Rational Modification of Protein Stability by the Mutation of Charged Surface Residues†

Shari Spector,‡,§ Minghui Wang,|| Stefan A. Carp,⊥ James Robblee,○ Zachary S. Hendsch,⊥ Robert Fairman,○ Bruce Tidor,*,⊥ and Daniel P. Raleigh*,||,⊗

Department of Physiology and Biophysics, State University of New York at Stony Brook, Stony Brook, New York 11794-8661, Department of Chemistry, State University of New York at Stony Brook, Stony Brook, New York 11794-3400, Department of Chemistry, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139-4307, Department of Molecular, Cellular and Developmental Biology, Haverford College, Haverford Pennsylvania 19041, and Graduate Programs in Biophysics and in Molecular and Cellular Biology, State University of New York at Stony Brook, Stony Brook, New York 11794

Received September 7, 1999; Revised Manuscript Received November 3, 1999

ABSTRACT: Continuum methods were used to calculate the electrostatic contributions of charged and polar side chains to the overall stability of a small 41-residue helical protein, the peripheral subunit-binding domain. The results of these calculations suggest several residues that are destabilizing, relative to hydrophobic isosteres. One position was chosen to test the results of these calculations. Arg8 is located on the surface of the protein in a region of positive electrostatic potential. The calculations suggest that Arg8 makes a significant, unfavorable electrostatic contribution to the overall stability. The experiments described in this paper represent the first direct experimental test of the theoretical methods, taking advantage of solid-phase peptide synthesis to incorporate approximately isosteric amino acid substitutions. Arg8 was replaced with norleucine (Nle), an amino acid that is hydrophobic and approximately isosteric, or with α-amino adipic acid (Aad), which is also approximately isosteric but oppositely charged. In this manner, it is possible to isolate electrostatic interactions from the effects of hydrophobic and van der Waals interactions. Both Arg8Nle and Arg8Aad are more thermostable than the wild-type sequence, testifying to the validity of the calculations. These replacements led to stability increases at 52.6 °C, the T_m of the wild-type, of 0.86 and 1.08 kcal mol^{-1}, respectively. The stability of Arg8Nle is particularly interesting as a rare case in which replacement of a surface charge with a hydrophobic residue leads to an increase in the stability of the protein.

The amino acid sequences of proteins include a wide variety of different residue types, including several acidic and basic groups. The charges on the surface of a protein are certainly important for its solubility, but what effect do electrostatic interactions have on the overall stability of the molecule? A number of recent experimental and theoretical studies have suggested that partially or completely buried salt bridges function at least in part to provide specificity to the fold, although they do not generally provide added stability beyond that of a hydrophobic bridge of similar geometry (*1, 2*). This conclusion is based on results showing that the favorable electrostatic interactions from the salt bridge are often insufficient to overcome the electrostatic desolvation penalty (*3*−*8*). Surface salt bridges appear to make only small contributions to protein stability (*9*−*12*). However, in T4 lysozyme, a partially exposed salt bridge appears to contribute 3−5 kcal mol^{-1} to the stability of the protein (*13*). Experimental studies have also been performed to examine the contribution of a single charged residue to the stability of a protein. The binding face of barstar, the inhibitor of the ribonuclease barnase, has four acidic residues. Replacement of any of these with alanine leads to an increase in the stability of the protein. On the basis of the ionic strength dependence, the increased stability of the barstar mutants is ascribed to the removal of unfavorable electrostatic interactions (*14*).

Comparison of these results is complicated by the choice of different reference states. In the T4 lysozyme study, the salt bridge in the wild-type protein is only partially exposed, and the mutation cycle involves changing each member of the salt bridge pair to asparagine, together and individually (*13*). In one barnase study investigating an existing, solvent-exposed salt bridge triad, a triple mutant cycle is used in which each residue is substituted with alanine (*10*). In another

† This research was supported by NIH grant GM 54233 to DPR who is a Pew Scholar in the Biomedical Sciences, and by NIH grants GM 55758 and GM 56552 to BT. SS was supported in part by a Graduate Council Fellowship from the State University of New York. SAC is a Beckman Scholar.

* To whom correspondence should be addressed. Bruce Tidor: telephone, 617-253-7258; fax, 617-252-1816; e-mail, tidor@mit.edu. Daniel Raleigh: telephone, 516-632-9547; fax, 516-632-7960; e-mail, draleigh@notes.cc.sunysb.edu.

‡ Department of Physiology and Biophysics, State University of New York at Stony Brook.

§ Current address: Departments of Chemistry and Biology, Massachusetts Institute of Technology 68-565, 77 Massachusetts Avenue, Cambridge, MA 02139-4307.

|| Department of Chemistry, State University of New York at Stony Brook.

⊥ Department of Chemistry, Massachusetts Institute of Technology.

○ Department of Molecular, Cellular and Developmental Biology, Haverford College.

⊗ Graduate Programs in Biophysics and in Molecular and Cellular Biology, State University of New York at Stony Brook.

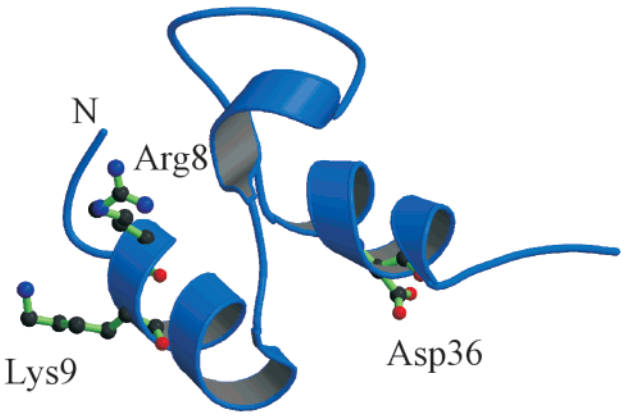
10.1021/bi992091m CCC: $19.00
Published on Web 01/13/2000

barnase study, the wild-type Asp12/Thr16 pair is compared to Ala/Thr, Asp/Arg, and Ala/Arg (*9*). Finally, in a third study of barnase, a salt bridge is engineered into an existing helix. In this case, the salt-bridging pair is compared to the wild-type sequence in a double mutant cycle in which Ser28 is replaced by glutatmate and Ala32 is replaced by lysine (*11*). Such mutant cycles assume that the interaction energy between the salt-bridging residues is purely electrostatic. It is possible for there to be hydrophobic and van der Waals interactions between residues in the wild-type sequence or in any of the mutants in the cycle, and different reference states make it difficult to compare the results of different studies. Moreover, conformational changes between mutants complicate the analysis further. These additional interactions are not accounted for in studies in which only single mutations are made.

Calculations using continuum electrostatics can estimate the contribution of a single charged residue to the overall stability of a protein. For example, such calculations can compare the stability of the protein with its native sequence to a variant in which an individual charged residue has been replaced by a hydrophobic isostere (*15*, *16*). In this manner, all other interactions within the protein are kept constant, and only the electrostatic interactions formed by a single residue with each of the other charged and polar groups in the protein are considered. Even so, there are a large number of interactions to account for in the calculations. For such calculations to be tractable, it is helpful to choose smaller proteins as models.

This study focuses on the peripheral subunit-binding domain, derived from the dihydrolipoamide acetyltransferase component (EC 2.3.1.12) of the pyruvate dehydrogenase multienzyme complex from *Bacillus stearothermophilus*. Because of its small size, 43 amino acids, it is an attractive target for such calculations. It adopts a stable, unique tertiary fold in the absence of any disulfide bridges or ligand binding (*17*−*19*). Its structure is comprised of two parallel alpha helices connected by a loop containing a short stretch of 3_{10}-helix (Figure 1). The loop is maintained in a unique conformation via a hydrogen bonding network between a buried, charged aspartate residue, Asp34, near the N-terminus of the second helix and several of the backbone amides in the loop. The variant used in the calculations and experiments described here is 41 amino acids long, corresponding to residues 3−43 of the peripheral subunit-binding domain (*18*, *19*) and will be referred to as the wild-type protein or psbd41.[1]

The small size of the peripheral subunit-binding domain offers another distinct advantage. The calculations compare the stability of variants in which the acidic and basic residues are replaced by hydrophobic isoteres. This is not often experimentally testable using natural amino acids, but some unnatural amino acids are close approximations. Since these proteins cannot be produced in high yield using traditional expression systems, solid-phase peptide synthesis must be used to prepare the large quantities of such variants required for biophysical characterization. The relatively short sequence of psbd41 means that it can easily be prepared by solid-phase peptide synthesis. The use of hydrophobic isosteres is expected to maintain all nonelectrostatic interactions between the residue of interest and the remainder of the protein and allows the isolation of electrostatic contributions to the stability of the protein. The use of isosteric amino acid substitutions makes this study the first direct experimental test of the theoretical methods.



1 11 21
VIAMPSV**R**KY AREKGVDIRL VQGTGKNGRV
31 41
LKEDIDAFLA GGA

FIGURE 1: Molscript diagram (*44*) and sequence of the peripheral subunit-binding domain. The side chains of Arg8, Lys9, and Asp36 are displayed on the ribbon diagram, the N-terminus of the protein is labeled, and the position of Arg8 in the sequence is emphasized in bold.

In this paper, we report the results of a set of continuum electrostatic calculations performed on the peripheral subunit-binding domain. The results of the calculations suggest that there are several residues located on the surface of the protein that provide a significant unfavorable electrostatic contribution to the overall stability of the domain, in part due to the asymmetry of electrostatic potential mapped to the surface of the protein. In particular, we have chosen Arg8 as a test case for the calculation. Replacement of Arg8 with norleucine, which approximates a hydrophobic isostere, leads to a significant increase in the thermal stability of the peripheral subunit-binding domain. Substitution of Arg8 with α-amino adipic acid, which is roughly isosteric but oppositely charged, leads to a further increase in thermal stability. The results of this study suggest a general strategy for increasing the stability of a protein by minimizing unfavorable surface interactions.

MATERIALS AND METHODS

Materials. Fmoc-PAL-PEG-PS resin was purchased from Perseptive Biosystems (Foster City, CA). HOBt and HBTU were purchased from Advanced ChemTech (Louisville, KY). Fmoc-L-α-amino adipic acid-*δ*-*tert*-butyl ester was from

[1] Abbreviations: Aad, α-amino adipic acid; CD, circular dichroism; Fmoc, 9- fluorenyl methoxy carbonyl; GdnHCl, guanidine hydrochloride; HBTU, 2-(1H-benzotriazole-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate; HOBt, *N*-hydroxybenzotriazole monohydrate; HPLC, high performance liquid chromatography; MALDI-TOF, matrix-assisted laser desorption and ionization time-of-flight mass spectrometry; Nle, norleucine; NMR, nuclear magnetic resonance; NOESY, nuclear Overhauser effect spectroscopy; PAL-PEG-PS, poly(ethylene glycol) polystyrene Fmoc support for peptide amides; psbd41, residues 3−43 of the peripheral subunit-binding domain; T_m, midpoint of thermal denaturation; TOCSY, total correlation spectroscopy; UV, ultraviolet.

Bachem Bioscience Inc. (King of Prussia, PA). All other Fmoc-protected amino acids were purchased from Perseptive Biosystems and Advanced ChemTech. D_2O and (trimethylsilyl)-proprionate were obtained from Cambridge Isotope Laboratories, Inc. (Andover, MA). All other solvents and reagents were obtained from Fisher Scientific (Springfield, NJ).

Calculations. The best representative (*20*) from the family of NMR structures of the peripheral subunit-binding domain, Protein Data Bank identifier 2pdd (*17*), was used for the calculations. Hydrogen atoms were placed using the HBUILD algorithm (*21*) in CHARMM (*22*) with standard pH 7 titration states for all amino acid side chains. Continuum electrostatic calculations were carried out with a modified version of the DELPHI computer program (*23*−*25*) using our previously published methods (*26*), except where differences are noted below. The PARSE parameter set (*27*) was used with protein and solvent dielectric constants of 4 and 80, respectively, a temperature of 300 K, and solvent ionic strength of 68 mM, corresponding to the experimental conditions. The protein−solvent boundary was defined as the analytic molecular surface of the protein with no dielectric smoothing applied. Computations were carried out for charging each amino acid side chain individually to permit the estimation of the electrostatic desolvation and the interaction contributions for each side chain.

Peptide Synthesis and Purification. Peptides were prepared by solid-phase synthesis using a Millipore 9050 Plus automated peptide synthesizer and standard Fmoc chemistry. Arg8Nle and Arg8Aad correspond to residues 3−43 of the peripheral subunit-binding domain in which Arg8 has been replaced by norleucine or α-amino adipic acid. Both peptides are N-terminally acetylated and C-terminally amidated. The peptides were purified by HPLC on a C18 reverse phase column (Vydac) in two steps. The solvent system used in the first step was a water−acetonitrile gradient containing 170 mM triethylamine phosphate. The second step used a water−acetonitrile gradient containing 0.1% (v/v) trifluoroacetic acid. Both peptides were greater than 95% pure as judged by HPLC.

The identity of each peptide was confirmed by matrix-assisted laser desorption and ionization time-of-flight mass spectrometry (MALDI-TOF). Arg8Nle had an experimental weight of 4383.1 Da (expected 4386.0), and Arg8Aad had an experimental weight of 4420.7 Da (expected 4416.2).

Analytical Utracentrifugation. Analytical ultracentrifugation was performed to test whether Arg8Nle and Arg8Aad are monomeric. Each sample was dialyzed against 2 mM phosphate, 2 mM borate, 2 mM citrate, 50 mM NaCl. Equilibrium experiments were performed at 25 °C with a Beckman Optima XL-A analytical ultracentrifuge using rotor speeds of 30 000, 40 000, and 50 000 rpm. Six-channel, 12 mm path length, charcoal-filled Epon cells with quartz windows were used. Ten scans were averaged. Partial specific volumes were calculated from the weighted average of the partial specific volumes of the individual amino acids and solution densities were calculated using standard tables listing coefficients for the power series approximation of density (*28*). This calculation was compared to a gravimetric determination of the solution system used. The HID program from the Analytical Ultracentrifugation Facility at the University of Connecticut was used for data analysis.

Circular Dichroism. All circular dichroism (CD) experiments described here were carried out on an Aviv 62A DS circular dichroism spectrophotometer using a buffer containing 2 mM sodium phosphate, 2 mM sodium borate, 2 mM sodium citrate, and 50 mM sodium chloride at pH 8.0. The protein concentrations for all experiments were obtained by measuring the absorbance at 276 nm in 6 M GdnHCl, 20 mM NaH_2PO_4, pH 6.5, using an extinction coefficient of 1450 M^{-1} cm^{-1}. The concentration dependence of the CD signal at 25 °C was monitored for both Arg8Nle and Arg8Aad at 222 nm, and the mean residue ellipticity was shown to be independent of concentration for all of our experimental conditions.

NMR Spectroscopy. All NMR experiments were performed on either a Varian Instruments Inova 500 MHz or Inova 600 MHz nuclear magnetic resonance spectrometer. The peptides were dissolved in 90% H_2O, 10% D_2O, pH 5.4, with (trimethylsilyl)- proprionate as a chemical shift standard. The concentrations of Arg8Nle and Arg8Aad were 1 mM and 4 mM, respectively. One-dimensional NMR spectra were acquired at 25 °C using standard presaturation methods. Two-dimensional data sets were also collected at 25 °C. For both Arg8Nle and Arg8Aad, TOCSY (total correlation spectroscopy)(*29, 30*) and NOESY (two-dimensional nuclear Overhauser enhancement spectroscopy)(*31, 32*) spectra were acquired with a spectral width of 6000.6 Hz on the 500 MHz spectrometer. The mixing times were 75 ms for the TOCSY and 250 ms for the NOESY. The collected data sets, with matrix sizes of 512 × 2048, were processed with Felix95.0 (Molecular Simulations Inc., 1995) on an SGI Indigo^2 workstation. All chemical shift assignments were made using standard procedures (*33*).

Thermal Denaturations. Thermal denaturations were monitored by far-UV CD (222 nm) and near-UV CD (280 nm) in a stirred 1 cm cuvette. The temperature was raised in 2 degree intervals from 2 to 98 °C for Arg8Nle and from 2 to 90 °C for Arg8Aad. The sample was allowed to equilibrate for 1.2 min, and the signal was averaged for 45 s. Reversibility was confirmed by comparing the ellipticity at 2 °C after a thermal denaturation to the initial ellipticity at 2 °C. Thermal denaturations of Arg8Nle were greater than 97% reversible, and for Arg8Aad, they were greater than 99% reversible. All thermal denaturations were analyzed by nonlinear least squares curve fitting using SigmaPlot (Jandel Scientific) as described previously (*18, 19*). Data are normalized to fraction unfolded. The errors in the thermodynamic parameters were analyzed using an F-test to determine the 95% confidence limits (*19, 34*).

RESULTS

Calculations. The analysis of the continuum electrostatic calculations is presented in Table 1. The electrostatic desolvation and intraprotein interaction contributions to the free energy of folding are listed for each polar or charged side chain. All of the side chains are computed to have essentially zero or net unfavorable electrostatic effects on the folding of the peripheral subunit-binding domain. The only exception is Asp17, which is buried in the structure and makes hydrogen bonds with the backbone NH groups of Arg19 and Leu20, as well as additional favorable electrostatic interactions with other neighboring groups. The

Table 1: Electrostatic Free Energies in the Peripheral Subunit-binding Domain[a]

sidechain	desolvation penalty (kcal mol^{-1})	interaction free energy (kcal mol^{-1})	total free energy (kcal mol^{-1})
Met4	0.1	0.4	0.5
Pro5	0.0	0.2	0.2
Ser6	1.8	1.0	2.8
Arg8	1.9	1.9	3.8
Lys9	1.9	2.5	4.4
Tyr10	1.5	0.3	1.8
Arg12	0.3	1.1	1.4
Glu13	0.9	−1.0	−0.1
Lys14	0.5	−1.0	−0.5
Asp17	3.3	−6.5	−3.2
Arg19	2.0	0.2	2.1
Gln22	0.3	0.4	0.6
Thr24	1.9	−0.7	1.2
Lys26	1.4	−1.0	0.4
Asn27	2.1	−0.7	1.4
Arg29	0.3	0.3	0.6
Lys32	0.2	−0.7	−0.5
Glu33	1.1	0.5	1.6
Asp34	11.2	−10.7	0.5
Asp36	1.5	0.6	2.1
Phe38	0.6	0.2	0.8

[a] The desolvation penalty describes the difference in solvation free energy of a residue in the native versus denatured state. The interaction free energy describes the energetics of the electrostatic interactions between a residue and the remainder of the protein. The total free energy is the sum of the desolvation penalty and the interaction free energy.

total effect of Asp17 is computed to be favorable by 3.2 kcal mol^{-1}. The most surprising result of the calculations, however, is the large number of amino acid side chains computed to have *unfavorable* interactions in the folded structure of the protein. One generally expects *favorable* folded state interactions that are offset by unfavorable desolvation penalties. Examination of the structure reveals that these repulsions include a grouping of positively charged residues near the surface, Arg8, Lys9, and Arg12, which lie along the exposed face of an α-helix. Figure 2 shows the resulting electrostatic potential. We are cautious about interpreting the results of calculations based on the best representative from a family of NMR structures determined for a small protein whose structure, particularly at the protein surface, is likely to be fluctuating. For this reasons, we feel that averaging over many conformations, which is beyond the scope of the current report, may improve the accuracy of the values in Table 1. Nevertheless, it is reasonable to expect unfavorable effects for the positive surface cluster because each side chain is clearly partially desolvated and there are certainly repulsions among the members of the set.

Amino Acid Substitutions. The calculations suggest that three residues, Arg8, Lys9, and Asp36, make a significant unfavorable electrostatic contribution to the overall stability of the peripheral subunit-binding domain. Of these, Arg8 was chosen to test the results of the calculations. It is located on the surface of the peripheral subunit-binding domain in a region of strong positive electrostatic potential (Figure 2). Arg8 is the second residue in a helix, and therefore, its charged guanidino group may interact unfavorably with backbone dipolar groups in the helix. In addition, there is another arginine on the same face of the helix at position 12, four residues away from Arg8, and these two residues could also interact unfavorably. Replacement of Arg8 with a hydrophobic residue should eliminate these unfavorable electrostatic interactions. Substitution with a negatively charged residue could provide further stability through favorable salt-bridge and backbone dipole interactions, as well as through other interactions with the local positive potential.

On this basis, Arg8 was replaced both with a hydrophobic residue and with a negatively charged residue, each similar to arginine in size and shape. For the hydrophobic substitution, Arg8 was replaced with norleucine (Arg8Nle), which has a straight chain aliphatic side chain four carbons long, an arginine analogue with a methyl group in place of the guanidino group. For the substitution of opposite charge, α-amino adipic acid (Arg8Aad) was chosen. This unnatural amino acid also has the same number of methylene groups as arginine, but the terminal guanidino group is replaced by a carboxylate.

Surface Charge Variants of Psbd41 are Monomeric. Both Arg8Nle and Arg8Aad remain monomeric throughout the concentration range of the experiments reported here. It was especially important to test this for Arg8Nle, since replacement of a surface charge with a hydrophobic residue could result in a sticky patch prone to association or aggregation. However, Arg8Nle is monomeric. The molar ellipticity at 222 nm is independent of concentration over the range 24−460 μM. Furthermore, the one-dimensional NMR spectrum is also identical for samples at 440 μM and 3.8 mM. If aggregation were to occur, broadening of the NMR lines would be expected, but the lines remain sharp at the higher concentration. Analytical ultracentrifugation also shows that Arg8Nle is monomeric. For a 242 μM sample, a single species fit gave a molecular weight of 4600 ± 200 Da, compared to a calculated molecular weight of 4386.0 Da, consistent with Arg8Nle remaining monmeric at this concentration. Fits with multiple species models were no better than the single species fit, as judged by the randomness of the residuals (data not shown). Finally, thermal denaturations were performed at 12 μM and 582 μM (see below), resulting, within the experimental uncertainty, in identical thermal denaturation midpoints and identical values of the enthalpy at the midpoint of the transition. All of these taken together provide strong evidence that Arg8Nle remains monomeric over the concentration range of interest.

Arg8Aad is also monomeric. The molar ellipticity at 222 nm is independent of concentration over the studied range of 58 μM to 1.42 mM. The one-dimensional NMR spectra of 500 μM and 4 mM Arg8Aad are also identical, which would be unlikely if aggregation were occurring. Analytical ultracentrifugation results confirm that at 343 μM Arg8Aad is monomeric. Using a single species analysis results in a molecular weight of 4600 ± 200 Da, compared to the calculated molecular weight of 4416.2 Da, and there is no improvement in the fit using multiple species models (data not shown). Finally, the midpoint of the thermal denaturation and the enthalpy at the midpoint are identical for 15 μM and 440 μM samples of Arg8Aad, providing additional evidence that the protein remains monomeric over the concentration range used for the experiments in this paper.

Arg8Nle and Arg8Aad adopt the same structure as wild-type psbd41. Spectroscopic evidence suggests that both Arg8Nle and Arg8Aad adopt essentially the same fold as the wild-type protein. The near-UV CD spectra have the same shape, and the signal at 280 nm at 25 °C is similar for the

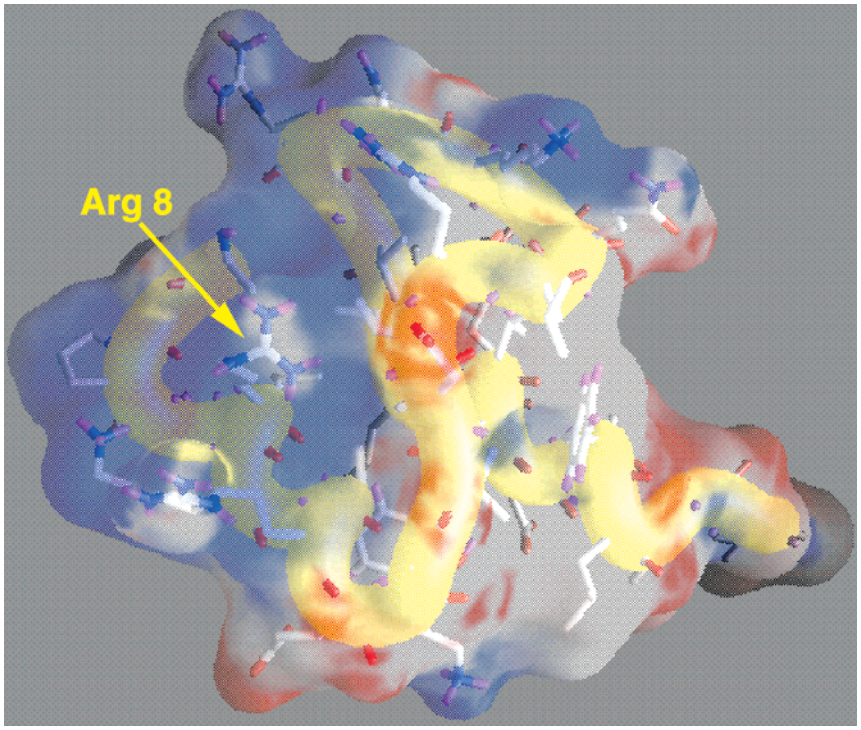


FIGURE 2: GRASP figure of the peripheral subunit-binding domain, shown in the same orientation as in Figure 1. Negative and positive values of electrostatic potential are indicated by linearly deepening shades of red and blue, respectively.

three proteins, with values of 75.9 deg cm^2 dmol^{-1} for Arg8Nle, 104.1 deg cm^2 dmol^{-1} for Arg8Aad, and 119.5 deg cm^2 dmol^{-1} for the wild-type. The far-UV CD spectra are also almost identical in shape, with $[\Theta]_{222}$ values at 25 °C of −10 600 deg cm^2 dmol^{-1} for Arg8Nle, −11 000 deg cm^2 dmol^{-1} for Arg8Aad, and −11 100 deg cm^2 dmol^{-1} for wild-type (data not shown).

Stronger evidence that the three proteins adopt the same structure comes from their one- and two-dimensional NMR spectra and their chemical shift assignments. The one-dimensional NMR spectra all show several very characteristic peaks. First are the two sets of ring-current-shifted methyl protons from Val16 and Val21, which are both clearly present in the three spectra (Figure 3A). For the wild-type protein, Val16 and Val21 appear at 0.42 and 0.23 ppm, respectively. They resonate at 0.41 and 0.23 ppm in Arg8Nle, and 0.42 and 0.23 ppm in Arg8Aad. Another characteristic resonance that is well-resolved in the one-dimensional spectrum arises from the amide proton of Thr24, which is hydrogen bonded to the buried, charged aspartate residue at position 34. As a result, it appears significantly downfield at 9.98 ppm in the wild-type protein. In Arg8Nle, this proton resonates at 9.94 ppm, and in Arg8Aad, its chemical shift is also 9.94 ppm. If the structure of Arg8Nle or Arg8Aad were significantly different from that of the wild-type protein, much larger differences in these chemical shifts would be expected.

Nearly complete resonance assignments were possible for both surface charge variants and are available from the authors upon request. The NMR assignments for the peptide backbone provide additional evidence that the surface charge variants adopt the same structure as the wild-type protein. For Arg8Nle, none of the assigned $C^{\alpha}H$ resonances has a chemical shift that differs from the wild-type protein by more than 0.1 ppm. In contrast, the $C^{\alpha}H$ chemical shifts differ from random coil values (*35*) by between −0.62 and 0.42 ppm (Figure 3B). Similarly, only one amide chemical shift, that of Met4, differs from the wild-type assignments by more than 0.20 ppm, and this residue is near the N-terminus. The assignments for Arg8Aad also agree well with those for the wild-type protein. Only two of the assigned C^{α} protons have a chemical shift different from the wild-type by more than 0.06 ppm (Figure 3C). In contrast, the C^{α} proton chemical shifts in Arg8Aad differ from average random coil values by between −0.41 and +0.64 ppm. The NH chemical shifts are also quite similar for Arg8Aad and wild-type psbd41, with only two residues having chemical shifts that differ by more than 0.11 ppm. Since the chemical shift of a proton is so sensitive to its environment, it is highly unlikely that the proteins could have such similar NMR spectra if they adopted different structures.

Stability Measurements. Thermal denaturation of the peripheral subunit-binding domain and the two surface charge variants shows that substitution of Arg8 results in an increase in thermal stability for both variants (Figure 4). In addition, all three proteins undergo two-state folding, as evidenced by the excellent agreement between the values of T_m and $\Delta H°$ (T_m) for thermal denaturations monitored by near- and far-UV CD spectroscopy. Averaged values are reported in Table 2. The uncertainties given here are determined by F-value analysis of the nonlinear regressions. Since such errors are often asymmetric, we report the larger limit on the error. The wild-type protein has a T_m of 53.1 ± 0.8 °C by far-UV CD and 52.1 ± 1.4 °C by near-UV CD, while $\Delta H°$ (T_m) is 33.4 ± 3.3 kcal mol^{-1} by far-UV CD and 30.1 ± 5.0 kcal mol^{-1} by near-UV CD. Replacement of Arg8 with norleucine results in an increase in the T_m to 61.9 ± 1.1 °C by far-UV CD and 61.4 ± 1.1 °C by near-UV CD. The $\Delta H°$ (T_m) is essentially unchanged, with values of 33.9 ± 4.4 kcal mol^{-1} determined by far-UV CD, and 33.7 ± 4.2 kcal mol^{-1} determined by near-UV CD. Arg8Aad is

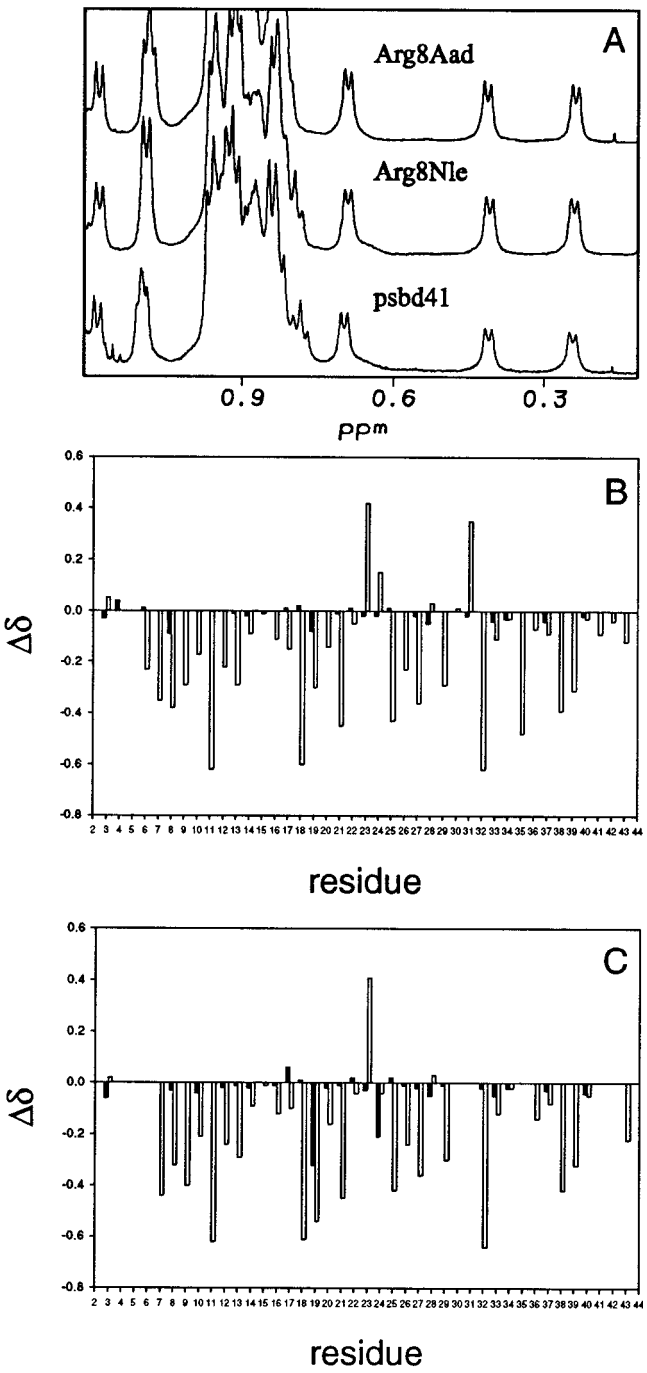


FIGURE 3: (A): One-dimensional NMR spectra of psbd41 (bottom), Arg8Nle (center), and Arg8Aad (top). (B, C): $C^{\alpha}H$ chemical shift differences between the surface charge variants and psbd41 (filled bars) or random coil values (open bars) (*35*). B, Arg8Nle; C, Arg8Aad.

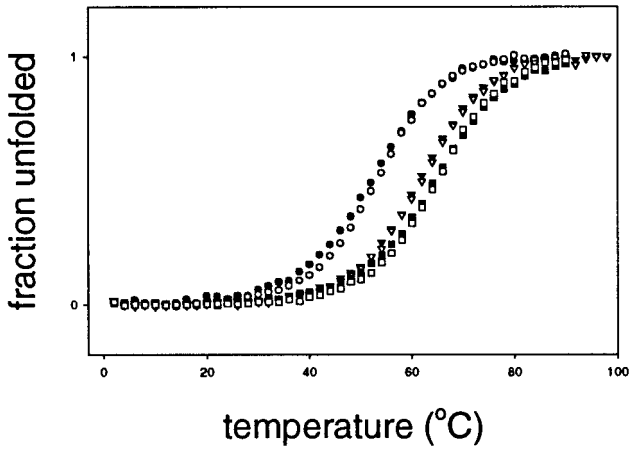


FIGURE 4: Thermal denaturations. Open symbols correspond to far-UV CD data collected at 222 nm, and solid symbols correspond to near-UV CD data collected at 280 nm. ○: 14.5 μM psbd41. ●: 435 μM psbd41. ▽: 12 μM Arg8Nle. ▼: 582 μM Arg8Nle. □: 15 μM Arg8Aad. ■: 440 μM Arg8Aad.

further stabilized, relative to Arg8Nle, with a T_m of 64.7 ± 1.7 °C by far-UV CD and 64.2 ± 2.1 °C by near-UV CD. The $\Delta H°$ (T_m) remains similar to the wild-type and to Arg8Nle, with values of 34.9 ± 4.6 kcal mol^{-1} determined by far-UV CD, and 31.5 ± 4.3 kcal mol^{-1} determined by near-UV CD.

Thermal denaturation experiments are analyzed using the Gibbs−Helmholtz equation. To do this requires knowledge of the heat capacity change, $\Delta C°_p$. We used the value determined previously for the peripheral subunit-binding domain, 0.43 kcal mol^{-1} K^{-1} (*19*), not only for the analysis of the data from the wild-type protein, but also for the two surface charge variants, and we make the additional assumption that $\Delta C°_p$ is independent of temperature. The heat capacity change is related to the difference in accessible surface area between the native and denatured states (*36*). On the basis of the NMR and CD data, Arg8Nle and Arg8Aad both adopt the same tertiary structure as the wild-type protein. Thus, they should have the same heat capacity change as the wild-type, unless the mutations perturb the structure of the denatured state ensemble. For Arg8Aad, this is unlikely, since the carboxylate on α-amino adipic acid will remain charged and solvent exposed at pH 8. However, the norleucine in Arg8Nle could potentially participate in a non-native hydrophobic cluster in the unfolded state, which might affect the heat capacity. The expected effect on $\Delta C°_p$ would be small, probably at least an order of magnitude less than the uncertainty in the measured value (0.43 ± 0.25 kcal mol^{-1} K^{-1}). The magnitude of the uncertainty results from the value being at the lower limit of what is measurable. The heat capacity could not be determined by differential scanning microcalorimetry, because the transition was too broad for an accurate analysis and instead was determined by the global analysis of thermal denaturations performed in the presence of a variety of chemical denaturant concentrations (*19*). The difference between the measured value and that for the wild-type is certain to be less than the uncertainty in the numbers. In addition, analysis of the thermal denaturations of the wild-type protein using a variety of values of $\Delta C°_p$ does not significantly affect the results. The T_m of psbd41 is insensitive to the value of the heat capacity used in the analysis, and the enthalpy at the midpoint does not change, within the uncertainty in the measurement, using heat capacities ranging from 0.30 to 0.65 kcal mol^{-1} K^{-1} (*18, 19*). The value of $\Delta C°_p$ determined for the wild-type, therefore, provides a reasonable estimate of the heat capacity of the surface charge variants.

Using the value of $\Delta C°_p$ for the wild-type protein and the measured values of $\Delta H°$ (T_m) and T_m from the surface charge variants, we have calculated $\Delta\Delta G°_{D-N}$ at the T_m of the wild-type protein. At 52.6 °C, $\Delta\Delta G°_{D-N}$ is equal to 0.86 kcal mol^{-1} for Arg8Nle, and 1.08 kcal mol^{-1} for Arg8Aad (Table 2). Evaluating the equation at 27 °C, the temperature assumed in the calculations, leads to $\Delta\Delta G°_{D-N}$ values of 0.65 kcal

Table 2: Thermodynamic Properties of the Peripheral Subunit-binding Domain and the Surface Charge Variants

	T_m[a] (°C)	$\Delta H°$ (T_m) (kcal mol^{-1})	$\Delta G°_{D-N}$(27 °C)[b] (kcal mol^{-1})	$\Delta\Delta G°_{D-N}$(27 °C) (kcal mol^{-1})	$\Delta\Delta G°_{D-N}$(52.6 °C) (kcal mol^{-1})
psbd41	52.6	31.8	2.06		
Arg8Nle	61.7	33.8	2.71	0.65	0.86
Arg8Aad	64.5	33.2	2.76	0.70	1.08

[a] T_m and $\Delta H°(T_m)$ are the average of the values obtained from the analysis of thermal denaturations monitored by near- and far-UV CD. [b] $\Delta G°_{D-N}$(27 °C) and $\Delta G°_{D-N}$(52.6 °C) are obtained using the Gibbs−Helmholtz equation, the values of T_m and $\Delta H°(T_m)$ from this table.

mol^{-1} for Arg8Nle and 0.70 kcal mol^{-1} for Arg8Aad (Table 2). However, estimating the stability at 27 °C requires a long extrapolation and leads to greater uncertainty in the values of $\Delta\Delta G°_{D-N}$. Calculations using the range of $\Delta H°$ (T_m) and $\Delta C°_p$ suggested by the uncertainties in the parameters result in only small changes in the stability at 53 °C, because the extrapolation is, at most, 11°, whereas there is much greater variation at 27 °C. Thus, as predicted by the calculations, substitution of Arg8 with a hydrophobic residue of approximately the same size and shape results in an increase in the stability of the protein. The difference in stability between the surface charge variants and the wild-type protein is greater at the T_m of the wild-type than at 27 °C, and these values are also included in Table 2.

Under ideal circumstances, the stabilities of the three proteins should also be compared by measurement using chemical denaturation. However, for this set of proteins there was no good choice of chemical denaturant. The two most common denaturants are guanidine hydrochloride and urea. It is not possible to obtain accurate thermodynamic parameters from a urea denaturation of the peripheral subunit-binding domain, because it is so small. Like the heat capacity, the *m*-value of chemical denaturation also depends on the difference in accessible surface area between the native and denatured states (*36*). Because psbd41 is so small, it has a low *m*-value, and therefore, very broad transitions. Guanidine hydrochloride, on the other hand, is a salt. Since ionic strength affects the stability of psbd41 (*19*), and because we are interested in electrostatic effects, guanidine is also not a suitable denaturant.

DISCUSSION

The continuum electrostatic methods used here work well to predict the role an individual charged residue plays in the stability of a protein. Arg8 is predicted to be destabilizing by 3.8 kcal mol^{-1}. Substitution of Arg8 with a hydrophobic residue of similar size and shape leads to an increase in stability at 27 °C of 0.65 kcal mol^{-1}, which is significant but much smaller than predicted. Part of this discrepancy could be due to the somewhat smaller size of norleucine relative to arginine. The difference in surface area buried could be responsible for roughly 1.5 kcal mol^{-1}, or half the total discrepancy (*37*). Static-structure continuum calculations with low internal dielectric may somewhat overestimate the size of mutational effects on stability (*38*). While increasing the value used for the internal dielectric may improve the agreement with experiment in some instances, it is more likely that explicitly sampling conformational degrees of freedom will be more appropriate. At 52.6 °C, the T_m of the wild-type, $\Delta\Delta G°_{D-N}$ is 0.86 kcal mol^{-1}. The replacement of a surface charge with a hydrophobic residue would not normally be expected to increase the stability of the protein, because it is more favorable for a hydrophobic side chain to be buried within the core of the protein, rather than exposed to solvent. However, because of its unusual environment the charge on Arg8 is unfavorable, and in this case the substitution leads to an increase in stability. This is in contrast to observations made with λ Cro protein, in which substitution of a tyrosine on the surface of the protein with smaller hydrophobic residues or with polar or charged residues leads to an increase in its stability. This phenomenon was dubbed the reverse hydrophobic effect because the residue in question becomes more exposed in the native structure than it is in the unfolded state (*39*). Obviously, any reverse hydrophobic effect occurring in Arg8Nle is compensated by the removal of unfavorable electrostatic interactions. The calculations do not take into account any changes in the denatured state, and a reverse hydrophobic effect would be expected to stabilize the denatured state of the hydrophobic mutant. This may account for some of the discrepancy between the calculated and experimental $\Delta\Delta G°_{D-N}$ values. In addition, the calculations assume a single static structure and do not take into account side chain dynamics. This may also contribute to the difference between the observed and calculated stability changes.

Substitution of Arg8 with α-amino adipic acid leads to a further increase in stability at 27 °C, with a $\Delta\Delta G°_{D-N}$ of 0.70 kcal mol^{-1} relative to wild-type, or 1.08 kcal mol^{-1} at 52.6 °C. Calculations similar to those described here suggest Arg8Aad to be about 1.5 kcal/mol more stable than Arg8Nle (data not shown); again, the correct direction but an overestimate of the magnitude. While in Arg8Nle a number of unfavorable electrostatic interactions are alleviated, Arg8Aad can overcome the unfavorable exposure of the norleucine side chain to solvent and may make several favorable electrostatic interactions with the remainder of the protein. The α-amino adipic acid side chain has the potential to interact favorably with helix backbone dipolar groups, and to form a salt bridge with Arg12.

Another interesting residue described in the calculations is Asp34. This residue is more than 95% buried, and its side chain takes part in hydrogen bonds to the backbone amides of Gly23, Thr24, Gly25, and Leu31 in the loop, and to the side chain hydroxyl group of Thr24. A buried charge is normally expected to be extremely unfavorable due to the large desolvation penalty (*3*). The calculations suggest that in this case the burial of Asp34 is only modestly destabilizing by 0.6 kcal mol^{-1}. This is likely due to the extensive hydrogen bonding network. Our previously reported studies of Asn and Val substitutions at this position have demonstrated that Asp34 is important for the specificity of the fold and also contributes to the stability of the domain (*18, 40*). These studies are in qualitative agreement with the calculations.

The amino acid changes chosen for this study have a distinct advantage, arising from the use of hydrophobic isosteres. It is difficult to choose a reference state such that only electrostatic interactions are considered. Although norleucine and α-amino adipic acid are not perfectly isosteric with arginine, they serve as excellent approximations. Through these changes we were able to overcome unfavorable electrostatic interactions and increase the stability of the peripheral subunit-binding domain.

The substitutions were made in the context of a model study. Nevertheless, it is instructive to consider their effect in a biological context. The peripheral subunit-binding domain is a piece of a much larger enzyme, the dihydrolipoamide acetyltransferase (E2). The domain's function within the pyruvate dehydrogenase multienzyme complex is to direct intermolecular interactions with each of the other two enzymes in the complex, the pyruvate decarboxylase (E1, EC 1.2.4.1) and the dihydrolipoamide dehydrogenase (E3, EC 1.8.1.4). The molecular details of the interactions between the peripheral subunit-binding domain and E1 have not been characterized, but there is a crystal structure available of the peripheral subunit-binding domain bound to E3 (*41*). In the complex, Arg8 is important for binding, forming a salt bridge with Glu431 of E3. Although the substitutions described in this paper would not serve this protein well in vivo, the methodology could nevertheless be applied to other proteins if care is taken to avoid residues involved in catalysis or intermolecular interactions.

Relatively little attention has been paid to the contributions of surface electrostatic interactions to the stability of globular proteins. This study provides a clear demonstration that alleviating unfavorable surface interactions can increase the stability of proteins. Many proteins contain clusters of positively or negatively charged residues, and the results presented here suggest that optimization of surface electrostatic interactions is likely to be a generally applicable strategy for enhancing protein stability. For example, in recent studies of ribonuclease T1 and ubiquitin, it was shown that relieving surface charge repulsion through mutation increased protein stability (*42*, *43*). These methods may prove useful, for example, in structural studies of marginally stable proteins, since surface mutations are much less likely to perturb the structure than a mutation to the core, or in membrane-associated proteins.

REFERENCES

1. Lumb, K. J., and Kim, P. S. (1995) *Biochemistry 34*, 8642−8648.
2. Raleigh, D. P., Betz, S. F., and DeGrado, W. F. (1995) *J. Am. Chem. Soc. 117*, 7558−7559.
3. Hendsch, Z. S., and Tidor, B. (1994) *Protein Sci. 3*, 211−226.
4. Honig, B., and Yang, A.-S. (1995) *Adv. Prot. Chem. 46*, 27−58.
5. Sindelar, C. V., Hendsch, Z. S., and Tidor, B. (1998) *Protein Sci. 7*, 1898−1914.
6. Waldburger, C. D., Schildbach, J. F., and Sauer, R. T. (1995) *Nat. Struct. Biol. 2*, 122−128.
7. Wang, L., O'Connell, T., Tropsha, A., and Hermans, J. (1996) *Biopolymers 39*, 479−489.
8. Wimley, W. C., Gawrisch, K., Creamer, T. P., and White, S. H. (1996) *Proc. Natl. Acad. Sci. U.S.A. 93*, 2985−2990.
9. Serrano, L., Horovitz, A., Avron, B., Bycroft, M., and Fersht, A. R. (1990) *Biochemistry 29*, 9343−9352.
10. Horovitz, A., Serrano, L., Avron, B., Bycroft, M., and Fersht, A. R. (1990) *J. Mol. Biol. 216*, 1031−1044.
11. Šali, D., Bycroft, M., and Fersht, A. R. (1991) *J. Mol. Biol. 220*, 779−788.
12. Xiao, L., and Honig, B. (1999) *J. Mol. Biol. 289*, 1435−1444.
13. Anderson, D. E., Becktel, W. J., and Dahlquist, F. W. (1990) *Biochemistry 29*, 2403−2408.
14. Schreiber, G., Buckle, A. M., and Fersht, A. R. (1994) *Structure 2*, 945−951.
15. Honig, B., and Nicholls, A. (1995) *Science 268*, 1144−1149.
16. Davis, M. E., and McCammon, J. A. (1990) *Chem. Rev. 90*, 509−521.
17. Kalia, Y. N., Brocklehurst, S. M., Hipps, D. S., Appella, E., Sakaguchi, K., and Perham, R. N. (1993) *J. Mol. Biol. 230*, 323−341.
18. Spector, S., Kuhlman, B., Fairman, R., Wong, E., Boice, J. A., and Raleigh, D. P. (1998) *J. Mol. Biol. 276*, 479−489.
19. Spector, S., Young, P., and Raleigh, D. P. (1999) *Biochemistry 38*, 4128−4136.
20. Kelley, L. A., Gardner, S. P., and Sutcliffe, M. J. (1996) *Protein Eng. 9*, 1063−1065.
21. Brünger, A. T., and Karplus, M. (1988) *Proteins: Struct., Funct., Genet. 4*, 148−156.
22. Brooks, B. R., Bruccoleri, R. E., Olafson, B. D., States, D. J., Swaminathan, S., and Karplus, M. (1983) *J. Comput. Chem. 4*, 187−217.
23. Gilson, M. K., and Honig, B. (1988) *Proteins: Struct., Funct., Genet. 4*, 7−18.
24. Gilson, M. K., Sharp, K. A., and Honig, B. (1988) *J. Comput. Chem. 9*, 327−335.
25. Sharp, K. A., and Honig, B. (1990) *Annu. Rev. Biophys. Biophys. Chem. 19*, 301−332.
26. Hendsch, Z. S., Sindelar, C. V., and Tidor, B. (1998) *J. Phys. Chem. B 102*, 4404−4410.
27. Sitkoff, D., Sharp, K. A., and Honig, B. (1994) *J. Phys. Chem. 98*, 1978−1988.
28. Laue, T. M., Shah, B. D., Ridgeway, T. M., and Pelletier, S. L. (1992) in *Analytical Ultracentrifugation in Biochemistry and Polymer Science* (Harding, S. E., Rowe, A. J., and Horton, J. C., Eds.) pp 90−125, The Royal Society of Chemistry, Cambridge.
29. Braunschweiler, L., and Ernst, R. R. (1983) *J. Magn. Reson. 53*, 521−528.
30. Davies, D. G., and Bax, A. (1985) *J. Am. Chem. Soc. 107*, 2820−2821.
31. Jeener, J., Meier, B. H., Bachmann, P., and Ernst, R. R. (1979) *J. Chem. Phys. 71*, 4546−4554.
32. Macura, S., Hyang, Y., Suter, D., and Ernst, R. R. (1981) *J. Magn. Reson. 53*, 259−281.
33. Wüthrich, K. (1986) *NMR of Proteins and Nucleic Acids*, Wiley and Sons, New York.
34. Shoemaker, D. P., Garland, C. W., and Nibler, J. W. (1989) *Experiments in Physical Chemistry*, 5th ed., McGraw-Hill Publishing Company, New York.
35. Wishart, D. S., Bigam, C. G., Holm, A., Hodges, R. S., and Sykes, B. D. (1995) *J. Biomolecular NMR 5*, 67−81.
36. Myers, J. K., Pace, C. N., and Scholtz, J. M. (1995) *Protein Sci. 4*, 2138−2148.
37. Pace, C. N. (1992) *J. Mol. Biol. 226*, 29−35.
38. Hendsch, Z. S., Jonsson, T., Sauer, R. T., and Tidor, B. (1996) *Biochemistry 35*, 7621−7625.
39. Pakula, A. A., and Sauer, R. T. (1990) *Nature 344*, 363−364.
40. Spector, S., Rosconi, M., and Raleigh, D. P. (1999) *Biopolymers 49*, 29−40.
41. Mande, S. S., Sarfaty, S., Allen, M. D., Perham, R. N., and Hol, W. G. J. (1996) *Structure 4*, 277−286.
42. Grimsley, G. R., Shaw, K. L., Fee, L. R., Alston, R. W., Huyghues-Despointes, B. M. P., Thurlkill, R. L., Scholtz, J. M., and Pace, C. N. (1999) *Protein Sci. 8*, 1843−1849.
43. Loladze, V. V., Ibarra-Molero, B., Sanchez-Ruiz, J. M., and Makhatadze, G. I. (1999) *Biochemistry 38*, 16419−16423.
44. Kraulis, P. J. (1991) *J. Appl. Crystallogr. 24*, 946−950.

BI992091M