Simulation Analysis of the Stability Mutant R96H of T4 Lysozyme†

Bruce Tidor‡ and Martin Karplus*
Department of Chemistry, Harvard University, Cambridge, Massachusetts 02138
Received July 17, 1990; Revised Manuscript Received November 27, 1990

ABSTRACT: Free energy simulation methods are used to analyze the effects of the mutation Arg 96 → His on the stability of T4 lysozyme. The calculated stability change and the lack of significant structural rearrangement in the folded state due to the mutation are in agreement with experimental studies [Kitamura, S., & Sturtevant, J. M. (1989) *Biochemistry 28*, 3788–3792; Weaver, L. H., et al. (1989) *Biochemistry 28*, 3793–3797]. By use of thermodynamic integration, the contributions of specific interactions to the free energy change are evaluated. It is shown that a number of contributions that stabilize the wild type or the mutant partially cancel in the overall free energy difference; some of these involve the unfolded state. Comparison of the results with conclusions based on structural and thermodynamic data leads to new insights into the origin of the stability difference between wild-type and mutant proteins. Of particular interest is the importance of the contributions of more distant residues, solvent water, and the covalent linkage of the mutated amino acid. Also, the analysis of the interactions of Arg/His 96 with the C-terminal end of a helix (residues 82–90) makes it clear that the nearby carbonyl groups (Tyr 88 and Asp 89) make the dominant contribution, that the amide groups do not contribute significantly, and that the helix–dipole model is inappropriate for this case.

A knowledge of the determinants of protein stability is essential for an understanding of the folding of native proteins, as well as for the engineering of mutants with specific properties (Oxender & Fox, 1987). The renewed interest in this area is due largely to the ability to make modified proteins almost at will by the techniques of site-directed mutagenesis. Such amino acid modifications, as well as those from random mutagenesis, are being combined with thermodynamic measurements of mutant and wild-type proteins of known structure to obtain experimental information concerning the contributions to protein stability (Shortle & Meeker, 1986; Matthews, 1987; Matouschek et al., 1989). Since the same factors are involved in other phenomena, such as molecular recognition and allostery, analyses of protein stability are likely to increase our understanding of structure–function relationships at the molecular level, which is important for rational design goals, such as the development of inhibitors to interfere with virus receptor binding (Rossman et al., 1985; Wiley & Skehel, 1987) and the alteration of enzyme activity (Knowles, 1987).

Although much is being learned from the experimental approach, it does have limitations (Alber, 1989). The focus on the crystal structure, per se, usually neglects possible entropic contributions due to internal motions and effects of the mutation on the unfolded state. Crystallographic temperature factors do have a dynamic contribution (Petsko & Ringe, 1984), but their use without corrections for rigid-body motions (Kuriyan & Weis, 1991), disorder (Ringe & Petsko, 1985; Sheriff et al., 1987), and the effects of packing constraints (Phillips, 1990) is problematic (Holbrook & Kim, 1984; Pjura et al., 1990). Longer range interactions, such as those due to electrostatics, are difficult to evaluate by visual inspection. Also, because the measured change in stability may involve a number of contributions, both positive and negative (Gao et al., 1989), they may be difficult to separate by experiment (Alber, 1989). These limitations suggest that it is of interest to use theoretical approaches, such as normal mode calculations (Brooks & Karplus, 1983), quasiharmonic dynamic analyses (Karplus & Kushick, 1981), and free energy simulations (Beveridge & DiCapua, 1989), to provide additional information concerning the factors that contribute to the thermodynamics.

It has been shown that molecular simulation techniques can estimate the changes in free energy responsible for a wide range of phenomena. Applications include the relative solvation free energy of ions and small molecules (Jorgensen & Ravimohan, 1985; Lybrand et al., 1985; Fleischman & Brooks, 1987; Bash et al., 1987a; Singh et al., 1987), the solvation contribution to the free energy change of organic reactions (Chandrasekhar et al., 1984), the difference in binding between enzyme inhibitors (Wong & McCammon, 1986; Warshel et al., 1986; Bash et al., 1987b), an analysis of cooperativity in hemoglobin (Gao et al., 1989), and the change in protein stability induced by mutations (Wong & McCammon, 1986; Bash et al., 1987a; Dang et al., 1989). So far, many of the calculations have focused more on demonstrating agreement with experiment than on providing insights concerning the interactions involved.

One protein that has been the object of a series of detailed studies by mutagenesis is the lysozyme of the bacteriophage T4 (Matthews, 1987). The molecule is a single polypeptide chain of 164 amino acid residues (molecular mass 18.6 kDa), which folds into two compact domains connected by a long α-helix; there are two cysteine residues but no disulfide bridges. When the bacteriophage enters the lytic phase of its life cycle, it is the function of the lysozyme to hydrolyze the glycosidic linkages in the cell wall of its bacterial host. The crystal structure for the wild-type enzyme has been refined to 1.7 Å (Weaver & Matthews, 1987) and a number of stability mutants have been produced and studied by crystallography (Grütter et al., 1979; Alber et al., 1987a,b; Matthews, 1987; Weaver et al., 1989) and thermodynamic measurements (Hawkes et al., 1984; Becktel & Baase, 1987; Becktel & Schellman, 1987; Kitamura & Sturtevant, 1989). Examples include the effects on stability of hydrophobic residues (Matsumura et al., 1988), helix–dipole interactions (Nicholson et al., 1988), cavities (Matsumura et al., 1988), disulfide bonds (Matsumura et al., 1989), the conformational freedom of

† Supported in part by grants from the National Science Foundation and the National Institutes of Health.
‡ Present address: Whitehead Institute for Biomedical Research, Nine Cambridge Center, Cambridge, MA 02142.

0006-2960/91/0430-3217$02.50/0

glycine and proline residues (Matthews et al., 1987), and the disruption of hydrogen bonds (Alber et al., 1987a). Analogous mutation studies have been made by Fersht and co-workers on barnase (Matouschek et al., 1989) and by Shortle and co-workers on staphylococcal nuclease (Shortle & Lin, 1985; Shortle & Meeker, 1986), although the thermodynamic measurements have not so far been complemented by crystallographic analyses.

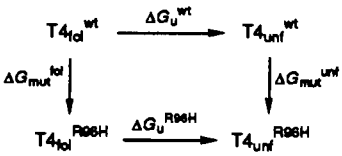
We have chosen the mutant Arg 96 → His (R96H) of T4 lysozyme for study by the free energy simulation method. This mutant is of particular interest because it does not fall into an obvious mutation class (at the pH of the experiments both residues are charged) and because of the relatively large effect of the mutation; the change in stability induced by the mutation is estimated to be 3.3 kcal/mol at pH 3 and 320 K (Hawkes et al., 1984). The crystal structure of the mutant shows that it is nearly identical with the wild type except for small localized changes in the neighborhood of the mutation. Examination of the wild-type and mutant crystal structures, with the assumption that all of the free energy difference arises in the folded protein, has led to the proposal that bond angle strain and the loss of helix–dipole interactions are the major source of the stability change (Weaver et al., 1989).

In the present work, two different local conformations for His 96 are considered in the folded state; this was done because the mutant crystal structure was not available to us when the simulations were made and because we believed that a comparison of the two histidine orientations would yield insights into the origin of the effects of the mutation. In each case, the simulations were begun from the wild-type crystal structure, and the free energy change for introducing the mutation was calculated. For the unfolded state a simple extended chain model was used. This is likely to be sufficient in the present problem because both the wild-type Arg and the mutant His residue are expected to be most stable with the charge of the side chain fully solvated. Analysis of the free energy simulation results, which provide an estimate of the observed stability change in approximate agreement with experiment, indicates that there is a significant effect of the mutation on the stability of the unfolded, as well as the folded, state. By use of thermodynamic integration, the contributions from the solvent and different parts of the protein environment were isolated. These provide detailed information for comparison with the experimental analysis and increase our understanding of the specific interactions that are responsible for the calculated free energy difference.

Section I describes the methods used in the free energy simulation and section II presents the results; they are discussed in section III. The conclusions are summarized in section IV.

I. Methods

A brief outline of the free energy simulation method used here is presented. A full description is given elsewhere [S. H. Fleischman, B. Tidor, C. L. Brooks, III, and M. Karplus, manuscript in preparation; see also the review in Beveridge and DiCapua (1989)]. To simplify the problem (i.e., to avoid the need for simulating the transition from the folded to the unfolded state), we make use of the thermodynamic cycle (Tembe & McCammon, 1984; Brooks et al., 1988):



(1)

Experiment measures the free energies of denaturation, ΔG_u^{wt} and ΔG_u^{R96H}, of the wild type and mutant, respectively. By contrast, a free energy simulation in which the Arg is changed to a His during the calculation for the unfolded and folded state yields values for ΔG_{mut}^{unf} and ΔG_{mut}^{fol}, respectively. The simulation results can be related to the experimental measurements via the thermodynamic cycle shown in (1); that is

$$\Delta(\Delta G) = \Delta G_u^{R96H} - \Delta G_u^{wt} = \Delta G_{mut}^{unf} - \Delta G_{mut}^{fol} \quad (2)$$

To obtain the free energy differences, ΔG_{mut}^{fol} and ΔG_{mut}^{unf}, a hybrid potential function of the form

$$U(\lambda) = (1 - \lambda)U_0 + \lambda U_1 \quad (3)$$

is used, where U_0 is the potential energy function for the wild-type system (with Arg 96) and U_1 is that for the mutant system (with His 96). The free energy changes ΔG_{mut}^{fol} and ΔG_{mut}^{unf} were calculated in two ways. The exponential formula (Zwanzig, 1954)

$$\Delta G(\lambda_i \rightarrow \lambda_{i+1}) = -\frac{1}{\beta} \ln \langle e^{-\beta[U(\lambda_{i+1})-U(\lambda_i)]} \rangle_{\lambda_i} \quad (4)$$

was used with $\beta = 1/k_B T$ and $\langle\ \rangle_\lambda$ an ensemble average determined with the potential function $U(\lambda)$. In principle, the free energy change from wild type to mutant can be obtained with eq 4 in a single step. In practice, this is not efficient because the change is sufficiently drastic that the ensemble average in eq 4 would require a very long simulation to converge. As others have done, we have carried out the mutation in a number (10 in the present case) of smaller, discrete steps such that in each step there is a relatively small perturbation. The total free energy is then the sum of the individual changes; that is

$$\Delta G = \sum_i \Delta G(\lambda_i \rightarrow \lambda_{i+1}) \quad (5)$$

with the λ_i spanning the range from 0 to 1. Thermodynamic integration (Kirkwood, 1968) was also employed to obtain ΔG from the integral

$$\Delta G = \int_0^1 \langle \Delta U \rangle_\lambda \, d\lambda \approx \sum_i \langle \Delta U \rangle_{\lambda_i} \Delta_{\lambda_i} \quad (6)$$

with $\Delta U = U_1 - U_0$.

Both eqs 4 and 6 are exact, in principle, since no approximations are introduced in their derivation (Brooks et al., 1988; Beveridge & DiCapua, 1989). The approximate nature of the calculated results is due to the use of an empirical potential function, $U(\lambda)$, to model the system, the finite length molecular dynamics simulations to evaluate the ensemble averages, and the quadrature introduced in eq 6.[1]

Molecular dynamics simulations were carried out with the stochastic boundary method (Brooks & Karplus, 1989), using a version of the program CHARMM (Brooks et al., 1983) with standard parameters for the polar hydrogen protein model (param19) and the CHARMM-adapted TIP3P water model (Jorgensen et al., 1983). The CHARMM potential consists of a sum of bond length, bond angle, improper torsion, dihedral angle, van der Waals, and electrostatic terms. Figure 1 shows the arginine and protonated histidine side chains with their partial atomic charges. Simulations were run by using the total Hamiltonian $H(\lambda)$ for the system (i.e., the total kinetic energy

[1] The experiments were done at constant pressure and thus measure ΔG, whereas the calculations used a stochastic boundary and yield a value intermediate between ΔG and ΔA. For the mutation under discussion, the difference between ΔG and ΔA should be small.

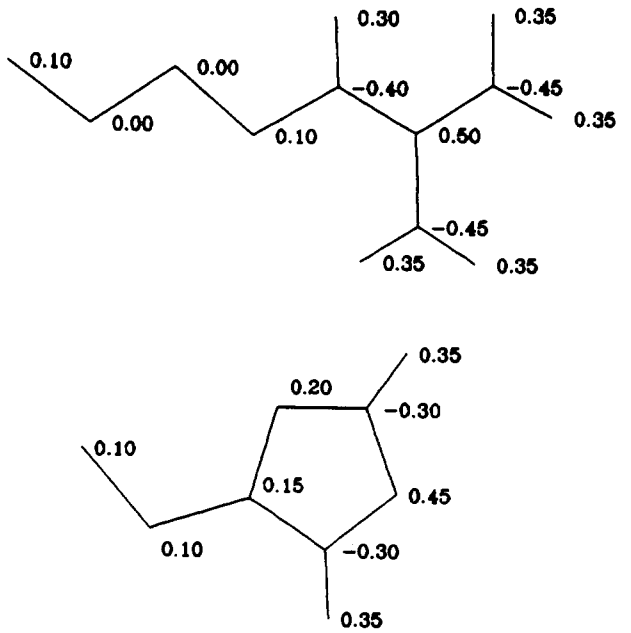


FIGURE 1: Partial atomic charges for the arginine and the charged histidine side chain.

and hybrid potential energy of the hybrid system are included), whereas only the potential energy portion, $U(\lambda)$, appears in the calculation of the free energy change (eqs 4 and 6). This is appropriate because the kinetic energy contributes an additive term to the free energy, which cancels for chemically allowed processes. Because the reactions $T4_{fol}^{wt} \rightarrow T4_{fol}^{R96H}$ and $T4_{unf}^{wt} \rightarrow T4_{unf}^{R96H}$ are not balanced (certain atoms are created and others destroyed in the alchemical simulations), physically meaningful results are obtained only when $\Delta\Delta G$ values, such as those in eq 2, are calculated. It is necessary to include the internal energy of the perturbed residue in $H(\lambda)$ and $U(\lambda)$ in the calculation of the free energy change. Although the internal energy difference may be unimportant in certain cases (Gao et al., 1989), it is significant if the mutation introduces (or removes) strain in one of the final states; in the present case, a portion of the free energy change was proposed to arise from strain associated with the histidine of the mutant protein in the folded state (Weaver et al., 1989). Such contributions have been neglected in some applications of the free energy simulation method (Jorgensen & Ravimohan, 1985; Singh et al., 1987).

Each state was simulated at 300 K by the introduction of a reduced system composed of part of the protein plus the surrounding solvent inside an 11-Å spherical boundary centered on the initial position of C_β of Arg 96. For the folded state, the crystal coordinates of the wild-type structure (Weaver & Matthews, 1987) were used to initiate the simulation; Figure 2a shows the simulation system with the mutant residue at position 96 and Figure 2b shows the same system without solvent, for clarity. There is no unique choice for the unfolded state (Wong & McCammon, 1986; Bash et al., 1987a; Dang et al., 1989). In the current work, the seven amino acid segment centered about residue 96 in the primary structure was initially represented as the extended conformation ($\psi = \phi = \omega = 180°$ with the χ angles set equal to 180° except His χ_2, which was set equal to −90°); boundary harmonic constraints were applied to the terminal residues, 93 and 99. Figure 2c, shows the extended structure used in the simulation, with and without solvent. An extended chain is one choice that exposes the charged residues to solvent in the denatured state. The choice of seven residues was based on the concept that the same sphere size should be used for the two types of simulations to obtain a cancellation of errors for the cutoffs used in the potential. Three residues on each side of the mutated residue were expected to include all the important interactions; this is borne out by the calculations. Each system was solvated by overlaying a reference set of water molecules from a simulation equilibrated at 1 atm pressure and 300 K (Jorgensen et al., 1983). Although there are four buried water molecules in T4 lysozyme, none is within the 11-Å simulation region. There were 64 (154) water molecules and 371 (79) protein atoms in the simulations of the folded (unfolded) state. Because His 96 in the mutant titrates at pH 6.8 (Becktel & Schellman, 1987; Weaver et al., 1989) and thermodynamic, as well as structural, data are available in solutions more acidic than this (Hawkes et al., 1984; Becktel & Schellman, 1987), a protonated His was used. The choice of ionization state for the Asp, Glu, Lys, Arg, and His corresponds to the pH range between 5 and 6. His 96 was built into the structure of the wild type (Weaver & Matthews, 1987) with the imidazolium ring in an orientation that approximates the one illustrated in Grütter et al. (1979). For comparison, two calculations were done for the folded state with the initial His 96 position differing by a 180° rotation about the C_β–C_γ bond (R96H and R96Hf). The one with $\chi_2 = -168°$ in the average structure is closest to the X-ray value of 157°, and the other ($\chi_2 = 33°$) is referred to as the "flipped" conformation; the same number of protein atoms and water molecules was used for the two conformations. The two folded systems and the one unfolded system were simulated at nine values of λ ($\lambda = 0.1, 0.2, ..., 0.9$). The $\lambda = 0.1$ simulations were begun from a 10-ps equilibration of the starting structures for each of the three simulations. A standard equilibration protocol was used; i.e., atomic velocities were selected at random from a Gaussian distribution at 300 K and equilibrated at 300 K. All subsequent simulations were carried out after a 5-ps equilibration started from the final step of the simulation for the previous value of λ. Nonbond interactions were smoothed to zero beyond 7.5 Å (switching for van der Waals and shift for electrostatics) (Brooks et al., 1983) and a constant dielectric ($\epsilon = 1$) was used. All bonds involving hydrogen atoms were constrained with the SHAKE algorithm (Ryckaert et al., 1977) so that a 1-fs timestep could be used for integrating the equation of motion for both equilibration and data collection in the simulations. The ensemble averages were determined from simulations of 10 ps with coordinates saved every 0.02 ps. Fifteen picoseconds of simulation required about 1 h of CPU time on a Cyber 205.

To decompose the calculated free energy changes into contributions from interactions within the mutant residue and from those with the environment, we make use of thermodynamic integration (eq 6). Because of the form of the potential energy function and the fact that $\langle \Delta U \rangle_\lambda$ appears linear in the integral (eq 6), ΔG can be decomposed into individual additive contributions; methods based on the exponential expression for the free energy (eq 4) lack this desirable feature. For the analysis we separated "self"-terms in ΔU from those that include interactions with other residues or with solvent. The self-terms are defined as any bond, bond angle, dihedral, improper, or nonbond term that involves *only* atoms in the mutant or wild-type side chain at position 96 (for this purpose C_α of residue 96 is considered part of the side chain). All other terms in ΔU represent "interactions" between the side chain 96 and the rest of the system; e.g., the bond angle term for $NC_\alpha C_\beta$ of residue 96 is included in the interaction contribution. The self-terms are of interest because they correspond to strain arising from distortion of the mutated or wild-type side chain.

Table I: Experimental and Theoretical Free Energy Changes (kcal/mol)

	$\Delta\Delta G$ for R96H	ΔG_{flip}[a] for R96H
experiment		
320 K, pH 3[b]	−3.3	
339 K, pH 6[c]	−2.82	
300 K, pH 3[d]	−3.2 ± 1.2	
calculation		
300 K (eq 5)	−1.6	4.9
300 K (eq 6)	−1.9	5.0

[a] This represents the free energy difference in the folded form between the observed and flipped structure for the mutant R96H (see text). [b] See Hawkes et al. (1984). [c] See Bectel and Schellman (1987). [d] Extrapolated from the data of Kitamura and Sturtevant (1989).

We make use of the distributive properties of the simulation averages and the integration in eq 6 and obtain

$$\Delta G = \int_0^1 \langle \Delta U_{self} + \Delta U_{int} \rangle_\lambda \, d\lambda = \int_0^1 \langle \Delta U_{self} \rangle_\lambda \, d\lambda + \int_0^1 \langle \Delta U_{int} \rangle_\lambda = \Delta G_{self} + \Delta G_{int} \quad (7)$$

where the self-term, ΔG_{self}, and the interaction term, ΔG_{int}, are, in general, a sum of covalent, electrostatic, and van der Waals contributions. In the empirical energy function used here, both hydrogen-bond and charge–charge interactions are included in the electrostatic term. The interaction free energy, ΔG_{int}, can be divided into contributions due to individual residues, as well as portions of residues (e.g., carbonyls, side chains), or otherwise separated into terms arising from different parts of the system (e.g., protein versus solvent) in such a way that the total free energy change is equal to the sum of its components.

To facilitate a structural analysis consistent with the free energy simulation, the average structures and other properties for the wild-type and mutant system (corresponding to the end state $\lambda = 0.0$ and $\lambda = 1.0$, respectively) were calculated with the perturbation method from the nearest states that were simulated ($\lambda = 0.1$ and $\lambda = 0.9$, respectively) by use of the exponential formula (Brooks, 1986)

$$\langle Q \rangle_{\lambda'} = \frac{\langle Q e^{-\beta[U(\lambda')-U(\lambda)]} \rangle_\lambda}{\langle e^{-\beta[U(\lambda')-U(\lambda)]} \rangle_\lambda} \quad (8)$$

where Q is a structural, dynamic, or energetic variable of interest and (λ, λ') are (0.1, 0.0) or (0.9, 1.0). To improve the quadrature in eq 6, total free energy and component free energy values were calculated by using eq 8 to evaluate $\langle \Delta U \rangle_{\lambda \pm \delta\lambda}$ for two values of $\delta\lambda$ (0.025 and 0.050) at each λ.

II. Results

Table I shows some experimental results and the values obtained from the free energy simulation by exponential averaging and thermodynamic integration. The two calculated values (−1.6 and −1.9 kcal/mol at acid pH and 300 K) agree within the estimated statistical error of the simulation. The statistical error in the simulations is on the order of a few tenths of a kilocalorie per mole; however, the systematic error is likely to be larger but is very difficult to estimate. The calculated results are of the observed sign (wild type more stable than mutant) and have a magnitude close to the experimental result [−3.2 ± 1.2 kcal/mol at pH 3 and 300 K; extrapolated from the data of Kitamura and Sturtevant (1989)]. What is important is that the free energy calculations appear to be sufficiently accurate to be used for obtaining information about the difference in stability of the protein conformations under consideration and to permit analysis of the origin of these differences.

The thermodynamic integration component decomposition method was used to evaluate the contribution of particular structural groups (individual side chains, backbone dipoles, and solvent water). The results are given in Table II. The self and interaction contributions defined in eq 7 have been computed and are listed. Each of these is divided into covalent, electrostatic, and van der Waals terms in accord with the form of the CHARMM potential function. It is apparent that the overall change in stability is due to the partial cancellation of many terms; i.e., the calculated value of −1.9 kcal/mol is composed of individual terms as large as 7 kcal/mol. Some

Table II: Contributions to the Free Energy Differences between Wild Type and R96H Mutant (kcal/mol)

	ΔG_{mut}^{fol}				ΔG_{mut}^{unf}				$\Delta\Delta G$			
contribution	total	cov	elec	vdW	total	cov	elec	vdW	total	cov	elec	vdW
total	−6.72	−6.99	1.61	−1.34	−8.65	−7.55	1.54	−2.64	−1.93	−0.56	−0.07	−1.30
self	−18.10	−7.12	−10.19	0.79	−18.15	−6.40	−10.35	−1.40	−0.06	0.72	−0.16	−0.62
interaction	11.37	0.13	11.79	−0.55	9.50	−1.15	11.88	−1.23	−1.86	−1.28	0.09	−0.68
Interaction Components[a]												
solvent	12.05		12.49	−0.44	14.40		14.28	0.12	2.36		1.79	0.56
protein	−0.68	0.13[b]	−0.70	−0.11	−4.90	−1.15[b]	−2.40	−1.35	−4.22	−1.28[b]	−1.70	−1.24
carbonyl groups												
Tyr 88	−2.42		−4.00	1.58					2.42		4.00	−1.58
Asp 89	3.54		3.38	0.16					−3.54		−3.38	−0.16
Leu 91	1.77		1.15	0.62					−1.77		−1.15	−0.62
Asp 92	−5.20		−4.47	−0.73					5.20		4.47	0.73
Ala 93	−1.93		−2.07	0.15	0.00		0.00	0.00	1.93		2.07	−0.15
Arg 95	0.48		0.88	−0.40	−2.91		−2.08	−0.83	−3.39		−2.96	−0.43
Res 96	−0.89		−0.79	−0.09	−2.19		−2.13	−0.06	−1.30		−1.34	0.04
Cys 97	0.66		0.66	0.00	0.74		0.88	−0.13	0.08		0.22	−0.13
C_αNH groups												
Ala 93	−1.37		−1.12	−0.25	0.00		0.00	0.00	1.37		1.19	0.25
Arg 95	−0.15		−0.15	0.00	0.39		0.33	0.05	0.54		0.48	0.05
Res 96	1.23		1.73	−0.50	0.41		0.54	−0.13	−0.81		−1.19	0.38
Cys 97	1.42		1.19	0.23	0.02		0.29	−0.27	−1.40		−0.90	−0.50
side chains												
Lys 85	−3.25		−3.26	0.02					3.25		3.26	−0.02
Tyr 88	−1.34		0.11	−1.45					1.34		−0.11	1.45
Asp 89	6.84		6.53	0.31					−6.84		−6.53	−0.31

[a] Only groups that contribute more than 0.5 kcal, in total, are included. [b] The covalent contribution is not broken down because it has three and four body terms.

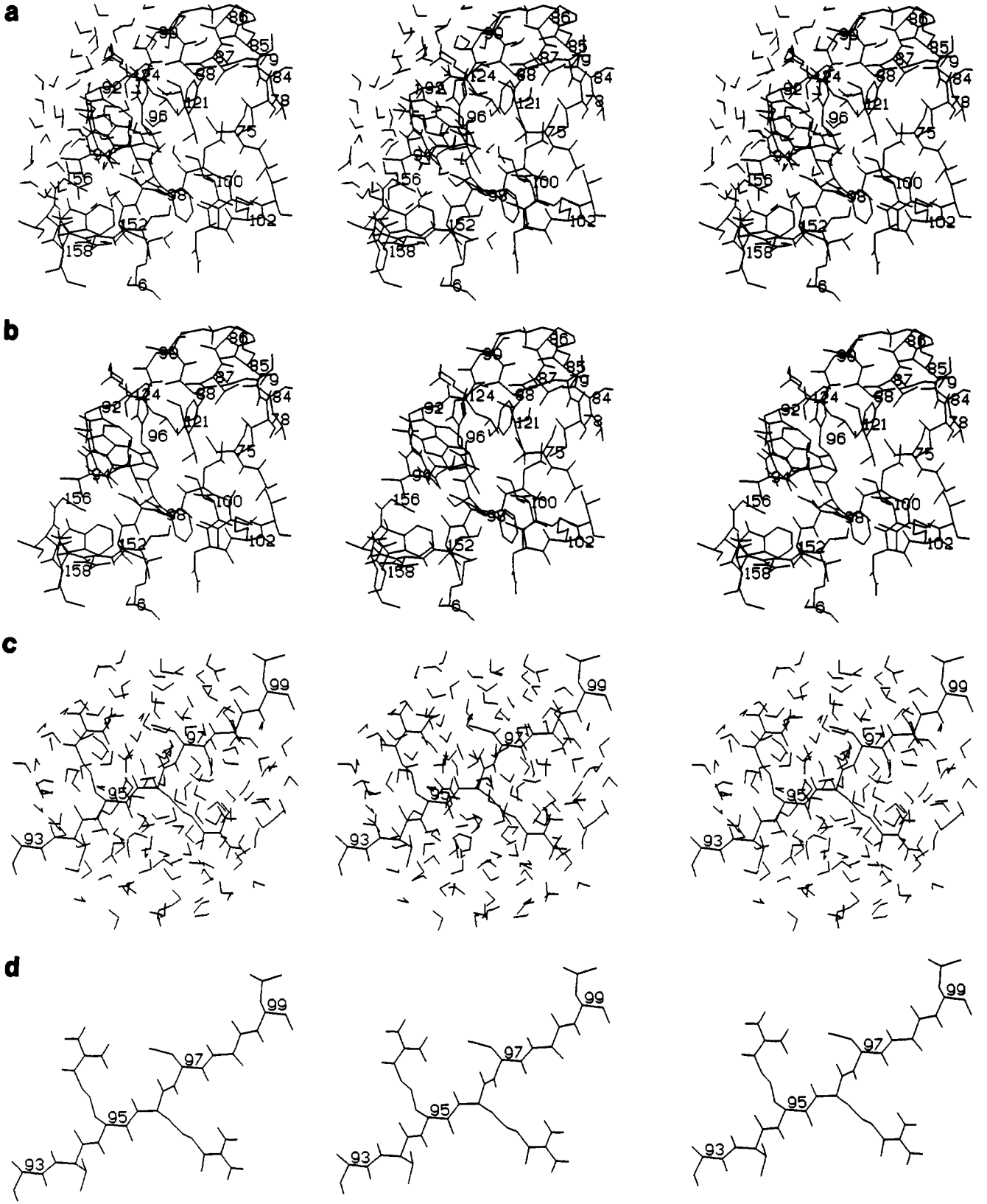


FIGURE 2: (a) Average simulation structure for the folded-state mutant system. (b) Average simulation structure for the folded-state mutant system, shown without solvent for clarity. (c) Average simulation structure for the unfolded-state wild-type system. (d) Average simulation structure for the unfolded-state wild-type system, shown without solvent for clarity. In each part, the three images are left-eye, right-eye, and left-eye views for stereoimaging.

of these stabilize the wild type and others the mutant, but the net effect ($\Delta\Delta G$) is to destabilize the mutant, relative to the wild type. Although the largest individual contributions come from electrostatic interactions, these virtually cancel in $\Delta\Delta G$. The large self-terms in ΔG_{mut}^{fol} and ΔG_{mut}^{unf} strongly favor the mutant in the folded and unfolded form. These contributions arise primarily from the nonconservation of particles, which cancels exactly in $\Delta\Delta G$. Thus, the dominant *net* effects on $\Delta\Delta G$ result from covalent interaction terms (−1.3 kcal/mol) and van der Waals interactions (−0.7 kcal/mol). As discussed below, the former is due mainly to a dihedral angle contribution and the latter to numerous partly cancelling interactions

with both main-chain and side-chain atoms.

Analysis of the covalent interactions shows that the $\Delta\Delta G$ contribution arises mainly from dihedral terms. The Arg side chain packs into the wild-type structure in an unstrained rotamer conformation. Both the crystal structure and the average simulation structure are in agreement and place χ_1 near gauche minus (−81°, X-ray; −64°, simulation) and χ_2 at trans (179°, X-ray; 175°, simulation). This is the (−, t) conformation, which Ponder and Richards (1987) found as the most common rotameric state (average $\chi_1 = -67.6°$, standard deviation = 13.3°; average $\chi_2 = 176.9°$, standard deviation = 20.0°) adopted by arginine side chains in a survey of X-ray crystal structures. In the unfolded state the Arg 96 side chain samples both the (−, t) and the (t, t) conformational minima, which leads to the average values $\chi_1 = -129°$, $\chi_2 = -177°$. Histidine residues show a tendency for χ_1 to be at ±60°, −180° and χ_2 to be near ±90° (Ponder & Richards, 1987). Since there is neither room nor flexibility in the structure for His 96 to adopt this conformation, both the X-ray structure and simulations have unusual dihedral angles; χ_1 is close to −90° (−88°, X-ray; −94°, simulation; −90° for simulation in the flipped orientation) and χ_2 causes the ring to be nearly eclipsed with C_α (158°, X-ray; −169°, simulation; 33°, flipped simulation). This strained conformation is unfavorable and simulations show that, in the unfolded state, a conformation near the ideal values (t, −90) is adopted (−170°, −104°). Thus, arguing from the average structures, there would be a destabilization of the mutant by 1.9 kcal/mol in the folded state. The corresponding free energy component is about 0.8 kcal/mol (0.7 folded, −0.1 unfolded). The total dihedral covalent contribution to the interaction free energy change is 0.83 kcal/mol in the folded state and −0.23 kcal/mol in the unfolded; this corresponds to a contribution to $\Delta\Delta G$ of 1.06 kcal/mol.

From Table II the solvent contribution *destabilizes* the mutant in both the folded and unfolded protein. Since the solvent interaction is more destabilizing in the unfolded form (14.4 versus 12.1 kcal), it leads to a net *stabilization* of the folded structure of the *mutant* protein, relative to the wild type. It is the strong hydrogen bonds between the charged amino acid side chain and water that dominate the interaction with the solvent. The solubility of charged atoms in water increases roughly as the square of the charge (Roux et al., 1990). Because arginine is represented by more point charges with greater magnitudes than the protonated histidine (see Figure 1), its solvation free energy is larger. Further, the Arg side chain is stabilized more in the unfolded than in the folded state because of its increased exposure to solvent; from the Lee–Richards algorithm (Lee & Richards, 1971) with a probe size of 1.4 Å, the exposure of the guanidinium group in the folded state is 36% of that in the unfolded state. Comparing Arg 96 and His 96, there are 8 (12) and 6 (8) waters, respectively, interacting strongly in the folded (unfolded) state. In a simulation of the mutation Asp 99β → Ala in oxyhemoglobin (Gao et al., 1989), the solvent contribution to ΔG was found to be on the order of 70 kcal/mol, near the full solvation free energy of a carboxylate (Pearson, 1986). For the R96H mutation, the charge is not neutralized but only redistributed.

There are many contributions to $\Delta\Delta G$ from protein interaction terms. In contrast to the solvent terms, the protein contributions *stabilize* the mutant in both the folded and unfolded form. Since the stabilization is greater in the unfolded than the folded form, the net contribution to $\Delta\Delta G$ is to stabilize the wild type. The stabilization is sufficiently large (−4.2 kcal/mol) to overcome the solvent contribution (2.36 kcal/mol), which is in the opposite direction (see above). Both the guanidinium group of Arg 96 and the imidazolium group of His 96 interact strongly with polar groups of the protein (Figure 3a,b). Most important are several backbone carbonyls that point toward the center of positive charge density of the guanidium group and make hydrogen bonds to its protons. Some of these carbonyls (88, 89, and 90) are at the C-terminus of an α-helix that consists of residues 82–90. These interactions might, therefore, be regarded as a local manifestation of the helix–dipole effect (Hol et al., 1978). Replacement of the arginine by a histidine, which is also positively charged, results in some small changes in the positions and orientations of the carbonyls in the folded protein (the calculated Cartesian coordinate differences are 1.0 Å or less but the local coordinate shifts are much smaller). These adjustments promote favorable electrostatic interactions between the imidazolium ring and the carbonyls. Overall, as shown in Table II, these carbonyls, which do "solvate" the arginine, are more effective in stabilizing the histidine mutant. The combined effect of carbonyls 88, 89, 91, and 92 is to favor histidine by 2.3 kcal/mol in the folded state; of particular importance for stabilizing the mutant are the carbonyls of Tyr 88 and Asp 92. All of the carbonyls involved are in residues that are not included in the model of the unfolded state because they are separated by more than three amino acids. They are presumably sufficiently far away from Arg (or His) and appropriately solvated so as not to contribute to the free energy difference; the effect of residues 93 and 99, which are included, is negligible. However, the backbone interactions of residues 95 and 97, as well as 96, stabilize the wild type. Both the dipolar carbonyl and amide groups are involved, with the dominant contribution from the carbonyls coming from the unfolded state. Taken together, the main chain of the mutated residue and its two neighbors contribute −6.3 kcal/mol (noncovalent) and −1.28 kcal/mol (covalent) to the $\Delta\Delta G$ stabilizing the wild type.

The side chains of Lys 85, Tyr 88, and Asp 89 also contribute to the relative stability of the wild-type and mutant enzymes (Figure 3a,b). The two charged side chains are not in direct contact with residues 96 and exert their effect in the folded state primarily through longer range electrostatic interactions; they are not included in the unfolded state model. Tyr 88 packs against Arg/His 96 and makes a stability contribution to the mutant through van der Waals interactions. Use of the end-point (crystal) structures to evaluate the interaction energy difference between the mutant and wild type yields −0.1 kcal/mol (0.0 kcal/mol) for Lys 85 and 3.1 kcal/mol (4.7 kcal/mol) for Asp 89; the average interaction energies are similar. The difference between the interaction energies and the free energy simulation results (−3.3 and 6.8 kcal/mol, respectively) shows the importance of the latter for a meaningful analysis (see also section III).

Flipped Orientation. We now consider the interactions that stabilize the observed orientation of His 96 rather than the flipped orientation (see Table III). The 180° rotation about χ_2 causes the center of positive charge to move somewhat and changes the locations and orientations of the two hydrogen-bond-donating imidazolium NH groups (see Figure 3b,c). These changes in the electrostatic field at residue 96 result in large electrostatic contributions to the free energy. As was the case for the unfolding equilibrium, the partial cancellation of many terms, some larger than 4 kcal/mol, is responsible for a total free energy change of the same magnitude (5 kcal/mol). However, in this case the overall effect is mainly electrostatic, although there is a self-covalent contribution due

Table III: Contributions to the Free Energy Difference for Flipping His 96 in R96H Mutant of T4 Lysozyme (kcal/mol)

contribution	total	cov	elec	vdW
total	5.06	0.83	4.75	−0.53
self	2.00	1.72	0.55	−0.27
interaction	3.06	−0.89	4.20	−0.26
	Interaction Components			
solvent	−0.40		−0.25	−0.15
protein	3.46	−0.89[a]	4.45	−0.11
carbonyl groups				
Tyr 88	−0.95		−0.33	−0.62
Asp 89	4.16		4.19	−0.03
Leu 91	4.18		4.59	−0.04
Asp 92	−2.37		−2.82	0.45
Ala 93	−0.35		−0.15	−0.20
Arg 95	0.53		0.46	0.06
Res 96	0.01		0.06	−0.05
Cys 97	−0.02		−0.03	0.00
C_αNH groups				
Ala 93	0.46		−0.15	0.61
Arg 95	0.01		0.02	−0.01
Res 96	−0.12		−0.21	0.08
Cys 97	−0.28		−0.08	−0.20
side chains				
Lys 85	2.77		2.78	−0.02
Tyr 88	0.02		0.03	−0.02
Asp 89	−2.80		−2.86	0.05
Asp 92	−0.08		−0.10	0.03
Arg 95	0.02		0.02	0.00
Cys 97	0.13		0.12	0.01
Lys 124	−1.79		−1.79	0.01

[a] See footnote *b* of Table II.

to strain in the flipped conformation. The favorable electrostatic interactions between the histidine side chain and its environment in the observed conformation are due primarily to the series of carbonyl groups considered above. Most of the carbonyl groups shift position to stabilize the flipped orientation and the number and types of interactions remain virtually unchanged. The solvent reorganizes to stabilize the flipped histidine orientation with a free energy change of only 0.40 kcal/mol (favoring the flipped orientation), the packing against the side chain of Tyr 88 changes not at all, and the hydrogen bond to the carbonyl of Tyr 88 changes by only 0.9 kcal/mol, in favor of the flipped form.

Some of the contributions that change significantly between the two orientations involve both a favorable interaction energy and a "strain" energy resulting from the structural perturbation required for the interaction. An example is the carbonyl of Leu 91, which stabilizes the observed conformation by 4.2 kcal/mol. Figure 3b,c shows that a fairly distant (3.5 Å) interaction in the observed state with the hydrogen of N_{ϵ_2} is replaced by a good hydrogen bond (2.2 Å) in the flipped state to the hydrogen of N_δ. This would suggest that the carbonyl should favor the flipped conformation. However, the carbonyl trades solvation by water in the observed conformation for the hydrogen bond in the flipped state. Although the new hydrogen bond energy is larger than that form the 3.5-Å interaction, it is smaller than combined contributions from this interaction and the solvation energy of the carbonyl.

III. Discussion

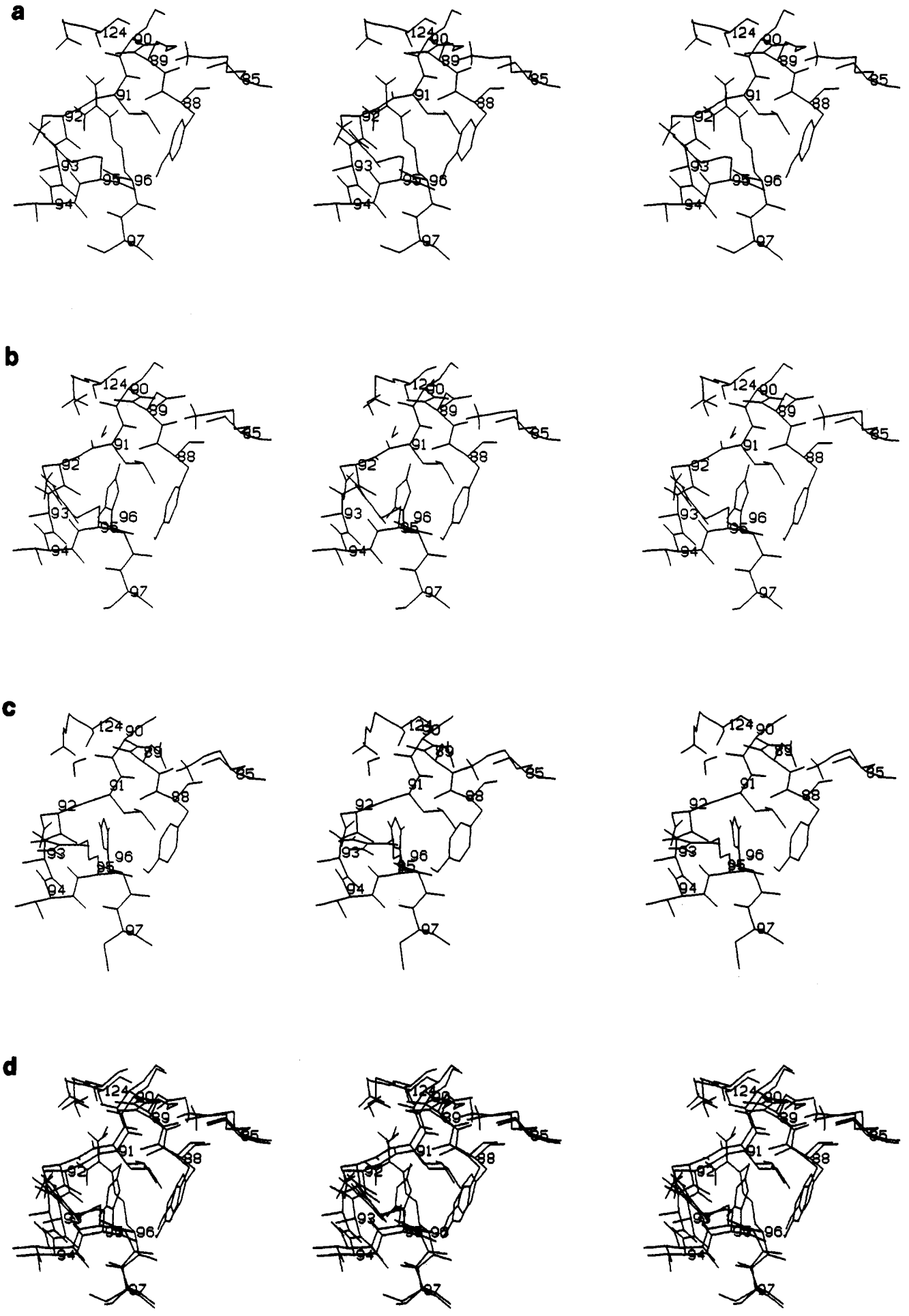
An important result of the simulation is the finding that many groups interact significantly with residue 96 to produce the overall rather small stability change induced by the mutation. This is in accord with the simulation of a cooperativity mutant in hemoglobin (Gao et al., 1989) but differs from the usual experimental interpretations, which tend to focus on a few terms (Alber, 1989). In the present case, some of the interactions *destabilize* the mutant, as one would expect, but a number actually *stabilize* the mutant. Also, effects in the unfolded, as well as the folded, protein are significant. Only by considering all of these contributions is a meaningful free energy change obtained. Moreover, while the stability change is calculated to be about 2 kcal/mol, many of the contributions are of the same magnitude or larger. Given that there is a partial cancellation of these terms, it is not surprising that the theoretical calculations have difficulty approaching the accuracy of the experiments. However, in spite of limitations inherent in such free energy simulations, they can isolate the contributing factors, which are not directly accessible from the experimental results (Alber, 1989).

It should be noted that, in addition to the approximations inherent in the use of an empirical potential function, there are other aspects of the simulation model (extended unfolded state, 11-Å cutoff) that could introduce errors. Some of these could be tested by doing comparison calculations (e.g., a larger sphere and cutoff, other geometries of the unfolded state). Examination of the contributions to the free energy change from more distant groups suggests that the sphere size and cutoff distance used in the model are adequate.

Given the above, it is important to compare the calculated results with the careful analysis of Weaver et al. (1989) and Kitamura and Sturtevant (1989), who used structural and thermodynamic data to analyze the origin of the stability change in R96H; these recent studies go beyond that published in the original characterization of the mutant (Grütter et al., 1979). Since the calculations had been completed before the high-resolution crystal structure was available to us, we reexamined our results to be able to compare them with the experimental analysis. Weaver et al. assumed that all of the observed free energy difference arises in the folded protein for which structural data are available. The simulation results indicate that the unfolded chain also plays a role; e.g., from Table II, as already discussed, there are significant interaction contributions to the difference between the mutant and wild type from the unfolded protein. These come not only from solvent and nonconservation of particles; the interactions with the protein result in a 4.9 kcal/mol [total interaction (9.5 kcal/mol) minus solvent contribution (14.4 kcal/mol)] preference for histidine in the unfolded state. This effect comes mainly through interactions between residue 96 and the local backbone. Neighboring side chains make small free energy contributions (less than 1 kcal/mol). This indicates that, even if the simulations are significantly in error (e.g., due to the simple model of the unfolded state), examination of the folded state alone is not sufficient.

For the folded state, Weaver et al. (1989) suggest two dominant contributions to the free energy change. Bond angle strain (particularly the angle $C–C_\alpha–C_\beta$ of His 96), was estimated to yield 2.0 kcal/mol and the less effective stabilization of His 96 than Arg 96 by interaction with the dipole of helix 82–90 was stated to contribute most of the rest, though no quantitative estimate was given. In addition, they consider a number of other possible contributions but conclude that none is important.

There is some indication that His 96 is distorted in the simulation of the folded state, though the simulation values for the bond angles discussed by Weaver et al. [see Table 2 of Weaver et al. (1989)] differ from the CHARMM optima by less than 4°; in particular, the average values of $N–C_\alpha–C$, $N–C_\alpha–C_\beta$, $C–C_\alpha–C_\beta$, and $C_\alpha–C_\beta–C_\gamma$ in His 96 are 1° smaller, 3.5° larger, 1.8° smaller, and 3.4° larger than the optima, while for Arg 96 the deviations are considerably less except for $N–C_\alpha–C_\beta$ (2.1° smaller). The strain energy associated with


a
b
c
d
124
90
89
85
91
88
92
93
95
96
94
97

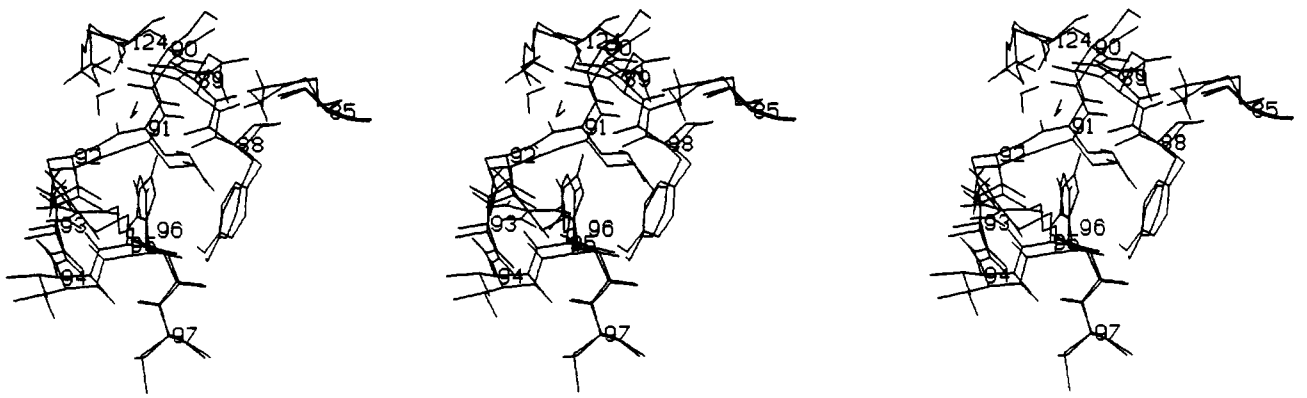


FIGURE 3: (a) Detail of the average folded-state wild-type structure is shown. Arg 96 is hydrogen-bonded to the backbone carbonyls of Tyr 88 and Asp 89 (at the C-terminus of an α-helix) and packs against the side chain of Tyr 88. (b) Detail of the average folded-state R96H mutant structure is shown. His 96 is hydrogen bonded to the backbone carbonyl of Tyr 88 and packs against its side chain but does not reach to the backbone carbonyl of Asp 89 (at the C-terminus of an α-helix). A bound water sits where the end of Arg 96 would be and accepts a hydrogen bond from His 96 $N_\epsilon H$ (see part d) and it and a second bound water bridge the backbone carbonyls of Asp 89 and Leu 91. (c) Detail of the average folded-state R96H mutant structure in the flipped orientation is shown. His 96 is hydrogen-bonded to the backbone carbonyl of Tyr 88 and packs against its side chain but does not reach to the backbone carbonyl of Asp 89 (at the C-terminus of an α-helix). A bound water bridges the backbone carbonyls of Asp 89 and Leu 91. (d) Superposition of parts a and b shows differences between wild-type and R96H mutant T4 lysozyme from simulations. One bound water in the mutant structure lies in the volume vacated by Arg 96 and the backbone carbonyls of Asp 89 and Leu 91 have adjusted somewhat to accommodate bridging hydrogen bonds from another bound water. (e) Superposition of parts b and c shows differences between the observed and flipped orientations of R96H mutant T4 lysozyme from simulations.

such static distortions would be 0.6 kcal/mol for His and 0.1 kcal/mol for Arg 96. For the corresponding X-ray values we find distortion energies equal to 2.2 kcal/mol for His and 0.1 kcal/mol for Arg with the CHARMM force field. The largest X-ray deviation from optimum values (the observed His 96 $C-C_\alpha-C_\beta$ angle of 99.9°) seems unlikely; the standard parameters in refinement programs range from 109.1° to 111°. Also, in spite of the deviations in the average structures, the free energy simulations show no significant bond angle contribution. The calculated value is 0.46 kcal/mol, in part due to the fact that there is strain in the unfolded structure. It is possible that the deviation between the X-ray and simulation structure arises from (a) errors in the calculation, (b) errors in the refinement, or (c) a real difference between the two due to crystal contacts; e.g., Arg 96 is involved in an intermolecular salt bridge but His 96 apparently is not (Weaver & Matthews, 1987; Weaver et al., 1989).

Interactions with the residues in the α-helix are important, as described above, in both the structural interpretation and the simulation results. However, a helix–dipole model is not appropriate because the guanidinium and imidazolium charged groups are close to the end of the helix. The validity of a dipole model requires that the distance of the charges from the center of the dipole be larger relative to the length of the dipole (Karplus & Porter, 1970). Consequently, it is better to focus on individual interactions with the C=O and $C_\alpha NH$ groups of the helical residues. With the parameter set used in the model, the backbone carbonyl is electrostatically neutral, whereas the backbone NH is not; therefore, the analysis was done in terms of carbonyl and $C_\alpha NH$ groups, which are both neutral. All the C=O and $C_\alpha NH$ groups of residues 82–90 are too far from the mutant site in the model of the unfolded state to contribute so that an analysis of the folded structure alone, in accord with that of Weaver et al., is appropriate here. Even in the folded structure, only the carbonyls of Tyr 88 and Asp 89 in the helix make a significant contribution in the simulation (though residues 84–90 were included). Both of these make good hydrogen bonds to Arg 96 (see Figure 3a,b). Weaver et al. argue that for His 96 the latter hydrogen bond is absent and the former is significantly weaker. We find, in agreement with their analysis, that the C=O of Asp 89 is important for stabilizing the wild type (3.5 kcal/mol), but that the C=O of Tyr 88 favors the mutant by 2.4 kcal/mol. Table IV shows the X-ray and simulation results for the essential geometric parameters for the Tyr 88 carbonyl hydrogen bond in both the mutant and wild type. The wild-type and mutant values for the parameters and their change on mutation are in excellent agreement; the only difference is in the wild-type dihedral angle C=O ··· H—N, which is not an energy-sensitive hydrogen-bonding parameter (Reiher, 1985). Weaver et al. argue than in spite of the short carbonyl oxygen to N_ϵ distance [2.73 Å in the mutant crystal, an unusually short value (Thanki et al., 1988) versus 2.80 Å in the simulation] the hydrogen bond is weak because the $N_{\epsilon_2}—H\cdots O$ angle (assuming a standard H atom position) is 129°; it is even less (126.5°) in the simulation. Although, as pointed out by Weaver et al., the N—H···O angle in hydrogen bonds is rarely less than 140°, ab initio calculations (Reiher, 1985) indicate that it is only for angles less than 130° that the hydrogen bond energy begins to decrease significantly. Moreover, the difference of 2.4 kcal/mol arises from the entire side chain–carbonyl interaction rather than the individual hydrogen bond and is a result of configurational averaging. The interaction energy calculated from the end-point structures is slightly better (0.7 kcal/mol) in the wild type. This illustrates the essential difference between static energies and dynamic free energy.

Table IV: Structural Parameters for Tyr 88 C=O Hydrogen Bonded to the Mutant Site

	X-ray		simulation	
	WT	R96H	WT	R96H
d_1: O···H (Å)	2.08	2.01	1.94	2.12
d_2: O···N (Å)	2.96	2.73	2.79	2.80
a_1: O···H—N (deg)	148.12	129.09[a]	145.63	126.45
a_2: C—O···H (deg)	153.84	138.65	160.41	133.12
a_3: C—O···N (deg)	143.87	136.67	151.10	133.28
dihedral: C—O···H—N (deg)	9.19	−91.90	−42.53	−98.13

[a] Value listed in Weaver et al. (1989) is 131°; the discrepancy results from the different procedures used to generate the hydrogen positions.

Weaver et al. noted that there is bound water present in the mutant that is absent in the wild type; most bound waters are the same in the two structures. They argue, by analogy with other mutant data (Alber et al., 1987a), that this water con-

tributes little to the free energy difference. In the simulation a bound water is found in the same position as the one described in the X-ray structure (see Figure 3b). The water occupies space in the mutant that corresponds to a "deleted" portion of the Arg side chain. It, as well as a second bound water, makes hydrogen bonds with the carbonyls of Asp 89 and Leu 91, but its direct effect on the free energy differences has not been determined.

An early report of the R96H mutant (Grütter et al., 1979) suggests that His 96 destabilizes the hydrophobic core of the folded protein; in the paper of Weaver et al. it is argued that this is not true because of the similarity in position of the buried portions of His and Arg at position 96 (Weaver et al., 1989). The hydrophobic portion of the side chain of Arg 96 forms part of the hydrophobic core of the C-terminal lobe of the molecule, and the charged portion is partly on the surface. Since the histidine side chain has a shorter hydrophobic portion, there is a resulting burial of charge (see Figures 1 and 3a). From the simulation the interactions between solvent and C_α, C_β, C_γ, and C_δ ($C_{\delta 2}$ for His) of residue 96 favored the wild type by 13.3 kcal/mol, in accord with the crystallographic analysis. However, the solvent interaction with the remaining atoms of the side chain contribute 15.7 kcal/mol to the free energy, destabilizing the wild type. This is due almost entirely to electrostatic terms and represents the unfavorable effect of moving Arg 96 from a fully exposed environment in the unfolded state to a partially exposed one in the folded state, relative to the corresponding transfer for histidine.

To further elucidate the difference between a structural analysis and thermodynamic simulation, we consider the backbone carbonyl of Asp 89 in more detail. One might expect that the strength of the interaction of the carbonyl of Asp 89 with Arg 96 and with His 96 can be calculated from end-point structures, which corresponds to the results from a crystallographic analysis. The value obtained is 13.4 kcal/mol, much larger but in the same direction as the 3.5 kcal/mol stabilization of the wild type calculated with the free energy simulations; averaging the interaction energy over the end-point simulations yields essentially the same value (13.0 kcal/mol) as that from the average structures. Examination of the average structures shows them to be very similar, which is consistent with crystallographic results (Grütter et al., 1979; Weaver et al., 1989); the largest differences are in the positions and orientations of solvent molecules. Visual inspection does not show any obvious structural source of the 10 kcal/mol difference between the energy and free energy. However, the difference of the interaction energy between backbone carbonyl 89 and the rest of the system (omitting the side chain of residue 96) is −8.6 kcal/mol from the average structures and −8.9 kcal/mol from the average over the end-point simulations. This essentially accounts for the discrepancy between the free energy and interaction energy results. Nearly all of this comes from electrostatic terms that derive from the difference in interactions with the solvent. From Figure 3a, carbonyl 89 makes a strong hydrogen bond to an H of an NH_2 group of Arg 96, while it is 4.3 Å from His 96. Thus, the direct effect of losing the interaction between side chain 96 and backbone carbonyl 89 due to the mutation is partially compensated by the indirect effect of improved interaction between solvent and carbonyl 89 in the mutant. This is in accord with the analysis of Fersht (1987) of the partial compensation of protein–protein and protein–water hydrogen bonds.

IV. Conclusion

Free energy simulations have been used to investigate the effect of the mutation R96H on the stability of T4 lysozyme. Since the calculated results are in approximate agreement with experiment (calculated, −1.6 to −1.9 kcal/mol; experimental, −3.2 ± 1.2 kcal/mol), a detailed examination of the origin of the stability changes was made; the lack of error estimates for the calculation is not supposed to imply that it is exact but rather that statistical and systematic errors are difficult to estimate (see text). The calculated change in stability is due to a number of fairly large effects (up to 7 kcal/mol), which mainly cancel to produce the overall value of ~2 kcal/mol. Some of the contributions to the free energy differences appear to be difficult to determine from examination of crystal and average dynamics structures. Also, evaluation of changes in the interaction energies from individual structures or from averages over simulations often do not match the calculated free energy contributions. This is likely to be due, in part, to entropy effects, but the simulations also suggest that rather small structural alterations, not obvious from the X-ray coordinates, may lead to significant changes in the free energy. This is particularly important in the present case because the mutation introduces a redistribution of charge. The resulting changes in charge–charge interactions operate over long distances so that numerous small alterations in the protein and the solvent structure can make significant contributions.

A detailed comparison of the interpretations that result from an examination of the high-resolution crystal structure of the wild-type and mutant proteins and from the free energy simulation analysis shows important differences. This is true even though the local structure in the wild-type and mutant simulations is close to the X-ray results. While the crystallographic analysis focuses on a small number of interactions that destabilize the mutant in the folded state, the simulations show significant contributions from many more interactions, some of which actually stabilize the mutant. These include changes in interactions with nearby side chains (Tyr 88) and more distant charged side chains (Lys 85 and Asp 89), backbone carbonyls (Tyr 88, Asp 89, Asp 92, and Arg 95), and solvent water and the covalent linkage to the backbone of residue 96. Moreover, neglect of interactions other than those with the solvent in the unfolded state would have changed the results by 4.9 kcal/mol and led to the conclusion that the mutant is more stable than the wild type.

A possible conceptual decomposition of the free energy change, for each "mutation" (i.e., that in the folded state, the unfolded state, and the folded state going to the flipped orientation) considers four terms: (1) changes in the nonbonded interactions between the mutation and the rest of the protein, (2) changes in the bonded interactions between the mutation and the rest of the protein, (3) changes in the interactions between the mutation and solvent (which involves only nonbonded terms), and (4) changes in the interactions *within* the mutation site (self contributions that include both bonded and nonbonded interactions). For a stability mutation, it would be convenient to assume that the covalent interaction term, (2), and the self-term, (4), have the same value for the folded and the unfolded state and so do not contribute to the double free energy difference, $\Delta\Delta G$. Moreover, if the unfolded state nonbonded protein interaction term, (1), can be assumed to be negligible, the only remaining unfolded state contribution is due to solvent, (3), and so could be approximated by free energy of transfer data [e.g., those of Wolfenden et al. (1981) and Pearson (1986)]. Kellis et al. (1988) have discussed the results of mutations in the hydrophobic core of barnase using such a model. Our results show that, at least for the R96H mutation in T4 lysozyme, such an approach appears not to be valid. Although the self-term, (4), is negligible due to a

cancellation between covalent and nonbonded terms, the covalent interaction term is significant because the histidine adopts an unfavorable rotameric state in the folded protein, and the nonbonded protein interaction terms are not negligible in the unfolded state; in fact, they contribute −3.75 kcal/mol, favoring His relative to Arg in the unfolded state. Since the barnase mutants (Ile 96 → Val and Ile 96 → Ala) considered by Kellis et al. (1988) involve the deletion of methyl groups, their interactions may be significantly different. A simulation of these mutations has been performed and is being analyzed (M. Prêvost, S. Wodak, B. Tidor, and M. Karplus, unpublished results).

The present simulations and their comparisons with structural and thermodynamic data have demonstrated that the experimental analysis is limited by the difficulty of isolating the contributing factors; theoretical calculations are limited by their inherent errors. Thus, the two approaches are complementary. It is now clear from crystallographic results, in accord with an early perturbation analysis (Shih et al., 1985), that structural changes induced in stable mutants are often very small (Mathews, 1987). Also, consistent with suggestions from early X-ray work (Grütter et al., 1979), it is emerging from theoretical studies that such small structural changes can have large effects on the free energy.

Acknowledgments

We thank B. W. Matthews for supplying the X-ray coordinates of the wild-type and mutant structures and T. Alber, C. L. Brooks III, S. H. Fleischman, J. Gao, S. C. Harrison, K. Kuczera, B. W. Matthews, M. Saper, and D. C. Wiley for stimulating discussions. Most of the calculations were done at the John von Neumann National Supercomputer Center.

Registry No. Arg, 74-79-3; His, 71-00-1; Tyr, 60-18-4; Asp, 56-84-8; lysozyme, 9001-63-2.

References

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

Alber, T., Dao-Pin, S., Wilson, K., Wozniak, J. A., Cook, S. P., & Matthews, B. W. (1987a) *Nature (London) 330*, 41–46.

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

Bash, P. A., Singh, U. C., Langridge, R., & Kollman, P. A. (1987a) *Science (Washington, D.C.) 236*, 564–568.

Bash, P. A., Singh, U. C., Brown, F. K., Langridge, R., & Kollman, P. A. (1987b) *Science (Washington, D.C.) 235*, 574–576.

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

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

Beveridge, D. L., & DiCapua, F. M. (1989) *Annu. Rev. Biophys. Biophys. Chem. 18*, 431–492.

Brooks, B. R., & Karplus, M. (1983) *Proc. Natl. Acad. Sci. U.S.A. 80*, 6571–6575.

Brooks, B. R., Bruccoleri, R. E., Olafson, B. D., States, D. J., Swaminathan, S., & Karplus, M. (1983) *J. Comput. Chem. 4*, 187–217.

Brooks, C. L., III (1986) *J. Phys. Chem. 90*, 6680–6684.

Brooks, C. L., III, & Karplus, M. (1989) *J. Mol. Biol. 208*, 159–181.

Brooks, C. L., III, Karplus, M., & Pettitt, B. M. (1988) *Adv. Chem. Phys. 71*, 1–249.

Chandrasekhar, J., Smith, S. F., & Jorgensen, W. L. (1984) *J. Am. Chem. Soc. 106*, 3049–3050.

Dang, L. X., Merz, K. M., Jr., & Kollman, P. A. (1989) *J. Am. Chem. Soc. 111*, 8505–8508.

Fersht, A. R. (1987) *Trends Biochem. Sci. 12*, 301–304.

Fleischman, S. H., & Brooks, C. L., III (1987) *J. Chem. Phys. 87*, 3029–3037.

Gao, J., Kuczera, K., Tidor, B., & Karplus, M. (1989) *Science (Washington, D.C.) 244*, 1069–1072.

Grütter, M. G., Hawkes, R. B., & Matthews, B. M. (1979) *Nature (London) 277*, 667–669.

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

Hol, W. G. J., van Duijnen, P. T., & Berendsen, H. J. C. (1978) *Nature (London) 273*, 443–446.

Holbrook, S. R., & Kim, S.-H. (1984) *J. Mol. Biol. 173*, 361–388.

Jorgensen, W. L., & Ravimohan, C. (1985) *J. Chem. Phys. 83*, 3050–3054.

Jorgensen, W. L., Chandrasekhar, J., Madura, J. D., Impey, R. W., & Klein, M. L. (1983) *J. Chem. Phys. 79*, 926–935.

Karplus, M., & Porter, R. N. (1970) *Atoms & Molecules: An Introduction for Students of Physical Chemistry*, Benjamin/Cummings, Menlo Park, CA.

Karplus, M., & Kushick, J. N. (1981) *Macromolecules 14*, 325–332.

Kellis, J. T., Jr., Nyberg, K., Sali, D., & Fersht, A. R. (1988) *Nature (London) 333*, 784–786.

Kirkwood, J. G. (1968) in *Theory of Liquids* (Alder, B. J., Ed.) Gordon and Breach, New York.

Kitamura, S., & Sturtevant, J. M. (1989) *Biochemistry 28*, 3788–3792.

Knowles, J. R. (1987) *Science (Washington, D.C.) 236*, 1252–1258.

Kuriyan, J., & Weis, W. I. (1991) *Proc. Natl. Acad. Sci. U.S.A.* (in press).

Lee, B., & Richards, F. M. (1971) *J. Mol. Biol. 55*, 379–400.

Lybrand, T. P., Ghosh, I., & McCammon, J. A. (1985) *J. Am. Chem. Soc. 107*, 7793–7794.

Matouschek, A., Kellis, J. T., Jr., Serrano, L., & Fersht, A. R. (1989) *Nature (London) 340*, 122–126.

Matsumura, M., Becktel, W. J., & Matthews, B. W. (1988) *Nature (London) 334*, 406–410.

Matsumura, M., Becktel, W. J., Levitt, M., & Matthews, B. W. (1989) *Proc. Natl. Acad. Sci. U.S.A. 86*, 6562–6566.

Matthews, B. W. (1987) *Biochemistry 26*, 6885–6888.

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

Nicholson, H., Becktel, W. J., & Matthews, B. W. (1988) *Nature (London) 336*, 651–656.

Oxender, D. L., & Fox, C. F., Eds. (1987) *Protein Engineering*, Alan R. Liss, Inc., New York.

Pearson, R. G. (1986) *J. Am. Chem. Soc. 108*, 6109–6114.

Petsko, G. A., & Ringe, D. (1984) *Annu. Rev. Biophys. Bioeng. 13*, 331–371.

Phillips, Jr., G. N. (1990) *Biophys. J. 57*, 381–383.

Pjura, P. E., Matsumura, M., Wozniak, J. A., & Matthews, B. W. (1990) *Biochemistry 29*, 2592–2598.

Ponder, J. W., & Richards, F. M. (1987) *J. Mol. Biol. 193*, 775–791.

Reiher, W. E., III (1985) Theoretical Studies of Hydrogen Bonding, Ph.D. Thesis, Department of Chemistry, Harvard University.

Ringe, D., & Petsko, G. A. (1985) *Prog. Biophys. Mol. Biol. 45*, 197–235.

Rossman, M. G., Arnold, E., Erickson, J. W., Frankenberger, E. A., Griffith, J. P., Hecht, H.-J., Johnson, J. E., Kamer,

G., Luo, M., Mosser, A. G., Rueckert, R. R., Sherry, B., & Vriend, G. (1985) *Nature (London) 317*, 145-153.
Roux, B., Yu, H.-A., & Karplus, M. (1990) *J. Phys. Chem. 94*, 4683-4688.
Ryckaert, J.-P., Ciccotti, G., & Berendsen, H. J. C. (1977) *J. Comput. Phys. 23*, 327-341.
Sheriff, S., Hendrickson, W. A., & Smith, J. L. (1987) *J. Mol. Biol.* 197, 273-296.
Shih, H. H.-L., Brady, J., & Karplus, M. (1985) *Proc. Natl. Acad. Sci. U.S.A. 82*, 1697-1700.
Shortle, D., & Lin, B. (1985) *Genetics 110*, 539-555.
Shortle, D., & Meeker, A. K. (1986) *Proteins 1*, 81-89.
Singh, U. C., Brown, F. K., Bash, P. A., & Kollman, P. A. (1987) *J. Am. Chem. Soc. 109*, 1607-1614.
Tembe, B. L., & McCammon, J. A. (1984) *Comput. Chem. 8*, 281-283.
Thanki, N., Thornton, J. M., & Goodfellow, J. M. (1988) *J. Mol. Biol. 202*, 637-657.
Warshel, A., Sussman, F., & King, G. (1986) *Biochemistry 25*, 8368-8372.
Weaver, L. H., & Matthews, B. W. (1987) *J. Mol. Biol. 193*, 189-199.
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.
Wiley, D. C., & Skehel, J. J. (1987) *Annu. Rev. Biochem. 56*, 365-394.
Wolfenden, R., Anderson, L., Cullis, P. M., & Southgate, C. C. B. (1981) *Biochemistry 20*, 849-855.
Wong, C. F., & McCammon, J. A. (1986) *J. Am. Chem. Soc. 108*, 3830-3832.
Zwanzig, R. W. (1954) *J. Chem. Phys. 22*, 1420-1426.

Effects of Inhibitors of N-Linked Oligosaccharide Processing on the Secretion, Stability, and Activity of Lecithin:Cholesterol Acyltransferase†

Xavier Collet and Christopher J. Fielding*
Cardiovascular Research Institute and Department of Physiology, University of California Medical Center, San Francisco, California 94143

Received May 18, 1990; Revised Manuscript Received October 12, 1990

ABSTRACT: The structure and function of the carbohydrate moiety of human lecithin:cholesterol acyltransferase (LCAT) were determined by using several glycosidases in reaction with the isolated plasma protein or by using specific inhibitors of glycoprotein assembly with cultured cells secreting LCAT activity. Analysis of the plasma enzyme indicated that almost all of the large carbohydrate moiety of LCAT (approximately 25% w/w) was N-linked with part of the high-mannose and part of the complex type. This analysis was confirmed with metabolic inhibitors of carbohydrate processing by using CHO cells stably transfected with the human LCAT gene. Inhibitors of the subsequent processing of the N-linked high-mannose chains formed by glucosidase activity were without effect on either the secretion rate or the catalytic activity of LCAT. The inhibition of catalytic activity by glucosidase inhibitors applied to both the phospholipase and the acyltransferase activities of LCAT. The reduction of the LCAT catalytic rate by terminal glycosidase inhibitors was without effect on apparent K_m and did not affect enzyme stability. These data indicate an unusual specific role for high-mannose carbohydrates in the catalytic mechanism of LCAT.

Lecithin:cholesterol acyltransferase (LCAT), a glycoprotein of M_r 63 000 synthesized and secreted by the liver, is an integral part of several high-density lipoprotein (HDL) species (Fielding & Fielding, 1981; Cheung et al., 1986; Francone et al., 1989). The enzyme catalyzes the transfer of an acyl group from the 2-position of phosphatidylcholine to the 3-position hydroxyl group of cholesterol and as such is believed to be responsible for generating a major fraction of plasma cholesteryl esters (Glomset, 1983).

The primary structure of human LCAT has been determined by cDNA and amino acid sequencing (Mclean et al., 1986; Yang et al., 1987). This reveals several regions of clustered hydrophobic residues, which may be related to the lipid-binding functions of the enzyme. The LCAT sequence also includes four potential N-glycosylation sites (Asn-X-Ser/Thr) at positions 20, 84, 272, and 384 of the 416-residue protein moiety. Total carbohydrate content has been estimated at about 25% of the total LCAT mass (Chung et al., 1979; Chong et al., 1983). In vitro modification of the sialic acid content of LCAT was shown to increase catalytic activity about 1.5-fold (Doi & Nishida, 1983). Otherwise, there has been little research on the potential role of the LCAT carbohydrate moiety on the secretion, stability, or functions of the enzyme protein.

The initial transfer of dolichol-linked glucose-capped high-mannose assemblies to polypeptide asparagine residues, and finally the modification of these units into the oligosaccharide sequences of the mature protein, proceeds first in the endoplasmic reticulum (ER) and subsequently within the Golgi apparatus (Snider, 1984; Kornfeld & Kornfeld, 1985). Within the ER, sequential glucosidase activities generate

† This research was supported by the National Institutes of Health through Arteriosclerosis SCOR Grant HL 14237.
* Address correspondence to this author at the Cardiovascular Research Institute.

0006-2960/91/0430-3228$02.50/0