Stabilization of *Escherichia coli* Ribonuclease HI by Cavity-Filling Mutations within a Hydrophobic Core†

Kohki Ishikawa, Haruki Nakamura, Kosuke Morikawa, and Shigenori Kanaya*

Protein Engineering Research Institute, 6-2-3 Furuedai, Suita, Osaka 565, Japan

Received March 12, 1993; Revised Manuscript Received April 9, 1993

ABSTRACT: The crystal structure of *Escherichia coli* ribonuclease HI has a cavity near Val-74 within the protein core. In order to fill the cavity space, we constructed two mutant proteins, V74L and V74I, in which Val-74 was replaced with either Leu or Ile, respectively. The mutant proteins are stabilized, as revealed by a 2.1–3.7 °C increase in the T_m values, as compared to the wild-type protein at pH values of 3.0 and 5.5. The mutant protein V74A, in which Val-74 is replaced with Ala, was also constructed to analyze the reverse effect. The stability of V74A decreases by 7.6 °C at pH 3.0 and 12.7 °C at pH 5.5 in T_m as compared to those values for the wild-type protein. None of the three mutations significantly affect the enzymatic activity. The crystal structures of V74L and V74I, determined at 1.8-Å resolution, are almost identical to that of the wild-type protein, except for the mutation site. In the two mutant proteins, calculation by the Voronoi procedure shows that the cavity volumes around the individual mutation sites are remarkably reduced as compared to that in the wild-type protein. These results indicate that the introduction of a methylene group into the cavity, without causing steric clash, contributes to an increase in the hydrophobic interaction within the protein core and thereby enhances protein stability. We also discuss the role of the Leu side chain, which can assume many different local conformations on a helix without sacrificing thermostability.

In folded proteins, the side chains of hydrophobic amino acid residues are generally buried inside the protein molecules, are packed efficiently with one another, and contribute to protein stability through hydrophobic interactions (Dill, 1990; Sharp, 1991). The role of such hydrophobic interactions in stabilizing folded proteins has been extensively studied by site-directed mutagenesis of T4 lysozyme (Matsumura *et al.*, 1988, 1989; Daopin *et al.*, 1991; Hurley *et al.*, 1992; Eriksson *et al.*, 1992), the α-subunit of tryptophan synthase (Yutani *et al.*, 1987), barnase (Kellis *et al.*, 1988, 1989), λ-repressor protein (Lim & Sauer, 1991), gene V protein from bacteriophage f1 (Sandberg & Terwilliger, 1989, 1991), and staphylococcal nuclease (Shortle *et al.*, 1990). On the basis of the results from these studies, Pace (1992) has made the empirical measure that a single methylene group buried contributes 1.3 ± 0.5 kcal/mol to the protein stability.

Although side chains are efficiently packed in folded proteins (Sneddon & Tobias, 1992), some cavities normally exist even in hydrophobic cores (Connolly, 1986). Therefore, it has been anticipated that the replacement of the hydrophobic residues facing the cavities with bulkier residues could enhance the protein stability if the replaced side chains could fill such cavities. However, few attempts have been made to increase protein stability by cavity-filling mutations. Karpusas *et al.* (1989) found two large cavities in T4 lysozyme and designed two mutations, Leu-133 → Phe and Ala-129 → Val, to fill them. Eijsink *et al.* (1992) introduced various mutations to fill cavities in *Bacillus stearothermophilus* neutral protease. None of these mutations resulted in a significant increase in the protein stability. The determination of the crystal structures of the mutant T4 lysozyme proteins (Karpusas *et al.*, 1989) revealed that the hydrophobic effects gained by the burial of the bulkier side chains might be cancelled by the strains observed in the nonoptimal dihedral angle χ1 and the unfavorable van der Waals contacts. In contrast, Mendel *et al.* (1992) have shown that the replacement of Leu-133 buried in T4 lysozyme with unnatural amino acid residues with bulkier side chains can increase the protein stability. More information is necessary to evaluate the effectiveness of cavity-filling mutations as an established strategy to increase protein stability.

Escherichia coli RNase HI has been used in our laboratory as a model protein to investigate the mechanism of protein stabilization (Kimura *et al.*, 1992a–c; Ishikawa *et al.*, 1993a). This protein is suitable for this purpose for the following reasons: (1) the crystal structure has been determined at 1.48-Å resolution (Katayanagi *et al.*, 1990, 1992) and at 2.0-Å resolution (Yang *et al.*, 1990), (2) the availability of the structural gene and the establishment of an overproduction system allow us to easily construct mutant proteins (Kanaya *et al.*, 1993), and (3) this protein reversibly unfolds in a single cooperative manner (Kanaya *et al.*, 1991).

Previously, we constructed a thermostabilized mutant *E. coli* RNase HI by replacing Ser-68 with Val to fill a cavity near the protein surface (Kimura *et al.*, 1992a). However, the hydroxyl group of Ser-68 in the wild-type protein forms a hydrogen bond with the amide group of Asn-130. In addition, Ser-68 is not completely shielded from the solvent. Therefore, the question remains as to whether an increase in protein stability by the Ser-68 → Val mutation is due to the embedding of the methyl group into the cavity or to the rearrangement of the hydrogen bonds.

In the current study, we have replaced the nonpolar residue Val-74, which faces a cavity in the hydrophobic core, by other nonpolar residues to examine the effects of the mutations on the stability and the structure of the protein.

EXPERIMENTAL PROCEDURES

Materials. Restriction enzymes and modifying enzymes were from Takara Shuzo Co., Ltd. Ultrapure-grade guanidine

† Dedicated to the memory of the late Dr. Tatsuo Miyazawa, Director, Protein Engineering Research Institute.

* Author to whom correspondence should be addressed.

0006-2960/93/0432-6171$04.00/0

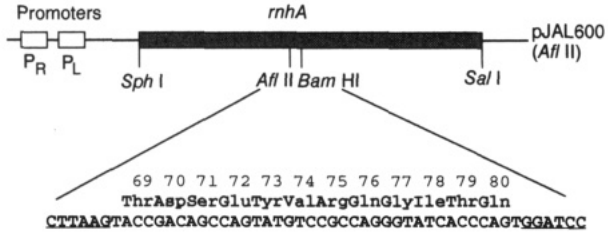

FIGURE 1: Structure of plasmid pJAL600(*Afl*II). In the plasmid, the structural gene of *E. coli* RNase HI (*rnhA*) is under the control of the bacteriophage λ promoters P_R and P_L. The nucleotide sequences and the amino acid sequences around the region where the mutagenized cassette was introduced are shown. Numbers represent the positions of amino acids in the primary structure of *E. coli* RNase HI.

hydrochloride (GdnHCl) was from Schwarz/Mann. Other chemicals were of reagent grade.

Cells and Plasmids. Competent cells of *E. coli* HB101 [F⁻, *hsdS20* (rB⁻, mB⁻), *recA13*, *ara13*, *proA2*, *lacY1*, *galK2*, *rspL20* (Sm^r), *xyl-5*, *mtl-1*, *supE44*, λ⁻] were obtained from Takara Shuzo Co., Ltd. For the overproduction of *E. coli* RNase HI, plasmid pJAL600, in which the *rnhA* gene is controlled by the bacteriophage λ promoters P_R and P_L, was used (Kanaya *et al.*, 1993). *E. coli* HB101 transformants with pJAL600 derivatives were grown in LB medium containing 100 mg/L ampicillin.

Site-Directed Mutagenesis. An *Afl*II site was introduced into the *rnhA* gene by changing the ATT codon of Ile-66 to ATC and the TTG codon of Leu-67 to TTA. The construction of this silent mutation was carried out as described previously (Kanaya *et al.*, 1993). The sequence of the 5′-mutagenic primer (complementary to the 3′-mutagenic primer) is 5′-CATTGCGAAGTCATcTTaAGTACCGACAGC-3′, where the lowercase letters show the locations of the mismatches and the underline shows the *Afl*II site. The resultant plasmid pJAL600(*Afl*II) contains unique *Afl*II and *Bam*HI sites encompassing the sequences encoding amino acid residues 67–82 (see Figure 1). Mutations of Val-74 to Ala, Leu, and Ile were introduced by replacing the 44-base-pair *Afl*II–*Bam*HI fragment of pJAL600(*Afl*II) with chemically synthesized fragments, in which the GTC codon of Val-74 was changed to GCG, CTG, or ATC, respectively. All oligodeoxyribonucleotides used in this experiment were synthesized with an Applied Biosystems Model 380A automatic synthesizer.

Overproduction and Purification. An overproducing strain for each mutant protein was constructed by transforming *E. coli* HB101 with each of the pJAL600 derivatives. The induction and the purification of the mutant proteins were described previously (Kanaya *et al.*, 1993). Usually 5–10 mg of the purified proteins was obtained from a 200-mL culture. Each mutation was confirmed by determining the amino acid sequence of the peptide LEP-4, which covers the entire region replaced by the cassette mutagenesis (Kanaya *et al.*, 1990). Amino acid sequence analysis was carried out with a gas-phase automated sequencer (Applied Biosystems Model 477A) equipped with an on-line HPLC (Applied Biosystems Model 120A). The mutant proteins in which Val-74 is replaced by either Ala, Leu, or Ile are designated as V74A, V74L, and V74I, respectively.

Determination of Kinetic Parameters. The kinetic parameters were determined as described previously (Kanaya *et al.*, 1991), using a ^{3}H-labeled M13 DNA/RNA hybrid as a substrate, except that the enzymatic activity was measured at 30 °C instead of 37 °C. One unit is defined as the amount of enzyme producing 1 μmol (RNA nucleotide phosphate) of acid-soluble material per minute. The concentrations of the mutant proteins were determined by UV absorption assuming that they have the same absorption coefficient as that of the wild-type protein ($A_{280}^{0.1\%}$ = 2.02) (Kanaya *et al.*, 1990).

Table I: Unit Cell Dimensions and Refinement Statistics[a]

protein	V74L	V74I
cell dimensions		
a (Å)	44.15	44.06
b (Å)	86.93	86.84
c (Å)	35.46	35.55
resolution (Å)	6–1.8	6–1.8
no. of reflns	10201	9916
R-factor	0.183	0.188
Δ bond (Å)	0.015	0.015
Δ angle (Å)	0.035	0.034

[a] Space group is $P2_12_12_1$. *R*-factor = $|F_o| - |F_c|/|F_o|$. Δ bond and Δ angle are the root mean square deviations of bond lengths and angles from ideal values, respectively.

Circular Dichroism Spectra. The CD spectra were measured at 25 °C on a J-600 spectropolarimeter (Japan Spectroscopic Co., Ltd.). Samples were dissolved either in 10 mM sodium acetate buffer, pH 5.5, containing 0.1 M NaCl, or in 10 mM Gly-HCl buffer, pH 3.0. The protein concentration and the optical path length were 0.15 mg/mL and 2 mm for the far-ultraviolet CD spectra (200–250 nm), and 0.5 mg/mL and 10 mm for the near-ultraviolet CD spectra (250–320 nm), respectively. The mean residue ellipticity, θ, which has units of deg cm^2 dmol^{-1}, was calculated using an average amino acid molecular weight of 110.

Thermal Denaturation. The thermal denaturation curve and the temperature of the midpoint of the transition, T_m, were determined as described previously (Kimura *et al.*, 1992a) by monitoring the CD value at 220 nm as the temperature was increased by 0.7 °C/min. Samples were dissolved to a concentration of 0.1–0.15 mg/mL in 10 mM Gly-HCl buffer, pH 3.0, or 20 mM sodium acetate buffer, pH 5.5, containing 0.1 M NaCl and 1.0 M GdnHCl. The thermal unfoldings of all mutant proteins were reversible under these conditions.

Crystallization, X-ray Data Collection, and Structure Determination. Crystals of V74L and V74I were obtained according to the procedure of Ishikawa *et al.* (1993a). Intensity data up to 1.8 Å were collected at room temperature on a four-circle diffractometer (Enraf-Nonius, CAD4) operated with a sealed copper tube. The crystallographic refinement was performed using the restrained least-squares refinement program PROLSQ (Hendrickson & Konnert, 1980), as previously reported by Ishikawa *et al.* (1993a). The refinement statistics are shown in Table I. Coordinates will be deposited in the Brookhaven Protein Data Bank (Bernstein *et al.*, 1977). The V74A mutant has not yielded diffraction-quality crystals to date.

RESULTS

Design of Amino Acid Replacements. Pores in the core structure of *E. coli* RNase HI were inspected to find the low packing-density region, following the protocol of Finkelstein and Nakamura (1993). A space including the wild-type protein was first divided into 3-Å cubic meshes, and the vacant volume not shared by the van der Waals spheres of any protein atom was computed. Figure 2 shows the center positions of these cubes, where the vacant volume exceeded 70%. It is obvious that the pores are localized near Val-74 and constitute a remarkable cavity within the hydrophobic core. According to the crystal structure of *E. coli* RNase HI (Katayanagi *et al.*, 1990, 1992), Val-74 is located in the αII helix. The accessible surface area of Val-74 is only 0.6 Å^2, indicating

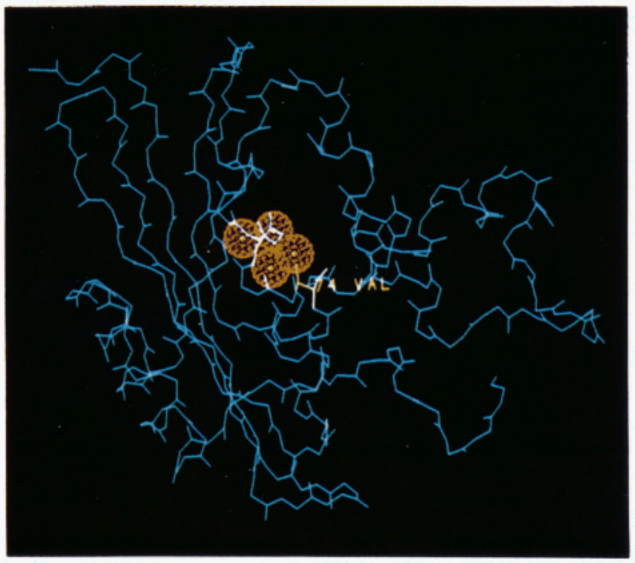

FIGURE 2: Localization of pores within the core of *E. coli* RNase HI. Pores are represented by red spheres. The backbone of the whole molecule (blue lines) and the side chain of Val-74 (yellow lines) are also shown.

Table II: Kinetic and Thermodynamic Parameters of Mutant RNases HI[a]

	relative		pH 3.0		pH 5.5	
protein	V_{max}	K_m	ΔT_m (°C)	$\Delta\Delta G$ (kcal/mol)	ΔT_m (°C)	$\Delta\Delta G$ (kcal/mol)
WT	1.0	1.0				
V74L	1.1	1.0	3.7	1.1	3.3	0.9
V74I	0.9	1.3	2.4	0.7	2.1	0.6
V74A	1.0	1.0	−7.6	−2.2	−12.7	−3.4

[a] Kinetic parameters were determined using the M13 DNA/RNA hybrid as a substrate. Relative V_{max} and K_m values were calculated by dividing these values of the mutant proteins by those of the wild-type protein, which were previously determined to be 20 units/mg and 0.11 μM, respectively (Oda *et al.*, 1993). Errors, which represent the 67% confidence limits, are within 30% of the values reported. Thermal denaturation curves for the mutant proteins were measured at both pH 3.0 and pH 5.5, as described under Experimental Procedures. ΔT_m is the change in the melting temperature, T_m, relative to that of the wild-type protein, which was previously determined to be 49.8 °C at pH 3.0 and 52.0 °C at pH 5.5 (Kimura *et al.*, 1992b). The change in the free energy of unfolding of the mutant protein relative to that of the wild-type protein was estimated by the relationship given by Becktel and Schellman (1987), $\Delta\Delta G = \Delta T_m \Delta S_m$ (wild type), where ΔS_m is the entropy change of the wild-type protein at T_m. The ΔS_m (wild type) was previously determined to be 0.304 kcal/(mol·K) at pH 3.0 and 0.275 kcal/(mol·K) at pH 5.5 (Kimura *et al.*, 1992b). Errors, which represent 67% confidence limits, are within ±0.3 °C in T_m and ±0.03 kcal/(mol·K) in ΔS_m at both pH values.

that this residue is almost completely shielded from the solvent. In order to examine the effect of the cavity volume on the protein stability, we replaced Val-74 with three different amino acid residues, Leu, Ile, and Ala. Two replacements, Val-74 → Leu and Val-74 → Ile, were expected to reduce the volume of the cavity. The other mutation, Val-74 → Ala, was designed to estimate the reduction in stability that might accompany the expansion of the cavity.

Kinetic Parameters. The kinetic parameters of the mutant proteins, V74A, V74L, and V74I, were determined using the M13 DNA/RNA hybrid as a substrate (Table II). The K_m and V_{max} values of all the mutant proteins are almost indistinguishable from those of the wild-type protein. The differences in each value between the wild-type and the mutant proteins are within experimental error. These results indicate that none of the mutations affect the enzymatic activity.

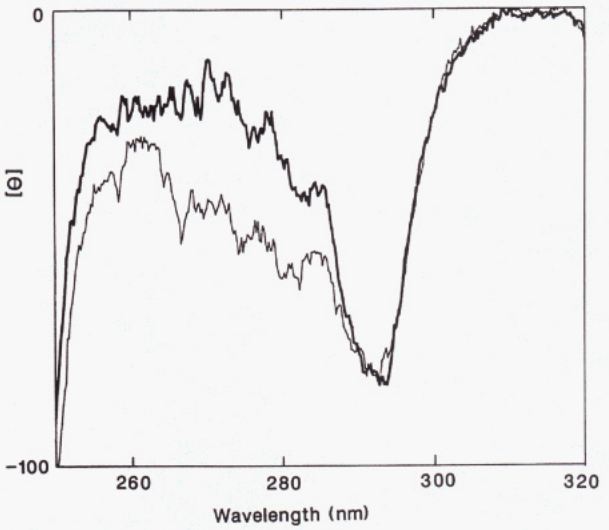

FIGURE 3: CD spectra of the *E. coli* RNase HI mutant. The near-ultraviolet CD spectrum of the mutant protein V74A (thick line) is compared to that of the wild-type protein (thin line). The spectra were measured at both pH 3.0 and pH 5.5. However, only the spectra at pH 5.5 are shown because of the similarities with those at pH 3.0.

Thermal Stability. Thermal denaturation curves of the mutant proteins were determined by monitoring the changes in CD values at pH 3.0 and 5.5. All mutant proteins unfold through a single cooperative transition at either pH (data not shown). The T_m value was determined from each thermal denaturation curve, and the difference in T_m (ΔT_m) of the mutant protein relative to the wild-type protein is listed in Table II. It is evident that V74L and V74I have higher T_m values than the wild-type protein, but the T_m of V74A dramatically decreases at either pH. Among all the mutant proteins and the wild-type protein, V74L is the most thermostable.

CD Spectra. In the far-ultraviolet region, the CD spectra from all the mutant proteins are identical to that of the wild-type protein at either pH 3.0 or pH 5.5 (data not shown). On the other hand, in the near-ultraviolet region, the CD spectrum of V74A is slightly different from that of the wild-type protein at pH 3.0, as well as at pH 5.5 (Figure 3). A gently sloping trough at 260–285 nm almost disappears in the CD spectrum of V74A, whereas CD spectra of V74L and V74I are similar to that of the wild-type protein in this region (data not shown). These results suggest that only the Val-74 → Ala mutation causes a local conformational change, but only to a small extent.

Overall Crystal Structures. To elucidate the mechanism of protein stabilization, the crystal structures of V74L and V74I were determined at 1.8-Å resolution. The overall structures of these mutant proteins are essentially the same as that of the wild-type protein. The root-mean-square displacements between the wild-type and the mutant proteins, V74L and V74I, are 0.16 and 0.17 Å, respectively, for the coordinates of all α-carbon atoms. The structural differences between the wild-type protein and these two mutant proteins are confined to the mutation site. The side chains of both mutated residues are well-defined in the electron density map. The accessible surface areas of Leu-74 and Ile-74 in the mutant proteins are 0.0 and 0.3 Å^2, respectively, indicating that these residues are almost buried as well as Val-74 in the wild-type protein.

Structure around Leu-74 in V74L. The refined crystal structure of V74L around residue 74 is compared to that of the wild-type protein in Figure 4a. The most striking structural feature in V74L is that the χ1 dihedral angle of the introduced

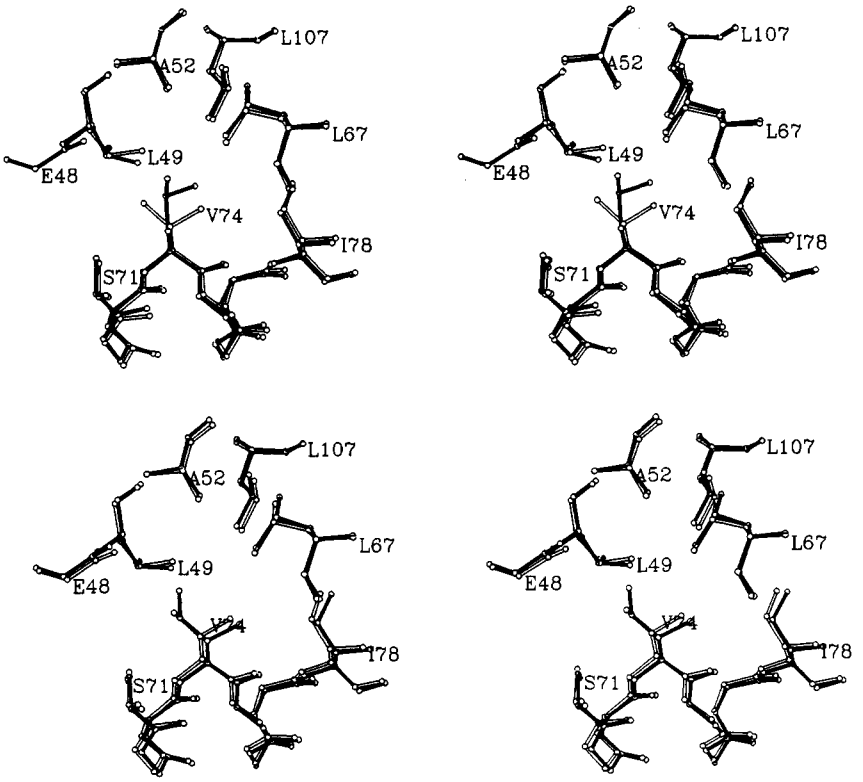

FIGURE 4: Comparison of the refined structures of the mutant proteins to that of the wild-type protein. The refined structures of (a, top) V74L and (b, bottom) V74I are represented by solid bonds and that of the wild-type protein is represented by open bonds.

Table III: Side-Chain Dihedral Angles of Residue 74

	protein and residue		
	wild type Val	V74L Leu	V74I Ile
χ1 (deg)	−72[a]	−116	−64
χ2 (deg)		175	−171

[a] Here, the χ1 value of Val is defined by the successive four atoms N, Cα, Cβ, and Cγ2 according to the definition of Janin *et al.* (1978). It is different from the conventional value defined by the N, Cα, Cβ, and Cγ1 atoms.

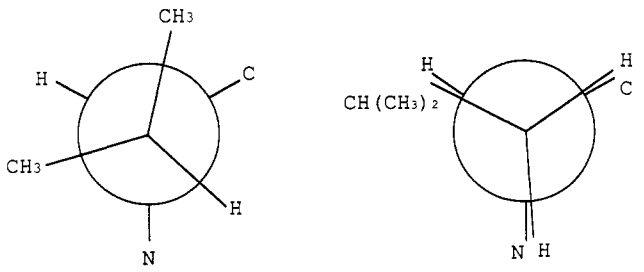

FIGURE 5: Newman projection for the Cα–Cβ bonds of residue 74. (a, left) Val in the wild-type protein; (b, right) Leu in V74L.

residue greatly deviates from that of the parent residue. The χ1 value of Val-74 is −72° in the wild-type structure, whereas the χ1 value of Leu-74 is −116° in the structure of V74L (Table III). The γ-atoms of Leu-74 assume an eclipsed conformation, while those of Val-74 assume a low-energy g^+ conformation (Figure 5). The χ2 value of Leu-74 is 175°, which indicates that the δ-atoms of this residue adopt a stable staggered conformation.

Structure around Ile-74 in V74I. The refined crystal structure of V74I around residue 74 is compared to that of the wild-type protein in Figure 4b. In contrast to Leu-74 in V74L, the side chain conformation of Ile-74 does not significantly deviate from that of the parent residue. The χ1 value of Ile-74 is −64° (Table III), which differs by only 8° as compared to that of Val-74. The χ2 value of Ile-74 is observed to be −171°. These results indicate that both the γ- and δ-atoms of Ile-74 assume a stable staggered conformation (low-energy g^+ conformation for γ-atoms).

The side chain atoms of Ile-74 slightly shift by 0.4 Å when compared to those of the parent residue, so that steric clash can be avoided between the δ-carbon of Ile-74 and the carbonyl oxygen of Glu-48. In the structure of V74I, the distance between these atoms is 3.5 Å. Consequently, the van der Waals interactions around the mutated residue are all acceptable, except for the one between the γ-carbon and the main-chain carbonyl oxygen of Ile-74. The distance between these atoms is 3.2 Å, which seems to be somewhat close and hence may cause an unfavorable van der Waals interaction.

Cavity Volume. In order to visualize the shape and location of the cavity around residue 74, we drew the Connolly surface (the contact and reentrant surface calculated with the 1.4-Å radius of a spherical water molecule) of the wild-type protein, from which the side chain atoms of Val-74 (Cβ, Cγ1, and Cγ2) were removed. In Figure 6, this surface is displayed together with the van der Waals surfaces of the side chain atoms of residue 74 in (a) the wild-type protein, (b) V74L, and (c) V74I, respectively. Obviously, the side chains of Leu-74 and Ile-74 fill the cavity space better than that of Val-74. A Leu residue has two δ-carbons, while an Ile residue has only one δ-carbon. Since the cavity extends above the two γ-carbon atoms of the Val-74 side chain in the wild-type protein, Leu fills the cavity more efficiently than Ile. Actually, the γ-carbon atom of Leu-74 is located within the middle of the cavity, such that the two δ-carbon atoms cover both sides of the cavity. The δ-carbon of Ile-74 is located at the edge of the cavity, and so expansion of the surrounding region may be required to avoid steric clash.

For a more quantitative analysis, the changes in the cavity volume around residue 74 of the mutant proteins relative to the wild-type protein were calculated with the Voronoi procedure [reviewed by Richards (1985)]. This procedure

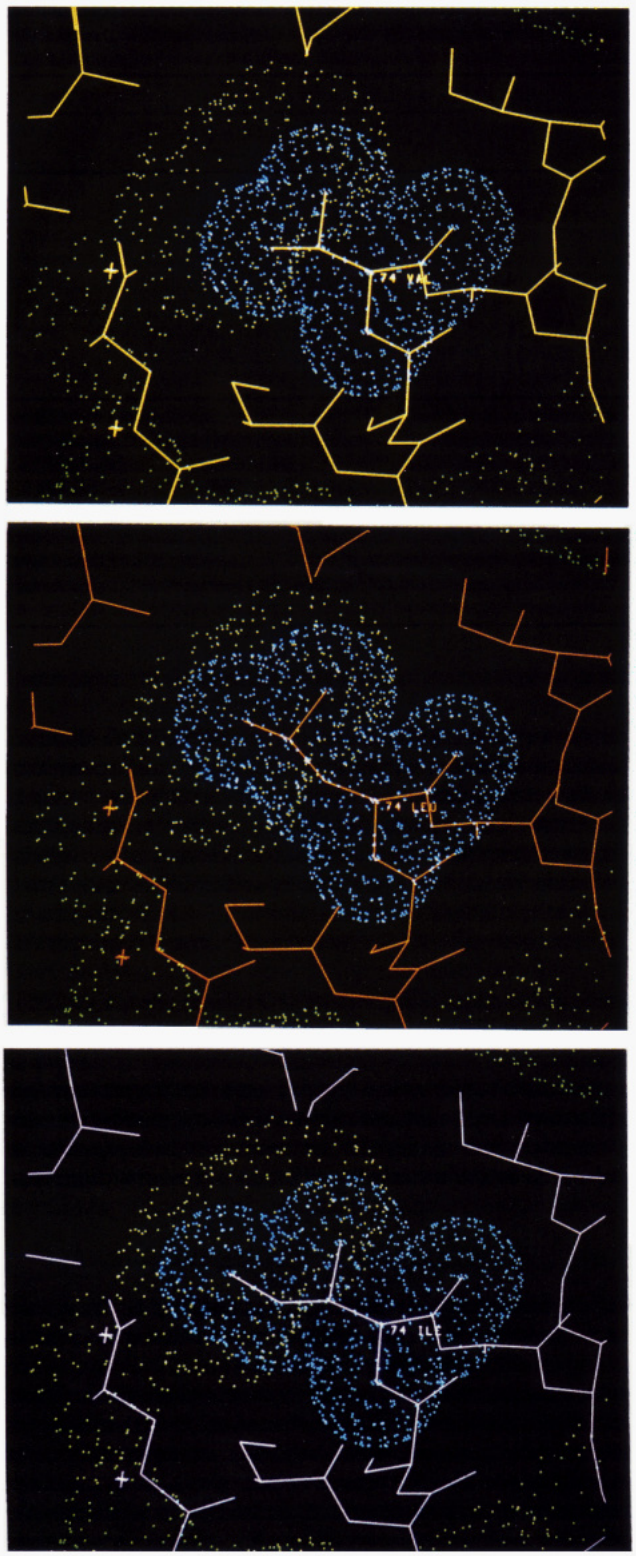

FIGURE 6: Visualization of the cavity determined in the absence of the Val-74 side chain in the wild-type structure. The cavity is represented by a Connolly surface (Connolly, 1986) with green dots. Also shown are the van der Waals surfaces (blue) of residue 74: (a, top) Val in the wild-type protein (yellow); (b, middle) Leu in V74L (red); (c, bottom) Ile in V74I (purple).

was recently applied to the cavity volume estimation of ribonuclease S variants by Varadarajan and Richards (1992). The cavity volume around residue 74 in the present work was estimated from the Voronoi volumes of residue 74 and the adjacent 14 residues, using the computer program written by S. Miyazawa, where the radius of a water molecule was assumed to be 1.4 Å. The results for the wild-type protein and the two mutant proteins, V74L and V74I, are listed in Table IV.

Table IV: Voronoi Values (Å^3) and Related Data Calculated from the Refined Structures of RNase HI Variants

	protein and residue 74		
	wild type Val	V74L Leu	V74I Ile
VOLSUR[a]	2660	2642	2653
ΔVOLSUR[f]		−18	−7
VOL74[b]	149	166	169
ΔVOL74[f]		17	20
VOLTOTAL[c]	2809	2808	2822
ΔVOLTOTAL[f]		−1	13
VDWVOL74[d]	105	124	124
ΔVDWVOL74[f]		19	19
ΔCAVVOL[e]		−20	−6

[a] VOLSUR: the summed Voronoi volume of the 14 residues that surround the cavity near residue 74 (residues 9, 48, 49, 52, 67, 69, 71, 73, 75, 78, 104, 107, 118, 120). [b] VOL74: the Voronoi volume of residue 74. [c] VOLTOTAL: VOLSUR + VOL74. [d] VDWVOL: the volume inside the van der Waals envelope, which was determined by Varadarajan and Richards (1992). [e] ΔCAVVOL: the difference in the cavity volume between the wild-type protein and each mutant protein (=ΔVOLTOTAL − ΔVDWVOL). [f] The difference in the respective values between the wild-type protein and each mutant protein.

The summed volume of the residues (VOLTOTAL) in the V74L structure is essentially the same as that in the wild-type structure, indicating that the introduced Leu-74 side chain is accommodated in the protein core without changing the structures of surrounding residues. Actually, the difference in the cavity volume between V74L and the wild-type protein (−ΔCAVVOL) is 20 Å^2, which is almost the same as the difference in the van der Waals volume (ΔVDWVOL74) between Val and Leu (19 Å^2).

On the other hand, the VOLTOTAL value of V74I is slightly larger than that of the wild-type protein, indicating that the δ-carbon atom of the Ile-74 side chain pushes out the neighboring atoms, as mentioned above (see Figure 4b). Consequently, the Val-74 → Ile replacement does not decrease the cavity volume as much as the Val-74 → Leu replacement does, although the van der Waals volume of Ile is essentially the same as that of Leu.

DISCUSSION

Hydrophobic stabilization by the filling of a cavity is generally attained by two different mechanisms (Sandberg & Terwilliger, 1989). The first is by the shielding of the additional methyl group from the external aqueous solvent. The second is by an increase in the favorable van der Waals interactions. Eriksson *et al.* (1992) reported that the former is constant, regardless of the context within the three-dimensional structure, while the latter increases in proportion to the change in the size of the cavity. In the following section, the mechanisms of the thermal stabilization in RNase HI mutant proteins are discussed in detail.

Val-74 → Ile Replacement. The mutant protein V74I is more thermostable, by 0.6–0.7 kcal/mol in ΔG, than the wild-type protein. The side chain of Ile contains a single additional δ-carbon as compared to that of Val. Since the major difference between V74I and the wild-type protein is only the presence or absence of this δ-carbon (Figure 4b), the $\Delta\Delta G$ value of 0.6–0.7 kcal/mol is considered to be a net hydrophobic effect gained by the burial of the δ-carbon. This value is slightly smaller than the empirically derived value of 1.3 ± 0.5 kcal/mol reported previously for the contribution of the burial of a single methylene group to the protein stability (Pace, 1992).

All the mutations analyzed by Pace (1992) were cavity-creating ones, such as Ile → Val and Val → Ala. Since the methylene groups in wild-type proteins are well packed, the change in van der Waals interactions is almost maximized. The same is true for the reported empirical $\Delta\Delta G$ value, which includes the van der Waals effect. V74I would have less of a contribution from the van der Waals interaction, because the cavity is not completely filled by the Ile-74 side chain (Figure 6c).

Val-74 → Leu Replacement. The Val-74 → Leu mutation results in a greater increase in the ΔG value than the Val-74 → Ile mutation. The $\Delta\Delta G$ value is 0.9–1.1 kcal/mol, which is similar to the value reported by Pace (1992). Since a single extra carbon atom is introduced into both mutant proteins, V74I and V74L, the increase in transfer free energy (from solvent to protein interior) must be essentially the same. The difference in the stability between these two mutant proteins (0.3–0.4 kcal/mol in ΔG) probably reflects the difference in the favorable van der Waals interaction energy. Actually, the cavity volume around residue 74 in the V74L structure is smaller than that in the V74I structure (Table IV).

The most striking feature in the conformation of Leu-74 is that the γ-atoms of Leu-74 assume an eclipsed conformation, which must detract from the protein stability (Figure 5b). However, the fact that V74L is more stable than the wild-type protein indicates that the stabilizing effect caused by the burial of the additional methylene group dominates the destabilizing effect caused by the introduction of an unfavorable $\chi 1$ angle into Leu-74. In fact, the eclipsed conformation of the Leu-74 side chain fills the cavity most efficiently without any steric clash (Figures 4a, 6b). Staggered conformations, with either $\chi 1 \approx -60°$ or $\chi 1 \approx -180°$, would leave large unfilled spaces in either side of the cavity.

It should be noted that, in addition to the hydrophobic effect, two other effects can be considered with respect to the difference in the $\Delta\Delta G$ values between V74I and V74L. One is the steric clash between the backbone carbonyl and the γ-carbon of the β-branched residues, Val or Ile. In the wild-type and V74I structures, the close distance between these associated atoms may negatively affect the stability to some extent. The other is the energy loss due to the eclipsed conformation of Leu-74 in the V74L mutant protein, as compared to the staggered conformation. Statistical and energetic analyses of the relative frequency of occurrence of this type of local side chain conformation can give us insight to this problem, as discussed below.

Statistical distributions of $\chi 1$ angles in natural proteins are shown for Val, Leu, and Ile in the helical structures (Figure 7). The eclipsed conformation with the $\chi 1$ angle of $\approx -120°$, which is found in V74L, is sometimes observed for Leu but is rarely observed for the β-branched residues, Val or Ile, suggesting that it is not strictly prohibited for Leu. For a Leu side chain, this conformation has only a small energy difference from the staggered conformation, which is estimated to 1.4 kcal/mol from the statistic frequency of occurrence in $\chi 1$ angles, probably because each pair of overlapping atoms in Leu has a hydrogen atom (Figure 5b). On the other hand, the eclipsed conformations with $\chi 1$ angles of $\approx 0°$ and $\approx 120°$ are infrequently observed for any of the three residues, suggesting that these conformations are quite unstable and are not favored to exist in natural proteins. Similar results were reported by McGregor *et al.* (1987).

In the same figure, the Boltzmann probability, based upon the molecular mechanics calculation, is also indicated for each side chain conformation. Energy minimization was carried out by the molecular mechanics program PRESTO (Morikami

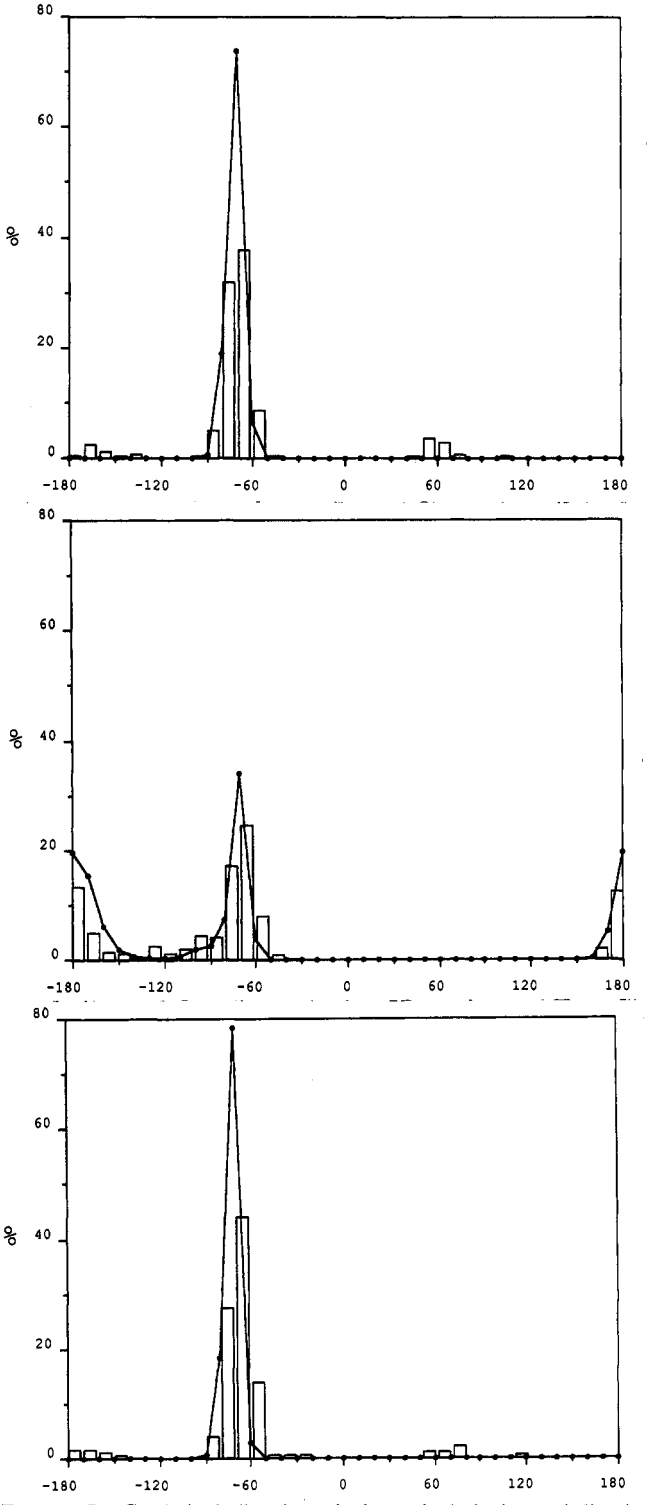

FIGURE 7: Statistical (bars) and theoretical (polygonal lines) distribution of the relative frequency of occurrence of $\chi 1$ angles in α-helices for (a, top) Val residues, (b, middle) Leu residues, and (c, bottom) Ile residues. The statistical analysis was performed on the basis of the 80 crystal structures refined at better than 1.8-Å resolution. The α-helices were defined by the program DSSP (Kabsch & Sander, 1983). The structures were obtained from the Brookhaven Protein Data Bank (Bernstein *et al.*, 1977). The Brookhaven identifications for the included structures are 1AAP, 1APB, 1BP2, 1CDP, 1CHO, 1COX, 1CRN, 1CSE, 1CTF, 1FKF, 1GCR, 1GD1, 1LZ1, 1MBA, 1PAL, 1PAZ, 1PCY, 1PGX, 1PPT, 1PSG, 1RDG, 1ROP, 1SAR, 1SGT, 1SN3, 1SNC, 1SRN, 1TGN, 1TON, 1UBQ, 1UTG, 256B, 2ACT, 2ALP, 2APR, 2AZA, 2CDV, 2CGA, 2CPP, 2CSC, 2CYP, 2ER7, 2FCR, 2FGF, 2GCT, 2HMQ, 2LTN, 2OVO, 2PAB, 2PRK, 2RHE, 2RNT, 2SAR, 2TMN, 2TRX, 2UTG, 2WRP, 2XIS, 351C, 3APP, 3B5C, 3CLA, 3CYT, 3EBX, 3ER5, 3EST, 3GRS, 3RN3, 3SGB, 4BP2, 4DFR, 4PEP, 4PTI, 4RXN, 5CPA, 5P21, 5RUB, 6ABP, 9PAP, and 9WGA. See the text for the theoretical analysis. The definition of the $\chi 1$ value of Val is the same as that in Table III.

et al., 1992), keeping the positions of the backbone atoms and the Cβ and Cγ atoms of Val, Leu, and Ile in the middle of the computed right-handed α-helix of a 15-mer polyalanine. The all-atom AMBER force field (Weiner *et al.*, 1986), excluding the 1–5 electrostatic energy term, was used. For Leu and Ile, the initial positions of the Cδ atoms were first set at three different positions (χ2 angle = 180°, −60°, and +60°). The Boltzmann probabilities were taken from these minimized energies, changing the χ1 angles of Val, Leu, and Ile, respectively. The computed results agree very well with the statistical frequencies observed in natural proteins, as shown in Figure 7. Similar results were reported for typical rotamers by Piela *et al.* (1987).

Structural Comparison of Leu-74 in V74L and That in Thermus thermophilus RNase H. An amino acid sequence alignment of *E. coli* RNase HI and its thermophilic counterpart, *T. thermophilus* RNase H, has revealed that Val-74 of the mesophilic enzyme is substituted by Leu in the thermophlic enzyme (Kanaya & Itaya, 1992). The crystal structure of *T. thermophilus* RNase H determined at 2.8-Å resolution (Ishikawa *et al.*, 1993b) shows that the χ1 angle of Leu-74 is −98°. This conformation also deviates from the staggered conformation with $\chi1 \approx -60°$, suggesting that Leu actually has a wide acceptable range for its χ1 value, from $\chi1 \approx -60°$ to $\chi1 \approx -180°$. In *T. thermophilus* RNase H, Leu-74 interacts with amino acid residues different from those in *E. coli* RNase HI. For example, Ile-78, Trp-118, and Trp-120 in *E. coli* RNase HI are all substituted by Phe in *T. thermophilus* RNase H. Therefore, it is difficult to conclude that Leu-74 in *T. thermophilus* RNase H plays the same role on its high thermostability.

Destabilization by Val-74 → Ala. The mutant protein V74A is less stable than the wild-type protein by 2.2 kcal/mol at pH 3.0 and 3.4 kcal/mol at pH 5.5 in the ΔG values. The removal of two buried carbon atoms probably creates a larger cavity space than in the wild-type protein, with slight shifts of the surrounding atoms toward the vacated space, as shown for a variety of T4 lysozyme mutant proteins (Eriksson *et al.*, 1992). The creation of a cavity is supposed to be responsible for the decrease in the thermal stability. It remains to be determined why the absolute $\Delta\Delta G$ value at pH 5.5 is larger than that at pH 3.0. The CD spectra of V74A suggest that the Val-74 → Ala mutation causes a slight local conformational change. It should be noted that both the far- and near-ultraviolet spectra of V74A are almost identical to those of the mutant protein S68G, in which Ser-68 is replaced by Gly (Kimura *et al.*, 1992a). The NMR spectra of S68G suggest that the tertiary structures around Trp-118 and Trp-120 are affected by the Ser-68 → Gly mutation. Since this mutation is also expected to create a cavity, and because the β-E strand, in which Trp-118 and Trp-120 are located, seems to be adjacent to the cavity created by either the Ser-68 → Gly or the Val-74 → Ala mutations, it seems likely that similar conformational changes are introduced by these cavity-creating mutations.

Conclusion. One of the main purposes of protein engineering is to develop methods that improve protein stability. Such a rational design of proteins allows one to construct enzymes resistant to heat inactivation and thereby make them tolerant to industrial uses. Various approaches have so far been tested for this purpose, and several useful strategies to enhance protein stability have been proposed (Branden & Tooze, 1991). Our results demonstrate that substitutions of small hydrophobic amino acid residues with bulkier ones, which face cavities located within the interior of the protein molecule, should be considered as a promising strategy to enhance protein stability. A Leu residue seems to be a better amino acid residue than Val or Ile to be substituted within a hydrophobic core, because it can adopt a wide range of χ1 angles without seriously sacrificing stability.

ACKNOWLEDGMENT

We thank M. Oobatake for the calculation of accessible surface areas and helpful discussions, S. Miyazawa for the donation of the program calculating the Voronoi volume, K. Katayanagi for providing the coordinates of *E. coli* RNase HI before publication, and T. Miyazawa for helpful discussions.

REFERENCES

Becktel, W. J., & Schellman, J. A. (1987) *Biopolymers 26*, 1859–1877.
Bernstein, F. C., Koetzle, T. F., Williams, G. J. B., Meyer, E. F., Brice, M. D., Rodgers, J. R., Kennard, O., Shimanouchi, T., & Tasumi, M. (1977) *J. Mol. Biol. 112*, 535–542.
Branden, C., & Tooze, J. (1991) in *Introduction to Protein Structure*, pp 247–268, Gerland Publishing, Inc., New York.
Connolly, M. L. (1986) *Int. J. Pept. Protein Res. 28*, 360–363.
Daopin, S., Alber, T., Baase, W. A., Wozniak, J. A., & Matthews, B. W. (1991) *J. Mol. Biol. 221*, 647–667.
Dill, K. (1990) *Biochemistry 29*, 7133–7155.
Eijsink, V. G. H., Dijkstra, B. W., Vriend, G., van der Zee, J. R., Veltman, O. R., van der Vinne, B., van den Burg, B., Kempe, S., & Venema, G. (1992) *Protein Eng. 5*, 421–426.
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.
Finkelstein, A. V., & Nakamura, H. (1993) *Protein Eng.* (in press).
Hendrickson, W. A., & Konnert, J. H. (1980) in *Computing in Crystallography* (Diamond, R., Ramaseshan, S., & Venkatesan, K., Ed.) pp 13.01–13.23, Indian Academy of Science, International Union of Crystallography, Bangalore.
Hurley, J. H., Baase, W. A., & Matthews, B. W. (1992) *J. Mol. Biol. 224*, 1143–1159.
Ishikawa, K., Kimura, S., Kanaya, S., Morikawa, K., & Nakamura, H. (1993a) *Protein Eng. 6*, 85–91.
Ishikawa, K., Okumura, M., Katayanagi, K., Kimura, S., Kanaya, S., Nakamura, H., & Morikawa, K. (1993b) *J. Mol. Biol. 230*, 529–542.
Janin, J., Wodak, S., Levitt, M., & Maigret, B. (1978) *J. Mol. Biol. 125*, 357–386.
Kabsch, W., & Sander, C. (1983) *Biopolymers 22*, 2577–2637.
Kanaya, S., & Itaya, M. (1992) *J. Biol. Chem. 267*, 10184–10192.
Kanaya, S., Kimura, S., Katsuda, C., & Ikehara, M. (1990) *Biochem. J. 271*, 59–66.
Kanaya, S., Katsuda, C., Kimura, S., Nakai, T., Kitakuni, E., Nakamura, H., Katayanagi, K., Morikawa, K., & Ikehara, M. (1991) *J. Biol. Chem. 266*, 6038–6044.
Kanaya, S., Oobatake, M., Nakamura, H., & Ikehara, M. (1993) *J. Biotechnol. 28*, 117–136.
Karpusas, M., Baase, W. A., Matsumura, M., & Matthews, B. W. (1989) *Proc. Natl. Acad. Sci. U.S.A. 86*, 8237–8241.
Katayanagi, K., Miyagawa, M., Matsushima, M., Ishikawa, M., Kanaya, S., Ikehara, M., Matsuzaki, T., & Morikawa, K. (1990) *Nature 347*, 306–309.
Katayanagi, K., Miyagawa, M., Matsushima, M., Ishikawa, M., Kanaya, S., Nakamura, H., Ikehara, M., Matsuzaki, T., & Morikawa, K. (1992) *J. Mol. Biol. 223*, 1029–1052.
Kellis, J. T., Jr., Nyberg, K., Sali, D., & Fersht, A. R. (1988) *Nature 333*, 784–786.
Kellis, J. T., Jr., Nyberg, K., & Fersht, A. R. (1989) *Biochemistry 28*, 4914–4922.
Kimura, S., Oda, Y., Nakai, T., Katayanagi, K., Kitakuni, E., Nakai, C., Nakamura, H., Ikehara, M., & Kanaya, S. (1992a) *Eur. J. Biochem. 206*, 337–343.
Kimura, S., Nakamura, H., Hashimoto, T., Oobatake, M., & Kanaya, S. (1992b) *J. Biol. Chem. 267*, 21535–21542.

Kimura, S., Kanaya, S., & Nakamura, H. (1992c) *J. Biol. Chem. 267*, 22014–22017.
Lim, W. A., & Sauer, R. T. (1991) *J. Mol. Biol. 219*, 359–376.
Matsumura, M., Becktel, W. J., & Matthews, B. W. (1988) *Nature 334*, 406–410.
Matsumura, M., Wozniak, J. A., Dao-pin, S., & Matthews, B. W. (1989) *J. Biol. Chem. 264*, 16059–16066.
McGregor, M. J., Islam, S. A., & Sternberg, M. J. E. (1987) *J. Mol. Biol. 198*, 295–310.
Mendel, D., Ellman, J. A., Chang, Z., Veenstra, D. L., Kollman, P. A., & Schultz, P. G. (1992) *Science 256*, 1798–1802.
Morikami, K., Nakai, T., Kidera, A., Saito, M., & Nakamura, H. (1992) *Comput. Chem. 16*, 243–248.
Oda, Y., Yoshida, M., & Kanaya, S. (1993) *J. Biol. Chem. 268*, 88–92.
Pace, C. N. (1992) *J. Mol. Biol. 226*, 29–35.
Piela, L., Nemethy, G., & Scheraga, H. A. (1987) *Biopolymers 26*, 1273–1286.
Richards, F. M. (1985) *Methods Enzymol. 115*, 440–464.
Sandberg, W. S., & Terwilliger, T. C. (1989) *Science 245*, 54–57.
Sandberg, W. S., & Terwilliger, T. C. (1991) *Proc. Natl. Acad. Sci. U.S.A. 88*, 1706–1710.
Sharp, K. A. (1991) *Curr. Opin. Struct. Biol. 1*, 171–174.
Shortle, D., Stites, W. E., & Meeker, A. K. (1990) *Biochemistry 29*, 8033–8041.
Sneddon, S. F., & Tobias, D. J. (1992) *Biochemistry 31*, 2842–2846.
Varadarajan, R., & Richards, F. M. (1992) *Biochemistry 31*, 12315–12327.
Weiner, S. J., Kollman, P. A., Nguyen, D. T., & Case, D. A. (1986) *J. Comput. Chem. 7*, 230–252.
Yang, W., Hendrickson, W. A., Crouch, R. J., & Satow, Y. (1990) *Science 249*, 1398–1405.
Yutani, K., Ogasahara, K., Tsujita, T., & Sugino, Y. (1987) *Proc. Natl. Acad. Sci. U.S.A. 84*, 4441–4444.