Fung, B. K.-K., Hurley, J. B., & Stryer, L. (1981) *Proc. Natl. Acad. Sci. U.S.A. 78*, 152–156.
Gaub, H., Buschl, R., Ringsdorf, H., & Sackmann, E. (1985) *Chem. Phys. Lipids 37*, 19–43.
Gros, L., Ringsdorf, H., & Schupp, H. (1981) *Angew. Chem., Int. Ed. Engl. 20*, 305–325.
Hargrave, P., & Fong, S. L. (1977) *J. Supramol. Struct. 6*, 99.
Hargrave, P., Adamus, G., Arendt, A., McDowell, J. H., Wang, J., Szary, A., Curtis, D., & Jackson, R. (1986) *Exp. Eye Res. 42*, 363–373.
Hong, K., & Hubbell, W. L. (1973) *Biochemistry 12*, 4517–4523.
Hub, H. H., Hupfer, B., Koch, H., & Ringsdorf, H. (1980) *Angew. Chem., Int. Ed. Engl. 19*, 938–940.
Hubbell, W. L. (1975) *Acc. Chem. Res. 8*, 85–91.
Jackson, M. L., Schmidt. C. F., Lichtenberg, D., Litman, B. J., & Albert, A. D. (1982) *Biochemistry 21*, 4576–4582.
Johnston, D. S., Sanghera, S., Pons, M., & Chapman, D. (1980) *Biochim. Biophys. Acta 602*, 57–69.
Kuhn, H. (1982) *Methods Enzymol. 81*, 556–564.
Laemmli, U. K. (1970) *Nature (London) 227*, 680–682.
Liebman, P. A., & Evanczsuk, A. T. (1982) *Methods Enzymol. 81*, 532–542.
Litman, B. J. (1982) *Methods Enzymol. 81*, 150–154.
O'Brien, D. F. (1982) *Methods Enzymol. 81*, 378–384.
O'Brien, D. F., Costa, L. F., & Ott, R. A. (1977) *Biochemistry 16*, 1295–1303.
O'Brien, D. F., Whitesides, T. H., & Klingbiel, R. T. (1981) *J. Polym. Sci., Polym. Lett. Ed. 19*, 95–101.
O'Brien, D. F., Klingbiel, R. T., Specht, D. P., & Tyminski, P. N. (1986) *Ann. N.Y. Acad. Sci. 446*, 282–295.
Regen, S. L., Czech, B., & Singh, A. (1980) *J. Am. Chem. Soc. 102*, 6638–6639.
Regen, S. L., Singh, A., Oehme, G., & Singh, M. (1982) *J. Am. Chem. Soc. 104*, 791–795.
Ringsdorf, H., & Schupp, H. (1981) *J. Macromol. Sci., Chem. A15*, 1015–1026.
Scotto, A., & Zakim, D. (1985) *Biochemistry 24*, 4066–4075.
Stubbs, G. W., & Litman, B. J. (1978) *Biochemistry 17*, 215–219.
Tieke, B. (1984) *J. Polym. Sci., Polym. Chem. Ed. 22*, 391–406.
Tieke, B., & Wegner, G. (1981) *Makromol. Chem., Rapid Commun. 2*, 543.
Tieke, B., & Chapuis, G. (1984) *J. Polym. Sci., Polym. Chem. Ed. 22*, 2895–2921.
Tsuchida, E., Seki, N., & Ohno, H. (1986) *Makromol. Chem. 187*, 1351–1359.
Tundo, P., Kippenberger, D. J., Klahn, P. L., Prieto, N. E., Jao, T. C., & Fendler, J. (1982) *J. Am. Chem. Soc. 104*, 456–461.
Tyminski, P. N., & O'Brien, D. F. (1984) *Biochemistry 23*, 3986–3993.
Tyminski, P. N., Latimer, L. H., & O'Brien, D. F. (1985) *J. Am. Chem. Soc. 107*, 7769–7770.
Yau, K.-W., & Nakatani, K. (1985) *Biophys. J. 47*, 356a.
Yee, R., & Liebman, P. A. (1978) *J. Biol. Chem. 253*, 8902–8909.

Structure of Phosphate-Free Ribonuclease A Refined at 1.26 Å

Alexander Wlodawer*
Center for Chemical Physics, National Bureau of Standards, Gaithersburg, Maryland 20899
L. Anders Svensson and Lennart Sjölin
Department of Inorganic Chemistry, Chalmers Institute of Technology, S-41294 Göteborg, Sweden
Gary L. Gilliland
Center for Chemical Physics, National Bureau of Standards, Gaithersburg, Maryland 20899
Received September 15, 1987; Revised Manuscript Received December 18, 1987

ABSTRACT: The structure of phosphate-free bovine ribonuclease A has been refined at 1.26-Å resolution by a restrained least-squares procedure to a final R factor of 0.15. X-ray diffraction data were collected with an electronic position-sensitive detector. The final model consists of all atoms in the polypeptide chain including hydrogens, 188 water sites with full or partial occupancy, and a single molecule of 2-methyl-2-propanol. Thirteen side chains were modeled with two alternate conformations. Major changes to the active site include the addition of two waters in the phosphate-binding pocket, disordering of Gln-11, and tilting of the imidazole ring of His-119. The structure of the protein and of the associated solvent was extensively compared with three other high-resolution, refined structures of this enzyme.

Refinement of a protein structure appears to be a never-ending task, since it is not as yet possible to describe the structure sufficiently well to ensure complete agreement of the model with the diffraction data. Even for the best refined protein structures, and final crystallographic R factors ($R = \sum|F_{obsd} - F_{calcd}|/\sum F_{obsd}$) are over 10% (for example, for crambin 11.1% at 0.945-Å resolution—Hendrickson, personal communication, 1986), although the quality of experimental data should allow the final numbers to be only half as large. Only a few proteins have been refined at high resolution with

* Address correspondence to this author at the Crystallography Laboratory, BRI-FCRF, National Cancer Institute, P.O. Box B, Frederick, MD 21701.

0006-2960/88/0427-2705$01.50/0

the final R factors on the order of 11–15%. Examples include insulin (Sakabe et al., 1981) and several proteases (James et al., 1980; Fujinaga et al., 1982; Read et al., 1983; Delbaere & Brayer, 1985), but in most cases the refinement is considered to be complete if the R factor is lower than about 20% and the difference Fourier maps do not show any more interpretable features.

Recent advances in the methods of data collection, in particular the use of synchrotron radiation sources and/or electronic area detectors, enable accurate measurement of diffraction intensities at very high resolution even though such data could not have been easily collected by conventional diffractometry. New methods of structure refinement utilizing fast Fourier techniques (Ten Eyck, 1973; Agarwal, 1978; Finzel, 1987) and the availability of interactive computer graphics and the appropriate software (Jones, 1978) made it possible to significantly decrease the amount of time needed to complete a structure determination. Finally, the development of the field of protein engineering introduced new requirements of the accuracy of protein structures necessary to interpret the changes introduced by point mutations. The movement of groups of atoms by a few tenths of an angstrom might be interpreted as being due to the substitution, while in reality such changes are sometimes artifacts due to errors in the data and overinterpretation of the refinement. Knowledge of the expected error level is very necessary.

For the reasons discussed above, a number of protein structures which were originally solved long ago have been recently reinterpreted and refined at high resolution. Examples include bovine pancreatic trypsin inhibitor (Wlodawer et al., 1987b), (carbonmonoxy)myoglobin (Kuriyan et al., 1986), and γ-chymotrypsin (Blevins & Tulinsky, 1985).

Bovine ribonuclease has played a very special role in the last 50 years as a model protein and was used in a large number of pioneering studies. Methods such as protein sequencing (Smyth et al., 1963) and protein NMR (Saunders et al., 1957) were developed with RNase as a test case. RNase was crystallized early (Kunitz, 1940), and the crystals were shown to diffract (Fankuchen, 1941). It was only the third enzyme whose structure was solved by X-ray diffraction. The native structure (RNase A) was solved at 5.5-Å resolution by Avey et al. (1967) and at 2.0 Å by Kartha et al. (1967). A proteolytic modification, RNase S, was solved by Wyckoff et al. (1967, 1970). Altogether, at least 12 different laboratories have published crystallographic studies of RNase, its modifications, and its complexes. Diffraction studies of RNase were recently reviewed by Wlodawer (1985).

Despite the large number of published structures of RNase, many questions were not fully answered. Even the best published models either were based on data at a relatively modest resolution of 2 Å (Wlodawer & Sjölin, 1983) or had a relatively high R factor of 22.3% with 1.45-Å data (Borkakoti et al., 1982). Comparison of the two above-mentioned models (Wlodawer et al., 1986) showed the presence of a number of discrepancies which were not resolved and which appeared to be due to the differences in interpretation, in addition to differences which may have been due to the variation in crystallization conditions. The structure of RNase has been apparently refined at high resolution with data collected at low temperatures (Dewan & Tilton, 1987), but the results of that work have not been reported as yet. It was clear that an improved structure of this important enzyme was needed to address these problems.

The success in collecting high-quality diffraction data with the area detector system at Genex (Howard et al., 1987) convinced us that the time has come for yet another look at the structure of RNase A. Our goal was to collect data to the limit of resolution imposed by the disorder within the crystals and to carry out the refinement as far as possible (with the limitation that only isotropic temperature factors would be utilized). Among our specific aims were the following:

(a) Recognition and modeling of side chains with discrete multiple conformations. This was a matter of some controversy, since Borkakoti et al. (1982) reported two distinct positions for the active site residue His-119, while only a single conformer was found by Wlodawer and Sjölin (1983). No other amino acid side chains with discrete multiple conformations were located in the previous studies.

(b) Location of bound waters. Only half of the solvent positions reported in the two refinements of RNase were common (Wlodawer et al., 1986), and it was likely that considerable errors in the description of solvent were still present.

(c) Location of alcohol used for crystallization (2-methyl-2-propanol in this study). Results reported for other proteins crystallized by precipitation with alcohols (Lehmann et al., 1985; Almassy et al., 1983; Walter et al., 1982) suggested that a few alcohol molecules should have been found near the protein surface, but none were reported so far.

(d) Subtle movements of the residues in the active site due to the presence or absence of bound anions. All previously reported structures of RNase A included a bound sulfate (or phosphate) in the active site, while Weber et al. (1985) reported a small change in the position of the catalytically important residue, His-119, in RNase X, in which the side chains of Lys-41 and Lys-7 were cross-linked. It was postulated that the shift could have been caused by the absence of an anion in the active site.

The results of the investigation of the crystal structure of phosphate-free RNase A at 1.26-Å resolution are reported here.

Experimental Procedures[1]

Crystallization. Crystals used in this 1.26-Å structure determination were grown from a phosphate-free form of bovine pancreatic ribonuclease, type XII-A, purchased from Sigma. The crystals were isomorphous with previously reported crystals containing phosphate, or perhaps sulfate (Borkakoti et al., 1982; Wlodawer & Sjölin, 1983). They belong to space group $P2_1$ with a = 30.18 Å, b = 38.4 Å, c = 53.32 Å, and β = 105.85°. This enzyme was crystallized according to a method similar to the one described by us previously (Wlodawer & Sjölin, 1983). Lyophilized protein was dissolved in distilled water at a concentration of 50 mg/mL, and the pH was adjusted to 5.3 by the addition of 0.3 M sodium hydroxide. A total of 1 mL of this solution was added to a 2-dram Kimax glass vial, and 2-methyl-2-propanol was slowly added to the final concentration of 43% (v/v). These vials were seeded with 1 μL of a microseeding solution which had been found to produce only one or two crystals per vial. The microseeding solution was produced by serial dilution of synthetic mother liquor in which an RNase A crystal had been crushed. At room temperature crystals appeared in a few days as thick plates and grew to their final size after several months. One crystal of a total volume of over 3 mm^3 was cut into smaller pieces, and two of these fragments, each

[1] Certain commercial equipment, instruments, and materials are identified in this paper in order to specify the experimental procedure. Such identification does not imply recommendation or endorsement by the National Bureau of Standards, nor does it imply that the materials or equipment identified are necessarily the best available for the purpose.

Table I: Statistical Summary of Crystallographic Data for Phosphate-Free Ribonuclease A at 1.26 Å

shell lower limit (Å)	av reflection, I	av $I/\sigma(I)$	no. of Bragg reflections			R_w[a]	R_{uw}[b]
			possible	collected	missing		
2.29	6592	153.7	5357	5355	2	4.55	3.91
1.82	2020	61.6	5297	5204	93	6.51	5.85
1.59	678	22.1	5242	4516	726	9.39	9.05
1.44	353	10.6	5258	4305	953	14.18	15.35
1.34	199	5.2	5247	3943	1304	20.95	23.57
1.26	139	2.9	4546	2273	2273	26.74	28.54
totals:	2020	51.2	30947	25732	5351	5.12	4.95

[a] R_w is the weighted-squared R factor on intensity times 100. [b] R_{uw} is the unweighted absolute value R factor on intensity times 100.

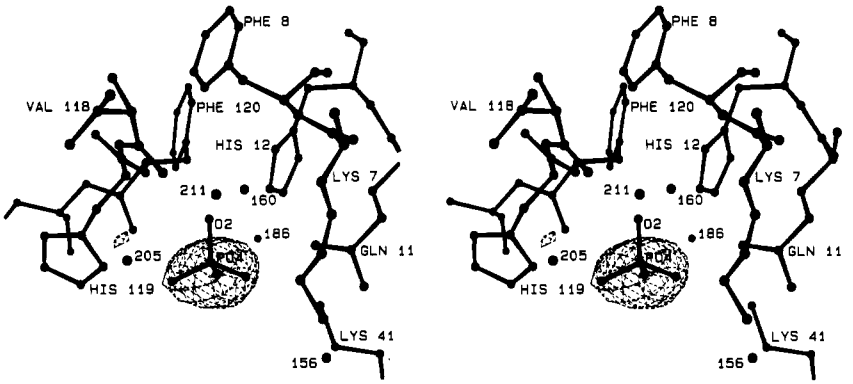

FIGURE 1: Stereoscopic view of the difference electron density in the active site region. The map was calculated with the coefficients $(F_A - F_G)\alpha_A$, where the subscripts A and G refer to the 2-Å data of Wlodawer and Sjölin (1983) and data collected here, respectively. The contour level is -4σ. The coordinates are those from the model of Wlodawer and Sjölin (1983), used to initiate the work described here.

about 0.5 mm^3 in volume, were used for data collection. Crystals were mounted in thin-walled glass capillaries in the usual way.

X-ray Data Collection. X-ray diffraction data were collected by a Nicolet imaging proportional counter (IPC), an electronic area detector, mounted on a Supper oscillation camera controlled by a Cadmus 9000 microcomputer. During data collection the area detector chamber was mounted 10 cm from the crystal. The carriage angle was varied from 0° to 50°, enabling the detector to intercept data from ∞ to 1.26 Å, depending on its position. Diffraction data collected by the IPC detector are recorded as a series of discrete frames or electronic images, each comprising a 0.25° oscillation counted for between 30 and 110 s, depending upon the carriage angle. The individual frames were contiguous in that the beginning of each small oscillation range coincided with the end of the previous range. Usually 400 data frames, corresponding to 100° of crystal rotation, were accumulated on the hard disk of the Cadmus during each run. Subsequently, the data were transferred by an Ethernet link to a Digital Equipment Corp. VAX 11/780 computer for processing. During the course of data collection, several data sets from different crystal orientations were recorded. The crystals were repositioned in the X-ray beam by adjustment of the goniometer arcs and the x, y, and z translations.

The X-ray source used to generate Cu Kα radiation was an Elliot GX-21 rotating anode, operating at 70 mA and 40 kV with a 0.3 × 3.0 mm focal spot and a 0.3-mm collimator. Monochromatization was provided by a Huber graphite monochromator. All data collection was performed at well-controlled room temperature (16–18 °C).

X-ray Data Processing. The determination of crystal orientation and the integration of reflection intensities was performed with the XENGEN program system (Howard et al., 1987). The X-ray diffraction data measured from the two crystal fragments included 129 694 observations of 25 732 unique reflections out of the 30 947 possible at 1.26 Å. The data scaled with a weighted least-squares R factor on intensity for symmetry-related observations of 0.05; 23 398 of the measured unique reflections had significant intensity [$F > 2\sigma(F)$]. A statistical summary of the final processing results of the crystallographic data are presented in Table I. Since only half of the total number of reflections present in the outermost shell were measured, the effective resolution is somewhat lower. This problem is common for data collected with area detectors, which have large blind regions.

Structure Solution. The structure of phosphate-free ribonuclease was solved by a straightforward analysis of difference Fourier maps, followed by least-squares refinement. Indeed, since no special precautions against the exposure to phosphate were taken and the commercial ribonuclease often contains some phosphate, we confirmed that we were dealing with the phosphate-free enzyme only after the interpretation of the initial difference Fourier map. This initial map was calculated using as coefficients the differences between 2-Å X-ray data from the previous refinement (Wlodawer & Sjölin, 1983) and the data collected as described above. These two data sets scaled with an R factor of 0.139 for 6559 common structure amplitudes. A difference Fourier map was essentially featureless, except for a peak of -11σ density observed at the location of the phosphate ion in the original model (Figure 1). The difference density, however, did not account for the complete phosphate ion, since the O2 oxygen of the phosphate group was located outside of the difference peak, indicating that a water molecule was located near that position. The subsequent refinement confirmed this assumption.

Restrained Least-Squares Refinement. The starting model used in the structure determination was derived from the results of the 2.0-Å joint X-ray/neutron study of Wlodawer and Sjölin (1983). That structure includes hydrogens, a phosphate group bound to the active site, and 128 water molecules, but there are no amino acid residues with multiple conformations nor bound 2-methyl-2-propanol molecules. For the reasons described above, the initial model omitted the

Table II: Details of Structure Determination Differences for the 1.26-Å Phosphate-Free Ribonuclease A Model (G), the 2.0-Å Joint X-ray/Neutron Starting Model (A), the 1.4-Å Model (B), and the 2.0-Å Cross-Linked Model (X)

	model			
	G	A	B	X
crystallization conditions	43% 2-methyl-2-propanol pH 5.3	43% 2-methyl-2-propanol pH 5.3	40% ethanol pH 5.2–5.7	30% ethanol pH 8.0 50 mM Tris-HCl or imidazole
space group	$P2_1$	$P2_1$	$P2_1$	$P2_12_12_1$
unit cell parameters	a = 30.18 Å b = 38.40 Å c = 53.32 Å β = 105.85°	a = 30.18 Å b = 38.40 Å c = 53.32 Å β = 105.85°	a = 30.31 Å b = 38.40 Å c = 52.91 Å β = 105.91°	a = 37.05 Å b = 41.26 Å c = 75.64 Å
refinement procedure	Finzel, 1987	Wlodawer & Hendrickson, 1982	Morffew & Moss, 1982	Hendrickson & Konnert, 1980
refinement resolution limit (Å)	10.0–1.26	10.0–2.0	?–1.45	20.0–2.0
no. of reflections used in refinement	23 398 ($F > 2\sigma F$)	7708 (X-ray) 4132 (neutron) ($I > 3\sigma I$)	19098 ($I > 0$)	7853 ($I > \sigma I$)
crystallographic R factor	0.150	0.159 (X-ray) 0.183 (neutron)	0.223	0.184
rms deviation from				
(1) bond length (Å)	0.024	0.022	0.018	0.024
(2) angle distance (Å)	0.038	0.062	0.038	
(3) planarity (Å)	0.017	0.018		0.022
(4) ideal chirality (Å^3)	0.188	0.207		0.283
protein data bank code	7RSA	5RSA	1RN3	1RSM

phosphate group bound to the active site but included the rest of the solvent oxygen atoms, as well as hydrogens of the protein which were very tightly constrained during the subsequent refinement to effectively ride on the heavier atoms to which they were bound.

Structure refinement was performed with the restrained least-squares procedure of Hendrickson (1985), which has been modified by Finzel (1987) to incorporate the fast Fourier transform algorithms of Ten Eyck (1973) and Agarwal (1978). Program FRODO (Jones, 1978) was used on an Evans and Sutherland PS300 graphics system to examine $2F_o - F_c$ and $F_o - F_c$ difference Fourier maps and to adjust the original model, to add solvent molecules, and to position alternate side-chain conformations. During the course of the refinement, each cycle took only from 8 to 15 min on a VAX 11/780, and four to eight cycles of refinement were performed between map calculations and rebuilding.

Modeling of side chains with multiple conformations involved treating one conformation with the normal restraints, while the other was included with no restraints on the interactions with other atoms belonging to the protein, and only its own bond lengths, angles, and planar groups (if present) were restrained. Thus the two conformations do not interact with each other. Relative occupancies of the alternate side-chain models were estimated on the basis of the appearance of electron density and were adjusted in such a way that the average temperature factors of the two models were similar. This was done in rather coarse steps, as we did not feel that the extent of the data would justify the refinement of this parameter. While it is possible that more than two orientations might be present for some amino acid side chains, we did not attempt to model more than two, since this could represent overinterpretation of the available data. We did not find it necessary, for any side chain, to assume that it shared the same volume with a solvent molecule. Results of the analysis of multiple conformations were reported previously (Svensson et al., 1987) on the basis of the model after refinement cycle 86. An additional 17 cycles of refinement performed since then resulted in only minor modification of the model.

Solvent structure was analyzed in each of the difference Fourier maps, and waters with high temperature factors, with low occupancies, or not clearly visible in the $F_o - F_c$ maps were removed from the model, while new waters were added where indicated by map peaks. Near the end of refinement waters were renumbered on the basis of the occupancies and temperature factors, from highest occupancy and lowest temperature factor to the lowest occupancy and the highest temperature factor. This order was not kept exactly to the end of refinement, since we changed one putative 2-methyl-2-propanol molecule to a water after cycle 86 and relative temperatures and occupancies changed slightly, but the order of waters is still roughly descending according to their quality. The final model of RNase contains 188 water molecules with occupancies ranging from 1.0 to 0.3 and temperature factors in the range of 12–66 Å^2.

Altogether, 103 cycles of least-squares refinement were performed interspersed with 16 calculations of electron density maps followed by model adjustment. The final crystallographic R factor of the struture of RNase A resulting from this work was 0.15 at 1.26-Å resolution, with root mean square deviation of bonded distances from ideality of 0.024 Å. A summary of the refinement statistics is presented in Table II. The coordinates and structure factors have been deposited with the Brookhaven Protein Data Bank (code 7RSA).

Results

The final refined structure of phosphate-free RNase A has a number of major differences from the starting model (Table V). The phosphate group located in the active site is absent, resulting in subtle changes in the geometry of the active site with an altered solvent structure in this region. Thirteen of the amino acid residues were modeled with two alternate side-chain positions each. A 2-methyl-2-propanol molecule was also added to the bound solvent. Several side chains were positioned with conformations quite different from the starting model. Details of each of these changes are described below along with a detailed comparison of the 1.26-Å RNase A model (referred to as model G) with the 2.0-Å starting model from the joint X-ray/neutron structure determination of Wlodawer and Sjölin (1983) (model A), the 1.4-Å RNase A model refined by Borkakoti et al. (1982) (model B), and the 2.0-Å cross-linked RNase A model of Weber et al. (1985) (model X). Protein Data Bank identifications of these models are listed in Table II.

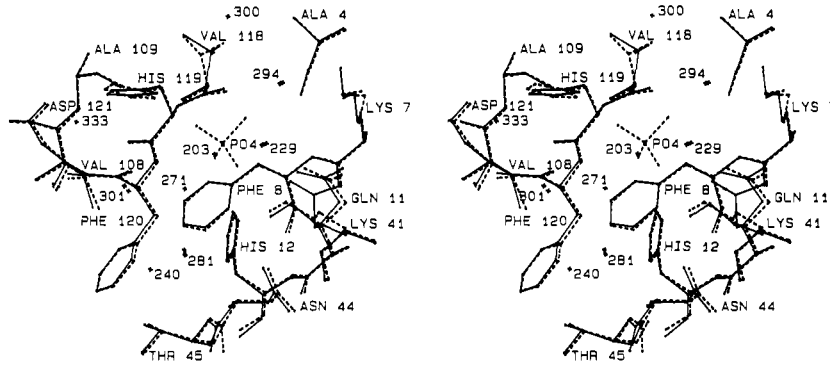

FIGURE 2: Stereoscopic view of the RNase A active site. Atoms connected by dashed lines represent the starting model (A) of Wlodawer and Sjölin (1983); atoms connected by solid lines represent the final model of phosphate-free RNase A refined at 1.26-Å resolution (model G). Waters are numbered as in model G.

Active Site. A comparison of the active sites of structures G and A has revealed overall similarity, but with several significant changes (Figure 2). The His-12 side chain atoms have an rms difference of 0.094 Å between the starting and final models indicating virtually no change in position in the absence of the phosphate. His-119, however, does move (rms difference of 0.230) by a χ_2 rotation of −12° which takes the plane of the histidine ring away from the location of the phosphate ion binding site. In the presence of phosphate a hydrogen bond is formed between ND1 of His-119 and O1 of the phosphate group, an interaction impossible in the absence of the anion. The tilt of the histidine ring allows a more favorable hydrogen-bonding geometry between NE2 of the histidine and OD1 of Asp-121. The Asp-121 side chain adjusts only slightly (rms difference of 0.129) to accommodate the altered hydrogen bond. The side chains of Lys-7 and Lys-41 are oriented in the same directions as in the starting model. The Lys-7 side-chain atoms deviate more than those of Lys-41 with an rms distance between models for these atoms of 0.68 and 0.36 Å, respectively. In both structures the NZ of Lys-41 forms a hydrogen bond to OD1 of Asn-44; however, the NZ of Lys-7 has close interactions neither with other residues nor with the solvent.

Gln-11 is the only residue in the region of the phosphate ion binding site which was modeled with multiple conformations (see below). This residue did not show any indication of multiple conformations in the starting model, while it is clearly found in two, closely related, conformations in the phosphate-free enzyme. In this case it is highly probable that the multiple conformers are induced by the removal of the phosphate from the active site, since the side-chain amide group is no longer able to make a hydrogen bond to an oxygen of the phosphate.

As expected, the active site has an altered solvent structure in the absence of the phosphate ion (Figure 2). In the phosphate-free form of the enzyme two additional water molecules are found in the vicinity of the phosphate ion binding site. Only one of these water molecules, Wat-203, is actually found within the space occupied by the phosphate ion, near its O2 position. The second water, Wat-271, is located adjacent to the ion binding site, but it is not near enough to Wat-203 to form a hydrogen bond. Wat-271 does form hydrogen bonds with Wat-281 and Wat-301. Wat-229, Wat-281, Wat-294, and Wat-301, which correspond respectively to Wat-160, Wat-286, Wat-211, and Wat-186 of model A, maintain their positions in the absence of the phosphate ion. Wat-333 and Wat-240 are other additions to the solvent in the active site region. Wat-240 hydrogen bonds to Wat-281, and Wat-333 hydrogen bonds to OD1 and O of Asp-121. Wat-300, corresponding to Wat-205 of the starting model, is located 4.2 Å from the CB of His-119 in the location interpreted by Borkakoti et al. (1982) as a second conformer of His-119. The distance between CB and Wat-205 in A was only 3.2 Å, indicating that this might possibly be a trace of an alternative conformation of His-119, rather than a solvent molecule, but Wat-300 in G is further removed by another 1.0 Å, making such an interpretation very unlikely. We must conclude that under the conditions of our experiment only a single site is occupied by the side chain of His-119.

Multiple Conformations of Amino Acid Residues. During the course of the refinement, 13 of the amino acid residues were modeled with alternate side-chain positions. The final model, when compared to the model previously reported after refinement cycle 86 (Svensson et al., 1987), has no significant changes for those residues which were modeled with multiple conformations although slight adjustments of atomic positions have altered several of the torsion angles of the side chains and the temperature factors have been modified from those values reported earlier. The multiple conformations of the side chains did not involve unusual torsion angles, and the averaged thermal parameters of the side chains of corresponding conformers have highly correlated and reasonable values as seen in Table III.

The residues with multiple conformations are almost all either polar or charged and are distributed over the surface of the molecule. In general, all of the multiple conformer atoms of the asparagine, glutamine, aspartate, serine, and arginine side chains which are capable of hydrogen bonding are engaged in intramolecular or solvent interactions. The exceptions are ND2 of Asp-67A and the NH1 and NH2 atoms of Arg-85A, which have no detectable interactions with other protein or solvent atoms.

Four of the 13 residues modeled with multiple conformations are lysines. Surprisingly, with the exception of Lys-61, none of the conformers of these lysines make discernible hydrogen bonds to solvent, or inter- or intramolecular hydrogen bonds. Not only is Lys-61 the only one of the four which forms a salt bridge, it is also the only residue of the 13 which is involved in crystal packing interactions. One of the conformers of Lys-61 forms a salt bridge with the side chain of Glu-9 of another molecule in the crystal lattice. Since only a single conformer of the 13 amino acid residues is involved in intermolecular interactions, crystal packing does not appear to be responsible for the ordered multiple conformations. Crystal packing, on the other hand, may preferentially stabilize a particular conformation out of several which are possible, as was observed in multiple crystal forms of bovine pancreatic trypsin inhibitor (Wlodawer et al., 1987b). Therefore, several high-resolution structures of the same protein in different crystal forms, with different crystal packing arrangements,

Table III: Parameters for the Final Refined Model of Side Chains with Multiple Conformations in Phosphate-Free Ribonuclease A at 1.26-Å Resolution[a]

residue	occupancy	temp factor (Å²)	degrees χ_1	χ_2	χ_3	χ_4	χ_5
Gln-11	0.67	14.7	−59	−168	−37		
Gln-11A	0.33	13.9	−159	160	89		
Ser-32	0.67	14.4	73				
Ser-32A	0.33	11.8	−72				
Asn-34	0.33	13.8	−145	43			
Asn-34A	0.67	12.7	−55	132			
Val-43	0.40	12.5	−167				
Val-43A	0.60	13.0	−70				
Ser-50	0.67	10.8	71				
Ser-50A	0.33	11.7	164				
Lys-61	0.50	23.1	−169	171	−152	164	
Lys-61A	0.50	19.6	−176	65	−179	177	
Asn-67	0.67	14.5	79	−21			
Asn-67A	0.33	16.2	−70	−20			
Ser-77	0.67	22.8	−70				
Ser-77A	0.33	18.4	−29				
Asp-83	0.50	16.7	−110	51			
Asp-83A	0.50	16.7	178	73			
Arg-85	0.50	24.5	−149	94	59	−169	−2
Arg-85A	0.50	25.2	143	−82	150	77	4
Lys-91	0.50	22.9	175	155	−176	−74	
Lys-91A	0.50	22.3	−160	−179	179	61	
Lys-98	0.50	18.5	177	74	−167	169	
Lys-98A	0.50	18.6	179	−173	−176	−56	
Lys-104	0.50	19.1	−74	−171	172	−64	
Lys-104A	0.50	20.0	−49	−83	−166	78	

[a] Temperature factors were averaged for all non-hydrogen atoms of the side chains, separately for each conformation. Torsion angles are denoted χ_1–χ_5.

might prove useful in examining the behavior of surface residues.

Three residues, Val-43, Asp-83, and Arg-85, for which conformer positions of each residue would appear to influence the preferred conformation of the other two, are located in the active site cleft on the surface of an antiparallel β-sheet. Two primary sets of conformation can be envisioned. In one conformation Asp-83 and Arg-85 form a salt bridge, and Val-43 is oriented to avoid too close van der Waals contact with Arg-85. This corresponds to conformers 83, 85, and 43A of Table III. The other set of conformers, 83A, 85A and 43, would have Asp-83 and Arg-85 directed away from one another, hydrogen bonding to OG of Thr-45 and O of Glu-86, respectively. This would allow Val-43 to orient as either of the modeled conformers.

As mentioned by Svensson et al. (1987), multiple conformations of Val-43 are not unique to this crystal form. Two positions of Val-43 were observed during the refinement of an RNase A derivative cross-linked between Lys-7 and Lys-41. The predominant position was that of Val-43A in the current refinement. Thus, the results of two independent studies show that very similar alternate conformations are sometimes present irrespective of crystal form and even of covalent modification of the protein.

Water Structure. The final model of RNase resulting from this refinement contains 188 water molecules. They have been modeled as oxygen atoms, and unlike in the starting model (A), they do not contain hydrogens. Of that number, 134 water molecules are in the first hydration shell and make direct hydrogen bonds with the atoms belonging to one protein molecule only. A further 20 first-shell waters are intermolecular. Thirty-two waters are identified as belonging to the second layer of hydration, while the connections are unclear for two waters. Of these, Wat-278 is in a clear peak of density but does not make good connections to any other atoms and may be an artifact, while Wat-388 occupies a site which was originally modeled as a second 2-methyl-2-propanol, and we are not certain whether it is indeed a water molecule.

Ninety-three waters correspond to those found in model A (conversely, 35 water molecules present in model A were not supported by the new refinement, and many of them were most probably incorrect before). Comparison with the 88 solvents present in model X yields 44 waters closer than 1 Å apart, with an additional 8 waters clearly making similar interactions, while up to 1.35 Å from each other. Similarly, out of 78 solvent molecules present in model B, 32 are closer than 1 Å to corresponding waters in model G, with a further 5 waters less than 1.3 Å away. Comparison of the positions of water molecules with the three models of RNase shows that 13 of them are found in similar positions in all four models, while 38 are found in three models. Water molecules which are present in all four models are probably bound quite strongly, since they are practically integral parts of the protein molecule. Due to different crystal packing in form X, they are also not likely to be intermolecular. A summary showing temperature factors, occupanices, hydrogen bonds, and corresponding waters from other models is given in Table IV. As can be seen from this summary, all of the closely bound water molecules are in the first shell of hydration, usually make more than one hydrogen bond to the protein, and are, with one exception, intramolecular. Two of the 13 waters have occupancies lower than one, but in view of their low temperature factors, this appears to be an artifact of refinement.

Analysis of the solvent shows quite clearly the lack of buried water sites in ribonuclease. Such sites were reported for other proteins, such as bovine pancreatic trypsin inhibitor, in which

Table IV: Waters Common in This Refinement and in the Three Other Models of RNase A[a]

model G	model A	model B	model X	B (Å²)	occupancy	hydrogen-bond interactions
201	200	250	208	12.0	1.0	Ser-23 O–Tyr-97 O–Thr-99 OG1
204	184	212	212	13.1	1.0	Ser-50 N–Glu-49 OE2–Asp-53 OD2
206	166	218	202	14.4	1.0	Ala-5 O–Pro-117 O
208	187	216	213	15.4	1.0	Gln-60 NE2–Asp-53 O
213	167	244	255	17.1	1.0	Ala-6 O–(Wat-205, Wat-323, Wat-345)
215	172	214	265	18.9	1.0	Ala-52 N–Ser-50A OG
218	196	266	253	19.6	1.0	Glu-111 O–(Wat-337, Wat-256, Wat-228)
223	179	213	276*	23.3	1.0	Ser-23 OG–Ser-15* O–Thr-17* O
229	160	202	204	24.6	1.0	Gln-11A OE1–(Wat-203, Wat-294)
240	252	211	231	32.5	1.0	Thr-45 OE1–(Wat-281, Wat-303)
246	213	267	224	33.2	1.0	Gly-112 N–(Wat-367*)
293	190	251	207	15.6	0.9	Thr-99 N–Ser-23 O
338	138	255	209	19.2	0.7	Asn-27 ND2–Tyr-97 O

[a] Water interactions shown if only a single hydrogen bond to protein is present. An asterisk (*) indicates a symmetry-related molecule.

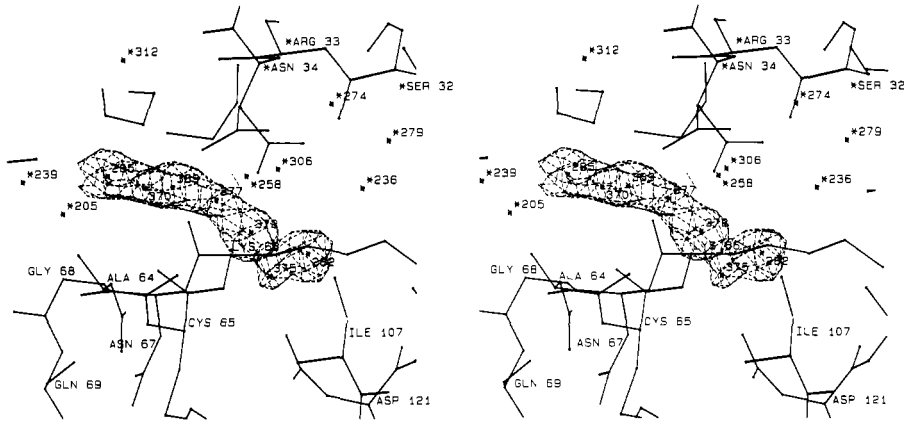

FIGURE 3: Electron density contoured at 1σ level and atomic positions of a string of eight water molecules which most likely form two interpenetrating networks with partial occupancy (see text for details).

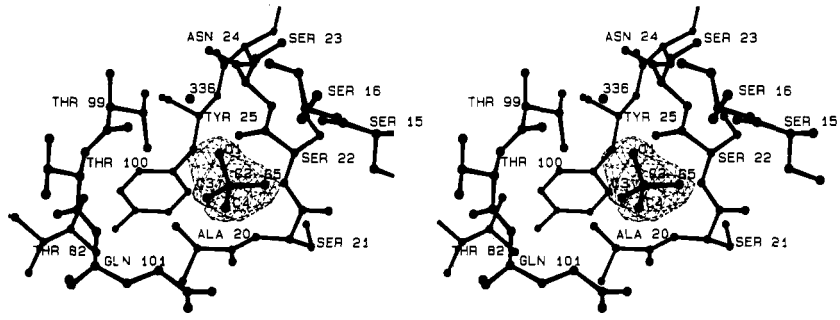

FIGURE 4: Stereoscopic view of the 2-methyl-2-propanol binding site on the surface of phosphate-free RNase A. The 2-methyl-2-propanol molecule is shown inside the electron density of the final $2F_o - F_c$ map which was contoured at 1σ. Density for the amino acid residues and for other solvent has been omitted for clarity.

three water molecules are buried in a deep pocket and do not interact with any bulk solvent regardless of the crystal packing (Wlodawer et al., 1987a,b). The only water in RNase which is in a shallow pocket is Wat-201. That water makes excellent hydrogen bonds to the carbonyl oxygens of residues 23 and 97 (presumably as a donor) and is an acceptor of a hydrogen bond from a side-chain hydroxyl of Thr-99. The arrangement of hydrogen bonds is nicely tetrahedral, and the temperature factor of this water is 12 Å^2, the lowest of all solvents. There are, however, three water molecules (Wat-293, Wat-302, and Wat-338) in the distance of 4–4.5 Å from Wat-201, and thus this water is not considered to be buried. As mentioned before, an equivalent water molecule is present in every model of RNase compared here. Some other water molecules (for example, Wat-284, Wat-290, Wat-314, Wat-348, and Wat-353) do not make clear contacts with other ordered water molecules, but they all are on the surface of the protein, are hydrogen bonded to polar groups, and are not buried.

During the process of refinement we found an interesting string of water sites forming two interpenetrating networks. Positions of these waters were very clear in difference Fourier maps, since positive peaks were persistently present between the original locations of the waters, at distances of about 1.5 Å apart. This string of waters consists of Wat-282, Wat-375, Wat-372, Wat-377, Wat-369, Wat-370, Wat-285, and the symmetry mate of Wat-239 (Figure 3). Most of these waters make hydrogen bonds with either one or another of the alternate positions of the side chain of Asn-34, and this may be the driving force in their arrangement. The distances between the molecules in the center of the string and their nearest neighbors are between 1.5 and 1.8 Å, while the distances between the second nearest neighbors are between 2.8 and 3.2 Å. Initially only waters Wat-282, Wat-372, Wat-369, Wat-285, and Wat-239 were placed in the density, but extra density was always found in the difference Fourier maps, indicating the presence of the other waters described above. All of the waters in this string make good hydrogen bonds to protein atoms, but at this stage we would not like to speculate if two or more alternate water networks are present and how the water molecules are partitioned between them (Savage & Wlodawer, 1986).

2-Methyl-2-propanol Binding Site. A single 2-methyl-2-propanol binding site was located during the course of the high-resolution refinement. The starting model (A) had two water molecules, Wat-198 and Wat-253, positioned 1.8 Å apart at this location. These waters had full occupancies with relatively high temperature factors, 39.45 and 30.65 Å^2, respectively. The two waters also had no close interactions with the RNase A molecule or other solvent molecules. The early $F_o - F_c$ difference maps revealed positive density at this location, suggesting that the original model was incorrect. Omission of these two waters from the refinement produced positive difference density with a shape of a 2-methyl-2-propanol. A 2-methyl-2-propanol molecule was positioned in the density in such a way as to maximize the hydroxyl group's interaction with the surrounding solvent. Near the end of the refinement, the 2-methyl-2-propanol molecule was omitted from the model, and four cycles of restrained least-squares refinement were completed. An $F_o - F_c$ difference map was calculated, and again, strong positive difference density (>3σ) confirmed the presence of a 2-methyl-2-propanol molecule.

The 2-methyl-2-propanol binding site is located on the face of the molecule opposite from the active site. A small hydrophobic pocket is formed by residues 20–25, a segment of the linker region connecting helices H1 and H2, and residues 99–101, a segment of the antiparallel β-strand. Tyr-25 and Thr-82 are found at the back of the pocket (Figure 4). Ser-15 and Ser-16, two residues of the type IV turn formed by residues

Table V: Root-Mean-Square Differences (Å) between the 1.26-Å Model (G), the Joint 2.0-Å Neutron/X-ray Starting Model (A), the Birkbeck 1.4-Å Model (B), and the Genex 2.0-Å Cross-Linked Model (X)[a]

		rms difference (Å)		
	rms difference for	G	X	B
A	Cα	0.154	0.485	0.188
	main chain	0.164	0.494	0.224
	side chain	0.969	1.356	0.982
	total atoms	0.680	1.003	0.698
B	Cα	0.199	0.483	
	main chain	0.240	0.518	
	side chain	1.105	1.439	
	total atoms	0.784	1.063	
X	Cα	0.444		
	main chain	0.451		
	side chain	1.325		
	total atoms	0.972		

[a] The number of atoms averaged: 124 for Cα; 496 for main chain; 454 for side chains; 950 total.

15–18 of a neighboring molecule, are also associated with the binding site.

The 2-methyl-2-propanol is oriented with the C3 methyl positioned at the back of the pocket. This group is less than 4.0 Å from atoms of Ala-20, Thr-99, and Gln-101. The C5 methyl is less than 4.0 Å from atoms of Ser-22 and Thr-23 and Ser-15 of a neighboring RNase A molecule. The C4 methyl group has no close contacts with the protein molecule and is within 4.0 Å of only water molecules, Wat-309, Wat-329, Wat-355, and Wat-329. The hydroxyl group, O1, is 2.63 Å from Wat-336 and less than 4.0 Å from Wat-99 and Wat-299. The angle between C2, O1, and Wat-336 is 169.7°, indicating poor hydrogen-bond geometry.

There may actually be two different orientations of the 2-methyl-2-propanol molecule, the one which is presently modeled and described above and another in which the C4 position contains the hydroxyl group. Although the waters near position C4 are not close enough for strong hydrogen bonds, the geometry of the alternate orientation is reasonable. There is a bulge in the electron density between methyl groups C4 and C5 which could be due to partial occupancy of the binding site by one or more waters.

Comparison with Other Models. The final 1.26-Å phosphate-free enzyme structure (model G), besides being compared with the 2.0-Å starting model from the joint X-ray/neutron structure (model A) of Wlodawer and Sjölin (1983), has been also compared with both the 1.4-Å RNase A structure (model B) of Borkakoti et al. (1982) and the 2.0-Å cross-linked RNase A structure (model X) of Weber et al. (1985). A synopsis of the crystallization conditions, space groups, unit cell dimensions, and refinement parameters for each of the structures is presented in Table II. Model A has previously been compared with model X (Weber et al., 1985) and with model B (Wlodawer et al., 1986). The comparison of the structures presented here was done after a least-squares minimization of the root-mean-square (rms) differences between model G and models A, B, and X.

The differences between each of the models are summarized in Table V. The results indicate that all of the high-resolution RNase A structures compared here are very similar to one another with the best structural correlation between models G and A. This is perhaps not too surprising since model A was the starting point in deriving model G. Model A was also the point of departure in the X structure determination. What is worthwhile noting is that the RNase A structure appears to be highly conserved even with different crystallization conditions and with different crystal symmetry.

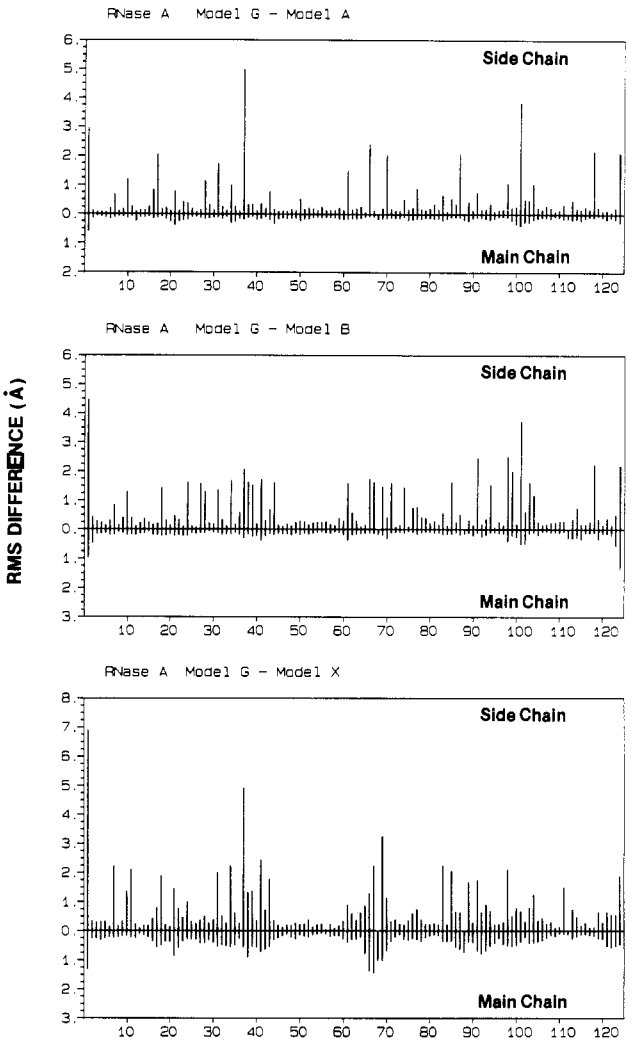

FIGURE 5: Comparison of the refined structures of RNase A model G with models A, B, and X. The rms positional differences are averaged for the side-chain and the main-chain atoms of each amino acid residue.

The rms differences between model G and models A, B, and X for main- and side-chain atoms for each residue are plotted in Figure 5. The qualitative differences between residues with greater than 1.0-Å rms differences in coordinate positions are summarized in Table VI. Residues with rms differences greater than 2.5 Å for side-chain atoms include Lys-1, Lys-37, Gln-69, and Gln-101, and residues with rms differences greater than 1.0 Å for main-chain atoms include Lys-1, Lys-66, Asn-67, Gly-68, Gln-69, and Val-124. The differences between the models for each of these residues are discussed below.

Lys-1 side chains of models G, A, and B have similar orientations through CB. Corresponding side-chain atoms beyond CB diverge, however, with the atoms for model A between those of G and B. The N-terminal nitrogen atom for model X is located near CB for the other models. Thus, the side chain of model X is located 120° away from the side chains of the other models. This accounts for the discrepancy of the main-chain atoms of model X with those of other models. When model X is superimposed on the other models, the Lys-1 side chain penetrates a symmetry-related molecule in the $P2_1$ cell. This difference in orientation must be a result of the different packing for molecules in the various crystal forms.

All of the models have Lys-37 oriented differently. The CB's of models G, A, and B are in the same relative position with the side chain in G aligned closest to that of model B. The model A side chain is between those of models B and X,

Table VI: Comparison of the 1.26-Å Phosphate-Free Ribonuclease A Model (G), the 2.0-Å Joint X-ray/Neutron Starting Model (A), the 1.4-Å Model (B), and the 2.0-Å Cross-Linked Model (X)

residue	>1.0 Å[a] A	B	X	differences
Lys-1	S	S	R	very different in all models; CB's of G, A, and B at similar positions; rest of side chain diverges; CB of X is 120° from others
Lys-7			S	all models have similar orientations through CG; G and A have similar conformations; NZ of B is oriented differently; X very different due to cross-linking
Arg-10	S	S	S	G and A have similar conformations; X is very similar; χ_5 difference between B and other models is 180°
Gln-11			S	nearly identical conformations for G, A, and B; X quite different from others; G amide plane tilted slightly
Gln-11A	S	S	S	CG for A between X and other models; the amide plane is perpendicular to X's
Thr-17	S			identical G, X, B; A is rotated 180° about CA–CB bond
Ser-18	S	S		identical G, A; OG was turned 120° in B and 240° in X
Ser-21			S	main chain shifted in X; OG of G near CB in X
Asn-24		S	S	all very similar; X CG shifted between CG and ND2 of G; CG of B shifted between CG and OD1 of G
Asn-27		S		OD1 and ND2 of B assigned the opposite from others
Gln-28	S	S		all similar to CB; G and X are similar; A and B very close, but CG positions differ
Lys-31	S	S	S	all models have CB's in similar location; general direction of B and A similar; X quite different; G between X and B
Ser-32				all four have similar conformations
Ser-32A	S	S	S	G OG is 120° from others
Asn-34		S	S	A and B are similar; CG of G between CG of A and ND2 of B
Asn-34A	S	S	S	G closest to X with ND2 of G near CG of X
Lsy-37	S	S	S	all conformations are different; CB's of G, A, and B are close; G and B are oriented in opposite direction from A and X; Wat-274 near CE of A
Asp-38		S	S	G and A are identical; OD1 and OD2 assignments of B opposite from G and A; X atoms are parallel but shifted with the carboxylate group rotated 45° relative to others
Arg-39		S	S	G and A similar; X has similar χ angles; B radically different near NE
Lys-41		S	S	G and A similar; B and X similar up to CG; B between G and X; X has altered orientation due to cross-linking
Val-43			S	A and B similar; G closest to A and B
Val-43A	S	S		G similar to X; CG1 of G near CG2 of A and B
Asn-44		S		OD1 and ND2 of B assigned opposite to others
Ser-50				G, A, B, and X identical
Ser-50A			S	OG of G is rotated 90° from others
Lys-61	S	S		G and X are similar, but diverge from each other; A and B have similar orientations with CG's varying
Lys-61A	S	S	S	G different from all models
Lys-66	S	S	R	all models different; G and X point in same direction; X is translated from others; Wat-333 near NZ of A and Wat of X; NZ of B near CE of G
Asn-67		S	R	G, A, and B correspond very closely; X shifted but similar conformation to other three; OD1 and ND2 assignments for B opposite from others
Asn-67A	S	S	S	χ_1 deviates by about 150° from others
Gly-68			M	A, G, and B identical; X shifted
Gln-69		S	R	identical in A, B, and G; very different in X (may be due to crystal packing); B OE1 and NE2 assigned opposite from others
Thr-70	S		S	general orientation of G, B, and X agree; A OG1 near CG2 of others with a slight shift of CB
Asn-71		S		B OD1 and ND2 assigned opposite from others
Gln-74			S	X OE1 and NE2 assigned opposite from others
Ser-77				A and B have similar orientations; OG of G is rotated ≈60° from A and B; X is similar to G but translated
Ser-77A	S	S		G is rotated ≈120° from A and B
Asp-83			S	G similar to A and B
Asp-83A	S	S		G similar to X
Arg-85		S	S	G almost identical with A and similar to B; X in the same general direction, but not identical
Arg-85A	S	S	S	G has no correspondence to the other models
Thr-87	S			G, B, and X are similar; A OG1 close to CG2 of others
Ser-89			S	G, A, and B are similar; X OG turned 120° from others
Lys-91		S	S	all models close through CG; NZ of B near Wat-220; X similar but CE and NZ shifted
Lys-91A	S	S	S	NZ 180° from NZ of G Lys-91
Lys-98	S	S	S	all models are similar through CG then diverge; G similar to B and X; NZ of G points away from others
Lys-98As		S	S	all models are similar through CG then diverge; G like A
Thr-99		S		G, A, and X are nearly identical; CG2 of B at OG1 of G
Glu-101	S	S		G and X are similar; A and B are similar; G Wat-202 and Wat-209 at OE1 and NE2 of A and B
Asn-103			S	G, A, and B are similar; X translated from others but has a similar conformation
Lys-104	S	S	S	G, A, B, and X oriented in similar direction
Lys-104A	S	S		NZ of G oriented opposite from G Lys-104
Glu-111			S	OE1 and OE2 of X assigned opposite from others
Val-118	S	S		G and X are similar; A and B similar with CG1 and CG2 corresponding to CG2 and CG1 of G and X
Val-124	S	R	S	A and B similar with CG1 and CG2 corresponding to CG2 and CG1 of G; CG2 of X at CG1 of G

[a] The root-mean-square distance in angstroms of side-chain atoms (S), main-chain atoms (M), and all atoms of the residue (R).

which are oriented in the opposite direction from that in model G. The thermal parameters for Lys-37 side-chain atoms are among the highest for any residue in all of the models (Figure 6), indicating a higher degree of uncertainty for positioning the atoms correctly than in most of the other residues of the molecule. In fact, Lys-37 is in a region of residues 34–39 which was completely rebuilt during the 1.4-Å refinement (Borkakoti et al., 1982). Even though the side-chain electron density was poorly defined in the maps of phosphate-free RNase, there was no indication that Lys-37 could be modeled with alternate side-chain conformations.

The conformations of Gln-69 are identical in models G, A, and B, except that model B has the assignments for OE1 and NE2 reversed. The assignment in model A was made on the basis of the appearance of neutron density and is more likely to be correct than any assignment made on the basis of X-ray

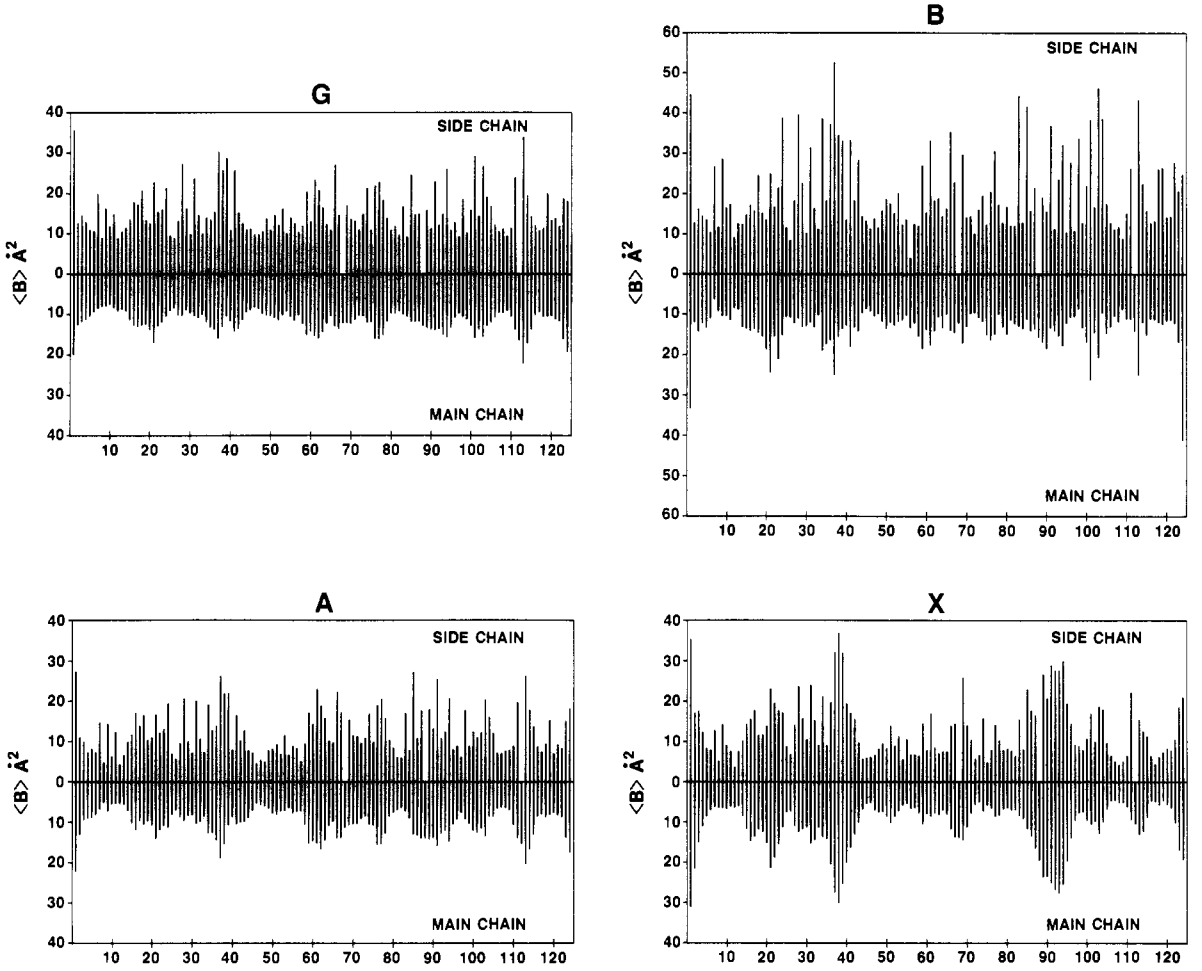

FIGURE 6: Temperature factors averaged for the main chain (above center line) and for the side chains (below center lines) for the four models of RNase compared here.

diffraction data only, as was the case with structure B. In model, X, Gln-69 is shifted along with residues 66–68; all of which, including Gln-69, have deviations from the other models of more than 1.0 Å for main-chain atoms. Residues 66–68 form part of a type IV turn, a loop of polypeptide chain which extends away from the body of the enzyme. The alteration of position for these residues in model X is due primarily to crystal packing differences.

The side-chain atoms of Gln-101 were repositioned during the refinement of the 1.26-Å structure. The side-chain atoms were rotated about χ_1 by $\approx 180°$. The new conformation positioned NE2 near Wat-235 of model A. Wat-202 and Wat-209 of model G were positioned near OE1 and NE2 of model A, respectively. Attempts were made to model this residue with alternate conformations; however, the $2F_o - F_c$ and $F_o - F_c$ difference maps could only be interpreted with a single conformation. Seven additional cycles were run after the final cycle of refinement. In this test, Gln-101 side-chain atoms, as well as Wat-202 and Wat-209, were omitted from the model. The resulting $2F_o - F_c$ and $F_o - F_c$ difference maps, when contoured respectively at 0.5σ and 3σ, suggested a possibility of two conformations. However, attempts at modeling this residue with two alternate conformers produced similar results to those found earlier in the refinement. There was no trace of electron density in the CD position of the conformation found in model A. We therefore have to conclude that there is only one predominent conformation for Gln-101 in the phosphate-free form of the enzyme.

The atomic models for Gln-101 in G and A, along with their respective maps, are shown in Figure 7. This figure also shows a comparison of all residues in the vicinity of Gln-101. There is a visible shift of positions of the backbone atoms of residues 99–104; these atoms shifted an average rms difference of 0.294 Å. Unbeknownst to the authors at the time of the 1.26-Å structure refinement, Gln-101 of model X was positioned in the same orientation as the final orientation in model G. Recently, an examination of the hydrogen-bond geometry of the 1.4-Å RNase A structure (Borkakoti et al., 1982) by Harris et al. (1987) has revealed an anomaly in the hydrogen-bond interactions of this residue, caused by tight packing of the side chain. While we initially suspected that this anomaly could be due to misinterpretation of the electron density, we are now convinced that Gln-101 is indeed found in two distinct positions, depending on crystallization conditions.

The C-terminal residue, Val-124, has a similar conformation in models A and B. These are similar to that found in model G except that there is an inversion about the CB atom; thus CG1 and CG2 of model G correspond to CG2 and CG1 of models A and B, respectively. CG2 of model X is located near CG1 of G, but there is no correspondence to the other side-chain atoms. The N- and C-terminal residues of many protein molecules have been difficult to model, and this appears to be true also for RNase A.

A common feature of all of the rms difference plots (Figure 5) is the low deviation observed for the amino acid residues 45–60. The average rms differences for residues in this region for the amino acid side-chain atoms are 0.181, 0.211, and 0.201 Å, while for the main-chain atoms they are 0.141, 0.131, and 0.165 Å for the comparisons of G with A, B, and X, respec-

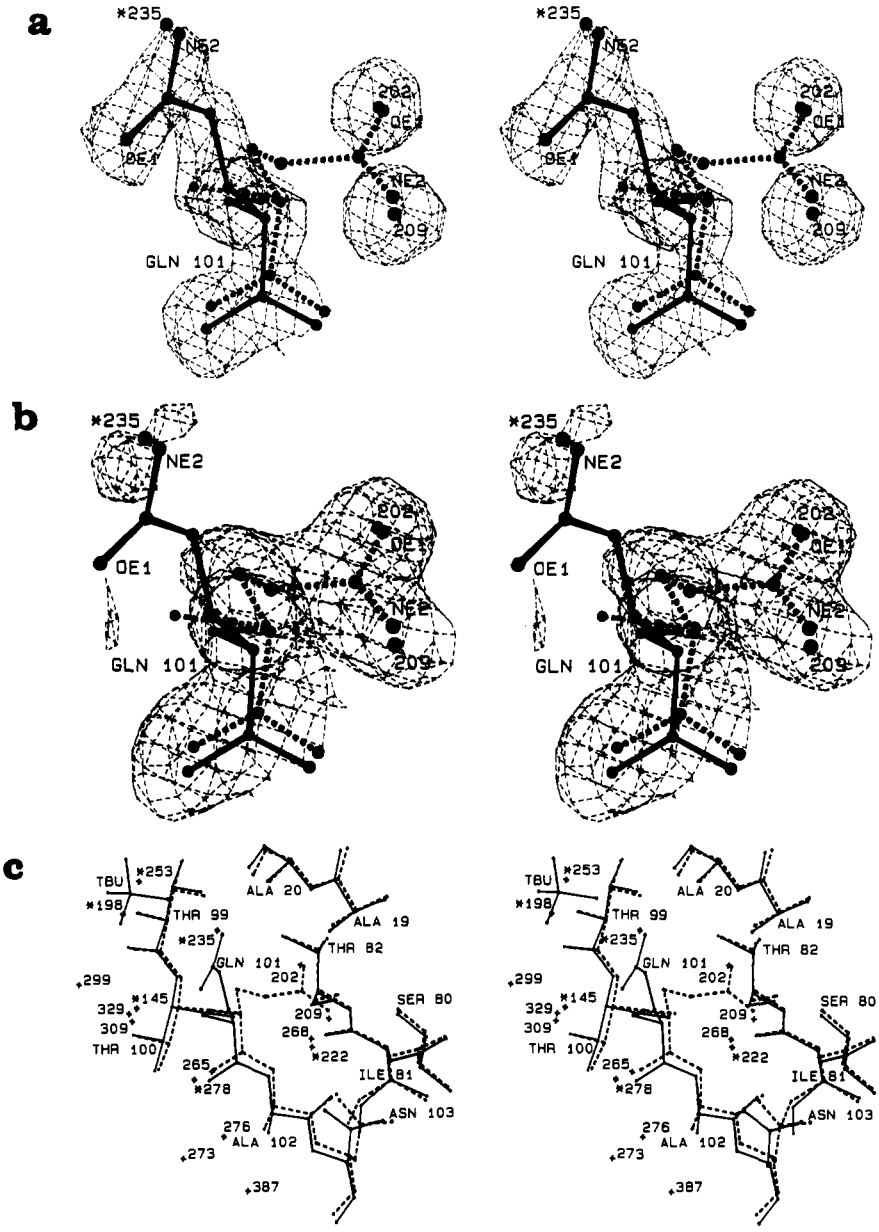

FIGURE 7: Stereoscopic views of (a) Gln-101 of model G (solid) and of the starting model A (dashed) with the final $2F_o - F_c$ map contoured at 1σ, (b) Gln-101 of model A (dashed) and the final model G (solid) in the final $2F_o - F_c$ X-ray map of Wlodawer and Sjölin (1983), and (c) model G (solid) and model A (dashed) in the region surrounding Gln-101. Water molecules labeled with asterisks are from model A.

tively. These values are substantially lower than the values given in Table V as differences for the entire structure. Residues of this region are integral parts of the secondary structure of RNase A: residues 45–48 form part of the antiparallel β-strand composed of residues 41–48, and residues 50–60 form the helix H3. Residues 50–55 form a partially distorted α-helix, while residues 56–60 progressively distort toward a 3_{10} helix. Besides the conformations being highly conserved in all RNase models, the atoms of this region have markedly lower thermal factors in all of the structures (Figure 6).

In addition to the secondary structural elements, the region spanning residues 45–60 contains Cys-58, which forms a disulfide bridge with Cys-110. An examination of the local region of the polypeptide chain around Cys-110 reveals another segment (residues 105–117) in which the conformations of the amino acids are highly conserved. Residues in this region also participate in the formation of secondary structural elements. Residues 105–110 form half of β-strand 94–110, and residues 112–115 form a type VI turn, due to the presence of a *cis*-proline. The thermal parameters for atoms in this region, however, are not noticeably lower than those for other residues of the polypeptide chain. A large difference seen for Glu-111 in the plot (Figure 5) is caused by renaming OE1 and OE2 (Table VI).

These two regions, which are adjacent to one another in the three-dimensional structure of RNase and together make up only 23% of the total sequence, contain more than 50% of the residues with fully protected amide hydrogens (7 out of 12) (Wlodawer & Sjölin, 1982). These regions also have only a single residue, Ser-50, which has been modeled with multiple conformations. The only difference in the two conformers is the position of OG.

Discussion

The high-resolution structure of phosphate-free RNase reported here, besides providing new information on the fine details of the three-dimensional structure, provides some insight into the dynamic properties of the enzyme. This is achieved both directly by the analysis of disordered side chains and

solvent molecules and indirectly by the comparison of the structure reported here with previously reported refined structures of RNase.

The active site has rather few differences compared with the phosphate-containing RNase (model A). Most of the active site residues are located in similar positions. There have been suggestions that some of the side chains involved in catalysis, such as Lys-41, move by as much as several angstroms (Alber et al., 1983), but our comparison of phosphate-free and phosphate- (sulfate-) containing active sites does not support this view. Of the residues directly involved in catalytic action, only the ring of His-119 moves, and only by a few tenths of an angstrom. This new location seems to be stabilized by a hydrogen bond to OD1 of Asp-121. Gln-11 can occupy multiple conformations in the phosphate-free enzyme, and this phenomenon is probably due to its inability to form hydrogen bonds in the absence of phosphate or sulfate ions.

With the removal of the phosphate group it is not surprising that the solvent structure in the active site is altered. At least two waters occupy positions in the vicinity of the phosphate binding site. Removal of the phosphate apparently also triggers higher mobility of the active site residues that can be seen when the temperature factors are analyzed. Whether the absence of phosphate in the active site induces a different dynamic motion pattern in a more global sense cannot be decided yet and could only be checked by computer simulations.

Currently only a few reports of protein structure determinations offer descriptions of discrete alternate positions of side chains, and with few exceptions no corresponding electron densities have been published [for discussion, see Svensson et al. (1987) and references cited therein]. We reported earlier (Svensson et al., 1987) alternate positions for 13 side chains in RNase A. As expected, residues with discrete multiple conformations are found predominantly on protein surfaces, where the barriers to reorientation are the lowest. While a few such residues are involved in intermolecular contacts, the vast majority are not, indicating that such phenomena are to be expected in solution as well. Most of the amino acid residues having multiple conformations are either polar or charged, with serines, lysines, and arginines common among them. Since multiple orientations of side chains are now becoming a common feature of carefully refined structures based on high-resolution data, discrete multiple conformations might be treated as a norm rather than an exception for a certain fraction of surface amino acids.

The comparison of the four models of RNase identified those water molecules that are common between them and are practically integral parts of the protein molecule. All of these waters are in the first shell of hydration, and they frequently form more than one hydrogen bond to the protein. This analysis also shows that there are no buried water sites in ribonuclease. An interesting feature that was found during the course of refinement is a string of waters which may be forming two interpenetrating networks (Figure 3). Most of these waters make hydrogen bonds with either one or another of the alternate positions of the side chain of Asn-34.

High-resolution investigation of proteins ought to establish the water structure with a high degree of significance. For instance, in the structure of crambin to atomic resolution, Teeter (1984) found waters in pentagonal clusters around such hydrophobic groups as side-chain methyls. We have not, as yet, seen any such arrangement in the structure of ribonuclease. It should be kept in mind, though, that crambin is strongly hydrophobic and is crystallized by precipitation with water, while ribonuclease is hydrophilic and highly soluble.

A site of solvent other than water, in this case a 2-methyl-2-propanol molecule, was also identified in the high-resolution structure of phosphate-free ribonuclease. In the studies done at lower resolution, potential solvent molecules most frequently are modeled as waters, but in this investigation it has become clear that the site indeed contains an alcohol. The maps even suggested that butanol at this site might have multiple conformations, but in the final model we have not utilized this indication. The solvent molecule is bound in a small hydrophobic pocket opposite to the active site.

One of the most puzzling features of this investigation is finding the side chain of Gln-101 in the orientation very different than that observed in the presence of the phosphate. Harris et al. (1987) noticed that this side chain makes as many as eight hydrogen bonds in the model containing sulfate in the active site. Since the conformation of Gln-101 was so drastically different in the phosphate-free enzyme, we suspected a possibility of an experimental error and reinvestigated very carefully the model of Wlodawer and Sjölin (1983). We are now convinced that in the presence of an ion the side chain of Gln-101 indeed packs into the center of the molecule. Our only remaining disagreement with the results of the London group is the identification of OE1 and NE2. Our assignment was made on the basis of neutron density maps, in which the differences between an oxygen and an amide group are very pronounced, while the X-ray model can be built only on the basis of the putative hydrogen bonds. Exchanging OE1 and NE2 results in more probable hydrogen bonds in any case. While the difference between the phosphate-containing and phosphate-free crystals is very clear, its origin is by no means certain. We cannot point to the mechanism which would, in the absence of phosphate, drive the side chain outside, replacing it with two water molecules. Also the importance, if any, of this phenomenon to the catalytic mechanism is not clear.

The unfolding properties of ribonuclease A have been investigated by many laboratories using different experimental techniques, and subsequently, various models have been suggested. Generally, crystallographic investigations do not shed much light on folding or unfolding of the proteins. However, when the structure of phosphate-free ribonuclease A is compared with three other models and the results are viewed in the context of other structural biochemical studies, insight into the folding pathway is obtained.

Our current investigation supports the hypothesis that the region defined by residues 37 and 92 and their closest neighbors unwinds early in the unfolding process. We base our conclusion on results summarized in Figures 5 and 6. From Figure 6, model X, it is clearly seen that in the cross-linked RNase this region (around Lys-37 and Tyr-92) has the highest mobility, and this can also be noticed in the other models, although not as clearly. Figure 5 shows the rms differences between all the models, and it is clearly evident that the very same region contains large differences, indicating more conformational freedom in this area of the enzyme structure. Further support for this hypothesis comes from our previous study of hydrogen exchange in ribonuclease (Wlodawer & Sjölin, 1982). At that time a somewhat puzzling feature of the enzyme structure was reported, namely, that the distribution of protected amides was found not to be uniform throughout the β-sheet. Of the two distinct areas surrounding the cleft near the active site, one appeared to be more flexible than the other. In particular, that part of the backbone containing residues 75, 106–109, 116, and 118 appeared to

be much more highly protected from exchange than any other region of the protein. Additional support is given by Matheson and Scheraga (1979), who suggested that Tyr-92 was indeed involved in this primary step of unfolding. The unfolding in this region probably initiates a cooperative unzippering of the β-sheet region. The lack of protection against hydrogen–deuterium exchange would support this view.

The folding nucleation site for ribonuclease A, on the other hand, should then be found in the other part of the molecule, where the protected residues are located, since a nucleation site by definition should unfold late. Matheson and Scheraga (1979) suggested that residues 106–118 formed that nucleation site. As mentioned previously, residues 45–60 and 105–110 (Figure 5) have the most highly conserved conformations of all residues in the four models and residues 45–60 have the lowest temperature factors. Thus, while our findings cannot conclusively prove that the folding nucleation site is within this region of the enzyme structure, the evidence is in general agreement with previous studies.

Added in Proof

Crystal structure of anion-free ribonuclease was recently refined by Campbell and Petsko (1987). They found that "... the overall *B* factor has increased by 3.9 Å^2 after the loss of the sulfate, indicating an overall loosening of the protein". Results presented here do not support such a conclusion, as no overall change in the rigidity of the structure was detected.

Acknowledgments

This project was initiated while G.L.G. was employed at Genex Corp., Gaithersburg, MD, and we acknowledge the permission to use their X-ray facility. We thank Dr. Andrew Howard for help in data collection and processing and Dr. Barry Finzel for his assistance in the early stages of structure refinement.

Registry No. RNase A, 9001-99-4.

References

Agarwal, R. C. (1978) *Acta Crystallogr., Sect. A: Cryst. Phys., Diffr., Theor. Gen. Crystallogr. A34*, 791–909.

Alber, T., Gilbert, W. A., Ponzi, D. R., & Petsko, G. A. (1983) *Ciba Symp. 93*, 4–24.

Almassy, R. J., Fontecilla-Camps, J. C., Suddath, F. L., & Bugg, C. E. (1983) *J. Mol. Biol. 170*, 497–527.

Avey, H. P., Boles, M. O., Carlisle, C. H., Evans, S. A., Morris, S. J., Palmer, R. A., Woolhouse, B. A., & Shall, S. (1967) *Nature (London) 213*, 557–556.

Blevins, R. A., & Tulinsky, A. (1985) *J. Biol. Chem. 260*, 8865–8872.

Borkakoti, N., Moss, D. S., & Palmer, R. A. (1982) *Acta Crystallogr., Sect. B: Struct. Crystallogr. Cryst. Chem. B38*, 2210–2217.

Campbell, R. L., & Petsko, G. A. (1987) *Biochemistry 26*, 8579–8584.

Delbaere, L. T. J., & Brayer, G. D. (1985) *J. Mol. Biol. 183*, 89–103.

Dewan, J. C., & Tilton, R. F. (1987) *J. Appl. Crystallogr. 20*, 130–132.

Fankuchen, I. (1941) *J. Gen. Physiol. 24*, 315–316.

Finzel, B. (1987) *J. Appl. Crystallogr. 20*, 53–55.

Fujinaga, M., Read, R. J., Sielecki, A., Ardelt, W., Laskowski, M., & James, M. N. G. (1982) *Proc. Natl. Acad. Sci. U.S.A. 79*, 4868–4872.

Harris, G. W., Borkakoti, N., Moss, D. S., Palmer, R. A., & Howlin, B. (1987) *Biochim. Biophys. Acta 912*, 348–356.

Hendrickson, W. A. (1985) *Methods Enzymol. 115*, 252–270.

Hendrickson, W. A., & Konnert, J. H. (1980) in *Computing in Crystallography* (Diamond, R., Ramseshan, S., & Venkatesan, K., Eds.) pp 13.01–13.23, Indian Institute of Science, Bangalore.

Howard, A. J., Gilliland, G. L., Finzel, B. C., Poulos, T. L., Ohlendorf, D. H., & Salemme, F. R. (1987) *J. Appl. Crystallogr. 20*, 383–387.

James, M. N. G., Sielecki, A. R., Brayer, G. D., Delbaere, L. T. J., & Bauer, C.-A. (1980) *J. Mol. Biol. 144*, 43–88.

Jones, T. A. (1978) *J. Appl. Crystallogr. 11*, 268–272.

Kartha, G., Bello, J., & Harker, D. (1967) *Nature (London) 213*, 862–865.

Kunitz, M. (1940) *J. Gen. Physiol. 24*, 15–32.

Kuriyan, J., Wilz, S., Karplus, M., & Petsko, G. A. (1986) *J. Mol. Biol. 192*, 133–154.

Lehmann, M. S., Mason, S. A., & McIntyre, G. J. (1985) *Biochemistry 24*, 5862–5869.

Matheson, R. R., & Scheraga, H. A. (1979) *Biochemistry 18*, 2438–2445.

Morffew, A., & Moss, D. S. (1982) *Comput. Chem. 6*, 1–3.

Read, R. J., Fujinaga, M., Sielecki, A. R., & James, M. N. G. (1983) *Biochemistry 22*, 4420–4433.

Sakabe, N., Sakabe, K., & Sasaki, K. (1981) in *Structural Studies on Molecules of Biological Interest* (Dodson, G., Glusker, J. P., & Sayre, D., Eds.) pp 509–526, Clarendon, Oxford.

Saunders, M., Wishnia, A., & Kirkwood, J. (1957) *J. Am. Chem. Soc. 79*, 3289–3290.

Savage, H., & Wlodawer, A. (1986) *Methods Enzymol. 127*, 162–183.

Smyth, D. G., Stein, W. H., & Moore, S. (1963) *J. Biol. Chem. 238*, 227–234.

Svensson, L. A., Sjölin, L., Gilliland, G. L., Finzel, B. C., & Wlodawer, A. (1987) *Proteins: Struc. Func. Gen. 1*, 370–375.

Ten Eyck, L. F. (1973) *Acta Crystallogr., Sect. A: Cryst. Phys., Diffr., Theor. Gen. Crystallogr. A29*, 183–191.

Walter, J., Steigemann, W., Singh, T. P., Bartunik, H., Bode, W., & Huber, R. (1982) *Acta Crystallogr., Sect. B: Struct. Crystallogr. Cryst. Chem. B38*, 1462–1472.

Weber, P. C., Sheriff, S., Ohlendorf, D. H., Finzel, B. C., & Salemme, F. R. (1985) *Proc. Natl. Acad. Sci. U.S.A. 82*, 8473–8477.

Wlodawer, A. (1985) *Biol. Macromol. Assem. 2*, 394–493.

Wlodawer, A., & Hendrickson, W. A. (1982) *Acta Crystallogr., Sect. A: Cryst. Phys., Diffr., Theor. Gen. Crystallogr. A38*, 239–247.

Wlodawer, A., & Sjölin, L. (1982) *Proc. Natl. Acad. Sci. U.S.A. 79*, 1418–1422.

Wlodawer, A., & Sjölin, L. (1983) *Biochemistry 22*, 2720–2728.

Wlodawer, A., Borkakoti, N., Moss, D. S., & Howlin, B. (1986) *Acta Crystallogr., Sect. B: Struct. Sci. B42*, 379–387.

Wlodawer, A., Deisenhofer, J., & Huber, R. (1987a) *J. Mol. Biol. 193*, 145–156.

Wlodawer, A., Nachman, J., Gilliland, G. L., Gallagher, W., & Woodward, C. (1987b) *J. Mol. Biol. 198*, 469–480.

Wyckoff, H. W., Hardman, K. D., Allewell, N. M., Inagami, T., Tsernoglou, D., Johnson, L. N., & Richards, F. M. (1967) *J. Biol. Chem. 242*, 3749–3753.

Wyckoff, H. W., Tsernoglou, D., Hanson, A. W., Knox, J. R., Lee, B., & Richards, F. M. (1970) *J. Biol. Chem. 245*, 305–328.