High Resistance of *Escherichia coli* Ribonuclease HI Variant with Quintuple Thermostabilizing Mutations to Thermal Denaturation, Acid Denaturation, and Proteolytic Degradation†

Atsuko Akasako, Mitsuru Haruki, Motohisa Oobatake, and Shigenori Kanaya*

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

Received January 27, 1995; Revised Manuscript Received April 4, 1995®

ABSTRACT: To test whether the combination of multiple thermostabilizing mutations is a useful strategy to generate a hyperstable mutant protein, five mutations, Gly23→Ala, His62→Pro, Val74→Leu, Lys95→Gly, and Asp134→His or Asn, were simultaneously introduced into *Escherichia coli* ribonuclease HI. The enzymatic activities of the resultant quintuple mutant proteins, 5H- and 5N-RNases HI, which have His and Asn at position 134, respectively, were 35 and 55% of that of the wild-type protein. The far-UV and near-UV CD spectra of these mutant proteins were similar to those of the wild-type protein, suggesting that the mutations did not seriously affect the tertiary structure of the protein. The differences in the free energy change of unfolding between the wild-type and mutant proteins, $\Delta\Delta G$, were estimated by analyzing the thermal denaturation of the proteins by CD. The 5H-RNase HI protein, which was slightly more stable than the 5N-RNase HI, was more stable than the wild-type protein by 20.2 °C in T_m and 5.6 kcal/mol in ΔG at pH 5.5. In addition, the 5H-RNase HI was highly resistant to proteolysis and acid denaturation. The effects of each mutation on the thermal stability and the susceptibility to chymotryptic digestion were nearly cumulative, and the 5H-RNase HI undergoes chymotryptic digestion at a rate that is 41 times slower than that of the wild-type protein. Good correlation was observed between the thermal stability and the resistance to chymotryptic digestion for all proteins examined. These results suggest that the thermostabilizing mutations contribute to shift the equilibrium between the folded and unfolded states of the protein so that the fraction of the folded state increases.

The establishment of general methods to increase protein stability is one of the main purposes of protein engineering. Various factors important for protein stability have been identified by introducing a series of mutations in a given protein and analyzing the effect of the mutation(s) on the protein stability (Alber, 1989; Dill, 1990; Fersht & Serrano, 1993). However, the contribution of a single amino acid substitution to protein thermostabilization is not very high, and is generally at most 1−2 kcal/mol in the free energy change of unfolding, ΔG. The combination of multiple thermostabilizing mutations has therefore been expected to be a useful strategy to generate a hyperthermostable mutant protein. The effects of mutations on enzymatic activity, protein−protein interactions, and protein−DNA interactions, as well as on protein stability, have been shown to be additive, when the mutation sites are independent of one another and the conformational change caused by the mutation is localized to the mutation site (Wells, 1990).

Dramatic increases in protein stability by combining more than three individual mutations have been observed in subtilisin BPN′ (Pantoliano *et al.*, 1989) and λ repressor (Stearman *et al.*, 1988). The combination of the six stabilizing mutations, which individually contribute to the stability of subtilisin BPN′ by 0.3−1.3 kcal/mol in ΔG (1.1−4.7 °C in T_m), increased the thermal stability of this protein by 3.8 kcal/mol in ΔG (14.3 °C in T_m) as compared to that of the wild-type protein. Likewise, the combination of the three stabilizing mutations increased the thermal stability of λ repressor by 5.5 kcal/mol in ΔG (16 °C in T_m) as compared to that of the wild-type protein. An additive effect of four stabilizing mutations has also been observed in T4 lysozyme (Zhang *et al.*, 1992). To determine whether this strategy is generally useful to dramatically increase protein stability and whether the enzyme can be hyperstabilized without seriously affecting the enzymatic activity, we simultaneously introduced multiple thermostabilizing mutations into *Escherichia coli* ribonuclease HI (RNase HI))[1] and analyzed the stability of the resultant mutant protein.

E. coli RNase HI, which is a monomeric enzyme (Kanaya *et al.*, 1989) composed of 155 amino acid residues (Kanaya & Crouch, 1983), endonucleolytically cleaves only the RNA strand (P−O3′ bond) of a DNA/RNA hybrid in the presence of Mg^{2+} ions (Berkower *et al.*, 1973). The structure and function of the enzyme have been extensively studied, and it has been shown to be evolutionary related to the RNase H domains of reverse transcriptases from various retroviruses, including HIV [for a review, see Hostomsky *et al.* (1993) and Kanaya and Ikehara (1994)]. It was recently shown that the crystal structures of *E. coli* RuvC resolvase (Ariyoshi *et al.*, 1994) and HIV-1 integrase (Dyda *et al.*, 1994) have similar topologies to that of *E. coli* RNase HI (Katayanagi *et al.*, 1990, 1992; Yang *et al.*, 1990). The topology seen

† The crystal structures of mutant proteins H62P, V74L, and K95G and the wild-type protein have been deposited in the Brookhaven Protein Data Bank under Accession Numbers 1RBR, 1LAV, 1RBT, and 2RN2, respectively.

* To whom correspondence and reprint requests should be addressed. Telephone: 06-872-8204. Fax: 06-872-8210.

® Abstract published in *Advance ACS Abstracts,* June 1, 1995.

[1] Abbreviations: RNase HI, ribonuclease HI; SDS−PAGE, SDS−polyacrylamide gel electrophoresis; HPLC, high-performance liquid chromatography; GdnHCl, guanidine hydrochloride; ribonuclease H (EC 3.1.26.4).

0006-2960/95/0434-8115$09.00/0

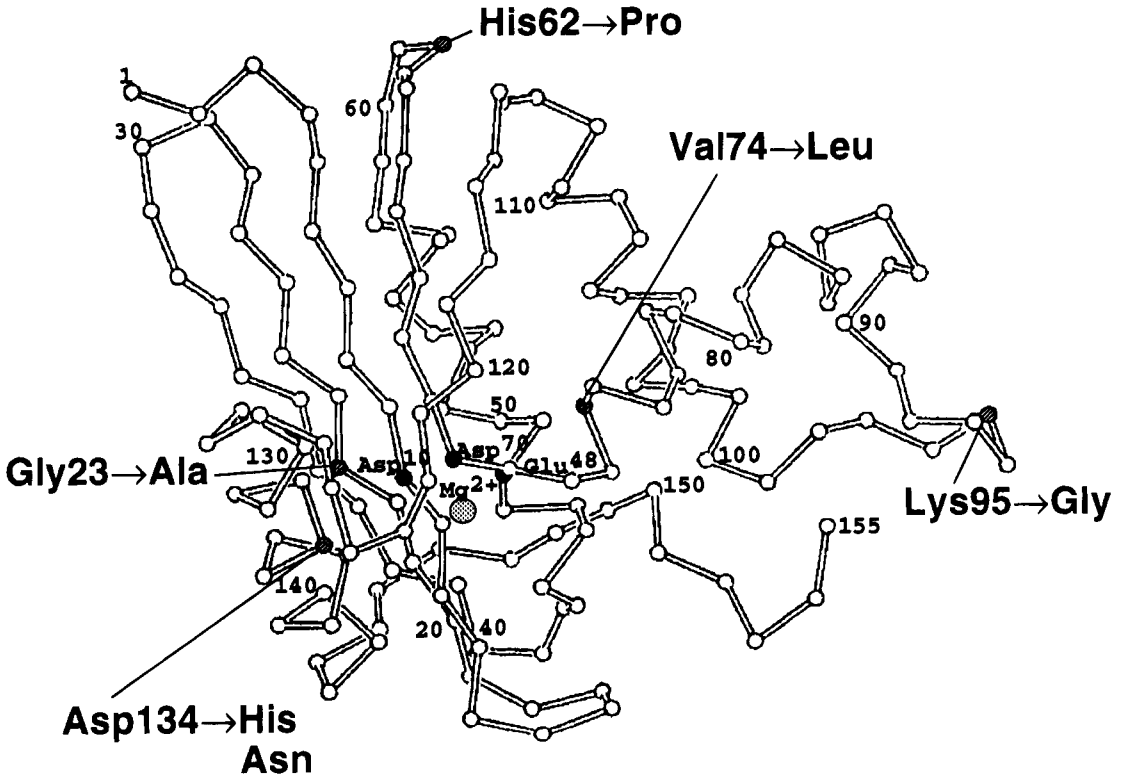

FIGURE 1: Backbone structure of *E. coli* RNase HI. The types and locations of the thermostabilizing mutations that were combined to generate the quintuple mutant proteins are shown. The mutated residues are indicated by hatched circles in the crystal structure of *E. coli* RNase HI (Katayanagi *et al.*, 1990, 1992), in which open circles represent the positions of the α-carbons. The active site residues (Asp10, Glu48, and Asp70) are indicated by solid circles, and the position where the catalytically essential Mg^{2+} ion binds is also shown. Numbers represent the positions of the amino acid residues relative to the initiator methionine of the protein.

in the crystal structure of *E. coli* RNase HI might therefore represent that of a superfamily of nucleotidyl transferases, which are involved in DNA repair or replication. Holm and Sander (1994) classified this topology as an RNase H-like fold.

The stability of the enzyme has been studied as well. We have previously shown that simultaneous replacement of the amino acid sequences in four regions (R_4−R_7) with the corresponding regions from the thermophilic counterpart increased the protein stability by 16.9 °C in T_m at pH 5.5 (Kimura *et al.*, 1992a). The effect of each replacement on the protein stability was independent of the others and was cumulative. However, the resultant mutant protein, $R_4/R_5/R_6/R_7$-RNase H, in which 19 residues are mutated and 1 residue is inserted, exhibits only 5% of the enzymatic activity of the wild-type protein at 37 °C. We have therefore decided to combine single mutations that had been shown to increase the protein stability without seriously affecting the enzymatic activity. Such mutations had been identified by using the following strategies: (1) computer-assisted design; (2) replacement of amino acid residues with those from the thermophilic counterpart; (3) random mutagenesis; and (4) serendipity. They include Gly23→Ala (Haruki *et al.*, 1994b), His62→Pro (Kimura *et al.*, 1992a), Val74→Leu (Ishikawa *et al.*, 1993b), Lys95→Gly (Kimura *et al.*, 1992b), and Asp134→His and Asn (Haruki *et al.*, 1994a). The locations of these mutation sites are shown in Figure 1. The effects of the mutations at these five sites are expected to be additive, because these sites are located far from one another in the crystal structure of the enzyme (even the nearest two residues, Gly23 and Asp134, are located 6.6 Å apart from each other). The mutant proteins with these single mutations, G23A, H62P, V74L, K95G, D134H, and D134N, are more stable than the wild-type protein by 0.5−1.9 kcal/mol in ΔG and 1.8−7.0 °C in T_m at pH 5.5.

The crystal structures of the mutant proteins H62P (Ishikawa *et al.*, 1993a), V74L (Ishikawa *et al.*, 1993b), K95G (Ishikawa *et al.*, 1993a), and D134N and D134H (T. Kashiwagi, personal communication) have been determined. The root-mean-square (rms) displacements between the wild-type and mutant proteins, H62P, V74L, and K95G, are 0.198, 0.160, and 0.156 Å, respectively. This means that the main chain foldings of these mutant proteins are almost identical with that of the wild-type protein. As a typical example, the main chain folding of the mutant protein H62P is compared with that of the wild-type protein in Figure 2. The crystal structures of these mutant proteins have been deposited in the Brookhaven Protein Data Bank under Accession Numbers of 1RBR for H62P, 1LAV for V74L, and 1RBT for K95G.

Determination of the crystal structures of the mutant proteins thus indicated that only the local conformational change around each mutation site is responsible for the increase in the thermostabilization. The H62P protein is stabilized probably because the entropy of the unfolded state of this protein is decreased. The V74L protein is stabilized because the volume of the cavity within the hydrophobic core is reduced by the introduction of an additional methylene group (cavity-filling mutation). The K95G protein is stabilized because the conformational strain caused by the left-handed backbone structure in the typical 3:5 type loop is eliminated. The D134H and D134N proteins are stabilized by reducing the negative-charge repulsion and/or by enhancing the stability of the α-helix. The G23A protein, for which

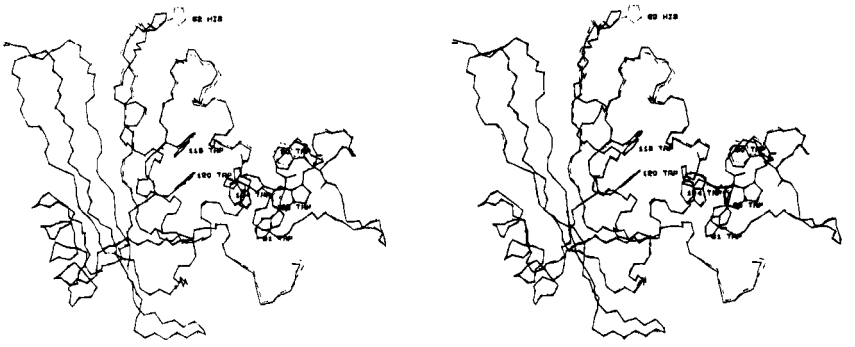

FIGURE 2: Comparison of the main chain foldings between H62P and wild-type proteins. Stereopair of the main chain folding for the mutant protein H62P (solid line) is superimposed to that for the wild-type protein (broken line). The N- and C-termini in these structures point to the same directions as those shown in Figure 1. The side chains at position 62 (Pro for H62P and His for wild-type), as well as those of all Trp residues at positions 81, 85, 90, 104, 118, and 120, are indicated. The crystal structures of the mutant protein H62P (Ishikawa *et al.*, 1993a) and the wild-type protein (Katayanagi *et al.*, 1990, 1992) are deposited in the Brookhaven Protein Data Bank under Accession Numbers 1RBR and 2RN2, respectively. This figure was kindly provided by Mr. T. Kashiwagi.

the crystal structure is not available, seems to be stabilized by filling a cavity and/or by decreasing the entropy of unfolding.

We report here that the simultaneous introduction of the five mutations, Gly23→Ala, His62→Pro, Val74→Leu, Lys95→Gly, and Asp134→His or Asn, dramatically increases the thermostability of *E. coli* RNase HI without considerably decreasing the enzymatic activity. We also report that the resultant quintuple mutant protein exhibits high resistance to acid denaturation and proteolytic degradation.

EXPERIMENTAL PROCEDURES

Materials. The wild-type *E. coli* RNase HI protein (Kanaya *et al.*, 1989) and its variants G23A (Haruki *et al.*, 1994b), H62P (Kimura *et al.*, 1992a), V74L (Ishikawa *et al.*, 1993b), K95G (Kimura *et al.*, 1992b), and D134H and D134N (Haruki *et al.*, 1994a), in which Gly23, His62, Val74, Lys95, Asp134, and Asp134 are replaced by Ala, Pro, Leu, Gly, His, and Asn, respectively, were previously constructed. In this report, these mutant proteins are represented by "single mutant proteins". Restriction enzymes and modifying enzymes were from Takara Shuzo Co., Ltd. Ultrapure-grade guanidine hydrochloride (GdnHCl) was from Schwarz/Mann. TLCK-treated chymotrypsin was from Sigma, and lysyl endopeptidase was from Wako Pure Chemicals Co., Ltd. Other chemicals were of reagent grade.

Cells and Plasmids. The plasmids pJAL600 (Kanaya *et al.*, 1993), pJAL74L (Ishikawa *et al.*, 1993b), and pJAL134N and pJAL134H (Haruki *et al.*, 1994a) for the overproduction of the wild-type protein and the mutant proteins V74L, D134N, and D134H, respectively, were previously constructed. These plasmids bear the wild-type or the mutant *rnhA* gene under the control of the bacteriophage λ promoters P_R and P_L, the cI^{ts857} gene, and the bacteriophage fd transcription terminator. Competent cells of *E. coli* HB101 [F^-, *hsdS20* (r_B^-,m_B^-) *recA13*, *ara-13*, *proA2*, *lacY1*, *galK2*, *rspL20* (S_m^r), *xyl-5*, *mtl-1*, *supE44*, λ^-] were obtained from Takara Shuzo Co., Ltd. *E. coli* HB101 transformants with pJAL600 derivatives were grown in Luria−Bertani medium (Miller, 1972) containing 100 mg/L ampicillin at 30 °C, unless specifically noted.

Mutations. Two plasmids, pJAL5N and pJAL5H, for the overproduction of the quintuple mutant proteins 5N-RNase HI and 5H-RNase HI, respectively, were constructed according to the following procedures. In the 5N-RNase HI protein, the Gly23→Ala, His62→Pro, Val74→Leu, Lys95→Gly, and Asp134→Asn mutations were simultaneously introduced. The 5H-RNase HI protein is identical to the 5N-RNase HI protein, except that it has His, instead of Asn, at position 134. The plasmid pJAL600 derivatives (pJAL74L, pJAL134N, and pJAL134H) contain the unique *Bam*HI site within the mutant *rnhA* genes, which encompasses the sequences encoding the amino acid residues at positions 81−83, and the unique *Sal*I site, which is located approximately 50 base pairs downstream of the termination codon of the mutant *rnhA* gene. The plasmid pJAL600 derivatives for the overproduction of the double mutant proteins V74L/D134N and V74L/D134H were first constructed by replacing the small *Bam*HI−*Sal*I fragment of pJAL74L with the corresponding fragments of pJAL134N and pJAL134H. The additional three mutations were then individually introduced into the mutant *rnhA* genes in these plasmids by PCR, as described previously (Kanaya *et al.*, 1993). The DNA oligomers used as the 5′- and 3′-mutagenic primers (30−40 bases long) were synthesized by Sawady Technology Co., Ltd. Consequently, in the mutant *rnhA* genes encoding either 5N-RNase HI or 5H-RNase HI, the codons for Gly23, His62, Val74, Lys95, and Asp134 in the wild-type *rnhA* gene are changed from GGC to GCC, CAT to CCA, GTC to CTG, AAA to GGC, and GAT to AAT or CAT, respectively. The nucleotide sequences of these mutant *rnhA* genes were determined using the dideoxy chain termination method (Sanger *et al.*, 1977).

Biochemical Characterizations. The mutant 5N- and 5H-RNase HI proteins were overproduced in *E. coli* HB101 harboring plasmids pJAL5N and pJAL5H by raising the temperature of the growth medium from 30 °C to 42 °C, and were purified as described previously (Kanaya *et al.*, 1993). The RNase H activity was determined at 30 °C in 10 mM Tris-HCl (pH 8.0) containing 10 mM $MgCl_2$, 50 mM NaCl, 1 mM 2-mercaptoethanol, and 10 μg/mL bovine serum albumin, using ^{3}H-labeled M13 DNA/RNA hybrid as a substrate (Kanaya *et al.*, 1991). The protein concentration was determined from the UV absorption, assuming that the mutant 5N- and 5H-RNase HI proteins have the same $A_{280}^{0.1\%}$ value of 2.0 as that of the wild-type protein (Kanaya *et al.*, 1990). The far- and near-UV circular dichroism (CD) spectra were measured in 10 mM Gly-HCl (pH 3.2) at 10 °C on a J-720 spectropolarimeter (Japan Spectroscopic Co., Ltd.), as

described previously (Ishikawa *et al.*, 1993b). The results were expressed as the mean residue ellipticity ([θ], deg cm^2 dmol^{-1}), which was calculated by using an average molecular weight of the amino acid of 110. For the analysis of the pH dependence of the CD spectra, which were measured at 25 °C, proteins were dissolved in 10 mM sodium acetate over the pH range of 4.3−5.0, in 10 mM Gly-HCl over the pH range of 1.01−3.00, or in an appropriate concentration of HCl over the pH range of 0.22−0.96. SDS−PAGE was carried out with a 15% polyacrylamide gel, as described previously (Laemmli, 1970). The gel was stained with Coomassie Brilliant Blue.

Thermal Denaturation. Thermal denaturation curves and the temperature of the midpoint of the transition (T_m) were determined by monitoring the change in the CD value at 220 nm, as described previously (Kimura *et al.*, 1992a). Proteins were dissolved either in 10 mM Gly-HCl (pH 3.0) or in 10 mM sodium acetate (pH 5.5) containing 0.1 M NaCl and 1 M GdnHCl. The enthalpy change of unfolding at the T_m (ΔH_m) was calculated by van't Hoff analysis. The difference in the free energy change of unfolding between the mutant and wild-type proteins at the T_m of the wild-type protein ($\Delta\Delta G_m$) was estimated by the relationship given by Becktel and Schellman (1987); $\Delta\Delta G_m = \Delta T_m \Delta S_m$ where ΔT_m is the change in T_m of a mutant protein relative to that of the wild-type protein and ΔS_m is the entropy change of the wild-type protein at the T_m, which was previously determined as 0.304 kcal/(mol·K) at pH 3.0 and as 0.275 kcal/(mol·K) at pH 5.5 (Kimura *et al.*, 1992a). The theoretical curves for unfolding were drawn on the assumption that the protein unfolds in a two-state mechanism, by using the T_m and ΔH_m values experimentally determined and the ΔC_p value of 1.79 kcal/mol (M. Oobatake, personal communication).

Protease Digestion. Digestion of the protein (0.1 mg/mL) with chymotrypsin was carried out at 37 °C in 0.1 M Tris-HCl, pH 8.0, with a substrate:enzyme ratio of 20:1. Proteolytic degradation was assayed by measuring the change in the amount of the uncleaved protein, which was separated by reverse-phase HPLC, as described previously (Kanaya *et al.*, 1990).

RESULTS AND DISCUSSION

Enzymatic Activity. It was previously shown that the single mutant protein D134H is more stable by 3.8 °C in T_m at pH 5.5, but enzymatically less active by 40%, than D134N (Haruki *et al.*, 1994a). We have therefore decided to construct two quintuple mutant proteins, 5H- and 5N-RNases HI, to examine whether the differences in the thermal stability and the enzymatic activity between these quintuple mutant proteins reflect those between D134H and D134N. The cellular production levels of the quintuple mutant proteins were higher than that of the wild-type protein, and the amount of protein purified from a 1-L culture, which is roughly 50 mg for the wild-type protein, was 95 mg for 5H-RNase HI and 70 mg for 5N-RNase HI.

The 5N-RNase HI protein exhibited 55% of the enzymatic activity of the wild-type protein. None of the constituent mutations alone has been reported to significantly affect the enzymatic activity. However, the RNase H activities of the single mutant proteins have been determined with relatively large errors (±20%), and therefore it is possible that each

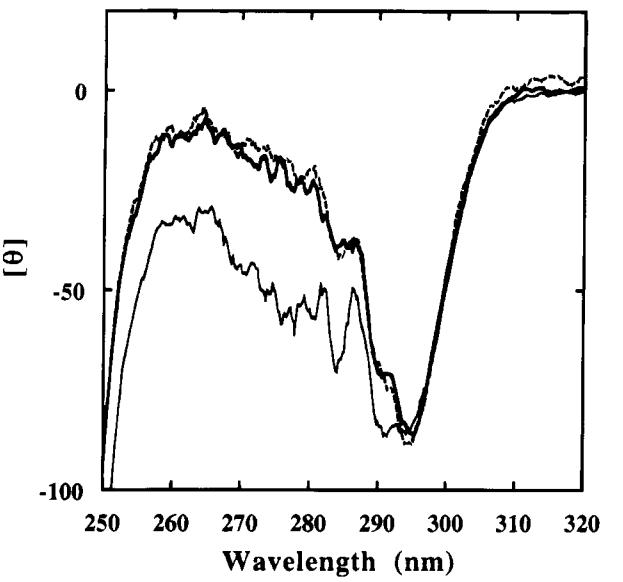

FIGURE 3: CD spectra of the quintuple mutant proteins. The near-UV CD spectra of 5N-RNase HI (broken line) and 5H-RNase HI (thick line), which were determined at pH 3.2 and 10 °C, are shown in comparison with that of *E. coli* RNase HI (thin line).

mutation slightly reduces the enzymatic activity. The cumulative effect of the subtle reduction in the enzymatic activity by each mutation may be able to account for the reduced enzymatic activity of 5N-RNase HI. The 5H-RNase HI protein exhibited 35% of the enzymatic activity of the wild-type protein. The decrease in the enzymatic activity of 5H-RNase HI relative to 5N-RNase HI is comparable to that of D134H relative to D134N. It has been suggested that the polar atom at the δ position of residue 134 is involved in the catalytic function of the enzyme (Haruki *et al.*, 1994a). The enzymatic activity of the mutant protein D134H is lower than D134N, probably because the N atom at the δ position of His134 is not located exactly at the same position, in which the O atom at the δ position of Asn134 is located.

CD Spectra. The far-UV CD spectra of the 5N- and 5H-RNases HI were identical to that of the wild-type protein (data not shown). However, the near-UV CD spectra of these mutant proteins, which were almost identical to each other, were slightly different from that of the wild-type protein (Figure 3). The near-UV CD spectrum of the wild-type protein exhibited double minima at 290 and 294 nm, with a [θ] value of approximately −90 deg cm^2 dmol^{-1} and a series of less intense minima around 270−285 nm. In contrast, those of the quintuple mutant proteins had a single minimum at 294 nm and lacked a series of less intense minima around 270−285 nm. To identify the amino acid substitution(s) that is (are) responsible for such alterations in the CD spectra, the near-UV CD spectra of the single mutant proteins were measured (Figure 4). The near-UV CD spectra of all the single mutant proteins, except for G23A and H62P, were almost identical to that of the wild-type protein (data not shown). It seems clear that the His62→Pro mutation is responsible for the disappearance of the minimum at 290 nm, and the Gly23→Ala and His62→Pro mutations are both responsible for the disappearance of a series of less intense minima around 270−285 nm.

According to the previously determined crystal structure (Ishikawa *et al.*, 1993a), the His62→Pro mutation only

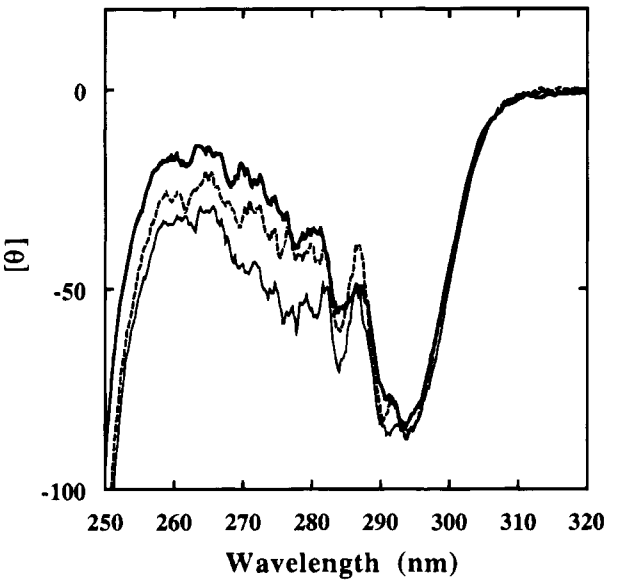

FIGURE 4: CD spectra of the single mutant proteins. The near-UV CD spectra of the mutant proteins G23A (broken line) and H62P (thick line), which were determined at pH 3.2 and 10 °C, are shown in comparison with that of *E. coli* RNase HI (thin line). The spectra of V74L, K95G, D134N, and D134H are not shown, because they are almost identical to that of the wild-type protein.

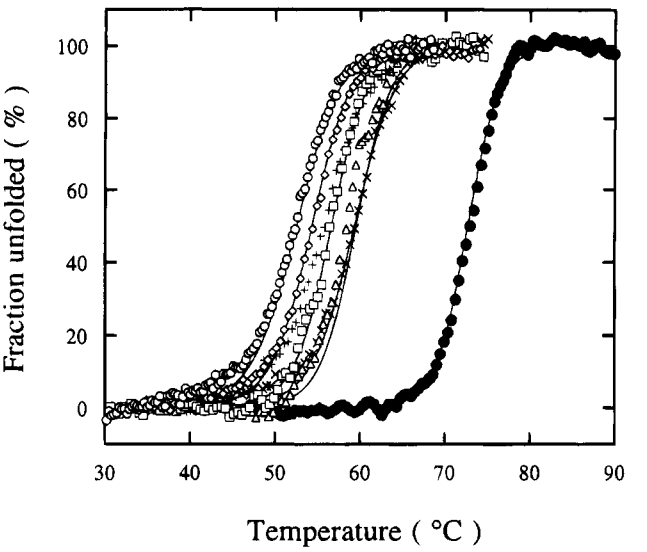

FIGURE 5: Thermal denaturation curves of the wild-type and mutant proteins. The apparent fraction of unfolded protein is shown as a function of temperature. (○) Wild-type; (◇) G23A; (□) H62P; (+) V74L; (△) K95G; (×) D134H; (●) 5H-RNase HI. Thermal denaturation curves were determined in the presence of 1.0 M GdnHCl at pH 5.5 by monitoring the change in the CD value at 220 nm, as described under Experimental Procedures. The theoretical curves for unfolding were drawn as described under Experimental Procedures.

seriously affects the structure at the mutation site (Figure 2). This mutation slightly affects the conformations of Tyr and Trp residues as well. The Val74→Leu and Lys95→Gly mutations also affect the conformations of these aromatic residues, but with much less extent. The conformations of the Trp residues of the mutant protein H62P are compared with those of the wild-type protein in Figure 2. All the Trp residues slightly change their conformations (at most 0.57 Å for Trp90), when the His62→Pro mutation is introduced. These subtle conformational changes of the Tyr and/or Trp residues, which may also be introduced by the Gly23→Ala mutation, might be responsible for the alteration in the near-UV CD spectrum. Nevertheless, the resemblance of the near-UV CD spectra of the quintuple mutant proteins to that of the wild-type protein strongly suggests that the structures of the quintuple mutant proteins are basically the same as that of the wild-type protein, except at the mutation sites.

Thermal Stability. Like the wild-type protein, the quintuple mutant proteins reversibly unfolded in a single cooperative fashion at pH 3.0, and at pH 5.5 in the presence of 1 M GdnHCl, with thermal denaturation. The thermal denaturation curves of the wild-type and 5H-RNase HI proteins determined at pH 5.5 in the presence of 1 M GdnHCl, as well as those of the five single mutant proteins with the constituent amino acid substitutions, are shown in Figure 5. These curves fit well to the theoretical curves for unfolding, which are obtained by assuming that these proteins unfold in a two-state mechanism. The parameters characterizing the thermal denaturation of the quintuple mutant proteins are summarized in Table 1. The hypothetical $\Delta\Delta G_m$ values for the quintuple mutant proteins were obtained by simply adding the $\Delta\Delta G_m$ values determined for the five single mutant proteins with the constituent amino acid

Table 1: Parameters Characterizing the Thermal Denaturation of the Wild-Type and Mutant *E. coli* RNase HI Proteins[a]

	pH 3.0				pH 5.5			
protein	ΔT_m °C	$\Delta\Delta G_m$ (kcal/mol)	$\Delta\Delta G_m$ (sum) (kcal/mol)	ΔH_m (kcal/mol)	ΔT_m °C	$\Delta\Delta G_m$ (kcal/mol)	$\Delta\Delta G_m$ (sum) (kcal/mol)	ΔH_m (kcal/mol)
wild-type				105.8				81.8
G23A	2.3	0.7		99.3	1.8	0.5		91.8
H62P	3.4	1.0		109.5	4.1	1.1		102.0
V74L	3.7	1.1		118.6	3.3	0.9		91.2
K95G	5.7	1.7		102.4	6.8	1.9		90.4
D134N	−0.3	−0.1		110.5	3.2	0.9		97.0
D134H	1.7	0.5		106.5	7.0	1.9		106.6
5N-RNase HI	12.5	3.8	4.4	130.9	17.6	4.8	5.3	129.6
5H-RNase HI	14.2	4.3	5.0	108.8	20.2	5.6	6.3	127.7

[a] Thermal denaturation curves were measured at 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 50.9 °C at pH 3.0 and 52.5 °C at pH 5.5. ΔH_m is the enthalpy change of unfolding at T_m, which was calculated by van't Hoff analysis. $\Delta\Delta G_m$, which is the change in the free energy change of unfolding at T_m, ΔG_m, relative to that of the wild-type protein, was estimated by the relationship given by Becktel and Schellman (1987) as described under Experimental Procedures. $\Delta\Delta G_m$ (sum) represents the sum of the $\Delta\Delta G_m$ values of the single mutant proteins with constituent substitutions. Data for the single mutant proteins G23A (Haruki *et al.*, 1994b), H62P (Kimura *et al.*, 1992a), V74L (Ishikawa *et al.*, 1993b), K95G (Kimura *et al.*, 1992b), and D134N and D134H (Haruki *et al.*, 1994a) were obtained from the literature. Errors are within ±0.3 °C for T_m, ±12 kcal/mol for ΔH_m, and 0.09 kcal/mol for $\Delta\Delta G_m$.

substitutions. The $\Delta\Delta G_m$ values of the 5H- and 5N-RNase HI proteins were slightly smaller than the hypothetical values at either pH 3.0 or pH 5.5, but only by 10−14% of the hypothetical values. These results indicate that the effect of each mutation on protein thermostability is nearly additive.

The 5H-RNase HI protein, which is more stable than the wild-type protein by 14.2 °C in T_m and 4.3 kcal/mol in ΔG at pH 3.0, and by 20.2 °C in T_m and 5.6 kcal/mol in ΔG at pH 5.5, is more stable than 5N-RNase HI at either pH 3.0 or pH 5.5. The difference in the T_m values between these proteins reflects that between the single mutant proteins D134N and D134H. The 5H-RNase HI protein was further analyzed for stability against acid denaturation and proteolytic degradation.

Stability against Acid Denaturation. E. coli RNase HI is stable over the pH range of 1.66−10.12 at 10 °C, but is denatured at extremely low pH (pH 1.2), as judged by the change in the far-UV CD spectra (Kanaya *et al.*, 1993). It was later shown that the acid-denatured state at pH 1.0 has the characteristics of a molten globule (Dabora & Marqusee, 1994). Dabora and Marqusee (1994) have obtained the far- and near-UV CD spectra of the Cys-free mutant RNase HI protein, in which all three Cys residues are replaced by Ala, in both the native and acid states. We obtained similar spectra for the wild-type protein as well (data not shown). At 25 °C, the far-UV CD spectrum of the native protein exhibits a broad minimum at 216 nm with a $[\theta]$ value of $-13\,000$ deg cm^2 dmol^{-1}, whereas that of the acid-denatured state shows a minimum at 205 nm with a $[\theta]$ value of $-11\,000$ deg cm^2 dmol^{-1}. The negative peak in the spectrum of the acid state is accompanied by a broad shoulder with a $[\theta]$ value of ~-6000 deg cm^2 dmol^{-1} at 215−220 nm. The near-UV CD spectrum of the protein in the acid-denatured state is similar to that of the unfolded state, rather than that of the native state. Thus, the CD values of the protein at 220 and 294 nm increased from $-13\,000$ to -6000 deg cm^2 dmol^{-1} and from -90 to -20 deg cm^2 dmol^{-1} upon acid denaturation, respectively.

The pH dependencies of the CD values at 294 nm (Figure 6a) and 220 nm (Figure 6b) are shown for the wild-type and 5H-RNase HI proteins. Both CD values of the wild-type protein at 294 nm (Figure 6a) and 220 nm (Figure 6b) increased as the pH was decreased below 2.5, and reached the highest values at around pH 1.6. In contrast, those of the 5H-RNase HI were unchanged at an extremely low pH, even at pH 0.22, indicating that 5H-RNase HI is quite resistant to acid denaturation. According to the classifications made by Fink *et al.* (1994), which are based on the behaviors of proteins upon acid denaturation, the wild-type protein and 5H-RNase HI belong to the type I and type III protein classes, respectively. Upon acid titration, the wild-type protein initially unfolds to the acid-unfolded state (U_A-state), which retains a low level of secondary structure and minimal tertiary structure, at a pH below 2.5. At this state, the protein shows cooperative temperature denaturation with heat absorption (Oobatake *et al.*, 1993). The apparent pH of the midpoint of this transition was around 2.1. Then, the protein probably refolds to the molten globule-like acid-denatured state (A-state), which retains a high level of secondary structure and minimal tertiary structure, at a pH below 1.6.

The hyperresistance of 5H-RNase HI to acid denaturation is not attributable to the change in the net charge of the protein caused by the mutations, because an increase in the number of positive charges relative to that of the negative charges contributes to destabilizing the protein at an acidic pH. Among the acidic amino acid residues in *E. coli* RNase HI, Asp102 and Asp148 have unusually low pK_a values (<2) (Oda *et al.*, 1994). These residues are located inside the protein molecule and form a salt bridge network with Arg46 (Katayanagi *et al.*, 1992). The elimination of such a salt bridge network, due to the protonation of these aspartates at a pH below 2, is therefore expected to seriously affect the conformation of the protein. However, the fact that the 5H-RNase HI retains a native-like conformation at pH 0.22 excludes this possibility, because it is unlikely that this salt bridge network is formed at pH 0.22, at which all carboxylates must be protonated. It has been reported that the protein stability of *E. coli* RNase HI decreased in accordance with a decrease of the pH below 4, which favors the protonation of carboxylates, probably due to an increase in unfavorable electrostatic interactions among the positive charges (Kanaya *et al.*, 1993). The transition of this protein to the U_A-state might be initiated when the protein stability is decreased below a certain level. The 5H-RNase HI protein did not unfold with acid treatment, probably because the protonation of all the carboxylates does not further decrease the protein stability.

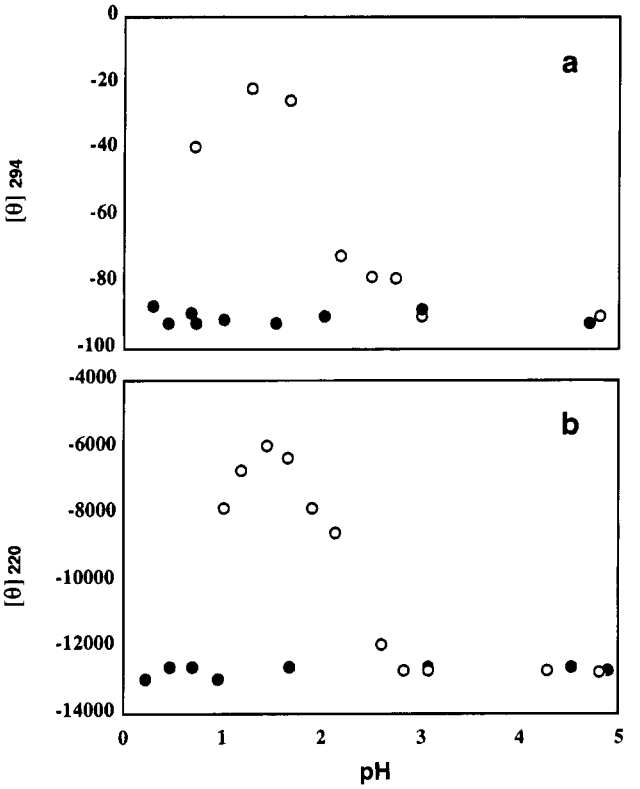

FIGURE 6: Effect of acid on protein conformation. The pH dependencies of the CD values at 294 nm (a) and 220 nm (b) for 5H-RNase HI (solid circles) at 25 °C are shown in comparison with those for the wild-type protein (open circles). The CD spectra were measured as described under Experimental Procedures.

Stability against Proteolysis. The susceptibility of the 5H-RNase HI to chymotryptic digestion was analyzed by monitoring the change in the amount of the uncleaved protein, as compared to the wild-type and single mutant proteins. The amount of the remaining uncleaved protein, which was determined by HPLC, was consistent with estimates from SDS−PAGE and the residual enzymatic

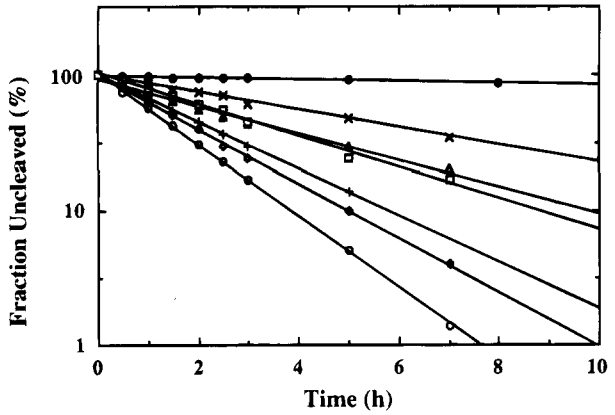

FIGURE 7: Susceptibility of the wild-type and mutant proteins to proteolysis by chymotrypsin. The fraction of uncleaved protein is shown as a function of incubation time. Proteins are designated by the same symbols as those in Figure 5. The amount of uncleaved protein was determined by subjecting the chymotryptic digest to reverse-phase HPLC, by which the uncleaved protein could be separated from the chymotryptic fragments.

Table 2: Susceptibilities of Proteins to Chymotryptic Digestion[a]

protein	k (h^{-1})	relative k
wild-type	0.60	1.0
G23A	0.46	0.77
H62P	0.25	0.42
V74L	0.39	0.65
K95G	0.23	0.39
D134H	0.15	0.25
5H-RNase HI	0.015	0.025

[a] Proteins were digested with chymotrypsin at 37 °C as described under Experimental Procedures. At different time intervals (20 min−2 h), an aliquot was withdrawn and subjected to reverse-phase HPLC to determine the amount of uncleaved protein. Rate constants (k) were determined from the slope of semilog plots of the fraction uncleaved *versus* the incubation time. In the absence of chymotrypsin, the amount of uncleaved protein determined by HPLC was unchanged during the 10 h incubation for all of the wild-type and mutant proteins examined. Relative k values were calculated by dividing the k values of the mutant proteins by that of the wild-type protein.

activity (data not shown). Well-defined chymotryptic fragments were not detected on SDS−PAGE for all proteins examined. In addition, the elution profile of the chymotryptic peptides, generated after incomplete digestion, on reverse-phase HPLC was almost identical to that generated after complete digestion, with a few exceptions, for all proteins examined. These results suggest that the intermediates generated after the first proteolytic cleavage were highly susceptible to further cleavage.

Figure 7 shows semilog plots of the fraction of uncleaved protein *versus* the incubation time for the wild-type, single, and quintuple mutant proteins. In all cases, the rate of decrease in the amount of uncleaved protein was first-order for more than 90% of the degradation. Among all the proteins examined, the wild-type protein was most sensitive to proteolytic degradation. The susceptibility of the proteins to proteolytic degradation decreased in the order of wild-type, G23A, V74L, H62P, K95G, D134H, and 5H-RNase HI. The rate constants (k) were determined from the slopes shown in Figure 7, and are summarized in Table 2. The relative k values were calculated by dividing the k values for the mutant proteins by that for the wild-type protein. The relative k value for 5H-RNase HI (0.025) is slightly larger than, but similar to, the hypothetical value (0.021), which is

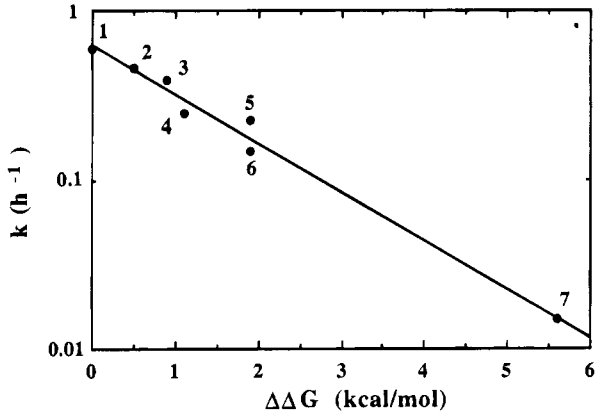

FIGURE 8: Inverse correlation between the susceptibility of proteins to proteolysis by chymotrypsin and the thermal stability. The rate constants (k) for the proteolytic degradation of wild-type (1), G23A (2), V74L (3), H62P (4), K95G (5), D134H (6), and 5H-RNase HI (7), which were determined at pH 8.0 from the slopes of semilog plots of the fractions uncleaved *versus* time shown in Figure 7, are shown as functions of the $\Delta\Delta G$ values at pH 5.5. The line was obtained by linear regression of the data.

obtained by simply multiplying the relative k values for the five single mutant proteins with the constituent amino acid substitutions. These results indicate that the effect of each thermostabilizing mutation on the susceptibility to chymotryptic digestion is nearly cumulative. Similar results were obtained when the susceptibility of the proteins to digestion with lysyl endopeptidase was examined (data not shown). Chymotrypsin recognizes hydrophobic amino acid residues as substrates, whereas lysyl endopeptidase recognizes Lys as a substrate. Therefore, the hyperresistance of the quintuple mutant protein to proteolytic degradation seems to be independent of the substrate specificity of the protease.

Semilog plots of the k values *versus* the $\Delta\Delta G$ values for all mutant proteins examined revealed a good inverse correlation (correlation coefficient of 0.984) between the susceptibility to chymotryptic digestion and the thermal stability (Figure 8). It has been reported that digestion by proteases proceeds mainly *via* the unfolded state of proteins (Imoto *et al.*, 1986). Therefore, it seems likely that each thermostabilizing mutation cumulatively contributes to shift the equilibrium between the folded and unfolded states of the protein so that the fraction of the folded state increases.

Stability−Activity Relationship. All the mutations used to generate the quintuple mutant proteins were designed so that they independently reinforce the various loci with conformational defects affecting the thermostability. There is controversy as to whether the enzymatic activity decreases as the protein stability increases. It has been reported for barnase that conformational flexibility in the active site is responsible for the enzymatic activity, and there is an inverse relationship between the stability and the activity for many of the active site mutations (Meiering *et al.*, 1992). The enzymatic activities of the 5N- and 5H-RNases HI are lower than that of the wild-type protein, probably because the reinforcement of the loci with conformational defects affecting the thermostability influences the conformational flexibility in the active site. However, the extent of enzymatic activity reduction was not very large. Enzyme molecules may generally have a number of loci outside the active site with conformational defects that affect the thermostability. Therefore, reinforcing these loci by intro-

ducing multiple mutations would be generally effective to dramatically increase the protein stability without seriously affecting the enzymatic activity. It should be noted, however, that the low correlation between protein stability and loss of enzymatic activity for *E. coli* RNase HI may be due in part to the relatively flat surface of its active site. We do not know whether this relationship is generally applicable to proteins that have much more intimate contact with their substrates.

The quintuple mutant protein 5H-RNase HI might be more useful than *E. coli* RNase HI or *T. thermophilus* RNase H. The enzyme has been used as a tool for recombinant DNA technologies. In addition, it would be useful for the oligoDNA-directed cleavage of RNA (*in vitro* RNA editing). A significant increase in the stability of the enzyme against heat and proteolysis might extend its applications. The *T. thermophilus* RNase H is more stable than the *E. coli* RNase HI by 33.9 °C in T_m (Kanaya & Itaya, 1992), and the 5H-RNase Hi is less stable than the *T. thermophilus* enzyme by 13 °C in T_m. However, this mutant protein is 5−7 times more active than the *T. thermophilus* RNase H over the temperature range of 30−60 °C (data not shown). The introduction of a few additional thermostabilizing mutations into the 5H-RNase HI may generate an *E. coli* RNase HI variant that is more stable and active than the *T. thermophilus* RNase H.

ACKNOWLEDGMENT

We thank Drs. E. Kanaya and K. Yamasaki for helpful discussions.

REFERENCES

Alber, T. (1989) *Annu. Rev. Biochem. 58*, 765−798.
Ariyoshi, M., Vassylyev, D. G., Iwasaki, H., Nakamura, H., Shinagawa, H., & Morikawa, K. (1994) *Cell 78*, 1063−1072.
Becktel, W. J., & Schellman, J. A. (1987) *Biopolymers 26*, 1856−1877.
Berkower, I., Leis, J., & Hurwitz, J. (1973) *J. Biol. Chem. 248*, 5914−5921.
Dabora, J. M., & Marqusee, S. (1994) *Protein Sci. 3*, 1401−1408.
Dill, K. A. (1990) *Biochemistry 29*, 7133−7155.
Dyda, F., Hickman, A. B., Jenkins, T. M., Engelman, A., Craigie, R., & Davies, D. R. (1994) *Science 266*, 1981−1986.
Fersht, A. R., & Serrano, L. (1993) *Curr. Opin. Struct. Biol. 3*, 75−83.
Fink, A. L., Calciano, L. J., Goto, T., Kurotsu, T., & Palleros, D. R. (1994) *Biochemistry 33*, 12504−12511.
Haruki, M., Noguchi, E., Nakai, C., Liu, Y.-Y., Oobatake, M., & Kanaya, S. (1994a) *Eur. J. Biochem. 220*, 623−631.
Haruki, M., Noguchi, E., Akasako, A., Oobatake, M., Itaya, M., & Kanaya, S. (1994b) *J. Biol. Chem. 269*, 26904−26911.
Holm, L., & Sander, C. (1994) *Proteins: Struct., Funct., Genet. 19*, 165−173.
Hostomsky, Z., Hostomska, Z., & Matthews, D. A. (1993) in *Nucleases* (Linn, S. M., & Roberts, R. J., Eds.) 2nd ed., pp 341−376, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.
Imoto, T., Yamada, H., & Ueda, T. (1986) *J. Mol. Biol. 190*, 647−649.
Ishikawa, K., Kimura, S., Kanaya, S., Morikawa, K., & Nakamura, H. (1993a) *Protein Eng. 6*, 85−91.
Ishikawa, K., Nakamura, H., Morikawa, K., & Kanaya, S. (1993b) *Biochemistry 32*, 6171−6178.
Kanaya, S., & Crouch, R. J. (1983) *J. Biol. Chem. 258*, 1276−1281.
Kanaya, S., & Itaya, M. (1992) *J. Biol. Chem. 267*, 10184−10192.
Kanaya, S., & Ikehara, M. (1994) *Subcell. Biochem. 24*, 377−422.
Kanaya, S., Kohara, A., Miyagawa, M., Matsuzaki, T., Morikawa, K., & Ikehara, M. (1989) *J. Biol. Chem. 264*, 11546−11549.
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.
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.
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.
Laemmli, U. K. (1970) *Nature 227*, 680−685.
Meiering, E. M., Serrano, L., & Fersht, A. R. (1992) *J. Mol. Biol. 225*, 585−589.
Miller, J. H. (1972) in *Experiments in Molecular Genetics* (Miller, J. H., Ed.) pp 433, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.
Oda, Y., Yamazaki, T., Nagayama, K., Kanaya, S., Kuroda, Y., & Nakamura, H. (1994) *Biochemistry 33*, 5275−5284.
Oobatake, M., Hashimoto, T., Nakamura, H., & Kanaya, S. (1993) *Protein Eng. 6*, 1023.
Pantoliano, M. W., Whitlow, M., Wood, J. F., Dodd, S. W., Hardman, K. D., Rollence, M. L., & Bryan, P. N. (1989) *Biochemistry 28*, 7205−7213.
Sanger, F., Nicklen, S., & Coulson, A. R. (1977) *Proc. Natl. Acad. Sci. U.S.A. 74*, 5463−5467.
Stearman, R. S., Frankel, A. D., Freire, E., Liu, B., & Pabo, C. O. (1988) *Biochemistry 27*, 7571−7574.
Wells, J. A. (1990) *Biochemistry 29*, 8509−8517.
Yang, W., Hendrickson, W. A., Crouch, R. J., & Satow, Y. (1990) *Science 249*, 1398−1405.
Zhang, X.-J., Baase, W. A., & Matthews, B. W. (1992) *Protein Sci. 1*, 761−776.

BI950191W