Crystal Structure of *Ustilago sphaerogena* Ribonuclease U_2 at 1.8 Å Resolution†

Shuji Noguchi,‡ Yoshinori Satow,‡,* Tsuneko Uchida,§ Chizuko Sasaki,|| and Takao Matsuzaki||

Faculty of Pharmaceutical Sciences, University of Tokyo, Hongo, Bunkyo-ku, Tokyo 113, Japan, Mitsubishi Kagaku Institute for Life Sciences, Minamiooya, Machida, Tokyo 195, Japan, and Research Center, Mitsubishi Kagaku Corporation, Midori-ku, Yokohama, Kanagawa 227, Japan

Received May 30, 1995; Revised Manuscript Received September 11, 1995⊗

ABSTRACT: The crystal structure of purine-specific ribonuclease (RNase) U_2 from *Ustilago sphaerogena* has been solved by the molecular replacement method using RNase T_1 as a search model. The structure, with 114 amino acid residues, 141 water molecules, and a sulfate ion, is refined to an *R* factor of 0.143 at 1.8 Å resolution. As evidenced by the electron densities, residues 49 and 50 are revised to Glu 49 and Asp 50, respectively, and also Asp 45 is identified as a β-isomerized form to L-isoaspartate with a β-peptide linkage. RNase U_2 consists of a β-hairpin at residues from 7 to 14, a 4.4-turn α-helix from 16 to 32, a central β-sheet with five strands, and a protruding β-turn from 74 to 77. As for the catalytic site residues, His 41, Glu 62, and Arg 85 are located as constituents of the central β-sheet, and Tyr 39 and His 101 are situated at either end of the β-sheet. The side chains of Tyr 39, Glu 62, Arg 85, and His 101 are hydrogen-bonded to the sulfate ion which marks the RNA phosphate position. Though the side chain of His 41 is pointing away from the sulfate, small conformational adjustments of His 41 enable the side chain to interact with either the phosphate or the ribose group of RNA. The loop region from Tyr 44 to Asp 50 is ascribed to the base recognition site where Glu 49 is involved in adenine recognition. β-Isomerized Asp 45 suggests that this region is conformationally flexible and alterable.

Ribonuclease (RNase)[1] U_2 is an endoribonuclease secreted by the smut fungus, *Ustilago sphaerogena* (Arima *et al.*, 1968a). It consists of 114 amino acids with three disulfide bonds (Sato & Uchida, 1975a; Kanaya & Uchida, 1986), and preferentially cleaves the 3′-phosphodiester bond of adenylic acid in RNA. This preference is in contrast to other homologous microbial RNases, such as RNase T_1 from *Aspergillus oryzae*, which share a common preference for guanylic acid. Figure 1 shows the alignment of the amino acid sequence of RNase U_2 compared to that of RNase T_1 (Takahashi, 1985), based on this structural study. The identity in these sequences is 28%. In RNase T_1, residues relevant to the RNA hydrolysis are His 40, Glu 58, Arg 77, and His 92. These are conserved also in RNase U_2, each corresponding to His 41, Glu 62, Arg 85, and His 101. Therefore, the mechanism of the RNA hydrolysis by RNase U_2 could be compared to that of RNase T_1.

It has been proposed that the hydrolysis by RNase T_1 takes place in two steps (Takahashi & Moore, 1982; Heinemann & Hahn, 1989). The first step of the hydrolysis is a transphosphorylation which yields a 2′,3′-cyclic phosphodiester and a 5′-hydroxy product. In this step, Glu 58 acts as a general base (Steyaert *et al.*, 1990) and His 92 acts as a general acid, while the involvement of His 40 as a general base has been proposed (Nishikawa *et al.*, 1986, 1987; Gohda *et al.*, 1994). In the second step, the roles of Glu 58 and His 92 are reversed, and the 2′,3′-cyclic phosphodiester is hydrolyzed to 3′-nucleotide. The molecular activity of RNase U_2 for 2′,3′-cyclic AMP is 100 times smaller than that for the dinucleotide ApC (Egami *et al.*, 1980), that is, the second step is much slower than the first step in the case of RNase U_2. As for the RNA cleavage by bovine RNase A, Cuchillo *et al.* (1993) proposed that the 2′,3′-cyclic phosphodiester is a true product. This could be also applicable to RNase U_2.

RNase U_2 is catalytically active at the pH range from 3.5 to 5.5 (Arima *et al.*, 1968b). The hydrolysis susceptibility of four nucleotide bases to RNase U_2 is reported as adenine > guanine ≫ cytosine > uracil (Uchida *et al.*, 1970). The dissociation constant of 2′-AMP is 2.3×10^{-5} M, and that of 2′-GMP is 1.3×10^{-4} M, at pH 4.5 and 30 °C (Sato & Uchida, 1975b). In RNase T_1, amino acid residues recognizing guanine bases are located at the residues from 43 to 46, in which Glu 46 forms two hydrogen bonds with the bound guanine (Arni *et al.*, 1988; Koepke *et al.*, 1989; Lenz *et al.*, 1991; Gohda *et al.*, 1994). The base specificity of RNase U_2 is originating from the three-dimensional structure supposedly around the region from Tyr 43 to Asp 50, although this region of the sequence is much different from that of RNase T_1, as is shown in Figure 1.

RNase U_2 exists in the culture broth of *Ustilago sphaerogena* as a mixture of two isoforms, U_2A and U_2B (Uchida & Shibata, 1981). Isoform U_2A is referred to the original peptide, and is readily converted to isoform U_2B in the alkali condition. This conversion is considered as isoaspartate formation at Asn 32 *via* succinimide (Kanaya & Uchida, 1986). For the ApA substrate, isoform U_2B displays a dissociation constant similar to that of U_2A, but its molecular

† This work was supported by a Grant-in-Aid for Scientific Research A (No. 02402051) and partly by a Grant-in-Aid for Scientific Research in Priority Areas (No. 0524402) to Y.S. from the Ministry of Education, Science, Sports and Culture, Japan.

* Author to whom correspondence should be addressed.
‡ Faculty of Pharmaceutical Sciences.
§ Mitsubishi Kagaku Institute for Life Sciences.
|| Research Center, Mitsubishi Kagaku Corporation.

⊗ Abstract published in *Advance ACS Abstracts,* November 1, 1995.

[1] Abbreviations: RNase, ribonuclease; 2′,3′-cyclic AMP, cyclic 2′,3′-adenosine monophosphate; ApC, adenylyl-3′,5′-cytidine; 2′-AMP, adenosine 2′-monophosphate; 2′-GMP, guanosine 2′-monophosphate; 3′-GMP, guanosine 3′-monophosphate; isoAsp, L-isoaspartate.

0006-2960/95/0434-15583$09.00/0

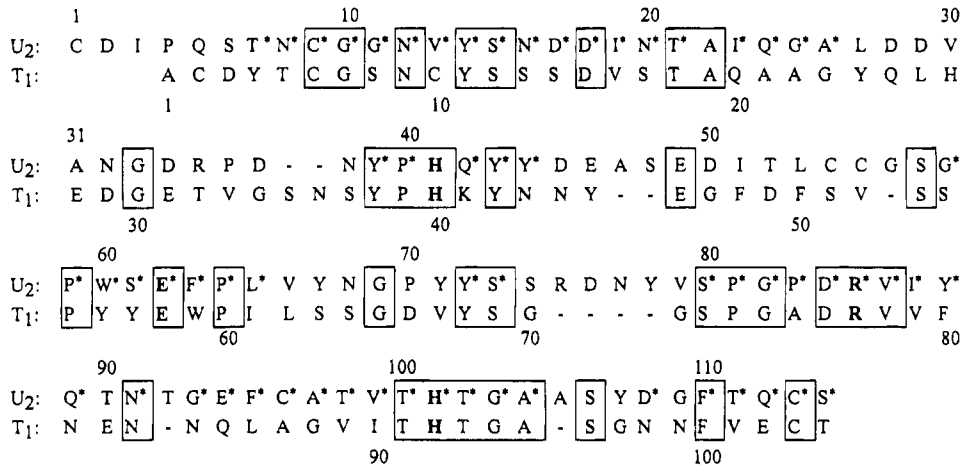

FIGURE 1: Comparison of the amino acid sequences of RNase U_2 and RNase T_1. The sequences are aligned on the basis of this study. Amino acid residues commonly placed in the sequences are boxed. The residues constituting the catalytic site are set in bold face. Each RNase U_2 residue is marked with an asterisk if its C_α position is deviated less than 1 Å from the corresponding atom of RNase T_1.

activity is three-quarters that of U_2A (Uchida & Shibata, 1981). This suggests that the isoaspartate formation involves structural changes mostly in the catalytic site.

Although the preliminary study on the RNase U_2 structure was carried out on type I crystal (Matsuzaki *et al.*, 1989), a full elucidation of its three-dimensional structure has yet to be made. Here we report the RNase U_2 structure of type II crystal[2] at 1.8 Å resolution. The structure has revealed isoaspartate formation at Asp 45, and hence is now attributed to another isoform, U_2C.

MATERIALS AND METHODS

Crystallization and Diffraction Data Collection. RNase U_2 of isoform U_2A was isolated and purified from the culture broth of *Ustilago sphaerogena* by the method of Uchida and Shibata (1981). Crystals were prepared at 20 °C from a 20 mg/mL protein solution containing 1.5 mg/mL of 2′-deoxy-2′-fluoro ApC and 0.4 M ammonium sulfate, equilibrated against a reservoir of 0.9 M ammonium sulfate containing 0.1 M acetate buffer (pH 4.5). 2′-Deoxy-2′-fluoro ApC was added in order to obtain crystals complexed with this inhibitor. Crystals grew to a typical size of 30 × 30 × 50 μm^3, and occasionally grew to 80 × 80 × 150 μm^3 over a year by the vapor diffusion method. The crystals, referred to as type II, belong to the space group $P2_12_12_1$ with cell dimensions a = 49.32 Å, b = 61.27 Å, and c = 34.95 Å. The larger crystals placed in the mother liquor of 1.0 M ammonium sulfate and 1.5 mg/mL 2′-deoxy-2′-fluoro ApC (pH 4.5) were subjected to diffraction experiments.

Diffraction data sets were collected at Photon Factory BL-14A of National Laboratory for High Energy Physics, Tsukuba, Japan (Satow & Iitaka, 1989). The four-circle diffractometer was used as an oscillation camera. Diffraction patterns from a single crystal cooled at 11 °C were recorded on Imaging Plates BAS-III (Fuji Photo Film Co. Ltd.), and then were digitized with an Imaging Plate scanner (Amemiya *et al.*, 1988). The wavelength was set to 0.9000 Å. Diffraction intensities were integrated with our modified version of program OSCMGR (Rossmann, 1979). Scaling and postrefinement of the intensities were carried out with programs TOMOKO and MARIKO (Satow *et al.*, 1986).

From 39 316 measured reflections between 12 Å and 1.80 Å spacings, 9 872 unique reflections yielded an R_{merge},

$$\sum_h \sum_i |I_i(\boldsymbol{h}) - \overline{I(\boldsymbol{h})}| / \sum_h \sum_i I_i(\boldsymbol{h})$$

where intensities $I_i(\boldsymbol{h})$ are meashred N_h times for reflection $\boldsymbol{h}$, and $\overline{I(\boldsymbol{h})}$ is an intensity averaged from them, of 0.034, covering 96.1% of the possible reflections in this spacing shell. In the outermost shell between 1.87 Å and 1.80 Å spacings, the R_{merge} was 0.077, and 86.4% of the possible reflections were covered. The average value of the ratios of intensities to their estimated errors was 14.9.

Structure Determination by the Molecular Replacement Method. The crystal structure was determined by the molecular replacement method. A search model was constructed from the crystal structure of RNase T_1 complexed with 2′-GMP (Arni *et al.*, 1988). RNase T_1 residues common to RNase U_2 were kept into the search model. Each residue homologously mathced to that of RNase U_2 was replaced with alanine, or with glycine when either of each matched pair was glycine. Unmatched residues as well as their neighboring residues were removed from the model; these were residues 34 to 37, 47 to 49, 62 to 71, 81 to 82, and 97 to 98 in the RNase T_1 numbering. The N-terminal two residues were also removed because the disulfide bond between Cys 2 and Cys 10 does not exist in RNase U_2. The search model had 84 residues and consisted of 515 atoms. Its overall temperature factor was set to 15 Å^2.

A rotation search, a Patterson correlation (PC) refinement, and a translation search were performed with program X-PLOR (Brünger, 1991). The rotation search was carried out by selecting Patterson peaks with an origin-to-peak range of 4−5 Å. In the subsequent PC-refinement, one orientation with a relatively high PC coefficient of 0.142 was obtained when reflections in a 5−4 Å spacing shell were used. Even when spacing shells were varied, this orientation was always associated with the highest rank in PC coefficients.

The PC-refined orientation was then subjected to the translation search. Correlation coefficients between the squares of normalized amplitudes calculated from the search

[2] Structural Data: Atomic coordinates and structure factors for RNase U_2 type II crystal have been deposited with the Brookhaven Protein Data Bank. The PDB file names are 1RTU for the coordinates and R1RTUSF for the structure factors.

model and those observed in a 5−4 Å shell were calculated at grid intervals of 1 Å. The highest correlation coefficient appeared as a low value of 0.361, but was 5.6 σ higher than the average value [σ is the root mean square (rms) value of the correlation coefficients]. This solution was verified by checking the absence of severe molecular intrusion in crystal packing. The solution was subsequently subjected to the rigid-body refinement with X-PLOR. Although the refinement yielded an R-factor, $\sum|F_o - F_c|/\sum F_o$ where F_o and F_c are observed and calculated structure amplitudes, respectively, of 0.496 in a 6−3 Å shell, further verification of the solution was made on the basis of the Fourier maps, by checking the presence of electron densities for residues that were not included in the model.

Model Building and Crystallographic Refinement. Following the rigid-body refinement, a great number of the refinement rounds involving model building procedures and crystallographic refinements were undertaken. The model building was carried out using program TURBO-FRODO (Roussel & Cambillau, 1990). The crystallographic refinements were performed with X-PLOR, and with PROLSQ (Hendrickson, 1985) in the final stages. As working structural models adopted in the initial rounds contained only 60% of the protein atoms, thorough examination of the electron density maps and elaborating model building were required. Most parts of the model were rebuilt several times in the course of the refinement. Particularly, long loops from Cys 1 to Ser 6, from Asp 45 to Ser 57, and from Tyr 67 to Val 79, and a region from Leu 27 to Asp 37, which includes a C-terminal part of an α-helix, were completely rebuilt. Then, solvent atoms including a sulfate ion were incorporated into the model with particular care. After a total of 23 rounds, the R factor for the model with the solvent and sulfate atoms dropped to 0.148 for an 8.0−1.8 Å shell.

At this stage, the electron density indicated structural features of the RNase U_2 crystal in further detail. Two side-chain conformers each for Val 86 and for Tyr 107 were noticed in the electron density and were incorporated into the model. Occupancies for these conformers were determined so that their average temperature-factors became comparable. Residues 49 and 50, reported respectively as Asp 49 and Gln 50 (Kanaya & Uchida, 1986), did not fit into the electron density (shown in Figure 3), and the stereochemistry for these regions with an Asp−Gln structure remained poorly defined. On the basis of the electron density and the sequence matching to RNase T_1, Asp 49 was replaced with a glutamic acid, and Gln 50 with an aspartic acid. Then the R factor dropped to 0.146.

Asp 45, however, did not fit into its electron density even at this stage. The stereochemistry for its main- and side-chains remained poorly defined. Any side-chain conformation of a regular aspartate structure was too long to be accommodated into the density, although Tyr 44 and Glu 46 fitted into the electron density quite well. We attempted to place disordered or multiple conformers of an aspartate residue as well as disordered water molecules in the Tyr 44−Glu 46 region that is involved in interaction with the neighboring protein molecule (shown in Figure 5). None of the attempts yielded well-fitted densities nor reasonable stereochemical parameters. The clear electron density (shown in Figure 4), which spans the locations representing the Asp 45 α-carbon atom and the carbonyl oxygen atom connected to the Glu 46 N atoms, is about 3 Å long, and is curved by about 110° at its midpoint. Furthermore, the density indicates that residue 45 should lack a carbon atom between the α-carbon and the side-chain carboxyl carbon. The density was interpreted as that for L-isoaspartate, and an isoaspartate model from the β-L-aspartyl-L-alanine (Görbitz, 1987) was finally built into the protein model. Residual electron densities in this region disappeared from the $F_o - F_c$ difference Fourier map, which was calculated from the model finally built and was contoured at 0.11 times the rms value for electron densities calculated with the F_o synthesis. A small decrease in free-R values (Brünger, 1992) for 10% of the reflection data, from 0.1781 for the normal aspartate model to 0.1778 for the isoaspartate model, was obtained with the PROLSQ refinements.

Though Pt derivative crystals were prepared by soaking native crystals into the mother liquor containing 7.5 mM K_2PtCl_4, they were not used in the phasing. The structure for the Pt derivative, however, was investigated so as to examine the model obtained from the native data. The diffraction data collected at wavelength 1.0713 Å yielded an R_{merge} of 0.044 in the 12.0−2.1 Å shell. The model transferred from the native model was refined to the R factor of 0.171 for the derivative data in the 8.0−2.1 Å shell. The electron density which appeared virtually identical to that of the native crystal confirmed the interpretation of isoaspartate at residue 45. A Pt atom with a low occupancy of 0.12 was located near the S_γ position of Cys 1 in the difference Fourier map.

RESULTS

Overall Structure of RNase U_2. The final model of the RNase U_2 structure for type II crystal consists of 114 amino acids, 141 water molecules, and 1 sulfate ion. The structure revealing isoaspartate at the Asp 45 position is referred to as isoform U_2C. The final R-factor is 0.143 for 9 807 reflections in the 8.0−1.8 Å shell. When the final model is superimposed to the initial model, the positional rms difference between these models is 2.72 Å. The overall structure is shown in Figure 2. The rms deviation from the ideal bond lengths is 0.014 Å, and that from the ideal bond angles is 1.1°. The main-chain torsion angles (φ, ψ) are clustered in the allowed regions of the Ramachandran plot. Of the nonglycine residues, Asp 37, Asn 38, and Asn 68 are in the region for the left-handed helical conformations and exist as tight turns. Three prolines—Pro 40, Pro 59, and Pro 70—out of eight, are in *cis* conformation. The mean error in atomic positions was estimated to be 0.13 Å by the method of Luzzati (1952). The mean positional difference between the native and the Pt derivative structures is 0.16 Å for all protein atoms, and is comparable to the mean error. No electron density corresponding to the 2′-deoxy-2′-fluoro ApC inhibitor was located. Possible binding-sites of the inhibitor are interfaced by a neighboring protein molecule. This crystal packing and the sulfate ion in the high concentration might have prevented the inhibitor from binding to RNase U_2 in the crystal.

The structural model is examined by locating sulfur atoms of the protein. Fourier maps with Bijvoet anomalous differences are calculated, using phases from the model with all the sulfur atoms excluded. Peaks 2.5 σ higher than average (σ is an rms value of residual densities) are located at the positions corresponding to three disulfide bridges, although the imaginary part of anomalous scattering factors of sulfur is as small as 0.198 electrons at the short wavelength

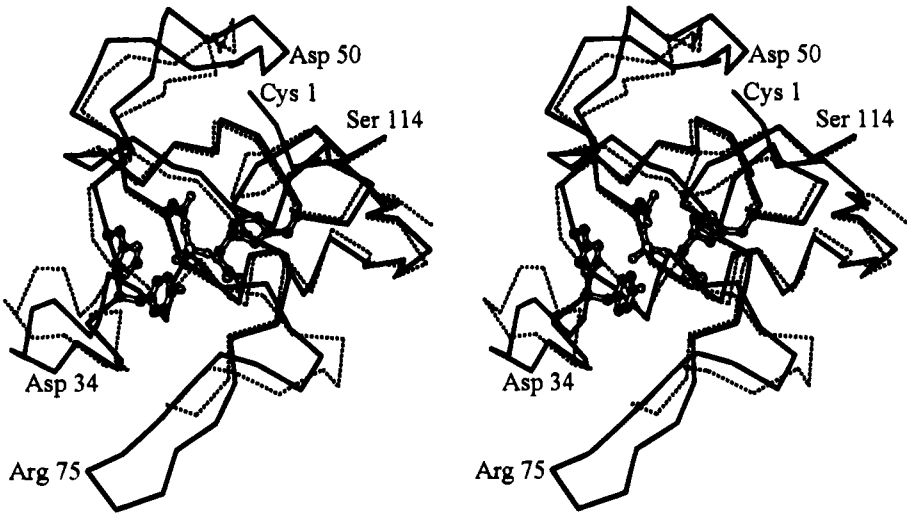

FIGURE 2: Stereo view of the C_α trace of the RNase U_2 structure. The C_α trace (solid line) for the RNase U_2 structure is superposed on the C_α trace (dotted line) for the RNase T_1 structure (Arni *et al.*, 1988) for comparison. The RNase U_2 side-chains of the catalytic site residues are shown as well as the sulfate ion. Figures 2, 6, and 8 are drawn by MOLSCRIPT (Kraulis, 1991).

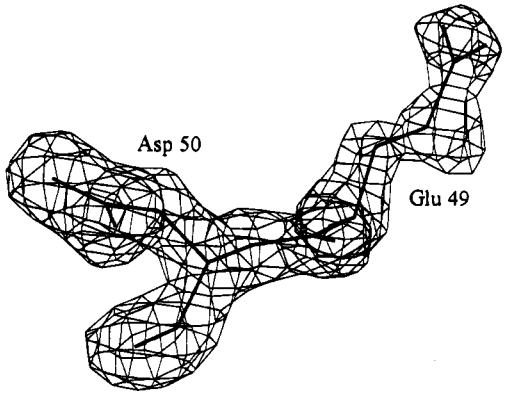

FIGURE 3: Difference Fourier map at Glu 49−Asp 50 of RNase U_2. The Glu 49−Asp 50 model is superposed on the $F_o - F_c$ difference Fourier map that is contoured at 2.5 times the rms value of electron densities. In the calculation of the map, a model without residues 49 and 50 was subjected to 5 cycles of the PROLSQ refinements. The phases from the model thus refined were used in the $F_o - F_c$ Fourier synthesis. Figures 3, 4, and 7 are generated using TURBO-FRODO (Roussel & Cambillau, 1990).

of 0.9000 Å (Sasaki, 1989). No peak is located at the sulfate position, however. This absence of the sulfur peak is supposedly due to the high temperature-factor of 23 Å^2 for the sulfate S atom. The Bijvoet difference ratio originating from anomalous scatterers (Yang *et al.*, 1990) is expected to be 0.0035 for 6 sulfur atoms. The observed ratio is very small—0.0168 for the diffraction data in the 8.0−3.5 Å shell—and is equivalent to an R_{merge} of 0.020 for intensities of Bijvoet pairs. The disulfide bridges are confirmed as Cys 1−Cys 54, Cys 55−Cys 96, and Cys 9−Cys 113.

Figure 3 shows the $F_o - F_c$ difference Fourier map obtained through 5 cycles of the PROLSQ refinements of a structure model without residues 49 and 50. The map clearly indicates that residue 49 should have an extra γ carbon atom, and that residue 50 should lack an atom at the C_γ position. In the final model, the Glu 49−Asp 50 structure yields a good stereochemistry and fits well into the electron density, as shown in Figure 3. The side chain of Glu 49 faces the solvent region, and its carboxyl oxygen is hydrogen-bonded to water 336. The average temperature factor of the side-chain atoms is obtained as a higher value of 31 Å^2. The electron density at Glu 49 also suggests that a small amount of rotational conformers around the $C_\beta-C_\gamma$ bond may exist. These findings suggest that the side chain of Glu 49 is flexible. The carboxyl oxygen of the Asp 50 side-chain is hydrogen-bonded to the O atom of the Phe 110 main-chain. Prompted by these results, reexamination of the amino-acid sequence for the Tyr 44−Leu 53 peptide was carried out by Kanaya and Uchida (1995). The study confirmed the proper sequence to be Glu 49−Asp 50.

RNase U_2 has a β-hairpin structure at residues from Thr 7 to Tyr 14, a 4.4-turn α-helix from Asn 16 to Asn 32, and a central β-sheet with five strands from His 41 to Tyr 43, from Trp 60 to Leu 65, from Asp 84 to Gln 89, from Glu 94 to His 101, and from Phe 110 to Gln 112. Except for Gly 10 and Gly 11, these residues constitute a structural core. A long loop from Tyr 67 to Gly 82 is protruding outwardly from the central β-sheet, and is distinctive of the RNase U_2 structure. Three hydrogen bonds in the main chain, between Ser 73 O_γ and Val 79 N, Ser 74 N and Asn 77 O, and Ser 74 O and Asn 77 N, have a role in stabilizing this protruding structure. The α-helix is bent at its middle by about 10° toward the β-turn from Ser 74 to Asn 77. This bend is ascribable to its succeeding Gly 33−Asn 38 loop, which is shorter by two residues than that of RNase T_1. In the sequence alignment shown in Figure 1, each RNase U_2 residue is marked with an asterisk if its C_α position is deviated less than 1 Å from the corresponding atom of RNase T_1; most of the marked residues are situated in the structural core. As is shown in Figure 2, large deviations are noticeable at the loop regions, from Tyr 44 to Gly 56 and from Tyr 67 to Gly 82, and also at the C-terminal half of the α-helix. The positional rms deviation of the main-chain atoms in the structural core is 1.02 Å. The deviation increases to 2.50 Å when side-chain atoms are included.

Each RNase U_2 molecule is related closely to neighboring molecules *via* the crystallographic 2_1 screw axis situated along the z axis which has the shortest cell constant of 34.95 Å. This crystal packing involves contacts between the side chains of the residues from Tyr 43 to Asp 50 and those from Asn 77′ to Gly 82′ of the neighboring molecule, whose residues are indicated hereafter with primed numbers.

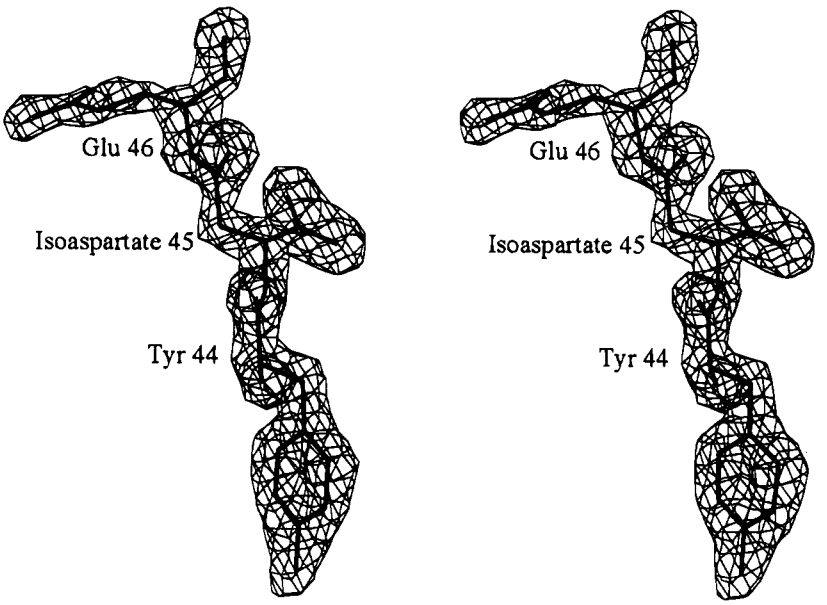

FIGURE 4: Stereo view of the $F_o - F_c$ difference Fourier map at residues 44 to 46 of RNase U_2. In the calculation of the map, coordinates for a model without residues 44 to 46 were shaken by moving all atoms at random up to 0.25 Å. Then the shaken model was subjected to 5 cycles of the PROLSQ refinements. The map was phased with the model thus refined, and was contoured at 3.0 times the rms value of electron densities. The final model with the isoaspartate structure is superposed on the map.

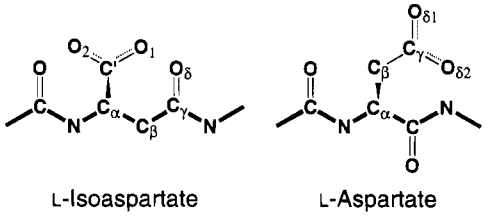

FIGURE 5: Schematic diagrams comparing chemical structures of isoaspartate and normal aspartate. Atoms are indicated with their corresponding labels. The bonds in the main chains are drawn with bold lines. Note the stereochemical configurations at the α carbon positions.

Each side chain of Val 86 and Tyr 107 has two conformers. Val 86 is located in a hydrophobic environment inside the protein molecule, and its side chain is in a cavity surrounded by the side chains of Ala 22, Ile 23, Leu 65, and Thr 98. In either conformer, the Val 86 side-chain exists at reasonable van der Waals distances from these side-chain atoms. The Tyr 107 side-chain in one conformer is pointing toward the solvent region, with an occupancy of 0.5. The other is in contact with the side chain of Pro 59′ of the neighboring molecule. These residues are located near the 2_1 screw axis.

Structure of Isomerized Asp 45. The electron density around isoaspartate (isoAsp) 45, which was calculated through the application of the shaking procedure (McRee, 1993) and the PROLSQ refinements to a model without residues 44−46, is shown in Figure 4, where the final model with the isoaspartate structure is superposed. Figure 5 shows schematic diagrams comparing chemical structures for normal aspartate and isoaspartate. The isoAsp 45 structure fits into the electron density quite well and yields good stereochemical parameters, of which the bond distances and the torsion angles are listed in Table 1. The rms deviation of the bond distances from those of the β-L-aspartyl-L-alanine structure (Görbitz, 1987) is as low as 0.011 Å. The temperature factors for the isoAsp 45 atoms are obtained as low values; the average values are 14.3 $Å^2$ and 12.1 $Å^2$ for the main- and the side-chain atoms, respectively. The main-chain torsion angles of θ_1 and θ_2 are also close to those for the β-L-aspartyl-L-alanine structure (Görbitz, 1987). The structure around isoAsp 45 is shown in Figure 6. The region from Tyr 44 to Asp 50 is located on the surface of the molecule, but is defined in clear electron densities. As for isoAsp 45, the side-chain carboxyl group, which originated from the main-chain carbonyl group of Asp 45, makes four hydrogen-bonds in reasonable geometry. One carboxyl oxygen atom, O_1, is hydrogen-bonded to the neighboring molecule, both to Asp 37′ $O_{\delta 2}$ with a distance of 2.5 Å and to Tyr 78′ O_η with 2.6 Å. The other carboxyl O_2 is hydrogen-bonded both to Ala 47 N with 3.0 Å and to water 348 with 3.2 Å; water 348 is further hydrogen-bonded to Tyr 44 O and to water 347′. The Tyr 44−Asp 50 region has few van der Waals contacts with neighboring regions. These hydrogen bonds along with the one formed between

Table 1: Bond Distances and Torsion Angles for Isoaspartate 45[a] in RNase U_2

bond	bond distance (Å)	(target value[b])
44C−45N	1.304	(1.320)
45N−45C_α	1.499	(1.489)
45C_α−45C_β	1.516	(1.535)
45C_β−45C_γ	1.512	(1.518)
45C_γ−45O_δ	1.232	(1.235)
45C_γ−46N	1.353	(1.344)
45C_α−45C′	1.528	(1.530)
45C′−45O_1	1.251	(1.256)
45C′−45O_2	1.247	(1.251)
rms deviation	0.011	

	torsion angle (°)	(reported values[c])
ϕ (44C−45N−45Cα−45C_β)	69.7	
θ_1 (45N−45C_α−45C_β−45C_γ)	−171.0	(−168.2)
θ_2 (45C_α−45C_β−45C_γ−46N)	−132.8	(−122.5)
θ' (45N−45C_α−45C−45O_1)	44.5	(−12.2)

[a] Atom names for isoaspartate are indicated in Figure 5. [b] Target values used in the structure refinement are derived from the crystal structure of β-L-aspartyl-L-alanine (Görbitz, 1987). [c] Reported values for the torsion angles are also from the β-L-aspartyl-L-alanine structure.

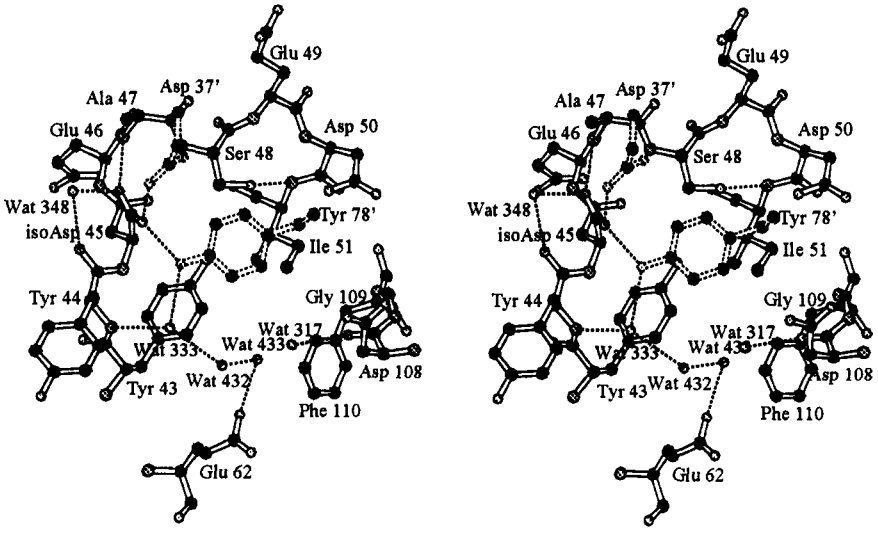

FIGURE 6: Stereo view of the base-binding site of RNase U_2. Residues in the vicinity of the base binding site are shown as well as the water molecules located there. C, N, and O atoms are drawn with black, dotted gray, and small gray circles, respectively. Hydrogen bonds are drawn with dotted lines. Residues of the neighboring protein molecule are indicated with primed numbers and are drawn with dotted lines. The carboxyl oxygen atoms of the side-chain of isoaspartate 45 are hydrogen-bonded to water 348, to Ala 47 N, to Asp 37′ $O_{\delta 1}$, and to Tyr 78′ O_{η}.

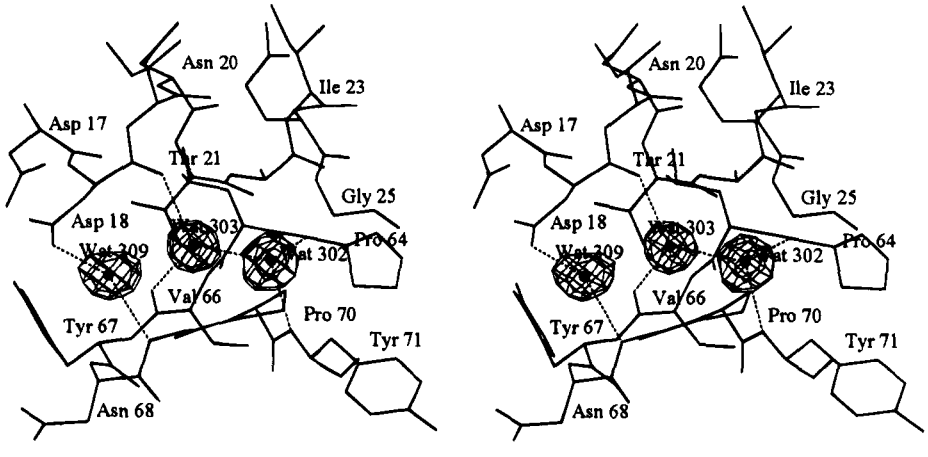

FIGURE 7: Stereo view of the water molecules deeply buried inside the RNase U_2 molecule. Hydrogen bonds are drawn with dotted lines. Waters 302, 303, and 309 are located in the crevice between the α-helix (upper side) and the long loop from Pro 64 to Tyr 71 (lower side). The model with the buried water molecules are superposed on the $F_o - F_c$ difference Fourier map contoured at 3.5 times the rms value. In the preparation of the map, the model without the water molecules was subjected to 5 cycles of the PROLSQ refinements, and resultant phases were used in the Fourier synthesis.

Ser 48 O_{γ} and Ile 51 N seem to contribute to possible rigidification of this region in the crystal.

Bound Waters to the Protein. Out of 141 water molecules located, 98 water molecules are hydrogen-bonded directly to one protein molecule, and 23 water molecules are hydrogen-bonded, in a bridging manner, to two protein molecules in the unit cell. Figure 7 shows some of the water molecules deeply buried inside the protein molecule. Waters 302, 303, and 309 exist in a channel-like crevice located between the N-terminal half of the α-helix and the residues from Pro 64 to Tyr 71 where the end of the central β-sheet is positioned. Water 302 is located at the bottom of the crevice, and is about 6 Å away from the end of the crevice. It is hydrogen-bonded to Pro 64 O, Tyr 71 N, and water 303, and appears to be maintaining the conformation of the loop at the residues from Val 66 to Tyr 71. Water 303, located in the middle of the crevice, is hydrogen-bonded to Val 66 O, Asp 18 O, and water 302. Water 309, located at the end of the crevice, is hydrogen-bonded to Asp 18 $O_{\delta 1}$ and Gly 69 N. These water molecules contribute to the spatial arrangement of the α-helix which is laid over the β-sheet.

Structure of the Catalytic Site. Sulfate ions bound to RNases often provide information both on the location where phosphate groups of RNA bind and on the residues that form catalytic sites. In the crystal structure of RNase U_2, the sulfate ion is located in the close proximity to the 2_1 screw axis which runs along the *z*-axis and relates each adjacent pair of protein molecules. As shown in Figure 8, the sulfate ion is hydrogen-bonded to the side chains of Tyr 39, Glu 62, Arg 85, and His 101, and also to water 396 and water 445. Of the oxygen atoms of the sulfate ion, sulfate O_1 has three hydrogen bonds to Tyr 39 O_{η}, to Glu 62 $O_{\epsilon 2}$, and to Arg 85 N_{ϵ}. Sulfate O_2 is hydrogen-bonded to water 396. Sulfate O_3, which is pointing toward the solvent region, is also hydrogen-bonded to Tyr 39 O_{η}. Sulfate O_4 has three hydrogen-bonds: to Arg 85 N_{η}, to His 101 $N_{\epsilon 2}$, and to water 445. The Arg 85 side-chain is located at the bottom of the catalytic site and is not accessible to the solvent region. It is surrounded by the sulfate and the side chain of Glu 62. The

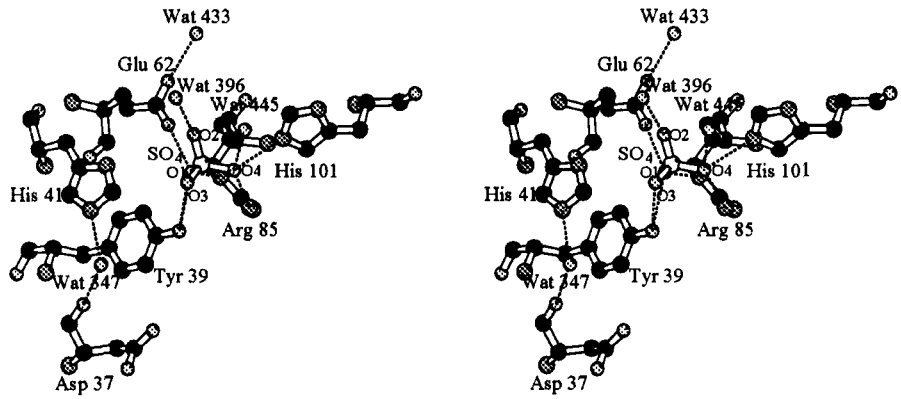

FIGURE 8: Stereo view of the catalytic site of RNase U_2. C, N, O, and S atoms are drawn with black, dotted gray, small gray, and white circles, respectively. Hydrogen bonds are drawn with dotted lines.

distance between the Arg 85 N_ϵ and Glu 62 $O_{\epsilon 2}$ atoms is 3.5 Å, suggesting a weak electrostatic interaction. The imidazole ring of His 101 is in van der Waals contact with the phenyl ring of Phe 110. The interplanar angle between these rings is 27°, and the shortest interatomic distance is 3.3 Å for the pair of His 101 $C_{\epsilon 1}$ and Phe 110 C_η.

The side chain of His 41 is also located near the sulfate, but is not hydrogen-bonded to it. The $N_{\epsilon 2}$ atom of His 41 is hydrogen-bonded to water 347, with a distance of 2.8 Å. Then this water 347 is tightly hydrogen-bonded both to Asp 37 O with 2.5 Å, and to water 348 with 2.7 Å. No hydrogen-bond is formed at the $N_{\delta 1}$ atom of His 41. The imidazole ring of His 41 is in van der Waals contact with the phenol ring of Tyr 39. The interplanar angle between these rings is 13°, and the shortest distance is 3.3 Å for the pair of Tyr 39 $C_{\epsilon 1}$ and His 41 $C_{\epsilon 1}$.

Structural Environment for the Adenine-Specific Recognition. The recognition site for an adenine base is ascribed to the region surrounded by Tyr 44, Glu 49, and Asp 108, on the basis of superposition of the three-dimensional structure of RNase U_2 to that of RNase T_1 complexed with 3′-GMP (Gohda *et al.*, 1994). Figure 6 shows the structure of the adenine recognition site. The site consists of the long loop from Tyr 44 to Asp 50 and the residues from Asp 108 to Phe 110. The distance between the N atom of Tyr 44 and the O atom of Asp 108 is 11.4 Å, and is wide enough to accommodate the adenine base. Hydrophobic side-chains of Tyr 43, Ile 51, and Phe 110 are located at the bottom of the base-recognition site. The aromatic side-chain of Tyr 78′ of the neighboring protein molecule is situated on this hydrophobic surface. This suggests that the binding of the adenine base would occur on this hydrophobic surface. Water molecules 317, 333, 432, and 433 are located in the vicinity of the base-recognition site. The acidic side-chains of Glu 46, Glu 49, and Asp 108 protrude toward the solvent region and extend away from the base-recognition site. The formation of isoaspartate at residue 45 suggests that the loop is conformationally flexible and alterable in solution.

DISCUSSION

The structural motif composed of a β-hairpin, an α-helix, and a central β-sheet is common to RNases U_2 and T_1. Two additional remarks can be made on the structures common to these RNases. One is on the valine residues buried inside the protein molecule. In RNase U_2, two conformers of Val 86 are observed. In RNase T_1, the side chain of Val 77 is reported to be disordered in the absence of the ribose group in its binding site (Kostrewa *et al.*, 1989; Ding *et al.*, 1991; Martinez-Oyanedel, 1991; Lenz *et al.*, 1993). The other remark is on the existence of the water molecules also buried inside the protein molecule. As shown in Figure 7, RNase U_2 has three water molecules in the crevice near the N-terminal half of the α-helix. In RNase T_1, the locations corresponding to these water molecules are filled with four water molecules (Malin *et at.*, 1991; Pletinckx *et al.*, 1994). It seems that not only the structural motif but also the states of some of the bound waters and the buried valine residues are conserved as structurally essential constituents through the evolution.

Isomerization of aspartate to isoaspartate via succinimide formation have been reported to occur in peptides (Geiger & Clarke, 1987). The succinimide formation proceeds rapidly near pH 4.0 in the case of the hexapeptide (Oliyai & Borchardt, 1993). The isomerizations chemically identified in proteins include calmodulin (Ota & Clarke, 1989), αB-crystallin (Fujii *et al.*, 1994), and lysozyme (Tomizawa *et al.*, 1994). It is unknown whether isoform U_2A isomerized into isoform U_2C in solution or in crystal. It should be pointed out that it took about one year for the crystals to grow at pH 4.5. We think that the isomerization occurred when the protein molecules became aggregated to the crystal over this long equilibration period, because of the environment of isoAsp 45 and its neighbors, and also because of the absence of detectable fractions of isoform U_2C in solution. It seems that a normal Asp 45 structure could not be packed into the lattice of type II crystal, nor could it be isomerized in the environment with the definitive conformation and ordered network of the hydrogen bonds. Conformational differences from the normal Asp 45 structure would exist around the isoAsp 45 residue, and supposedly involve a part of the adenine-recognition loop located on the surface of the protein molecule.

It has been reported that the treatment of isoform U_2A in the alkali condition yields isoaspartate only at Asn 32 (Kanaya & Uchida, 1986). There are two Asn-Gly sequences at 32−33 and at 68−69. In the formation of isoaspartate from asparagine *via* succinimide formation, a covalent bond is formed between the C_γ atom of asparagine and the N atom of glycine. Our simple model-building showed that typical values of the torsion angles prerequisite to succinimide formation for asparagine and aspartate are about 120° for χ_1, 90° or −90° for χ_2, and −120° for ψ. In type II crystal,

Asn 32 has torsion angles of $\chi_1 = -66°$, $\chi_2 = -51°$, and $\psi = 3°$, and Asn 68 has $\chi_1 = -63°$, $\chi_2 = -59°$, and $\psi = 49°$. For both of the Asn−Gly pairs, it seems possible to rotate their χ_1 and χ_2 angles so as to adopt the prerequisite torsion angles since their side-chains are exposed to the solvent region and are stereochemically less restrained. The Asn 32−Gly 33 pair precedes a flexible loop from Asp 34 to Asp 37, which is also situated on the surface of the protein. Therefore, the conformation of the Asn 32−Gly 33 main-chain is not tightly restrained. On the other hand, the ψ rotation of Asn 68 is highly restrained by the conformation of the Gly 69−Pro 70 main-chain. This explains the occurrence of the isoaspartate formation only at Asn 32. Asn 32 is located far away from the base-recognition site but is close to the catalytic site residues, Tyr 39 and His 41. Thus, the isoaspartate formation at Asn 32 could impose some changes in the conformation or in the dynamic nature on the catalytic-site residues. The structural environment around Asn 32 partially explains the observation that the molecular activity of isoform U_2B decreased while its dissociation constant for ApA remained unchanged (Uchida & Shibata, 1981).

The side chains of Tyr 39, Glu 62, Arg 85, and His 101 are hydrogen-bonded to the sulfate. These polar side-chains are supposedly protonated in this crystallization condition of pH 4.5. The pK_a values for His 41 and His 101 were reported to be around 8 (Uchida, 1985). The His 101 side-chain bound to the sulfate is less accessible to the solvent. Hence its protonated form would be preserved in the interaction with the RNA phosphate moiety, and could act as a general acid that donates a proton to the $O_{5'}$ atom of the leaving nucleotide in the RNA cleavage, as in RNase T_1 (Takahashi & Moore, 1982; Heinemann & Hahn, 1989). Though His 41 has no interaction with the sulfate, about 100° rotation of its side-chain around the $C_\alpha-C_\beta$ bond makes it possible for the $N_{\epsilon 2}$ atom to interact with the sulfate. Thus, the structure reported here does not exclude the possibility that His 41 interacts with the RNA phosphate group. A small additional adjustment of the His 41 main-chain is necessary for the His 41 side-chain to adapt itself to the arrangement of the His 40 residue of RNase T_1 complexed with 3′-GMP (Gohda *et al.*, 1994), in which the His 41 $N_{\epsilon 2}$ atom is hydrogen-bonded to the $O_{2'}$ atom of ribose. Therefore, it also seems possible that the His 41 side-chain would be involved in the interaction with the ribose $O_{2'}$ atom. As the side chain of Glu 62 is located near the position corresponding to the $O_{2'}$ position in the 3′-GMP complexed structure of RNase T_1, the involvement of the Glu 62 side-chain as a general base is also possible. The side chain of His 101 and the $O_{2'}$ position are on opposite sides of the bound sulfate. This location of His 101 is again consistent with its presumed role as a general acid.

When the C_α atoms of the catalytic-site residues Tyr 39, His 41, Glu 62, Arg 85, and His 101, and the S atom of the sulfate in the RNase U_2 structure are fitted to the C_α atoms of the corresponding residues and the P atom in the RNase T_1−3′-GMP structure (Gohda *et al.*, 1994), the positional rms deviation between these atoms is 0.62 Å. Of these atoms, the C_α atom of His 41 shows the largest positional displacement of 1.16 Å. The rms deviation becomes 0.34 Å when the C_α atom of His 41 is excluded from the calculation. This large deivation, together with the side-chain conformation of His 41, suggests that the role of His 41 in catalysis might be different from that of RNase T_1, although His 41 is conserved in the sequences and is spatially located in the vicinity of the corresponding histidines of the homologous RNases.

In the oligonucleotide cleavages by RNase U_2, residues with pK_a values 3.8, 4.0, and 5.0 are suggested to participate (Yasuda & Inoue, 1982). A residue with pK_a of 5.5 is also reported to participate in the binding of 2′-AMP (Sato & Uchida, 1975b). These suggest that amino acids with carboxyl groups take part in the nucleotide binding to RNase U_2. RNase U_2 hydroylzes guanosine 3′-phosphate esters, though the susceptibility is lower than the case of adenosine 3′-phosphate esters. Among homologous fungal RNases guanine-specific T_1, F_1, and less specific Ms, the glutamic acid in the base-recognition site is conserved; it is Glu 46 in RNase T_1 (Arni *et al.*, 1988), Glu 45 in Ms (Nonaka *et al.*, 1991), and Glu 46 in F_1 (Vassylyev *et al.*, 1993). It is also conserved in less homologous bacterial RNases Sa (Sevcik *et al.*, 1991) and Barnase (Guillet *et al.*, 1993). The recognition mode of the guanine base is virtually common to these RNases: the carboxyl side-chain of the glutamic acid is hydrogen-bonded to the guanine base, with one carboxyl oxygen bound to the guanine N_1 atom and with the other to the N_2 atom. The alignment of the amino-acid sequence gives two candidates for the glutamic acid. They are Glu 46 and Glu 49; the protruding side-chains of these residues are located at a moderate distances from the base-recognition site. In the absence of the substrate, and possibly as affected by the isomerized Asp 45 structure, these side-chains may not appear facing the base-recognition site. In our modeling of an adenine moiety of adenosine onto the base-recognition site, the N atom of Tyr 44 that is hydrogen-bonded to water 333 is superposable to the position corresponding to the Asn 43 N atom that is hydrogen-bonded to the N_7 atom of the guanine in the 3′-GMP complexed structure of RNase T_1 (Gohda *et al.*, 1994). Placing the adenine N_7 atom at the location of water 333 in the RNase U_2 structure, the adenine base is accommodated into the region surrounded by the residues Tyr 44, Glu 49, and Asp 108. In this arrangement of the adenine base, its N_7 atom is hydrogen-bonded to Tyr 44 N atom, and its N_1 atom is sited toward Glu 49. Glu 49 in the surface region is less restrained by van der Waals contact with the rest of the protein molecule, and is capable of interacting with the adenine base through its main- and side-chain conformational changes. Glu 49 now matches the glutamic acid conserved in the RNases, and is the likeliest residue to interact with the base.

The crystal packing involves van der Waals contacts between the region of Tyr 43 and Tyr 44, and the region from Tyr 78′ to Gly 82′. As shown in Figure 6, the aromatic side-chain of Tyr 78′ of the neighboring protein molecule is located on the hydrophobic surface of the base-recognition site. The side-chain of Tyr 44, in turn, is in contact with the residues Pro 81′ to Pro 83′. It was noted on RNase U_2 that the dependence of catalytic activities on the lengths of nucleotides suggests the existence of the subsite where the 5′- nucleotide moiety downstream of the scissile phosphodiester bond is bound (Yasuda & Inoue, 1982). In RNase T_1, the nucleotide binding subsite is located near Gly 74 (Koepke *et al.*, 1989; Lenz *et al.*, 1991). This Gly 74 position is superposable on the Gly 82 position in the RNase U_2 structure. Based on the intermolecular contacts between the side chain of Tyr 44 and the region of Pro 81′−Pro 83′, it is

proposed that the subiste of RNase U_2 could be located near the Gly 82 position, where one end of the long protruding loop from Tyr 67 to Gly 82 is joined to the central β-sheet.

REFERENCES

Amemiya, Y., Matsushita, T., Nakagawa, A., Satow, Y., Miyahara, J., & Chikawa, J. (1988) *Nucl. Instrum. Methods Phys. Res. Sect. A 266*, 645−653.

Arima, T., Uchida, T., & Egami, F. (1968a) *Biochem. J. 106*, 601−607.

Arima, T., Uchida, T., & Egami, F. (1968b) *Biochem. J. 106*, 609−613.

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

Brünger, A. T. (1991) *Acta Crystallogr. Sect. A 47*, 195−204.

Brünger, A. T. (1992) *Nature 355*, 472−475.

Cuchillo, C. M., Parés, X., Guasch, A., Barman, T., Travers, F., & Nogués, M. V. (1993) *FEBS Lett. 333*, 207−210.

Ding, J., Koellner, G., Grunert, H. P., & Saenger, W. (1991) *J. Biol. Chem. 266*, 15128−15314.

Egami, F., Oshima, T., & Uchida, T. (1980) *Mol. Biol. Biochem. Biophys. 32*, 250−270.

Fujii, N., Ishibashi, Y., Satoh, K., Fujino, M., & Harada, K. (1994) *Biochim. Biophys. Acta 1204*, 157−163.

Geiger, T., & Clarke, S. (1987) *J. Biol. Chem. 262*, 785−794.

Gohda, K., Oka, K., Tomita, K., & Hakoshima, T. (1994) *J. Biol. Chem. 269*, 17531−17536.

Görbitz, C. H. (1987) *Acta Chem. Scand. B 41*, 679−685.

Guillet, V., Laphtorn, A., & Mauguen, Y. (1993) *FEBS Lett. 330*, 137−140.

Heinemann, U., & Hahn, U. (1989) in *Protein−Nucleic Acid Interaction* (Saenger, W., & Heinemann, U., Eds.), pp 111−141, Macmillan, London.

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

Kanaya, S., & Uchida, T. (1986) *Biochem. J. 240*, 163−170.

Kanaya, S., & Uchida, T. (1995) *J. Biochem. (Tokyo) 118*, 681−682.

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.

Kraulis, P. J. (1991) *J. Appl. Crystallogr. 24*, 946−950.

Lenz, A., Choe, H. W., Granzin, J., Heinemann, U., & Saenger, W. (1993) *Eur. J. Biochem. 211*, 311−316.

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

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

McRee, D. E. (1993) in *Practical Protein Crystallography*, pp 245−275, Academic Press, San Diego.

Malin, R., Zielenkiewicz, P., & Saenger, W. (1991) *J. Biol. Chem. 266*, 4848−4852.

Matsuzaki, T., Sasaki, C., Okumura, C., & Uchida, T. (1989) Proc. of the first international meeting, in *Structure and Chemistry of Ribonucleases* (Pavlovsky, A., & Polyakov, K., Eds.) pp 286−290, Academy of Sciences of the USSR, Moscow.

Martinez-Oyanedel, J., Choe, H. W., Heinemann, U., & Saenger, W. (1991) *J. Mol. Biol. 222*, 335−352.

Nishikawa, S., Morioka, H., Fuchimura, K., Tanaka, T., Uesugi, S., Hakoshima, T., Tomita, K., Ohtsuka, E., & Ikehara, M. (1986) *Biochem. Biophys. Res. Commun. 138*, 789−794.

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

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

Oliyai, C., & Borchardt, R. T. (1993) *Pharm. Res. 10*, 95−102.

Ota, I. M., & Clarke, S. (1989) *J. Biol. Chem. 246*, 54−60.

Pletinckx, J., Steyaert, J., Zegers, I., Choe, H. W., Heinemann, U., & Wyns, L. (1994) *Biochemistry 33*, 1654−1662.

Rossmann, M. G. (1979) *J. Appl. Crystallogr. 12*, 225−238.

Roussel, A., & Cambillau, C. (1990) TURBO-FRODO version 1.1 release b, Faculté de Medecine Nord, Marseille, France.

Sasaki, S. (1989) *KEK report 88-14*, KEK Natl. Lab. for High Energy Phys., Tsukuba, Japan.

Satow, Y., Cohen, G. H., Padlan, E. A., & Davies, D. R. (1986) *J. Mol. Biol. 190*, 593−604.

Satow, Y., & Iitaka, Y. (1989) *Rev. Sci. Instrum. 60*, 2390−2393.

Sato, S., & Uchida, T. (1975a) *Biochem. J. 145*, 353−360.

Sato, S., & Uchida, T. (1975b) *Biochim. Biophys. Acta 383*, 168−177.

Sevcik, J., Dodson, E. J., & Dodson, G. G. (1991) *Acta. Crystallogr. Sect. B 47*, 240−253.

Steyaert, J., Hallenga, K., Wyns, L., & Stanssens, P. (1990) *Biochemistry 29*, 9064−9072.

Takahashi, K. (1985) *J. Biochem. (Tokyo) 98*, 815−817.

Takahashi, K., & Moore, S. (1982) *The Enzymes* (Boyer, P. D. Ed.) pp Vol. 15, 435−467, Academic Press, New York.

Tomizawa, H., Yamada, H., Ueda, T., & Imoto, T. (1994) *Biochemistry 33*, 8770−8774.

Uchida, T. (1985) *Protein, Nucleic Acid Enzyme 30*, 604−616.

Uchida, T., & Shibata, Y. (1981) *J. Biochem. (Tokyo) 90*, 463−471.

Uchida, T., Arima, T., & Egami, F. (1970) *J. Biochem. (Tokyo) 67*, 91−102.

Vassylyev, D. G., Katayanagi, K., Ishikawa, K., Tshujimoto-Hirano, M., Danno, M., Pähler, A., Matsumoto, O., Matsushima, M., Yoshida, H., & Morikawa, K. (1993) *J. Mol. Biol. 230*, 979−996.

Yang, W., Hendrickson, W. A., Crouch, R. J., & Satow, Y. (1990) *Science 249*, 1398−1405.

Yasuda, T., & Inoue, Y. (1982) *Biochemistry 21*, 364−369.

BI951196W