

Memória descritiva referente à patente de invenção de HOECHST AKTIENGESELLSCHAFT, alemã, industrial e comercial, com sede em D-6230 Frankfurt am Main 80, República Federal Alemã, (inventores: Dr. Michael Dörschug e Dr. Gerhard Seipke, residentes na Alemanha Ocidental), para "PROCESSO PARA A CISÃO SELETIVA DE PROTEÍNAS DE FUSÃO".

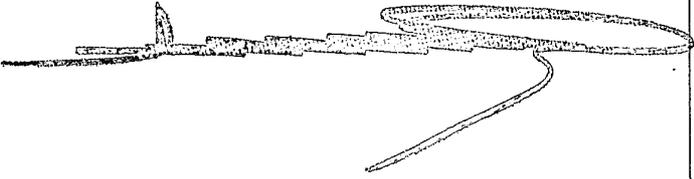
Memória descritiva

A invenção refere-se a um processo para a preparação de polipéptidos ou de proteínas por meio de cisão enzimática de uma proteína de fusão.

O significado crescente da tecnologia de DNA recombinante para a obtenção de polipéptidos ou de proteínas requer o desenvolvimento de novos processos para o enriquecimento e purificação dos produtos que se ajustem aos materiais de partida modificados.

Actualmente sintetiza-se um grande número de proteínas em microorganismos na forma de proteínas de fusão, isto é, antes da sequência de aminoácidos de polipéptido desejado introduz-se a sequência de uma proteína estranha (F.A.O. Harston, Biochem. J. 240(1986) 1-12). As proteínas de fusão precipitam em geral a célula devido à sua baixa solubilidade ou mesmo insolubilidade, na forma dos chamados corpúsculos de inclusão, encontrando-se por esta razão protegidos da degradação proteolítica. Conseguem-se assim elevados rendimentos e facilidade de isolamento deste produto genético primário.

Para se recuperar o polipéptido desejado tem contudo de se separar este da proteína de fusão por meio de cisão enzimática ou química. Efectua-se frequentemente a cisão por méto-



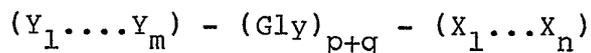
dos químicos, porque estes se adaptam mais facilmente ao carácter dificilmente solúvel da proteína de fusão. Verificam-se contudo cisões incompletas ou formação de produtos secundários por derivatização irreversível de cadeias laterais de aminoácidos, em praticamente todos os métodos químicos. Existe também sempre o perigo da degradação não específica das cadeias polipeptídicas (K.-K. Han et al., Int. J. Biochem. 15, (1983)875-884). O emprego de processos químicos é além disso limitado uma vez que a especificidade de cisão é determinada essencialmente por um único aminoácido que existe também muitas vezes naturalmente no polipéptido desejado.

As fragmentações enzimáticas podem efectuar-se em condições muito mais convenientes. Contudo, neste caso, as dificuldades gerais resultam do facto de a proteína de fusão pouco solúvel ser posta e mantida em solução com detergentes, ureia ou cloridrato de guanidina, isto é, em condições nas quais os enzimas se inactivam frequentemente. As proteases que reconhecem apenas um único aminoácido específico, não se podem empregar muitas vezes, dado que esse aminoácido também se encontra presente na proteína desejada. Por outro lado, existe apenas um pequeno número de proteases, que reconhecem apenas determinadas e raras sequências de entre vários aminoácidos. Para cada produto será assim necessário arranjar um processo de cisão específico. É deste modo desejável conseguir um método de cisão de aplicação universal, que efectue a cisão apenas de determinadas e raras sequências de aminoácidos, sem degradar a proteína, e quase possa aplicar também a proteínas de fusão dificilmente solúveis.

Da literatura conhece-se a lisostafina como enzima degradador da parede celular, que é segregado para o meio a partir do *staphylococcus simulans* (NRRL B-2628; Sloan et al., Int. J. of Systematic Bacteriology, Vol. 32, Nº. 2 (1982) 170-174). Este enzima promove a lisagem de praticamente todos os tipos conhecidos de *Staphylococcus*, mas não de qualquer outra espécie de bactéria. Até agora considerava-se que a lisostafina-endoprotease efectuava a cisão de um modo muito selectivo apenas das pontes de poliglicina no Sacculus mureínico dos Sta-

phylococcus (Iversen e Grov., Eur. J. Biochem. 38 (1973) 293-300). Além disso, conhecem-se experiências de transpeptidinação, sob catálise por lisostafina, com péptidos glicílicos sintéticos de cadeia curta (G.L. Sloan et al., Biochem. J. 167 (1977) 293-296). Descobrimos então surpreendentemente que a lisostafina-endoprotease também efectua a cisão de proteínas de fusão com uma sequência de oligo- ou de poli-glicina. A posição desta sequência no arranjo estereoquímico específico de uma parede celular bacteriana, não é manifestamente determinante para a selectividade da cisão. Tal facto possibilita a cisão exacta desejada de uma determinada proteína a partir de uma proteína de fusão.

O objectivo da presente invenção é um processo para a preparação de polipéptidos ou de proteínas por meio de cisão enzimática, caracterizado por se cindir uma proteína de fusão com a sequência



na qual

$(Y_1 \dots Y_m) - (Gly)_p$ representa a sequência a cindir e

$(Gly)_q - (X_1 \dots X_n)$ representa o polipéptido ou a proteína,

X e Y, independentemente um do outro, significam aminoácidos naturais,

m e n são números maiores do que 1,

p ou q são um número maior do que 0, e

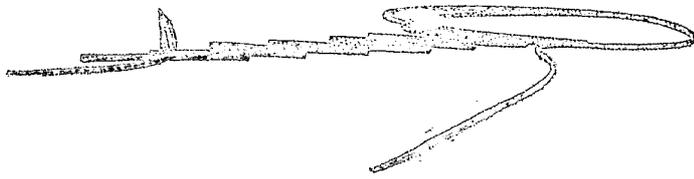
p+q significam um número natural entre 2 e 100,

no caso em que Y_m seja Gly, p+q pode também ser < 2 e p = 0 e

no caso em que X_1 seja Gly, p+q pode também ser < 2 e q = 0,

com uma endoprotease específica para sequências de oligo- ou poli-glicina, e se submeter o polipéptido ou a proteína assim libertado, conforme o caso, a um tratamento posterior químico ou enzimático, ou se continuar a processar directamente.

Conforme a grandeza dos índices p ou q, submete-se a sequência a cindir, ou o polipéptido ou a proteína, conforme o caso, a posteriores cisões enzimáticas.



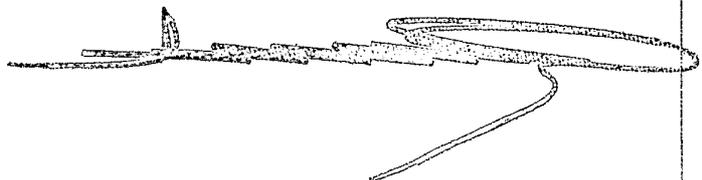
$(Y_1 \dots Y_m)$ é uma sequência de proteína natural ou artificial, como as habitualmente empregues para a preparação de proteínas de fusão. Referem-se por exemplo β -galactosidase, enzimas do metabolismo do triptofano ou partes destas moléculas de proteína, que conduzem em geral a produtos insolúveis, e ainda sequências de polipéptido que possibilitam um rápido enriquecimento de um produto de fermentação solúvel (por exemplo, anticorpos).

$(X_1 \dots X_n)$ é um polipéptido ou proteína farmacologicamente activo ou um precursor de mais elevado peso molecular, a partir do qual se pode obter a forma biologicamente activa desejada, através de um processamento posterior como dobragem com formação de pontes de dissulfeto correctas e/ou da referida cisão das cadeias de polipéptido. Um exemplo disto será a pré-proinsulina, a partir da qual se obtém a insulina.

Os grupos X e Y são, independentemente um do outro, aminoácidos naturalmente ocorrentes (vide por exemplo Schröder, Lübke "The Peptides" vol. I, New York 1965, pags. 137-270 e Houben-Weyl "Methoden der organischen Chemie", vol. 15/1 e 2 (Síntese de péptidos), Georg Thieme Ed., Stuttgart 1974, Suplemento). Referem-se em particular: Gly, Ala, Ser, Thr, Val, Leu, Ile, Asp, Asn, Glu, Gln, Cys, Met, Arg, Lys, Hyl, Orn, Cit, Tyr, Phe, Trp, Mis, Pro e Hyp.

Por uma endoprotease específica para sequências de oligo- ou poli-glicina entende-se em especial a lisostafina (Fabricante: ^(R)SIGMA Chemie GmbH, Deisenhofen; vide também Recsei et al., Proc. Natl. Acad. Sci. USA, vol 84, (1987) 1127-1131).

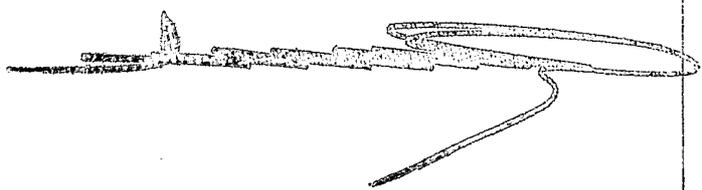
O produto desejado $(X_1 \dots X_n)$ liga-se com uma proteína $(Y_1 \dots Y_m)$ através de uma sequência de glicina adequada $(Gly)_{p+q}$. Preferem-se as proteínas de fusão em que p+q significam em conjunto um número natural entre 2 e 30. Um maior número de grupos glicina consecutivos pode influenciar positivamente a velocidade da reacção devido à melhor acessibilidade da posição de cisão. Consegue-se o mesmo efeito quando curtas sequências de oligo-glicina se encontram flanqueadas por um ou mais outros amino-



ácidos que não impedem estereoquimicamente o enzima na reacção de cisão. O comprimento do troço de ligação está assim relacionado com as propriedades estruturais dos parceiros de fusão.

As condições reaccionais de cisão podem variar numa gama muito larga conforme as propriedades da proteína de fusão. Deste modo a proporção enzima/substrato pode situar-se por exemplo entre 1:1 e 1:1 000 000 e efectuar-se a reacção numa gama de pH de 6 a 9, de preferência na gama de 7 a 8, de preferência a uma temperatura de 20 a 60°C, em especial de 32 a 42°C. Eventualmente e conforme o grau de insolubilidade da proteína de fusão, pode ainda adicionar-se um aditivo, por exemplo ureia, detergentes ou cloridrato de guanidina, que ajude a manter a proteína de fusão em solução. A cisão ocorre mais rapidamente contudo, quando se consegue uma solubilização quase completa da proteína de fusão, sem adição de agentes de desnaturação, e embora a lisostafina-endoprotease não seja desactivada pela presença de ureia, a reacção enzimática torna-se mais lenta. Este fenómeno pode aproveitar-se para a cisão de proteínas de fusão particularmente pouco solúveis. Por outro lado, podem efectuar-se com sucesso, durante um tempo de reacção mais longo, fragmentações dos corpúsculos de inclusão em suspensão. Em soluções da proteína de fusão - com ou sem agente de desnaturação - pode utilizar-se com vantagem a lisostafina-endoprotease numa forma ligada a um suporte (enzima immobilizado) e recuperar-se de novo para posterior utilização. Como suportes de enzima referem-se por exemplo suportes inorgânicos como silicatos de alumínio ou então polímeros, suportes orgânicos como agaroses, por exemplo ^(R)Affi-Gel 10 (Firma Bio Rad), celulosas, gels de poliacrilamida modificados contendo grupos amino ou hidroxilo, ou ainda copolímeros orgânicos de acrilamida, metacrilatos ou metacrilamida e anidrido maleico. A preparação dos respectivos enzimas suportados encontra-se descrita por exemplo em Halwachs et al., *Biotechnology and Bioengineering* XIX (1977) 1667-1677 (fixação em silicato de alumínio) ou na DE-OS 29 27 534 (fixação em celulose).

O processo de acordo com a invenção é adequado não só para a cisão selectiva de proteínas de fusão, mas também para



aplicação aos respectivos polipéptidos em geral.

Os exemplos seguintes destinam-se a ilustrar a presente invenção, sem contudo a eles a limitarem. Os exemplos 3 e 4 mostram que as cisões de proteína de fusão obtidas nos exemplos 1 e 2 se devem atribuir à construção de uma sequência de poliglicina e não à degradação de $(X_1 \dots X_n)$ ou de $(Y_1 \dots Y_m)$.

Exemplo 1

Cisão de uma proteína de fusão contendo poliglicina

Utiliza-se uma construção em que um segmento de β -galactosidase se encontra ligado à proinsulina por meio de um péptido $(Gly)_{18}$ e de um hexapéptido sintético. A proinsulina forma o extremo carboxilado da proteína de fusão. A proteína está enriquecida em até cerca de 40%.

Dissolve-se esta proteína com 20 mg/ml de tampão (ureia 8M; Tris/HCl 50 mM; pH 7,5) e, por diluição lenta com 50 mM Tris/HCl; pH 7,5 e ureia 8M; pH 7,5, leva-se a diferentes concentrações de ureia (vide Tabela 1) e a uma concentração de proteína de 2 mg/ml ou 10 mg/ml. Às soluções ligeiramente turvas adiciona-se lisostafina (^(R)SIGMA) numa proporção enzima/substrato de 1:100 ou 1:1000, e conservam-se a 37°C. A intervalos determinados retiram-se amostras e analisam-se por meio de electroforese SDS. A degradação selectiva da proteína de fusão na sequência de poliglicina pode determinar-se pelo desaparecimento da banda da proteína de fusão e pelo aparecimento de uma nova banda para o fragmento de galactosidase com um peso molecular mais baixo de 1 0000 Dalton.

Tabela 1

Concentração de proteína (mg/ml)	Proporção enzima/substrato	Concentração de ureia (M)	Tempo de reacção (>95% de desaparecimen- to da banda da proteí- na de fusão) (horas)
2	1 : 100	4	> 20
2	1 : 100	3	20
2	1 : 100	2	5
2	1 : 100	1	2
10	1 : 1000	2	20

Na Tabela 1 indicam-se os tempos, após os quais por determinação densitométrica da electroforese SDS, a área da banda da proteína de fusão original, se tinha reduzido a menos de 5%. Os valores obtidos situam-se entre 1 e 20 horas conforme a concentração de ureia e a proporção enzima/substrato. A presença de um agente de redução, por exemplo ditioeritritol (DTE, reagente de Cleland), que é vantajoso para a dissolução da proteína de fusão, não tem qualquer influência significativa sobre a actividade do enzima.

Exemplo 2

Cisão de uma proteína de fusão em suspensão

Utiliza-se a mesma construção do exemplo 1, mas não na forma seca isolada. Utiliza-se antes uma suspensão, como a que ocorre normalmente no isolamento dos corpúsculos de inclusão e que contém, em cerca de 200 g/l de substância seca (90% de proteína), cerca de 50 g/l de proteína de fusão.

Diluem-se 10 ml desta suspensão com 50 mM Tris/HCl; pH 7,5 a 200 ml e misturam-se com 20 mg de lisostafina (^(R)SIGMA). Após 20 horas a 37°C a electroforese SDS apresenta apenas leves vestígios de proteína de fusão não fragmentada.

Exemplo 3

Incubação de proinsulina com lisostafina

Dissolve-se 1 mg/ml de proinsulina (humana) em 50 mM Tris/HCl; pH 7,5 e adiciona-se lisostafina numa proporção enzima/substrato de 1:100. Conserva-se a solução a 37°C. Após 1,3, 5 e 20 horas retiram-se amostras e analisam-se por meio de HPLC de fase reversa. Estas análises não revelam qualquer degradação da proinsulina pela lisostafina.

Exemplo 4

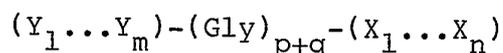
Ensaio de cisão de uma proteína de fusão sem sequência de poliglicina

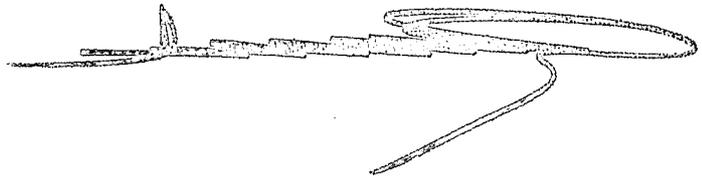
Utiliza-se uma construção na qual um segmento de B-galactosidase se encontra ligado à proinsulina por meio de um hexapéptido sintético. A proinsulina forma o extremo carboxilado da proteína de fusão. A proteína de fusão está enriquecida até cerca de 80%. Dissolve-se esta proteína com 20 mg/ml de tampão (ureia 8M; 50 mM Tris/HCl; pH 7,5) e, por diluição lenta com 50 mM Tris/HCl; pH 7,5, leva-se a uma concentração de proteína de 2 mg/ml. À solução ligeiramente turva adiciona-se lisostafina (^(R)SIGMA) na proporção enzima/substrato de 1:100 e conserva-se a 37°C. A análise de amostras por electroforese SDS não revela qualquer degradação da proteína de fusão, mesmo após 20 horas.

REIVINDICAÇÕES

- 1a -

Processo para a preparação de um polipéptido ou de uma proteína por meio de cisão enzimática, caracterizado por se cindir uma proteína de fusão com a sequência





na qual

$(Y_1 \dots Y_m) - (\text{Gly})_p$ representa a sequência a cindir e

$(\text{Gly})_q - (X_1 \dots X_n)$ representa o polioéptido ou a proteína X e Y, independentemente entre si, significam aminoácidos naturais,

m e n significam números superiores a 1,

p ou q significam um número superior a 1 e

p + q significam em conjunto um número natural entre 2 e 100, caso Y seja Gly, p + q pode também ser < 2 e p=0 e caso X_1 seja Gly, p + q pode também ser < 2 e q=0,

com uma endoprotease específica para sequências de oligo- ou poli-glicina, empregando uma razão enzima/substrato entre 1:1 e 1:1000 000 e efectuando a reacção numa gama de pH entre 6 e 9 e a uma temperatura entre 20 e 60°C,

e se sujeitar, conforme o caso, o polipéptido ou a proteína assim libertados a um tratamento posterior químico ou enzimático, ou se utilizá-los directamente.

- 2a -

Processo de acordo com a reivindicação 1, caracterizado por se empregar como endoprotease a lisostafina.

- 3a -

Processo de acordo com a reivindicação 1, caracterizado por p+q significarem em conjunto um número inteiro entre 2 e 30.

- 4a -

Processo para a preparação de uma composição farmacêutica, caracterizado por se incorporar como ingrediente farmacologicamente efectivo um polipéptido obtido a partir de uma proteína de fusão, quando preparada de acordo com qualquer das reivindicações anteriores, com veículos e/ou aditivos farmacologicamente adequados.

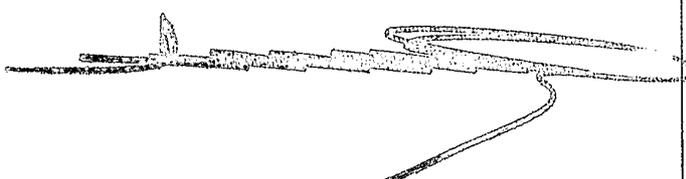
- 9 -

A requerente declara que o primeiro pedido desta patente foi apresentado na República Federal Alemã em 20 de Novembro de 1987, sob o Nº. P 37 39 347.2.

Lisboa, 17 de Novembro de 1988

SECRETARIA DE ESTADO DA INDÚSTRIA

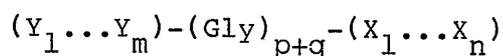
A handwritten signature in black ink, consisting of several loops and a long horizontal stroke, positioned below the typed text.



RESUMO

"PROCESSO PARA A CISÃO SELECTIVA DE PROTEÍNAS DE FUSÃO"

A invenção refere-se a um processo para a preparação de um polipéptido ou de uma proteína por meio de cisão enzimática, que compreende cindir-se uma proteína de fusão com a sequência



na qual

$(Y_1 \dots Y_m) - (\text{Gly})_p$ representa a sequência a cindir e

$(\text{Gly})_q - (X_1 \dots X_n)$ representa o polipéptido ou a proteína

X e Y, independentemente entre si, significam aminoácidos naturais,

m e n significam números superiores a 1,

p ou q significam um número superior a 1 e

p + q significam em conjunto um número natural entre 2 e 100, caso Y seja Gly, p + q pode também ser < 2 e p=0 e caso X_1 seja Gly, p + q pode também ser < 2 e q = 0,

com uma endoprotease específica para sequências de oligo- ou poli-glicina, empregando uma razão enzima/substrato entre 1:1 e 1:1 000 000 e efectuando a reacção numa gama de pH entre 6 e 9 e a uma temperatura entre 20 e 60°C, e se sujeitar conforme o caso, o polipéptido ou a proteína assim libertados a um tratamento posterior químico ou enzimático, ou se utilizá-los directamente.