μ l (0,18 μ g) da proteína de fusão Myc produzida por bactérias contendo os aminoácidos 353-439 de Myc (faixa 12) ou 4 μ l (0,36 μ g) da mesma proteína de fusão Myc produzida por bactérias (faixa 13). A sonda foi ERP1/2.

Fig. 6. Anticorpos contra c-Myc interactuam com os complexos C1 e C2. As reacções de EMSA foram estabelecidas com as preparações de proteína Myc indicadas (0,37 μ g de c-Myc produzido por baculovírus ou 0,47 μ g de c-Myc produzido por CHO). Estas reacções foram pré-incubadas 30 minutos em gelo na presença do anticorpo indicado (monoclonal α -Myc 1F7 ou um monoclonal dirigido contra o repressor de lambda, cI). 1 ng da sonda SMS ou a sonda contendo μ E2 número 7 (ver Fig. 7) foi adicionado subsequentemente e a ligação e a electroforese realizadas como habitualmente.

Fig. 7. Oligonucleótidos seleccionados a partir da sequência ao acaso após 8 ciclos de EMSA. As sequências foram seleccionadas a partir de oligonucleótidos contendo 20 pares de bases de sequência ao acaso usando um processo EMSA reiterativo descrito nos materiais e métodos. Os nucleótidos sublinhados são dos sítios dos iniciadores de PCR. As Tabelas abaixo mostram as

sequências alinhadas contendo a frequência de cada uma das bases nas 6 posições delimitantes à volta dos motivos CACGTG. Apenas as bases a seguir a uma correspondência correcta de CACGTG foram apresentadas uma vez que as sequências sem este núcleo central não funcionam como sítios de alta afinidade de ligação (Fig. 8, e resultados não apresentados). Os números carragados adjacentes a sequências individuais indicam os oligonucleótidos que foram testados individualmente por EMSA na Fig. 8. Os asteriscos indicam mais sequências que foram testadas individualmente (resultados não apresentados).

Fig. 8. Sítios seleccionados formam complexos previstos. Realizou-se EMSA usando 2,8 ng da sonda SMS ou iguais quantidades (1 ng) das sondas 1-11 indicadas na Fig. 7. As sondas 1-11 foram marcadas e paralelamente isoladas em gel e tinham actividades específicas aproximadamente iguais. As reacções de ligação não continham proteína adicional (-), 0,37 μ g de c-Myc produzido por baculovírus (B) ou 0,47 μ g de c-Myc produzido por CHO (C). A sonda livre é visível no fim do gel.

Fig. 9. Ensaaios de competição com os complexos C1 e C2. A reacção padrão de EMSA foi aumentada para 11 amostras contendo 0,4 μ g de c-Myc purificado produzido por baculovírus por amostra. A sonda e competidor foram (USE)₃. Após deixar 20 minutos para a ligação 20 μ l foi aplicado num gel de EMSA, previamente corrido, como meio de medir a quantidade de partida do complexo (ST) e adicionou-se suficiente competidor frio à restante amostra para se conseguir um excesso molar de 250 vezes relativamente à sonda. Imediatamente quando da adição do competidor a amostra foi suavemente misturada e amostras de 20 μ l foram aplicadas nos tempos indicados (0, 30 s, 1 min, 4 min, etc.) Uma amostra testemunha (C) foi criada individualmente na qual o competidor foi adicionado antes do início da ligação para demonstrar

competição completa. Esta amostra foi aplicada na mesma altura da amostra TS. Todas as amostras foram aplicadas num gel continuamente a correr de forma que o complexo na faixa de partida migra à frente do complexo equivalente nas faixas aplicadas mais tarde.

DESCRIÇÃO DAS REALIZAÇÕES PREFERIDAS

Na descrição que se segue usou-se extensivamente uma série de termos correntemente utilizados na tecnologia de DNA recombinante. Para proporcionar uma compreensão mais clara e consistente da especificação e reivindicações, incluindo o âmbito a ser dado a tais termos, são dadas as definições que se seguem.

Oligómero de interesse. Conforme aqui usado, um "oligómero de interesse" refere-se a qualquer um dos tipos que se seguem de proteínas oligoméricas: primeiro, oligómeros contendo Myc incluindo homo-oligómeros de peptídeos Myc (um complexo C1) e hetero-oligómeros contendo pelo menos um peptídeo de Myc e um peptídeo de um "parceiro" de Myc (um complexo C2); segundo, que formam na presença de homo-oligómeros contendo Myc ou de hetero-oligómeros contendo Myc mas que eles próprios não contêm o peptídeo Myc, tais oligómeros incluindo homo-oligómeros não contendo Myc que se associam na presença de hetero-oligómeros não contendo Myc que se associam na presença de Myc (um complexo C2').

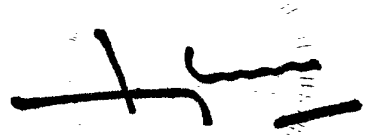
Oligómero. Um "oligómero" referindo-se a proteínas significa uma proteína composta por mais de uma subunidade peptídica, como sejam dímeros, trímeros, tetrâmeros, etc. Tal proteína oligomérica pode ser um homo-oligómero, ou seja, composto totalmente por duas ou mais subunidades idênticas; como alternativa tal proteína oligomérica pode ser um hetero-oligómero, ou seja, composto por pelo menos dois peptídeos diferentes.

Oligômeros contendo três ou mais peptídeos podem conter mais de uma cópia de um peptídeo.

Proteína(s) C2' Como aqui é usado, por conveniência, uma "proteína C2'" é uma proteína ou peptídeo que é um membro da segunda classe dos "oligômeros com interesse", ou seja, uma proteína que forma oligômeros na presença de Myc, homo-oligômeros de c-Myc ou hetero-oligômeros contendo Myc de forma a ligar-se a uma sequência específica de DNA, mas que não contem um peptídeo Myc, tais oligômeros incluindo homo-oligômeros não contendo Myc que se associam na presença de hetero-oligômeros contendo ou não Myc que se associam na presença de Myc.

Operacionalmente ligado. Tal como aqui é usado, dois elementos macromoleculares estão ligados operacionalmente quando os dois elementos macromoleculares estão fisicamente arranjados de forma a que tais factores que influenciam a actividade do primeiro elemento fazem com que este induza um efeito sobre o segundo elemento. Por exemplo, a transcrição de uma sequência codificadora que está operacionalmente ligada a um elemento promotor é induzida por factores que "activam" a actividade do promotor; a transcrição de uma sequência codificadora que está operacionalemnte ligada a um elemento promotor é inibida por factores que "reprimem" a actividade do promotor. Assim, uma região de promotor estaria operacionalmente ligada à sequência codificadora de uma proteína se a transcrição da actividade da sequência codificadora for influenciada pela actividade do promotor.

Resposta. Tal como aqui é usado o termo "resposta" destina-se a referir uma alteração em qualquer parâmetro que possa ser usado para medir, indicar ou de outra forma descrever a



acção de c-Myc ou a formação de oligómeros (homo-oligómero (complexo C1) ou hetero-oligómero (complexo C2)), incluindo a formação de hetero-oligómeros dependente de c-Myc (complexo C2'). A resposta pode ser revelada por uma alteração física (como seja uma alteração do fenótipo) ou pode ser revelada como uma alteração molecular (como seja uma alteração numa velocidade de reacção ou numa constante de afinidade). A detecção da resposta pode ser realizada por qualquer meio adequado. "Detecção" refere-se a qualquer método pelo qual tal resposta pode ser avaliada de forma a proporcionar um indício significativo sobre a ocorrência do acontecimento.

Composto. O termo "composto" destina-se a referir uma entidade química, na fase sólida, líquida ou gasosa. O termo deve ser entendido como incluindo compostos sintéticos, produtos naturais e entidade macromoleculares tais como polipeptídeos, polinucleótidos ou lípidos e também pequenas entidades tais como neurotransmissores, ligandos, hormonas ou compostos elementares.

Composto bioactivo. O termo "composto bioactivo" destina-se a referir qualquer composto que induza uma resposta detectável ou mesurável nos métodos do invento.

Promotor. Um "promotor" é uma sequência de DNA localizada próximo do início da transcrição no extremo 5' da sequência transcrita. O promotor pode conter múltiplos elementos reguladores que interactivam na modulação da transcrição do gene operacionalmente ligado.

Expressão. A expressão é um processo pelo qual a informação codificada dentro de um gene é revelada. se o gene codificar uma proteína, a expressão envolve a transcrição do DNA

para mRNA, o processamento do mRNA (se necessário) para dar um produto mRNA maduro e tradução do mRNA maduro para dar proteína.

Uma molécula de ácido nucleico, como seja um DNA ou gene é referida como sendo "susceptível de expressar" um polipeptídeo se o DNA possuir as sequências codificadoras do polipeptídeo e as sequências de controle da expressão que, no ambiente hospedeiro adequado, proporcionam a capacidade de transcrever, processar e traduzir a informação genética contida no DNA para dar origem a um produto proteico e se tais sequências de controle da expressão estiverem operacionalmente ligadas à sequência de nucleótidos que codifica o polipeptídeo.

Veículo de clonagem. Um "veículo de clonagem" é qualquer entidade molecular que seja capaz de libertar uma sequência de ácido nucleico para o interior de uma célula hospedeira para fins de clonagem. Exemplos de veículos de clonagem incluem plasmídeos ou genomas fágicos. Um plasmídeo que se pode replicar autonomamente na célula hospedeira é especialmente pretendido. como alternativa, uma molécula de ácido nucleico que pode ser inserida no DNA cromossômico da célula hospedeira é especialmente útil.

Os veículos de clonagem são muitas vezes caracterizados por um ou mais sítios de reconhecimento por endonucleases nos quais tais sequências de DNA podem ser cortadas numa forma determinável sem perda de uma função biológica essencial do veículo e no qual o DNA possa ser cortado e ligado de forma a poder sofrer replicação e clonagem.

O veículo de clonagem pode ainda conter uma marca adequada para usar na identificação de células transformadas com o veículo de clonagem. As marcas são por exemplo resistência à

tetraciclina ou resistência à ampicilina. A palavra "vector" é por vezes usada para "veículo de clonagem".

Veículo de expressão. Um "veículo de expressão" é um veículo ou vector semelhante a um veículo de clonagem mas é especialmente destinado a proporcionar sequências capazes de expressarem o gene clonado após transformação de um hospedeiro.

Num veículo de expressão, o gene a ser clonado é geralmente operacionalmente ligado a certas sequências de controle tais como sequências de promotor. As sequências de controle da expressão variarão dependendo se o vector se destina a expressar o gene operacionalmente ligado num hospedeiro procariótico ou eucariótico e podem adicionalmente conter elementos transcritoriais tais como elementos potenciadores, sequências terminadoras, elementos específicos de tecido e/ou sítios de iniciação e terminação da tradução.

Hospedeiro. Por "hospedeiro" pretende-se significar qualquer organismo que seja o recipiente de um veículo de clonagem ou expressão.

- a. Isolamento de proteína c-Myc a partir de células de mamífero e preparação de fracções contendo actividade de ligação dos complexos C2 e C2'

Se bem que tenha havido anteriormente descrições de proteína Myc purificada, os presentes inventores encontraram que as preparações de proteína Myc descritas e os métodos usados para isolar aquela proteína não conseguiam a quantidade necessária para caracterizar a acção de Myc em mamíferos. Os inventores ultrapassaram este problema e descrevem pela primeira vez um método original e útil para o isolamento de proteína c-Myc de

mamífero em grande quantidade necessário à caracterização da ligação a DNA dirigida por c-Myc os inventores conseguiram também purificar grandes quantidades de Myc a partir de um sistema recombinante de células de insectos. A proteína Myc purificada do invento apresenta a única actividade bioquímica conhecida de c-Myc, uma capacidade para de ligar à sequência CACGTG. Como resultado directo do método do invento para o isolamento da proteína c-Myc, os inventores foram capazes de identificar peptídeos que naturalmente se associam a c-Myc em hetero-oligómeros ou peptídeos que naturalmente se associam uns aos outros como resultado da acção de c-Myc, tais peptídeos estando presentes em certas fracções de cromatografia em coluna do esquema de purificação de c-Myc.

De acordo com o invento, a purificação de Myc a partir de uma fonte de mamífero é de preferência conseguido usando uma linha celular de mamífero que expressa c-Myc recombinante ou não recombinante e é realizada totalmente em gelo ou temperaturas equivalentes de 0-5°C, usando reagentes e tampões à mesma temperatura. Por exemplo, a linha celular de ovário de hamster chinês (CHO) 5A que expressa Myc em grande quantidade é útil para tal purificação. Nas células CHO5, c-Myc recombinante de murganho está sob o controle de um promotor regulável e foi integrado e amplificado no genoma da linha celular parental CHO para estabilidade e produção máximas. Excepto sempre que de outra forma seja notado, para os métodos e ensaios do invento, Myc nativo ou recombinante deverá incluir pelo menos dois exões codificadores de Myc.

Após colheita das células por centrifugação usando técnicas conhecidas e antes da lise da membrana celular externa, as células deverão ser lavadas pelo menos uma vez num tampão neutro de baixo teor salino como seja 0,9% NaCl em fosfato 10-50

mM, pH 7,0-7,5 (tampão de fosfatos salino, PBS) para remover o restante meio de crescimento.

A lise das células lavadas é também conseguida num tampão de lise neutro ou ligeiramente ácido de baixa concentração salina, de preferência à volta de pH 6,8, contendo pelo menos um inibidor de proteases, como seja aprotinina ou fluoreto de fenilmetilsulfonilo (PMSF), de preferência contendo uma combinação de tais inibidores. Sais como os de potássio (na forma de KCl) e magnésio (na forma de $MgCl_2$) são também preferencialmente adicionados. Ainda, detergentes não iónicos tais como NP40 (0,5% v/v) e Na-desoxicolato (0,1%) deverão ser adicionados.

A lise da membrana celular externa deverá ser realizada em condições que lisem a célula hospedeira sem lisar o núcleo ou induzir perdas significativas através da membrana nuclear. As células podem ser deixadas em repouso durante um pequeno período, por exemplo, 10 minutos, no tampão de lise contendo detergente antes da intervenção mecânica ser usada no passo de lise. A intervenção mecânica é melhor realizada com disrupção suave das células tratadas com detergente, por exemplo, usando 40 estocadas num homogenizador Dounce com um pistão tipo A ou o equivalente de tal tratamento.

Os núcleos podem ser colhidos a partir da preparação celular lisada usando técnicas conhecidas, como seja por exemplo, centrifugação a 1000xg durante 5 minutos a 4°C e lavados pelo menos uma vez no mesmo tampão de lise com baixa concentração de sal usado para lisar a membrana celular externa.

Os núcleos foram então ressuspensos no tampão de lise de baixa concentração salina que contém ainda DNase I suficiente e incubado durante um tempo suficiente para degradar eficazmente

o DNA de tais núcleos até um tamanho e viscosidade que permite subsequente purificação da c-Myc a partir desta preparação como descrito abaixo.

Após tratamento com DNase I, a amostra foi diluída com um tampão neutro de alta concentração salina o que leva a concentração salina (como NaCl) da amostra para pelo menos 2M. Tal tampão de concentração salina elevada de preferência também contém quantidades de $MgCl_2$ suficientes para manter a mesma concentração deste sal na preparação final diluída e também mais detergente NP40 de forma a reter níveis eficazes após diluição da amostra.

Nas células hospedeiras de mamífero, c-Myc está geralmente estreitamente associado aos núcleos. Assim, é necessário solubilizar c-Myc de forma a não destruir a sua actividade biológica ou a sua capacidade para renaturar numa forma biologicamente activa. O material nuclear residual é primeiro removido por centrifugação e depois o sedimento ressuspenso para solubilização de c-Myc. A solubilização da proteína c-Myc de forma a destruir esta associação pode ser conseguida com dodecilsulfato de sódio (SDS) ou ureia em concentrações superiores a 4M. De preferência utiliza-se ureia 5M. Os núcleos não lisados residuais podem também ser solubilizados nesta altura por agitação forte durante cerca de 30 minutos. A solução é então centrifugada para sedimentar qualquer material insolúvel residual antes de passos cromatográficos subsequentes, por exemplo a 5000xg durante cerca de 10 minutos.

A fracção dosobrenadante recuperado do passo de centrifugação é aplicada a uma coluna de DEAE Sepharose CL-6B equilibrada em tampão contendo ureia como descrito atrás e a coluna bem lavada com tal tampão para remover a proteína não ligada. Uma

segunda lavagem foi realizada com a adição de uma quantidade intermédia de NaCl, 0,1 M NaCl ao tampão. Finalmente, a proteína c-Myc foi eluída aumentando a concentração salina no tampão para 0,35M.

Toda a proteína que eluiu com a lavagem de sal 0,35M foi colhida e aplicada numa coluna de FPLC Mono-Q. A coluna foi lavada e aplicado um gradiente de 0,10 M NaCl a 0,35 M NaCl, seguido de um passo de lavagem com 2M NaCl. Mantendo gradiente em concentrações salinas intermédias, por exemplo a cerca de 0,19 M NaCl até a cauda final da proteína contaminante ter acabado de eluir aumentará a pureza da proteína Myc subsequentemente eluída.

Myc pode ser identificado no eluente da coluna por qualquer técnica que reconheça especificamente a proteína Myc ou a sua actividade. Por exemplo, um anticorpo monoclonal como seja 1F7 pode ser usado num imunoensaio para a presença da proteína Myc. Como alternativa, a actividade de ligação a DNA a um oligonucleótido contendo a sequência 5'-CACGTG-3' pode ser usada para controlar a purificação. O anticorpo monoclonal 1F7 é dirigido contra a sequência peptídica dos aminoácidos 305-317 em c-Myc murino. Outros anticorpos monoclonais contra Myc úteis em tais ensaios podem ser adquiridos comercialmente.

Conjuntos de fracções desta coluna continha as actividades de ligação de C2 e C2' descritas abaixo e a presença de peptídeos capazes de entrar nos hetero-oligómeros C2 e C2' e especialmente hetero-oligómeros C2 e C2', podem ser testados pela capacidade de tais hetero-oligómeros em se ligarem às sequências de DNA 5'-CACGTG-3' e 5'-CAGCTG-3', respectivamente. Myc purificado a partir de células CHO surgiu como várias bandas em imunotransferência.

b. Purificação de c-Myc e seus parceiros a partir de uma fonte de baculovírus

c-Myc humana pode também ser purificada usando o sistema de expressão abundante de Myc em baculovírus. Para a purificação células Sf9 que tinham sido infectadas com baculovírus recombinante portador do gene c-Myc, usando técnicas conhecidas foram colhidas antes da lise (≈ 48 horas pós-infecção). A solubilização e purificação da c-Myc recombinante foram realizadas como para Myc produzido por CHO resultando num rendimento de $2,5 \text{ mg}/8 \times 10^8$ células. Myc purificado a partir destas células de insecto estava aparentemente homogêneo por coloração com prata e migrou numa electroforese como uma banda difusa de $\approx 60 \text{ KD}$. Isto estava em oposição com as múltiplas bandas observadas com Myc de mamífero por imunotransferência (Fig. 1B).

c. Detecção da actividade de ligação a DNA específica de sequência

As preparações atrás contêm duas actividades de ligação a DNA específicas de sequência possuindo as duas a proteína Myc. A primeira actividade contem apenas Myc (i.e., forma o homo-oligómero Myc) e liga-se muito fracamente a sequências com o núcleo CACGTG. A ligação foi testada por determinação a velocidade de exclusão e por ensaios de competição, ambas técnicas conhecidas. A ligação de homo-oligómeros de c-Myc é caracterizada por uma velocidade de exclusão alta e pela observação de que é quase impossível adicionar suficiente sequência competidora fria para retirar por competição este complexo em ensaios de alteração da mobilidade electroforética (EMSA). Esta última observação implica que pode não ser possível aumentar as concentrações de nucleótidos acima do K_D , impossibilitando assim a determinação exacta de

qual a fracção das preparações finais de Myc são activas na ligação específica de sequência pelos homo-oligómeros de Myc.

Um processo de selecção de sítios de ligação pode ser usado para determinar o sítio de ligação óptimo para Myc. Os sítios podem ser seleccionados a partir de um conjunto de oligómeros ao acaso, como sejam 20-meros, de forma a diminuir as hipóteses de erro na determinação de um sítio de ligação óptimo. Pode ser usada uma sequência consensus de 12 bases de GACCACGTGCTC [SEQ ID Nº 1], com a caixa central E de CACGTG surgindo como a mais conservada. Halazonetis e Kandil (Halazonetis e Kandil, Proc. Natl. Acad. Sci. USA 88:6162-6166 (1991)) assumiram que as sequências delimitantes devem ser simétricas e descreveram uma sequência óptima de GACCACGTGGTC [SEQ ID Nº2]. Esta sequência é muito semelhante ao consensus que é preferido aqui, diferindo apenas na posição 10 (onde predominantemente um C foi utilizado no invento, se bem que G esteja significativamente representado Fig. 7, Grupo I). De acordo com o invento, é possível seleccionar uma sequência consensus de 12 bases a partir de um conjunto de complexidade prevista de 4^{20} ($\approx 10^{12}$) indicando assim que Myc tem uma preferência forte relativamente a uma sequência apesar da sua aparente afinidade de ligação fraca.

O segundo complexo de ligação a DNA contendo Myc proporcionado pelas preparações do invento também se liga a sequências com um núcleo de CACGTG, mas liga-se mais fortemente do Myc sózinho. Este complexo (o complexo C2) requiere um factor de 26-29 KD para além de Myc. Este factor adicional copurificou com Myc, presumivelmente devido às propriedades cromatográficas semelhantes e não via associação com Myc pois a cromatografia efectuada em ureia 5M desnaturaria esta associação. Este factor adicional assemelha-se a Max, uma proteína cujo gene foi

recentemente isolado a partir de células de mamífero, por não se ligar eficazmente a DNA por si só mas poder hetero-oligomerizar com Myc para se ligar fortemente à sequência CACGTG. No entanto, o factor do invento difere de Max no seu tamanho aparente (Max está descrito como migrando como 21 KD). Ainda, o hetero-oligómero Myc/Max parece migrar pelo menos tão lentamente como um complexo de apenas Myc em EMSAs, enquanto que o complexo C2do invento migra mais rapidamente do Myc sózinho.

Para além do factor de 26-29 KD, foi identificado um segundo factor de 40-50 KD. Os sítios seleccionados por complexos contendo este factor (aqui designados complexos C2') continham um núcleo CAGCTG (o motivo da sequência μ E2) assim como sequências delimitantes muito semelhantes a um sítio de ligação recentemente descrito para homo-oligómeros de miogenina (Wright *et al.*, Mol. Cell. Biol. 11:4104-4110 (1991)). A miogenina é uma proteína contendo HLH com um peso molecular previsto de 32,5 KD cujo sítio de ligação óptimo é AACAGT/CTGT [SEQ ID Nº3]. Nenhum dos sítios seleccionados (0/36) pelos complexos C2 ou C2' do invento continha um motivo CAGTTG enquanto aproximadamente metade dos sítios seleccionados pela miogenina continham tais sequências.

d. Ensaio para um composto que inibe a acção de Myc

Para facilidade na descrição destes ensaios, a associação do complexo C1 e/ou ligação a DNA, a associação do complexo C2 e/ou a ligação a DNA e a associação do complexo C2' e/ou ligação a DNA são todos referidos como actividade de c-Myc.

Os ensaios para a actividade de c-Myc podem ser realizados in vitro ou in vivo. Os ensaios in vitro podem ser realizados como descrito nos Exemplos, por exemplo avaliando os efeitos do composto pretendido ou várias quantidades de tal composto

sobre os resultados do ensaio de alteração da mobilidade electroforética e de técnicas de selecção de sítios que revelarão se a ligação do oligómero com interesse a um motivo específico de uma sequência de DNA ocorreu na presença do composto.

Para o ensaio in vivo de um composto que inibe a actividade de c-Myc pretendida são utilizadas pelo menos duas construções genéticas. Primeiro é necessário uma construção recombinante capaz de expressar Myc; segundo é necessário um gene repórter cuja expressão está operacionalmente ligada à actividade Myc e especialmente à ligação do oligómero pretendido à sequência de DNA ou motivo específico.

Caso se pretenda, pode também ser usada uma construção recombinante capaz de expressar uma proteína do complexo C2 ou uma proteína do complexo C2'. Como alternativa, pode ser escolhido um hospedeiro que expresse naturalmente tal proteína.

As construções recombinantes que são capazes de expressar a proteína Myc podem ser construídas usando as linhas gerais como descrito abaixo ou adquiridas comercialmente.

A sequência de ligação a DNA pretendida pode estar operacionalmente ligada a qualquer gene que confira uma marca seleccionável no sistema. Numa realização preferida, usa-se um gene de uma marca que permita a selecção fenotípica em levedura e especialmente em Saccharomyces cerevisiae.

A levedura que foi cotransformada com um gene myc expressável e com a sequência de ligação a DNA pretendida pode ser usada para (1) identificar a presença ou ausência de proteínas endógenas hospedeiras que interactuam com Myc num complexo C2 ou C2' (2) classificar uma proteína como uma proteína do

complexo C1 ou como uma proteína do complexo C2'; e (3) identificar e classificar compostos como agentes que destroem tal actividade Myc. As proteínas do complexo C2 foram anteriormente designadas proteínas "parceiras" de Myc.

Estas três aplicações baseiam-se no mesmo principio: na presença de actividade biológica de c-Myc, uma de três coisas acontecerão: formam-se os complexos C1; formam-se os complexos C2; ou formam-se os complexos C2'. Os complexos proteicos assim formados e especialmente os complexos oligoméricos, ligar-se-ão a um motivo de DNA específico, a ligação a tal motivo estará operacionalmente ligado ao gene da marca e a expressão do gene da marca será alterada, de preferência estimulada, como resposta a tal ligação a DNA. Na ausência de de tal oligomerização, a formação do complexo oligómero-DNA não ocorrerá e a expressão da proteína marca não será alterada.

Nos ensaios do invento, pode haver algum nível de ligação à sequência de ligação a DNA pretendida mesmo na ausência de c-Myc. No entanto quando c-Myc está presente na célula, a quantidade e força da ligação específica a DNA é aumentada.

Os hospedeiros que foram cotransformados com um gene c-Myc expressável e com a sequência de ligação a DNA pretendida podem ser usados para testar a presença ou ausência de proteínas hospedeiras endógenas que interactuem com a actividade de c-Myc. Se tais análises revelarem que o hospedeiro contem proteínas de ligação a c-Myc ou oligómeros dependentes de c-Myc que, na presença de c-Myc se ligam especificamente a uma sequência de DNA pretendida, tal proteína parceira de c-Myc ou proteína dependente de oligómero pode ser isolada usando técnicas conhecidas como sejam a análise de alteração da mobilidade em géis, vectores de clonagem de expressão de cDNA, tais como, por exemplo lambda gt10

e lambda gt11 ou outros sistemas de clonagem especificamente destinados a clonagem de alta eficiência e expressão de cDNA de tamanho completo em levedura, como sejam por exemplo pG1 e pTRP56, todos eles podendo ser adquiridos comercialmente (Clontech, Palo Alto, California).

Não é necessário que o hospedeiro seja totalmente deficiente em proteínas do complexo C2 (proteínas parceiras de c-Myc) ou proteínas do complexo C2' para serem úteis no método do invento. Como descrito abaixo, se c-Myc for expresso em níveis muito superiores aos encontrados no hospedeiro, a transcrição do gene repórter a partir de proteínas parceiras endógenas pode ser negligível ou em quantidade tão pequena que não altera a utilidade dos métodos do invento.

Se a expressão de c-Myc for transcrita com um promotor forte e/ou se a cassete de expressão de c-Myc for fornecida num vector de elevado número de cópias, os níveis de c-Myc serão suficientemente altos para ultrapassarem um fundo baixo e tais construções de c-Myc poderão ser usadas para analisar a capacidade das parceiras de c-Myc clonadas para influenciar a ligação a DNA de c-Myc. Os familiarizados com a matéria podem adaptar o sistema de expressão para o nível de expressão pretendido usando métodos conhecidos.

A proteína do complexo C2 (a proteína parceira) ou a proteína do complexo C2', se dada como uma construção recombinante à célula hospedeira, deverá ser expressa em níveis comparáveis aos da proteína c-Myc. As proteínas do complexo C2 podem ser identificadas usando um ensaio de placas fágicas, como descrito no pedido de patente copendente U.S. "Protein Partner Screening Assays and Uses Thereof", Pedido de Patente Nº 510,254, entregue em 19 de Abril, 1990 e aqui incluído como referência. As

proteínas identificadas por tal ensaio de despiste podem ser subclonadas nos vectores de expressão eucarióticos conhecidos e que podem ser adquiridos comercialmente de forma a proporcionar uma fonte recombinante de expressão do gene da proteína parceira.

As construções genéticas do invento podem ser colocadas em diferentes plasmídeos ou combinadas num plasmídeo. Uma construção pode também ser inserida no genoma de uma célula hospedeira. De preferência, a construção codificadora da proteína c-Myc e a construção codificadora da proteína do complexo C2 ou da proteína do complexo C2' são proporcionadas ao hospedeiro em dois plasmídeos diferentes.

É importante estabelecer que o efeito do composto é devido a um efeito na actividade c-Myc e não a um efeito na actividade do produto repórter per se. Tal efeito pode ser estabelecido por comparação dos resultados encontrados em hospedeiros que não têm o vector de expressão de c-Myc ou o vector de expressão da proteína C2 ou C2' ou ambos.

O motivo de ligação a DNA pretendido pode estar situado em qualquer sítio na cassete de transcrição do gene repórter o que permite a transcrição daquele gene operacionalmente ligada à ligação do oligómero pretendido. Assim, tal motivo pode estar situado 5' relativamente ao sítio de iniciação da transcrição ou 3' relativamente ao sítio de iniciação da transcrição, por exemplo, num intrão, semelhante à sua localização relativamente à região do promotor nos genes das imunoglobulinas.

O gene repórter cuja expressão está operacionalmente ligada à actividade c-Myc e especialmente à ligação a DNA do oligómero pode ser qualquer gene cuja expressão possa ser controlada. Qualquer alteração detectável do fenótipo pode servir como

base aos métodos do invento. Numa realização preferida, o gene repórter é um gene que não é normalmente expresso pelo hospedeiro, ou um gene que substitui o gene endógeno do hospedeiro. Pode ser usado qualquer gene repórter que possa estar operacionalmente ligado a um promotor capaz de responder à ligação do oligómero com interesse à sequência de DNA alvo específica.

São especialmente úteis por exemplo genes que transmitam ao hospedeiro capacidade para crescer num meio selectivo. Por exemplo, em levedura, a utilização do gene LEU2 de levedura como gene repórter em estirpes que normalmente não possuam LEU2 permite que tais leveduras cresçam em leucina como única fonte de carbono. A expressão do gene repórter é controlada por mera observação do crescimento do hospedeiro em leucina. De forma semelhante, a utilização do gene suc2 como gene repórter permitirá o crescimento do hospedeiro levedura suc2⁻ em sucrose a ser usado como método de detecção. Em ambos os exemplos, o crescimento no substrato indicado poderá ser usado para indicar a ligação específica a DNA do oligómero com interesse e ausência de tal crescimento poderá ser usado para indicar a ausência de ligação ou ausência da formação de oligómeros.

Num outro exemplo, uma construção (e hospedeiro) que seja gal1⁺gal10⁻ responderá à galactose no meio; uma construção (e hospedeiro) que seja lac2⁺gal1⁺ será sensível à lactose. Outros genes repórteres incluem his3, ura3 e trp5. Os familiarizados com a matéria poderão imaginar muitos outros sistemas repórter adequados que revelem a presença ou inibição da ligação a DNA ou actividade biológica do oligómero com interesse.

As construções repóter em que o motivo da sequência de DNA pretendida e o gene repórter lacZ estão operacionalmente ligados expressarão β -galactosidase como resposta à ligação de

uma actividade c-Myc induzida pela ligação de oligómeros a tal sequência de DNA. Tal expressão poder ser facilmente avaliada controlando a capacidade do hospedeiro para produzir β -galactosidase (Maniatis, T. et al., Molecular Cloning (A Laboratory Manual), 2ª edição, Cold Spring Harbor Laboratory, 1989). A produção de β -galactosidase pode ser controlada visualmente pela detecção da sua actividade de redução do corante cromofórico X-gal (comercialmente disponível na International Biotechnologies, Inc., New Haven, CT). A β -galactosidase reduz X-gal para uma forma que possui uma cor azul. Numa outra realização, a sequência codificadora da cloranfenicol-acetiltransferase (CAT) é usada como gene repórter.

Pode ser usado qualquer método de detecção que possa identificar a expressão do gene repórter. Por exemplo, os níveis do produto do gene repórter podem ser directamente testados com um imunoensaio. Tais imunossaios incluem aqueles em que o anticorpo está numa fase líquida ou ligado a um veículo de fase sólida. Em adição, o gene repórter pode ser detectavelmente marcado de várias formas para usar em imunossaios. Os imunossaios preferidos para a detecção de uma proteína repórter incluem ensaios imunossorventes ligados a enzimas (ELISA) ou outros ensaios conhecidos, como sejam ensaios imunofluorescentes, ensaios quimioluminescentes ou bioluminescentes.

Num ensaio para despistar a capacidade de um composto para alterar a ligação do oligómero com interesse, estirpes de levedura que expressam tal peptídeo ou peptídeos pretendidos e que contêm o motivo da sequência de ligação a DNA relacionada, podem ser semeadas e crescidas como tapetes e o composto a ser testado pode ser aplicado sobre as placas num disco de papel de filtro que está impregnado com tal composto. Como alternativa, o

composto pode ser incorporado no meio em que as células hospedeiras estejam a crescer.

Pode-se detectar a capacidade de um composto para alterar a actividade de c-Myc pelo aparecimento de uma zona, a qual muitas vezes se assemelha a um halo, à volta de um disco impregnado de composto. Se por exemplo, o composto for tóxico para a sobrevivência do hospedeiro per se, o hospedeiro não crescerá na zona contendo o composto.

Os métodos do invento podem ser usados para fazer o despiste compostos na sua forma pura, numa variedade de concentrações e também na sua forma impura. Os métodos do invento podem também ser usados para identificar a presença de tais inibidores em extractos brutos e para seguir a purificação dos seus inibidores. Os métodos do invento são também úteis para avaliar a estabilidade dos inibidores identificados como atrás, para avaliar a eficácia de várias preparações.

A permeabilidade das células a vários compostos pode ser aumentada, se necessário, pela utilização de uma estirpe celular mutante que possui uma maior permeabilidade ou pela utilização de compostos que sabe aumentarem a permeabilidade. Por exemplo, em levedura compostos tais como o nonapeptídeo polimixina B podem ser usados para aumentar a permeabilidade de levedura a pequenos compostos. Em células dos eucariotas superiores, o dimetilsulfóxido (DMSO) pode ser usado para aumentar a permeabilidade. Os análogos de tais compostos que sejam mais permeáveis através das membranas de levedura podem ser igualmente usados. Por exemplo, os derivados dibutirilo muitas vezes apresentam uma maior permeabilidade.

Numa realização preferida, as construções genéticas e os métodos para os usar são utilizados em hospedeiros eucarióticos e especialmente em levedura, células de insecto e células de levedura. A sequência introduzida é incorporada num plasmídeo ou vector capaz de replicação autónoma ou de actividade integradora.

A sequência de c-Myc é conhecida (Battey, J. et al., Cell 34:779-787 (1983)) e sondas que sejam capazes de identificar um clone de c-Myc podem ser adquiridas comercialmente (New England Nuclear/Dupont Biotechnology, Boston, MA).

A sequência de DNA do gene pretendido pode ser construída por síntese química caso não se pretenda usar um clone do genoma ou mRNA como fonte da informação genética. Os métodos de síntese química do DNA são bem conhecidos (Oligonucleotide Synthesis, A Practical Approach, M.J. Gail, ed. IRL Press, Washington, D.C., 1994; Synthesis and Applications of DNA and RNA, S.A. Narang, ed. Academic Press, San Diego, CA, 1987) Devido ao código genético ser degenerado mais de um codão pode ser usado para construir a sequência de DNA codificadora de um aminoácido particular (Watson, J.D., Em: Molecular Biology of the Gene, 3ª edição, W.A. Benjamin, Inc., Menlo Park, CA, 1977, pp. 356-357).

Para expressar as construções recombinantes do invento, são necessários sinais de transcrição e tradução reconhecíveis pelo hospedeiro. Uma sequência clonada codificadora de uma proteína, obtida através dos métodos descritos atrás, (de preferência numa forma de cadeia dupla), pode ser operacionalmente ligada a sequências que controlem a transcrição num vector de expressão e introduzida por exemplo por transformação numa célula hospedeira para produzir proteínas recombinantes úteis nos métodos do invento, ou seus derivados funcionais. Tais técnicas são bem conhecidas (Recombinant DNA Technology, Wu, R. et al.,

eds., Academic Press, (1989); Maniatis, T. et al., Molecular Cloning (A Laboratory Manual) segunda edição, Cold Spring Harbor Laboratory, 1989).

Os sinais reguladores da iniciação da transcrição podem ser seleccionados de forma a permitir a expressão ou activação da expressão da construção c-Myc ou da construção do peptídeo do complexo C2 recombinante (ou do peptídeo C2'), ou de ambos, de forma a que a expressão de tais construções possam ser moduladas caso se pretenda. São particularmente interessantes os sinais reguladores que sejam sensíveis à temperatura de forma a que ao variar-se a temperatura, a expressão possa se reprimida ou iniciada ou que sejam sujeitos a regulação química, por exemplo por um metabolito, sal ou substrato adicionado ao meio de crescimento.

Sempre que as sequências de controle da expressão nativas não funcionem satisfatoriamente na célula hospedeira, então podem ser substituídas por sequências funcionais na célula hospedeira.

A expressão das construções do invento em diferentes hospedeiros pode resultar em diferentes modificações pós-tradução que possam alterar as proteínas expressas por estas construções. É necessário expressar as proteínas num hospedeiro em que a capacidade da proteína para reter a sua função biológica não seja alterada. A expressão das proteínas em hospedeiros levedura é de preferência conseguido usando sinais reguladores de levedura. Os vectores do invento podem conter elementos reguladores operacionalmente ligados tais como sequência activadoras a montante em levedura ou elementos de DNA que confirmam expressão específica de espécie, tecido ou tipo celular num gene operacionalmente ligado.

Em geral os vectores de expressão contendo sequências reguladoras da transcrição, como sejam sequências de promotores e sequências de terminação da transcrição são usadas em ligação com um hospedeiro. Estas sequências facilitam a transcrição eficaz do fragmento de gene que lhe esteja operacionalmente ligado. Em adição, os vectores de expressão também contêm tipicamente elementos de DNA discretos tais como por exemplo (a) uma origem de replicação que permita a replicação autónoma do vector ou elementos que promovam a inserção do vector no cromossoma do hospedeiro de forma estável e (b) genes específicos que sejam capazes de proporcionar selecção fenotípica em células transformadas. Os vectores de expressão eucarióticos podem também conter elementos que lhe permitam ser mantidos nos hospedeiros procarióticos; tais vectores são conhecidos como vectores vai-vem.

A natureza precisa das regiões reguladoras necessárias para a expressão do gene variarão entre as espécies ou tipos celulares e existem muitos sistemas vectores de expressão adequados que podem ser adquiridos comercialmente.

Numa realização altamente preferida, as leveduras são usadas como células hospedeiras. Os elementos necessários à transcrição de um gene em levedura foram recentemente revistas (Struhl, K. Ann. Rev. Biochem. 58:1051-1077 (1989)). Em levedura, a maior parte dos promotores contem três elementos de DNA básicos: (1) uma sequência activadora a montante (UAS); (2) um elemento TATA; e, (3) um elemento de iniciação (I). Alguns promotores podem conter elementos operadores. Os métodos usados em genética de leveduras são bem conhecidos (Struhl, K. Nature 305:391-397 (1983); Sherman, et al., Methods in Yeasts Genetics, Cold Spring Harbor Laboratory (1983)).

Numa outra realização as células de mamífero são usadas como células hospedeiras. Uma larga variedade de sinais reguladores da transcrição e da tradução podem ser derivados para a expressão das proteínas em células de mamífero e especialmente a partir de sequências genômicas de vírus que infectem células eucarióticas.

Uma vez preparado o vector ou sequência de DNA contendo a ou as construções para expressão, a ou as construções de DNA são introduzidas numa célula hospedeira adequada por qualquer uma de uma variedade de métodos adequados, por exemplo por transformação. Após a introdução do vector, as células recipientes são crescidas num meio selectivo o qual selecciona o crescimento de células contendo o vector. A expressão das sequências do gene clonado resulta na produção da proteína. Esta expressão pode ter lugar numa forma contínua nas células transformadas ou numa forma controlada.

Pode-se construir transformantes geneticamente estáveis com sistemas vectores episomais ou com sistemas vectores integrados pelo que o DNA da proteína de fusão é integrado no cromossoma do hospedeiro. Tal integração pode ocorrer de novo dentro da célula ou ser auxiliada pela transformação com um vector que se insere funcionalmente no cromossoma do hospedeiro, por exemplo, com vectores retrovirais, transposões ou outros elementos de DNA que promovam a integração de sequências de DNA nos cromossomas.

As células que foram transformadas com os vectores de DNA do invento, são seleccionadas também através da introdução de uma ou mais marcas que permitam a selecção das células hospedeiras que contêm o vector, por exemplo, a marca pode proporcionar resistência a biocidas, e.g., resistência a antibióticos ou a metais pesados, tais como cobre ou similares.

A célula hospdeira transformada pode ser fermentada ou cultivada de acordo com meios conhecidos para se conseguir crescimento celular óptimo e também para se conseguir expressão óptima dos fragmentos das sequências proteicas clonadas. Como aqui descrito um nível elevado de expressão de proteína recombinante para as sequências clonadas das proteínas pode ser conseguido de acordo com um processo preferido deste invento.

Os métodos do invento não se destinam a ser limitados a c-Myc e possuem utilidade para a caracterização de inibidores contra qualquer proteína Myc, como seja por exemplo, N-Myc e L-Myc. Os peptídeos do complexo C2 do invento podem interactuar com mais de uma proteína Myc e os peptídeos do complexo C2' do invento podem formar-se como resultado da actividade de mais de uma proteína Myc.

Os exemplos que se seguem descrevem ainda os materiais e métodos usados na realização do invento. Os exemplos não se destinam a limitar de forma alguma o invento.

EXEMPLOS

EXEMPLO 1

Materiais e Métodos

Crescimento celular e expressão abundante de Myc: A linha celular 5A foi mantida em culturas em suspensão sob seleção com metotrexato 80 μM . A purificação da proteína começou com aproximadamente 6 litros de células a $8 \times 10^5/\text{ml}$ crescidas sem seleção. A indução do promotor de choque térmico foi conseguida por ressuspensão e meio fresco previamente aquecido (43°C) a $1/3$ do volume original. As células foram incubadas com agitação a 43°C durante 1 hora. Para permitir a tradução do mRNA acumulado, as células foram transferidas para condições de cultura a 37°C durante 3 horas. As células foram então sujeitas à purificação descrita abaixo.

O vector de expressão em baculovírus foi construído por inserção do fragmento BamHI/BclII de pGEMMycB [Halazonetis e Kandil, Proc. Natl. Acad. Sci. USA 88:6162-6166 (1991)] no sítio BamHI de um vector de expressão para baculovírus, pVL941, obtido do laboratório de Max Summers (Texas A&M University, College Station, Texas). O plasmídeo resultante continha toda a sequência codificadora do gene Myc humano incluindo 6 nucleótidos 5' relativamente ao codão de iniciação e 3' relativamente à sequência não traduzida estendendo-se até ao sítio genómico RsaI. As células Sf9 foram cultivadas e infectadas com baculovírus recombinante de acordo com os métodos de Summers [Summers e Smith, A Manual of Methods for Baculovirus Vectors and Insect Cell Culture Procedures, Texas Agricultural Experiment Station Bulletin Nº 1555] com pequenas alterações. As células foram passadas em cultura em suspensão e semeadas em placas de cultura

de tecidos de 150 mm de diâmetro para a produção de proteína. As células foram infectadas e colhidas aproximadamente 48 horas pós infecção por raspagem. As células foram então lavadas em PBS e sujeitas à purificação descrita abaixo.

A proteína de fusão Proteína A-c-Myc foi expressa na estirpe AR68 de E. coli a partir de um vector anteriormente publicado pRIT2T [Dang, C.V., Anal. Biochem. 174:313-317 (1988)] que fundiu a fracção de ligação a Ig da proteína A aos aminoácidos 353-439 ou aminoácidos 372-439 de c-Myc. O crescimento e indução das células foi como descrito por Dang et al. [Anal. Biochem. 174:313-317 (1988)].

Purificação da proteína: Todos os passos de purificação foram realizados em gelo ou com tampões arrefecidos em gelo a menos que de outra forma seja estabelecido. As células podem ser usadas frescas ou guardadas por arrefecimento rápido em azoto líquido para as preparações em grandes lotes. As células 5A ou Sf9 foram lavadas em tampão de fosfatos salino (PBS) e ressuspensas a $2,1 \times 10^7$ células/ml em Tampão de Lise de Baixa Concentração Salina (20 mM HEPES pH 6,8, 5 mM KCl, 5 mM MgCl₂, 0,5% NP40, 0,1% Na-desoxicolato, 1 µg/ml de aprotinina e 0,1 mM PMSF) [Evan e Hancock, Cell 43:253-261 (1985)]. Após 10 minutos as células foram sujeitas a 40 estocadas num homogenizador Dounce com um pistão tipo A. Os núcleos foram sedimentados a 1000xg, 5 min, 4°C, lavados uma vez em 50 ml de Tampão de Lise de Baixa Concentração Salina, ressuspensos a $2,5 \times 10^8$ núcleos/ml em Tampão de Lise de Baixa Concentração Salina contendo 50 µg/ml de DNase I e incubados a 4°C durante 1 hora. Um volume igual de Tampão de Alta Concentração Salina 2x arrefecido em gelo (concentrações 2x: pH 7,4, 4M NaCl, 1 mM MgCl₂, e 0,1% NP40) [Evan e Hancock, Cell 43:253-261 (1985)] foi então adicionado, misturado suavemente e incubado durante 10 minutos. O material nuclear residual

(incluindo a proteína c-Myc) foi sedimentado (2000xg, 10 minutos, 4°C) e ressuspensão para solubilização a $5,5 \times 10^7$ equivalentes de núcleos/ml em Tampão A (50 mM Tris, pH 8,0, 2 mM EDTA, 5% glicerol, 0,1 mM DTT, e 0,1 mM PMSF). [Watt et al., Mol. Cell. Biol. 5:448-456 (1985)] contendo ureia 5M (referido como tampão A ureia 5M) conseguido por diluição de um stock de ureia 6 M recentemente desionizada. Este e todos os tampões usados nas colunas foram passados através de unidades de filtração de poros de 0,2 μ m. Os núcleos residuais foram solubilizados por agitação forte em gelo durante 30 minutos. Esta solução de proteína foi centrifugada (10 minutos, 5000xg, 4°C) para sedimentar qualquer material insolúvel antes da cromatografia. O sobrenadante foi aplicado numa coluna de 10 ml de DEAE Sepharose CL-6B (Pharmacia) equilibrada com 5 volumes da coluna de Tampão A ureia 5M. A aplicação da amostra foi a 0,1 ml/min e a lavagem da coluna e eluição foram a 0,4 ml/min. Após aplicação, a coluna foi lavada com 3 volumes de Tampão A ureia 5M não contendo mais sal seguido de 4 volumes do mesmo tampão contendo 0,1M NaCl. A proteína Myc foi eluída no passo de eluição seguinte a 0,35M NaCl. As fracções contendo proteína deste passo de 0,35 M NaCl foram reunidos e diluídos com Tampão A ureia 5M para 0,1M NaCl e aplicado numa coluna Mono-Q de FPLC (Pharmacia) corrida a 0,5 ml/min. A coluna de Mono-Q foi eluída com um gradiente programado de 5 ml indo de 0,10 M NaCl até 0,35M NaCl seguido de um passo de 2M NaCl. Para uma maior pureza o gradiente foi mantido manualmente a aproximadamente 0,19 M até a principal proteína contaminante acabar de eluir conforme determinado por um monitor de UV ligado em série. No desenvolvimento inicial do protocolo de purificação as fracções das colunas foram testadas relativamente a Myc por transferência em barras transversais ("slot blotting") seguido de visualização usando o anticorpo monoclonal 1F7 e o segundo anticorpo marcado com ^{125} I. Para outras preparações a coloração com prata de SDS-PAGE permitiu a identificação suficientemente clara das proteínas Myc e

permitiu avaliar a pureza das fracções. As fracções contendo Myc foram reunidas com base na pureza e dialisadas contra tampão contendo 20 mM Tris pH 7,8, 50 mM KCl, 10% glicerol, 0,1 mM DTT e 0,1 mM PMSF (referido como Tampão de Diálise) em sacos membrana SpectroPor 2 para 3 mudas, 2 litros cada, durante um mínimo de 3 horas cada. Os conjuntos de fracções preparados desta forma continham actividades de ligação de C1 e C2 (e C2'). Para se obter actividade de ligação C1 pura as fracções contendo Myc da coluna de MnoQ foram testadas por EMSA e as que não possuíam actividade de ligação C2 foram reunidas e dialisadas separadamente.

A proteína de fusão Proteína A-c-Myc produzida por bactérias foi parcialmente purificada por centrifugação diferenciada e solubilizada em ureia 5M de acordo com Watt et al. [Bagchi et al., Mol. Cell. Biol. 7:4151-4158 (1987)] com as seguintes pequenas modificações: Inibidores de proteases estavam presentes no tampão de lise inicial (10 µg/ml de pepstatina, 1 mM PMSF, 50 µg/ml de aprotinina, 2 µg/ml de leupeptina, 10 mM Na-metabissulfito e 1 mM benzamidina) e as células foram destruídas por 6 pulsos de 15 s cada num Cuisinart MiniMate em gelo. O material solubilizado com ureia foi separado do material insolúvel por centrifugação (10000 xg, 10 min., 4°C) e dialisado contra Tampão de Diálise contendo 0,5 mM DTT. O material precipitado foi removido por centrifugação (15000xg, 20 minutos, 4°C). A proteína de fusão Proteína A-Myc foi purificada a partir do sobrenadante por afinidade com Ig essencialmente de acordo com Nilsson et al. [EMBO J. 4:1075-1080 (1985)]. Uma amostra de 1 ml de sobrenadante foi incubado com 0,1 ml de uma pasta a 50% de IgG Sepharose 6 fluxo rápido (Pharmacia) agitando durante 1 hora a 4°C. O sedimento foi lavado duas vezes com Tampão A e a proteína de fusão eluída com diiodo-salicilato de lítio 0,3 M (LIS). O eluato foi então dialisado extensivamente para remover LIS (inicialmente

contra Tampão A à temperatura ambiente para evitar a precipitação de LIS, depois contra Tampão de Diálise 4°C). As duas preparações de Myc expressas por bactérias foram comparadas por coloração com azul Coomassie de SDS-PAGE para assegurar que quantidades iguais das proteínas de fusão foram usadas nas experiências.

Sequenciação N-terminal: As 3 bandas de Myc purificada a partir de células 5A foram individualmente isoladas por electroeluição de acordo com Hunkapiller et al. [Meth. Enz. 91:227-236 (1983)] SDS-PAGE preparativo foi realizado e as bandas de proteína removidas após visualização com Coomassie Brilliant Blue R-250. Após electroeluição o material foi precipitado 2 vezes com metanol/acetona e submetido a sequenciação N-terminal por degradação de Edman.

Anticorpos: O anticorpo monoclonal 1F7 (uma oferta generosa de R. Chizzonite, Hoffman LaRoche), é dirigido contra a sequência peptídica compreendendo os aminoácidos 305-317 em c-Myc murina [Miyamoto et al., Proc. Natl. Acad. Sci. USA 82:7232-7236 (1985)]. O anticorpo dirigido contra cI foi o anticorpo monoclonal 51F [Breyer e Sauer, J. Biol. Chem. 264:13348-13354 (1989)] o qual foi purificado por precipitação com sulfato de amônio e cromatografia em QAE Sephadex.

Ensaio de Alteração da Mobilidade Electroforética (EMSA): Sondas marcadas radioativamente foram produzidas via preenchimento com Klenow dos oligonucleótidos emparelhados contendo 4 bases nos extremos protuberantes 5' (ver tabela abaixo para as sequências). As reacções de ligação tiveram lugar num volume final de 20 µl contendo 2 ng de sonda marcada, 125 ng poli d(IC), uma quantidade indicada de proteína e as seguintes condições finais de tampão: 10 mM Tris, pH 7,5, 50 mM KCl, 0,1 mM EDTA, 1 mM DTT, 1 mM MgCl₂ e 5% de glicerol. As reacções de

ligação prosseguiram durante 20 minutos à temperatura ambiente e foram então aplicadas imediatamente num gel de 4% de poliacrilamida o qual tinha sido previamente corrido pelo menos 1 h a 10V/cm. A electroforese decorreu durante 1,5 hora a 10 V/cm em 0,5x TBE.

Corte e Renaturação: O método de Bagchi et al. [Bagchi et al., Mol. Cell. Biol. 7:4151-4158 (1987)] foi seguido excepto no passo final de diálise. As amostras de proteína precipitada contendo BSA como proteína veículo foram solubilizadas em 6M hidrocloreto de guanidina (200 μ l a não ser que de outro modo seja indicado) de acordo com Bagchi et al. [Mol. Cell. Biol. 7:4151-4158 (1987)]. Directamente antes da análise por EMSA as amostras foram sujeitas a diálise sózinhas ou em combinação com uma outra amostra num volume total de 15 μ l. Volumes iguais de cada amostra foram usados numa determinada experiência e o volume foi levado a 15 μ l usando HGuHCl 6M contendo 0,1 mg/ml de BSA. A diálise foi contra 40 ml de Tampão de Diálise durante 1 h a 4°C em discos de membrana de 13 mm flutuantes (Millipore #VSWP-013, poro 0,025 μ m).

Seleção de Sítios a partir de Sequências ao Acaso: O processo que se segue foi pensado com base no método de Pollock e treisman [Nucl. Acids Res. 18:6197-6204 (1990)]. Um oligómero de 52 bases "randomer" (ver tabela abaixo) foi emparelhado com o seguinte iniciador de 16 bases: iniciador Xho I 5' CCGATATCTCGAGACGG3', [SEQ ID N°4]. O iniciador emparelhado foi prolongado usando Klenow e nucleótidos (0,2 mM dNTPS frios e 0,4 μ M α ³²P-dCTP 800 Ci/mmol) para criar um conjunto de sondas de cadeia dupla representando aproximadamente 4²⁰ sequências. O ciclo inicial de selecção de sítios de ligação por EMSA usou 200 ng deste conjunto e 0,37 μ g de c-Myc produzido por baculovírus ou 0,5 μ g de c-Myc produzido por CHO. Os outros parâmetros foram

como anteriormente descrito para EMSA. As faixas contendo sondas ao acaso foram alternadas com faixas de referência contendo 2 ng de sonda (USE)₃ e 0,37 µg de c-Myc de baculovírus. O gel de EMSA completo foi electrotransferido para membrana NA45 (200 mA, 2,5 horas) e a membrana molhada foi embrulhada em plástico e exposta durante pelo menos 1,5 horas. As regiões das faixas de "randomer" correspondendo aos complexos visíveis C1 e C2 das faixas de referência foram retirados e eluídos com 100 µl de solução de eluição (10 mM Tris, pH 8,0, 0,1 mM EDTA, 1M NaCl) 30 minutos a 68°C. O líquido foi transferido para um tubo novo e a membrana foi lavada com 100 µl de TE que foi adicionado a este eluato. Após sedimentação dos detritos, o DNA foi precipitado com a adição de 10 µg de glicogénio, 2 µl de 1M MgCl₂ e 2,5 volumes de etanol. O sedimento foi lavado com etanol 70%, seco e a recuperação avaliada num contador de cintilação. A totalidade do sedimento de cada amostra (≈29-57 pg) foi ressuspensa em 10 µl de tampão PCR 10x (500 mM KCl, 100 mM Tris, pH 8,4 1 mg/ml de gelatina, 15 mM MgCl₂) e 32 µl de água. Após adição de 1 µl do iniciador Xho I 100 µM e do iniciador Xba I (5' GGACGATCTAGATTCG3', SEQ ID N°5]), 5 µl de mistura de nucleótidos (2 mM dNTPs e 4 µM α³²P-dCTP 800 Ci/mmol) e 1 U de polimerase Taq as reacções foram cobertas com óleo de parafina e sujeitas a 20 ciclos de PCR numa máquina Ericomp: 2 min 94°C, 20x (15 seg 95°C, 15 seg 55°C), 10 minutos 72°C. Os produtos foram purificados em gel em 10% de acrilamida e precipitados usando 10 µg de glicogénio como veículo. A recuperação foi medida num contador de cintilação e após ressuspensão no tampão de reacção de EMSA (10 mM Tris, pH 7,5, 50 mM KCl, 1 mM EDTA, 1 mM MgCl₂ e 5% glicerol) esta sonda foi usada para um novo ciclo de selecção por EMSA. Os ciclos subsequentes foram basicamente como este, no entanto, usou-se 50 ng de sonda. Completaram-se oito ciclos de selecção e amplificação para c-Myc de baculovírus e sete ciclos para c-Myc de CHO. Após a reacção de PCR final os produtos foram extraídos duas

vezes com fenol e duas vezes com éter e precipitados antes da digestão com XhoI e XbaI. Após isolamento em gel o fragmento adequado foi subclonado no vector Bluescript SK (Stratagene) e sequenciado por processos convencionais.

Oligonucleótidos Usados: As sequências de oligonucleótidos que foram usadas estão apresentadas abaixo, com as sequências da caixa E sublinhadas:

SEQ ID NO. 6:

(μ e2)₃ 5'GATCTCTGCAGCAGCAGCTGGCAGCAGCTGGCAGCAGCTGGCG3';

SEQ ID NO. 7:

(μ e3)₃ 5'GATCTGCAGTCATGTGGCGTCATGTGGCGTCATGTGGGCAG3';

SEQ ID NO. 8:

(USE)₃ 5'GATCTGCAGTCACGTGGCGTCACGTGGCGTCACGTGGCAG3';

SEQ ID NO. 9:

MLC-A 5'TCGACGTCGCAGCAGGTTGCAG3';

SEQ ID NO.10:

MLC-B 5'TCGACCCACCAGCTGGCGAG3';

SEQ ID NO.11:

ERP1/2 5'AGCTTCGAACACCTGCAGCAGCTGGCAGGAGCAGGCCTA 3';

SEQ ID NO.12:

ERP3/4 5'AGCTTTAAATCCCCACCAGCTGGCGAAGCAACAGGTGCA3';

SEQ ID NO.13:

HSE 5'AATTGCGAAACCCCTGGAATATTCGACCTGGCAGCCTC3';

SEQ ID NO.14:

SMS 5'TCGACTTTAGACCAGTGGTCCCCTCGA3';

SEQ ID NO.15:

"Randomer"5'GGACGATCTAGATTCG(n)₂₀CCGTCTCGAGTATCGG3'.

EXEMPLO 2

Purificação da Proteína c-Myc

Um dos principais objectivos deste trabalho foi purificar e caracterizar Myc derivada de uma fonte de mamífero. Foi utilizado um sistema de mamífero para expressão abundante anteriormente descrito (Wurm et al., Proc. Natl. Acad. Sci. 83:5414-5418 (1986)). Resumidamente, os dois exões codificadores do gene c-Myc de murganho sob o controle de um promotor de choque térmico de Drosophila foi integrado e amplificado no genoma de uma linha celular de ovário de hamster chinês (CHO). Esta linha celular de expressão abundante, 5A, foi adaptada à cultura em suspensão. O choque térmico (43°C) induz a transcrição dos genes myc amplificados enquanto que um período de recuperação subsequente de 2 horas à temperatura normal de crescimento (37°C) permite a tradução. Os produtos resultantes foram fosfoproteínas de 60, 62 e 72 KD que foram imunoprecipitadas com anticorpos monoclonais específicos contra Myc (Wurm et al. Proc. Natl. Acad. Sci. USA 83:5414-5418 (1986)). A c-Myc produzida estava fortemente associada ao núcleo a tentativas para a solubilizar usando um número de detergentes, sais e agentes redutores foram infrutíferos (resultados não apresentados). Conseguiu-se solubilização significativa com SDS ou com ureia em concentrações superiores a 4M. Para purificação, a Myc foi solubilizada com ureia 5M e cromatografada em resina DEAE e FPLC Mono-Q como descrito em materiais e métodos. A presença de Myc nas fracções da coluna foi testado por imunotransferência usando um anticorpo monoclonal anti-peptídeo, 1F7 (Miyamoto et al., Proc. Natl. Acad. Sci. USA 82:7232-7236 (1985)). Este processo de purificação deu 150 µg de c-myc por litro de células em suspensão (8×10^8 células). O Myc surgiu como 95% homogêneo avaliando por coloração com prata (Fig. 1A).

Um sítio de iniciação da tradução alternativo para c-myc poderá explicar alguma da heterogeneidade molecular de c-Myc traduzido in vitro e expresso em várias linhas celulares (Hann et al., Cell 52:185-195 (1988)). Este sítio alternativo está a montante do sítio de iniciação canónico, no entanto não está presente no nosso gene abundantemente expresso. A análise de sequência N-terminais de cada uma das três bandas proeminentes descritas atrás revelou conforme esperado a sequência prevista pelo sítio de iniciação canónico (resultados não apresentados), se bem que a metionina N-terminal não estivesse presente, presumivelmente devido ao processamento N-terminal. Portanto as diferenças potencialmente importantes no peso molecular aparente que são observadas devem ser atribuídas a modificações após tradução e não a heterogeneidade N-terminal.

C-Myc humana foi também purificada usando o sistema de expressão abundante em baculovírus. Para purificação, as células Sf9 que foram infectadas com vírus recombinante foram colhidas antes do estabelecimento da lise (≈ 48 horas após infecção). Myc produzido usando o sistema baculovírus foi anteriormente descrito como sendo fosforilado e como estando fortemente associado ao núcleo (Miyamoto et al., Mol. Cell. Biol. 5:2860-2865 (1985)).

A solubilização e purificação foram realizadas com Myc produzido em células CHO resultando num rendimento de 2,5 mg/ 8×10^8 células. Myc purificado a partir destas células de insecto estava aparentemente homogêneo por coloração com prata e migrou como uma só banda difusa de ≈ 60 KD (Fig. 1B). Isto está em oposição com as múltiplas bandas observadas com Myc de mamífero por imunotransferência (Fig. 1B).

DISCUSSÃO:

Myc foi purificado até quase à homogeneidade a partir de células de mamífero expressando abundantemente e a partir de células infectadas com baculovírus. A proteína derivada de mamífero parece estar altamente modificada ao contrário de Myc expresso e purificado a partir de células de insecto. Até 19 espécies distintas de c-Myc podem ser identificadas por electroforese bidimensional (Fig. 1). Estas espécies diferem em tamanho (MRs aproximadas de 60000, 62000 e 72000, apesar desta estimativa de tamanho poder variar com diferentes condições do gel) e no pI. Estas diferenças no pI devem em parte ser atribuídas a diferenças na fosforilação, como se sabe que c-myc é fosforilado e a alteração no pI da espécie é consistente com adições de fosfato. Se bem que Myc produzido pelo sistema de expressão de baculovírus não apresente a mesma heterogeneidade de pesos moleculares da proteína de mamífero (Miyamoto *et al.*, Mol. Cell. Biol. 5:2860-2865 (1985)). Os sítios específicos de fosforilação não foram determinados para a preparação de Myc e outras modificações ainda não identificadas podem distinguir estas duas preparações de Myc.

EXEMPLO 3

Actividade Específica de Ligação a DNA Presente em c-Myc Purificado

A presença de um domínio B-HLH em c-Myc sugeriu que ela se ligaria a uma sequência tipo caixa E do padrão geral CANNTG. Estes sítios foram primeiro identificados em potenciadores de imunoglobulinas mas têm sido encontrados em muitos outros potenciadores específicos de tecidos. Determinou-se primeiro se qualquer um destes seria ligado por proteínas c-Myc purificadas descritas no Exemplo 2. Um grande número de sequências

relacionadas com a caixa E foram encontradas por ensaios de alteração da mobilidade electroforética (EMSA). Os apresentados na Fig.2 incluem oligonucleótidos sintéticos contendo trímeros dos sítios μ E2 (CAGCTG) ou μ E3 (CATGTG) do potenciador da imunoglobulina e um trímero do elemento a montante (USE) do principal promotor tardio dos Adenovírus (CACGTG). Também são apresentados dois sítios do potenciador da cadeia leve da miosina (MLC): o sítio A (CAGGTG) que se assemelha ao sítio potenciador da imunoglobulina μ E2 e o sítio B (CAGCTG) que tem a mesma sequência nuclear do sítio μ E2. O elemento de choque térmico (HSE) serviu como testemunha uma vez que a sua sequência não se assemelha a um núcleo da caixa E.

Foram detectadas três actividades de ligação específicas neste ensaio formando complexos referidos como C1 (específico de USE), C2 (específico de USE) e C2' (específico de μ E2). Conforme demonstrado abaixo, a pesar da comigração de C2 e de C2', estes representam complexos separados baseados nas diferenças observadas na composição de proteínas assim como especificidade de ligação. Os resultados apresentados defendem que os complexos C1 sejam formados por homo-oligómeros de Myc enquanto que a formação dos complexos C2 e C2' requerem uma proteína adicional. O complexo que migra lentamente (C1) formou-se mais rapidamente no USE (Fig. 2, faixas 5, 11 e 12), menos bem no sítio semelhante μ E3 (Fig. 2, faixas 2 e 8) e não se formou de todo noutros sítios testados da caixa E e sem serem da caixa E. Preparações de Myc derivadas de CHO e de báculo vírus surgiram semelhantes relativamente ao complexo C1, no entanto eles diferiram relativamente aos complexos de migração mais rápida. Nos ensaios de Myc mamífero o complexo C2' formou-se no sítio μ E2 do potenciador da imunoglobulina e na sequência tipo μ E2 do sítio MLC-B (Fig. 2, faixas 1 e 4). Myc de báculo vírus não continha actividade de ligação com esta especificidade (Fig. 2, faixas 7 e

10). Pelo contrário, a formação do complexo C22 foi detectada usando qualquer uma das preparações de Myc. O complexo C2 formou-se mais rapidamente no sítio USE (Fig. 2, faixas 5, 11 e 12) e menos bem na sequência semelhante $\mu F3$ (Fig. 2, faixas 2 e 8). Detectou-se pouca ou nenhuma ligação nas sequências tipo KE2 (MLC-A Fig. 2, faixas 3 e 9 e $\pi\mu E5$, resultados não apresentados). Não se encontrou ligação específica nas sequências que não fossem caixa E como seja a HSE (Fig. 2, faixa 6 e 13).

As experiências de competição foram realizadas nas três actividades de ligação C1, C2 e C2' para melhor caracterizar a sua especificidade (resultados não apresentados). Nas experiências usando as sequências $\mu E2$, $\mu E3$, USE, $\mu E5$ ou HSE como competidores, a competição do complexo C2' formado nas sondas $\mu E2$ foi mais facilmente conseguida com os oligos $\mu E2$ enquanto que os complexos C2 competiram preferencialmente para a sequência USE. O complexo C1 competiu também mais eficazmente para a sequência USE. Apresenta-se abaixo uma análise detalhada das especificidades de ligação destes complexos.

EXEMPLO 4

Proteínas Responsáveis pela Formação dos Complexos C1, C2 e C2'

Um cenário sugerido pelas diferenças na ligação é que Myc possa não ser a única proteína envolvida na formação dos três complexos. Para distinguir o papel de c-Myc das outras proteínas que copurificam com ela na formação dos complexos observados realizaram-se experiências de corte e renaturação como se segue. Quantidades preparativas de Myc foram separadas por SDS-PAGE. As proteínas foram electroeluidas a partir de fatias de diferentes pesos moleculares, precipitadas, solubilizadas em hidrócloro de guanidina e dialisadas para renaturar para análise por EMSA. A

actividade de ligação do complexo C1 pode ser renaturada a partir de fatias contendo Myc de preparações de baculovírus ou de mamífero (Fig. 3) enquanto que nenhuma das outras fatias do gel continham actividade C1 (resultados não apresentados). Estes resultados suportam Myc sózinho seja a proteína responsável pelo complexo C1 e que a proteína Myc de tamanho completo tal como é expressa em células eucarióticas se possa ligar especificamente a sítios com a sequência CACGTG.

A análise das proteínas responsáveis pela formação dos complexos C2 e C2' foi conseguida com outras experiências de corte e renaturação realizadas como descrito abaixo. EMSA usando a sonda USE não revelou uma fatia única do gel de preparações derivadas de CHO ou de baculovírus as quais continham quantidades detectáveis de ligação C2 (resultados não apresentados). No entanto, esta actividade foi recuperada por renaturação das proteínas a partir da fatia de 26-29 KD juntamente com proteínas na fatia de 60-70 KD contendo Myc (Fig. 4, faixas 1-8). O componente de 26-29 estava presente nos géis com c-Myc produzido a partir de CHO ou a partir de baculovírus e quando renaturadas com Myc, demonstraram a mesma especificidade do complexo C2 no material aplicado. A renaturação da fatia de 26-29 KD com BSA ou proteína A não deu a actividade de ligação a USE sugerindo que Myc desempenha um papel específico na recuperação da actividade de ligação C2. Para examinar melhor os papéis da copurificação de proteínas e de modificações de Myc na ligação observada, Myc foi também purificado a partir de um sistema de expressão bacteriano. O sistema de expressão e método de purificação usados foram os de Chi Dang e colaboradores (ver materiais e métodos). A proteína produzida por bactérias contém o segmento de ligação a IgG da proteína A fundido com os 85 aminoácidos C-terminais de Myc, o segmento de Myc que contém os motivos B-HLH e de fecho de correr com leucina. Para muitas das proteínas B-HLH, a pequena região da

proteína contendo o motivo B-HLH não só é necessária mas também suficiente para a ligação a DNA se estiver presente o parceiro de oligomerização correcto. Esta proteína foi capaz de formar o complexo C1 nas sondas USE (Fig. 4, faixa 9) e de se combinar com o factor de 26-29 KD para criar o complexo C2 (Fig. 4, faixa 10). As experiências de competição confirmaram a especificidade deste complexo C2 reconstituído. Os complexos C1 e C2 formados usando esta proteína de fusão bacteriana migraram mais rapidamente do que os formados usando c-Myc de tamanho completo (comparar a Fig. 4, faixa 8 com faixas 9 e 10). Isto pode ser devido à diferença no tamanho entre c-Myc de tamanho completo (60-72 KD) e a proteína de fusão Proteína A-Myc (≈ 38 KD) e portanto a mobilidade do complexo C2 pode ser interpretada como indicadora de que Myc está fisicamente presente no complexo C2, presumivelmente como parte de um hetero-oligómero com o factor de 26-29 KD.

Experiências análogas foram realizadas usando uma proteína de fusão bacteriana semelhante contendo apenas os 67 aminoácidos C-terminais de c-Myc. Esta proteína contém a maior parte do domínio HLH e todo o domínio em fecho de correr de leucina mas não a região básica. Se bem que esta proteína seja capaz de formar homo-oligómeros em solução (Gentz *et al.*, Science 243:1695-1699 (1989)), foi incapaz de se ligar a DNA para formar o complexo C1 e foi também incapaz de se combinar com o factor de 26-29 KD para criar qualquer actividade de ligação a USE (Fig. 4, faixa 12). Estes resultados apoiam a hipótese do papel de Myc na formação de hetero-oligómeros requerer uma região básica intacta, a região responsável pelos contactos específicos com DNA nas outras proteínas B-HLH.

Usando experiências de corte e renaturação a actividade de ligação a μ E2 responsável pelo complexo C2' pode ser identificada. Uma pequena quantidade do complexo C2' foi frequentemente

observada com proteínas da fatia que englobava a gama de pesos moleculares 40-50 KD das preparações de Myc de mamífero (Fig. 5A). Se bem que o complexo C2' fosse sempre visto com a preparação contendo apenas Myc, a renaturação da proteína a partir da fatia de Myc com a fatia de 40-50 KD reprodutivelmente aumentou a quantidade de complexo C2' formado. Tanto Myc produzido por baculovírus como a proteína de fusão expressa por bactérias contendo a região básica, a qual não forma complexos por si só com sondas μ E2, foram também capazes de aumentar a quantidade de complexo formado pela fatia de 40-50 KD obtida a partir de preparações de mamífero (Fig. 5B e C). Surpreendentemente Myc produzido por bactérias sem a região básica pode reconstituir também a actividade C2', enquanto várias outras proteínas tentadas incluindo BSA, imunoglobulinas e proteína A não o conseguiram. A aparente ausência de um papel para esta região básica sugere que o envolvimento de Myc na formação deste complexo pode ser outro que não o contacto com o DNA.

Para melhor determinar se Myc estava presente nos complexos analisados, as preparações de Myc foram incubadas com um anticorpo monoclonal específico contra Myc antes de EMSA. A sonda usada nesta experiência (SMS) continha um único sítio com a sequência de USE, CACGTG. O anticorpo monoclonal específico contra Myc eliminou tanto os complexos C1 como C2 e produziu um complexo proeminente de mobilidade mais baixa (Fig. 9). Não é claro a partir destes dados quais dos dois complexos foi super-alterado mas a presença de um complexo predominantemente alterado quando o anticorpo está presente e de dois complexos na ausência de anticorpo suporta a hipótese de o anticorpo específico de Myc ter também destruído um dos complexos originais. Não houve efeito com um anticorpo testemunha na formação do complexo C1 ou C2. O anticorpo específico de Myc não alterou o complexo C2', sugerindo que Myc não está presente neste complexo.

A partir destas experiências pode-se concluir que o complexo C1 é formado por Myc sózinho, que o complexo C2 contém Myc e um factor de 26-29 KD e que o complexo C2' contém um factor de 40-50 KD mas não contém Myc. É intrigante que o complexo C2' necessite da presença de Myc para a sua formação, mas aparentemente não contém Myc. Myc parece portanto ser capaz de afectar a capacidade do factor de 40-50 KD para formar o complexo C2' sem ser um membro do complexo. Qualquer que seja o mecanismo, o aumento da actividade de ligação a μ E2 do factor de 40-50 KD parece ser específico de Myc uma vez que as quatro proteínas Myc diferentes aumentaram a quantidade de complexo C2' observado enquanto que várias outras proteínas não o fizeram.

A proteína Max pode ser imunoprecipitada a partir de células aviárias e humanas e análise Southern de baixa restrição sugeriu que existe um único gene Max ou uma pequena família de genes existe igualmente noutros vertebrados (Blackwood e Eisenmann, Science 251:1211-1217 (1991)). É possível que células de hamster e de insecto tenham uma Max equivalente. A recuperação de uma actividade tipo Max a partir de células de insecto é particularmente interessante uma vez que não se encontrou até hoje homólogos de Myc em insectos. Drosophila claramente usa heterodímeros de B-HLH para regular aspectos do desenvolvimento e permanece a possibilidade de que o parceiro natural da proteína 26-29 KD em células de insecto seja uma proteína B-HLH ainda não identificado que funcione como Myc.

A presença do factor 26-29 KD nestas preparações deve limitar a sua utilidade em certas experiências. Ao reunir-se as fracções contendo Myc baseadas num ensaio EMSA, pode-se obter fracções que contêm apenas actividade C1 e que não contêm a actividade C2, se bem que esta modificação reduza o rendimento final em aproximadamente 80%.

EXEMPLO 5

Seleccção de Sítios de Ligação para Myc a partir de Sequências de DNA ao Acaso

Para determinar os sítios de ligação óptimos para os três complexos nas preparações de Myc descritas atrás, usou-se uma modificação de uma técnica recentemente descrita para o isolamento de sítios de ligação preferidos a partir de sequências de DNA ao acaso (Pollock e Treisman, Nucl. Acids Res. 18:6197-6204 (1990)). Resumidamente, foi criado um conjunto de oligonucleótidos de cadeia dupla que consistiu em regiões flangeantes de 16 bases de sequência definida envolvendo uma região de 20 bases de sequência completamente ao acaso. Cada uma das preparações de Myc eucariótico descritas atrás foi misturada com este conjunto de sequências e os complexos de DNA proteína que se formaram foram separados por EMSA. O DNA que migrou na posição dos complexos C1 ou de C2 (e comigrando com C2') foi isolado, amplificado pela reacção em cadeia com polimerase (PCR) e usado num segundo ciclo de selecção por EMSA. Sete ciclos de selecção (preparação CHO) ou oito (preparação de baculovírus) no total foram efectuadas antes da subclonagem de membros individuais das sequências seleccionadas. Como se esperava que cada ciclo enriquecesse em melhores sítios de ligação, esperava-se que os oligonucleótidos finais subclonados contivessem sítios de ligação de alta afinidade para os complexos C1, C2 e C2'.

As sequências seleccionadas foram colocadas em três grupos separados para análise (Fig. 7). O grupo I contém sequências que foram seleccionadas pelo complexo C1 a partir de preparações de mamífero ou de baculovírus. Estas sequências foram reunidas para análise pois com ambas as preparações a formação do complexo C1 requiere apenas proteína Myc e porque as duas séries

de sequências (que foram isoladas com Myc de mamífero e com Myc de baculovírus) eram semelhantes uma à outra. A maior parte das sequências seleccionadas neste grupo continham a sequência CACGTG (21/27 dos subclones sequenciados). Ao alinhar-se todas as sequências que continham esta sequência central, encontrou-se que as sequências flanqueantes deste núcleo também não eram ao acaso. Determinou-se uma sequência consensus de 12 bases GACCACGTGCTC [SEQ ID Nº 1] para sítios seleccionados pelo complexo C1 (ver tabela na Fig. 7 para frequências em cada uma das posições; para uma base a ser incluída no consensus teve de se encontrar 10 de 21 sequências com um núcleo CACGTG). O complexo C2 derivado de preparações de baculovírus seleccionou sequências semelhantes às seleccionadas pelo complexo C1 (Fig. 7, grupo II). A maior parte destas sequências seleccionadas também continham CACGTG (19/22). Estas sequências tinham sequências flanqueantes semelhantes adjacentes ao hexâmero central relativamente à encontrada com o complexo C1, se bem que houvesse uma ligeira preferência para GCC relativamente a CTC na região flanqueante 3' (ver Tabela para o grupo II na Fig. 7).

Conforme esperado, os complexos que migraram na posição de C2 que foram seleccionados pelas preparações de Myc de mamífero têm uma maior diversidade de sequências (Fig. 7, Grupo III). Várias sequências (8/36) continham CACGTG. Estas sequências foram presumivelmente seleccionadas pelo complexo C2 de mamífero (compreendido por Myc e pelo factor 26-29 KD) e demonstraram as mesmas preferências relativamente às sequências flanqueantes que o complexo C1. Vários outros sítios seleccionados (9/36) continham uma sequência CAGCTG presumivelmente seleccionada pelo complexo C2'. Em adição, 8 das 36 sequências eram muito ricas em AT e muitas das sequências nos três grupos continham segmentos ricos em AT. Este enriquecimento em sequências ricas em AT deve refletir uma preferência de Myc para estas sequências ou poderá

apenas simplesmente indicar uma tendência devida ao protocolo usado. É interessante notar, no entanto que em experiências de ligação a filtros anteriores, a preparação de Myc de mamífero demonstrou uma preferência para a ligação a sequências ricas em AT dentro de DNA de vários plasmídeos ou de DNA genômico de lambda.

Para confirmar a validade do nosso processo de selecção de sítios foi testada uma série de sítios seleccionados individualmente por EMSA (Fig. 8). Conforme esperado, encontrou-se que sequências contendo CACGTG constituíam tanto o complexo C1 como o complexo C2 (Fig. 8, grupos de sondas 1, 2, 5 e 6) enquanto que sequências contendo CAGCTG formaram apenas o complexo C2' (Fig. 8, grupos de sondas 7 e 8). Note-se que a actividade C2' está apenas presente nas preparações Myc derivadas de CHO. Não se formaram complexos quando as sequências seleccionadas que continham uma caixa canónica E foram testadas (Fig. 8 grupos de sondas 3, 4, 9, 10 e 11). Portanto estas últimas sequências não representam sítios de elevada afinidade para proteínas nas preparações de Myc.

EXEMPLO 6

Velocidade de Formação dos Complexos C1 e C2

As velocidades de formação para os complexos contendo Myc foram medidas como um meio de comparar as suas afinidades. A velocidade de formação do complexo C2 na sonda USE foi de aproximadamente 1-2 minutos (Fig.9, Myc de baculovírus; resultados semelhantes foram obtidos com Myc de CHO, resultados não apresentados). O complexo C1 não competiu totalmente nesta experiência usando um excesso de 250 vezes de competidor USE. Se bem que a competição não fosse completa, a quantidade de complexo C1

restante no tempo mais precoce susceptível de ser avaliado ("0") foi significativamente menos do que a quantidade de partida e virtualmente igual à competição máxima conseguida nestas experiências. Estes resultados são indicadores de uma proteína de ligação fraca abundante com uma velocidade de associação rápida não mensurável. Assim, Myc sózinho parece ligar-se significativamente mais fracamente do que Myc e do que o factor 26-29 KD.

EXEMPLO 7

Identificação de um Inibidor da Actividade de Complexo C2 de c-Myc em Células de Levedura

Células hospedeiras de levedura foram transformadas com plasmídeos portadores de um vector de expressão de c-Myc (hospedeiro "a"); ou o vector de expressão com c-Myc e uma proteína do complexo C2 de 26-29 KD identificada como atrás (hospedeiro "b"). Ainda todas as estirpes de levedura foram cotransformadas com um plasmídeo que contem a sequência codificadora da β -galactosidase operacionalmente ligada à sequência do motivo CACGTG descrito atrás.

Um tapete de cada uma das estirpes de levedura transformadas foi espalhado em placas de agar contendo X-gal no meio e colocados no tapete pequenos discos de papel de filtro contendo o composto W, X, Y ou Z. A levedura cresceu e as placas foram controladas por crescimento das colónias e observação visual da cor das colónias. Os resultados típicos de uma destas experiências estão apresentados na Tabela 1.

Tabela 1: Identificação de Inibidores da Actividade do Complexo C2

Composto	Levedura	Crescimento de colónias	Cor do Ensaio de β -gal com X-gal
nenhum	a	+	Branco
	b	+	Azul
W	a	+	Branco
	b	+	Branco
X	a	-	-
	b	-	-
Y	a	+	Branco
	b	+	Azul
Z	a	+	Azul
	b	+	Azul

Os resultados da tabela acima indicam que o composto W evita a indução de β -galactosidase nas células hospedeiras "b". Portanto, o composto W é um inibidor da formação de hetero-oligómeros do complexo C2 e um inibidor da actividade biológica de c-Myc. O composto X inibe o crescimento de levedura per se e portanto não será um composto com interesse.

O composto Y não evita a indução da actividade de β -galactosidase nas células hospedeiras "b". Assim, o composto Y

não é um inibidor da formação de hetero-oligómeros do complexo C2.

O composto Z mostra um efeito interessante na indução da actividade de β -galactosidase nas células hospedeiras "a" que não contem a proteína do complexo C2 usada nos hospedeiros "b", não evitando a formação de hetero-oligómeros. Isto sugere que o composto Z possa induzir a síntese de uma proteína parceira que de outra forma não está presente nas células hospedeiras de levedura ou que ele possa ser (ou imite) tal proteína.

A partir destes resultados, o composto W seria identificado como um inibidor da formação do complexo C2 e/ou ligação a DNA e assim da actividade de transcrição de c-Myc in vivo.

EXEMPLO 8

Identificação de um Inibidor da Actividade de Complexo C2' de c-Myc em Células de Levedura

Células hospedeiras de levedura foram transformadas com dois plasmídeos, cada um dos plasmídeos sendo portador de um vector de expressão do complexo C2' codificador de pelo menos um peptídeo C2' de 40-50 Kilodalton (hospedeiro "a"); ou o vector de expressão de c-Myc para além dos vectores codificadores das proteínas do complexo C2' identificadas como atrás (hospedeiro "b"). Ainda todas as estirpes de levedura foram cotransformadas com um plasmídeo que contem a sequência codificadora da β -galactosidase operacionalmente ligada ao motivo da sequência CAGCTG como descrito atrás.

Um tapete de cada uma das estirpes de levedura transformadas foi espalhado em placas de agar contendo X-gal no meio e

pequenos discos de filtros contendo o composto W, X, Y ou Z foram colocados sobre o tapete. Deixou-se as leveduras crescerem e as placas forma controladas relativamente ao crescimento de colónias e cor das colónias por observação visual. Resultados típicos de cada experiência estão apresentados na Tabela 2.

Tabela 2: Identificação de Inibidores da Actividade do Complexo C2'

Composto	Levedura	Crescimento de colónias	Cor do Ensaio de β -gal com X-gal
nenhum	a	+	Branco
	b	+	Azul
W	a	+	Branco
	b	+	Branco
X	a	-	-
	b	-	-
Y	a	+	Branco
	b	+	Azul
Z	a	+	Azul
	b	+	Azul

Os resultados da tabela acima indicam que o composto W evita a indução de β -galactosidase nas células hospedeiras "b". Portanto, o composto W é um inibidor da formação de hetero-oligómeros do complexo C2' e um inibidor da actividade biológica de c-Myc que é dirigida no sentido de promover a formação de hetero-oligómeros do complexo C2'. O composto X inibe o crescimento de levedura per se e portanto não será um composto com interesse.

O composto Y não evita a indução da actividade de β -galactosidase nas células hospedeiras "b". Assim, o composto Y não é um inibidor da formação de hetero-oligómeros do complexo C2.

O composto Z mostra um efeito interessante na indução da actividade de β -galactosidase nas células hospedeiras "a" que não contem a proteína c-Myc usada nos hospedeiros "b", não evitando a formação de hetero-oligómeros. Isto sugere que o composto Z possa induzir a síntese de uma proteína que c-Myc possa substituir na promoção da formação do complexo C2' que de outra forma não está presente nas células hospedeiras de levedura ou que ele possa ser (ou imite) tal proteína.

A partir destes resultados, o composto W seria identificado como um inibidor da formação do complexo C2' e/ou ligação a DNA e assim da actividade de transcrição de c-Myc in vivo.

Todas as referências aqui citadas são incluídas na sua totalidade por referência. Tendo descrito detalhadamente o invento, será entendido pelos familiarizados com a matéria que o invento pode ser executado dentro de uma larga gama de condições, parâmetros e similares equivalentes, sem afectar o espírito ou âmbito do invento ou de qualquer realização dele.

LISTAGEM DE SEQUÊNCIAS

(1) INFORMAÇÃO GERAL:

(i) REQUERENTE: Kingston, Robert E
Papoulas, Ophelia

(ii) TÍTULO DO INVENTO: PARCEIROS, MOTIVOS, ENSAIOS DE
DESPISTE E UTILIZAÇÕES DA LIGAÇÃO DE C-MYC A DNA

(iii) NÚMERO DE SEQUÊNCIAS: 101

(iv) ENDEREÇO PARA CORRESPONDÊNCIA:

- (A) DESTINATÁRIO: Stern, Kessler, Goldstein & Fox
- (B) RUA: 1225 Connecticut Avenue, N.W., Suite 300
- (C) CIDADE: Washington
- (D) ESTADO: DC
- (E) PAÍS: USA
- (E) CÓDIGO: 20036

(v) FORMA LIDA PELO COMPUTADOR:

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

(vi) DADOS DE PATENTE :

- (A) NÚMERO DO PEDIDO: US
- (B) DATA DE ENTREGA:
- (C) CLASSIFICAÇÃO:

(viii) INFORMAÇÃO SOBRE O SOLICITADOR/AGENTE:

- (A) NOME: Cimbala, Michele A
- (B) NÚMERO DE REGISTO: 33,851

(C) NÚMERO DA REFERÊNCIA/DOCUMENTO: 0609.3440004

(ix) INFORMAÇÃO PARA TELECOMUNICAÇÃO:

(A) TELEFONE: (202)833-7533

(B) TELEFAX: (202)833-8716

(2) INFORMAÇÃO PARA SEQ ID NO.1:

(i) CARACTERÍSTICAS DA SEQUÊNCIA:

(A) COMPRIMENTO: 12 pares de bases

(B) TIPO: ácido nucleico

(C) TIPO DE CADEIA: simples

(D) TOPOLOGIA: linear

(ii) TIPO DE MOLÉCULA: DNA

(xi) DESCRIÇÃO DA SEQUÊNCIA: SEQ ID NO. 1:

GACCACGTGC TC

(2) INFORMAÇÃO PARA SEQ ID NO.2:

(i) CARACTERÍSTICAS DA SEQUÊNCIA:

(A) COMPRIMENTO: 12 pares de bases

(B) TIPO: ácido nucleico

(C) TIPO DE CADEIA: simples

(D) TOPOLOGIA: linear

(ii) TIPO DE MOLÉCULA: DNA

(xi) DESCRIÇÃO DA SEQUÊNCIA: SEQ ID NO. 2:

GACCACGTGG TC

(2) INFORMAÇÃO PARA SEQ ID NO.3:

(i) CARACTERÍSTICAS DA SEQUÊNCIA:

- (A) COMPRIMENTO: 12 pares de bases
- (B) TIPO: ácido nucleico
- (C) TIPO DE CADEIA: simples
- (D) TOPOLOGIA: linear

(ii) TIPO DE MOLÉCULA: DNA

(xi) DESCRIÇÃO DA SEQUÊNCIA: SEQ ID NO. 3:

AACAGTYCTG TT

(2) INFORMAÇÃO PARA SEQ ID NO.4:

(i) CARACTERÍSTICAS DA SEQUÊNCIA:

- (A) COMPRIMENTO: 17 pares de bases
- (B) TIPO: ácido nucleico
- (C) TIPO DE CADEIA: simples
- (D) TOPOLOGIA: linear

(ii) TIPO DE MOLÉCULA: DNA

(xi) DESCRIÇÃO DA SEQUÊNCIA: SEQ ID NO. 4:

CCGATATCTC GAGACGG

(2) INFORMAÇÃO PARA SEQ ID NO.5:

(i) CARACTERÍSTICAS DA SEQUÊNCIA:

- (A) COMPRIMENTO: 16 pares de bases
- (B) TIPO: ácido nucleico

(C) TIPO DE CADEIA: simples

(D) TOPOLOGIA: linear

(ii) TIPO DE MOLÉCULA: DNA

(xi) DESCRIÇÃO DA SEQUÊNCIA: SEQ ID NO. 5:

GGACGATCTA GATTCG

(2) INFORMAÇÃO PARA SEQ ID NO.6:

(i) CARACTERÍSTICAS DA SEQUÊNCIA:

(A) COMPRIMENTO: 40 pares de bases

(B) TIPO: ácido nucleico

(C) TIPO DE CADEIA: simples

(D) TOPOLOGIA: linear

(ii) TIPO DE MOLÉCULA: DNA

(xi) DESCRIÇÃO DA SEQUÊNCIA: SEQ ID NO. 6:

GATCTCTGCA GCAGCTGGCA GCAGCTGGCA GCAGCTGGCG

(2) INFORMAÇÃO PARA SEQ ID NO.7:

(i) CARACTERÍSTICAS DA SEQUÊNCIA:

(A) COMPRIMENTO: 40 pares de bases

(B) TIPO: ácido nucleico

(C) TIPO DE CADEIA: simples

(D) TOPOLOGIA: linear

(ii) TIPO DE MOLÉCULA: DNA

(xi) DESCRIÇÃO DA SEQUÊNCIA: SEQ ID NO. 7:

GATCTGCAGT CATGTGGCGT CATGTGGCGT CATGTGGCAG

(2) INFORMAÇÃO PARA SEQ ID NO.8:

(i) CARACTERÍSTICAS DA SEQUÊNCIA:

- (A) COMPRIMENTO: 40 pares de bases
- (B) TIPO: ácido nucleico
- (C) TIPO DE CADEIA: simples
- (D) TOPOLOGIA: linear

(ii) TIPO DE MOLÉCULA: DNA

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

GATCTGCAGT CACGTGGCGT CACGTGGCGT CACGTGGCAG

(2) INFORMAÇÃO PARA SEQ ID NO.9:

(i) CARACTERÍSTICAS DA SEQUÊNCIA:

- (A) COMPRIMENTO: 21 pares de bases
- (B) TIPO: ácido nucleico
- (C) TIPO DE CADEIA: simples
- (D) TOPOLOGIA: linear

(ii) TIPO DE MOLÉCULA: DNA

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

TCGACGTCGC AGCAGGTGCA G

(2) INFORMAÇÃO PARA SEQ ID NO. 10:

(i) CARACTERÍSTICAS DA SEQUÊNCIA:

- (A) COMPRIMENTO: 21 pares de bases
- (B) TIPO: ácido nucleico
- (C) TIPO DE CADEIA: simples
- (D) TOPOLOGIA: linear

(ii) TIPO DE MOLÉCULA: DNA

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

TCGACCCAC CAGCTGGCGA G

(2) INFORMAÇÃO PARA SEQ ID NO.11:

(i) CARACTERÍSTICAS DA SEQUÊNCIA:

- (A) COMPRIMENTO: 40 pares de bases
- (B) TIPO: ácido nucleico
- (C) TIPO DE CADEIA: simples
- (D) TOPOLOGIA: linear

(ii) TIPO DE MOLÉCULA: DNA

(xi) DESCRIÇÃO DA SEQUÊNCIA: SEQ ID NO. 11:

AGCTTCGAAC ACCTGCAGCA GCTGGCAGGA GCTGGCAGGA AGCAGGCCTA

(2) INFORMAÇÃO PARA SEQ ID NO.12:

(i) CARACTERÍSTICAS DA SEQUÊNCIA:

- (A) COMPRIMENTO: 40 pares de bases
- (B) TIPO: ácido nucleico
- (C) TIPO DE CADEIA: simples

(D) TOPOLOGIA: linear

(ii) TIPO DE MOLÉCULA: DNA

(xi) DESCRIÇÃO DA SEQUÊNCIA: SEQ ID NO. 12:

AGCTTTAAAA TCCCCACCAG CTGGCGAAGC AACAGGTGCA

(2) INFORMAÇÃO PARA SEQ ID NO.13:

(i) CARACTERÍSTICAS DA SEQUÊNCIA:

- (A) COMPRIMENTO: 39 pares de bases
- (B) TIPO: ácido nucleico
- (C) TIPO DE CADEIA: simples
- (D) TOPOLOGIA: linear

(ii) TIPO DE MOLÉCULA: DNA

(xi) DESCRIÇÃO DA SEQUÊNCIA: SEQ ID NO. 13:

AATTGCGAAA CCCCTGGAAT ATTCCGACCT ATTCCGACCT GGCAGCCTC

(2) INFORMAÇÃO PARA SEQ ID NO.14:

(i) CARACTERÍSTICAS DA SEQUÊNCIA:

- (A) COMPRIMENTO: 28 pares de bases
- (B) TIPO: ácido nucleico
- (C) TIPO DE CADEIA: simples
- (D) TOPOLOGIA: linear

(ii) TIPO DE MOLÉCULA: DNA

(xi) DESCRIÇÃO DA SEQUÊNCIA: SEQ ID NO. 14:

TCGACTTTAG ACCACGTGGT CCCCTCGA

(2) INFORMAÇÃO PARA SEQ ID NO.15:

(i) CARACTERÍSTICAS DA SEQUÊNCIA:

- (A) COMPRIMENTO: 52 pares de bases
- (B) TIPO: ácido nucleico
- (C) TIPO DE CADEIA: simples
- (D) TOPOLOGIA: linear

(ii) TIPO DE MOLÉCULA: DNA

(xi) DESCRIÇÃO DA SEQUÊNCIA: SEQ ID NO. 15:

GGACGATCTA GATTCGNNNN NNNNNNNNNN NNNNNNCCGT CTCGAGTATC GG

(2) INFORMAÇÃO PARA SEQ ID NO.16:

(i) CARACTERÍSTICAS DA SEQUÊNCIA:

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

(ii) TIPO DE MOLÉCULA: DNA

(xi) DESCRIÇÃO DA SEQUÊNCIA: SEQ ID NO. 16:

NNCANNTGNN

(2) INFORMAÇÃO PARA SEQ ID NO.17:

(i) CARACTERÍSTICAS DA SEQUÊNCIA:

- (A) COMPRIMENTO: 21 pares de bases
- (B) TIPO: ácido nucleico
- (C) TIPO DE CADEIA: simples
- (D) TOPOLOGIA: linear

(ii) TIPO DE MOLÉCULA: DNA

(xi) DESCRIÇÃO DA SEQUÊNCIA: SEQ ID NO. 17:

GCAGAATCTA CCACGTGCTC C

(2) INFORMAÇÃO PARA SEQ ID NO.18:

(i) CARACTERÍSTICAS DA SEQUÊNCIA:

- (A) COMPRIMENTO: 20 pares de bases
- (B) TIPO: ácido nucleico
- (C) TIPO DE CADEIA: simples
- (D) TOPOLOGIA: linear

(ii) TIPO DE MOLÉCULA: DNA

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

GGGGCTACCA CGTGCTTATG

(2) INFORMAÇÃO PARA SEQ ID NO.19:

(i) CARACTERÍSTICAS DA SEQUÊNCIA:

- (A) COMPRIMENTO: 20 pares de bases
- (B) TIPO: ácido nucleico
- (C) TIPO DE CADEIA: simples
- (D) TOPOLOGIA: linear

(ii) TIPO DE MOLÉCULA: DNA

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

GGACGAAAGC ACGTGCTCCG

(2) INFORMAÇÃO PARA SEQ ID NO.20:

(i) CARACTERÍSTICAS DA SEQUÊNCIA:

- (A) COMPRIMENTO: 20 pares de bases
- (B) TIPO: ácido nucleico
- (C) TIPO DE CADEIA: simples
- (D) TOPOLOGIA: linear

(ii) TIPO DE MOLÉCULA: DNA

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

GCACATGACC ACGTGCTCTG

(2) INFORMAÇÃO PARA SEQ ID NO.21:

(i) CARACTERÍSTICAS DA SEQUÊNCIA:

- (A) COMPRIMENTO: 20 pares de bases
- (B) TIPO: ácido nucleico
- (C) TIPO DE CADEIA: simples
- (D) TOPOLOGIA: linear

(ii) TIPO DE MOLÉCULA: DNA

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

GGCAGAGACA CGTGCCCTGG

(2) INFORMAÇÃO PARA SEQ ID NO.22:

(i) CARACTERÍSTICAS DA SEQUÊNCIA:

- (A) COMPRIMENTO: 20 pares de bases
- (B) TIPO: ácido nucleico
- (C) TIPO DE CADEIA: simples
- (D) TOPOLOGIA: linear

(ii) TIPO DE MOLÉCULA: DNA

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

GGCAAACCAC GTGTTATGTG

(2) INFORMAÇÃO PARA SEQ ID NO.23:

(i) CARACTERÍSTICAS DA SEQUÊNCIA:

- (A) COMPRIMENTO: 22 pares de bases
- (B) TIPO: ácido nucleico
- (C) TIPO DE CADEIA: simples
- (D) TOPOLOGIA: linear

(ii) TIPO DE MOLÉCULA: DNA

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

CGACCACGTG CTCTTCGACT TG

(2) INFORMAÇÃO PARA SEQ ID NO.24:

(i) CARACTERÍSTICAS DA SEQUÊNCIA:

- (A) COMPRIMENTO: 23 pares de bases
- (B) TIPO: ácido nucleico

(C) TIPO DE CADEIA: simples

(D) TOPOLOGIA: linear

(ii) TIPO DE MOLÉCULA: DNA

(xi) DESCRIÇÃO DA SEQUÊNCIA: SEQ ID NO. 24:

GCACAATTTG TACCACGTGG CCG

(2) INFORMAÇÃO PARA SEQ ID NO.25:

(i) CARACTERÍSTICAS DA SEQUÊNCIA:

(A) COMPRIMENTO: 23 pares de bases

(B) TIPO: ácido nucleico

(C) TIPO DE CADEIA: simples

(D) TOPOLOGIA: linear

(ii) TIPO DE MOLÉCULA: DNA

(xi) DESCRIÇÃO DA SEQUÊNCIA: SEQ ID NO. 25:

GGACAACATC GACCACGTGG CCG

(2) INFORMAÇÃO PARA SEQ ID NO. 26:

(i) CARACTERÍSTICAS DA SEQUÊNCIA:

(A) COMPRIMENTO: 21 pares de bases

(B) TIPO: ácido nucleico

(C) TIPO DE CADEIA: simples

(D) TOPOLOGIA: linear

(ii) TIPO DE MOLÉCULA: DNA

(xi) DESCRIÇÃO DA SEQUÊNCIA: SEQ ID NO. 26:

GCCTGCATGA CCACGTGGAC C

(2) INFORMAÇÃO PARA SEQ ID NO.27:

(i) CARACTERÍSTICAS DA SEQUÊNCIA:

- (A) COMPRIMENTO: 21 pares de bases
- (B) TIPO: ácido nucleico
- (C) TIPO DE CADEIA: simples
- (D) TOPOLOGIA: linear

(ii) TIPO DE MOLÉCULA: DNA

(xi) DESCRIÇÃO DA SEQUÊNCIA: SEQ ID NO. 27:

GCAAATATGA CCACGTGGTA C

(2) INFORMAÇÃO PARA SEQ ID NO.28:

(i) CARACTERÍSTICAS DA SEQUÊNCIA:

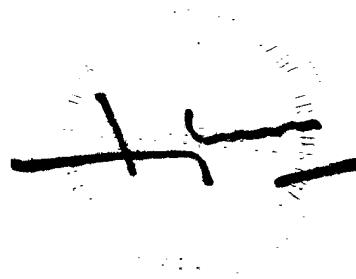
- (A) COMPRIMENTO: 20 pares de bases
- (B) TIPO: ácido nucleico
- (C) TIPO DE CADEIA: simples
- (D) TOPOLOGIA: linear

(ii) TIPO DE MOLÉCULA: DNA

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

GGACCACGTG CTCTTTTGTG

(2) INFORMAÇÃO PARA SEQ ID NO.29:



(i) CARACTERÍSTICAS DA SEQUÊNCIA:

- (A) COMPRIMENTO: 21 pares de bases
- (B) TIPO: ácido nucleico
- (C) TIPO DE CADEIA: simples
- (D) TOPOLOGIA: linear

(ii) TIPO DE MOLÉCULA: DNA

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

GGCATAAACT CCACGTGGTC C

(2) INFORMAÇÃO PARA SEQ ID NO.30:

(i) CARACTERÍSTICAS DA SEQUÊNCIA:

- (A) COMPRIMENTO: 22 pares de bases
- (B) TIPO: ácido nucleico
- (C) TIPO DE CADEIA: simples
- (D) TOPOLOGIA: linear

(ii) TIPO DE MOLÉCULA: DNA

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

CGGGCACGTG CTCCTCGGAC TG

(2) INFORMAÇÃO PARA SEQ ID NO.31:

(i) CARACTERÍSTICAS DA SEQUÊNCIA:

- (A) COMPRIMENTO: 22 pares de bases
- (B) TIPO: ácido nucleico
- (C) TIPO DE CADEIA: simples

(D) TOPOLOGIA: linear

(ii) TIPO DE MOLÉCULA: DNA

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

GGTAGCAAAA AGCACGTGCC CG

(2) INFORMAÇÃO PARA SEQ ID NO.32:

(i) CARACTERÍSTICAS DA SEQUÊNCIA:

- (A) COMPRIMENTO: 22 pares de bases
- (B) TIPO: ácido nucleico
- (C) TIPO DE CADEIA: simples
- (D) TOPOLOGIA: linear

(ii) TIPO DE MOLÉCULA: DNA

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

GGGGGATTTA AGCACGTGCT CC

(2) INFORMAÇÃO PARA SEQ ID NO.33:

(i) CARACTERÍSTICAS DA SEQUÊNCIA:

- (A) COMPRIMENTO: 12 pares de bases
- (B) TIPO: ácido nucleico
- (C) TIPO DE CADEIA: simples
- (D) TOPOLOGIA: linear

(ii) TIPO DE MOLÉCULA: DNA

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

CACCTATTAA CCACGTGGTA C

(2) INFORMAÇÃO PARA SEQ ID NO.34:

(i) CARACTERÍSTICAS DA SEQUÊNCIA:

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

(ii) TIPO DE MOLÉCULA: DNA

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

GACCACGCGG CATCCACGTG CCGT

(2) INFORMAÇÃO PARA SEQ ID NO.35:

(i) CARACTERÍSTICAS DA SEQUÊNCIA:

- (A) COMPRIMENTO: 20 pares de bases
- (B) TIPO: ácido nucleico
- (C) TIPO DE CADEIA: simples
- (D) TOPOLOGIA: linear

(ii) TIPO DE MOLÉCULA: DNA

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

GGGGACCACG TGCTCGGTTG

(2) INFORMAÇÃO PARA SEQ ID NO.36:

(i) CARACTERÍSTICAS DA SEQUÊNCIA:

- (A) COMPRIMENTO: 22 pares de bases
- (B) TIPO: ácido nucleico
- (C) TIPO DE CADEIA: simples
- (D) TOPOLOGIA: linear

(ii) TIPO DE MOLÉCULA: DNA

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

CACATATTAG ACCACGTGCT CC

(2) INFORMAÇÃO PARA SEQ ID NO.37:

(i) CARACTERÍSTICAS DA SEQUÊNCIA:

- (A) COMPRIMENTO: 23 pares de bases
- (B) TIPO: ácido nucleico
- (C) TIPO DE CADEIA: simples
- (D) TOPOLOGIA: linear

(ii) TIPO DE MOLÉCULA: DNA

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

CGGCCACGTG CTCACTGTCT ACC

(2) INFORMAÇÃO PARA SEQ ID NO.38:

(i) CARACTERÍSTICAS DA SEQUÊNCIA:

- (A) COMPRIMENTO: 20 pares de bases
- (B) TIPO: ácido nucleico
- (C) TIPO DE CADEIA: simples
- (D) TOPOLOGIA: linear

(ii) TIPO DE MOLÉCULA: DNA

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

GGATGGACAG CTTCTTCCTG

(2) INFORMAÇÃO PARA SEQ ID NO.39:

(i) CARACTERÍSTICAS DA SEQUÊNCIA:

- (A) COMPRIMENTO: 20 pares de bases
- (B) TIPO: ácido nucleico
- (C) TIPO DE CADEIA: simples
- (D) TOPOLOGIA: linear

(ii) TIPO DE MOLÉCULA: DNA

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

GCAATCCCCC GCTGCTCGCC

(2) INFORMAÇÃO PARA SEQ ID NO.40:

(i) CARACTERÍSTICAS DA SEQUÊNCIA:

- (A) COMPRIMENTO: 22 pares de bases
- (B) TIPO: ácido nucleico
- (C) TIPO DE CADEIA: simples
- (D) TOPOLOGIA: linear

(ii) TIPO DE MOLÉCULA: DNA

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

GCCAAAAATG TACAGCTGTG CC

(2) INFORMAÇÃO PARA SEQ ID NO.41:

(i) CARACTERÍSTICAS DA SEQUÊNCIA:

- (A) COMPRIMENTO: 23 pares de bases
- (B) TIPO: ácido nucleico
- (C) TIPO DE CADEIA: simples
- (D) TOPOLOGIA: linear

(ii) TIPO DE MOLÉCULA: DNA

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

CGGCCACGCG GTCATGAATG TGC

(2) INFORMAÇÃO PARA SEQ ID NO.42

(i) CARACTERÍSTICAS DA SEQUÊNCIA:

- (A) COMPRIMENTO: 20 pares de bases
- (B) TIPO: ácido nucleico
- (C) TIPO DE CADEIA: simples
- (D) TOPOLOGIA: linear

(ii) TIPO DE MOLÉCULA: DNA

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

GCAGGCTGTA' CGTGACTTGG

(2) INFORMAÇÃO PARA SEQ ID NO.43:

(i) CARACTERÍSTICAS DA SEQUÊNCIA:

- (A) COMPRIMENTO: 20 pares de bases
- (B) TIPO: ácido nucleico

- (C) TIPO DE CADEIA: simples
- (D) TOPOLOGIA: linear

(ii) TIPO DE MOLÉCULA: DNA

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

CCGCAGTCCT GGTGCTCGC

(2) INFORMAÇÃO PARA SEQ ID NO.44:

(i) CARACTERÍSTICAS DA SEQUÊNCIA:

- (A) COMPRIMENTO: 23 pares de bases
- (B) TIPO: ácido nucleico
- (C) TIPO DE CADEIA: simples
- (D) TOPOLOGIA: linear

(ii) TIPO DE MOLÉCULA: DNA

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

CACTAAGAAA TACCACGTGG CCG

(2) INFORMAÇÃO PARA SEQ ID NO.45:

(i) CARACTERÍSTICAS DA SEQUÊNCIA:

- (A) COMPRIMENTO: 21 pares de bases
- (B) TIPO: ácido nucleico
- (C) TIPO DE CADEIA: simples
- (D) TOPOLOGIA: linear

(ii) TIPO DE MOLÉCULA: DNA

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

GGGGATTAA GCACGTGCTC C

(2) INFORMAÇÃO PARA SEQ ID NO.46:

(i) CARACTERÍSTICAS DA SEQUÊNCIA:

- (A) COMPRIMENTO: 23 pares de bases
- (B) TIPO: ácido nucleico
- (C) TIPO DE CADEIA: simples
- (D) TOPOLOGIA: linear

(ii) TIPO DE MOLÉCULA: DNA

(xi) DESCRIÇÃO DA SEQUÊNCIA: SEQ ID NO. 46:

CGGCCACGTG CCTTCTTTCT CCG

(2) INFORMAÇÃO PARA SEQ ID NO.47:

(i) CARACTERÍSTICAS DA SEQUÊNCIA:

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

(ii) TIPO DE MOLÉCULA: DNA

(xi) DESCRIÇÃO DA SEQUÊNCIA: SEQ ID NO. 47:

CATAGTCGAG AGAGCACGTG CGAA

(2) INFORMAÇÃO PARA SEQ ID NO.48:

(i) CARACTERÍSTICAS DA SEQUÊNCIA:

- (A) COMPRIMENTO: 22 pares de bases
- (B) TIPO: ácido nucleico
- (C) TIPO DE CADEIA: simples
- (D) TOPOLOGIA: linear

(ii) TIPO DE MOLÉCULA: DNA

(xi) DESCRIÇÃO DA SEQUÊNCIA: SEQ ID NO. 48:

CATAAGTCAG ACCACGTGGC CG

(2) INFORMAÇÃO PARA SEQ ID NO.49:

(i) CARACTERÍSTICAS DA SEQUÊNCIA:

- (A) COMPRIMENTO: 22 pares de bases
- (B) TIPO: ácido nucleico
- (C) TIPO DE CADEIA: simples
- (D) TOPOLOGIA: linear

(ii) TIPO DE MOLÉCULA: DNA

(xi) DESCRIÇÃO DA SEQUÊNCIA: SEQ ID NO. 49:

CCCAACTAAG ACCACGTGGC CG

(2) INFORMAÇÃO PARA SEQ ID NO.50:

(i) CARACTERÍSTICAS DA SEQUÊNCIA:

- (A) COMPRIMENTO: 23 pares de bases
- (B) TIPO: ácido nucleico
- (C) TIPO DE CADEIA: simples

(D) TOPOLOGIA: linear

(ii) TIPO DE MOLÉCULA: DNA

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

CAGTCGAAGA GGCCACGTGG CGA

(2) INFORMAÇÃO PARA SEQ ID NO.51:

(i) CARACTERÍSTICAS DA SEQUÊNCIA:

(A) COMPRIMENTO: 23 pares de bases

(B) TIPO: ácido nucleico

(C) TIPO DE CADEIA: simples

(D) TOPOLOGIA: linear

(ii) TIPO DE MOLÉCULA: DNA

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

CGTAGGTTAT TCCCACGTGG CCG

(2) INFORMAÇÃO PARA SEQ ID NO.52:

(i) CARACTERÍSTICAS DA SEQUÊNCIA:

(A) COMPRIMENTO: 21 pares de bases

(B) TIPO: ácido nucleico

(C) TIPO DE CADEIA: simples

(D) TOPOLOGIA: linear

(ii) TIPO DE MOLÉCULA: DNA

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

CATAAATAGG CCACGTGCTC C

(2) INFORMAÇÃO PARA SEQ ID NO.53:

(i) CARACTERÍSTICAS DA SEQUÊNCIA:

- (A) COMPRIMENTO: 21 pares de bases
- (B) TIPO: ácido nucleico
- (C) TIPO DE CADEIA: simples
- (D) TOPOLOGIA: linear

(ii) TIPO DE MOLÉCULA: DNA

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

GGAAAATGTA CCACGTGCTC C

(2) INFORMAÇÃO PARA SEQ ID NO.54:

(i) CARACTERÍSTICAS DA SEQUÊNCIA:

- (A) COMPRIMENTO: 20 pares de bases
- (B) TIPO: ácido nucleico
- (C) TIPO DE CADEIA: simples
- (D) TOPOLOGIA: linear

(ii) TIPO DE MOLÉCULA: DNA

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

GGAACAGACC ACGTGGCTTG

(2) INFORMAÇÃO PARA SEQ ID NO.55:

(i) CARACTERÍSTICAS DA SEQUÊNCIA:

- (A) COMPRIMENTO: 20 pares de bases
- (B) TIPO: ácido nucleico
- (C) TIPO DE CADEIA: simples
- (D) TOPOLOGIA: linear

(ii) TIPO DE MOLÉCULA: DNA

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

GTACCACGTG CTTTTTTGGC

(2) INFORMAÇÃO PARA SEQ ID NO.56:

(i) CARACTERÍSTICAS DA SEQUÊNCIA:

- (A) COMPRIMENTO: 22 pares de bases
- (B) TIPO: ácido nucleico
- (C) TIPO DE CADEIA: simples
- (D) TOPOLOGIA: linear

(ii) TIPO DE MOLÉCULA: DNA

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

CAGTCCGACC AGCACTGTCC CG

(2) INFORMAÇÃO PARA SEQ ID NO.57:

(i) CARACTERÍSTICAS DA SEQUÊNCIA:

- (A) COMPRIMENTO: 23 pares de bases
- (B) TIPO: ácido nucleico
- (C) TIPO DE CADEIA: simples
- (D) TOPOLOGIA: linear

(ii) TIPO DE MOLÉCULA: DNA

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

CGGCCACGTG TCGAGCATGA GTC

(2) INFORMAÇÃO PARA SEQ ID NO.58:

(i) CARACTERÍSTICAS DA SEQUÊNCIA:

- (A) COMPRIMENTO: 23 pares de bases
- (B) TIPO: ácido nucleico
- (C) TIPO DE CADEIA: simples
- (D) TOPOLOGIA: linear

(ii) TIPO DE MOLÉCULA: DNA

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

CGGCCACGTG CTCGTAAATT TGC

(2) INFORMAÇÃO PARA SEQ ID NO.59:

(i) CARACTERÍSTICAS DA SEQUÊNCIA:

- (A) COMPRIMENTO: 23 pares de bases
- (B) TIPO: ácido nucleico
- (C) TIPO DE CADEIA: simples
- (D) TOPOLOGIA: linear

(ii) TIPO DE MOLÉCULA: DNA

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

GCGACAAAAT TACCACGTGG CCG

(2) INFORMAÇÃO PARA SEQ ID NO.60:

(i) CARACTERÍSTICAS DA SEQUÊNCIA:

- (A) COMPRIMENTO: 22 pares de bases
- (B) TIPO: ácido nucleico
- (C) TIPO DE CADEIA: simples
- (D) TOPOLOGIA: linear

(ii) TIPO DE MOLÉCULA: DNA

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

CGCAAAATCG ACCACGTGGT CC

(2) INFORMAÇÃO PARA SEQ ID NO.61:

(i) CARACTERÍSTICAS DA SEQUÊNCIA:

- (A) COMPRIMENTO: 23 pares de bases
- (B) TIPO: ácido nucleico
- (C) TIPO DE CADEIA: simples
- (D) TOPOLOGIA: linear

(ii) TIPO DE MOLÉCULA: DNA

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

GCATAAGTAA TACCACGTGG CCG

(2) INFORMAÇÃO PARA SEQ ID NO.62:

(i) CARACTERÍSTICAS DA SEQUÊNCIA:

- (A) COMPRIMENTO: 20 pares de bases
- (B) TIPO: ácido nucleico

(C) TIPO DE CADEIA: simples

(D) TOPOLOGIA: linear

(ii) TIPO DE MOLÉCULA: DNA

(xi) DESCRIÇÃO DA SEQUÊNCIA: SEQ ID NO.62:

GCAAAAAAAC CACGTGGTCC

(2) INFORMAÇÃO PARA SEQ ID NO.63:

(i) CARACTERÍSTICAS DA SEQUÊNCIA:

(A) COMPRIMENTO: 20 pares de bases

(B) TIPO: ácido nucleico

(C) TIPO DE CADEIA: simples

(D) TOPOLOGIA: linear

(ii) TIPO DE MOLÉCULA: DNA

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

GGGGGCGGAA CTCCGTTGTC

(2) INFORMAÇÃO PARA SEQ ID NO.64:

(i) CARACTERÍSTICAS DA SEQUÊNCIA:

(A) COMPRIMENTO: 20 pares de bases

(B) TIPO: ácido nucleico

(C) TIPO DE CADEIA: simples

(D) TOPOLOGIA: linear

(ii) TIPO DE MOLÉCULA: DNA

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

GGGGACCCGA TCTCTCGCTG

(2) INFORMAÇÃO PARA SEQ ID NO.65:

(i) CARACTERÍSTICAS DA SEQUÊNCIA:

- (A) COMPRIMENTO: 20 pares de bases
- (B) TIPO: ácido nucleico
- (C) TIPO DE CADEIA: simples
- (D) TOPOLOGIA: linear

(ii) TIPO DE MOLÉCULA: DNA

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

CAATAATATT TCCTTTCCTG

(2) INFORMAÇÃO PARA SEQ ID NO.66:

(i) CARACTERÍSTICAS DA SEQUÊNCIA:

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

(ii) TIPO DE MOLÉCULA: DNA

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

GTCCACGCGG CATCCACGTG CCGT

(2) INFORMAÇÃO PARA SEQ ID NO.67:

(i) CARACTERÍSTICAS DA SEQUÊNCIA:

- (A) COMPRIMENTO: 23 pares de bases
- (B) TIPO: ácido nucleico
- (C) TIPO DE CADEIA: simples
- (D) TOPOLOGIA: linear

(ii) TIPO DE MOLÉCULA: DNA

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

CGGCCACGTG CTCTATACAT GCC

(2) INFORMAÇÃO PARA SEQ ID NO.68:

(i) CARACTERÍSTICAS DA SEQUÊNCIA:

- (A) COMPRIMENTO: 20 pares de bases
- (B) TIPO: ácido nucleico
- (C) TIPO DE CADEIA: simples
- (D) TOPOLOGIA: linear

(ii) TIPO DE MOLÉCULA: DNA

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

GGACCACGTG CTTATCTTTG

(2) INFORMAÇÃO PARA SEQ ID NO.69:

(i) CARACTERÍSTICAS DA SEQUÊNCIA:

- (A) COMPRIMENTO: 22 pares de bases
- (B) TIPO: ácido nucleico
- (C) TIPO DE CADEIA: simples

(D) TOPOLOGIA: linear

(ii) TIPO DE MOLÉCULA: DNA

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

CGACCACGTG TTCCGCTACT CG

(2) INFORMAÇÃO PARA SEQ ID NO.70:

(i) CARACTERÍSTICAS DA SEQUÊNCIA:

- (A) COMPRIMENTO: 21 pares de bases
- (B) TIPO: ácido nucleico
- (C) TIPO DE CADEIA: simples
- (D) TOPOLOGIA: linear

(ii) TIPO DE MOLÉCULA: DNA

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

CGAGTAGCGA GCACGTGTTG C

(2) INFORMAÇÃO PARA SEQ ID NO.71

(i) CARACTERÍSTICAS DA SEQUÊNCIA:

- (A) COMPRIMENTO: 20 pares de bases
- (B) TIPO: ácido nucleico
- (C) TIPO DE CADEIA: simples
- (D) TOPOLOGIA: linear

(ii) TIPO DE MOLÉCULA: DNA

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

GCACCACGTG CTTACCATGT C

(2) INFORMAÇÃO PARA SEQ ID NO.72:

(i) CARACTERÍSTICAS DA SEQUÊNCIA:

- (A) COMPRIMENTO: 20 pares de bases
- (B) TIPO: ácido nucleico
- (C) TIPO DE CADEIA: simples
- (D) TOPOLOGIA: linear

(ii) TIPO DE MOLÉCULA: DNA

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

CCGCAGTCCT GGTGCTCGC

(2) INFORMAÇÃO PARA SEQ ID NO.44:

(i) CARACTERÍSTICAS DA SEQUÊNCIA:

- (A) COMPRIMENTO: 23 pares de bases
- (B) TIPO: ácido nucleico
- (C) TIPO DE CADEIA: simples
- (D) TOPOLOGIA: linear

(ii) TIPO DE MOLÉCULA: DNA

(xi) DESCRIÇÃO DA SEQUÊNCIA: SEQ ID NO.72:

GGACAAAAAG CACGTGCTAC

(2) INFORMAÇÃO PARA SEQ ID NO.73:

(i) CARACTERÍSTICAS DA SEQUÊNCIA:

- (A) COMPRIMENTO: 20 pares de bases
- (B) TIPO: ácido nucleico
- (C) TIPO DE CADEIA: simples
- (D) TOPOLOGIA: linear

(ii) TIPO DE MOLÉCULA: DNA

(xi) DESCRIÇÃO DA SEQUÊNCIA: SEQ ID NO.73:

GCAAAACTCC ACGTGGTCGG

(2) INFORMAÇÃO PARA SEQ ID NO.74:

(i) CARACTERÍSTICAS DA SEQUÊNCIA:

- (A) COMPRIMENTO: 22 pares de bases
- (B) TIPO: ácido nucleico
- (C) TIPO DE CADEIA: simples
- (D) TOPOLOGIA: linear

(ii) TIPO DE MOLÉCULA: DNA

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

GGGCAAAAAC AACAGCTGTG CG

(2) INFORMAÇÃO PARA SEQ ID NO.75:

(i) CARACTERÍSTICAS DA SEQUÊNCIA:

- (A) COMPRIMENTO: 21 pares de bases
- (B) TIPO: ácido nucleico
- (C) TIPO DE CADEIA: simples
- (D) TOPOLOGIA: linear

(ii) TIPO DE MOLÉCULA: DNA

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

GGGAAAGAGA TCAGCTGTGC G

(2) INFORMAÇÃO PARA SEQ ID NO.76:

(i) CARACTERÍSTICAS DA SEQUÊNCIA:

- (A) COMPRIMENTO: 21 pares de bases
- (B) TIPO: ácido nucleico
- (C) TIPO DE CADEIA: simples
- (D) TOPOLOGIA: linear

(ii) TIPO DE MOLÉCULA: DNA

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

GGAGAATTGA ACAGCTGACC C

(2) INFORMAÇÃO PARA SEQ ID NO.77:

(i) CARACTERÍSTICAS DA SEQUÊNCIA:

- (A) COMPRIMENTO: 23 pares de bases
- (B) TIPO: ácido nucleico
- (C) TIPO DE CADEIA: simples
- (D) TOPOLOGIA: linear

(ii) TIPO DE MOLÉCULA: DNA

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

GGGACAAACC AGTCAGCTGG

(2) INFORMAÇÃO PARA SEQ ID NO.78:

(i) CARACTERÍSTICAS DA SEQUÊNCIA:

- (A) COMPRIMENTO: 20 pares de bases
- (B) TIPO: ácido nucleico
- (C) TIPO DE CADEIA: simples
- (D) TOPOLOGIA: linear

(ii) TIPO DE MOLÉCULA: DNA

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

GGGCACAGCT GTTTAGTGGG

(2) INFORMAÇÃO PARA SEQ ID NO.79:

(i) CARACTERÍSTICAS DA SEQUÊNCIA:

- (A) COMPRIMENTO: 20 pares de bases
- (B) TIPO: ácido nucleico
- (C) TIPO DE CADEIA: simples
- (D) TOPOLOGIA: linear

(ii) TIPO DE MOLÉCULA: DNA

(xi) DESCRIÇÃO DA SEQUÊNCIA: SEQ ID NO.79:

GGCAAGCCGGA CAGCTGTTCC

(2) INFORMAÇÃO PARA SEQ ID NO.80:

(i) CARACTERÍSTICAS DA SEQUÊNCIA:

- (A) COMPRIMENTO: 20 pares de bases
- (B) TIPO: ácido nucleico

(C) TIPO DE CADEIA: simples

(D) TOPOLOGIA: linear

(ii) TIPO DE MOLÉCULA: DNA

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

GGCATTGATC AGCTGTGTGG

(2) INFORMAÇÃO PARA SEQ ID NO.81:

(i) CARACTERÍSTICAS DA SEQUÊNCIA:

(A) COMPRIMENTO: 20 pares de bases

(B) TIPO: ácido nucleico

(C) TIPO DE CADEIA: simples

(D) TOPOLOGIA: linear

(ii) TIPO DE MOLÉCULA: DNA

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

GCAAAAACCA GCTGGTCCCC

(2) INFORMAÇÃO PARA SEQ ID NO.82:

(i) CARACTERÍSTICAS DA SEQUÊNCIA:

(A) COMPRIMENTO: 21 pares de bases

(B) TIPO: ácido nucleico

(C) TIPO DE CADEIA: simples

(D) TOPOLOGIA: linear

(ii) TIPO DE MOLÉCULA: DNA

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

CGCAAGTGTA ACAGCTGGTG C

(2) INFORMAÇÃO PARA SEQ ID NO.83:

(i) CARACTERÍSTICAS DA SEQUÊNCIA:

- (A) COMPRIMENTO: 20 pares de bases
- (B) TIPO: ácido nucleico
- (C) TIPO DE CADEIA: simples
- (D) TOPOLOGIA: linear

(ii) TIPO DE MOLÉCULA: DNA

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

GGATGGTTTT TTTTTGTAC

(2) INFORMAÇÃO PARA SEQ ID NO.84:

(i) CARACTERÍSTICAS DA SEQUÊNCIA:

- (A) COMPRIMENTO: 20 pares de bases
- (B) TIPO: ácido nucleico
- (C) TIPO DE CADEIA: simples
- (D) TOPOLOGIA: linear

(ii) TIPO DE MOLÉCULA: DNA

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

GCATGATTTT CTTTTGTCC

(2) INFORMAÇÃO PARA SEQ ID NO.85:

(i) CARACTERÍSTICAS DA SEQUÊNCIA:

- (A) COMPRIMENTO: 20 pares de bases
- (B) TIPO: ácido nucleico
- (C) TIPO DE CADEIA: simples
- (D) TOPOLOGIA: linear

(ii) TIPO DE MOLÉCULA: DNA

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

CAGAGTTTTT TTGAGCCCCC

(2) INFORMAÇÃO PARA SEQ ID NO.86:

(i) CARACTERÍSTICAS DA SEQUÊNCIA:

- (A) COMPRIMENTO: 20 pares de bases
- (B) TIPO: ácido nucleico
- (C) TIPO DE CADEIA: simples
- (D) TOPOLOGIA: linear

(ii) TIPO DE MOLÉCULA: DNA

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

GCAAAAAATA AAAATACATC

(2) INFORMAÇÃO PARA SEQ ID NO.87:

(i) CARACTERÍSTICAS DA SEQUÊNCIA:

- (A) COMPRIMENTO: 20 pares de bases
- (B) TIPO: ácido nucleico
- (C) TIPO DE CADEIA: simples

(D) TOPOLOGIA: linear

(ii) TIPO DE MOLÉCULA: DNA

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

GGCAAAAAAG TCAAAATACG

(2) INFORMAÇÃO PARA SEQ ID NO.88:

(i) CARACTERÍSTICAS DA SEQUÊNCIA:

(A) COMPRIMENTO: 20 pares de bases

(B) TIPO: ácido nucleico

(C) TIPO DE CADEIA: simples

(D) TOPOLOGIA: linear

(ii) TIPO DE MOLÉCULA: DNA

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

GCACAATAAA AAACCTTGCG

(2) INFORMAÇÃO PARA SEQ ID NO.89:

(i) CARACTERÍSTICAS DA SEQUÊNCIA:

(A) COMPRIMENTO: 20 pares de bases

(B) TIPO: ácido nucleico

(C) TIPO DE CADEIA: simples

(D) TOPOLOGIA: linear

(ii) TIPO DE MOLÉCULA: DNA

(xi) DESCRIÇÃO DA SEQUÊNCIA: SEQ ID NO.89:

CCATATGTTC ATTGTTGTCC

(2) INFORMAÇÃO PARA SEQ ID NO.90:

(i) CARACTERÍSTICAS DA SEQUÊNCIA:

- (A) COMPRIMENTO: 20 pares de bases
- (B) TIPO: ácido nucleico
- (C) TIPO DE CADEIA: simples
- (D) TOPOLOGIA: linear

(ii) TIPO DE MOLÉCULA: DNA

(xi) DESCRIÇÃO DA SEQUÊNCIA: SEQ ID NO.90:

CACAAAAATT TAGTGTGTGC

(2) INFORMAÇÃO PARA SEQ ID NO.91:

(i) CARACTERÍSTICAS DA SEQUÊNCIA:

- (A) COMPRIMENTO: 23 pares de bases
- (B) TIPO: ácido nucleico
- (C) TIPO DE CADEIA: simples
- (D) TOPOLOGIA: linear

(ii) TIPO DE MOLÉCULA: DNA

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

CGGCCCCGTG CTCTAGCCCA TGC

(2) INFORMAÇÃO PARA SEQ ID NO.92:

(i) CARACTERÍSTICAS DA SEQUÊNCIA:

- (A) COMPRIMENTO: 21 pares de bases
- (B) TIPO: ácido nucleico
- (C) TIPO DE CADEIA: simples
- (D) TOPOLOGIA: linear

(ii) TIPO DE MOLÉCULA: DNA

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

CGGGAAGTC CCAAGTGCCC C

(2) INFORMAÇÃO PARA SEQ ID NO.93:

(i) CARACTERÍSTICAS DA SEQUÊNCIA:

- (A) COMPRIMENTO: 22 pares de bases
- (B) TIPO: ácido nucleico
- (C) TIPO DE CADEIA: simples
- (D) TOPOLOGIA: linear

(ii) TIPO DE MOLÉCULA: DNA

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

CACAGGAACA TACACGGGCC CG

(2) INFORMAÇÃO PARA SEQ ID NO.94:

(i) CARACTERÍSTICAS DA SEQUÊNCIA:

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

(ii) TIPO DE MOLÉCULA: DNA

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

GGGACGGGAT GATTGACGTG CCGT

(2) INFORMAÇÃO PARA SEQ ID NO.95:

(i) CARACTERÍSTICAS DA SEQUÊNCIA:

- (A) COMPRIMENTO: 20 pares de bases
- (B) TIPO: ácido nucleico
- (C) TIPO DE CADEIA: simples
- (D) TOPOLOGIA: linear

(ii) TIPO DE MOLÉCULA: DNA

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

CGCAAGCGAC GTCAGTCCTG

(2) INFORMAÇÃO PARA SEQ ID NO.96:

(i) CARACTERÍSTICAS DA SEQUÊNCIA:

- (A) COMPRIMENTO: 20 pares de bases
- (B) TIPO: ácido nucleico
- (C) TIPO DE CADEIA: simples
- (D) TOPOLOGIA: linear

(ii) TIPO DE MOLÉCULA: DNA

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

CACCTACCAC TGATCGCGGC

(2) INFORMAÇÃO PARA SEQ ID NO.97:

(i) CARACTERÍSTICAS DA SEQUÊNCIA:

- (A) COMPRIMENTO: 20 pares de bases
- (B) TIPO: ácido nucleico
- (C) TIPO DE CADEIA: simples
- (D) TOPOLOGIA: linear

(ii) TIPO DE MOLÉCULA: DNA

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

GGACAAACAT CCCATTACCC

(2) INFORMAÇÃO PARA SEQ ID NO.98

(i) CARACTERÍSTICAS DA SEQUÊNCIA:

- (A) COMPRIMENTO: 20 pares de bases
- (B) TIPO: ácido nucleico
- (C) TIPO DE CADEIA: simples
- (D) TOPOLOGIA: linear

(ii) TIPO DE MOLÉCULA: DNA

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

GGGGATGGAA CATCGCGCTG

(2) INFORMAÇÃO PARA SEQ ID NO.99

(i) CARACTERÍSTICAS DA SEQUÊNCIA:

- (A) COMPRIMENTO: 20 pares de bases
- (B) TIPO: ácido nucleico

(C) TIPO DE CADEIA: simples

(D) TOPOLOGIA: linear

(ii) TIPO DE MOLÉCULA: DNA

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

GCACAATAAA AAACCTTGCG

(2) INFORMAÇÃO PARA SEQ ID NO.100:

(i) CARACTERÍSTICAS DA SEQUÊNCIA:

(A) COMPRIMENTO: 20 pares de bases

(B) TIPO: ácido nucleico

(C) TIPO DE CADEIA: simples

(D) TOPOLOGIA: linear

(ii) TIPO DE MOLÉCULA: DNA

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

GGGAGCCATC GACGCCGGTG

(2) INFORMAÇÃO PARA SEQ ID NO.101:

(i) CARACTERÍSTICAS DA SEQUÊNCIA:

(A) COMPRIMENTO: 20 pares de bases

(B) TIPO: ácido nucleico

(C) TIPO DE CADEIA: simples

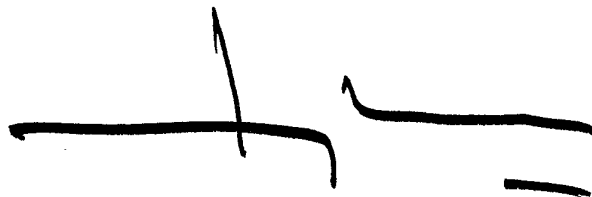
(D) TOPOLOGIA: linear

(ii) TIPO DE MOLÉCULA: DNA

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

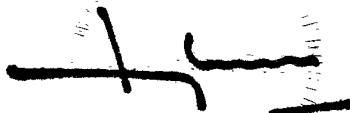
CCATAGGGGA GTTGACAGCC

Lisboa, 28 de Outubro de 1992

A handwritten signature in black ink, consisting of a long horizontal stroke with a vertical line crossing it, and a smaller horizontal stroke to the right.

J. PEREIRA DA CRUZ
Agente Oficial da Propriedade Industrial
RUA VICTOR CORDON, 10 - A 3.º
1200 LISBOA

11 19



REIVINDICAÇÕES

1ª. - Método para a purificação de Myc a partir de uma fonte de mamífero, caracterizado por compreender:

- (a) o crescimento de células de mamífero capazes de expressarem c-Myc;
- (b) a indução da expressão de c-Myc nas referidas células;
- (c) a lise da membrana das referidas células de mamífero e purificação dos núcleos a partir delas;
- (d) o tratamento dos referidos núcleos num tampão contendo DNaseI;
- (e) o tratamento dos referidos núcleos num tampão compreendendo dodecilsulfato de sódio ou ureia em concentrações superiores a 4M e separação do sedimento de núcleos da fracção do sobrenadante;
- (f) a aplicação da referida fracção do sobrenadante do passo (e) numa coluna de Sepharose CL-6B e eluição de c-Myc ligado a partir da referida coluna de DEAE Sepharose CL-6B com um gradiente de sal;
- (g) aplicação do referido c-Myc do passo (f) numa coluna de FPLC Mono-Q e eluição do c-Myc ligado com um gradiente de sal.

2ª. - Método para a detecção de complexos C1 numa amostra, caracterizado por compreender a detecção da ligação a DNA de homo-oligómeros contendo c-Myc ao motivo de DNA 5'-CACGTG-3', na sua forma de DNA de cadeia dupla.

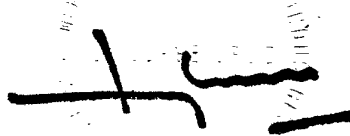
3ª. - Método para a detecção de complexos C2 numa amostra, caracterizado por compreender a detecção da ligação a DNA de hetero-oligómeros c-Myc motivo de DNA 5'-CACGTG-3', na sua forma de DNA de cadeia dupla.

4ª. - Método para a detecção de complexos C2' numa amostra, caracterizado por compreender a detecção da ligação a DNA dirigida por c-Myc ao motivo de DNA 5'-CACGTG-3', na sua forma de DNA de cadeia dupla.

5ª. - Composição de proteínas, caracterizada por compreender pelo menos um peptídeo capaz de formar um complexo C2, em que o referido peptídeo capaz de formar um complexo C2 se encontra numa fracção proteica de 26-29 KD purificada a partir de células de ovário de hamster Chinês ou de baculovírus.

6ª. - Composição de proteínas de acordo com a reivindicação 5, caracterizada por ser preparada a partir de células de ovário de hamster Chinês por um método compreendendo os passos de:

- (a) o crescimento das referidas células;
- (b) a lise da membrana das referidas células e purificação dos seus núcleos;
- (c) o tratamento dos referidos núcleos num tampão contendo DNase I;

- 
- (d) a solubilização dos referidos núcleos num tampão compreendendo dodecilsulfato de sódio ou ureia em concentrações superiores a 4M e separação do sedimento de núcleos da fracção do sobrenadante;
 - (e) a aplicação da referida fracção do sobrenadante do passo (e) numa coluna de DEAE Sepharose CL-6B e a proteína ligada do complexo C2 da referida coluna de DEAE Sepharose CL-6B eluída com um gradiente de sal; e
 - (g) a aplicação da proteína eluída do complexo C2 do passo (f) numa coluna de FPLC Mono Q e eluição da proteína ligada do complexo C2 com um gradiente de sal.

7ª. - Composição de proteína de acordo com a reivindicação 5, caracterizada por ser preparada a partir de baculovírus.

8ª. - Composição de proteínas, caracterizada por compreender pelo menos um peptídeo capaz de formar um complexo C2' na presença de c-Myc, em que o referido peptídeo capaz de formar um complexo C2' na presença de c-Myc se encontra numa fracção proteica de 40-50 KD purificada a partir de células CHO.

9ª. - Composição de proteínas de acordo com a reivindicação 8, caracterizada por ser preparada a partir de células de ovário de hamster Chinês por um método compreendendo os passos de:

- (a) o crescimento das referidas células;

- (b) a lise da membrana das referidas células e purificação dos seus núcleos;
- (c) o tratamento dos referidos núcleos num tampão contendo DNase I;
- (d) a solubilização dos referidos núcleos num tampão compreendendo dodecilsulfato de sódio ou ureia em concentrações superiores a 4M e separação do sedimento de núcleos da fracção do sobrenadante;
- (e) a aplicação da referida fracção do sobrenadante do passo (e) numa coluna de DEAE Sepharose CL-6B e a proteína ligada do complexo C2' da referida coluna de DEAE Sepharose CL-6B eluída com um gradiente de sal; e
- (g) a aplicação da proteína eluída do complexo C2 do passo (f) numa coluna de FPLC Mono Q e eluição da proteína ligada do complexo C2 com um gradiente de sal.

10^a. - Método para a classificação objectiva de compostos, incluindo fármacos para utilização em humanos, como inibidores da actividade c-Myc, caracterizado por compreender a detecção da capacidade do referido composto para inibir a formação do complexo C1, a formação do complexo C2 ou a formação do complexo C2'.

11^a. - Método de acordo com a reivindicação 10, caracterizado por a formação do referido complexo ser a formação do complexo C1.

12ª. - Método da reivindicação 10, caracterizado por a formação do referido complexo ser a formação do complexo C2.

13ª. - Método da reivindicação 10, caracterizado por a formação do referido complexo ser a formação do complexo C2'.

14ª. - Método para classificação objectiva de compostos, incluindo fármacos para utilização em humanos, como inibidores da actividade c-Myc, caracterizado por compreender a detecção da capacidade do referido composto para inibir a ligação a DNA do complexo C1, a ligação a DNA do complexo C2 ou a ligação a DNA do complexo C2'.

15ª. - Método de acordo com a reivindicação 14, caracterizado por a referida ligação a DNA ser a ligação a DNA do complexo C1.

16ª. - Método de acordo com a reivindicação 15, caracterizado por a referida ligação a DNA ser a um oligonucleótido compreendendo a sequência 5'-CACGTG-3'.

17ª. - Método de acordo com a reivindicação 15, caracterizado por a referida ligação a DNA ser a um oligonucleótido compreendendo a sequência 5'-CATGTG-3'.

18ª. - Método de acordo com a reivindicação 14, caracterizado por a referida ligação a DNA ser a ligação a DNA do complexo C2.

19ª. - Método de acordo com a reivindicação 18, caracterizado por a referida ligação a DNA ser a um oligonucleótido compreendendo a sequência 5'-CACGTG-3'.

20ª. - Método de acordo com a reivindicação 18, caracterizado por a referida ligação a DNA ser a um oligonucleótido compreendendo a sequência 5'-CATGTG-3'.

21ª. - Método de acordo com a reivindicação 14, caracterizado por a referida ligação a DNA ser a ligação a DNA do complexo C2'.

22ª. - Método de acordo com a reivindicação 21, caracterizado por a referida ligação a DNA ser a um oligonucleótido compreendendo a sequência 5'-CAGCTG-3'.

23ª. - Método para a purificação de um peptídeo capaz de formar um complexo C2 ou C2' ou uma mistura de tais peptídeos a partir de uma preparação bruta, caracterizado por compreender a extracção do referido peptídeo a partir de células de ovário de hamster Chinês e o respectivo ensaio através da detecção da capacidade do referido peptídeo para formar o referido complexo C2 ou C2'.

24ª. - Método para identificação e classificação de um composto como um inibidor da ligação a DNA do hetero-oligómero de c-Myc, caracterizado por compreender a avaliação da capacidade do referido composto para alterar a expressão de um gene repórter numa célula hospedeira, em que a expressão do referido gene repórter está operacionalmente ligado à ligação a DNA pelo referido hetero-oligómero a um oligonucleótido compreendendo a sequência 5'-CACGTG-3'.

25ª. - Método de acordo com a reivindicação 24, caracterizado por a expressão do referido gene repórter induzir uma alteração fenotípica numa célula hospedeira.

26ª. - Método de acordo com a reivindicação 24, caracterizado por o referido gene repórter ser lacZ.

27ª. - Método de acordo com a reivindicação 24, caracterizado por o referido gene repórter ser CAT.

28ª. - Método de acordo com a reivindicação 24, caracterizado por o referido gene repórter ser LEU2.

29ª. - Método de acordo com a reivindicação 24, caracterizado por a referida alteração fenotípica ser detectada pela inspecção visual da célula hospedeira.

30ª. - Método de acordo com a reivindicação 24, caracterizado por o referido hospedeiro ser S. cerevisiae.

31ª. - Método de acordo com a reivindicação 24, caracterizado por o referido hospedeiro ser uma célula de mamífero.

32ª. - Método para a identificação e classificação de um composto como um inibidor da ligação a DNA de um hetero-oligómero C'2 dirigida por c-Myc, caracterizado por compreender a avaliação da capacidade do referido composto em alterar a expressão de um gene repórter numa célula hospedeira, em que a expressão do referido gene repórter está operacionalmente ligado à ligação a DNA pelo referido hetero-oligómero a um oligonucleótido compreendendo a sequência 5'-CAGCTG-3'.

33ª. - Método de acordo com a reivindicação 32, caracterizado por a expressão do referido gene repórter induzir uma alteração fenotípica numa célula hospedeira.

34ª. - Método de acordo com a reivindicação 32, caracterizado por o referido gene repórter ser o gene LacZ.

35ª. - Método de acordo com a reivindicação 32, caracterizado por o referido gene repórter ser CAT.

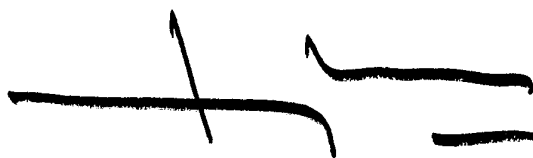
36ª. - Método de acordo com a reivindicação 32, caracterizado por o referido gene repórter ser LEU2.

37ª. - Método de acordo com a reivindicação 32, caracterizado por a referida alteração fenotípica ser detectada pela inspecção visual da célula hospedeira.

38ª. - Método de acordo com a reivindicação 32, caracterizado por a referida célula hospedeira ser S. cerevisiae.

39ª. - Método de acordo com a reivindicação 32, caracterizado por o referido hospedeiro ser uma célula de mamífero.

Lisboa, 28 de Outubro de 1992



J. PEREIRA DA CRUZ
Agente Oficial da Propriedade Industrial
RUA VICTOR CORDON, 10-A 3.ª
1200 LISBOA

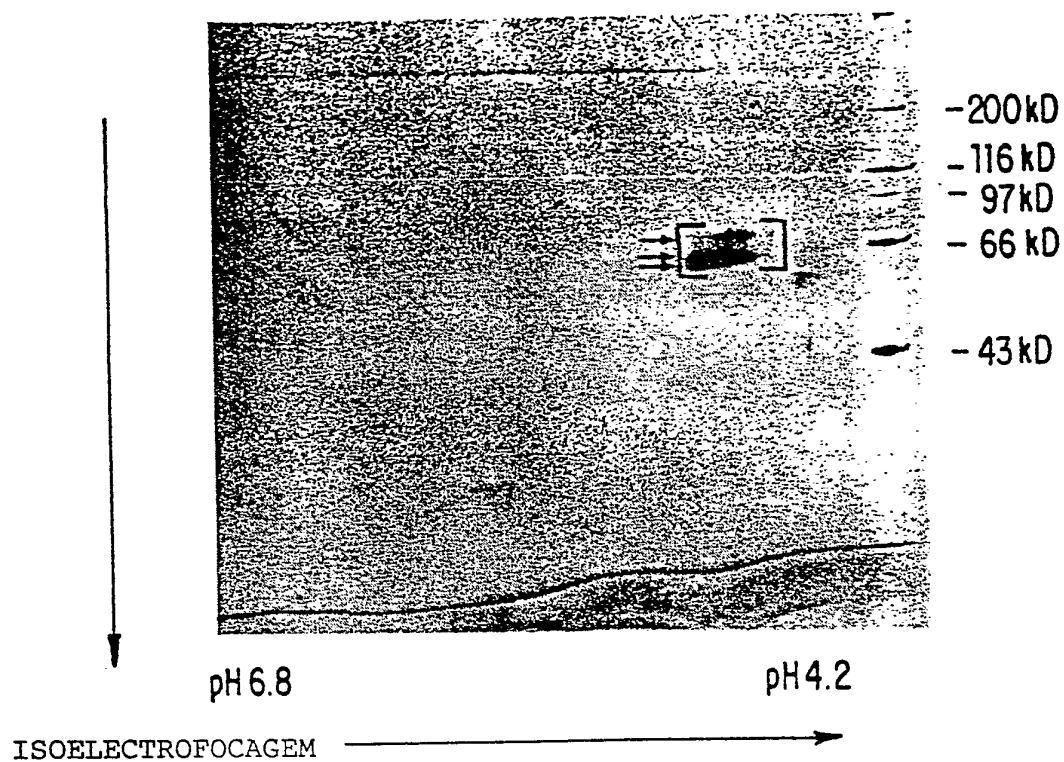


FIG. 1A

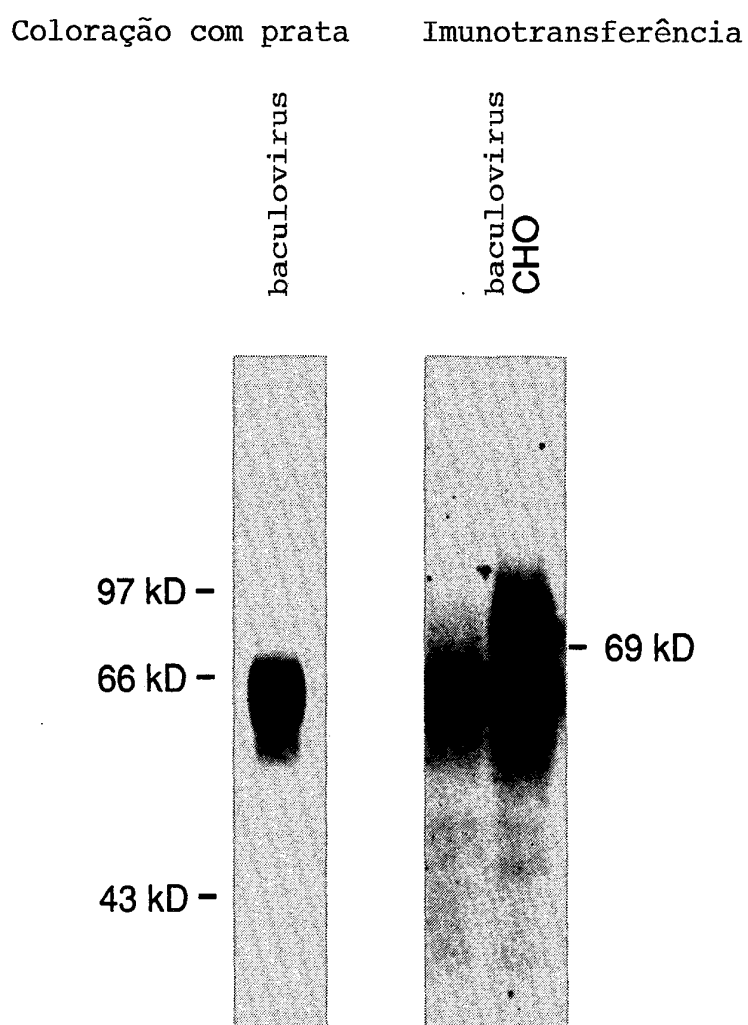
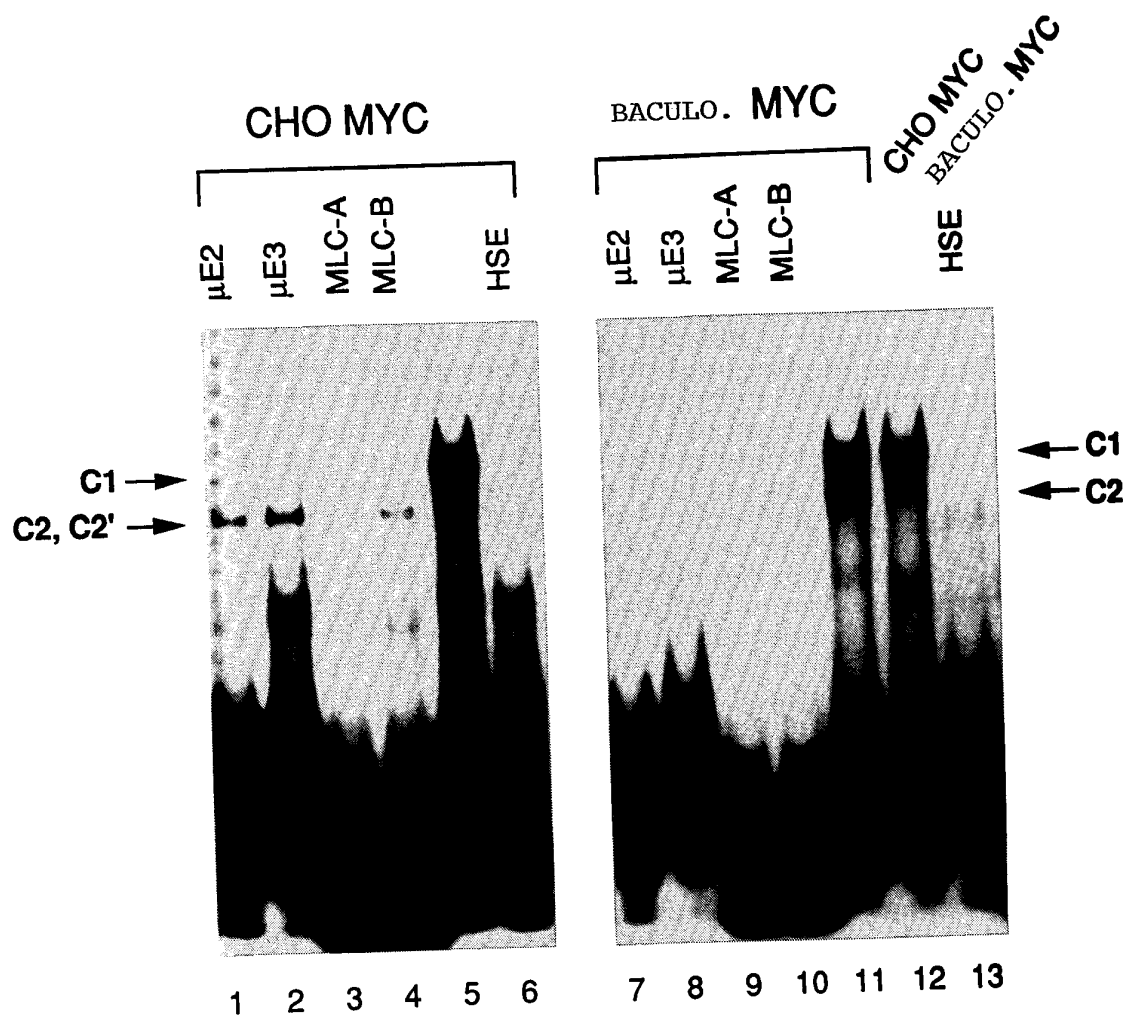


FIG. 1B



MLC-B (μ E2)	CAGCTG
μ E3	CATGTG
	CACGTG
MLC-A (KE2)	CAGGTG

FIG. 2

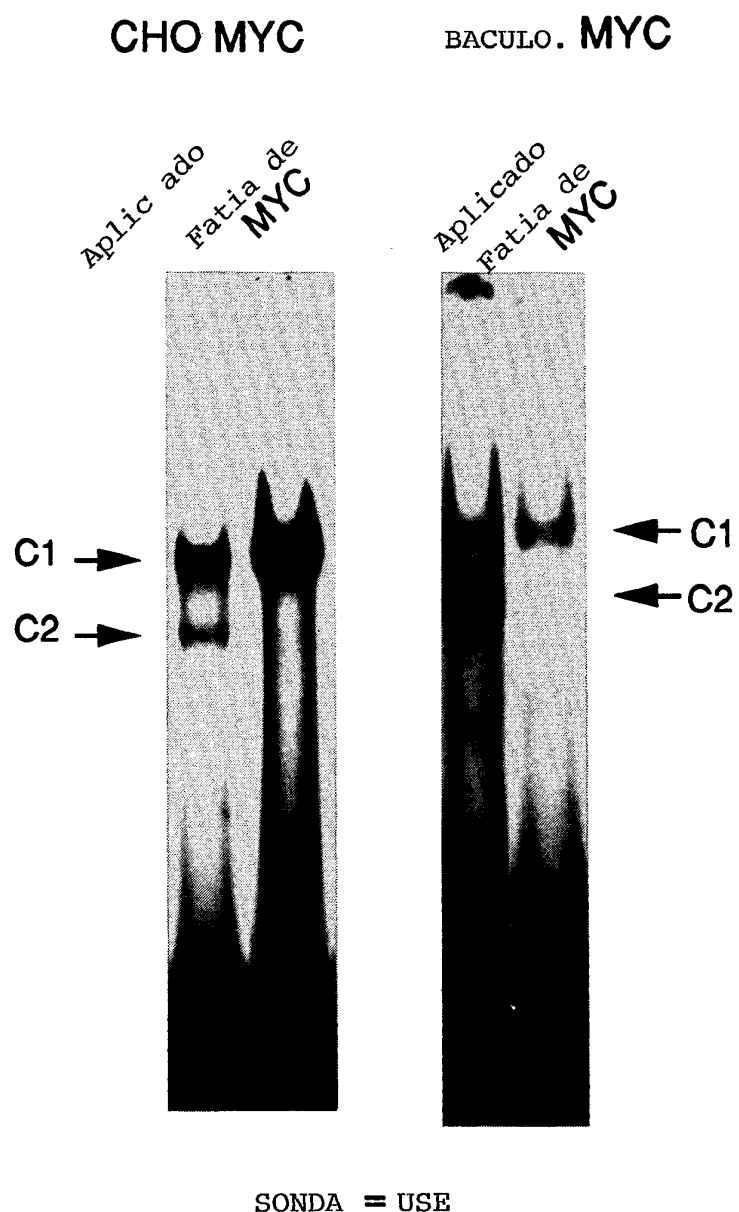
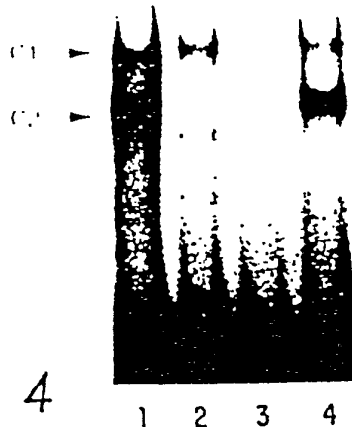


FIG. 3

[Handwritten signature]

Aplicado MYC
Fatia

26-29 KD+Myc FATIA
26-29 KD
APLICADO
FATIA Myc



CHO MYC

26-29KD+Myc(372-439)
Myc (372-439)
26-29KD+Myc(353-439)
Myc(353-439)
26-29KD+MYC FATIA
26-29 KD
FATIAMYC

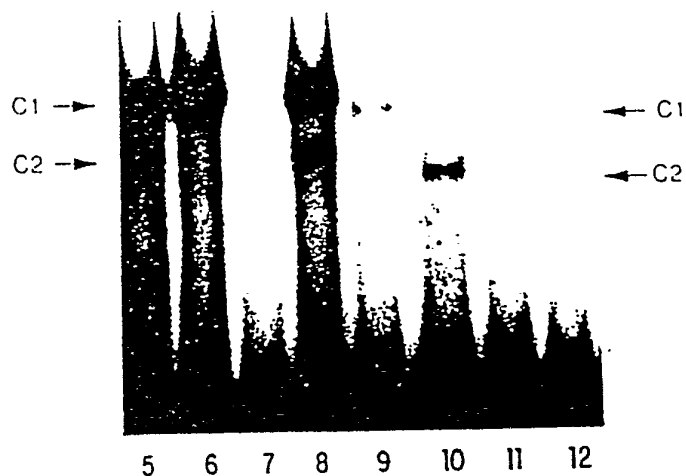
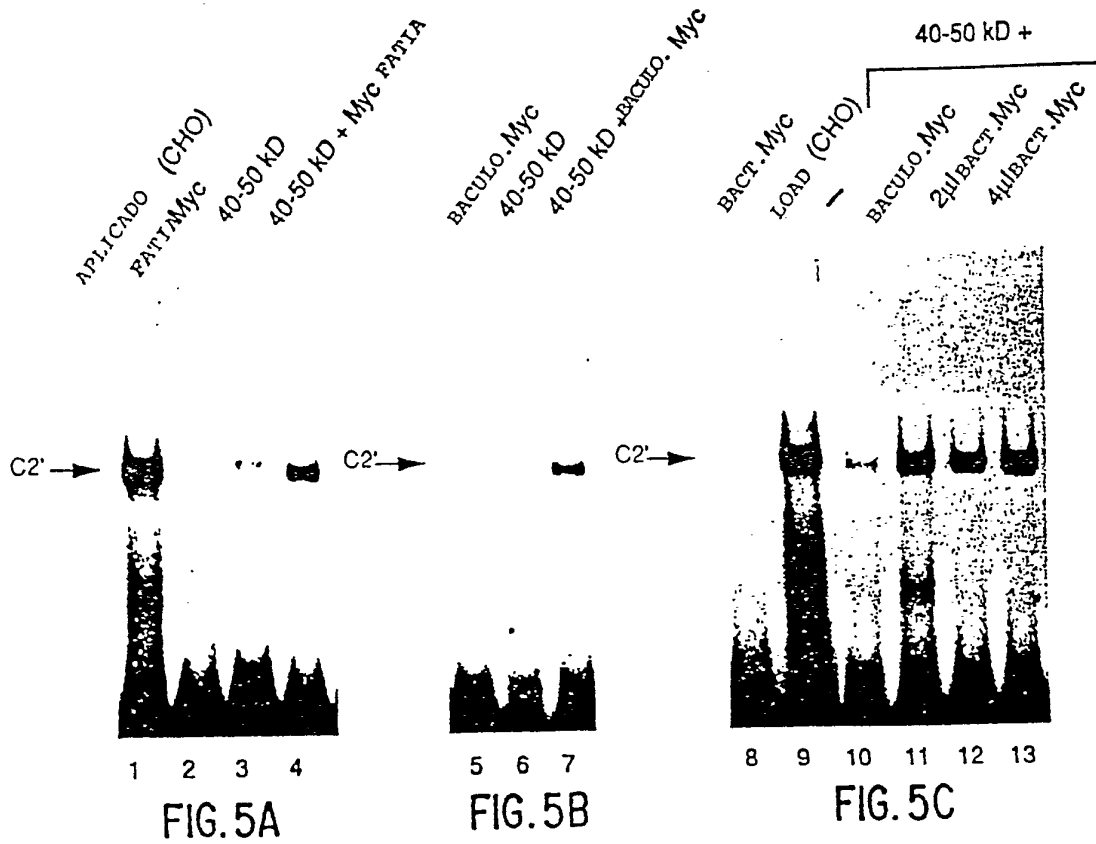
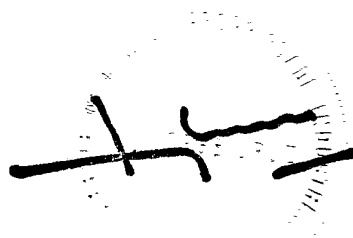


FIG. 4

SONDA = USE



SONDA = μ E2

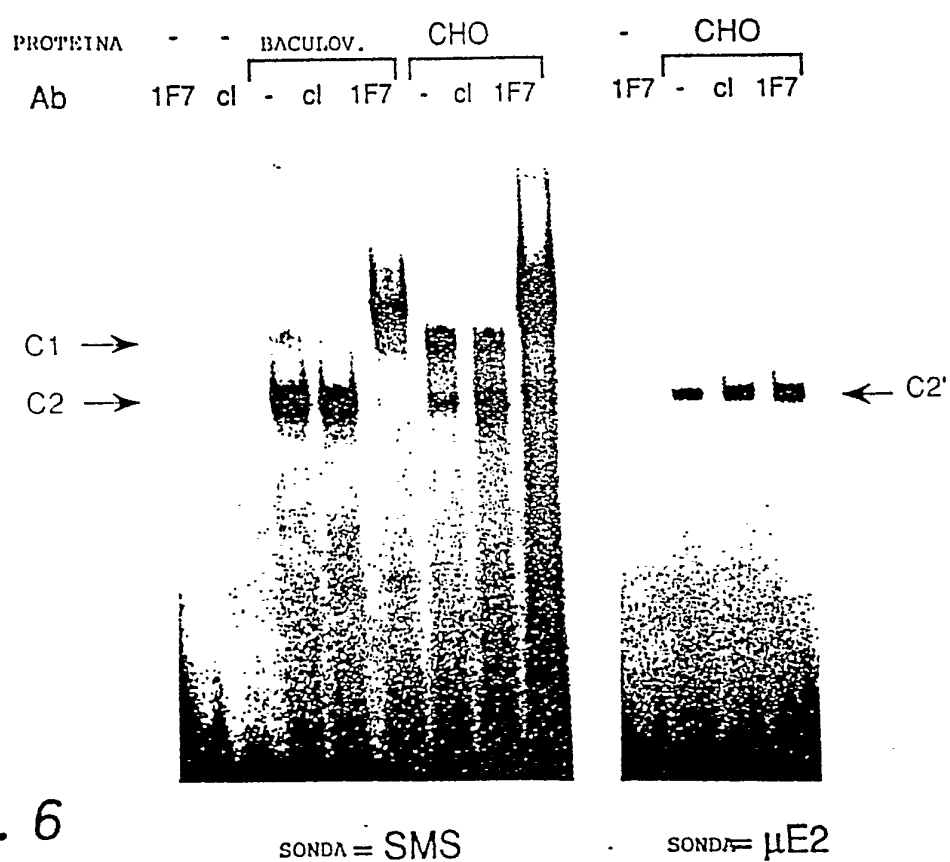


FIG. 6

GRUPO I Complexos C1 de CHO Baculovirus
Seleccinoam a sequência CACGTG

GCAGAATCTAC CACGTG CTCC
GGGGCTAC CACGTG CTTATG
GGACGAAAG CACGTG CTCCG
GCACATGAC CACGTG CTCTG 1
GGCAGAGA CACGTG CCCTGG
GGCAAA CACGTG TTATGTG *
CGAC CACGTG CTCTTCGACTTG
GCACAATTTGTAC CACGTG GCGG
GGACAACATCGAC CACGTG GCGG
GCCTGCATGAC CACGTG GACC
GCAAAATATGAC CACGTG GTAC
GGAC CACGTG CTCTTTTGTG
GGCATAAACTC CACGTG GTCC
CGGG CACGTG CTCTTCGGACTG
GGTAGCAAAAAG CACGTG CCCC 2
GGGGGATTTAAG CACGTG CTCC
CACCTATTAAC CACGTG GTAC
GACCACGGCGCATE CACGTG CCGT
GGGGAC CACGTG CTGGTTG
CACATATTAGAC CACGTG CTCC
CGGC CACGTG CTCCTGTCTACC

GGATGGA CAGCTT CTTCCTG
GCAATCCC CCGTG CTGCCC 3
GCCAAAAATGTA CAGCTG TGCC
CGGC CACGAG GTCATGAATGTG
GCAGGCTG TACGTG ACTTGG
CCGCAGTC CTGGTG CTCTGC

GRUPO II Complexos C2 Baculovirus
seleccinoam a sequência CACGTG

CACTAAGAAATAC CACGTG GCGG
GGGGATTTAAG CACGTG CTCC
CGGC CACGTG CTTCTTTCTCCG
CATAGTCGAGAGAG CACGTG CGAA
CATAAGTCAGAC CACGTG GCGG
CCCAACTAAGAC CACGTG GCGG *
CAGTCGAAGAGGC CACGTG GCGA
CGTAGGTTATTCC CACGTG GCGG
CATAAATAGGC CACGTG CTCC *
GGAAAATGTAC CACGTG CTCC
GGAACAGAC CACGTG GCTTG
GTAC CACGTG CTTTTTTGGC
CAGTCCGAGGAG CACGTG CCCC
CGGC CACGTG TCGAGCATGAGTC
CGGC CACGTG CTCGTAAATTTGC
GCGACAAAATTAC CACGTG GCGG
CGCAAAATCGAC CACGTG GTCC
GCATAAGTAATAC CACGTG GCGG
GCAAAAAAAC CACGTG GTCC

GGGGGCGGAACTCCGTTGTC 4
GGGGACCCGATCTCTCGCTG
CAATAATATTTCTTTCTCTG

	-3	-2	-1	CACGTG	1	2	3
C	1	0	16		14	5	16
G	10	3	4		6	0	1
A	7	16	1		0	1	3
T	3	2	0		1	15	1

	-3	-2	-1	CACGTG	1	2	3
C	0	1	16		8	11	13
G	11	5	3		10	1	2
A	2	13	0		0	0	1
T	6	0	0		1	7	3

FIG.7A

Grupo 3/ Três classes Sequência $\mu E2$
de sítios são seleccionadas pelos complexos
de CHO que migram como
C2 e C2'

(CAGCTG)

Sequência USE (CAGCTG)

GTCCACGCGGCATC CAGCTG CCGT
CGGC CAGCTG CTCTATACATGCC
GGAC CAGCTG CTTATCTTTG 5
CGAC CAGCTG TTCCGCTACTCG *
CGAGTAGCGAG CAGCTG TTGC
GCAC CAGCTG CTTACCATGTC
GGACAAAAAG CAGCTG CTAC
GCAAACTC CAGCTG GTCGG 6

	-3	-2	-1	CAGCTG	1	2	3
C	2	0	6		5	1	3
G	4	1	2		1	0	2
A	2	5	0		0	0	1
T	0	2	0		2	7	2

GGGCAAAAACAA CAGCTG TGCG
GGGAAAGAGAT CAGCTG TGCG
GGAGAATTGAA CAGCTG ACCC
GGGACAAACCAGT CAGCTG GCGG
GGGCA CAGCTG TTTAGTGGG
GGCAAGCGGA CAGCTG TTCC 7
GGCATTGAT CAGCTG TGTGG *
GCAAAAAC CAGCTG GTCCCC 8
CGCAAGTGTA CAGCTG GTGC

GGATGGTTTTTTTTTTGTAC
GCATGATTTTCTTTTGTCC
CAGAGTTTTTTTGAGCCCC
GCAAAAATAAAAATACATC
GGCAAAAAGTCAAAATACG
GCACAATAAAAACTTTGCG 9
CCATATGTTTCATTGTTGTC
CACAAAAATTAGTGTGTGC 10

Outros

CGGC CCGCTG CTCTAGCCCATGC
CGGGGAAGTCC CAAGTG CCCC *
CACAGGAACATA CACGGG CCGG
GGGACGGGATGATT GACGTG CCGT
CGCAAGC GACGTC AGTCCTG
CACCTACCACTGATCGGGC 11
GGACAAACATCCCATACCC *
GGGGATGGAACATCGCGCTG *
CCAGTCGGGCTAACC GGCC
GGGAGCCATCGACGCCGTG
CCATAGGGGAGTTGACAGCC

FIG. 7B

[Handwritten signature]

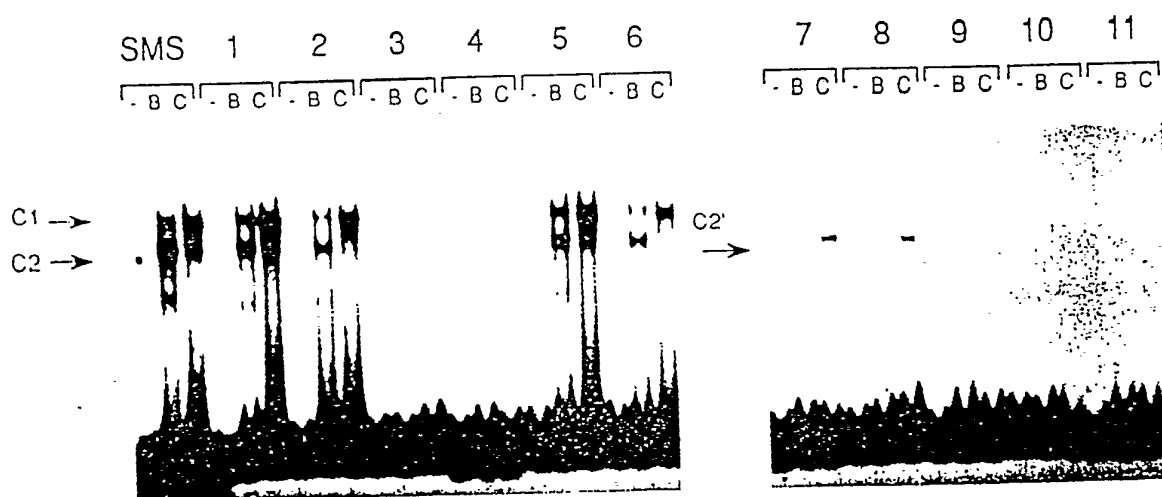


FIG. 8

Handwritten signature or mark.

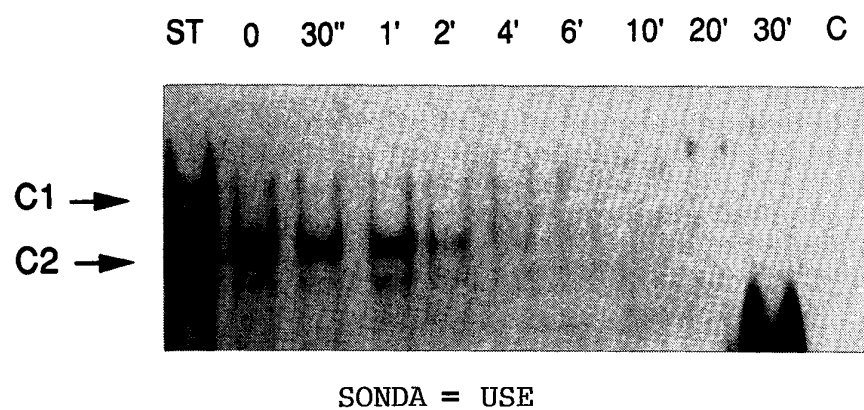


FIG. 9