Quantifying Protein Microstructure and Electrostatic Effects on the Change in Gibbs Free Energy of Binding in Immobilized Metal Affinity Chromatography

ARTICLE *in* ANALYTICAL CHEMISTRY · APRIL 2008

Impact Factor: 5.64 · DOI: 10.1021/ac7023188 · Source: PubMed

3 AUTHORS, INCLUDING:

Chenming Zhang

Virginia Polytechnic Institute and State Univ…

59 PUBLICATIONS **1,048** CITATIONS

SEE PROFILE

Anal. Chem. **2008,** *80,* 1628−1640

Quantifying Protein Microstructure and Electrostatic Effects on the Change in Gibbs Free Energy of Binding in Immobilized Metal Affinity Chromatography

Lakshmi P Pathange,[†,‡] David R Bevan,[§] and Chenming Zhang*[,†]

Departments of Biological Systems Engineering and Biochemistry, Virginia Polytechnic Institute and State University, Blacksburg, Virginia 24061

Electrostatic forces play a major role in maintaining both structural and functional properties of proteins. A major component of protein electrostatics is the interactions between the charged or titratable amino acid residues (e.g., Glu, Lys, and His), whose pK_a (or the change of the pK_a) value could be used to study protein electrostatics. Here, we report the study of electrostatic forces through experiments using a well-controlled model protein (T4 lysozyme) and its variants. We generated 10 T4 lysozyme variants, in which the electrostatic environment of the histidine residue was perturbed by altering charged and neutral amino acid residues at various distances from the histidine (probe) residue. The electrostatic perturbations were theoretically quantified by calculating the change in free energy ($\Delta\Delta G_E$) using Coulomb's law. On the other hand, immobilized metal affinity chromatography (IMAC) was used to quantify these perturbations in terms of protein binding strength or change in free energy of binding ($\Delta\Delta G_B$), which varies from −0.53 to 0.99 kcal/mol. For most of the variants, there is a good correlation ($R^2 = 0.97$) between the theoretical $\Delta\Delta G_E$ and experimental $\Delta\Delta G_B$ values. However, there are three deviant variants, whose histidine residue was found to be involved in site-specific interactions (e.g., ion pair and steric hindrance), which were further investigated by molecular dynamics simulation. This report demonstrates that the electrostatic ($\Delta\Delta G_{Elec}$) and microstructural effects ($\Delta\Delta G_{Micro}$) in a protein can be quantified by IMAC through surface histidine mediated protein−metal ion interaction and that the unique microstructure around a histidine residue can be identified by identifying the abnormal binding behaviors during IMAC.

In general, almost all important biochemical processes in living systems require proteins. Proteins are large, heterogeneous, flexible, and complex molecules.[1−4] It is believed that there is an intricate relationship between protein structure and function. Consequently any subtle change that might occur in a protein's structure can affect its function, which is generally modulated by the amino acid residues on its surface.[5] One way to characterize the role played by protein surfaces is by calculating its electrostatic energy (or potential).[5] In fact, electrostatics[6] play a major role in maintaining both the structural and functional properties of proteins.[7−11] For example, molecular level interactions such as salt bridges and hydrogen bonds, which are primarily electrostatic in nature,[2,4,12] are critical not only in maintaining the structure of a protein but also in defining its specific function by stabilizing the charged transition states during catalytic reactions.[13−15] Therefore, it is essential to understand the role played by electrostatic forces to decipher protein structure−function relationships.[9,16]

Analyzing electrostatic forces not only elucidates structure−function relationships but also aids in engineering proteins rationally.[6] However, theoretical assessment of electrostatic forces has been challenging due to the inability to extend current models to analyze the heterogeneous nature of protein structures.[7,17−19]

* Corresponding author. Dr. Chenming (Mike) Zhang. E-mail: cmzhang@vt.edu. Phone: (540) 231-7601. Fax: (540) 231-3199.

† Department of Biological Systems Engineering.

‡ Current address: Bayer Healthcare, 800 Dwight Way, Berkeley, CA 94710.

§ Department of Biochemistry.

(1) Teague, S. J. *Nat. Rev. Drug Discovery* **2003**, *2*, 527−541.
(2) Kumar, S.; Nussinov, R. *Biophys. J.* **2002**, *83*, 1595−1612.
(3) Schellekens, H. *Nat. Biotechnol.* **2004**, *22*, 1357−1359.
(4) Kumar, S.; Nussinov, R. *ChemBioChem* **2002**, *3*, 604−617.
(5) Sivasankar, S.; Gumbiner, B.; Leckband, D. *Biophys. J.* **2001**, *80*, 1758−1768.
(6) Gitlin, I.; Carbeck, J. D.; Whitesides, G. M. *Angew. Chem., Int. Ed.* **2006**, *45*, 3022−3060.
(7) Russell, A. J.; Thomas, P. G.; Fersht, A. R. *J. Mol. Biol.* **1987**, *193*, 803−813.
(8) Louro, R. O.; Catarino, T.; Paquete, C. M.; Turner, D. L. *FEBS Lett.* **2004**, *576*, 77−80.
(9) Lee, K. K.; Fitch, C. A.; Garcia-Moreno, B. *Protein Sci.* **2002**, *11*, 1004−1016.
(10) Loewenthal, R.; Sancho, J.; Reinikainen, T.; Fersht, A. R. *J. Mol. Biol.* **1993**, *232*, 574−583.
(11) Allewell, N. M.; Oberoi, H. *Methods Enzymol.* **1991**, *202*, 3−19.
(12) Jeffrey, G. A. *Food Chem.* **1996**, *56*, 241−246.
(13) Thomas, P. G.; Russell, A. J.; Fersht, A. R. *Nature* **1985**, *318*, 375−376.
(14) Jackson, S. E.; Fersht, A. R. *Biochemistry* **1993**, *32*, 13909−13916.
(15) Laurents, D. V.; Huyghues-Despointes, B. M. P.; Bruix, M.; Thurlkill, R. L.; Schell, D.; Newsom, S.; Grimsley, G. R.; Shaw, K. L.; Trevino, S.; Rico, M.; Briggs, J. M.; Antosiewicz, J. M.; Scholtz, J. M.; Pace, C. N. *J. Mol. Biol.* **2003**, *325*, 1077−1092.
(16) Warshel, A. *Biochemistry* **1981**, *20*, 3167−3177.
(17) Russell, A. J.; Fersht, A. R. *Nature* **1987**, *328*, 496−500.
(18) de Kreij, A.; van den Burg, B.; Venema, G.; Vriend, G.; Eijsink, V. G. H.; Nielsen, J. E. *J. Biol. Chem.* **2002**, *277*, 15432−15438.
(19) Schutz, C. N.; Warshel, A. *Proteins: Struct., Funct., Genet.* **2001**, *44*, 400−417.

10.1021/ac7023188 CCC: $40.75
Published on Web 01/30/2008

Therefore a combined strategy of utilizing both theory (to predict) and experimentation (to verify) is required to understand protein electrostatics.[6,20−22]

One way to study the effects of electrostatic forces is by measuring the pK_a value of the titratable amino acid residues.[9,19,21] Because of environmental effects in structured proteins, the titratable amino acid residues (here histidine) display different pK_a values from those in peptides or free amino acids.[6,9,15,21,23−27] The environmental factors that influence the pK_a values of histidine are (1) their surface accessibility (hydrophobic/hydrophilic effect),[6,21,26,28] (2) their microenvironment[29] (which involves (a) the involvement of histidine residues in intramolecular interactions such as salt bridges and hydrogen bonds,[21,26,30−32] (b) histidine's microstructure which consists of short-range charge−charge interactions that depend on the local zonal amino (polar/nonpolar) acid distribution,[33] and (c) long-range charge−charge interactions and overall protein net charge[7,14,15,26,32]), (3) the type, pH, and ionic strength of the buffer solution,[6,9,26] and (4) other factors, which are generally neglected, such as histidine residue's conformational effects[7,34] including its direction[33] and its side chain flexibility.[35,36] Understanding and quantifying the effects of these factors on the pK_a of the probe-histidine (a histidine residue directly participates in protein−metal ion interaction) would allow possible inference of the microenvironment around the histidine in proteins, which is otherwise difficult to obtain by other analytical methods.

Currently, the experimental techniques that are used to measure pK_a values of the titratable groups are NMR and FT-IR.[21,31] However, even though these techniques are reliable, they are costly and complicated. Other simpler and indirect techniques such as kinetic analysis using pH dependent catalytic activity have been used to measure the changes in pK_a (ΔpK_a) values.[10,17,21] For example, Fersht and co-workers used enzyme activity studies to calculate the ΔpK_a values of histidine residues in subtilisin variants and to quantify the electrostatic effects due to charge variation.[10,17] However, these indirect techniques cannot be extended to study the pK_a values of the titratable groups other than those involved in the catalytic process. Also, it has been reported that in a few cases the changes in the catalytic site are so complex that straightforward site-directed mutagenesis may not yield desired results.[31] Such problems can be addressed by utilizing techniques wherein the histidine group is not involved in the catalytic process. For example, Chicz and Regnier used site-directed protein mutagenesis to induce charge variation in the microenvironment of His-64 in subtilisin to illustrate that immobilized metal affinity chromatography (IMAC)[37−39] can be used to qualitatively distinguish these subtilisin variants via the differences in their retention time.[24] Other studies also demonstrated that histidine's microenvironment influences the binding behavior of the proteins.[40,41] However, because of the intrinsic complexity of a protein structure, so far there are no attempts to quantify the microenvironmental contributions to a protein's interaction with metal ions.

Previously we were able to establish a correlation between the protein−metal ion binding strength (K_{aP}) and the histidine surface accessibility.[29] In addition, we were able to demonstrate quantitatively the effect of intramolecular bonds, which is one of the factors comprising the histidine's microenvironment, through the protein−metal ion interaction in IMAC when there is only one histidine on the surface of a protein.[29] Our current goal is to understand the effect of the other factor, namely, the electrostatic environment, on the histidine's electrostatic property (such as its pK_a) that is manifested through the protein's binding strength on IMAC.[29] Our hypothesis is that any subtle variation in the histidine's electrostatic environment affects its binding strength values via a change in its pK_a.

More specifically, the overall goal of our present work is to quantify the effects of electrostatic perturbation (charge variation), via change in Gibbs free energy of binding, in the probe-histidine's electrostatic environment. Fersht and co-workers quantified long-range electrostatic perturbations (>12 Å) using enzyme kinetics.[10] Such long-range studies elucidate only charge effects but not any associated structural effects, unless the perturbed amino acid is involved in a network contributing to the catalytic activity. In this report, all the sites chosen for charge modification around the histidine residue were within short or medium range (<15 Å). A major advantage in perturbing short-range residues is that they not only introduce strong electrostatic effects due to the proximity of the charge variation but also may impart site-specific microstructural effects on the to-be probed histidine residue.[42]

Theoretically, the electrostatic perturbation was quantified by calculating the change in net electrostatic energy (ΔE), which represents the cumulative effect of both charge and distance of the perturbation from the histidine. The ΔE values were used to calculate the change in Gibbs free energy due to the change in net electrostatic energy($\Delta\Delta G_E$). Experimentally, IMAC was used as a quantitative tool to obtain protein binding strength.[29] The

(20) Alvaro, G.; Russell, A. J. *Methods Enzymol.* **1991**, *202*, 620−643.
(21) Joshi, M. D.; Sidhu, G.; Nielsen, J. E.; Brayer, G. D.; Withers, S. G.; McIntosh, L. P. *Biochemistry* **2001**, *40*, 10115−10139.
(22) Sundd, M.; Iverson, N.; Ibarra-Molero, B.; Sanchez-Ruiz, J. M.; Robertson, A. D. *Biochemistry* **2002**, *41*, 7586−7596.
(23) Mehler, E. L.; Fuxreiter, M.; Simon, I.; Garcia-Moreno, E. B. *Proteins: Struct., Funct., Genet.* **2002**, *48*, 283−292.
(24) Chicz, R. M.; Regnier, F. E. *Anal. Chem.* **1989**, *61*, 1742−1749.
(25) Chicz, R. M.; Regnier, F. E. *J. Chromatogr., A* **1990**, *500*, 503−518.
(26) Huyghues-Despointes, B. M. P.; Thurlkill, R. L.; Daily, M. D.; Schell, D.; Briggs, J. M.; Antosiewicz, J. M.; Pace, C. N.; Scholtz, J. M. *J. Mol. Biol.* **2003**, *325*, 1093−1105.
(27) Mehler, E. L.; Eichele, G. *Biochemistry* **1984**, *23*, 3887−3891.
(28) Edgcomb, S. P.; Murphy, K. P. *Proteins: Struct., Funct., Genet.* **2002**, *49*, 1−6.
(29) Pathange, L. P.; Bevan, D. R.; Larson, T. J.; Zhang, C. M. *Anal. Chem.* **2006**, *78*, 4443−4449.
(30) Bastyns, K.; Froeyen, M.; Diaz, J. F.; Volckaert, G.; Engelborghs, Y. *Proteins: Struct., Funct., Genet.* **1996**, *24*, 370−378.
(31) Garcia-Mayoral, M. F.; Perez-Canadillas, J. M.; Santoro, J.; Ibarra-Molero, B.; Sanchez-Ruiz, J. M.; Lacadena, J.; del Pozo, L. M.; Gavilanes, J. G.; Rico, M.; Bruix, M. *Biochemistry* **2003**, *42*, 13122−13133.
(32) Li, J. W. *Chromatographia* **2005**, *61*, 479−492.
(33) Mehler, E. L.; Guarnieri, F. *Biophys. J.* **1999**, *77*, 3−22.
(34) Georgescu, R. E.; Alexov, E. G.; Gunner, M. R. *Biophys. J.* **2002**, *83*, 1731−1748.
(35) Mehler, E. L. *J. Phys. Chem.* **1996**, *100*, 16006−16018.
(36) Whitten, S. T.; Garcia-Moreno, B.; Hilser, V. J. *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 4282−4287.
(37) Porath, J.; Carlsson, J.; Olsson, I.; Belfrage, G. *Nature* **1975**, *258*, 598−599.
(38) Porath, J. *TrAC, Trends Anal. Chem.* **1988**, *7*, 254−259.
(39) Porath, J. *Protein Expression Purif.* **1992**, *3*, 263−281.
(40) Berna, P. P.; Mrabet, N. T.; VanBeeumen, J.; Devreese, B.; Porath, J.; Vijayalakshmi, M. A. *Biochemistry* **1997**, *36*, 6896−6905.
(41) Zhao, Y. J.; Sulkowski, E.; Porath, J. *Eur. J. Biochem.* **1991**, *202*, 1115−1119.
(42) Liu, R. J.; Baase, W. A.; Matthews, B. W. *J. Mol. Biol.* **2000**, *295*, 127−145.

protein binding strength values were used to calculate the equivalent ΔpK_a and change in Gibbs free energy of binding due to electrostatic pertubation ($\Delta\Delta G_B$). NMR spectroscopy was also used to measure the ΔpK_a of the histidine.[9,31,43,44] The specific objectives of this article are (1) to obtain a possible correlation between ($\Delta\Delta G_E$) and($\Delta\Delta G_B$), and to use this correlation as a gauge to discover possible intramolecular interactions involving the binding histidine and the microstructure around it, and (2) to dissect the contribution of microstructural and electrostatic effects to ΔpK_a and $\Delta\Delta G_B$ values.

MATERIALS AND METHODS

Site Selection. Previously, we generated two variants, K135H and R80H, from cysteine-free T4 lysozyme.[29] For this report, these two variants were treated as pseudo wild-types (pWT-135 and pWT-80). Two sets of T4-lysozyme variants were generated by performing site-directed protein mutagenesis on the pWT proteins. The first set of variants has a histidine residue at position 135 (pWT-135), while the second set of variants has a histidine residue at position 80 (pWT-80). The two sites His135 and His80 were chosen because they have similar surface accessibility with high binding strength on IMAC−Cu^{2+} columns.[29] Three sites Asp-127, Ser-136, and Arg-137 around His135 (Figure 1a) and two sites Arg-76 and Asp-108 around His80 (Figure 1b) were chosen for site-directed mutagenesis (Table 1), and the resulting lysozyme variants vary both in charge and in distance from the histidine residue. The distances reported are the $C_\alpha-C_\alpha$ of original amino acids and the histidine. Instead of the side chain atoms, C_α atoms were chosen because the orientation of the side chain in replaced amino acids is hard to predict using molecular modeling tools.[45] All these sites were mutated to two different residues except for site Ser-136 (three mutations) and Glu-108 (one mutation) (Table 1). As a result, 10 variants were generated from 5 sites, and all variants have a single mutation around the sole histidine residue (Table 1).

In Silico Molecular Modeling Methods. All lysozyme varaints were also generated in silico using SYBYL 7.1 software (Tripos Inc., St. Louis, MO). In general, the orientation of the replacing residue is similar to the original residue in the wild-type, unless some specific steric constraints are present in the microstructure of the replacing residues.[46,47] All variant structures were energy-minimized using the SYBYL program.[29] The surface accessibility of all the sites was obtained by using a probe with a radius of 1.4 Å (Table 2). To visualize and generate figures, the UCSF Chimera and Swiss-Pdb Viewer programs were utilized.[48,49]

Characterization of Protein Microstructure. The relative hydrophobicity/hydrophilicity value (rHy), which characterizes the microenvironment around all the sites, was calculated using a modification of a published method.[23,33] The sum of all the hydrophobic/hydrophilic constants (sHy) of the moieties around these sites (4.25 Å) was calculated manually from the X-ray crystal

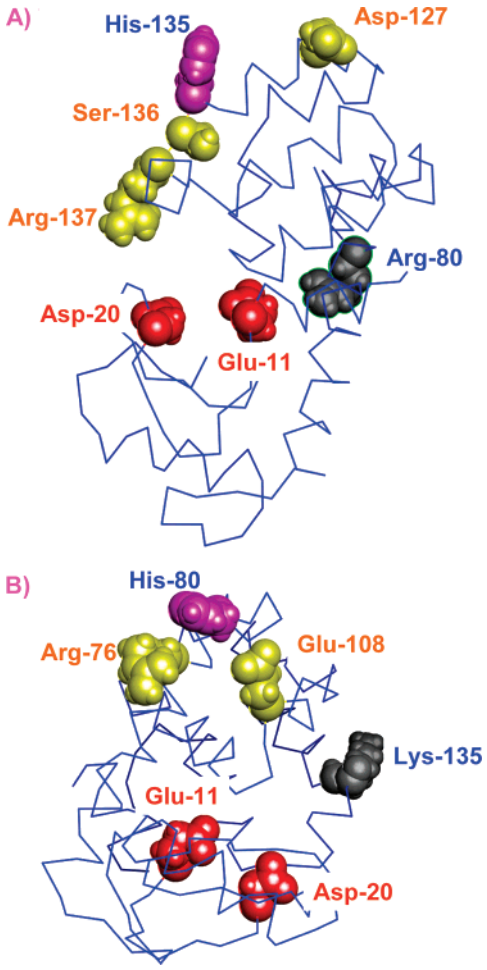

Figure 1. (A and B) The sites shown in red are the residues involved in catalysis. (A) Structure of pWT-135. The three sites (yellow) chosen around the histidine residue, His135 (magenta). Site 80 is shown in gray. (B) Structure of pWT-80. The two sites (yellow) chosen around the histidine residue, His80 (magenta). Site 135 is shown in gray.

Table 1. The Change in Net Charge and the Distance from the Histidine Residue in Both Control as Well as Test Site, for All the 10 Lysozyme Variants. The pWT-135 and the pWT-80 Represent the Pseudo Wild-Type Lysozymes with Histidine at Site 135 and Site 80, Respectively.

protein variant	change in net charge (Units)	$C_\alpha-C_\alpha$ distance between the site and the histidine residue (Å)
control variant (pWT-135)	0	N/A
D127S	+1	12.96
D127K	+2	12.96
S136D	−1	3.83
S136K	+1	3.83
S136M	−0.15	3.83
R137S	−1	7.35
R137D	−2	7.35
control variant (pWT-80)	0	N/A
R76S	−1	5.97
R76D	−2	5.97
E108K	+2	9.62

(43) Daopin, S.; Sauer, U.; Nicholson, H.; Matthews, B. W. *Biochemistry* **1991**, *30*, 7142−7153.
(44) Anderson, D. E.; Becktel, W. J.; Dahlquist, F. W. *Biochemistry* **1990**, *29*, 2403−2408.
(45) Najmanovich, R.; Kuttner, J.; Sobolev, V.; Edelman, M. *Proteins: Struct., Funct., Genet.* **2000**, *39*, 261−268.
(46) Baldwin, E. P.; Hajiseyedjavadi, O.; Baase, W. A.; Matthews, B. W. *Science* **1993**, *262*, 1715−1718.
(47) *SYBYL*, version 7.1; Tripos Associates Inc.: St. Louis, MO.
(48) Pettersen, E. F.; Goddard, T. D.; Huang, C. C.; Couch, G. S.; Greenblatt, D. M.; Meng, E. C.; Ferrin, T. E. *J. Comput. Chem.* **2004**, *25*, 1605−1612.
(49) Guex, N.; Peitsch, M. C. *Electrophoresis* **1997**, *18*, 2714−2723.

Table 2. Relative Surface Accessibility[a] and Relative Hydrophilicity Values (rHy)[b] of All the Five Sites Chosen for Single Point Mutations

sites	relative surface accessibility	relative hydrophilicity (rHy)
Arg76	0.60	0.84
Glu108	0.42	0.55
Asp127	1.00	1.02
Ser136[c]	0.10	0.70
Arg137	0.63	0.78

[a] The relative surface area of each amino acid side was directly obtained from the SYBYL program. [b] Microenvironments with rHy values (1) >0.25 are fairly hydrophilic, (2) <0.25 are hydrophobic, and (3) <0.0 are extremely hydrophobic. [c] The aHy value of tyrosine was used to calculate the rHy of Ser136.

structure of the cysteine free T4 lysozyme (1L63).[33] The rHy values of all the sites were obtained from the following equations[33] and are reported in Table 2:

$$\mathrm{tHy} = (1 - \mathrm{SA})\mathrm{sHy} + (\mathrm{SA})\mathrm{aHy} \quad (1)$$

$$\mathrm{rHy} = \frac{\mathrm{tHy}}{\mathrm{aHy}} \quad (1a)$$

where, SA is surface accessibility, tHy is the total hydrophobicity, and aHy is the absolute hydrophilic value in water obtained from a method by Mehler et al.[33]

Molecular Dynamics Simulation. Molecular dynamics (MD) was used to demonstrate the flexibility of the side chain. MD simulations were done using AMBER8. Structures were prepared for simulation using Leap. Counterions (Cl^-) were added to neutralize the charge of the protein, and the protein was immersed in a box of TIP3P water, with no protein atom being closer than 9 Å to the edge of the water box. The system was equilibrated by running MD at 300 K of the water molecules and counterions for 100 ps, after which the energy of the system was minimized, first minimizing the counterions and water only, and then the entire system. MD of the entire system was begun by heating from 0 to 300 K over 30 ps and then equilibrating the system for an additional 50 ps, all under constant volume conditions. Constant pressure conditions were then imposed for the remainder of the simulation, which was carried out for a total of 3000 ps. MD and energy minimizations were done using the sander module of AMBER8 running on the System X supercomputer at Virginia Tech.

Experimental. The production and purification of all lysozyme variants with aforementioned mutations (Table 1) was performed following the published method.[29] Mutations in the T4 lysozyme gene were carried out by single mutagenic[50] and overlap extension[50−52] PCR techniques. The point mutation in the lysozyme gene for all 10 variants was verified by DNA sequencing. All lysozyme variants were overexpressed in *Escherichia coli* cells and purified using an XK 16/20 carboxyl methyl sepharose fast flow ion exchange column (GE Health Care, Piscataway, NJ).[29] The final protein purity of all variants was determined by SDS-PAGE (>95%). The protein sequence of all the variants was verified by trypsin digest followed by LC−MS, except for R76S, R137S, and E108K (Table 3).

Table 3. The Peptide Sequences of the Seven Lysozyme Variants Obtained from Electrospray Ionization Mass Spectrometry, the Peptide Sequences Were Identified by SEQUEST, the Histidine and the Replaced Residue (Point Mutation) Are Shown in Bold Letters

protein variant	peptide position	peptide sequence
D127S	126−137	W**K**EAAVNLA**H**SR
D127K	126−137	W**K**EAAVNLA**H**SR
S136D	125−137	RWDEAAVNLA**HD**R
S136K	126−137	RWDEAAVNLA**HK**R
S136M	125−137	RWDEAAVNLA**HM**R
R137D	125−145	RWDEAAVNLA**HSD**WYNQTPNR
R76D	61−83	DEAEKLFNQDVDAAV**D**GIL**H**NAK

Table 4. Enzymatic Activities of Purified T4 Lysozyme Variants, the Activities of All the Lysozyme Variants Are Compared with the Wild Type

protein variant	(%) activity
wild type	100
pWT-135	162
D127S	132
D127K	77
S136D	70
S136K	82
S136M	89
R137S	86
R137D	109
pWT-80	126
R76S	120
R76D	124
E108K	102

The final protein solution (in 20 mM sodium phosphate, pH 7.0, 0.5 M NaCl, and 0.01% sodium azide) used for conducting IMAC experiments was obtained by a buffer exchange step with Buffer B in an Econo-pac desalting column (Bio-Rad Laboratories, Hercules, CA). The absorbance value at 280 nm ($A^{1\mathrm{mg/mL}} = 1.28$) was used to determine the concentration of all lysozyme variants.[53] The enzyme activity of all 10 lysozyme variants was determined by following the published method.[29] Briefly, a lysozyme sample (1 μg) was added to a cell suspension (0.5 mg dry weight/mL) of *Micrococcus lysodeiktikus* in 20 mM sodium phosphate and 50 mM NaCl (pH 6.0). The enzyme activity is defined as the rate of clearance of *M. lysodeiktikus* cells, which was measured by the rate of change in absorbance at 540 nm (in a microplate reader, Bio-Tek Instruments Inc., Winooski, VT). The enzyme activities of all 10 variants and the wild-type lysozyme are reported in Table 4.

CD and NMR Spectroscopy. All CD experiments were conducted at room temperature using a J-720 spectropolarimeter (JASCO Incorporated, Easton, MD). The CD spectra were

(50) Qiu, D.; Shenkin, P. S.; Hollinger, F. P.; Still, W. C. *J. Phys. Chem. A* **1997**, *101*, 3005−3014.
(51) Higuchi, R., Ed. *Using PCR to Engineer DNA*; Stockton Press: New York, 1989.
(52) Yang, B.; Gerhardt, S. G.; Larson, T. J. *Mol. Microbiol.* **1997**, *24*, 511−521.
(53) Rennell, D.; Bouvier, S. E.; Hardy, L. W.; Poteete, A. R. *J. Mol. Biol.* **1991**, *222*, 67−87.

measured from wavelength 190−350 nm in quartz cuvettes of 1 mm path length. The protein concentration of all the samples was around 50 μg/mL. The sample buffer used was 20 mM sodium phosphate (pH 7.0).

For all NMR experiments, the sample buffer solution used was a D_2O solution containing 10 mM D_3PO_4 and 10 mM KCl. Protein samples were prepared using a modification of published methods.[54,55] The samples containing 6−10 mg of lysozyme were concentrated from 4 to 1 mL by centrifugation at 7500*g* using disposable Amicon ultra tubes (Millipore, Billerica, MA) having membranes with a 5 kDa molecular-weight-cutoff. Then 3 mL of D_2O was added and the above centrifugation step was repeated, followed by addition of 3 mL of sample buffer solution, and then the centrifugation step was repeated. This step was repeated three more times. Then the protein samples were heated at 55 °C for 30 min to undergo reversible denaturation.[54] The samples were then slowly cooled and stored at 4 °C. Two protein stock solutions were prepared by adding NaOD (base stock) and DCl (acid stock). Then by adding varying amounts of protein acid−base stock solution, NMR samples with varying pH values (5.5−8.5) were generated. The pH of all the samples was measured before and after the NMR experiments. The maximum variation in the pH values was ±0.03. All NMR experiments were conducted at room temperature using a 400 MHz NMR spectrometer.

Determination of Protein Binding Strength (K_{aP}). IMAC experiments were conducted at room temperature on a TSK gel column (Tosoh Bioscience, Philadelphia, PA) mounted onto an HPLC system (HP 1050). The binding strength value of all the lysozyme variants (including the control variants) in the IMAC column was determined using an isocratic elution method known as zonal analysis.[56]

By varying the inhibitor (imidazole) concentration in the protein sample and the mobile phase, the association constant was determined using the following equation:[56]

$$\frac{1}{V - V_0} = \frac{1}{K_{aP}(V_0 - V_m)B_t} + \frac{K_{aI}[I]}{K_{aP}(V_0 - V_m)B_t} \qquad (2)$$

where V = elution volume to elute half of the protein (peak maximum), V_0 = elution volume in the absence of interaction (0.79 mL), V_m = void volume in the column determined by blue dextran (0.44 mL), [I] = concentration of the inhibitor (imidazole) in the mobile phase, K_{aI} = the association constant of the imidazole−copper complex, K_{aP} = association constant for the interaction between the protein and the copper ion in IMAC (protein binding strength), and B_t = total immobilized copper ions (23 μmol/mL). The elution volume of the protein from the IMAC column varies with the inhibitor concentration [I]. Therefore, a plot between $1/(V - V_0)$ vs [I] would be linear. K_{aI} is equal to the ratio of the slope to the intercept, and $1/K_{aP}$ can be obtained from the intercept and the independently obtainable parameters V_0, V_m, and B_t.

Theoretical Components Contributing to Gibbs Free Energy of Binding in IMAC. The interaction or the binding phenomena between lysozyme and the immobilized copper ion in IMAC can be considered as a protein−ligand interaction.[57] The protein−ligand interaction is a thermodynamic process, which can be characterized by a change in Gibbs free energy of binding (ΔG_B). In general ΔG_B can be defined by the following equation,

$$\Delta G_B = \Delta H - T\Delta S \qquad (3)$$

where, ΔH is the change in enthalpy and ΔS is the change in entropy of the interaction. The specific components that contribute to ΔG_B (in eq 3) are given in eq 4.[58] On the right-hand side of eq 4, the first three ΔG terms contribute to ΔH, while the next two ΔG terms (inside the brackets) contribute to $T\Delta S$.

$$\Delta G_B = \Delta G_{Dir} + \Delta G_{Micro} + \Delta G_{SA} - (\Delta G_{Des} + \Delta G_{Con}) \qquad (4)$$

ΔG_{Dir} is the ΔG due to direct interactions between protein and ligand (copper ions here), which includes forces such as electrostatic, van der Waals, and intermolecular bonds (hydrogen bonds). ΔG_{Micro} is the ΔG due to the specific microenvironment of the amino acid residues (interactions that do not directly involve the reporter histidine) around the histidine group such as intramolecular interactions, hydrophobic interactions, and steric effects, which might enhance or hinder the binding process. All these microenvironment effects (hereafter these effects are specifically referred to as *microstructure or microstructural effects*) may also affect the pK_a of a reporter histidine and thus the protein−metal interaction. ΔG_{SA} is the ΔG due to the ability of the surface accessible side chain (of histidine residue including its flexibility and orientation) to interact with the immobilized ligand. ΔG_{Des} is the loss of free energy due to desolvation of the protein and ligand to release the solvent molecules. ΔG_{Con} is the loss of free energy due to the restrictions of protein and ligand conformations during the binding process. The first three terms in eq 4 can be further dissected as

$$\Delta G_{Dir} = \Delta G_{Elec} + \Delta G_{Vdw} + \Delta G_{Inter} \qquad (4a)$$

$$\Delta G_{Micro} = \Delta G_{Int} + \Delta G_{Hydro} + \Delta G_{Steric} \qquad (4b)$$

$$\Delta G_{SA} = \Delta G_{Ori} + \Delta G_{Flex} \qquad (4c)$$

In eq 4a, ΔG_{Elec} describes the charge perturbation around a histidine reporter group, and the perturbation directly affects the pK_a value of the histidine and thus the protein−metal ion interaction. It is one of the main parameters studied in this article. As stated above, K_{aP} is used to quantify the adsorption affinities of all variants in the IMAC column. Since the adsorption process is a thermodynamic equilibrium process, the relation between K_{aP} and $\Delta\Delta G_B$ (change in Gibbs free energy of binding due to electrostatic pertubation) can be derived, and their relationship

(54) Barbas, C. F.; Heine, A.; Zhong, G. F.; Hoffmann, T.; Gramatikova, S.; Bjornestedt, R.; List, B.; Anderson, J.; Stura, E. A.; Wilson, I. A.; Lerner, R. A. *Science* **1997**, *278*, 2085−2092.
(55) Tishmack, P. A.; Bashford, D.; Harms, E.; VanEtten, R. L. *Biochemistry* **1997**, *36*, 11984−11994.
(56) Dunn, B. M.; Chaiken, I. M. *Proc. Natl. Acad. Sci. U.S.A.* **1974**, *71*, 2382−2385.
(57) Finette, G. M. S.; Mao, Q. M.; Hearn, M. T. W. *J. Chromatogr., A* **1997**, *763*, 71−90.
(58) Lazaridis, T. *Curr. Org. Chem.* **2002**, *6*, 1319−1332.

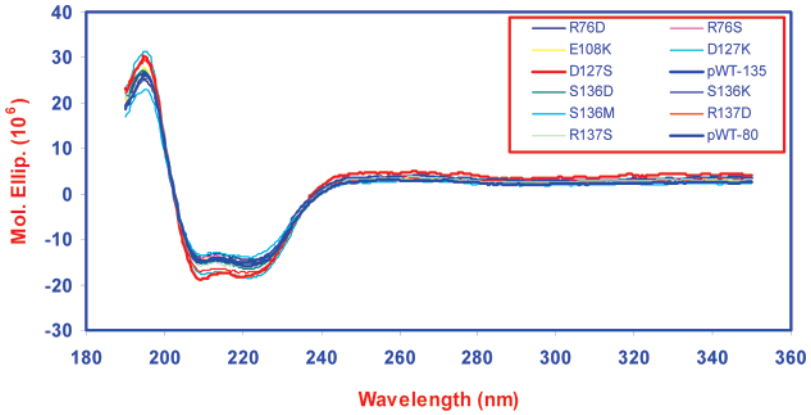

Figure 2. CD spectrum (wavelength vs molecular ellipticity) for all of the 10 variants and both the pseudo wild type lysozyme proteins, pWT-135 and pWT-80.

Table 5. Calculated $\Delta\Delta G_E$ values from Coulomb's Law, IMAC Protein Binding Strength (K_{aP}) Values, and Experimental ΔpK_a and $\Delta\Delta G_B$ Values Obtained from the K_{aP} Values

protein variant (change in net charge)	$\Delta\Delta G_E$ (kcal/mol)	protein binding strength (K_{aPv}) (10^1 M^{-1})	change in pK_a	$\Delta\Delta G_B$ (kcal/mol)
control variant (His135)	0	92.1	0	0
D127S (+1)	−0.29	163.4	−0.25	−0.34
D127K (+2)	−0.58	138.6	−0.17	−0.24
S136D (−1)	0.98	54.8	0.22	0.31
S136K (+1)	−0.98	63.4	0.16	0.22
S136M (−0.15)	0.15	70.5	0.12	0.16
R137S (−1)	0.51	50.6	0.26	0.35
R137D (−2)	1.02	34.8	0.42	0.58
control variant (His80)	0	106.7	0	0
R76S (−1)	0.63	48.1	0.34	0.47
R76D (−2)	1.25	20.1	0.73	0.99
E108K (+2)	−0.78	261.8	−0.39	−0.53

with the change of pK_a (ΔpK_a) can be described by the following equation:[10]

$$1.364\Delta pK_a = -RT\ln\frac{K_{aPv}}{K_{aPc}} = \Delta\Delta G_B \quad (5)$$

wherein, K_{aPc} is the protein binding strength of the control variants (pWT-135 and pWT-80) and K_{aPv} is the protein binding strength of all variants in which there is a perturbation in the histidine microenvironment.

Components Contributing to ΔG_B Due to Perturbation ($\Delta\Delta G_B$). In the present set of IMAC experiments, the major factors that affect the formation of the coordinate covalent bond between the sole histidine residue on lysozyme and the copper ion are (1) the histidine's pK_a, which is dependent on all the aforementioned factors listed in the introduction section, (2) surface accessibility, flexibility, and orientation of the histidine's side chain, (3) histidine's microenvironment, and (4) flexibility of the immobilized ligand. Here, all variants have similar, if not identical, histidine surface accessibility, combining with the fact that the changes of ΔG_{Des} and ΔG_{Con} are negligible due to the way the variants were generated, $\Delta\Delta G_B$ can be determined by the following equations for a particular variant before and after perturbations take place:

$$\Delta\Delta G_B = \Delta G_{Dir1} + \Delta G_{Micro1} - (\Delta G_{Dir2} + \Delta G_{Micro2}) \quad (6)$$

Assuming the values of ΔG_{Vdw} and ΔG_{Inter} are negligible for all variants (for IMAC experiments, this is especially true) (as in eq 4a), the following equations can be obtained,

$$\Delta\Delta G_B = \Delta\Delta G_{Elec} + \Delta\Delta G_{Micro} \quad (7)$$

Thus, when there are no significant microstructural effects ($\Delta\Delta G_{Micro} = 0$) around the histidine group, $\Delta\Delta G_B = \Delta\Delta G_{Elec}$. Under this circumstance, eq 5 can be extended to the following equation,

$$1.364\Delta pK_a = -RT\ln\frac{K_{aPv}}{K_{aPc}} = \Delta\Delta G_B = \Delta\Delta G_{Elec} \quad (5a)$$

Moreover, if the $\Delta\Delta G_{Elec}$ value is known, then $\Delta\Delta G_{Micro}$ value can be calculated from eq 6 for those variants in which there is a significant structural contribution to the capability of binding of the histidine residue. Then from the obtained $\Delta\Delta G_{Micro}$ value, with the assistance of molecular modeling, it is possible to discover the specific microstructural effects, such as intramolecular interactions, hydrophobic interaction, or steric effects, around the reporter histidine residue.

On the other hand, $\Delta\Delta G_{Elec}$ can be theoretically derived from the basic Coulomb's law.[59] The change in electrostatic force between two charges q_1 and Δq_2 can be obtained from Coulomb's law (eq 8):[60]

$$\Delta F = -\frac{1}{\epsilon 4\Pi\epsilon_0}\frac{q_1\Delta q_2}{rr} \quad (8)$$

where, q_1 is the charge of histidine residue at pH 7.0 (0.5 for free histidine), Δq_2 is the net change in charge due to the perturbation

(59) Warshel, A.; Aqvist, J. *Chem. Scr.* **1989**, *29A*, 75−83.
(60) Pace, C. N.; Huyghues-Despointes, B. M. P.; Briggs, J. M.; Grimsley, G. R.; Scholtz, J. M. *Biophys. Chem.* **2002**, *101*, 211−219.

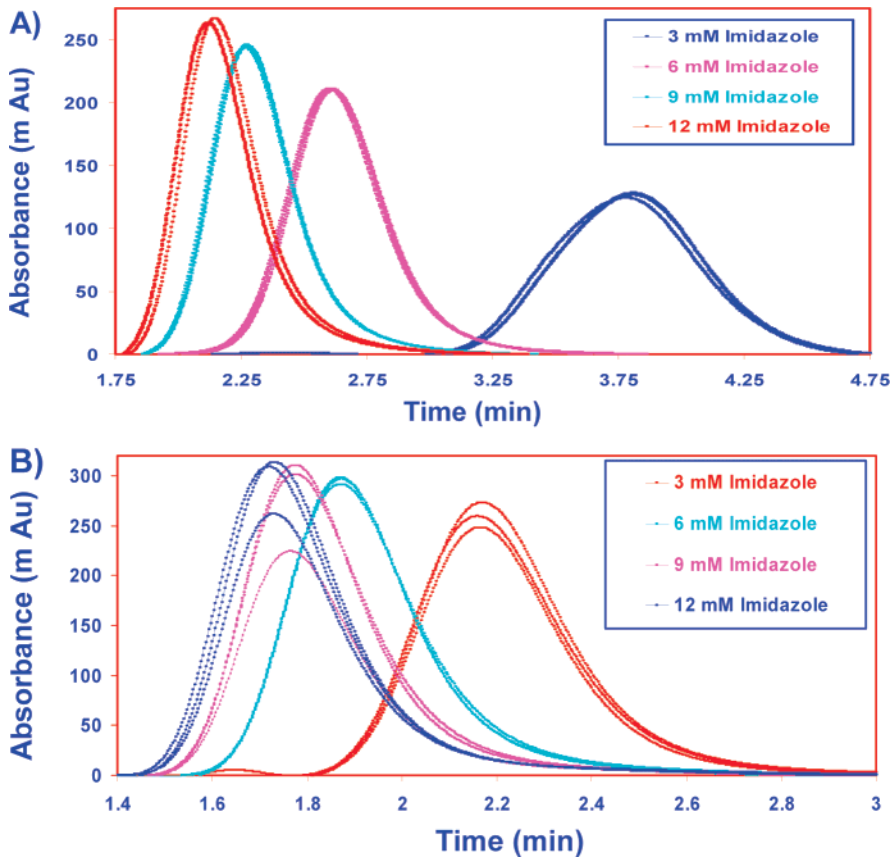

Figure 3. Elution profiles of the lysozyme variants (A) D127K and (B) R76D under four different concentrations of imidazole ([I]).

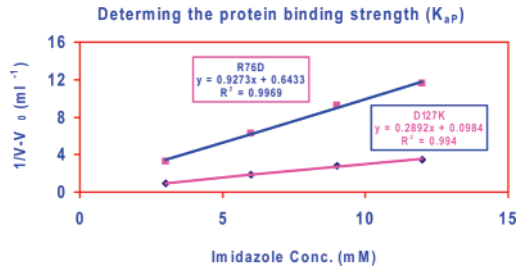

Figure 4. Determination of the intercept required to calculate the binding strength (K_{aP}) of variants R76D and D127K.

in the histidine's microenvironment at pH 7.0, r is the distance between q_1 and Δq_2 (from C_α to C_α), ϵ_0 is the permittivity of free space, and ϵ is the dielectric constant of the protein surface, which equals 45.[15] The change in electrostatic free energy (ΔE) can be defined as the product of the net force (ΔF) and the distance between the two charges q_1 and Δq_2 (eq 9)[35] and is given by the following equation,

$$\Delta E = \Delta F r = -\frac{1}{\epsilon 4\Pi\epsilon_0}\frac{q_1 \Delta q_2}{r} \qquad (9)$$

The change in free energy due to change in electrostatic free energy, ($\Delta\Delta G'_{Elec}$) is given in eq 10,[27,59]

$$\Delta\Delta G'_{Elec} = \Delta E N \qquad (10)$$

where N is Avogadro's number.

RESULTS

All the sites chosen for single point mutations were not critical for the protein's structure or activity because none of these sites were involved in any extensive amino acid network including those involved in the protein's catalytic activity.[53,61,62] Choosing such isolated residues was necessary because the aim of this study was to understand that the structural effects, if any, on the sole probe-histidine were exclusively due to the engineered single point mutation. Circular dichroism (CD) results (Figure 2) and enzymatic assays (Table 4) indicate that the 3D structure of all the lysozyme variants was intact. The $\Delta\Delta G_E$ values of all the variants are reported in Table 5. The elution volume required to half-elute the protein, obtained from the elution profiles (Figure 3), was used to obtain the intercept by plotting $1/(V - V_0)$ vs [I] (Figure 4).[29] From the intercept, the protein binding strength (K_{aP}) value for each variant was calculated (Table 5).

IMAC Derived ΔpK_a and $\Delta\Delta G_B$ Values. The relationship between K_{aPv} and the histidine's pK_a is given in eq 5. From eq 5, we can infer that as the K_{aPv} value of a variant increases or decreases, the ΔpK_a value of the histidine will proportionally decrease (−) or increase (+).[41] The ΔpK_a values of all the variants vary from −0.4 to +0.7 units (Table 5). With the exception of two variants, S136K and D127K (which are later shown to have profound microstructural effects), the approximate ΔpK_a value for an electrostatic perturbation of ±1 and ±2 units is ∓0.25 and ∓0.5

(61) Wray, J. W.; Baase, W. A.; Lindstrom, J. D.; Weaver, L. H.; Poteete, A. R.; Matthews, B. W. *J. Mol. Biol.* **1999**, *292*, 1111−1120.
(62) Matthews, B. W. In *Advances in Protein Chemistry*;Eisenberg, D. S., Richards, F. M., Eds.; Academic Press: San Diego, CA, 1995; Vol. 46, pp 249−278.

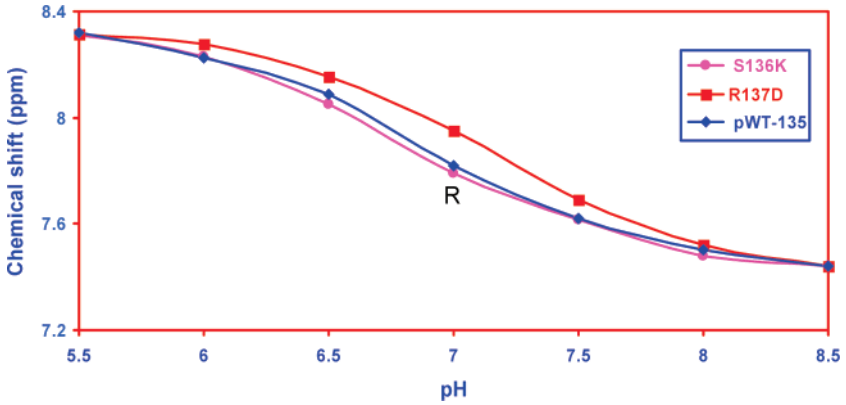

Figure 5. The NMR profiles of three variants R137D (■), pWT-135 (◆), and S136K (●).

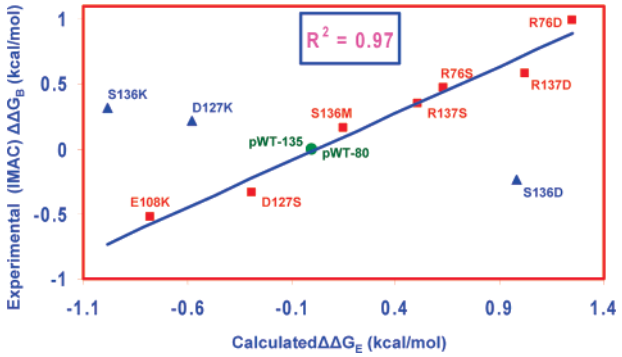

Figure 6. Correlation between the calculated ($\Delta\Delta G_E$) values from Coulomb's law and experimental ($\Delta\Delta G_B$) values obtained from the K_{aP} values obtained from IMAC, excluding the three variants S136K, S136D, and D127K (shown in blue). These three excluded variants (outliers) deviate from the general trend due to the presence of strong site-specific structural (microstructural) effects. The data point indicated in green represents the control variants (pWT-135 and pWT-80).

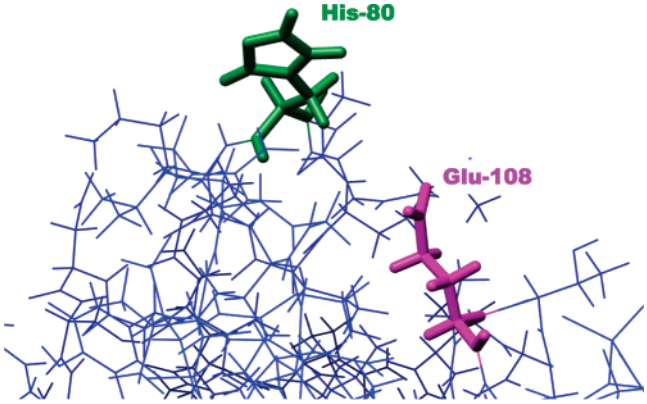

Figure 7. The microstructure of pWT-80, where the side chain of Glu108 is oriented toward His80.

units, respectively. These values are within the range of ΔpK_a values reported by Fersht and co-workers.[10,17] They reported that when aspartate residue was replaced by serine and lysine, the ΔpK_a value of the histidine residue was −0.4 and −0.6 units, respectively. The ΔpK_a values reported in Table 5 support the general trend observed, wherein as the change in net charge varies from a positive to a negative value (due to electrostatic perturbation), the histidine's pK_a decreases or increases, respectively.[63] The causation for this trend can be theorized as the following: when the net positive charge on a protein increases, there is an increase in the repulsive force experienced by the histidine residue's proton, and consequently, an apparent decrease in the histidine's pK_a is observed.[17,26] Conversely, when there is a net increase in the negative charge, the proton on the histidine is stabilized resulting in a higher pK_a value of the histidine residue.[13,14,17,26] In addition, the NMR profiles[8,9,31] of the two variants R137D and S136K also support the above-mentioned general trend (Figure 5). The profile of variant R137D indicates that it has a higher histidine pK_a value when compared to the control histidine (pWT-135), which in turn has a higher histidine pK_a than variant S136K.

The relationship between $\Delta\Delta G_B$ and K_{aPv} is also given in eq 5, which indicates that as the K_{aPv} value increases or decreases, the $\Delta\Delta G_B$ value proportionally decreases (−) or increases (+), respectively. This general trend is supported by the values reported in Table 5. In addition, as shown in Table 5 and Figure 6, the calculated $\Delta\Delta G_{Elec}$ values tightly correlate with the $\Delta\Delta G_B$ values ($R^2 = 0.96$), when the three variants S136K, S136D, and D127K are excluded. These exceptional variants warrant separate discussions due to their unique structural features. However, to appreciate the subtle structural variations in the exceptional variants, we have included a brief description of variants that follow the correlation.

Microstructure Around the Three Sites 108, 76, and 137. As shown in Figure 7, the side chain of the glutamic acid at site 108 (before any site-directed mutagenesis) is not involved in any intramolecular bonds and its orientation is pointing toward the histidine residue. The distance between the nitrogen in the imidazole ring of the histidine and the carboxyl carbon of glutamic acid ($N_\epsilon - C_\gamma$ is 8.8 Å) is much closer than the $C_\alpha - C_\alpha$ distance (9.6 Å). In cases of sites 76 and 137, the side chain (amino moieties) of the arginine residue at these two sites is involved in a salt bridge with Asp72 ($N_\epsilon - O_{\delta 2}$ is 1.7 Å, $N_\epsilon - O_{\delta 1}$ is 2.3 Å, and $N_{1H} - O_{\delta 1}$ is 3.2 Å) and Glu22 ($N_{1H} - O_{\epsilon 1}$ is 1.7 Å and $N_{2H} - O_{\epsilon 2}$ is 1.7 Å), respectively (parts a and b of Figure 8). The residues that replace the original amino acid at these three sites are all surface accessible.

IMAC Derived ΔpK_a and $\Delta\Delta G_B$ Values at Sites 108, 76, and 137. In the case of variant E108K, the ΔpK_a and $\Delta\Delta G_B$ values

(63) Sternberg, M. J. E.; Hayes, F. R. F.; Russell, A. J.; Thomas, P. G.; Fersht, A. R. *Nature* **1987**, *330*, 86−88.

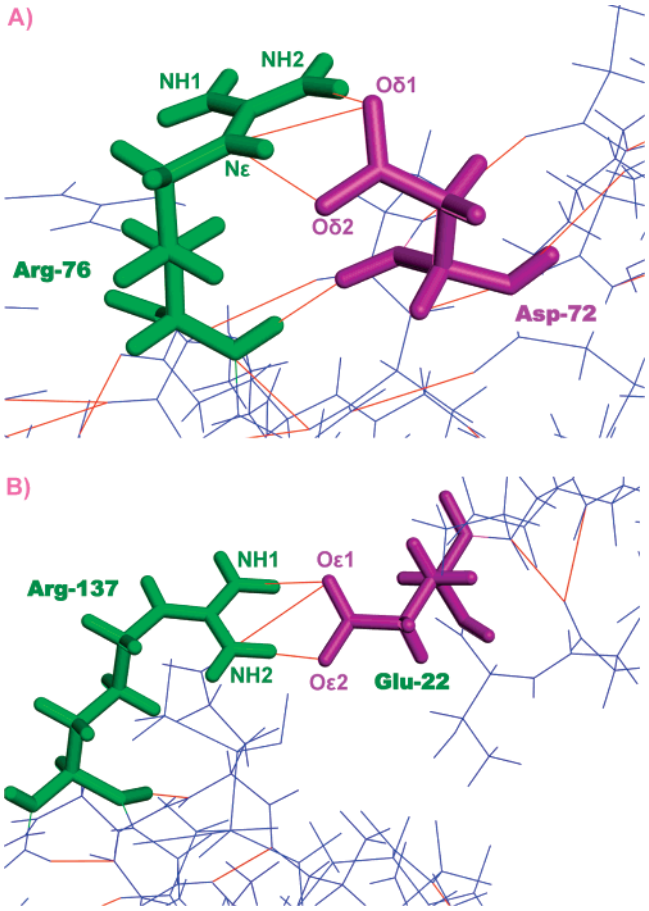

Figure 8. The microstructure of (A) pWT-80, where the side chain of Arg76 is involved in a salt bridge (red lines) with Asp72 and (B) pWT-135, around the site Arg137, where the side chain is involved in a salt bridge (red lines) with Glu22.

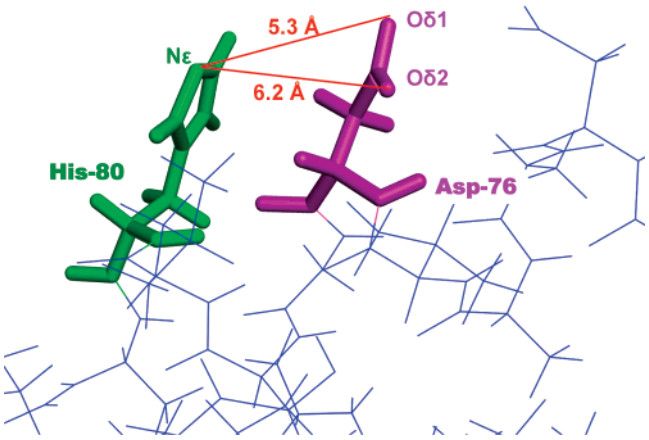

Figure 9. Microstructure of variant R76D, where the side chain of Asp76 is involved in a (weak) salt bridge with the side chain of His80.

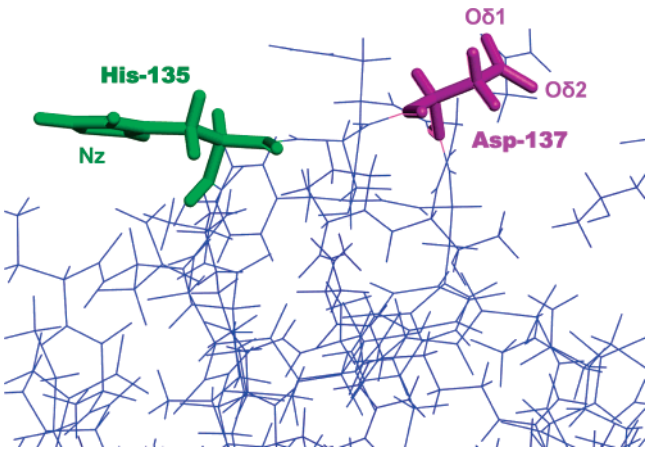

Figure 10. Microstructure of variant R137D, where the two charged moieties of His135 and Asp137 are oriented in opposite direction and are 16.7 Å apart.

of the histidine residue are −0.39 units and −0.55 kcal/mol, respectively. From Figure 6, the $\Delta\Delta G_B$ value is primarily due to the ΔpK_a of the histidine and therefore $\Delta\Delta G_B = \Delta\Delta G_{Elec}$ (eq 5a). In the case of the two sites, 76 and 137, although the generated variants follow the correlation (Figure 6), there are some notable differences between the $\Delta\Delta G_B$ values for the two pairs of variants with the same mutation (R137S vs R76S and R137D vs R76D, Table 5). These two sites can be characterized as having a similar microenvironment due to their (1) equivalent relative surface accessible area, (2) involvement in a salt bridge, and (3) equivalent microstructure (rHy) (Table 2). However, the ΔpK_a value of the variants at these two sites varies as much as 30% for a serine (−1) substitution and 75% for an aspartate (−2) substitution (Table 5). This difference in the ΔpK_a values for the serine substitution can be attributed to the difference in the orientation of the original arginine residue at these two sites. To consider the effects of orientation, it is important to employ the distance between the charged moieties of the probe-histidine ($N_{\epsilon 2}$) and the arginine (N_{H1}, N_{H2}) instead of the $C_\alpha - C_\alpha$ distance. Consequently, the distance between Arg137 and His135 is approximately 10.5 Å more than the distance between Arg76 and His80. In the case of aspartate substitution, assuming that the replaced amino acid side chain mimics the orientation of the side chain of the original amino acid, the distance between His80 and Asp76 ($N_{\epsilon 2} - O_{\delta 1}$ is 5.2 Å and $N_{\epsilon 2} - O_{\delta 2}$ is 6.2 Å) allows the formation of a weak salt bridge (Figure 9), which results in the histidine's ΔpK_a value to be 75% higher than that of His135, whose distance from Asp137 is 16.7 Å in R137D (Figure 10).

Microstructure Around Site 127. In general, to stabilize the charge between a pair of oppositely charged residues, the positively charged residue acts as an acid while the negatively charged residue acts as a base.[30] During this stabilization process, the pK_a of the positively charged residue is raised while that of the negatively charged is lowered (e.g., hydrogen bond).[21,64] On the other hand, to stabilize the charge between a pair of identically charged residues, the amino acids involved have either lower (for positively charged residues) or higher (for negatively charged residues) pK_a values.[21,65,66] Thus in both cases, the pK_a of the amino acids involved is altered.

Asp127 is located next ($C_\alpha - C_\alpha$ is 3.86 Å) to another negatively charged residue Glu128, whose distance from the histidine is ($C_\alpha - C_\alpha$) 11.3 Å, equivalent to that from Asp127 (12.96 Å). In fact, the side chains of the three residues (Asp127 C_δ, His135 N_ϵ, and Glu128 C_γ) form a triangle with an angle of 65.8° at His135 N_ϵ

(64) Li, H.; Robertson, A. D.; Jensen, J. H. *Proteins: Struct., Funct., Bioinf.* **2004**, *55*, 689−704.
(65) Highbarger, L. A.; Gerlt, J. A.; Kenyon, G. L. *Biochemistry* **1996**, *35*, 41−46.
(66) Harris, T. K.; Turner, G. J. *IUBMB Life* **2002**, *53*, 85−98.

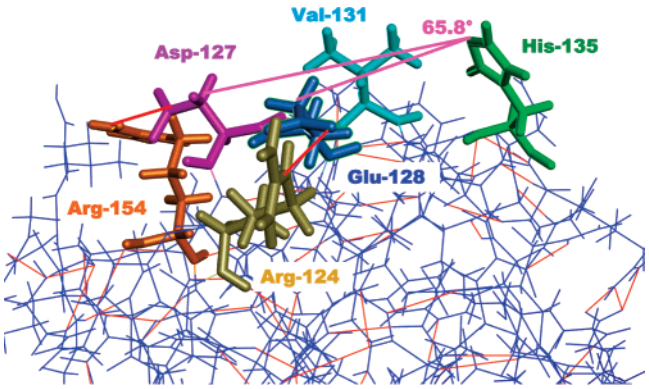

Figure 11. The microstructure of pWT-135 around the site Asp127. The angle between three residues His135, Asp127, and Glu128 is 65.8° (shown in pink). The ion pair formation between the charged moieties of (1) Asp127 and Arg154 and (2) Glu128 and Arg124 is shown by the line highlighted in red.

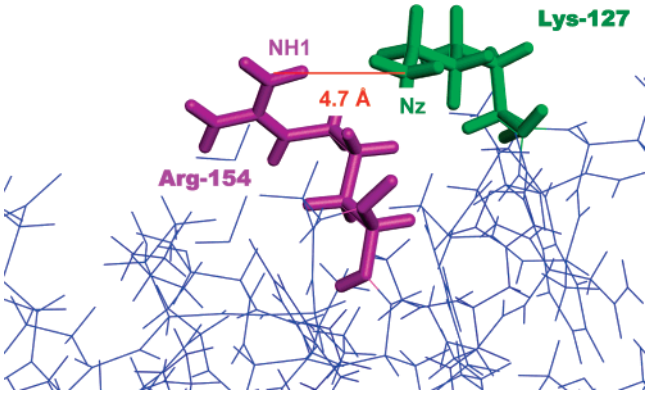

Figure 12. Microstructure of variant D127K, where the side chain of Lys127 is involved in an ion pair of similar charges. The distance between the charged moieties is shown by the red line.

(Figure 11). Therefore, these two negatively charged residues might be involved in an ion pair of similar charges resulting in an increase in their pK_a value. However, from the X-ray crystal structure, these two negatively charged residues are involved in salt bridges with their adjacent arginine residues (Figure 11), and thus the mobility of their side chains is restricted. Therefore, it is hard to rationalize their charge effect on the histidine's pK_a value.[22] In addition, the presence of Val131, which may act as a screening agent between Asp127 and the histidine,[7] further complicates the effect of Asp127 on the histidine's pK_a value (Figure 11).

When a polar serine replaces the negative charged Asp127, there will be a net increase in the total positive charge on the protein by + 1 unit. Though the charge of Glu128 is affected by the variation in the protein's net charge, its effect on His135 is minimal due to its immobilized side chain, which is involved in a salt bridge with Arg125. Thus there should be a net decrease in the pK_a value of the histidine.

On the other hand, replacing Asp127 with lysine results in the formation of an ion pair of similar charges with Arg154 ($Lys127N_z$–$Arg154N_{1H}$ is 4.7 Å) (Figure 12). This ion pair formation might lower the pK_a value of Lys127 (by as many as 4 pH units).[65] Also, assuming the orientation of these residues is similar to that of the original residue, the distance between the charged moieties of Lys127 (N_z) and the histidine ($N_{\epsilon 2}$) is 15.8 Å compared to 12.7 Å for serine. Therefore, the effect of lysine replacement on the

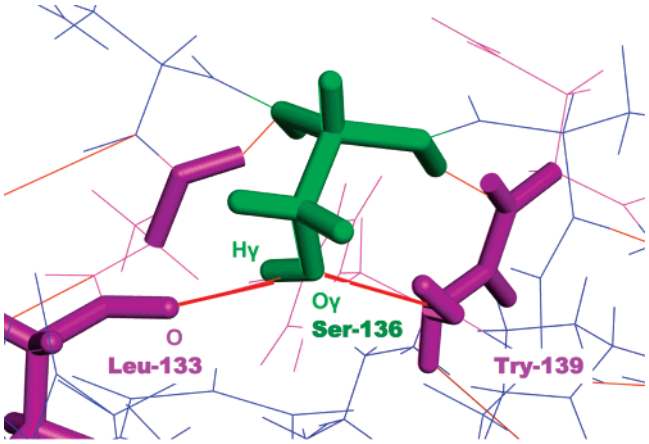

Figure 13. The microstructure of pWT-135, where the side chain of Ser136 (OH) is involved in two hydrogen bonds (highlighted in red) with the Try139 (N) and Leu133 (O).

net protein charge should be much lower than + 2 units.[25] In addition, the NMR profile of the two variants S127K and S127S is almost overlapping (data not shown), thus indicating an equivalent contribution to the pK_a for the histidine residue. Thus, from the microstructure of Lys127, it can be inferred that the effect of the lysine and serine substitutions will have similar effects on histidne's pK_a value.

IMAC Derived ΔpK_a and $\Delta\Delta G_B$ Values at Site 127. As expected, variant D127S has a greater $\Delta\Delta G_B$ value than pWT-135 (Table 5, indicated by a negative value). On the other hand, variant D127K has a lower $\Delta\Delta G_B$ value than both E108K and D127S variants (Table 5). This is an unexpected result because the net charge variation for variant D127K is +2 similar to E108K, while for D127S is +1. Nevertheless, this low $\Delta\Delta G_B$ of D127K can be explained based on the microstructure of Lys127. Because of a lower pK_a of Lys127 and its distance (15.8 Å) from the imidazole ring of histidine, the ΔpK_a of the histidine in variant D127K could be lower than that in variant D127S. Thus, it can be concluded that unlike E108K ($\Delta\Delta G_B = \Delta\Delta G_{Elec}$), the D127K $\Delta\Delta G_B$ value depends on both the $\Delta\Delta G_{Elec}$ and $\Delta\Delta G_{Micro}$ (eq 7).

Microstructure Around Site 136. The side chain of serine at site 136 has very low surface accessibility (Table 2) because its side chain is buried inside the protein interior (or a cavity).[7] The cavity around Ser136, due to the presence of the predominantly nonpolar residues such as Ala134, Leu133, Phe114, Trp138, and Tyr139, can be characterized as hydrophobic.[54] However, the rHy value at site 136 indicates a hydrophilic microstructure (Table 2), and its side chain is involved in two hydrogen bonds with its adjacent residues (Figure 13). Consequently, the mobility or the flexibility of the serine residue is restricted, and thus it can be concluded that the orientation of the side chain (even in solvent) pointing into the protein interior is most probable.

The SYBYL generated S136D variant model also orients the side chain of Asp136 into the protein interior. The microstructure of Asp136 is hydrophilic (rHy = 0.28) and its side chain is involved in a single hydrogen bond. Thus the orientation of the Asp136 side chain should be similar to the side chain orientation of the serine residue in pWT-135 (Figure 13 and 14).

On the other hand, when lysine replaces Ser136 in S136K, there are two possible orientations that need to be considered.[30] In case 1, because of the presence of the cavity and the absence

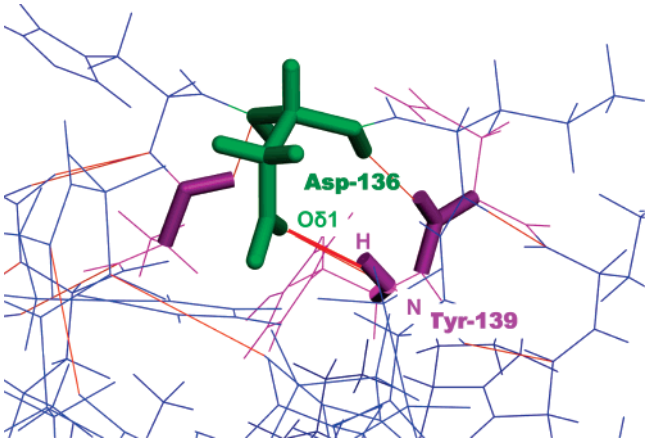

Figure 14. The microstructure of variant S136D. The hydrogen bonds (highlighted in red) are formed by the aspartate (O) with Try139 (H−N).

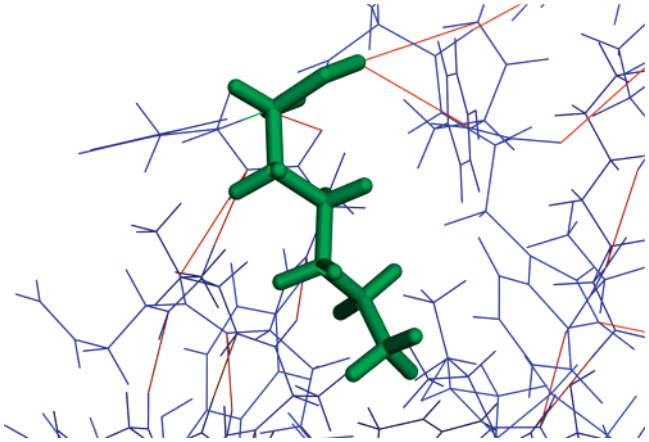

Figure 15. Microstructure of variant S136K, where the side chain of the lysine residue (Lys136) has no hydrogen bonds with its adjacent amino acid residues and is oriented inward into the protein.

of any site-specific steric clashes around the serine residue, the lysine side chain may be orientated into the protein interior,[42] as modeled by the SYBYL software (Figure 15). In case 2, because of (a) the high desolvation energy costs from a hydrophobic microstructure (cavity) around the lysine residue (rHy = 0.08), (b) the high *B*-values (flexibility) of the lysine residue,[45,67−69] and (c) the absence of any hydrogen bonds with its adjacent residues, the side chain of Lys136 may be very mobile as demonstrated in Figure 16. This figure clearly demonstrated that during the course of the MD simulation, the N_z atom of lysine is closer (most of the time) to the $N_{\delta 1}$ atom of His135 than is the Ser O_G atom to the His135 $N_{\delta 1}$ atom. Therefore the Lys136 side chain possesses the capability to orient itself in a different direction from the serine residue outward toward the probe-histidine at site 135.[70,71] This

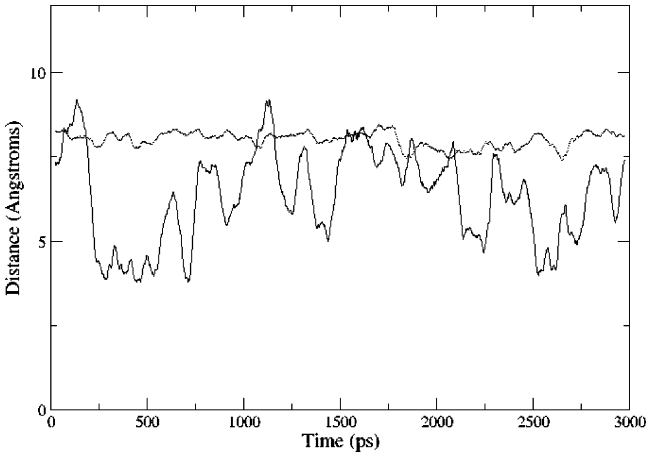

Figure 16. Distance between His135 and residue 136 (in pWT-135 and S136K) during the MD simulations. The lower solid curve depicts the distance between atom N_{D1} of His135 and atom N_Z of Lys136 in the S136K simulation. The upper dotted line is the distance between atom N_{D1} of His135 and atom O_G of Ser136 in the pWT-135 simulation.

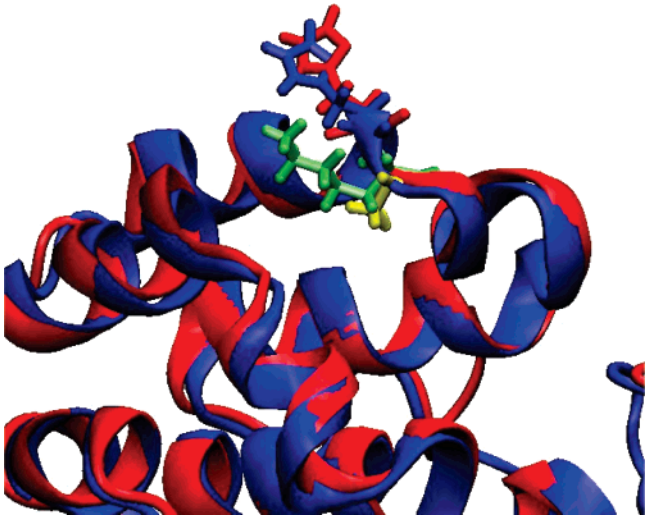

Figure 17. The structures are snapshots at the 3 ns time point in the MD simulations. Structures are illustrated in ribbon format for pWT-135 in blue and S136K in red. The His135 residue in each molecule is colored the same as the ribbon backbone. The residue Ser136 in pWT-135 is colored in yellow, and the lysine that replaces serine in S136K is colored green.

hypothesis is supported by the molecular dynamic simulation result shown in Figure 17. This figure clearly illustrates that unlike the orientation of the serine side chain in pWT-135, the orientation of lysine in S136K is pointing away from the protein's interior (in the orientation similar to probe-histidine at site 135).

IMAC Derived ΔpK_a and $\Delta\Delta G_B$ Values for Variants at Site 136. In case of variant S136D, the possibility of the aspartate residue orienting itself in the direction of serine is very high because (1) aspartic acid is similar in size to serine and is involved in a hydrogen bond (Figure 14 compared to Figure 13) and (2) the microstructure of aspartate is more hydrophilic (rHy = 0.28 in S136D) than lysine (rHy = 0.08 in S136K). However, on the other hand, the relative hydrophilicity of aspartic acid is much less than serine (rHy = 0.70) and thus might increase its pK_a value significantly. Consequently, this higher pK_a value of Asp136 (low net negative charge) lowers the total contribution of Asp136

(67) Eyal, E.; Najmanovich, R.; Edelman, M.; Sobolev, V. *Proteins: Struct., Funct., Genet.* **2003**, *50*, 272−282.
(68) Yuan, Z.; Bailey, T. L.; Teasdale, R. D. *Proteins: Struct., Funct. Bioinf.* **2005**, *58*, 905−912.
(69) Vieira, E.; Binggeli, A.; Breu, V.; Bur, D.; Fischli, W.; Guller, R.; Hirth, G.; Marki, H. P.; Muller, M.; Oefner, C.; Scalone, M.; Stadler, H.; Wilhelm, M.; Wostl, W. *Bioorg. Med. Chem. Lett.* **1999**, *9*, 1397−1402.
(70) McHaourab, H. S.; Kalai, T.; Hideg, K.; Hubbell, W. L. *Biochemistry* **1999**, *38*, 2947−2955.
(71) McHaourab, H. S.; Lietzow, M. A.; Hideg, K.; Hubbell, W. L. *Biochemistry* **1996**, *35*, 7692−7704.

to the ΔpK_a value of His135 in S136D. This lower contribution of Asp136 is captured by Figure 6, wherein the $\Delta\Delta G_B$ value of S136D is shown as an outlier. Despite the low pK_a value, the $\Delta\Delta G_B$ value of S136D is equivalent to the other −1 charged variants (R137S or R76S), due to its short distance from His135, and most significantly still a positive value.

In the case of variant S136K, there is a net positive increase in the $\Delta\Delta G_B$ value opposing the expected net negative increase (Table 5). This variation in the $\Delta\Delta G_B$ value may not be exclusively due to the ΔpK_a value of variant S136K because none of the aforementioned factors (see the introduction section) affecting the ΔpK_a differ from other variants. In fact, the NMR profile of S136K is lower than the control variant pWT-135 (Figure 5), indicating that the pK_a of the histidine residue of variant S136K is lower than pWT-135. Therefore, the variation in the ΔpK_a value might be due to the additional microstructural effects contributing to the $\Delta\Delta G_B$ value (eq 7). The $\Delta\Delta G_B$ value of S136K varies by 0.56 kcal/mol when compared with any other +1 overall net charged variant (D127S or R76S or R137S). This variation can be explained by considering the two following scenarios.

In case 1, the basic assumption for this scenario is that the orientation of the lysine side chain is similar to the serine residue that points into the cavity (Figure 15). Then because of the strong hydrophobic nature of the cavity, the pK_a of Lys136 might be lowered.[54] If this scenario is true, the $\Delta\Delta G_B$ value of S136K (0.22 kcal/mol, which is a positive value) should be approximately equivalent to that (or in range) of the D127S variant (a negative value). In addition, this scenario is not supported by the NMR profile of variant S136K, which indicates a lower pK_a than pWT-135 (Figure 5), and thus a negative $\Delta\Delta G_B$ value (eq 6) was expected. Therefore in order to understand this "unreasonably" high deviation in the $\Delta\Delta G_B$ value of S136K, case 2 was considered.

In case 2, contrary to case 1, the orientation of the lysine residue may be outward toward His135, into the protein exterior (Figure 17). Because of the bulky nature of lysine and its location (His135 C_α−Lys136 C_α is 3.83 Å), the surface accessibility of the histidine residue might be reduced due to (1) steric hindrance or (2) change in its orientation.[46] As the surface accessibility of the histidine decreases, the binding strength of the protein decreases,[29] yielding a decrease in its $\Delta\Delta G_B$ value (eq 7). Thus it can be concluded that even though there might be a considerable electrostatic energy generated due to the charge perturbation (effect on pK_a), the major (or dominant) contribution for $\Delta\Delta G_B$ for variant S136K can be accredited to the steric hindrance effect ($\Delta\Delta G_{Micro}$ = 0.56 kcal/mol) generated by the adjacent lysine residue. On the other hand, lowering of the pK_a value due to inward orientation may be partly responsible (as discussed in case 1) for such a strong deviation in the $\Delta\Delta G_B$ value of S136K.

In order to further investigate the presence of the steric hindrance effect, variant S136M was generated. The methionine residue was selected because it resembles the bulky nature of the lysine residue and is not charged. Also from the hydrophobic parameters, the microstructure around methionine, sHy = −4.7 (rHy was not determined), is characterized to be similar to serine (sHy = −5.6) and therefore is hydrophilic. Thus, because of the nonpolar nature of methionine and the hydrophilic microstructure around it, the orientation of the methionine might be more likely to be outward then inward. Moreover, because of the pair of donor electrons on the sulfur atom and the hydrophobic nature of the methyl moiety in the end of the side chain, the charge of methionine was assigned a value of −0.15. Though the experimental results indicate that the $\Delta\Delta G_B$ value for S136M follows the trend (Table 5), our hypothesis is that both the hydrophobic nature of the methionine (case 1 for S136K) and the steric hindrance effect (case 2 for S136K, due to the bulky adjacent methionine) were still a contributing factor for the $\Delta\Delta G_B$ value of S136M.

Dissection of $\Delta\Delta G_B$. Understanding the importance of each thermodynamic factor that contributes to change in Gibbs free energy (eqs 4 and 7) is important because they provide significant insight into the protein ligand binding (here, interactions in IMAC). The $\Delta\Delta G_B$ values for variants E108K, D127S, R137S, and R137D are primarily due to $\Delta\Delta G_{Elec}$ since no microstructural effects were present. At site 127, the two variants D127K and D127S do not follow the additivity in their $\Delta\Delta G_B$ values.[17] This discrepancy in $\Delta\Delta G_B$ values ($\Delta\Delta G_{Micro}$ = 0.21 kcal/mol) can be attributed to (1) the orientation of the long side chain of Lys127, which augments the distance between its charged moiety from His135 by 3 Å when compared with that from Ser127 to His135 and (2) the presence of the ion pair between two identically charged residues Arg154 and Lys127, which might lower the pK_a of Lys127 (and thus lowering its positive charge). In the case of variant S136K, the perturbation is short ranged and bulky. It induces a significant $\Delta\Delta G_{Micro}$ = 0.56 kcal/mol and can be attributed to $\Delta\Delta G_{Hydro}$ (for case 1) and $\Delta\Delta G_{Steric}$ (for case 2). Thus, dissection of $\Delta\Delta G_B$ provides a quantitative contribution or the relative importance of each thermodynamic factor to the binding process in IMAC.

DISCUSSION

IMAC can be used as an effective technique to study short-range perturbation effects because of the specific nature of the interaction between the histidine's imidazole nitrogen and IMAC−Cu^{2+} ions by forming a coordinate covalent bond. Coordinate covalent bond formation in an IMAC column is a thermodynamic equilibrium process and can be quantified using a parameter called the dissociation constant. On the other hand analyzing these short-range perturbations by enzyme activity assays is challenging because the catalytic activity is generally dependent on a network of residues, which can be both short- and long-range. Therefore, the perturbation of a residue in the catalytic network residue might trigger a cascade effect[6] instead of an exclusive single perturbation effect. Also these cascade effects are protein dependent, and the values generated thereof are not generally applicable.

The dissociation constant, which reflects the protein behavior in an IMAC column, was previously used to measure the structural perturbations in a protein having one surface accessible histidine.[29] In this report, the variation in the dissociation constant value was used to quantify the electrostatic perturbations in the microenvironment of a lone surface histidine. This variation in the dissociation constant was attributed to the ΔpK_a of the histidine and was used to calculate the $\Delta\Delta G_B$ value (eq 6). However, since all the electrostatic charge perturbations around the sole histidine residue are in short and medium range, there might be some associated microstructural effects ($\Delta\Delta G_{Micro}$), in addition to electrostatic effects ($\Delta\Delta G_{Elec}$), that might contribute to $\Delta\Delta G_B$ (e.g., site 136). To study the contribution of these individual factors to the $\Delta\Delta G_B$

value, quantitative assessment of both the electrostatic forces ($\Delta\Delta G_{Elec}$) and the microstructural effects ($\Delta\Delta G_{Micro}$) on the behavior of the histidine residue were evaluated by employing eq 6. In addition, the factor $\Delta\Delta G_{Micro}$ is comprised of three factors: (1) the involvement of the histidine in intramolecular bonds with adjacent amino acids ($\Delta\Delta G_{Int}$), (2) the microenvironment of the replaced residue by an unfavorable hydrophobic- ($\Delta\Delta G_{Hydro}$) environment, which may (a) alter its pK_a value and (b) introduce structural strain,[42] and (3) site-specific structural effects that might affect the histidine ($\Delta\Delta G_{Steric}$) by (i) the orientation and flexibility of the adjacent residue's side chain that affects the histidine's pK_a,[2,65,66,72] and (ii) the size of the adjacent amino acids that might induce a steric hindrance (blocking) effect.[65] Studying and quantifying these individual factors influencing the microstructural effects, due to single point mutations, on the IMAC binding process is very challenging. Here, we have used molecular modeling tools, $\Delta\Delta G'_{Elec}$, and dissected $\Delta\Delta G_B$ values to quantify these factors.

Our study indicates that there are significant microstructural effects that affect the IMAC binding process. In addition, we were able to demonstrate that other factors such as flexibility and orientation of not only the histidine residue (side chain) but also neighboring amino acid residues are important in mediating the protein−metal ion interactions. These results further demonstrate the importance of a flexible or an ensemble protein structure that aids in its function.[36,73,74] Quantifying these subtle microstructural effects is important because they play a critical role in important physiological process such as molecular recognition, protein−ligand and protein−protein interactions.[6,71] Thus, the $\Delta\Delta G_B$ (or the variation of the dissociation constant) values can be used to quantify the change in electrostatic and microstructure of any lone surface histidine. Also it can be concluded that when there were no major microstructural effects influencing the $\Delta\Delta G_B$ value, there might be additivity in the effects due to charge perturbations.[17] On the other hand, this report illustrates the complexity involved in reverse engineering of ion pairs and catalytic sites.[75]

In summary, we have focused our investigation primarily on two directly obtainable parameters, the $\Delta\Delta G_B$ values that were calculated from the measurable protein−metal ion interaction in IMAC, and $\Delta\Delta G_{Elec}$ from theoretical calculations (obtained from Coulomb's law) from the controlled electrostatic perturbation. A direct correlation between these two parameters should exist if all variants have similar microstructural effects ($\Delta\Delta G_{Micro} \approx 0$ in eq 6). If a significant deviation is observed for a particular variant, $\Delta\Delta G_{Micro}$ can be calculated from the correlation developed. Subsequently, with the assistance of molecular modeling, the specific microstructural effects, such as intramolecular interactions, hydrophobic interaction, or steric effects, around the reporter histidine residue can be determined for the variant. Consequently, the protein microstructures around the reporter histidine (electrostatic interaction involving histidine and the microstructure around it) are quantitatively characterized.

CONCLUSIONS

Short and medium range charge perturbations around the sole histidine residue in T4 lysozyme, which is not involved in any catalytic or other network of residues, were engineered within a 15 Å distance. Theoretical $\Delta\Delta G_E$ values were calculated using Coulomb's law. Experimental ΔpK_a and $\Delta\Delta G_B$ values were obtained by measuring the protein binding strength values using IMAC. The direct correlation ($R^2 = 0.96$) between $\Delta\Delta G_E$ and $\Delta\Delta G_B$ indicates that it is possible to predict the electrostatic effect due to mutations ($\Delta\Delta G_{Elec}$), when no other structural effects are involved. The $\Delta\Delta G_B$ value of the three variants (S136K, S136D, and D127K) differ from the other similarly charged variants and their own $\Delta\Delta G_E$ value because of the additional site-specific microstructural effects ($\Delta\Delta G_{Micro}$) that contribute to their $\Delta\Delta G_B$ value. The contribution of these individual factors (e.g., electrostatic, microstructural) to $\Delta\Delta G_B$ was dissected by comparing the experimental $\Delta\Delta G_B$ values with the other variants of identical charge perturbations. Electrostatic and microstructural effects due to a single point mutation on the $\Delta\Delta G_B$ value of the probe-histidine were quantified. In addition, the contribution of the relevant $\Delta\Delta G_{Micro}$ factor to the $\Delta\Delta G_B$ value was identified and quantified. Quantifying these components of $\Delta\Delta G_B$ will help in not only understanding the structure−function relationship but also engineering or designing tailor-made proteins.

ACKNOWLEDGMENT

We gratefully acknowledge computing time on the System X supercomputer in the terascale computing facility at Virginia Tech. The work is partly funded by a grant (Grant No. J-706) from the Jefferess Memorial Trust.

Received for review November 9, 2007. Accepted December 10, 2007.

AC7023188

(72) Tran, N. L.; Colvin, M. E.; Gronert, S.; Wu, W. M. *Bioorg. Chem.* **2003**, *31*, 271−277.
(73) Laberge, M. *Biochim. Biophys. Acta*: *Protein Struct. Mol.ecular Enzymol.* **1998**, *1386*, 305−330.
(74) Price, N. C. *Biotechnol. Appl. Biochem.* **2000**, *31*, 29−40.
(75) Hwang, J. K.; Warshel, A. *Nature* **1988**, *334*, 270−272.