Conformational Stabilities of *Escherichia coli* RNase HI Variants with a Series of Amino Acid Substitutions at a Cavity within the Hydrophobic Core*

(Received for publication, March 14, 1997, and in revised form, May 5, 1997)

**Atsuko Akasako‡, Mitsuru Haruki§, Motohisa Oobatake‡, and Shigenori Kanaya§¶**

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

---

*Escherichia coli* ribonuclease HI has a cavity within the hydrophobic core. Two core residues, Ala52 and Val74, resided at both ends of this cavity. We have constructed a series of single mutant proteins at Ala52 and double mutant proteins, in which Ala52 was replaced by Gly, Val, Ile, Leu, or Phe, and Val74 was replaced by Ala or Leu. All of these mutant proteins, except for A52W, A52R, and A52G/V74A, were overproduced and purified. Measurement of the thermal denaturations of the proteins at pH 3.2 by CD suggests that the cavity is large enough to accommodate three methyl or methylene groups without creating serious strains. A correlation was observed between the protein stability and the hydrophobicity of the substituted residue. As a result, a number of the mutant proteins were more stable than the wild-type protein. The stabilities of the mutant proteins with charged or extremely bulky residues at the cavity were lower than those expected from the hydrophobicities of the substituted residues, suggesting that considerable strains are created at the mutation sites in these mutant proteins. However, examination of the far- and near-UV CD spectra and the enzymatic activities suggest that all of the mutant proteins have structures similar to that of the wild-type protein. These results suggest that the cavity in the hydrophobic core of *E. coli* RNase HI is conformationally fairly stable. This may be the reason why the cavity-filling mutations effectively increase the thermal stability of this protein.

In the interiors of globular proteins, and especially in the protein cores, hydrophobic side chains are generally well packed. Since the formation of a protein core governs the entire folding process of a protein, and because general methods for designing amino acid sequences that fold into the desired core structures have not been established, efforts have been made to evaluate the role of core residues in the stability and function of natural proteins by site-directed mutagenesis (1–21), random mutagenesis (22–24), and theoretical approaches (25–29). The protein stability decreases as the volume of such a cavity increases. In this case, the cost of reduced stability, because they serve to reduce the strains caused by dense packing of the hydrophobic side chains, and to provide conformational flexibility to proteins. Therefore, it seems likely that the tolerance of proteins to core mutations increases as the volume of the naturally existing cavity at the core increases.

*Escherichia coli* RNase HI, which hydrolyzes only the RNA strand of RNA/DNA hybrids, is a small globular protein with 155 amino acid residues. We have used this protein for studies of protein stability (14, 35–43) and folding (44, 45). This protein is suitable for these studies for the following reasons: (i) highly refined coordinates of this protein are available (46), (ii) an overproduction system for this protein is available (36), and (iii) this protein reversibly unfolds in a single cooperative fashion with thermal and chemical denaturations (38).

In the hydrophobic core of this protein, a cavity exists, which is not occupied by water molecules. We have previously shown that cavity-filling mutations, such as Val74 → Leu and Ile, enhanced the protein stability by roughly 1 kcal/mol in *Ag* (14). We also found by random mutagenesis experiments that the mutation of Ala52 → Val also enhanced the protein stability by 1.7 kcal/mol, probably due to the reduction in the cavity volume around the mutation site (42). These results allowed us to propose that the replacement of a hydrophobic residue facing the cavity with a bulkier and more hydrophobic residue is a general method to increase protein stability. However, proteins are not always stabilized by this method. Karpusas et al. (1) found two large cavities in T4 lysozyme and designed two mutations (Leu139 → Phe and Ala129 → Val) to fill them. Eij-sink et al. (8) introduced various mutations to fill the cavities in the neutral protease from *Bacillus stearothermophilus*. None of

---

*The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

‡ Present address: Biomolecular Engineering Research Institute, 6-2-3, Furuedai, Suita, Osaka 565, Japan.

§ Present address: Dept. of Material and Life Science, Graduate School of Engineering, Osaka University, 2-1, Yamadaoka, Suita, Osaka 565, Japan.

¶ To whom correspondence should be addressed: Dept. of Material and Life Science, Graduate School of Engineering, Osaka University, 2-1, Yamadaoka, Suita, Osaka 565, Japan. Tel./Fax: 81-6-879-7938; E-mail: kanaya@chem.eng.osaka-u.ac.jp.

---

1 The abbreviation used is: RNase HI, ribonuclease HI.
Stabilities of E. coli RNase HI Mutants at the Cavity

These mutagenesis results in a significant increase in the protein stability. It was shown for the T4 lysozyme mutants that the hydrophobic effects gained by the mutation were canceled by the hindrance. The thermal stabilities of a series of the mutant proteins suggest that this cavity is more effectively filled by the methyl or methylene groups without causing significant steric hindrance. The thermal stabilities of a series of the mutant proteins at Ala52 and Val74, and have analyzed their conformational stabilities and enzymatic activities.

Here we report that the size of the cavity in the hydrophobic core of E. coli RNase HI is large enough to introduce three methyl or methylene groups without causing significant steric hindrance. The thermal stabilities of a series of the mutant proteins suggest that this cavity is more effectively filled by the double mutations (Ala52→ Val and Val74→ Leu) than by the single mutation (Ala52→ Ile). We also report that this cavity seems to be conformationally fairly stable and the protein is tolerant to mutations that seriously affect the hydrophobicity and the packing of the cavity residues. Thus, the structure of this protein was not severely damaged when the cavity residues were replaced by either aromatic or charged residues, although these mutations greatly reduced the protein stability and the enzymatic activity.

EXPERIMENTAL PROCEDURES

Materials—The wild-type E. coli RNase HI protein (36) and the mutant protein A52V (42), in which Ala52 was replaced by Val, were prepared previously. Restriction enzymes and modifying enzymes for recombinant DNA technology were from Takara Shuzo Co., Ltd. Other chemicals were of reagent grade.

Cells and Plasmids—The plasmids pJAL600 (36), pJAL52V (42), and pJAL74L and pJAL74A (14) for the overproduction of the wild-type protein, and the mutant proteins A52V, V74L, and V74A, respectively, were constructed previously. These plasmids bear the wild-type or the mutant gene under the control of the bacteriophage λ promoters P0 and P3, the class 5 gene, and the bacteriophage fd transcription terminator. Competent cells of E. coli HB101(×′, RΔ92, recA13, ara-14, proA2, lac Y1, gal K2, rps L20(str), xyl-5, mtl-1, sup E44, leuB6, thi-1) were from Takara Shuzo Co., Ltd. E. coli HB101 transformants with pJAL600 derivatives were grown in Luria-Bertani (LB) medium (47) containing 100 mg/liter ampicillin at 30 °C.

Mutagenesis—Alteration of the rhnA gene was carried out by site-directed mutagenesis using polymerase chain reaction, as described previously (36). The DNA oligomers used as the 5′ - and 3′-mutagenic primers (30–40 bases long) were synthesized by Sawady Technology Co., Ltd. These DNA oligomers were designed such that the codon for Ala52 was changed from GCT to ATC for Ile, CTG for Leu, TGT for Cys, ATG for Met, TTC for Phe, ACT for Thr, CAG for Gin, GAA for Glu, CCG for Pro, TCT for Ser, AAC for Asn, GAT for Asp, TAC for Tyr, GTT for Gly, CAT for His, AAA for Lys, TGG for Trp, and CTT for Arg. For the construction of the single mutant proteins at position 52, the wild-type rhnA gene in the plasmid pJAL600 was used as a template. For the construction of the double mutant proteins at positions 52 and 74, the mutant rhnA gene in either plasmid pJAL74A or pJAL74L was used as a template. The resultant plasmids for the overproduction of the mutant proteins were designated as either pJAL52X or pJAL52X74X, in which X and X′ represent the amino acid residues (one-letter notation) substituted for Ala52 and Val74, respectively. Likewise, the mutant proteins were designated as either A52X or A52X/V74X.

Overproduction and Purification—The mutant proteins were overproduced in E. coli HB101 cells harboring the pJAL600 plasmid deriv-
Stabilities of E. coli RNase HI Mutants at the Cavity

Yield represents the amount of the protein purified from 1-liter culture. The enzymatic activity was determined at 30 °C for 15 min in 10 mM Tris-HCl (pH 8.0) containing 10 mM MgCl₂, 50 mM NaCl, 1 mM 2-mercaptoethanol, and 10 μg/ml bovine serum albumin, by using an M13-RNA/DNA hybrid as a substrate. Errors, which represent the 67% confidence limits, are within 30% of the values reported. Thermal denaturation curves of the mutant proteins were measured at pH 3.2, as described under “Experimental Procedures.” ∆Tm is the change in the melting temperature, Tm, relative to that of the wild-type protein, which is 53.0 °C. ΔHm is the enthalpy change of unfolding at Tm calculated by van’t Hoff analysis. The change in the free energy of unfolding of the mutant protein relative to that of the wild-type protein (ΔGm), was estimated by the relationship given by Becktel and Shellenberg (51): 

\[ \Delta G_m = \Delta G_w - \Delta G_m \text{ (wild-type)} \]

The ΔGm (wild-type) value of 0.304 kcal/(mol · K), which was previously determined at pH 3.0 (38), was used for the calculation of the ΔGm values. Errors, which represent the 67% confidence limits, are within 0.3 °C in Tm.

| Mutant | Yield (mg/liter) | % | ΔTm °C | ΔGm kcal/mol | ΔHm kcal/mol |
|--------|------------------|---|--------|---------------|--------------|
| A52I   | 52               | 38 | 6.2    | 1.58          | 84.8         |
| A52V   | 30               | 69 | 5.5    | 1.67          | 128.0        |
| A52N   | 47               | 112| 2.5    | 0.76          | 192.2        |
| A52C   | 47               | 112| 2.5    | 0.76          | 192.2        |
| A52M   | 32               | 97 | 1.6    | 0.49          | 111.9        |
| WT     | 50               | 100| 0      | 99.2          |
| A52E   | 10               | 0.3| -1.5   | 0.46          | 89.0         |
| A52T   | 48               | 102| -2.7   | -0.82         | 107.9        |
| A52Q   | 47               | 5  | -3.9   | -1.19         | 90.3         |
| A52E   | 6                | 2  | -5.0   | -1.52         | 100.6        |
| A52P   | 62               | 107| -5.4   | -1.64         | 93.0         |
| A52S   | 27               | 97 | -5.8   | 1.76          | 117.6        |
| A52N   | 28               | 52 | -5.9   | 1.79          | 95.6         |
| A52D   | 10               | 27 | -6.1   | 1.85          | 103.5        |
| A52Y   | 3                | 0.1| -7.6   | -2.31         | 70.8         |
| A52G   | 27               | 67 | -8.9   | -2.71         | 91.0         |
| A52H   | 32               | 0.4| -11.8  | -3.59         | 84.7         |
| A52K   | 5                | 10 | -19.5  | -5.93         | 70.8         |

FIG. 2. CD spectra of the wild-type and single mutant proteins. The CD spectra of the wild-type and all single mutant proteins were measured as described under “Experimental Procedures.” a, the far-UV CD spectra of the wild-type protein (thin line) and the mutant protein A52K (thin line) are shown. The spectra of all of the other mutant proteins are almost identical to that of the wild-type protein. b, the near-UV CD spectra of the wild-type protein (thin line), the A52I mutant (thin line), and the A52F mutant (broken line) are shown as representatives of types A, B, and C, respectively.

The simultaneous introduction of the two cavity-creating mutations probably dramatically destabilizes the protein.

Stabilities and Activities of Single Mutant Proteins—The far-UV CD spectra of all of the single mutant proteins are basically the same as that of the wild-type protein (Fig. 2a). In contrast, the near-UV CD spectra of these mutant proteins differed from one another. They were roughly classified into three groups (types A-C), based on the CD values at 260–280 nm (Fig. 2b). The wild-type protein and the mutant proteins A52C, A52P, and A52G gave the type A spectrum. In the type B and type C spectra, the CD values at 260–280 nm increased as compared with those in the type A spectrum. The extent of this increase is relatively small for the type B spectrum and is relatively large for the type C spectrum. The mutant proteins A52I, A52L, A52V, A52N, A52D, A52Q, A52E, A52S, and A52T gave the type B spectrum. The mutant proteins A52F, A52Y, A52M, A52H, and A52K gave the type C spectrum. Thus, the extent of the increase in the CD values at 260–280 nm seems to be correlated with the volume of the side chain of the replaced residue. Apparently, the shape of the near-UV CD spectrum of the protein, especially that around 290 nm, is not markedly changed by the mutations. These results suggest that the mutations of Ala52 cause a local conformational change, but only to a small extent, even when an aromatic or ionized group is introduced into the cavity by the mutation.

The thermal denaturation curves of the wild-type and mutant proteins are shown in Fig. 3. All of the mutant proteins reversibly unfolded in a single cooperative manner with thermal denaturation. The parameters characterizing the thermal denaturation of the mutant proteins, which were determined based on the assumption that these proteins unfold in a two-state mechanism, are summarized in Table I. Of the 17 mutant proteins located in the αI helix, Leu107 is located in the β strand, Leu is located in the αIV helix, and Trp118 is located in the βE strand. All of these residues are almost fully buried inside the protein molecule. The hydroxyl group of Thr23 does not face the cavity but forms the hydrogen bond with the hydroxyl group of Thr23 (46). For a comprehensive analysis of the effect of the mutation at the cavity on the protein stability, we have constructed a series of single mutant proteins, in which Ala52 is replaced by the 19 other amino acid residues. Position 52 was chosen as the site for the introduction of the series of mutations, because Ala52 is the smallest residue among those forming the cavity and therefore it is possible to introduce into the cavity a variety of side chains, that greatly differ in size and hydrophobicity. In addition, we have constructed 10 double mutant proteins, in which Ala52 is replaced by Gly, Val, Leu, Ile, or Phe and Val74 is replaced by Ala or Leu. These mutant proteins were constructed to estimate the number of methylene groups that could be introduced into the cavity without creating serious strains, and to examine whether effects other than hydrophobicity, such as packing, at the cavity contribute to the protein stability.

Overproduction and Purification—All of the single mutant proteins, except for A52W and A52R, were overproduced and purified in an amount sufficient for biochemical characterization. The mutant proteins A52W and A52R could not be purified, because of the extremely low production levels in cells (data not shown). The mutations of Ala52→Trp and Arg may dramatically destabilize the protein and thereby increase the susceptibility to proteolytic degradation, probably because these residues are too large to fill the cavity without causing significant steric hindrance. The ionized group in the Arg side chain may also contribute to altering the protein conformation, because it is unlikely that an ionized group remains alone without a hydrogen-bonding or ion-pair partner within the hydrophobic core of a protein. In fact, the cellular production levels of the mutant proteins A52D, A52E, A52K, A52Y, and A52F, in which Ala52 is replaced by ion or aromatic residues, were considerably lower than that of the wild-type protein (data not shown). In contrast, the cellular production levels of the other mutant proteins were similar to that of the wild-type protein. Consequently, the yields of the mutant proteins A52D, A52E, A52K, A52Y, and A52F, which are prepared from 1-liter cultures were 3–10 mg, and those of the other mutant proteins were 27–62 mg (Table I). Likewise, of the 10 double mutant proteins, only the mutant protein A52G/V74A could not be purified, because of the extremely low production level in cells (data not shown).
proteins, those in which Ala52 is replaced by more hydrophobic aliphatic or sulfur-containing residues were more stable than the wild-type protein by 0.5–1.9 kcal/mol in AG. In contrast, the other mutant proteins were less stable than the wild-type protein by 0.5–5.9 kcal/mol. We have analyzed the thermal stabilities of these mutant proteins at pH 5.5 in the presence of 1 M guanidine hydrochloride, conditions under which the wild-type protein reversibly unfolds (38). However, except for the apparent $T_m$ values, the thermodynamic parameters could not be determined, because the thermal denaturations of all mutant proteins were not reversible under these conditions. The $T_m$ values of the mutant proteins relative to that of the wild-type protein ($\Delta T_m$) at pH 5.5 were comparable with those determined at pH 3.2, except for those of the mutant proteins A52D and A52E (data not shown). The stabilities of these mutant proteins relative to that of the wild-type protein at pH 5.5 were much lower than those at pH 3.2, probably because Asp and Glu are not ionized at pH 3.2, but are at least partially ionized at pH 5.5. Because the mutant proteins unfold irreversibly at pH 5.5, only the thermodynamic values at pH 3.2, at which all the mutant proteins unfold reversibly, are discussed in this report.

The enzymatic activities of the mutant proteins varied from 0.1 to 112% of that of the wild-type protein (Table I). To determine whether the mutation affects the catalytic efficiency or substrate binding, the kinetic parameters were determined for some of the mutant proteins with poor enzymatic activities. The results are summarized in Table II. The $K_m$ values of these mutant proteins increased only by at most 2.5-fold, whereas their $V_{max}$ values decreased by 12–44-fold, as compared with those of the wild-type protein A52X/V74L (Fig. 4). These results suggest that the tolerance of the protein conformation to the cavity mutations increases as the volume of the cavity increases. The enzymatic activities are summarized in Table III. The enzymatic activities of the double mutant proteins A52X/V74L were always lower than those of the corresponding mutants A52X/V74A. In addition, some of the double mutant proteins A52X/V74A, such as A52I/V74A and A52F/V74A, were more active than the corresponding single mutant proteins at position 52. Therefore, the increase in the cavity volume apparently contributes to reducing the strains caused by the cavity filling mutations that are unfavorable for the enzymatic activity.

The parameters characterizing the thermal denaturation of the double mutant proteins are also summarized in Table III. Among these parameters, $\Delta \Delta G_m$, instead of $\Delta G_m$, reflects the effect of the mutation at Ala52 on the stability of the mutant protein V74X', and $\Delta \Delta G_m$, which is calculated as $\Delta G_m'(A52X/V74X') - \Delta G_m(A52X)$, reflects the difference between the effect of the mutation at Ala52 on the stability of the wild-type protein ($\Delta \Delta G_m(A52X)$) and that on the stability of the mutant protein V74X' ($\Delta G_m'(A52X/V74X')$). The double mutant proteins A52X/V74X' should give positive $\Delta \Delta G_m$ values, if the mutation at Ala52 creates strains at the cavity of the wild-type protein, that are unfavorable for the protein stability, and if the mutation at Val74 contributes to eliminating these strains. In contrast, they should give negative $\Delta \Delta G_m$ values, if the mutation at Val74 contributes to creating additional strains. As shown in Table III, all of the double mutant proteins A52X/V74A gave positive $\Delta \Delta G_m$ values. Whereas, all of the double mutant proteins A52X/V74L, in which Ala52 is replaced by bulkier hydrophobic residues, gave negative $\Delta \Delta G_m$ values. These results suggest that the mutation of Ala52 to Val, Leu, Ile, or Phe creates strains within the cavity of the wild-type protein due to a collision between the substituted residue and the surrounding residues. These strains must be at least partially eliminated when Val74 is replaced by Ala, because of the

![Fig. 3. Thermal denaturation curves of the wild-type and single mutant proteins.](Image)

**Fig. 3.** Thermal denaturation curves of the wild-type and single mutant proteins. The apparent fraction of unfolded protein was determined as a function of temperature. The thermal denaturation curves of all mutant proteins were determined at pH 3.2 by monitoring the change in the CD value at 220 nm, as described under “Experimental Procedures.” The curves of the wild-type (○), A52I (●), A52C (○), A52Q (■), A52S (□), A52G (▲), A52H (△), and A52K (●) proteins are shown as representatives.

**Table II**

*Kinetic parameters of RNase HI mutants*

| Mutant   | $K_m$ (μM) | $V_{max}$ (units/mg) | Relative $V_{max}$ |
|----------|------------|----------------------|--------------------|
| WT       | 0.12       | 12.4                 | 1.0                |
| A52K     | 0.23       | 0.90                 | 0.07               |
| A52L     | 0.22       | 0.28                 | 0.02               |
| A52E     | 0.30       | 1.04                 | 0.08               |

![Fig. 4. CD spectra of the double mutant proteins.](Image)

**Fig. 4.** CD spectra of the double mutant proteins. The near-UV CD spectra of the double mutant proteins A52X/V74A (a) and A52X/V74L (b) are shown in comparison with those of the single mutant proteins V74A and V74L (thick line), respectively. X represents Val (thin line), Ile (thick broken line), Leu (thin broken line), and Phe (dotted line). All spectra were measured as described under “Experimental Procedures.”

**Stabilities of Double Mutant Proteins**—The cavity volumes of the mutant proteins V74A and V74L are larger and smaller than that of the wild-type protein, respectively (14). Accordingly, the mutant protein V74A must be much more tolerant of mutations at Ala52 to bulkier hydrophobic residues than the V74L protein. The conformational changes of the proteins caused by the mutations at Ala52 were analyzed by CD. Comparison of the near-UV CD spectra of the double mutant proteins with that of the parent single mutant protein (V74A or V74L) revealed that the changes in the CD spectra by the mutations at position 52 are much smaller for V74A than for V74L (Fig. 4). These results suggest that the tolerance of the protein conformation to the cavity mutations increases as the

![Table II](Image)
The enzymatic activities and thermal stabilities of the mutant proteins were determined as described in the legend for Table I. $\Delta G_m$ is the change in the free energy of unfolding of the double mutant protein $A52X/V74X$ relative to that of the single mutant protein $V74X$, instead of the wild-type protein. $\Delta \Delta G_m$ is calculated as $\Delta G_m$ (wild-type) $- \Delta G_m$ (mutant), where $\Delta \Delta G_m$ represents the $\Delta G_m$ value for $A52X/V74X$ and the $\Delta \Delta G_m$ (wild-type) value represents the $\Delta G_m$ value for $A52X$.

| Mutant | Relative activity | $\Delta T_m$ | $\Delta G_m$ | $\Delta \Delta G_m$ | $\Delta \Delta \Delta G_m$ | $\Delta \Delta \Delta \Delta G_m$ |
|--------|------------------|-------------|-------------|------------------|--------------------------|-----------------------------|
| WT     | 100              | 0           | 0           | 99.2             |                          |                             |
| V74A   | 100              | -7.6        | -2.31       | 0                | 101.1                    |                             |
| A52/V74A | 62                | -1.7        | -0.52       | 1.79             | 0.12                     | 106.9                       |
| A52L/V74A | 2.6             | 1.2         | 0.36        | 2.67             | 1.36                     | 105.2                       |
| A52I/V74A | 96                | 3.9         | 1.19        | 3.50             | 1.62                     | 101.9                       |
| A52F/V74A | 9                 | -4.6        | -1.40       | 0.91             | 1.37                     | 85.8                        |
| V74L   | 110              | 3.7         | 1.12        | 0                | 118.6                    |                             |
| A52G/V74L | 95                | -4.6        | -1.40       | -2.52            | 0.19                     | 97.8                        |
| A52V/V74L | 16                | 8.0         | 2.43        | 1.31             | -0.36                    | 125.5                       |
| A52L/V74L | 2.4              | 4.9         | 1.49        | 0.37             | -0.94                    | 118.1                       |
| A52I/V74L | 3.8              | 3.9         | 1.19        | 0.07             | -1.81                    | 122.3                       |
| A52F/V74L | 1.9              | 0.1         | 0.03        | 1.00             | -0.63                    | 99.8                        |

* Data from Ishikawa et al. (14).

increase in the cavity volume. In contrast, additional strains must be created when Val\(^{74}\) is replaced by Leu, because of the decrease in the cavity volume. It should be noted that the $A52V/V74A$ mutant gave a positive $\Delta G_m$ value, but of only $-0.1$ kcal/mol. Likewise, the $A52V/V74L$ mutant also gave a negative $\Delta G_m$ value, but of only $-0.4$ kcal/mol. Therefore, the mutation of Ala\(^{52}\) to Val probably does not create a significant strain at the cavity of the wild-type protein, even when its volume is reduced by the mutation of Val\(^{74}\) to Leu. Consequently, the double mutant protein $A52V/V74L$ was the most stable mutant protein ($\Delta G_m$ of 2.43 kcal/mol) among those constructed in this experiment, and it was more stable than the most stable single mutant protein, $A52I$, by 0.6 kcal/mol.

**DISCUSSION**

As previously reported for a series of tryptophan synthase $\alpha$ subunit mutants (52) and T4 lysozyme mutants (53) with multiple amino acid substitutions at a unique position in the protein interior, a correlation was observed between the changes in the thermal stability ($\Delta T_m$) and the hydrophobicities of the substituted residues for the RNase HI variants with a series of mutations at Ala\(^{52}\) (Fig. 5). However, the data for the RNase HI variants substituted with Tyr, Phe, His, Lys, Pro, and Gly do not fall on the straight line that was obtained from a least-square fit of the data for 12 proteins with Ile, Leu, Cys, Val, Met, Ala, Thr, Ser, Gln, Asn, Glu, and Asp at position 52. The stabilities of the mutant proteins $A52Y$, $A52F$, $A52H$, and $A52K$ are unexpectedly unstable, probably because the introduction of an aromatic or ionized side chain into the cavity destabilizes the protein. Moreover, the mutant proteins $A52E$ and $A52D$, and the normal cellular production level of $A52H$ (data not shown), also suggest that these mutant proteins are considerably unstable only when these residues are ionized. At around pH 7.0, at which the cells grow, Asp and Glu may be ionized, whereas His may not. Since the introduction of a polar group into the cavity neither unexpectedly destabilizes the protein nor seriously affects the near-UV CD spectrum of the protein, unless it is ionized, it seems likely that a polar group could be introduced into the cavity without creating significant strain. In addition to the hydrophobic residues that surround the cavity (Leu\(^{56}\), Val\(^{54}\), Leu\(^{107}\), and Trp\(^{118}\)), many hydrophobic residues, such as Val\(^{51}\), Ile\(^{57}\), Val\(^{59}\), Leu\(^{66}\), Leu\(^{67}\), Leu\(^{74}\), Leu\(^{97}\), Val\(^{165}\), Leu\(^{167}\), Leu\(^{169}\), Leu\(^{171}\), Ile\(^{176}\), and Trp\(^{120}\), form the hydrophobic core of the protein. The extensive hydrophobic interactions among these residues probably make the conformation of the cavity fairly stable and thereby make it tolerant of the mutations. In the wild-type protein, the cavity is vacant and is not occupied by water molecules. However, it remains to be determined whether the introduction of an aromatic group into the cavity is accompanied by the introduction of a water molecule.

In contrast to the mutations to Tyr, Phe, His, and Lys, the mutations to Pro and Gly did not seriously affect the protein conformation, suggesting that these mutations do not create unfavorable van der Waals contacts within the cavity. Ala\(^{52}\) is located in the $\alpha$ helix. Therefore, the instabilities of the mu-

---

2 R. Tanimura, personal communication.

---

FIG. 5. Correlation between the change in the free energy of unfolding of the mutant proteins relative to that of the wild-type protein ($\Delta G_m$) and the hydrophobicity of the substituted residues at position 52 in E. coli RNase HI. The data for the $A52X$ single mutant proteins are indicated by $X$ ("A" for the wild-type protein). The hydrophobic parameter $\sigma$ represents the free energy of transfer of individual amino acids from octanol to water (60). The straight line was obtained from a least-squares fit of the 12 points shown by solid circles.
tant proteins A52P and A52G must reflect the intrinsic helix-
destabilization associated with the mutation of Ala52 → Pro or
Gly. Statistical analyses in natural proteins suggest that Pro is
an α helix breaker, and the introduction of a Pro residue into an
α helix kinks it by an average of 26° ± 5° (54). However, the
similarities in the near-UV CD spectra and the enzymatic
activity between the mutant protein A52P and the wild-type
protein suggest that the conformation of the αI helix is not
seriously affected by the mutation of Ala 52 → Pro. Energy
minimization was carried out to see the effect of the mutation
of Ala52 → Pro on the protein structure by the molecular
mechanics program PRESTO (55). The result supports the
hypothesis that the αI helix is not appreciably kinked at the
position where Pro is introduced (data not shown). The crys-
tallographic (56) and computer (57) analyses for the structures
of the mutant proteins of T4 lysozyme, in which Pro is intro-
duced into α helices, also support this hypothesis. The long
interdomain α helix in T4 lysozyme, which is originally kinked by
8.5°, was shown to be additionally kinked by only 5.5° by the
mutation of Asp72 → Pro (56). All of the mutant proteins of T4
lysozyme with Pro in the helix are dramatically less stable than
the wild-type protein, by 2.7–8.2 kcal/mol. In contrast, the
RNase HI mutant A52P is less stable than the wild-type pro-
tein by only 1.6 kcal/mol. In addition, this mutant protein is
enzymatically fully active. The αI helix of E. coli RNase HI
seems to be more tolerant to Pro substitutions than the α
helices of T4 lysozyme, probably because the αI helix of E. coli
RNase HI is located in the interior of the protein molecule and
is highly stabilized through extensive hydrophobic interactions
with the rest of the protein. In fact, the mutation of either Ala51
or Ala55 → Pro within the αI helix neither seriously destabi-
lizes the protein nor dramatically affects the enzymatic
activity.3

The cavity-filling mutations effectively increase the stability

3 M. Haruki, personal communication.
the single and double mutations at positions 52 (Ala to Ile) and 74 (Val to Ala and Leu) are schematically illustrated in Fig. 6. The mutation of Val to Ala increases the cavity volume, whereas the mutation of Val to Leu decreases it. The former destabilizes the protein by 2.3 kcal/mol and the latter stabilizes it by 1.1 kcal/mol. The mutation of Ala to Ile increases the stability of the V74A mutant by 3.5 kcal/mol. However, the strains still contribute to destabilizing this mutant protein. If the introduction of the double mutation of Ala52 to Ile and Val74 to Leu into the wild-type protein did not create strain, it would increase the protein stability by 1.1 kcal/mol.

It has been reported that “swapped” mutant proteins, such as V35I/I47V of the gene V protein (3, 21) and L121A/A129L of T4 lysozyme (17), in which the core residues are reversed, are considerably less stable than the parent wild-type protein, by 2.9 and 1.1 kcal/mol, respectively. These results indicate that the packing effects are the major determinants of the stabilities of the protein variants with core mutations. However, the swapped RNase HI mutant, A52V/V74L, is less stable than the wild-type protein by only 0.5 kcal/mol. The packing effects may not be major determinants of the stabilities of the proteins with mutations at the cavity. A space in the cavity may serve to reduce the unfavorable contacts caused by a swapped mutation.

In this report, the ΔGm values estimated from the equation of Becktel and Schellman (51) were used to evaluate the effects of the mutations on the protein stability. These values are almost identical with those calculated by using the enthalpy change of unfolding (ΔHm) and the change in the heat capacity (ΔCp) as reported previously (58), except for that of the mutant protein A52K (data not shown). The estimated ΔGm value of the mutant protein A52K (−5.93 kcal/mol) is different from the calculated one (−5.35 kcal/mol), but only by 10%. These results indicate that the equation of Becktel and Schellman (51) is valid for the almost entire range of ΔGm values obtained.

Acknowledgments—We thank R. Tanimura for computer modeling of the structures of the single mutant proteins and S. Kawakita for providing the backbone structure of E. coli RNase HI.

REFERENCES

1. Karpusas, M., Baase, W. A., Matsumura, M. & Matthews, B. W. (1989) Proc. Natl. Acad. Sci. U. S. A. 86, 8237–8241
2. Kellis, J. T. Jr., Nyberg, K. & Fersht, A. R. (1989) Biochemistry 28, 4914–4922
3. Sandberg, W. S. & Terwilliger, T. C. (1989) Science 245, 54–57
4. Sandberg, W. S. & Terwilliger, T. C. (1991) Proc. Natl. Acad. Sci. U. S. A. 88, 1706–1710
5. Shorttle, D., Stites, W. E. & Meeker, A. K. (1990) Biochemistry 29, 8033–8041
6. Doao, S., Anderson, D. E., Baase, W. A., Dahlquist, F. W. & Matthews, B. W. (1991) Biochemistry 30, 11521–11529
7. Baldwin, E., Xu, J., Hajiseyedjavadi, O. & Fersht, A. R. (1992) Biochemistry 31, 4324–4333
8. Elspin, V. G. H., Dijkstra, B. W., Vriend, G., van der Zee, J. R., Veliman, O. R., van der Venne, B., van den Burg, B., Kempe, S. & Venema, G. (1992) Proteins Eng. 5, 421–426
9. Eriksson, A. K., Baase, W. A., Zhang, X.-J., Heinz, D. W., Blaber, M., Baldwin, E. P. & Matthews, B. W. (1992) Science 255, 178–183
10. Wilson, K. P., Malcolm, B. A. & Matthews, B. W. (1992) J. Biol. Chem. 267, 10942–10949
11. Hurley, J. H., Baase, W. A. & Matthews, B. W. (1992) J. Mol. Biol. 224, 1143–1159
12. Varadarajan, R. & Richards, F. M. (1992) Biochemistry 31, 12315–12327
13. Eleanor, K. S., Moracci, M., el-Masy, N., Johnson, C. M. & Fersht, A. R. (1993) Biochemistry 32, 11259–11269
14. Ishikawa, K., Nakamura, H., Morikawa, K. & Kanaya, S. (1993b) Biochemistry 32, 6171–6179
15. Takano, K., Ogashahara, K., Kanesa, H., Yamagata, Y., Fujii, S., Kanay, E., Kikuchi, M., Oobatake, M. & Yutani, K. (1995) J. Mol. Biol. 254, 62–76
16. Baldwin, E. P., Haisjeyedjavadi, O., Baase, W. A. & Matthews, B. W. (1993) Science 262, 1715–1718
17. Baldwin, E., Xu, J., Haisjeyedjavadi, O., Baase, W. A. & Matthews, B. W. (1996) J. Mol. Biol. 259, 542–559
18. Gassner, N. C., Baase, W. A. & Matthews, B. W. (1996) Proc. Natl. Acad. Sci. U. S. A. 93, 12155–12158
19. Buckle, A. M., Henrick, K. & Fersht, A. R. (1993) J. Mol. Biol. 234, 847–860
20. Buckle, A. M., Cramer, P. & Fersht, A. R. (1996) Biochemistry 35, 4288–4305
21. Zhang, H., Skinner, M. M., Sandberg, W. S., Wang, A. H.-J. & Terwilliger, T. C. (1996) J. Mol. Biol. 259, 148–159
22. Lim, W. A. & Sauer, R. T. (1989) Nature 340, 31–36
23. Lim, W. A. & Sauer, R. T. (1993) J. Mol. Biol. 219, 359–376
24. Ake, D. D., Foster, N. W. & Fersht, A. R. (1996) Proc. Natl. Acad. Sci. U. S. A. 93, 5590–5594
25. Lee, C. & Levitt, M. (1991) Nature 352, 448–451
26. Sneddon, S. F. & Tobias, D. J. (1992) Biochemistry 31, 2842–2846
27. van Gunsteren, W. F. & Mark, A. E. (1992) J. Mol. Biol. 227, 389–395
28. Sun, S., Brem, R., Chan, H. S. & Dill, K. A. (1995) Protein Eng. 8, 1265–1213
29. Song, L. & Hermans, J. (1996) Proteins Struct. Funct. Genet. 24, 433–438
30. Pace, C. N. (1992) J. Mol. Biol. 226, 29–35
31. Connolly, M. L. (1986) Int. J. Pept. Protein Res. 28, 360–363
32. Rashin, A., Infin, M. & Wells, J. D. (1993) Biochemistry 32, 3619–3625
33. Williams, M., Goodfellow, J. M. & Thornton, J. M. (1994) Protein Sci. 3, 1224–1235
34. Hubbard, S. I., Gross, K.-H. & Argus, P. (1994) Protein Eng. 7, 613–626
35. 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
36. Kanaya, S., Oobatake, M., Nakamura, H. & Ikehara, M. (1993) J. Biotechnol. 28, 117–136
37. Kimura, S., Oda, Y., Nakai, T., Katayanagi, K., Kitakuni, E., Nakai, C., Nakamura, H., Ikehara, M. & Kanaya, S. (1992a) Eur. J. Biochem. 206, 329–341
38. Kimura, S., Nakamura, H., Hashimoto, T., Oobatake, M. & Kanaya, S. (1992b) J. Biol. Chem. 267, 21535–21542
39. Kimura, S., Kanaya, S. & Nakamura, H. (1992c) J. Biol. Chem. 267, 22014–22017
40. Ishikawa, K., Kimura, S., Kanaya, S., Morikawa, K. & Nakamura, H. (1993a) Protein Eng. 6, 85–91
41. Ishikawa, K., Nakamura, H., Morikawa, K., Kimura, S. & Kanaya, S. (1993c) Biochemistry 32, 7136–7142
42. Haruki, M., Noguchi, E., Akasako, A., Oobatake, M., Itaya, M. & Kanaya, S. (1994) J. Biol. Chem. 269, 26904–26911
43. Akasako, A., Haruki, M., Oobatake, M. & Kanaya, S. (1995) Biochemistry 34, 8115–8122
44. Kanaya, E. & Kanaya, S. (1995) J. Biol. Chem. 270, 19853–19860
45. Yamasaki, K., Ogasahara, K., Yutani, K., Oobatake, M. & Kanaya, S. (1995) Biochemistry 34, 16552–16562
46. Katayama, K., Miyagawa, M., Matsushima, M., Ishikawa, M., Kanaya, S., Nakamura, H., Ikehara, M., Matsuzaki, T. & Morikawa, K. (1992) J. Mol. Biol. 223, 1029–1052
47. Miller, J. H. (1972) in Experiments in Molecular Genetics (Miller, J. H., ed) pp. 433, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY
48. Sanger, F., Nicklen, S. & Coulson, A. R. (1977) Proc. Natl. Acad. Sci. U. S. A. 74, 5463–5467
49. Kanaya, S., Kimura, S., Katsuda, C. & Ikehara, M. (1990) Biochem. J. 271, 59–66
50. Laemmli, U. K. (1970) Nature 227, 680–685
51. Becktel, W. J. & Schellman, J. A. (1987) Biopolymers 26, 1859–1877
52. Yutani, K., Ogasahara, K., Tsujita, T. & Sugino, Y. (1987) Proc. Natl. Acad. Sci. U. S. A. 84, 4441–4444
53. Matsumura, M., Becktel, W. J. & Matthews, B. W. (1988) Nature 334, 406–410
54. Barlow, D. J. & Thornton, J. M. (1988) J. Mol. Biol. 201, 601–619
55. Morikami, K., Nakai, T., Kidera, A., Saito, M. & Nakamura, H. (1992) Comput. Chem. 16, 243–248
56. Sauer, U. H., Dao-pin, S. & Matthews, B. W. (1992) J. Biol. Chem. 267, 2393–2399
57. Gray, T. M., Arnoys, E. J., Blankespoor, S., Born, T., Jagar, R., Everman, R., Plowman, D., Stair, A. & Zhang, D. (1996) Protein Sci. 5, 742–751
58. Kanaya, S., Oobatake, M. & Liu, Y. (1996) J. Biol. Chem. 271, 32729–32736
59. Kraulis, P. J. (1991) J. Appl. Crystallogr. 24, 946–950
60. Fauchere, J.-L. & Pliska, V. (1983) Eur. J. Med. Chem. 18, 369–375