Acta Cryst. (1998). D**54**, 1343–1352

Plasticity, Hydration and Accessibility in Ribonuclease A. The Structure of a New Crystal Form and its Low-Humidity Variant

C. Sadasivan,* H. G. Nagendra and M. Vijayan

Molecular Biophysics Unit, Indian Institute of Science, Bangalore 560 012, India. E-mail: mv@mbu.iisc.ernet.in

(*Received 15 January 1998; accepted 3 April 1998*)

Abstract

The structures of a new crystal form of ribonuclease A and its low-humidity variant, each containing two crystallographically independent molecules, have been determined and refined. A detailed comparison of these structures with those of the other known crystal forms of the enzyme, which have different packing arrangements and solvent composition, leads to a meaningful delineation of the rigid and flexible regions of the protein molecule and the nature of its plasticity. Many of the water molecules which are common to all the structures are involved in bridging different regions of the protein molecule, thus emphasizing the role of water in stabilizing the tertiary structure. The analysis of the structures shows that for a given N or O atom, the level of hydration increases with accessible surface area, but levels off at an area of about 10 Å^2. Generally, the hydration level tends to drop when the area increases beyond about 20 Å^2. This drop correlates with an increase in the displacement parameter. The analysis also suggests that the van der Waals radii and probe radius normally used in accessible surface area calculations are not appropriate for dealing with all situations.

1. Introduction

Mobility and hydration of proteins are related aspects of considerable general importance. It is known that different regions of a protein molecule are mobile to different extents, and that this is often related to its function. X-ray, NMR and molecular-dynamics studies have led to valuable insights into the nature of this mobility (Artymiuk *et al.*, 1979; Poole & Finney, 1983; Ringe & Petsko, 1985; Huber, 1988; Brooks & Karplus, 1989; Kodandapani *et al.*, 1990; Caspar & Badger, 1991; Madhusudan & Vijayan, 1991; Santoro *et al.*, 1993; Radha Kishan *et al.*, 1995; Nadig *et al.*, 1996; Nagendra *et al.*, 1996). Such studies have also provided valuable information on the water structure that is invariably associated with a protein molecule (Baker & Hubbard, 1984; Saenger, 1987; Thanki *et al.*, 1988; Madhusudan & Vijayan, 1991; Otting *et al.*, 1991; Teeter, 1991; Madhusudan *et al.*, 1993; Frey, 1994; Radha Kishan *et al.*, 1995). It is generally believed that a characteristic hydration shell surrounds and directly interacts with the protein molecule. We have been pursuing an approach involving water-mediated transformation (Salunke *et al.*, 1984, 1985; Kodandapani *et al.*, 1990; Madhusudan & Vijayan, 1991; Madhusudan *et al.*, 1993; Radha Kishan *et al.*, 1995; Nagendra *et al.*, 1995) with the well known enzymes lysozyme and ribonuclease A as models to address the problem of mobility and hydration. These reversible transformations, accompanied by abrupt changes in the solvent content, are caused by the systematic variation in humidity around the protein crystal. Removal of a few water molecules from the medium surrounding the protein is perhaps the gentlest way to cause a transformation and the changes that result are likely to correspond to the natural tendencies of the protein molecule. Comparison of the native structures and the structures of the low-humidity forms resulting from water mediated transformation has contributed to the delineation of the relatively rigid and flexible regions of the protein molecules, and to the identification of the variable and invariant features in their hydration shells (Kodandapani *et al.*, 1990; Madhusudan & Vijayan, 1991; Radha Kishan *et al.*, 1995). Furthermore, it has been demonstrated that the changes in the molecular structure of lysozyme and ribonuclease A that accompany water-mediated transformation are similar to those that occur during catalysis, thus establishing a relationship between protein hydration, mobility and enzyme action (Radha Kishan *et al.*, 1995; Nagendra *et al.*, 1996).

Unlike the case of lysozyme, in which several different crystal forms have been examined, most crystallographic studies on ribonuclease A, including those involving water-mediated transformations, have been on one crystal form grown in the presence of alcohol (Wlodawer *et al.*, 1988) and its variants (Wlodawer *et al.*, 1986; Howlin *et al.*, 1989). The only other known native structure of ribonuclease A is that of a trigonal form (Zegers *et al.*, 1994). In order to ensure that the general results obtained are averaged over environmental perturbations such as those resulting from crystal packing and the composition of the medium, it is important to also examine other crystal forms. With this in view, we report the crystal structures of the native and 93% relative humidity forms of crystals grown from Tris buffer, pH 7.6, with acetone as precipitant (Salunke *et*

ISSN 0907-4449

al., 1984). Both contain two molecules in the asymmetric unit and thus contribute four crystallographically independent molecules for detailed comparison both among themselves and with ribonuclease A molecules in other crystal forms. Such a comparison, attempted in this paper, is expected to provide a realistic picture of the invariant and the variable features of the enzyme molecule and its hydration shell.

2. Materials and methods

2.1. *Crystallization and preliminary studies*

Bovine pancreatic ribonuclease (RNAase A) was purchased from Sigma Chemical Company. The enzyme was crystallized by slow diffusion of acetone into a protein solution of concentration 12 mg ml^{-1} in Tris–HCl buffer pH 7.6 (Salunke *et al.*, 1984). The low-humidity form was obtained by maintaining the relative humidity at 93% by placing a saturated solution of sodium sulfate (West & Astle, 1980–1981) at a distance of $\sim$1 cm away from the crystal mounted in a closed capillary tube. The native and the low-humidity forms belong to monoclinic space group $P2_1$ with two molecules in the asymmetric unit and solvent contents of 41.9 and 36.6%, respectively (Matthews, 1968).

2.2. *Data collection*

Three-dimensional diffraction data from a native crystal were collected at 293 K using a Siemens area-detector system mounted on an Avionics rotating-anode X-ray generator using Cu $K\alpha$ radiation. The detector was kept at 10 cm from the crystal and at an angle of 22° to the direct beam. Data were processed using the *XENGEN* package (Howard *et al.*, 1987). Data from the 93% relative-humidity form were collected at 293 K on a MAR Research imaging plate with X-rays generated by a Rigaku rotating-anode generator. The crystal-to-detector distance was 12 cm. The data were processed using the *XDS* program package (Kabsch, 1988). Crystal data and data-collection statistics for both the native and low-humidity forms are given in Table 1.

2.3. *Structure solution and refinement*

The structure of the native form was solved by the molecular-replacement method employing *MERLOT* (Fitzgerald, 1988) with phosphate-free RNAase A (Wlodawer *et al.*, 1988) as the search model using 10–3 Å resolution data. The calculation led to a single solution in the rotation and translation search, even though the asymmetric unit contained two molecules. Rigid-body refinement of the model corresponding to this solution and the electron-density map calculated subsequently indicated the solution to be correct. The magnitudes of the differences between the observed and calculated structure factors for the single molecule were then used to compute a second set of rotation and translation functions which yielded the orientation and the position of the second molecule in the asymmetric unit. It turned out that the two molecules have nearly the same orientation and are separated by a quarter of the unique axis. A rigid-body refinement with the two molecules treated as independent rigid groups was then performed using *X-PLOR* (Brünger *et al.*, 1987). This model was refined further using *X-PLOR* with step-by-step addition of data to the maximum resolution, interspersed with manual model fitting to the $2F_o - F_c$ and $F_o - F_c$ electron-density maps using *FRODO* (Jones, 1978). A molecular-dynamics refinement using the simulated-annealing technique was also carried out. Water molecules were added during the course of the refinement and the positions were determined based on peaks of at least 2.5σ in $F_o - F_c$ and 1.0σ in $2F_o - F_c$ electron-density maps. Towards the end of the refinement, the model was checked against an omit map (Vijayan, 1980; Bhat & Cohen, 1984). The same procedure was also followed when refining the low-humidity form. The electron-density maps and Ramachandran plots (Ramachandran *et al.*, 1963) indicated the models to be of good quality. The final R factors, refinement parameters and r.m.s. deviation from ideal parameters are given in Table 2. Luzzati plots (Luzzati, 1952) indicated coordinate errors of 0.17 and 0.18 Å in the native and the low-humidity forms, respectively.†

3. Results and discussion

3.1. *The database*

The available crystal structures of RNAase A include five unliganded (except for bound phosphate or sulfate) forms, all grown with alcohol as the precipitant with the same space group ($P2_1$) and similar unit-cell dimensions. These alcohol forms are phosphate- and sulfate-free RNAase A (Wlodawer *et al.*, 1988; PDB code 7RSA; resolution 1.26 Å), the phosphate-bound enzyme (Wlodawer *et al.*, 1986; 5RSA; 2.0 Å), the sulfate-bound enzyme (Howlin *et al.*, 1989; 3RN3; 1.45 Å) and the 88 and 79% relative-humidity forms studied in this laboratory (Radha Kishan *et al.*, 1995; 1RHB and 1RHA; 1.5 and 1.8 Å). These will be referred to as INative, IPhosphate, ISulfate, I88 and I79 respectively. The present analysis contributes four more crystallographically independent enzyme molecules, two from the native crystals (IINativeA and IINativeB) and two from the 93% relative-humidity forms (II93A and II93B). These acetone forms have crystal structures different from those of the alcohol forms and are grown from a mother liquor with a different composition. The trigonal form (Zegers *et al.*, 1994; 1RPH; 2.2 Å) exhibits

† Atomic coordinates and structure factors have been deposited with the Protein Data Bank, Brookhaven National Laboratory (References: 1XPT, R1XPTSF, 1XPS, R1XPSSF).

Table 1. *Crystal data and data-collection statistics*

	Native	93% relative-humidity form
Space group	$P2_1$	$P2_1$
a (Å)	34.06	33.49
b (Å)	107.34	106.20
c (Å)	32.09	30.55
β (°)	98.4	102.0
Z	4	4
Unit-cell volume (Å^3)	116062	106293
Solvent content (%)	41.9	36.6
Data resolution (Å)	1.9	1.8
Number of observations	49377	43080
Number of unique reflections	16089	14420
Completeness of data (%)	88.9	74.6
Completeness (%) in the highest 0.1 Å resolution shell	48	51
Merging R for all reflections	5.54	3.84
Average $I/\sigma(I)$	31	34

Table 2. *Refinement parameters*

	Native	93% relative-humidity form
Resolution limit (Å) of reflections used in refinement	10.0–1.9	10.0–1.8
Number of reflections with $F > 2\sigma(F)$	15004	13564
Final R factor	0.162	0.175
R_{free}	0.222	0.248
R.m.s. deviations from ideal		
Bond lengths (Å)	0.01	0.01
Bond angles (°)	1.6	1.7
Dihedral angles (°)	25.8	25.9
Improper angles (°)	1.4	1.5
Number of protein atoms	1902	1902
Number of water molecules	280	246
Average B (Å^2)		
Protein atoms	13.3	15.5
Solvent atoms	30.9	33.5

yet another type of crystal packing and has a different mother-liquor composition. This form will be referred to as IIINative. The ten crystallographically independent molecules distributed among the eight crystal structures thus have differences in crystal environment, bound ion, composition of the solvent regions or/and level of hydration. A comparative study involving these should, therefore, be useful in delineating the invariant and the variable features in the enzyme molecule and the associated water.

3.2. *Intramolecular mobility and level of hydration*

It has been shown that the RNAase-A molecule is made up of two domains, domain *A* made up of residues 1–13, 49–79 and 105–124 and domain *B* consisting of residues 16–46 and 82–101, with residues 14, 15, 47, 48, 80, 81 and 102–104 constituting the hinge between them (Radha Kishan *et al.*, 1995). The hinge angle, defined as the angle subtended by the centres of mass of the two domains at the centre of mass of the hinge, in each of the ten crystallographically independent molecules is listed in Table 3. Clearly, all except I79 have very nearly the same value for this angle, in the range 94.3–95.0°, with an average of 94.7 + 0.2°. The angle in I79 is lower at 93.2°.

The r.m.s. deviation in C^α positions in each domain obtained on pairwise superposition of the ten molecules are listed in Table 4. In domain *A*, the deviations are lower when the alcohol forms are superposed among themselves. They are somewhat larger when the acetone forms are involved. The same is generally true in the case of domain *B* as well. However, in domain *B*, the deviations become substantially larger when I79 is involved in the superposition. The r.m.s. deviations in all other pairs range between 0.12 and 0.57 Å, while those involving I79 range from 0.65 to 0.90 Å.

The above results confirm that the mobility of RNAase A involves hinge-bending motion of the two domains and movement within the domains which are more pronounced in domain *B* (Radha Kishan *et al.*, 1995). The solvent content in all but I79, as estimated by Matthews' method (Matthews, 1968), range from 36.6 to 57.3% and the movements in them are comparatively low irrespective of differences in crystal packing and the composition of the mother liquor. In I79, the solvent content is as low as 30.9% and the movements are substantially larger. Thus, it would appear that the mobility of the enzyme molecules primarily involves hinge-bending motion and movements within domain *B*, and that the level of hydration is correlated to this mobility.

3.3. *Rigid and flexible regions of the molecule*

The availability of ten crystallographically independent molecules with different environmental features permit a realistic delineation of the rigid and flexible regions through comparison between them. Our earlier attempts of this nature (Madhusudan & Vijayan, 1991; Radha Kishan *et al.*, 1995) were based on manual examination of difference-distance plots (Nishikawa *et al.*, 1972). The rigid region was defined as that in which no vector among the C^α positions differs by more than 0.5 Å in any pair of structures. An approach of this type based on manual examination becomes untenable when the number of molecules is as large as ten. Therefore, the following simple computational approach was adopted for the purpose. If the distance between the C^α atom of the ith residue and that of the jth in a domain is R_{ij}, then the difference in R_{ij} between the kth and lth molecule may be specified as $d_{ijkl} = R_{ijk} - R_{ijl}$. The r.m.s. deviation for each residue in the domain is defined as r.m.s.$(i) = (\langle|d_{ijkl}|^2\rangle)^{1/2}$ where the average is taken over j, k and l. The residue with the highest r.m.s.(i) is taken as the most flexible residue. The calculation is repeated

Table 3. *Hinge angle (°) in different RNAase A structures*

INative	IPhosphate	ISulfate	I88	I79	IINativeA	IINativeB	II93A	II93B	IIINative
94.8	94.3	94.6	94.7	93.2	94.8	94.9	94.7	95.0	94.9

after eliminating the most flexible residue to identify the next most flexible residue. The calculation is repeated until the r.m.s.(i) of the most flexible residue of the set is less than a specified value. It was found that a value of 0.13 Å yielded the same proportion between residues in the rigid and the flexible regions as that obtained by the manual examination employed earlier for the five alcohol forms (Radha Kishan *et al.*, 1995). Using this cut-off value, calculations were performed separately for domain *A*, domain *B* and the hinge region.

The calculations outlined above indicated the rigid region of the molecule to be made up of residues 14, 15, 25–30, 46–48, 50, 52–61, 63, 72–75, 80–84, 97–100, 102–112 and 116–120 accounting for a little over 41% of the polypeptide chain. This region, which is contiguous, and the flexible regions are indicated in Fig. 1. In each domain, the residues in the rigid region are on an average closer to the centre of mass than those in the flexible regions. The same is true for the distances from the centre of mass of the hinge region. Furthermore, the average *B* value of atoms in the rigid region is substantially lower than that of atoms in the flexible regions in all the structures. Thus, the results obtained by superposing different structures are corroborated by the displacement parameters in individual structures.

Not surprisingly, the rigid region is almost entirely made up of residues in helices and sheets, although all residues in these secondary structural elements do not belong to the rigid region, while almost all residues in loops are in flexible regions. Three stretches of polypeptide chain, 41–64, 71–91 and 94–113 span the two domains and the hinge region, and form the backbone of the molecule. Of these, the 94–113 stretch forms part of both the mutually connected β-sheets in the molecule. Almost the entire stretch, except for a couple of residues at the termini and one in the middle, are in the rigid region. The 71–91 stretch has three segments, 71–75, 79–87 and 90–91, that form part of β-sheets. Most of the first two segments belong to the rigid region. The third stretch is made up of the 41–48 strand in the domain *B* β-sheet and the 50–60 helix and the 61–64 strand in domain *A*. The first strand is partly rigid while the helix and the second strand are substantially so. Parts of the 24–34 helix in domain *B* and 114–120 strand in the domain *A* β-sheet account for the remainder of the residues in the rigid region.

The N-terminal residues 1–13 and 16–24 belong to the flexible regions. This stretch encompasses the S peptide, which is largely helical, and the loop containing the scissile bond. Another long stretch, residues 31–45, involves the C-terminal residues of the 24–34 helix, the N-terminal residues of the 41–48 strand and the loop that connects the two. The flexible 85–96 stretch is only slightly shorter and it similarly connects the 79–87 and 94–104 strands. These two stretches together form a contiguous flexible blob in domain *B* at one end of the three long polypeptide stretches that span the entire molecule. The flexible stretches 64–71, 76–79 and 113–115 involve loops connecting strands and occur at the other end in domain *A*. The only other flexible stretch involving two or more residues is the C-terminal tetrapeptide.

3.4. *Protein hydration*

IINative and II93 contain 280 and 246 crystallographically independent water molecules, respectively. The hydration shell, made up of the water molecules that directly interact with the protein molecule, in IINativeA, IINativeB, II93A and II93B contain 115, 134, 122 and 112 water molecules, respectively. They are involved in 182, 214, 192 and 180 contacts (<3.6 Å) with

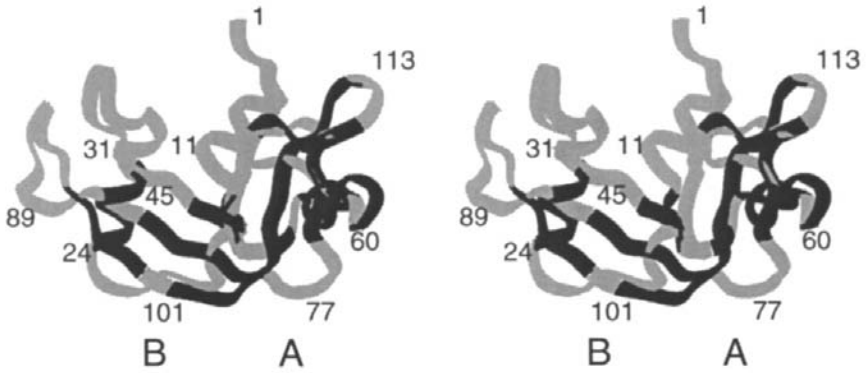

Fig. 1. Rigid (dark) and flexible (grey) regions in RNAase A delineated based on difference-distance matrices of nine crystallographically independent structures. This figure and Figs. 2, 3 and 6 were drawn using the program package *InsightII* (Biosym Technologies, 1992), and Figs. 4 and 5 were drawn using the program package *MATLAB* (MathWorks Inc., 1994).

Table 4. *The r.m.s. deviation (Å) of C^{α} atoms in domain A and domain B after domain superposition of different pairs of molecules*

	Domain	IPhosphate	ISulfate	I88	I79	IINative	IINativeB	II93A	II93B	IIINative
INative	*A*	0.13	0.12	0.13	0.22	0.21	0.40	0.45	0.414	0.28
	B	0.14	0.12	0.20	0.68	0.34	0.32	0.33	0.46	0.35
IPhosphate	*A*		0.13	0.19	0.23	0.25	0.39	0.43	0.43	0.31
	B		0.16	0.17	0.71	0.39	0.37	0.37	0.47	0.38
ISulfate	*A*			0.20	0.23	0.23	0.39	0.41	0.43	0.33
	B			0.21	0.65	0.39	0.36	0.36	0.49	0.37
I88	*A*				0.26	0.26	0.43	0.52	0.41	0.28
	B				0.68	0.41	0.40	0.38	0.50	0.40
I79	*A*					0.27	0.33	0.43	0.39	0.31
	B					0.80	0.76	0.77	0.90	0.81
IINativeA	*A*						0.33	0.40	0.40	0.29
	B						0.19	0.24	0.32	0.30
IINativeB	*A*							0.41	0.32	0.41
	B							0.25	0.39	0.30
II93A	*A*								0.53	0.57
	B								0.39	0.31
II93B	*A*									0.44
	B									0.48

O or N atoms in the protein. The average number of contacts per residue (2.3) is higher in loops than in helices (1.7) and sheets (1.4). Nearly half the water molecules in the hydration shell do not interact (<3.6 Å) with any other water molecules in the shell. Interacting pairs range from 13 to 18 in the four shells. Patches containing more than two interacting molecules are very few. The number of molecules in the largest patch in the four shells range between five and nine. Thus, as noted earlier in the case of lysozyme (Madhusudan *et al.*, 1993), the hydration shell is very discontinuous.

3.5. *Invariant water molecules and tertiary interactions*

The availability of several crystal structures of the same protein grown under different conditions permits the identification of water molecules which remain relatively invariant in the hydration shell. As in earlier work from this laboratory (Kodandapani *et al.*, 1990; Madhusudan & Vijayan, 1991; Radha Kishan *et al.*, 1995), a water molecule was considered invariant if the following conditions were satisfied when the molecules along with their hydration shells were superposed. First, there should be at least one interaction with the protein which is common to all structures and secondly, the distance between equivalent water molecules should be less than 1.8 Å after superposition. Using these criteria, 31 water molecules remain invariant in the five alcohol forms (Radha Kishan *et al.*, 1995). Naturally, the number was expected to fall when more structures were added in the analysis. The four molecules in the acetone forms reported here have another set of 32 invariant water molecules among them. The two sets and the trigonal form have 14 water molecules in common. They are listed in Table 5, together with the interactions in which they are involved. Interestingly, a substantial number of them are involved in bridging different regions of the polypeptide chain, as illustrated in Fig. 2. Thus, many water molecules involved in stabilizing the tertiary structure tend to remain invariant with respect to environmental changes.

Zegers *et al.* (1994) had earlier identified 17 water molecules common to two native forms, a cross-linked form and a chemically modified form. In the present study, only the native forms and their low-humidity variants have been used. In fact, only two crystal structures are common between the nine used by Zegers *et al.* in their analysis and the eight used in the present work. Yet, despite the considerable chemical and environmental differences between the two sets of structures used, ten out of the 14 invariant molecules identified in the present study are found among the 17 water molecules common to the structures considered by Zegers *et al.*

3.6. *Crystal structure and transformation in molecular packing*

The arrangement of molecules in IINative is illustrated in Fig. 3. That in II93 is also essentially the same except for differences in detail. The arrangement can be described as consisting of layers of molecules stacked along the *b* axis. Adjacent layers are crystallographically independent of each other and are separated by about *b*/4. Thus, there are alternating '*A* layers' and '*B* layers'. In each layer in the native structure, each protein molecule is surrounded by and is in contact (<3.6 Å) with four other molecules related by translations along *a*, *c* or both. This number increases to six in II93. In both the crystal forms each molecule is in contact with three molecules in the layer above and three molecules in the layer below. However, the number of contacts molecule

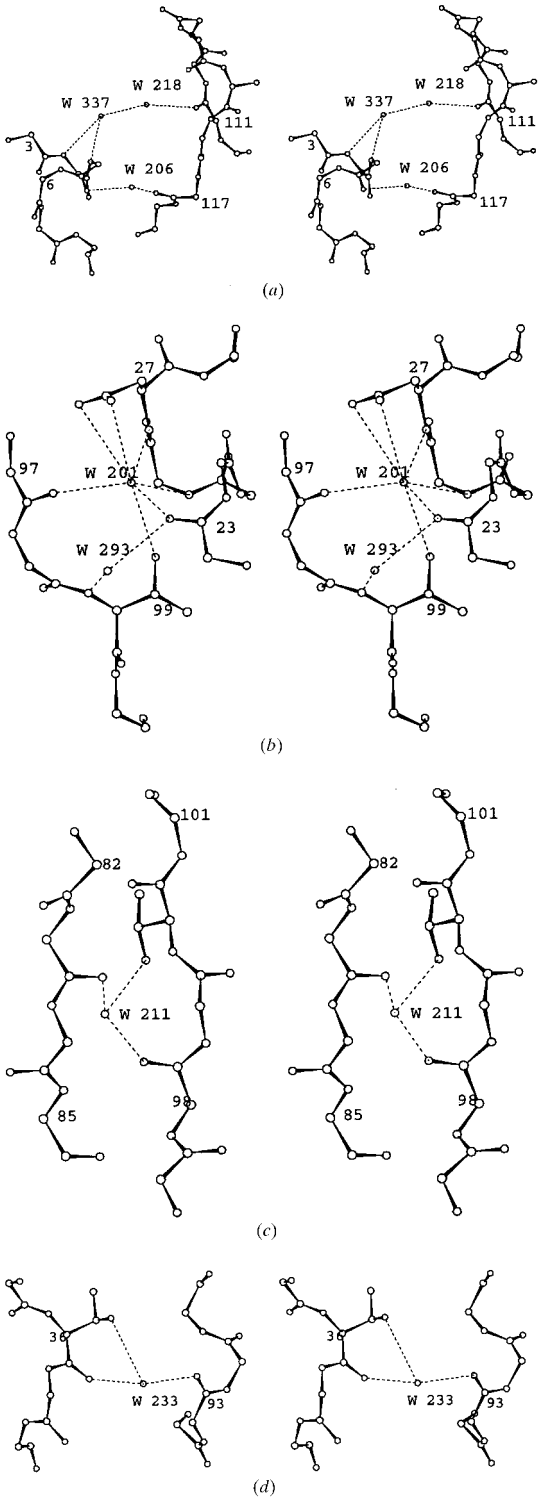

Fig. 2. Invariant water molecules involved in bridging (*a*) the N-terminal helix 3–13 and the C-terminal strands 105–113 and 114–119; (*b*) the helical region, 23–33 and the strand 94–101; (*c*) the strands 79–87 and 94–104; (*d*) two loop regions widely separated along the polypeptide chain. The numbering of water molecules is as in INative.

Table 5. *The invariant water molecules in the hydration shell of RNAase A (water numbering is as given in INative)*

Water	Protein atom(s) with which the water interacts in all the structures	Additional protein atom(s) with which the water interacts at least in one structure
201	23 O, 27 N, 97 O, 99 OG1	26 N, 27 OD1, 27 ND2
204	49 OE2, 50 N, 53 OD2	50 OG
206	117 O	5 O, 55 NE2
208	53 O, 60 NE2	
210	3 N	1 O, 3 OG1
211	83 O, 98 O, 100 OG1	83 OD2, 85 N, 85 NH1
213	6 O	
215	52 N	51 N
216	60 OE1, 76 N, 77 N	60 NE2, 77 OG
218	111 O	
233	36 O, 93 O	31 NZ, 36 OG1
293	23 O, 99 N	98 NZ
337	5 N	4 N, 3 OG1
348	14 OD2	14 OD1, 16 OG, 17 OG1, 86 OE2, 90 O

A has with molecules in layer *B* is 30 while the number of contacts with molecules in layer *B′* is only eight. Likewise, molecule *B* has 30 contacts with molecules in layer *A* and eight with those in layer *A′*. Thus, layers *A* and *B* form a distinct double layer. Naturally, the number of inter-layer contacts increases in II93. However, the *A*–*B′* (and the crystallographically equivalent *B*–*A′*) contacts increase from eight to 19, while the *A*–*B* contacts remain nearly the same. Thus, the double layers come closer in the low-humidity form but the distinctiveness of the double-layer is maintained. Such an arrangement does not occur in the alcohol forms.

The rotation required to superpose the two crystallographically independent molecules in the native form is 10.5°, while it is 5.9° in the low-humidity form. Thus, water-mediated transformation involves a rotation of the molecules even though, as seen earlier, structural changes within the molecule are marginal. As the alcohol forms contain only one molecule in the asymmetric unit, an indication of the rotation that might accompany the transformation cannot be readily obtained. Therefore, a different approach was adopted for this purpose. The two molecules in the unit cell are related by a 2_1 screw axis parallel to *b*. The existence of molecular rotation could be ascertained by comparing the two molecules in one form with the two in another. Firstly, the reference molecule in the two forms are superposed. The rotation necessary to superpose the two screw-related molecules is then evaluated. This value gives an indication of the rotation undergone by the molecule during the transformation between the two forms. This value is 0.2° when INative and I88 are compared, and 7.2° when the comparison is between I88 and I79. It may be recalled that the changes in the structure of the molecules are marginal in the transfor-

mation between the native and the 88% relative-humidity forms (Radha Kishan *et al.*, 1995), as is the rotation of the molecule as a whole. Both are substantial in the transformation between I88 and I79.

3.6.1. *Accessible surface area and hydration*. The alcohol and acetone forms, and the trigonal form of RNAase A, have different packing arrangements. Therefore, differences are expected between the two kinds of crystal structures in the regions of the molecule accessible to water. Convincing correlation between the loss of accessible surface area owing to crystal packing and the nature of hydration could not be established. However, this exploration led to the observation of an interesting relationship between atomic accessibility and hydration. Calculations (Connolly, 1983) using van der Waals radii given by Richards (1974) and a probe radius of 1.4 Å showed that nearly 60% of the N and O atoms in the ten molecules have non-zero accessible surface area. Of these, 32.1% belong to main-chain O atoms, 22.2% to main-chain N atoms, 26.7% to side-chain O atoms and 19% to side-chain N atoms. The percentage of atoms with non-zero accessibility among main-chain O atoms, main-chain N atoms, side-chain O atoms and side-chain N atoms are, however, 55.6, 38.6, 85.3 and 87.5%, respectively. The average number of water molecules interacting with atoms in different slabs with specified ranges of atomic surface area was calculated. This calculation leads to a relationship shown in Fig. 4. Hydration increases rapidly with accessible surface area, but levels off around a value a little over one at an accessible surface area of about 10 $Å^2$. Hydration appears to drop beyond an accessible surface of 20 $Å^2$. This trend appears to relate well with the increased displacement parameters of atoms with higher accessibility. The relationship between average displacement parameters and accessible surface area of N and O atoms in the structures is illustrated in Fig. 5. The higher the *B* value of an atom, the less likely it is to have a detectable water molecule attached to it. Water molecules attached to highly accessible protein atoms are unlikely to be restrained for steric reasons. They might be distributed in a number of sites with short residence times (Levitt & Park, 1993). This, along with their higher mobility, might account for the lower degree of observable hydration of protein atoms with higher accessibility.

Yet another observation in relation to accessibility and hydration deserves particular mention. In the protein molecules, several N and O atoms which are formally inaccessible were found to interact with well ordered water molecules in the hydration shell. All the structures, especially the alcohol and acetone forms, have been refined well at reasonably high resolution and the observation was unlikely to be an artifact. Most of them satisfied the hydrogen-bonding criteria employed in our earlier studies (Madhusudan *et al.*, 1993). Furthermore, most of the interactions occurred in more

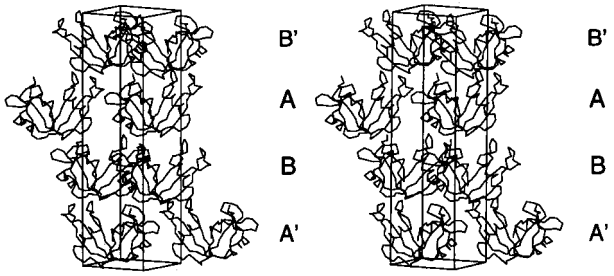

Fig. 3. The crystal structure of IINative. The layers parallel to the *ac* plane formed by the crystallographically independent molecules *A* and *B* and symmetry-related molecules *A'* and *B'* are indicated.

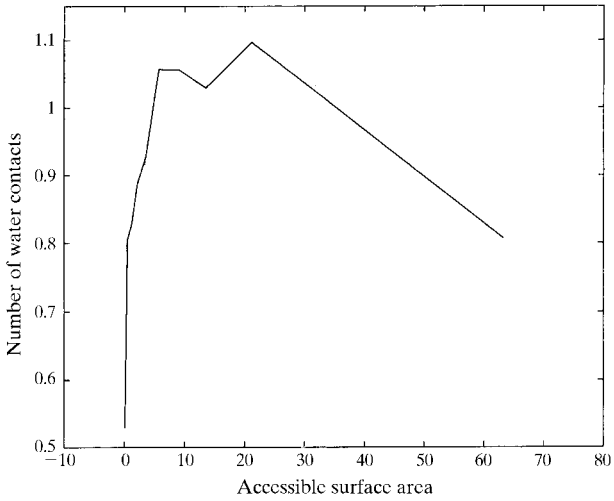

Fig. 4. The relation between the average number of water contacts of N and O atoms and the average accessible surface area of these atoms.

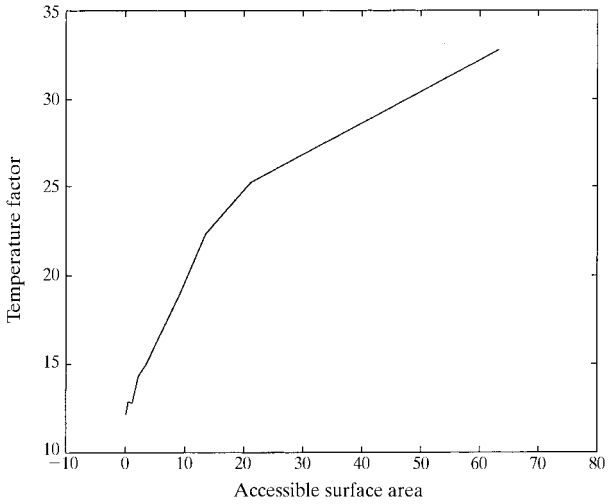

Fig. 5. The relation between the displacement parameters of N and O atoms and the average accessible surface area of these atoms.

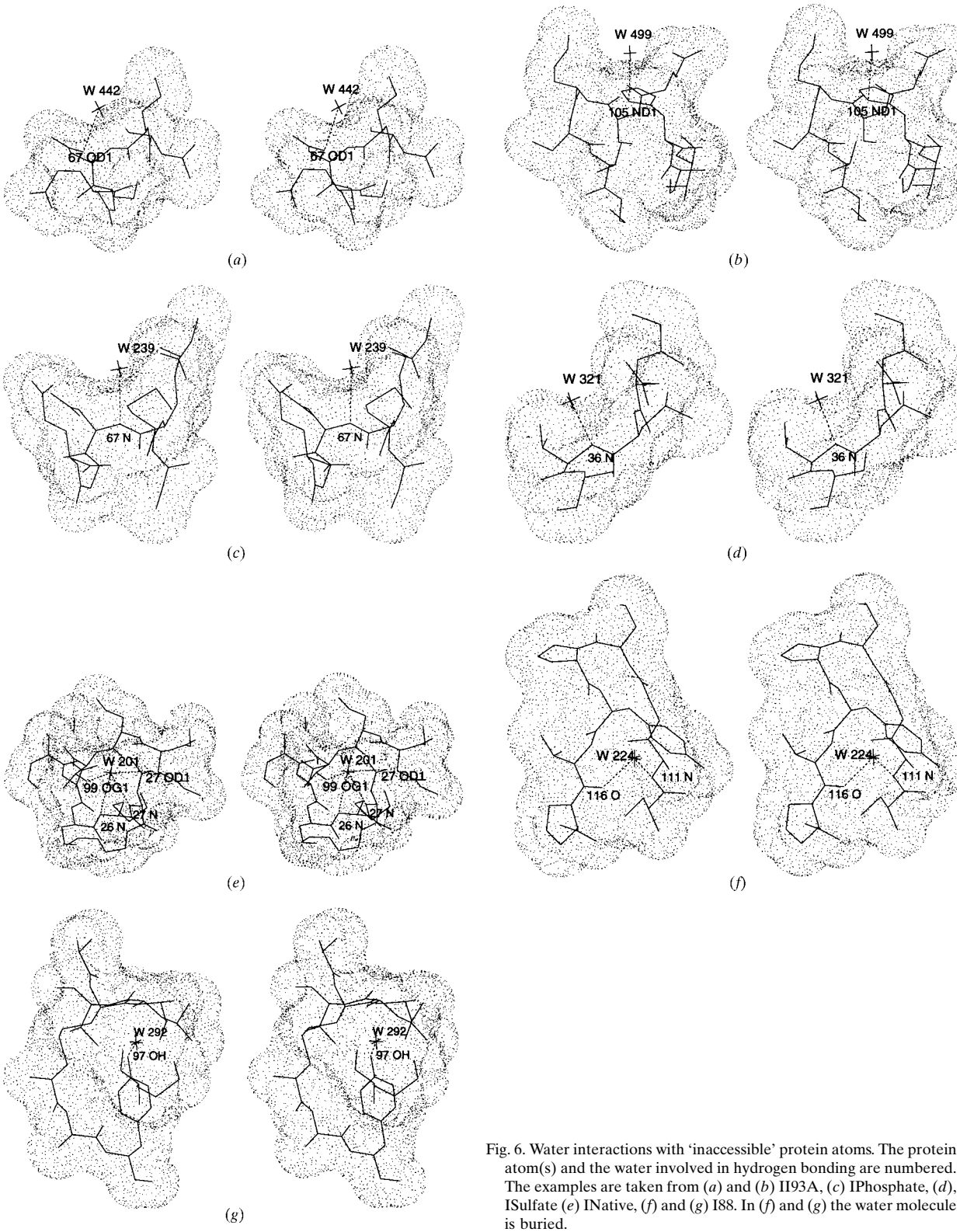

Fig. 6. Water interactions with 'inaccessible' protein atoms. The protein atom(s) and the water involved in hydrogen bonding are numbered. The examples are taken from (*a*) and (*b*) II93A, (*c*) IPhosphate, (*d*), ISulfate (*e*) INative, (*f*) and (*g*) I88. In (*f*) and (*g*) the water molecule is buried.

than one structure, adding to the confidence on the reliability of the observation, despite the expected errors in individual structures. A few examples of these forbidden interactions are given in Fig. 6. In most instances, the protein atom is buried according to the results of accessibility calculations (Connolly, 1983), while the water molecule is situated at or outside the accessibility surface of the protein molecule. There are also a few instances where the water molecule is also buried, although the accessibility calculations indicated insufficient space to accommodate a water molecule at the observed position.

A careful consideration suggests that the van der Waals radii normally used in accessibility calculations (Richards, 1974) are not appropriate for all situations. This becomes apparent when the range of hydrogen-bonded distances involving water is taken into account. On the basis of a careful analysis of the relevant crystal structures, Steiner (1995) indicates the typical C···O*W* distance in a C—H···O*W* hydrogen bond to be 3.3 Å, and can be as low as 3.1 Å. However, the van der Waals radius normally used and a probe radius of 1.4 Å do not permit this distance to be less than 3.4 Å. Similarly, N···O*W* and O···O*W* distances in N—H···O*W* and O—H···O*W* hydrogen bonds in proteins have been shown to generally be in the range 2.9–3.1 and 2.7–2.8 Å, respectively (Thanki *et al.*, 1991). Considerably smaller distances have also been observed in many proteins (Baker & Hubbard, 1984; Stickle *et al.*, 1992). The lower ends of these distance ranges cannot be attained on the basis of currently used van der Waals and probe radii. It was observed that most of the forbidden interactions in RNAase A structures become possible if the van der Waals radii are each reduced by 0.2 Å. Nearly the same effect can be achieved by reducing the probe radius by 0.2 Å.

The inadequacy of the parameters used currently for describing accurately the accessibility of protein atoms should be a matter of concern as they are widely used to deal with a variety of situations involving, for example, protein–protein, protein–ligand and inter-subunit interactions. Efforts to assess and modify them, wherever necessary, are warranted. In any case, the results of accessibility calculations should be interpreted with a certain degree of flexibility.

Intensity data were collected using the Area Detector/Imaging Plate Facility supported by the Department of Science and Technology (DST) and the Department of Biotechnology (DBT). The computations were carried out at the Supercomputer Education and Research Centre of the Institute and the DBT-supported Interactive Graphics Based Molecular Modelling Facility. The work forms part of a DST-funded project.

References

Artymiuk, P. J., Blake, C. C. F., Grace, D. E. P., Oatley, S. J., Phillips, D. C. & Sternberg, M. J. E. (1979). *Nature (London)*, **280**, 563–568.
Baker, E. N. & Hubbard, R. E. (1984). *Prog. Biophys. Mol. Biol.* **44**, 97–179.
Bhat, T. N. & Cohen, G. H. (1984). *J. Appl. Cryst.* **17**, 244–248.
Biosym Technologies (1992). *InsightII. Version 2.1.0*. Biosym Technologies Inc., 9685 Scranton Road, San Diego, CA 92121-2777, USA.
Brooks, C. L. III & Karplus, M. (1989). *J. Mol. Biol.* **208**, 159–181.
Brünger, A. T., Kuriyan, J. & Karplus, M. (1987). *Science*, **223**, 458–460.
Caspar, D. L. D. & Badger, J. (1991). *Curr. Opin. Struct. Biol.* **1**, 877–882.
Connolly, M. L. (1983). *J. Appl. Cryst.* **16**, 548–558.
Fitzgerald, P. M. D. (1988). *J. Appl. Cryst.* **21**, 273–278.
Frey, M. (1994). *Acta Cryst.* D**50**, 663–666.
Howard, A. J., Gilliland, G. L., Finzel, B. C., Poulos, T. L., Ohlendorf, D. H. & Salemme, F. R. (1987). *J. Appl. Cryst.* **20**, 383–387.
Howlin, B., Moss, D. S. & Harris, G. W. (1989). *Acta Cryst.* A**45**, 851–861.
Huber, R. (1988). *Angew. Chem. Int. Ed. Engl.* **27**, 79–88.
Jones, T. A. (1978). *J. Appl. Cryst.* **11**, 268–272.
Kabsch, W. (1988). *J. Appl. Cryst.* **21**, 916–924.
Kodandapani, R., Suresh, C. G. & Vijayan, M. (1990). *J. Biol. Chem.* **265**, 16126–16131.
Levitt, M. & Park, B. H. (1993). *Structure*, **1**, 223–226.
Luzzati, V. (1952). *Acta Cryst.* **5**, 802–810.
Madhusudan, Kodandapani, R. & Vijayan, M. (1993). *Acta Cryst.* D**49**, 234–245.
Madhusudan & Vijayan, M. (1991). *Curr. Sci.* **60**, 165–170.
MathWorks Inc. (1994). *MATLAB. Version 4.2c*. MathWorks Inc., 24 Prime Parkway, Natick, Massachussetts 01760-1500, USA.
Matthews, B. W. (1968). *J. Mol. Biol.* **33**, 491–497.
Nadig, G., Ratnaparkhi, G. S., Varadarajan, R. & Vishveshwara, S. (1996). *Protein Sci.* **5**, 2104–2114.
Nagendra, H. G., Sudarsanakumar, C. & Vijayan, M. (1995). *Acta Cryst.* D**51**, 390–392.
Nagendra, H. G., Sudarsanakumar, C. & Vijayan, M. (1996). *Acta Cryst.* D**52**, 1067–1074.
Nishikawa, K., Ooi, T., Isogai, Y. & Saito, N. (1972). *J. Phys. Soc. Jpn*, **32**, 1331–1337.
Otting, G., Liepinsh, E. & Wuthrich, K. (1991). *Science*, **254**, 974–980.
Poole, P. L. & Finney, J. L. (1983). *Int. J. Biol. Macromol.* **5**, 308–310.
Radha Kishan, K. V., Chandra, N. R., Sudarsanakumar, C., Suguna, K. & Vijayan, M. (1995). *Acta Cryst.* D**51**, 703–710.
Ramachandran, G. N., Ramakrishnan, C. & Sasisekharan, V. (1963). *J. Mol. Biol.* **7**, 95–99.
Richards, F. M. (1974). *J. Mol. Biol.* **82**, 1–14.
Ringe, D. & Petsko, G. A. (1985). *Prog. Biophys. Mol. Biol.* **45**, 197–235.
Saenger, W. (1987). *Annu. Rev. Biophys. Biophys. Chem.* **16**, 93–114.
Salunke, D. M., Veerapandian, B., Kodandapani, R. & Vijayan, M. (1985). *Acta Cryst.* B**41**, 431–436.

Salunke, D. M., Veerapandian, B. & Vijayan, M. (1984). *Curr. Sci.* **53**, 231–235.

Santoro, J., Gonzalez, C., Bruix, M., Neira, J. L., Nieto, J. L., Herranz, J. & Rico, M. (1993). *J. Mol. Biol.* **229**, 722–734.

Steiner, T. (1995). *Acta Cryst.* D**51**, 93–97.

Stickle, D. F., Presta, L. G., Dill, K. A. & Rose, G. D. (1992). *J. Mol. Biol.* **226**, 1143–1159.

Teeter, M. M. (1991). *Annu. Rev. Biophys. Biophys. Chem.* **20**, 577–600.

Thanki, N., Thornton, J. M. & Goodfellow, J. M. (1988). *J. Mol. Biol.* **202**, 637–657.

Thanki, N., Umrania, Y., Thornton, J. M. & Goodfellow, J. M. (1991). *J. Mol. Biol.* **221**, 669–691.

Vijayan, M. (1980). *Computing in Crystallography*, edited by R. Diamond, S. Ramaseshan & K. Venkateshan, pp. 19.01–19.26. Bangalore: Indian Academy of Sciences.

West, R. C. & Astle, M. J. (1980–1981). Editors. *CRC Handbook of Chemistry and Physics*, 61st ed., p. E-46. Florida: CRC Press.

Wlodawer, A., Borkakoti, N., Moss, D. S. & Howlin, B. (1986). *Acta Cryst.* B**42**, 379–387.

Wlodawer, A., Svensson, L. A., Sjolin, L. & Gilliland, G. L. (1988). *Biochemistry*, **27**, 2705–2717.

Zegers, I., Maes, D., Dao-Thi, M., Poortmans, F., Palmer, R. & Wyns, L. (1994). *Protein Sci.* **3**, 2322–2339.