Structure of a Ribonuclease B+d(pA)$_4$ Complex

Tzu-Ping Ko, Roger Williams and Alexander McPherson*

Department of Biochemistry, University of California, Riverside, California 92521, USA

Abstract

The structure of a tetragonal crystal of bovine pancreatic RNase B complexed with d(pA)$_4$ was determined by molecular replacement and difference Fourier methods. This crystal belongs to space group $P4_12_12$ and has unit-cell dimensions $a = b = 44.5$, $c = 156.5$ Å. The model consists of the enzyme and a tetranucleotide with fractional occupancies, suggesting multiple modes of oligonucleotide binding. It does not include any polysaccharide residues or solvent molecules. After refinement at 2.7 Å, the R value was 0.163 with acceptable stereochemistry. The model illustrates a set of well defined interactions for substrate binding, particularly between the central dinucleotide and the enzyme.

1. Introduction

The structure of RNase A (E.C. 3.1.27.5) with d(pA)$_4$ and with d(pT)$_4$ have been solved by X-ray diffraction analysis and used to model the association of an extended polynucleotide chain with the binding surface of the protein (McPherson, Brayer & Morrison, 1984, 1986; McPherson, Brayer, Cascio & Williams, 1986; Birdsall & McPherson, 1992). The crystals used for these studies were of orthorhombic space group $P2_12_12_1$ and the protein molecules, in the case of d(pA)$_4$, were seen to have four oligomers of the nucleic acid bound. The structure of RNase B from bovine pancreas, the glycosylated form of RNase A, in an unliganded state, was solved independently from a monoclinic crystal (Williams, Greene & McPherson, 1987). In that crystal the carbohydrate component was observed to be disordered and its disposition uncertain. RNase B has not previously been visualized with nucleic acid bound.

RNase B complexed with d(pA)$_4$ was crystallized in a unique tetragonal unit cell having as its asymmetric unit one molecule of RNase B plus an indeterminate number of d(pA)$_4$ oligomers (Brayer & McPherson, 1981). These crystals presumably contained protein, nucleic acid, and carbohydrate. We report here the solution by molecular replacement of the structure of these RNase B + d(pA)$_4$ crystals having space group $P4_12_12$ with $a = b = 44.5$ and $c = 156.5$ Å.

2. Methods

The growth of crystals of the complex from protein purchased from Sigma Biochemicals and oligonucleotides from Pharmacia Co. has beeen described previously (Brayer & McPherson, 1981). It was achieved by use of vapor diffusion (McPherson, 1982) at low ionic strength from polyethylene glycol solutions at 277 K and was generally complete within 24–48 h.

X-ray diffraction data were collected using an Enraf–Nonius CAD-4 automated diffractometer as described for the RNase A + d(pA)$_4$ complex crystals (McPherson *et al.*, 1984). The resolution of the data was 2.6 Å though the diffraction pattern extends to higher limits. Molecular-replacement procedures employed were those of Crowther & Blow (Crowther & Blow, 1967; Crowther, 1972), which had been included in an early version of *MERLOT* (Fitzgerald, 1988). Fourier calculations used the *FFT* of Ten Eyck (1985). Images were created and analyzed using the program *FRODO* (Jones, 1982, 1985) running on an Evans and Sutherland PS390 graphics system. All other calculations were performed on the UCR campus VAX network and an SGI 320 VGX workstation.

Initial rigid-body refinement was carried out using the program *CORELS* (Sussman, Holbrook, Church & Kim, 1977; Sussman, 1985). Subsequent refinement of the complete model employed simulated annealing by *X-PLOR* (Brünger, Kuriyan & Karplus, 1987; Brünger, 1991) and the restrained least-squares approach in *TNT* (Ten Eyck, Weaver & Matthews, 1976; Tronrud, Ten Eyck & Matthews, 1987). $2F_o - F_c$ and $F_o - F_c$ Fourier syntheses were used to improve the model, again using *FRODO*. The model used for the molecular-replacement searches was that of RNase A as determined and refined by Wlodawer, Borkakoti, Moss & Howlin (1986) using both X-ray and neutron diffraction data. Our assumption was that the approach would succeed even in the absence of the nucleic acid and carbohydrate components of the asymmetric unit. This was in fact the case.

3. Results

The molecular-replacement search was complicated somewhat because the space group of the crystals was

uncertain, both $P4_12_12$ and $P4_32_12$ being equally possible. While this had no effect on the rotation function search, it was of course important for a correct translation function and for all subsequent refinement.

With only one molecule of RNase in the asymmetric unit of the crystals, the rotation function was relatively straightforward. A single major peak appeared at $\alpha = 62.5$, $\beta = 90.0$, $\gamma = 180.0°$. The translation function produced one consistent solution, if space group $P4_12_12$ was assumed, at (0.83, 0.40, 0.075) and another at (0.60, 0.17, 0.425) if the alternate space group was used.

Rigid-body refinement of an RNase A molecule disposed according to the two possibilities in unit cells of the two symmetries was adequate to discriminate the correct from incorrect enantiomorphic space group. In $P4_12_12$ an R factor of 0.36 was obtained for all reflections in the range 20.0–2.6 Å while the corresponding value for the alternative was 0.44. The overall correlation coefficients were 0.62 and 0.41, respectively.

Calculated phase angles following rigid-body refinement were exploited in the subsequent calculation of $2F_o - F_c$ and $F_o - F_c$ difference Fourier maps. As shown in Fig. 1, two of the adenylate residues were clearly visible in the electron-density map. These were located near the active site and were flanked by the side chains of His12, Lys41 and His119. The purine rings appeared to be stacked next to the imidazole group of His119. At both 5′ and 3′ ends of this dinucleotide, one additional adenylate residue could be inferred from the maps, though the densities were weaker. Consequently, four nucleotides were modelled, numbered from 201 to 204.

The original model of RNase A (Wlodawer et al., 1986) was also examined in the context of the difference Fourier maps. Most of the polypeptide chain was very well defined. There was neither breakage nor weakening of the densities required for tracing the protein fold.

Some flexible side chains on the surface, such as Lys31 and Lys91, however, lacked corresponding strong densities. Guided by the Fourier maps, we made numerous, though generally minor, adjustments to the original model. Unfortunately, we were not able to identify density for the oligosaccharide moiety, which is attached to Asn34 and is unique to RNase B.

The reconstructed model consisted of 1035 non-H atoms, including those of the four deoxyadenosyl nucleotides, but no solvent molecules. Prior to crystallographic refinement, the model was regularized by FRODO to an overall deviation from ideal geometry of about 5%. Subsequent refinement employed a resolution range of 8.0–2.7 Å. It was initiated using the simulated-annealing procedure of X-PLOR, by slow-cooling the molecule from 3000 to 300 K. No variation in the thermal parameters was allowed and these remained $12\,\text{Å}^2$ for all atoms. This yielded an R value of 0.204. Calculated structure factors were again used to compute difference Fourier maps but at this stage few modifications of the model seemed necessary.

Initial refinement was followed by the conjugate-gradient procedure of TNT, in which the temperature factors were also refined. The final model yielded an R value of 0.163 and a correlation coefficient of 0.894, for 3071 reflections. Statistics from this refinement are presented in Table 1. The average thermal parameter was $15.7\,\text{Å}^2$, while 27 atoms had β-values of 30–$40\,\text{Å}^2$. The refined occupancies of the four nucleotides 201–204 were 0.367, 1.00, 0.833 and 0.484, respectively. Fig. 2 shows the final $F_o - F_c$ map superimposed on the tetranucleotide model. This was omitted when calculating the map. A Luzzati plot (Luzzati, 1952), shown in

Fig. 1. A portion of the $2F_o - F_c$ map calculated using the RNase A model following rigid-body refinement. The map showed strong densities for two adenosyl residues bound within the active site. The imidazole group of His119, which stacked with the base rings, is also shown.
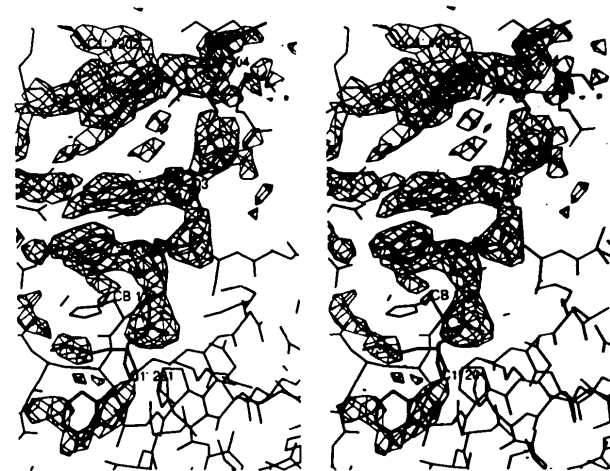

Fig. 2. A portion of the $F_o - F_c$ map calculated using the refined structure phases, but with the nucleic acid moiety omitted. The map shown in the figure encompasses the entire $d(pA)_4$ model. The densities for the central nucleotides are similar to those in Fig. 1. Also shown are two dyad-symmetry-related adenine rings, near the top-left corner.

Table 1. *TNT refinement of the tetragonal $P4_12_12$ RNase B + d(pA)$_4$ structure*

Unit cell $a = b = 44.5$. $c = 156.5$ Å. $\alpha = \beta = \gamma = 90.0°$. Resolution $= 8.0$–2.7 Å. No. of reflections $= 3071$. Native $K = 9.23$. Overall $B = 0.0$. $K_{solvent} = 0.69$. $B_{solvent} = 126$. Overall $R = 0.163$. Correlation coefficient $= 0.895$.

Resolution breakdown

d_{min} (Å)	5.27	4.40	3.92	3.60	3.36	3.17	3.02	2.90	2.79	2.70
No. of F_{obs}	401	391	392	362	335	298	279	229	224	160
% Complete	87.4	85.2	85.4	78.9	73.0	64.9	60.8	49.9	48.8	34.9
R_{shell}	0.234	0.148	0.142	0.135	0.136	0.168	0.155	0.184	0.179	0.230
R_{sphere}	0.234	0.185	0.169	0.161	0.156	0.158	0.158	0.160	0.161	0.163

Thermal parameter distribution
Total No. of atoms $= 1035$. Average $B_{iso} = 15.46$ Å^2. R.m.s. $= 2.19$ Å^2.

B_{max} (Å^2)	5	10	15	20	25	30	35	40
No. of atoms	38	157	338	265	157	53	23	4

Stereochemical deviation

Category	Bond length	Bond angle	Torsional angle	Trigonal atom	Planar group	Bad contact	Chiral center
No. of restraints	1066	1445	589	28	144	63	146
R.m.s. deviation (Å, °)	0.020	2.982	21.91	0.018	0.020	0.083	0

Fig. 3, indicates that the estimated error in the atomic coordinates is about 0.20–0.25 Å.

4. Discussion

It was reported by Williams *et al.* (1987) that the overall conformation of RNase A and the unliganded RNase B are almost identical. This is true in the case of RNase B complexed with d(pA)$_4$ as well. An overall r.m.s. deviation of 1.071 Å in 951 atomic coordinates was obtained when the refined model was superimposed on the original RNase A model. For 496 backbone atoms the r.m.s.d. was 0.662 Å and for the 124 Cα atoms it was 0.629 Å. Most of the differences were less than 0.5 Å while in five residues they were greater than 1 Å. These five residues, numbered 1, 21, 37, 67 and 89, are located at the N-terminus, in loops between α-helices and β-strands, and in the neighborhood of a β-bulge. They are all near the protein surface and might be expected to be flexible. The loop of residues 16–22, in which a
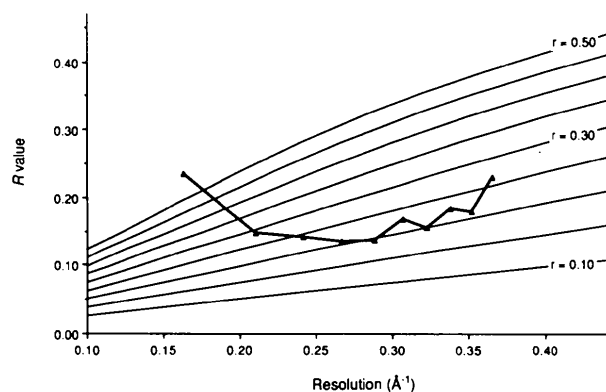
maximum deviation appeared, is the cleavage site that produces RNase S. Residues of significantly higher mobility were also implied by the distribution of temperature factors, which showed a strong correlation with the variation in coordinates.

Fig. 4 shows a packing diagram of RNase B in the tetragonal crystal. Packing of protein molecules alone in the crystal is quite tenuous, with only the barest of interfaces relating them. Crystal contacts can be grouped into four regions with no more than a total of 14 amino-acid residues involved. The interactions occur almost
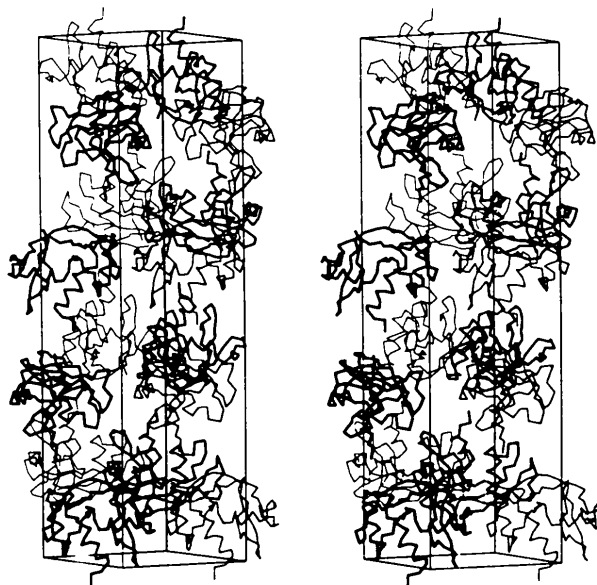
Fig. 4. Packing diagram of the RNase B molecules alone in the tetragonal crystal. The nucleic acid molecules are not shown. A single unit cell, outlined in the figure, contains eight molecular complexes. For clarity, eight additional symmetry-related RNase B molecules were included.

Fig. 3. A Luzzati plot (Luzzati, 1952) for the refined structure of the tetragonal RNase B + d(pA)$_4$ crystal. The R values for individual shells of reflections were calculated using *TNT*.

entirely between loop structures, and no specific salt bridges or hydrogen bonds could be clearly identified. The oligosaccharide moiety linked to Asn34, as noted above, could be localized only from diffuse density as a protrusion into a void in the unit cell. Thus, it is most likely disordered and heavily associated with water molecules. There seems to be little, if any, possibility for it to be in contact with anything except solvent.

Although the occupancy of residue Ade201 is low in the model, the 5′ phosphate group appears to interact with the ε-amino group of Lys66, which corresponds to the p_0 site. The sugar and base moiety binds to the R_1B_1 site of the enzyme. The adenine group does not fit into the pocket for optimal binding to a smaller pyrimidine ring. A similar observation was reported earlier for the RNase A + d(pA)$_4$ complex (McPherson et al., 1984). Yet the atom N6 of the base is in contact with OG1 of Thr45, probably forming a hydrogen bond.

Residue Ade202 occupies the p_1 site and the R_2B_2 site. The phosphate group is surrounded by the three catalytic groups of His12, Lys41 and His119, as observed in other RNase A–nucleotide complexes. The planar base (B_2) stacks with the imidazole group of His119 and hydrogen bonds to the side chains of Asn67 and Asn71, as shown in Fig. 5. Residue Ade203 binds to the p_2 and R_3B_3 sites. The phosphate is in contact with the ε-amino group of Lys7. The base B_3 makes a further stacking with the base B_2 and forms hydrogen bonds to the side chain of Gln69, which is also shown in Fig. 5. These interactions involving the two central adenylate residues are virtually identical of those of A^3 and A^4 in the model of RNase A + d(ApTpApApG) complex reported by Fontecilla-Camps, de Llorens, le Du & Cuchillo (1994). For these two crystal structures, the positions of 42 equivalent atoms in the dinucleotide corresponding to the $p_1R_2B_2p_2R_3B_3$ sites deviate by an r.m.s. of 0.504 Å.

The strong electron density seen in our difference Fourier map (Fig. 2) confirmed the disposition of the

second nucleoside (R_3B_3). If the base B_3 is replaced by a guanine, the side chain of Gln69 may rotate 180° about the CB—CG bond, flipping the amide group, and it can still form hydrogen bonds to the base. We also observed that the side chain of Glu111 would be in an optimal position for interacting with the 2-NH$_2$ group of the guanine, as proposed by Fontecilla-Camps et al. (1994). In contrast, if the base B_2 were a guanine, similar rotation of the side chain of Asn71 would be hindered, because the atom ND2 is hydrogen bonded to the main-chain CO group of Cys110. These may offer a possible explanation for the substrate preference of RNase A and B for an adenosine at the R_2B_2 site and a guanosine at the R_3B_3 site.

The electron density corresponding to the phosphate of Ade204 (the p_3 site) is strong and clearly visible in the difference Fourier map, as shown in Fig. 2. However, it is not involved in any direct interaction with the enzyme, except perhaps the ε-amino group of Lys1 which is 5 Å away. The sugar moiety and the adenine ring of Ade204 are in contact with the side chains of Asn67 and Gln69 of a twofold symmetry-related RNase molecule, and with the N6 atom of Ade203 related by the same dyad symmetry. The base is also in close proximity to the surface of a third neighboring, 4$_1$ symmetry-related enzyme molecule, near Asn103. In the crystal, these interactions may contribute to stabilizing forces between molecules. Besides, there are some residual densities about the above-mentioned dyad axis which relates two approaching d(pA)$_4$ molecules. We suspect that the oligonucleotide might be further extended at the 3′ end, reflecting multiple modes of binding. These could be disordered, yet they may provide additional intermolecular contacts in the crystal.

A model of the tetranucleotide bound to the RNase B is shown in Fig. 6. We did not observe other bound nucleotide molecules more distant from the active site. In this tetragonal crystal of RNase B + d(pA)$_4$, unlike the orthorhombic crystals of RNase A + d(pA)$_4$, we did not observe multiple, ordered oligonucleotide chains that
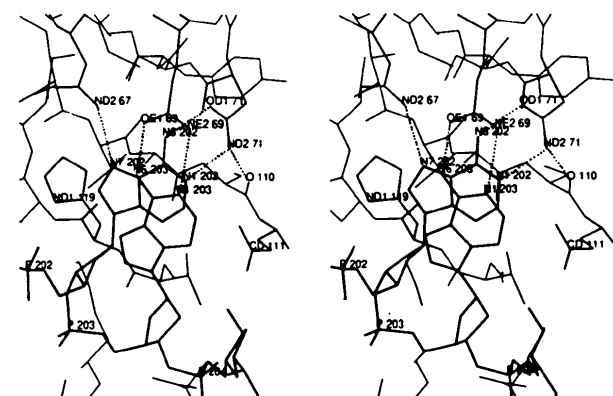

Fig. 5. Hydrogen-bond interactions between RNase B and the bases of Ade202 and 203. Residues Asn67, Gln69 and Asn71 are involved. Also labeled are the side chains of Glu111 and His119, and three phosphates in the oligonucleotide.
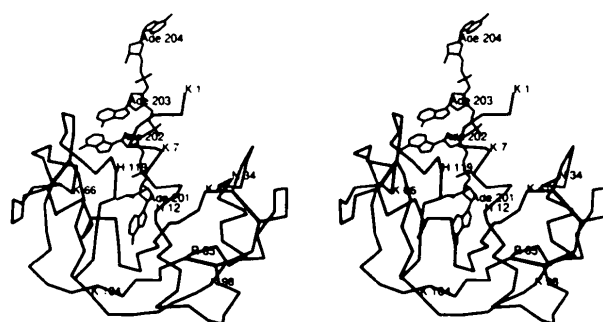

Fig. 6. The binding of a tetranucleotide d(pA)$_4$ to the RNase B backbone model, according to the refined crystal structure. Disulfide bonds are indicated by heavy lines. The Cα atoms of His12 and 119, Lys1, 7, 41, 66, 98 and 104, Arg85 and Asn34 are labelled.

appeared to 'wrap around' the enzyme molecule. This could be due to the different contacts between protein molecules in the two crystals, and/or to the presence of disordered oligosaccharide which could preclude the ordered binding of additional d(pA)$_4$ oligomers. Nevertheless, in the model presented in Fig. 6, addition of nucleotides at the 5′ end of the tetranucleotide strand could be made that would yield essentially the same path as observed in the RNase A + d(pA)$_4$ complex (McPherson *et al.*, 1986). That is, this model is consistent with extension toward Lys104, Arg85, Lys98 and further around the enzyme. There is, as noted above, however, no density in the difference Fourier maps of this crystal to support such a tracing.*

---

* Atomic coordinates and structure factors have been deposited with the Protein Data Bank, Brookhaven National Laboratory (Reference: 1RBJ, R1RBJSF). Free copies may be obtained through The Managing Editor, International Union of Crystallography, 5 Abbey Square, Chester CH1 2HU, England (Reference: GR0431).

References

Birdsall, D. & McPherson, A. (1992). *J. Biol. Chem.* **267**, 22230–22236.

Brayer, G. D. & McPherson, A. (1981). *J. Biol. Chem.* **257**, 3359–3361.

Brünger, A. T. (1991). *Ann. Rev. Phys. Chem.* **42**, 197–223.

Brünger, A. T., Kuriyan, J. & Karplus, M. (1987). *Science*, **235**, 458–460.

Crowther, R. A. (1972). *The Molecular Replacement Method*, edited by M. G. Rossmann, pp. 173–178. New York: Gordon & Breach.

Crowther, R. A. & Blow, D. M. (1967). *Acta Cryst.* **23**, 544–548.

Fitzgerald, P. M. D. (1988). *J. Appl. Cryst.* **21**, 273–278.

Fontecilla-Camps, J. C., de Llorens, R., le Du, M. H. & Cuchillo, C. M. (1994). *J. Biol. Chem.* **269**, 21526–21531.

Jones, T. A. (1982). *Computational Crystallography*, edited by D. Sayer, pp. 303–317. Oxford: Clarendon Press.

Jones, T. A. (1985). *Methods Enzymol.* **115**, 157–171.

Luzzati, P. A. (1952). *Acta Cryst.* **5**, 802–810.

McPherson, A. (1982). *Preparation and Analysis of Protein Crystals*, pp. 94–97. New York: John Wiley.

McPherson, A., Brayer, G., Cascio, D. & Williams, R. (1986). *Science*, **232**, 765–768.

McPherson, A., Brayer, G. D. & Morrison, R. D. (1984). *J. Mol. Biol.* **189**, 305–327.

McPherson, A., Brayer, G. & Morrison, R. (1986). *Biophys. J.* **49**, 209–219.

Sussman, J. L. (1985). *Methods Enzymol.* **115**, 271–303.

Sussman, J. L., Holbrook, S. R., Church, G. M. & Kim, S. H. (1977). *Acta Cryst.* **A33**, 800–824.

Ten Eyck, L. F. (1985). *Methods Enzymol.* **115**, 324–337.

Ten Eyck, L. F., Weaver, L. A. & Matthews, B. W. (1976). *Acta Cryst.* **A32**, 349–355.

Tronrud, D. E., Ten Eyck, L. F. & Matthews, B. W. (1987). *Acta Cryst.* **A43**, 489–501.

Williams, R. L., Greene, S. M. & McPherson, A. (1987). *J. Biol. Chem.* **262**, 16020–16031.

Wlodawer, A., Borkakoti, N., Moss, D. S. & Howlin, B. (1986). *Acta Cryst.* **B42**, 379–387.