Role of Proline Residues in Human Lysozyme Stability: A Scanning Calorimetric Study Combined with X-ray Structure Analysis of Proline Mutants†

Thierry Herning,‡,§ Katsuhide Yutani,‖ Koji Inaka,‡,⊥ Ryota Kuroki,‡,▽ Masaaki Matsushima,‡ and Masakazu Kikuchi*,‡

Protein Engineering Research Institute, 6-2-3, Furuedai, Suita, Osaka 565, Japan, and Institute for Protein Research, Osaka University, 3-2, Yamadaoka, Suita, Osaka 565, Japan

Received January 10, 1992; Revised Manuscript Received April 8, 1992

ABSTRACT: It has been shown that protein stability can be modulated from site-directed mutations that affect the entropy of protein unfolding [Matthews, B. W., Nicholson, H., & Becktel, W. J. (1987) *Proc. Natl. Acad. Sci. U.S.A. 84*, 6663–6667]. However, the effect of a specific amino acid replacement on stability highly depends on the location of the mutation site and its environment in the protein structure [Yutani, K., Hayashi, S., Sugisaki, Y., & Ogasahara, K. (1991) *Proteins Struct., Funct., Genet. 9*, 90–98). To clarify the role of specific proline residues in the thermostability of human lysozyme (h-lysozyme), a series of proline mutants were investigated by means of scanning calorimetry and high-resolution X-ray crystallography. The thermodynamic properties of the mutant and wild-type h-lysozymes are compared and discussed on the basis of their three-dimensional structure. h-Lysozyme contains two proline residues at positions 71 and 103. The Pro71 → Gly substitution was found to destabilize h-lysozyme by decreasing the entropic contribution of unfolding by about 2 kcal/mol at 68.8 °C. This is consistent with the theoretical expectations for such a substitution. However, the same substitution at position 103 (Pro103 → Gly) does not affect h-lysozyme stability, and the thermodynamic properties of the P71G/P103G and P71G mutants are essentially the same. Pro71 which is conserved among lysozymes from other species, appears to be important for stability, whereas Pro103, which is not conserved, does not. These differences are explained in terms of residue accessibility to the solvent and crystallographic *B*-factor, which reflects the amino acid mobility. Other mutants containing three prolines (A47P, D91P, and V110P) are also examined. These mutation sites were selected in order to minimize the conformational perturbations caused by the introduction of a trans proline in the protein sequence. All of these mutants are entropically stabilized, but the entropic effects are often masked by concomitant changes in the enthalpy of denaturation. The overall three-dimensional structures of the mutant and wild-type h-lysozymes are essentially identical; only slight local structural readjustments are found around some of the mutation sites.

Understanding the physical basis of the stability of proteins is crucial not only to the understanding of their structure and function but also to elucidate how the unique conformation of the folded state can be acquired. Such knowledge would be particularly useful for protein engineering or protein design where there is a considerable interest in enhancing protein stability.

However, the mechanism of protein stabilization remains one of the difficult problems of protein chemistry due to the great number of factors involved in that process and the lack of experimental methods allowing the evaluation of their respective contributions. The overall net free energy of stabilization of proteins is usually small; e.g., the native state of a typical globular protein is only about 10 kcal/mol more stable than the unfolded state. However, this marginal stability results from a delicate balance between a great number of weak interactions, globally corresponding to a stabilization of quite high energy (≈100 kcal/mol), and the destabilization arising from the loss of conformational entropy of the folded polypeptide chain. At the moment, there is no direct way to evaluate the contribution of all these interactions in the stabilization mechanism.

In spite of that complexity, structural and thermodynamic comparisons of genetically altered proteins have recently provided new insights about the role of individual amino acids in determining the structure and stability of globular proteins (Alber, 1989). For instance, considerations introduced several years ago about the influence of amino acid side chains on the entropy of protein unfolding (Nemethy et al., 1966) have been confirmed more recently on bacteriophage T4 lysozyme (Matthews et al., 1987) by showing that protein thermostability can be increased from site-directed mutations that decrease the entropy of unfolding. It was also reported that tryptophan synthase α subunit can be more or less destabilized by the replacement of conserved prolines by alanine or glycine (Yutani et al., 1991), depending on the backbone mobility of the folded protein at the mutation site. Increasing protein stability from selected amino acid substitutions that decrease the conformational entropy of unfolding is an attractive approach because it implies that the unfolded state is more affected by the mutation than the folded state. Ideally, the folded state, whose properties are essential to protein function, would be unchanged. The entropic contribution to the unfolded state

† The coordinates for human lysozyme mutants A47P, D91P, P103G, P71G, and V110P have been deposited in the Brookhaven Protein Data Bank.

* To whom correspondence should be addressed.

‡ Protein Engineering Research Institute.

§ Present address: Laboratoire de Cristallographie Biologique, Institut de Biologie Moléculaire et Cellulaire du CNRS, 15 rue René Descartes, 67084 Strasbourg Cedex, France.

‖ Osaka University.

⊥ Present address: Research Laboratory of Resources Utilization, Tokyo Institute of Technology, 4259 Nagatsuta, Midori-ku, Yokohama 227, Japan.

▽ Present address: Central Laboratories of Key Technology, Kirin Brewery Co. Ltd., 1-13-5, Fukuura, Kanazawa-Ku, Yokohama 236, Japan.

0006-2960/92/0431-7077$03.00/0

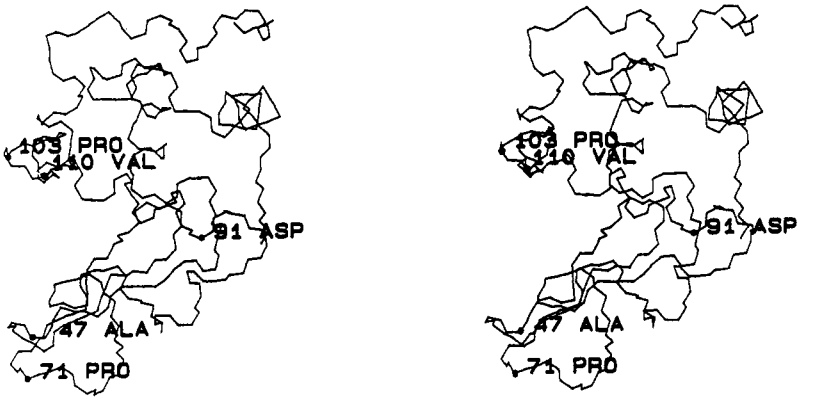

FIGURE 1: Stereo drawing of the human lysozyme structure. Amino acids that have been substituted are indicated.

free energy depends on the nature of amino acid side chains because the freedom of internal rotation of the backbone in the random coil form is influenced by the length or bulkiness of the side chains. Such effect is particularly important in the case of proline substitutions due to the pyrrolidine ring of proline, which restricts this residue to fewer conformations than are available to the other amino acids (Nemethy et al., 1966).

In the present study, we focus on the role of proline residues in the stability of a model protein: human lysozyme (h-lysozyme). h-Lysozyme is a typical globular protein (130 amino acids, 4 disulfide bonds) that contains only two proline residues at positions 71 and 103; both of them are in the trans conformation in the native state (Artymiuk & Blake, 1981). In order to investigate separately the role of specific prolines in h-lysozyme stability, a series of proline mutants were produced and characterized. The h-lysozyme mutants examined in this work are P71G (Pro71 is replaced by Gly), P103G (Pro103 is replaced by Gly), and P71G/P103G (Pro71 and Pro103 are replaced by Gly). Other mutants such as A47P, D91P, and V110P (Ala47, Asp91 and Val110 are replaced by Pro, respectively) were also investigated to extend our study. The mutation sites were selected in order to minimize the conformational perturbations caused by the introduction of a trans proline in the protein sequence. Only one proline remains in the P71G and P103G mutants and no proline in the P71G/P103G mutant, whereas the A47P, D91P, and V110P mutants contain three proline residues. As shown in Figure 1, Pro71 and Pro103 are located in loop regions, Asp91 and Val110 are located in α-helices near the N-terminal region, and Ala47 is in a β-turn. All these mutant proteins have already been introduced in a previous work relative to the role of proline residues on h-lysozyme folding kinetics (Herning et al., 1991). In this second part, thermodynamic properties of each protein are determined by means of scanning calorimetry and their three-dimensional structures are solved at high resolution by X-ray crystallography. The calorimetric technique provides the most accurate and reliable way for the determination of the basic thermodynamic functions associated with protein denaturation. The results are discussed in light of the high-resolution X-ray structural data of the crystalline proline mutant and wild-type h-lysozymes. The contribution of proline residues to configurational entropy is examined, taking into account their environment in the protein structure.

MATERIALS AND METHODS

Materials. Wild-type h-lysozyme was purchased from Sigma (St. Louis, MO). All chemicals were reagent grade and were used without further purification.

Purification of h-Lysozyme Variants. Mutations in the h-lysozyme gene were generated by oligonucleotide-directed mutagenesis. The mutants used for this study were P71G, P103G, P71G/P103G, A47P, D91P, and V110P. Mutant lysozymes were produced as described previously (Herning et al., 1991), using *Saccharomyces cerevisiae* AH22R$^-$ (a *leu2 his4 can1 pho80*) (Minohara et al., 1983) as a host strain. The genes encoding the signal sequence and mutated h-lysozymes were combined with a *Xho*I–*Sma*I large fragment from pERI8602 (Taniyama et al., 1988) to construct the expression plasmids. Secreted h-lysozyme mutants were purified as previously described (Kikuchi et al., 1988).

Calorimetric Measurements. Calorimetric measurements were carried out with a scanning microcalorimeter, DASM4, equipped with an NEC personal computer. The scan rate was 1.0 deg/min for all the experiments. Each memory unit in the computer stored one datum for 5 s. Protein solutions were prepared just before measurements by dissolving lyophilized proteins, stored at −20 °C, in 50 mM glycine hydrochloride buffer at various pHs ranging from 1.8 to 4.5. The pH of protein solutions was determined before measurements and was checked again after measurements. Protein concentration was carefully determined by measuring the absorbance at 280 nm, taking ϵ_{280} = 37 700. Protein concentrations were within the range 0.8–3.5 mg/mL. Calorimetric and van't Hoff enthalpies (ΔH_d^{cal} and ΔH_d^{VH}, respectively) were calculated using a computer program developed by Kidokoro and Wada. The calorimetric enthalpy is calculated without making any assumption on the nature of the phase transition. It is derived from the area of the heat absorption peak using an electric calibration [see Figure 2 in Privalov and Khechinashvili (1974)]. The peak area is taken as the area limited from above by the heat absorption curve and from below by the heat capacity values of the native and denatured protein obtained by linear extrapolation of heat capacity before and after the process to the middle of the transition at the transition temperature T_d. The van't Hoff enthalpy is also calculated simultaneously from the same calorimetric curve. However, for this it is necessary to assume that the denaturation is a two-state transition process. The van't Hoff enthalpy (ΔH_d^{VH}) is calculated from

$$\Delta H_d^{VH} = 4RT_d^2\Delta C_p/Q_d \quad (1)$$

where ΔC_p is the height of the heat capacity peak at denaturation temperature, T_d, and Q_d is the area of the peak of the excess heat capacity curves [eq 5 in Privalov and Khechinashvili (1974)]. By comparing the calorimetric and van't Hoff enthalpies one can estimate how close to a two-state transition the measured denaturation process is. More details

Table I: Crystal Structure Analysis of Wild-Type and Mutant h-Lysozymes[a]

(a) Data Processing

protein	cell dimensions (Å)			no. of reflections		R_{merg} (%)	R_{sym} (%)	completeness (%)
	a	*b*	*c*	independent	total			
wild type	56.50	60.89	33.83	11487	35801	4.98	2.80	87.0 (80.4)
P71G	56.59	60.71	33.83	9952	25987	5.92	3.99	87.6 (80.3)
P103G	56.46	61.21	33.84	9358	39581	3.40	2.13	82.0 (80.9)
A47P	56.55	61.00	33.89	10426	38695	5.64	3.93	91.4 (85.7)
D91P	56.68	61.04	33.86	10105	39225	4.80	2.51	88.3 (79.2)
V110P	56.82	60.81	33.85	10232	29639	5.90	3.80	81.4 (76.6)

(b) Refinement Statistics

protein	atom no.	*R* values (%)		average deviation (Å) from ideal values			average *B* factor (Å^2)		
		overall	highest shell	bond length	bond angle	planarity	main chain	side chain	solvent
wild type	1144	16.3	22.3	0.016	0.034	0.013	10.9	14.9	27.9
P71G	1149	15.6	19.4	0.017	0.037	0.013	13.4	17.4	32.1
P103G	1154	15.2	18.7	0.017	0.038	0.014	11.7	16.0	30.7
A47P	1146	15.4	18.5	0.013	0.034	0.012	11.6	15.7	30.3
D91P	1143	15.3	20.0	0.016	0.036	0.012	12.2	16.4	30.4
V110P	1130	16.0	21.9	0.016	0.038	0.013	10.0	13.4	23.4

[a] The highest shell goes from 1.89-Å to 1.8-Å resolution. The completeness in this shell is listed in parentheses for each protein. $R_{merg} = \Sigma(I - \langle I \rangle)/\Sigma\langle I \rangle$, $R_{sym} = \Sigma(\langle I^+ \rangle - \langle I^- \rangle)/\Sigma(\langle I^+ \rangle + \langle I^- \rangle)$, and $R = \Sigma||F_o| - k|F_c||/\Sigma|F_o|$, where I and $\langle I \rangle$ are the measured and equivalent average intensities, respectively, $\langle I^+ \rangle$ and $\langle I^- \rangle$ are the Bijvoet pair average intensities, and F_o and F_c are the observed and calculated structure factors, respectively. k is the scale factor.

concerning the data analysis procedure on which this program is based can be found in a previous study (Kidokoro & Wada, 1987).

Crystallography. The method used for the crystallization was the same for the wild-type and mutant h-lysozymes and has already been described (Inaka et al., 1991). Since pH is a key parameter for calorimetric study, we restate that the crystallization pH was 5.8. The three-dimensional X-ray structure of each mutant, except P71G/P103G, was determined at high resolution. All of the mutants crystallize in the space group $P2_12_12_1$, similarly to wild-type h-lysozyme. The diffraction data were collected at 9 °C up to 1.8-Å resolution with only one crystal for each mutant by an automated oscillation camera system (Dip-100, Mac Science) equipped with an "imaging plate" (Fujii et al., 1991). The X-ray radiation generated at 50 kV, 90 mA, was obtained from a Cu target and was monochromatized by graphite. The crystal was oscillated by 2.0 deg/frame at a rotation speed of 0.002 deg/s, which gives 33.3 min/frame. The overall exposure time per data set was about 28 h. The diffraction intensities were evaluated by the program system ELMS (Tanaka et al., 1990) and processed at 1.8-Å resolution by the program PROTEIN (Steigemann, 1974). The ratio $I/\Sigma I$ was 3, where I is estimated from the square root of the summation of the photon counts at peak pixels and background counts. To check if protein deterioration occurred within data collection, the scale factors relative to the different frames were compared. For D91P the scale factor for the intensities increased from 1.00 to 1.11 between the first and the last frame. For the other mutants the scale factor was basically unchanged between the first and the last frame (less than 3%). Therefore protein deterioration was considered as limited during data collection. The crystal structures were solved by the molecular replacement method, where the initial molecular model was the independently rerefined structure of wild-type h-lysozyme (Inaka et al., 1991). The model was fitted to the crystal structure of the mutants by rigid body refinements using the program TRAREF (Huber & Schneider, 1985). The substituted amino acid was identified and positioned with the computer graphics system [Evans & Sutherland PS390, operated by the program FRODO (Jones, 1978)]. Further refinements were carried out by the program package PROTIN/PROLSQ (Hendrickson, 1985). The total numbers of solvent molecules in the final structure models were 119, 120, 132, 125, and 101 for A47P, P71G, D91P, P103G, and V110P, respectively. The structures were refined against the structure factors between 5.0- and 1.8-Å resolutions and the *R*-factors were reduced to 0.154, 0.156, 0.153, 0.152, and 0.160 for the mutant structures of A47P, P71G, D91P, P103G, and V110P, respectively. The crystal data and refinement statistics are listed in Table I.

RESULTS

Calorimetric Results. To evaluate the contribution of specific amino acids to protein stability, one needs to compare the thermodynamic data relative to the denaturation of point-mutated proteins to those of the wild-type protein. Most point mutations are likely to induce relatively small changes in protein stability (Privalov, 1979). Therefore, it is necessary to determine very accurately the various thermodynamic functions describing the denaturation process of the protein. To date, the most direct and accurate way to determine these thermodynamic functions is provided by the scanning calorimetry technique. In this section, we present the thermodynamic results obtained by means of this technique on various mutant and wild-type h-lysozymes. Figure 2 shows typical excess heat capacity curves resulting from calorimetric recordings of mutant and wild-type h-lysozymes at different pHs in the vicinity of pH 2.8. Both the pre- and posttransitional baselines vary linearly with temperature, and the transition curves appear to be single peak and symmetrical. Each protein considered in this study gave an excess heat capacity curve with similar characteristics. Thermal denaturation of h-lysozyme was found to be reversible. This was confirmed by reheating the protein solution in the calorimeter cell immediately after cooling from the first run. Repeated thermal scans of the samples exhibited similar transition profiles. To determine the transition temperature, T_d, the transition enthalpy, ΔH_d^{cal}, and the change in the heat capacity, $\Delta_d C_p$, the excess heat capacity versus temperature profiles were analyzed as described under Materials and Methods.

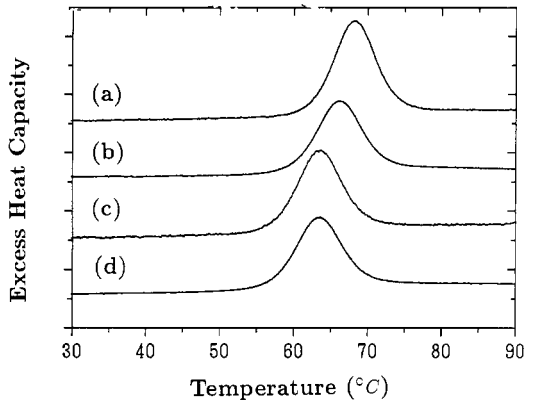

FIGURE 2: Typical excess heat capacity curves for (a) the wild-type h-lysozyme (pH 2.81, 2.57 mg/mL), (b) the P103G mutant (pH 2.72, 2.2 mg/mL), (c) the P71G mutant (pH 2.80, 1.43 mg/mL), and (d) the P71G/P103G mutant (pH 2.76, 2.69 mg/mL). The increment of excess heat capacity is 5 kcal·mol^{-1}·K^{-1}.

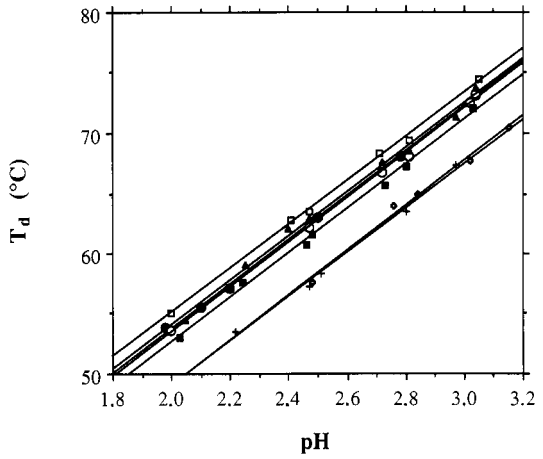

FIGURE 3: Plot of the denaturation temperature, T_d, versus pH for (●) the wild type, (+) P71G, (○) P103G, (◇) P71G/P103G, (■) D91P, (□) V110P, and (▲) A47P.

The validity of the two-state model could be tested because scanning calorimetric recordings provide simultaneously not only the real (calorimetric) enthalpy associated with the conformational transition but also the effective (van't Hoff) enthalpy of the transition. It has been shown (Privalov, 1979) that when the van't Hoff enthalpy, $\Delta_d H^{VH}(T)$, is smaller than the calorimetric enthalpy, $\Delta_d H^{cal}(T)$, over the entire temperature range, it indicates that the transition is a multistate one, while if $\Delta_d H^{VH}(T)$ is greater than $\Delta_d H^{cal}(T)$, it means that there exist some intermolecular interactions. As reported in Table II, in our experiments that ratio was close to unity for all of the proteins, indicating that we have a characteristic two-state denaturation process in every case (Privalov, 1979).

Calorimetric experiments have been carried out for each protein at various pHs ranging from 1.8 to 4.5. The pH range was chosen far from the isoelectric point of h-lysozyme (≈11) in order to avoid aggregation at high temperature. The buffer used in all experiments was glycine hydrochloride since the enthalpy of ionization of this buffer is close to that of the protein groups. This ensures practically complete compensation of the protein ionization heat effects by the buffer ionization heat effects during denaturation so that the measured enthalpy of transition is essentially the enthalpy of the conformational change of the protein (Privalov & Khechinashvili, 1974). The transition temperature is very sensitive to pH variations. For all the proteins it was found to vary linearly with the pH over the range 1.8–3.0, as shown in Figure 3. This increase has been reported on other globular proteins

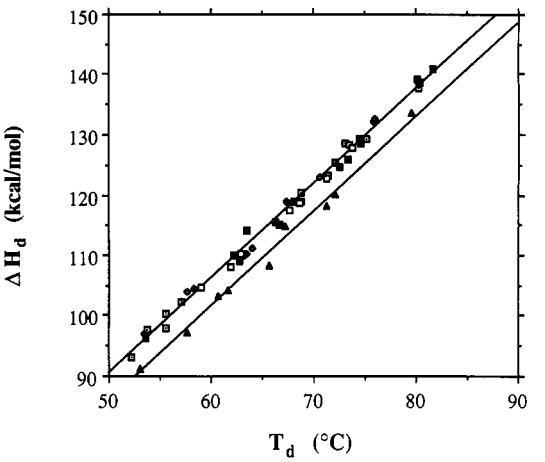

FIGURE 4: Plot of the calorimetric enthalpy of denaturation as a function of the denaturation temperature for (⧈) the wild type, (◆) P71G, (◘) P103G, (◈) P71G/P103G, (▲) D91P, (■) V110P, and (□) A47P.

and has been ascribed to the stabilization of the native protein brought about by the decrease in its net positive charge as the solution become less acidic (Tsong et al., 1970).

In principle, one could obtain all the basic thermodynamic functions describing the thermal denaturation of a protein from a single calorimetric recording at a given pH. The parameters that are directly measured from the recording are the transition temperature, T_d, the calorimetric enthalpy change, ΔH_d^{cal}, and the heat capacity change, $\Delta_d C_p$. If $\Delta_d C_p$ does not depend on temperature, changes in enthalpy, $\Delta_d H(T)$, entropy, $\Delta_d S(T)$, and Gibbs energy, $\Delta_d G(T)$, on thermal denaturation can be calculated as a function of temperature using the following equations (Privalov & Khechinashvili, 1974):

$$\Delta_d H(T) = \Delta H_d^{cal} - \Delta_d C_p(T_d - T) \quad (2)$$

$$\Delta_d S(T) = \Delta H_d^{cal}/T_d - \Delta_d C_p \ln (T_d/T) \quad (3)$$

$$\Delta_d G(T) = \Delta_d H(T) - T\Delta_d S(T) \quad (4)$$

where ΔH_d^{cal} is the measured calorimetric enthalpy change at the transition temperature T_d. However, the reliability of the calculated parameters highly depends on the accuracy of $\Delta_d C_p$, and to a lower extent on ΔH_d^{cal} and T_d. If one wishes to discuss small stability changes upon protein mutation, it is necessary to determine these thermodynamic parameters with the highest possible accuracy. This is done by performing calorimetric measurements at various pHs and plotting ΔH_d^{cal} versus T_d. ΔH_d^{cal} usually varies linearly with T_d and the slope determined by least-squares analysis gives $\Delta_d C_p$ with a reliability which increases with the data point number. In our experiments the relative error on the measured ΔH_d^{cal} was estimated to be approximately 2–4% from the standard deviations of the data points.

The plots of ΔH_d^{cal} versus T_d of the mutant h-lysozymes are compared to that of the wild type in Figure 4. Except for D91P, the T_d, ΔH_d^{cal} data points for all the mutant and wild-type h-lysozymes are superimposed on the same straight line. The experimental data obtained from calorimetric recordings at various pH are summarized in Table II. The thermodynamic parameters of denaturation for the wild-type and mutant proteins are compared in Table III at a reference temperature of 68.8 °C, the denaturation temperature of the wild-type protein at pH 2.80. For each mutant, T_d at pH 2.80 was

Table II: Thermodynamic Parameters Obtained by Calorimetry of Wild-Type and Substituted h-Lysozymes

protein	pH	T_d (°C)	ΔH_d^{cal} (kcal/mol)	ΔH_d^{VH} (kcal/mol)	$\Delta H_d^{cal}/\Delta H_d^{VH}$	average
wild type	4.51	80.3	138.51	145.80	0.95	0.94
	4.50	80.3	137.67	144.92	0.95	
	3.01	72.1	125.38	130.61	0.96	
	4.49	80.4	138.78	146.08	0.96	
	2.81	68.8	118.90	122.58	0.97	
	3.03	71.4	123.33	128.47	0.96	
	3.50	75.1	129.16	137.41	0.94	
	2.10	55.5	98.06	104.32	0.94	
	3.05	73.5	128.43	138.10	0.93	
	2.80	68.8	120.34	132.24	0.91	
	2.50	63.0	109.90	122.11	0.90	
	2.00	55.6	100.41	105.70	0.95	
	2.20	57.1	102.21	109.91	0.93	
	1.80	52.2	93.24	99.19	0.94	
	3.04	73.1	128.51	135.28	0.95	
	2.78	68.1	118.87	126.46	0.94	
	1.98	53.8	97.65	100.67	0.97	
P71G	2.97	67.3	118.96	120.17	0.99	0.98
	2.51	58.3	104.46	108.82	0.96	
	2.22	53.5	96.97	98.95	0.98	
	2.80	63.5	110.25	116.06	0.95	
	4.49	75.8	132.16	133.43	1.01	
P103G	3.02	72.5	125.75	122.31	1.02	0.97
	2.72	66.7	115.04	116.20	0.99	
	3.04	73.3	126.00	131.40	0.96	
	2.47	62.2	110.05	115.84	0.95	
	2.81	68.1	119.00	128.28	0.95	
	2.00	53.6	96.31	101.38	0.95	
	2.74	66.2	115.49	117.85	0.98	
	4.50	80.1	139.20	147.70	0.94	
P71G/ P103G	4.50	76.0	132.62	136.72	0.97	0.99
	3.15	70.5	123.01	119.42	1.03	
	2.76	64.0	111.30	112.42	0.99	
	2.48	57.6	104.00	108.97	0.95	
D91P	2.24	57.6	97.35	102.47	0.95	0.94
	2.03	53.0	91.26	98.13	0.93	
	3.03	72.1	120.06	123.77	0.97	
	2.80	67.2	114.83	120.87	0.95	
	2.48	61.6	104.23	115.81	0.90	
		71.3	118.08	121.73	0.97	
	2.73	65.7	108.20	115.11	0.94	
	2.46	60.7	103.25	111.02	0.93	
	4.47	79.6	133.50	146.70	0.91	
A47P	2.97	71.3	122.70	125.46	0.97	0.97
	2.25	59.0	104.79	112.68	0.93	
	2.81	68.6	118.63	122.30	0.97	
	3.04	73.7	127.90	128.44	0.98	
	2.47	62.9	110.17	113.30	0.99	
	2.72	67.6	117.53	122.43	0.96	
	2.40	62.0	108.19	114.78	0.96	
V110P	3.10	74.6	128.50	133.85	0.96	0.94
	2.41	62.8	109.12	118.61	0.92	
	3.05	74.4	129.21	133.21	0.97	
	2.47	63.5	114.00	122.58	0.93	
	4.50	81.6	140.79	156.43	0.90	

obtained from the least-squares fit of the data of Figure 3, which corresponds to the linear range of the curves. $\Delta_d C_p$ and ΔH_d^{cal} at pH 2.80 were derived from the linear least-squares fit of the data of Figure 4. Thermodynamic functions at T = 68.8 °C could then be calculated by using eqs 2–4. These results show that P71G and P71G/P103G mutants have the same stability and are both 1.6 kcal/mol less stable than wild-type h-lysozyme. This destabilization can be ascribed to the important increase in the entropic component of the free energy of denaturation caused by the mutations: about 2 kcal/mol at 68.8 °C for both mutants. The mutation Pro103 → Gly does not affect significantly the overall protein stability. The V110P mutant and to a lesser extent the A47P mutant are more stable than the wild-type protein, whereas the D91P mutant is a little bit destabilized as compared to the wild type. This last mutation greatly stabilizes the protein through entropic effects (about 4.3 kcal/mol at 68.8 °C), but this effect is completely offset by a greater enthalpic destabilization, resulting in a decrease in the overall stability of that mutant (Table III).

Crystallography. Comparison of the properties of the wild-type and mutant h-lysozymes depends critically on the assumption that the amino acid substitutions do not affect significantly the three-dimensional structure of the protein. The validity of this assumption was tested by solving at high resolution the X-ray structure of each mutant except P71G/P103G. The overall three-dimensional structure of all the mutants was found to be quite similar to that of the wild type. The root-mean-square discrepancies of the main-chain atoms (N, C_α, C, and O) between the wild-type and mutant h-lysozymes were found to be 0.192, 0.241, 0.092, 0.055, and 0.129 Å for P71G, P103G, D91P, A47P, and V110P, respectively. These values are very low. However, some local and small structural readjustments are observed at some of the mutation sites. The difference in the coordinates of the corresponding C_α atoms was slightly lower than 0.5 Å for Pro71 → Gly and Pro103 → Gly substitution sites, 0.3 Å for Val110 → Pro, and lower than 0.1 Å for the other mutations (data not shown). In Figure 5 are shown the stereo drawings of the mutant structure around their mutation site, superimposed on the wild-type structure. In Table IV are given the Φ, Ψ dihedral angles at the mutation site for each mutant along with their values in wild-type h-lysozyme. For the Xaa → Pro substitutions, the Φ, Ψ values are also given for the residue preceding the mutation site because a proline residue not only restricts the Φ, Ψ values at the proline itself but also limits the degree of freedom of the preceding residue. The Xaa → Pro substitutions were selected in such a way that they were compatible with these contraints. For the mutants A47P, D91P, and V110P, Pro47, Pro91, and Pro110 have Φ, Ψ values within the allowed range for a proline (Schimmel & Flory, 1968) and at the preceding sites the Φ, Ψ values are essentially the same as for the wild type.

In Figure 6 are shown the differences in the average crystallographic *B*-factors of the main-chain atoms between the mutant and wild-type protein for the P71G and P103G mutants. We can see that there is an increase in the *B*-factor around both of the Pro → Gly substitutions as compared with the wild type. However, the relative increase in the *B*-factor is significantly higher for P103G than for P71G.

DISCUSSION

It has been reported that thermostability of proteins can be modulated by modifying their backbone conformational flexibility by means of site-directed mutagenesis (Matthews et al., 1987). However, the effect of a particular amino acid replacement on protein stability highly depends on the location of the mutation site and its environment in the protein structure (Yutani et al., 1991). In many cases, point mutations that do have a large effect on protein stability occur at buried or rigid sites (Alber, 1989). Here, we compare a series of proline mutants of h-lysozyme and discuss their thermostability in the light of their high-resolution X-ray three-dimensional structures. Comparison of the properties of the wild-type and mutant proteins implies that amino acid substitutions do not alter significantly protein conformation. This requirement was found to be satisfied for all the mutants as shown by the X-ray data under Results. The overall structure of the mutants is basically identical to that of the wild type; only slight structural readjustments are observed at a very local scale

Table III: Thermodynamic Parameters Obtained by Calorimetry for the Wild-Type and Substituted h-Lysozymes at the Denaturation Temperature of the Wild Type at pH 2.80 (68.8 °C)[a]

protein	T_d (°C)	ΔT_d (°C)	$\Delta_d C_p$	$\Delta_d H$	$\Delta_d S$	$T\Delta_d S$	$\Delta_d G$	$\Delta\Delta_d H$	$T\Delta\Delta_d S$
wild type	68.8	0	1.575	120.1	351.4	120.1	0	0	0
P71G	64.1	−4.7	1.584	120.7	357.8	122.3	−1.6	+0.6	+2.2
P103G	68.7	−0.1	1.579	119.7	350.0	119.8	−0.1	−0.4	−0.3
P71G/P103G	64.3	−4.5	1.578	120.5	357.2	122.1	−1.6	+0.4	+2.0
D91P	67.7	−1.1	1.571	115.4	338.7	115.8	−0.4	−4.7	−4.3
A47P	69.1	+0.3	1.562	119.4	349.0	119.3	+0.1	−0.7	−0.8
V110P	70.0	+1.2	1.572	120.3	350.6	119.8	+0.5	+0.2	−0.3

[a] $\Delta_d H$, $T\Delta_d S$, and $\Delta_d G$ are calculated at T = 341.8 K by using eqs 2–4 and are given in kilocalories per mol. T_d at pH 2.80 was estimated from the pH dependence of T_d (Figure 3). The $\Delta_d H$ values at T_d and $\Delta_d C_p$ were obtained by a linear least-squares fit of the experimental data for each protein (Table II). $\Delta_d C_p$ and $\Delta_d S$ are in kilocalories per mole per kelvin, and $\Delta\Delta_d H$ and $T\Delta\Delta_d S$ are in kilocalories per mole and refer to $\Delta_d H' - \Delta_d H$ and $T(\Delta_d S' - \Delta_d S)$, respectively (the prime denotes a mutant).

around some of the mutation sites. In this section, we shall first discuss the results relative to the h-lysozyme mutants in which one of the two native prolines or both of them have been replaced by a glycine (Pro → Gly mutants). In the next section we shall focus on mutants into which a third proline was introduced (Xaa → Pro mutants). These substitutions were done at sites which could potentially accommodate the constraints introduced by a proline residue.

Pro → Gly Mutants: P71G, P103G, and P71G/P103G. To analyze the individual contribution on the entropy of unfolding of the two proline residues of h-lysozyme, the single-proline mutants, P71G and P103G, and the proline-free mutant, P71G/P103G, were designed. If no secondary effects are introduced by the mutations, such amino acid substitutions are expected to destabilize the protein due to the increase in the entropy of unfolding they induce. The entropic contribution to the unfolded state free energy depends on the nature of the amino acid side chains because the freedom of internal rotation of the backbone in the random coil form is influenced by the length or the bulkiness of the side chains. For instance, a glycine, which lacks a β-carbon, has more backbone conformational flexibility than any other amino acid residue, whereas the pyrrolidine ring of proline restricts this residue to fewer conformations than are available to the other amino acids. By considering the area accessible to a given amino acid in a Ramachandran conformational map (Ramachandran et al., 1963), the backbone contribution to the entropy of unfolding of a proline relative to a glycine was estimated to be about −1.9 kcal/mol at 25 °C (Nemethy et al., 1966).

Within the pH range investigated, the P71G h-lysozyme mutant is found to be less stable than the wild-type protein, which is in agreement with the theoretical expectations. At pH 2.80 the denaturation temperature of P71G is lower than that of the wild type by 4.7 °C, corresponding to a net decrease of 1.6 kcal/mol of the free energy of unfolding (Table III). The plots of ΔH_d^{cal} versus T_d give straight lines with essentially identical intercepts and slopes (e.g., similar $\Delta_d C_p$) for both the P71G and wild-type proteins, indicating that the stability difference has rather an entropic origin than enthalpic. This is confirmed by the calculated values of the thermodynamic functions summarized in Table III. At a reference temperature of 68.8 °C, corresponding to the denaturation temperature of the wild type at pH 2.80, the unfavorable entropic contribution to the net free energy, $T\Delta S$, is higher by 2.2 kcal/mol for the P71G mutant as compared to the wild type. It is precisely the value that is theoretically expected for the entropic energy variation induced by the replacement of a proline by a glycine (ΔS = 6.4 cal·mol^{-1}·K^{-1}) at a temperature of 68.8 °C (Nemethy et al., 1966). The entropic destabilization is accompanied with an enthalpic stabilization which is, however, small in comparison (0.6 kcal/mol).

Structural data obtained from the X-ray analysis are consistent with these results. The overall three-dimensional structure of the P71G mutant is basically identical to that of the wild-type protein. The root-mean-square deviation for the main-chain atoms between the wild type and P71G is 0.19 Å. Nevertheless, locally the mutation causes some structural relaxations which are reflected in the changes of the dihedral angles Φ, Ψ at site 71 between the wild-type (−71, +137) and P71G mutant (+99, +106) proteins. Such structural readjustments might explain the 0.6 kcal/mol difference in the enthalpic term.

Although the destabilization induced by replacing Pro71 by Gly can be explained by entropic effects, the fact that P103G is as stable as the wild type requires a more careful analysis. For the P103G mutant, the linear fit of ΔH_d^{cal} versus T_d gives similar coefficients to those of P71G or the wild-type proteins, and within experimental error, all the thermodynamic functions are essentially identical to those of the wild type at a reference temperature of 68.8 °C as shown in Table III. No significant changes in the three-dimensional structure of the P103G mutant are observed from the X-ray analysis, except locally as was the case for the P71G mutant. These results indicate that there is no important modification in the overall entropy of unfolding for Pro103 → Gly substitution; neither is the enthalpy of unfolding significantly affected. It is likely that the effect of the replacement of a proline by a glycine on the unfolded state free energy is similar whether the mutation occurs at site 71 or 103. Therefore, it is reasonable to speculate that the replacement of Pro by Gly affects differently the entropy of the folded state depending whether it takes place at position 71 or 103.

The relative entropy of unfolding of a different type of residue, z, relative to y, ΔS_{zy}, was estimated by using (Nemethy et al., 1966)

$$\Delta S_{zy} = \Delta S_{conf}(z) - \Delta S_{conf}(y) = R \ln (\gamma_z/\gamma_{nz}) - R \ln (\gamma_y/\gamma_{ny}) \quad (5)$$

where $\Delta S_{conf}(z)$ is the entropy of unfolding for residue z and γ_z and γ_{nz} are the numbers of conformations available to residue z in the unfolded and folded forms, respectively. R is the gas constant. If $\gamma_{nz} \approx \gamma_{ny}$, eq 5 becomes $\Delta S_{zy} = R \ln (\gamma_z/\gamma_y)$. The approximation $\gamma_{nz} \approx \gamma_{ny}$ is apparently not correct for the P103G mutant. It is likely that for the P103G mutant the folded state is entropically stabilized to the same extent as the unfolded state. Both states have a lower free energy but the net free energy difference is not changed and P103G remains as stable as the wild-type protein.

This interpretation is supported by the structural data. The accessible surface area (ASA) values, calculated from the wild-type X-ray structure of h-lysozyme, are 146.05 Å^2 for

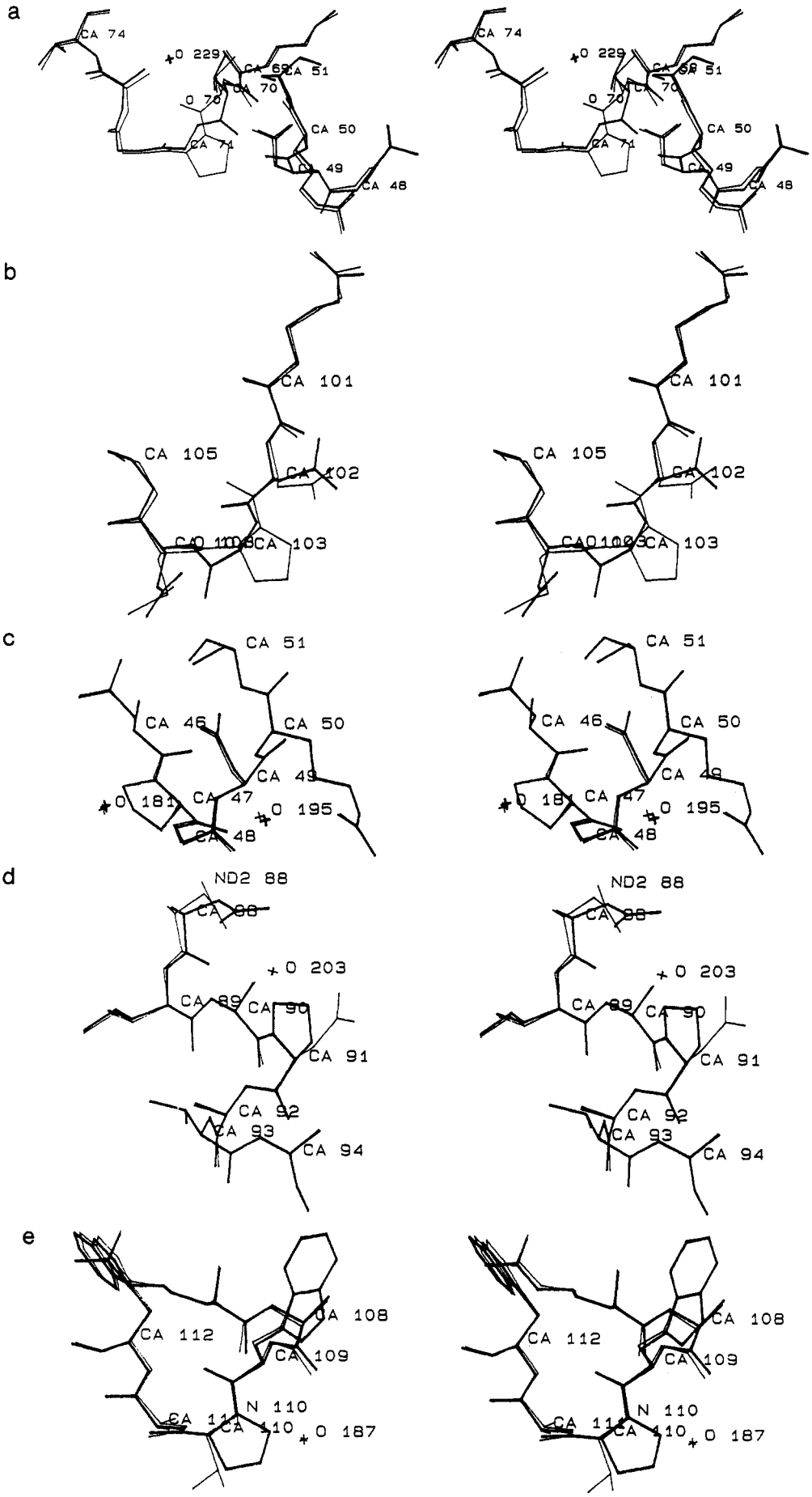

FIGURE 5: Stereo drawings showing the mutant structure in the vicinity of the mutation sites. Wild-type and mutant structures are superimposed. (a) P71G; (b) P103G; (c) A47P; (d) D91P; and (e) V110P. CA denotes α-carbon.

Table IV: List of the Dihedral Angles (Φ, Ψ) for the Wild-Type and Mutant h-Lysozymes in the Vicinity of the Substitution Sites

protein	site 71	site 103	site 90	site 91	site 46	site 47	site 109	site 110
wild type	(−71, +137)	(−52, −39)	(−51, −52)	(−69, −41)	(−100, +97)	(−56, −31)	(−99, +115)	(−63, −29)
P71G	(+99, +106)							
P103G		(−18, +38)						
D91P			(−53, −48)	(−64, −42)				
A47P					(−112, +95)	(−55, −34)		
V110P							(−107, +108)	(−49, −39)

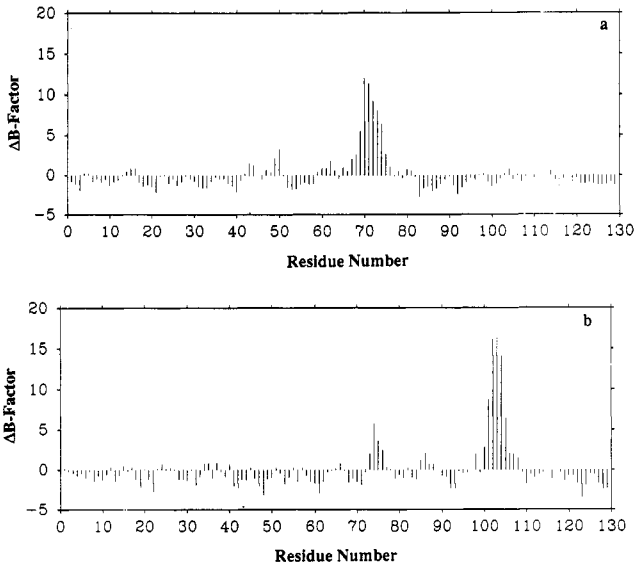

FIGURE 6: Crystallographic *B*-factors of the main-chain residues for (a) P71G and (b) P103G relative to that of the wild-type h-lysozyme.

Pro103 and 75.95 Å^2 for Pro71. The standard ASA value for a proline residue in a random coil conformation was estimated to be 155.86 Å^2 (Go & Miyazawa, 1980). Using this last value as a reference for the 100% random coil character for a proline, Pro103 has a 93.7% random coil character and is accessible to the solvent, whereas Pro71 has only a 48.7% random coil character. According to this estimation, the backbone conformation at site 103 is similar in both the unfolded and folded forms and close to a random conformation. Therefore, the replacement of Pro by a Gly at such a site is likely to affect both forms in the same way. On the other hand, the backbone conformation around position 71 appears to be more rigid. Pro71 is more buried inside the protein and has a lower random coil character than Pro103. Due to the rather rigid local conformation around site 71, even with a Gly instead of a Pro at position 71, the rigid character of the backbone is kept by the neighboring framework, and the configurational entropy of the folded forms is less affected by the mutation than it is for the unfolded form.

This is furthermore supported by the crystallographic data relative to the thermal factor (*B* value) of the atoms at the mutation sites. The thermal factor, which is related to the mobility of the atoms, is higher at the mutated sites Gly71 and Gly103 than it is for the wild type at Pro71 and Pro103 locations. It follows that both of these mutations affect the local flexibility of the folded state. However, the Pro71 → Gly substitution induces a significantly smaller *B*-factor increase than does the Pro103 → Gly substitution (Figure 6). This is, at least qualitatively, consistent with our interpretation and with conclusions drawn in a previous study about the contribution of various conserved prolines to the stability of the tryptophan synthase α subunit (Yutani et al., 1991), where the entropic stabilizing property of various prolines was suggested to be correlated to the local backbone mobility.

The thermostability of the mutant free of proline, P71G/P103G, is identical to that of P71G mutant. The effects of these two substitutions on h-lysozyme stability are additive. Pro103 → Gly replacement does not alter stability, whereas Pro71 → Gly substitution is responsible for the observed stability decrease. This additivity was also observed for the folding kinetic properties relative to those three mutants (Herning et al., 1991), and such an additivity rule is likely to be satisfied as well for the local structural perturbations introduced by each mutation.

Xaa → Pro Mutants: A47P, V110P, and D91P. Each mutant of this series contains three proline residues. The mutations were selected in order to minimize the conformational perturbations caused by the introduction of a trans proline in the protein sequence. For the A47P mutant, Ala was replaced by Pro at position 47, which is part of a β-turn. As reported under Results, the X-ray analysis shows that the backbone structure is only slightly affected by this mutation. The overall stability of this mutant is similar to that of the wild-type protein as measured over a wide temperature range by scanning calorimetry. However, a decrease in the entropy of unfolding is indeed induced by that mutation (0.8 kcal/mol), but its effect on the net free energy is partly compensated by an enthalpic decrease of almost the same magnitude (0.7 kcal/mol), resulting in only a small increase in the overall stability.

The mutant V110P is clearly more stable than the wild type. The free energy of denaturation of this mutant is 0.5 kcal/mol higher than that of the wild type. The structure of the V110P mutant was found to be similar to that of the wild type. Val110 is located at the N-terminal region of an α-helix which remains unchanged when the valine is replaced by a proline. The calorimetric data indicate that the increase of stability due to this mutation must be ascribed to both enthalpic and entropic effects, +0.2 and −0.3 kcal/mol, respectively, at the reference temperature of 68.8 °C and pH 2.80. Both Ala47 → Pro and Val110 → Pro substitutions are located inside secondary structure elements which are locked in the early stage of folding (Herning et al., 1991). The main-chain backbone is therefore kept rather rigid at those mutation sites in the folded state due to the secondary structure framework, and this whatever the intrinsic flexibility of the amino acid is.

Although proline is considered to be more rigid than aspartic acid, the overall stability of the D91P mutant is lower than that of the wild type by 0.4 kcal/mol. This apparent discrepancy with the entropic stabilization hypothesis can be explained from the calorimetric results combined with those obtained from the X-ray analysis. The plot of ΔH_d^{cal} versus T_d clearly indicates that there is a decrease in the enthalpy of denaturation for D91P as compared to the wild-type protein. The decrease in the denaturation enthalpy of this mutant is 4.7 kcal/mol, which is considerable. This effect can be explained from the X-ray analysis, which shows that an H-bond has been lost on the replacement of Asp91 by Pro. In wild-type h-lysozyme, the carboxylic group of Asp91 is involved

in a H-bond with the amine group of the glutamine at position 86 (not shown). The distance between H-bond acceptor and donor atoms is 2.8 Å. This H-bond is absent in the D91P mutant, resulting in the enthalpy decrease measured by calorimetry. However, the entropic effect led by the Asp91 → Pro substitution indeed corresponds to a quite important stabilization of about 4.3 kcal/mol, in fact, higher than for the two previous mutants. However, this entropic stabilization effect is offset by the larger enthalpic destabilization brought about by the mutation, resulting in a lower overall stability as compared to the wild-type protein.

To conclude, the results presented in this work are consistent with the hypothesis that stability of mutants involving proline largely involves entropic factors. However, this behavior can be greatly affected by the extent of solvent accessibility and the mobility of the substitution site. We have also shown that in some cases the expected behavior concerning entropic effects on mutation can be masked by important compensating enthalpic changes. Finally, we wish to emphasize the fact that our discussion concerning mutant stability was limited to the contributions involving the polypeptide chain. However, one has to keep in mind that the local solvent structure also changes upon mutation and that it is likely to contribute to both entropic and enthalpic effects as well.

ACKNOWLEDGMENT

We thank Dr. M. Ikehara for helpful advice and encouragement and Dr. Y. Taniyama for the useful discussions we had. We also thank Takeda Chemical Ind., Ltd., for providing plasmid pGEL125.

REFERENCES

Alber, T. (1989) *Annu. Rev. Biochem. 58*, 765–798.

Alber, T., Dao-Pin, S., Nye, J. A., Muchmore, D. C., & Matthews, B. W. (1987) *Biochemistry 26*, 3754–3758.

Alber, T., Bell, J. A., Dao-Pin, S., Nicholson, H., Wozniak, J. A., Cook, S., & Matthews, B. W. (1988) *Science 239*, 631–635.

Artymiuk, P. J., & Blake, C. C. F. (1981) *J. Mol. Biol. 152*, 737–762.

Fujii, I., Morimoto, Y., Higuchi, Y., & Yasuoka, N. (1991) *Acta Crystallogr., Sect. B. 47*, 137–144.

Go, M., & Miyazawa, S. (1980) *Int. J. Pept. Protein Res. 15*, 211–224.

Hendrickson, W. A. (1985) *Methods Enzymol. 115*, 252–270.

Herning, T., Yutani, K., Taniyama, Y., & Kikuchi, M. (1991) *Biochemistry 30*, 9882–9891.

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

Inaka, K., Taniyama, Y., Kikuchi, M., Morikawa, K., & Matsushima, M. (1991) *J. Biol. Chem. 266*, 12599–12603.

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

Kidokoro, S., & Wada, A. (1987) *Biopolymers 26*, 213–229.

Kikuchi, M., Yamamoto, Y., Taniyama, Y., Ishimaru, K., Yoshikawa, W., Kaisho, Y., & Ikehara, M. (1988) *Proc. Natl. Acad. Sci. U.S.A. 85*, 9411–9415.

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

Minohara, A., Toh-e, A., Nozaki, C., Hamada, F., Ohtomo, N., & Matsubara, K. (1983) *Proc. Natl. Acad. Sci. U.S.A. 80*, 1–5.

Nemethy, G., Leach, S. J., & Scheraga, H. A. (1966) *J. Phys. Chem. 70*, 998–1004.

Nicholson, H., Becktel, W. J., & Matthews, B. W. (1988) *Nature 366*, 651–656.

Privalov, P. L. (1979) *Adv. Protein Chem. 33*, 167–240.

Privalov, P. L., & Khechinashvili, N. N. (1974) *J. Mol. Biol. 86*, 665–684.

Ramachandran, G. N., Ramakrishnan, C., & Sasisekharan, V. (1963) *J. Mol. Biol. 7*, 95–99.

Schimmel, P. R., & Flory, P. J. (1968) *J. Mol. Biol. 34*, 105–120.

Shortle, D., Meeker, A. K., & Freire, E. (1988) *Biochemistry 27*, 4761–4768.

Steigemann, W. (1974) Ph.D. Thesis, Technische Universitat, Munchen, Germany.

Tanaka, I., Yao, M., Suzuki, M., Hikichi, K., Matsumoto, T., Kozasa, M., & Katayama, C. (1990) *J. Appl. Crystallogr. 23*, 334–339.

Taniyama, Y., Yamamoto, Y., Nakao, M., Kikuchi, M., & Ikehara, M. (1988) *Biochem. Biophys. Res. Commun. 152*, 962–967.

Tsong, T. Y., Hearn, R. P., Wrathall, D. P., & Sturtevant, J. M. (1970) *Biochemistry 9*, 2666–2675.

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