Acta Cryst. (1982). B**38**, 2210–2217

Ribonuclease-A: Least-Squares Refinement of the Structure at 1·45 Å Resolution

BY NIVEDITA BORKAKOTI,* DAVID S. MOSS AND REX A. PALMER†

Department of Crystallography, Birkbeck College, Malet Street, London WC1E 7HX, England

(*Received* 29 *April* 1981; *accepted* 15 *February* 1982)

Abstract

The crystal structure of bovine pancreatic ribonuclease-A has been refined by restrained least-squares analysis employing X-ray diffractometer data to 1·45 Å resolution. The current R factor for 19 238 reflections is 0·26 and 0·24 for 17 427 reflections with $I(hkl) > 0$. The r.m.s. deviation from ideality of bond lengths is 0·01 Å. Corrections, mostly of minor character, to previous models of the secondary structure have been made and a quantitative analysis of intramolecular hydrogen bonds is given. A total of 79 solvent molecules have been clearly identified in the first coordination sphere around the enzyme molecule and included in the least-squares analysis. A sulphate anion occurs in the active site and has also been refined as part of the structure. Further new features of the structure to emerge are: alternative sites for the His-119 side group with occupancies, refined in the analysis, of 0·80 and 0·20 respectively; a solvent molecule hydrogen bonded to the N-terminal amino group; and extensive disorder of the side chains in the region of residues 35–39. The r.m.s. deviation in atomic position between the current model and the starting model is 1·1 Å including some shifts of 7–8 Å where major rebuilding of side groups was necessary.

Introduction

Although ribonuclease-A was one of the first enzymes to be studied by X-ray crystallography, and two independent structure analyses were reported some time ago (Kartha, Bello & Harker, 1967; Carlisle, Palmer, Mazumdar, Gorinsky & Yeates, 1974), neither of these groups published atomic coordinates. A set of coordinates for ribonuclease-S (formed by subtilisin cleavage of ribonuclease-A between Ala-20 and Ser-21) was published by Wyckoff, Tsernoglou, Hanson, Knox, Lee & Richards (1970), but this work also lacked the precision which can be attained by least-squares refinement. Nevertheless, subsequent publications based on the ribonuclease-S structure described important features of the catalytic mechanism deduced by introducing inhibitors and covalent modifiers into the enzyme molecule (Wodak, Liu & Wyckoff, 1977; Mitsui, Urata, Torii & Irie, 1978; Holmes, Deiters & Galluci, 1978). During the past ten years a great deal of new information has become available on the mechanism and evolution of ribonuclease (see, for example, Huestis & Raftery, 1971; Haer, Maurer & Rüterjans, 1976; Welling, Lenstra & Beintema, 1976; Pares, Arus, Llorens & Cuchillo, 1978). Techniques have also been developed for refinement and coordinate analysis of protein structure models. These innovations have facilitated the structure refinement reported here and the subsequent analysis of the catalytic mechanism to be published later.

Recently Wlodawer (1980) has reported 2·5 Å resolution X-ray studies and 2·8 Å resolution neutron diffraction studies leading to a refined model of ribonuclease-A. The starting model for Wlodawer's analysis was a hybrid based on unpublished coordinates for ribonuclease-A (Kartha *et al.*, 1967; Carlisle *et al.*, 1974) and the ribonuclease-S coordinates of Wyckoff *et al.* (1970) and Powers (1976). Wlodawer (1980) reports extensive rebuilding of the structure in order to produce a refinable model with a final R factor of 0·25 and r.m.s. deviation of bond lengths from ideality of 0·042 Å.

The present studies were initiated from Wlodawer's (1980) atomic parameters and employed new X-ray intensity data to 1·45 Å resolution. We report now details of these analyses and the results including active-site geometry, location of the sulphate molecule, hydrogen bonding, intermolecular contacts and the solvent structure. Further publications in preparation include analysis of the thermal vibrations of the molecule, detailed study of the hydrogen bonds involving solvent molecules, and details of the catalytic mechanism.

Experimental

Crystals were prepared as described by Carlisle *et al.* (1974). Large crystals with dimensions up to about 0·7 × 1·0 × 0·7 mm were used for X-ray intensity measurements employing a 5 m diffracted-beam colli-

* SERC Postdoctoral Research Assistant.

† To whom correspondence should be addressed.

0567-7408/82/082210-08$01.00

mator. The unit-cell parameters were refined by least squares using values of θ, χ, φ for eight reflections measured on a Hilger & Watts Y290 four-circle diffractometer with Cu $K\alpha$ radiation (Tickle, 1979). Intensities were measured on the same instrument using ω step-scans, typically with $50 \times 0{\cdot}02°$ steps, the exact values used being adjusted to suit each crystal. The time per step was between 1 and 3 s. A total of five crystals were used to measure $I(\bar{h}k\bar{l})$ and $I(hkl)$ for $d(hkl) \geq 1{\cdot}45$ Å, allowing overlap in the range of θ covered for successive crystals to facilitate scaling. Intensities of 35 702 reflections were measured which after scaling and merging equivalents gave a data set of 17 427 intensities, with $I(hkl) > 0$.

Lorentz and polarization factors were applied in the usual manner and semi-empirical absorption corrections were applied using the method of North, Phillips & Mathews (1968), absorption curves being derived from measurements of selected $0k0$ reflections, **b** being parallel to the capillary mounting tube.

Refinement techniques

The method of restrained least squares has increasingly become a standard technique in the refinement of protein structures. Reviews of the technique have been given, for example, by Konnert (1976) and Hendrickson & Konnert (1980). Other proteins refined by this method include sickling deer type III hæmoglobin (Girling, Houston, Schmidt & Amma, 1980) and *Streptomyces griseus* serine protease (SGPA) (Sielecki, Hendrickson, Broughton, Delbaere, Brayer & James, 1979).

The structure of ribonuclease A, excluding H atoms, was refined by a least-squares procedure (Morffew & Moss, 1982) alternating with extensive model-building sessions, the latter using the program *FRODO* (Jones, 1978) implemented on an Evans and Sutherland Picture System 2 (computer graphics). The course of the analysis is indicated in Table 1.

Table 1. *Course of the refinement of ribonuclease-A*

R is the crystallographic R factor $\sum_{hkl}||F_o| - |F_c||/\sum_{hkl}|F_o|$ and δ is the r.m.s. shift in atomic position.

Pass 1	Pass 2	Pass 3	Pass 4
1(*a*) Wlodawer's (1980) coordinates. 2·3 Å, $R = 0{\cdot}273$.	2(*a*) Two cycles of refinement using parameters from step 1(*i*). 1·45 Å, $R = 0{\cdot}354$, $\delta = 0{\cdot}053$ Å.	3(*a*) Four cycles of refinement using parameters from step 2(*f*). 1·45 Å, $R = 0{\cdot}336$, $\delta = 0{\cdot}034$ Å.	4(*a*) Six cycles of refinement using parameters from step 3(*f*). 1·45 Å, $R = 0{\cdot}27$, $\delta = 0{\cdot}029$ Å.
1(*b*) Four cycles of least-squares refinement at 2·3 Å, $R = 0{\cdot}248$, $\delta = 0{\cdot}048$ Å.	2(*b*) Calculation of electron density maps (*FFT*).	3(*b*) Three solvent molecules with high U_{iso} omitted. One cycle of refinement, 1·45 Å, $R = 0{\cdot}31$, $\delta = 0{\cdot}034$ Å.	4(*b*) Lys-41 and His-119 side groups both removed.
1(*c*) Inclusion of 1·45 Å data, $R = 0{\cdot}385$, $\delta = 0{\cdot}108$ Å.	2(*c*) Graphics inspection of maps (*FRODO*).	3(*c*) Calculation of electron density maps (*FFT*).	4(*c*) Two cycles of refinement. 1·45 Å, $R = 0{\cdot}27$, $\delta = 0{\cdot}030$ Å.
1(*d*) Four cycles of least squares. 1·45 Å, $R = 0{\cdot}344$, $\delta = 0{\cdot}015$ Å.	2(*d*) Regions *A* and *B* still unclear [see 1(*g*)].	3(*d*) Graphics inspection of maps (*FRODO*).	4(*d*) Electron density maps recalculated.
1(*e*) Calculation of $(\|F_o\| - \|F_c\|)$ and $(2\|F_o\| - \|F_c\|)$ maps (*FFT*).	2(*e*) 79 solvent molecules located.	3(*e*) Rebuilding of region *A*, residues 17–24 and adjustment of residues 15 and 16. Rebuilding of region-*B* main chain and side chains 36 and 38, and approximate fitting of Leu-35, Lys-37 and Arg-39 side chains. Fitting of SO_4^{2-} O atoms in tetrahedral array and feasible orientation. Inclusion of solvent atom at N terminus. Possibility of multiple His-119 side-group sites and lack of clarity in Lys-41 side group noted. Lys-41 side group rebuilt.	4(*e*) Lys-41 side group rebuilt. His-119 side-group alternatives His-119*A* and *B* fitted to density. 1·45 Å, $R = 0{\cdot}27$, $\delta = 0{\cdot}081$ Å.
1(*f*) Graphics analysis (*FRODO*).	2(*f*) Structure factor calculation still omitting residues 17–24 and 35–39, and including S of SO_4^{2-} and 79 solvent molecules. 1·45 Å, $R = 0{\cdot}359$, $\delta = 0{\cdot}088$ Å.	3(*f*) Structure factor calculation, including all side groups. 1·45 Å, $R = 0{\cdot}34$, $\delta = 0{\cdot}035$ Å.	4(*f*) Structure factor calculation including rebuilt Lys-41 and multiple His-119 side groups. 1·45 Å, $R = 0{\cdot}027$, $\delta = 0{\cdot}048$ Å.
1(*g*) Region *A*: residues 17–24. Region *B*: residues 35–39. Poorly defined and omitted temporarily after failure to rebuild model in these regions.			Pass 5
1(*h*) SO_4^{2-} characterized and located near active site. S coordinates determined.			5(*a*) Five cycles of refinement using parameters from 4(*f*). His-119*A* and *B* refined separately including occupancy. 1·45 Å, $R = 0{\cdot}26$, $\delta = 0{\cdot}035$ Å.
1(*i*) Structure factor calculation omitting residues 17–24 and 35–39, and including S of SO_4^{2-}. 1·45 Å, $R = 0{\cdot}367$, $\delta = 0{\cdot}052$ Å.			

In his fast Fourier least-squares analysis of actinidin, Baker (1980) included H atoms in geometrically assigned positions for structure factor calculations. The present program does not have this facility which in the case of X-ray diffraction studies appears to offer only a marginal improvement relative to the additional effort involved. However, higher-resolution X-ray studies may indeed warrant inclusion of H positions as is customary with small-molecule analyses.

The function minimized by least squares in the program *RESTRAIN* was:

$$R = \sum_{hkl} w_F(|F_o| - |F_c|)^2 + \sum_i w_i(d_T^i - d_c^i)^2$$

where d_T^i and d_c^i are target and calculated values for the ith interatomic distance, and w_F and w_i are weights associated with the respective terms in the summation. The geometrical restraints were of two types, namely bonded distances and bond angles and distances, respectively. For each type of restraint the same weight was applied to all restraints.

The interatomic restraints only weakly restrained the planarity of phenyl and peptide groups. Deviations from planarity in the phenyl groups, which were of the order of 0·05 Å at the end of the refinement were corrected using the regulization procedure of Hermans & McQueen (1974).

No restraints were placed on hydrogen-bond interactions and thus the water molecules were subjected to unrestrained refinement. The normal equations of least-squares refinement were set up in the block-diagonal approximation as 3 × 3 blocks for the deviations of the structure amplitudes with respect to the positional parameters. The diagonal approximation was used for the thermal-parameter derivatives and the correlation of the latter with the scale factor was taken into account by using a 2 × 2 block containing derivatives with respect to scale factor and overall temperature factor in the manner described by Rollett (1965).

All derivatives of the restrained distances were fully included in the normal matrix. The normal equations were solved by the Gauss–Seidel method.

For the 2·3 Å resolution refinement, weights w_i were chosen so that deviations <0·2 Å from standard geometry were allowed. At 1·45 Å, weights w_i were chosen such that in each class, defined below, the average value of $w_i(d_T^i - d_c^i)^2$ was equal to the average of the weighted residuals in the structure-amplitude classes. Structure amplitudes were grouped into classes according to (*a*) sin θ and (*b*) magnitude. Weights were then generated according to: $w_F = a(\sin\theta)^b/(c + |F_o| + d|F_o|^2)$, where the coefficients a, b, c, d were chosen to minimize variation of the average value of $w_F(|F_o| - |F_c|)^2$ in the ranges of sin θ and $|F|$ respectively.

The parameters refined were an overall scale factor, an overall temperature factor, and for each atom the three Cartesian coordinates X_o, Y_o, Z_o and a nominal mean square vibration U_{iso}. In the case of His-119 where two positions were indicated from inspection of a $(2|F_o| - |F_c|)$ electron density map, side-group occupancies were refined for each position and the sum of the occupancies was constrained to be unity. Individual thermal vibration parameters were refined for each group.

In the final cycle of least squares* the values of a, b, c, d for the w_F weighting scheme were: $a = 0{\cdot}002$, $b = 1{\cdot}0$, $c = 200{\cdot}0$, $d = 0{\cdot}001$; final R values were 0·26 for all reflections and 0·24 for $I(hkl) > 0$; $\delta_{r.m.s.} = 0{\cdot}037$ Å (Table 1).

Further details of the course of the analysis are given below. Atomic scattering factors used were those given in *International Tables for X-ray Crystallography* (1974).

A résumé of the average peptide geometry obtained from the 1·45 Å refinement of ribonuclease-A is given below; the values quoted represent the average of all such values in the structure together with their standard deviations:

Cα–N	1·47 (1) Å	N–Cα–C(O)	110 (2)°
Cα–C(O)	1·52 (1)	Cα–C(O)=O	120 (2)
C(O)–N	1·32 (1)	O=C(O)–N	121 (2)
C(O)=O	1·24 (2)	C(O)–N–Cα	121 (2).

Refinement of the structure at 1·45 Å resolution

In the following discussion, unless otherwise stated, the crystallographic R factors are quoted for the complete 1·45 Å data set of 19 238 reflections. A summary of the course of the analysis is given in Table 1.

Pass 1(*A*): *initial refinement*

Restrained least-squares refinement was initiated with a 2·3 Å data set measured from one crystal. The atomic coordinates employed were from Wlodawer's (1980) 2·5 Å refinement. Four cycles of refinement were undertaken and in the final cycle the r.m.s. shift in atomic position was 0·049 Å. Four further cycles were then undertaken with the complete 1·45 Å data set, after which the r.m.s. shift from the original model coordinates was 0·34 Å. Fourier maps using $(2|F_o| - |F_c|)$ and $(|F_o| - |F_c|)$ as coefficients were calculated

* Atomic coordinates and structure factors have been deposited with the Protein Data Bank, Brookhaven National Laboratory (Reference: 1RN3, R1RN3SF), and are available in machine-readable form from the Protein Data Bank at Brookhaven or one of the affiliated centres at Cambridge, Melbourne or Osaka. The data have also been deposited with the British Library Lending Division as Supplementary Publication No. SUP 37005 (2 microfiche). Free copies may be obtained through The Executive Secretary, International Union of Crystallography, 5 Abbey Square, Chester CH1 2HU, England.

and examined on the interactive graphics system. The $(2|F_o| - |F_c|)$ maps proved generally more useful.

Analysis of the maps revealed two areas in the model which required extensive rebuilding. Region *A* extending from Thr-17 to Asn-24 and region *B* from Leu-35 to Arg-39 both exhibited poorly defined electron density in the main chain. In region *A* the small serine and alanine side chains added to the difficulty of rebuilding the model, and in region *B* side chains were almost totally absent in the electron density map.

Pass 1(*B*): *location of the sulphate anion*

Low-resolution studies of ribonuclease have shown that a polyanion, either phosphate (Kartha *et al.*, 1967; Wlodawer, 1980) or sulphate (Wyckoff *et al.*, 1970) may be bound in the vicinity of His-12 and His-119 in the active-site cleft of the molecule. The preliminary difference electron density map at 1·45 Å resolution, based on the protein structure alone, revealed a large peak in this region. As our crystals were prepared from phosphate-free buffer and EDAX (Energy Dispersive Analysis by X-rays) studies indicated no phosphorus in our single crystals, a further examination of the crystals became necessary in order firmly to establish the nature of the bound anion. It is well known that sulphate ions bind to ribonuclease-A (Saroff & Carroll, 1962). The Raman spectrum of solutions of ribonuclease-A has suggested that certain preparations contain five or six bound sulphate ions per enzyme molecule (Chen & Lord, 1976). Our own single-crystal laser Raman spectra indicate a peak at 982 cm^{-1} which corresponds to the totally symmetric stretch of the sulphate ion. It seems certain, therefore, that the prominent peak in our difference density is due to SO_4^{2-}. Preliminary attempts to fit a tetrahedral SO_4^{2-} group into the density, with S–O = 1·5 Å, were however unsuccessful due to the lack of density for the oxygen atoms. The S position alone was therefore included in the next refinement pass located at the centre of the observed density which was close to the P position of Wlodawer's (1980) PO_4^{2-} group. Subsequently it became clear that three of the O atoms of the SO_4^{2-} anion were hydrogen bonded to the enzyme molecule, as described below (pass 3). A libration tensor analysis of the SO_4^{2-} ion is now in progress and will be published separately. This will enable the SO_4^{2-} oxygens to be placed with more confidence.

Pass 2

Excluding all atoms in regions *A* and *B* (residues 17–24 and 35–39 respectively) and including the S of the SO_4^{2-} group, two cycles of refinement reduced *R* from 0·367 to 0·354. The r.m.s. shift per cycle was 0·053 Å. The electron density maps were recalculated and inspected on the computer graphics. Regions *A* and *B* had improved slightly but it was still not possible to rebuild the model in these areas with any degree of confidence. Detailed analysis of the maps did, however, reveal a number of possible solvent peaks within hydrogen-bond distance of the protein molecule. A total of 79 such peaks was observed, each clearly defined in the density and at a distance *r*, 2·5 Å $\leq r \leq$ 3·5 Å, from a N or O atom of the protein molecule. The 79 new sites were included as neutral O atoms in the next round of structure factor and least-squares calculations, the region *A* and *B* atoms still being omitted. No significant density was observed for $r <$ 2·5 Å.

Pass 3

Four further cycles of least squares reduced *R* from 0·359 to 0·336, at which stage three solvent atoms with high U_{iso} (>0·62 $Å^2$) were removed. All other solvent atoms were well behaved. After one more cycle *R* was 0·31 and the r.m.s. shift 0·034 Å. Electron density maps were again calculated and analysed on the computer graphics. Region *A* (residues 17–24) was now interpretable with reasonable main-chain density and, although in areas of fragmented density, the side-chain atoms could also be fitted; the main-chain and side-chain atoms of residues 15 and 16 also required manual adjustment. The electron density for the main chain in region *B*, residues 34–39, was also clearer than in previous maps, but the side chains of Leu-35, Lys-37 and Arg-39 were not apparent. The region *B* main chain was then built into the density. Side chains of the poorly defined groups were built in stereochemically feasible positions, avoiding short intermolecular contacts, as best as possible in the available density. Oxygens of the SO_4^{2-} ion were inserted tetrahedrally at 1·5 Å from the refined S position, orientated in such a way as to maximize the possibility of hydrogen-bond formation with neighbouring groups. This procedure was necessary in view of the absence of resolved O positions in the SO_4^{2-} density.

Three new features of interest were noticed in the electron density at this stage: (i) There was a peak at the N-terminal end of the main chain 2·5 Å from the N-peptide of Lys-1, resembling a possible solvent atom. (ii) Lys-41 was poorly defined beyond Cγ. The side chain was therefore rebuilt as best as possible in the available density. (iii) There was a region of significant density in the vicinity of His-119. Since the centroid of this density close to His-119 was only 2·4 Å from His-119 Cβ it could not be considered as a possible solvent peak. It seemed likely that the possibility of an alternative orientation of the imidazole ring of His-119 would have to be investigated. In fact Wyckoff *et al.* (1970) reported that the ring of His-119 in ribonuclease-S was poorly defined in their 2 Å electron density map and postulated statistical occupancy of two sites for this side group.

Pass 4

In view of the importance of Lys-41 and His-119 for the catalytic activity, a fourth refinement pass was undertaken. Six cycles of refinement were carried out and in the last two cycles the side groups of Lys-41 and His-119 were omitted beyond the β carbons. Electron density maps were recalculated and inspected as before on the graphics. Lys-41 was rebuilt in the new density which was quite clear except for the terminal N atom. The previously located imidazole ring of His-119 (subsequently designated His-119*A*) appeared very clearly in the new density. Additional density as noted under pass 3 also reappeared and an alternative position (His-119*B*) was found to fit. The His-119*A* and *B* side groups are related by rotation about the Cα–Cβ and Cβ–Cγ bonds of 142 and 38° respectively. It was also possible to rebuild the Lys-41 side chain, starting at Cγ. The new Lys-41 Nζ position was shifted about 5 Å from its previous position, with corresponding but smaller shifts for the other side-chain atoms.

Pass 5

The present phase of the refinement was completed by undertaking five further cycles of refinement with occupancy refinement for the His-119*A* group and including an occupancy factor for His-119*B* such that the sum of the two occupancies = 1·0 (Table 1). Further refinement is planned, as explained previously, and the analysis still proceeds with checking of such items as further solvent sites.

Assessment of errors

The estimation of errors from the least-squares refinement is beset by these problems:

(*a*) The full normal matrix is not calculated and inserted.

(*b*) Certain regions of the structure show appreciable static or dynamic disorder and there remains the possibility that the atoms are in a local minimum.

(*c*) The use of restraints biases the atomic positions to the target values which in some cases may be significantly in error.

The assessment of errors in the positions of the water O atoms is easiest because these atoms were not restrained. They are also further than 2 Å from other atoms and thus the off-diagonal terms in the normal matrix between adjacent atoms are smaller enabling an estimate of the standard deviation to be obtained from the inverse of the diagonal terms of the normal matrix. Standard deviations of coordinates obtained in this way may vary from 0·08 to 0·13 Å for oxygens with U_{iso} varying from 0·1 to 0·6 Å^2 respectively.

Estimates of the standard deviations obtained in the same way for restrained atoms yield figures ranging from 0·004 to 0·05 Å for atoms with U_{iso} ranging from 0·04 to 0·8 Å^2. However, these estimates do not necessarily indicate the likely errors in the atomic coordinates, for reasons stated earlier. Errors in the target values could give bias of up to 0·05 Å. In the badly disordered region from residues 35 to 39 errors of 2 Å could be present in the coordinates of atoms at the end of the side chains. However, in well defined regions where restraints have been employed, the errors in the atomic positions are likely to be smaller than for the solvent O atoms

Discussion of the structure at 1·45 Å resolution

Refinement of the structure of ribonuclease-A at 1·45 Å resolution has provided a detailed and precise model of this classical enzyme molecule. The geometry of the active site is now accurately known and this information will facilitate the studies now being undertaken to investigate the precise mechanism of the catalytic activity. Previous studies of ribonuclease-A

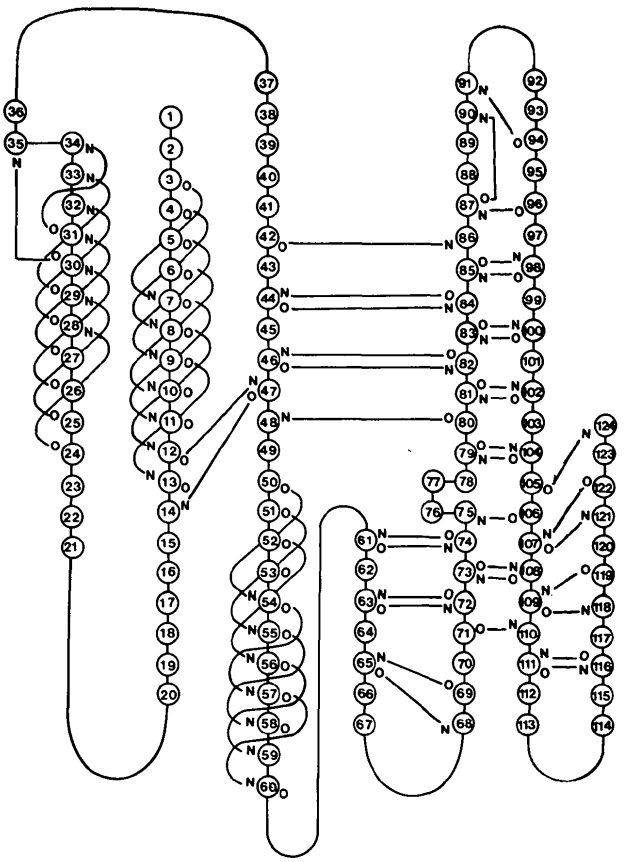

Fig. 1. Revised intramolecular hydrogen-bonding scheme in ribonuclease-A.

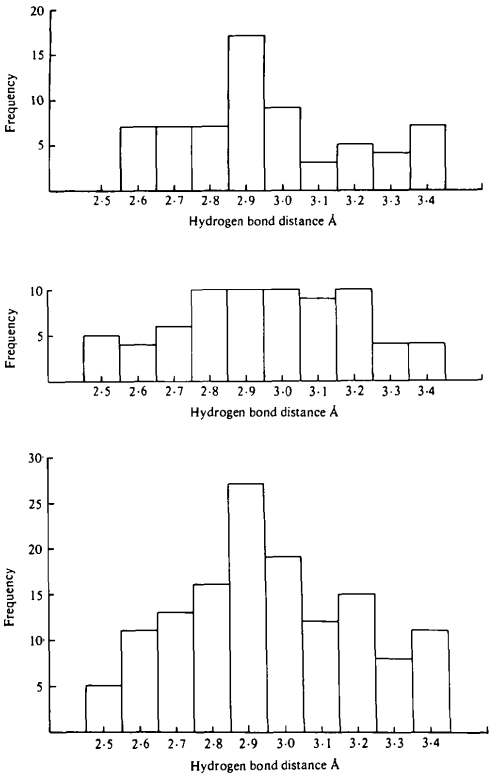

Fig. 2. Histograms showing the distribution of hydrogen-bond distances in ribonuclease-A determined in the 1·45 Å analysis: (*a*) involving solvent and protein atoms, (*b*) involving main-chain atoms only, (*c*) composite of (*a*) and (*b*). [The distance criterion for hydrogen bonds was 2·5 to 3·5 Å. In (*b*) the angle at the carboxyl oxygen was found to be in the range 120 to 180°.]

(Kartha *et al.*, 1967; Carlisle *et al.*, 1974; Wlodawer, 1980) and ribonuclease-S (Wyckoff *et al.*, 1970) have all contributed to the present refined model. We now present some details of the structure, including amendments where necessary to the previous models, although these will not always be specified as such in the discussion.

The ribonuclease molecule is U-shaped with approximate dimensions 35 × 45 × 31 Å. The backbone chain (Figs. 1 and 2) contains three sections of helix and an extensive region of β-pleated sheet consisting of three antiparallel chains. The most striking region of helix consists of some two to three turns involving residues 3 to 13 close to the N terminus of the polypeptide chain. This region defines one face of the cleft formed between the two arms of the U-shaped molecule. A second stretch of helix runs from Asn-24 to Arg-33 including one turn of 3_{10} helix involving a hydrogen bond between the carbonyl oxygen of residue 31 and the peptide N of residue 34. The third region of helix, about the same length as the other two, lies between residues 50–60. There are seven hydrogen bonds assigned to this stretch; three involving Ser-50 to Ala-56 form about one turn of α helix; the remaining four hydrogen bonds form a short section of 3_{10} helix (Figs. 1 and 2).

The three chains of antiparallel β sheet are formed by residues 42 to 48, 71 to 92 and 94 to 110. The *cis* Pro-93 and Pro-114 (Wlodawer, 1980) are involved in reversal of chain direction. The ω torsion angles are 19° for Pro-93 and 4° for Pro-114. For the 122 *trans* peptide groups the ω torsion angle distribution has a mean value of 180° with a standard deviation of 9°. The poor definition of the residues in region *B* (Leu-35 to Arg-39) is probably a feature of the ribonuclease-A structure. The groups involved being apparently non-functional and located near the surface of the molecule in a relatively uncongested region some way from the active site are simply disordered. It is of interest to note that region *B* is within the hypervariable segment of ribonuclease (Barnard, Cohen, Gold & Kim, 1972) which exhibits an unusual diversity of sequence between species.

The main-chain hydrogen-bonding scheme of ribonuclease-A (Fig. 1) differs from that of ribonuclease-S (Wyckoff *et al.*, 1970) near the site of subtilisin cleavage Ala-20...Ser-21. In ribonuclease-A the distance from the peptide N of Ser-16 to the carbonyl O of His-48 is 7·37 Å and the corresponding distance between Ala-20 and Ser-15 is 10·3 Å. Both of these distances are less than 3·2 Å in ribonuclease-S. In the region between residues 14–21 a least-squares fit between ribonuclease-A and ribonuclease-S shows an r.m.s. difference of about 7 Å in the coordinates of the main-chain atoms of the two molecules. Fig. 2 shows histograms of the observed hydrogen-bond distances in ribonuclease-A.

The active site

Since the X-ray analysis cannot distinguish between C and N, both His-12 and His-119 may require rotation through 180° about the Cβ–Cγ bonds in order to bring the imidazole rings into their true orientations. In fact examination of the active-site cleft region on the computer graphics does indicate that better hydrogen-bonding environments are available for the imidazole nitrogens of both rings upon such rotation (Fig. 5). It is hoped that these ambiguities may be resolved in the light of neutron diffraction studies (Wlodawer, 1980) and our own libration analysis of the SO_4^{2-} anion. The situation with respect to His-119 is further complicated by partial occupation of two alternative locations. These side-chain torsion angles corresponding to the two positions of His-119 are: χ_1(N–Cα–Cβ–Cγ) = 149° and χ_2(Cα–Cβ–Cγ–Cδ) = −101° for His-119*A* and $\chi_1 = -69°$, $\chi_2 = -63°$ for His-119*B*. The relative occupancies of the two sites are 0·80 and 0·20 respectively (Fig. 3). Average U_{iso} values for side-chain atoms are: His-119*A* 0·243, His-119*B* 0·418 Å^2 compared to 0·123 Å^2 for His-12.

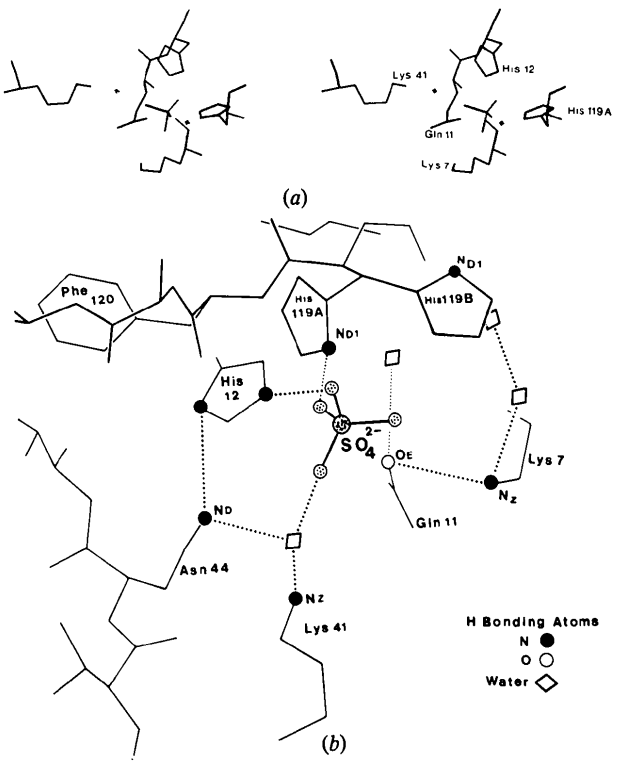

Fig. 3. (*a*) Stereoplot showing active-site geometry. (*b*) Hydrogen bonding within the active site. The absolute orientations of the imidazole ring of His-119 are not known with certainty but those shown are more favourable with respect to potential hydrogen-bond formation.

Two water molecules hydrogen bonded to Gln-11 and Lys-41 are located in the active-site cleft. Inhibitor studies (in progress) may indicate whether these solvent molecules are displaced by or play any part in substrate binding.

The role of Lys-41 in the catalytic mechanism of ribonuclease-S has been examined by Holmes *et al.* (1978). As in ribonuclease-S the ε-amino group of Lys-41 is not in the immediate vicinity of His-12 and His-119 and in ribonuclease-A is at a distance of 6·8 Å from the SO_4^{2-} S atom. Holmes *et al.* (1978) have suggested that the involvement of Lys-41 is likely to be in the stabilization of a cyclic pentacoordinate phosphate intermediate. Their proposed mechanism requires considerable movement of the phosphorus atom, the Lys-41 side chain and His-119 imidazole. Examination of the accessible volume in our model indicates that ribonuclease-A can accommodate such a mechanism, but it should be recalled that ribonuclease-S is not fully active and its precise mechanism may differ from that of ribonuclease-A. At present it is unclear as to whether Lys-7 plays any part in the enzyme activity. However, it is of interest to note that Nζ of Lys-7, which is about 5 Å from the SO_4^{2-} S position and hydrogen bonded to a water molecule, is well situated in the active site for

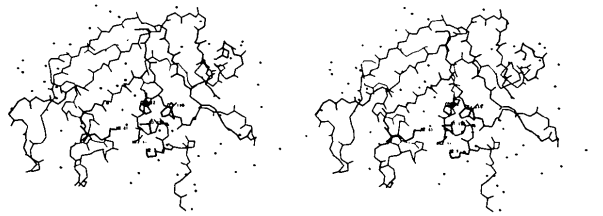

Fig. 4. Stereoview of the refined ribonuclease-A structure showing the 79 solvent molecules located in the 1·45 Å analysis. Side chains are shown only for the residues involved in the catalytic mechanism.

interaction with the substrate. Fig. 4 is a stereoview showing the polypeptide chain, active-site side groups and the 79 ordered solvent molecules located in the 1·45 Å analysis. Some important features of the active-site geometry are listed below; *H*12 and *H*119*A* are the centres of the His-12 imidazole and His-119*A* imidazole rings respectively:

*H*12···Lys-41(Nζ)	6·02 Å	Lys-41(Nζ)···*H*119*A*	1·24
*H*12···*H*119*A*	7·96	Lys-41···SO_4^{2-}(S)	6·58
*H*12···SO_4^{2-}(S)	4·54	*H*119*A*···SO_4^{2-}(S)	4·95

The dihedral angle between imidazole rings His-12···His-119*A* = 96°. The active-site geometry is depicted in Fig. 3. Fig. 5 shows the electron density in the region of His-119 indicating the alternative sites for the imidazole rings *A* (occupancy 0·80) and *B* (occupancy 0·20). The average value of U_{iso} for the 1·45 Å refinement is 0·15 Å^2 for the main-chain and about 0·35 Å^2 for side-chain atoms. The crystal packing is such that each ribonuclease molecule is in contact with six neighbours. Intermolecular hydrogen bonds are listed in Table 2.

Table 2. *Short intermolecular contacts in ribonuclease-A*

Molecule 1	Molecule 2	Distance (Å)	Symmetry relationship to generate molecule 2
Lys-1 N peptide	Tyr-92 Oε	2·72	2_1 axis ($\frac{1}{2}$,y,0)
Ser-15 Carbonyl O	Ser-23 N peptide	2·90	2_1 axis (1,y,$\frac{1}{2}$)
Ser-18 N peptide	Ser-23 Oγ	2·97	2_1 axis (1,y,$\frac{1}{2}$)
Asn-34 Nδ	Lys-66 Carbonyl O	3·36	Translation $-a$
Asn-67 Carbonyl O	Lys-91 Nζ	3·02	2_1 axis (1,y,0)
Thr-70 N peptide	Asp-38 Carbonyl O	2·98	2_1 axis (1,y,0)
Thr-87 Carbonyl O	Pro-114 Cγ	3·25	Translation $-b$

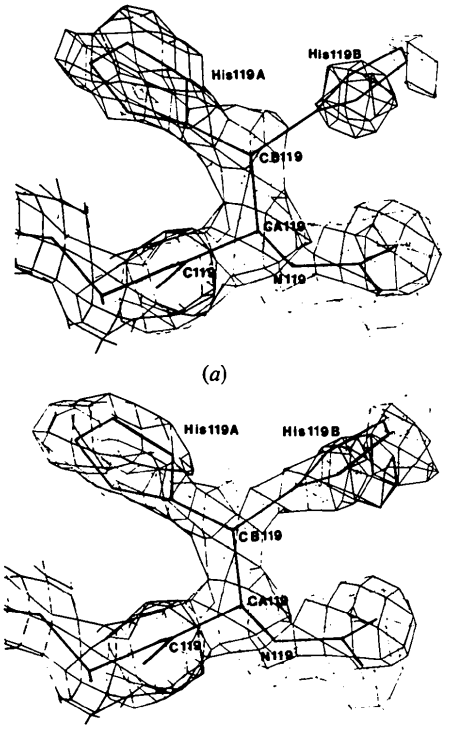

Fig. 5. (*a*) Electron density ($2|F_o| - |F_c|$) in the region of His-119 showing the alternative imidazole ring positions *A* and *B*. (*a*) Atoms for His-119 were included only up to Cβ in the F_c calculation. (*b*) As (*a*) but including in addition imidazole rings *A* and *B* with site occupancies of 0·80 for site *A* and 0·20 for site *B*. The two imidazole ring positions are related by rotation about bonds Cα–Cβ and Cβ–Cγ of 142 and 38° respectively. Contours are at the 0·6 e Å^{-3} level.

Conclusion

The combination of restrained least squares and interactive computer graphics for refinement, assessment and rebuilding of the protein model has provided an efficient strategy for optimizing the structure of ribonuclease-A. However, the effort spent in graphics work amounted to about 60 hours of one person's time. Most of this time was spent in locating the solvent molecules and fitting the side chains in the few regions of the map where the electron density was badly fragmented. The accuracy of the resulting molecular model is higher than could have been achieved by mechanical model building in a comparable or even longer period of time. The transfer of atomic coordinates and Fourier maps between the in-house PDP 11/60 and the IBM 370 at Daresbury Laboratory has been by posting magnetic tapes. This was often the rate-determining step in the work and it showed the desirability of either carrying out the graphic and refinement work on the same computer or having a wide bandwidth link between the two machines.

We thank Dr Alexander Wlodawer for fruitful discussions, the North Atlantic Treaty Organization for grant RG 040.80 and the SERC for providing support for NB.

References

BAKER, E. N. (1980) *J. Mol. Biol.* **141**, 441–484.
BARNARD, E. A., COHEN, M. S., GOLD, M. H. & KIM, J. (1972). *Nature (London)*, **240**, 395–398.
CARLISLE, C. H., PALMER, R. A., MAZUMDAR, S. I., GORINSKY, B. A. & YEATES, D. G. R. (1974). *J. Mol. Biol.* **85**, 1–18.
CHEN, M. C. & LORD, R. C. (1976). *Biochemistry*, **15**, 1889–1897.
GIRLING, R. L., HOUSTON, T. E., SCHMIDT, W. C. & AMMA, E. L. (1980). *Acta Cryst.* A**36**, 43–50.
HAER, W., MAURER, W. & RÜTERJANS, H. (1976). *Eur. J. Biochem.* **44**, 201–211.
HENDRICKSON, W. A. & KONNERT, J. H. (1980). *Computing in Crystallography*, edited by R. S. DIAMOND, S. RAMASESHAN & K. VENKATESAN, Ch. 13, pp. 13.01–13.25. Bangalore: Indian Academy of Sciences.
HERMANS, J. & MCQUEEN, J. D. (1974). *Acta Cryst.* A**30**, 730–739.
HOLMES, R. R., DEITERS, J. A. & GALLUCI, J. C. (1978). *J. Am. Chem. Soc.* **100**, 7394–7402.
HUESTIS, W. H. & RAFTERY, M. S. (1971). *Biochemistry*, **10**, 1181–1186.
International Tables for X-ray Crystallography (1974). Vol. IV. Birmingham: Kynoch Press.
JONES, T. A. (1978). *J. Appl. Cryst.* **11**, 268–272.
KARTHA, G., BELLO, J. & HARKER, D. (1967). *Nature (London)*, **213**, 862–865.
KONNERT, J. H. (1976). *Acta Cryst.* A**32**, 614–617.
MITSUI, Y., URATA, Y., TORII, K. & IRIE, M. (1978). *Biochim. Biophys. Acta*, **535**, 299–308.
MORFFEW, A. & MOSS, D. S. (1982). *Comput. Chem.* **6**(1), 1–3.
NORTH, A. C. T., PHILLIPS, D. C. & MATHEWS, F. S. (1968). *Acta Cryst.* A**24**, 351–359.
PARES, X., ARUS, C., LLORENS, R. & CUCHILLO, C. (1978). *Biochem. J.* **175**, 21–27.
POWERS, T. B. (1976). PhD Thesis, Yale Univ.
ROLLETT, J. S. (1965). *Computing Methods in Crystallography*. Oxford: Pergamon Press.
SAROFF, H. A. & CARROLL, W. R. (1962). *J. Biol. Chem.* **237**, 3384–3387.
SIELECKI, A. R., HENDRICKSON, W. A., BROUGHTON, C. G., DELBAERE, L. T. J., BRAYER, G. D. & JAMES, M. N. G. (1979). *J. Mol. Biol.* **134**, 781–804.
TICKLE, I. J. (1979). Personal communication.
WELLING, G. W., LENSTRA, J. A. & BEINTEMA, J. J. (1976). *FEBS Lett.* **63**, 89–94.
WLODAWER, A. (1980). *Acta Cryst.* B**36**, 1826–1831.
WODAK, S. Y., LIU, M. Y. & WYCKOFF, H. W. (1977). *J. Mol. Biol.* **116**, 855–875.
WYCKOFF, H. W., TSERNOGLOU, D., HANSON, A. W., KNOX, J. R., LEE, B. & RICHARDS, F. M. (1970). *J. Biol. Chem.* **245**, 305–328.