Cooperative Stabilization of *Escherichia coli* Ribonuclease HI by Insertion of Gly-80b and Gly-77 → Ala Substitution

Kokhi Ishikawa, Haruki Nakamura, Kosuke Morikawa, Shigenobu Kimura,§ and Shigenori Kanaya*

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

Received March 26, 1993; Revised Manuscript Received May 10, 1993

ABSTRACT: The insertion of a Gly residue (designated as Gly-80b) between the C-cap of the αII-helix (Gln-80) and the N-cap of the αIII-helix (Trp-81) in *Escherichia coli* ribonuclease HI enhances the protein stability by 0.4 kcal/mol in ΔG (Kimura, S., Nakamura, H., Hashimoto, T., Oobatake, M., & Kanaya, S. (1992) *J. Biol. Chem. 267*, 21535–21542). Another mutation within the αII-helix, Gly-77 → Ala, reduces the stability by 0.9 kcal/mol. Simultaneous introduction of these mutations enhances the stability by 0.8 kcal/mol, indicating that the effects of these mutations are cooperative and not simply independent. We determined the crystal structures of these three mutant proteins (G^{80b}-, A^{77}-, and A^{77}/G^{80b}-RNase H) to investigate this cooperative mechanism of the protein stabilization. The structures revealed that the inserted Gly-80b assumes a left-handed helical conformation in both the G^{80b}- and the A^{77}/G^{80b}-RNase H. This inserted glycine residue allows the formation of a "paperclip", which is a common motif at the C-termini of α-helices. Accompanying the formation of the paperclip motif, two intrahelical hydrogen bonds are formed between the backbone atoms (O78–N80b and O80b–N84). The stabilization caused by the insertion of Gly-80b can be ascribed to the formation of these hydrogen bonds. The Gly-77 → Ala substitution destabilizes the protein due to the deformed packing interactions in the hydrophobic core around Ala-77 and the stress in the wedged indole ring of Trp-81. These effects are alleviated by the insertion of Gly-80b, which relaxes the backbone structure. The A^{77}/G^{80b}-RNase H is more stable than the G^{80b}-RNase H because of the helix-stabilization and the hydrophobic effect associated with the Gly-77 → Ala replacement.

Establishment of methods to improve protein stability is one of the main purposes of protein engineering. Various strategies have been proposed to enhance protein stability (Branden & Tooze, 1991). However, our knowledge of the structure–stability relationships of proteins is too meager to rationally design amino acid replacements for the enhancement of protein stability. Further mutagenesis experiments, supported by solid structural backgrounds, will be necessary to make such design more reliable.

One promising strategy to yield greater amounts of information on structure–stability relationships of proteins is a comparative study of proteins from mesophilic and thermophilic origins. Ribonuclease H (RNase H), which endonucleolytically hydrolyzes only the RNA strand of DNA/RNA hybrids (Crouch & Dirksen, 1982), is an ideal protein for such a purpose, due to the following reasons: (1) The availability of the structural genes of the enzymes from both *Escherichia coli* (Kanaya & Crouch, 1983) and *Thermus thermophilus* (Itaya & Kondo, 1991) facilitates mutational experiments. (2) *E. coli* RNase HI and *T. thermophilus* RNase H are small in size and are composed of 155 and 164 amino acid residues, respectively. (3) Despite the relatively high amino acid sequence similarity (52% identity), the thermophilic enzyme is more stable than the *E. coli* enzyme by 33.9 °C in T_m (Kanaya & Itaya, 1992). (4) Both *E. coli* RNase HI (Kanaya *et al.*, 1991a) and *T. thermophilus* RNase H (Kanaya & Itaya, 1992) reversibly unfold in a single cooperative fashion with heat and guanidine hydrochloride denaturations. (5) The three-dimensional structures of both *E. coli* RNase HI (Katayanagi *et al.*, 1990, 1992; Yang *et al.*, 1990) and *T. thermophilus* RNase H (Ishikawa et al., 1993b) have been determined by X-ray analysis.

We have systematically replaced nine regions (R1–R9) in the amino acid sequence of *E. coli* RNase HI with the corresponding regions from *T. thermophilus* RNase H and have examined the effect of each mutation on the stability of *E. coli* RNase HI (Kimura *et al.*, 1992a). We found that replacements in the following three regions yielded more stable mutant proteins (R_4-, R_5-, and R_6-RNase H proteins) than the *E. coli* protein at pH values of 3.0 and 5.5. The stabilization mechanisms for two of the mutant proteins (R_4- and R_6-RNase H) have already been elucidated.

R_4-RNase H has a single amino acid substitution, His-62 → Pro, in a loop region. The crystal structure of this mutant protein (Ishikawa *et al.*, 1993a) provided evidence for previous proposals that Pro residues in loops contribute to protein stability by decreasing the entropy of unfolding (Matthews *et al.*, 1987; Yutani *et al.*, 1991). R_6-RNase H has three amino acid substitutions (residues 91, 94, and 95) in another loop region. Further mutagenesis experiments (Kimura *et al.*, 1992b), followed by X-ray crystallographic studies (Ishikawa *et al.*, 1993a), revealed that replacement of Lys-95, with a left-handed conformation, by Gly enhances protein stability. These results indicate that comparative studies between *E. coli* RNase HI and *T. thermophilus* RNase H are effective, not only to propose novel strategies but also to justify strategies previously proposed to enhance protein stability. In this article, we investigate the stabilization mechanism of R_5-RNase H.

R_5-RNase H has six amino acid substitutions (residues 74–78 and 80) in the αII-helix and an insertion of Gly-80b between αII and αIII. The amino acid sequence of R_5-RNase H is LKKAFTEG in the single-letter codes, while that of the wild-type protein is VRQGITQ. The hydrophobic side-chains in

* Author to whom correspondence should be addressed.
§ Present address: Basic Research Laboratories, Toray Industries, Inc., 1, 111, Tebiro, Kamakura, Kanagawa 248, Japan.

0006-2960/93/0432-7136$04.00/0

this region (R_5) are tightly packed into a hydrophobic core. The hydrophobic residues in *E. coli* RNase HI are substituted by more bulky residues in *T. thermophilus* RNase H, such as Val-74 → Leu and Ile-78 → Phe. Comparison of the crystal structures of the two enzymes indicates that the increased side-chain volumes in *T. thermophilus* RNase H force the αII-helix to shift away from the molecular center by more than 1.0 Å, as compared to that in *E. coli* RNase HI (Ishikawa *et al.*, 1993b). The strain caused by this shift seems to be alleviated by the insertion of Gly-80b. It was previously shown that either the insertion of Gly-80b alone (Kimura *et al.*, 1992a) or the Val-74 → Leu substitution alone (Ishikawa *et al.*, 1993c) increases the stability of *E. coli* RNase HI by 0.4 kcal/mol or 1.1 kcal/mol in ΔG at pH 3.0, respectively. Since R_5-RNase H is more stable than the wild-type protein by only 0.8 kcal/mol at this pH, it seems likely that the stability of R_5-RNase H reflects additional contributions from individual amino acid residues within the R_5 region.

We first started to replace amino acid residues in the R_5 region of *E. coli* RNase HI with the corresponding residues of *T. thermophilus* RNase H. Using this process, we found that the Gly-80b insertion and the Gly-77 → Ala substitution cooperatively increase the protein stability. The crystal structures of the three mutant proteins, in which these two mutations are individually or simultaneously introduced, allow us to discuss the mechanism by which the protein stability is altered.

EXPERIMENTAL PROCEDURES

Mutant Constructions. Plasmid pJAL600(A*fl*II) for the overproduction of *E. coli* RNase HI, in which the *rnh*A gene is controlled by the bacteriophage λ promoters P_R and P_L, was previously constructed (Ishikawa *et al.*, 1993c). This plasmid contains unique A*fl*II and *Bam*HI sites in the *rnh*A gene, encompassing the sequences encoding amino acid residues 67–82. The Gly-77 → Ala mutation with or without the insertion of Gly-80b, or the insertion of Ala-80b, was introduced by cassette mutagenesis by replacing the 44-base-pair A*fl*II–*Bam*HI fragment of pJAL600(A*fl*II) with either the 44- or 47-base-pair chemically synthesized A*fl*II–*Bam*HI fragment. The GGT codon of Gly-77 was changed to GCT. For the insertion of either Gly-80b or Ala-80b, a GGT or a GCT codon, respectively, was newly inserted. Accompanying these mutations, the TAT codon of Tyr-73 and the GTC codon of Val-74 were silently changed to TAC and GTA, respectively, to create a *Spl*I site within the gene. The mutants were initially screened by *Spl*I mapping on plasmid DNA. All oligonucleotides used in this experiment were synthesized with an Applied Biosystems automatic synthesizer, Model 380A.

Construction of an overproducing strain, overproduction and purification of the mutant proteins, and determination of the amino acid sequences to confirm mutations were carried out as described previously (Ishikawa *et al.*, 1993c). Usually 2–5 mg of the purified proteins was obtained from 200 mL of culture. The mutant proteins in which Gly-77 is replaced by Ala, or Ala-80b is inserted, are designated as A^{77}-RNase H and A^{80b}-RNase H, respectively. G^{80b}-RNase H, in which Gly-80b is inserted, was previously constructed (Kimura *et al.*, 1992a). The mutant protein with the double mutation is designated as A^{77}/G^{80b}-RNase H.

Determination of Kinetic Parameters. The kinetic parameters were determined as described previously (Kanaya *et al.*, 1991a), using ^{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 min. Protein 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).

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 temperature was increased by 0.7 °C/min. Samples were dissolved to a final concentration of 0.1–0.15 mg/mL in 10 mM Gly–HCl buffer, pH 3.0. The thermal unfoldings of all mutant proteins were reversible under these conditions.

Crystallization and X-ray Data Collection. Crystals of the three mutant proteins, A^{77}-RNase H, G^{80b}-RNase H, and A^{77}/G^{80b}-RNase H, were obtained according to the procedure of Ishikawa *et al.* (1993a). X-ray intensity data for A^{77}/G^{80b}-RNase H were collected with a FAST area-diffractometer on a GX21 rotating-anode X-ray generator (Enraf-Nonius). Intensity data for the other two mutant proteins were collected on a four-circle diffractometer (Enraf-Nonius, CAD4) operated with a sealed copper tube.

Structure Determination and Refinement. The crystallographic refinement of A^{77}-RNase H was performed using the restrained least-square refinement program PROLSQ (Hendrickson & Konnert, 1980), as previously reported by Ishikawa *et al.* (1993a).

The crystal structure of G^{80b}-RNase H was solved by the molecular replacement method. The coordinates of wild-type *E. coli* RNase HI refined at 1.48 Å (Katayanagi *et al.*, 1992) were placed in an orthogonal *P*1 unit cell with axes of 80-Å length. Triclinic structure factors from this hypothetical crystal were calculated between 10.0- and 2.5-Å resolution, assuming an overall temperature factor of 15 Å^2. The cross-rotation function was then calculated within a spherical shell bounded by radii of 3 and 20 Å, using the real space search routine of the program PROTEIN (Steigemann, 1974). The step size (in the angles notation (ψ, θ, ϕ) defined by R. Huber $(0 \le \psi \le 180°, 0 \le \theta \le 180°, 0 \le \phi \le 180°)$ for the ϕ rotation was initially 5° and was decreased to 0.2° near the highest correlation peak. The highest peak was observed at $\psi = 2.6°$, $\theta = 8.8°$, $\phi = -3.4°$.

The location of the correctly oriented molecule was determined using the translation function of Crowther & Blow (1967) with programs written by E. E. Lattman and modified by J. Deisenhofer and R. Huber. The model Fourier transform was calculated in an orthogonal cell with axes of 80-Å length using data from 10.0- and 2.5-Å resolution. The position of the center of mass was determined to be (–3.40, 16.38, 10.85) given in $^1/_{44}$, $^1/_{88}$, and $^1/_{36}$ fractions of the cell lengths, respectively.

The positional and orientational parameters obtained from the molecular replacement technique were further improved by rigid-body refinement, using the program TRAREF (Huber & Schneider, 1985). At this stage, the crystallographic *R*-factor was 0.421 for the intensity data from 10.0- to 2.5-Å resolution. The correctly oriented and positioned molecule was then refined by alternating rounds of XPLOR (Brünger *et al.*, 1987) simulated annealing refinement and manual rebuilding using the program FRODO (Jones, 1978). Finally, the model was refined at 1.9-Å resolution with the program PROLSQ.

Since the crystal of A^{77}/G^{80b}-RNase H is isomorphous with that of G^{80b}-RNase H, the crystal structure of A^{77}/G^{80b}-RNase H was solved on the basis of the refined structure of G^{80b}-RNase H using the programs TRAREF and PROLSQ.

Table I: Crystal Data and Refinement Statistics[a]

	G^{80b}	A^{77}	A^{77}/G^{80b}
cell dimens			
a (Å)	45.03	44.35	44.98
b (Å)	86.93	86.80	86.45
c (Å)	35.04	35.52	35.38
resolution (Å)	6~1.9	6~2.0	6~2.0
no. of refls	8658	6941	6964
R-factor	0.194	0.191	0.183
Δbond (Å)	0.015	0.015	0.015
Δangle (Å)	0.034	0.034	0.037

[a] The crystals belong to the space group $P2_12_12_1$. The unit cell parameters of the wild-type protein are a = 44.06 Å, b = 86.85 Å, c = 35.47 Å. R-factor = $|F_{obs}| - |F_{calc}|/|F_{obs}|$. Δbond and Δangle are the root mean square deviations of bond lengths and angles from ideal values, respectively.

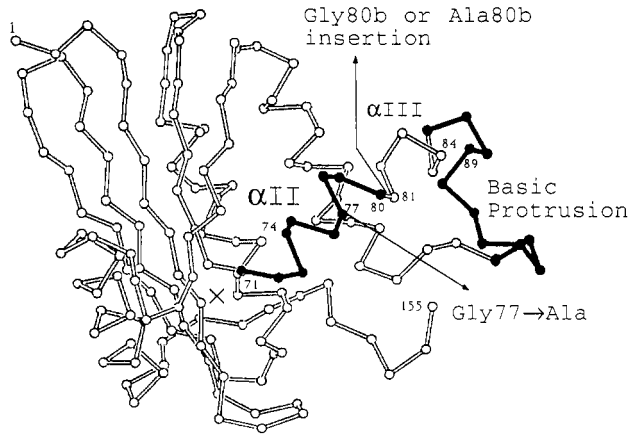

FIGURE 1: Backbone structures of *E. coli* RNase HI. The locations of the αII-helix (residues 71–80) and the αIII-helix (residues 81–89), as well as the positions and types of amino acid substitutions, are shown. The N- and C-termini are indicated by the residue numbers 1 and 155, respectively. All the residues involved in the αII-helix (SQYVRQGITQ) and the basic protrusion (KKRGWKTADKK) are indicated by solid bonds. × represents the catalytic site of the enzyme, in which Mg^{2+} binds.

The refinement statistics of the three mutant proteins are provided in Table I. Coordinates will be deposited in the Brookhaven Protein Data Bank.

RESULTS

Stabilities. The locations of the mutated residues in the wild-type structure are shown in Figure 1. Position 77 is located within the αII-helix, and position 80b is located at the junction between αII and αIII, the axes of which make a kink at an angle of 35° (Katayanagi *et al.*, 1992).

The differences in the melting temperatures (ΔT_m) and in the free energy change of unfolding ($\Delta\Delta G$) of the mutant proteins relative to the wild-type protein are summarized in Table II. The effects of the insertions of Gly or Ala at position 80b on the protein stability are opposite to each other. While the insertion of Gly-80b enhances the thermostability by 0.4 kcal/mol in ΔG, that of Ala-80b reduces it by 0.5 kcal/mol.

The replacement of Gly-77 by Ala reduces the protein stability by 0.9 kcal/mol. If the effects of the Gly-77 → Ala substitution and the Gly-80b insertion on the protein stability were independent of each other, the double-mutant protein, A^{77}/G^{80b}-RNase, would be less stable than the wild-type protein by 0.5 kcal/mol. However, A^{77}/G^{80b}-RNase H is more stable than the wild-type protein by 0.8 kcal/mol, indicating that the Gly-77 → Ala substitution and the Gly-80b insertion cooperatively increase the protein stability.

Enzymatic Activities. The kinetic parameters of the mutant proteins relative to those of the wild-type protein are summarized in Table II. The insertion of either Gly-80b or Ala-80b results in a decrease in the k_{cat} value along with a considerable increase in the K_m value, suggesting that these mutations seriously affect the substrate-binding site of the enzyme. In contrast, the Gly-77 → Ala substitution does not seriously affect the substrate binding, although the hydrolysis rate was significantly decreased by the mutation. Little difference is observed between the kinetic parameters of A^{77}/G^{80b}-RNase H and those of G^{80b}-RNase H, suggesting that the effects of the Gly-77 → Ala substitution and the Gly-80b insertion on the enzymatic activity are not additive.

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

	relative			
protein	k_{cat}	K_m	ΔT_m (°C)	$\Delta\Delta G$ (kcal/mol)
WT	1.0	1.0		
G^{80b}-RNase H	0.1	50	1.2[b]	0.4[b]
A^{80b}-RNase H	0.1	50	−1.5	−0.5
A^{77}-RNase H	0.3	1.5	−2.9	−0.9
A^{77}/G^{80b}-RNase H	0.1	50	2.7	0.8

[a] Kinetic parameters were determined using the M13 DNA/RNA hybrid as a substrate. Relative k_{cat} and K_m values were calculated by dividing these values of the mutant proteins by those of the wild-type protein, which are 20 units/mg and 0.11 μM, respectively. Errors, which represent the 67% confidence limits, are within 30% of the values reported. Thermal denaturation curves for the mutant proteins were measured at pH 3.0, 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 is 49.8 °C. 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 $\Delta\Delta G = \Delta T_m \Delta S_m$ (Becktel & Schellman, 1987). ΔS_m is the entropy change of the wild-type protein at T_m, which was previously determined as 0.304 kcal/(mol·K) (Kimura et al., 1992a). Errors, which represent the 67% confidence limits, are within ±0.3 °C in T_m. [b] Data from Kimura *et al.* (1992a).

Overall Crystal Structures. To elucidate the mechanism by which the Gly-77 → Ala substitution and the Gly-80b insertion cooperatively increase the protein stability, the crystal structures of G^{80b}-RNase H, A^{77}-RNase H, and A^{77}/G^{80b}-RNase H were determined by X-ray analysis. While A^{77}-RNase H crystallizes in an isomorphous form with the wild-type protein, the crystals of the other two mutant proteins are not isomorphous with that of the wild-type protein, as shown in Table I.

The conformations of the amino-termini, the carboxyl-termini, and some loops in the structures of G^{80b}-RNase H and A^{77}/G^{80b}-RNase H largely deviate from those in the wild-type protein because of the difference in the crystal-packing interactions. However, their overall structures are almost the same. Except for local structural differences, high geometrical similarities are observed between the mutant and wild-type proteins, as determined by least-squares fitting of the Cα positions (root mean square deviations of 0.45 Å for G^{80b}-RNase H and 0.51 Å for A^{77}/G^{80b}-RNase H) for the 92-residue core containing three α-helices (αI, αIV, and αV) and five β-strands (βA–βE). Residues located in αII and αIII were not used for the superposition in order to precisely understand the conformational changes caused by the mutations introduced in these regions. The superposition of the structure of A^{77}-RNase H on the wild-type structure was also performed using the same 92 α-carbons, giving the root mean-square deviation of 0.20 Å. No considerable difference was observed in temperature factors for residues in αII and αIII between the various mutant proteins and the wild-type protein.

Structure of G^{80b}-RNase H. The Gly-80b residue is inserted at the junction between the αII- and αIII-helices. This Gly residue is nicely accommodated between these two helices, as shown in Figures 2 and 3a. The high flexibility of the αIII-

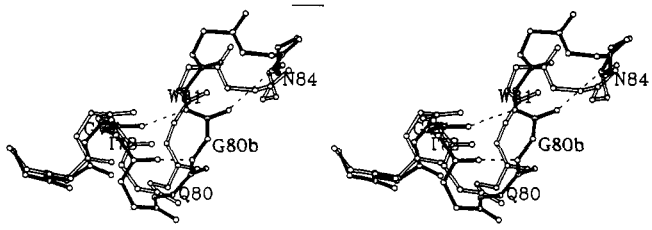

FIGURE 2: Comparison of backbone structures around the junction of the αII- and αIII-helices between the wild-type (open bonds) and G^{80b}-RNase H (solid bonds). Two hydrogen bonds characteristic for the paperclip motif, O78–N80b and O77–N81, as well as another hydrogen bond between O80b and N84, are represented in the structure of G^{80b}-RNase H by dashed lines.

helix (Katayanagi *et al.*, 1992; Yamazaki *et al.*, 1991) probably facilitates this accomodation. The backbone imino nitrogen of Gly-80b forms an intrahelical hydrogen bond with the backbone carbonyl oxygen of Ile-78, at a distance of 3.3 Å. Likewise, the backbone carbonyl oxygen of Gly-80b forms an intrahelical hydrogen bond with the backbone imino nitrogen of Asn-84, at a distance of 3.2 Å. Interestingly, Gly-80b is the C-cap residue of αII as well as the N-cap residue of αIII, where the N-cap and C-cap residues are specified as those located at the N- and C-termini of α-helices, in which the backbone carbonyl or the imino nitrogen forms an intrahelical hydrogen bond with the backbone carbonyl or imino nitrogen of the next turn (Presta & Rose, 1988).

The (ϕ, ψ) values of Gly-80b are (52°, 46°), which are classified as a left-handed helical (αL) conformation. An interesting structural motif called a "paperclip" (Milner-White, 1988), which is common at the C-terminal end of α-helices, is observed around Gly-80b. This structural motif has two characteristic hydrogen bonds, involving residues C-4 and C + 1 and C-3 with the C-cap. In G^{80b}-RNase H, a hydrogen bond is formed between the backbone carbonyl oxygen of Gly-77 and the backbone imino nitrogen of Trp-81, in addition to the intrahelical hydrogen bond between Ile-78 and Gly-80b (Figure 2). The former hydrogen bond is a substitute for the one between Gly-77 and Ile-82 present in the wild-type structure, and therefore it does not increase the total number of hydrogen bonds.

The effect of the Gly-80b insertion on the structure extends over residues that make no direct interaction with Gly-80b. The α-carbon of Ala-93, which is located 16 Å away from the Cα atom of Gly-80b, shifts by 2.2 Å as compared to that in the wild-type protein. Slight shifts in the position and orientation of αIII, relative to those of αII, cause a large shift in the loop containing Ala-93.

Without structural adjustments of residues around Gly-80b, shifts of main-chain atoms in αII and αIII would create some cavities inside the protein molecule. Consequently, the adjacent hydrophobic side-chains have to move and change conformations to fill the cavities. In fact, Leu-111 changes its conformation from (χ^1, χ^2) = (−72°, −178°) to (χ^1, χ^2) = (−102°, 45°). The side-chain conformation of Leu-111 in G^{80b}-RNase H is somewhat strained, while that in the wild-type protein is almost ideal. The movement of the side-chain of Leu-111 accompanies the Cα atom shift of the same residue by 1.0 Å, although this Cα atom is located 11 Å away from the Cα atom of Gly-80b.

The indole ring of Trp-81 is forced to change its orientation to avoid collision with the indole ring of Trp-104 (Figure 3a), and the χ^2 angle changes from −23° to 174°. In spite of this reorientation, the Cδ1 atom of Trp-81 still makes several close contacts with surrounding atoms, such as the Cη2 atom of Trp-85 (3.1 Å) and the Cα atom of Gly-77 (3.3 Å).

Structure of A^{77}-RNase H. The replaced Ala residue in A^{77}-RNase H is located between the two indole rings of Trp-81 and Trp-104 and is completely shielded from the solvent (Figure 3b). The β-carbon of Ala-77 would have excessively close contacts with the indole ring of Trp-104, particularly with the Cη2 atom at the distance of 1.9 Å, if the surrounding residues did not change their conformations. In the refined structure of A^{77}-RNase H, however, the Cβ atom of Ala-77 is pushed away from the indole ring of Trp-104, and this movement accompanies the shift of αII toward the solvent region. As a result of this movement, the Cβ atom of Ala-77 approaches the Cε3 atom of Trp-81 and thereby forces the indole ring of Trp-81 to move away from the Cβ atom of Ala-77. The efficient packing around residue 77 in the wild-type structure is thus deformed, causing a reduction in the protein stability. The consequential access of the Cδ1 atom of Trp-81 to the backbone atoms of Gln-80 and Trp-81 limits the movement of the indole ring of Trp-81, and therefore the close contact with the Cβ atom of Ala-77 is not completely alleviated. This wedged structure of the indole ring of Trp-81 seems to be stressed between the indole ring of Trp-104 and the backbone of Gln-80 and Trp-81. The close contacts of the side-chain atoms of Ala-77 and Trp-81 with the surrounding atoms are listed in Table III.

Structure of A^{77}/G^{80b}-RNase H. The backbone structure of A^{77}/G^{80b}-RNase H largely deviates from that of the wild-type protein and is similar to that of G^{80b}-RNase H (Figure 3c). The paperclip motif is also observed at the same position as that in G^{80b}-RNase H. However, the orientation of the indole ring of Trp-81 is essentially the same as that in the wild-type structure.

Comparison of the contact distances in A^{77}/G^{80b}-RNase H with those in A^{77}-RNase H indicates that the close contacts involving Ala-77 and Trp-81 in A^{77}-RNase H are alleviated in A^{77}/G^{80b}-RNase H (Table III). The introduction of Gly-80b alters the backbone structure, so that the backbone carbonyl group of Gln-80 moves away from the indole ring of Trp-81, as shown in Figure 3c. This movement of the backbone carbonyl group of Gln-80 allows the indole ring to move away from the Cβ atom of Ala-77 and thereby alleviates the close contacts involving Ala-77 and Trp-81. Actually, the spatial position of the Cβ atom of Ala-77 is similar to that of the Cα atom of Gly-77 in the wild-type protein. In addition, since the backbone has more structural freedom than that in A^{77}-RNase H, due to the insertion of Gly-80b, the side-chain atoms have greater freedom of movement to improve the hydrophobic packing interactions around Ala-77.

DISCUSSION

Cooperative Stabilization. Gly is a helix-destabilizing residue, since it lacks a β-carbon and has more flexibility in its backbone conformation than Ala. Therefore, a Gly → Ala substitution at an internal position of an α-helix is expected to enhance protein stability because of a decrease in the entropy of the unfolded state. In addition to the entropic effect, the burial of a methyl group of Ala in a hydrophobic core contributes to protein stability due to the hydrophobic effect. In fact, Matthews *et. al* (1987) stabilized bacteriophage T4 lysozyme by replacing Gly, which is located in an α-helix, with Ala. Serrano *et al.* (1992a) reported that the substitution of Ala for Gly stabilizes an α-helix by approximately 2 kcal/mol when it is located within an α-helix and buried in a hydrophobic core.

However, *E. coli* RNase HI is destabilized by the Gly-77 → Ala substitution, although Ala-77 is located at an internal position in αII and is completely buried within the hydrophobic

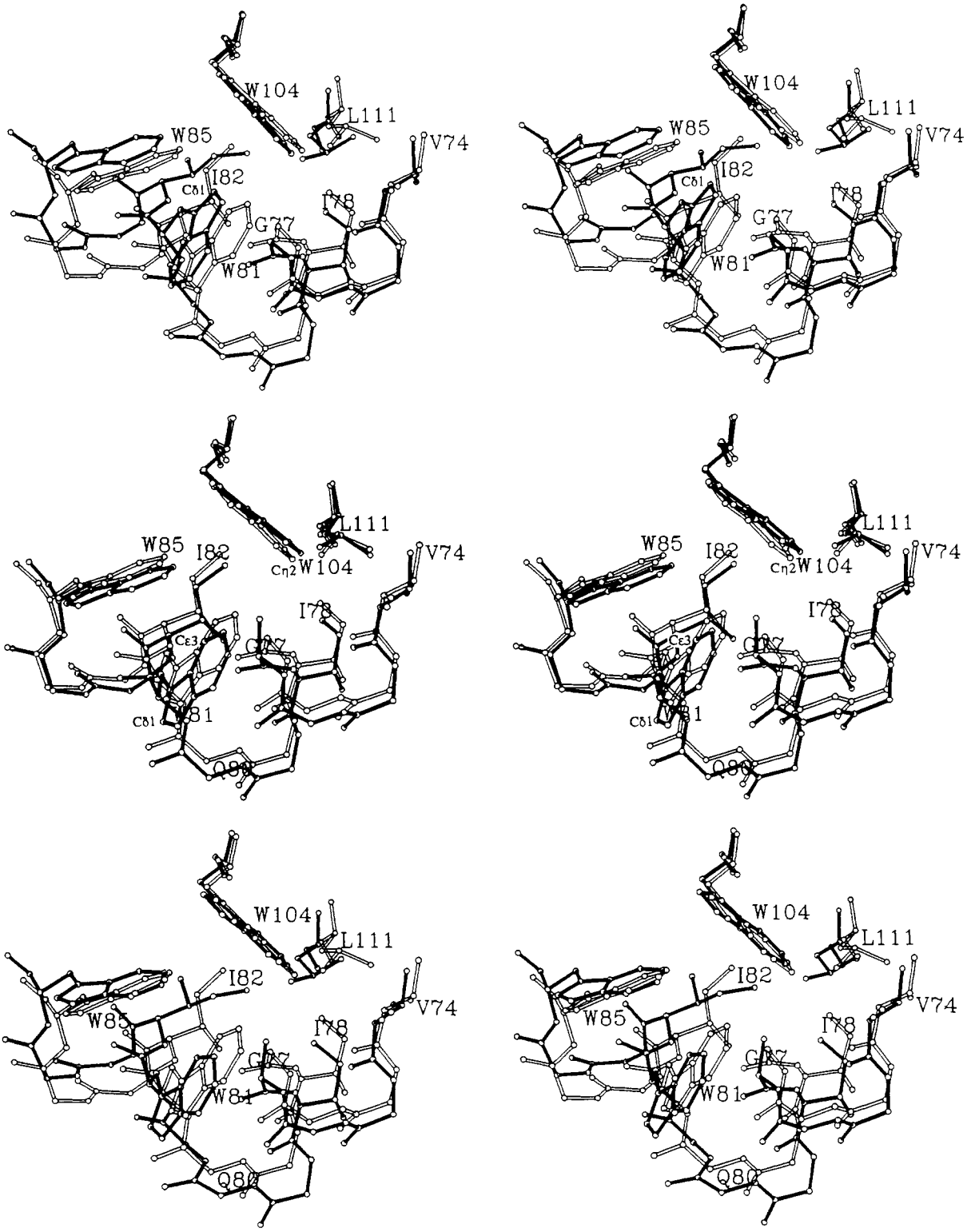

FIGURE 3: Comparison of backbone structures around the mutation sites between the wild-type and mutant proteins. The structures of (a, top) G^{80b}-RNase H, (b, middle) A^{77}-RNase H, and (c, bottom) A^{77}/G^{80b}-RNase H are superimposed on that of the wild-type protein. Open and solid bonds represent the structures of the wild-type and mutant proteins, respectively.

core. The intrinsic helix-stabilization associated with the Gly → Ala replacement is canceled and even destabilized by the deformed packing interactions around Ala-77 and by the stress at the wedged structure of the indole ring of Trp-81. However, the packing of the hydrophobic core is improved and the stress in the wedged indole ring is alleviated by the changes in the backbone structures of the αII- and αIII-helices caused by the insertion of Gly-80b. Consequently, in the structure of A^{77}/G^{80b}-RNase H, the Gly-77 → Ala substituent enhances the protein stability.

The energy of cooperativity, $\Delta\Delta G_{coop}$, was defined by Hurley *et al* (1992): $\Delta\Delta G_{coop} = \Delta\Delta G - \sum\Delta\Delta G_i$, where $\Delta\Delta G_i$ means the value of an individual change. They constructed a series of multiple mutants of bacteriophage T4 lysozyme, which were designed to modify the hydrophobic core packing arrangement. For their mutants in which two amino acid residues were replaced, $\Delta\Delta G_{coop}$ varied from 0.1 to 0.6 kcal/mol. In the present case, $\Delta\Delta G_{coop}$ of A^{77}/G^{80b}-RNase H is 1.3 kcal/mol. Such a large cooperativity can be ascribed to the change in the backbone conformation caused by the

Table III: Contact Distances for Atoms around Residues 77 and 81

close atoms		distance to close atoms (Å) wild type	A^{77}	A^{77}/G^{80b}
(a) Contacts with Cα of Gly-77 (Wild Type) or Cβ of Ala77 (A^{77}-RNase H and A^{77}/G^{80b}-RNase H)				
Trp-104	Cζ2	3.9	3.3	3.8
	Cη2	3.4	3.4	3.3
Trp-81	Cε3	5.0	3.3	3.6
(b) Contacts with Cδ1 of Trp-81				
Gln-80	C	3.7	3.3	4.0
	O	3.5	3.3	4.2
Trp-81	N	3.1	3.2	3.5

insertion of Gly-80b. This result suggests that the amino acid insertions may produce greater cooperativity than the amino acid substitutions.

It is noted that protein stability is a measure of the difference in free energy between the folded and unfolded states. Our rationalizations of relative protein stability focus almost exclusively on a comparison of folded three-dimensional structures and assume that the free energies of the unfolded mutant proteins are similar.

Stabilization by a Paperclip Motif. The Gly-80b insertion is of interest because the crystallographic analysis reveals that it creates a paperclip motif at the C-terminus of the αII-helix (Figure 2). The "paperclip" is a common motif in protein structures, particularly at the C-terminal ends of α-helices (Milner-White, 1988). This motif in an α-helix contains a left-handed helical (αL) residue at the C-cap. Since Cβ atoms in αL-conformations cause steric hindrance with carbonyl oxygen atoms in the same residues (Kimura *et al.*, 1992b; Ishikawa *et al.*, 1993a), non-glycine residues are not suitable for the C-cap. In fact, mutational and theoretical studies (Serrano *et al.*, 1992b) strongly support the preference for Gly with an αL-conformation at the C-cap of an α-helix. Our result that the replacement of Gly-80b by Ala decreases the protein stability (Table II) also supports this preference. The difference in the ΔG values (0.9 kcal/mol) between G^{80b}-RNase H and A^{80b}-RNase H is comparable to that (1.3 kcal/mol) predicted for the relative stabilizing effects of Gly *vs* Ala with an αL conformation at the C-cap (Serrano *et al.*, 1992b).

Two hydrogen bonds are additionally formed by the insertion of Gly-80b (Figure 2). One is characteristic for the "paperclip", and another is located at the N-terminal end of αIII. Shirley *et al.* (1992) reported that an intramolecular hydrogen bond contributes 1.3 (±0.6) kcal/mol to the conformational stability of globular proteins. Therefore, the creation of two hydrogen bonds involving Gly-80b is expected to increase the protein stability by 2.6(±1.2) kcal/mol. However, the protein stability is enhanced by only 0.4 kcal/mol. Most of the increase in stability is probably offset by the deformation of the excellent packing in the wild-type protein. The stabilization mechanism of the Gly-80b insertion is quite complicated, but it is attributable to the creation of the "paperclip" to some extent.

Decrease in Enzymatic Activity by the Insertion of Gly-80b. The observed kinetic parameters indicate that the insertion of Gly-80b seriously affects substrate binding. However, there is no observable difference in the structures of the catalytic residues between the mutant and the wild-type proteins.

In *E. coli* RNase HI, the αIII-helix directly follows the αII-helix with a kink. Such a kink allows α-helices to move independently, and it sometimes plays an important role in protein functions, such as facilitating an induced fit (Barlow & Thornton, 1988). A good example is seen in the kinked helix comprising residues 142–168 in adenylate kinase (Sachsenheimer & Schulz, 1977). This helix is one of the structures that shifts as part of the conformational change that occurs upon substrate binding (induced fit). The conformational shift occurs only in the N-terminal region of the kinked site, while the C-terminal region is not affected by substrate binding.

The αII-helix in *E. coli* RNase HI is not likely to undergo a conformational change, since the hydrophobic residues in this helix are packed very well in the hydrophobic core. Most of the residues in the basic protrusion, which ranges over the αIII-helix and the following loop (Figure 1), are hydrophilic, and hence these regions may adopt conformations to allow the enzyme to effectively bind the substrate. The substrate titration experiment using heteronuclear two-dimensional NMR supports this possibility, because the chemical shifts of His-83 and Asn-84 in the αIII-helix significantly change upon binding of the substrate (Nakamura *et al.*, 1991).

A striking feature in the paperclip motif observed in G^{80b}-RNase H is that another α-helix, αIII, follows the structural motif. Gly-80b seems to tighten the junction between αII and αIII by forming the two intrahelical hydrogen bonds with Ile-78 and Asn-84. The rigidity at this junction may impede the induced fit of the substrate-binding site.

Alternatively, slight shifts in the positions of the αII- and αIII-helices, due to the insertion of Gly-80b, may not allow the substrate to bind properly, so that the cleavage sites of the substrate effectively contact the active site of the enzyme. According to the model for the enzyme–substrate complex (Nakamura *et al.*, 1991), αII and αIII are involved in the substrate-binding site of the enzyme. Mutational studies support this model (Kanaya *et al.*, 1991b).

Stabilization Mechanism of R_5-RNase H. In this report, we have shown that the simultaneous introductions of the Gly-77 → Ala substitution and the Gly-80b insertion increase the protein stability by 2.7 °C in T_m and 0.8 kcal/mol in ΔG. This increase in the stability of the double mutant happens to be the same as that of R_5-RNase H. However, the stability of the triple-mutant protein (LAG-RNase H), in which the Val-74 → Leu substitution is introduced in addition to these two mutations, increased by 7.4 °C in T_m and 2.2 kcal/mol in ΔG at pH 3.0 (S. Kanaya, personal communication). Other mutations besides these three mutations must therefore decrease the protein stability, so that they compensate for an increase in the stability of LAG-RNase H. In fact, the Ile-78 → Phe substitution destabilizes *E. coli* RNase HI by 5.6 °C in T_m and 1.7 kcal/mol in ΔG (S. Kimura & S. Kanaya, personal communication). If the effects of these mutations were additive, the quadruplex mutations would increase the protein stability by only 0.5 kcal/mol, which is comparable to the increase in the stability of R_5-RNase H.

The crystal structure of LAG-RNase H is almost identical to that of A^{77}/G^{80b}-RNase H, except for the side-chain of residue 74 (K. Ishikawa, personal communication). This mutant protein is, therefore, stabilized by the hydrophobic effect of the Val-74 → Leu substitution (Ishikawa *et al.*, submitted), the α-helix stabilizing effect of the Gly-77 → Ala substitution, and the creation of the "paperclip" by the Gly-80b insertion. Destabilization by the Ile-78 → Phe substitution probably arises from the deformed packing interactions around the replaced Phe residue.

In *T. thermophilus* RNase H, Gly-80b also adopts a left-handed conformation, but the hydrogen bonds characteristic for the "paperclip" are not observed. This is probably because some hydrophobic residues are substituted in *T. thermophilus* RNase H, like the Ile-78 → Phe replacement. However, the

αII-helix is definitely stabilized relative to that in *E. coli* RNase HI by the Gly-77 → Ala replacement.

Conclusion. Earlier studies on the insertion of a single Gly or Ala residue at 20 randomly selected positions in staphylococcal nuclease (Sondek & Shortle, 1990), or at three different locations in an α-helix of T4 lysozyme (Heinz *et al.*, 1993), indicate that insertions are generally tolerated but drastically reduce the protein stability.

Our result, however, demonstrates that protein stability can be increased when the insertion of a Gly residue at the C-terminal end of an α-helix creates a paperclip motif. The C-cap residue of αII is changed from Gln-80 in the wild-type protein to Gly-80b in G^{80b}-RNase H, resulting in the extension of αII. This extension is reasonable, since Gly is extremely preferred at the C-cap, while Gln is favored at the C-1 position to the C-cap (Richardson & Richardson, 1988). It is also natural that Gly-80b appears as an α_L-conformation, since about 50% of α-helices have C-cap residues with positive ϕ and ψ angles (Serrano *et al.*, 1992b).

To date, the following strategies have been shown to be usable to enhance the stability of α-helices: (1) replacement of Gly located at an internal position of an α-helix with Ala (Hecht *et al*, 1986; Matthews *et al.*, 1987), (2) formation of salt bridges between side-chains at positions (i, $i + 3$), or (i, $i + 4$) (Serrano *et al.*, 1990; Sali *et al.*, 1991; Dao-pin *et al.*, 1991), and (3) stabilization of the dipole moment (Sali *et al.*, 1988; Nicholson *et al.*, 1988). The creation of a paperclip motif may now be considered as a strategy to enhance the stability of α-helices. Alternatively, an insertion of Gly may be used as a general strategy to relax proteins from local stress in a wedged structure, as seen for A^{77}/G^{80b}-RNase H.

ACKNOWLEDGMENT

We thank Dr. K. Katayanagi for providing the coordinates of *E. coli* RNase HI. We also thank Drs. Y. Oda, M. Oobatake, and T. Miyazawa for helpful discussions.

REFERENCES

Barlow, D. J., & Thornton, J. M. (1988) *J. Mol. Biol. 201*, 601–619.

Becktel, W. J., & Schellman, J. A. (1987) *Biopolymers 26*, 1859–1877.

Branden, C., & Tooze, J. (1991) in *Introduction to Protein Structure*, pp 247–268, Gerland Publishing, Inc., New York.

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

Crouch, R. J., & Dirksen, M.-L. (1982) in *Nuclease* (Linn, S. M., & Roberts, R. J., Eds). pp 211–241, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.

Crowther, R. A., & Blow, D. M. (1967) *Acta Crystallogr. 23*, 544–548.

Dao-pin, S., Sauer, U., Nicholson, H., & Matthews, B. W. (1991) *Biochemistry 30*, 7142–7153.

Hecht, M. H., Sturtevant, J. M., & Sauer, R. T. (1986) *Proteins: Struct., Funct., Genet. 1*, 43–46.

Heinz, D. W., Baase, W. A., Dahlquist, F. W., & Matthews, B. W. (1993) *Nature 361*, 561–564.

Hendrickson, W. A., & Konnert, J. H. (1980) in *Computing in Crystallography* (Diamond R., Ramaseshan, S., & Venkatesan, K., Eds.) pp 13.01–13.23, Indian Academy of Science, International Union of Crystallography, Banglore, India.

Huber, R., & Schneider, M. (1985) *J. Appl. Crystallogr. 18*, 165–169.

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.

Ishikawa, K., Nakamura, H., Morikawa, K., & Kanaya, S. (1993c) *Biochemistry* (in press).

Itaya, M., & Kondo, K. (1991) *Nucleic Acids Res. 19*, 4443–4449.

Jones, T. A. (1978) *J. Appl. Crystallogr. 11*, 268–272.

Kanaya, S., & Couch, R. J. (1983) *J. Biol. Chem. 258*, 1276–1281.

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. (1991a) *J. Biol. Chem. 266*, 6038–6044.

Kanaya, S., Katsuda-Nakai, C., & Ikehara, M. (1991b) *J. Biol. Chem. 266*, 11621–11627.

Katayanagi, K., Miyagawa, M., Matsushima, M., Ishikawa, M., Kanaya, S., Ikehara, M., Matsuzuki, T., & Morikawa, K. (1990) *Nature 347*, 306–309.

Katayanagi, K., Miyagawa, M., Matsushima, M., Ishikawa, M., Kanaya, S., Nakamura, H., Ikehara, M., Matsuzuki, T., & Morikawa, K. (1992) *J. Mol. Biol. 223*, 1029–1052.

Kimura, S., Nakamura, H., Hashimoto, T., Oobatake, M., & Kanaya, S. (1992a) *J. Biol. Chem. 267*, 21535–21542.

Kimura, S., Kanaya, S., & Nakamura, H. (1992b) *J. Biol. Chem. 267*, 22014–22017.

Matthews, B. W., Nicholson, H., & Becktel, W. J. (1987) *Proc. Natl. Acad. Sci. U.S.A. 84*, 6663–6667.

Milner-White, E. J. (1988) *J. Mol. Biol. 199*, 503–511.

Nakamura, H., Oda, Y., Iwai, S., Inoue, H., Ohtsuka, E., Kanaya, S., Kimura, S., Katsuda, C., Katayanagi, K., Morikawa, K., Miyashiro, H., & Ikehara, M. (1991) *Proc. Natl. Acad. Sci. U.S.A. 88*, 11535–11539.

Nicholson, H., Becktel, W. J., & Matthews, B. W. (1988) *Nature 336*, 651–656.

Presta, L. G., & Rose, G. D. (1988) *Science 240*, 1632–1641.

Richardson, J. S., & Richardson, D. C. (1988) *Science 240*, 1648–1652.

Sachsenheimer, W., & Schulz, G. E. (1977) *J. Mol. Biol. 114*, 23–36.

Sali, D., Bycroft, M., & Fersht, A. R. (1988) *Nature 335*, 740–743.

Sali, D., Bycroft, M., & Fersht, A. R. (1991) *J. Mol. Biol. 220*, 779–788.

Serrano, L., Horovitz, A., Avron, B., Bycroft, M., & Fersht, A. R. (1990) *Biochemistry 29*, 9343–9352.

Serrano, L., Neira, J. L., Sancho, J., & Fersht, A. R. (1992a) *Nature 356*, 453–455.

Serrano, L., Sancho, J., Hirshberg, M., & Fersht, A. R. (1992b) *J. Mol. Biol. 227*, 544–559.

Shirley, B. A., Stanssens, P., Hahn, U., & Pace, C. N. (1992) *Biochemistry 31*, 725–732.

Sondek, J., & Shortle, D. (1990) *Proteins: Struct., Funct., Genet. 7*, 299–305.

Steigemann, W. (1974) Ph.D. Thesis, Technical University of Munich.

Yamazaki, T., Yoshida, M., Kanaya, S., Nakamura, H., & Nagayama, K. (1991) *Biochemistry 30*, 6036–6047.

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

Yutani, K., Hayashi, S., Sugisaki, Y., & Ogasawara, K. (1991) *Proteins: Struct., Funct., Genet. 9*, 90–98.