Volume 333, number 3, 257–260 FEBS 13177 November 1993

The three-dimensional structure of the colicin E3 immunity protein by distance geometry calculation

Shunsuke Yajima[a], Yutaka Muto[b], Souichi Morikawa[c], Haruki Nakamura[c], Shigeyuki Yokoyama[b], Haruhiko Masaki[a,*], Takeshi Uozumi[a]

[a]*Department of Biotechnology, The University of Tokyo, Yayoi 1-1-1, Bunkyo-ku, Tokyo 113, Japan*
[b]*Department of Biophysics and Biochemistry, Graduate School of Science, The University of Tokyo, Bunkyo-ku, Tokyo 113, Japan*
[c]*Protein Engineering Research Institute, Furuedai, Suita, Osaka 565, Japan*

Received 2 September 1993

The three-dimensional solution structure of the colicin E3 immunity protein (84 residues) was determined by distance geometry calculations. The hydrophilic side of a four-stranded antiparallel β-sheet constitutes a part of the surface of the protein, and two loops lie on the hydrophobic side of the sheet. All the three specificity-determining residues, which are included in the center of the β-sheet, display their side groups on the protein surface.

Nuclear magnetic resonance; Colicin E3; Distance geometry; Protein–protein interaction

1. INTRODUCTION

Colicin E3 is a type of bacteriocin and can kill sensitive *Escherichia coli* cells [1]. E3 is a special RNase which cuts 16 S ribosomal RNA at the specific site in the 70 S ribosome. This activity is exclusively located on the C-terminal domain, T2A (97 residues) [2–5]. The ColE3 plasmid encodes colicin E3 and also its inhibitor protein (ImmE3, 84 residues) which protects the host cell from both endogenous and exogenous colicin action, a phenomenon referred to as immunity, by binding to the T2A domain [6–8].

Colicin E6 and cloacin DF13, encoded by plasmids ColE6 and CloDF13, respectively, are colicin E3 homologues, and have been demonstrated to show the same killing activity as that of E3 ([9]; unpublished data). Although DF13 does not kill standard *E. coli* strains in vivo due to the difference in the cell surface receptors, it has the same immunity as E6, but not E3. The differences in amino acid sequences between E3 and E6 types are observed almost only in T2A's and Imm's; 10 residues in T2A's and 12 residues in Imm's, but immunity specificities are strictly defined; ImmE3 inhibits only the E3-T2A activity, and both ImmE6 and ImmDF13 inhibit E6-T2A and DF13-T2A [10]. This implies that a small number of residues play an important role in the recognition of their cognate proteins. In fact, our genetic studies revealed that the ImmE3 and ImmE6 specificities are dominantly determined by two residues in each molecule [11].

We believe that T2A and Imm, comprising a small heterodimer with only a few residues of specificity determinants, represent a promising model for both genetical and physicochemical studies of protein–protein interactions.

Three-dimensional structures have not been known for the nuclease-type colicins nor their Imm proteins. We recently revealed the secondary structure of ImmE3, which suggests that the β-sheet, including the specificity determinant recognizing T2A. [12]. In this paper, we report the three-dimensional structure of the ImmE3 protein in solution determined by distance geometry and simulated annealing calculations on the basis of isotope-aided NMR data.

2. MATERIALS AND METHODS

A full description of the experimental procedures for NMR measurements and assignments were given previously [12]. Distance geometry calculation was performed using program EMBOSS ver. 4.1 [13], and final refinement of the structures was done with the program, PRESTO [14].

NOEs were classified on the basis of the 80 ms NOESY spectra recorded at 30 and 17°C into strong, medium and weak, corresponding to inter-proton distance upper restraints of 3.0, 4.0 and 5.0 Å, respectively. All lower limits were set to the sum of the van der Waals' radii. The upper limits of the distance restraints involving methylene, methyl and aromatic protons were corrected according to the method of Wüthrich [15]. Backbone ϕ torsion angle restraints were derived from the HMQC-J spectrum. $^3J_{HN\alpha}$ coupling constants over 9 Hz were converted into the range $-160° \leq \phi \leq -80°$, and those less than 4.5

*Corresponding author. Fax: (81) (3) 5684-0387.

Abbreviations: NMR, nuclear magnetic resonance; NOE, nuclear Overhauser effect.

Hz were into the range $-90° \leq \phi \leq -40°$ [15]. All calculations were performed on a Fujitsu VP2600 and an IRIS Indigo.

3. RESULTS

A total of 733 NOEs, which are composed of 172 intra-residual, 336 short-range NOEs ($1 < |i-j| \leq 5$), and 259 long-range NOEs ($5 < |i-j| \leq$), were collected from the NOESY spectra. 22 distance restraints for hydrogen bonds [12] were also included for calculation. Stereo-specific assignments were obtained for the methyl groups of five valines (V31, V37, V50, V55 and V75) and the β-methylene protons of two residues (F25 and N71). Furthermore, 29 distance restraints from ϕ torsion angles were constructed.

Initially, based on these 784 experimental restraints, the distance geometry calculation was performed with the program, EMBOSS. Random coil conformations of the initial coordinates were optimized. Out of 70 calculated structures, 22 with total distance violations less than 4.0 Å were selected.

The final refinement was carried out by PRESTO. In PRESTO, the structure was finally improved by 1,000 steps of energy minimization with the distance and torsion angle restraints, using the AMBER all-atom force field [16] with the 6–12 Lennard-Jones, hydrogen bonding and electrostatic potentials instead of the simple repulsive force. Among the 22 refined structures, we selected 10 structures which have very small violations. The RMSD value of 10 structures was 1.76 ± 0.18 Å for the backbone atoms of residues 3–78, and, in particular, 0.94 ± 0.16 Å for those of residues 3–9, 19–22, 44–49, and 73–78 which make up a β-sheet (Fig. 1). Fig. 2 shows the number of distance restraints (Fig. 2A) and the average RMSD of the backbone atoms (N, Cα, C) among the 10 structures calculated above (Fig. 2B) at each residue. The residues in the β-sheet showed lower RMSD values. The statistics of the final structure are shown in Tables I and II.

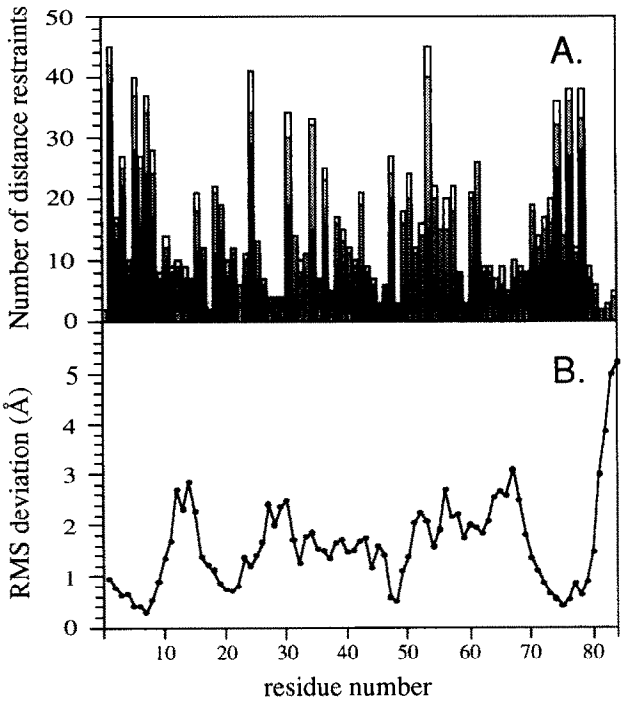

Fig. 2. Plots of the number of the distance restraints, and average RMSD from the final 10 structures after PRESTO at each residue. (A) Solid, dotted and open bars represent long-range ($|i-j| > 5$), short-range ($1 < |i-j| \leq 5$), and intra-residue restraints, respectively. (B) Average RMSD of the backbone atoms (N, Cα, C) between the 10 structures calculated by PRESTO.

4. DISCUSSION

4.1. *Structural feature of the ImmE3 protein*

Fig. 3 shows the backbone conformation of the ImmE3 protein after the final refinement. ImmE3 has a four stranded antiparallel β-sheet as a main motif. One side of the β-sheet forms a part of the hydrophilic surface of the protein and the other hydrophobic side is involved in formation of the core of the protein. The

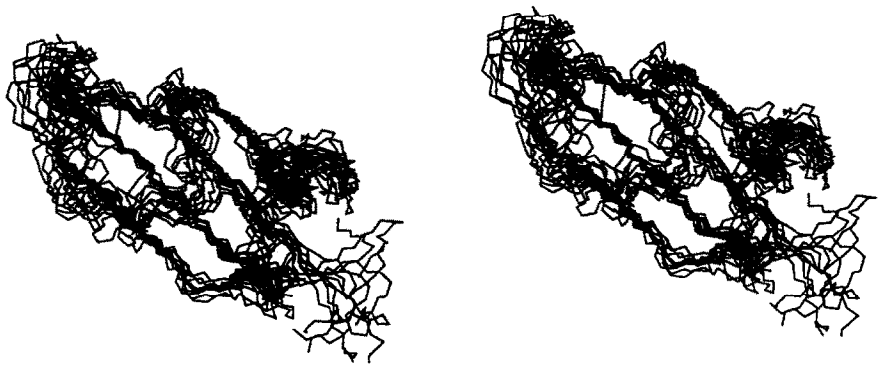

Fig. 1. Stereoview of the 10 refined structures of ImmE3. The structures were superimposed to minimize the RMSD value of the backbone atoms (N, Cα, C) of residues 3–78.

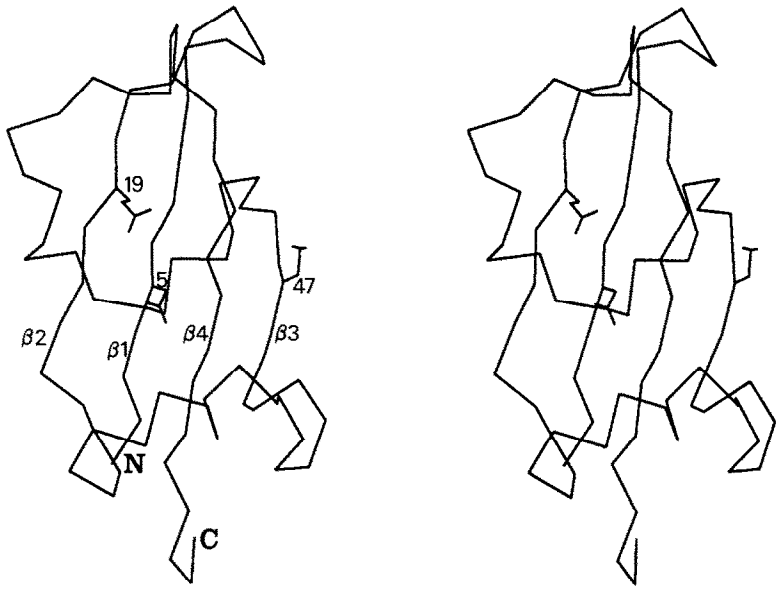

Fig. 3. The Cα-atom representation of the mean structure of ImmE3, which was minimized with restraints. Several side chains at 5, 19, 47, which are the specificity determinants, are also presented.

β-strands are composed of β1 (residue 2–10), β2 (residue 19–23), β3 (residue 44–49) and β4 (residue 73–79). The secondary structure is arranged as follows; β1 and β2 are connected by a short loop, l1; β2 and β3 are connected by a long loop, l2; and β3 and β4 are connected by a longer loop, l3. A short α-helix (residue 31–36) are in the center of l2, and no other conspicuous secondary structure is observed. Both l2 and l3 are located on the hydrophobic surface of the β-sheet. The hydrophobic core is mainly made up of F25, V31, L35 and V37 in the l2; W54, L57, L58, Y61 and F62 in the l3; and of L4, L6, F48, V50, V75 and F77 in the β-sheet.

4.2. *Specificity determination in ImmE3*

Although ImmE3, ImmE6 and ImmDF13 are highly homologous proteins, ImmE3 shows a distinct immunity specificity from the other two, which share the ImmE6 immunity [10]. Thus, specificity determinants should be focused on non-homologous residues between ImmE3 and ImmE6 (or ImmDF13). In fact, our previous genetical experiment demonstrated that a single amino acid replacement of Trp-47 in ImmE6 by Cys-47 of the ImmE3 type acquired the ImmE3 immunity. In addition, Glu-19 in ImmE3 was suggested to be an auxiliary ImmE3 determinant, and in ImmE6, both His-5 and Trp-47 were inferred to be the specificity determinants. On the basis of the ImmE3 structure, these three positions are located near the center of the β-sheet, and their side groups protrude into the solvent (Fig. 3). This structure supports the idea that ImmE3, and possibly ImmE6, interacts with T2A mainly on the surface of the β-sheet. Our preliminary chemical shift perturbation experiment suggested that the β-sheet is included in the contact surface (unpublished results).

4.3. *Relation with RNA binding proteins*

There are no homologous proteins of ImmE3 so far except ImmE6 and ImmDF13, the higher order structures of which are not known. The β-sheet motif is

Table I

Statistics of the distance violation for the 10 refined structure

Distance (Å)	No. of violation	Average violation (Å)
0.0–0.1	64.9 ± 3.21	0.033 ± 0.026
0.1–0.2	12.8 ± 2.97	0.139 ± 0.029
0.2–0.3	2.7 ± 1.25	0.246 ± 0.031
0.3–0.4	0.9 ± 1.29	0.354 ± 0.026
> 0.4	1	0.410
Total	81.4 ± 4.033	0.060 ± 0.067

Table II

Structural statistics for the 10 refined structure

Parameter	Value
Bond length	0.0061 ± 0.00018 (Å)
Angle	0.79 ± 0.026 (degree)
Angle	0.014 ± 0.00054 (Å)
Improper	0.57 ± 0.064 (degree)
Chiral volume	0.066 ± 0.0032 (Å^3)
E_{dist} [a]	20.89 ± 3.60 (kcal/mol)
E_{tors} [a]	2.43 ± 1.71 (kcal/mol)
E_{L-J} [b]	−346.34 ± 13.91 (kcal/mol)

[a] The distance (E_{dist}) and torsion (E_{tors}) energies are calculated with force constants of 2 kcal/Å^4/mol and 200 kcal/rad^2 mol, respectively.
[b] E_{L-J} is the 6–12 Lennard-Jones potential energy calculated with AMBER all-atom force field parameters.

frequently observed in such proteins that interact with other proteins [17], however, the overall folding of ImmE3 seems to resemble the RNA binding domain of RNA binding proteins, which has a four-stranded antiparallel β-sheet, two α-helices lying on one side of the sheet, with basic residues located around the β-sheet and aromatic residues contained on the surface of the sheet [18–20]. Interestingly, in ImmE3, which has almost the same size as the RNA binding domain, basic residues (K3, K11, K17, K23, K40 and R82) and three aromatic residues, F9, F16 and F74 in the β-sheet, seem to correspond to those in the RNA binding domain described above. While the β-sheet of the RNA binding domain forms the contact surface to RNA [21], ImmE3 recognizes T2A using the β-sheet, as described above. Colicin E3 uniquely cuts the 16 S RNA when integrated in the 70 S ribosome, and this mode of action gave us the idea that E3 may recognize the higher-order structure of ribosomes, in particular its proteinaceous region. Therefore, the structure of ImmE3 possibly imitates that of a ribosomal protein, a kind of RNA binding protein, which E3 recognizes.

We have found that ImmE6 has almost the same secondary structure as ImmE3 (unpublished data), and an analysis of the T2A–Imm complex structures is in progress in our group. Many Imm variants with various specificities that we have obtained are also being examined by genetic and physicochemical experiments.

Acknowledgements: This work was supported by a Grant-in-Aid from the Ministry of Education, Science and Culture of Japan, and JSPS Fellowships for Japanese Junior Scientists (to S.Y.).

REFERENCES

[1] Pugsly, A.P. and Oudega, B. (1987) Plasmids: A Practical Approach, IRL Press, Oxford.
[2] Boon, T. (1971) Proc. Natl. Acad. Sci. USA 68, 2421–2425.
[3] Bowman, C.M., Sidikaro, J. and Nomura, M. (1971) Nature New Biol. 234, 133–137.
[4] Suzuki, K. and Imahori, K. (1978) J. Biochem. 84, 1637–1640.
[5] Ohno-Iwashita, Y. and Imahori, K. (1980) Biochemistry 19, 652–659.
[6] Jakes, K.S. and Zinder, N.D. (1974) Proc. Natl. Acad. Sci. USA 74, 3380–3384.
[7] Mochitate, K., Suzuki, K. and Imahori, K. (1981) J. Biochem. 89, 1609–1618.
[8] Masaki, H. and Ohta, T. (1985) J. Mol. Biol. 182, 217–227.
[9] de Graaf, F.K., Niekus, H.G.D. and Klootwijk, J. (1973) FEBS Lett. 35, 161–165.
[10] Akutsu, A. and Masaki, H. and Ohta, T. (1989) J. Bacteriol. 171, 6430–6436.
[11] Masaki, H., Akutsu, A. and Ohta, T. (1991) Gene 107, 133–138.
[12] Yajima, S., Muto, Y., Yokoyama, S., Masaki, H. and Uozumi, T. (1992) Biochemistry 31, 5578–5586.
[13] Nakai, T., Kidera, A. and Nakamura, H. (1993) J. Biomol. NMR 3, 19–40.
[14] Morikami, K., Nakai, T., Kidera, A., Saito, M. and Nakamura, H. (1992) Comput. Chem. 16, 243–248.
[15] Wüthrich, K. (1986) NMR of protein and nucleic acids, Wiley, New York.
[16] Weiner, S.J., Kollman, P.A., Nguyen, D.T. and Case, D.A. (1986) J. Comp. Chem. 7, 230–252.
[17] Branden, C. and Tooze, J. (1991) Introduction to Protein Structure, Garland.
[18] Wittekind, M., Görlach, M., Friedrichs, M., Dreyfuss, G. and Mueller, L. (1992) Biochemistry 31, 6254–6265.
[19] Nagai, K., Oubridge, C., Jessen, T.H., Li, J. and Evans, P.R. (1990) Nature 348, 515–520.
[20] Ramakrishnan, V. and White, S.W. (1992) Nature 358, 768–771.
[21] Görlach, M., Wittekind, M., Beckman, R., A., Mueller, L. and Dreyfuss, G. (1992) EMBO J. 11, 3289–3295.