Amino acid sequence determination of guanyl-specific ribonuclease Sa from *Streptomyces aureofaciens*

S.V. Shlyapnikov+, V. Both*, V.A. Kulikov+, A.A. Dementiev+, J. Ševčík and J. Zelinka

Institute of Molecular Biology, Slovak Academy of Sciences, Dúbravská cesta 21, 842 51 Bratislava, Czechoslovakia and +*Institute of Molecular Biology of the USSR Academy of Sciences, Vavilova 32, Moscow, USSR*

Received 21 October 1986

Using automated Edman degradation of two nonfractionated peptide mixtures of tryptic and staphylococcal protease digests of the protein, the complete amino acid sequence of the guanyl-specific ribonuclease Sa from *Streptomyces aureofaciens* was established. Ribonuclease Sa contains 96 amino acid residues (M_r 10 566). A 50% sequence homology of ribonuclease Sa to the guanyl-specific ribonuclease St from *S. erythreus* was found.

RNase (Streptomyces aureofaciens) *Amino acid sequence*

1. INTRODUCTION

RNase Sa (EC 3.1.4.8) isolated from *Streptomyces aureofaciens* is a guanyl-specific endoribonuclease. The enzyme catalyses the splitting of RNA only at sites where a guanosine nucleotide residue is located in the 3′-position of the 3′–5′-phosphodiester bond of RNA [1]. Its amino acid composition [2] differs considerably from that of RNase T_1, which is a well-known representative of guanyl-specific RNases.

Investigations on the RNase Sa active site by chemical modifications and kinetic measurements [3,4] and NMR spectroscopy [5] showed some similarities but mainly distinctions from the active-site functional topochemistry of other guanyl-specific RNases.

In searching for the relationship between functional homology of enzymes and the homology of their molecular structure the primary structures of many guanyl-specific RNases have been determined lately. To contribute to the comparative study of primary structures of these enzymes we established the complete amino acid sequence of RNase Sa using an express method successfully applied to two fungal RNase [6,7]. The sequence of 30 N-terminal residues of RNase Sa has been determined previously [8].

2. MATERIALS AND METHODS

RNase Sa was isolated and purified to a chromatographically pure state according to [9] slightly modified by applying rechromatography on phosphocellulose. Reduction and *S*-carboxymethylation of the enzyme were performed as in [10].

Automated Edman degradations were carried out on a Beckman 890C sequencer according to the 0.1 M Quadrol program in the presence of polybrene and Gly-Gly [11,12]. Proline N-terminal residues were split twice in a course of sequential protein analysis at 12 and 13 cycles. Reagents and solvents for Edman degradation were from Beckman.

PTH-amino acid derivatives were identified by three methods: gas-liquid and thin-layer chromatography [13,14] as well as by amino acid analysis after hydrolysis of the derivatives with 5 N HI at

* To whom correspondence should be addressed

Published by Elsevier Science Publishers B.V. (Biomedical Division)
00145793/86/$3.50

150°C for 24 h. Edman degradation products were converted in 1 N HCl [15].

Digestions with staphylococcal protease and trypsin were performed in the reaction cell of the sequencer on reduced *S*-carboxymethylated RNase (S-Cm-RNase) preparations preliminarily subjected to 21 complete cycles of automated Edman degradations (des-(1–21)-RNase). Des-(1–21)-RNase was modified with fluorescamine [16] and after performing one additional cycle of Edman degradation it was cleaved with staphylococcal protease or trypsin in 0.1 M ethylmorpholine acetate buffer, pH 8.1, at 37°C for 8 or 4 h and at an enzyme/substrate ratio of 1:30 or 1:50, respectively.

The C-terminal sequence of RNase was followed by digestion of the protein with carboxypeptidase Y in 50 mM ethylmorpholine acetate, pH 5.5, and at an enzyme/substrate ratio of 1:50.

3. RESULTS AND DISCUSSION

The present method of determining the primary structure of RNase Sa is based on using the technique of automated Edman polypeptide sequencing and provided mainly two types of results. The first type is the results from amino acid sequencing of the whole carboxymethylated protein (fig.1). In this case (repetitive yield of Edman degradation is about 97–98%) we can estimate the position of the 73 N-terminal amino acid residues of the protein (fig.2).

The second kind is the results of amino acid sequencing of the nonfractionated peptide mixture produced by tryptic digests of S-Cm-RNase preliminarily subjected to 21 complete steps of automatic Edman degradation and treated thereafter with fluorescamine (table 1). This latter modification made it possible to restrict the general quantity of analysing peptides produced by enzymatic cleavage and allowed reduction of the background of sequencing.

The data presented in table 1 detail the sequencing of five or six peptides (see also fig.2); moreover two pairs of peptides presented duplicate copies because the polypeptide chain of RNase contains two adjacent Arg residues (Arg^{68} and Arg^{69}). According to the specific manner of the tryptic digestion, incomplete cleavage of the RNase polypeptide chain may result in the formation of two different peptide variants. The positions of all Arg residues of the polypeptide chain were determined in the course of protein sequencing (fig.1). At the same time the determination of the complete amino acid sequence of the protein chain of RNase containing the Arg^{40}, Arg^{63} and Arg^{65} positions as well as the partial recognition of the amino acid sequence containing the Arg^{40} and Arg^{69} positions (fig.1) makes it possible to continue the identification of the C-terminal protein structure (table 1, fig.2).

Some additional information which supports the above results arose from the consideration of the kinetics of RNase hydrolysis with carboxypeptidase Y (fig.2) and from the products generated in Edman degradation of the fluorescamine-modified des(1–21)-RNase peptide hydrolysate produced by staphylococcal protease (table 1).

Four peptides undergo Edman degradation in the original mixture of the staphylococcal protease des-(1–21)-RNase hydrolysate. The structure of the RNase fragment 42–73 was deduced earlier (fig.1). Combined analysis of the amino acid sequences of all peptides of RNase digestion with staphylococcal protease gave results corresponding completely with the data from tryptic peptide sequencing. These results were sufficient for unambiguous establishment of the protein primary structure.

Thus, RNase Sa contains 96 amino acid residues: Asp_7, Asn_2, Thr_{12}, Ser_7, Glu_5, Gln_5, Pro_6, Gly_8, Ala_6, Cys_3, Val_9, Ile_5, Leu_6, Tyr_8, Phe_3, His_2 and Arg_5 (M_r 10566).

The amino acid sequence of RNase Sa differs considerably from that of the fungal guanyl-specific RNases completely sequenced so far ([6,7] and see references in [6]). On the other hand, the primary structure of RNase Sa shows a very close

Fig.2. (1–3) Structures determined in sequencing of *S*-carboxymethylated protein and products of proteolysis of des-(1–21)-RNase with trypsin and staphylococcal protease, respectively. (4) Data from studying the kinetics of RNase hydrolysis with carboxypeptidase Y. The sites of protein cleavage with staphylococcal protease (Sp) and trypsin (Tr) are indicated by vertical arrows.

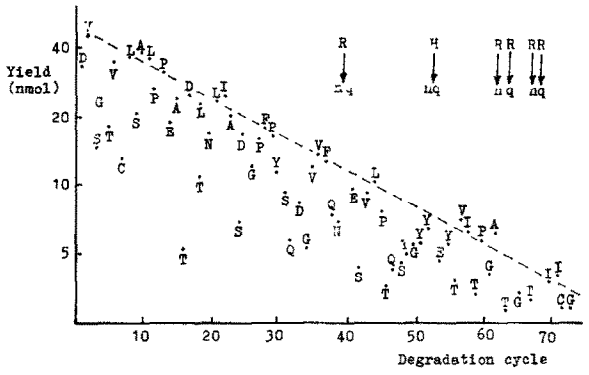

Fig.1. Automated Edman degradation of RNase Sa. The one-letter code for amino acids is used. Cysteine was determined as the carboxymethylated derivative. nq, not quantified.

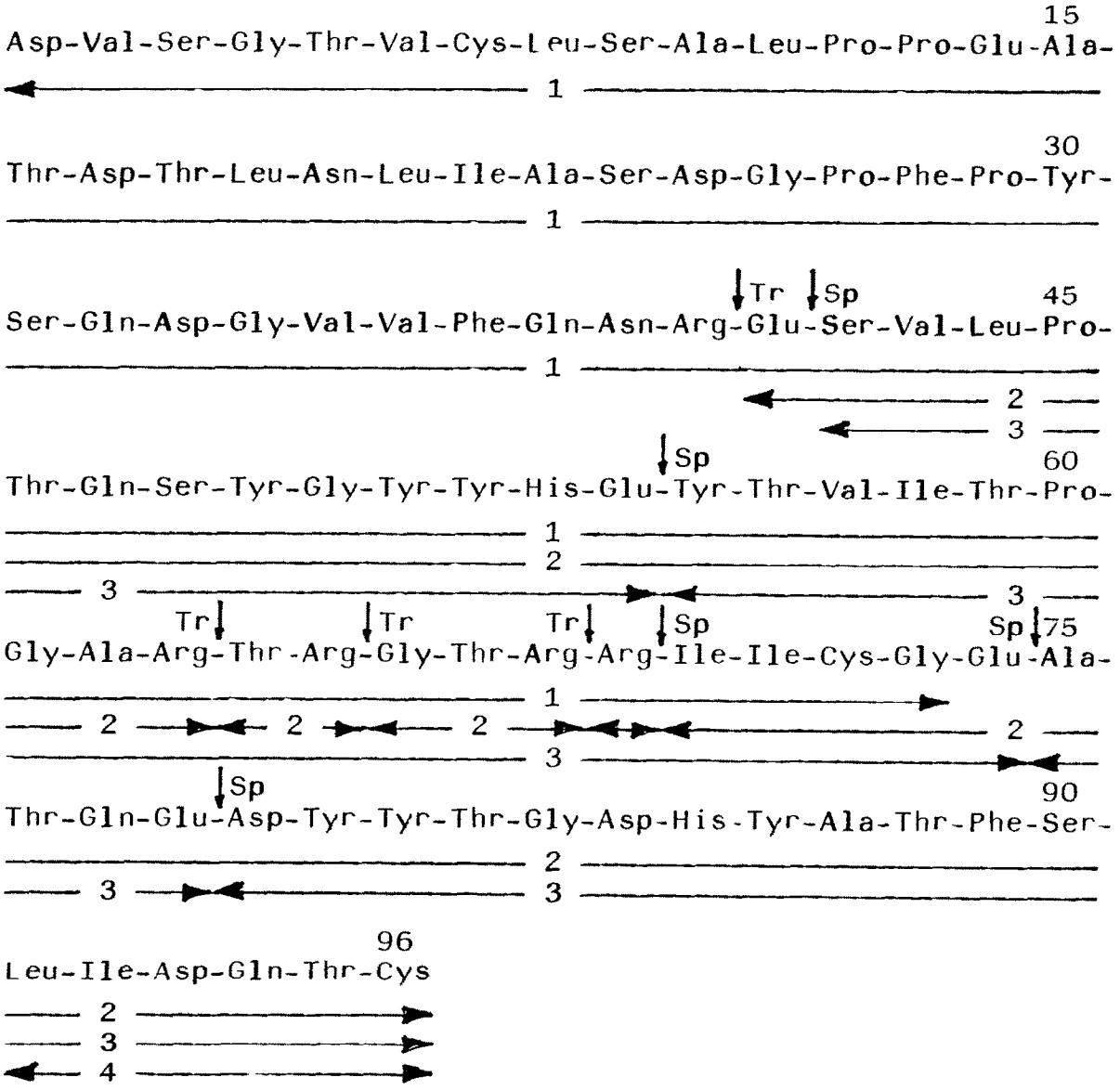

Table 1

Automated Edman degradation of fluorescamine-modified des-(1–21)-S-Cm-RNase Sa peptide hydrolysates

Degradation cycle	Yield of amino acids (nmol)								
	Tryptic hydrolysate of des-(1–21)-RNase					Hydrolysate of des-(1–21)-RNase by staphylococcal protease			
1	E,25	T,33	G,42	R,nq	I,49	S,19	Y,22	A,41	D,28
2	S,14	R,nq	T,36	I,62[a]	I,62[a]	V,34	T,35[a]	T,35[a]	Y,29
3	V,33		R,nq	I,10	C,25	L,31	V,29	Q,17	Y,32
4	L,35		(R)nq	C,5	C,28	P,26	I,25	E,9	T,21
5	P,29			G,6	E,26	T,27[a]	T,27		G,25
6	T,15			E,5	A,35	Q,15	P,21		D,21
7	Q,19			A,7	T,20	S,11	G,15		H,nq
8	S,9			T,4	Q,25	Y,27[a]	A,18		Y,27[a]
9	Y,19			Q,31[a]	E,31[a]	G,16	R,nq		A,27
10	G,12			E,5	D,23	Y,14	T,23[a]		T,23[a]
11	Y,38[a]			D,5	Y,38[a]	Y,16	R,nq		F,21
12	Y,45[a]			Y,45[a]	Y,45[a]	H,nq	G,8		S,9
13	H,nq			Y,8	T,16	E,6	T,5		L,18
14	Q,11			T,3	G,17		R,nq		I,19
15	Y,12			G,3	D,19		R,nq		D,9
16	T,5			D,4	H,nq		I,11		Q,8
17	V,14			H,nq	Y,14		I,13		T,4
18	I,10			Y,3	A,21		C,12[a]		C,12[a]
19	T,13[a]			A,4	T,13[a]		G,4		
20	P,4			nd	F,17		E,3		
21	G,3			F,3	S,8				
22	A,5			nd	L,13				
23	nd			L,3	I,15				
24				I,3	D,11				
25				nd	Q,8				
26				nd	T,5				
27				nd	C,6				
Position at the RNase chain	41–63	64–65	66–68/69	69–96	70–96	42–54	55–74	75–78	79–96

[a] The yields of amino acids were not differentiated for the individual peptides

nd, not determined (other abbreviations as in fig.1)

sequence homology to guanyl-specific RNase St from *S. erythreus* [17]. The sequences of these RNases contain 50 identical amino acid residues.

REFERENCES

[1] Zelinková, E., Bačová, M. and Zelinka, J. (1971) Biochim. Biophys. Acta 235, 343–352.

[2] Ševčík, J., Gašperík, J. and Zelinka, J. (1982) Gen. Physiol. Biophys. 1, 255–259.

[3] Both, V., Witzel, H.H. and Zelinková, E. (1982) Gen. Physiol. Biophys. 1, 261–268.

[4] Kéry, V., Both, V., Ševčík, J. and Zelinka, J. (1986) Gen. Physiol. Biophys. 5, 405–414.

[5] Both, V., Zachar, J. and Zelinka, J. (1983) Gen. Physiol. Biophys. 2, 269–278.

[6] Shlyapnikov, S.V., Kulikov, V.A. and Yakovlev, G.I. (1984) FEBS Lett. 177, 246–248.

[7] Shlyapnikov, S.V., Bezborodova, S.I., Kulikov, V.A. and Yakovlev, G.I. (1986) FEBS Lett. 196, 29–33.

[8] Gašperík, J., Murányová, M., Šuranová, B., Zelinka, J. and Kostka, V. (1982) Biológia, 37, 769–775.
[9] Bačová, M., Zelinková, E. and Zelinka, J. (1971) Biochim. Biophys. Acta 235, 335–342.
[10] Crestfield, A.M., Moore, S. and Stein, W.H. (1963) J. Biol. Chem. 238, 622–634.
[11] Klapper, D.G., Wilde, C.E. iii and Capra, J.D. (1978) Anal. Biochem. 85, 126–131.
[12] Kulikov, V.A., Shlyapnikov, S.V. and Yakovlev, G.I. (1985) Bioorg. Khim. 11, 334–344.
[13] Pisano, J.J., Bronzert, T.J. and Brewer, H.B. (1972) Anal. Biochem. 45, 43–59.
[14] Alakhov, Yu.B., Motuz, L.P., Stengrevicz, O.A. and Vinokurov, L.M. (1982) Bioorg. Khim. 4, 1301–1313.
[15] Edman, P. and Begg, G. (1967) Eur. J. Biochem. 1, 80–91.
[16] Bhown, A.S., Bennett, J.C., Morgan, P.H. and Mole, J.E. (1981) Anal. Biochem. 112, 158–162.
[17] Yoshida, N., Sasaki, A., Rashid, M.A. and Otsuka, H. (1976) FEBS Lett. 64, 122–125.