Crystallographic Structures of Ribonuclease S Variants with Nonpolar Substitution at Position 13: Packing and Cavities‡

R. Varadarajan§ and F. M. Richards*

Department of Molecular Biophysics and Biochemistry, Yale University, New Haven, Connecticut 06511

Received March 26, 1992; Revised Manuscript Received July 21, 1992

ABSTRACT: Seven hydrophobic residues ranging in size from glycine to phenylalanine have been substituted for the wild-type methionine residue at position 13 in a 15-residue truncated version (S15) of S-peptide, the small component of ribonuclease S. Complexes of both S-15 and the seven variants with S-protein yielded isomorphous crystals. The structures of all eight complexes have been refined to final *R*-factors in the range of 17–19%. [See Kim, E. E., Varadarajan, R., Wyckoff, H. W., and Richards, F. M. (1992) *Biochemistry* (preceding paper in this issue) for the description of the reference S-15 complex.] Multiple side-chain conformations were seen for six residues in all of the complexes and for two to three additional residues in at least some of the complexes. Three of the complexes, Gly, Ala, and α-amino-*n*-butyric acid (ANB), contained a single water molecule in the cavity near residue 13 that makes three hydrogen bonds to protein atoms. Although space is available, no evidence for additional water in this region, ordered or disordered, was found. The atoms in the cavity wall tend to shrink the cavity by moving in on the small residues and to swell the cavity by moving out for the larger Phe substitution. A swelling seen with leucine was attributed to a shape effect since Leu, Ile, and Met all have the same volume. A slight volume contraction of the collection of interior residues outside of the region of position 13 was also noted. (All changes noted are in the direction to maintain a constant packing density averaged over the whole protein.) Leu51, a surface hydrophobic residue, moved considerably in the G, A, and ANB complexes in directions which would tend to decrease the cavity volume. The only other major change in position, 1.5 Å, was the 66–69 loop, which is about 25 Å from position 13. His12, Phe120, and Asp121 appear to be involved in this movement, but the connection with position 13 is not clear at all. The thermodynamic data on the association reaction for all of these complexes have been previously reported [Connelly, P. R., Varadarajan, R., Sturtevant, J. M., & Richards, F. M. (1990) *Biochemistry 29*, 6108–6114; Varadarajan, R., Connelly, P. R., Sturtevant, J. M., & Richards, F. M. (1992) *Biochemistry 31*, 1421–1426]. Some comments are offered on our initial attempts to correlate the structural changes with the changes in the thermodynamic parameters. Now that both are in hand, a serious effort to connect the thermodynamic and structural data would appear useful.

To understand the relationship between protein sequence and structure, it is important to know in quantitative detail the various interactions that stabilize a folded protein relative to its unfolded state. A characteristic feature of the folded states of all globular proteins is that they are stabilized by interactions between buried nonpolar groups. A number of studies have examined the effects of substitutions of buried hydrophobic amino acids on the stability of globular proteins (Kellis et al., 1989; Shortle et al., 1990; Eriksson et al., 1992). In a different approach, we have studied the effects of specific substitutions of buried hydrophobic residues on the bimolecular binding reaction between S-protein and S-peptide analogs using titration calorimetry. The advantages of such a system are isothermal measurements, no solvent perturbation, and no required extrapolations (Connelly et al., 1990). This association can be considered as a model for a late stage in the folding process or as an example of an effector–receptor interaction.

Bovine pancreatic ribonuclease A (RNase-A) may be cleaved with subtilisin to give two fragments, S-peptide (residues 1–20) and S-protein (residues 21–124) (Richards & Vithyathil, 1959). These fragments can be reconstituted to give the catalytically active complex ribonuclease S (RNase-S), having a structure very similar to that of native RNase-A (Wyckoff et al., 1970). Since residues 16–20 of S-peptide are not important for binding to S-protein, we have used a truncated version of S-peptide consisting of the first 15 residues with a C-terminal amide designated, S15. Seven hydrophobic residues ranging in size from Gly to Phe have been substituted for Met13 in S15. We have reported estimates of the free energies, enthalpies, entropies, and specific heats of binding of the various peptide analogs for S-protein in the temperature range 5–25 °C (Connelly et al., 1990; Varadarajan et al., 1992). There were significant changes in binding parameters as a result of these substitutions. Structural data are essential in any attempt to interpret these thermodynamic changes, and fortunately, all of the peptide–S-protein complexes have been successfully crystallized. The structure of the reference S15 complex is reported in the accompanying paper (Kim et al., 1992). We focus here on the structural differences between the S15 and mutant complexes.

Examination of a large number of protein crystal structures shows that interior residues are packed together in such a manner as to ensure that there are no steric overlaps between residues, no large empty spaces, and few buried charges in the interior of a protein. Furthermore, the conformations of the

‡ Both the model coordinate data and the structure factors for the mutant structures discussed in this paper have been deposited in the Protein Data Bank at Brookhaven with no restrictions. The PDB codes for the mutant, coordinate file, and structure factor file are as follows: M13A, 1RBC, 1RBC-SF; M13ANB, 1RBD, 1RBD-SF; M13F, 1RBE, 1RBE-SF; M13G, 1RBF, 1RBF-SF; M13I, 1RBG, 1RBG-SF; M13L, 1RBH, 1RBH-SF; M13V, 1RBI, 1RBI-SF.

* Author to whom correspondence should be addressed.

§ Present address: Molecular Biophysics Unit, Indian Institute of Science, Bangalore 560 012, India.

0006-2960/92/0431-12315$03.00/0

residue side chains are largely restricted to the staggered positions expected for the single bond dihedral angles. These observations led to the suggestion that packing interactions between interior residues were crucial in determining the protein fold, and the above constraints were used to devise a computer program to predict possible lists of sequences compatible with a given protein tertiary structure (Ponder & Richards, 1987). Recent work by Shortle et al. (1990) and Lim and Sauer (1991) has shown that proteins are remarkably tolerant to a wide range of single-site substitutions. This has led to the suggestion that specific packing interactions between residues are not what determines the final folded structure (Behe et al., 1991) but that it is the sequence pattern of hydrophobicity that is important. However, since no structural data are currently available for these mutants, it is not possible to verify this hypothesis.

With regard to permitted sets of interior residues, it is important to know (a) the energetics of cavity formation within a protein, (b) to what extent a cavity can be reduced by rearrangements of surrounding protein side chains or solvent to fill the space, and (c) the extent and nature of main-chain rearrangements that are allowed. Recent studies of cavity formation in mutants of T4 lysozyme by Eriksson et al. (1992) as well as the work described here address some of these questions.

MATERIALS AND METHODS

(*a*) *Purification and Crystallization.* Bovine pancreatic RNase-S and cytidine 3′-monophosphate (3′CMP) were obtained from Sigma Chemical Co. in the form of lyophilized powders. Stock solutions of 20 mM 3′CMP in 0.1 M acetate buffer, pH 5.5, were prepared and stored at −20 °C. S-Protein and analogs of S-peptide were obtained, purified, and quantitated as described earlier (Connelly et al., 1990).

The following crystallization procedure was used for the M13G and M13A complexes. The S-protein and S-peptide analog in water were mixed together in a molar ratio of 1:5. The final S-protein concentration was approximately 1 mg/mL. The complexes were lyophilized and dissolved in 6.0 M CsCl, 0.1 M acetate buffer, pH 5.75, and 4 mM 3′CMP at 4 °C. The final protein concentration was about 8–12 mg/mL in a total volume of about 15 μL. To this was added an equal volume of 60–70% saturated $(NH_4)_2SO_4$ at 4 °C. The solutions were mixed, and four 7-μL drops were aliquoted to adjacent wells of a 96-well tissue culture plate and sealed with a cover slip. The tissue culture plate was then removed from the cold, and usable crystals were obtained after a few days at room temperature. For the remaining complexes the procedure followed was identical except that no 3′CMP was used and all steps were carried out at room temperature. 3′CMP was added to the A and G mutants in an attempt to ensure association of the peptide and S-protein components whose binding constants are the weakest in the series. The binding constants for all the other mutants were sufficiently high so that no addition of 3′CMP was necessary.

(*b*) *Data Collection and Refinement.* The crystals were stabilized in 75% $(NH_4)_2SO_4$ at pH 4.75, and CsCl and 3′CMP were washed out by repeated buffer changes. X-ray data were collected from capillary-mounted crystals with a San Diego Multiwire System detector on a Rigaku RU300 rotating anode generator operated at 50 kV and 200 mA. The beam passed through a graphite monochromator and 0.5-mm collimator. Each frame was collected for 20 s with a 0.1° sweep. Refinements were carried out by simulated annealing with the slow cooling protocol of X-PLOR (Brünger, 1990). Further details on data collection and reduction, cutoff values, and so forth are given by Kim et al. (1992). All calculations were carried out on a VAX 8800 and Convex C-2 at the Center for Structural Biology at Yale. The graphics terminal was an E&S PS 390.

The refined wild-type coordinates without water, sulfate, or alternate conformations were used as the starting model for all refinements. $2F_o - F_c$ maps using the calculated phases and amplitudes obtained from the starting model were used to position the substituted side chain. Refinement, including addition of water, sulfate, and multiple conformations, was then carried out primarily using the program X-PLOR (Brünger, 1990) as described in Kim et al. (1992).

(*c*) *Cavity Volume Estimation.* The description of cavities is a complex problem. Although bounded by concave spherical surfaces, there are no fixed sizes and shapes as there are with the atoms themselves. Approaches to the problem have been provided by Rashin et al. (1986), Connolly (1991), and Alard and Wodak (1991). For the purposes of this study we have focused on changes in volume using either the Voronoi procedure or, in small regions, a grid search approach. One should note that, regardless of the method used, cavities are undefinable in surface regions in the absence of a completely defined solvent layer. The latter is elusive. Only in one instance are the mean solvent positions completely defined in a protein X-ray structure (Teeter, 1984).

The Voronoi Procedure. For a bounded set of points, the Voronoi procedure provides a clearly defined and mathematically accurate method for dividing up space. A plane is constructed between each atom pair. The intersection of such planes provides a polyhedron around each atom. Collectively, these polyhedra account for all of the space occupied by the set of points. There are two ways of defining how the planes are drawn. As suggested by Gellatly and Finney (1982), we have used the radical plane modification which allows one to take into account the differing intrinsic atomic sizes. The application of this method to macromolecules has been reviewed by Richards (1977).

The faces of the polyhedron which the procedure constructs around each atom lie outside of the van der Waals sphere around the atom except where the sphere overlaps those of neighboring atoms. Such overlaps occur regularly between bonded atoms, between next nearest neighbors, and often between one to four atom pairs. Atoms more distant in covalent connection will overlap only rarely. Such occurrences may be due to coordinate errors or to situations where the energy penalty can be tolerated.

In addition to enclosing the van der Waals envelope of the atom, each polyhedron includes a share of the cavity space between the atoms. Collectively, the polyhedra include all of the cavity space. While the van der Waals volume of each atom is essentially independent of conformation, the Voronoi volume will change with the nature of the local packing. Changes in the Voronoi volume of a given atom in different structures reflect solely the changes in the associated cavities.

A modified version of the program VOLUME has been used in the current study. [The earlier versions are reviewed by Richards (1985).] In addition to providing the Voronoi volume and atom-pair cavity volumes, the program puts out the centroid of the polyhedral plane separating each atom pair and the distance of separation of their van der Waals surfaces. These data have been used, as described later, for the estimation of the mean electron density at locations in the unit cell that are certain *not* to contain water molecules.

Table I: Summary of Data and Refinement of Peptide–S-Protein Complexes

	M13G	M13A	M13ANB	M13V	M13I	M13L	S15[a]	M13F
$a = b$ (Å)	44.45	44.51	44.45	44.29	44.30	44.31	44.45	44.76
c (Å)	97.28	97.04	97.18	97.63	97.65	97.94	97.60	96.62
resolution (Å)	1.8	2.0	1.7	1.8	1.8	1.7	1.6	1.75
% unique refl measd	97	98	95	96	95	94	95	94
av $I/\sigma(I)$	12.3	8.9	10.4	9.3	13.0	13.9	13.1	13.3
R_{merge} (%)	3.9	6.0	4.3	5.7	4.6	5.0	5.7	4.6
R_{final} (%)	19.4	16.7	17.7	17.1	17.6	18.2	18.6	18.6
rms deviations in								
bond lengths (Å)	0.013	0.012	0.012	0.013	0.013	0.012	0.013	0.013
bond angles (deg)	2.86	2.82	2.74	2.78	2.81	2.72	2.79	2.74
av B-factor for main chain (Å^2)	21.3	21.3	21.0	20.6	19.5	18.5	18.0	19.2
no. of waters	62	56	57	60	57	58	70	63
no. of residues in two conf	9	8	9	9	9	8	8	7

[a] Data taken from Kim et al. (1992).

In VOLUME the protein is surrounded on a coarse grid with SHELL, a collection of positions needed to produce closed polyhedra for the surface atoms. While satisfying the mathematical requirement, these positions have no physical significance since they do not represent a realistic solvent layer. The volumes of the surface atoms, so calculated, are useful for some purposes but less so in the present study, unless otherwise indicated. Only atoms whose polyhedra are defined by two or less SHELL positions in addition to other protein atoms have been included in the summations. These provide the "core" of the protein, and their geometric properties are essentially unaffected by the uncertainties at the surface. The average atom will have about 16 neighbors leading to faces in the Voronoi polyhedron.

(*d*) *Water Grid Calculation.* The cavity space around residue 13 has also been examined with the help of a 0.5-Å grid centered at the position of the sulfur atom of the methionine in the "wild-type" variant and extending 6 Å in each coordinate direction. Grid positions outside of the van der Waals envelope of any atom were flagged, and those which could serve without steric overlap as the center of a "mini water molecule", and edge-truncated 5 × 5 × 5 cube, were identified. A 5 × 5 × 5 cube in the grid with the 44 edge positions removed served as a rough approximation of a 1.3-Å radius sphere. The cube-center positions occurred in clearly separated clusters in the full grid. The cluster related to the residue in position 13 was selected by visual examination on a graphics terminal with the program FRODO (Jones, 1982).

RESULTS AND DISCUSSION

(*a*) *Structure and Refinement.* The results of the final refinements are summarized in Table I. Detailed descriptions of the wild-type complex can be found in Kim et al. (1992), which presents new data and refinement of the original structure reported by Wyckoff et al. (1970), Richards et al. (1971), and Richards and Wyckoff (1971, 1973). All of the mutant complexes crystallized in the same space group as the wild type and with similar unit cell parameters. With the exception of M13A at 2.0 Å, all of the mutant structures have nominal resolutions between 1.6 and 1.8 Å and R-factors between 17.1 and 19.4. The location of Met13 in a ribbon diagram of the wild-type structure is shown in Figure 1. Met13 occurs at the C-terminal end of the α-helix that extends from residues 3 to 13. The side chain is 94% buried and has an accessible area of 10 Å^2 (Lee & Richards, 1971). The side chain packs against main-chain atoms from residues 9, 12, 14, and 49–51 and side chains of Ser15, Leu51, and Val54. The final models for the mutants include between 919 and 926 protein non-hydrogen atoms, 56–63 water molecules, and

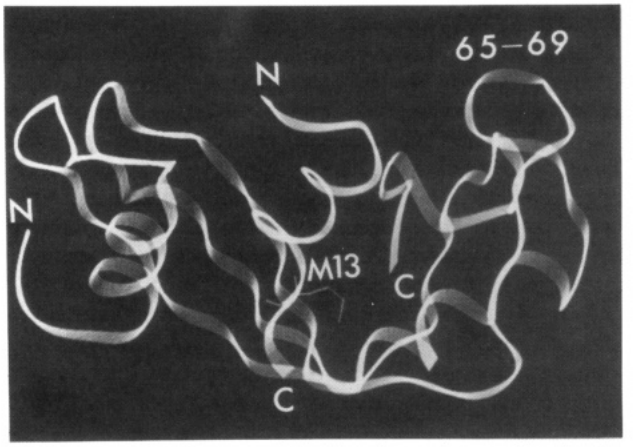


FIGURE 1: Ribbon diagram of the S15–S-protein complex showing the location of Met13 and the loop region which shows large differences in some mutants relativew to the wild type.

1 sulfate ion. The side chains of six residues were modeled in two conformations in all eight complexes (Table IIa). In addition, a few other residues are modeled in two conformations in some of the complexes (Table IIb). Structural differences between the various complexes are described in detail below.

(*b*) *Main-Chain Shifts.* Differences in the main-chain coordinates (C, O, N, Cα) in the mutant complexes relative to the wild type have been analyzed both by difference matrices (Nishikawa et al., 1972; Richards & Kundrot, 1988) and by least squares superposition with the routines in X-PLOR.

The difference distance maps comparing the reference structure S15 and four of the mutant structures are shown in Figure 2. Similar maps for the other three mutants, V, I, and L, are essentially featureless at the contour levels shown, indicating that these structures are identical to S15 within 0.3 Å for each Cα position.

The rms match of each mutant structure to S15 was established using only main-chain atoms with the omissions listed in the legend. The magnitudes of the difference in position of these atoms for three mutants are shown in Figure 3 as a function of the distance of the atom from the Cα of residue 13 in S15. The conformations of the main chain are also reflected in the shifts in the ϕ and ψ angles, shown for the Gly mutant in Figure 4. The standard deviations of these angles in all eight complexes are shown in Figure 5.

Residues 1 and 124. These residues are disordered in the final electron density maps. No significance should be attached to the structural parameters provided by the refinements for these two terminal residues.

The 66–69 Loop. The largest shifts in the main-chain position, up to 1.5 Å, are associated with this loop region.

Table II: Residues Found or Modeled in Two Conformations

(a) Residues Found in Two Conformations in All Eight Complexes[a]

residue	conf	*B*	*q*	χ_1	χ_2	χ_3	χ_4	χ_5
Ser32	1	33	0.42	−53 (3)				
	2	42	0.59	43 (6)				
Leu35	1	15	0.48	−56 (2)	−62 (2)			
	2	16	0.52	−72 (3)	−148 (6)			
Val43	1	20	0.49	56 (3)				
	2	21	0.51	−53 (5)				
Ser77	1	26	0.59	53 (7)				
	2	28	0.41	−168 (12)				
Asp83	1	20	0.47	177 (2)	82 (4)			
	2	26	0.53	−66 (7)	116 (7)			
Arg85	1	30	0.37	−170 (3)	158 (5)	−81 (21)	96 (4)	179 (0)
	2	22	0.63	187 (36)[b]	144 (57)[b]	9 (91)[b]	−150 (44)[b]	182 (2)

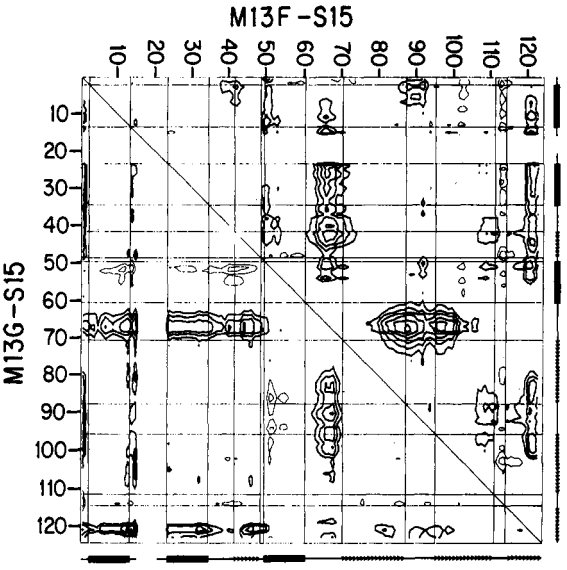
(b) Residues Modeled in Two Conformations in Some of the Complexes[c]

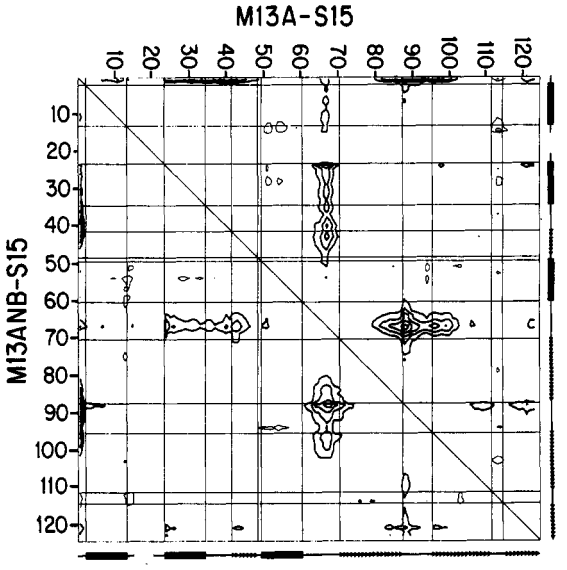
complex	residue	*B*	*q*	χ_1	χ_2	χ_3	χ_4	χ_5
M13G	Arg39	36	0.65	74	179	−178	−57	−177
		38	0.35	80	−94	−75	176	178
	Leu51	40	0.68	−96	178			
		19	0.32	−93	68			
	Phe120	27	0.54	−106	−54			
		54	0.46	−127	−116			
M13A	Gln28	42	0.56	−66	−68	151		
		29	0.44	−68	−180	81		
	Leu51	28	0.56	−174	172			
		49	0.44	−108	−54			
M13ANB	Gln28	35	0.58	−62	−77	167		
		41	0.42	−63	178	68		
	Arg39	39	0.52	58	154	175	159	178
		28	0.48	67	−179	71	80	178
	Leu51	35	0.56	−175	174			
		45	0.44	−99	−55			
M13V	Gln28	41	0.51	−66	−70	157		
		43	0.49	−62	178	63		
	Arg39	39	0.55	56	157	177	172	178
		9	0.45	83	−99	−70	173	178
	Val47	9	0.65	−177				
		22	0.35	61				
M13I	Gln28	33	0.35	162	−145	−160		
		48	0.65	−62	−178	73		
	Gln69	40	0.46	−113	−70	−127		
		47	0.54	−76	166	49		
	Lys98	27	0.7	178	−178	176	56	
		30	0.3	−76	−64	159	61	
M13L	Arg39	39	0.50	75	−169	72	90	179
		39	0.50	62	156	171	159	179
	Lys98	36	0.65	175	−175	179	62	
		32	0.35	−72	−62	164	62	
M13M	Met29	24	0.51	−164	60	141		
		15	0.49	−75	−73	−61		
	Lys98	31	0.24	−168	−131	−53	87	
		38	0.76	179	−169	170	66	
M13F	Val47	12	0.51	−171				
		14	0.49	58				

[a] The values of the temperature factor (*B*), the occupancy (*q*), and the χ values for each side chain listed are the averages of the values in all eight complexes. Standard deviations in each χ are also indicated in parentheses. [b] The large standard deviations arise not from genuine conformational differences between mutants but rather from the fact that Arg85 cannot be adequately represented by two alternate conformations. However, the quality of the maps is not sufficient to justify inclusion of a third conformation for this residue in the model. [c] The *B* values are averaged over all side-chain non-hydrogen atoms.

Within a single molecule this loop is approximately 25 Å from residue 13, the site of mutation. The extent of loop motion is not correlated with changes in unit cell parameters (Table I) but does appear to be correlated with the size difference between Met13 in S15 and the mutant residues at position 13. The extent of both overall structural rearrangement and loop motion decreases in the order M13F > M13G > M13A > M13ANB > M13V > M13L > M13I. Note that the loop shifts *away* from the rest of the molecule whether the substitution for M13 is G (smaller) or F (larger). The same direction of the loop shift is seen in all of the mutant structures (Figure 6). None produce a shift *toward* the neighboring part of the molecule. Residues 66–69 pack against residues 42, 85, and 115 on a symmetry-related molecule. This latter region does not move relative to the rest of the molecule (Figure 2). These facts suggest that the differences in the 66–69 loop are not artifacts caused by differences in crystal packing.

The origin of the loop movement remains unclear. The OD1 and OD2 of Asp121 are hydrogen bonded to N66 and form the only intramolecular contact (distance ≤4 Å) between residues 66–69 and the rest of the molecule. The movement of Asp121 appears to be associated with the movement of the 66–69 loop. Asp121 is affected by its main-chain connection to Phe120, which in turn is shifted, especially in the G and F mutants. The Cα movements in residues 120 and 121 are clearly visible in Figure 2A. Phe120 is adjacent to Val47,





M13G - M13M

M13F - M13M

	8	12	13	14	15	33	46	47	49	50	51	54
8		1	0	0	0	0	1	0	0	0	-2	-1
12	0		0	0	0	0	0	-1	-3	-2	-6	-3
13	-2	0		0	0	0	2	3	0	0	-5	0
14	-1	2	0		0	2	0	1	-3	-3	-6	-2
15	5	5	6	0		0	2	4	1	-2	-5	-1
33	0	1	0	0	0		0	0	-1	0	-2	-1
46	0	0	-1	1	0	1		0	-2	-1	-5	-1
47	-1	0	-1	2	0	0	0		-1	0	-5	-1
49	1	4	3	2	1	2	1	2		0	-3	0
50	5	7	7	3	6	4	3	2	0		0	0
51	4	5	3	-2	4	2	0	-1	-2	0		-2
54	1	2	0	-2	2	0	0	-2	-2	0	-1	

FIGURE 2: Difference distance plots between Cα atoms. Thick lines represent distance increases while thin lines represent distance decreases. The minimum contour levels are 0.3 Å; subsequent contours are in increments of 0.2 Å. The disordered residues 21 and 23 were omitted from the graphs. The upper and lower triangles each contain a different plot. (Top panel) Upper, M13F vs S15; lower, M13G vs S15. (Middle panel) Upper, M13A vs S15; lower, M13ANB vs S15. Plots for M13V, M13I, or M13L vs S15 are essentially blank at the 0.3-Å starting level used here and are not shown. (Bottom panel) Difference distance matrices for Cα atoms of the selected residues listed, representing the wall of the cavity for residue 13. The numbers are in units of 0.1 Å. The upper panel shows the changes in going from Met to Gly in position 13; the lower panel, Met to Phe.

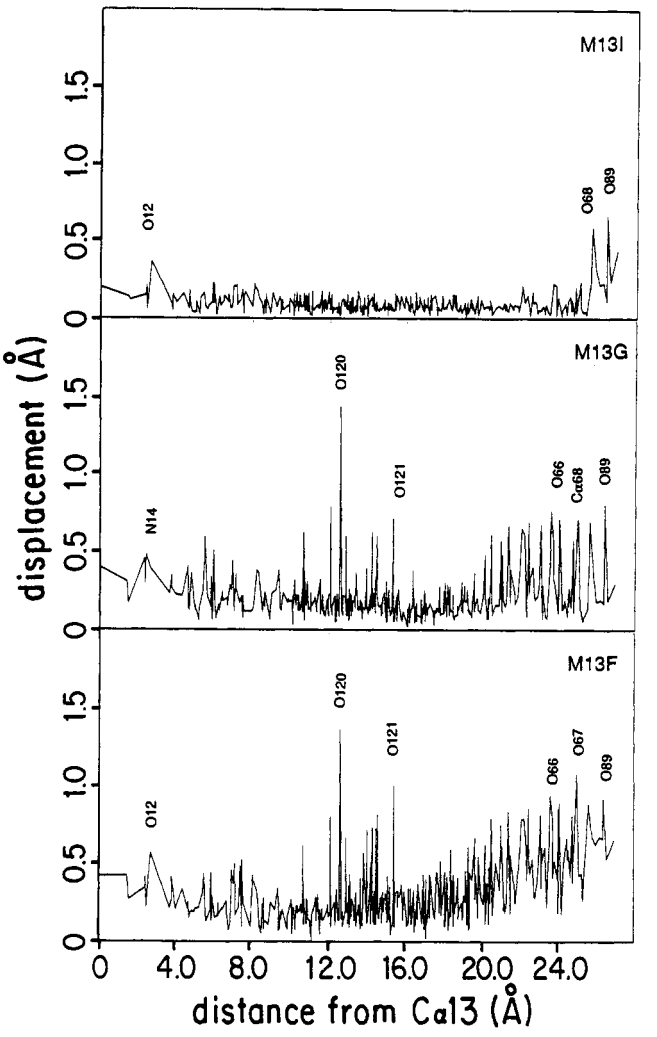


FIGURE 3: Absolute value of the difference in position of main-chain atoms in the M13F (bottom), M13G (middle), and M13I (top) complexes with respect to the S15 complex. The mutant and S15 complexes were first superposed by least squares so as to minimize the deviation between corresponding main-chain atoms. The disordered residues 1 and 21–24 and residue 13 were omitted from the superposition procedure.

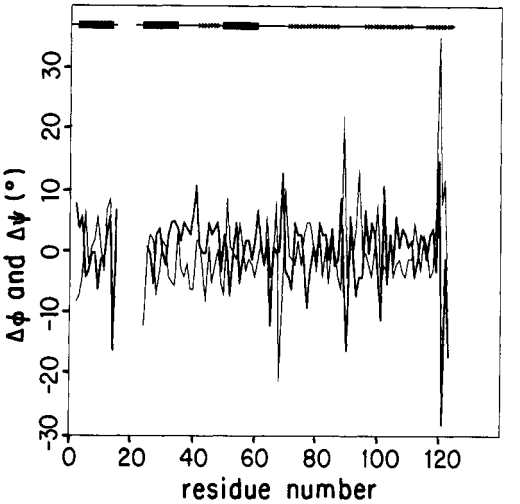


FIGURE 4: Differences in the main-chain dihedral angles ϕ (thick line) and ψ (thin line) in M13G relative to S15.

whose side chain in turn contacts the residue in position 13. However, the extent of the motion of the atoms in all of these residues is small, <0.5 Å, and no simple lever action model is particularly convincing.

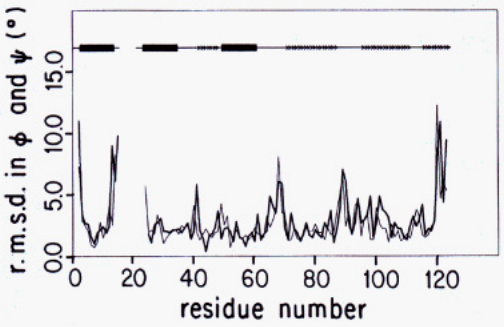


FIGURE 5: Standard deviations in ϕ (thick line) and ψ (thin line) calculated using all eight complexes.

In their study of the effect of temperature on the structure of RNase-A, Tilton et al. (1992) noted that the largest thermally induced changes occurred in the loop regions including the 66–69 loop. However, over the range from 98 to 320 °C that they surveyed, the maximum shifts for any of the loop atoms were less than 0.5 Å while the maximum shifts in the mutant series reported here are about 3 times as large.

There is little change in the position of the Cα 13 atom in any of the complexes though there is a change of 0.3–0.5 Å at Cα's 12, 14, and 15 in the M13G and M13F complexes. While the His12 motions appear to correlate with those of Phe120, the connection of the two residues is not obvious nor does His12 directly affect Asp121.

Other Parts of the Main Chain. All other main-chain shifts are smaller than those of the 66–69 loop. From the dihedral angle changes seen in Figure 4, the 88–90 loop and the C-terminal residues 120–124 are also prominent. From the rmsd in ϕ and ψ using all eight complexes (Figure 5), the largest changes occur in the two mentioned loop regions and at the chain terminii of both peptide and S-protein. However, the changes in ϕ_n and ψ_{n-1} compensate so that the Cα distance shifts are actually quite small. Nonetheless, most of these differences are real. Many of those plotted in Figure 3 are significantly greater than the error in main-chain coordinates which was estimated to range from 0.050 to 0.088 Å for the S15 complex (Kim et al., 1992).

A detailed difference distance matrix in numerical form is shown in Figure 2C. The entries are restricted to the Cα atoms of the residues surrounding position 13 and defining the surface of the cavity. For the Gly mutant most of the numbers are negative; that is, there is a contraction of the cavity compared to its size in the Met structure. The largest motion involves Leu51. For the Phe mutant the numbers are generally positive indicating an expansion, with Ser50 and Leu51 showing the biggest shifts. The maximum values for any of these changes are –0.6 and +0.7 Å. Even these small shifts produce significant changes in cavity size as discussed later.

(*c*) *Side-Chain Rearrangements.* A characteristic feature of all of the side-chain rearrangements, as commonly found, is that they are largest for atoms close to the surface of the protein as well as for those close to chain terminii. Atoms in the protein interior are constrained by packing interactions with the rest of the protein and undergo little or not rearrangement. Analysis of side-chain rearrangement is complicated by the fact that there are several disordered surface residues in each complex. Since the electron density is not clearly defined and does not suggest obvious alternate sites for these residues, the refinement procedure places them in different conformations in the various complexes. However, such conformational differences are likely to be artifactual. The principal focus in this paper is on the side chains that are in contact with or close to the site of mutation. With the caveats expressed above, all of the side chains in the complexes show little or no change in conformation with the exception of those specifically discussed below.

Rearrangments Close to the Site of the Mutation. Some 12 residues (8, 9, 12, 14, 15, 33, 46, 47, 49, 50, 51, 54) have at least one atom which is part of the surface of the cavity in which residue 13 is inserted. The surface atoms of the cavity are defined as those for which there is no other atom intervening between them and atoms of residue 13. Since the distances between these cavity surface atoms and the nearest atoms of residue 13 may vary considerably, the use of the word "contact" can be misleading. Only some of these residues undergo significant changes in position or conformation in the different mutant structures.

Listed in Table III are the dihedral angles for 7 of the 12 cavity side chains. The other residues show χ angle changes of less than 10° in all mutants. Although it is at least 7.6 Å from Met13 and not a cavity surface residue by the above definition, Phe120 is included in the table since it shows clear differences in conformation in the M13G and M13F complexes relative to the remaining complexes.

Ser15, the C-terminal residue in S-peptide, is characterized by a relatively high temperature factor, and it does not interact much with the rest of the complex. Hence, large motions of this residue are possible at relatively low energetic cost. However, it is important in helping to separate the cavity region from the solvent. In M13F the large Phe residue pushes Ser15, and its neighbor Asp14, out toward the solution (Figure 2). There is also an apparent change in the χ_1 rotamer, but

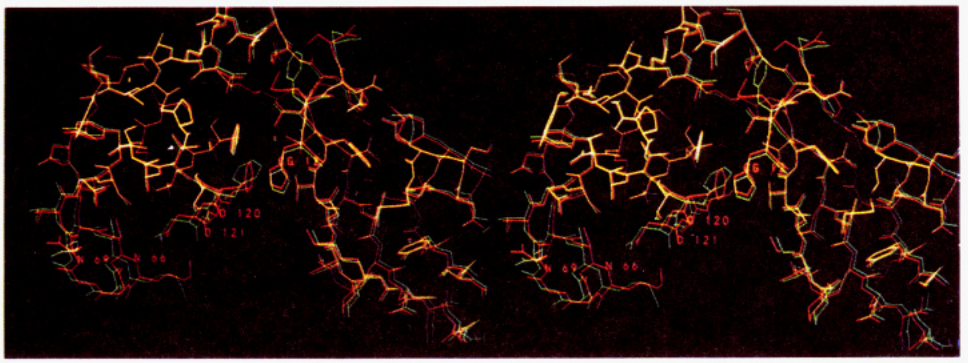


FIGURE 6: Motion of the 66–69 loop region. M13F (cyan) and S15 (red) complexes were superposed by least squares so as to minimize the deviation between corresponding main-chain atoms. The disordered residues 1 and 21–24 and residue 13 were omitted from the superposition procedure. Residue 13 as both Met and Phe occurs in the top center of the diagram, the induced Phe120 rotamer shift can be seen some distance below, and the loop region is at the bottom left.

Table III: Side-Chain Dihedral Angles of Residues near Residue 13[a]

residue	property	M13G	M13A	M13ANB	M13V	M13I	M13L	M13M	M13F
Glu9	χ_1 (−70)	−60	−58	−70	−62	−64	−67	−66	−63
	χ_2 (−177)	177	172	173	171	164	160	165	172
	χ_3 (−11)	−46	−31	−19	−15	−3	3	−11	−19
	rmsd[b] (mc)	0.17	0.12	0.14	0.13	0.10	0.09		0.21
	rmsd[b] (sc)	0.66	0.28	0.38	0.14	0.16	0.15		0.30
Gln11	χ_1 (−67)	−59	−57	−62	−57	−61	−62	−59	−59
	χ_2 (−179)	−165	−166	−167	−167	−169	−167	−170	−172
	χ_3	126	115	126	121	−48[c]	123	−51[c]	122
	rmsd (mc)	0.24	0.13	0.17	0.22	0.20	0.16		0.32
	rmsd (sc)	1.4[c]	1.4[c]	1.39[c]	1.41[c]	0.18	1.41[c]		1.42[c]
Res13	χ_1			−70	176	−61	−176	−69	177
	χ_2					−63	−163*	−57*	72
	χ_3							−53*	
Asp14	χ_1 (−169)	−165	−174	−179	178*	176*	179*	−179	179*
	χ_2 (4)	17	24	19	18	22	20	21	27
	rmsd (mc)	0.48	0.20	0.14	0.13	0.07	0.09		0.31
	rmsd (sc)	0.31	0.22	0.13	0.17	0.11	0.10		0.56
Ser15	χ_1[d]	−54*	−42*	−58	−50*	48	−57	−34*	30*
	rmsd (mc)	1.38	0.97	0.48	0.55	0.56	1.05		0.64
	rmsd (sc)	1.07	0.83	0.77	0.82	1.17	1.23		2.62
Val47	χ_1 (174)	−172*	−170*	−173*	−177	−177	−174*	−175*	−171*
	χ_1 (69)				61				58*
	rmsd (mc)	0.18	0.04	0.13	0.20	0.12	0.14		0.27
	rmsd (sc)	0.21	0.13	0.15	0.23	0.12	0.11		0.23
					1.95				1.98
Ser50	χ_1 (65)	91*	86*	72	81	72	81	71	74
	rmsd (mc)	0.25	0.21	0.24	0.23	0.15	0.16		0.46
	rmsd (sc)	0.26	0.29	0.18	0.33	0.18	0.15		0.42
Leu51	χ_1 (−165)	−96*	−174	−175	−179*	179*	179*	−175	−176*
	χ_2 (168)	+178	172	174	175	−174	−176	−172	171
	χ_1[e]	−93	−108	−99					
χ_2[e]	68	−54	−55						
	rmsd (mc)	0.43	0.33	0.32	0.25	0.15	0.13		0.41
	rmsd (sc)	2.23	0.45	0.34	0.22	0.27	0.56		0.55
		2.44	2.09	2.14					
Phe120	χ_1 (−179)	−106*	−156*	−158*	−164*	−168	−166*	−169	−108*
	χ_2 (−109)	−54*	−88*	−86*	−89*	−88*	−89*	−85*	−52*
	χ_1 (−179)	−127*							
	χ_2 (−109)	−116							
	rmsd (mc)	0.87	0.32	0.31	0.17	0.13	0.11		0.83
	rmsd (sc)	1.88	0.36	0.37	0.17	0.15	0.15		1.62
		0.73							

[a] Mean values from the rotamer library of Ponder and Richards are listed in parentheses in column 2. Dihedral angles that differ by more than 1 standard deviation from these mean values are marked with an asterisk. A residue is only listed if at least one atom is within 4.0 Å of some atoms of residue 13 in the S15 complex and if corresponding side-chain dihedrals for the residue in any two of the complexes differ by more than 10°. Hence, residues 8, 12, 33, 49, and 54, which are within 4.0 Å of residue 13, are not listed. The dihedral angles for these residues are close to the rotamer library values in all complexes. Although Phe120 is 7.6 Å from Met13, dihedral angles for this residue are listed because the conformation of this residue is clearly different in the M13G and M13F complexes. [b] Coordinates of the mutant complexes were superposed with the S15 coordinates using the main-chain atoms of all residues except 1, 13 and 21, 24. The root mean square deviations of the difference in position in the main-chain (mc) and side-chain (sc) atoms relative to the S15 complex were then calculated for each residue. [c] This apparent 180° change in χ_3 and large rmsd (sc) relative to the other complexes occur because the refinement cannot distinguish the O atom of the side-chain amide group from the N atom. [d] Rotamers with χ_1 angles close to either −60° or +60° are observed. [e] No similar rotamer in the library.

it should be noted that the OG atom is disordered. Similar changes are seen only in M13I.

Val47 was modeled in two conformations in the M13V and M13F complexes. The second conformation involves a change in χ_1 from −171° to 58°. This enables the CG1 atom to occupy some of the cavity. In the other complexes this conformation of Val47 would be destabilized by steric repulsion between the CG1 methyl group and either the side chain of residue 13 or, in the case of M13G, M13A and M13ANB, the buried water molecule.

Leu51 shows the greatest range of conformational behavior of all the cavity residues. For the structures with V, I, L, M, or F in position 13, Leu51 appeared in only one conformational region. The χ_1 and χ_2 values are reasonably close to those of the canonical trans,trans rotamer. In the M13A and M13ANB mutants Leu51 was modeled in two conformations. In both complexes one conformation is the trans,trans; the second corresponds to a rotamer not found in the proteins examined by Ponder and Richards (1987). A large displacement of the CG atom is effected by rotation about χ_1 and an approximate exchange in positions of the CD1 and CD2 atoms through a compensating rotation about χ_2. The new χ_1 introduces strain and produces an energetically unfavorable conformation (Table III).

Leu51 was also modeled in two conformations in M13G. However, in this case both rotamers are different from those found in the other complexes and have unfavorable χ_1 angles. Thus filling of the cavity in this complex occurs at the cost of some torsional strain (estimated to be approximately 2.0 kcal/mol using the dihedral angle potential in X-PLOR). The large movement of Leu51 is only possible because of the absence in M13G of the Cβ atom in position 13.

The positional changes of Leu51 are the largest of all the cavity residues. The following factors provide a qualitative explanation: constraints on interior residues due to tight packing; constraints on charged and polar surface residue motions through electrostatic and H-bonding interactions. Thus the residues near Met13 most likely to change confor-

Table IV: Voronoi Volumes (in Å^3) and Related Data Calculated from the Refined Structures of the Ribonuclease S Variants

		residue in position 13							
line	set name[a]	G	A	ANB	V	I	L	M	F
1	REFERENCE	66	92		142	169	168	171	203
2	-M	-105	-79		-29	-2	-3		+32
3	VDWVOL	50	68	87	105	124	124	119	141
4	-M	-69	-51	-32	-14	+5	+5		+22
5	CAVVOL	16	24		37	45	44	52	62
6	-M	-36	-28		-15	-7	-8		+10
7	PAKDEN	0.76	0.74		0.74	0.73	0.74	0.70	0.69
8	RESIDUE-13	84	113	135	156	174	180	176	201
9	-M	-92	-63	-41	-20	-2	+4		+25
10	CAVSURF	1358	1334	1322	1312	1300	1318	1298	1330
11	-M	+60	+36	+24	+14	+2	+20		+32
12	CAVSURF+13	1442	1447	1457	1468	1474	1498	1474	1531
13	-M	-32	-27	-17	-6	0	+24		+57
14	CORE	6617	6619	6630	6622	6651	6664	6678	6650
15	-M	-61	-59	-48	-56	-27	-14		-28
16	ΔCAV-13	+37	+24	+15	+8	-5	+19		+35

[a] Lines labeled -M list the differences of the values immediately above from the corresponding value for the wild-type complex in column M. REFERENCE (line 1): mean Voronoi volume of the various residue types as determined in a series of native protein structures and reviewed by Richards (1977). VDWVOL (line 3): volume inside the van der Waals envelope of each residue as determined by a rough numerical integration; assumed independent of conformation. CAVVOL (line 5): the cavity component of the Voronoi volume of each residue = line 1 - line 3. PAKDEN (line 7): mean packing density of these residue types = line 3/line 1. RESIDUE-13 (line 8): the Voronoi volume of residue 13 in the various complexes determined directly from the refined coordinate sets. CAVSURF (line 10): the summed Voronoi volumes of all of the atoms of residues 8, 9, 12, 14, 15, 33, 46, 47, 49, 50, 51, and 54 which have two or less of their polyhedral faces defined by SHELL positions (73 atoms of the total of 94 in these 12 residues). One or more of the atoms in each of these residues are part of the surface of the cavity surrounding residue 13. CAVSURF+13 (line 12): line 10 + line 8. CORE (line 14): the summed Voronoi volumes of all atoms *not* in the CAVSURF list or in residue 13 which have two or less of their polyhedral faces defined by SHELL positions (425 atoms of the total of 811 atoms in these 106 residues). ΔCAV-13 (line 16): the change in the volume of the cavity surrounding residue 13 with reference to the native M13 structure = line 13 - line 4.

mation in response to substitution are nonpolar surface residues and any polar ones that are not hydrogen bonded to the rest of the protein, i.e., Ser15 and Leu51. Leu51 is the only hydrophobic residue adjacent to Met13 that has nonzero accessibility. It is thus reasonable that the largest movements should occur for these two residues.

Phe120 was modeled in two conformations in the M13G complex. In the M13F complex the residue was modeled in a single conformation with χ values similar to those found in one of the M13G conformations. Both of these conformations are different from the one found in all the other structures. None of the Phe120 χ_2 values for any of the structures were close to the canonical rotamer values. For M13G and M13F the χ_1 values for the rotamer in common are severely distorted as well. The source of the compensation for the Phe120 strain energies in M13G and M13F is not obvious.

The Cavity and Position 13. The largest main-chain rearrangements close to the site of substitution occur for the M13F and M13G mutants. Phe is both larger than Met and has a very different shape. Consequently, an appreciable increase in cavity volume appears in the M13F complex in spite of the fact that F is larger than M. In M13G the contraction of the cavity due to the combined main- and side-chain shifts only goes part way to filling the extra space.

In M13L the cavity appears to expand relative to the Met or Ile complexes (see discussion of Table V below). This must reflect the effect of a difference in residue shape as the volumes and surface areas of these three residue types are essentially identical. In the other mutants the shifts are even smaller. β-Branched residues (M13V and M13I) are accommodated without appreciable main-chain shifts. These two peptides, in fact, form slightly tighter complexes than S15 itself (Varadarajan et al., 1992).

In all complexes residue 13 is observed in a single conformation that is commonly found in other protein structures. The main- and side-chain atoms for this residue have lower than average *B*-values and are in low-energy rotamer conformations that are commonly found in other protein structures (Ponder & Richards 1987). This statement is true for the cavity residues as well except for the cases noted above. In the M13ANB complex the ANB side chain appears to be in a single conformation with a low *B*-factor in which the CG methyl group occupies roughly the same position as the CG2 atom of Val13 in the M13V complex ($\chi_1 = -70°$). If no buried water (see below) is present, it is also sterically possible for the CG of ANB to occupy the position ($\chi_1 = 180°$) that is observed for the CG1 carbons of residue 13 in the M13V and M13I complexes. Occupation of the CG2 position is probably due to steric repulsion with the buried water molecule as well as the greater solvent accessibility (6 Å^2 vs 0) of the CG1 position.

(*d*) *Cavity Volumes and Distributions.* Various Voronoi volumes of sets of atoms are listed in Table IV. The mean Voronoi volumes of the various amino acids in known protein structures have been described previously (Richards, 1977). Lines 1–7 list these and related properties for seven of the eight residues inserted at position 13. No data currently exist for ANB, a nonnatural amino acid. The mean packing densities for these residues (line 7) are very similar, with glycine tending to be slightly higher and methionine and phenylalanine slightly lower than the others.

On line 8 the Voronoi volumes of the various residues in position 13, calculated from the structures reported in this study are listed. While the values for the larger residues are similar to the averages listed on line 1, the values for the smaller residues are distinctly larger, reflecting that appearance of larger than normal cavities. The summed volumes for the residues (CAVSURF) making up the surface of the space containing residue 13 (line 10) show an increase in volume in all cases. The identity of the atoms summed for these totals is the same in all of the eight structures. Since the volumes of their van der Waals envelopes do not change, the only explanation for the volume changes is a change in the cavity structure.

When the volumes for the CAVSURF residues (line 10) and residue 13 (line 8) are summed, the total is a number

proportional to the volume for this whole region or packing unit (line 12). The atom volumes and cavity volume are accounted for. If the only changes in structure between the complexes were the number of atoms in the side chain of residue 13, all other atoms remaining fixed, these totals would all be the same. As can be seen in line 12, this is not the case. The total volume tends to decrease slightly for the smaller residues and increase for the larger one (Phe). The compensation is not perfect, however (see line 16). There is a larger cavity space associated with the small residues and with Phe. Phe is too big to fit in the normal Met "hole", and expansion is necessary to avoid steric overlap. The residue shapes are such that extra cavity space appears. For the smaller residues the "collapse" of the CAVSURF residues to fill the cavity cannot be complete, again presumably for steric reasons. Views of the cavities in the M13G and M13F complexes are shown in Figure 7.

The interior part of the protein more remote from residue 13 than the CAVSURF residues, defined as the CORE, also changes its volume slightly between the different complexes. It tends to decrease its volume for both the large and small substituents at position 13. There is not perfect compensation, but there appears to be a tendency to maintain a constant total cavity volume for the whole protein. The actual volume changes in the CORE are all less than 1%. The very small positional shifts are not visible in the difference distance matrices (Figure 2) at the contour levels drawn.

To estimate cavity shape and size, both the volumes and the positions of the cavity components of the Voronoi volumes of the CAVSURF residues are required, in addition to comparable data for residue 13 itself. A procedure to use atom-pair cavity components for this purpose is under development by Richards and Hodel (unpublished results) but is not yet adequately tested. Thus, the actual size of the cavity around Met13 cannot be obtained from the numbers given in Table IV. However, it is certainly greater than 60 Å^3. This volume is distributed all around the side chain and particularly reflected in the Voronoi volumes of CE and SD. However, at no one point is there space to insert a sphere the size of a water molecule. The extra space required for Phe13 may be similarly described. Only with the other residues is the cavity volume sufficiently collected in one region to sterically permit the addition of other atoms. As a scale for the numbers in Table IV, a methylene group has an effective volume of about 26 Å^3. Under close packing conditions with 12 nearest neighbors, a water molecule would occupy about 15.5 Å^3 when modeled as a 1.4-Å radius sphere. In its tetrahedral arrangement in bulk water it occupies about 30 Å^3. The appropriate number for isolated molecules inside a protein is not entirely clear. See Savage and Finney (1986) for a discussion of this problem.

(*e*) *Internal Ordered Water Molecules.* Replacement of Met13 in the S15 complex with smaller side chains results in the formation of cavities. These cavities are partially filled by the shifts of Leu51. In the case of the M13G and M13A complexes an ordered water molecule in the cavity was identified by the presence of a peak greater than 4.5σ in the $F_o - F_c$ electron density map. A water oxygen placed at this position refined with a *B*-factor of 26 for the M13G complex and 38 for the M13A complex and clear electron density was observable at 0.8σ (the contouring level used to examine all data sets using FRODO) in the $2F_o - F_c$ maps. In the case of the M13ANB complex there is a peak of 2.5σ close to the same position in the $F_o - F_c$ map. However, a water molecule placed at this position refines with a *B*-factor of 60–80 (the precise value changed from one refinement run to another), and only a very small peak is observable in the $2F_o - F_c$ maps at 0.8σ.

Table V: Interactions of the Buried Water with Its Surroundings

(a) Repulsive VDW Interactions[a]

atom	complex	distance (Å)	VDW energy (kcal/mol)
Cα 13	M13G	3.4	0.25
	M13A	3.5	0.21
	M13ANB	3.3	0.77
Cβ 13	M13A	3.4	0.10
	M13ANB	3.1	1.4
Cγ 13	M13ANB	3.4	0.20

(b) Hydrogen Bonding[b]

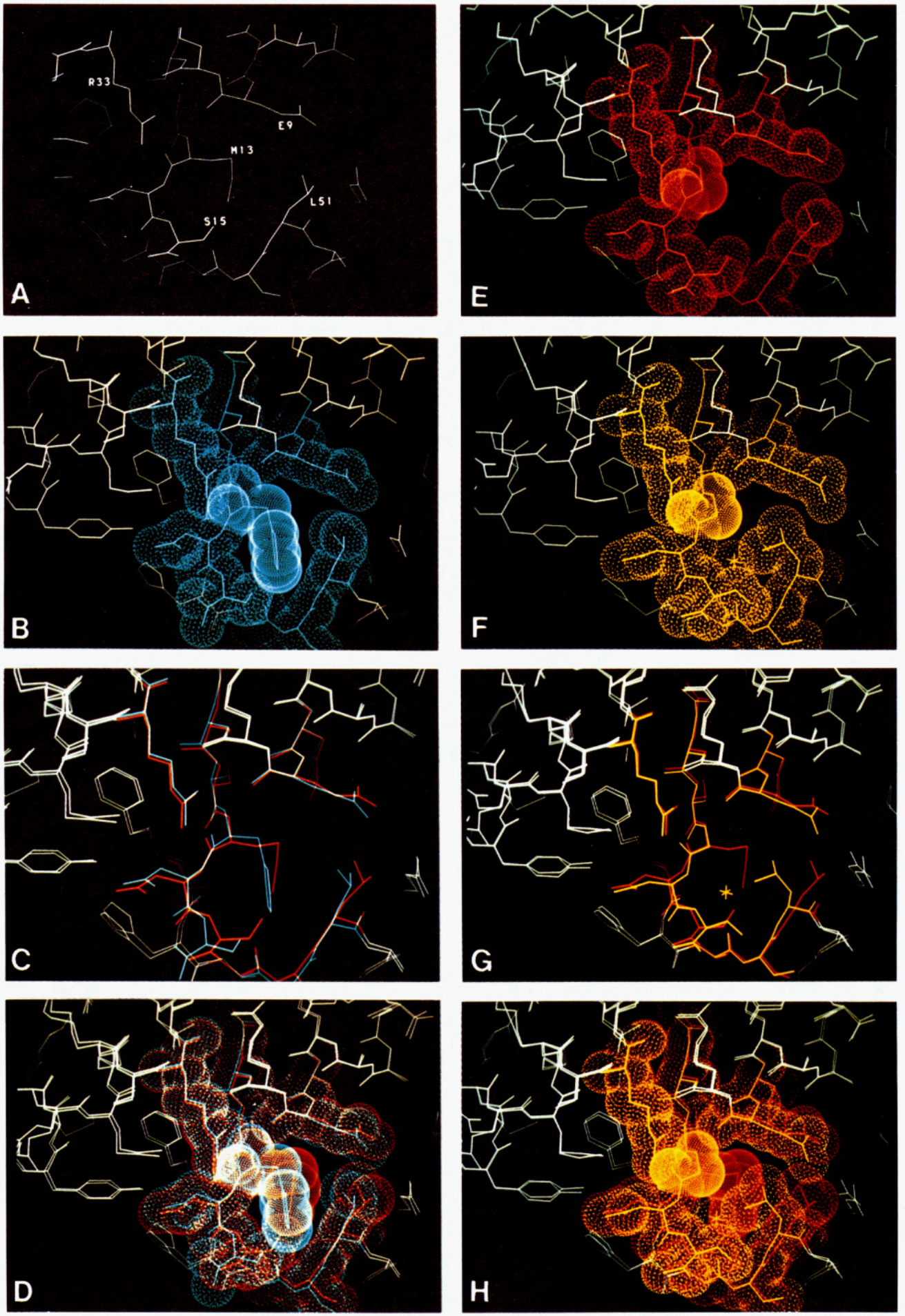
atom	complex	distance (Å)	angle (deg)
N 14	M13G	3.2	149
	M13A	3.6	129
	M13ANB	3.3	126
Oγ 15[c]	M13G	3.8	130
	M13A	3.4	141
	M13ANB	4.1	129
O 47	M13G	3.1	134
	M13A	3.2	128
	M13ANB	2.8	130
O 49	M13G	2.8	149
	M13A	2.8	141
	M13ANB	2.8	141

[a] VDW energies between atoms listed in the first column and the buried water oxygen atom were calculated using X-PLOR. [b] Listed in column 1 are hydrogen bond donors/acceptors that are within 3.5 Å of the buried water oxygen atom in at least one complex. Listed in column 3 is the distance between the atom in column 1 and the oxygen (O_w) of the buried water molecule. If this atom is N/Oγ, then the angle refers to the N/Oγ–H–O_w angle. If the atom is O, this angle is the C–O–O_w angle. For main-chain–water hydrogen bonds in proteins, these angles are 156 ± 15° and 129 ± 16°, respectively (Baker & Hubbard, 1984). [c] This atom has a high temperature factor and is not clearly defined in the electron density map. Hence, differences in hydrogen-bonding geometry involving this atom are not significant.

The presence of these waters was examined by generating maps after simulated annealing with the water molecule and atoms within a 5-Å sphere surrounding it omitted from the model. In all three cases a peak was visible at 1.5σ in the $F_o - F_c$ omit map, but the peak volume decreased in the order M13G > M13A > M13ANB. No peak in the vicinity of residue 13 is observed in the $F_o - F_c$ maps of the remaining complexes, and it is also sterically impossible to fit in a water molecule at the above position in the remaining complexes.

Since the cavity volume decreases on going from M13G to M13ANB, the apparent increase in *B*-factor of the bound water molecule is likely to be due to a decrease in occupancy. If the *B*-factors for the M13A and M13ANB water oxygens are constrained to have the same value (26) as for the M13G protein and the occupancies of these atoms refined instead, then occupancy values of 0.98, 0.73, and 0.54 are obtained for the water oxygens of M13G, M13A, and M13ANB, respectively. Table V describes the steric and hydrogen-bonding interactions of the buried water with its surroundings in the three complexes.

In an attempt to predict the position of the bound water observed in the M13G, M13A, and M13ANB complexes, a grid was constructed within the cavity as described in Materials and Methods section d. We selected those grid points in the cavity that were within 3.5 Å of at least one N and one O atom of the protein and then calculated the angles that the oxygen atom (O_w) of a hypothetical water molecule placed at these grid positions would make with such nearby N and O atoms within the protein. A definition of these angles is provided in Table V, footnote *b*. Only those waters which could make


R33
E9
M13
L51
S15
A
E
B
F
C
G
D
H

angles close to the values listed in Baker and Hubbard (1984) were considered acceptable. This grid procedure was successfully able to predict the position of the bound water molecule within the cavity. The difference between the allowed grid positions and the water position experimentally observed in the M13G complex was 0.3 Å.

The buried water is stabilized in M13G relative to M13A and M13ANB because of changes in the main-chain dihedral angles of residues 12 and 13 which enable N14 to form a better hydrogen bond to the buried water oxygen (Table Vb). In M13A and M13ANB these dihedral angles are inaccessible because the Cβ 13 would make bad steric contacts with O 9 (2.87 Å) and N 14 (3.19 Å). The water molecule in the M13ANB complex is destabilized relative to the M13A complex because of an unfavorable van der Waals interaction with Cβ 13 (Table Va).

The observed water molecule makes three hydrogen bonds to the protein. At least in the Gly mutant there is certainly cavity space for an additional molecule. However, this second molecule could only make a single bond to the first water since there are no appropriately placed protein donors or acceptors. Obviously that is not energetically favorable. In their study of cavity-generating mutations in T4-lysozyme, Eriksson et al. (1992) found no solvent in any of six different cavities of varying size. Presumably, insufficient protein H-bonding groups were available in any of these structures. We have not yet found a description of the minimum size of a water cluster which would be intrinsically stable in a cavity with a hydrocarbon wall where the "reactants" are bulk water and a large hydrocarbon bath with the empty cavity. Clearly, the five to six waters which the largest cavity in T4-lysozyme might have contained are not enough.

(f) Search for Internal Disordered Water. In an attempt to determine whether additional amounts of disordered water were present in the cavities surrounding residue 13, we used the following procedure. The coordinates of cavity centers were determined as described in the Materials and Methods section for all internal cavities within the protein. Two subsets of these coordinates were selected for further examination. The first set of approximately 1600 coordinates consisted of those internal cavity centers that were not adjacent to any atom belonging to residue 13 or the cavity boundary residues 8, 9, 12, 14, 15, 46, 47, 49, 50, 51, and 54. All of this first set of cavities are too small to contain any solvent. The second set, which contained between 16 and 46 members, consisted of those cavities which lay between an atom of residue 13 and an atom of one of the residues listed above. $F_o - F_c$ difference electron density maps were calculated for each complex using all reflections of resolution higher than 8 Å. Each map consisted of the difference electron density calculated at a series of grid points spaced approximately 0.6 Å (one-third of the high-resolution limit) apart. Maps were scaled so that the overall rmsd of the difference electron density for all grid points in the map is 1.0. The difference electron density at each cavity center was determined from the map by approximating it to the density at the grid point closest to the cavity center.

For the M13G, M13A, and M13ANB complexes two difference maps were calculated. In the first map the model included the buried water molecule while in the second map this molecule was omitted.

The average rmsd of the difference electron density for both sets of cavity coordinates was determined for each mutant. The average values for the two sets were identical within the rmsd (approximately 1.0) in all cases. The only peaks greater than 2.5σ were observed in the M13G, M13A, and M13ANB maps, where the internal water had been omitted from the model. These peaks ranged from 3.0σ to 5.2σ and occurred at cavity positions very close to the position of the internal water molecule.

In order to sample the cavity region more extensively, a set of grid points was generated as described in Materials and Methods section d. The origin of the grid was placed at the position occupied by the SD atom of Met13 in the S15 structure. The number of grid points where it was sterically possible to put in a 1.25-Å radius "miniwater" water molecule ranged from 125 for the M13G cavity to 0 for the wild-type complex. The difference density at these grid points was calculated as described above. As in the case of the cavity centers, the only grid points with difference densities greater than 2.5σ were those close to the internal water in M13G, M13A, and M13ANB.

Difference densities were also calculated for a series of grid positions in a region of the unit cell located in bulk solvent. The average values calculated for the bulk solvent region were indistinguishable within the rmsd from the average values calculated for grid positions within the cavity. It should be noted that no bulk solvent had been included in the model used to calculate the $F_o - F_c$ map.

The fact that the average electron density for the bulk solvent region was indistinguishable from the average within the cavity suggested that it would not be possible to state whether or not disordered water was present within the cavity. However, the extent to which such internally bound water could be disordered will depend on the cavity volume accessible to additional solvent molecules. Approximate estimates of the volumes of cavities that could be occupied by the centers of solvent molecules were made by visual inspection of the allowed grid positions using FRODO. For the M13G, M13A, and M13ANB mutants the ordered water was now included in the protein model before additional allowed grid positions were calculated. The largest set of contiguous grid positions was found in the M13G complex and occupied a volume of approximately 3.4 $Å^3$. Such a volume could not be occupied by more than a single water molecule. Furthermore, we expect that a water molecule whose center was restricted to such a small volume would be clearly visible in the difference electron density map as was the case for the single ordered water observed in M13G,

FIGURE 7: (A) Stick diagram for orientation and identification of selected residues. This is a thick section where the bottom of the cavity enclosing residue 13 has been clipped off. All other panels have the same viewing direction and orientation. (B) Variant M13F. Residue 13 is shown with a high dot density surface; the surrounding residues are shown with a low dot density van der Waals surface. (C) Variant M13F (blue) and wild-type S15M (red) overlayed, showing the approximate atomic shifts. The shifts correspond to a small expansion of the cavity to provide room for Phe13. Residues not forming part of the cavity wall are shown in white. (D) Same as (C) with dot surfaces shown for both structures. (E) Hypothetical Gly mutant. The wild-type structure is shown with the side chain of Met13 removed back to the α-carbon atom to show the appearance of the cavity if there were no change in structure on mutation at this position. (F) Actual M13G structure shown with dot surfaces. The change in the Leu51 position and the water molecule at the + position are an attempt to fill the cavity. (G) Overlay of M13G (yellow) and S15M (red). (The red structure is identical to that shown in the left panel.) The shifts are small but generally in the opposite direction to M13F and are tending to collapse the cavity. The rotamer change in L51 is clearly visible. (H) Same as (G) with dot surfaces shown for both structures. M13 can be seen under L51 and shows some of the cavity *not* filled by either L51 or the water molecules.

M13A, and M13ANB. The fact that no additional peaks greater than 2.5σ are observed thus strongly suggests that no more water, ordered or otherwise, is present within the cavity.

(*g*) *Other Located Water Molecules.* The total number of water molecules included in the models for the mutant complexes ranges from 56 (M13A) to 63 (M13F). Of these, 45 are found in similar positions in all seven mutant structures. The overall rmsd in the positions of these 45 waters in the different complexes is 0.33 Å, and the average *B*-factor is 38 Å^2. The average *B*-factor of the waters not common to all structures is 55 Å^2. Most of this latter group occur in regions of poorly defined density. The most important difference in water structure is the internal ordered water described in section d above.

(*h*) *Relation between Thermodynamic and Structural Changes.* We have previously described measurements of the thermodynamic binding parameters (ΔH, $\Delta G°$, ΔC_p) of the various S-peptide analogs to S-protein in the temperature range 5–25 °C (Connelly et al., 1990; Varadarajan et al., 1992). We showed that the binding parameters were not simply correlated with residue size or hydrophobicity and argued that the observed free energy differences could be broken up into three components: (a) changes in the hydration of the free peptides relative to the S15 peptide, (b) changes in the hydration of the complexes relative to the S15 complex, and (c) changes in the packing interactions in the various complexes. With the structures now determined, we made preliminary attempts to calculate these components, and these efforts as well as the problems encountered are described below.

There is no experimental way of directly measuring the contributions due to (a). However, linear free energy correlations between free energies of transfer and nonpolar accessible area have been observed for a variety of nonpolar molecules (Chothia, 1974; Spolar et al., 1989). It may therefore be reasonable to assume that term (a) will be linearly proportional to the difference in the total nonpolar accessible area of the free peptides relative to the S15 peptide. These values have been estimated previously (Varadarajan et al., 1992).

The only significant changes in the water structure of the complexes are the internal waters found in the M13G, M13A, and M13ANB complexes. Estimation of the contribution of this internal water to the binding thermodynamics is difficult and thus complicates estimation of the magnitude of term b for these three complexes.

Since the substituents at residue 13 are all hydrophobic amino acids, we initially attempted to calculate term c by assuming that it was proportional to the difference in the total van der Waals energies of the complexes. A similar approach was used by Lee and Levitt (1991) to predict the stabilities of several mutants of the λ repressor. However, there are two major problems with such an approach. First, the total van der Waals energies will tend to scale with the number of atoms in the complex and so the mutants with larger substituents will generally have a more negative van der Waals energy than those with smaller substituents regardless of whether or not cavities are formed. Second, there are several disordered surface residues in each complex. The crystallographic refinement procedure places such residues in different conformations in the different complexes. Such differences also introduce artifactual differences in the van der Waals energies.

Although there are large and often compensating changes in $\Delta\Delta H$ (25 °C) and $\Delta\Delta S°$ (25 °C) (Varadarajan et al., 1992) in several mutants, we have not attempted to interpret these. Here the $\Delta\Delta S°$ refer to values of the parameters relative to the corresponding values for the S15 complex. This compensation is particularly noticeable for the M13I, M13L, and M13V peptides which have $\Delta\Delta G°$ (25 °C) close to 0 but a $\Delta\Delta H$ (25 °C) of 5–6 kcal/mol. The S15 peptide thus has a significantly more negative ΔH (25 °C). There are no obvious differences in the crystal structures of the other three complexes which would account for the large $\Delta\Delta H$. In order to determine whether this difference is due to the presence of the S atom in Met, we are examining the thermodynamics and structure of a variant in which Met13 has been replaced by norleucine. An added complication in the interpretation of $\Delta\Delta H$'s and $\Delta\Delta S°$'s is that these differences are temperature dependent. This is a consequence of the fact that the ΔC_p's for the various peptide binding reactions are different (Varadarajan et al., 1992). Thus a complete understanding of the binding thermodynamics will not be possible until the determinants of the differences in ΔC_p have been identified. This remains to be done.

ACKNOWLEDGMENT

The authors express their thanks to Drs. E. Kim and H. W. Wyckoff for many helpful discussions during the data collection, reduction, and refinement of the seven mutant structures. Thanks are due to A. Perlo and D. Keller and the Center for Structural Biology for the facilities, both X-ray and computational, used in this work. A. Brünger has contributed to this work through provision of, and advice on, the program X-PLOR. Our thanks are also due to J. L. Mouning, T. Mouning, and A. Johnson for technical and secretarial assistance in the preparation of the manuscript. The authors also thank P. R. Connelly for all of his work on the thermodynamic studies and for discussions on the current attempts to correlate that data with the structural observations.

REFERENCES

Alard, P., & Wodak, S. J. (1991) *J. Comput. Chem. 12*, 918–922.

Baker, E. N., & Hubbard, R. E. (1984) *Prog. Biophys. Mol. Biol. 44*, 97–179.

Behe, M. J., Lattman, E. E., & Rose, G. D. (1991) *Proc. Natl. Acad. Sci. U.S.A. 88*, 4195–4199.

Brünger, A. T. (1990) *X-PLOR, Version 2.1*, Yale University, New Haven, CT.

Chothia, C. (1974) *Nature 248*, 338–339.

Connolly, M. L. (1991) *Comput. Chem. 15*, 37–45.

Connelly, P. R., Varadarajan, R., Sturtevant, J. M., & Richards, F. M. (1990) *Biochemistry 29*, 6108–6114.

Eriksson, A. E., Baase, W. A., Zhang, X.-J., Heinz, D. W., Blaber, M., Baldwin, E. P., & Matthews, B. W. (1992) *Science 255*, 178–183.

Gellatly, B. J., & Finney, J. L. (1982) *J. Mol. Biol. 161*, 305.

Jones, T. A. (1982) in *Computational Crystallography* (Sayre, D., Ed.) p 303, Clarendon Press, Oxford.

Kellis, J. T., Nyberg, K., & Fersht, A. R. (1989) *Biochemistry 28*, 4914–4922.

Kim, E. E., Varadarajan, R., Wyckoff, H. W., & Richards, F. M. (1992) *Biochemistry* (preceding paper in this issue).

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

Lee, C., & Levitt, M. (1991) *Nature 352*, 448–451.

Lim, W. A., & Sauer, R. T. (1991) *J. Mol. Biol. 219*, 359–376.

Nishikawa, K., Ooi, T., Ysogai, Y., & Saito, N. (1972) *J. Phys. Soc. Jpn. 32*, 1331–1337.

Ponder, J. W., & Richards, F. M. (1987) *J. Mol. Biol. 193*, 775–791.

Rashin, A. A., Iofin, M., & Honig, B. (1986) *Biochemistry 25*, 3619–3625.

Richards, F. M. (1977) *Annu. Rev. Biophys. Bioeng. 6*, 151–176.
Richards, F. M. (1985) *Methods Enzymol. 115*, 145–154.
Richards, F. M., & Vithayathil, P. J. (1959) *J. Biol. Chem. 234*, 1459–1465.
Richards, F. M., & Wyckoff, H. W. (1971) *Enzymes 4*, 647–806.
Richards, F. M., & Wyckoff, H. W. (1973) in *Atlas of Molecular Structures in Biology* (Phillips, D. C., & Richards, F. M., Eds.) Vol. I, Clarendon Press, Oxford.
Richards, F. M., & Kundrot, C. E. (1988) *Proteins 3*, 71–84.
Richards, F. M., Wyckoff, H. W., Carlson, W. A., Allewell, N. M., Lee, B., & Mitsui, Y. (1971) *Cold Spring Harbor Symp. Quant. Biol. 36*, 35–43.
Savage, H. F. J., & Finney, J. L. (1986) *Nature 322*, 717–720.
Shortle, D., Stites, W. E., & Meeker, A. K .(1990) *Biochemistry 29*, 8033–8041.
Spolar, R. S., Ha, J.-H., & Record, M. T. Jr. (1989) *Proc. Natl. Acad. Sci. U.S.A. 86*, 8382–8385.
Teeter, M. M. (1984) *Proc. Natl. Acad. Sci. U.S.A. 81*, 6014–6018.
Tilton, R. F., Jr., Dewan, J. C., & Petsko, G. A. (1992) *Biochemistry 31*, 2469–2481.
Varadarajan, R., Richards, F. M., & Connelly, P. R. (1990) *Curr. Sci. 59*, 819–824.
Varadarajan, R., Connelly, P. R., Sturtevant, J. M., & Richards, F. M. (1992) *Biochemistry 31*, 1421–1426.
Wyckoff, H. W., Tsernoglou, D., Hanson, A. W., Knox, J. R., Lee, B., & Richards, F. M. (1970) *J. Biol. Chem. 245*, 305–328.

Registry No. Gly, 56-40-6; Ala, 56-41-7; Val, 72-18-4; Ile, 73-32-5; Leu, 61-90-5; Phe, 63-91-2; ANB, 80-60-4; Met, 63-68-3; RNase S, 9001-99-4.