



(11) Número de Publicação: **PT 857211 E**

(51) Classificação Internacional:

**C12N 15/12** (2006.01) **C12N 15/85** (2006.01)  
**C07K 14/47** (2006.01) **C12Q 1/68** (2006.01)

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

(22) Data de pedido: **1996.10.24**

(30) Prioridade(s): **1995.10.24 DE 1953949**

(43) Data de publicação do pedido: **1998.08.12**

(45) Data e BPI da concessão: **2007.02.28**  
**003/2007**

(73) Titular(es):

**BOEHRINGER INGELHEIM PHARMA GMBH & CO.KG**  
**BINGER STRASSE 173 55216 INGELHEIM AM RHEIN** **DE**

(72) Inventor(es):

**WOLFGANG NOE** **DE**  
**BARBARA ENENKEL** **DE**  
**FRANK GANNON** **DE**  
**KLAUS BERGEMANN** **DE**

(74) Mandatário:

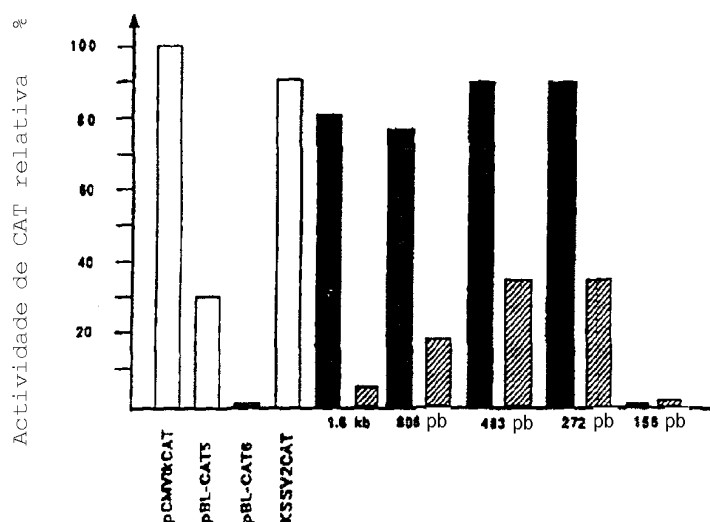
**PEDRO DA SILVA ALVES MOREIRA**  
**RUA DO PATROCÍNIO, N.º 94 1350-232 LISBOA** **PT**

(54) Epígrafe: **PROMOTOR HOMÓLOGO FORTE OBTIDO A PARTIR DE HAMSTER**

(57) Resumo:

## RESUMO

### "PROMOTOR HOMÓLOGO FORTE OBTIDO A PARTIR DE HAMSTER"



A presente invenção refere-se a um promotor homólogo forte, obtido a partir de hamster, em particular o promotor do gene que codifica para a proteína de fusão da ubiquitina-S27a. O promotor pode ser utilizado em processos para a preparação de produtos genéticos heterólogos em células de cultura, em particular células de ovários de hamster chinês.

## DESCRIÇÃO

### "PROMOTOR HOMÓLOGO FORTE OBTIDO A PARTIR DE HAMSTER"

A invenção refere-se a um promotor homólogo forte obtido a partir de hamster, a ácidos nucleicos que contêm as sequências do promotor e/ou reguladoras do gene da ubiquitina-S27a, bem como a um processo para a preparação de produtos genéticos heterólogos recorrendo a estes ácidos nucleicos.

No caso da preparação economicamente significativa de proteínas por meio de expressão de genes recombinantes em células hospedeiras eucarióticas, por exemplo células CHO, têm-se utilizado até hoje sistemas de expressão heterólogos, isto é, que são, por exemplo, os elementos promotor/intensificador, bem como elementos de terminação de origem viral. A utilização de promotores não virais em vez de sequências virais em sistemas de expressão seria vantajosa quanto à aceitação pelo público da engenharia genética e biotecnologia e contribuiria ainda para a biossegurança dos sistemas de vectores que são utilizados para a expressão de genes em culturas celulares animais.

No que se refere à ubiquitina, trata-se de um polipeptídeo de 76 aminoácidos com elevado grau de conservação que se encontra em todas as células eucarióticas em grande número e que é codificado por uma família genética diferente (Recensão: Jentsch *et al.*, 1991; Schlesinger & Bond, 1987). A ubiquitina desempenha um papel importante numa multiplicidade de processos biológicos por meio da modificação de proteínas alvo, como, p. ex., a degradação de proteínas dependente do ATP através da via

dos proteossomas da ubiquitina (Ciechanover, 1994). Devido à sua estrutura, pode-se classificar o gene da ubiquitina em dois grupos. No primeiro grupo, encontram-se reunidos os genes da poliubiquitina, nos quais se sucedem as unidades codificadoras da ubiquitina com a dimensão de 228 pb (= 76 aminoácidos) numa sequência cabeça para cauda (Jentsch *et al.*, 1991; Schlesinger & Bond, 1987). O número de unidades varia de espécie para espécie, encontrando-se na maioria dos organismos dois genes de poliubiquitina de comprimento diferente (Fornace *et al.*, 1989; Wiborg *et al.*, 1985). As regiões promotoras destes genes contêm uma caixa TATA e elementos promotor/intensificador para um indutor de choque térmico (Schlesinger & Bond, 1987).

Os genes de fusão da ubiquitina pertencem ao segundo grupo. Trata-se neste caso de fusões entre uma unidade de ubiquitina e uma proteína ribossomal (Jentsch *et al.*, 1991; Schlesinger & Bond, 1987). Ambos os genes de fusão da ubiquitina conhecidos podem ser identificados devido às diferenças no comprimento e na sequência da proteína ribossomal. Num dos casos, trata-se de uma proteína ribossomal da subunidade grande do ribossoma com um comprimento de 52 aminoácidos (Baker *et al.*, 1991) e no outro caso de uma proteína ribossomal da subunidade pequena do ribossoma (nos mamíferos designada por proteína S27a) com um comprimento dependente da espécie de 76 a 81 aminoácidos (Redman & Rechsteiner, 1989). A fracção de ubiquitina nestas proteínas de fusão suporta assim, aparentemente, a integração eficiente das proteínas ribossomais no ribossoma (Finley *et al.*, 1989).

Os genes da ubiquitina individuais são expressos nos mais variados tecidos de um organismo e em diferentes estádios de desenvolvimento em diferente escala. Assim, os genes da poliubiquitina são expressos constitutivamente em nível reduzido

que só sob o efeito de tensão aumenta fortemente (Fornace *et al.*, 1989; Schlesinger & Bond, 1987). Os genes de fusão da ubiquitina são expressos principalmente com maior intensidade na fase de crescimento exponencial. Em células terminais diferenciadas e de crescimento parado, a expressão é, pelo contrário, regulada no sentido da redução (Schlesinger & Bond, 1987; Shimbara *et al.*, 1993; Wong *et al.*, 1993).

O objectivo, cuja solução é o objecto da presente invenção, foi o de proporcionar promotores não virais fortes para processo para a preparação de produtos genéticos heterólogos em culturas de células, especialmente células de hamster.

Surpreendentemente encontrou-se um promotor de um gene, que em especial em células CHO tem uma actividade equivalente ao promotor de SV40 viral. Este gene codifica para a proteína de fusão da ubiquitina Ub/S27a. O promotor Ub/S27a de hamster é o objecto da invenção.

Um outro objecto da presente invenção é uma molécula de ácido nucleico que contém sequências promotoras e/ou outras sequências reguladoras do gene Ub/S27a. Neste caso, as sequências promotoras e/ou outras sequências reguladoras derivam do gene Ub/S27a do hamster. Em especial, o objecto da invenção é uma molécula de ácido nucleico recombinante que contém as sequências promotoras e/ou outras sequências reguladoras que estão contidas na sequência de acordo com a Fig. 5.

Um objecto preferido da invenção são moléculas de ácidos nucleicos que contém sequências do domínio que corresponde às posições -161 até -45 de acordo com a Fig. 5. É do conhecimento dos especialistas a preparação, com os métodos que são descritos

no exemplo 4, de moléculas de ácidos nucleicos que contêm sequências parciais das sequências de acordo com a invenção, em especial uma sequência parcial do domínio de -161 até -45, que pode igualmente conferir forte actividade promotora. Por conseguinte, estas sequências parciais são igualmente objecto da invenção.

É também do conhecimento do especialista a alteração da sequência promotora por meio de substituição, inserção, eliminação ou adição de uma, duas ou três bases, em relação à sequência de acordo com a Fig. 5, sem com isso diminuir significativamente a actividade promotora, que pode ser medida com os métodos descritos no exemplo 4. Por uma diminuição significativa da actividade promotora deve-se entender uma diminuição de mais de 50% do valor que é obtido para o fragmento de eliminação de 272 pb do quadro 1, na análise de CAT de acordo com o exemplo 4, em condições comparáveis. Estas variantes da sequência promotora são, por conseguinte, incluídas no âmbito das reivindicações na invenção.

Nas moléculas de ácidos nucleicos de acordo com a invenção, as sequências promotoras e/ou reguladoras são vantajosamente acopladas a um gene funcional, de forma que este gene possa ser expresso sob o controlo destas sequências. Um gene destes pode, por exemplo, codificar para um activador de plasminogénio de tecidos (documento EP 0093619), um activador de plasminogénio de *second-generation* p. ex. tnk-t-PA (documento WO 93/24635), interferão, p. ex. interferão- $\alpha$  (documento EP 0595241), interferão- $\beta$  (documentos EP 0041313, EP 0287075), Interferão- $\gamma$  (documento EP 01463 54) ou interferão- $\omega$  (documento EP 0170204), factor de necrose de tumores (documento EP 0168214), eritropoietina (documento WO 86/03520), factor de estimulação de

colónias de granulócitos (documentos EP 0220520, EP 0237545) ou superoxidodismutase de manganês (documento EP 0282899) que são conhecidos do estado da técnica. Pode-se tratar também de um gene que codifica para uma cadeia de imunoglobulina, os domínios variáveis de uma cadeia de imunoglobulina, um anticorpo humanizado (documentos EP 0230400, EP 0451216), um anticorpo de cadeia simples, etc.. Em especial, um gene destes pode codificar para uma cadeia de imunoglobulina humanizada que é específica para variante CD44 (documento WO 95/33771). Adequadamente, uma molécula de ácido nucleico destas pode ser um vector de expressão (Sambrook et al., 16.3-16.29, 1989). São também objectos da invenção em especial os vectores de expressão tais que depois da introdução numa célula hospedeira eucariótica são integrados no seu genoma por recombinação.

Um outro aspecto da presente invenção é uma célula hospedeira na qual foi introduzida uma das moléculas de ácidos nucleicos mencionadas. De preferência, introduz-se numa célula hospedeira destas um vector de expressão que contém o gene para um produto genético heterólogo acoplado ao promotor do gene Ub/S27a e/ou outras sequências reguladoras. A célula hospedeira de acordo com a invenção é de um modo preferido uma célula de mamífero. Em especial, a célula hospedeira pode ser uma célula de hamster, p. ex. uma célula CHO (CHO = *Chinese hamster ovary*, Urlaub et Chasin, 1980; cf. também Kaufman, 1987 bem como referências incluídas, bem como Sambrook et al., 16.28-16.29, 1989) BHK- (BHK = *Baby hamster kidney*) ou célula hibridoma, sendo especialmente preferida uma célula CHO.

Um outro aspecto da presente invenção é um processo para a preparação de um produto genético heterólogo numa célula hospedeira eucariótica, de um modo preferido uma célula de

hamster, de um modo especialmente preferido uma célula CHO, caracterizado por se introduzir uma das moléculas de ácido nucleico mencionadas na célula hospedeira eucariótica, se cultivar a célula hospedeira e se isolar o produto genético sintetizado. No processo de acordo com a invenção, o produto genético heterólogo é expresso sob o controlo de sequências promotoras e/ou sequências reguladoras do gene Ub/S27a, de um modo preferido de hamster. Vantajosamente podem-se utilizar para um processo destes moléculas de ácidos nucleicos que contêm as sequências promotoras e/ou sequências reguladoras que são contidas na Fig. 5. É especialmente preferido um processo no qual as sequências promotoras estão contidas na sequência que corresponde às posições -161 até -45 de acordo com a Fig. 5. Também neste caso, é do conhecimento do especialista a preparação de sequências parciais com actividade promotora ou variantes equivalentes das sequências divulgadas.

Um forte promotor homólogo do hamster, que é muito vantajoso para a preparação de produtos genéticos heterólogos em células de hamster, em especial células CHO, é agora apresentado pela primeira vez. Em especial, a invenção refere-se a um promotor homólogo forte do hamster que apresenta nas células CHO em análise de CAT de acordo com o exemplo 4 uma maior actividade do que o promotor de cinase de timidina de *Herpes simplex*. Vantajosamente, um promotor destes apresenta uma actividade de transcrição que se situa pelo menos na mesma ordem de grandeza que a do promotor SV40. Na mesma ordem de grandeza significa neste caso que o promotor de acordo com a invenção tem pelo menos 50%, melhor pelo menos 80%, de um modo ainda mais preferido pelo menos 90% da actividade do promotor SV40 em análise de CAT de acordo com o exemplo 4. É preferido um promotor destes caracterizado por possuir pelo menos uma das

características: apresenta um domínio de sequência rico em GC, local de ligação Sp1, elemento polipirimidina, ausência de uma caixa TATA. É especialmente preferido um promotor destes que apresenta um local de ligação Sp1 na ausência de uma caixa TATA. Além disso, é preferido um promotor deste tipo que seja activado constitutivamente e em especial que seja igualmente activo em condições de cultura de células com soro, pobres em soro e isentas de soro. Numa outra forma de realização, trata-se de um promotor indutível, em especial de um promotor que é activado pela remoção do soro. Uma forma de realização especialmente vantajosa é um promotor com uma sequência que se encontra na Fig. 5. É especialmente preferida neste caso uma sequência que está contida na sequência que corresponde às posições -161 a -45 de acordo com a Fig. 5.

Um outro aspecto da presente invenção é um processo para a expressão de um produto genético heterólogo em células de hamster no âmbito das reivindicações da patente, de um modo preferido células CHO, que é caracterizado por o produto genético ser expresso sob o controlo de um promotor homólogo forte do hamster. Em formas de realização preferidas, um promotor deste tipo é caracterizado pelas características descritas no parágrafo anterior.

Um outro aspecto da presente invenção é a utilização de um promotor, tal como descrito anteriormente, para a preparação de um produto genético heterólogo em células de hamster, de um modo preferido células CHO ou BHK.

As sequências promotoras descritas podem ser introduzidas numa cassette de expressão em conexão funcional com outras sequências reguladoras. Por exemplo, podem estar funcionalmente

acopladas com sequências intensificadoras e aumentarem assim a actividade de transcrição. Neste caso, pode-se tratar de um ou vários intensificadores e/ou várias cópias de uma sequência intensificadora. Por exemplo, pode-se utilizar neste caso um intensificador CMV ou SV40. O intensificador CMV humano faz parte neste caso dos intensificadores mais fortes identificados até hoje. Um exemplo de um intensificador indutível é o intensificador de metalotioneína, que pode ser estimulado por glucocorticóides ou metais pesados. Uma outra modificação possível seria, p. ex., a introdução de múltiplos locais de ligação Sp1. As sequências promotoras podem ainda ser combinadas com sequências reguladoras que constituem um controlo/regulação da actividade de transcrição. Assim, o promotor pode ser tornado reprimível/indutível. Tal pode acontecer, por exemplo, por meio de acoplagem com sequências que constituem locais de ligação para factores de transcrição reguladores positivos ou negativos. O factor de transcrição acima mencionado SP-1, por exemplo, tem uma influência positiva na actividade de transcrição. Um outro exemplo é o local de ligação para a proteína activadora AP-1, que pode exercer um efeito de modo tanto positivo como negativo sobre a transcrição. A actividade do AP-1 pode ser controlada pelos mais variados factores como, p. ex., factores de crescimento, citocinas e soro (Faisst et Meyer, 1992, e referências respectivas). A eficácia de transcrição pode também ser aumentada por modificação da sequência promotora por meio de mutação (substituição, inserção, eliminação) de uma, duas, três ou mais bases e medindo depois no teste CAT de acordo com o exemplo 4 se a actividade promotora aumentou desta forma. Por meio das medidas descritas neste parágrafo pode-se obter uma cassette de expressão optimizada, que é de grande utilidade para a expressão de produtos genéticos heterólogos, em especial células CHO. Portanto, uma cassette de expressão obtida por meio

de uma ou várias destas medidas é igualmente objecto da invenção.

Em análises de *ADNaseI-Footprint* e de mutação pode-se examinar quais os factores que influenciam a expressão e se a actividade promotora pode ser mais aumentada por meio da eliminação de elementos de controlo negativos eventualmente presentes e pela introdução de outros elementos de controlo positivos. Exames de outros grupos de trabalho têm, além disso, mostrado que a expressão do gene Ub/S27a pode ser regulada muito visivelmente por meio de diferentes factores. Assim, o grupo de Shimbara revelou que numa diferenciação de terminais *in vitro* de linhas de células de leucemia humana (HL-60, K562, U937, TH01) por meio da adição de diferentes substâncias como TPA (acetato de 12-*O*-tetra-decanoilforfol-13), DMSO, ácido retinóico e 1,25-di-hidroxivitamina D3 no meio de cultura se regulou no sentido da diminuição a expressão do gene Ub/S27a (Shimbara *et al.*; 1993). E o grupo de Wong estabeleceu uma sobreexpressão do gene Ub/S27a em células de carcinoma do cólon (Wong *et al.*, 1993). A expressão genética correlacionou-se neste caso com os estádios clínicos do tumor com uma maior expressão num cancro avançado.

A realização da invenção é possível para o especialista com o conhecimento do conteúdo divulgado deste pedido de patente, por meio de métodos conhecidos em si, bem como de acordo com os exemplos.

O gene Ub/S27a completo, a região não traduzida 5' do gene Ub/S27a ou fragmentos seleccionados desta podem ser obtidos por meio de vários métodos padrão com o conhecimento das sequências de acordo com Fig. 1, 2 e 5. A partir da sequência de acordo com

a Fig. 5 pode-se, por exemplo, seleccionar uma secção adequada e sintetizar quimicamente uma sonda de oligonucleótidos que tem a sequência desta secção (Sambrook *et al.*, 11.3-11.44, 1989). Com uma sonda destas pode-se clonar por meio de hibridação a partir de uma biblioteca genómica (Saambrook *et al.*, 9 4-9.62, 11.45-11.61, 1989) do hamster o gene Ub/S27a ou a sua região não traduzida a 5'. A região a 5' não traduzida ou fragmentos especiais desta podem também ser facilmente obtidos por meio de amplificação por PCR com os iniciadores correspondentes, a partir de uma biblioteca genómica (Sambrook *et al.*, 14.5-14.35, 1989). Estes métodos são especialmente adequados para a preparação de fragmentos seleccionados da região promotora por exemplo do domínio de -161 a -45 ou de uma secção deste domínio. Os fragmentos da região não traduzida a 5', tal como os indicados por exemplo no quadro 1, podem também ser obtidos por meio de digestão limitada com exonuclease III a partir de fragmentos de ADN maiores (Sambrook *et al.*, 15.14-15.19; 5.84-5.85, 1989). Moléculas de ADN deste tipo podem também ser quimicamente sintetizadas ou geradas a partir de fragmentos quimicamente sintetizados por meio de ligação. O gene Ub/S27a de uma outra espécie, de um modo preferido de uma espécie mamífera, ou a sua região não traduzida a 5' pode ser isolado por meio de hibridação cruzada com sondas do domínio não traduzido a 5' do gene Ub/S27a de hamster ou possivelmente sondas da fracção S27a do gene Ub/S27a de hamster.

Existe no estado da técnica uma variedade de vectores de expressão à disposição que podem ser utilizados no contexto da presente invenção. As sequências promotora e/ou reguladora de acordo com a invenção podem ser integradas no lugar dos elementos promotores existentes nestes vectores e controlar assim a expressão do respectivo gene, que deve ser expresso com

este vector. Os vectores disponíveis no comércio, que se adequam bem para a integração do promotor de acordo com a invenção, são por exemplo o vector pCAT-Basic (Promega; sequência compilada acessível pelo n.º de acesso EMBL X65322) ou o vector pCAT-Enhancer que contém ainda um intensificador de SV40 (Promega; sequência compilada acessível pelo n.º de acesso EMBL X65319). Um exemplo para um vector sem promotor é o plasmídeo pBL-CAT6 (Boshart *et al.*, 1992; ver também Luckow *et Schutz*, 1987). As sequências promotoras de acordo com a invenção podem ser ligadas aos sítios de clivagem de restrição de HindIII, SphI, PstI, SalI, XbaI, BamHI, BglII ou XhoI deste vector e ser assim acopladas ao gene da transferase de cloranfenicol (CAT) contido neste vector de forma funcional. Em vez do gene CAT pode-se também integrar neste vector um outro gene, possivelmente para o activador de plasminogénio de tecido. O gene CAT pode, p. ex., ser removido do vector pBL-CAT6 por meio de uma dupla digestão com as enzimas de restrição XhoI e ClaI. A região não traduzida a 3' de SV40 eliminada nesta operação, que contém a sequência intrão e o sinal de poliadenilação, pode eventualmente (isto é, desde que a sua função não tenha sido absorvida pelas sequências a 3' próprias do gene) depois ser novamente incorporada, amplificando-se, p. ex., esta região de SV40 por meio de PCR e dotando assim com sítios de clivagem de restrição adequados nas duas extremidades, a fim de facilitar a recloneagem subsequente, p. ex. com a introdução de um outro gene desejado. Uma outra estratégia de substituição do gene CAT por um outro gene consiste na introdução, em primeiro lugar, por meio de PCR e oligonucleótidos em mutagénese adequados, de um sítio de clivagem de restrição único no terminal 5' da região de SV40 no vector pBL-CAT6, que possibilita subsequentemente a remoção orientada do gene CAT. Uma outra possibilidade consiste na utilização do vector pLUC, que em princípio é construído como

pBL-CAT6, mas que, em vez do gene CAT, contém um gene relator de luciferase. Este é facilmente removido por uma dupla digestão por XhoI/EcoNI, mantendo-se a sequência de SV40 no vector.

Para gerar linhas de células estáveis com elevado grau de expressão do gene heterólogo, escolhe-se vantajosamente um vector que confirma a selecção de transformantes que integram o vector no seu cromossoma e que amplificaram o gene heterólogo integrado. Os vectores para este processo de selecção/amplificação tal como o processo DHFR/metotrexato em linhas de células deficientes de DHFR (p. ex., CHO-DUKX) estão igualmente contidas no estado da técnica (Sambrook *et al.*, 16.9-16.15; 16.28-16.29, 1989).

O processo para a inserção dos vectores obtidos em células hospedeiras, tal como a selecção e cultura dos transformantes, pode ser deduzido pelo especialista de obras padrão (p. ex. Sambrook *et al.*, 16.3-16.81, 1989)

À parte da optimização do promotor existe ainda um ponto de partida completamente diferente para o melhoramento do rendimento do produto, o qual é por meio de um efeito com dose genética. Com um crescente número de cópias da construção do vector integrada no genoma deveria também aumentar a quantidade de transcrições produzidas. Uma amplificação espontânea das construções introduzidas, que numa integração estável tem como resultado várias cópias, pode ser conseguida por meio da utilização de uma sequência com propriedades de promoção da amplificação, as chamadas "*amplification promoting sequences*". Uma sequência assim, com um teor de AT muito elevado, foi p. ex. isolada da região intergenes não transcrita do gene ribossomal de ratinho (Wegner *et al.*, 1989) e foi já utilizada com êxito

para a inibição estável e eficaz da replicação de HIV-1 por ARN anti-sentido (Meyer *et al.*, 1993). Um domínio de sequência com a dimensão de 49 pb da região não traduzida a 5' de Ub/S27a (posição -1477 até -1429; Fig. 4) com um elevado teor de AT de 88% apresenta grande homologia com as sequências promotoras da amplificação acima descritas do ratinho. Pode-se provar que este domínio de sequência de CHO possui igualmente estas propriedades por meio da utilização de construções promotoras de Ub/S27a, às quais esta sequência de CHO está ligada.

### ***Descrição das figuras***

Fig. 1: Comparação da sequência de ADN do clone completo de ADNc de ubiquitina Ub/S27a a partir de células CHO com a sequência de ADNc humana. Comparação da sequência de ADNc Ub/S27a de CHO com a sequência de ADNc humana (Adams *et al.*, 1992). As identidades no interior da sequência são salientadas por "\*". A fracção de ubiquitina é caracterizada por linhas duplas sobre a sequência. As posições dos três intrões são marcadas com triângulos.

\_\_\_\_\_ Sinal poli A        ::: codão de início +++codão de paragem

Fig. 2: Sequência de aminoácidos da proteína de fusão da ubiquitina Ub/S27a. Comparação da sequência de aminoácidos Ub/S27a derivada da sequência de ADNc de CHO com a sequência de aminoácidos humana (Adams *et al.*, 1992). As identidades no interior da sequência são caracterizadas por "\*". A unidade de ubiquitina de 76 aminoácidos é salientada por linhas duplas sobre a sequência.

*Fig. 3: Análise do nível de transcrição de Ub/S27a em células CHO.* O ARN total citoplasmático desnaturado (10 µg) de células CHO cultivadas em soro e sem soro foi separado por electroforese num gel de agarose que continha formaldeído e transferido para uma membrana de nylon. Como sonda de hibridação utilizou-se o ADNc de Ub/S27a (508 pb) de CHO marcado com <sup>32</sup>P. O tempo de exposição da película de raios X foi de 3 horas.

- 1 + 3 células CHO cultivadas em soro
- 2 + 4 células CHO cultivadas sem soro
- 5 células CHO cultivadas com soro foram cultivadas durante 24 horas em meio isento de soro
- 6 células CHO cultivadas sem soro foram cultivadas durante 24 horas em meio com soro

*Fig. 4: Estratégia de sequenciação da região não traduzida a 5' de Ub/S27a.* A região não traduzida a 5' com 2,5 kb de comprimento do gene Ub/S27a é representada de forma esquemática. Foram sequenciados tanto fragmentos de restrição subclonados como também clones de eliminação gerados por digestão por exonuclease III, sendo a escala e a orientação das sequenciações indicadas por setas.

*Fig. 5: Sequência de ADN genómico da região não traduzida a 5' de Ub/S27a.* Elucidação das figuras:

	Sítios de clivagem de restrição
- >	Extremidade 5' do clone de eliminação do promotor
*	Extremidade 3' do clone de eliminação do promotor
^^^^^^	Homologia com sequências estimuladoras da amplificação.
""""""	Domínio de sequência rica em polipirimidina

```

:::          Codão de início
+1, +13      Sítios de início de transcrição determinados por
              mapeamento por nuclease S1
+ 85         Extremidade 5' obtida através de extensão do
              iniciador
>>>>>      Sítio de ligação Sp 1

```

Os nucleótidos estão numerados em relação ao sítio de início da transcrição, que é designada com +1. Os sítios de clivagem de restrição para SacII e EagI, que são específicos de domínios de sequência ricos em GC, bem como os sítios de clivagem de restrição destacados com fins de subclonagem para EcoRI, HincII e HindIII estão destacados na sequência com um sublinhado. Para identificação da sequência do intrão utilizaram-se letras minúsculas. A sequência de aminoácidos é reproduzida sob a sequência de ADN.

*Fig. 6: Determinação do ponto de início da transcrição por meio de extensão do iniciador.* A análise de extensão do iniciador foi realizada com 5 µg de ARN total citoplasmático de células CHO cultivadas em soro (Pista 1). Como controlo utilizou-se 5 µg de ARNt de levedura (Pista 2). O iniciador com marcação na extremidade por <sup>32</sup>P utilizado hibridou com a fracção S27a do ARNm de Ub/S27a (nucleótidos +256 a +276 na sequência de ADNc). Os produtos de extensão foram separados num gel de poliacrilamida desnaturante a 6%. Como marcador de dimensão utilizou-se uma reacção de sequenciação. O comprimento do produto de extensão, cuja posição é indicada por uma seta, era de 304 nucleótidos.




*Fig. 7: Determinação do ponto de início da transcrição por meio de mapeamento por nuclease S1.* Uma das amostras de cadeia

simples com marcação final  $^{32}\text{P}$  correspondente à sequência de cadeia contrária, que abrange toda a área da região não traduzida a 5' (-2289 a +176) foi hibridada com 5  $\mu\text{g}$  de ARN total (pista 2) ou ARN total citoplasmático (pista 3) de células CHO cultivadas com soro. Como controlo utilizou-se 5  $\mu\text{g}$  de ARNt de levedura (Pista 1). Os fragmentos protegidos antes da decomposição por nuclease S1 foram separados num gel de poliacrilamida desnaturante a 8%. Como marcador de dimensão utilizou-se uma reacção de sequenciação, na qual foi sequenciada a cadeia de ADN complementar da amostra de hibridação utilizada. O iniciador utilizado para a sequenciação hibridou neste caso com a sequência de nucleótidos de +157 até +176. As posições dos fragmentos de ADN protegidos da decomposição estão marcadas por setas na sequência destacada. Ao ponto distal de início da transcrição foi atribuída a posição de nucleótido +1.



*Fig. 8: Análise funcional da actividade promotora de Ub/S27a.* Construções de vector, em que as eliminações em série da região de flanco a 5' do gene Ub/S27a em ambas as orientações tinham sido fundidas com o gene relator CAT, foram utilizadas para a transfecção transitória de células CHO cultivadas com soro. Como controlo foram utilizados plasmídeos em que o gene relator CAT estava sob o controlo de um promotor viral. No total foram realizadas quatro experiências de transfecção independentes entre si.

A actividade de CAT relativa das diferentes construções de vectores é indicada como percentagem da actividade de CAT em células CHO transfectadas por pCMVtkCAT, que foi assumida como 100%, e representa o valor médio (desvio padrão em todos os casos  $\leq 5\%$ ) das quatro experiências de transfecção. Várias actividades de CAT foram corrigidas, relativamente às

quantidades de proteínas utilizadas bem como à eficácia da transfecção, que foi determinada por meio da medição da actividade de  $\beta$ -galactosidase do plasmídeo de controlo pCMVtklacZ co-transfectado.

-  Plasmídeo de controlo: sem promotor ou promotores virais (tk, SV40)
-  região a 5' não traduzida de Ub/S27a: orientação 5'-3' em pBL-CAT6
-  região a 5' não traduzida de Ub/S27a: orientação 3'-5' em pBL-CAT6

*Fig. 9: Análise funcional da actividade do promotor de Ub/S27a em células BHK transfectadas transitoriamente (BHK21).*

-  Plasmídeo de controlo: sem promotor ou promotores virais (tk, SV40)
-  região a 5' não traduzida de Ub/S27a: orientação 5'-3' em pBL-CAT6

*Fig. 10: Análise funcional da actividade do promotor de Ub/S27a em células CHO transfectadas de modo estável (I). É representada a actividade de CAT em células que foram transfectadas de forma estável com o vector pCMVtkCAT (promotor tk/intensificador CMV) ou com o vector pBL-CAT5 (promotor tk) (cf. exemplos/plasmídeo, exemplo 6). São mostrados em todos os casos vários clones diferentes. A actividade de CAT é normalizada pela actividade do clone pCMVtkCAT 6, que tinha a*

actividade mais elevada, (= 100%). O gráfico destina-se à comparação com as figuras 11 e 12.

*Fig. 11: Análise funcional da actividade do promotor de Ub/S27a em células CHO transfectadas de modo estável (II).* É mostrada a actividade de CAT de diferentes construções de vectores, que é representada como percentagem da actividade de CAT do clone 6 das células CHO transfectadas com pCMVtkCAT da Fig. 10. Fragmentos da região promotora de acordo com o quadro 1 foram integrados em pBL-CAT6 na orientação 5' -> 3'. As abcissas indicam qual o fragmento que foi utilizado. É mostrada a actividade de cada um dos vários clones diferentes.

*Fig. 12: Análise funcional da actividade do promotor de Ub/S27a em células CHO transfectadas de modo estável (III).* É mostrada a actividade de CAT de diferentes construções de vectores, que é representada como percentagem da actividade de CAT do clone 6 das células CHO transfectadas com pCMVtkCAT da Fig. 10. Fragmentos da região promotora de acordo com o quadro 1 foram integrados em pBL-CAT6 na orientação 5' -> 3'. As abcissas indicam qual o fragmento que foi utilizado. É mostrada a actividade de cada um dos vários clones diferentes.

*Fig. 13: Análise funcional da actividade do promotor de Ub/S27a em células CHO transfectadas de modo estável (IV).* Esta figura representa um resumo dos dados das figuras 10 a 12. São mostrados os clones de células com a maior e a menor expressão de CAT para cada construção de vector, que foi utilizada nesta experiência (só plasmídeos com orientação 5' -> 3' do promotor de Ub/S27a)

## **Exemplos**

### **Plasmídeos**

Para a realização de várias manipulações genéticas, foi utilizado o vector pBluescript SK- (Stratagene Cloning Systems, Stratagene GmbH, Heidelberg, Alemanha). Para a análise de expressão utilizaram-se vectores de expressão eucarióticos que continham o gene da acetiltransferase de cloranfenicol (CAT) bacteriana, utilizado como gene relator. O gene CAT é então fundido com diferentes promotores virais. Em pBL-CAT5 é o promotor da cinase de timidina (tk) do vírus *Herpes simplex* (Boshart *et al.*, 1992; sequência compilada acessível pelo n.º de acesso do GenBank M80483), em pCMVtkCAT o promotor tk em combinação com um intensificador do citomegalovírus humano (CMV) (Christoph Winkler, Würzburg) e no pKSSV2CAT o promotor SV40 (Tsonis *et al.* 1988). O pBL-CAT6 serve como controlo negativo e contém um gene CAT sem promotor (Boshart *et al.*, 1992; sequência compilada acessível pelo n.º de acesso do GenBank M80484). No vector pCMVtklacZ encontra-se a expressão do gene da  $\beta$ -galactosidase bacteriana sob o controlo da combinação intensificador de CMV/promotor tk (Christoph Winkle, Würzburg).

### **Cultura de células**

As células CHO-DUKX (Urlaub e Chasin, 1980) foram cultivadas em meio de Eagle modificado por Dulbecco/meio de Ham F12 (1:1), suplementado com 5% de soro fetal de bovino em frascos de Roux numa incubadora adaptada para ambiente de CO<sub>2</sub> a 90% de humidade do ar, 37 °C e 5% de CO<sub>2</sub>. Para o cultivo isento de soro utilizaram-se células CHO-DUKX, que tinham sido adaptadas lentamente a um desenvolvimento em meio sem soro e que foram então cultivadas permanentemente num meio sem soro. Estas

células foram cultivadas como culturas de suspensão em frascos com agitação em meio de Dulbecco modificado por Iscove/meio de Ham F12 (1:1), que foi suplementado com peptona de baixo peso molecular (Aldag, Hamburgo) insulina e transferrina (Canada Packars).

***Exemplo 1: Hibridação diferencial de bancos de genes de ADNc de CHO recombinantes***

Recorrendo a ARNm poliadenilado de células CHO, que tinham sido cultivadas com soro ou sem soro, prepararam-se em primeiro lugar dois bancos de genes de ADNc em  $\lambda$ ZAPII, consistindo em  $4,2 \times 10^6$  ou  $2,9 \times 10^5$  clones de fagos recombinantes.

O ARN citoplasmático foi preparado de células lisadas NP40 e purificado por meio de extracções com fenol/clorofórmio (Nicolaidis e Stoeckert, 1990). O ARNm poliadenilado de células CHO, que tinham sido cultivadas com ou sem soro, foi obtido por meio de cromatografia de afinidade numa coluna de oligo(dT)-celulose (Sambrook et al., 1989). Para a preparação do ADNc utilizou-se o conjunto de síntese de ADNc da Pharmacia LKB, tendo-se utilizado um iniciador oligo(dT) para a síntese da primeira cadeia. O ADNc foi clonado no vector  $\lambda$ ZAPII digerido com EcoRI (Stratagene Cloning Systems). Os duplicados do filtro do banco de genes de ADNc de CHO não amplificado, a partir de células CHO cultivadas sem soro, foram seleccionados por meio de hibridação diferencial. Como sonda utilizou-se ADNc total de células CHO cultivadas com soro e sem soro que tinham sido marcadas por meio de "random priming" (Prime a Gene labelling Kit; Promega Cooperation, Promega Biotec, Madison, WI, EUA) com ( $\alpha$ - $^{32}$ P) dCTP (6000 Ci/mmol; Amersham International plc, Amersham-

Buchler, Braunschweig, Alemanha). A hibridação foi realizada tal como descrito na análise da mancha de Northern (cf. Exemplo 2). Placas de fagos, que revelaram uma forte hibridação com ambas as sondas de ADNc, foram isoladas e submetidas uma vez mais a uma hibridação diferencial. A partir dos fagos recombinantes  $\lambda$ ZAPII obtiveram-se fagemídeos por meio de excisão *in vivo* com utilização de fagos adjuvantes R408 (Protocolo do conjunto de clonagem  $\lambda$ ZAPII; Stratagene Cloning Systems).

Cerca de 6000 clones de fagos foram seleccionados desta forma. No total foram isolados 12 clones de fagos recombinantes que revelaram uma hibridação especialmente forte com as duas sondas de ADNc conjuntas.

As sequências de ADN foram analisadas segundo o método dideoxi, com utilização de conjuntos de sequenciação T7 da Pharmacia LKB. Como iniciador utilizaram-se tanto o iniciador do promotor T3 e T7 como também o iniciador específico do gene. Em todos os casos, as sondas de ADN foram marcadas com dATP [ $\alpha$ - $^{35}$ S] ou dCTP [ $\alpha$ - $^{35}$ S] (10  $\mu$ Ci; Amersham International plc). Os produtos da reacção foram separados por electroforese num gel de sequenciação de poliácridamida a 6%. Para a análise das sequências destacaram-se os bancos de dados de GenBank e EMBL.

Um dos clones de ADNc isolados codifica para uma proteína de fusão, Ub/S27a, constituída por uma unidade monomérica de ubiquitina e uma proteína ribossomal da sub-unidade pequena S27a (Fig. 1). A maior homologia existe com a sequência Ub/S27a humana (Adams *et al.*, 1992) com 92,2% de homologia ao nível do ADNc e 100% de homologia ao nível dos aminoácidos. O ADNc de Ub/S27a CHO isolado tem 508 pb de comprimento e abrange toda a região de codificação, bem como um sinal de poliadenilação na

área não traduzida a 3' e dois elementos de iniciação da tradução sobrepostos na região não traduzida a 5' (Fig. 1). Com base na sequência de ADNc é também evidente que se trata de um autêntico gene de fusão, isto é, ambas as fracções de proteína são codificadas por um único gene. A sequência proteica da proteína de fusão resultante é altamente conservada, abrangendo os 76 primeiros aminoácidos da proteína com 156 aminoácidos de comprimento da fracção da ubiquitina (Fig. 2).

### ***Exemplo 2: Análise da expressão genética de Ub/S27a.***

A escala da expressão genética de Ub/S27a foi investigada em experiências com a mancha de Northern, em que foi utilizado o ARN conjunto citoplasmático de células CHO cultivadas em soro e sem soro.

O ARN citoplasmático foi preparado a partir de células lisadas de NP40 e purificado por meio de extracções com fenol/clorofórmio (Nicolaidis e Stoeckert, 1990). O ARN conjunto (10 µg) foi separado por electroforese num gel de formaldeído-agarose, transferido para uma membrana de nylon (Hybond N, Amersham International plc) (Sambrook et al., 1989) e reticulado por meio de irradiação durante 5 minutos com UV (254 nm) para formar uma ligação covalente com a membrana de nylon. Os filtros de ARN foram hibridados durante a noite a 65 °C numa solução constituída por 4 x SSC, 10 x Denhardt, 0,1% de SDS e 1 x 10<sup>6</sup> cpm/ml da sonda de ADNc marcada com <sup>32</sup>P. Os fragmentos de ADNc foram marcados recorrendo a "random primer" (Prime a Gene labeling kit; Promega Corporation), em que a actividade específica das amostras de ADN per fez 4 - 8 x 10<sup>8</sup> cpm/µg de ADN. Depois da hibridação, os filtros foram lavados por duas vezes em

0,2 x SSC/0,1% de SDS a 65 °C durante 20 minutos. Sobre o filtro enrolado em película de uso doméstico colocou-se uma película de raios X (Curix; Agfa-Gevaert N.V.) e realizou-se a auto-radiografia durante 3 horas a -70 °C.

Tanto nas condições padrão, com células CHO cultivadas em soro, como também nas células CHO adaptadas ao crescimento em meio sem soro foi possível detectar transcrições de Ub/S27a com um comprimento de cerca de 600 nucleótidos em grande número e quantidades iguais (Fig. 3). Em ambas as transcrições adicionais de cerca de 1500 e 2800 nucleótidos trata-se de transcrições de poliubiquitina em que estão fundidas entre si várias unidades monoméricas de ubiquitina. As transcrições de poliubiquitina, que são expressas num nível consideravelmente reduzido nestas condições de cultura, hibridam com a fracção de ubiquitina do ADNc de Ub/S27a utilizado nas análises da mancha de Northern como sonda. O nível de transcrição de Ub/S27a manteve-se invariavelmente elevado quando se cultivaram as células CHO, que cresceram permanentemente em meio sem soro, durante 24 h em meio com soro (Fig. 3) Deparou-se com uma outra imagem quando as células CHO cultivadas em soro em condições padrão durante 24 h foram transferidas para meio sem soro. O número de transcrições de Ub/S27a diminuiu consideravelmente (Fig. 3). Este facto deve-se a que a parte principal das células CHO que durante esta fase de cultivo de 24 horas em meio sem soro não se dividiram mais e além disso apresentaram uma redução do volume de células. Em células que atingiram um estágio destes, tem lugar apenas ainda uma reduzida transcrição.

**Exemplo 3: Isolamento e análise da região promotora de Ub/S27a.**

Para isolar a região promotora do gene de Ub/S27a preparou-se em primeiro lugar um banco de genes de CHO genómico com mais de um milhão de clones de fagos recombinantes, utilizando-se o ADN genómico de células CHO cultivadas em soro.

O ADN genómico foi obtidos de células lisadas NP40 segundo um protocolo de Nicolaides e Stoeckert (Nicolaides & Stoeckert, 1990). Ao contrário do método de salinização descrito, extraiu-se porém o ADN depois da digestão com proteinase K1, por três vezes com fenol/clorofórmio.

As extremidades de ADN de ADN genómico parcialmente digerido com Sau3A1, com um comprimento médio de fragmento de 20 kb, foram preenchidos de um só lado recorrendo a dGTP e dATP e ligadas com vector  $\lambda$ XGEM-12 digerido com XhoI (extremidades de ADN preenchidas de um só lado com dTTP e dCTP; Promega Corporation). Para empacotar utilizaram-se extractos disponíveis no comércio (Gigapack II Plus packaging extracts; Stratagene Cloning Systems).

Este banco de genes genómico foi hibridado simplesmente com a fracção S27a do ADNc de Ub/S27a a fim de evitar uma hibridação cruzada com os genes da poliubiquitina. Um dos clones genómicos isolados com um comprimento total de 14 kb continha, entre outras, a região de codificação completa e 2,5 kb da região não traduzida a 5'. A região de codificação está interrompida por três intrões com sequências consensuais correctas na transição exão/intrão e intrão/exão (Breathnach & Chambon, 1981), estando

dois destes intrões localizados na fracção de ubiquitina e o terceiro intrão localizado na fracção S27a (Fig. 1).

Ambas as cadeias de ADN da região não traduzida a 5' foram completamente sequenciadas (método vide exemplo 1). Isto teve lugar por meio de sequenciação de fragmentos de restrição subclonados, bem como por meio da sequenciação de clones de eliminação sobrepostos que tinham sido gerados por meio de digestão com exonuclease III (Fig. 4). A região promotora potencial não contém qualquer caixa TATA, mas sim um domínio de sequência rico em CpG (67% de GC de -144 até +129), onde existem também os locais de ligação singulares para o factor de transcrição Sp1 (Dyran & Tjian, 1983) e cada um dos locais de clivagem de restrição para EagI e SacII, que são específicos destas regiões ricas em GC, bem como domínios de sequência ricos em polipirimidina (Fig. 5).

Para a localização do ponto de início da transcrição, realizaram-se tanto a extensão do iniciador como também o mapeamento de nuclease S1 no ARN conjunto de células CHO cultivadas em soro. A fim de evitar uma extensão das transcrições da poliubiquitina, utilizou-se, para a extensão do iniciador um iniciador que hibridou com a fracção S27a do ARNm de Ub/S27a (complementar dos nucleótidos +256 - 276 pb na sequência de ADNc).

Um oligonucleótido da sequência 5'-GTGGTGTAGGACTTCTTCTTC-3', complementar dos nucleótidos +256 a +276 na sequência de ADNc de Ub/S27a foi hibridado com 5 µg de ARN conjunto citoplasmático de células CHO cultivadas em soro e prolongado (Ausubel et al., 1987).

A sonda de cadeia simples utilizada para o mapeamento por nuclease S1, que abrange o domínio de -2289 até +176 da sequência genômica de Ub/S27a, foi obtida através de PCR da seguinte maneira. A região não traduzida a 5' da sequência genômica de Ub/S27a (-2289 a +240) foi clonada na orientação 5' - 3' no vector pBluescript SK (Stratagene Cloning Systems). Este plasmídeo híbrido foi utilizado como matriz na PCR. Para a amplificação utilizou-se um iniciador de promotor T3 biotinilado e um iniciador específico de Ub/S27a marcado com ATP [ $\gamma$ -<sup>32</sup>P] (5'-CTCGAGCGTGATCGTTTTCC-3', complementar dos nucleótidos +157 a +176 da sequência genômica de Ub/S27a). A obtenção da cadeia simples complementar da sequência de ARN foi realizada por meio de desnaturação alcalina do produto de PCR ligado a esferas de estreptavidina magnéticas (Protocolo Dynabeads M-280 Streptavidin; Dynal A. S., Noruega). 2 x 10<sup>5</sup> cpm da amostra de cadeia simples foram hibridados com 5 µg de ARN conjunto de células CHO cultivadas em soro durante a noite a 55 °C e os produtos da hibridação foram depois tratados com nuclease S1 (Ausubel *et al.*, 1987).

O produto obtido por meio de extensão do iniciador apresentava um comprimento de 304 nucleótidos (Fig. 6), onde o ponto de início da transcrição se situaria 44 pb a montante do codão de início no interior de um elemento de polipirimidina (Fig. 5). Não foi possível, no entanto, verificar este ponto de início pelo mapeamento com nuclease S1. Por meio do mapeamento com nuclease S1 determinaram-se dois pontos de início da transcrição (Fig. 7) que se encontram 128 pb ou 116 pb a montante do codão de início, no interior de uma sequência de polipirimidina, e as suas posições são subsequentemente indicadas como posição +1 e +13 (Fig. 5, Fig. 7). A explicação mais provável para a discrepância observada entre os dois

métodos de mapeamento é uma terminação prematura da reacção de extensão do iniciador. Condicionada pelo posicionamento do iniciador na sequência S27a, o comprimento do produto da extensão esperado, com mais de 300 bases, encontra-se acima do comprimento óptimo de cerca de 100 bases. Quanto maior o afastamento entre a posição do iniciador e o sítio de início da transcrição procurado, maior será a probabilidade de uma ruptura prematura da transcrição inversa por meio de sequências ricas em GC e formações de estruturas secundárias no interior do ARNm. Para o mapeamento por nuclease S1, utilizou-se pelo contrário uma amostra de cadeia simples abrangendo o domínio não traduzido a 5' completo (domínio da sequência -2289 a +176; Fig. 5) que foram obtidos a partir de um produto de PCR, pelo que são contornados os problemas da transcrição inversa de sequências ricas em GC, que surgem numa extensão do iniciador.

As nossas investigações da região não traduzida a 5' de 2,5 kb de comprimento revelaram que o promotor possível não possuía nem uma caixa TATA nem uma caixa CAAT. Para este fim, no entanto, a sequência a montante do codão de início caracteriza-se por um elevado teor de GC. No interior desta sequência encontrou-se um sítio singular de ligação para o factor de transcrição Sp1. Por meio de mapeamento por nuclease S1, identificaram-se dois pontos de início de transcrição proeminentes que estavam situados no interior de uma sequência de polipirimidina a 128 ou 116 pb a montante do codão de início.

**Exemplo 4: Identificação e limitação do domínio de sequência com actividade promotora**

**Preparação de clones de eliminação por meio de digestão com exonuclease III**

Por meio de digestão com exonuclease III preparou-se uma série de clones de eliminação a 5' (Quadro 1). A região não traduzida a 5' do gene de Ub/S27a (-2289 até +240) foi para esse fim clonada em ambas as orientações no sítio de clivagem HincII do vector pBluescript SK- (Stratagene Cloning Systems). Para a introdução de eliminações unidireccionais, digeriram-se estes plasmídeos híbridos com KpnI e XhoI e tratou-se com exonuclease III, tal como descrito no protocolo do conjunto "Erase-a-base" utilizado (Promega Corporation). Os fragmentos de ADN obtidos foram integrados como fragmentos BamHI no sítio de clivagem BamHI singular do vector pBL-CAT6.

Antes do início da digestão com exonuclease III, o vector pBluescript Sk- pode também ser modificado por meio da introdução de adaptadores que contêm sítios de clivagem de restrição adequados, a fim de facilitar as experiências subsequentes de clonagem dos genes. Assim, a par de BamHI podem-se também ainda destacar outros sítios de clivagem de restrição enzimática com fins de clonagem, p. ex. segundo o esquema

NotI-XbaI-SpeI-BamHI-SmaI-EcoRI-extremidade 3'.....extremidade 5'-SalI/HincII-BamHI

**Fragmento de eliminação 5'**

**Quadro 1: Composição do clone de eliminação 5' da região não traduzida a 5' de Ub/S27a preparado por digestão com exonuclease III**

Fragmento de eliminação a 5'	Extremidade 5' do fragmento	Extremidade 3' do fragmento
1612 pb	Posição - 1501	Posição + 111
806 pb	Posição - 695	Posição + 111
483 pb	Posição - 372	Posição + 111
272 pb	Posição - 161	Posição + 111
156 pb	Posição - 45	Posição + 111

Os dados da posição referem-se à numeração da sequência Ub/S27a genómica representada na Fig. 5.

Todos os clones de eliminação têm uma extremidade 3' comum (posição +111), que se situa entre o sítio de início da transcrição e o codão de início (Fig. 5). O fragmento maior abrange 1,7 kb e o fragmento menor abrange 150 pb (Quadro 1). O último é o único fragmento que já não contém o sítio de ligação Sp1 singular. Estas regiões promotoras potenciais foram clonadas no vector de expressão pBL-CAT6 eucariótico antes do gene CAT sem promotor, que servia de gene relator, e foram utilizadas para a transfecção temporária de células CHO cultivadas em soro.

#### ***Transfecção celular mediada por ADN e análise de CAT***

Na véspera da transfecção foram semeadas  $2 \times 10^5$  células por 20 cm<sup>2</sup> de placa. As células foram transfectadas com 10 µg de ADN plasmídico (construção relatora de CAT) e 500 ng do plasmídeo pCMVtklacZ, recorrendo a um método de precipitação de fosfato de cálcio modificado (Chen & Okayama, 1987). A actividade de

$\beta$ -galactosidase do vector de controlo pCMVtklacZ foi neste caso levada em conta para a determinação da eficácia de transfecção. O precipitado de ADN excedentário foi removido depois de incubação de 4 horas a 37 °C e CO<sub>2</sub> a 5%, por meio de lavagem com PBS.

Depois de um período de incubação de 48 horas, lavaram-se as células em primeiro lugar com PBS. Depois, deitou-se 1 mL de tampão NTE gelado (NaCl 0,1 M, tris-HCl 10 mM, pH 7,8, EDTA 1 mM) em cada placa e as células foram retiradas das placas com auxílio de um raspador e transferidas para um recipiente de Eppendorf. As células foram sedimentadas por meio de 3 minutos de centrifugação a 9000 rpm, ressuspensas em 70  $\mu$ L de tris-HCl 0,25 M, pH 7,8 e guardadas a -70 °C. Para a determinação de actividade de acetiltransferase de cloranfenicol segundo o método de Sleight utilizaram-se respectivamente 30  $\mu$ L de suspensão de células (Sleight, 1986). Uma normalização da actividade CAT relativa de cada transfecção foi executada com base na actividade de  $\beta$ -galactosidase determinada segundo o método de Norton e Coffin (Norton & Coffin, 1985). Para este teste utilizaram-se em cada caso 10  $\mu$ L a suspensão de células. Em todos os casos executou-se ainda uma correcção relativamente à quantidade de proteínas utilizada. A concentração de proteínas foi neste caso determinada segundo o método de Bradford (Ausubel *et al.*, 1987).

No histograma na Figura 8 encontram-se representados os resultados de quatro experiências de expressão transitória realizadas independentemente umas das outras. Para o controlo foram utilizados plasmídeos em que a expressão genética de CAT foi controlada por meio de um promotor viral constitutivo. Em pBL-CAT5 é o promotor da cinase de timidina do vírus do *Herpes*

*Simplex*. Em pCMVtkCAT, este promotor encontra-se em combinação com um intensificador do citomegalovírus humano e em pKSSV2CAT trata-se do promotor de SV40. O pBL-CAT6 serviu como controlo negativo e continha um gene de CAT sem promotor.

Com excepção do fragmento de 156 pb, todos os fragmentos de Ub/S27a que tinham sido clonados na orientação 5'-3' antes do gene relator de CAT sem promotor revelaram uma actividade promotora forte que era 2,5 a 3 vezes maior que a do promotor tk do vírus do *Herpes simplex* (Fig. 8). A actividade promotora mais forte, comparável à do promotor de SV40 em pKSSV2CAT, foi apresentada neste caso pelos fragmentos de 483 pb e 272 pb. Só a combinação intensificador viral de CMV/promotor-tk em pCVtkCAT teve como resultado uma actividade de CAT ainda maior em cerca de 10%. Numa eliminação que se estende até à posição -45 (fragmento de 156 pb), que abrangia também o sítio de ligação singular Sp1, já não foi possível detectar qualquer actividade de CAT (Fig. 8). A região não traduzida a 5' de Ub/S27a, que é suficiente para veicular uma actividade promotora forte, pode ser assim delimitada no domínio de -161, a extremidade 5' do fragmento de 272 pb, até -45, a extremidade 5' do fragmento de 156 pb. No interior deste domínio de 117 pb encontra-se também o sítio de ligação singular de Sp1.

Foi um resultado inesperado observar que o fragmento mais curto, também na orientação 3'-5', apresentava uma actividade promotora (Fig. 8). Além disso, este promotor localizado na contracadeia não é tão forte quanto o promotor de Ub/S27a, mas sim reduzido em 42%. Também neste caso os fragmentos de 483 pb e de 272 pb revelavam a maior actividade. Esta actividade era comparável à do promotor tk em pBL-CAT5. O gene correspondente, cuja expressão é controlada por este promotor de contracadeia,

ainda não tinha podido ser identificado até à data. Portanto, no que se refere ao domínio do promotor de 117 pb acima mencionado, podia tratar-se de uma região promotora bidireccional.

Em resumo pode-se dizer que com o promotor de Ub/S27a isolado pelos autores da presente invenção de células CHO se apresentou pela primeira vez um promotor do hamster homólogo constitutivo muito forte. As investigações em células CHO cultivadas em soro, transfectadas estáveis documentam que o promotor de Ub/S27a é ainda muito activo, mesmo depois da integração no genoma da célula e proporciona uma expressão muito forte do gene relator de CAT.

***Exemplo 5: Expressão temporária de um gene relator sob controlo do promotor de Ub/S27a em células BHK***

Construções do gene relator de CAT, que continham diferentes fragmentos da região promotora de Ub/S27a foram introduzidas em células BHK-21 (ATCC CCL 10) de forma análoga e a actividade de CAT dos transfectantes foi medida. A Fig. 9 mostra que, com as sequências promotoras de acordo com a invenção, também nas células BHK pode ser atingido uma taxa de expressão extraordinariamente alta.

***Exemplo 6: Expressão estável de um gene relator sob controlo do promotor de Ub/S27a em células CHO***

Transfectantes estáveis foram preparados da seguinte forma. Na véspera da transfecção foram semeadas 200.000 células por 20 cm<sup>2</sup> de placa. As células foram transfectadas com 10 µg de ADN

plasmídico (construção relatora de CAT) e 500 ng do plasmídeo pSV2pac (Vara *et al.*, 1986), recorrendo a um método de precipitação com fosfato de cálcio modificado (Chen & Okayama, 1987). O plasmídeo pSV2pac possui o gene que codifica para a acetiltransferase de puromicina-N (pac) e veicula assim resistência contra o antibiótico inibidor da biossíntese de proteínas. O precipitado de ADN excedentário foi removido depois de incubação de 4 horas a 37 °C e CO<sub>2</sub> a 5%, por meio de lavagem com PBS. Passadas mais 24 horas foi iniciado com a selecção dos transfectantes por meio de adição de 10 µg/mL de puromicina. A todos os dois a três dias procedeu-se a uma troca do meio com meio contendo puromicina. Por meio de clonagem de diluição dos transfectantes seleccionados obtiveram-se clones de célula única.

As Figs. 10 a 12 revelam taxas de expressão de CAT de diferentes clones de transfectantes estáveis relativamente a pCMVtkCAT cl.6 (clone 6 do clone celular transfectado pCMVtkCAT com a maior actividade de CAT).

### **Literatura**

Adams, S.M., Sharp, M.G., Walker, R.A., Brammar, W.J. & Varley, J.M. (1992)

Differential expression of translation-associated genes in benign and malignant human breast tumours

Br.J.Cancer 65, 65 - 71

Ausubel, F.M., Brent, R., Kingston, R.E., Moore, D.D., Smith, J.A., Seidman, J.G. & Struhl, K (1987)

Current protocols in molecular biology

Greene Publishing Associates and Wiley-Interscience

Baker, R.T. & Board, P.G. (1991)

The human ubiquitin-52 amino acid fusion protein gene shares several structural features with mammalian ribosomal protein genes

Nucleic Acids Research 19, 1035 - 1040

Boshart, M., Klüppel, M., Schmidt, A. , Schütz, G. & Luckow, B. (1992)

Reporter constructs with low background activity utilizing the cat gene

Gene 110, 129 - 130

Breathnach, R & Chambon, P. (1981)

Organization and expression of eukaryotic split genes coding for proteins

Ann.Rev.Biochem. 50, 349 - 383

Chen, C. & Okayama, H. (1987)

High-efficiency transformation of mammalian cells by plasmid DNA

J. Mol. Cell. Biol. 7, 2745 - 2752

Ciechanover, A. (1994)

The ubiquitin-proteasome proteolytic pathway

Cell 79, 13 - 21

Dynan, W.S. & Tjian, R. (1983)

The promoter-specific transcription factor Sp1 binds to upstream sequences in the SV40 early promoter

Cell 35, 79 - 87

Faisst, S. & Meyer, S. (1992)  
Compilation of vertebrate-encoded transcription factors  
Nucleic Acids Research 20, 3 - 26

Finley, D., Bartel, B. & Varshavsky, A. (1989)  
The tails of ubiquitin precursors are ribosomal proteins whose  
fusion to ubiquitin facilitates ribosome biogenesis  
Nature 333, 394 - 401

Fornace, A.J., Alama, I., Hollander, M.C. & Lamoreaux, E.  
(1989)  
Ubiquitin mRNA is a major stress-induced transcript in mammalian  
cells  
Nucleic Acids Research 17, 1215 - 1230

Huxley, C. & Fried, M. (1990)  
The mouse rpL7a gene is typical of other ribosomal protein genes  
in it's 5' region but differs in being located in a tight  
cluster of CpG-rich islands  
Nucleic Acids Research 18, 5353 - 5357

Jentsch, S., Seufert, W. & Hauser, H.-P. (1991)  
Genetic analysis of the ubiquitin system  
Biochimica et Biophysica Acta 1089, 127 - 139

Kaufman R. J. (1987)  
High level production of proteins in mammalian cells  
In: Genetic Engineering: Principles and methods (Hrsg. J K  
Setlow), Bd. 9, S. 155  
Plenum Publishing, New York

Luckow, B., Schütz, G. (1987)

CAT constructions with multiple unique restriction sites for the functional analysis of eucaryotic promoters and regulatory elements

Nucleic Acids Res. 15, 5490

Luo, X. & Kim, K.-H. (1990)

An enhancer element in the house-keeping promoter for acetyl-CoA carboxylase gene

Nucleic Acids Research 18, 3249 - 3254

Meyer, J., Nick, S., Stamminger, T., Grummt, F., Jahn, G. & Lipps, H.J. (1993)

Inhibition of HIV-1 replication by a high-copy-number vector expressing antisense RNA for reverse transcriptase

Gene 129, 263 - 268

Nicolaides, N.C. & Stoeckert, C.J. (1990)

A simple, efficient method for the separate isolation of RNA and DNA from the same cells

BioTechniques 8, 154 - 155

Norton, P.A. & Coffin, J.M. (1985)

Bacterial  $\beta$ -galactosidase as a marker of Rous Sarcoma virus gene expression and replication

Mol. Cell. Biol. 5, 281 - 290

Redman, K.L. & Rechsteiner, M. (1989)

Identification of the long ubiquitin extension as ribosomal protein S27a

Nature 338, 438 - 440

Sambrook, J., Fritsch, E.F. & Maniatis, T. (1989)

Molecular cloning A Laboratory Manual

Cold Spring Harbor Laboratory, Cold Spring Harbor, NY

Schlesinger, M.J. & Bond, U. (1987)

Ubiquitin genes

Oxf. Survey Euk. Genes 4, 77 - 91

Shimbara, N., Sato, C., Takashina, M., Tanaka, T., Tanaka, K. & Ichihara, A. (1993)

Down-regulation of ubiquitin gene expression during differentiation of human leukemia cells

FEBS Letters 322, 235 - 239

Sleigh, M.J. (1986)

A nonchromatographic assay for expression of the chloramphenicol acetyl transferase gene in eukaryotic cells

Anal. Biochem. 156, 252 - 256

Urlaub, G. & Chasin, L.A. (1980)

Isolation of chinese hamster cell mutants deficient in dihydrofolate reductase activity

Proc. Natl. Acad. Sci. USA 77, 4216 - 4220

Tsonis, P.A., Manes, T., Millan, J.L. & Goetinck, P.F. (1988)

CAT constructs with convenient sites for cloning and generating deletions

Nucleic Acids Research 16, 7745

Vara, J.A., Portela, A., Ortin, J. & Jiménez, A. (1986)

Expression in mammalian cells of a gene from Streptomyces alboniger conferring puromycin resistance

Nucleic Acids Research 14, 4617 - 4624

Wegner, M., Zastrow, G., Klavinius, A., Schwender, S.,  
Müller, F., Luksza, H., Hoppe, J., Wienberg, J. & Grummt, F.  
(1989)

Cis-acting sequences from mouse rDNA promote plasmid DNA  
amplification and persistence in mouse cells: implication of  
HMG-I in their function

Nucleic Acids Research 17, 9909 - 9932

Wiborg, O., Pedersen, M.S., Wind, A., Berglund, L.E.,  
Marcker, K.A. & Vuust, J. (1985)

The human ubiquitin multigene family: some genes contain  
multiple directly repeated ubiquitin coding sequences

EMBO J. 4, 755 - 759

Wong, J.M., Mafune, K., Yow, H., Rivers, E.N., Ravikumar,  
T.S., Steele, G.D. & Chen, L.B. (1993)

Ubiquitin-ribosomal protein S27a gene overexpressed in human  
colorectal carcinoma is an early growth response gene

Cancer Research 53, 1916 - 1920

## PROTOCOLO DE SEQUÊNCIAS

### (1) DADOS GERAIS

#### (i) REQUERENTE

- (A) NOME: Boehringer Ingelheim International GmbH
- (B) RUA: Binger Str.
- (C) LOCALIDADE: Ingelheim
- (E) PAÍS: Alemanha
- (F) CÓDIGO POSTAL: 55216
- (G) TELEFONE: 06132-772770
- (H) TELEFAX: 06132-774377

(ii) EPÍGRAFE DA INVENÇÃO: Promotor homólogo forte obtido a partir de hamster

(iii) NÚMERO DAS SEQUÊNCIAS: 8

#### (iv) FORMA DE LEITURA INFORMÁTICA:

- (A) SUPORTE DOS DADOS: Disquete
- (B) COMPUTADOR: Compatível IBM PC
- (C) SISTEMA OPERATIVO: PC-DOS/MS-DOS
- (D) PROGRAMA: PatentIn Release #1.0, Versão #1.30 (EPA)

### (2) DADOS PARA SEQ ID N.º: 1:

#### (i) CARACTERÍSTICAS DA SEQUÊNCIA:

- (A) COMPRIMENTO: 508 pares de bases
- (B) TIPO: nucleótido
- (C) FORMA DE CADEIA: ambas
- (D) TOPOLOGIA: ambas

(ii) TIPO DE MOLÉCULA: ADNc para ARNm

(xi) DESCRIÇÃO DA SEQUÊNCIA: SEQ ID N.º: 1:

```

TGAACCGCC GCCAAGATGC AGATTTTCGT GAAGACCCTT ACGGGGAAAA CGATCACGCT 60
CGAGGTTGAA CCCTCGGACA CTATAGAAAA TGTAAGGCC AAGATCCAGG ATAAGGAAGG 120
AATTCCTCCT GACCAGCAGA GGCTGATCTT TGCTGGTAAG CAACTGGAAG ATGGCCGTAC 180
TTTGTCTGAC TACAACATCC AAAAGGAGTC CACCCTTCAT CTTGTGTTGA GACTTCGTGG 240
TGGTGCTAAG AAGAGGAAGA AGAAGTCCTA CACCACTCCC AAGAAGAATA AGCATAAGAG 300
AAAGAAGGTT AAGTTGGCTG TGCTGAAGTA CTATAAGGTG GATGAAAATG GCAAAATTAG 360
TCGCCTTCGT CGAGAGTGTC CATCTGATGA GTGTGGTGCT GGAGTTTTCA TGGCTAGCCA 420
TTTTGACAGA CATTACTGTG GCAAGTGTG TCTGACTTAC TGCTTCAACA AACCAGAAGA 480
CAAGTAGTTG TGTATGAATA AATAAAAA 508

```

(2) DADOS PARA SEQ ID N.º: 2:

(i) CARACTERÍSTICAS DA SEQUÊNCIA:

- (A) COMPRIMENTO: 487 pares de bases
- (B) TIPO: nucleótido
- (C) FORMA DE CADEIA: ambas
- (D) TOPOLOGIA: ambas

(ii) TIPO DE MOLÉCULA: ADNc

(xi) DESCRIÇÃO DA SEQUÊNCIA: SEQ ID N.º: 2:

```

TGGAGCCGCA ACCAAAATGC AGATTTTCGT GAAAACCCTT ACGGGGAAGA CCATCACCCCT 60
CGAGGTTGAA CCCTCGGATA CGATAGAAAA TGTAAGGCC AAGATCCAGG ATAAGGAAGG 120
AATTCCTCCT GATCAGCAGA GACTGATCTT TGCTGGCAAG CAGCTAGAAG ATGGACGTAC 180
TTTGTCTGAC TACAATATTC AAAAGGAGTC TACTCTTCAT CTTGTGTTGA GACTTCGTGG 240
TGGTGCTAAG AAAAGGAAGA AGAAGTCTTA CACCACTCCC AAGAAGAATA AGCACAAGAG 300
AAAGAAGGTT AAGCTGGCTG TCCTGAAATA TTATAAGGTG GATGAGAATG GCAAAATTAG 360
TCGCCTTCGT CGAGAGTGCC CTTCTGATGA ATGTGGTGCT GGGGTGTTTA TGGCAAGTCA 420
CTTTGACAGA CATTATTGTG GCAAATGTTG TCTGACTTAC TGTTTCAACA AACCAGAAGA 480
CAAGTAA 487

```

(2) DADOS PARA SEQ ID N.º: 3:

(i) CARACTERÍSTICAS DA SEQUÊNCIA:

- (A) COMPRIMENTO: 156 aminoácidos
- (B) TIPO: aminoácido
- (C) FORMA DE CADEIA: cadeia simples
- (D) TOPOLOGIA: linear

(ii) TIPO DE MOLÉCULA: Proteína

(xi) DESCRIÇÃO DA SEQUÊNCIA: SEQ ID N.º: 3:

```
Met Gln Ile Phe Val Lys Thr Leu Thr Gly Lys Thr Ile Thr Leu Glu
1           5           10           15

Val Glu Pro Ser Asp Thr Ile Glu Asn Val Lys Ala Lys Ile Gln Asp
          20           25           30

Lys Glu Gly Ile Pro Pro Asp Gln Gln Arg Leu Ile Phe Ala Gly Lys
          35           40           45

Gln Leu Glu Asp Gly Arg Thr Leu Ser Asp Tyr Asn Ile Gln Lys Glu
50           55           60

Ser Thr Leu His Leu Val Leu Arg Leu Arg Gly Gly Ala Lys Lys Arg
65           70           75           80

Lys Lys Lys Ser Tyr Thr Thr Pro Lys Lys Asn Lys His Lys Arg Lys
          85           90           95

Lys Val Lys Leu Ala Val Leu Lys Tyr Tyr Lys Val Asp Glu Asn Gly
          100          105          110

Lys Ile Ser Arg Leu Arg Arg Glu Cys Pro Ser Asp Glu Cys Gly Ala
          115          120          125

Gly Val Phe Met Ala Ser His Phe Asp Arg His Tyr Cys Gly Lys Cys
          130          135          140

Cys Leu Thr Tyr Cys Phe Asn Lys Pro Glu Asp Lys
145           150           155
```

(2) DADOS PARA SEQ ID N.º: 4:

(i) CARACTERÍSTICAS DA SEQUÊNCIA:

- (A) COMPRIMENTO: 156 aminoácidos
- (B) TIPO: ácidos aminados
- (C) FORMA DE CADEIA: cadeia simples
- (D) TOPOLOGIA: linear

(ii) TIPO DE MOLÉCULA: Proteína

(xi) DESCRIÇÃO DA SEQUÊNCIA: SEQ ID N.º: 4:

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

	100		105		110										
Lys	Ile	Ser	Arg	Leu	Arg	Arg	Glu	Cys	Pro	Ser	Asp	Glu	Cys	Gly	Ala
	115						120					125			
Gly	Val	Phe	Met	Ala	Ser	His	Phe	Asp	Arg	His	Tyr	Cys	Gly	Lys	Cys
	130					135					140				
Cys	Leu	Thr	Tyr	Cys	Phe	Asn	Lys	Pro	Glu	Asp	Lys				
145					150					155					

(2) DADOS PARA SEQ ID N.º: 5:

(i) CARACTERÍSTICAS DA SEQUÊNCIA:

(A) COMPRIMENTO: 2529 pares de bases

(B) TIPO: nucleótido

(C) FORMA DE CADEIA: ambas

(D) TOPOLOGIA: linear

(ii) TIPO DE MOLÉCULA: ADN genómico

(xi) DESCRIÇÃO DA SEQUÊNCIA: SEQ ID N.º: 5:

```
GATCTCCAGG ACAGCCATGG CTATTACACA GAGAAACCCT GTCTGGAAAA ACAAAAATT 60
AGTGTCCATG TGTAATGTG TGGAGTATGC TTGTCATGCC ACATACAGAG GTAGAGGGCA 120
GTTTATGGGA GTCAGTTCCT ATTCTTCCTT TATGGGGGAC CTGGGGACTG AACTCAGGTC 180
ATCAGGCTTG GCAGAAAGTG CATTAGCTCA CGGAGCCTTA TCATTGGCGA AAGCTCTCTC 240
AAGTAGAAAA TCAATGTGTT TGCTCATAGT GCAATCATTA TGTTTCGAGA GGGGAAGGGT 300
ACAATCGTTG GGGCATGTGT GGTACATCT GAATAGCAGT AGCTCCCTAG GAGAATTCCA 360
AGTTCTTTGG TGGTGTATCA ATGCCCTTAA AGGGGTCAAC AACTTTTTTT CCCTCTGACA 420
AAACTATCTT CTTATGTCCT TGTCCCTCAT ATTTGAAGTA TTTTATTCTT TGCAGTGTG 480
AATATCAATT CTAGCACCTC AGACATGTTA GGTAAGTACC CTACAACCTA GGTTAACATA 540
TTTAATTTAA CTAATTTAAC CCCAACACTT TTTCTTTGTT TATCCACATT TGTGGAGTGT 600
GTGTGTGTGT GTGTGTGTGT GTGTGTGTGT GTGTGTGTGT GTGTGTGTGT GTGTGTGTGC 660
CGCGCGCGCT CGGATCATTC TACCTTTTGT TTAAAAATG TTAGTCCAGG GGTGGGGTGC 720
```

ACTGTGAAAG TCTGAGGGTA ACTTGCTGGG GTCAGTTCTT TCCACTATAG GACAGAACTC 780  
 CAGGTGTCAA CTCTTTACTG ACAGAACCAT CCAAATAGCC CTATCTAATT TTAGTTTTTT 840  
 ATTTATTTAT TTTTGTGTTT TCGAGACAGG GTTCTCTGT GGCTTTGGAG GCTGTCCTGG 900  
 AACTAGCTCT TGTAGACCAG GCTGGTCTCG AACTCAGAGA TCCACCTGCC TCTGCCTCCT 960  
 GAGTGCTGGG ATTAAAGGCA TGCGCCACCA ACGCTTGGCT CTACCTAATT TAAAAGAGA 1020  
 TTGTGTGTCA CAAGGGTGTC ATGTCGCCCT GCAACCACCC CCCCCCAA AAAAAAAAAA 1080  
 AAAAATTCA CTGAAGCTGA AGCAGCATGA TTTGGTTACT CTGGCTGGCC AATGAGCTCT 1140  
 AGGGAGTCTC CTGTCAAACA GAATCTCAAC AGGCGCAGCA GTCTTTTTTA AAGTGGGGTT 1200  
 ACAACACAGG TTTTGCATA TCAGGCATTT TATCTAAGCT ATTTCCAGC CAAAAATGTG 1260  
 TATTTTGGAG GCAGCAGAGC TAATAGATTA AAATGAGGGA AGAGCCCACA CAGGTTATTA 1320  
 GGAAGATAAG CATCTTCTTT ATATAAACA AAACCAAACC AACTGGAGG AGGTCTACCT 1380  
 TTAGGGATGG AAGAAAAGAC ATTTAGAGGG TGCAATAGAA AGGGCACTGA GTTTGTGAGG 1440  
 TGGAGGACTG GGAGAGGGCG CAACCGCTTT AACTGTCCTG TTTTGCCTAT TTTTGGGGA 1500  
 CAGCACATGT TCCTATTTTT CCCAGGATGG GCAATCTCCA CGTCCAACT TCGGTCGAG 1560  
 GACTACAGTC ATTTTGCAGG TTTCCTTACT GTATGGCTTT TAAACGTGC AAAGGTGACC 1620  
 ATTAACCGTT TCACGCTGGG AGGGCACGTG CGGCTCAGAT GCTTCCTCTG ACTGAGGGCC 1680  
 AGGAGGGGGC TACACGGAAG AGGCCACACC CGCACTTGGG AAGACTCGAT TTGGGCTTCA 1740  
 GCTGGCTGAG ACGCCCCAGC AGGCTCCTCG GCTACACCTT CAGCCCCGAA TGCCTTCCGG 1800  
 CCCATAACCC TTCCCTTCTA GGCATTTCCG GCGAGGACCC ACCCTCGCGC CAAACATTCG 1860  
 GCCCCATCCC CCGGTCCTCA CCTGAATCTC TAACTCTGGA CTCCAGAGTT TAGAGACTAT 1920  
 AACCAGATAG CCCGGATGTG TGGAAGTGA TCTTGGGACG AGTAGTTTTA GCAAAAAGAA 1980  
 AGCGACGAAA AACTACAATT CCCAGACAGA CTTGTGTTAC CTCTCTTCTC ATGCTAAACA 2040  
 AGCCCCCTTT AAAGGAAAGC CCCTCTTAGT CGCATCGACT GTGTAAGAAA GCGGTTTGAA 2100  
 ACATTTTAAT GTTGGGCACA CCGTTTCGAG GACCGAAATG AGAAAGAGCA TAGGGAAACG 2160  
 GAGCGCCCGA GCTAGTCTGG CACTGCGTTA GACAGCCGCG GTCGTTGCAG CGGGCAGGCA 2220  
 CTTGCGTGGA CGCCAAGGGG CGGGTCTTTC GGCCGGGAAG CCCCCTTGGT CCGCGCGGCT 2280  
 CTTCTTTTCC GATCCGCCAT CCGTGGTGAG TGTGTGCTGC GGGCTGCCGC TCCGGCTTGG 2340

```

GGCTTCCCGC GTCGCTCTCA CCCTGGTCGG CGGCTCTAAT CCGTCTCTTT TCGAATGTAG 2400
GTGGAACCGC CGCCAAGATG CAGATTTTCG TGAAGACCCT TACGGGGAAA ACGATCACGC 2460
TCGAGGTACG AACCAGGTGG CGTGAGAAGC GAAGGCCTGC CAGAGGCCCT CTATGCTCGC 2520
TTAAAGCTT 2529

```

(2) DADOS PARA SEQ ID N.º: 6:

(i) CARACTERÍSTICAS DA SEQUÊNCIA:

- (A) COMPRIMENTO: 117 pares de bases
- (B) TIPO: nucleótido
- (C) FORMA DE CADEIA: ambas
- (D) TOPOLOGIA: linear

(ii) TIPO DE MOLÉCULA: ADN genómico

(xi) DESCRIÇÃO DA SEQUÊNCIA: SEQ ID N.º: 6:

```

AGGACCGAAA TGAGAAAGAG CATAGGGAAA CGGAGCGCCC GAGCTAGTCT GGCCTGCGT 60
TAGACAGCCG CGGTCGTTGC AGCGGGCAGG CACTTGCGTG GACGGGAAGG GCGGGT 117

```

(2) DADOS PARA SEQ ID N.º: 7:

(i) CARACTERÍSTICAS DA SEQUÊNCIA:

- (A) COMPRIMENTO: 21 pares de bases
- (B) TIPO: nucleótido
- (C) FORMA DE CADEIA: cadeia simples
- (D) TOPOLOGIA: linear

(ii) TIPO DE MOLÉCULA: Outros ácidos nucleicos

- (A) DESCRIÇÃO: /desc = "iniciador"

(xi) DESCRIÇÃO DA SEQUÊNCIA: SEQ ID N.º: 7:

GTGGGTAGG ACTTCTTCTT C

21

(2) DADOS PARA SEQ ID N.º: 8:

(i) CARACTERÍSTICAS DA SEQUÊNCIA:

- (A) COMPRIMENTO: 20 pares de bases
- (B) TIPO: nucleótido
- (C) FORMA DE CADEIA: cadeia simples
- (D) TOPOLOGIA: linear

(ii) TIPO DE MOLÉCULA: Outros ácidos nucleicos

(A) DESCRIÇÃO: /desc = "iniciador"

(xi) DESCRIÇÃO DA SEQUÊNCIA: SEQ ID N.º: 8:

CTCGAGCGTG ATCGTTTTCC

20

Lisboa, 9 de Março de 2007

## **REIVINDICAÇÕES**

1. Molécula de ácido nucleico que contém a sequência promotora do gene da ubiquitina-S27a de hamster da posição 2129 até 2400 de acordo com a SEQ ID NO:5 ou uma sequência que pode ser derivada desta por meio de substituição, inserção ou eliminação de uma, duas ou três bases, sem que por esta substituição, inserção ou eliminação diminua a actividade promotora em mais de 50% do valor que é obtido para um promotor com a sequência da posição 2129 até 2400 de acordo com a SEQ ID NO:5.
2. Molécula de ácido nucleico de acordo com a reivindicação 1, caracterizada por conter a sequência promotora do gene da ubiquitina-S27a de hamster da posição 1918 até 2400 de acordo com a SEQ ID NO:5 ou uma sequência que pode ser derivada desta por meio de substituição, inserção ou eliminação de uma, duas ou três bases, sem que por esta substituição, inserção ou eliminação diminua a actividade promotora em mais de 50% do valor que é obtido para um promotor com a sequência da posição 2129 até 2400 de acordo com a SEQ ID NO:5.
3. Molécula de ácido nucleico de acordo com a reivindicação 1 ou 2, caracterizada por conter a sequência promotora do gene da ubiquitina-S27a de hamster de acordo com a SEQ ID NO:5 ou uma sequência que pode ser derivada desta por meio de substituição, inserção ou eliminação de uma, duas ou três bases, sem que por esta substituição, inserção ou eliminação diminua a actividade promotora em mais de 50% do

valor que é obtido para um promotor com a sequência da posição 2129 até 2400 de acordo com a SEQ ID NO:5.

4. Molécula de ácido nucleico de acordo com uma das reivindicações 1 a 3, caracterizada por conter um ou mais intensificadores que está/estão acoplados de forma funcional à sequência promotora do gene de ubiquitina-S27a de hamster.
5. Molécula de ácido nucleico de acordo com uma das reivindicações 1 a 4, caracterizada por as sequências promotoras estarem acopladas a um gene de forma funcional.
6. Molécula de ácido nucleico de acordo com a reivindicação 5, caracterizada por o gene mencionado ser expresso sob o controlo das sequências promotoras.
7. Molécula de ácido nucleico de acordo com uma das reivindicações 5 ou 6, caracterizada por o gene mencionado codificar para o activador de plasminogénio de tecido, activador de plasminogénio de tecido de *second generation*, interferão, factor de necrose de tumores, eritropoietina, factor de estimulação da colónia de granulócitos, superoxidodismutase de manganês, uma cadeia de imunoglobulina, a região variável de uma cadeia de imunoglobulina, uma cadeia de imunoglobulina humanizada, a região variável de uma cadeia de imunoglobulina humanizada, um anticorpo de *single-chain* e/ou um anticorpo que é específico para CD44 variante.

8. Molécula de ácido nucleico de acordo com uma das reivindicações 1 a 7, caracterizado por ser um vector de expressão.
9. Célula hospedeira na qual foi introduzida uma molécula de ácidos nucleicos de acordo com uma das reivindicações 1 a 8.
10. Processo para a preparação de um produto genético heterólogo numa célula hospedeira eucariótica, caracterizado por se introduzir uma molécula de ácidos nucleicos de acordo com uma das reivindicações 1 a 8 na célula hospedeira eucariótica, por se cultivar a célula hospedeira e se isolar o produto genético sintetizado.
11. Processo de acordo com a reivindicação 10, caracterizado por a célula hospedeira ser uma célula de hamster.
12. Processo de acordo com a reivindicação 11, caracterizado por a célula de hamster ser uma célula CHO.
13. Processo de acordo com uma das reivindicações 10 a 12, caracterizado por o gene heterólogo codificar para o activador de plasminogénio de tecido, activador de plasminogénio de tecido de *second generation*, interferão, factor de necrose de tumor, eritropoietina, factor de estimulação da colónia de granulócitos, superoxidodismutase de manganês, uma cadeia de imunoglobulina, a região variável de uma cadeia de imunoglobulina, uma cadeia de imunoglobulina humanizada, a região variável de uma cadeia de imunoglobulina humanizada, um anticorpo de *single-chain* e/ou um anticorpo que é específico para CD44 variante.

14. Utilização de uma molécula de ácidos nucleicos contendo uma sequência promotora de acordo com uma das reivindicações 1 a 8 para a preparação de um produto genético heterólogo em células de hamster cultivadas, de um modo preferido células CHO.

Lisboa, 9 de Março de 2007

Ser Humano	TGGAGCCGCAACCAAAATGCAGATTTTCGTGAAAACCCCTTACGGGGAAGACCATCACCT	82
CHO	TGGAACCGCCGCCAAGATGCAGATTTTCGTGAAGACCCCTTACGGGGAACGATCAGCT	60
	***	
Ser Humano	CGAGGTTGAACCCCTCGGATACGATAGAAAATGTAAAGGCCAAGATCCAGGATAAGGAAGG	142
CHO	CGAGGTTGAACCCCTCGGACACTATAGAAAATGTAAAGGCCAAGATCCAGGATAAGGAAGG	120
	▲	
Ser Humano	AATTCCTCCTGATCAGCAGAGACTGATCTTTGCTGGCAAGCAGCTAGAAGATGGACGTAC	202
CHO	AATTCCTCCTGACCAGCAGAGGCTGATCTTTGCTGGTAAGCAACTGGAAGATGGCCGTAC	180
	***	
Ser Humano	TTTGTCTGACTACAATAATTCAAAAGGAGTCTACTCTTCATCTTGTGTTGAGACTTTCGTGG	262
CHO	TTTGTCTGACTACAACATCCAAAAGGAGTCCACCCTTCATCTTGTGTTGAGACTTCGTGG	240
	▲	
Ser Humano	TGGTGCTAAGAAAAGGAAGAAGAAGTCTTACACCACTCCCAAGAAGAATAAGCACAAGAG	322
CHO	TGGTGCTAAGAAGAGGAAGAAGAAGTCTTACACCACTCCCAAGAAGAATAAGCATAAGAG	300
	***	
Ser Humano	AAAGAAGGTTAAGCTGGCTGTCTGAAATATTATAAGGTGGATGAGAATGGCAAAATTAG	382
CHO	AAAGAAGGTTAAGTTGGCTGTCTGAAATATTATAAGGTGGATGAAAATGGCAAAATTAG	360
	▲	
Ser Humano	TGCCTTCGTGAGAGTGCCCTTCTGATGAATGTGGTGCTGGGGTGTATTATGGCAAGTCA	442
CHO	TGCCTTCGTGAGAGTGTCATCTGATGAGTGTGGTGCTGGAGTTTTCATGGCTAGCCA	420
	***	
Ser Humano	CTTTGACAGACATTATGTGGCAAATGTTGTCTGACTTACTGTTCAACAAACCAGAAGA	502
CHO	TTTTGACAGACATTACTGTGGCAAGTGTGTCTGACTTACTGCTTCAACAAACCAGAAGA	480
	***	
Ser Humano	CAAGTAA	509
CHO	CAAGTAGTTGTGTATGAATAAATAAAAA	508

92.2% de homologia

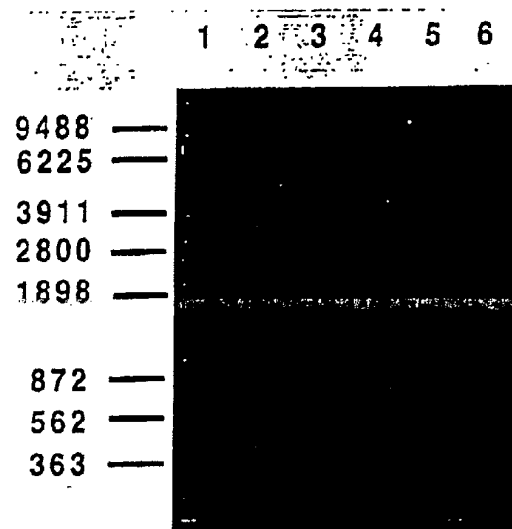
— Sinal poli A    ::::codão de início    \*\*\*    codão de paragem

Fig. 1

Ser Humano	MQIFVKTLTGKTITLEVEPSDTIENVKAKIQDKEG *****	35
CHO	MQIFVKTLTGKTITLEVEPSDTIENVKAKIQDKEG	35
Ser Humano	IPPDQQRLLIFAGKQLEDGRTLSDYNIQKESTLHLV *****	70
CHO	IPPDQQRLLIFAGKQLEDGRTLSDYNIQKESTLHLV	70
Ser Humano	LRLRGGAKKRKKKSYTTPKKNKHKRKKVKLAVLKY *****	105
CHO	LRLRGGAKKRKKKSYTTPKKNKHKRKKVKLAVLKY	105
Ser Humano	YKVDENGKISRLRRECPSDECGAGVFMASHFDRHY *****	140
CHO	YKVDENGKISRLRRECPSDECGAGVFMASHFDRHY	140
Ser Humano	CGKCCLTYCFNKPEDK *****	156
CHO	CGKCCLTYCFNKPEDK	156

100% de homologia

Fig. 2



*Fig 3*

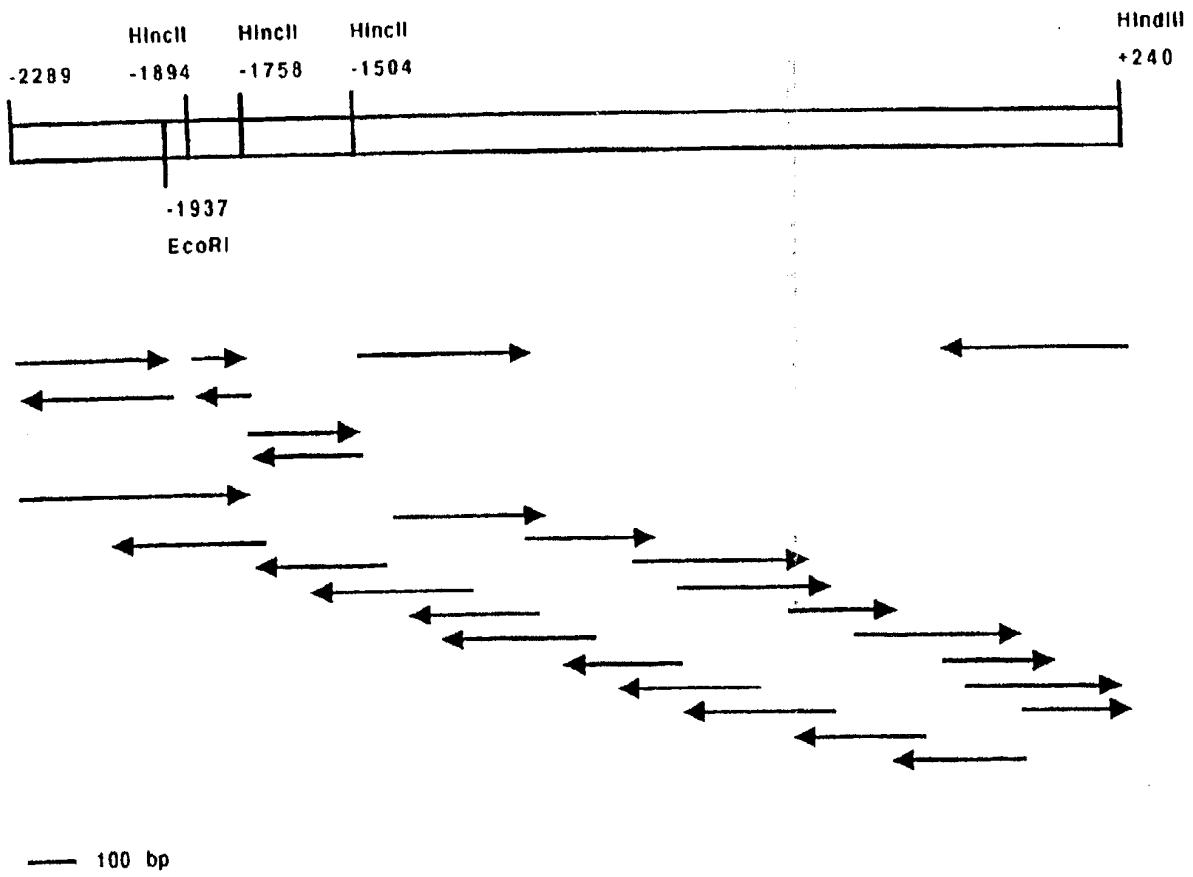


Fig. 4

GATCTCCAGGACAGCCATGGCTATTACACAGAGAAACCTGTCTGGAAAA	- 2240
ACAAAAAATTAGTGTCATGTGTAAATGTGTGGAGTATGCTTGTTCATGCC	- 2190
ACATACAGAGGTAGAGGGCAGTTTATGGGAGTCAGTTCCTATTCTTCCTT	- 2140
TATGGGGGACCTGGGGACTGAACTCAGGTCATCAGGCTTGGCAGAAAGTG	- 2090
CATTAGCTCACGGAGCCTTATCATTGGCGAAAGCTCTCTCAAGTAGAAAA	- 2040
TCAATGTGTTTGCTCATAGTGCAATCATTATGTTTCGAGAGGGGAAGGGT	- 1990
ACAATCGTTGGGGCATGTGTGGTCACATCTGAATAGCAGTAGCTCCCTAG	- 1940
<u>GAGAATTCCAAGTTCTTTGGTGGTGTATCAATGCCCTTAAAGGGGTCAAC</u>	- 1890
EcoRI	HincII
AACTTTTTTCCCTCTGACAAACTATCTTCTTATGTCCTTGTCCCTCAT	- 1840
ATTGGAAGTATTTTATCTTTGTCAGTGTGGAATATCAATTCTAGCACCTC	- 1790
AGACATGTTAGGTAAGTACCCTACAACCTCAGGTTAACTAATTTAATTTAA	- 1740
HincII	
CTAATTTAACCCCAACACTTTTTCTTTGTTTATCCACATTTGTGGAGTGT	- 1690
GT	- 1640
GTGTGTGTGCCGCGCGCGCTCGGATCATTCTACCTTTTGTTTAAAAAATG	- 1590
TTAGTCCAGGGGTGGGTGCACTGTGAAAGTCTGAGGGTAACTTGCTGGG	- 1540
-> 1612 bp	
GTCAGTTCTTTCCACTATAGGACAGAACTCCAGGTGTCAACTCTTTACTG	- 1490
HincII	
ACAGAACCATCCAAATAGCCCTATCTAATTTTAGTTTTTTATTTATTTAT	- 1440
*****	
TTTTTGTTTTTCGAGACAGGGTTTCTCTGTGGCTTTGGAGGCTGTCCTGG	- 1390
*****	
AACTAGCTCTGTAGACCAGGCTGGTCTCGAACTCAGAGATCCACCTGCC	- 1340
TCTGCCTCCTGAGTGCTGGGATTAAAGGCATGCGCCACCAACGCTTGGCT	- 1290
CTACCTAATTTTAAAAGAGATTGTGTGTCAAGGGTGTCTGTGCGCCCT	- 1240
GCAACCAACCCCCCCCCCAAAAAAAAAAAAAAAAAAACTTCACTGAAGCTGA	- 1190
AGCACGATGATTTGGTTACTCTGGCTGGCCAATGAGCTCTAGGGAGTCTC	- 1140
CTGTCAAACAGAATCTCAACAGGCGCAGCAGTCTTTTTTAAAGTGGGGTT	- 1090
ACAACACAGGTTTTTGCATATCAGGCATTTTATCTAAGCTATTTCCCAGC	- 1040
CAAAAATGTGTATTTTGGAGGCAGCAGAGCTAATAGATTAAAATGAGGGA	- 990
AGAGCCCACACAGGTTATTAGGAAGATAAGCATCTTCTTTATATAAAACA	- 940
AAACCAAAACCAACTGGAGGAGGTCTACCTTTAGGGATGGAAGAAAAGAC	- 890
ATTTAGAGGGTGCAATAGAAAGGGCACTGAGTTTGTGAGGTGGAGGACTG	- 840
GGAGAGGGCGCAACCGCTTTAACTGTCTGTGTTTGCCTATTTTTTGGGGA	- 790

Fig. 5a

CAGCACATGTTCCCTATTTTTCCAGGATGGGCAATCTCCACGTCCTCAAACT - 740

| -> 806 bp

TGCGGTTCGAGGACTACAGTCATTTTTGCAGGTTTCTTACTGTATGGCTTT - 690

TAAACCGTGCAAAGGTGACCATTAACCGTTTCACGCTGGGAGGGCACGTG - 640

CGGCTCAGATGCTTCCTCTGACTGAGGGCCAGGAGGGGGCTACACGGAAG - 590

AGGCCACACCCGCACTTGGGAAGACTCGATTTGGGCTTCAGCTGGCTGAG - 540

ACGCCCCAGCAGGCTCCTCGGCTACACCTTCAGCCCCGAATGCCTTCCGG - 490

CCCATAACCCCTTCCCTTCTAGGCATTTCCGGCGAGGACCCACCCTCGCGC - 440

CAAACATTTCGGCCCCATCCCCCGTCTCCTCACCTGAATCTCTAACTCTGGA - 390

| -> 483 bp

CTCCAGAGTTTAGAGACTATAAGCAGATAGCCCGGATGTGTGGAACTGCA - 340

TCTTGGGACGAGTAGTTTATAGCAAAAAGAAAGCGACGAAAACTACAATT - 290

CCCAGACAGACTTGTGTTACCTCTCTTCTCATGCTAAACAAGCCCCCTTT - 240

AAAGGAAAGCCCCCTCTTAGTCGCATCGACTGTGTAAGAAAGGCGTTTGAA - 190

| -> 272 bp

ACATTTTAATGTTGGGCACACCGTTTTGAGGACCGAAATGAGAAAGAGCA - 140

TAGGGAAACGGAGCGCCCGAGCTAGTCTGGCACTGCGTTAGACAGCCGCG - 90

SacII

| -> 156 bp

GTGCTTGCAGCGGGCAGGCACTTTCGTGGACGCCAAGGGGCGGGTCTTTTC - 40

>>>>>

+1

GGCCGGGAAGCCCCGTTGGTCCGCGCGGCTCTTCTTTCCGATCCGCCAT + 11

EagI

+13

CCGTGGTGAGTGTGTGCTGCGGGCTGCCGCTCCGGCTTGGGGCTTCCCGC + 61

+85

GTGCTCTCACCCCTGGTCGGCGGCTCTAATCCGTCTCTTTTTCGAATGTAG + 111

+161

GTGGAACCGCCGCCAAGATGCAGATTTTCGTGAAGACCCCTTACGGGAAA + 161

M Q I F V K T L T G K

ACGATCACGCTCGAGgtacgaaccaggtggcgtgagaagcgaaggcctgc + 211

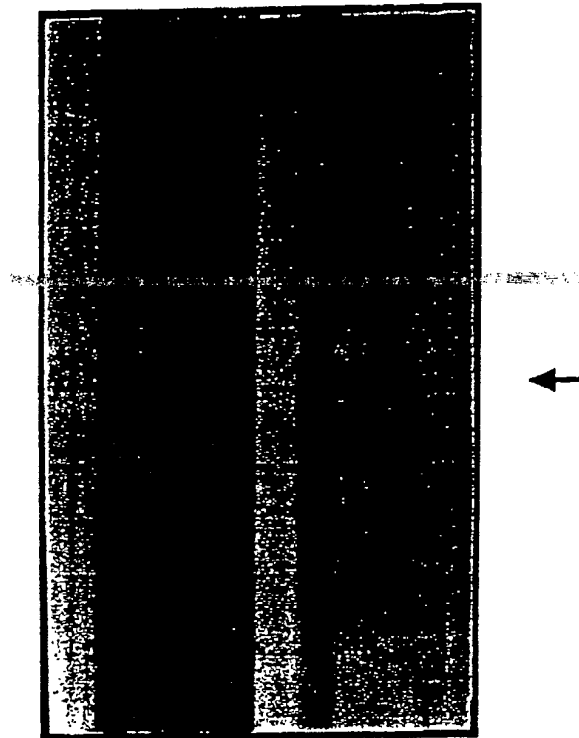
T I T L E

cagaggccctctatgctcgcttaaaagctt - 240

HindIII

*Fig. 5b*

1 2



*Fig. 6*

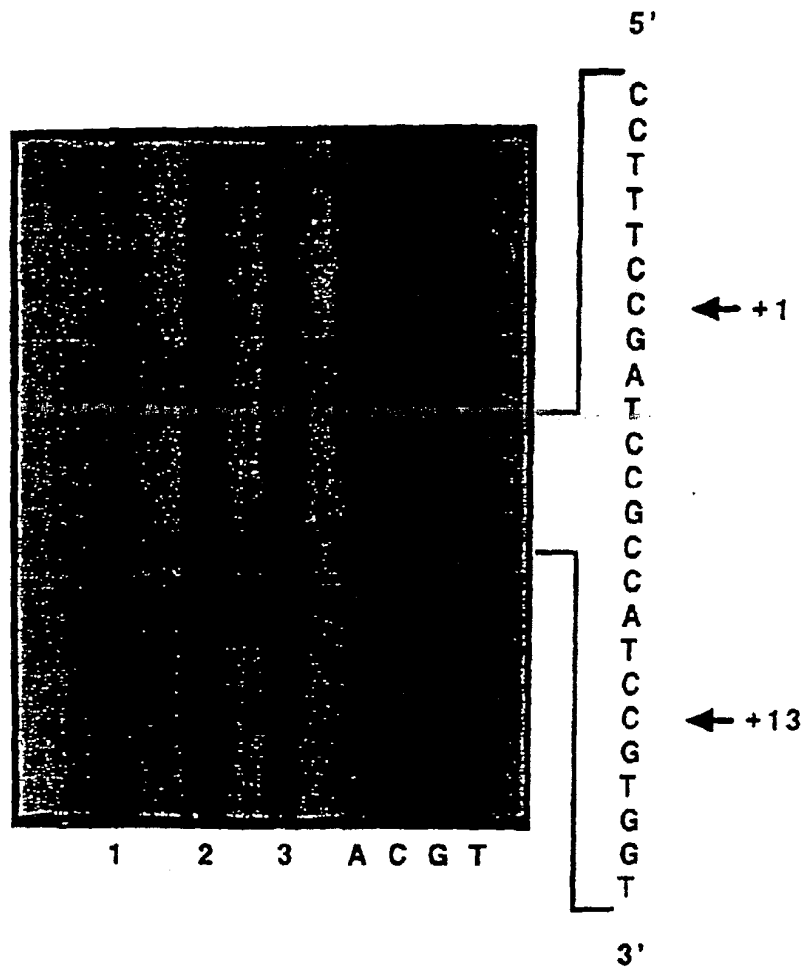
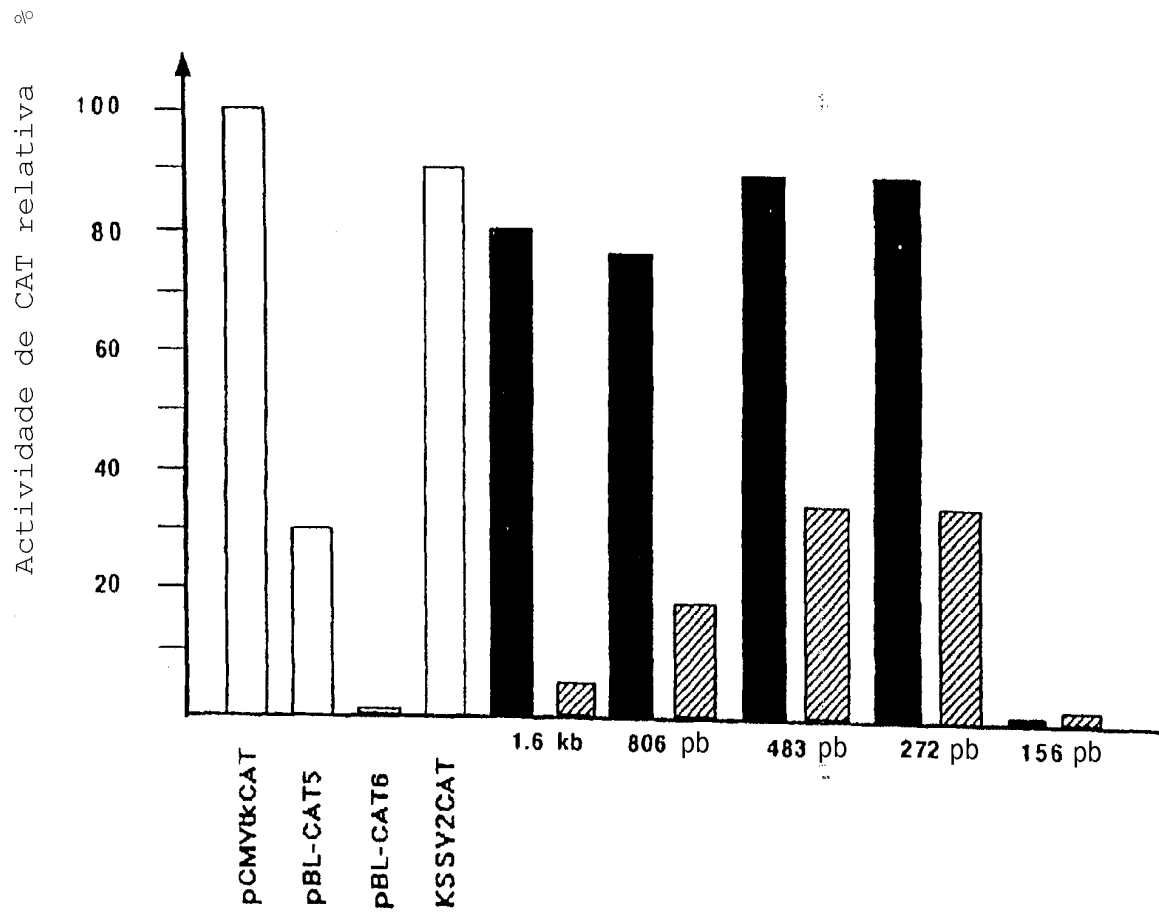


Fig. 7

*Fig. 8*

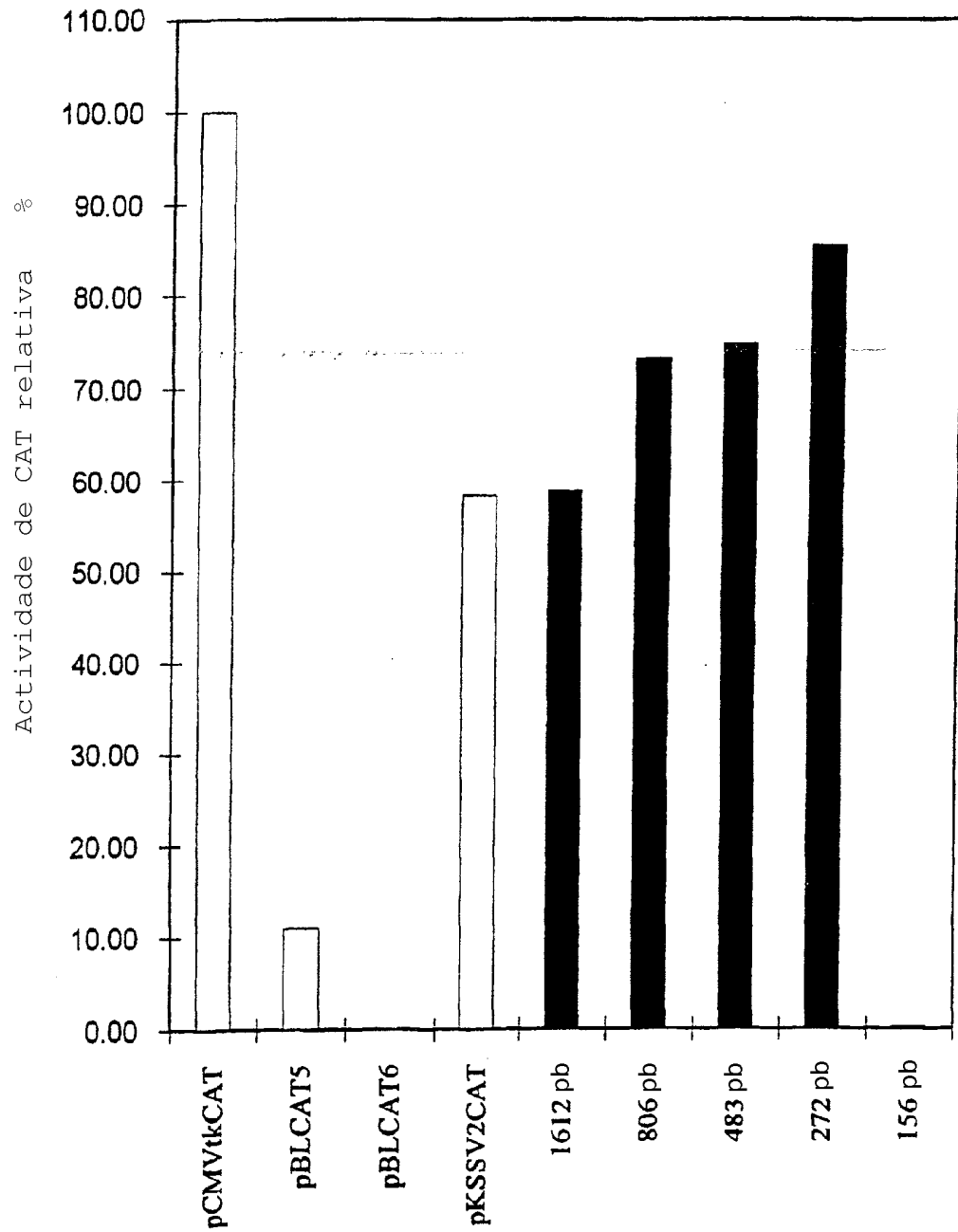


Fig. 9

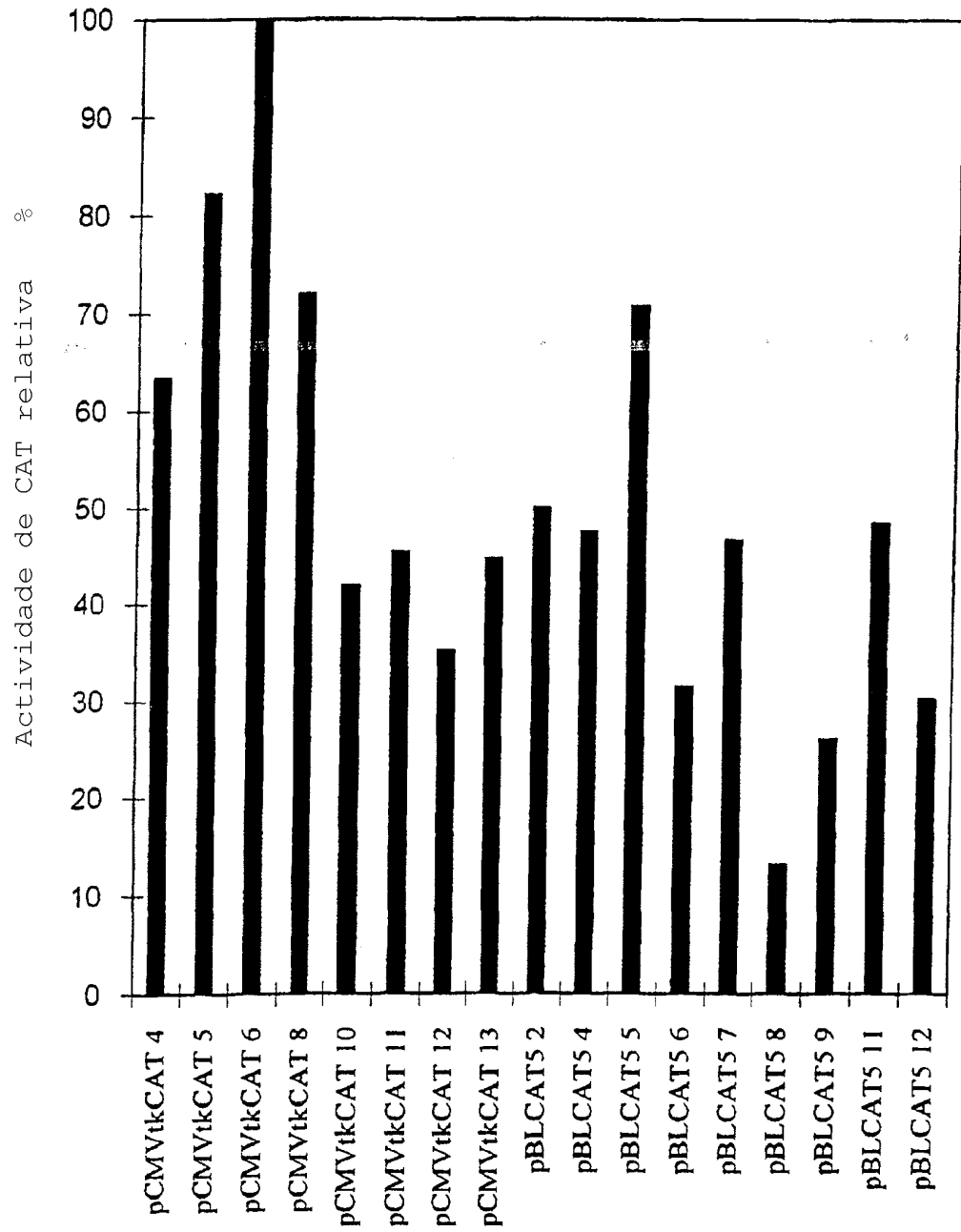


Fig. 10

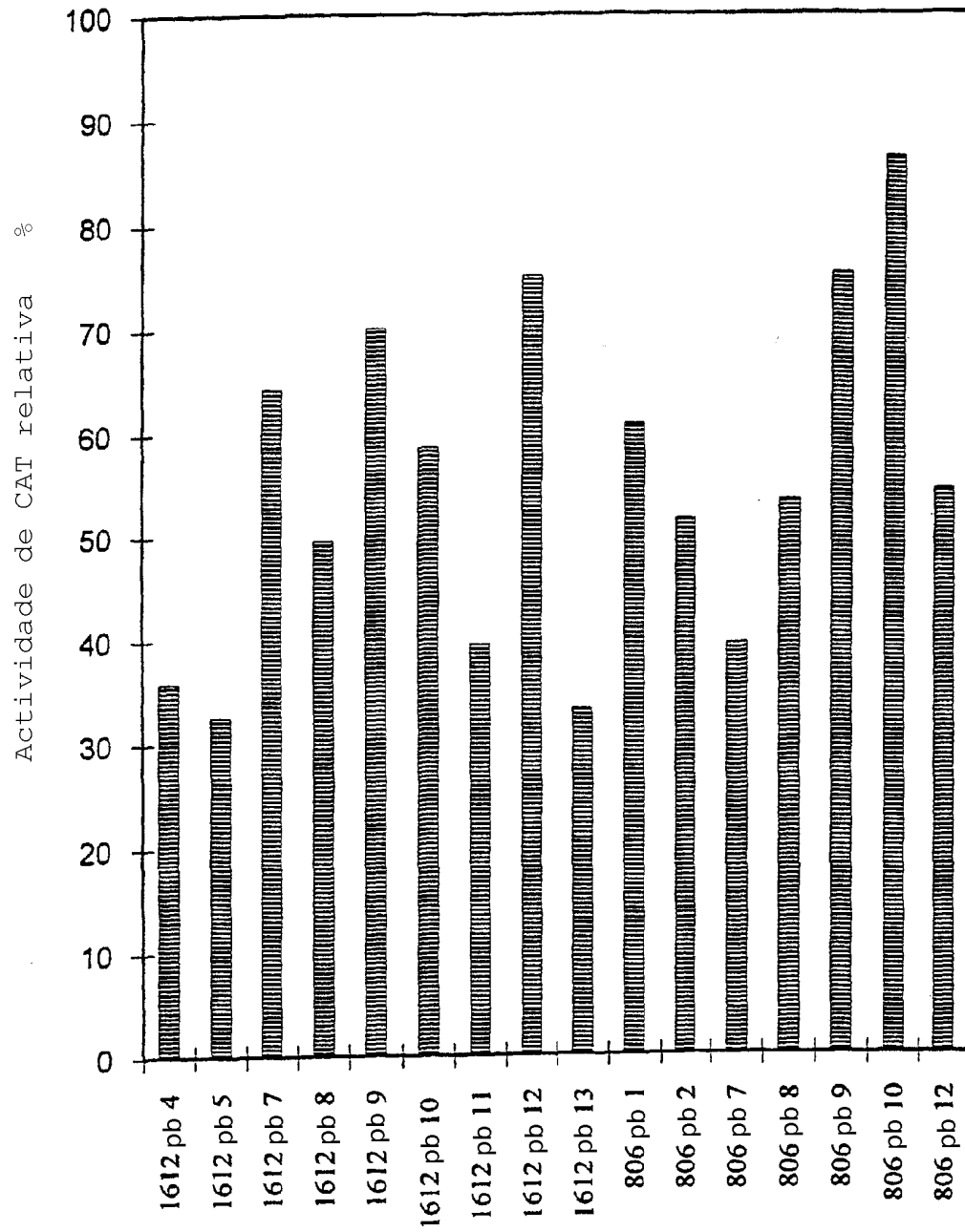


Fig. 11

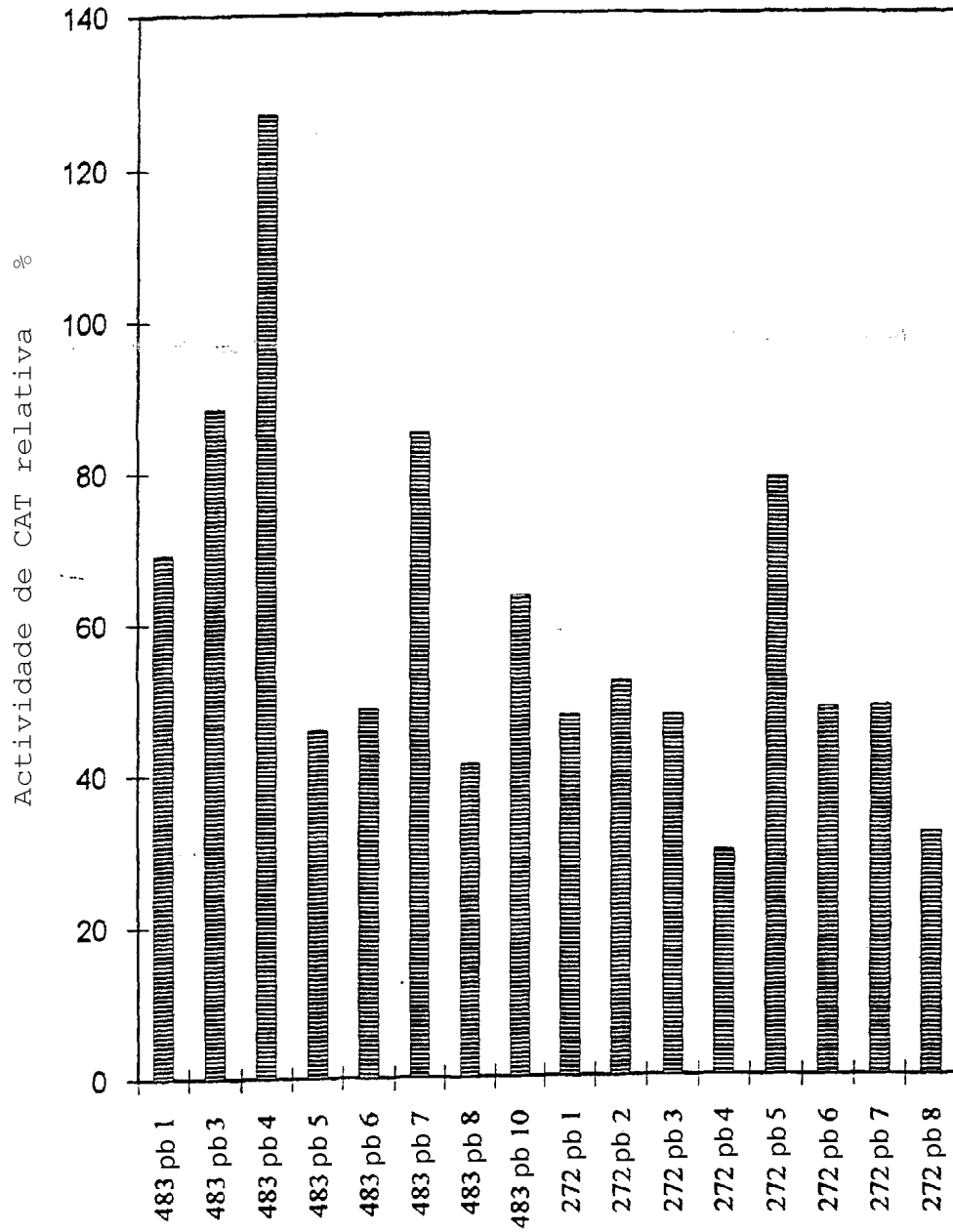


Fig. 12

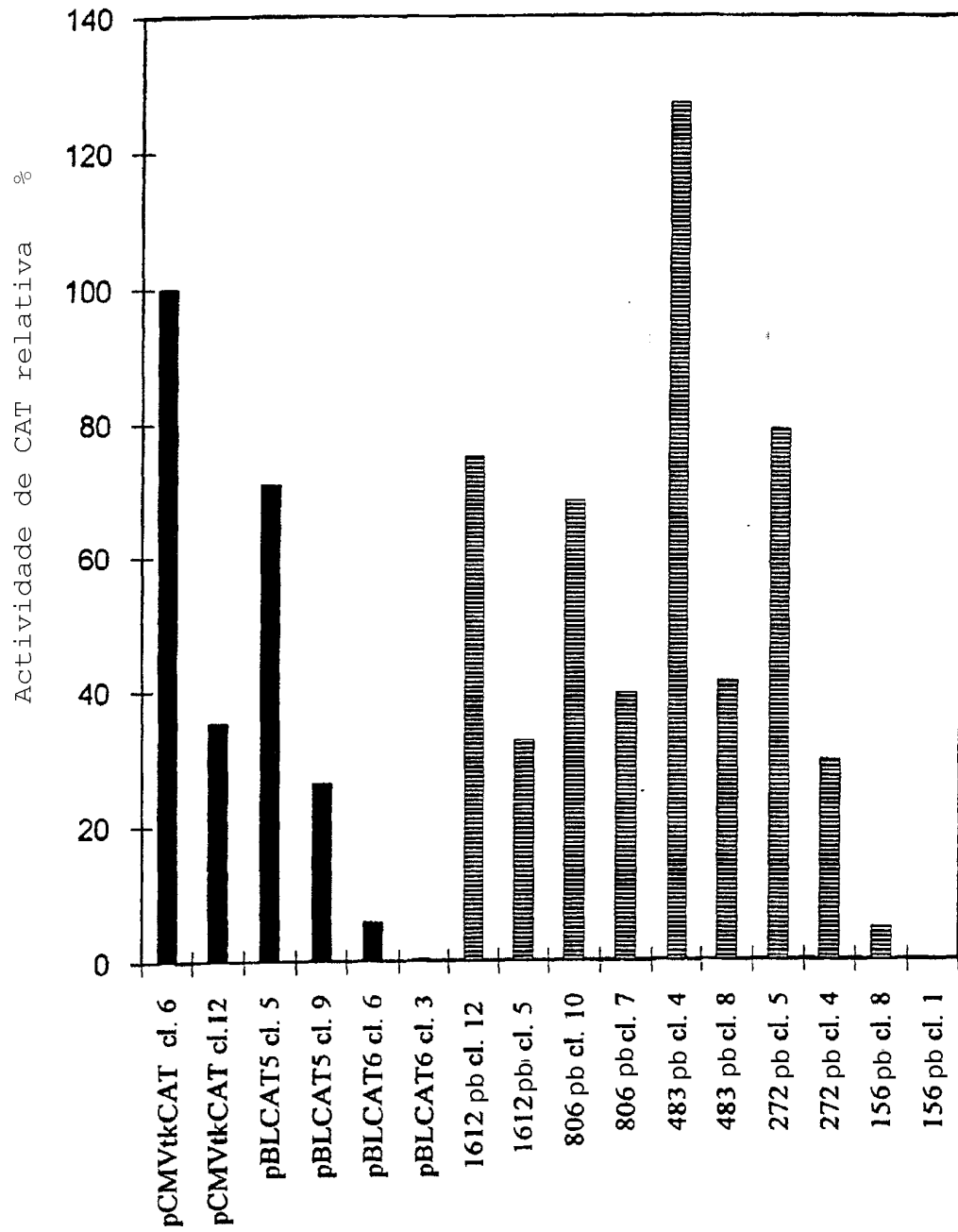


Fig. 13