Articles

Analysis of the Interaction between Charged Side Chains and the α-Helix Dipole Using Designed Thermostable Mutants of Phage T4 Lysozyme†,‡

H. Nicholson, D. E. Anderson, S. Dao-pin,§ and B. W. Matthews*

Institute of Molecular Biology, Howard Hughes Medical Institute, and Department of Physics, University of Oregon, Eugene, Oregon 97403

Received April 30, 1991; Revised Manuscript Received August 14, 1991

ABSTRACT: It was shown previously that the introduction of a negatively charged amino acid at the N-terminus of an α-helix could increase the thermostability of phage T4 lysozyme via an electrostatic interaction with the "helix dipole" [Nicholson, H., Becktel, W. J., & Matthews, B. W. (1988) *Nature 336*, 651–656]. The prior report focused on the two stabilizing substitutions Ser 38 → Asp (S38D) and Asn 144 → Asp (N144D). Two additional examples of stabilizing mutants, T109D and N116D, are presented here. Both show the pH-dependent increase in thermal stability expected for the interaction of an aspartic acid with an α-helix dipole. Control mutants were also constructed to further characterize the nature of the interaction with the α-helix dipole. High-resolution crystal structure analysis was used to determine the nature of the interaction of the substituted amino acids with the end of the α-helix in both the primary and the control mutants. Control mutant S38N has stability essentially the same as that of wild-type lysozyme but hydrogen bonding similar to that of the stabilizing mutant S38D. This confirms that it is the electrostatic interaction between Asp 38 and the helix dipole, rather than a change in hydrogen-bonding geometry, that gives enhanced stability. Structural and thermodynamic analysis of mutant T109N provide a similar control for the stabilizing replacement T109D. In the case of mutant N116D, there was concern that the enhanced stability might be due to a favorable salt-bridge interaction between the introduced aspartate and Arg 119, rather than an interaction with the α-helix dipole. The additivity of the stabilities of N116D and R119M seen in the double mutant N116D/R119M indicates that favorable interactions are largely independent of residue 119. As a further control, Asp 92, a presumed helix-stabilizing residue in wild-type lysozyme, was replaced with Asn. This *decreased* the stability of the protein in the manner expected for the loss of a favorable helix dipole interaction. In total, five mutations have been identified that increase the thermostability of T4 lysozyme and appear to do so by favorable interactions with α-helix dipoles. As measured by the pH dependence of stability, the strength of the electrostatic interaction between the charged groups studied here and the helix dipole ranges from 0.6 to 1.3 kcal/mol in 150 mM KCl. In the case of mutants S38D and N144H, NMR titration was used to measure the pK_a's of Asp 38 and His 144 in the folded structures. These provide independent estimates, under different experimental conditions, of 1.4 and 0.9 kcal/mol in 100 mM KCl for the respective interaction energies in these two instances. NMR also indicates that the favorable electrostatic interaction energy for Asp 92 in wild-type lysozyme is ~2.1 kcal/mol. As evidenced by the mutants T109D, N116D, N144H, and N144D, electrostatic interactions between the second residue within an α-helix and the N-terminus appear to be especially important for stability. These results provide a rationale for the previous empirical observation that aspartic acid is the most commonly observed amino acid at this position while histidine is rarely observed [Richardson, J. S., & Richardson, D. C. (1988) *Science 240*, 1648–1652]. The strength of the interaction between a charged amino acid and an α-helix dipole does not depend strongly on the length of the α-helix. This is consistent with a model in which the charge-stabilization effects are attributed to the partial charges on the unsatisfied amides and carbonyls in the first and last turn of the helix.

Although the dipolar nature of α-helices has long been appreciated (Blagdon & Goodman, 1975; Wada, 1976; Hol et al., 1978), it has taken some time to develop a full understanding of the importance of the so-called α-helix dipole in protein folding and function (Ihara et al., 1982; Hol, 1985; Perutz et al., 1985; Shoemaker et al., 1985, 1987; Mitchison & Baldwin, 1986).

We previously demonstrated that the stability of T4 lysozyme could be increased by the introduction of aspartates that interact with the N-termini of α-helices (Nicholson et al., 1988). Electrostatic stabilization of approximately 0.7 kcal/mol was attributable to each introduced charge–dipole interaction. The pH dependence of the stability of the mutants corresponded to that expected for favorable electrostatic interaction. The crystal structures of the mutant proteins also showed that the interaction was between the introduced charge and the helix dipole and did not require hydrogen bonding with the amide groups at the end of the helix. Other studies based on mutant or modified peptides and proteins have also indicated the significance of helix–dipole interactions (Shoemaker

† This work was supported in part by grants from the NIH (GM21967) and from the Lucille P. Markey Charitable Trust.

‡ The atomic coordinates in this paper have been deposited with Brookhaven Protein Data Bank.

§ Present address: National Institute of Diabetes, Digestive and Kidney Diseases, National Institutes of Health, Bethesda, MD 20892.

0006-2960/91/0430-9816$02.50/0

Table I: Amino-Termini of α-Helices in T4 Lysozyme[a]

helix	residues	N_{cap}	N1	N2
A	3–10	Asn 2	Ile 3	Phe 4
B	39–50	Ser 38 ↑ (Asp)	Leu 39	Asn 40 ↑ (Asp)
C	60–79	Thr 59	Lys 60	Asp 61
D	82–90	Asn 81	Ala 82	Lys 83
E	93–106	Asp 92 ↓ (Asn)	Ala 93	Val 94
F	108–113	Gly 107	Glu 108	Thr 109 ↑ (Asp)
G	115–123	Phe 114	Thr 115 ↑ (Glu)	Asn 116 ↑ (Asp)
H	126–134	Arg 125	Trp 126	Asp 127
I	137–141	Ser 136	Arg 137	Trp 138
J	143–155	Thr 142	Pro 143	Asn 144 ↑ (Asp)

[a] The table gives the N_{cap} residue and the first two amino acids in each α-helix of T4 lysozyme (Remington et al., 1978; Weaver & Matthews, 1987). The "N_{cap}", N1, and N2 residues are as defined by Richardson and Richardson (1988). Amino acids in parentheses indicate mutants (from Table II) that involve changes in charge. An upward arrow indicates that the mutation is stabilizing at pH ~6; a downward arrow indicates that the mutation is destabilizing at pH ~6. The mutation Asn 40 → Asp was identified by P. Pjura and M. Matsumura (personal communication).

et al., 1985, 1987; Sali et al., 1988; Weaver et al., 1989; Serrano & Fersht, 1989).

In order to further investigate the structural and thermodynamic characteristics of interactions between charged amino acid side chains and the ends of α-helices, several new mutations were introduced into T4 lysozyme. Table I summarizes the locations of the α-helices in T4 lysozyme and gives the identities of the residues at the amino terminus of each helix. The stabilizing replacements described previously were Ser 38 → Asp and Asn 144 → Asp (Nicholson et al., 1988). In this paper we describe variants that fall into three classes. First, there are mutations designed to introduce stabilizing replacements at two new sites (i.e., at positions 109 and 116). Second, there are control mutations constructed to confirm the nature of the stabilization observed both at the original sites (38 and 144) as well as the new sites (109 and 116). Third, a final control was carried out by replacing Asp 92, a residue that appears to participate in a helix dipole interaction in wild-type lysozyme. Figure 1 summarizes the lysozyme variants that have been constructed.

Choice of Mutations. In the initial attempt to increase the stability of T4 lysozyme by introducing charged residues designed to interact with α-helix dipoles, we focused on the substitution of aspartic acid residues close to the amino termini of helices. The two main requirements for selection of sites were (1) that a negatively charged group should not already be located close to the helix terminus and (2) that model building of the proposed substitution indicated that it would not introduce unfavorable steric interference with the rest of the protein (Nicholson et al., 1988). On this basis residues 38, 116, and 144 were selected as potential sites for amino acid replacement. The mutations S38D and N144D were constructed and shown to be stabilizing. Position 116 was, however, avoided because it was thought that it might interfere with crystal contacts. Despite this potential difficulty, we still wanted to determine the effectiveness of the design procedure and so have now constructed the mutant N116D. Because of the concern that the introduced aspartate at position 116 might interact with Arg 119 in the next turn of the α-helix, the controls R119M and the double mutant N116D/R119M were also constructed.

With optimal selection of sites, an aspartate would not be introduced at a helix N-terminus if other negatively charged side chains were nearby. Slight relaxation of this criteria, however, suggested that Thr 109 might be a candidate for replacement with aspartic acid. The neighboring residue, Glu

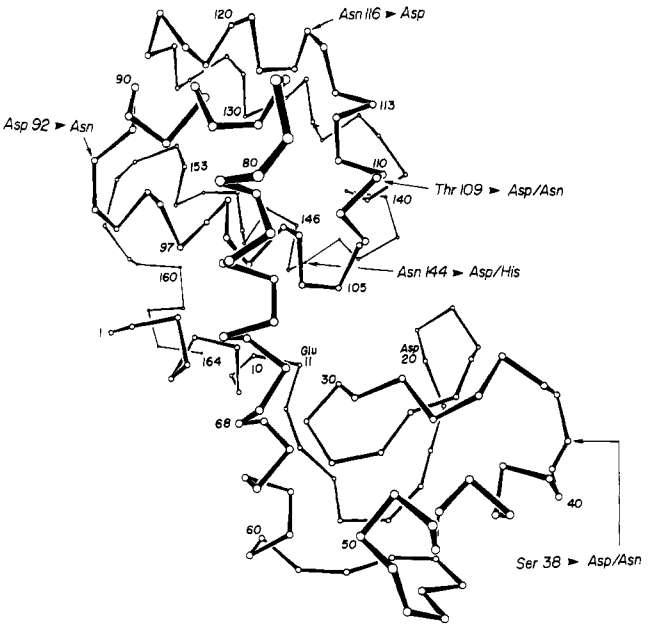

FIGURE 1: Backbone structure of T4 lysozyme showing the locations of the mutations described in the text.

108, is acidic, but the carboxylate extends perpendicular to the helix axis and is involved in a hydrogen bond with the side chain of residue Asn 81. It therefore appeared that Glu 108 would not prevent interaction between Asp 109 and the N-terminus of the 108-113 helix. The distance from the carboxylate of Glu 108 to the side chain of Thr 109 in wild-type lysozyme is about 5.5 Å. Residues 108 and 109 are also solvent-exposed, suggesting that repulsion between the aspartate and the glutamate side chains at these positions would be attenuated. With all these factors in mind, it was decided to construct the mutant Thr 109 → Asp. As a control, Thr 109 was also replaced with asparagine.

As noted above, the replacements S38D and N144D have previously been shown to be stabilizing (Nicholson et al., 1988). To further characterize these replacements, Ser 38 was replaced by asparagine and Asn 144 replaced by histidine. The pK_a's of both Asp 38 and His 144 have been determined by NMR.

To further quantitate the charge–dipole interaction, asparagine was introduced in place of Asp 92. Asp 92 is located at the end of α-helix 93–106 and appears to have a favorable interaction with the α-helix dipole. The substitution places an asparagine at the N-terminus of the α-helix, which would be expected to destabilize the protein at neutral pH where the mutant side chain lacks the negative charge that would be present on the aspartate.

Materials and Methods

Mutant Lysozymes. The mutant proteins were produced by oligonucleotide-directed mutagenesis (Zoller & Smith, 1984; Kunkel et al., 1987). All mutations were produced in a T4 lysozyme derivative of M13mp18. N116D, R119M, and N116D/R119M were produced on a wild-type template, while S38N, D92N, T109N, and T109D were produced in the "cysteine-free wild-type" (i.e., C54T/C97A or WT*) background kindly provided by Dr. Masazumi Matsumura (Matsumura & Matthews, 1989; Matsumura et al., 1989). After purification, the entire lysozyme gene was sequenced in each case to ensure that no alterations other than those expected had been introduced. Restriction digests using *Bam*H1 and *Hin*dIII were employed to move the mutant T4

Table II: Data Collection and Refinement Statistics[a]

protein	WT*	S38N*	D92N*	T109D*	T109N*	N116D
data collection						
cell dimensions						
a,b (Å)	60.9	61.0	61.4	61.0	61.0	61.3
c (Å)	96.8	97.1	96.1	96.9	96.9	96.6
total reflections measured	14345	14358	12110	16191	14487	13266
R_{merge} (%)	4.7	6.5	8.2	5.7	5.3	7.1
R_{mutant} (%)		13.9	24.0	11.5	12.4	15.8
refinement						
unique reflections	13879	13808	11169	15679	14015	12852
resolution (Å)	1.75	1.80	1.90	1.70	1.75	1.85
R (%)	14.8	15.7	15.3	15.2	15.7	16.3
$\Delta_{bond\ length}$ (Å)	0.016	0.017	0.018	0.016	0.016	0.018
$\Delta_{bond\ angle}$ (deg)	2.08	2.21	2.33	2.14	2.08	2.20

[a] R_{merge} is the agreement between intensities measured on different films. R_{mutant} is the average difference between the observed structure amplitudes for the mutant and wild-type data. *R* is the crystallographic residual after refinement. The coordinate shift is the rms difference between the mutant and WT coordinates for all backbone atoms in residues 1–161 after a least-squares superposition of the molecules. $\Delta_{bond\ length}$ and $\Delta_{bond\ angle}$ are the rms deviations of the bond lengths and angles in the refined models relative to their "ideal" values. Cysteine-free proteins are denoted by an asterisk and can be compared to the pseudo-wild-type (WT*). N116D data can be compared with WT values from Bell et al. (1991). Data for WT* were measured by Dr. Elisabeth Eriksson.

lysozyme gene from M13 into the plasmid expression vector pHN1403 (Poteete et al., 1990) based on an expression system created by D. C. Muchmore (Muchmore et al., 1989). The protein was purified as previously described (Poteete et al., 1990). All proteins were shown to be at least 95% pure by analytical reverse-phase HPLC and SDS–polyacrylamide gel electrophoresis.

Thermal Stability. Thermal denaturation of each protein was monitored by following the change in dichroism at 223 nm, as has been described (Becktel & Baase, 1987). The protein (20 μg/mL) was heated at the rate of 1 °C/min until the protein was completely unfolded. Then the temperature was decreased to ensure that the transition was reversible. At pH 2.0, thermal denaturation was carried out in 200 mM KCl and 0.01 N HCl, and the pH was adjusted with HCl. Above pH 5.5, a solution of 150 mM KCl and 10 mM phosphate was used.

NMR Titrations. Histidine titration of N144H was performed as previously described (Weaver et al., 1989; Anderson et al., 1990). The single histidine in wild-type T4 lysozyme has an abnormally high pK_a (Anderson et al., 1990) and is easily distinguishable from the additional histidine introduced in N144H.

The carboxyl titrations of S38D and of wild-type lysozyme were performed by using ^{13}C NMR on samples of T4 lysozyme that were specifically enriched with ^{13}C at the side-chain γ-carbon of all aspartate and asparagine residues [cf. Muchmore et al. (1989) and Anderson et al. (1990)]. Buffer conditions for the titrations were 30 mM phosphate and 100 mM KCl. The temperature was 10 °C. Titration parameters were obtained by using a nonlinear least-squares fitting routine.

Comparison of the NMR spectra of S38D and WT* at different pH values revealed the presence of one new peak with no significant changes in the chemical shifts of any of the peaks that had previously been assigned to the aspartates in the wild-type protein (D. E. Anderson and F. W. Dahlquist, personal communication). By this criteria the new peak was assigned to Asp 38.

The assigned amide peaks in the ^{15}N-labeled spectra of T4 lysozyme (McIntosh et al., 1990) were used for preliminary inferential titration of Asp 92, and the result was verified by comparison of the ^{13}C spectra of the WT and D92N (D. E. Anderson and F. W. Dahlquist, unpublished results).

Crystallography. Crystallization of all of the mutants was attempted under conditions similar to those used to grow wild-type T4 lysozyme crystals (Weaver & Matthews, 1987). All mutant proteins except N144H gave crystals isomorphous with the native enzyme. This includes the mutant N116D even though this residue is involved in a crystal contact. The crystals of D92N were difficult to obtain and macroseeding had to be used to provide crystals sufficiently large for data collection. Before X-ray exposure, the crystals were equilibrated with the standard mother liquor containing 1.05 M K_2HPO_4, 1.26 M NaH_2PO_4, 0.23 M NaCl, and 1.4 mM 2-mercaptoethanol at pH 6.7. High-resolution X-ray data were collected by oscillation photography (Rossmann, 1979; Schmid et al., 1981).

Least-squares refinement was carried out with the TNT package of programs (Tronrud et al., 1987). The most current wild-type structure (Bell et al., 1991) was used as a starting model both for the refinement of N116D as well as for refinement of the cysteine-free pseudo-WT structure C54T/C97A. (Data for the WT* structure were kindly provided by Dr. Elisabeth Eriksson.) The refined model for pseudo-wild-type lysozyme supercedes that reported by Pjura et al. (1990) and was used as the initial model for all mutant structures, except N116D. Whenever a conformation was difficult to determine, the atoms with weak density were deleted from the model and the rest of the structure refined to improve phasing until the "best" position for the side chain became evident in a difference map. All difference electron density features greater than 4σ were monitored through the course of refinement as potential solvent sites. New solvent atoms were required to have at least one reasonable hydrogen bond and no unacceptable van der Waals contacts with the protein molecule. All solvent atoms were also required to have thermal factors of approximately 70 Å^2 or less. The number of solvent atoms in the refined models of the WT and WT* lysozyme structures is about 150. Almost all of these are retained in the respective mutant structures. Thermal factors for the protein were refined on an individual atom-by-atom basis. Additional details of the typical procedure used to refine T4 lysozyme mutants are given by Dao-pin et al. (1991a). Data processing and refinement statistics are listed in Table II. Coordinates of the mutant lysozymes have been deposited in the Brookhaven Protein Data Bank.

RESULTS

Thermal Stability. Mutant S38N, which was made as a control for the stabilizing mutant S38D (Nicholson et al., 1988), has the same stability as wild-type lysozyme at both pH 2.0 and 6.7 (Table III).

Mutant D92N was designed to replace Asp 92, which is a

Table III: Thermodynamic Stability of Mutant Lysozymes[a]

mutant	ref protein	pH 2.0		pH 5.7–6.9		$\Delta\Delta G_E$ (kcal/mol)	mutational effect
		ΔT_m (deg)	$\Delta\Delta G$ (kcal/mol)	ΔT_m (deg)	$\Delta\Delta G$ (kcal/mol)		
S38D[b]	WT	−0.3 ± 0.5	−0.1 ± 0.15	1.5 ± 0.5	0.6 ± 0.2[c]	0.7	substitute Asp at N_{cap}
S38N	WT*	−0.3 ± 0.4	−0.1 ± 0.1	−0.1 ± 0.6	0.0 ± 0.2[d]	0.1	Asn control
D92N	WT*	−0.3 ± 0.4	−0.1 ± 0.1	−3.7 ± 0.5	−1.4 ± 0.2[d]	−1.3	replace Asp at N_{cap}
T109D	WT*	−0.8 ± 0.4	−0.3 ± 0.1	1.5 ± 0.6	0.6 ± 0.2[d]	0.9	substitute Asp at N2
T109N	WT*	−0.1 ± 0.4	0.0 ± 0.1	0.3 ± 0.5	0.1 ± 0.2[d]	0.1	Asn control
T115E[e]	WT	−1.7 ± 0.5	−0.5 ± 0.15	0.7 ± 0.5	0.3 ± 0.2[f]	0.8	substitute Glu at N1
N116D	WT	−0.2 ± 0.3	−0.1 ± 0.1	1.6 ± 0.4	0.6 ± 0.15[g]	0.7	substitute Asp at N2
R119M	WT	−0.9 ± 0.3	−0.3 ± 0.1	0.3 ± 0.5	0.1 ± 0.20[g]	0.4	control
N116D/R119M	WT	−0.8 ± 0.3	−0.3 ± 0.1	1.6 ± 0.4	0.6 ± 0.15[g]	0.9	control
N144D[b]	WT	−0.2 ± 0.5	−0.1 ± 0.15	1.4 ± 0.5	0.5 ± 0.2[c]	0.6	substitute Asp at N2
N144H	WT*	−2.3 ± 0.4	−0.6 ± 0.1	0.7 ± 0.6	0.3 ± 0.2[d]	0.9	substitute His at N2

[a] T_m is the melting temperature of the mutant protein, and ΔT_m is the change in melting temperature of the mutant protein relative to its wild-type parent, either normal wild-type lysozyme (WT) or cysteine-free pseudo-wild-type lysozyme (WT*). $\Delta\Delta G$ is the change in the free energy of unfolding of the mutant relative to the corresponding wild-type, estimated from ΔT_m by the relationship of Becktel and Schellman (1987). A positive value of $\Delta\Delta G$ indicates that the mutant protein is more stable than wild-type. Measurements for the reference protein (WT or WT*) were carried out in parallel with each mutant under identical experimental conditions. This provides the most reliable estimate of ΔT_m and $\Delta\Delta G$. The melting temperature of WT was 40.9 ● 0.1 °C at pH 2.0 and 66.1 ± 0.2 °C at pH 5.7. For WT* the melting temperature was 39.5 ± 0.2 °C at pH 2.0 and 62.4 ± 0.2 °C at pH 6.7. $\Delta\Delta G_E$ is the inferred electrostatic free energy of stabilization, defined as the difference between $\Delta\Delta G$ at pH 5.7–6.9 and $\Delta\Delta G$ at pH 2.0. The positive values of $\Delta\Delta G_E$ for the Asp, Glu, and His residues indicate that electrostatic interactions are favorable for Asp and Glu, which are negatively charged at neutral pH, but unfavorable for His because the histidine is positively charged at low pH. [b] From Nicholson et al. (1988). [c] pH 6.9. [d] pH 6.7. [e] From Dao-pin et al. (1991b). [f] pH 6.5. [g] pH 5.7.

helix-capping residue at the amino-terminus of α-helix 93–106. The mutant shows a decrease in stability relative to wild-type at pH 6.7, suggesting that the carboxylate of Asp 92 does indeed exert a stabilizing interaction when it is charged. At pH 2.0, where the aspartate loses most of its charge, its replacement with asparagine causes almost no change of stability (Table III).

Mutant T109D, designed to interact with the amino-terminus of α-helix 108–113, is found to increase stability at pH 6.5, as expected. At pH 2.0, Asp 109 is expected to be largely uncharged, causing a loss of favorable electrostatic interaction. At this pH, a slight decrease in stability is in fact observed. In addition, the control mutant protein T109N has stability essentially the same as that of wild-type at both pH 2.0 and 6.7, providing further evidence that it is the charge on Asp 109 that is essential for the increase in stability (Table III).

As anticipated, the replacement of Asn 144 with histidine destabilizes the protein at pH 2.0, where the imidazole is fully protonated, but not near neutral pH, where the histidine in the folded protein is mostly uncharged.

The replacement Asn 116 → Asp was designed to create a favorable interaction with α-helix 115–123. The thermodynamic data (Table III) show that the stability of this variant relative to wild-type does increase with pH, as anticipated. Because it was possible that the apparent favorable electrostatic interaction of Asp 116 might be with Arg 119 rather than the α-helix dipole, Arg 119 was replaced with methionine. The pH-dependent changes in the stability of the single mutants N116D and R119M are additive in the double mutant N116D/R119M (Table III). This shows that the interaction of Asp 116 with the Arg 119 is weak.

NMR Titrations. At 10 °C the pK_a's of Asp 38 and Asp 92 are 3.0 ± 0.1 and 2.5 ± 0.2, respectively. These pK_a shifts correspond to electrostatic interaction energies of 1.4 and 2.1 kcal/mol, assuming a nonperturbed pK_a value of 4.0 for aspartic acid. The pK_a of His 144 is 6.2 ± 0.1, which is 0.6 pH unit (0.9 kcal/mol) below the expected pK_a of 6.8 for a histidine in unfolded T4 lysozyme under these conditions (Anderson et al., 1990).

Structures of Mutant Lysozymes. (A) Position 38. The crystallographic data collection and refinement statistics for S38N and the other mutant crystal structures are summarized in Table II.

Serine 38 is the capping residue for the helix that includes residues 39–50. No charged residues are within 9 Å of residue 38, and it is also free of crystal contacts. We previously showed that an aspartic acid replacement at residue 38 interacts favorably with the peptide dipoles of the N-terminal peptide groups (Nicholson et al., 1988). In the present study, asparagine was introduced in place of the wild-type serine. Asparagine serves as a good control because its hydrogen-bonding groups have similar geometry to the stabilizing aspartic acid.

The map showing the difference in electron density between S38N and WT* (Figure 2a) has a positive feature due to the extra bulk of the asparagine side chain. For S38N, as was the case with S38D (Nicholson et al., 1988), the positions of all side-chain atoms except one terminal atom of residue 38 are apparent in a $(2F_{o_{mut}} - F_{o_{wt}})$ map (Figure 2b). Rather than showing the superposition of S38N on wild-type lysozyme, we show in Figure 2c the superposition of S38N and S38D. It is the comparison of these two mutants that is of special interest. [The superposition of S38D on wild-type lysozyme was shown previously. See Figure 3b of Nicholson et al. (1988).] Figure 2c suggests that the structures of S38N and S38D are very similar. There does appear to be a difference in the orientation of the side chain of Asp 38 relative to Asn 38, but this may be of dubious significance because the terminal atoms of both side chains have high crystallographic thermal factors (respective values of 65 and 77 Å^2). To investigate whether the difference in side-chain orientation of Asp 38 and Asn 38 was significant, a map with coefficients $(F_{S38D,obs} - F_{S38N,obs})$ and phases from the refined S38N model was calculated (not shown) and had no significant features (3σ or larger) in the vicinity of residue 38. Lack of difference density indicates that the conformations of Asp 38 and Asn 38 are very similar. The crystallographic data therefore suggest that the δ-oxygens of Asp 38 and Asn 38 occupy roughly the same position, are reasonably well ordered, and form similar if not identical hydrogen bonds to the amides in the first turn of the α-helix. In contrast, the second δ-oxygen of Asp 38 and the δ-nitrogen of Asn 38 are both solvent-exposed and mobile.

(B) Position 92. In wild-type lysozyme, Asp 92 (Figure 3a) is the charged N-terminal capping residue for the α-helix that includes amino acids 93–106. Residue 92 also participates in a 2.9-Å ion pair interaction with Arg 95. In turn, Arg 95

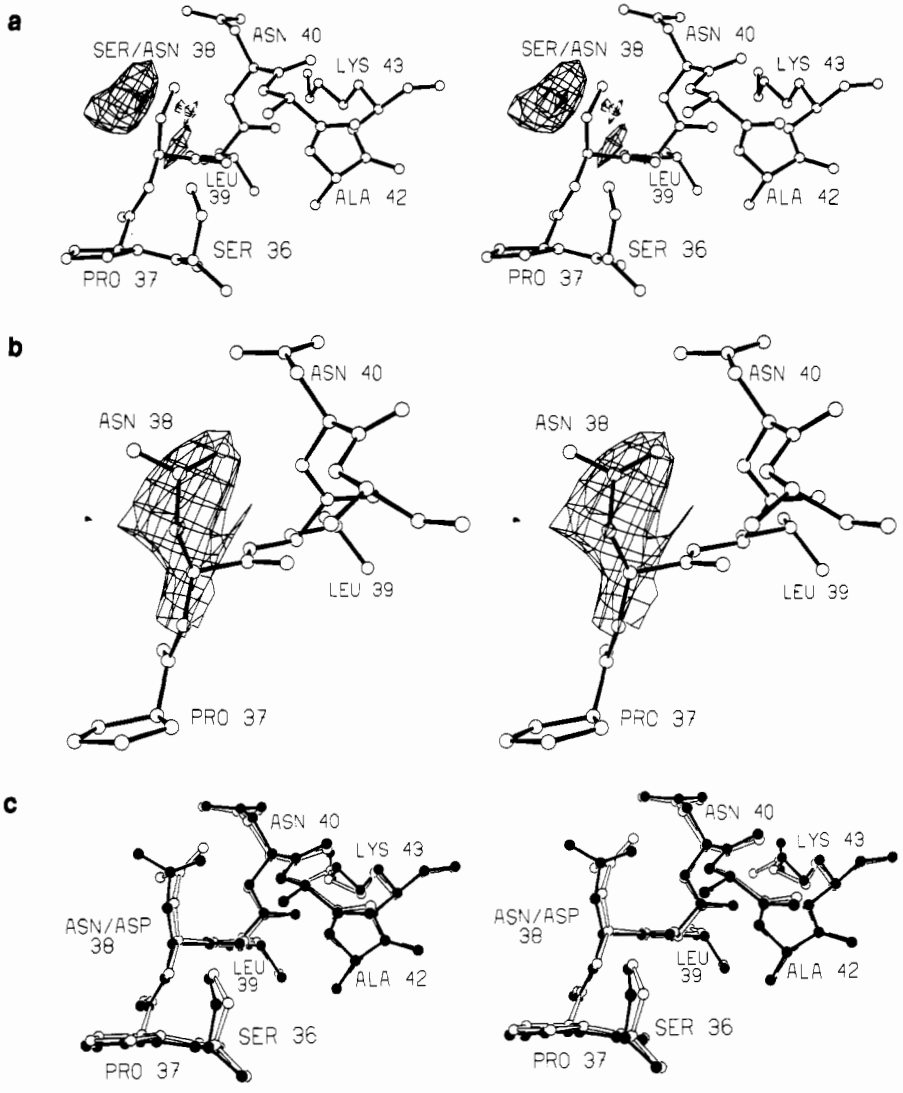

FIGURE 2: (a) Stereofigure showing the structure of WT* lysozyme in the vicinity of residue 38. Superimposed is the difference map for the mutant S38N. Coefficients are ($F_{S38N,obs} - F_{WT*,obs}$) and phases are from the refined WT* structure. The resolution is 1.8 Å, and the contours are at $+3.5\sigma$ (solid) and -3.5σ (broken) where σ is the root-mean-square value of the difference density throughout the unit cell. (b) Map calculated with amplitudes ($2F_{S38N,obs} - F_{WT*,obs}$) and phases from the refined model of WT* superimposed on the refined model of S38N. The resolution is 1.80 Å. The map is contoured at 1σ. (c) Superposition of the refined structure of mutant S38N (open bonds and atoms) on that of mutant S38D (solid bonds and atoms). Because the densities for the side chains of Asp 38 and Asn 38 are weak, the apparent difference in the side chains of these two residues may not be significant. The apparent movement in the side chain of residue 43 is also of dubious significance because the density is weak.

interacts with the C-terminus of another α-helix and stacks against Trp 126. The solvent-mediated crystal contacts between Asp 92 and a symmetry-related molecule are shown in the figure.

The mutation D92N affected crystal growth. With the use of crystal seeding it was possible to obtain crystals large enough to provide data to 1.9-Å resolution. The *c* axis of the unit cell of the mutant crystals was, however, 0.8 Å shorter than the wild-type crystals, complicating interpretation of the crystallographic data. Because of the change in *c*, a special procedure was used to calculate the difference map for this derivative (see the legend to Figure 3b). The difference density map for D92N (Figure 3b) has a pair of positive and negative features indicating rotation and outward movement of the mutant side chain of Asn 92 that leaves the position of one terminal atom relatively invariant. The δ-oxygen of the asparagine appears to remain close to the amino-terminus of the 93–106 α-helix in order to accept hydrogen bonds from the amide nitrogens. The δ-nitrogen of Asn 92, a proton donor, is no longer able to accept a hydrogen bond from Arg 95 and swings toward the solvent. This rotation also allows the dipole of the side chain of Asn 92 to align with the peptide dipoles of the α-helix that it caps. The second positive and negative density features in Figure 3b indicate a solvent molecule that moves to accept the available hydrogen bond from Arg 95. Although these changes can be rationalized in terms of intramolecular hydrogen bonding and electrostatic considerations, we cannot rule out the possibility that intermolecular crystal contacts may play some role in the changes observed. Additional features in the difference map in the vicinity of the crystal contact were also observed (data not shown). Refinement indicates that two new intermolecular contacts are formed. Gln 141 in the symmetry-related molecule forms a new hydrogen bond with $N^{\delta 2}$ of Asn 92, and Thr 21, also in a symmetry-related molecule, assumes a new χ_1 angle that allows it to hydrogen bond with Trp 126. A smaller feature indicates increased mobility in the side chain of Lys 124, which becomes less ordered but retains its conformation. Also, several solvent molecules in the crystal contact region are rearranged in the mutant structure. The overall rms difference between the mutant structure and wild-type is 0.20 Å. This suggests relatively large conformational changes, but it has

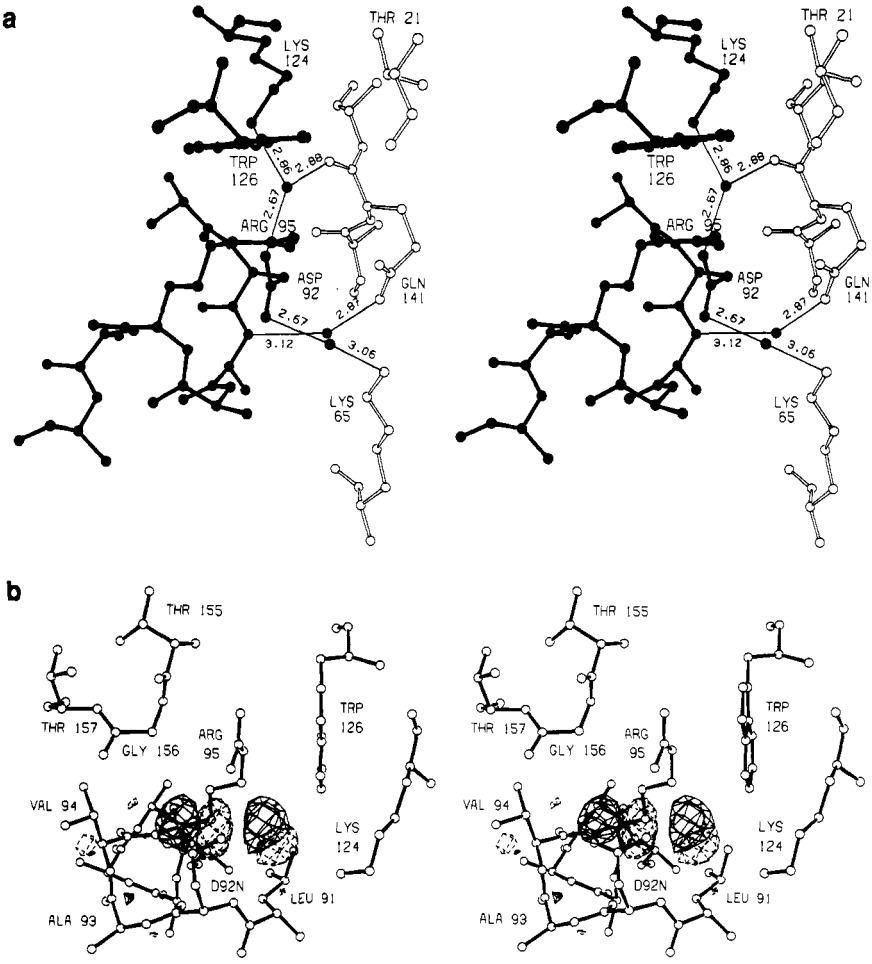

FIGURE 3: (a) The crystal structure of wild-type lysozyme in the vicinity of residue 92. The filled bonds represent the "reference" lysozyme molecule. The structure drawn with open bonds is part of a neighboring lysozyme molecule in the crystal. Solvent-mediated crystal contacts involving Asp 92 and Gln 141 are also shown. The thin lines and distances indicate hydrogen bonds. (b) Map showing the difference in density between mutant D92N and wild-type lysozyme. Because of the change in cell dimensions associated with this mutation, the following procedure was followed. First, the refined structure of WT* lysozyme, including solvent, was placed in the mutant cell. One cycle of rigid-body refinement using TNT (Tronrud et al., 1987) was used to reduce the crystallographic R factor at 3-Å resolution from 29.3% to 19.0%. Structure amplitudes, F_c, and phases from this refined model were then used to calculate a map with coefficients $(F_{D92N,obs} - F_c)$, as shown in the figure. The resolution is 1.9 Å, and the map is contoured at $\pm 3.5\sigma$.

to be remembered that there are structural adjustments not only near residue 92 but also because of the crystal contact.

(C) Position 109. Thr 109 is located at the second helical position of a short helix that includes residues 108–113. The only charged group near this site is Glu 108, which is involved in a hydrogen bond with Asn 81. Atoms within the side chain of Thr 109 in wild-type lysozyme have an average thermal factor of 35 Å^2, indicating that the side chain is reasonably well located. There are two solvent molecules that make good hydrogen bonds to the amide nitrogens in the first turn of the 108–113 helix (Figure 4a). The site of the mutation is free of crystal contacts.

For both T109D and T109N, the X-ray data show that the structure is very similar to wild-type except at the site of the substitution. Maps showing the difference between both T109D and WT* (Figure 4a) and T109N and WT* (Figure 4b) are essentially featureless except near residue 109. The negative feature superimposed on the side chain of Thr 109 in both maps is in part due to the loss of the β-branch but also suggests that the wild-type threonine is better ordered than either mutant side chain. An adjacent dumbbell-shaped positive feature indicates new density for the carboxylate oxygens of Asp 109 (Figure 4a). A smaller feature indicates a similar position for the side chain of Asn 109 (Figure 4b). Although refinement suggests the side chain of Asp 109 is disordered, a map with coefficients $(2F_{mutant} - F_{WT})$ (Figure 4c) clearly shows the conformation of Asp 109 that is most highly occupied. The $2F_{mutant} - F_{WT}$ map for Asn 109 (not shown) indicates partial occupancy of two χ_1 rotamers. The most highly occupied conformer is the same as that for Asp 109, while the conformer that is farthest from the 108–113 helix terminus is also partially occupied. The second rotamer of Asn 109 directs the side chain toward the neighboring ϵ-amino group of Lys 83. Only the conformation with highest occupancy was refined; for this reason, the side-chain dihedral angles are not very reliable. The refined structures of the T109D and T109N mutants are shown in Figure 4d,e). The χ_1 conformation adopted by Asp 109 is such that the side chain is directed predominantly toward the end of the helix (and Glu 108) and away from Lys 83. This suggests that, if a favorable charge–charge interaction does exist between the negatively charged aspartate side chain and its most approachable positively charged neighbor, it is not of significant magnitude to dominate the choice of χ_1 rotamer for residue 109. In further support of an interaction between Asp 109 and the N-terminus of the 108–113 helix, refinement of the T109D structure indicates that Asp 109 has a partial occupancy (approximately 10–20%) of the rotamer that places the carboxyl group directly over the helix N-terminus. In summary, the side chains of Asn 109 and Asp 109 in the respective mutant structures are more mobile than Thr 109 in WT lysozyme. Asp 109 adopts rotamers closer to the helix N-terminus (and to Glu 108) and

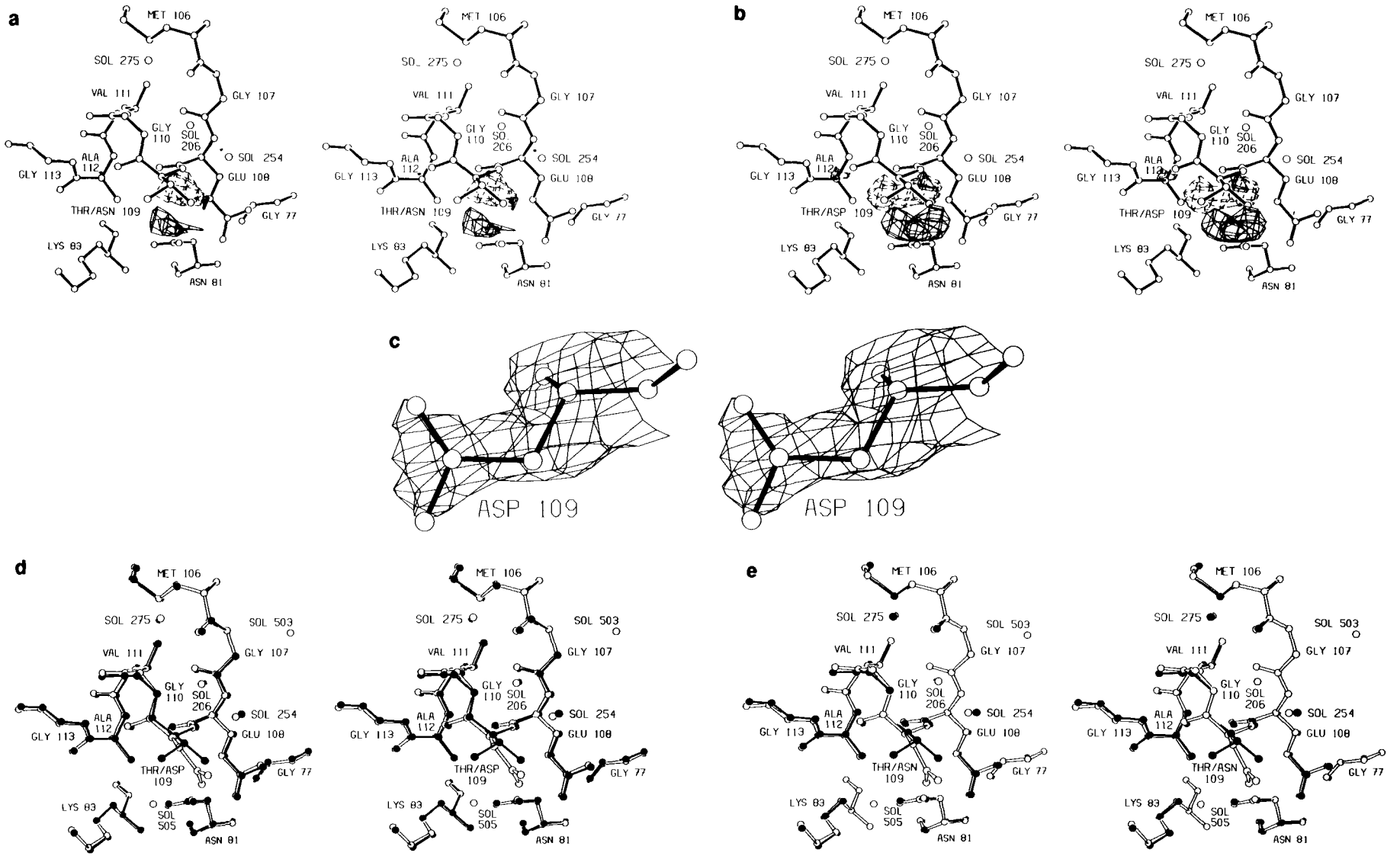

FIGURE 4: (a) Stereodrawing showing the wild-type structure in the vicinity of residue 109. Superimposed is the difference in electron density between the T109D mutant and pseudo-wild-type lysozyme. Amplitudes are $F_{T109D,obs} - F_{WT^*,obs}$, and the phases are calculated from the WT* model. The map is contoured at $\pm 3.5\sigma$, where σ is the root-mean-square density throughout the unit cell. Positive and negative density are represented by solid and broken lines, respectively. The resolution is 1.7 Å. (b) Map showing the difference in density between T109N and pseudo-wild-type lysozyme. All conventions are the same as in panel a except that the resolution is 1.75 Å. (c) Electron density for Asp 109 in the T109D mutant. Coefficients are $2F_{T109D,obs} - F_{WT^*,obs}$, and phases are from the refined model of WT*. The map is superimposed on the refined mutant structure. Resolution is 1.7 Å, and the map is contoured at 1σ. (d) Superposition of the refined structure of mutant T109D (open bonds and open atoms) on pseudo-wild-type lysozyme (solid bonds and solid atoms). (e) Superposition of the refined structure of mutant T109N (open bonds and open atoms) on pseudo-wild-type lysozyme (solid bonds and solid atoms).

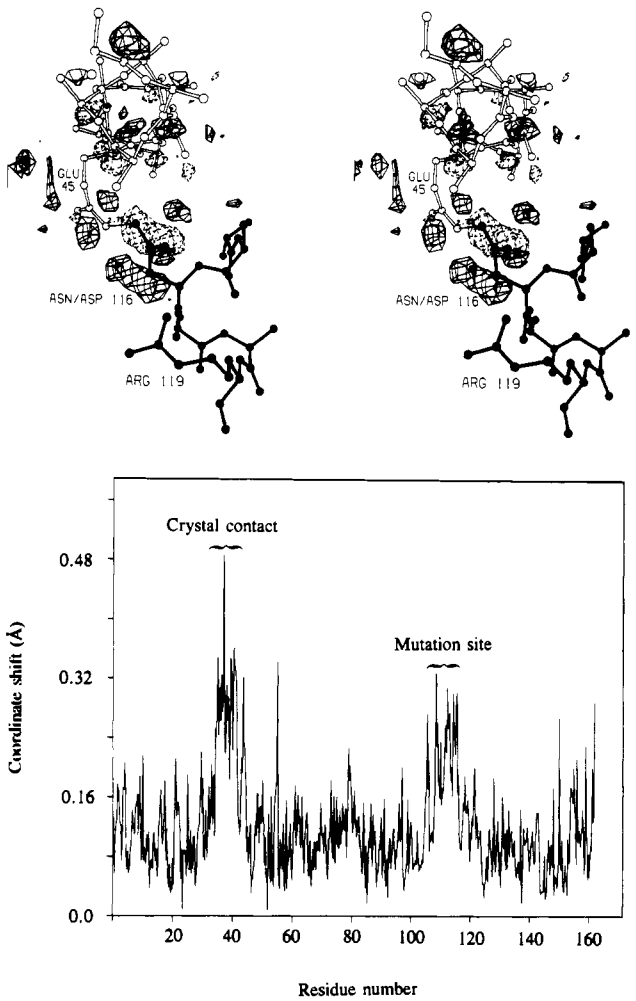

FIGURE 5: (a) The structure of wild-type lysozyme in the vicinity of Asn 116. The Asn 116 → Asp mutation site is near a symmetry contact in the crystals. The atoms and bonds of the "reference" and symmetry-related molecules are drawn solid and open, respectively. The thin line and distance indicate a hydrogen bond between the side chain of Asn 116 and Glu 45. Superimposed is the difference electron density between mutant and wild-type structures. Coefficients are $F_{N116D} - F_{WT}$, where the structure factor amplitudes were observed from mutant and WT crystals and the phase angles are from the refined wild-type model. The 1.85-Å resolution electron density map is contoured at $\pm 3.5\sigma$. (b) "Shift plot" showing the displacement between the backbone atoms of N116D lysozyme and that of wild-type. All backbone N, CA, C, and O atoms for residues 1–161 are included. Prior to making the plot, the respective structures were superimposed so as to minimize the root-mean-square displacement between all atoms. The relative movements in a symmetry-related α-helix (residues 39–50) that results from crystal contacts is the largest observed.

farther from Lys 83 than is the case for Asn 109. The partial disorder of residue 109 in the mutant structures makes analysis of hydrogen bonding difficult, although it does indicate that neither Asp 109 nor Asn 109 forms strong hydrogen bonds. There are no significant movements of the neighboring charged groups of Glu 108, Arg 80, Lys 83, and Lys 35, which are, respectively, 4.7, 8.4, 8.6, and 9.4 Å from the closest carboxylate oxygen of Asp 109.

(*D*) *Position 116.* Asparagine 116 is at the second helical position of the helix that includes residues 115–123. The average difference between the structure amplitudes observed for the mutant and wild-type crystals is 15.8% and the rms coordinate shift is 0.13 Å (Table II). The difference map for the mutant N116D, superimposed on the WT model, is shown in Figure 5a. In the wild-type structure, there is a 2.8 Å hydrogen bond between Asn 116 and Glu 45 in a symmetry-related molecule. The presence of aspartic acid at position 116 creates an unfavorable intermolecular interaction between these groups. This results in a movement in the symmetry-related molecule that is indicated by paired positive and negative features along the helix containing Glu 45. In fact, the coordinate shift plot (Figure 5b) shows that shifts near symmetry-related residue 45 are larger than shifts around the site of mutation. Thus, most of the changes in electron density, and the associated coordinate shifts, are the result of changes in crystal contacts, rather than the Asn 116 → Asp mutation, per se. Possibly as a result of the altered crystal contacts the side chain of Asp 116 moves toward the side chain of Arg 119 with which it can have a favorable electrostatic interaction. The lack of any significant difference density features in the vicinity of Arg 119 (Figure 5a) indicates that this residue remains essentially unmoved. Only a limited structural description of this mutant is presented because the intermolecular contact compromises the relevance of the structure in the crystal to that in solution. This mutation was initially avoided (Nicholson et al., 1988) in part because of this very reason.

(*E*) *Position 144.* The crystal structure of the primary mutant in which Asn 144 was replaced with Asp has been described previously (Nicholson et al., 1988). The mutant constructed here as a control, Asn 144 → His, could not be crystallized and no details of this structure can be provided.

Discussion

We should, perhaps, begin with a comment on the use of the term "α-helix dipole". The role of the α-helix dipole in protein folding, stability, and function has been discussed extensively (Wada, 1976; Hol et al., 1978; Hol, 1985; Perutz et al., 1985; Shoemaker et al., 1987). The dipolar nature of an α-helix can be considered either as a macrodipole or as a clustering of positive charge at the amino-terminus and of negative charge at the carboxy-terminus of the helix. When we use the terms "helix dipole" and "charge–dipole interaction" in the text, we do not mean to discriminate between these two alternatives. We will, however, return to this point at the end of the discussion.

Effectiveness of Designed Charge–Dipole Interactions. In order to facilitate the discussion, we have collected together in Table I the relevant stability data for the variants that are the subject of this paper, as well as lysozyme variants described elsewhere. These include the two stabilizing substitutions S38D and N144D (Nicholson et al., 1988), as well as the replacement T115E, constructed by Dao-pin et al. (1991b) to investigate the contribution of salt bridges to protein stability.

The signature of an electrostatic interaction is its pH dependence. We are primarily concerned here with interactions between aspartate or glutamate side chains and the positive charge at the N-terminus of an α-helix. The pK_a's for the side-chain carboxylates of Asp and Glu in solution are approximately 4. This value may, of course, be perturbed in the folded protein. At pH 2.0 most aspartates and glutamates in the protein will be largely noncharged and will not be expected to have significant electrostatic interactions. Above pH 5.7, however, most aspartates and glutamates will be negatively charged, and any electrostatic interaction in which they participate will be manifest. In the absence of any other effects, a mutation that introduces a favorable interaction of an Asp or Glu with an N-terminal helix dipole is therefore expected to have little effect on stability at pH 2.0 and to increase stability near neutral pH. Five of the replacements shown in Table I, S38D, T109D, T115E, N116D, and N144D, have this characteristic. What is remarkable is that four of these are the replacements that were designed to interact with α-helix dipoles and all have been successful. Each designed mutant

increases the melting temperature of the protein by 1.4–1.6 °C at pH 5.7–6.9. The replacements S38D, N116D, and N144D were chosen in the initial selection (Nicholson et al., 1988). T109D was obtained by slightly relaxing the initial criteria (this work). T115E was generated for other purposes, but crystallographic and thermodynamic studies indicate that most of its favorable electrostatic interaction is with an α-helix (Dao-pin et al., 1991b).

The crystal structures of the mutant proteins, together with the stability data for the primary and control mutants (Table III), show quite convincingly that the observed stabilization near neutral pH is due to interaction of the introduced Asp or Glu with the α-helix dipole. At position 38, the Ser → Asp replacement increases stability at pH 6.7 but not at pH 2.0. The isostructural replacement Ser → Asn does not increase stability at either pH. This strongly suggests that it is the negative charge on Asp 38 that is essential for enhanced stability. Conversely, the replacement of the presumed stabilizing WT residue Asp 92 with Asn results in a decrease in stability at pH 6.7 but essentially no change at pH 2.0. This also indicates that it is the negative charge on Asp 92 that is essential for the favorable interaction. Replacement of Thr 109 with Asp decreases stability at pH 2.0 but increases stability at pH 6.7, again consistent with the requirement that the aspartate be charged to interact favorably with the α-helix dipole. The control mutant T109N has stability essentially the same as that of the wild-type. Although T115E was not designed as a charge–dipole stabilizing mutant, it shows increased stability at neutral pH relative to acid pH (Dao-pin et al., 1991b). At position 116 the mutant N116D increases stability in a pH-dependent manner consistent with an electrostatic interaction. Because of the concern that this apparent favorable electrostatic interaction might be with Arg 119, this residue was replaced with methionine. The double mutant N116D/R119M retains the same pH-dependent increase in stability relative to R119M as does N116D relative to WT. This indicates that the stabilization seen in each case is due to electrostatic interaction between Asp 116 and the helix dipole and does not depend on an interaction between Asp 116 and Arg 119. Finally, the replacement of Asn 144 with Asp also increases stability at pH 6.9 but not at pH 2.0. The control mutant, Asn 144 → His, was designed to introduce a residue of the "wrong" charge at the end of an α-helix. At low pH, where His 144 is positively charged, the mutant protein is indeed destabilized (by 0.6 kcal/mol). Surprisingly, at pH 6.7, the variant is slightly stabilized relative to wild-type (0.3 kcal/mol) although not to the same degree as the designed charge–dipole mutants. Because N144H could not be crystallized, possible interactions of His 144 in the mutant structure remain uncertain.

Nature the of Charge–Dipole Interaction. It was previously suggested that stabilization of T4 lysozyme due to the interaction of a charged amino acid with an α-helix dipole is electrostatic in nature and does not require specific hydrogen bonding between the amino acid and the peptide backbone at the end of the helix [Nicholson et al., 1988; see also Shoemaker et al. (1987)]. This was based in part on the crystal structure of N144D, which showed that neither Asn 144 in wild-type lysozyme nor Asp 144 in the mutant structure makes H-bonds to the amide groups in the first turn of the helix. Also, in the case of the stabilizing mutant S38D, the serine had better H-bonding geometry than the aspartate, yet the aspartate gives a more stable protein above pH ~3 (Nicholson et al., 1988). All of the structural data now available support the importance of enhanced electrostatic interaction rather than improved hydrogen bonding in accounting for the enhanced stability of the mutant structures. In the case of mutant S38N, for example, the Asn 38 side chain forms essentially the same hydrogen bonds with the end of α-helix 39–50 as does Asp 38 (Figure 2c), yet S38D is more stable than S38N (at pH 6.7). The difference in stability cannot reasonably be attributed to differences in H-bonding. Rather, it must be due to the difference in charge. The side chains of Asp 109 in T109D, Asp 116 in N116D, and Asp 144 in N144D all occur at the N2 position of an α-helix (Table I) and all have similar conformations relative to the backbone of the helix within which they are located. In all three cases, the side chain of the introduced aspartate is not located directly over the end of the α-helix but is substantially off-axis. Steric restraints prevent hydrogen bonding between the side chain and any of the amides within the first turn of the α-helix. (A δ-oxygen of each introduced aspartate is 3.2–3.6 Å from its own amide, but the geometry for a hydrogen bond is very poor, with the angle C^{γ}–O^{δ}···N ranging from 52° to 74°.) Similarly, in the proteins with asparagine at each of these three sites, there is no hydrogen bonding between the side chain and the end of the helix. For both the aspartic acid and asparagine variants, the amides in the first turn of the helix appear to hydrogen bond only to solvent. These observations all support the idea that stabilization of the helix dipole does not *require* hydrogen bonding to the amides or carbonyls in the first or last turn of the helix. This is not to say, however, that such hydrogen bonds, when they occur, do not contribute to the interaction with the helix dipole (see the section below on the energetics of helix–dipole interactions).

We believe that we have ruled out the possibility that the stabilization arising from the introduction of aspartates or glutamates might be due to adventitious electrostatic interactions with positively charged groups in the vicinity. For replacements at position 38, there is no positively charged group within a distance of 9 Å. In the case of position 109, the closest charged group is the carboxylate of Glu 108, which is within 5.5 Å. Notwithstanding a possible unfavorable electrostatic interaction with this group, T109D is still a stabilizing substitution (at pH 6.7), indicating that the interaction of Asp 109 with the helix dipole is much stronger than the repulsion between Asp 109 and Glu 108. In the case of replacements at position 116, there was concern that the introduced aspartate might interact favorably with Arg 119. When Arg 119 was replaced with a noncharged methionine, stability was essentially unaffected, showing that it is the interaction of Asp 116 with the helix dipole that is the origin of enhanced stability. A related control also supports interaction with α-helix dipoles as the basis of the stabilization observed for T115E (Dao-pin et al., 1991b).

It might be argued that the increased stability of aspartate- or glutamate-substituted mutants of T4 lysozyme is due to neutralization of the overall positive charge on the molecule. To test this possibility, Dao-pin et al. (1991c) created multiply mutated lysozymes in which the overall formal charge on the molecule was sequentially reduced from +9 to +1 units at neutral pH. Rather than an increase, a slight loss of stability was observed for most of the single and multiple replacements. This seems to eliminate the possibility that reduction of the overall positive charge on T4 lysozyme is responsible for the increased stability observed for aspartic acid and glutamic acid substitutions.

Location of Substitutions. Table I summarizes the locations of all the substitutions in Table III that involve a change in charge. This includes all the substitutions that result in a

significant change in protein stability.

Three of the four stabilizing aspartates were introduced at the second (or "N2") position in an α-helix. Interestingly, statistical analysis of amino acid frequencies at various positions in α-helices indicates that the most frequent residue at the N2 position is aspartate (Richardson & Richardson, 1988). As anticipated by Blagdon and Goodman (1975), the interaction involving helical dipoles demonstrated here gives an energetic rationale for this empirical observation. Consistent with the idea that the second residue in an α-helix tends to interact with the helix dipole, one of the least frequently observed amino acids at this position is histidine (Richardson & Richardson, 1988). The mutant protein N144H directly demonstrates the unfavorable energetics associated with a charged histidine at the second position. Thus, the unfavorable charge–dipole interaction provides a rationale for the low frequency of histidine at the N2 position in α-helices.

At the N2 position of most helices in T4 lysozyme, either an aspartate appears to have favorable electrostatic interactions with the helix dipole or the side chain has hydrophobic interactions with the core. Helix A is capped by Asn 2. Glu 5, the N3 residue, is hydrogen bonded to Asn 2 and may help satisfy the electrostatic potential at the N-terminus of the helix. Phe 4, the N2 residue, has hydrophobic contacts with Phe 67, Asn 68, and Val 71. Presumably its replacement with Asp would be destabilizing. The stabilizing mutant S38D indicates that helix B in the wild-type protein has unsatisfied electrostatic potential. A genetic screen of randomly generated mutants (P. Pjura and M. Matsumura, unpublished results) yielded a mutant that introduces aspartate at position 40, the N2 position, which has either increased activity or stability, suggesting a favorable electrostatic interaction might exist with the helix. Helix C, the longest helix, has Asp 61 at the N2 position in the wild-type structure in a conformation similar to that displayed by the introduced aspartate in T109D, N116D, and N144D. The absence of positively charged side-chains within 7.5 Å of the carboxylate at position 61 in the wild-type model suggests that any favorable electrostatic interactions would be with the N-terminus of the helix. Helix D is distorted by a proline (Alber et al., 1988). Its N-terminal electrostatic potential appears to be satisfied by Glu 108, which hydrogen bonds to the capping residue, Asn 81 (J. Bell, unpublished results). It is unclear whether Lys 83, the N2 residue, is important for stability. The N-terminal electrostatic potential of helix E is well satisfied by a tight interaction with Asp 92 at the capping position. The pH-dependent stability changes associated with replacement of the side chain by Asn indicate favorable electrostatic interactions in the wild-type structure. Thus, a charged group at the N2 position is unnecessary. Stabilizing interactions between the amino-terminus of helix F and the charged side chain of Asp 109 are demonstrated above. Our data also suggest that Asp 116 stabilizes helix G by virtue of charge–dipole interactions. Asp 127, at the N2 position of helix H in wild-type lysozyme, is in a similar conformation to that observed for aspartates at positions 61, 109, 116, and 144. Helix I has no negatively charged side chain at its amino-terminus, but the terminus is very close to the C-terminus of helix H such that the two helices stabilize each other by interaction of their respective dipoles. The N2 residue of helix I is a Trp, which is essential to core stability and was not replaced. Finally, mutant N144D demonstrates that helix J can be stabilized by charge–dipole interactions with Asp at the N2 position. The last helix, K, is omitted from consideration because it is short and distorted (Weaver & Matthews, 1987).

In summary, of the nine ordered helices in T4 lysozyme, six (B, C, F, G, H, and J) either have Asp at the N2 position or appear to be stabilized by mutations that introduce charge–dipole interactions involving the N2 position. Two of the remaining helices, A and I, require other side chains (Phe and Trp) at the N2 position, while helix E is already capped by aspartate. The number of charge–dipole interactions in T4 lysozyme involving the N2 position suggests such interactions are important for stability and are a dominant reason for the high occurrence of aspartate at the N2 position in α-helices in general.

Energetics of Helix–Dipole Interactions. The contribution of an electrostatic interaction to the stability of a protein can be measured in two ways, either from the pH dependence of stability or from the shift in pK_a of the interacting group.

There is a thermodynamic linkage between the acid dissociation constant of a titratable group in a protein and the contribution of that group to the pH-dependent stability of that protein. The equation that defines this relationship (Cantor & Schimmel, 1980) is

$$\Delta\Delta G_E = 2.303RT(pK_a^F - pK_a^U)$$

where ($pK_a^F - pK_a^U$) is the difference between the pK_a values of the titratable group in the folded and unfolded forms of the protein and $\Delta\Delta G_E$ is the contribution to the free energy of stabilization of the protein due to the electrostatic interaction of the titratable group. We use $\Delta\Delta G_E$ rather than ΔG_E to indicate that the equation gives the contribution to stability of a single amino acid. The equation is valid at pH values where the group is completely titrated (a reasonable approximation 2 pH units or more from the pK_a of the ionizing group). Direct measurement of shifts in pK_a, for example by NMR, therefore provide a direct assessment of the electrostatic contribution of the ionizable group to the stability of the folded state (Anderson et al., 1990). pH-dependent changes in unfolding free energy and shifts in pK_a that occur upon folding are different aspects of the same physical phenomena and can be compared. These comparisons are approximate, because the conditions of the measurements are different, especially the temperature of the determination.

For the mutant S38D, the shift in pK_a is from 4.0 in the unfolded state to 3.0 in the folded protein, corresponding to an energy of interaction of 1.4 kcal/mol. Unfolding studies of S38D were carried out at several pH values (Nicholson et al., 1988). Comparing values of ΔT_m at pH 2 with those at pH 6.9, and using the perturbation theory approach (Becktel & Schellman, 1987) to convert to free energy, gives a value of 0.7 kcal/mol for electrostatic stabilization by Asp 38. The stabilization measured by thermal denaturation is about half that calculated from the shift in pK_a (Figure 6). An explanation for this apparent discrepancy might be that residue 38 is at the N_{cap} position and in both the wild-type and mutant structures hydrogen bonds directly to the amides within the first turn of the α-helix [Figure 2c and Figure 4b of Nicholson et al. (1988)]. The stability measurements, which compare the Asp variant with wild-type (Ser), will reflect the enhanced electrostatic interaction of Asp 38 relative to Ser 38, but the hydrogen-bonding and van der Waals interactions in the two cases will largely cancel. The same is true if one compares the stability of the Asp 38 variant with the Asn 38 protein. To some degree the NMR measurement may reflect the combined influence on Asp 38 of its electrostatic and hydrogen-bonding interactions, whereas the thermodynamic stability measurement may represent the incremental electrostatic interaction experienced by Asp 38 relative to Ser 38 (or Asn 38). At the same time the apparent differences between the NMR

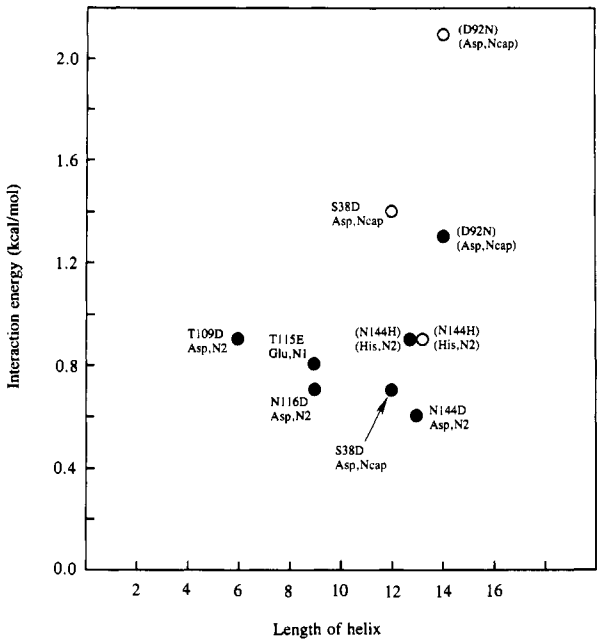

FIGURE 6: Helix–dipole interaction energies for different amino acids in T4 lysozyme. Solid circles indicate interaction energies measured thermodynamically (Table II). Open circles indicate interaction energies estimated via NMR (see text). Note that the thermodynamically determined energies are derived from differences in stability as a function of pH ($\Delta\Delta G_E$, Table II) and therefore measure the electrostatic interaction energy but do not include possible contributions from hydrogen bonding. The NMR measurements are under different experimental conditions and do not measure the same quantity, so the energy values cannot be compared directly (see text). In each case, the mutant on which the measurement is based is identified (e.g., T109D) as well as the type of amino acid (Asp) and its location in the α-helix (N2 position). Mutants enclosed within parentheses correspond to destabilizing substitutions (D29N removes Asp 92 from its helix-cap position in wild-type lysozyme; N144H introduces a histidine at the "wrong" end of an α-helix). The other substitutions all increase the stability of T4 lysozyme.

and thermodynamic measurements may simply reflect the fact that the two techniques are measuring different quantities under different experimental conditions. (See also the discussion for Asp 92, below.)

Mutant N144H was also analyzed by NMR, and the pK_a of the histidine was found to decrease by 0.6 units in the folded protein relative to the solution value measured in the unfolded protein (Anderson et al., 1990). This implies a destabilization of 0.9 kcal/mol. Although a crystal structure is not available for the mutant N144H, both the pH dependence of stability and the shift in the pK_a of His 144 are consistent with the expected unfavorable interaction between the helix dipole and the protonated form of the imidazole. The X-ray crystallographic and stability studies of same-site mutation N144D indicate directly that Asp 144 does participate in charge–dipole interactions and suggest by analogy that His 144 might do likewise. In this case the difference of 0.9 kcal/mol between the protein thermostability at pH 2.0 and 6.7 (Table III) is in good agreement with the value of 0.9 kcal/mol from NMR. The magnitude and sign of the electrostatic interaction energy is consistent with a model in which an unfavorable electrostatic interaction between the charged imidazole and the peptide groups at the N-terminus of the α-helix is the dominant factor in the destabilization at low pH associated with the mutation N144H.

Because pK_a values are not yet available for the aspartates associated with the mutations T109D, N116D, and N144D, thermodynamic stability measurements are the only basis on which to estimate apparent electrostatic interaction energies. These estimates are approximately 0.9 kcal/mol for T109D, 0.6 kcal/mol for N116D, and 0.7 kcal/mol for N144D. The apparent electrostatic interaction energy of the Glu 115 in T115E is 0.8 kcal/mol (Dao-pin et al., 1991b).

The shift in pK_a observed for the Asp 92 in wild-type lysozyme indicates an interaction energy of about 2.1 kcal/mol. Thermostability measurements indicate a value of 1.3 kcal/mol. As was the case with residue 38, residue 92 is at the N_{cap} position so that Asp 92 not only interacts electrostatically with the helix dipole but also makes hydrogen bonds to the amide nitrogens at the ends of the α-helix. Asp 92 also forms a salt bridge with Arg 95, which might explain why its interaction energy is larger than the other examples shown in Figure 6. The stability measurements also suggest that Asp 92 interacts favorably with the helix dipole via both electrostatic interaction and hydrogen bonding. The NMR measurement is at 10 °C in 100 mM KCl and 10 mM phosphate buffer and reflects the difference between a protonated and nonprotonated carboxylate group. The thermodynamic measurement is at the melting temperature of the mutant protein (i.e., about 62 °C at pH 6.7) in 150 mM KCl and 10 mM phosphate buffer and relies on the comparison of a mutant (Asn) and a reference (Asp) structure at two pH values. Because of these severalfold differences, it is not possible to make a direct comparison between the energy values estimated from the two techniques.

Serrano and Fersht (1989) made a series of different replacements for Thr 6 and Thr 26, two N_{cap} residues in barnase, and found that the change in the free energy of unfolding ranged from a 2.53 kcal/mol decrease to an increase in stability of 0.11 kcal/mol. Asp and Glu were found to provide $\sim$1.3 kcal/mol of electrostatic stabilization as compared to Asn and Gln. Related results have been obtained with multiple replacements of Thr 59 in T4 lysozyme (J. A. Bell, W. J. Becktel, U. Sauer, and B.W.M., unpublished observations). These results, as well as those summarized in Figure 6, show that the overall increase (or decrease) in stability of a protein resulting from mutations designed to interact with helix dipoles depend not only on the amino acid introduced but also on the amino acid replaced and the context. Nevertheless, the observation that stabilization can be achieved by electrostatic interaction and does not require hydrogen bonding to the end of the α-helix greatly simplifies the design of stabilizing replacements.

Although the introduction of acidic groups close to the amino-termini of α-helices consistently increase stability, mutations that were designed to introduce new salt bridges on the surface of T4 lysozyme caused minimal changes in stability (Dao-pin et al., 1991b). It appears that the electrostatic interaction energy between mobile side chains on the surface of a protein is offset by the entropy cost of immobilizing or restricting the motion of the interacting partners (Dao-pin et al., 1991b). In the case of an α-helix dipole, however, the entropy cost of localizing the partial positive charges at the N-terminus of the helix and the partial negative charge at the C-terminus is paid for during the folding of the protein. It may be for this reason that interactions of charged groups (especially aspartate, with its limited conformational flexibility) with α-helix dipoles are consistently effective in increasing protein stability, whereas electrostatic interactions between mobile solvent-exposed charged groups on the surface of the protein are ineffectual.

It is always possible that a given mutation may change the stability of the protein by an effect on the unfolded rather than the folded form. In the present case, four different mutants

were designed on the basis of the wild-type crystal structure to increase stability via helix–dipole interactions, and all proved to be successful. The pH dependence of stability of these mutants as well as a number of additional control mutants can be understood on the basis of their observed crystal structures. In no case is there a large unexplained change in stability. It suggests that if these mutants do change the stability of the unfolded form of the protein, then the effects must be small relative to the changes in energy in the folded structure.

Nature of the α-Helix Dipole. Although the manifestations of the "α-helix dipole" are now well established, the origin of the helix dipole remains a subject of debate. Should the α-helix dipole be considered as a macrodipole arising from the sum of the individual dipoles associated with each peptide group within the α-helix (Blagdon & Goodman, 1975; Hol et al., 1978, 1981)? Alternatively, is the α-helix dipole simply a manifestation of the concentration of local unsatisified charges on the amides and carbonyls within the first and last turn of the helix (Pflugrath & Quiocho, 1985; Aqvist et al., 1991)? According to the "macrodipole model" the strength of the helix dipole should become greater as the helix increases in length (Hol et al., 1978, 1981). However, once there are about 10 residues in the α-helix, the calculated potential on the helix axis becomes practically independent of length (Hol et al., 1978). Several authors have suggested that the electrostatic effect of the helix dipole can be well-represented by placing half a positive unit charge on the helix axis near the N-terminus and half a negative unit charge near the C-terminus (Hol et al., 1978, 1985; Sheridan & Allen, 1980; Sheridan et al., 1982). One of the referees has kindly used this model to estimate that the expected interaction energy of a negatively charged group 5 Å from the N-terminus of a six-residue helix is about 0.7 kcal/mol and only increases to about 0.9 kcal/mol for a 13-residue helix. This weak dependence on length is very similar to that estimated by Hol et al. (1978) and Sheridan and Allen (1980).

Figure 6 summarizes the apparent helix–dipole interactions observed in T4 lysozyme, plotted as a function of helix length. Although Figure 6 includes all the interaction energies measured in this paper, it is not appropriate that each of these values should be treated equally in assessing possible length dependence. In the first place, as described in the preceding section, the energies derived from NMR (open circles) do not necessarily measure the same interaction energy as those determined from thermodynamics (solid circles). Therefore these values cannot be directly compared. Also it is not appropriate to compare energies for residues at different locations in the α-helix. There is, for example, a suggestion that residues at the N_{cap} position tend to have higher interaction energies than those at the N1 or N2 position.

It does seem appropriate, however, that the energies estimated for the three mutants T109D, N116D, and N144D be compared. For each of these three mutants the local geometry is extremely similar. In all three cases an aspartate is introduced at the N2 position; in all three cases the residue is fully solvent exposed; and in all three cases the introduced acidic groups have virtually identical conformations. It might be added that in each case one can also directly compare the stability and structural data for the asparagine/aspartate pair. Although Figure 6 shows the stability of T109D relative to wild type, this is virtually identical with the stability of T109D relative to T109N (Table III). For the three structurally similar examples, T109D, N116D, and N144D, the interaction energy decreases from 0.9 to 0.6 kcal/mol while the length of the α-helix increases from 6 to 13 residues. By allowing some variability in the estimated energies due to context dependence and to possible experimental error, Figure 6 suggests that the interaction energies do not depend strongly on the length of the helix. This is consistent with a model in which the effects of an α-helix dipole can be represented by positive and negative charges at the N- and C-termini of the α-helix.

The focus of the present study has been to investigate the interaction of acidic groups with the amino-termini of α-helices. Within an α-helix the C^{α}–C^{β} bonds are tilted toward the amino-terminus. For this reason the side chains emanating from the N_{cap}, N1, and N2 positions are likely to interact with the α-helix dipole. In contrast, the side chains toward the C-terminus of an α-helix tend to be directed toward the middle rather than the end of the helix and may be less likely to interact strongly with the α-helix dipole. Aspartic and glutamic acid are observed to occur with higher than average frequency toward the amino-termini of α-helices in known protein structures. Lysine and arginine occur with above-average frequency toward the carboxy-termini of helices, although the preference is not as pronounced as for the acidic residues (Richardson & Richardson, 1988). As shown in the present study, amino acid substitutions that enhance interaction with the helix dipole at the amino-terminus of an α-helix can often be located within the helix itself. This may be less effective at the carboxy-terminus, although it needs to be tested. There are examples of strong interactions between basic groups and the dipole at the carboxy-termini of α-helices. Such interacting groups tend to be located just beyond the end of the helix [e.g., Perutz et al. (1985) and Sali et al. (1988)] or in other parts of the polypeptide backbone [e.g., Weaver et al. (1989)]. There may, therefore, be less candidates available for substitutions designed to interact with the carboxy-termini of α-helices than is the case with the amino-termini.

Acknowledgments

We are most grateful to J. Wozniak for technical assistance. We thank Dr. E. Eriksson for the X-ray data of pseudo-wild-type lysozyme, Drs. J. Bell, L. H. Weaver, W. Baase, and E. Baldwin for advice and helpful discussions, Dr. P. Pjura for access to unpublished data, Drs. F. W. Dahlquist and J. A. Schellman for providing facilities for the NMR and thermodynamic studies, and Dr. W. J. Becktel for preliminary thermodynamic analyses of some of the lysozyme variants. We also thank one of the referees for an enlightening, detailed exposition on the length dependence of the helix–dipole interaction.

References

Aqvist, J., Luecke, H., Quiocho, F. A., & Warshell, A. (1991) *Proc. Natl. Acad. Sci. U.S.A. 88*, 2026–2030.

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

Anderson, D. E., Becktel, W. J., & Dahlquist, F. W. (1990) *Biochemistry 29*, 2403–2408.

Becktel, W. J., & Baase, W. A. (1987) *Biopolymers 26*, 619–623.

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

Bell, J. A., Wilson, K. P., Zhang, X.-J., Faber, H. R., Nicholson, H., & Matthews, B. W. (1991) *Proteins: Struct., Funct., Genet. 10*, 10–21.

Blagdon, D. E., & Goodman, M. (1975) *Biopolymers 14*, 241–245.

Cantor, C. R., & Schimmel, P. R. (1980) *Biophysical Chemistry Part 1: The Conformation of Biological Mac-*

romolecules, pp 44–45, W. H. Freeman, San Francisco.
Dao-pin, S., Alber, T., Baase, W. A., Wozniak, J. A., & Matthews, B. W. (1991a) *J. Mol. Biol.* (in press).
Dao-pin, S., Sauer, U., Nicholson, H., & Matthews, B. W. (1991b) *Biochemistry 30*, 7142–7153.
Dao-pin, S., Soderlind, E., Baase, W. A., Wozniak, J. A., Sauer, U., & Matthews, B. W. (1991c) *J. Mol. Biol.* (in press).
Hol, W. G. J. (1981) *Nature 294*, 532–536.
Hol, W. G. J. (1985) *Prog. Biophys. Mol. Biol. 45*, 149–195.
Hol, W. G. J., van Duijnen, P. T., & Berendeson, H. J. C. (1978) *Nature 273*, 443–446.
Ihara, S., Ooi, T., & Takahashi, S. (1982) *Biopolymers 21*, 131–145.
Kunkel, T. A., Roberts, J. D., & Zakour, R. A. (1987) *Methods Enzymol. 154*, 367–382.
Matsumura, M., & Matthews, B. W. (1989) *Science 243*, 792–794.
Matsumura, M., Becktel, W. J., Levitt, M., & Matthews, B. W. (1989) *Proc. Natl. Acad. Sci. U.S.A. 86*, 6562–6566.
McIntosh, L. P., Wand, A. J., Lowry, D. F., Redfield, A. G., & Dahlquist, F. W. (1990) *Biochemistry 29*, 6341–6362.
Mitchison, C., & Baldwin, R. L. (1986) *Proteins: Struct., Funct., Genet. 1*, 23–33.
Muchmore, D. C., McIntosh, L. P., Russell, C. B., Anderson, D. E., & Dahlquist, F. W. (1989) *Methods Enzymol. 177*, 44–73.
Nicholson, H., Becktel, W. J., & Matthews, B. W. (1988) *Nature 336*, 651–656.
Perutz, M. F., Gronenborn, A. M., Clore, G. M., Fogg, J. H., & Shih, D. T.-b. (1985) *J. Mol. Biol. 183*, 491–498.
Pflugrath, J. W., & Quiocho, F. A. (1985) *Nature 314*, 257–260.
Pjura, P., Matsumura, M., Wozniak, J., & Matthews, B. W. (1990) *Biochemistry 29*, 2592–2598.
Poteete, A. R., Dao-pin, S., Nicholson, H., & Matthews, B. W. (1990) *Biochemistry 30*, 1425–1432.
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.
Richardson, J. S., & Richardson, D. C. (1988) *Science 240*, 1648–1652.
Rossmann, M. G. (1979) *J. Appl. Crystallogr. 12*, 225–239.
Sali, D., Bycroft, M., & Fersht, A. R. (1988) *Nature 335*, 740–743.
Schmid, M. F., Weaver, L. H., Holmes, M. A., Grutter, M. G., Ohlendorf, D. H., Reynolds, R. A., Remington, S. J., & Matthews, B. W. (1981) *Acta Crystallogr. A37*, 701–710.
Serrano, L., & Fersht, A. R. (1989) *Nature 342*, 296–299.
Sheridan, R. P., & Allen, L. C. (1980) *Biophys. Chem. 11*, 133–136.
Sheridan, R. P., Levy, R. M., & Salemme, F. R. (1982) *Proc. Natl. Acad. Sci. U.S.A. 79*, 4545–4549.
Shoemaker, K. R., Kim, P. S., Brems, D. N., Marqusee, S., York, E. J., Chaiken, I. M., Stewart, J. M., & Baldwin, R. L. (1985) *Proc. Natl. Acad. Sci. U.S.A. 79*, 2349–2353.
Shoemaker, K. R., Kim, P. S., York, E. J., Stewart, J. M., & Baldwin, R. L. (1987) *Nature 326*, 563–567.
Tronrud, D. E., Ten Eyck, L. F., & Matthews, B. W. (1987) *Acta Crystallogr. A43*, 489–503.
Wada, A. (1976) *Adv. Biophys. 9*, 1–63.
Weaver, L. H., Gray, T. M., Grütter, M. G., Anderson, D. E., Wozniak, J. A., Dahlquist, F. W., & Matthews, B. W. (1989) *Biochemistry 28*, 3793–3797.
Weaver, L. H., & Matthews, B. W. (1987) *J. Mol. Biol. 193*, 189–199.
Zoller, M. J., & Smith, M. (1984) *DNA 3*, 479–488.