Effects of Charge-to-Alanine Substitutions on the Stability of Ribosomal Protein L30e from *Thermococcus celer*†

Chi-Fung Lee,‡ George I. Makhatadze,*,§ and Kam-Bo Wong*,‡

Department of Biochemistry and Molecular Biotechnology Programme, Centre for Protein Science and Crystallography, The Chinese University of Hong Kong, Hong Kong SAR, China, and Department of Biochemistry and Molecular Biology, Penn State University, College of Medicine, Hershey, Pennsylvania 17033

Received September 28, 2005; Revised Manuscript Received October 28, 2005

ABSTRACT: The ability to rationally engineer a protein with altered stability depends upon the detailed understanding of the role of noncovalent interactions in defining thermodynamic properties of proteins. In this paper, we used *T. celer* L30e as a model to address the question of the role of charge−charge interactions in defining the stability of this protein. A total of 26 single-site charge-to-alanine variants of this protein were generated, and the stability of these proteins was determined using thermal- and denaturant-induced unfolding. It was found that, although L30e is isolated from a thermophilic organism and is highly thermostable, some of the substitutions lead to a further increase in the transition temperature. Analysis of the effects of high ionic strength on the stabilities of L30e variants shows that the long-range charge−charge interactions are as important as the short-range (salt bridge) interactions. The changes in stabilities of the *T. celer* L30e protein variants were compared with the changes in the energy of charge−charge interactions calculated using different computational models. It was found that there is a good qualitative agreement between experimental and calculated data: for 70−80% (19−21 of 26, confidence $p < 0.003$) of the variants, computational models predict correctly the sign of the stability changes. In particular, computational models identify correctly those charged amino acid residue substitutions of which led to enhancement in thermostability. Thus, optimization of the charge−charge interactions might be a useful approach for the rational increase in protein stability.

Understanding the contributions of different noncovalent interactions to the stability of proteins represents one of the important aspects of deciphering the protein-folding problem. A test for such understanding is the ability to predict the effects of small perturbations in the sequence because of a single-site substitution on the stability of a protein. Among different interactions that contribute to the protein stability (hydrophobic effect, hydrogen bonding, packing interactions, and hydration), the effects of interactions between ionizable groups in the protein surface (charge−charge interactions) attracted little attention until recently. However, in the last 5 years, there have been a number of reports showing that charge−charge interactions on the protein surface are capable of modulating protein stability (e.g., see refs *1*−*5*). Furthermore, the effects of changes in charge−charge interactions upon amino acid substitutions on the stability of several proteins were predicted rather well using different computational models (*2*, *5*−*16*). In this paper, we analyzed 26 variants of a thermophilic ribosomal protein L30e from *Thermococcus celer* with single-site alanine substitutions in ionizable (K, R, H, D, or E) amino acid side chains. The *T. celer*[1] L30e protein, is a small globular protein of 100 amino acid residues (*17*). It does not contain any disulfide bonds or bound cofactors. The three-dimensional structure of *T. celer* L30e is known from both X-ray (*18*) and solution nuclear magnetic resonance (NMR) (*19*) experiments and thus allows a structure−stability relationship analysis using protein-engineering approaches (*20*). The effects of alanine substitutions that lead to charge neutralization in *T. celer* L30e on the stability are compared with the effects predicted by computational modeling of energy of charge−charge interactions. This comparison shows that current computational models are rather efficient in predicting the effects of charge−charge interactions on the protein stability.

MATERIALS AND METHODS

Alanine-Scanning Mutagenesis. All alanine mutations were introduced to the coding sequence of *T. celer* L30e by the

† This work was supported by a grant from the Research Grants Council of the Hong Kong Special Administrative Region, China (project number CUHK4254/02M to K.-B.W.) a grant from Research Committee of the Chinese University of Hong Kong (project number 2030253 to K.-B.W.), and the National Science Foundation Grant MCB-0416746 (to G.I.M.).

* To whom correspondence should be addressed: Department of Biochemistry and Molecular Biology, Pennsylvania State College of Medicine, Hershey, PA 17033. E-mail: makhatadze@psu.edu. Telephone: (717) 531-0712. Fax: (717) 531-7072 (G.I.M.); Room 507B Mong Man Wai Building, Department of Biochemistry, The Chinese University of Hong Kong, Shatin, Hong Kong SAR, China. E-mail: kbwong@cuhk.edu.hk. Telephone: (852)-2609-8024. Fax: (852)-2603-5123 (K.-B.W.).

‡ Chinese University of Hong Kong.

§ Penn State University.

[1] Abbreviations: *T. celer*, *Thermococcus celer*; L30e, ribosomal protein L30e; TK-SA, surface-accessibility-corrected Tanford−Kirkwood model; FDPB, finite difference Poisson−Boltzmann; NMR, nuclear magnetic resonance; GdnHCl, guanidine hydrochloride; T_m, melting temperature; ΔG_{H_2O}, free-energy change of unfolding; WT, wild type.

10.1021/bi0519654 CCC: $30.25
Published on Web 12/01/2005

polymerase chain reaction using mutagenic primers. The coding sequence for the mutants was cloned into pET3d (Novagen), and their sequences were subsequently confirmed by DNA sequencing.

Protein Expression and Purification. The vectors containing the coding sequences of *T. celer* L30e variants were transformed to *Escherichia coli* BL21 (DE3) pLysS (Novagen) for overexpression. The bacterial culture was grown in M9ZB medium until A_{600} reached ~0.4, and protein expression was induced by the addition of 0.4 mM isopropyl-β-D-thiogalactopyranoside. The bacterial culture was harvested after 16 h, resuspended in 20 mM sodium acetate buffer at pH 5.4 (buffer B), and lysed by sonication. The lysate was centrifuged at 15000*g* for 30 min, and the supernatant was loaded to a Hi-Trap SP Sepharose HP column (Amersham Biosciences) pre-equilibrated with buffer B. The protein was then eluted at ~0.4 M NaCl using a linear gradient of 0.2−0.7 M NaCl in buffer B over a volume of 225 mL. The eluted protein was loaded to a Hi-Trap Heparin HP column (Amersham Biosciences) pre-equilibrated with 0.2 M NaCl in buffer B. The protein was eluted at ~0.4 M NaCl using a linear gradient of 0.2−0.7 M NaCl in buffer B over a volume of 120 mL. The eluted protein was then concentrated to ~5 mL and loaded to a Superdex G-75 column HiLoad 26/60 (Amersham Biosciences) gel-filtration column pre-equilibrated with 0.2 M Na_2SO_4 in buffer B. The purified *T. celer* L30e was eluted at ~200 mL.

Stability Measurements from Guanidine-Induced Denaturation. Protein samples (~20 μM) were equilibrated with 0−7.2 M guanidine hydrochloride (GdnHCl) in 10 mM sodium phosphate buffer at pH 7.4 and 25 °C for 30 min before circular dichroism measurements. The concentration of the guanidine hydrochloride solution was determined from refractive index measurements (*21*) using a Leica AR200 refractometer. Mean residue ellipticity at 222 nm was measured at 25 °C using a 1-mm path-length cuvette with a JASCO J810 spectropolarimeter equipped with a peltier-type temperature control unit. The data were fitted by a nonlinear regression to a two-state model using (*22*) $y_{obs} = \{(y_n + m_n[D]) + (y_u + m_u[D])\, e^{-\Delta G_{(D)}/RT}\}/(1 + e^{-\Delta G_{(D)}/RT})$, where y_{obs} is the observed mean residue ellipticity at 222 nm, y_n and m_n are the *y* intercept and slope of the linear baseline before the transition, y_u and m_u are the *y* intercept and slope of the linear baseline after the transition, *R* is the gas constant, *T* is the temperature in Kelvin, [D] is the concentration of GdnHCl, and $\Delta G_{(D)}$ is the free energy of unfolding at [D]. The free energy of unfolding without the denaturant, $\Delta G_{(H_2O)}$, was obtained by the linear extrapolation model (*22*): $\Delta G_{(D)} = \Delta G_{H_2O} - m[D]$. Average values and standard deviations of ΔG_{H_2O}, midpoint of transition, and *m* value over three independent experiments were reported.

Stability Measurements from Thermal Denaturation. Thermal denaturation was followed by molar ellipticity at 222 nm using a JASCO J810 spectropolarimeter equipped with a peltier-type temperature control unit. All protein samples were dialyzed in 10 mM sodium phosphate buffer at pH 7.4 and were thoroughly degassed before circular dichroism measurements. The samples were heated in a 1-mm path-length cuvette from 25 to 110 °C at a heating rate of 1° min^{-1}. The cuvette was securely stoppered to ensure that there was no loss in the volume of the protein solution because of evaporation. The thermal denaturation data were analyzed by a two-state model: $K_{(T)} = \{y_{obs} - (y_n + m_nT)\}/\{(y_u + m_uT) - y_{obs}\}$, where $K_{(T)}$ is the equilibrium constant of unfolding at temperature *T*. $K_{(T)}$ values within the transition zone were used to obtain ΔG values by $\Delta G = -RT \ln K_{(T)}$. The melting temperature, T_m, was determined as the temperature at which $\Delta G = 0$.

Thermal unfolding of several protein variants, most of them with the substitutions in the negatively charged residues (D2A, E6A, D12A, D44A, K46A, E47A, D48A, E50A, E62A, E64A, E69A, and D87A), was irreversible. In these cases, only apparent melting temperatures were estimated as follows. Because the values of molar ellipticity of all variants and wild-type *T. celer* L30e proteins were similar ($-6.9 \pm 0.2 \times 10^5$ deg cm^2 mol^{-1}) at their corresponding T_m values, the apparent melting temperatures were estimated as the temperatures at which the values of the molar ellipticity were equal to -6.9×10^5 deg cm^2 mol^{-1}. T_m measurements were repeated twice for all variants and wild-type *T. celer* L30e.

For all protein variants, the ellipticity at 25 °C was very similar, suggesting that the single-site substitutions did not have any significant effect of the structure of the protein. Recently solved X-ray structures of the position K9A and R92A variants of the *T. celer* L30e protein (*20*) further support this conclusion.

Calculations of the Energy of Charge−Charge Interactions. The energies of the charge−charge interaction were determined from the changes in the pK_a values relative to the model compound values. The pK_a values of ionazable residues were calculated using several different computational models on the PDB entries 1H7M [X-ray (*18*)] and 1GO0 [NMR (*19*)] for *T. celer* L30e. The computational models include both continuum (TK-SA, FDPB/UHBD, MCCE, MM_SCP) and semimicroscopic (PDLD) approaches to treat electrostatic effects. (1) The TK-SA procedure implementation is described elsewhere (*1*, *2*, *23*). In this model, the energy of pairwise interactions between unit charges was calculated according to the Tanford−Kirkwood algorithm (*24*) with the solvent accessibility correction as proposed by Gurd et al. (*25*, *26*). Mean field approximation was used to calculated the effect of charge−charge interactions on the pK_a of ionizable groups from their model compound values (Asp, 4.0; Glu, 4.5; Lys, 10.5; Arg, 12.0; His, 6.3; Tyr, 10.5; N terminus, 7.7; C terminus, 3.6). To estimate the uncertainties of calculation for the X-ray structure, 10 structural models were built using MODELLER (*27*). The standard deviations from the mean obtained using these structures are reported for the energies calculated using TK-SA for the X-ray data. (2) The finite difference Poisson−Boltzmann (FDPB) method as implemented in the UHBD software package (*28*, *29*) was described previously (*30*). (3) The multiconformer continuum electrostatics (MCCE) software package uses Monte Carlo sampling of different side-chain rotamers in conjunction with the FDPB calculations using the DELPHI software package and PARSE solvation (*31*, *32*). (4) The microenvironment-modulated screened Coulomb potential (MM_SCP) approach introduced by Mehler et al. (*33*, *34*) was used as described previously (*30*). (5) A semimicroscopic model of protein dipoles Langevin dipoles (PDLD) as implemented in the MOLARIS/POLARIS software package was introduced by Warshel et al. (*35*−*37*).

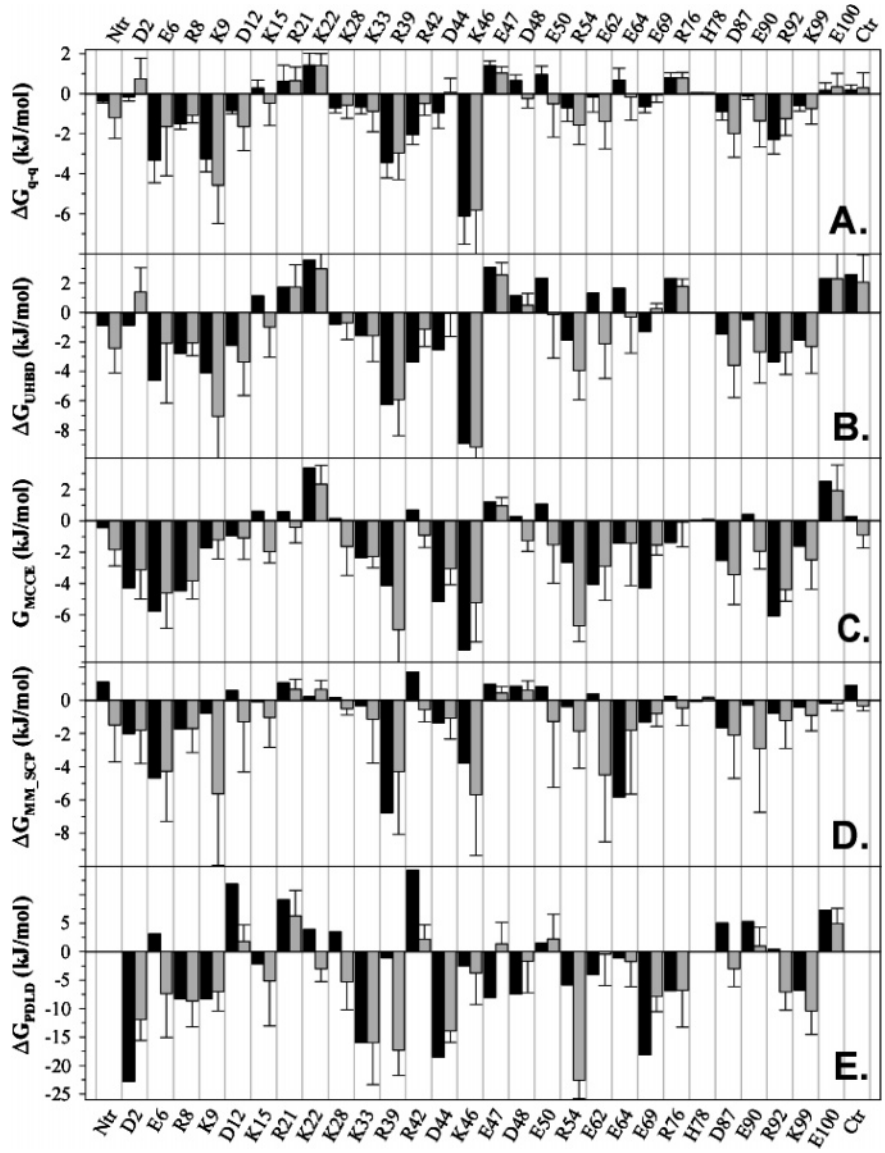

FIGURE 1: Comparison of the energies of the charge−charge interactions in *T. celer* L30e calculated according to different computational models using X-ray (1H7M, black bars) or NMR (1GO0, gray bars). (A) TK-SA (*62*). (B) FDPB method as implemented in the UHBD software package FDPB/UHBD (*28*, *29*). (C) MCCE software package (*31*, *32*). (D) MM_SCP software package (*33*, *34*). (E) PDLD software package (*35*, *36*). Positive energies show that a given residue has an unfavorable overall energy of charge−charge interactions, while a negative value indicates a favorable overall energy. The error bars are calculated as described in the Materials and Methods.

Charge−charge interactions were calculated for each of the 10 structural models given in the NMR Protein Data Bank entry file, and averaged values and standard deviations are reported. For the comparison of the wild-type protein with the L30e variants containing amino acid substitutions, it is assumed that there are no differences in the charge−charge interactions in the unfolded states of these proteins. Alternatively, the energy of charge−charge interactions in the unfolded state was calculated using a Gaussian chain model as introduced by Zhou (*38*).

RESULTS AND DISCUSSION

Charge−Charge Interactions in T. celer L30e. Figure 1 shows the results of the calculations for energies of the charge−charge interactions in the *T. celer* L30e protein using five different computational models: surface-accessibility-corrected Tanford−Kirkwood (TK-SA) model, finite difference Poisson Boltzmann (FDPB/UHBD) model, multiconformer continuum electrostatic (MCCE) model, microenvironment-modulated screened Coulomb potential (MM_SCP) model, and protein dipoles/Langevin dipoles (PDLD) model. Calculations done on both X-ray and NMR-based structural models are shown. Both structural models give statistically indistinguishable results for the energies of charge−charge interactions per individual ionizable residue basis. Correspondingly, in all further discussions, the averaged energies are used. A comparison also shows that qualitatively the results of the calculations are very similar for different computational models. Such agreement is expected for the residues on the protein surface (*35*, *39*).

Calculations presented in Figure 1 show that certain residues have very strong favorable charge−charge interactions, such as E6, K9, R39, and K46. It is thus expected that neutralization of charges at these positions upon substitutions with alanine will dramatically decrease the stability of these variants. Several other positions also appear to have favorable charge−charge interactions (e.g., R8, D12, K28, K33, D87, and R92), and thus, substitutions at these

Table 1: Thermodynamic Parameters of the *T. celer* L30e Protein and Its Variants[a]

T. celer L30e variants	T_m[b] (°C)	ΔT_m[c] (°C)	$\Delta\Delta T_m$[d] (°C)	$D_{1/2}$ (M)	ΔG_{H_2O} (kJ/mol)	$\Delta\Delta G_{H_2O}$ (kJ/mol)
wild type	93.8 ± 0.1			4.49 ± 0.02	47.5 ± 0.4	
D2A	91.3 ± 0.1*	−2.5*	1.4	4.29 ± 0.02	45.3 ± 0.2	−2.2
E6A	82.3 ± 0.2*	−8.6*	5.4	4.19 ± 0.01	44.3 ± 0.2	−3.2
R8A	86.8 ± 0.2	−7.0	2.8	4.20 ± 0.02	44.3 ± 0.2	−3.2
K9A	85.4 ± 0.3	−8.4	4.2	4.39 ± 0.01	46.3 ± 0.3	−1.1
D12A	94.2 ± 0.2*	0.4*	2.6	4.59 ± 0.01	48.5 ± 0.2	1.1
K15A	87.7 ± 0.1	−6.0	2.9	4.14 ± 0.01	43.7 ± 0.2	−3.7
R21A	92.3 ± 0.3	−1.4	3.3	4.53 ± 0.03	47.8 ± 0.2	0.4
K22A	94.8 ± 0.1	1.1	1.9	4.67 ± 0.04	49.3 ± 0.3	1.9
K28A	89.4 ± 0.1	−4.3	2.1	4.03 ± 0.04	42.5 ± 0.3	−4.9
K33A	87.4 ± 0.3	−6.3	4.0	4.55 ± 0.05	48.0 ± 0.3	0.5
R39A	89.0 ± 0.1	−4.8	3.6	4.47 ± 0.04	47.2 ± 0.3	−0.2
R42A	94.0 ± 0.1	0.3	1.6	4.63 ± 0.03	48.9 ± 0.3	1.4
D44A	93.9 ± 0.1*	0.1*	−1.2	4.51 ± 0.07	47.7 ± 0.4	0.2
K46A	86.4 ± 0.2*	−7.4*	1.8	4.06 ± 0.05	42.9 ± 0.3	−4.6
E47A	95.8 ± 0.4*	1.6*	1.1	4.48 ± 0.03	47.3 ± 0.2	−0.1
D48A	97.0 ± 0.2*	3.2*	2.3	4.65 ± 0.03	49.0 ± 0.2	1.6
E50A	96.1 ± 0.4*	2.3*	−1.0	4.73 ± 0.03	49.9 ± 0.2	2.4
R54A	88.4 ± 0.1	−5.3	3.4	4.40 ± 0.02	46.4 ± 0.2	−1.1
E62A	88.3 ± 0.1*	−5.7*	3.0	3.98 ± 0.01	42.1 ± 0.2	−5.4
E64A	93.4 ± 0.1*	0.5*	0.8	4.40 ± 0.02	46.4 ± 0.2	−1.1
E69A	91.8 ± 0.2*	−2.0*	1.4	4.20 ± 0.02	44.4 ± 0.2	−3.1
R76A	92.3 ± 0.1	−1.5	0.8	4.57 ± 0.03	48.3 ± 0.2	0.9
H78A	91.6 ± 0.2	−2.2	3.8	4.30 ± 0.02	45.4 ± 0.2	−2.1
D87A	84.6 ± 0.2*	−9.2*	−0.4	3.96 ± 0.05	41.8 ± 0.3	−5.6
E90A	92.9 ± 0.1	−0.8	2.3	4.34 ± 0.01	45.8 ± 0.2	−1.7
R92A	90.7 ± 0.1	−3.7	1.4	4.38 ± 0.03	46.2 ± 0.2	−1.2

[a] Thermal denaturation of these variants was irreversible, and only apparent melting temperatures (indicated by asterisks) estimated as described in the Materials and Methods are reported. [b] T_m = transition temperature in 0 M NaCl. [c] $\Delta T_m = T_m$(mutant) − T_m(WT). [d] $\Delta\Delta T_m = \Delta T_m$(200 mM NaCl) − ΔT_m(0 M NaCl).

positions are also predicted to decrease the stability of the variants. Importantly, amino acid residues at several positions appear to have an unfavorable overall energy of charge−charge interactions (e.g., R21, K22, E47, and R76). Substitutions in these positions that lead to charge neutralization are predicted to be stabilizing because of the removal of these unfavorable interactions. These predictions are tested experimentally by measuring the effects of substitutions on the relative changes in stability of *T. celer* L30e protein variants.

Scanning Charge-to-Alanine Mutagenesis of T. celer L30e. To study the role of surface charges in the stability of *T. celer* L30e, 26 charge-to-alanine variants of *T. celer* L30e have been constructed. Their melting temperature (T_m) and free-energy change of unfolding at 25 °C (ΔG_{H_2O}) were measured by thermal- and guanidine-induced denaturation, respectively (Table 1). Most of the charge-to-alanine mutations (14 of 26: D2A, E6A, R8A, K9A, K15A, K28A, K46A, R54A, E62A, E69A, H78A, D87A, E90A, and R92A) resulted in significant decreases in both T_m and ΔG_{H_2O} values ($\Delta T_m > 0.5$, $\Delta\Delta G_{H_2O} > 1$ kJ/mol; Table 1). Three variants (K22A, D48A, and E50A) have increased T_m and ΔG_{H_2O} values, suggesting that these variants are more stable than the wild-type protein. On the other hand, the D44A substitution did not affect the stability of *T. celer* L30e, because there was no change in either T_m or ΔG_{H_2O} values. The changes in T_m and ΔG_{H_2O} were not always consistent. For example, K33A and K39A substitutions produced a large decrease in the T_m, yet their ΔG_{H_2O} values were similar to that of the wild type. However, the ΔG_{H_2O} values were measured from the GdnHCl-induced unfolding. Thus, it is possible that because of the ionic nature of this denaturant the Gibbs energy estimates might reflect not only the differences in intrinsic stability but also the differences in the ionic strength dependence of the stability (*40*). Indeed, it is known that the stability of the L30e strongly depends upon the ionic strength (*20*). For example, the T_m for the wild-type protein increases by ∼4 °C (from 92.5 to 97 °C) upon addition of 200 mM NaCl, while the T_m for the K9A variant increases by ∼8 °C. These effects appear to vary depending upon the mutation. Figure 2 shows the ionic strength dependency of ΔT_m: ΔT_m(0 M NaCl) = T_m(mutant) − T_m(WT) for 12 significantly destabilized variants (E6A, R8A, K9A, K15A, K28A, K33A, R39A, K46A, R54A, E62A, D87A, and R92A). It is apparent that the biggest decreases in T_m values for charge-to-alanine variants relative to the wild type are observed in the absence of salt (ΔT_m ranges from −3.7 to −9.2 °C; Table 1 and Figure 2). The ΔT_m values became less negative at an increasing salt concentration (Figure 2) and level off at the NaCl concentrations above ∼200 mM. This observation suggests that the destabilization because of the charge-to-alanine mutations was contributed by electrostatic interactions, which can be screened efficiently by salt. As the ΔT_m reaches a plateau value at 200 mM NaCl, we define the ionic strength sensitivity on ΔT_m as $\Delta\Delta T_m = \Delta T_m$(200 mM NaCl) − ΔT_m(0 M NaCl). We have measured the $\Delta\Delta T_m$ values for all 26 charge-to-alanine variants (Table 1). A nonzero value of $\Delta\Delta T_m$ indicates the contribution of electrostatic interactions to the stability of the variants. For example, the ΔT_m value of R8A is ∼ −7 °C at 0 M NaCl but levels off to −4 °C at 200 mM NaCl. In this case, $\Delta\Delta T_m$ is ∼3 °C, which represents the electrostatic contribution of the destabilization caused by the R8A substitution.

Although *T. celer* L30e is already a very stable protein (T_m ∼ 94 °C, ΔG_{H_2O} ∼ 48 kJ/mol), it could be further stabilized. Removal of charges at positions E47, D48, and

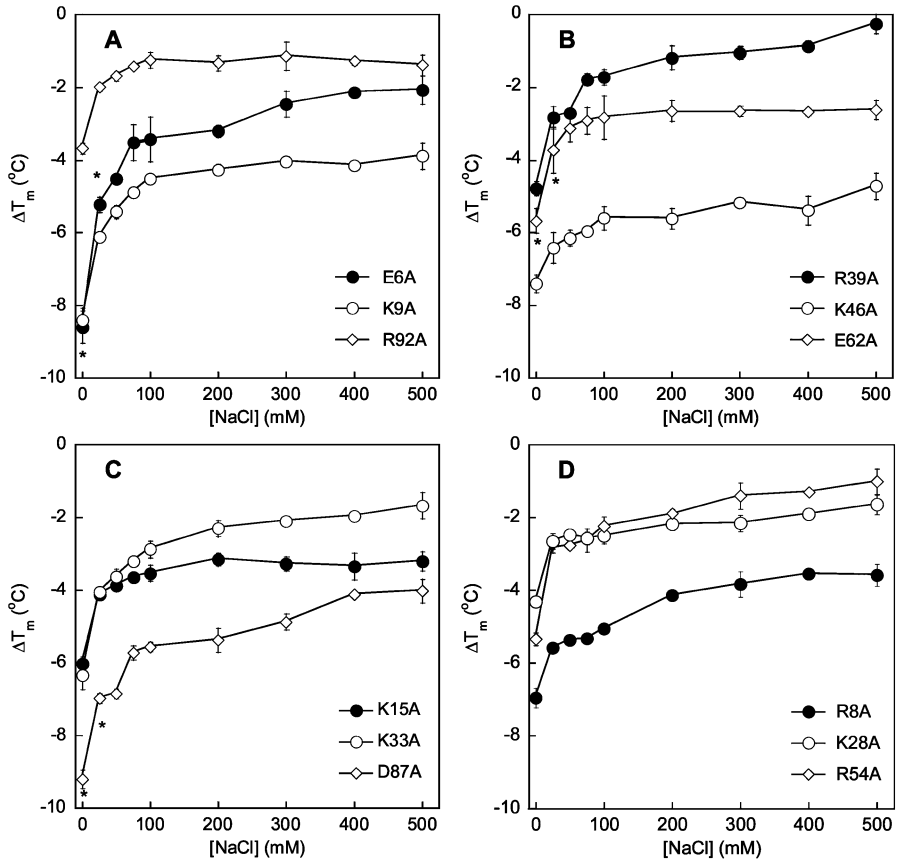

FIGURE 2: Salt dependence of the thermostability of charge-to-alanine variants. ΔT_m, of *T. celer* L30e and its variants, was measured in 10 mM sodium phosphate buffer at pH 7.4 and 0−0.5 M NaCl. ΔT_m, calculated as T_m(mutant) − T_m(WT), was found to be dependent upon the salt concentration. At a low salt concentration, thermal denaturation of E6A, E62A, and D87A was irreversible. In these cases, only apparent temperatures, indicated by asterisks, were estimated.

E50 resulted in increases in T_m and/or ΔG_{H_2O} values. All of these residues are located in the helix 3 of L30e (Figure 3). However, this increase in stability cannot be due to only better helical propensity of Ala versus Asp or Glu residues because similar substitutions in the helix 1 (E6 and D12) or helix 4 (E64) are destabilizing. An additional factor, namely, the removal of unfavorable charge−charge interactions, appears to be also consistent with the observed increase in stability (*41*).

Comparison with the Computational Models. Experimental data on the effects of 26 single-site substitutions on the stability changes in the *T. celer* L30e protein can be compared with the results of the predictions based on the computational assessment of the energetics of charge−charge interactions in this protein. A comparison of experimental data with two different computational models (TKSA and UHBD) is presented in Figure 4. Inspection of Figure 4 reveals two important features. First, *qualitatively*, changes in stability are predicted rather well; most of the points in Figure 4 are in the either lower left or upper right quadrants of the correlation plot. Second, the relative ranking of the effects of substitutions on stability are also predicted relatively well. The significance of these observations is particularly important because alanine substitutions eliminate not only charges. There are differences between alanine and charged residues (aspartate, glutamate, lysine, and arginine) also in terms of configurational entropy, hydrogen-bonding potential, hydrophobicity, and secondary-structure propensity to name just the major ones. However, it appears that, for the surface residues, charge−charge interactions dominate the contributions to the protein stability from these other factors. If so, the results of computational modeling could provide further insights into how charge−charge interactions contribute to the protein stability. In particular, the issue of contribution of short-range (salt bridges) versus long-range charge−charge interactions may be possible to analyze.

The residues for which calculations of charge−charge interactions were unable to predict the changes in stability are also of interest because they can reveal why the prediction failed. There is a possibility that substitutions in charge residues have an effect on the charge−charge interactions in the unfolded state (*13*, *42*−*47*). Figure 4D compares the experimental data on stability of L30e with the results of the calculation done using the UHBD model corrected for the charge−charge interactions in the unfolded state modeled as a Gaussian chain (*38*). It appears however that inclusion of the charge−charge interactions in the unfolded state does not improve the observed correlation. Similar results (data not shown) are observed using other models for predicting the charge−charge interactions in the unfolded state (*7*, *48*). Another possibility is related to the fact that, as we already discussed above, alanine substitutions remove not only charge−charge interactions. For all residues that are mispredicted (i.e., in the upper left and lower right quadrant of Figure 4), the contribution of charge−charge interactions are predicted to be small (within ±2 kJ/mol), and in these cases, misprediction could be due to larger changes in the nonelectrostatic factors than the changes in the charge−charge

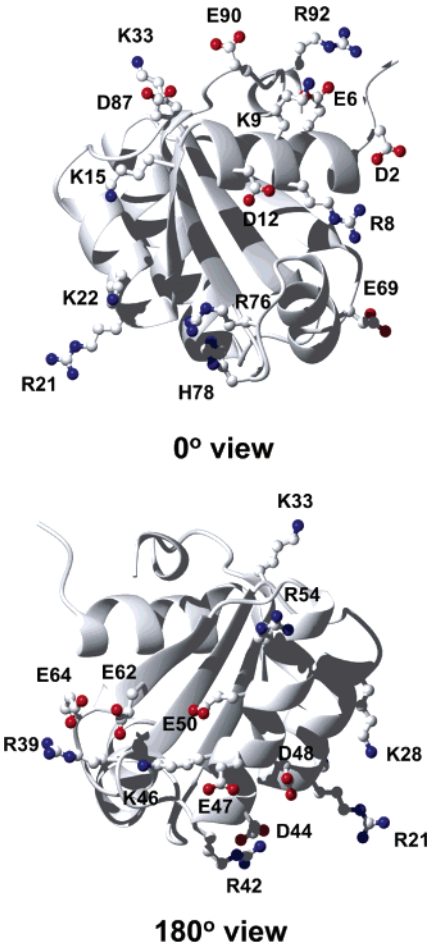

FIGURE 3: Cartoon representation of the three-dimensional structure and locations of the surface-charged residues of *T. celer* L30e. Ionizable residues are shown in a ball-and-stick representation and labeled.

interactions. Looking for these nonelectrostatic factors by structural analysis of *T. celer* L30e reveals an interesting trend (see Figure 3): all of these residues (D2, D12, R21, R42, R44, and R76) are located either at the N or C terminus of α helices. Such clustering suggests that the residues at these positions might be engaged in the structure-specific interactions at the helix ends.

Indeed, several so-called helix-capping motifs have been identified from the analysis of frequencies of residues in the sequences at the ends of α helices (*49*−*51*). Subsequent experimental studies of the capping interactions in protein helices confirmed that different residues at a given position in the capping motif contribute differently to the thermodynamic stability of proteins (*52*−*57*). According to the notation of Rose and Aurora (*51*), the residues in the helix are labeled as

... −N″−N′−Ncap−N1−N2−N3−N4− ... −C4−C3−C2−C1−Ccap−C′−C″− ...

where the numbered residues belong to helix, Ncap and Ccap are the first residues before and after helical residues, respectively, and primed residues are the residues outside the helix. Using this notation, D2 is an Ncap residue and D12 is a C1 residue for the first helix (residues 3−12), R21 is an N2 residue of helix 2 (residues 20−28), R42 is an Ncap and D44 is an N2 residue for the third helix (residues 43−56), and R76 is a C″ residue for the fourth helix (residues 67−73). Looking at the correlation plot in Figure 4, substitutions at residues D2, D12, R42, and D44 are mispredicted, although they are within the experimental error of the "correct" (lower left or upper right) quadrants. The two residues that are mispredicted even taking into consideration the experimental error and the difference in using X-ray versus NMR structures for calculations are R21 and R76. The effect of alanine substitution in these two residues is predicted to be stabilizing, while experimental data show destabilization. Arg21 is located at the N2 position of helix 2. Statistical analysis of the residue frequencies at the N2 position shows that alanine has almost twice the normalized frequency than arginine (*51*, *52*, *58*). This would suggest that the observed decrease instead of increase in stability (as predicted by calculations of the charge−charge interactions) is due to the nonelectrostatic factors. The C terminus of helix 4 is capped by a Schellman motif, with L74, G75, and R76 at the Ccap, C′, and C″ positions. Residues with long alkyl side chain are favored at the C″ position of a Schellman motif because they can form hydrophobic interactions with C3 and Ccap residues (*51*, *56*). Again, a nonelectrostatic factor such as the removal of a hydrophobic interaction might be contributing to the observed decrease in stability upon R76A substitution and the misprediction.

Short- and Long-Range Charge−*Charge Interactions.* Analysis of the X-ray structure of *T. celer* L30e shows that there are three salt bridges, defined as the distance between charged groups less than 5 Å (*59*): E6/R92, R39/E64, and K46/E62. Analysis of the distance matrix in the NMR structural models also identifies, in addition to these three pairs, a potential salt bridge between K9 and E90. Interestingly, substitution of individual partners in the salt bridge leads to different changes in stability. For example, for the E6/R92 salt bridge, the E6A variant is less stable than the R92A variant, while for the K46/E62 salt bridge, the K46A variant is less stable than the E62A variant. The most dramatic difference is for the R39/E64 salt bridge, where the R39A variant is less stable than the wild type, while E64A is more stable. Keeping in mind that E6/R92, K9/E90, R39/E64, and K46/E62 salt bridges are located on the surface of *T. celer* L30e and thus the desolvation penalty is probably minimal, these experimental results support an earlier finding (*60*) that the contribution of the salt bridge to the protein stability is defined not only by the net strength of the salt bridge but also by the long-range charge−charge interactions of a given partner in the salt bridge with the rest of the ionizable groups. Indeed, our experimental results can be rationalized by computational modeling, which shows that there is a qualitative correlation in the relative effects of substitutions in the residues forming salt bridges and the strength of the long-range charge−charge interactions of these residues, i.e., the sum of all charge−charge interactions except with the salt-bridge partner (Figure 5). For example, in the absence of the salt-bridge partner, K9 has overall favorable charge−charge interactions, while E90 has overall unfavorable charge−charge interactions. Thus, the substitution in K9 removes both a favorable salt bridge and favorable long-range interactions and therefore leads to a larger destabilization than the substitution in E90 that removes not only a favorable salt bridge but also unfavorable long-range interactions.

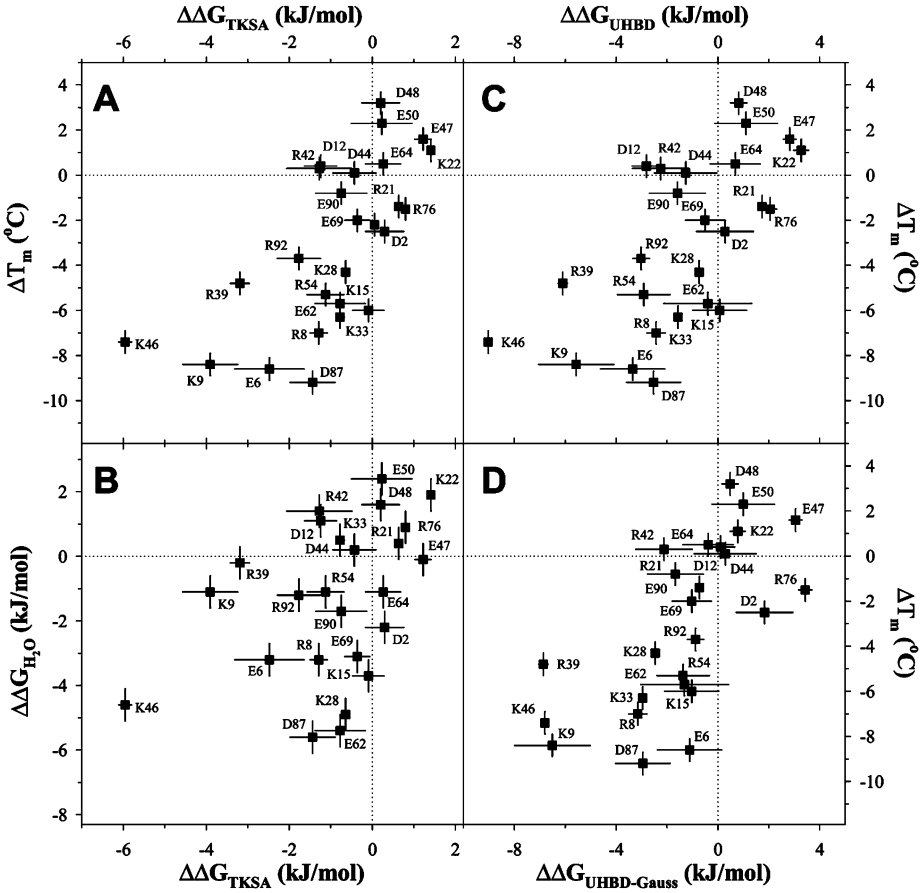

FIGURE 4: Correlation between the experimentally observed changes in stability of the *T. celer* L30e variants and the calculated changes in charge−charge interactions. (A) Changes in the transition temperature versus the expected changes in the energies of charge−charge interactions according to the TKSA model, $\Delta\Delta G_{TKSA}$. (B) Changes in Gibbs energy, $\Delta\Delta G_{H_2O}$, versus the expected changes in the energies of charge−charge interactions according to the TKSA model, $\Delta\Delta G_{TKSA}$. (C) Changes in the transition temperature, ΔT_m, versus the expected changes in the energies of charge−charge interactions according to the UHBD model, $\Delta\Delta G_{UHBD}$. (D) Changes in the transition temperature, ΔT_m, versus the expected changes in the energies of charge−charge interactions according to the UHBD model, corrected for the interactions in the unfolded state using the Gaussian model, $\Delta\Delta G_{UHBD-Gauss}$. Vertical bars at each data point show the estimated experimental uncertainty in ΔT_m at ±0.5 °C and in $\Delta\Delta G_{H_2O}$ at ±0.4 kJ/mol. Horizontal bars at each data point show the spread in $\Delta\Delta G$ as calculated using an X-ray or NMR structural model. See the Materials and Methods on the details of the calculations.

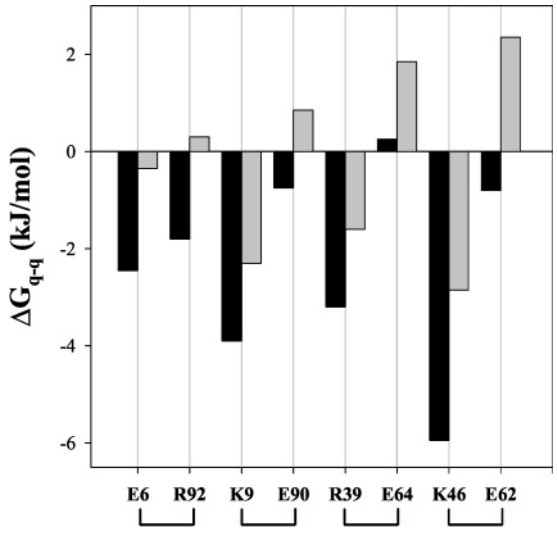

FIGURE 5: Comparison of the contributions from the total (black bars) and long-range only (gray bars) interactions for the four salt bridges in *T. celer* L30e as estimated by (A) TK-SA and (B) UHBD calculations.

The significance of long-range charge−charge interactions is further supported by the experimental data on the effect of ionic strength on protein stability. It is known that the

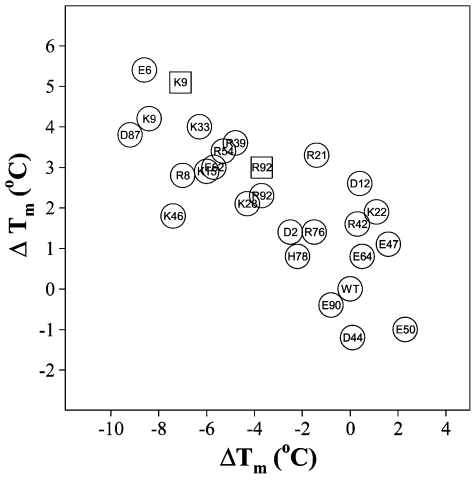

FIGURE 6: Correlation between the thermostability of the *T. celer* L30e variants, ΔT_m, and the ionic strength sensitivity, $\Delta\Delta T_m$. Data for K9M and R92M variants shown with open squares were taken from ref *20*.

increase in the ionic strength affects the strength of charge−charge interactions. However, this effect appears to be different for the short-range (e.g., salt bridges) and long-range interactions. Both experimental data (*61*) and theoretical calculations (*7*) show that the long-range charge−charge

interactions are strongly dependent upon the ionic strength, while salt bridges are weakly modulated by the changes in the salt concentration. Figure 6 shows the relationship between the difference in ionic-strength sensitivity on T_m, $\Delta\Delta T_m$, and the difference in stability of the variant and WT at low ionic strength, ΔT_m(0 M NaCl). It reveals an interesting trend: a decrease in stability leads to an increase in the ionic-strength sensitivity. Keeping in mind the notion that the changes in the ionic strength have an effect mostly on the long-range interactions, inverse correlation between ΔT_m and $\Delta\Delta T_m$ is indicative that the observed destabilization is largely due to removal of long-range charge−charge interactions.

CONCLUDING REMARKS

Analysis of the effects of charge elimination on the stability of 26 single-site variants of the *T. celer* L30e protein shows that the interactions between ionizable side chains can provide a significant contribution to the protein stability. The magnitude of these effects can be qualitatively predicted rather well using a computational model for calculating charge−charge interactions. From the practical standpoint, the effects of surface charge−charge interactions on the protein stability can be predicted with a rather good confidence (70−80%, $p < 0.003$). Importantly, the use of any computational model will produce reasonable results, and thus, this approach can find broad applications in modulating the protein stability.

ACKNOWLEDGMENT

We thank Dr. Emil Alexov for the help and advice with MCCE software, Drs. Arieh Warshel for providing temporary license and Zhen Chu for the help with installation of the MOLARIS software package to perform PDLD calculations, and Drs. Marimar Lopez and Alexey Gribenko for the discussions of the manuscript.

REFERENCES

1. Grimsley, G. R., Shaw, K. L., Fee, L. R., Alston, R. W., Huyghues-Despointes, B. M., Thurlkill, R. L., Scholtz, J. M., and Pace, C. N. (1999) Increasing protein stability by altering long-range Coulombic interactions, *Protein Sci. 8*, 1843−1849.
2. Spector, S., Wang, M., Carp, S. A., Robblee, J., Hendsch, Z. S., Fairman, R., Tidor, B., and Raleigh, D. P. (2000) Rational modification of protein stability by the mutation of charged surface residues, *Biochemistry 39*, 872−879.
3. Perl, D., Mueller, U., Heinemann, U., and Schmid, F. X. (2000) Two exposed amino acid residues confer thermostability on a cold shock protein, *Nat. Struct. Biol. 7*, 380−383.
4. Koide, A., Jordan, M. R., Horner, S. R., Batori, V., and Koide, S. (2001) Stabilization of a fibronectin type III domain by the removal of unfavorable electrostatic interactions on the protein surface, *Biochemistry 40*, 10326−10333.
5. Loladze, V. V., Ibarra-Molero, B., Sanchez-Ruiz, J. M., and Makhatadze, G. I. (1999) Engineering a thermostable protein via optimization of charge−charge interactions on the protein surface, *Biochemistry 38*, 16419−16423.
6. Sanchez-Ruiz, J. M., and Makhatadze, G. I. (2001) To charge or not to charge? *Trends Biotechnol. 19*, 132−135.
7. Dominy, B. N., Perl, D., Schmid, F. X., and Brooks, C. L., III (2002) The effects of ionic strength on protein stability: The cold shock protein family, *J. Mol. Biol. 319*, 541−554.
8. Elcock, A. H. (1998) The stability of salt bridges at high temperatures: Implications for hyperthermophilic proteins, *J. Mol. Biol. 284*, 489−502.
9. Xiao, L., and Honig, B. (1999) Electrostatic contributions to the stability of hyperthermophilic proteins, *J. Mol. Biol. 289*, 1435−1444.
10. Kumar, S., and Nussinov, R. (2001) Fluctuations in ion pairs and their stabilities in proteins, *Proteins 43*, 433−454.
11. Kumar, S., and Nussinov, R. (1999) Salt bridge stability in monomeric proteins, *J. Mol. Biol. 293*, 1241−1255.
12. Kumar, S., Ma, B., Tsai, C. J., and Nussinov, R. (2000) Electrostatic strengths of salt bridges in thermophilic and mesophilic glutamate dehydrogenase monomers, *Proteins 38*, 368−383.
13. Zhou, H. X., and Dong, F. (2003) Electrostatic contributions to the stability of a thermophilic cold shock protein, *Biophys. J. 84*, 2216−2222.
14. Dong, F., and Zhou, H. X. (2002) Electrostatic contributions to T4 lysozyme stability: Solvent-exposed charges versus semi-buried salt bridges, *Biophys. J. 83*, 1341−1347.
15. Kumar, S., and Nussinov, R. (2002) Close-range electrostatic interactions in proteins, *Chembiochem 3*, 604−617.
16. Warshel, A. (2003) Computer simulations of enzyme catalysis: Methods, progress, and insights, *Annu. Rev. Biophys. Biomol. Struct. 32*, 425−443.
17. Wong, K. B., Wang, W. K., Proctor, M. R., Bycroft, M., and Chen, Y. W. (2001) Crystallization and preliminary crystallographic studies of a ribosomal protein L30e from the hyperthermophilic archaeon *Thermococcus celer*, *Acta Crystallogr., Sect. D: Biol. Crystallogr. 57*, 865−866.
18. Chen, Y. W., Bycroft, M., and Wong, K. B. (2003) Crystal structure of ribosomal protein L30e from the extreme thermophile *Thermococcus celer*: Thermal stability and RNA binding, *Biochemistry 42*, 2857−2865.
19. Wong, K. B., Lee, C. F., Chan, S. H., Leung, T. Y., Chen, Y. W., and Bycroft, M. (2003) Solution structure and thermal stability of ribosomal protein L30e from hyperthermophilic archaeon *Thermococcus celer*, *Protein Sci. 12*, 1483−1495.
20. Lee, C. F., Allen, M. D., Bycroft, M., and Wong, K. B. (2005) Electrostatic interactions contribute to reduced heat capacity change of unfolding in a thermophilic ribosomal protein L30e, *J. Mol. Biol. 348*, 419−431.
21. Pace, C. N. (1986) Determination and analysis of urea and guanidine hydrochloride denaturation curves, *Methods Enzymol. 131*, 266−280.
22. Santoro, M. M., and Bolen, D. W. (1988) Unfolding free energy changes determined by the linear extrapolation method. 1. Unfolding of phenylmethanesulfonyl α-chymotrypsin using different denaturants, *Biochemistry 27*, 8063−8068.
23. Loladze, V. V., and Makhatadze, G. I. (2002) Removal of surface charge−charge interactions from ubiquitin leaves the protein folded and very stable, *Protein Sci. 11*, 174−177.
24. Tanford, C., and Kirkwood, J. G. (1957) Theory of protein titration curves. I. General equations for impenetrable spheres, *J. Am. Chem. Soc. 79*, 5333−5339.
25. Matthew, J. B., Gurd, F. R., Garcia-Moreno, B., Flanagan, M. A., March, K. L., and Shire, S. J. (1985) pH-dependent processes in proteins, *CRC Crit. Rev. Biochem. 18*, 91−197.
26. Matthew, J. B., and Gurd, F. R. (1986) Calculation of electrostatic interactions in proteins, *Methods Enzymol. 130*, 413−436.
27. Sali, A., and Blundell, T. L. (1993) Comparative protein modelling by satisfaction of spatial restraints, *J. Mol. Biol. 234*, 779−815.
28. Antosiewicz, J., McCammon, J. A., and Gilson, M. K. (1994) Prediction of pH-dependent properties of proteins, *J. Mol. Biol. 238*, 415−436.
29. Antosiewicz, J., McCammon, J. A., and Gilson, M. K. (1996) The determinants of pK_as in proteins, *Biochemistry 35*, 7819−7833.
30. Makhatadze, G. I., Loladze, V. V., Gribenko, A. V., and Lopez, M. M. (2004) Mechanism of thermostabilization in a designed cold shock protein with optimized surface electrostatic interactions, *J. Mol. Biol. 336*, 929−942.
31. Alexov, E. G., and Gunner, M. R. (1997) Incorporating protein conformational flexibility into the calculation of pH-dependent protein properties, *Biophys. J. 72*, 2075−2093.
32. Georgescu, R. E., Alexov, E. G., and Gunner, M. R. (2002) Combining conformational flexibility and continuum electrostatics for calculating pK_as in proteins, *Biophys. J. 83*, 1731−1748.
33. Mehler, E. L., and Guarnieri, F. (1999) A self-consistent, microenvironment modulated screened coulomb potential approximation to calculate pH-dependent electrostatic effects in proteins, *Biophys. J. 77*, 3−22.

34. Mehler, E. L., Fuxreiter, M., Simon, I., and Garcia-Moreno, E. B. (2002) The role of hydrophobic microenvironments in modulating pK_a shifts in proteins, *Proteins 48*, 283−292.
35. Schutz, C. N., and Warshel, A. (2001) What are the dielectric "constants" of proteins and how to validate electrostatic models? *Proteins 44*, 400−417.
36. Lee, P. S., Chu, Z. T., and Warshel, A. (1993) Microscopic and semimicroscopic calculations of electrostatic energies in proteins by the POLARIS and ENZYMIX programs, *J. Comput. Chem. 14*, 161−185.
37. Sham, Y. Y., Muegge, I., and Warshel, A. (1998) The effect of protein relaxation on charge−charge interactions and dielectric constants of proteins, *Biophys. J. 74*, 1744−1753.
38. Zhou, H. X. (2002) A Gaussian-chain model for treating residual charge−charge interactions in the unfolded state of proteins, *Proc. Natl. Acad. Sci. U.S.A. 99*, 3569−3574.
39. Garcia-Moreno, E. B., and Fitch, C. A. (2004) Structural interpretation of pH and salt-dependent processes in proteins with computational methods, *Methods Enzymol. 380*, 20−51.
40. Makhatadze, G. I. (1999) Thermodynamics of protein interactions with urea and guanidinium chloride, *J. Phys. Chem. 103*, 4781−4785.
41. Spassov, V. Z., Ladenstein, R., and Karshikoff, A. D. (1997) Optimization of the electrostatic interactions between ionazable groups and peptide dipoles in proteins, *Protein Sci. 6*, 1190−1196.
42. Cho, J. H., and Raleigh, D. P. (2005) Mutational analysis demonstrates that specific electrostatic interactions can play a key role in the denatured state ensemble of proteins, *J. Mol. Biol. 353*, 174−185.
43. Cho, J. H., Sato, S., and Raleigh, D. P. (2004) Thermodynamics and kinetics of non-native interactions in protein folding: A single point mutant significantly stabilizes the N-terminal domain of L9 by modulating non-native interactions in the denatured state, *J. Mol. Biol. 338*.
44. Oliveberg, M., Arcus, V. L., and Fersht, A. R. (1995) pK_a values of carboxyl groups in the native and denatured states of barnase: the pK_a values of the denatured state are on average 0.4 units lower than those of model compounds, *Biochemistry 34*, 9424−9433.
45. Trefethen, J. M., Pace, C. N., Scholtz, J. M., and Brems, D. N. (2005) Charge−charge interactions in the denatured state influence the folding kinetics of ribonuclease Sa, *Protein Sci. 14*, 1934−1938.
46. Kuhlman, B., Luisi, D. L., Young, P., and Raleigh, D. P. (1999) pK_a values and the pH dependent stability of the N-terminal domain of L9 as probes of electrostatic interactions in the denatured state. Differentiation between local and nonlocal interactions, *Biochemistry 38*, 4896−4903.
47. Pace, C. N., Alston, R. W., and Shaw, K. L. (2000) Charge−charge interactions influence the denatured state ensemble and contribute to protein stability, *Protein Sci. 9*, 1395−1398.
48. Elcock, A. H. (1999) Realistic modeling of the denatured states of proteins allows accurate calculations of the pH dependence of protein stability, *J. Mol. Biol. 294*, 1051−1062.
49. Richardson, J. S., and Richardson, D. C. (1988) Amino acid preferences for specific locations at the ends of α helices, *Science 240*, 1648−1652.
50. Presta, L. G., and Rose, G. D. (1988) Helix signals in proteins, *Science 240*, 1632−1641.
51. Aurora, R., and Rose, G. D. (1998) Helix capping, *Protein Sci. 7*, 21−38.
52. Doig, A. J., and Baldwin, R. L. (1995) N- and C-capping preferences for all 20 amino acids in α-helical peptides, *Protein Sci. 4*, 1325−1336.
53. Chakrabartty, A., Doig, A. J., and Baldwin, R. L. (1993) Helix capping propensities in peptides parallel those in proteins, *Proc. Natl. Acad. Sci. U.S.A. 90*, 11332−11336.
54. Viguera, A. R., and Serrano, L. (1995) Experimental analysis of the Schellman motif, *J. Mol. Biol. 251*, 150−160.
55. Thomas, S. T., Loladze, V. V., and Makhatadze, G. I. (2001) Hydration of the peptide backbone largely defines the thermodynamic propensity scale of residues at the C′ position of the C-capping box of α-helices, *Proc. Natl. Acad. Sci. U.S.A. 98*, 10670−10675.
56. Ermolenko, D. N., Thomas, S. T., Aurora, R., Gronenborn, A. M., and Makhatadze, G. I. (2002) Hydrophobic interactions at the Ccap position of the C-capping motif of α-helices, *J. Mol. Biol. 322*, 123−135.
57. Gong, Y., Zhou, H. X., Guo, M., and Kallenbach, N. R. (1995) Structural analysis of the N- and C-termini in a peptide with consensus sequence, *Protein Sci. 4*, 1446−1456.
58. Engel, D. E., and DeGrado, W. F. (2004) Amino acid propensities are position-dependent throughout the length of α-helices, *J. Mol. Biol. 337*, 1195−1205.
59. Barlow, D. J., and Thornton, J. M. (1983) Ion-pairs in proteins, *J. Mol. Biol. 168*, 867−885.
60. Makhatadze, G. I., Loladze, V. V., Ermolenko, D. N., Chen, X., and Thomas, S. T. (2003) Contribution of surface salt bridges to protein stability: Guidelines for protein engineering, *J. Mol. Biol. 327*, 1135−1148.
61. Lee, K. K., Fitch, C. A., and Garcia-Moreno, E. B. (2002) Distance dependence and salt sensitivity of pairwise, Coulombic interactions in a protein, *Protein Sci. 11*, 1004−1016.
62. Ibarra-Molero, B., Loladze, V. V., Makhatadze, G. I., and Sanchez-Ruiz, J. M. (1999) Thermal versus guanidine-induced unfolding of ubiquitin. An analysis in terms of the contributions from charge−charge interactions to protein stability, *Biochemistry 38*, 8138−8149.

BI0519654