Crystal Structure of Ribonuclease Ms (as a Ribonuclease T_1 Homologue) Complexed with a Guanylyl-3′,5′-Cytidine Analogue[†,‡]

Takamasa Nonaka,[§] Kazuo T. Nakamura,[§,∥] Seiichi Uesugi,[⊥] Morio Ikehara,[#] Masachika Irie,[△] and Yukio Mitsui*[,§]

Department of BioEngineering, Nagaoka University of Technology, Kamitomioka, Nagaoka, Niigata 940-21, Japan, Faculty of Engineering, Yokohama National University, Hodogaya, Yokohama, Kanagawa 240, Japan, Protein Engineering Research Institute, Furuedai, Suita, Osaka 565, Japan, and Department of Microbiology, Hoshi College of Pharmacy, Ebara, Shinagawa, Tokyo 142, Japan

Received May 24, 1993; Revised Manuscript Received August 19, 1993[●]

ABSTRACT: A ribonuclease T_1 homologue, ribonuclease Ms (RNase Ms) from *Aspergillus saitoi*, has been crystallized as a complex with a substrate analogue GfpC where the 2′-hydroxyl (2′-OH) group of guanosine in guanylyl-3′,5′-cytidine (GpC) is replaced by the 2′-fluorine (2′-F) atom to prevent transesterification. The crystal structure of the complex was solved at 1.8-Å resolution to a final *R*-factor of 0.204. The role of His*92* (RNase T_1 numbering) as the general acid catalyst was confirmed. Of the two alternative candidates for a general base to abstract a proton from the 2′-OH group, His*40* and Glu*58* were found close to the 2′-F atom, making the decision between the two groups difficult. We then superposed the active site of the RNase Ms/GfpC complex with that of pancreatic ribonuclease S (RNase S) complexed with a substrate analogue UpcA, a phosphonate analogue of uridylyl-3′,5′-adenosine (UpA), and found that His12 and His119 of RNase A almost exactly coincided with Glu*58* and His*92*, respectively, of RNase Ms. Similar superposition with a prokaryotic microbial ribonuclease, RNase St [Nakamura, K. T., Iwahashi, K., Yamamoto, Y., Iitaka, Y., Yoshida, N., & Mitsui, Y. (1982) *Nature 299*, 564–566], also indicated Glu*58* as a general base. Thus the present comparative geometrical studies consistently favor, albeit indirectly, the traditional as well as the most recent notion [Steyaert, J., Hallenga, K., Wyns, L., & Stanssens, P. (1990) *Biochemistry 29*, 9064–9072] that Glu*58*, rather than His*40*, must be the general base catalyst in the *intact* enzymes of the RNase T_1 family.

Ribonuclease Ms (RNase Ms; EC 3.1.4.23)[1] from the fungus *Aspergillus saitoi* is a microbial eukaryotic RNase (Ohgi & Irie, 1975) and consists of a single polypeptide chain (M_r 11 401) of 105 amino acid residues (Watanabe et al., 1982) having two disulfide bridges. RNase Ms exhibits 65% amino acid sequence identity with ribonuclease T_1 (RNase T_1; EC 3.1.27.3) from *Aspergillus oryzae*. Due to very high homology, the alignment between the RNase Ms and RNase T_1 sequences is easy, as shown in Figure 1 of Nonaka et al. (1991) (minor corrections to this scheme will be given later in this paper). Hereafter, the T_1 numbering shown in italics (and in parentheses) will be used, when necessary, to facilitate comparison between RNases within the T_1 subfamily or comparison with RNases of other families or subfamilies.

RNase T_1 is the leading member of a family of microbial RNases that cleave single-stranded RNA by a reaction mechanism involving formation of the intermediate 2′,3′-cyclic phosphate (the transesterification step) followed by hydrolysis to yield a terminal 3′-phosphate (Takahashi et al., 1967; Takahashi & Moore, 1982). The RNase T_1 family of enzymes (consisting of RNases T_1, Ms, and a few other closely related RNases) belongs to one (*a smaller microbial RNase family* having molecular weights of ca. 11 000) of the two major families covering all the known microbial RNAses of bacterial (prokaryotic) and fungal (eukaryotic) origins. The other family (*a larger microbial RNase family* having molecular weights of ca. 20 000) involves RNases T_2, M, Rh, and a few other closely related enzymes. The enzymes of the smaller microbial RNase family are mutually related by at least partial sequence and spatial homology (Hartley, 1980; Hill et al., 1983). However, they are completely unrelated to the RNase A-type mammalian RNases (Richards & Wyckoff, 1971) with respect to both primary and tertiary structures (Blackburn & Moore, 1982; Nakamura et al., 1982). The three-dimensional structure of smaller microbial RNases was first established for a prokaryotic enzyme, RNase St from *Streptomyces erithreus* (Yamamoto et al., 1981; Nakamura et al., 1982). Nakamura et al. (1982) called attention to the existence of at least three kinds of subfamilies (the T_1, barnase, and St families) among the smaller microbial RNase family. Since then, three-dimensional structures have been established for a few other smaller microbial RNases including RNase T_1 (Heinemann & Saenger, 1982; Sugio et al., 1985a,b), barnase (Mauguen et al., 1982), RNase Sa (Sevcik et al., 1991), and RNase F_1 (Vassylyev et al., 1993). The early concerted efforts to compare the five three-dimensional

[†] This work was supported, in part, by the Ministry of Education, Science, and Culture.

[‡] Crystallographic coordinates for ribonuclease Ms complexed with 3′-guanylic acid have been deposited with the Brookhaven Protein Data Bank as entry 1RMS and those for ribonuclease Ms complexed with a guanylyl-3′,5′-cytidine analogue as entry 1RDS.

* To whom correspondence should be addressed.

[§] Nagaoka University of Technology.

[∥] Present address: Faculty of Pharmaceutical Sciences, Showa University, Hatanodai, Shinagawa, Tokyo 142, Japan.

[⊥] Yokohama National University.

[#] Protein Engineering Research Institute.

[△] Hoshi College of Pharmacy.

[●] Abstract published in *Advance ACS Abstracts*, October 1, 1993.

[1] Abbreviations: 2′-F, 2′-fluorine atom; GfpC, 2′-deoxy-2′-fluoroguanylyl-(3′-5′)-cytidine (a 2′-fluoro analogue of GpC); GfpU, 2′-deoxy-2′-fluoroguanylyl-(3′-5′)-uridine (a 2′-fluoro analogue of GpU); 2′-GMP, 2′-guanylic acid; 3′-GMP, 3′-guanylic acid; GpC, guanylyl-3′,5′-cytidine; G2′p5′G, guanylyl-2′,5′-guanosine; GpU, guanylyl-3′,5′-uridine; NMR, nuclear magnetic resonance; 2′-OH, 2′-hydroxyl group; rms, root mean square; RNase, ribonuclease; UpA, uridylyl-3′,5′-adenosine; UpcA, a phosphonate analogue of UpA; UVC, uridine vanadate complex.

0006-2960/93/0432-11825$04.00/0

structures of smaller microbial RNases have been described (Hill et al., 1983). Recently the first three-dimensional structures of *larger* microbial RNases have been established for RNase Rh from *Rhizopus niveus* (Kurihara et al., 1992), establishing that the enzymes of the larger microbial RNase family are related neither to the smaller microbial RNases nor to the RNase A-type mammalian RNases with respect to both primary and tertiary structures.

The catalytic residues of smaller microbial RNases have long been a matter of controversy, as studied by various methods of chemical modification and NMR spectroscopy [see the literature cited in Arni et al. (1988) for history]. The targets of the controversy have been the identity of the general base and acid residues which respectively participate in the abstraction of a proton from the 2′-hydroxyl group of the substrate ribose and in the protonation of the 5′-oxy group of a leaving substrate. For RNases of the T_1 family, Glu*58* and His*92* were respectively assigned as the base and acid catalysts on the basis of the three-dimensional structure of RNase T_1 (Heinemann & Saenger, 1982) in agreement with the earliest speculation based on extensive chemical modification studies (Takahashi et al., 1967; Takahashi, 1970). Recently, however, Nishikawa et al. (1987b) found that the His*40*Ala and His*92*Ala mutants of RNase T_1 are virtually inactive while the mutants in which Glu*58* is replaced by Gln, Asp, or Ala retain considerable activity. On the basis of these observations, they proposed a new mechanism for the reaction catalyzed by RNase T_1 in which His*40* (rather than Glu*58*) and His*92* act as the base and acid catalysts, respectively. Crystallographically it has been recognized that the decision between *His40* and Glu*58* as the true base catalyst is not easy if it is attempted merely on the basis of the three-dimensional structures of RNase T_1 in various liganded states [for example, see Arni et al. (1988)]. This difficulty is due to the fact that both the relevant imidazole ring and the carboxylate group are closely located to the proton-donating 2′-hydroxyl group. More recently, Steyaert et al. (1990) confirmed the original mutation experiments of Nishikawa et al. (1987b) made for RNase T_1. However, on the basis of a few other additional mutation experiments and extensive kinetic studies, they interpreted essentially the same observation in a different context: Glu*58* must be playing the role of a general base catalyst (as traditionally conceived) in the *intact* RNase T_1, while His*40* takes at least part of that role (only) when Glu*58* has been converted to other amino acid residues.

In the present laboratory, RNase Ms (an RNase T_1 homologue) complexed with 3′-GMP was crystallized (Nonaka et al., 1989), and a preliminary report on the structure solution has appeared (Nonaka et al., 1991). Furthermore, in view of the fact that none of the ligands attached to RNase T_1 in a series of crystallographic studies have been good mimetics of the true substrate for the transesterification step, we have synthesized an analogue of guanylyl-3′,5′-cytidine (GpC), GfpC, in which the 2′-hydroxyl group (2′-OH) of guanosine is replaced by a 2′-fluorine atom (2′-F) to prevent the transesterification step. In the hope that detailed examination of the geometrical environment of the 2′-F group bound to RNase Ms would provide a decisive clue as to the true base catalyst, we crystallized the RNase Ms/GfpC complex. Here we report on (1) a molecular replacement solution and refinement of the RNase Ms/GfpC complex and (2) further refinement of the structure of the RNase Ms/3′-GMP complex. Since even the above new approach to determine the base catalyst turned out to give inconclusive results, we then decided to compare, in three-dimensional space, the active sites of the RNase Ms complexes with those of the enzymes of T_1 and St subfamilies of the smaller microbial RNase family and those of the mammalian RNase family. Each of the above comparisons exhibited a remarkable similarity between the two active sites with respect to both the enzyme active site and bound ligand structures. Taking advantage of the fact that the base and acid catalysts had almost been established for the latter two enzymes [RNase St (Nakamura et al., 1982) and RNase A (or S) (Richards & Wyckoff, 1971)], we were able to identify Glu57 (or Glu*58* in RNase T_1) rather than His39 (or His*40* in RNase T_1) as the most likely general base catalyst in the *intact* enzymes of the RNase T_1 family, apparently in agreement with the recent notion put forward by Steyaert et al. (1990) (as cited above).

EXPERIMENTAL PROCEDURES

Materials. RNase Ms was prepared as described before (Watanabe et al., 1982). 3′-GMP was purchased from Sigma. GfpC (a 2′-fluoro analogue of GpC) was synthesized by a procedure similar to that used for preparing GfpU (Ikehara & Imura, 1981).

Crystallization. The RNase Ms/3′-GMP complex was crystallized as described before (Nonaka et al., 1989). A similar procedure was applied for preparing crystals of the RNase Ms/GfpC complex. The 3′-GMP complex crystals belong to a space group $P2_12_12_1$ with unit cell dimensions of $a = 47.0$ Å, $b = 62.8$ Å, and $c = 37.9$ Å. The GfpC complex crystals belong to the same space group with slightly different cell dimensions of $a = 46.5$ Å, $b = 60.6$ Å, and $c = 35.0$ Å. The solvent contents of 3′-GMP complex crystals and GfpC complex crystals were 50% and 43%, respectively.

X-ray Diffraction Data Acquisition. In order to improve the initial molecular replacement model of the 3′-GMP complex structure at 2.5 Å resolution (Nonaka et al., 1991), a new data set was collected to 1.9-Å resolution from one crystal using a Rigaku four-circle diffractometer modified as described earlier (Hirono et al., 1979). A total of 10 035 reflections were collected with an R_{merge} of 11.6% on observed structure amplitudes and reduced to 9223 unique reflections. The moderate quality of the diffraction data is most probably due to an unusually broad peak profile of the reflections, the average half-width along ω being in the range of 0.2–0.3 deg. For collecting data from GfpC complex, a macromolecule-oriented Weissenberg camera devised by N. Sakabe (Sakabe, 1983, 1991; Mitsui et al., 1993) installed at Beam Line $6A_2$ of the Synchrotron Radiation Source at the National Laboratory for High Energy Physics (Tsukuba, Japan) was utilized. Two data sets were collected at about 20 °C from two crystals having the approximate dimensions of 1.0 mm × 0.3 mm × 0.25 mm, one with the *a*-axis and the other with the *c*-axis chosen as the rotation axis, using Fuji imaging plates (Miyahara et al., 1986). The conditions of data collection are summarized in Table I. The data sets were processed using the program package WEIS (Higashi, 1989) as fully described by Takeuchi et al. (1991). The total number of measured reflections from the two data sets was 23 821. The two data sets were merged to yield 7554 unique reflections. The overall merging *R*-factor ($I \geq 2\sigma$) was 6.45%, which is considerably better than the above-mentioned value (11.6%) for the 3′-GMP complex, most probably due to two-dimensional recording despite a considerably broad peak profile similar to the 3′-GMP complex.

Refinement of the Structure of the RNase Ms/3′-GMP Complex. The 2.5-Å resolution model of the 3′-GMP complex (Nonaka et al., 1991) was the starting model. The restrained

Table I: Experimental Conditions of Weissenberg Data Collection for the RNase Ms/GfpC Complex Using Synchrotron Radiation[a]

	crystal 1	crystal 2
rotation axis	*a*	*c*
wavelength (Å)	1.00[b]	1.04
nominal resolution limit (Å)	1.73	1.80
total number of IPs[c]	6	6
total exposure time (s)	540	540

[a] The following common conditions were employed for both crystals 1 and 2: collimator aperture = 0.2 × 0.2 mm, crystal-to-IP distance = 430 mm, oscillation angle ω/IP = 18.0°, ω rotation/IP movement = 2.0°/mm, number of oscillations/IP = 5, ω scan speed = 2.0°/s. [b] To save the beam time, we continued the measurement without resetting the wavelength adopted by the preceding users. [c] Fuji imaging plates (IP) were used in place of conventional X-ray films.

least-squares refinement program PROLSQ (Hendrickson, 1985) was used to refine the structure against 5863 independent reflections ($|F| \geq 3\sigma$) in the 6.0–1.9-Å resolution range. The 3′-GMP dictionary was prepared with the aid of the program CHARMm (Brooks et al., 1983). Water molecules were incorporated at places where the electron density in the ($|F_o|$–$|F_c|$) difference Fourier map exhibited round peaks higher than 4σ. The rather conservative criterion for identifying water molecules is due to the moderate quality of the present data (see X-ray Diffraction Data Acquisition section). At the 247th cycle, the peptide linkage Gly53–Thr54 automatically switched into the *cis* conformation. Similar automatic *trans–cis* conversion was also observed in an additional confirmative experiment employing X-PLOR (Brünger et al., 1987) despite the initial *trans* conformation with the much lower dihedral force constant imposed on the peptide linkage (Stewart et al., 1990). In fact, the *cis* conformer for the Gly53–Thr54 peptide linkage, rather than the *trans* conformer, exhibited a better fit to the electron density map in each of the three RNase Ms complexes with 3′-GMP, GfpC, and 2′-GMP (unpublished work).

Refinement of the Structure of the RNase Ms/GfpC Complex. For the structure determination of the GfpC complex, the 2.5-Å resolution model of the 3′-GMP complex (Nonaka et al., 1991) served as the initial search model. Fast rotation functions (Crowther, 1972) were calculated for the rotation search. After the orientation of the search model was determined, a translation search was done using the program BRUTE (Fujinaga & Read, 1987). A total of 233 cycles of restrained least-squares refinement was performed in the next stage as described above using 6838 independent reflections in the 6.0–1.8-Å resolution range ($|F| \geq 3\sigma$). At the 176th cycle, another conformer of GfpC was incorporated following the method of Koepke et al. (1989) devised for dealing with alternative conformations of G2′p5′G bound with RNase T_1. The GfpC dictionary was prepared using the program CHARMm.

RESULTS AND DISCUSSION

Description of Enzyme Structure

(i) Quality and Accuracy of the Refined Models. The rms deviations from ideal bond lengths are 0.016 Å for both the 3′-GMP and GfpC complexes. When the *R*-factor values were examined as a function of scattering angle (Table II), the expected coordinate errors were estimated to be ca. 0.20 and 0.25 Å for the 3′-GMP and GfpC complexes, respectively, following the methods of Luzzati (1952). Statistics related to *B*-factors are summarized in Table III. As for the Ramachandran plot, we found that Phe49 and Ser104 in the 3′-GMP complex and Ser104 in the GfpC complex are in a conformationally forbidden area. Of the two apparently problematic residues, Ser104 is adjacent to the C-terminus and slightly disordered in both crystal forms, while Phe49 in the 3′-GMP complex exhibits well-defined electron densities.

Table II: Resolution Breakdown of the *R*-Factor

	RNase Ms/3′-GMP		RNase Ms/GfpC	
resolution range (Å)	no. of reflections	shell *R*-factor	no. of reflections	shell *R*-factor
6.00–2.93	2082	0.160	1910	0.168
2.93–2.37	1670	0.201	1680	0.215
2.37–2.08	1242	0.211	1404	0.241
2.08–1.90	869	0.242	1145	0.265
1.90–1.80			699	0.295
total	5863	0.185	6838	0.204

(ii) Cis Peptide Bonds. RNase Ms has two *cis* peptide bonds. One is a *cis* imide (Tyr37–Pro38) peptide bond and the other is a *cis* amide (Gly53–Thr54) bond. For the former peptide bond, *cis* peptide restraints were imposed during the refinement of both the complex structures because the corresponding Tyr*38*–Pro*39* bond in RNase T_1 has been established to be in *cis* imide geometry (Arni et al., 1988).

For the Gly53–Thr54 linkage, on the other hand, a *trans* peptide restraint (ω = 180°) was initially imposed for the 3′-GMP complex although the corresponding Ser*54*–Pro*55* bond in RNase T_1 was found to be *cis* imide. During the restrained least-squares refinement, however, the peptide bond spontaneously switched to a *cis* conformation. As shown in Figure 1, the pentapeptide Val51(*52*)–Tyr55(*56*) segment, including the above *cis* amide bond, is in an extensively curved conformation. There is a hydrogen-bond interaction between Tyr55(*56*)O–H and Val51(*52*)C=O. Also, this highly bent region is stabilized by two hydrogen bonds, Tyr41(*42*)N–H···Tyr55(*56*)C=O and Asn80(*81*)ND2–H···Ser52(*53*)C=O. These three hydrogen bonds are also observed in the structures of RNase T_1 complexes with 2′-GMP (Arni et al., 1988), G2′p5′G (Koepke et al., 1989), and vanadate (Kostrewa et al., 1989). For this pentapeptide region, the rms deviation between the main chain of RNase Ms complexed with 3′-GMP and that of the RNase T_1 complexed with 2′-GMP is only 0.30 Å.

(iii) Secondary Structure Analyses. For assigning secondary structure elements, we adopted the following criteria for hydrogen-bond assignments: (1) the hydrogen–acceptor distance is less than 2.60 Å, (2) the donor–acceptor distance is less than 3.35 Å, (3) the N–H···O angle is more than 100°, (4) the C=O···H angle is more than 90°, and (5) the C=O···N angle is more than 100°. By inspecting thus assigned hydrogen bonds and the values of ϕ, ψ angles, we have assigned α-helix and β-sheets in the RNase Ms structure (Figure 2).

There is only one α-helix with 4.7 helical turns (Ser14–Ser30). On the N- and C-terminal sides, respectively, of the α-helix, there are a two-stranded antiparallel β-sheet (β1 and β2) and a five-stranded antiparallel β-sheet (β3–β7) (Figure 2). Arni et al. (1988) designated the above two β-sheets as "minor" and "major" sheets, respectively.

(iv) Solvent Structure. The GfpC and 3′-GMP complexes respectively contain 32 and 25 water molecules, all of which have *B*-factors below 61 Å^2 and densities above 1σ in the final $(2|F_o| - |F_c|)\exp(i\alpha_c)$ map. The number of water molecules identified is moderate due to a rather conservative criterion described in the Experimental Procedures. Among these, 18 water molecules occupy mutually common positions within 1 Å deviation when all the protein non-hydrogen atoms are superposed by a least-squares method. The average *B*-factors

Table III: Analysis of the *B*-Factors

	RNase Ms/3′-GMP			RNase Ms/GfpC		
	no. of atoms	average (Å²)	standard deviation (Å²)	no. of atoms	average (Å²)	standard deviation (Å²)
overall	852	18.6	12.2	895	22.1	10.5
protein	803	18.1	12.0	803	21.4	10.5
main chain	420	17.7	11.2	420	20.4	9.6
side chain	383	18.5	12.8	383	22.5	11.4
ligand	24	25.2	10.2	60	29.2	6.4
major conformer[a]				40	27.5	5.9
minor conformer[a]				40	28.1	7.1
water	25	29.1	14.5	32	25.9	8.8

[a] The moiety involving the 40 atoms of the GfpC ligand was assumed to be composed of one major and one minor conformer.

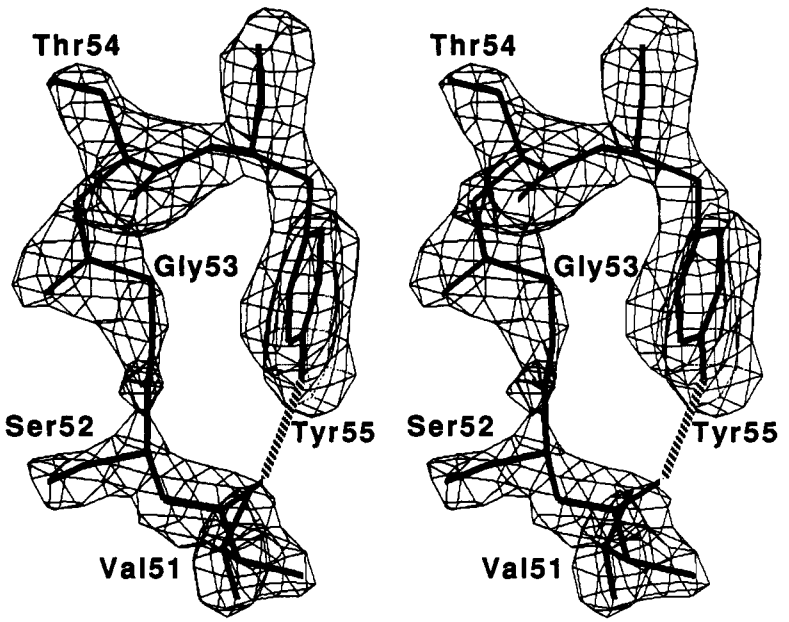

FIGURE 1: Stereo pair showing the conformation of the Val51–Tyr55 pentapeptide segment involving the Gly53–Thr54 *cis* peptide linkage as it appeared in the ($|F_o| - |F_c|$) omit Fourier map of the RNase Ms/GfpC complex specifically prepared for this segment. The contour level is 2.5σ. The hydrogen bond formed between Val51C=O and Tyr55O–H is shown by a broken line.

and accessible surface area (Lee & Richards, 1971) of these common water molecules are smaller than those of noncommon water molecules.

(*v*) *Crystal Packing.* Because the RNase Ms molecules are more densely packed in the GfpC complex crystals than in the 3′-GMP complex crystals as manifested by uniformly smaller cell dimensions of the RNase Ms/3′-GMP complex (see Crystallization under Experimental Procedures), the number (N) of intermolecular contacts less than or equal to 3.5 Å is much larger in the GfpC complex (N = 50) than that in the 3′-GMP complex (N = 27).

(*vi*) *Comments on the Further Refined Structure of the 3′-GMP Complex.* During the refinement, we found serious errors in the former interpretation (Nonaka et al., 1991) of solvent structure around the 3′-GMP ligand. We now conclude that the 3′-GMP sugar puckering is C2′-*exo* (P = 332°) rather than C2′-*endo* (Figure 4, thin line). The glycosidic bond is *syn* with the χ torsion angle (O4′–C1′–N9–C4) being 24.1°.

(*vii*) *Revision of the Amino Acid Sequence Alignment between RNases Ms and T_1.* After complete refinement of the 3′-GMP complex, we found it necessary to revise the old sequence alignment scheme [Watanabe et al. (1982); revised by Nonaka et al. (1991)]. The rms deviation between the superposed Cα-positions of RNase Ms and RNase T_1 (Arni et al., 1988) is 0.55 Å, excluding the four N-terminal residues. Significant deviation occurs only at the N-terminus except for a loop-out portion which indicates that the RNase T_1 sequence Gly*34*-Ser*35*-Asn*36*-Ser*37*, should be aligned with the RNase Ms sequence, Asp35-gap-gap-Asp36.

(*viii*) *Structure of the GfpC ligand.* The GfpC molecule takes two alternative conformations starting from the phosphate atom and ending at the 3′-terminus. The conformer lying closer to the protein surface is the major conformer (Figure 4, thick line) having an occupancy of 0.6 and an average *B*-factor of 27.5 Å². The minor conformer, as shown in Figure 3, protrudes from the protein surface and has an occupancy of 0.4 and an average *B*-factor of 28.1 Å². The glycosidic bond conformations are *syn* (χ = 47.3°), *anti* (χ = –133.4°), and *anti* (χ = 120.9°) for the guanosine, major conformer cytidine, and minor conformer cytidine, respectively. The sugar puckering modes are respectively C4′-*exo* (P = 49°), C1′-*exo* (P = 129°), and C2′-*endo*-C3′-*exo* (P = 175°). Since the electron densities for the two alternative cytosine bases are both clearly seen in the corresponding omit map (Figure 3), the glycosidic bond conformations could be determined without ambiguity. It appears that the electron density for the minor conformer cytosine is clearly seen despite its protrusion into the solvent because of a hydrogen bond between N4 and O (Cys102) of the symmetrically contiguous molecule. The puckering modes of the cytidines, however, are only tentatively assigned because of the ambiguous electron density of the ribose moieties (Figure 3).

The guanine base of GfpC is bound with RNase Ms through six strong hydrogen bonds and stacking interactions with the

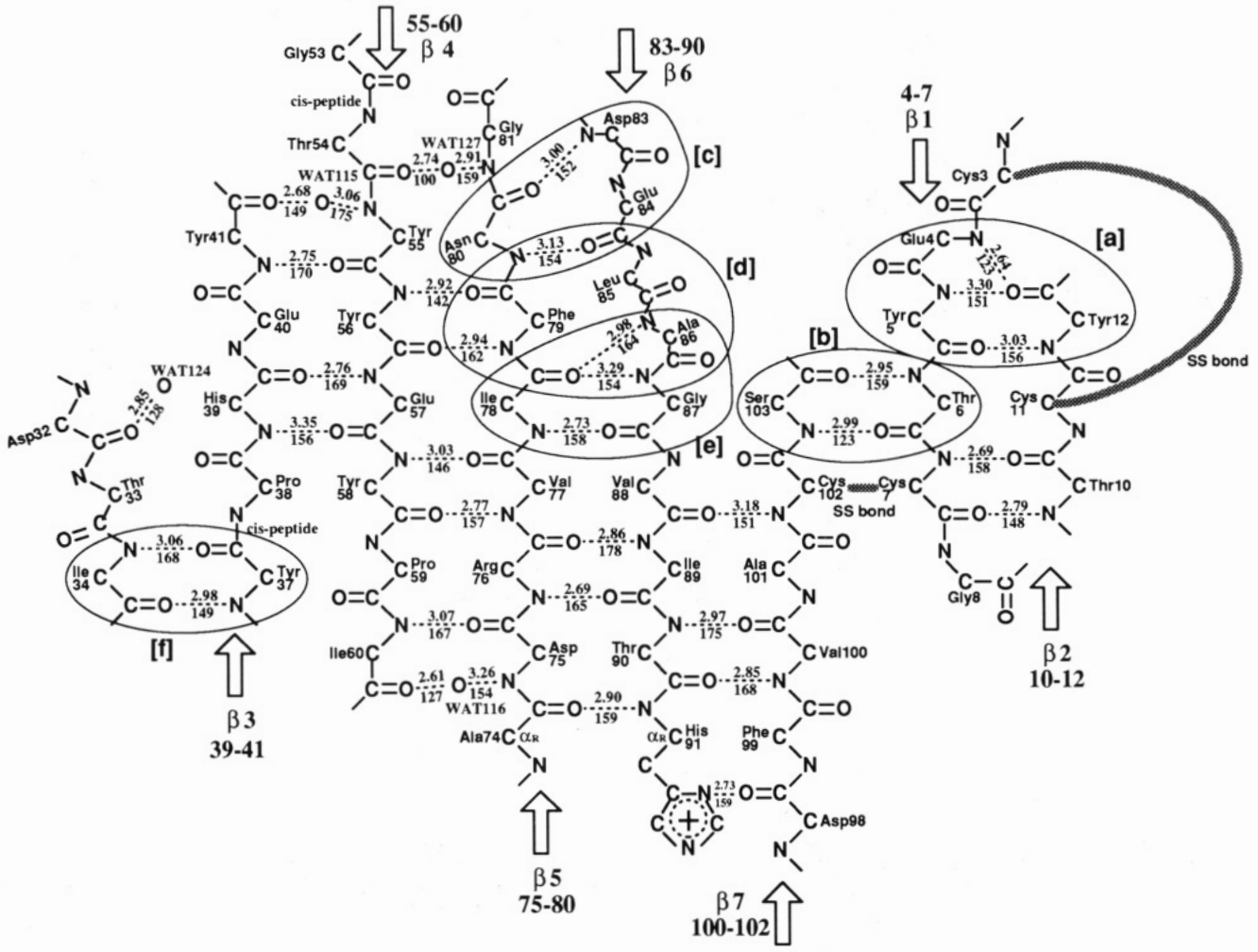

FIGURE 2: β-Structures found in the RNase Ms/3′-GMP complex and the related hydrogen-bond geometries. For each N–H···O hydrogen bond, the distance N···O and the angle N–H···O are given. The disulfide linkages are indicated by thick lines. The β-bulges (Richardson et al., 1978) and β-bridges (Kabsch & Sander, 1983) are encircled and designated as [a] through [f]. These can be classified as a "classic" β-bulge ([a] and [e]), a β-bridge ([b] and [f]), a "G1" β-bulge ([c]) and a "wide" β-bulge ([d]).

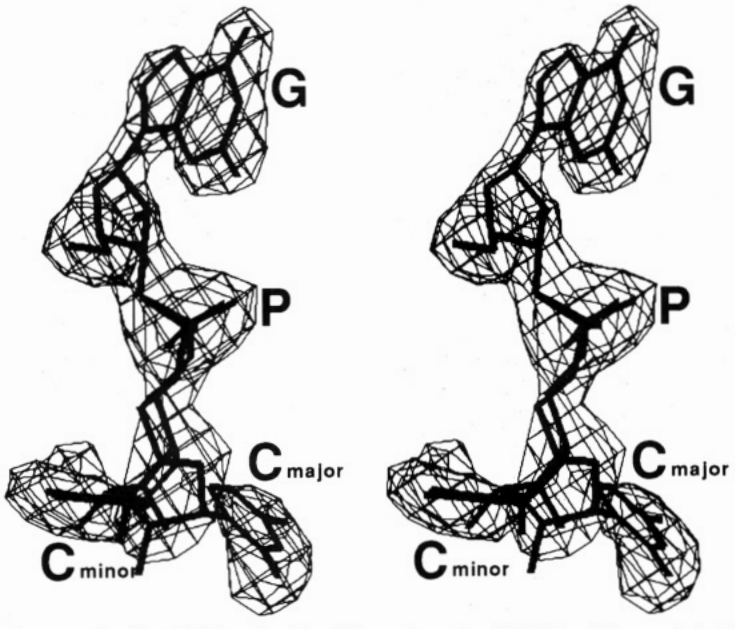

FIGURE 3: ($|F_o| - |F_c|$) omit Fourier map for the GfpC molecule. The contour level is 2.5σ. The symbols G and P designate the guanosine and phosphate moieties, respectively. The cytosine moieties appear as a superposition of one major (C_{major}) and one minor (C_{minor}) conformer.

phenol rings of Tyr41 and Tyr44 (Figure 4) in a similar way as the guanine base of 2′-GMP is bound with RNase T_1 (Arni et al., 1988). The cytosine base of the major conformer piles up on the stack of Phe99 and His91 (Figure 4) as found in the RNase T_1/G2′p5′G complex (Koepke et al., 1989), is fixed by a weak hydrogen bond (N4···Gly96O) (Table IV),

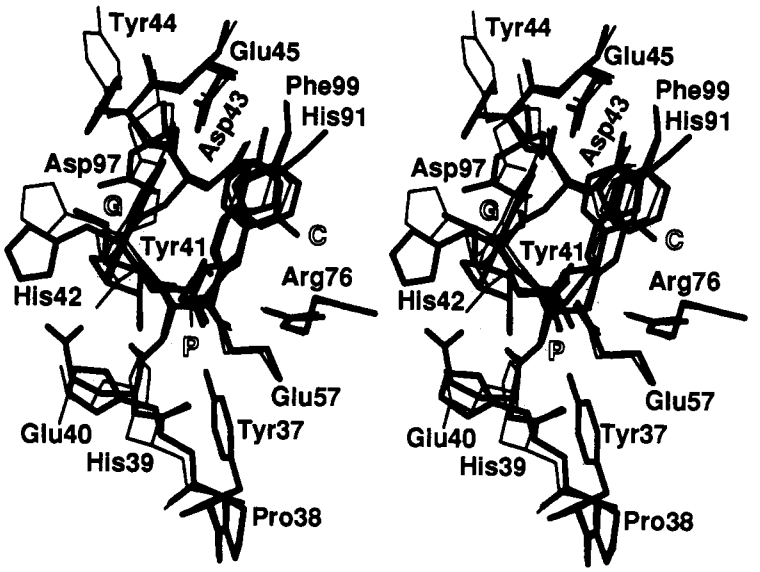

FIGURE 4: Superposition of the active site of the GfpC complex (thick lines) with that of the 3′-GMP complex (thin lines). The open letters G, P, and C designate the guanine base, the phosphate moiety, and the cytidine base (in the major conformer), respectively.

Table IV: Possible Hydrogen Bonds and Some Other Close Contacts between GfpC and RNase Ms[a]

GfpC atom[b] (X or X-H)	protein or water atom (Y or Y-H)		distance (Å) X···Y	distance (Å) X···H or H···Y	angle (deg) X···H–Y or X–H···Y	hydrogen bond evaluation[c]
GN1-H	OE1	Glu45	2.90	1.95	155.3	S
GN2-H	OE2	Glu45	2.98	1.98	171.5	S
	OH	Tyr44	3.28	3.10	91.3	
	O	Asp97	2.75	1.80	154.3	S
GN2	OH–H	Tyr44	3.28	3.67	59.5	
GO6	N-H	Tyr44	2.82	1.85	170.9	S
	N-H	Asp43	2.88	2.23	122.2	S
	N-H	His42	3.49	3.15	102.0	
GN7	N-H	His42	2.85	1.86	172.0	S
GC8	OE1	Glu1[d]	3.06			
GN9	OE1	Glu1[d]	3.32			
GC1′	OE1	Glu1[d]	3.28			
GF2′	ND1	His39	3.08			
GO4′	O5′-H	GfpC	2.95	3.78	26.3	
O1PA	OH-H	Tyr37	2.68	3.49	28.5	
	OE1-H	Glu57	2.82	2.31	111.9	W
	NH2-H	Arg76	3.08	2.15	153.1	S
	NE-H	Arg76	3.12	2.22	153.1	S
O2PA	NE2-H	His91	3.10	2.30	138.5	S
CO5′A	NH2-H	Arg76	3.16	2.35	136.7	W
	NE2-H	His91	3.49	2.64	145.7	
CO4′A	NH2-H	Arg76	3.07	2.49	116.8	W
CO3′A	O-H	Wat135	3.15	2.39	136.5	W
CO2′A	O-H	Wat135	3.08	2.32	136.7	W
CN4A-H	O	Gly96	3.20	2.36	140.6	W
O1PB	OH-H	Tyr37	2.56	3.10	48.3	
	NH2-H	Arg76	3.27	2.28	172.5	W
O2PB	OE1-H	Glu57	2.60	1.70	154.3	S
	NH2-H	Arg76	3.44	2.71	130.0	
	NH-H	Arg76	3.21	2.51	128.5	W
CO5′B	NE2-H	His91	2.73	1.98	132.7	S
CO3′B	O-H	Wat135	2.82	1.87	173.9	S
CN4B-H	O	Cys102[e]	3.10	2.20	148.5	S

[a] Hydrogen atom positions were calculated by the program X-PLOR (Brünger & Karplus, 1988). Atomic name with "-H" means that the group was regarded as a hydrogen donor. [b] The first character G, C, or O indicates the guanosine, cytidine, or phosphate moiety, respectively, while the last character A or B means that the atom belongs to the major or minor conformer, respectively. [c] S means a strong hydrogen bond satisfying the following three criteria: donor–acceptor distance ≤ 3.2 Å, hydrogen–acceptor distance ≤ 2.3 Å, and donor–H···acceptor angle ≥ 110°. W refers to weak hydrogen bond satisfying the following three criteria: donor–acceptor distance ≤ 3.35 Å, hydrogen–acceptor distance ≤ 2.6 Å, and donor–H···acceptor angle ≥ 100°. [d] Glu1 is derived from the symmetry-related $(1 - x, -1/2 + y, 3/2 - z)$ adjacent molecule. [e] Cys102 is derived from the symmetry-related $(1/2 + x, 3/2 - y, 2 - z)$ adjacent molecule.

and is surrounded by the Gly73 and Asp97 residues. We may, therefore, reasonably conclude that Gly96, Glu73, and Asp97 are related to the "N1 subsite" (Lenz et al., 1991a) [or the "B2 subsite" according to the nomenclature of RNase A (Richards & Wyckoff, 1971)]. The ribose moiety of guanosine interacts with Glu1 of the symmetrically contiguous protein molecule, His39 and His91 (Table IV). The phosphate moiety interacts with Glu57, Phe99, Tyr37, His91, and Arg76. The close contacts of Phe99 with the phosphate moiety may explain the specifically large chemical shifts observed for the aromatic

protons of the equivalent residue, Phe*100*, in RNase T_1 upon binding a series of ligands (Shimada & Inagaki, 1990).

(*ix*) *Comparison between the 3′-GMP and GfpC Complexes.* When global least-squares superposition is done for the two complexes, the rms deviation between the two complexes is 1.03 Å for 803 non-hydrogen protein atoms. Among these, 69 atoms from 20 residues deviate by more than 1.0 Å. However, when these 20 residues are excluded, the rms deviation for the residual 642 atoms is only 0.27 Å. The conformation of His39(*40*) is very much different between the two complexes: the ϕ, ψ, χ_1, and χ_2 torsion angles are respectively −153°, 156°, 71°, and 76° for the 3′-GMP complex and −121°, 149°, −158°, and −94° for the GfpC complex. The conformation change in the GfpC complex appears to be needed to avoid the short contacts between the 2′-fluorine atom of GfpC and CE1 or NE2 of His39. As a result of the swinging of the imidazole ring of His39 into the solvent region, the catalytic center in the GfpC complex is widely open, while in the 3′-GMP complex, the entrance to the catalytic center is half-shielded by the imidazole ring. It is noteworthy that His91(*92*) adopts virtually the same conformation in the two complexes.

Identification of Acid and Base Catalysts

(*i*) *Role of His91(92) as Acid Catalyst.* In the GfpC complex, His91(*92*) and Arg76(*77*) are in the vicinity of the 5′-O of the cytidine (Figure 4, thick line). Of these, Arg76-(*77*) cannot be the acid catalyst because of its high pK_a value, although one of its guanidyl nitrogen atom is within hydrogen-bond distance of the 5′-O atom (Table IV). The NE2 atom of His91(*92*), on the other hand, is toward one of the lone pairs of 5′-O and within (in the case of the minor conformer) or close to (in the case of the major conformer) hydrogen-bond distance (Table IV). Thus, as Arni et al. (1988) described for the equivalent residue His*92* of RNase T_1, it is clear that His91 of RNase Ms is the acid catalyst in the transesterification step.

(*ii*) *State of the Carboxylate Group of Glu57(58).* In the GfpC complex, the carboxylate group of Glu57(*58*) is almost perfectly shielded from the solvent, the accessible surface area being zero. The carboxylate OE1 appears to form strong hydrogen bonds with the non-ester oxygen atom (O2P) of the phosphate in the GfpC complex (Table IV). Since the pH range for crystallization was 6.0–6.1, the GfpC phosphate should carry one negative charge and no hydrogen atoms. Thus the carboxylate group of Glu75(*58*) must be protonated. A similar strong hydrogen bond was also observed in the 3′-GMP complex, which was also crystallized at pH 6.0 (Nonaka et al., 1989). A similar conclusion was drawn for the equivalent residue Glu*58* in the RNase T_1 complexes with 2′-GMP (Arni et al., 1988) and with G2′p5′G (Koepke et al., 1989). In fact, Iida and Ooi (1969) pointed out that the pK_a of a carboxylate group (presumably Glu*58*) in RNase T_1 shifts from 4.9 to 7.8 upon binding 2′-GMP.

(*iii*) *Two Conformations of His39(40).* In the RNase Ms/GfpC complex, the side chain of His39(*40*) is swung to the opposite direction from that in the RNase Ms/3′-GMP complex (Figure 4). This alternative conformation of the histidine side chain has never been reported for RNase T_1 complexed with various ligands. It is intriguing to speculate that the present RNase Ms/GfpC complex structure reflects the structure in the initial stage of catalytic reaction and the RNase Ms/3′-GMP may be the structure which appears immediately before and/or after the reaction. Thus, to determine whether the true general base is Glu57 or His39, we must always take into account this conformation change of His39.

(*iv*) *Environment of the 2′-Fluorine Atom as a Substitute for 2′-Hydroxyl Group.* In the GfpC ligand utilized in the present study, the 2′-hydroxyl group (2′-OH) in the guanosine moiety is replaced by a fluorine atom (2′-F) to prohibit the transesterification step. It appears important to examine the geometrical environment of 2′-F to identify the residue playing the role of general base in the transesterification reaction. It turns out that the side chains of both His39(*40*) and Glu57-(*58*) are close to 2′-F: 3.08 Å for ND1 of His39, 3.57 and 3.59 Å for OE2 and OE1, respectively, of Glu57. However the line F···ND1 forms an angle of 82° with the imidazole ring of His39. Also, both the lines F···OE1 and F···OE2 form angles of 65° with the side-chain carbonyl plane of Glu57. Thus, all of the 2′-F-related contacts observed in the present GfpC complex are far from ideal hydrogen bonds in view of the angles made with the plane of the acceptor groups. Therefore, assigning His39(*40*) rather than Glu57(*58*) as the base catalyst cannot be justified simply because of its significantly shorter (by ca. 0.5 Å as given above) ND1···F distance.

In the case of the RNase A/UVC complex, the 2′-O atom is approximately in the same plane as the His12 imidazole ring. This structural feature would facilitate the extraction of a proton from the 2′-OH group by the imidazole ring. Here it should be noted that NZ of Lys41 is located closer to the 2′-oxygen of UVC (2.76 Å) than the imidazole nitrogen of His12 (3.00 Å). Thus, if the identification of the base catalyst had been done solely on the basis of hydrogen-bonding geometry, the base catalyst in RNase A would have been wrongly assigned to Lys41.

The above observations taken together with the realization that the replacement of 2′-OH with 2′-F is a serious perturbation of the environment with respect to both geometric and electronic differences between the two groups, we feel that deciding between His39(*40*) and Glu57(*58*) as the base catalyst in the transesterification step is not justified by simply examining the 2′-F environment.

Comparison between Active-Site Structure of RNase Ms and Those of Other RNases

(*i*) *Rationale.* Given the inconclusive results as outlined above even after the structure of RNase Ms complexed with an apparently ideal substrate analogue, the 2′-fluoro analogue of a dinucleoside monophosphate, has been fully refined, we came to the point that the present active-site structures should be compared with those of other microbial and mammalian RNases for which a three-dimensional structure has been established. It seems reasonable to expect that at least some of them exhibit an active-site geometry similar to that of RNase Ms (as a member of the T_1 subfamily in the smaller microbial RNase family). The reasons for this expectation are the following: (1) the (3′→5′) phosphate diester linkage of the RNA substrate is common to all the above RNases, (2) the reaction scheme with the 2′,3′-cyclic nucleotide phosphate as an intermediate is the same with all the above RNases, and thus (3) the disposition of the general base and acid catalysts (such as His12 and His119 of RNase A) *relative to the substrate (analogue) molecule* is likely to be similar among at least certain types of RNases with respect to both the distance and the orientation parameters, because the reaction catalyzed by a general base or acid catalyst is essentially a quantum chemical process and thus must be very sensitive to both their distance and orientation relative to the proton donating and accepting groups on the substrate. In contrast,

such strict geometrical requirements will not be imposed on other kinds of active-site amino acid residues whose roles are either a mere electrostatic stabilization of substrates, reaction intermediates or products (such as Lys41 of RNase A most probably stabilizing the pentacovalent intermediate), or a mere building block to constitute appropriate binding sites.

Fortunately, for the mammalian RNases (represented by RNase A or its modified form RNase S), the general base and acid catalysts have long been established as His12 and His119, respectively (Richards & Wyckoff, 1971). Moreover, in the case of RNase St (Yamamaoto et al., 1981) representing the St subfamily of a smaller microbial RNase family, the His91 and Glu61 pair is the most likely catalytic pair according to chemical modification studies (Ohgi et al., 1981) and consideration of the three-dimensional structure. Here it should be emphasized that the only other histidine residue, His60, in the entire RNase St sequence cannot be the catalyst, because the imidazole ring is oriented away from the active site (Nakamura et al., 1982). Furthermore the sequence alignment between RNases St and T_1 clearly indicates that the His91, Glu61 pair of RNase St is evolutionarily equivalent to the His*92*, Glu*58* pair of RNase T_1 (Nakamura et al., 1982). With these points in mind, we performed structural comparisons as outlined in the following three sections. For this purpose, we used the coordinates of the RNase T_1/2′-GMP complex (Arni et al., 1988), the RNase St complex with a sulfate anion bound to the active site [the RNase St/SO_4^{2-} complex (Nakamura et al., 1982)], the RNase A complex with a transition-state analogue, UVC [the RNase A/UVC complex (Borah et al., 1985)], and the RNase S complex with a phosphonate analogue of uridylyl-3′,5′-adenosine (UpA), UpcA [the RNase S/UpcA complex (Richards & Wyckoff, 1973)]. The conclusion of the comparative studies will be given in the final section.

(*ii*) *RNase Ms vs RNase T_1*. Figure 5a shows a least-squares superposition of the RNase Ms/3′-GMP complex and the RNase T_1/2′-GMP complex. The superposition was done using all the Cα atoms except for three N-terminal and two C-terminal atoms (99 Cα atoms in total) resulting in the rms deviation of 0.55 Å. As seen, the disposition of His39(*40*), Glu57(*58*), and His91(*92*) is almost perfectly conserved. Thus it is clear that RNase Ms and RNase T_1 (both belonging to the T_1 subfamily of a smaller microbial RNase family) share a common catalytic geometry. The present observation also indicates that, most probably, the crystal packing does not affect the catalytic geometry to such an extent as to spoil the quantitative comparisons outlined below, since the mode of crystal packing is completely different between the RNase T_1 and RNase Ms crystals: in the RNase T_1 crystal, the major sites of intermolecular interaction are the two regions, one involving residues *1*, *2*, *7*, *9*, *10*, and *63* and the other involving residues *45*, *46*, *92*, *94*, *95*, *96*, *97*, and *100*, whereas in the RNase Ms crystal, there are four regions, the first involving residue *41*, the second involving residues *15* and *65*, the third involving residues *32*, *34*, and *69*, and the last involving residues *13*, *17*, *20*, and *86*. Furthermore, there are indications showing that the extent of dynamic fluctuation of these RNase molecules (in solution) may not be so large as to spoil the conclusion obtained from the quantitative comparisons described below. Thus, molecular dynamics calculations on RNase T_1/2′-GMP complex (Hirono & Kollman, 1990) showed an rms deviation of only 0.9 Å from the X-ray structure. Clearly, more detailed computational studies on various RNase active sites are necessary to solidify the above argument since there is an indication that certain side-chain atoms can deviate from the X-ray structure in a much more conspicuous way (Hirono & Kollman, 1990).

Another important structural feature exhibited by Figure 5a is that the mode of specific interaction between the guanine base and the enzyme active site through six hydrogen bonds (Table IV) is exactly the same between the two enzymes. This is a rather surprising feature in view of the fact that RNase Ms can hydrolyze a series of dinucleoside monophosphate substrates, ApX, where X is A, G, C, or U, as well as the corresponding series of substrates, GpX, although the rate as measured by V_{max} is ca. 280–1300-fold (depending on X) slower for the adenosine-containing substrates (Irie & Ohgi, 1976). In contrast, RNase T_1 is known to be virtually GpX specific. The above contrast in substrate specificity between RNase T_1 and RNase Ms has recently been reconfirmed by an independent study using ^{32}P-labeled GpC and ApC substrates (Nishikawa et al., 1987a). Clearly some of the hydrogen-bond interactions between the guanine base and the enzyme active sites cannot be realized when the base moiety is replaced by an adenine base. Thus the structural origin of the additional activity (albeit very weak) of RNase Ms for the adenosine-containing substrates is not clear. Further crystallographic studies aiming at finding the possible alternative binding mode for the adenosine-containing substrates would be necessary for clarifying this point. It should be recalled here that a classical example of such an apparent conflict between the crystal structure and the additional weak substrate specificity exists in the case of pancreatic RNase A (or S) as well. Thus, despite the well-known pyrimidine specificity (as a B1 base) and an apparently perfectly consistent crystal structure observed for a series of RNase S/ligand complexes (Richards & Wyckoff, 1971), it is also known that RNase A (or S) can definitely hydrolyze poly(A), although much more slowly (Beers, 1960; Imura et al., 1965).

(*iii*) *RNase Ms vs RNase St*. Figure 5b shows a least-squares superposition of the RNase Ms/GfpC complex (MS) and the RNase St/SO_4^{2-} complex (ST). The superposition was done using the 20 Cα atoms on the major β-sheet selected on the basis of good correspondence between the two structures. The rms deviation was 0.58 Å. As seen, His39(*40*), Glu57(*58*), Arg76(*77*), and His91(*92*) of RNase Ms correspond to Thr42, Glu61, Arg76, and His91 of Rnase St, respectively. Moreover, the phosphate group in the RNase Ms complex is almost exactly superposed on the sulfate anion in the RNase St complex. Thus, in agreement with the conclusion based on mere sequence alignment, the most likely catalytic pair (Glu61 and His91) in RNase St corresponds to the Glu57(*58*) [rather than His39(*40*)] and His91(*92*) pair of RNase Ms. Apart from the above residues, moreover, similarity seems to exist in the active-site geometry. For example, Phe99(MS) and Tyr37(MS) respectively occupy positions similar to those of Tyr92(ST) and Arg72(ST).

(*iv*) *RNase Ms vs RNase A (or S)*. Since there is no homology in the overall chain-folding topology between the microbial ribonuclease, RNase Ms, and mammalian pancreatic ribonuclease, RNase A (or S), the superpositions were done using the bound ligands as the reference frame. For the superposition of the RNase Ms/GfpC complex (MS) and the RNase A/UVC complex (A), the first least-squares superposition was done using the following 14 atom pairs: N1 of uracil (A) and N9 of guanine (MS), nine atom pairs between the corresponding ribose atoms, and four atom pairs between the corresponding atoms in the vanadate (A) and phosphate (MS) moieties. The second superposition was done excluding the following five atom pairs which deviated by more than 1.0 Å in the first superposition step: 2′-O (A) and 2′-F (MS), the

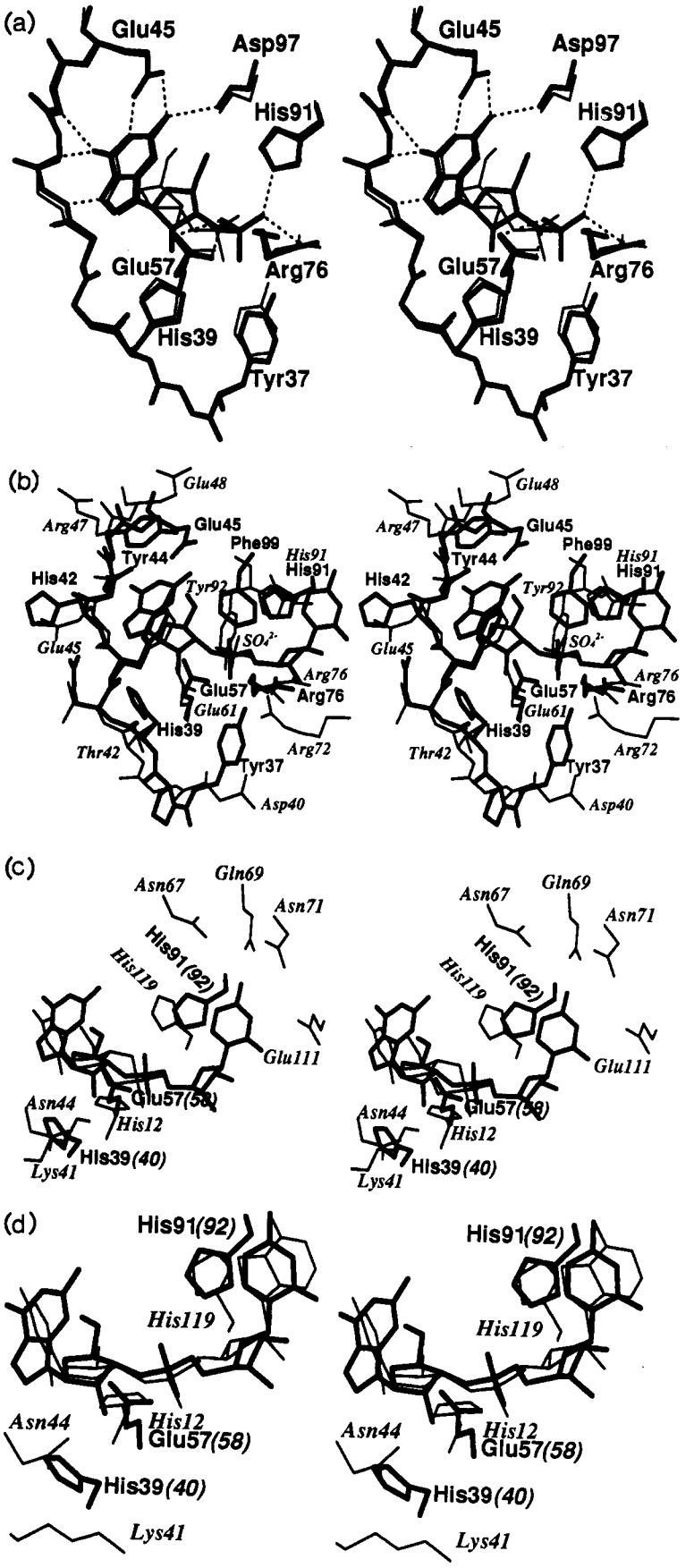

FIGURE 5: Superposition of the active sites of RNase Ms complexes with those of mammalian ribonucleases (RNase A and RNase S) and (smaller) microbial ribonucleases (RNase T_1 and RNase St) in various liganded states. (a) RNase Ms/3′-GMP complex (bold lines) vs RNase T_1/2′-GMP complex (thin lines). Strong hydrogen bonds as assigned by the described method (see Table IV, footnote *c*) are shown by broken lines. (b) RNase Ms/GfpC complex (bold lines) vs RNase St/SO_4^{2-} complex (thin lines, italic residue identification). (c) RNase Ms/GfpC complex (bold lines) vs RNase A/UVC complex (thin lines, italic residue identification). (d) RNase Ms/GfpC complex (bold lines) vs RNase S/UpcA complex (thin lines, italic residue identification). In (a) and (b) the amino acid residues of RNase Ms are labeled only by RNase Ms numbering, while in (c) and (d), they are followed by italic RNase T_1 numbering shown in parentheses.

corresponding 5′-C atoms and 3′-O atoms, O3V (A) and O2P (MS), and O1V(A) and O1P (MS). The second superposition yielded the rms deviation of 0.59 Å and the result is shown in Figure 5c. This shows that (1) among the three candidates for the catalytic residues of RNase Ms, the side chains of Glu57(*58*) and His91(*92*) are respectively closely located to the imidazole rings of His12 (established as the base catalyst) and His119 (established as the acidic catalyst) of RNase A, (2) the imidazole ring of His39(*40*) lies between the side chains of Lys41 and Asn44 in RNase A, ND1 of His39(MS) being adjacent to both NZ of Lys41 (A) and OD1 of Asn44 (A), (3) the cytosine base of the GfpC ligand bound to RNase Ms apparently nicely fit to the B2 pocket (Richards & Wyckoff, 1971) of RNase A, which is formed by the side chains of Asn67, Gln69, Asn71, and Glu111, and moreover (4) the residual portion of the GfpC ligand gives rise to no serious atomic collision with the surface of RNase A except for the six-membered ring portion of the guanine base. The above exception is understandable since the B1 pocket of RNase A is known to be pyrimidine specific.

The intriguing observations 1–4 described above prompted us to compare the RNase Ms/GfpC structure with that of a dinucleoside monophosphate complex of pancreatic RNase. The superposition between the RNase Ms/GfpC complex (MS) and the RNase S/UpcA complex (S) was initially done in a way similar to that described above. The final least-squares superposition was done using the following 18 atom pairs: N1 of uracil (S) and N9 of guanine (MS); five ribose atom pairs of uridine (S) and guanosine (MS) moieties, excluding the largely deviating 5′-C atom; five phosphate atom pairs; five ribose atom pairs of adenosine (S) and cytidine (MS), excluding largely deviating 2′-O, 3′-C, and 3′-O; and N9 of adenine (S) and N1 of cytosine (MS). [The coordinates for the 5′-O of uridine were not found in the list of Richards & Wyckoff (1973).] The final superposition yielded an rms deviation of 0.52 Å and the result is shown in Figure 5d. As seen, the side chains of Glu57(*58*) and His91(*92*) in RNase Ms almost exactly overlap the imidazole rings of His12 (established as the base catalyst) and His119 (established as the acid catalyst), respectively. Again the imidazole ring of His39(*40*) in RNase Ms lies between the side chains of Lys41 and Asn44, as in the above comparison with the RNase A/UVC complex.

(*v*) *More Quantitative Comparisons between the Active-Site Geometries of Mammalian RNases and Those of T_1 Subfamily RNases.* For this purpose, we defined group I RNases and group II RNases as follows. The group I RNases contain mammalian pancreatic RNase A, its modified form RNase S (Richards & Wyckoff, 1971), and a glycosylated form, RNase B. The group II RNases contain RNase T_1 and highly homologous RNase Ms and RNase F_1 (Vassylyev et al., 1993), all of which belong to the T_1 subfamily of the smaller microbial RNase family. In addition, we included a microbial ribonuclease RNase Rh from *R. niveus* to Group I. RNase Rh belongs to the larger microbial RNase family (see introduction) and has a unique chain-folding topology distinct from both smaller microbial RNases and mammalian RNases (Kurihara et al., 1992). However the disposition of the potential catalytic pair (His46 and His109) in RNase Rh (Ohgi et al., 1992) is very similar to that of mammalian RNases (Kurihara, 1991; Kurihara et al., 1992).

In Figure 6, the "potential catalytic groups" A, B, and C are introduced and the identity of the actual amino acid residue assigned to each group is listed in Table V. It should be noticed here that the functional assignment of the groups A, B, and C as the "acid catalyst", "base catalyst", and "cationic stabilizer" is based on many lines of established facts in the case of group I RNases (except for RNase Rh), but for group II RNases, the assignment as proposed in Table V is only the consequence of the following statistical considerations. In other words, at the beginning, the functions of the groups A, B, and C were regarded as unknown for group II RNases.

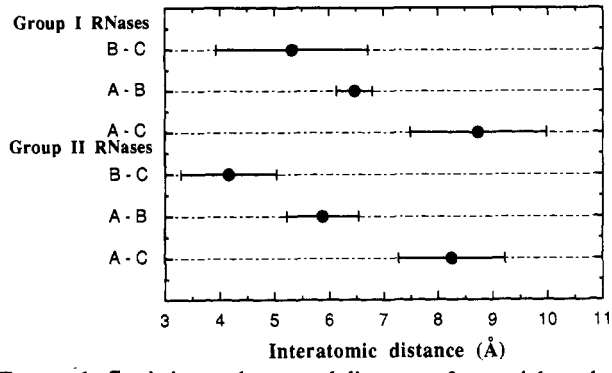

FIGURE 6: Statistics on the mutual distances of potential catalytic groups. Mutual distances have been calculated for the "relevant atoms" of the three "potential catalytic groups" A, B, and C, which are respectively the acid catalyst, the base catalyst, and cationic stabilizer as defined in Table V. For a His residue, either ND1 or NE2 [whichever is closer to the relevant atoms (5′-O or 2′-O) of the substrate analogues] was chosen as the relevant atom, while for a Glu residue, either OE1 or OE2 [whichever is closer to the relevant atom (2′-O) of the substrate analogues] was chosen as the relevant atom. For a Lys residue, NZ was chosen. The intergroup distances, A–B, B–C, and A–C were separately averaged for Group I and Group II RNases (see text for definition). For each intergroup distance for each RNase group, a bar representing 1 σ deviation from the average value (solid circle) is shown. The calculations are based on either the coordinates deposited in the Protein Data Bank (Bernstein et al., 1977; code numbers in parentheses) or those derived in the present laboratory as shwon below. Group I: RNase Rh (Kurihara et al., 1992); RNase A [Howlin et al., 1989 (3RN3); Aguilar et al., 1992 (1RNC, 1RND); Martin et al., 1987 (1SRN); Wlodawer et al., 1986 (5RSA); Borah et al., 1985 (6RSA); Wlodawer et al., 1988 (7RSA); Nachman et al., 1990 (8RSA, 9RSA); Tilton et al., 1992 (1RAT ~ 9RAT); Weber et al., 1985 (1RSM)]; RNase S [Fletterick & Wyckoff, 1975 (1RNS); Kim et al., 1992 (2RNS); Varadarajan & Richards, 1992 (1RBC ~ 1RBI); Richards & Wyckoff, 1973 (UpcA complex)]; RNase B [Williams et al., 1987 (1RBB)]. Group II: RNase T_1 [Arnie et al., 1988 (1RNT); Koepke et al., 1989 (2RNT); Kostrewa et al., 1989 (3RNT); Koellner et al., 1992 (4RNT); Lenz et al., 1991b (5RNT); Ding et al., 1991 (6RNT); Koellner et al., 1991 (7RNT); Ding et al., 1992 (8RNT); Oyanedel et al., 1991 (9RNT); Granzin et al., 1992 (1RGL); Zegers et al., 1992 (2AAD)]; RNase Ms (present study (1RMS, 1RDS) and unpublished work (2′-GMP complex)]. As for 1RNT, 5RNT, and 2AAD, His*92*CE1 atoms were chosen as the relevant atom instead of His*92*NE2 atom, because the CE1 atoms occupy the same positions as His*92*NE2 atoms of other (2RNT, 3RNT, 4RNT, 6RNT, 7RNT, 8RNT, 9RNT, and 1RGL) RNase T_1 molecules.

Examining the statistics for group I RNases (upper half of Figure 6), we note that the distance A–B [the distance between the acid catalyst (e.g., His119 in RNase A) and the base catalyst (e.g., His12 in RNase A); see the legend to Figure 6 for detailed definition] has an average value of ca. 6.5 Å with very small dispersion (ca. 0.3 Å), while the other pairs A–C and B–C involving group C (cationic stabilizer, e.g., Lys41 in RNase A) respectively give significantly smaller (ca. 5.3 Å) and larger (ca. 8.7 Å) average values with much larger dispersion (ca. 1.4 and 1.2 Å, respectively). The much smaller dispersion of the A–B distance is easy to understand considering the geometrical precision required by the quantum chemical procedures exerted by the acid and base catalysts (see Rationale above).

When we turn to the statistics for group II RNases (lower half of Figure 6), we note that the pair A–B exhibits

Table V: Structurally Conserved Residues in the Catalytic Centers of Various Ribonucleases

region[a]	symbol[b]	RNase A	RNase Rh[c]	RNase T_1	RNase Ms	barnase[d]	RNase St
1	A	His119	His46	His92	His91	His102	His91
2	B	His12	His109	Glu58	Glu57	Glu73	Glu61
3	C	Lys41	Lys108			Lys27	
3	C			His40	His39		
3		Asn44				Asp54	Thr42
4		Gln11	Glu105	Arg77	Arg76	?	Arg76
5		Lys7	His104	Tyr38	Tyr37	?	Arg72
6				Phe100	Phe99	?	Tyr92
6		Phe120	Trp49				
6				Tyr42	Tyr41	?	Phe44

[a] Each region represents an approximately equivalent three-dimensional region in the active-site superpositions. [b] Each of these symbols implies a functionally equivalent group of residues (see text). The same symbols are also used in Figure 6. [c] Temporary assignments (Kurihara, 1991; Kurihara et al., 1992; Ohgi et al., 1992). [d] This superposition was done by Steyaert et al. (1990).

significantly smaller dispersion (ca. 0.7 Å) and an average value (ca. 5.9 Å) closest to the A–B distance (ca. 6.5 Å) of group I RNases. The other pairs B–C and A–C exhibit much more deviated average values (ca. 4.2 and 8.2 Å, respectively) and much larger dispersion (ca. 0.9 and 1.0 Å, respectively). Thus it appears most likely that the pair A–B in group II RNases also represents the acid–base catalytic pair. Moreover, among His*92* and Glu*58* (of RNase T_1) constituting the A–B pair, His*92* has already been established as the acid catalyst [see Role of His91(*92*) as Acid Catalyst above]. Therefore, here again, it is very tempting to conclude that the base catalyst must be Glu*58* (rather than His*40*) in RNase T_1 [or Glu57 (rather than His39) in RNase Ms] in agreement with the previous conclusion [see RNase Ms vs RNase A (or S) above]. It might be argued that the above considerations can be irrelevant because they are solely based on the static crystal structures, which may be significantly distorted through crystal packing interactions. Such a possible argument, however, is most probably irrelevant as indicated through crystal packing comparisons and molecular dynamics calculations carried out for RNase T_1 and RNase Ms systems (see RNase Ms vs RNase T_1 above).

(*vi*) *Other Comparisons, Further Evolutionary Considerations, and Final Comments.* Apart from the above four kinds of active-site comparison, Steyaert et al. (1990) superimposed the active site of RNase T_1 with that of barnase, a prokaryotic microbial RNase from *Bacillus amyloliquefaciens*. In the present classification scheme (see introduction), barnase is the leading member of the barnase subfamily of the smaller microbial RNase family. The three-dimensional structure was solved by Mauguen et al. (1982). On the basis of this structure, Steyaert et al. (1990) found that Glu73 and His102 of barnase are respectively the functional analogues of Glu*58* and His*92* in RNase T_1. This conclusion is incorporated in Table V. As in the case of RNase St (see RNase Ms vs RNase St), the only other histidine residue in the entire barnase molecule, His18, cannot be the catalytic residue of any function, the imidazole ring being located far from the active site. Indeed, in another enzyme of the barnase subfamily, RNase Bi from *Bacillus intermedius* (having 80% sequence identity with barnase), His18 is replaced by lysine. It should be stressed here that, in RNase Bi, His101 (equivalent to His102 of barnase) is the only histidine residue in the entire sequence of RNase Bi. In considering a similar argument on RNase St (given in RNase Ms vs RNase St above), it can be argued, from geometrical considerations alone, that the base and acid catalytic pair cannot be His–His, at least in the RNase St and barnase subfamilies of the smaller microbial RNase family. Indeed detailed enzyme kinetic studies have indicated the pairs Glu73, His102 and Glu58, His92 as the base, acid catalytic pairs for barnase (Mossakowska et al., 1989) and (intact) RNase T_1 (Steyaert et al., 1990), respectively.

The sequence alignment of smaller microbial RNases (Hill et al., 1983) shows that Glu*58* and His*92* (of RNase T_1) are conserved throughout all the smaller microbial RNases, but His*40* (of RNase T_1) is conserved only among the eukaryotic enzymes of the RNase T_1 subfamily (RNases T_1, Ms, F_1, U2, and C2), the corresponding residues in RNases St (belonging to the St subfamily), barnase, and RNase Bi (both belonging to the barnase subfamily) being Thr, Asp, and Asp, respectively. Here it should be noticed that the RNases of the St and barnase subfamilies are of prokaryotic origin [see Table I of Hill et al. (1983)]. Therefore, using such a classification, it can be argued that, in *eukaryotic* smaller microbial RNases, both His*40* and Glu*58* (of RNase T_1) are the possible base catalysts (forming either a His–His or a Glu–His catalytic pair), while in *prokaryotic* smaller microbial RNases, only the glutamic residue corresponding to Glu*58* of RNase T_1 can be the base catalyst (forming a Glu–His catalytic pair). In this respect, it is evolutionarily not unreasonable to interpret, as Nishikawa et al. (1987b) did, the results of the mutation experiments on RNase T_1 (see introduction) as indicating a His*40*–His*92* catalytic pair in the limited kingdom of *eukaryotic* smaller microbial RNases. The existence of an almost established His12–His119 catalytic pair in mammalian pancreatic RNases (Richards & Wyckoff, 1971) may well be a strong chemical support for the His–His catalytic pair in certain microbial RNases.

However, the detailed comparison of the three-dimensional structures of various RNase active sites as outlined above consistently favors the Glu–His catalytic pair even in the enzymes of the *eukaryotic* smaller microbial RNase family. Detailed examination of the kinetic parameters for intact (wild-type) RNase T_1 and a series of mutated RNase T_1 (Steyaert et al., 1990) also supported the Glu–His catalytic pair in the *intact* RNase T_1. Then how can the apparently unusual results of the mutation experiments on RNase T_1 (Nishikawa et al., 1987b; Steyaert et al., 1990; see introduction) be interpreted? The present authors have no unique mechanistic idea as to this question. However, we feel that the idea that His*40* plays the role of the base catalyst in the absence of Glu*58* (Steyaert et al., 1990) is at least structurally reasonable and appears to be based on sufficient enzyme kinetic evidence. It is intriguing to note that, if this were the case, it means that even such a small enzyme as RNase T_1 retains a flexibility, apparently to guard against perturbations imposed from outside. Such flexibility seems to be one of the characteristics of biological (as opposed to mechanical) systems. Finally we would like to stress that the three-dimensional structure of a mutated RNase T_1 (or its homologue) in which Glu*58* (but not His*40*)

is spoiled has to be crystallographically elucidated before we arrive at a final conclusion.

ACKNOWLEDGMENT

We thank Dr. H. Mizuno for access to his four-circle diffractometer. We are indebted to Dr. A. Takenaka for advice on the use of the fast rotation function program. Thanks are also due to Professor N. Sakabe and Dr. A. Nakagawa for help with data collection and processing at the Photon Factory, Tsukuba, Japan. We also thank Drs. K. Ohgi and H. Watanabe for help in preparing the RNase Ms samples. The reviewers' comments were extremely useful in improving the manuscript.

REFERENCES

Aguilar, C. F., Thomas, P. J., Mills, A., Moss, D. S., & Palmer, R.A. (1992) *J. Mol. Biol. 224*, 265.

Arni, R, Heinemann, U., Tokuoka, R., & Saenger, W. (1988) *J. Biol. Chem. 263*, 15358–15368.

Beers, R. F. (1960) *J. Biol. Chem. 235*, 2393–2399.

Bernstein, F. C., Koetzle, T. F., Williams, G. J. B., Meyer, E. F., Jr., Brice, M. D., Rodgers, J. R., Kennard, O., Shimanouchi, T., & Tasumi, M. (1977) *J. Mol. Biol. 112*, 535–542.

Blackburn, P., & Moore, S. (1982) *Enzymes (3rd Ed.) 15*, 317–433.

Borah, B., Chen, C.-W., Egan, W., Miller, M., Wlodawer, A., & Cohen, J. S. (1985) *Biochemistry 24*, 2058–2067.

Brooks, B. R., Bruccoleri, R. E., Olafson, B. D., States, D. J., Swaminathan, S., & Karplus, M. (1983) *J. Comput. Chem. 4*, 187–217.

Brünger, A. T., & Karplus, M. (1988) *Proteins: Struct., Funct., Genet. 4*, 148–156.

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

Crowther, R. A. (1972) in *The Molecular Replacement Method* (Rossmann, M. G., Ed.) pp 173–178, Gordon and Breach, New York.

Ding, J., Koellner, G., Grunert, H.-P., & Saenger, W. (1991) *J. Biol. Chem. 266*, 15128.

Ding, J., Choe, H.-W., Granzin, J., & Saenger, W. (1992) *Acta Crystallogr. B48*, 185–191.

Fletterick, R. J., & Wyckoff, H. W. (1975) *Acta Crystallogr. A31*, 698.

Fujinaga, M., & Read, R. J. (1987) *J. Appl. Crystallogr. 20*, 517–521.

Granzin, J., Lutzke, R. P., Lanndt, O., Grunert, H.-P., Heinemann, U., Saenger, W., & Hahn, U. (1992) *J. Mol. Biol. 225*, 533.

Hartley, R. W. (1980) *J. Mol. Evol. 15*, 355–358.

Heinemann, U., & Saenger, W. (1982) *Nature (London) 299*, 27–31.

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

Higashi, T. (1989) *J. Appl. Crystallogr. 22*, 9–18.

Hill, C., Dodson, G., Heinemann, U., Saenger, W., Mitsui, Y., Nakamura, K., Borisov, S., Tischenko, G., Polyakov, K., & Pavlovsky, S. (1983) *Trends Biochem. Sci. 8*, 364–369.

Hirono, S., & Kollman, P. A. (1990) *J. Mol. Biol. 212*, 197–209.

Hirono, S., Nakamura, K. T., Iitaka, Y., & Mitsui, Y. (1979) *J. Mol. Biol. 131*, 855–869.

Howlin, B., Moss, D. S., & Harris, G. W. (1989) *Acta Crystallogr. A45*, 851.

Iida, S., & Ooi, T. (1969) *Biochemistry 8*, 3897–3902.

Ikehara, M., & Imura, J. (1981) *Chem. Pharm. Bull. 29*, 2408–2412.

Imura, N., Irie, N., & Ukita, C. (1965) *J. Biochem. (Tokyo) 58*, 264–269.

Irie, M., & Ohgi, K. (1976) *J. Biochem. (Tokyo) 80*, 39–43.

Kabsch, W., & Sander, C. (1983) *Biopolymers 22*, 2577–2637.

Kim, E. E., Varadarajan, R., Wyckoff, H. W., & Richards, F. M. (1992) *Biochemistry 31*, 12304–12314.

Koellner, G., Grunert, H.-P., Landt, O., & Saenger, W. (1991) *Eur. J. Biochem. 201*, 199.

Koellner, G., Choe, H.-W., Heinemann, Y., Grunert, H.-P., Zouni, A., Hahn, U., & Saenger, W. (1992) *J. Mol. Biol. 224*, 701–713.

Koepke, J., Maslowska, M., Heinemann, U., & Saenger, W. (1989) *J. Mol. Biol. 206*, 475–488.

Kostrewa, D., Choe, H.-W., Heinemann, U., & Saenger, W. (1989) *Biochemistry 28*, 7592–7600.

Kurihara, H. (1991) Ph.D. Thesis, The University of Tokyo, Tokyo, Japan.

Kurihara, H., Mitsui, Y., Ohgi, K., Irie, M., Mizuno, H., & Nakamura, K. T. (1992) *FEBS Lett. 306*, 189–192.

Lee, B., & Richards, F. M. (1971) *J. Mol. Biol. 55*, 379–400.

Lenz, A., Cordes, F., Heinemann, Y., & Saenger, W. (1991a) *J. Biol. Chem. 266*, 7661–7667.

Lenz, A., Heinemann, U., Maslowska, M., & Saenger, W. (1991b) *Acta Crystallogr. B47*, 521.

Luzzati, P. V. (1952) *Acta Crystallogr. 5*, 802–810.

Martin, P. D., Doscher, M. S., & Edwards, B. F. P. (1987) *J. Biol. Chem. 262*, 15930–15938.

Mauguen, Y., Hartley, R. W., Dodson, E. J., Dodson, G. G., Bricogne, G., Chothia, C., & Jack, A. (1982) *Nature (London) 297*, 162–164.

Mitsui, Y., Nonaka, T., Senda, T., & Nakamura, K. T. (1993) *Synchrotron Radiation in Biology* (Helliwell, J. R., Ed.) Oxford University Press, Oxford, England. (In press).

Miyahara, J., Takahashi, K., Amemiya, Y., Kamiya, N., & Satow, Y. (1986) *Nucl. Instrum. Methods A246*, 572–578.

Mossakowska, D. E., Nyberg, K., & Fersht, A. R. (1989) *Biochemistry 28*, 3843–3850.

Nachman, J., Miller, M., Gilliland, G. L., Carty, R., Pincus, M., & Wlodawer, A. (1990) *Biochemistry 29*, 928–937.

Nakamura, K. T., Iwahashi, K., Yamamoto, Y., Iitaka, Y., Yoshida, N., & Mitsui, Y. (1982) *Nature (London) 299*, 564–566.

Nishikawa, S., Morioka, H., & Ikehara, M. (1987a) in *Tampakusitsu Kogaku Nyumon* (Introduction to Protein Engineering) (Sakiyama, F., Ed.), pp 146–164, Shujunsha, Tokyo, Japan (in Japanese).

Nishikawa, S., Morioka, H., Kim, H. J., Fuchimura, K., Tanaka, T., Uesugi, S., Hakoshima, T., Tomita, K., Ohtsuka, E., & Ikehara, M. (1987b) *Biochemistry 26*, 8620–8624.

Nonaka, T., Mitsui, Y., Nakamura, K. T., Watanabe, H., Ohgi, K., & Irie, M. (1989) *J. Mol. Biol. 207*, 853–854.

Nonaka, T., Mitsui, Y., Irie, M., & Nakamura, K. T. (1991) *FEBS Lett. 283*, 207–209.

Ohgi, K., & Irie, M. (1975) *J. Biochem. (Tokyo) 77*, 1085–1094.

Ohgi, K., Watanabe, H., Emman, K., Yoshida, N., & Irie, M. (1981) *J. Biochem. (Tokyo) 90*, 113–123.

Ohgi, K., Horiuchi, H., Watanabe, H., Iwama, M., Takagi, M., & Irie, M. (1992) *J. Biochem. (Tokyo) 112*, 132–138.

Oyanedel, J. M., Choe, H.-W., & Saenger, W. (1991) *J. Mol. Biol. 222*, 335.

Richards, F. M., & Wyckoff, H. W. (1971) *Enzymes (3rd Ed.) 4*, 647–806.

Richards, F. M., & Wyckoff, H. W. (1973) in *Atlas of Molecular Structures in Biology 1. Ribonuclease S* (Phillips, D. C., & Richards, F. M., Ed.) pp 7–14, Clarendon Press, Oxford, England.

Richardson, J. S., Getzoff, D., & Richardson, D. C. (1978) *Proc. Natl. Acad. Sci. U.S.A. 75*, 2574–2578.

Sakabe, N. (1983) *J. Appl. Crystallogr. 16*, 542–545.

Sakabe, N. (1991) *Nucl. Instr. Methods A303*, 448–463.

Sevcik, E., Dodson, E. J., & Dodson, G. G. (1991) *Acta Crystallogr. B47*, 240.

Shimada, I., & Inagaki, F. (1990) *Biochemistry 29*, 757–764.

Stewart, D. E., Sarkar, A., & Wampler, E. (1990) *J. Mol. Biol. 214*, 253–260.
Steyaert, J., Hallenga, K., Wyns, L., & Stanssens, P. (1990) *Biochemistry 29*, 9064–9072.
Sugio, S., Amisaki, T., Ohishi, H., Tomita, K., Heinemann, U., & Saenger, W. (1985a) *FEBS Letters 181*, 129–132.
Sugio, S., Oka, K.-I., Ohishi, H., Tomita, K., & Saenger, W. (1985b) *FEBS Lett. 183*, 115–118.
Takahashi, K. (1970) *J. Biochem. (Tokyo) 67*, 833–839.
Takahashi, K., & Moore, S. (1982) *Enzymes (3rd Ed.) 15–B*, 435–468.
Takahashi, K., Stein, W. H., & Moore, S. (1967) *J. Biol. Chem. 242*, 4682.
Takeuchi, Y., Satow, Y., Nakamura, K. T., & Mitsui, Y. (1991) *J. Mol. Biol. 221*, 309–325.
Tilton, R. F., Jr., Dewan, J. C., & Petsko, G. A. (1992) *Biochemistry 31*, 2469.
Varadarajan, R., & Richards, F. M. (1992) *Biochemistry 31*, 12315–12327.
Vassylyev, D. G., Katayanagi, K., Ishikawa, K., Tsujimoto-Hirano, M., Danno, M., Pähler, A., Matsumoto, O., Matsushima, M., Yoshida, H., & Morikawa, K. (1993) *J. Mol. Biol. 230*, 979–996.
Watanabe, H., Ohgi, K., & Irie, M. (1982) *J. Biochem. (Tokyo) 91*, 1495–1509.
Weber, P. C., Sheriff, S., Ohlendorf, D. H., Finzel, B. C., & Salemme, F. R. (1985) *Proc. Natl. Acad. Sci. U.S.A. 82*, 8473.
Williams, R. L., Greene, S. M., & McPherson, A. (1987) *J. Biol. Chem. 262*, 16020–16031.
Wlodawer, A., Borkakoti, N., Moss, D. S. & Howlin, B. (1986) *Acta Crystallogr. B42*, 379–387.
Wlodawer, A., Svensson, L. A., Sjolin L., & Gilliland, G. L. (1988) *Biochemistry 27*, 2705–2717.
Yamamoto, Y., Iwahashi, K., Nakamura, K. T., Iitaka, Y., & Mitsui, Y. (1981) *Nucleic Acid Res. Symp. Ser. 10*, 227–230.
Zegers, I., Verhelst, P., Choe, H.-W., Steyaert, J., Heinemann, U., Saenger, W., & Wyns, L. (1992) *Biochemistry 31*, 11317–11325.