Structural basis for the catalytic mechanism and substrate specificity of the ribonuclease α-sarcin

Ramón Campos-Olivas[a], Marta Bruix[a], Jorge Santoro[a], Alvaro Martínez del Pozo[b], Javier Lacadena[b], José G. Gavilanes[b], Manuel Rico[a,*]

[a] *Instituto de Estructura de la Materia, Consejo Superior de Investigaciones Científicas, Serrano 119, 28006 Madrid, Spain*
[b] *Departamento de Bioquímica y Biología Molecular I, Facultad de Química, Universidad Complutense, 28040 Madrid, Spain*

Received 23 August 1996; revised version received 21 October 1996

Abstract **α-Sarcin is a ribosome-inactivating protein which selectively cleaves a single phosphodiester bond in a universally conserved sequence of the major rRNA. The solution structure of α-sarcin has been determined on the basis of 1898 distance and angular experimental constraints from NMR spectroscopy. It reveals a catalytic mechanism analogous to that of the T_1 family of ribonucleases while its exquisite specificity resides in the contacts provided by its distinctive loops.**

Key words: Ribosome inactivating protein; α-Sarcin; Three-dimensional structure; NMR; Ribonuclease specificity

1. Introduction

α-Sarcin (αS) is a 150-residue ribosome-inactivating protein secreted by the mold *Aspergillus giganteus*. It selectively cleaves a single phosphodiester bond in a universally conserved sequence of ~14 nucleotides in the major rRNA [1] which is implicated in translocation during protein biosynthesis [2]. Together with restrictocin and mitogillin, also secreted by aspergilli, and with which it shares ~85% sequence identity, αS constitutes a distinctive family of fungal ribonucleases, related to the T_1 family of ribonucleases (~100 residues), but with differential characteristics, namely, its exquisite substrate specificity and ability to interact with and translocate across acid phospholipid bilayers [3]. Therefore, elucidation of the three-dimensional (3D) structure of αS has been pursued, without success [1,4], from its sequence determination in 1983 [5]. We have recently assigned the ^{1}H- and ^{15}N-NMR spectra of the protein and identified the elements of secondary structure [6], and here report the 3D structure in solution of αS by NMR methods, as well as an analysis of the structural basis of its ribonucleolytic activity.

2. Materials and methods

The program DIANA [7] was used for obtaining conformers compatible with the NMR-derived structural constraints. Initially, only unambiguous distance constraints from a NOESY spectrum in D_2O at 50 ms mixing time, together with restraints for 25 H-bonds for which clear evidence existed and those for the two disulphide bridges were introduced. 50 conformers were obtained which were used to further assign NOEs from the same spectrum and from a NOESY spectrum in H_2O at 150 ms mixing time using the program ASNO [8]. φ angular restrictions from $^3J_{HNH\alpha}$ measurements were introduced at this stage and again the ASNO strategy was iteratively applied. The final set of NMR constraints consists of 491 intraresidual and sequential, 300 short-range (i–j < 5), and 996 long-range upper distance limits, 56 lower limit distance constraints, and 45 φ angular restraints. The best 20 conformers of the final 50 calculated are used for analysis of the structure of αS in solution. The average value for the sum of violations of the 20 model conformers is 7.5 Å, 0.5 Å, 8.9 Å, and 8.3° for upper limits, lower limits, van der Waals limits, and dihedral angle constraints, respectively, and the corresponding average value for the maximum violation is 0.40 Å, 0.12 Å, 0.26 Å, and 3.6°.

3. Results

A backbone superposition of 20 structures calculated on the basis of the NMR data is shown in Fig. 1a. β-strands are shown in dark blue, the α-helix in red, and loops connecting them in magenta (turn1), yellow (loop1), green (loop2), blue (loop3), pink (turn2), and orange (turn3). N- and C-terminal residues are coloured in light green and red, respectively. If all residues are used in the superposition the average pairwise root mean square deviation (rmsd) is 2.86 Å for the backbone heavy atoms, and 3.66 Å for all heavy atoms. However, when considering exclusively the well-defined residues (global average pairwise backbone rmsd < 2 Å in the complete superposition), the corresponding values decrease to 1.24 and 1.88 Å, respectively. The well-defined residues are 105: 1–12, 17–44, 48–63, 67–72, 75–82, 92–98, 120–138, and 142–150, which amount to 70% of the total. Therefore, residues implicated in secondary structure and composing the core of the molecule have been delineated with satisfactory precision while some others forming part of the long loops are not so well-defined. Fig. 1b shows a schematic representation of the structure which highlights its secondary structure elements. The first two β-strands define an N-terminal β-hairpin comprising residues 3–12 (β_1) and 17–26 (β_2). Residues 27–36 form a short α-helix with scarcely three turns. The C-terminal major β-sheet of αS is composed of 5 short strands, β_3 (49–52), β_4 (93–98), β_5 (120–126), β_6 (131–137), and β_7 (145–147), arranged in a +1,+1,+1,+1 topology; the rest constitute the numerous and long loops of the protein structure (turn1, 13–16; loop1, 37–48; loop2, 53–92; loop3, 99–119; turn2, 127–130; turn3, 138–144). The α-helix packs nearly orthogonally against strands β_3, β_4, β_5, and β_6 in one face of the major sheet. The last strand (β_7) is linked to β_1 by the disulfide bond 148–6, bringing together the N- and C-termini of the protein and connecting the two sheets so that the terminal residues in the β-hairpin pack against β_6 and β_7 in the same face of the major sheet. The other face of the major β-sheet is more exposed and is surrounded by the long loops which project out toward the solvent, defining a depression in the structure where the catalytic residues are located (see below). This

*Corresponding author. Fax: (34) (1) 5642431.
E-mail: emcampos@roca.csic.es

PII S0014-5793(96)01320-8

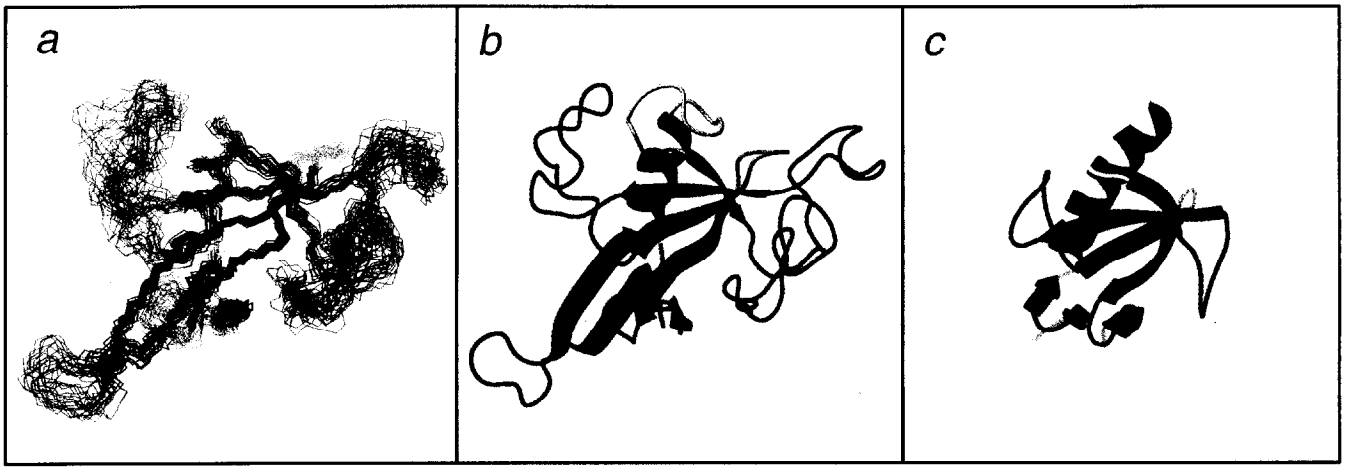

Fig. 1. (a) Backbone superposition of the 20 conformers used for the analysis of the three-dimensional structure of αS in solution. (b) Ribbon plot of the three-dimensional structure of αS (one of the 20 conformers) and (c) ribonuclease T_1 (pdb entry 1rga).

structure already allows several structure-function relationships to be established.

β_3	S47 Y48 P49 H50 W51 F52 T53 N54 G55 Y56 D57 G58	αS
	S37 Y38 P39 H40 K41 Y42 N43 N44 Y45 R46 G47	T_1
β_4	Y93 L94 L95 E96 F97 P98 T99	αS
	P55 Y56 Y57 E58 W59 P60 I61	T_1
β_5	P117 G118 P119 A120 R121 V122 I123 Y124 T125 Y126	αS
	P73 G74 A75 D76 R77 V78 V79 F80 N81 N82	T_1
β_6	K129 V130 F131 C132 G133 I134 I135 A136 H137 T138	αS
	N84 Q85 L86 A87 G88 V89 I90 T91 H92 T93	T_1

Fig. 2. Sequence alignment of the residues composing the strands β_3, β_4, β_5, β_6 in αS and ribonuclease T_1 [10], where the catalytic residues in T_1 are located.

4. Discussion

Comparison of the structure of αS with that of the ribonuclease T_1 [9,10] reveals an identical architecture in the nature and topology of connection of the secondary structure elements in the two proteins (Fig. 1b,c). The main differences are located in the length of the N-terminal β-hairpin (left-bottom) and the loops connecting the secondary structure elements; in αS the loops are generally much larger and exposed than in T_1 (i.e. number of residues in N-terminal hairpin, 24 vs. 11; in loop1 (yellow), 12 vs. 10; in loop2 (green), 40 vs. 13; and in loop3 (cyan), 21 vs. 12). In contrast, the α-

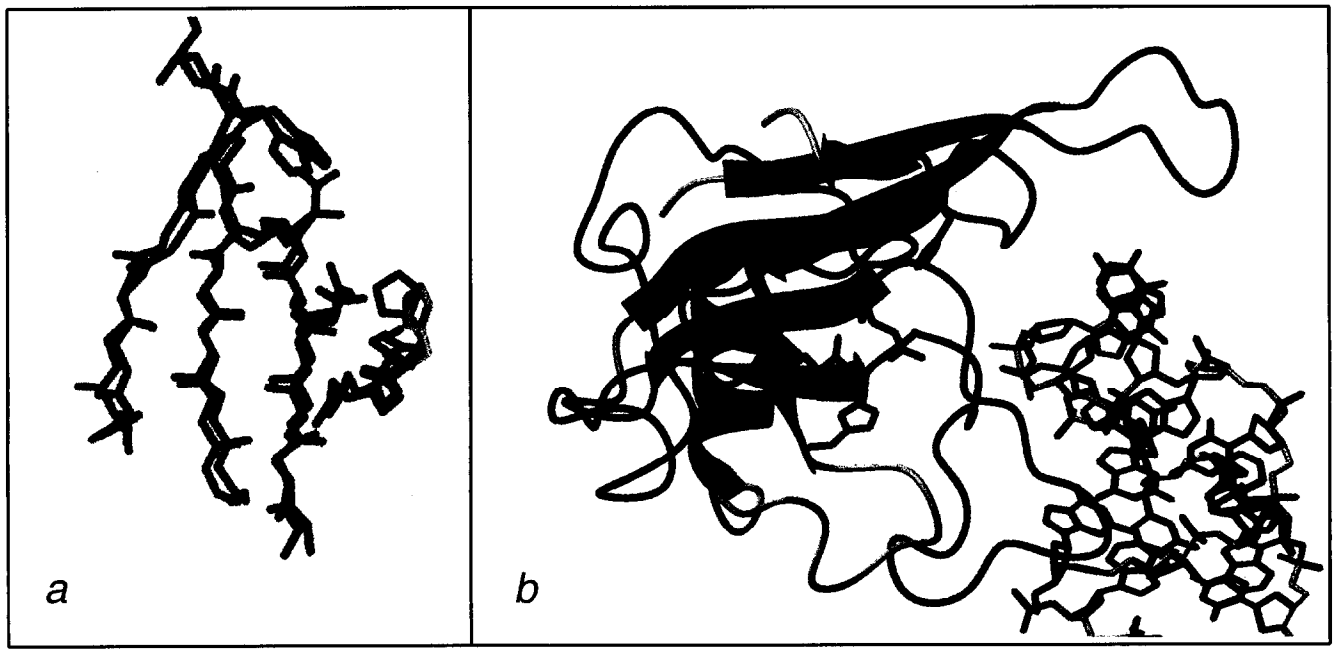

Fig. 3. (a) Detail of the active site residues in T_1 and the corresponding ones in αS after superposition of the backbone atoms of the residues in strands β_3, β_4, β_5, β_6 (from right to left). αS is shown in yellow and T_1 in magenta. The residues shown are, from left-top to right-bottom, H137 and H92 (dark blue), R121 and R77 (red), E96 and E58 (orange), and H50 and H40 (grey), for αS and T_1, respectively. (b) Model for the interaction of αS (one conformer, left) with its RNA substrate (one model from pdb entry 1scl, right). The two molecules have been separated for the sake of clarity and only the catalytic residues in αS and the crucial base for recognition (G10, red) as well as the scissile bond (yellow) between the two nucleotides (G16, pink; A17, pale green) in the RNA have been highlighted. In the RNA the backbone is shown in white and the bases in purple.

helix (red) in αS is one turn shorter than in T_1. Superposition of the backbone heavy atoms constituting the regions of secondary structure in both proteins (44 residues in total) reveals a close similarity of the two structures (backbone rmsd=1.64 Å). This superposition allows the alignment of the sequences of αS and T_1 on the basis of their 3D structures, as shown in Fig. 2. The catalytic residues (in bold) of T_1 are strictly conserved in equivalent positions in αS and other residues involved in recognition of the base in mononucleotide-T_1 complexes (in italics) and constituting the β-strands (underlined) are also highly conserved in the two proteins. More remarkably, the geometry of the four β-strands where the catalytic residues in T_1 are located, as well as the nature and orientation of the catalytic residues themselves, is strictly conserved in αS (Fig. 3a). This is a clear indication that the catalytic mechanism of RNA hydrolysis by the α-sarcin family of ribonucleases is very similar to that performed by the ribonucleases of the T_1 family. Therefore, the mode in which αS recognizes the phosphate group of the scissile bond in its target RNA should be analogous to that found in the complex of ribonuclease T_1 and other family members with 3′GMP [9]. Consequently, the binding geometry found in T_1 was imposed to direct the docking of the structure of the minimal RNA oligonucleotide recognized by αS [11] with that of the protein. In the resulting model for the complex (Fig. 3b), which is exclusively inferred from the spatial matching of the catalytic sidechains Glu-96, Arg-121, and His-137 in the protein with atoms G16 C2′, P, and A17 O5′ in the RNA, respectively, the protein embraces the single-stranded turn of the RNA structure revealing a large number of potential contacts between the RNA and the long and extended loops of the protein. When the two molecules approach the interaction distances for the above residue-atom pairs, loop3 (cyan) and loop1 (yellow) contact a base (G10, red) which is six residues away from the scissile bond (yellow) and was previously identified as the crucial residue for the specific recognition [12] whereas specific residues in loop2 (green) may recognize the base of the nucleotide (A16, pink) at 3′ to the scissile phosphodiester bond. The role of loop3 in conferring the substrate specificity is further supported by deletion mutants of mitogillin which lack part of it (K106–K113) and only retain nonspecific ribonucleolytic activity [13]. Obviously, conformational changes of both the protein and its RNA substrate cannot be ruled out, although the basic geometry of the interaction is very likely that reflected by our model for the complex. Consequently, the exquisite specificity of αS to recognize and cut its target RNA lies in the contacts provided by the loops connecting its secondary structure elements, which constitute the major difference between αS and the members of the T_1 family of ribonucleases.

Finally, we should like to add that, during the preparation of this communication, a paper appeared reporting the crystal structure of restrictocin [14]. The results obtained were in good general agreement with our conclusions, although differences in the structure of the loops should be carefully examined once the coordinates of restrictocin become available from the Brookhaven Protein Data Bank

Acknowledgements: This work was supported by the Dirección General de Investigación Científica y Técnica (Spain) (PB93-0189). R.C.O. was recipient of a predoctoral grant (FP92-02882770) from the Ministerio de Educación y Ciencia (Spain).

References

[1] Wool, I.G. (1984) Trends Biochem. Sci. 9, 14–17.
[2] Wool, I.G., Glück, A. and Endo, Y. (1992) Trends Biochem. Sci. 17, 266–269.
[3] Gasset, M., Mancheño, J.M., Lacadena, J., Turnay, J., Olmo, N., Lizarbe, M.A., Martínez del Pozo, A., Oñaderra, M. and Gavilanes J.G. (1994) Curr. Top. Pept. Protein Res. 1, 99–104.
[4] Martínez, S.E. and Smith, J.L. (1991) J. Mol. Biol. 218, 489–492.
[5] Sacco, G., Drickamer, K. and Wool I.G. (1983) J. Biol. Chem. 258, 5811–5818.
[6] Campos-Olivas, R., Bruix, M., Santoro, J., Martínez del Pozo, A., Lacadena, J., Gavilanes, J.G. and Rico, M. (1996) Protein Sci. 5, 969–972.
[7] Güntert, P., Braun, W. and Wüthrich, K. (1991) J. Mol. Biol. 217, 517–530.
[8] Güntert, P., Berndt, K. and Wüthrich, K. (1993) J. Biomol. NMR 3, 601–606.
[9] Pace, C.N., Heinemann, U., Hahn, U. and Saenger, W. (1991) Angew. Chem. Int. Ed. Engl. 30, 343–360.
[10] Hoffmann, E. and Rüterjans, H. (1988) Eur. J. Biochem. 117, 539–560.
[11] Szewczak, A.A. and Moore, P.B. (1995) J. Mol. Biol. 247, 81–98.
[12] Glück, A. and Wool, I.G. (1996) J. Mol. Biol. 256, 838–848.
[13] Kao, R. and Davies, J. (1995) Biochem. Cell Biol. 73, 1151–1159.
[14] Yang, X. and Moffat, K. (1996) Structure 4, 837–852.