8842 *Biochemistry* **2009,** *48,* 8842–8851
DOI: 10.1021/bi900685j

BIOCHEMISTRY
including biophysical chemistry & molecular biology
Article

Use of Stabilizing Mutations To Engineer a Charged Group within a Ligand-Binding Hydrophobic Cavity in T4 Lysozyme[†,‡]

Lijun Liu, Walter A. Baase, Miya M. Michael, and Brian W. Matthews*

Institute of Molecular Biology and Department of Physics, 1229 University of Oregon, Eugene, Oregon 97403-1229

Received April 21, 2009; Revised Manuscript Received July 31, 2009

ABSTRACT: Both large-to-small and nonpolar-to-polar mutations in the hydrophobic core of T4 lysozyme cause significant loss in stability. By including supplementary stabilizing mutations we constructed a variant that combines the cavity-creating substitution Leu99 → Ala with the buried charge mutant Met102 → Glu. Crystal structure determination confirmed that this variant has a large cavity with the side chain of Glu102 located within the cavity wall. The cavity includes a large disk-shaped region plus a bulge. The disk-like region is essentially nonpolar, similar to L99A, while the Glu102 substituent is located in the vicinity of the bulge. Three ordered water molecules bind within this part of the cavity and appear to stabilize the conformation of Glu102. Glu102 has an estimated pK_a of about 5.5−6.5, suggesting that it is at least partially charged in the crystal structure. The polar ligands pyridine, phenol and aniline bind within the cavity, and crystal structures of the complexes show one or two water molecules to be retained. Nonpolar ligands of appropriate shape can also bind in the cavity and in some cases exclude all three water molecules. This disrupts the hydrogen-bond network and causes the Glu102 side chain to move away from the ligand by up to 0.8 Å where it remains buried in a completely nonpolar environment. Isothermal titration calorimetry revealed that the binding of these compounds stabilizes the protein by 4−6 kcal/mol. For both polar and nonpolar ligands the binding is enthalpically driven. Large negative changes in entropy adversely balance the binding of the polar ligands, whereas entropy has little effect on the nonpolar ligand binding.

Soluble globular proteins are characterized by tight nonpolar packing of the interior to form a hydrophobic core (*1*). However, internal cavities do exist (*2*). In some cases these cavities are polar in character and accommodate water molecules (*2*−*4*). In other cases there can be cavities that are largely hydrophobic and typically appear to be empty (*5*). When charged residues occur within proteins, they are often (*6*) but not always (*7*) buried as ion pairs. If a normally charged residue is buried and noncompensated, it often has a critical role in protein function and, in most cases, will be present in the neutral state (*7*−*9*).

In general, the free energy of folding for a globular protein ranges from 5 to 15 kcal/mol and is independent of size (*10*). This implies that there is only a small window within which destabilizing mutations are allowed. For example, the hydrophobic interiors of protein can sometimes be repacked by making replacements with residues of equivalent hydrophobicity (*11*). In contrast, large-to-small mutations in the hydrophobic core typically cause a larger destabilization (*12*). So too does burying polar or charged groups in the hydrophobic core (*13*−*19*). The combination of creating a large hydrophobic cavity and then introducing a charged group into the cavity is expected to be especially deleterious.

In T4 lysozyme, the large-to-small mutation Leu99 to Ala (L99A) creates an essentially nonpolar cavity and decreases the stability of the protein ~5 kcal/mol (*12*). A series of ligands has been identified which bind in this cavity, mostly stabilizing the protein (*20*−*22*).

In another T4 lysozyme variant the mutation of Met102 to lysine (M102K) buries the lysine and causes a drop in stability of 2.2 and 8.9 kcal/mol at pH 3.0 and 10.4, respectively (*15*). Similarly, losses in stabilization energy of 4−5 kcal/mol were caused by the mutation of Leu133 to aspartate in the hydrophobic core (*15*).

It has been shown that the mutant M102Q could be used to introduce polarity into the L99A cavity and change its ligand-binding specificity (*23*). Here we explore the consequences of the more demanding combination, L99A/M102E. Is it possible to introduce a potentially charged group into an otherwise nonpolar cavity? If so, what are the consequences for ligand binding specificity?

As will be shown, the L99A/M102E combination, introduced into pseudo-wild-type lysozyme, is prohibitively destabilizing. By adding supplemental stabilizing substitutions it was possible to obtain a folded protein which permitted structural and ligand-binding analysis. We refer to the final construct as L99A/M102E/St[1] where the "St" stands for "stabilized".

EXPERIMENTAL PROCEDURES

Gene Cloning, Protein Expression and Purification. As a starting point, the mutant L99A/M102E was made in cysteine-free

[†]This project was supported in part by NIH grant GM21967 to B.W.M.

[‡]PDB access codes: 3GUI, 3GUJ, 3GUK, 3GUM, 3GUL, 3GUP, 3GUO, 3GUN.

*Corresponding author: Brian W. Matthews, Institute of Molecular Biology, 1229 University of Oregon, Eugene, OR 97403-1229. Phone (541)346-2572. Fax: (541)346-5870. E-mail: brian@uoregon.edu.

[1]Abbreviations: CD, circular dichroism; T_m, protein melting temperature; L99A/M102E/St, stabilized T4 lysozyme variant including the mutations Leu99 to Ala and Met102 to Glu.

 Published on Web 08/10/2009

Table 1: Statistics for X-ray Data Collection, Processing and Structure Refinement

crystal	ligand-free	benzene	toluene	*p*-xylene	ethylbenzene	pyridine	phenol	aniline
space group	$P4_12_12$	$P4_12_12$	$P4_3$	$P4_3$	$P4_3$	$P4_3$	$P4_3$	$P4_3$
unit cell								
a, *b* (Å)	48.77	48.81	49.49	49.71	49.27	48.96	49.51	48.70
c (Å)	128.51	128.37	129.51	129.71	129.33	129.31	130.25	128.65
resolution (Å)	50.0−1.45 (1.50−1.45)	50.0−1.60 (1.66−1.60)	50.0−1.85 (1.92−1.85)	50.0−2.25 (2.33−2.25)	50.0−2.07 (2.14−2.07)	50.0−1.50 (1.55−1.50)	50.0−2.15 (2.23−2.15)	50.0−1.50 (1.55−1.50)
completeness[a] (%)	99.8 (100.0)	92.5 (96.2)	99.5 (99.9)	99.5 (99.0)	99.9 (99.9)	97.5 (99.5)	99.8 (99.9)	98.2 (99.9)
redundancy	6.7 (6.6)	6.0 (6.2)	3.7 (3.7)	7.2 (7.1)	3.7 (3.6)	3.7 (3.6)	7.5 (7.3)	3.7 (3.6)
$I/\sigma(I)$	42.0 (6.3)	19.1 (3.9)	26.1 (3.2)	39.0 (6.2)	27.0 (3.5)	20.0 (3.8)	42.8 (5.7)	22.0 (3.2)
R_{merge}[b] (%)	3.6 (28.0)	6.5 (39.8)	4.4 (36.9)	6.1 (39.7)	4.6 (33.1)	4.7 (46.3)	4.4 (39.8)	4.4 (40.1)
R (%)	18.5	20.2	22.5	21.8	23.7	21.1	21.4	22.1
R_{free}[c] (%)	20.7	23.9	27.3	26.3	29.3	25.3	27.3	25.1
rmsd of bond								
length (Å)	0.01	0.01	0.01	0.02	0.01	0.02	0.02	0.02
angle (deg)	1.35	1.15	1.23	1.58	1.17	1.64	1.49	1.83
Ramachandran distribution								
preferred (%)	98.6	98.7	96.6	95.5	96.8	98.7	97.1	97.5
allowed (%)	1.4	1.3	3.1	3.2	3.2	1.3	2.6	2.5
mean *B* (Å^2)	19	26	44	43	44	30	46	31
PDB code	3GUI	3GUJ	3GUK	3GUM	3GUL	3GUP	3GUO	3GUN

[a]Values in parentheses correspond to the outer resolution shell. [b]$R_{merge} = \sum_{hkl,i}|I_{hkl,i} - \langle I_{hkl}\rangle|/\sum_{hkl,i}I_{hkl,i}$. [c]5% of the unique reflections were randomly selected for the calculation of R_{free}.

pseudo-wild-type lysozyme (*24*) following the standard protocol (*25*, *26*). The resultant protein was highly unstable and did not permit structural study (data not shown). To enhance stability the mutations T21C, S38D, E108V, S117V, T142C and N144D were introduced, and the original two cysteine residues, C54 and C97, were restored. Thus, relative to wild-type T4 lysozyme the final construct includes the mutations T21C/S38D/L99A/M102E/E108V/S117V/T142C/N144D. It will be referred to as L99A/M102E/St, where L99A is the cavity-creating mutant, M102E introduces the charge and "St" is for "stabilized".

The gene was cloned into the pHS1403 vector and the protein expressed with *Escherichia coli* RR1 strain (*26*). The protein was forced to overexpress into inclusion bodies for 3 h at 37 °C by adding 1.0 mM IPTG when the OD_{600} reached 1.2. Purification was performed at 4 °C. For a 4 L preparation, the collected cell pellet was suspended in 100 mL of buffer solution (50 mM Tris HCl, pH 8.0, 10 mM sodium EDTA and 50 mM NaCl) with protease inhibitors (Roche) and disrupted by sonication. After centrifugation, the pellet was resuspended and stirred overnight with 100 mL of wash buffer solution (50 mM Tris HCl, pH 8.0, 50 mM sodium EDTA and 2.0% Triton X-100). Again, the pellet was collected and resuspended and stirred with 200 mL of wash buffer (50 mM Tris HCl, pH 8.0 and 2.5% n-octyl-β-D-glucopyranoside) for at least 3 h.

After centrifugation, the purified inclusion body was suspended in 25 mL of 10 mM glycine. Solid urea was added to a final concentration of 4.0 M with intense stirring. The solution was quickly titrated to pH 3.0 with phosphoric acid. After centrifugation, the supernatant was saved and dialyzed overnight against 4 L of a solution of 50 mM sodium citrate, pH 3.0, 15% glycerol and 5 mM β-cyclodextrin, followed by another one-day dialysis against 4 L of the same solution except the pH was adjusted to 5.0. After removal of precipitate, the solution with refolded protein was diluted to a conductivity less than 4 mmho/cm, loaded onto a CM Sepharose ion exchange column (CL-6B, Sigma), and eluted with a linear gradient of 0 to 400 mM NaCl in 10 mM sodium citrate, pH 5.0. The protein-containing fractions were pooled and dialyzed at least twice at a 1:50 ratio against a harvesting buffer of 10 mM sodium citrate, pH 5.0 and 200 mM NaCl. The dialysis apparently removed small-molecule contaminants and made it possible to concentrate the protein to 40 mg/mL.

Crystallization. Crystals were obtained by the vapor-diffusion sitting-drop method using a protein concentration of 8.0 mg/mL (*25*). Two crystal forms were obtained with different morphology and distinct space groups, typically taking six months to grow. Conditions that gave crystals included 30% PEG 8000, 0.20 M $(NH_4)_2SO_4$, 0.10 M Tris HCl, pH 8.0, and 30% PEG 4000, 0.20 M $(NH_4)_2SO_4$, 0.10 M HEPES, pH 7.0. The inclusion of Mg^{2+} or Ca^{2+} made crystallization more reproducible, e.g. 30% PEG 8000, 0.14 M calcium acetate, 0.10 M PIPES, pH 6.5, but did not shorten the time for crystal growth. Without these ions the crystals grew with better shape and larger size but with very low frequency. In addition, crystals obtained with Mg^{2+} or Ca^{2+} were poorer looking but had comparable diffraction quality. Crystallization setups typically included 50 mM oxidized β-mercaptoethanol and 50 mM reduced β-mercaptoethanol, additives that have been helpful in crystallizing a number of T4 lysozyme mutants.

Ligand Complexes. For volatile ligands (benzene, toluene, ethylbenzene, *p*-xylene and pyridine), the protein−ligand complexes were obtained by vapor diffusion (*20*) over a period of three days at 4 °C. For the less volatile ligands (aniline, phenol, benzylamine, phenylethylamine, phenylpropylamine and phenylbutylamine), the L99A/M102E/St crystals were soaked overnight in the reservoir solution containing the ligand at a concentration of 50 mM or saturated if the solubility was less than 50 mM. Binding was first tested crystallographically and, where appropriate, by isothermal titration calorimetry.

Data Collection and Processing. Diffraction data were recorded on the ADSC Quantum 315 CCD detectors at beamlines 8.2.1 and 8.2.2 of the Advanced Light Source (Berkeley, CA) using a wavelength of 1.000 Å. Crystals were frozen to 100 K to reduce radiation damage. For most crystallization conditions, crystals could be directly frozen without additional cryoprotectants or, if needed, 10% glycerol was added. Data were processed with HKL2000 (*27*), and statistics are shown in Table 1.

Table 2: Thermal Unfolding of Mutant Lysozymes[a]

protein	unfolding at pH 3.0			unfolding at pH 5.3		
	T_m (°C)	ΔH (kcal/mol)	$\Delta\Delta G$ (kcal/mol)	T_m (°C)	ΔH (kcal/mol)	$\Delta\Delta G$ (kcal/mol)
WT	53.5	122	0	66.9	147	0
WT*	51.8	110	−0.76	65.2	127	−0.8
L99A/WT*	36.6	79	−5.4	53.4	98	−4.8
L99A/M102E/St reduced	37.7	59	−4.6	ND	ND	ND
L99A/M102E/St oxidized	53.3	89	−0.1	58.1	83	−3.2

[a] $\Delta\Delta G$ values are computed as ΔG(mut) − ΔG(WT) where ΔG are standard free energies of unfolding computed at 45 °C and 60 °C for pH 3 and pH 5.3 data, respectively. ΔC_p values of 1.8 kcal/mol-deg were used for the pH 3 data and 2.5 kcal/mol-deg for the pH 5.3 data. At these reference temperatures uncertainties in ΔC_p contribute no more than ±5% to $\Delta\Delta G$. Reference data are included both for wild-type lysozyme (WT), which includes Cys54 and Cys97, and for pseudo-wild-type (WT*) in which these cysteines have been replaced (C54T, C97A). Reliable data for the reduced form of L99A/M102E/St at pH 5.3 could not be obtained (see text).

The ligand-free and the benzene-binding structures of L99A/M102E/St were determined with crystals obtained in the absence of Mg^{2+} or Ca^{2+}. Their space group is $P4_12_12$, with one molecule per symmetric unit (Table 1). Crystals obtained in the presence of these divalent ions were used for structure determination of the complexes with the other ligands. Although their cell dimensions are close to those of the other crystal form, their space group was determined to be $P4_3$, with two molecules per asymmetric unit. The $P4_3$ crystals have pseudo $P4_32_12$ symmetry with a local 2-fold axis running through the Mg^{2+}/Ca^{2+} binding site. This packing is not related to that in the $P4_12_12$ space group mentioned above.

Structure Determination. Structures were determined by molecular replacement (*28*), and refinement was with *REFMAC* (*28*). In brief, rigid-body refinement was done first and as needed thereafter, followed by restrained refinement with TLS refinement included. Modeling was done with COOT (*29*). The identification and placement of ligands was based on $F_{obs} - F_{obs}$ maps (for benzene) or $F_{obs} - F_{calc}$ maps (for other ligands). To reduce possible bias from the model, features of interest inside the cavity, including ligands, water molecules and the side chain of Glu102, were omitted in the initial refinement, and were included only after the other solvent molecules had been placed. The TLS refinement was also performed within *REFMAC* (*28*). The *B*-factors noted in the Results and Discussion have been corrected by the program TLSANL in the CCP4 package (*28*). In the different structures the solvent-exposed cysteine at site 97 was typically seen to be derivatized with β-mercaptoethanol, but this was not the case for the less exposed Cys54.

With knowledge of the noncrystallographic symmetry and its possible implications for twinning, the $P4_3$ crystals were subjected to additional indexing and twinning tests. Refinements were also performed in parallel in space group $P4_32_12$, which typically resulted in R_{free} being 4% higher when the comparable refinement protocols were applied. Inclusion of twinning factors did not improve structure refinement. Refinement statistics are summarized in Table 1.

Protein Stability and Ligand Binding. Thermal unfolding measurements (Table 2) were made using in-cell stirring and an in-cell thermal probe as outlined in Eriksson et al. (*25*) except that a JASCO model 810 spectropolarimeter and the model PDF-425S thermal controller were used. Melting temperatures, T_m, and enthalpies of unfolding at T_m were determined using the JASCO Spectrum Manager "denatured protein" software. The pH 3.0 buffer was 25 mM KCl, 3 mM H_3PO_4, 17 mM KH_2PO_4, while the pH 5.3 buffer was 0.10 M NaCl, 10 mM $H_{0.14}Na_{0.86}OAc$.

Binding constants and enthalpies of binding (Table 3) were estimated using a VP-ITC isothermal titration calorimeter (MicroCal, Northampton, MA). Operating in multiple injection mode, a 1−2 mM solution of L99A/M102E/St was injected stepwise into buffer to determine a baseline injection heat profile, and then into buffer containing ligands dissolved overnight to about 0.1 mM with stirring on ice. A 1 μL, 29 × 10 μL injection scheme with 180 s between injections was used. Baseline heats were subtracted from reaction heats and the resulting profile fit to a single-site binding model using Origin version 7E (OriginLab Corp., Northampton, MA). Titrations were done at 4 °C with stirring at 307 rpm. The buffer was 200 mM NaCl, 10 mM sodium citrate, pH 4.7. Proteins were dialyzed into outgassed buffer. Concentrations were determined with an HP-8453 spectrophotometer (Agilent Technologies, Santa Clara, CA) and the molar absorptivity at 280 nm of 24,170 L mol^{-1} cm^{-1} for T4 lysozyme determined by Elwell and Schellman (*30*). In modeling the binding reaction, ligand concentrations were adjusted to give unit binding stoichiometries. Reductions of 10−20% in nominal concentrations were usual for the more volatile compounds.

Table 3: Thermodynamic Analysis of Ligand Binding[a]

ligand	$\Delta H°$ (kcal mol^{-1})	$\Delta S°$ (cal K^{-1} mol^{-1})	$\Delta G°$ (kcal mol^{-1})	K (mol^{-1})
benzene	−6.0	−0.7	−5.8	37200
toluene	−5.3	−0.1	−5.3	14100
p-xylene	−3.7	4.2	−4.9	6610
ethylbenzene	−5.0	2.1	−5.5	22800
phenol	−9.2	−15.0	−5.0	9150
aniline	−8.1	−14.8	−4.0	1360

[a] $\Delta H°$, $\Delta S°$ and $\Delta G°$ are standard enthalpies, entropies and free energies of binding for the reaction [protein + ligand ↔ complex]. K is the equilibrium constant for binding and yields $\Delta G° = -RT \ln K$ where, in this case, temperature was 4 °C. Heats per injection were based on the protein concentration. The protein was always in the injection syringe. Aromatic ligands, in the sample cell, were more or less volatile. Stoichiometries ranged from 1.07 to 0.78. Error estimates were based on repeated titrations done for some ligands. Generally, the highest binding constants had the least error with ±30% in K seen for the lower values of K where titration profiles were relatively featureless.

Measurement of pK_a. We anticipated that measurement of the pK_a of Glu102 would be difficult because L99A/M102E/St precipitates above neutral pH, at least at 10 mg/mL. In order to have a reference protein that lacked Glu102, we made the construct L99A/St. This is identical to L99A/M102E/St but has the wild-type methionine at position 102. The protein was purified as described (*25*, *26*).

Circular dichroism (CD) spectra were taken using a JASCO model 810 spectropolarimeter from 340 to 200 nm at 2.5 °C in

Hellma QS-111 without stirring at various pH values. L99A/St was initially at 1.65 μM in 1.6 mL of 5 mM citric acid, pH 2.86. L99A/M102E/St was initially 3% lower in concentration. pH values were determined in parallel at 2.5 °C using the Radiometer PHM240 meter, the GK2401C electrode, and method A with the meter standardized also at 2.5 °C using the Radiometer IUPAC standards labeled 1.679, 4.005, 6.865, and 10.01. Aqueous K_2HPO_4 was delivered via two identical micrometer-driven syringes to both the pH and optical cells. The pH cell was stirred and the QS-111 was mixed via a single inversion at each pH. Total time to collect data was 2.5 h per protein. Spectra were corrected for volumetric dilution and the signal at 223 nm displayed as corrected CD.

RESULTS AND DISCUSSION

Inclusion of the Stabilizing Substitutions. As noted above, the cavity-containing mutant L99A and the buried charge mutant M102E are both substantially destabilizing and the double mutant L99A/M102E does not appear to fold.

In order to obtain protein amenable to structural analysis we had to include a series of stabilizing substitutions. The final construct is T21C/S38D/L99A/M102E/E108V/S117V/T142C/N144D. None of the supplemental mutations is located close to the charge-containing cavity.

The stability data are summarized in Table 2. Each of the mutations S38D, E108V, S117V and N144D is expected to increase the melting temperature of T4 lysozyme by 1.5−5 °C (*31*−*33*). These contributions, which are expected to be additive (*31*), total 12 °C at pH 3.0. At the same pH the melting temperature of reduced L99A/M102E/St is about the same as that of L99A, which suggests that the combined stabilization from S38D, E108V, S117V plus N144D roughly offsets the loss of stability due to the introduction of the buried glutamic acid (M102E). Closure of the 21−142 disulfide bridge (*34*), however, in oxidized L99A/M102E/St had a more dramatic effect, increasing T_m (relative to the reduced form) by 15.6 °C. This increased the stability of L99A/M102E/St to be essentially equal to that of wild-type T4 lysozyme at pH 3.0 and somewhat less stable at pH 5.3.

To check on the status of the Cys21−Cys142 disulfide bridge, the L99A/M102E/St protein at 15 mg/mL (0.8 mM) was incubated with 25 mM DTT at room temperature and the stability monitored as a function of time (data not shown). At zero time there appeared to be a single species present with a melting temperature of 53.3 °C and ΔH of 89 kcal/mol monitored at pH 3.0. After 18.5 h about 23% of the protein had converted to the low T_m (reduced) species, and this gradually increased to 52% after 72 h. This reduction reaction was also monitored by thermal unfolding at pH 5.3. In this case no distinct species with lower melting temperature was seen. Rather, as the reaction proceeded, a low enthalpy form of the protein became increasingly apparent essentially co-unfolding with the remaining high enthalpy form. Both sets of data support the idea that the 21−142 disulfide bridge was fully formed after protein purification and refolding.

The overall structure of L99A/M102E/St remained essentially the same in the apo- and ligand-bound forms. Superimposition of the C^α atoms revealed a rmsd of ~1.5 Å. For those crystals with space group $P4_3$, which were used for most of the ligand-bound structures, the two molecules in the asymmetric unit are also closely similar, although the loop regions associated with the Cys21/Cys142 disulfide bridge showed relatively large difference in thermal factors. These regions are associated with crystal contacts. In one copy of L99A/M102E/St, the disulfide bond is highly ordered, while this region shows higher flexibility in the other copy.

Ligand binding adversely affected crystal diffraction, especially for the nonpolar ligands of larger size (Table 1).

The Polarized Cavity Contains Ordered Water. The cavity resulting from the original L99A mutation has a donut or disk-like shape with an additional bulge in the vicinity of Met102 (*22*). When Met102 is replaced by glutamic acid (and the stabilizing mutations are included elsewhere in the protein), the overall shape of the cavity is conserved. This also implies that Glu102 occupies the site vacated by Met102 and remains well-ordered within the core of the protein (Figure 1).

Ordered water molecules are seen at three sites in the bulge region of the cavity, one site being partially occupied. The first fully occupied site, Wat601, is 2.8 Å from $O^{\varepsilon 1}$ of Glu102. The second, Wat602, is bound by the carbonyl oxygen atoms of Val111 (2.6 Å) and Phe114 (3.0 Å). The distance between the two water molecules is 2.9 Å, and their *B*-factors are 36 and 28 Å^2, respectively. These values are 10−20 Å^2 higher than those of the oxygen atoms of Glu102 (20 Å^2) and overall averaged *B*-factor (19 Å^2), suggesting enhanced mobility of these water molecules. The third partially occupied water molecule site (Wat603) is 2.7 Å and 3.3 Å, respectively, from the $O^{\varepsilon 2}$ atom of Glu102 and Wat602. There is also a short distance of 2.7 Å to the side chain of Phe114 which is partially disordered (Figure 1). Perhaps as a consequence of the disordering of Phe114, Wat603 has a refined occupancy of 0.35. It is likely that Phe114 could act as a "gate" between the cavity and the bulk solvent.

The observation that water molecules are bound in the vicinity of Glu102 has some similarity to the behavior observed for the Val66 → Glu, Val66 → Asp and Ile92 → Glu mutants of staphylococcal nuclease (*17*−*19*). In each of these cases the buried Asp or Glu side chain hydrogen bonds to one or more internal water molecules.

In contrast to the bulge region, the main nonpolar part of the L99A/M102E/St cavity does not show any indication of bound solvent. In the cavity of the previously reported T4 lysozyme variant, L99A/M102Q (which lacks the stabilizing mutations), the main region of the cavity was found to bind β-mercaptoethanol, with a single ordered water molecule in the bulge region hydrogen-bonded to Gln102 (*23*). Recently a careful, low-temperature analysis was made of the electron density in the mutant L99A/M102L (*35*). In that variant the cavity was similar in size to that of L99A, and also largely nonpolar. It appeared to contain 1−2 poorly ordered water molecules, hydrogen bonded to each other and to the partially exposed carbonyl group of Ala99.

In summary, the mutation of M102E introduces polarity into the bulge region of the L99A cavity. Three ordered waters bind in this region, stabilized by the side-chain oxygens of Glu102 and the carbonyl groups of Val111 and Phe114. Otherwise, the nonpolarity of the main region of the cavity is conserved, and no ordered water molecules are seen in this region. The cavity serves as a useful model for testing the binding of ligands of varying polarity. Brenk et al. (*36*) have explored the use of an engineered cavity of comparable volume in cytochrome *c* peroxidase that also includes a negative charge (Asp23). The cavity binds four ordered water molecules and a potassium ion and is more polar than that in T4 lysozyme.

pK_a of Glu102. To try to estimate the pK_a of Glu102 we monitored circular dichroism (CD) as a function of pH for

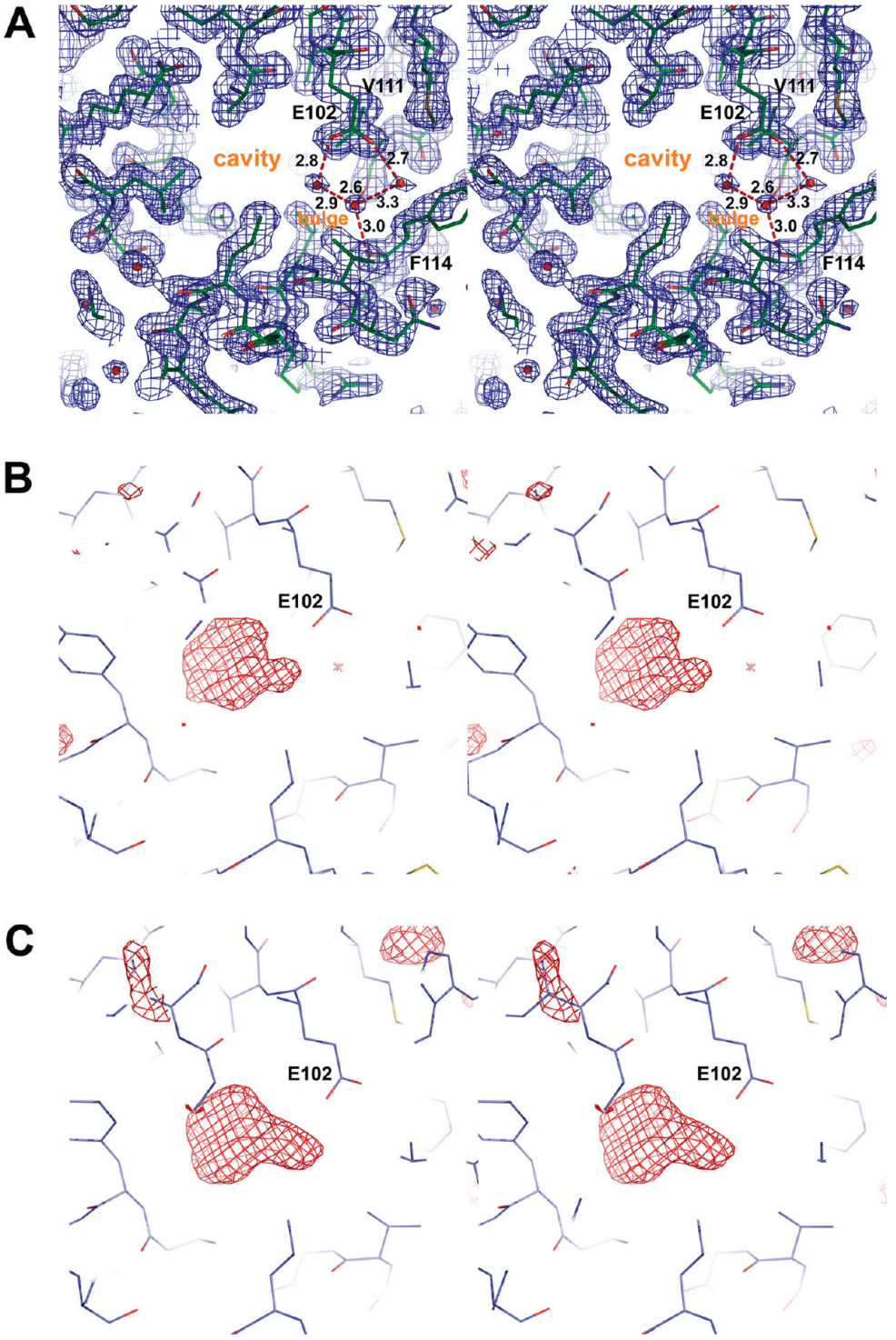

FIGURE 1: Stereoviews of the electron density in the vicinity of the L99A/M102E/St cavity. (A) The apo-L99A/M102E/St structure. The map coefficients are $(2F_o - F_c)$ where F_o and F_c are the observed and calculated structure factors for the L99A/M102E/St crystals. The resolution is 1.45 Å and the map contoured at 1σ. The central region of the cavity is labeled "cavity", and the introduced glutamic acid, Glu102, points to the "bulge" region. Three ordered water molecules bound in this region are, from left to right, Wat601, Wat602 and Wat603. Hydrogen bonds are shown as dashed red lines. (B) The unbiased $F_o - F_c$ electron-density map of the cavity region for the complex with aniline. (C) The unbiased $F_o - F_c$ electron-density map of the cavity region for the complex with ethylbenzene. In (B) and (C), the maps were calculated using the calculated phases from the corresponding refined model excluding the ligand. The $F_o - F_c$ maps were contoured at 3.0σ. All figures were generated using PyMOL (*38*).

L99A/M102E/St and the reference protein L99A/St which has methionine at site 102.

Increasing pH (and ionic strength) resulted in some loss of CD signal for L99A/St at 223 nm even after correction for dilution (Figure 2). However, the data are fit well by a straight line at least from pH 3 to just over pH 8. In the case of L99A/M102E/St, CD decreases initially in a parallel manner until about pH 5. After that the CD at 223 nm decreases more rapidly. While it is tempting to force-fit the data to the Henderson−Hasselbalch equation for a single titratable group, these data are noisy. In addition, subsequent titrations, done more rapidly, show the same effect but of diminished intensity. This indicates this

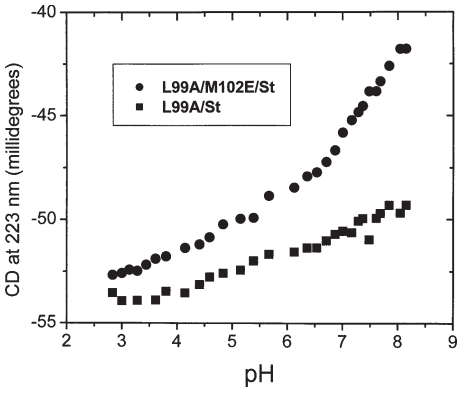

FIGURE 2: Changes in circular dichroism as a function of pH for L99A/M102E/St (solid circles) and L99A/St (solid squares).

titration has a kinetic dimension at least for L99A/M102E/St, a protein that tends to precipitate above neutral pH. For these reasons we conclude that the pK_a of Glu102 is probably between 5.5 and 6.5 but cannot be more accurately defined based on this technique. (Attempts to measure the pK_a using fluorescence were unsuccessful because L99A/St has pH-dependent changes in fluorescence which prevent it being used as a reference. Data not shown.)

Because the pH of the crystals is in the range 6.5−7.0, Glu102 is presumably partially charged in the crystal structure.

Binding of Polar Ligands: Pyridine, Phenol and Aniline. X-ray crystallographic analysis was used to evaluate the binding to the L99A/M102E/St cavity of the polar ligands pyridine, phenol, aniline, benzylamine, phenylethylamine, phenylpropylamine, and phenylbutylamine. None of the phenylalkylamines were detected to bind, whereas pyridine, phenol and aniline did so. The difference density map for aniline is shown in Figure 1B.

(*a*) *Pyridine*. The structure of pyridine bound to L99A/M102E/St lysozyme is shown in Figure 3A. Water molecule Wat603, which has an occupancy of 0.35 in the apo structure, is displaced, and the remaining two solvent molecules are reorganized. As a result Wat601 hydrogen bonds to each of the carbonyl oxygens of Glu102 as well as to the other bound solvent molecule, Wat602. The pyridine nitrogen is shown to have a close contact (3.0 Å) to Wat601 (not directly to the side chain of Glu102).

The *B*-factors of the carboxylate oxygens of Glu102 are 31 Å^2 and 32 Å^2, those for Wat601 and 602 are 35 Å^2 and 36 Å^2, and that for the pyridine nitrogen is 28 Å^2. The overall *B*-factor for the structure is 30 Å^2. These values suggest that the hydrogen-bonding network is strengthened in the pyridine complex relative to the apo structure.

(*b*) *Phenol*. The larger ring substituent of phenol relative to pyridine, together with its expanded hydrogen-bonding potential, lead to a further change in the solvent structure within the cavity (Figure 3B). The phenolic oxygen hydrogen bonds to Wat601, but also makes a direct hydrogen bond to one of the oxygens of Glu102. Wat601 is still hydrogen-bonded to both of the oxygens of Glu102 but is pushed deeper into the bulge region, displacing Wat602. Presumably because the single remaining solvent molecule Wat601 has three hydrogen-bonds relative to four in the pyridine complex, its *B*-factor increases from 35 Å^2 to 57 Å^2.

(*c*) *Aniline*. Generally speaking, aniline binds to the cavity in a manner very similar to that for phenol. The amine group makes direct hydrogen bonds to the side chain of Glu102 and to Wat601 (Figure 3C). Wat601 adjusts its position to lose a hydrogen bond to $O^{\varepsilon 1}$ of Glu102 (3.7 Å), maintain a hydrogen bond to $O^{\varepsilon 2}$ of Glu102 (3.0 Å), and make a new hydrogen bond to the backbone carbonyl of Val111 (3.0 Å). All of the hydrogen-bonding atoms including Wat601 have *B*-factors in the range 25−34 Å^2, indicative of a well-ordered network.

The six-membered rings of pyridine, phenol and aniline all bind in a very similar fashion and occupy essentially the same location as that of benzene bound to L99A (*20*). The phenyl rings of phenol and aniline superimpose on that of pyridine with a root-mean-square deviation of less than 0.2 Å, this discrepancy being due mostly to rotations of less than 10° around the 2-fold or pseudo-2-fold axes of the respective ligands.

The binding affinity of the polar ligands is expected to be dependent on their size, shape and basicity. The pK_a of aniline (protonated), pyridine (protonated) and phenol are 4.63, 5.25 and 9.89, respectively (*37*). In the neutral state, their basicity follows the order of phenol < aniline < pyridine. As determined by ITC the binding constants of phenol and aniline are 9150 M^{-1} and 1360 M^{-1}, respectively (Table 3). Due to low heat of binding a value could not be obtained for pyridine. Because phenol and aniline have similar shapes, they can be compared directly. Aniline is more basic but binds less well. This suggests that electrostatic interactions with Glu102 are not the dominant factor in binding. In this context it can also be noted that pyridine is even more basic than aniline but does not make a direct interaction with Glu102 (Figure 2A). In no case does there appear to be a fully developed "ion pair" interaction between Glu102 and the bound ligand.

Binding of Nonpolar Ligands: Benzene, Toluene, p-Xylene and Ethylbenzene. As with the polar compounds, X-ray crystallography was used to analyze the binding of a set of representative nonpolar ligands of different shape and size. The difference density for ethylbenzene is shown in Figure 1C.

(*a*) *Benzene*. When benzene (at 10 mM) binds to L99A, it increases the melting temperature by 5−6 °C (*20*). A similar thermal upshift occurs when this ligand binds to L99A/M102E/St (not shown). The geometry of benzene binding to L99A/M102E/St is also similar to that for L99A (*20*). The overall size and shape of benzene matches fairly well the main, nonpolar region of the cavity. As with the pyridine complex, two water molecules, Wat601 and Wat602, remain bound (Figure 3A). They have *B*-factors of 42 Å^2 and 37 Å^2. The closest distance from the benzene ring to the side chain of Glu102 and to the closest of the bound water molecules is 3.7 Å in each case.

(*b*) *Toluene*. The difference electron density map for toluene (not shown) clearly indicated that this ligand bound within the cavity, although the alignment was not obvious. Crystallographic refinement (Table 1) indicated that toluene has two alternative modes of binding (Figure 4B). In the main mode (occupancy 0.6), the methyl group is closest to Wat601. In the alternative mode (occupancy 0.4), the ring is azimuthally rotated by about 80° and the ring centroid shifts 0.6 Å. The occupancy of Wat601 was set at unity and its *B*-factor refined to 51 Å^2. Wat602, which is retained in the complex with benzene, is expelled when toluene binds.

(*c*) *p-Xylene*. Notwithstanding its apparently unfavorable shape, *p*-xylene bound in the L99A/M102E/St cavity (Figure 4C). The addition of the extra methyl group at the *para* position caused the phenyl ring to translate 0.7 Å toward the bulge region. This in turn caused displacement of all of the water molecules from the cavity. Even though the side chain of Glu102 is thereby buried in a completely polar region with no apparent hydrogen-bonding partners, it remains quite well ordered. The *B*-factors of

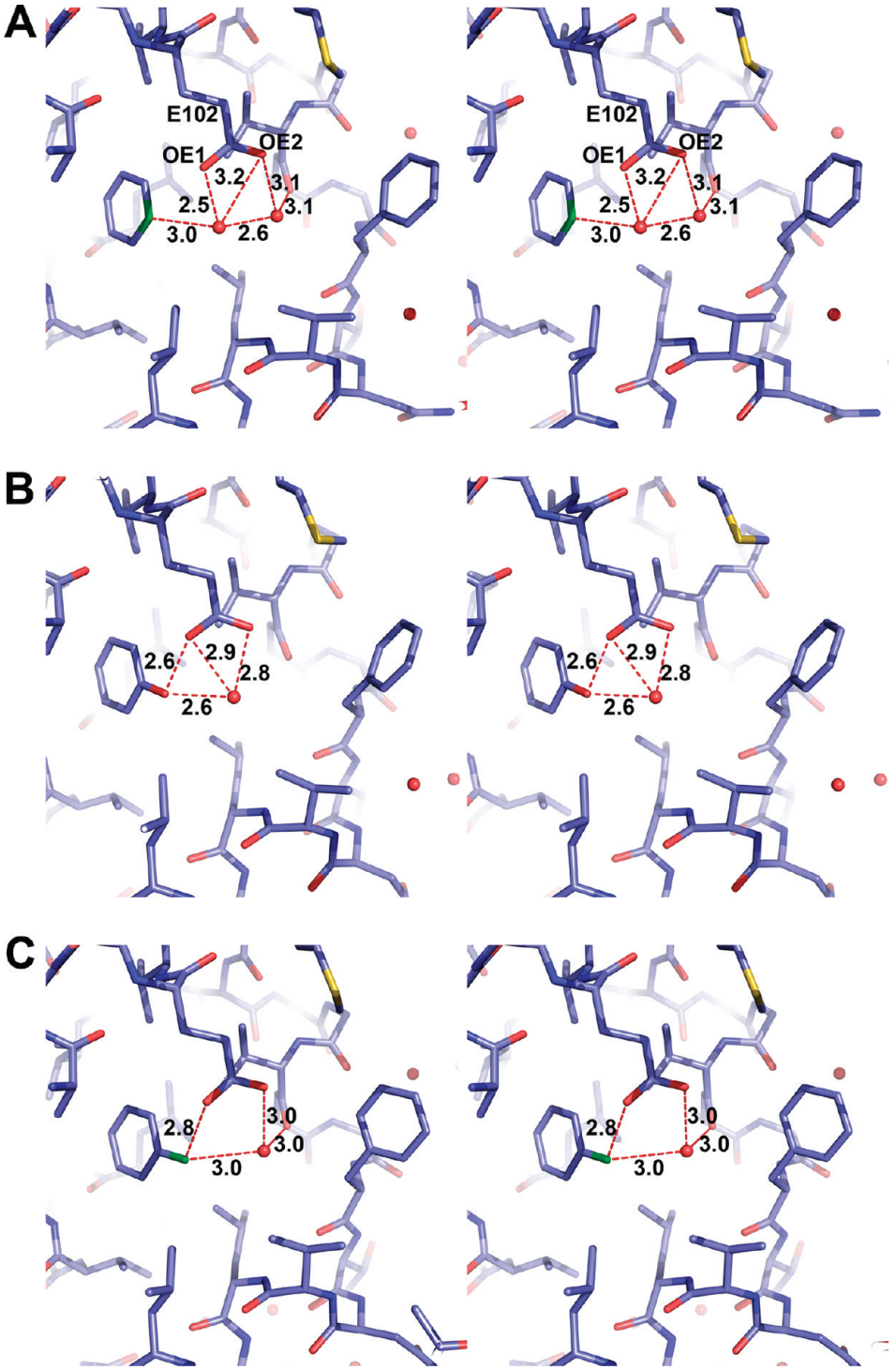

FIGURE 3: Stereoviews showing polar ligands bound within the L99A/M102E/St cavity. Solvent molecules are shown as red spheres with hydrogen bonds drawn as dashed lines. (A) Pyridine. (B) Phenol. (C) Aniline.

$O^{\varepsilon 1}$ and $O^{\varepsilon 2}$ are 44 Å^2 and 47 Å^2, respectively. The burial of an acidic group in such a fashion is highly unusual. As shown in Figure 5, the side chain of Glu102 moves somewhat in response to the binding of the different ligands, the largest displacement being for *p*-xylene.

(*d*) *Ethylbenzene*. From a geometrical point of view, the overall shape of ethylbenzene matches quite well the overall shape of the cavity. The benzene moiety occupies the central part of the cavity while the ethyl group fills the bulge region (Figure 4D). As for *p*-xylene, ethylbenzene displaces all bound solvent molecules from the cavity. Also, as in the complex with *p*-xylene, Glu102 maintains its position in a strictly nonpolar non-hydrogen-bonded environment.

The geometries of binding of all ligands, both polar and nonpolar, are compared in Figure 5.

Energetics of Binding. The thermodynamic ligand-binding constants measured by isothermal titration calorimetry are given in Table 3. Although benzene is the smallest ligand and is non-hydrogen-bonding, it binds most tightly ($\Delta G^\circ = -5.8$ kcal/mol). This is presumably because it occupies the nonpolar part of the L99A/M102E/St cavity and does not have any unfavorable short contacts with Glu102 or the two highly occupied water molecules

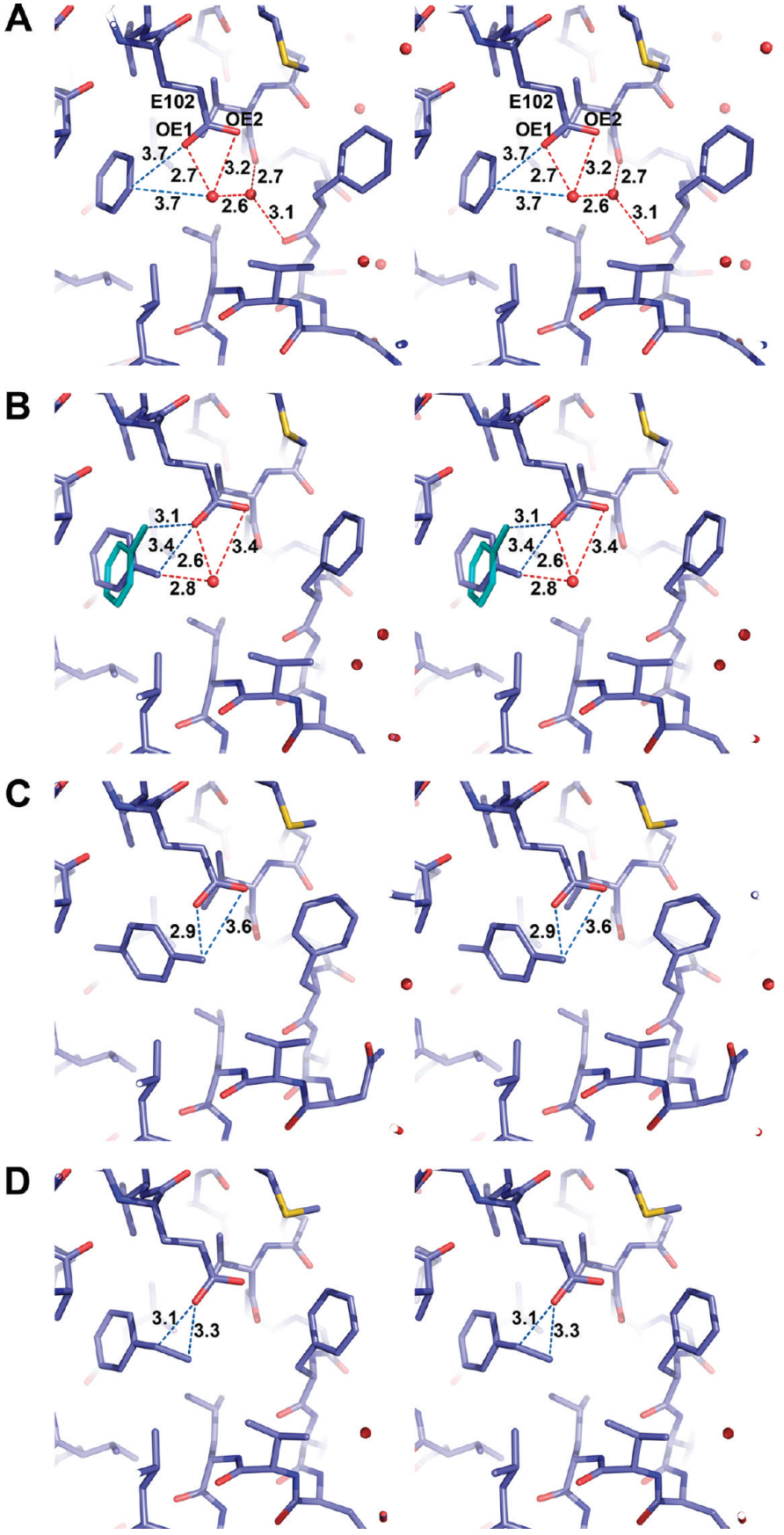

FIGURE 4: Representative nonpolar ligands bound within the L99A/M102E/St cavity. Solvent molecules are shown in red with hydrogen bonds as red dashed lines. Some key close approaches (non-H-bonding) are indicated by blue dashed lines. (A) Benzene. (B) Toluene. The higher occupancy (0.6) site of binding is shown in blue while the lower occupancy (0.4) site is in cyan. (C) *p*-Xylene. (D) Ethylbenzene.

(Figure 4A). Of the nonpolar ligands, *p*-xylene binds least well, no doubt because of steric clashes associated with its different shape. The difference in the binding of *p*-xylene relative to the other ligands is shown in Figure 5. Ethylbenzene binds almost as

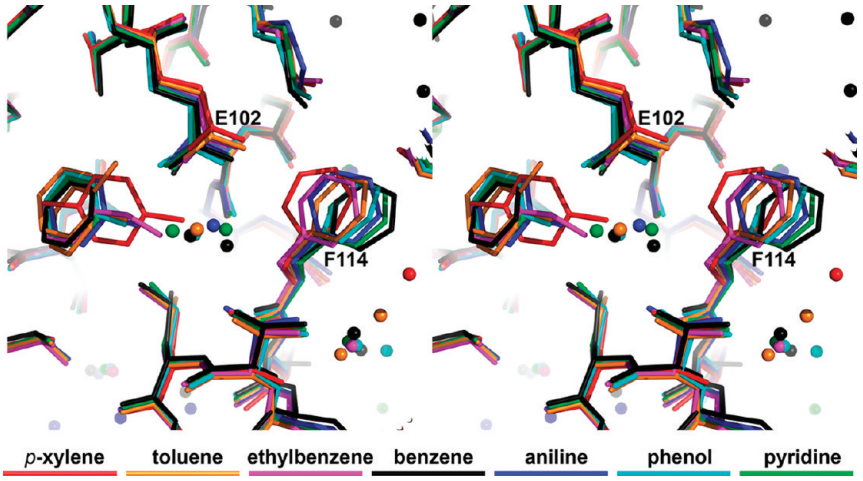

FIGURE 5: Comparison of different ligands bound within the L99A/M102E/St cavity. The alignment is based on the superposition of residues 70−125 of chain A, these being the most rigid part of the C-terminal domain of T4 lysozyme. Residues that undergo the largest conformational changes are labeled. The ligand and water molecules associated with a given structure are all shown in the same color, and the color coding is given at the bottom of the figure.

well as benzene. This presumably reflects a compromise. On one hand, its shape matches that of the cavity better than benzene. On the other hand, its binding necessitates the displacement of two solvent molecules from the cavity. For the four nonpolar ligands in Table 3 the affinities of binding to L99A/M102E/St (−5.8 to −4.9 kcal/mol) are comparable to those for L99A (−5.8 to −4.7 kcal/mol) (*21*). This suggests that there is relatively little energy cost for displacing the bound solvent molecules from the cavity and loss of the associated hydrogen bonds. In other words, the hydrogen-bond network in the polar part of the cavity does not outweigh in energy the binding of a highly hydrophobic ligand like ethylbenzene.

Insofar as measurements could be made, the polar ligands generally bind slightly less well than the nonpolar ones (Table 3). The contributions from enthalpy and entropy, however, are very different. For the nonpolar ligands the binding enthalpy is around −4 to −6 kcal/mol, which roughly corresponds to the overall free energy of binding. The binding reactions of these compounds are exothermic and largely enthalpically driven. As shown in Figure 4, the phenyl rings of benzene, toluene and ethylbenzene all bind in essentially the same location and all three ligands fit snugly in the cavity. In contrast, *p*-xylene fits less well and increases the mobility of some atoms in the cavity wall, which could explain the modest decrease in enthalpy and increase in entropy.

For the polar ligands entropic effects make a dominant contribution to binding. Presumably part of this entropy term is reflected in the stabilization of the hydrogen bond network. There could also be an entropy cost associated with the removal of the polar ligand from bulk solvent. Overall, the binding of the polar ligands in the L99A/M102E/St cavity is enthalpy driven.

As will be apparent from Figure 5, when the three larger nonpolar ligands, toluene, ethylbenzene and *p*-xylene, are bound, the side chain of Glu102 rotates away. In contrast, when the polar compounds are bound, the side chain of Glu102 moves closer. If the benzene-bound structure is taken as a reference, these shifts of Glu102 toward or away from the ligand range from 0.2 Å to 0.8 Å and have associated rotations of up to about 10°. There are also significant shifts in the position of Phe114. In cases where the ligand displaces all solvent from the cavity (e.g., *p*-xylene and ethylbenzene), the side chain of Phe114 moves toward the cavity to partially occupy this vacated space. This shows that the walls of the cavity have flexibility and helps explain the easy access of ligands and solvent into and out of the cavity.

ACKNOWLEDGMENT

We would like to thank Andy Fields for help in protein mutagenesis and purification and Drs. Corie Ralston, Nathan Smith and Jeff Dickert at the Advanced Light Source for their generous help in data collection.

REFERENCES

1. Chothia, C. (1974) Hydrophobic bonding and accessible surface area in proteins. *Nature 248*, 338–339.
2. Rashin, A. A., Iofin, M., and Honig, B. (1986) Internal cavities and buried waters in globular proteins. *Biochemistry 25*, 3619–3625.
3. Williams, M. A., Goodfellow, J. M., and Thornton, J. M. (1994) Buried waters and internal cavities in monomeric proteins. *Protein Sci. 3*, 1224–1235.
4. Buckle, A. M., Cramer, P., and Fersht, A. R. (1996) Structural and energetic responses to cavity-creating mutations in hydrophobic cores: observation of a buried water molecule and the hydrophilic nature of such hydrophobic cavities. *Biochemistry 35*, 4298–4305.
5. Matthews, B. W., and Liu, L. (2009) A review about nothing: Are apolar cavities in proteins really empty? *Protein Sci. 18*, 494–502.
6. Barlow, D. J., and Thornton, J. M. (1983) Ion-pairs in proteins. *J. Mol. Biol. 168*, 867–885.
7. Giletto, A., and Pace, C. N. (1999) Buried, charged, non ion paired aspartic acid 76 contributes favorably to the conformational stability of ribonuclease T1. *Biochemistry 38*, 13379–13384.
8. Xie, A., Kelemen, L., Hendriks, J., White, B. J., Hellingwerf, K. J., and Hoff, W. D. (2001) Formation of a new buried charge drives a large amplitude protein quake in photoreceptor activation. *Biochemistry 40*, 1510–1517.
9. Honig, B., Ebrey, T., Callender, R. H., Dinur, U., and Ottolenghi, M. (1979) Photoisomerization, energy storage, and charge separation: a model for light energy transduction in visual pigments and bacteriorhodopsin. *Proc. Natl. Acad. Sci. U.S.A. 76*, 2503–2507.
10. Makhatadze, G. I., and Privalov, P. L. (1995) Energetics of protein structure. *Adv. Protein Chem. 47*, 307–425.
11. Baldwin, E. P., Hajiseyedjavadi, O., Baase, W. A., and Matthews, B. W. (1993) The role of backbone flexibility in the accommodation of variants that repack the core of T4 lysozyme. *Science 262*, 1715–1718.
12. Eriksson, A. E., Baase, W. A., Zhang, X. J., Heinz, D. W., Blaber, M., Baldwin, E. P., and Matthews, B. W. (1992) Response of a protein

structure to cavity creating mutations and its relation to the hydrophobic effect. *Science 255*, 178–183.
13. Lim, W. A., Farruggio, D. C., and Sauer, R. T. (1992) Structural and energetic consequences of disruptive mutations in a protein core. *Biochemistry 31*, 4324–4333.
14. Green, S. M., Meeker, A. K., and Shortle, D. (1992) Contributions of the polar, uncharged amino acids to the stability of staphylococcal nuclease - evidence for mutational effects on the free energy of the denatured state. *Biochemistry 31*, 5717–5728.
15. Daopin, S., Anderson, D. E., Baase, W. A., Dahlquist, F. W., and Matthews, B. W. (1991) Structural and thermodynamic consequences of burying a charged residue within the hydrophobic core of T4 lysozyme. *Biochemistry 30*, 11521–11529.
16. Blaber, M., Lindstrom, J. D., Gassner, N., Xu, J., Heinz, D., and Matthews, B. W. (1993) Energetic cost and structural consequences of burying a hydroxyl group within the core of a protein determined from Ala→Ser and Val→Thr substitutions in T4 lysozyme. *Biochemistry 32*, 11363–11373.
17. Dwyer, J. J., Gittis, A. G., Karp, D. A., Lattman, E. E., Spencer, D. S., Stites, W. E., and Garcia-Moreno E., B. (2000) High apparent dielectric constants in the interior of a protein reflect water penetration. *Biophys. J. 79*, 1610–1620.
18. Nguyen, D. M., Reynald, L. R., Gittis, A. G., and Lattman, E. E. (2004) X ray and thermodynamic studies of staphylococcal nuclease variants I92E and I92K: insights into polarity of the protein interior. *J. Mol. Biol. 341*, 565–574.
19. Schlessman, J. L., Abe, C., Gittis, A., Karp, D. A., Dolan, M. A., and Garcia Moreno E., B. (2008) Crystallographic study of hydration of an internal cavity in engineered proteins with buried polar or ionizable groups. *Biophys. J. 94*, 3208–3216.
20. Eriksson, A. E., Baase, W. A., Wozniak, J. A., and Matthews, B. W. (1992) A cavity containing mutant of T4 lysozyme is stabilized by buried benzene. *Nature 355*, 371–373.
21. Morton, A., Baase, W. A., and Matthews, B. W. (1995) Energetic origins of specificity of ligand binding in an interior nonpolar cavity of T4 lysozyme. *Biochemistry 34*, 8564–8575.
22. Morton, A., and Matthews, B. W. (1995) Specificity of ligand binding in a buried nonpolar cavity of T4 lysozyme linkage of dynamics and structural plasticity. *Biochemistry 34*, 8576–8588.
23. Wei, B. Q., Baase, W. A., Weaver, L. H., Matthews, B. W., and Shoichet, B. K. (2002) A model binding site for testing scoring functions in molecular docking. *J. Mol. Biol. 322*, 339–355.
24. Matsumura, M., and Matthews, B. W. (1989) Control of enzyme activity by an engineered disulfide bond. *Science 243*, 792–794.
25. Eriksson, A. E., Baase, W. A., and Matthews, B. W. (1993) Similar hydrophobic replacements of Leu99 and Phe153 within the core of T4 lysozyme have different structural and thermodynamic consequences. *J. Mol. Biol. 229*, 747–769.
26. Poteete, A. R., Sun, D. P., Nicholson, H., and Matthews, B. W. (1991) Second site revertants of an inactive T4 lysozyme mutant restore activity by restructuring the active site cleft. *Biochemistry 30*, 1425–1432.
27. Otwinowski, Z., and Minor, W. (1997) Processing of X ray diffraction data collected in oscillation mode. *Methods Enzymol. 276*, 307–326.
28. Collaborative Computational Project, Number 4. (1994) The CCP4 Suite: Programs for Protein Crystallography. Acta Crystallogr. D50, 760−763.
29. Emsley, P., and Cowtan, K. (2004) Coot: model building tools for molecular graphics. *Acta Crystallogr. D60*, 2126–2132.
30. Elwell, M .L., and Schellman, J. A. (1977) Stability of phage T4 lysozymes: I. Native properties and thermal stability of wild-type and two mutant lysozymes. *Biochim. Biophys. Acta 494*, 367–383.
31. Zhang, X.-J., Baase, W. A., Shoichet, B. K., Wilson, K. P., and Matthews, B. W. (1995) Enhancement of protein stability by the combination of point mutations in T4 lysozyme is additive. *Protein Eng. 8*, 1017–1022.
32. Shoichet, B. K., Baase, W. A., Kuroki, R., and Matthews, B. W. (1995) A relationship between protein stability and protein function. *Proc. Natl. Acad. Sci. U.S.A. 92*, 452–456.
33. Wray, J. W., Baase, W. A., Lindstrom, J. D., Weaver, L. H., Poteete, A. R., and Matthews, B. W. (1999) Structural analysis of a non-contiguous second-site revertant in T4 lysozyme shows that increasing the rigidity of a protein can enhance its stability. *J. Mol. Biol. 292*, 1111–1120.
34. Matsumura, M., Signor, G., and Matthews, B. W. (1989) Substantial increase of protein stability by multiple disulphide bonds. *Nature 342*, 291–293.
35. Liu, L., Quillin, M. L., and Matthews, B. W. (2008) Use of experimental crystallographic phases to examine the hydration of polar and nonpolar cavities in T4 lysozyme. *Proc. Natl. Acad. Sci. U.S.A. 105*, 14406–14411.
36. Brenk, R., Vetter, S. W., Boyce, S. E., Goodin, D. B., and Shoichet, B. K. (2006) Probing molecular docking in a charged model binding site. *J. Mol. Biol. 357*, 1449–1470.
37. CRC Handbook of Chemistry and Physics (1996) 77th ed., CRC Press, Boca Raton.
38. DeLano, W. L. (2002) The PyMOL Molecular Graphics System, on the World Wide Web, http://www.pymol.org.