Kruijff, B. (1985) *Biochim. Biophys. Acta 818*, 398–409.
Rizo, J., Blanco, F. J., Kobe, B., & Gierasch, L. M., (1991) *J. Cell. Biochem. 15G*, 89.
Roder, H., & Elöve, G. A. (1991) *J. Cell. Biochem. 15G*, 158.
Roder, H., Elöve, G. A., & Englander, S. W. (1988) *Nature 335*, 700–704.
Smith, M., Leung, D. W., Gillian, S., Astell, C. R., Montgomery, D. L., & Hall, B. D. (1979) *Cell 16*, 753–761.
Snel, M. M. E., Kaptein, R., & de Kruijff, B. (1991) *Biochemistry 30*, 3387–3394.
Spooner, P. J. R., & Watts, A. (1991a) *Biochemistry 30*, 3871–3879.
Spooner, P. J. R., & Watts, A. (1991b) *Biochemistry 30*, 3880–3884.
Stuart, R. A., Nicholson, D. W., & Neupert, W. (1990) *Cell 60*, 31–43.
Walter, A., Margolis, D., Mohan, R., & Blumenthal, R. (1987) *Membr. Biochem. 6*, 217–237.
Wand, A. J., Di Stefano, D. L., Feng, Y., Roder, H., & Englander, S. W. (1989) *Biochemistry 28*, 186–194.
Wojtczak, L., & Sottocasa, G. L. (1972) *J. Membr. Biol. 7*, 313–324.
Wüthrich, K. (1986) in *NMR of Proteins and Nucleic Acids*, pp 13–25, Wiley, New York.

Differential Scanning Calorimetric Study of the Thermal Unfolding of Mutant Forms of Phage T4 Lysozyme†

Cui-Qing Hu,‡ Shinichi Kitamura,§ Akiyoshi Tanaka,‖ and Julian M. Sturtevant*

Department of Chemistry and Department of Molecular Biophysics and Biochemistry, Yale University, New Haven, Connecticut 06511

Received August 28, 1991; Revised Manuscript Received November 8, 1991

ABSTRACT: In two recent papers, we reported the effects of several point mutations on the thermodynamics of the thermal unfolding of the lysozyme of phage T4 as determined by differential scanning calorimetry. The mutants studied were R96H [Kitamura, S., & Sturtevant, J. M. (1989) *Biochemistry 28*, 3788–3792] and T157 replaced by A, E, I, L, N, R, and V [Connelly, P., Ghosaini, L., Hu, C.-Q., Kitamura, S., Tanaka, A., & Sturtevant, J. M. (1991) *Biochemistry 30*, 1887–1891]. Here we report the results of a similar study of the single mutations A82P, A93P, and G113A and the double mutation C54T:C97A. The three single mutants all show small apparent stabilization at pH 2.5 and 46.2 °C (the denaturational temperature of the wild-type protein), amounting to -0.5 ± 0.4 kcal mol^{-1} in free energy, whereas the double mutant shows a weak apparent destabilization, $+0.8 \pm 0.4$ kcal mol^{-1}. As in all our previous studies of mutant proteins, the enthalpy changes produced by these mutations are in general of much larger magnitude than the corresponding free energy changes and frequently of opposite sign.

It has often been remarked that globular proteins in solution, despite their highly organized and nearly perfectly cooperative structures, nevertheless are only marginally stable, their stabilization free energies of 5–20 kcal mol^{-1} being the results of cancellations of many positive and negative contributions. Consequently, as pointed out by Hawkes et al. (1984), thermodynamic parameters for the reversible thermal denaturation of proteins afford very little basis for understanding the wide range of forces involved in protein structures. Nevertheless, such data have proven to be useful in assessing the effects resulting from known amino acid replacements and in attempting to correlate these effects with observed changes in structure as determined by X-ray crystallography or 3-D NMR, and in supplying numbers with which the results of theoretical calculations can be compared.

A recent paper by Tidor and Karplus (1991) illustrates very clearly the formidable difficulties involved in reaching a detailed understanding of the energetic results of a single amino acid replacement in a protein. These authors used free energy simulation methods to analyze the effects of the R96H mutation in T4 lysozyme, one of the mutations on which we have run DSC experiments (Kitamura & Sturtevant, 1989). Limiting their treatment to interactions involving 7 residues, 3 on each side of residue 96, they tabulated some 35 free energy contributions involving the change of the folded WT to the folded mutant and 21 involving the unfolded forms, some positive and some negative with magnitudes up to 7 kcal mol^{-1}. These contributions added up to a total destabilization of 1.9 kcal mol^{-1} at 300 K compared with the extrapolated DSC result of 3.2 ± 1.2 kcal mol^{-1}.

The lysozyme of bacteriophage T4 is a single-chain polypeptide containing 164 residues and having a molecular weight of 18 700. Its crystal structure has been accurately determined (Weaver & Matthews, 1987), and numerous known amino acid replacements have been produced by appropriate mutagenesis and cloning. Although the apparent stabilities of many of these mutants have been estimated by means of optical melting curves, mainly employing circular dichroism and fluorescence emission, to date only eight have been studied by DSC. We believe it is important to subject as many as possible

† Partial support of this research was supplied by Grant GM-04725 from the National Institutes of Health and by Grant DMB-8810329 from the National Science Foundation.

* Address correspondence to this author at the Department of Chemistry, Yale University.
‡ Permanent address: Institute of Chemistry, Academia Sinica, Beijing, China.
§ Permanent address: Department of Agricultural Chemistry, Kyoto Prefectural University, Kyoto 606, Japan.
‖ Permanent address: Department of Chemistry, Mie University, Mie 514, Japan.

0006-2960/92/0431-1643$03.00/0

Table I: Variation with pH of the Temperature of Half-Denaturation

protein	pH range	no. of pH values	no. of expts	A (°C)[a]	B (°C pH^{-1})[a]	SD (°C)
WT	1.60–2.84	9	43	9.13	14.81	±0.6
A82P	1.77–2.98	8	23	2.26	18.13	±1.1
A93P	1.77–2.98	8	22	3.90	17.47	±0.9
G113A	1.77–2.98	7	20	11.54	14.44	±1.4
C54T:C97A	1.76–4.54	9	23	12.59	12.49	±2.1

[a] $t_{1/2} = A + BpH$.

of these mutants to calorimetric investigation because it appears that DSC affords the method of choice for the study of denaturational thermodynamics, despite the relatively high protein concentrations required by the method with the attendant increased likelihood of complications arising from aggregation. Three important advantages of DSC are (1) both the van't Hoff and calorimetric enthalpies can be evaluated, (2) the heat capacity change can be directly determined, and (3) due allowance for the frequently large variation of enthalpy with temperature can be readily included in the procedure for data analysis.

The three single mutations studied here, A82P, A93P, and G113A, were constructed by Matthews et al. (1987) to test their prediction that certain residues replacing other residues, at positions where no strain or other unfavorable effect would be introduced by the replacement, would decrease the configurational entropy of the unfolded form and thus lead to apparent stabilization of the native form of the protein. The double mutation was produced (Matsumura & Matthews, 1989) to give a "pseudo-wild-type" protein into which various engineered disulfide bridges were introduced.

Materials and Methods

Materials. The proteins used in this research were prepared at the University of Oregon and shipped at ice temperature to New Haven. They were supplied at high concentration, were dialyzed before use against a 20 mM potassium phosphate buffer containing 25 mM KCl and 0.5 mM dithiothreitol, and were used in our experiments at concentrations of 1–13 mg mL^{-1}. Protein concentrations were determined spectrophotometrically, employing an absorption coefficient of 1.28 cm^2 mg^{-1} at 280 nm. Since none of the mutations involved aromatic residues, it was assumed that they caused no significant changes in absorbance.

Calorimetry. The DSC experiments were performed in either the MC-2 instrument (Microcal, Inc., Northampton, MA) or the DASM-4 instrument (Biopribor, Pushchino, Moscow region, USSR; Privalov, 1980). The two calorimeters gave concordant results. A scan rate of 1 K min^{-1} was used throughout. Instrumental base lines were determined with both cells filled with dialysate, with which the reference cell was filled during protein scans. The denaturations were found to be fully reversible provided the first heating was not carried above about 95% completion.

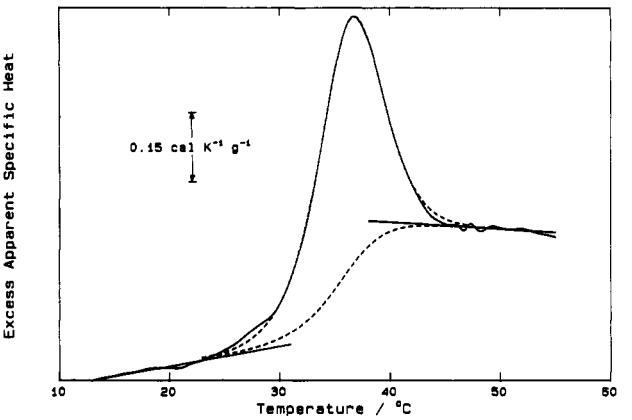

FIGURE 1: DSC curve observed with C54T:C97A at a protein concentration of 0.98 mg mL^{-1} at pH 1.97. Solid lines, pre- and post-transition base lines; solid curve, observed data; dashed curves, calculated base line and theoretical curve calculated with $t_{1/2}$ = 35.58 °C, ΔH_{cal} = 77.6 kcal mol^{-1}, $\Delta H_{vH}/\Delta H_{cal}$ = 1.18, and Δc_p = 0.268 cal K^{-1} g^{-1}. The value of ΔC_p, amounting to about 5 kcal K^{-1} mol^{-1}, is 2–3 times the value usually observed for the unfolding of globular proteins. The calculated data deviated from the observed data with a standard deviation of 0.0068 cal K^{-1} g^{-1} (0.9% of the maximal excess specific heat).

Data Analysis. The data for all the scans, after subtraction of the instrumental base line, were fitted by a least-squares procedure to a modified two-state model in which the van't Hoff enthalpy and the calorimetric enthalpy are allowed to be varied independently (Sturtevant, 1987; Kitamura & Sturtevant, 1989; Connelly et al., 1991). The adequacy of fit, measured by the standard deviation of calculated from observed points, and expressed as percent of the maximal value of the excess heat capacity, is indicated in the last column of Table I.

Results

A typical DSC scan is shown in Figure 1. This was run with the mutant C54T:C97A at pH 1.97 and a protein concentration of 0.98 mg mL^{-1}. The dashed curves are the calculated base line and the theoretical curve which best fits the observed (solid) curve. The parameters for the theoretical curve are $t_{1/2}$ = 35.48 °C, ΔH_{cal} = 77.6 kcal mol^{-1}, $\Delta H_{vH}/\Delta H_{cal}$ = 1.18, and Δc_p = 0.268 cal K^{-1} g^{-1}. The standard deviation of the calculated from the observed points is 0.9% of the maximal value of the excess specific heat.

The data obtained are summarized in Tables I and II. The data for the wild-type protein are taken from previous publications (Kitamura & Sturtevant, 1989; Connelly et al., 1991). Table I lists the number of different pH values used for each protein and the number of experiments performed. The fifth and sixth columns show the variation with pH of $t_{1/2}$, the temperature (in degrees centigrade) at which the denaturation is half-completed, in the form of the constants A and B obtained by linear least squaring, with the standard deviation shown in the last column.

Table II: Variation with Temperature of the Enthalpy of Denaturation

protein	ΔH_0 (kcal mol^{-1})	ΔC_p (kcal K^{-1} mol^{-1})[a]	SD (kcal mol^{-1})	ΔC_p(mean) (kcal K^{-1} mol^{-1})[b]	$\Delta H_{vH}/\Delta H_{cal}$	mean SD (curve fitting, %)[c]
WT	−10.51	2.57	±5.7	2.41 ± 0.08		
A82P	−0.27	2.46	±5.9	2.39 ± 0.09	0.93 ± 0.02	1.2 ± 0.2
A93P	20.10	2.04	±5.9	2.42 ± 0.14	0.96 ± 0.02	0.9 ± 0.1
G113A	−19.48	2.88	±5.1	2.52 ± 0.11	1.00 ± 0.03	0.8 ± 0.1
C54T:C97A	16.38	1.94	±4.0	3.20 ± 0.16	0.98 ± 0.02	0.8 ± 0.2

[a] $\Delta H_{cal} = \Delta H_0 + \Delta C_p t_{1/2}$. [b] Mean of the values of ΔC_p evaluated at $t_{1/2}$ in each scan. [c] Percent of the maximal excess heat capacity.

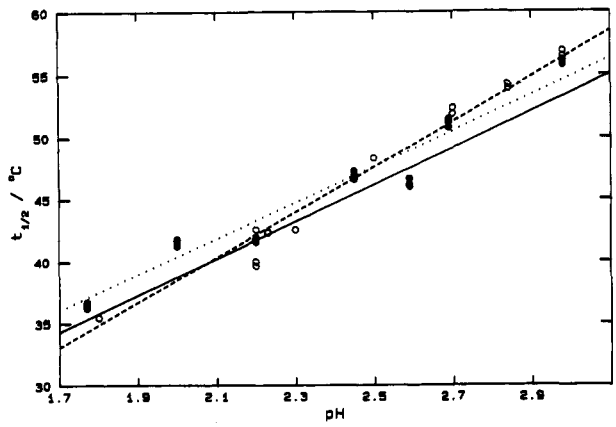

FIGURE 2: Variation with pH of the temperature of half-denaturation, $t_{1/2}$, of the mutants A82P (●) and G113A (○). Included in the figure are the least-squared lines for G113A (---), for A82P (···), and for WT protein (—).

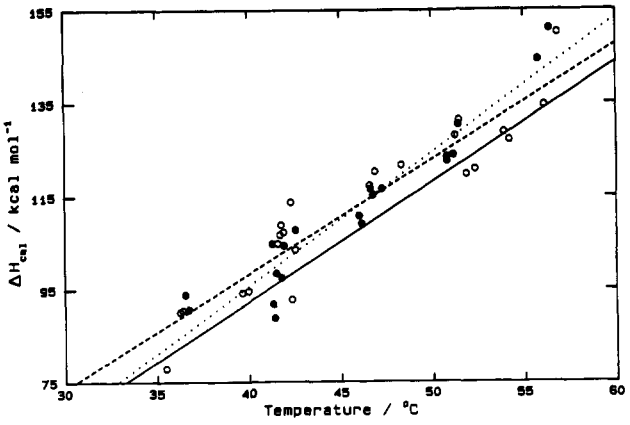

FIGURE 3: Variation with temperature of the enthalpies of denaturation of the mutants A82P (●; 23 experiments) and G113A (○; 20 experiments). Included in the figure are the least-squared lines for G113A (---), for A82P (···), and for WT protein (—; 43 experiments).

The dependence of $t_{1/2}$ on pH is illustrated for A82P and G113A in Figure 2, which also includes the least-squared line for WT. Figure 3 shows the strong dependence of ΔH_{cal} on temperature for these same three proteins. Although the plots in Figure 3 are actually of values of ΔH_{cal} at $t_{1/2}$ vs $t_{1/2}$, it seems reasonable to assume that they describe the general dependence of denaturational enthalpy on temperature. The least-squared results for all four mutants are summarized in Table II. The fifth column lists the means (±SE) of the values for ΔC_p obtained in the curve fitting for each scan. The mean values (±SE) for the ratio of van't Hoff to calorimetric enthalpies are shown in column six and the mean (±SE) standard deviations in the curve fitting in column seven.

Discussion

As we have observed in previous DSC work with T4 lysozyme and its mutants, and in numerous scans of staphylococcal nuclease and its mutants (A. Tanaka et al., unpublished data), the apparent specific heats before the unfolding transitions increase with temperature with a coefficient averaging approximately 0.002 cal K^{-2} g^{-1}, while these quantities after the transition are nearly independent of temperature. This indication of a variation of ΔC_p with temperature is incompatible with the temperature independence of ΔC_p shown in the variation of ΔH_{cal} with temperature (Figure 3). A possible rationalization of this anomaly (P. L. Privalov, personal communication) involves the contributions to the apparent heat capacity of a protein from low-frequency combination modes which, in contrast to ordinary vibrational modes in small molecules, can be excited to states above the ground state at ordinary temperatures. It seems likely that there will be a decrease in such combination modes when a protein unfolds with concomitant decreases in intramolecular interactions and that this may account for the reduced temperature dependence of the heat capacity of the unfolded form.[1]

The variation of $t_{1/2}$ with pH is similar to that observed with other T4 mutants. The standard deviation found in this variation corresponds to a variation of ±0.08 pH unit, suggesting that uncertainties in pH determinations account for a significant fraction of the uncertainties in the values for $t_{1/2}$. The increase in $t_{1/2}$ with pH shows that protons are taken up from the buffer during denaturation. The constants listed in Tables I and II, used in conjunction with the equation

$$\Delta\nu = \frac{1000\Delta H_{cal}}{2.303RT_{1/2}^2}\frac{dt_{1/2}}{dpH} \quad (1)$$

where $T_{1/2} = 273.15 + t_{1/2}$, give for the mean value of $\Delta\nu$

$$\Delta\nu = 1.54 + 0.86pH \quad (2)$$

over the pH range covered in the present work. Since the isoelectric pH of T4 lysozyme is approximately 10, the protein carries a large net positive charge at low pH, which could lead to some of the carboxyl groups having abnormally low values of pK (Anderson et al., 1990) in the native state such that they become increasingly protonated in the denaturated state. This could account for the positive value of the slope in eq 2.

In view of the substantial uptake of protons during these denaturations, the question arises concerning corrections to the DSC enthalpies for the heat of ionization of the buffer. We have not usually applied such corrections because fully valid corrections could not be applied in the absence of values for the heats of ionization of the appropriate groups on the protein. In the case of C54T:C97A where two different buffers were used, to bring the data obtained in acetate buffer to the same basis as those in phosphate buffer requires deducting 3.6 kcal mol^{-1} at pH 3.52 where ΔH_{cal} averaged 127.2 kcal mol^{-1}, and 4.7 kcal mol^{-1} at pH 4.54 where ΔH_{cal} averaged 139.9 kcal mol^{-1}. These corrections, though hardly significant, have been applied.

It was found in earlier work on T4 mutants that the ratio $\Delta H_{vH}/\Delta H_{cal}$ tended to be >1 at low pH and to become <1 at higher pH. In contrast, as shown in Table II, this ratio is unity within experimental uncertainty for the four mutants studied here.

A convenient measure of the change produced by a mutation in the apparent stability of a protein is the quantity

$$\Delta\Delta G_d^\circ = \Delta G_d^\circ(\text{WT}) - \Delta G_d^\circ(\text{mutant}) \quad (3)$$

the difference in the standard free energies of unfolding for the wild-type and mutant proteins, evaluated at $t_{1/2}$ for the wild-type protein by means of the Gibbs–Helmholtz equation:

$$\Delta G^\circ(\text{at } T_1) = \Delta H_2\frac{T_2 - T_1}{T_2} - \Delta C_p\left(T_2 - T_1 + T_1 \ln\frac{T_1}{T_2}\right) \quad (4)$$

where T_1 and T_2 are $t_{1/2}$ + 273.15 for wild type and mutant,

[1] It should be noted here that it was proposed earlier (Sturtevant, 1977) that part of the usually observed increase in C_p accompanying protein unfolding is due to an *increase* in the contribution of low-frequency modes. This view does not seem to be tenable since the low-frequency modes result primarily from intramolecular interactions many of which are lost during the unfolding.

Table III: Changes in Free Energy, Enthalpy, and Entropy of Denaturation, Evaluated at $t_{1/2}$ for the Wild-Type Protein, Produced by the Various Mutations[a]

	pH 2.0, 38.8 °C			pH 2.5, 46.2 °C			pH 3.0, 53.6 °C		
protein	$\Delta\Delta G_d^\circ$	$\Delta\Delta H_d$	$\Delta\Delta S_d^\circ$	$\Delta\Delta G_d^\circ$	$\Delta\Delta H_d$	$\Delta\Delta S_d^\circ$	$\Delta\Delta G_d^\circ$	$\Delta\Delta H_d$	$\Delta\Delta S_d^\circ$
A82P	+0.07	−6.0	−19	−0.51	−5.2	−15	−1.27	−4.4	−10
A93P	−0.03	−10.1	−32	−0.51	−6.2	−18	−1.10	−2.2	−3
G113A	−0.50	−3.0	−8	−0.54	−5.3	−15	−0.54	−7.6	−22
C54T:C97A	+0.34	−2.5	−9	+0.77	+2.2	+4	+1.27	+6.9	+17

[a] $\Delta\Delta G_d^\circ$ and $\Delta\Delta H_d$, kcal mol^{-1}; $\Delta\Delta S_d^\circ$, cal K^{-1} mol^{-1}. Estimated uncertainties: $\Delta\Delta G_d^\circ$, ±0.4 kcal mol^{-1}; $\Delta\Delta H_d$, ±4 kcal mol^{-1} (average value); $\Delta\Delta S_d^\circ$, ±10 cal K^{-1} mol^{-1}.

respectively, ΔH_2 is the denaturational enthalpy in kilocalories per mole of the mutant at T_2, and ΔC_p is the denaturational heat capacity change in kilocalories per degree kelvin per mole of the mutant. It is obvious that in the present application of eq 3 the term $\Delta G_d^\circ(\mathrm{WT}) = 0$, so that destabilization of the native structure, or stabilization of the denatured structure, is indicated by a positive value for $\Delta\Delta G_d^\circ$. Table III lists the values for $\Delta\Delta G_d^\circ$ at pH 2.0, 2.5, and 3.0 calculated from the constants given in Tables I and II and also lists the values for

$$\Delta\Delta H_d = \Delta H_d(\mathrm{WT}) - \Delta H_d(\mathrm{mutant}) \qquad (5)$$

at $t_{1/2}$ for the wild type. We estimate the uncertainty in the values for $\Delta\Delta G_d^\circ$ to be ±0.4 kcal mol^{-1} and in the values for $\Delta\Delta H_d$ to average about ±4 kcal mol^{-1}.

It is evident that the mutations considered here have only small effects on the thermodynamic stability of T4 lysozyme. A result which might not have been predicted is the apparent stabilization caused by the substitution A82P at least at pH 2.5 or higher. Residues 81–91 form an α-helix in T4 lysozyme (Remington et al., 1978), and it would be expected that replacement of Ala, a good helix-former, by Pro, the worst helix-former, would lead to destabilization. On the other hand, as pointed out by Matthews et al. (1987), the introduction of proline should decrease the configurational entropy of the unfolded form, thereby producing an apparent stabilization of the native form. It is interesting that in this same helix, replacement of Pro-86 by Ala led to a small apparent destabilization (Alber et al., 1988). It is noted by Alber et al., that residues 81–91 do not form an ideal helix in wild-type T4 lysozyme or in any of its mutant forms for which structures have been determined, so that the usual scales of helix-forming potential may be irrelevant.

It was to test the prediction noted above, that in the absence of other effects the introduction of proline should reduce the entropy of unfolding [see also Suzuki (1989)], that the mutant A82P was first prepared by Matthews et al. (1987). This effect should also contribute to the small apparent stabilization found for A93P. However, comparison of the observed values of $\Delta\Delta S_d^\circ$ for A82P and A93P (Table III) with the value for the entropic contribution due to the Ala to Pro replacement, estimated by Matthews et al. (1987) to amount to −4 kcal K^{-1} mol^{-1} at pH 2, indicates that additional contributions undoubtedly complicate the situation.

Gly-113 is identified by Remington et al. (1978) as being at the end of an imperfect α_{II}-helix. Perhaps this helix is somewhat stabilized by the substitution G113A. The double mutation C54T:C97A causes a small apparent destabilization of the molecule.

According to the usually accepted scales of helical propensities (Chou & Fasman, 1978; Scheraga, 1978), Gly is approximately equivalent to Pro. Replacements of Gly by other residues in an α-helix of the N-terminal domain of λ-repressor have been found to follow the simple prediction of significant stabilizations, amounting in the case of the double mutant G46A:G48A to an increase in the denaturational

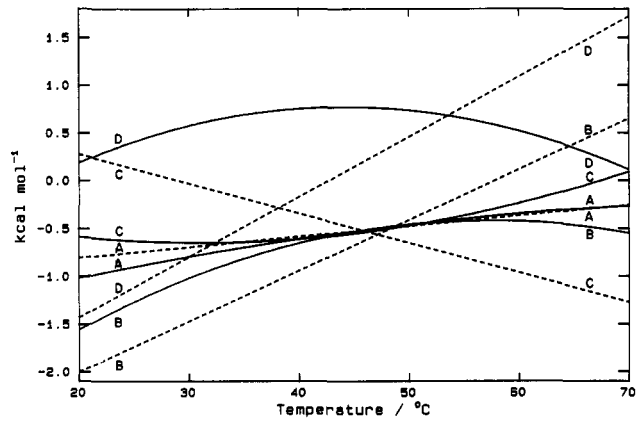

FIGURE 4: Variation with temperature at pH 2.5 of $\Delta\Delta G_d^\circ$ (—) and $\Delta\Delta H_d/10$ (---) for (A) A82P, (B) A93P, (C) G113A, and (D) C54T:C97A calculated from the constants given in Tables I and II and the Gibbs–Helmholtz equation.

temperature of 6.2 °C (Hecht et al., 1986).

The procedure outlined above for calculating $\Delta\Delta G_d^\circ$ and $\Delta\Delta H_d$ is the most reliable one available since it involves the shortest possible extrapolations. It is nevertheless of interest to consider the variation of these quantities over a wider temperature range in view of the large values of ΔC_p characteristic of protein denaturations. The values of $\Delta\Delta G_d^\circ$ (—) and $\Delta\Delta H_d/10$ (---) calculated for the four mutants at pH 2.5 are shown in Figure 4. When it is noted that the enthalpies have been divided by 10, the figure emphasizes the lack of correlation between free energies and enthalpies and illustrates that significant variations of $\Delta\Delta G_d^\circ$ with temperature can occur.

ACKNOWLEDGMENTS

We are greatly indebted to Joan Wozniak in the laboratory of Brian Matthews at the University of Oregon for the generous supplies of the proteins consumed in this research. The helpful comments of Professor Matthews are appreciated. We are also indebted to Professor John Brandts of the University of Massachusetts for the long-term loan of the MC-2 microcalorimeter used in most of our DSC experiments.

REFERENCES

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

Chou, P. Y., & Fasman, G. D. (1978) *Annu. Rev. Biochem. 47*, 251–276.

Connelly, P., Ghosaini, L., Hu, C.-Q., Kitamura, S., Tanaka, A., & Sturtevant, J. M. (1991) *Biochemistry 30*, 1887–1891.

Hawkes, R., Grutter, M. G., & Schellman, J. (1984) *J. Mol. Biol. 175*, 195–212.

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

Kitamura, S., & Sturtevant, J. M. (1989) *Biochemistry 28*, 3788-3792.
Matsumura, M., & Matthews, B. W. (1989) *Science 243*, 792-794.
Matthews, B. W., Nicholson, H., & Becktel, W. J. (1987) *Proc. Natl. Acad. Sci. U.S.A. 84*, 6663-6667.
Privalov, P. L. (1980) *Pure Appl. Chem. 52*, 479-497.
Remington, S. J., Anderson, W. F., Owen, J., Ten Eyck, L. F., Grainger, C. T., & Matthews, B. W. (1978) *J. Mol. Biol. 118*, 81-98.
Scheraga, H. A. (1978) *Pure Appl. Chem. 50*, 315.
Sturtevant, J. M. (1977) *Proc. Natl. Acad. Sci. U.S.A. 74*, 2236-2240.
Suzuki, Y. (1989) *Proc. Jpn. Acad., Ser. B 65*, 146-148.
Tidor, B., & Karplus, M. (1991) *Biochemistry 30*, 3217-3228.
Weaver, L. H., & Matthews, B. W. (1987) *J. Mol. Biol. 193*, 189-199.

The Chemical Shift Index: A Fast and Simple Method for the Assignment of Protein Secondary Structure through NMR Spectroscopy†

D. S. Wishart,*,‡,§ B. D. Sykes,§ and F. M. Richards‡

Department of Molecular Biophysics and Biochemistry, Yale University, New Haven, Connecticut 06511, and MRC Group in Protein Structure and Function, Department of Biochemistry, University of Alberta, Edmonton, Alberta T6G 2H7, Canada

Received August 20, 1991; Revised Manuscript Received November 4, 1991

ABSTRACT: Previous studies by Wishart et al. [Wishart, D. S., Sykes, B. D., & Richards, F. M. (1991) *J. Mol. Biol.* (in press)] have demonstrated that ^{1}H NMR chemical shifts are strongly dependent on the character and nature of protein secondary structure. In particular, it has been found that the ^{1}H NMR chemical shift of the α-CH proton of all 20 naturally occurring amino acids experiences an upfield shift (with respect to the random coil value) when in a helical configuration and a comparable downfield shift when in a β-strand extended configuration. On the basis of these observations, a technique is described for rapidly and quantitatively determining the identity, extent, and location of secondary structural elements in proteins based on the simple inspection of the α-CH ^{1}H resonance assignments. A number of examples are provided to demonstrate both the simplicity and the accuracy of the technique. This new method is found to be almost as accurate as the more traditional NOE-based methods of determining secondary structure and could prove to be particularly useful in light of the recent development of sequential assignment techniques which are now almost NOE-independent [Ikura, M., Kay, L. E., & Bax, A. (1990) *Biochemistry 29*, 4659-4667]. We suggest that this new procedure should not necessarily be seen as a substitute to existing rigorous methods for secondary structure determination but, rather, should be viewed as a complement to these approaches.

For more than 20 years NMR spectroscopists have been attempting to apply chemical shift information to conformational problems of biological significance. Early efforts in this regard were first begun by Sternlicht and Wilson (1967) and Markley et al. (1967). Both groups were interested in studying the chemical shift tendencies of α-CH ^{1}H NMR[1] resonances in amino acid homopolymers, particularly with respect to the systematic changes in proton chemical shifts that were associated with helix formation and helix disruption in these compounds. Subsequent studies by Clayden and Williams (1982) and Dalgarno et al. (1983), based on accumulated data from naturally occurring proteins, suggested that reasonably strong conformationally dependent chemical shift tendencies existed in β-strands as well as in α-helices and that these trends were not confined to certain homopolymers or to selected solvent conditions. More recent work by Szilagyi and Jardetzky (1989) and Wishart et al. (1991) have confirmed these early observations by placing them on a more solid statistical basis. In fact, these workers have clearly demonstrated that a strong relationship exists between α-CH ^{1}H NMR chemical shifts and protein secondary structure for all 20 amino acids.

Some of these observations have already begun to be put to use. Pastore and Saudek (1990) have recently described a useful method for displaying chemical shift information which permits the qualitative identification of secondary structure in proteins. This procedure is based on plotting "smoothed" chemical shift differences (with respect to random coil values) versus protein sequence and using the resulting curve to approximate the location and identity of secondary structures. However, because of the qualitative nature of these plots it is often difficult to identify the exact limits as well as the true identity of all significant secondary structures in the protein of interest.

We describe a new method for secondary structure determination based on chemical shift propensity that was developed quite independently of Pastore and Saudek's work. This particular technique can be used for the precise identification of protein secondary structure from chemical shift information alone. It is fast, simple, and accurate and can be used either

† Financial support by the MRC Group in Protein Structure and Function and by the National Institutes of Health (NIH Grant GM-22778) is gratefully acknowledged.
* To whom correspondence should be addressed.
‡ Yale University.
§ University of Alberta.

[1] Abbreviations: BPTI, bovine pancreatic trypsin inhibitor; CSI, chemical shift index; DSS, 2,2-dimethyl-2-silapentane-5-sulfonic acid; NMR, nuclear magnetic resonance; NOE, nuclear Overhauser effect; TMS, tetramethylsilane; TSP, sodium 3-(trimethylsilyl)propionate.

0006-2960/92/0431-1647$03.00/0