

# The primary structure of the glutamic acid-specific protease of *Streptomyces griseus*

Ib Svendsen, Martin Roland Jensen and Klaus Breddam

Carlsberg Laboratory, Department of Chemistry, Gamle Carlsberg Vej 10, DK-2500 Copenhagen Valby, Denmark

Received 9 September 1991

The amino acid sequence and part of the DNA sequence of a glutamic acid-specific serine protease from *Streptomyces griseus* is reported. This protease is shown to be homologous with other serine proteases. An improved purification protocol for this enzyme is described.

Serine-protease (*Streptomyces griseus*); Glutamic acid-specific enzyme; Amino acid sequence

## 1. INTRODUCTION

*Streptomyces griseus* is used for the production of the commercial product Pronase, a culture filtrate from which a serine protease which preferentially cleaves at the carboxyl side of glutamic acid has been isolated [1]. We here describe an improved purification procedure for this enzyme and the complete amino acid sequence, partially determined from the DNA sequence of part of the gene. The amino acid sequence shows homologies with proteases A and B from *S. griseus* [2–4] and  $\alpha$ -lytic proteinase of *Lysobacter enzymogenes* [5].

## 2. EXPERIMENTAL

### 2.1. Purification of SGPE

200 g of pronase (Actinase, Kaken Seiyaku, Tokyo, Japan), a culture filtrate of *S. griseus*, was dissolved in 2700 ml water and the pH adjusted to 5.25 by addition of 10% acetic acid. This sample was applied to a 10  $\times$  17 cm CM-52 cellulose column equilibrated with 10 mM sodium acetate pH 5.25. The column was washed with 1.5 l of the same buffer until  $A_{280}$  was below 0.4 and the enzyme was then eluted with a salt gradient from 0 to 0.2 M NaCl in the same buffer (2  $\times$  7 liter). The fractions with activity towards Z-Glu-NH-Ph-NO<sub>2</sub> were pooled, concentrated and diafiltrated towards 10 mM MES, pH 6.0. The concentrate was applied to a 2.6  $\times$  19 cm CABS-Sepharose column [6] equilibrated with 10 mM MES, pH 6.0. The column was washed with 110 ml of the same buffer and the enzyme was eluted with a salt gradient from 0 to 0.35 M NaCl in the same buffer (2  $\times$  600 ml). The fractions with activity towards Z-Glu-NH-Ph-NO<sub>2</sub> were concentrated and diafiltrated against 50 mM HEPES, pH 7.5. To remove residual activity towards Ac-Arg-NH-Ph-NO<sub>2</sub> this sample was applied to a 1.6  $\times$  5 cm Arginine-Sepharose 4 B (Pharmacia) column equilibra-

ted with 50 mM HEPES, pH 7.5. The flowthrough was concentrated, diafiltrated against water and frozen at  $-18^{\circ}\text{C}$ .

Enzyme activity was determined spectrophotometrically at 410 nm using the following assay mixture: 965  $\mu\text{l}$  50 mM Bicine, pH 8.25 + 25  $\mu\text{l}$  40 mM Z-Glu-NH-Ph-NO<sub>2</sub> (gift from Dr E. Dey) in methanol or 25  $\mu\text{l}$  8 mM Ac-Arg-NH-Ph-NO<sub>2</sub> (Bachem) in water or 25  $\mu\text{l}$  Ac-Phe-NH-Ph-NO<sub>2</sub> (Bachem) in methanol + 10–50  $\mu\text{l}$  enzyme solution. Protein concentrations were determined by measurements of the absorbance at 280 nm using  $\epsilon_{280} = 18\,100\text{ M}^{-1}\text{cm}^{-1}$  determined from the amino acid composition. The homogeneity and molecular weight was determined by SDS gel electrophoresis [7]. The amino acid composition was determined after hydrolysis in 6 M HCl at  $110^{\circ}\text{C}$  for 24 h.

### 2.2. Amino acid sequencing

Amino acid sequence analysis was performed using an Applied Biosystems gas phase sequencer model 470A. Reduction and vinylpyridination, CNBr cleavage, digestion with clostripain, chymotrypsin, trypsin and EndoLysC protease (Boehringer) were carried out as previously described [8]. Separation of peptides was made either by gel chromatography on Biogel P6, Sephadex G-50 super fine in 30% acetic acid or by HPLC on a Vydac C<sub>4</sub> column. The C-terminal sequence was determined by digestion with carboxypeptidase Y: 0.5 mg reduced and alkylated SGPE was dissolved in 50  $\mu\text{l}$  0.05 M acetic acid, 1 mM EDTA, 50  $\mu\text{l}$  sodium acetate, 1 mM EDTA was added (resulting pH around 4.5) followed by 2  $\mu\text{l}$  carboxypeptidase Y (3.2 mg/ml). 25  $\mu\text{l}$  aliquots were removed at different reaction times, acidified by addition of 5  $\mu\text{l}$  0.4 M HCl and then applied to the amino acid analyzer.

### 2.3. DNA procedures

*Streptomyces griseus* IMRU 3499 was obtained from the Waksman Institute of microbiology, Rutgers University, Piscataway, NJ. Growth of *S. griseus* mycelium, and chromosomal DNA isolation were essentially performed as previously described [9]. Two mixed oligonucleotides were synthesized using an Applied Biosystems 394 DNA/RNA Synthesizer, based on the amino acid sequence using the known *Streptomyces* codon bias [10] (Fig. 1). Southern blots of *S. griseus* endonuclease digested chromosomal DNA were hybridized with both oligonucleotides at low stringency ( $37^{\circ}\text{C}$ ,  $6\times\text{SSC}$ ,  $1\times\text{Denhardt's}$  solution, 0.5% SDS). The oligonucleotides were end-labelled by 5'-phosphorylation with ( $\gamma$ -<sup>32</sup>P)ATP using T4 polynucleotide kinase. The two oligonucleotides were used as PCR primers, and purified *S. griseus* chromosomal DNA was used as template. The interspersing 240 bp fragment isolated. Double stranded DNA sequencing on both strands was performed on the resulting product. The DNA sequences were generated on an Applied Biosystems Model 373A DNA sequencer using the ABI dye-labeled DyeDeoxy Terminators.

**Abbreviations:** Ac, N-acetyl; CABS-sepharose, [*N*-( $\epsilon$ -aminocaproyl)-*p*-amino-benzyl]succinyl-Sepharose; PCR, Polymerase Chain Reaction; SGPA, SGPB, *Streptomyces griseus* protease A and B, respectively; SGPE, *Streptomyces griseus* protease Glu (E) specific; SSC, 0.15 M NaCl, 0.015 M Na-citrate, pH 7.6; Z, N-carbobenzyloxy.

**Correspondence address:** I. Svendsen, Carlsberg Laboratory, Department of Chemistry, Gamle Carlsberg Vej 10, DK-2500 Copenhagen Valby, Denmark. Fax: (45) (31) 21 13 66.

Table I  
Purification of SGPE

Purification step	Total protein	Enzymatic activity (total units)			Specific activity <sup>a</sup> (U/A <sub>280</sub> )
		Ac-Phe-NH-Ph-NO <sub>2</sub>	Ac-Arg-NH-Ph-NO <sub>2</sub>	Z-Glu-NH-Ph-NO <sub>2</sub>	
Dissolved sample	55	19	57.000	22	0.00039
CM-52 Cellulose	0.45	0	2.900	10	0.022
CABS-Sepharose	0.079	0	3	8.4	1.1
Arginine-Sepharose	0.055	0	0.1	8.4	1.6

<sup>a</sup>specific activity towards Z-Glu-NH-Ph-NO<sub>2</sub>

### 3. RESULTS AND DISCUSSION

The purification of SGPE was followed by activity measurements towards 3 substrates to monitor the separation of SGPE (Z-Glu-NH-Ph-NO<sub>2</sub>) from *S. griseus* trypsin (Ac-Arg-NH-Ph-NO<sub>2</sub>) and *S. griseus* proteases A and B (Ac-Phe-NH-Ph-NO<sub>2</sub>). The ion exchange chromatography on CM-52 cellulose separated SGPE (eluting around 0.13 M NaCl) from proteases A and B (eluting around 0.03 M NaCl) and partially from the trypsin (eluting around 0.11 M NaCl). Further removal of the trypsin was achieved by chromatography on CABS-Sepharose and finally, by affinity chromatography on Arg-Sepharose. The overall purification is 400-fold and the yield is 39% (see Table I). The employed purification scheme represents an improvement relative to that previously described [1]. The isolated enzyme showed a single band by SDS polyacrylamide gel electrophoresis with a mobility corresponding to the previously published values (20 000–22 000) [1].

The amino acid sequence of SGPE was obtained by N-terminal sequencing of the protein and peptides derived by cleavage by CNBr, hydroxylamine, clostripain, chymotrypsin, trypsin and EndoLysC protease as shown in Fig. 2. The sequence from residue 46 to 138 was obtained by DNA sequencing of a 240 bp PCR product. It was ascertained that this was not a part of other known genes from *S. griseus*. The DNA sequencing was done without subcloning directly on the PCR product purified from an agarose gel.

The enzyme consists of a single polypeptide chain of 188 residues, and has a molecular weight of 18 267 calculated from the amino acid sequence. The amino

T A G H C T N I S A N W  
#F2: 5' AC<sub>G</sub><sup>C</sup> GC<sub>G</sub><sup>C</sup> GGC CAC TGC AC<sub>G</sub><sup>C</sup> AAC ATC AGC GC<sub>G</sub><sup>C</sup> AAC TGG 3'

N Y G D G P V Y N M G R  
(5' AAC TAC GC<sub>G</sub><sup>C</sup> GAC GC<sub>G</sub><sup>C</sup> CC<sub>G</sub><sup>C</sup> GT<sub>G</sub><sup>C</sup> TAC AAC ATG GC<sub>G</sub><sup>C</sup> CG 3')  
#R1: 3' TTG ATG CC<sub>G</sub><sup>C</sup> CTG CC<sub>G</sub><sup>C</sup> GC<sub>G</sub><sup>C</sup> CA<sub>G</sub><sup>C</sup> ATG TTG TAC CC<sub>G</sub><sup>C</sup> GC 5'

Fig. 1. Oligonucleotide probes/primers. The mixed oligonucleotides were made so #F2 would hybridize to the non-coding strand and R1 would hybridize to the coding strand, allowing their use for PCR.

acid composition [1] corresponds to that determined from the amino acid sequence.

When due allowances are made for gaps in order to obtain maximal correspondence the amino acid sequence of SGPE was found to be very similar to those of the *S. griseus* proteases A (59% identity) and B (56% identity) [2–4] and  $\alpha$ -lytic proteinase from *Lysobacter enzymogenes* (36%) [5] (see Fig. 3). This suggests that divergent evolution in *S. griseus* have given rise to the 3 proteases (SGPA, SGPB and SGPE) with different specificities. This is the first report of a sequence of a carboxylic acid specific protease being homologous with proteases of other specificities. However, no homo-

```

V L G G G A I Y G G G S R C S A A F N V T K G G A R Y
c- - - - - c- - - - - c- - - - -
N- - - - -

OCC ACC TCC GGC GGT TCC GTC GTC GGA GTG CGG GAG
F V T A G H C T N I S A N W S A S S G G S V V G V R E
c- - - - - c- - - - - c- - - - -
cl-

OCC ACC AAT TTC CCG ACC AAC GAC TAC GAT ATC GTC CAC TAC ACC GAC GGC TCG CCG GGC GGC ACC GTC GAC GTC TAC
G T S F P T N D Y G I V R Y T D G S S P A G T V D L Y
c- - - - - c- - - - -
tr-

AAC GGG TCC ACC CAG GAC ATC TTC TCC GGC ACC AAC GGC GTC GTC GGC CAG GGC ATC AAC AAC AAC GGA TCC ACC ACC AAC
N G S T Q D I S S A A N A V V G Q A I K K S G S T T K
c- - - - -

GTG ACC TCC GGC ACC GTC ACC GGC GTC AAC GTC ACC GTC
V T S G T V T A V N V T V N Y G D G P V Y N M G B T T
cc- - - - -
ch- - - - -
tr- - - - -

A C S A G G D S G G A H F A G S V A L G I H S G S S G
c- - - - - c- - - - - c- - - - -
c- - - - -

C S G T A G S A I H Q P V T K A L S A Y G V T V Y L
c- - - - - tr- - - - - tr- - - - -
cc- - - - -

```

Fig. 2. Amino acid sequence and part of the DNA sequence of SGPE. The N-terminal sequencing of reduced and vinylpyridinated SGPE was performed on: N, Whole protein; c, Chymotrypsin fragments; cl, Clostripain fragments; tr, Trypsin fragments; ng, Hydroxylamine fragments; cn, Cyanogenbromide fragment; and ec, EndoLys-C fragments. C = C-terminal sequencing. The amino acid sequences on which the two probe/primer oligonucleotides (F2 and R1) were based are underlined.

	1									10										20										30				
SGPE		V	L	A	C	G	G	G	A	I	Y	V	-	G	C	G	S	R	C	S	A	L	F	N	V	S	V	N	G	S	A	R	Y	
SGPA		I	A	G	G	E	A	I	T	-	-	-	-	G	C	G	S	R	C	S	A	L	F	N	V	S	V	N	G	S	A	R	Y	
SGPB		I	S	G	G	D	A	I	V	-	-	-	-	G	C	G	S	R	C	S	A	L	F	N	V	S	V	N	G	S	A	R	Y	
$\alpha$ -ly.		A	N	I	V	G	G	I	E	Y	S	I	N	N	A	S	L	C	S	V	G	F	S	V	T	R	G	A	T	K	G	-	-	
SGPE		F	V	T	A	G	H	C	T	N	I	S	A	S	N	W	S	A	S	-	-	-	-	-	-	-	-	-	-	-	-	-	40	
SGPA		A	L	T	A	G	H	C	T	N	I	S	A	S	N	W	S	A	S	-	-	-	-	-	-	-	-	-	-	-	-	-	50	
SGPB		F	L	T	A	G	H	C	T	N	I	S	A	S	N	W	S	A	S	-	-	-	-	-	-	-	-	-	-	-	-	-	60	
$\alpha$ -ly.		F	V	T	A	G	H	C	T	N	I	S	A	S	N	W	S	A	S	-	-	-	-	-	-	-	-	-	-	-	-	-	70	
SGPE		S	F	P	N	N	D	Y	G	I	V	R	Y	T	D	G	S	S	F	A	-	-	-	-	-	-	-	-	-	-	-	-	80	
SGPA		S	F	P	N	N	D	Y	G	I	V	R	Y	T	D	G	S	S	F	A	-	-	-	-	-	-	-	-	-	-	-	-	90	
SGPB		S	F	P	N	N	D	Y	G	I	V	R	Y	T	D	G	S	S	F	A	-	-	-	-	-	-	-	-	-	-	-	-	100	
$\alpha$ -ly.		V	F	P	G	N	D	R	A	W	V	S	L	T	S	A	Q	T	-	-	-	-	-	-	-	-	-	-	-	-	-	-	110	
SGPE		-	-	T	Q	D	I	S	S	A	N	A	F	V	V	G	Q	A	I	K	K	S	G	S	T	T	K	V	T	H	S	G	120	
SGPA		-	-	V	Q	D	I	T	T	A	N	A	F	V	V	G	Q	A	I	K	K	S	G	S	T	T	K	V	T	H	S	G	130	
SGPB		-	-	G	Q	D	I	T	S	A	N	A	F	V	V	G	H	A	V	T	R	K	G	S	T	T	G	T	H	S	G	140		
$\alpha$ -ly.		F	V	T	V	R	G	S	T	E	A	N	A	F	V	G	A	V	T	R	K	G	S	T	T	G	T	H	S	G	150			
SGPE		T	V	T	A	V	N	V	T	V	N	V	G	S	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	160	
SGPA		S	F	P	N	N	D	Y	G	I	V	R	Y	T	D	G	S	S	F	A	-	-	-	-	-	-	-	-	-	-	-	-	170	
SGPB		S	F	P	N	N	D	Y	G	I	V	R	Y	T	D	G	S	S	F	A	-	-	-	-	-	-	-	-	-	-	-	-	180	
$\alpha$ -ly.		T	T	T	A	K	N	V	T	A	N	Y	A	E	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	190	
SGPE		D	S	G	G	A	H	F	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	200	
SGPA		D	S	G	G	S	L	F	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	210	
SGPB		D	S	G	G	P	L	Y	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	220	
$\alpha$ -ly.		D	S	G	G	S	W	I	T	S	A	G	Q	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	230	
SGPE		S	G	T	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	240	
SGPA		S	G	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	250	
SGPB		S	G	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	260	
$\alpha$ -ly.		G	I	P	A	S	Q	R	S	S	L	F	E	R	L	Q	P	I	L	S	G	V	G	L	S	L	V	T	G	-	-	-	-	270

Fig. 3. Comparison of the amino acid sequence of SGPE with proteases A (SGPA) and B (SGPB), from *S. griseus* [2-4] and  $\alpha$ -lytic protease of *Lyobacter enzymogenes* [5]. Residues from SGPE in identical positions with residues in SGPA and SGPB, respectively, are boxed, and gaps are introduced to maximize identity.

logy was found with the carboxylic acid specific protease from *Streptococcus aureus* V8 [11].

The three-dimensional structures of SGPA, SGPB [12] and the  $\alpha$ -lytic protease [13] are known and have been shown to be similar, strongly indicating a similar structure for SGPE. SGPE is a serine protease [1] and based on the sequence homologies (Fig. 3) Ser<sup>152</sup>, His<sup>36</sup> and Asp<sup>66</sup> are probably the catalytically essential amino acid residues. SGPE has 4 Cys residues (No. 17, 37, 146 and 180 in Fig. 3) in positions identical to those in

SGPA and SGPB. In SGPA disulfide bridges are formed from Cys<sup>17</sup> to Cys<sup>37</sup> and Cys<sup>146</sup> to Cys<sup>180</sup>. No attempt has been made to identify the disulfide bridges in SGPE but a chymotryptic digest of intact SGPE contained the 2 peptides Gly<sup>12</sup>-Phe<sup>21</sup> and Val<sup>32</sup>-Trp<sup>44</sup> (in Fig. 3) both of which had a Cys. This finding together with the strong similarity to the other enzymes strongly indicates the disulfide bridges to be identical to those of SGPA and SGPB. As seen in Fig. 3  $\alpha$ -lytic protease has one additional bridge from Cys<sup>109</sup> to Cys<sup>119</sup>.

## REFERENCES

- [1] Yoshida, N., Tsuruyama, S., Nagata, K., Hirayama, K., Noda, K. and Makisumi, S. (1988) *J. Biochem.* 104, 451-456.
- [2] Henderson, G., Krygsman, P., Liu, C.J., Davey, C.C. and Malek, L.T. (1987) *J. Bacteriol.* 169, 3778-3784.
- [3] Johnson P. and Smillie, L.B. (1974) *FEBS Lett.* 47, 1-6.
- [4] Jurasek L., Carpenter, M.R., Smillie, L.B., Gertler, A., Levy, S. and Ericsson, L.H. (1974) *Biochem. Biophys. Res. Commun.* 61, 1095-1100.
- [5] Olson M.O.J., Nagabhushan, M., Dzwiniel, M. and Smillie, L.B. (1970) *Nature* 228, 438-442.
- [6] Breddam, K., Sørensen, S.B. and Ottesen, M. (1983) *Carlsberg Res. Commun.* 48, 217-230.
- [7] Weber, K., Pringle, J.R. and Osborn, M. (1972) *Methods Enzymol.* 26, 3-27.
- [8] Ballance, G.M. and Svendsen, I. (1988) *Carlsberg Res. Commun.* 53, 411-420.
- [9] Hopwood D.A., Bibb, M.J., Chater, K.F., Kieser, T., Bruton, C.J., Kieser, H.M., Lydiate, D.J., Smith, C.P., Ward, J.M. and Schrepf, H. (1985) *Genetic Manipulation of Streptomyces. A Laboratory Manual*, The John Innes Foundation, Norwich, UK.
- [10] Bibb, M.J., Findlay, P.R. and Johnson, M.W. (1984) *Gene* 30, 157-166.
- [11] Drapeau G.R. (1978) *Can. J. Biochem.* 56, 534-544.
- [12] Brayer, G.D., Delbaere, L.T.J. and James, M.N.G. (1978) *J. Mol. Biol.* 124, 216-283.
- [13] Brayer, G.D., Delbaere, L.T.J. and James, M.N.G. (1979) *J. Mol. Biol.* 131, 743-775.