Rational Design, Structural and Thermodynamic Characterization of a Hyperstable Variant of the Villin Headpiece Helical Subdomain†

Yuan Bi,‡ Jae-Hyun Cho,§ Eun-Young Kim,§ Bing Shan,‡ Hermann Schindelin,§,|| and Daniel P. Raleigh*,‡,§

Department of Chemistry, State University of New York at Stony Brook, Stony Brook, New York 11794-3400, Graduate Program in Biochemistry and Structural Biology, State University of New York at Stony Brook, Stony Brook, New York 11794-5230, and Rudolf Virchow Center for Experimental Biomedicine and Institute of Structural Biology, Univeristy of Würzburg, Versbacher Str. 9, 97078 Würzburg, Germany

Received December 21, 2006; Revised Manuscript Received March 22, 2007

ABSTRACT: A hyperstable variant of the small independently folded helical subdomain (HP36) derived from the F-actin binding villin headpiece was designed by targeting surface electrostatic interactions and helical propensity. A double mutant N68A, K70M was significantly more stable than wild type. The T_m of wild type in aqueous buffer is 73.0 °C, whereas the double mutant did not display a complete unfolding transition. The double mutant could not be completely unfolded even by 10 M urea. In 3 M urea, the T_m of wild type is 54.8 °C while that of the N68AK70M double mutant is 73.9 °C. Amide H/^{2}H exchange studies show that the pattern of exchange is very similar for wild type and the double mutant. The structures of a K70M single mutant and the double mutant were determined by X-ray crystallography and are identical to that of the wild type. Analytical ultracentrifugation demonstrates that the proteins are monomeric. The hyperstable mutant described here is expected to be useful for folding studies of HP36 because studies of the wild type domain have sometimes been limited by its marginal stability. The results provide direct evidence that naturally occurring miniature protein domains have not been evolutionarily optimized for global stability. The stabilizing effect of this double mutant could not be predicted by sequence analysis because K70 is conserved in the larger intact headpiece for functional reasons.

A common goal of protein engineering studies is to increase protein stability. Increased protein stability is important for applications in biotechnology because the performance of protein pharmaceuticals and proteins used as biocatalysts can often be improved by stabilization (*1*−*3*). Enhanced stability can also be extremely useful in biophysical studies, particularly for the study of marginally stable proteins and protein domains. In this case, protein engineering experiments are often of limited utility because even modestly destabilizing mutations can drastically depopulate the folded state. There are also theoretical reasons why developing more stable proteins is of interest, particularly for proteins that fold rapidly. For example, recent work in protein folding has led to the hypothesis that stabilizing a rapidly folding protein can lead to downhill folding, that is, folding that does not involve crossing a free energy barrier (*4*). The analysis predicts that stabilizing a protein that folds at a rate near the speed limit for folding could lead to downhill folding. A wide variety of approaches have been applied to increase protein stability, sometimes with mixed success (*5*−*9*). Here, we describe the rational design of a hyperstable variant of the small, fast folding villin headpiece helical subdomain (HP36[1]).

HP36 is the helical C-terminal subdomain from the villin headpiece and can fold in isolation. The subdomain is one of the smallest known naturally occurring proteins that folds to a compact native state. The complete villin headpiece domain is a 76 residue F-actin binding domain that is found at the extreme C-terminus of villin and related actin binding proteins involved in cytoskeleton remodeling (*10*−*14*). Headpiece domains are not related to any other F-actin binding proteins. Intact villin is a multidomain protein comprising a set of gelsolin domains and headpiece. It has the interesting property of being able to act as both an actin-bundling protein and an actin-severing protein. Actin bundling requires two independent actin binding sites, and the

† This work was supported by a grant from the NSF (MCB-0614365) to D.P.R.

* To whom correspondence should be addressed. Tel: 631-632-9547. Fax: 631-632-7960. E-mail: draleigh@notes.cc.sunysb.edu.

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

§ Graduate Program in Biochemistry and Structural Biology, State University of New York at Stony Brook.

|| Univeristy of Würzburg.

[1] Abbreviations: AUC, analytical ultracentrifugation; CD, circular dichroism; C_M, the midpoint of the denaturant induced unfolding transition; ESI MS, electro-spray ionization mass spectrometry; $\Delta Cp°$, the change in heat capacity upon unfolding; $\Delta H°$, the change in enthalpy upon unfolding; $\Delta G°(T)$, the change in free energy upon unfolding measured at a temperature T; HP36, the villin headpiece helical subdomain corresponding to residues Leu42−Phe76 of the headpiece with an additional Met at the N-terminus; HP36 N68A, the Asn68 to Ala mutant of HP36; HP36 K70M, the Lys70 to Met mutant of HP36; HP36 N68AK70M, the Asn68 to Ala and the Lys70 to Met double mutant of HP36; HP67, the C-terminal 67 residues of the villin headpiece (this is the smallest characterized headpiece construct which exhibits full affinity actin binding); HSQC, heteronuclear single quantum coherence; m, the m-value for unfolding; MALDI-TOF, matrix-assisted laser desorption and ionization time-of-flight mass spectrometry; NMR, nuclear magnetic resonance; rmsd, root mean square deviation.

10.1021/bi6026314 CCC: $37.00
Published on Web 05/31/2007

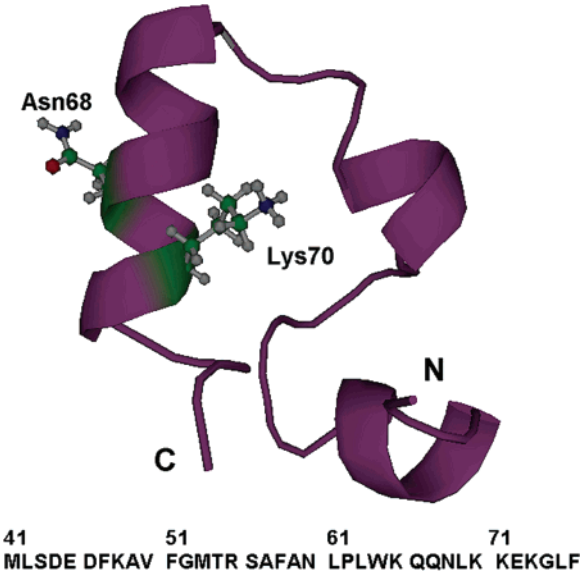

FIGURE 1: Ribbon diagram and primary sequence of the villin headpiece subdomain (HP36). The side chains of Asn68 and Lys70 are shown in ball-and-stick representation, created by PyMOL, version 0.99 (*51*).

headpiece domain provides the second F-actin binding site (*10*, *11*). The first nine residues of the headpiece can be removed without compressing actin binding activity, and the resulting construct is commonly referred to as HP67 (*12*). The N-terminal portion of the protein is made up of a set of loops, turns, and a short helix, whereas the C-terminal portion, here designated HP36, comprises a small three helix bundle (*12*, *13*, *15*). HP67 is the smallest characterized fragment of the headpiece that retains full actin binding activity. Extensive structural and mutational studies from the McKnight group have led to a model of the binding site (*12*−*14*). Key residues include K38, E39, K65, and the sequence comprising residues 70−74 and possibly L75 and F76. E39 and K70 are involved in a structurally important buried salt bridge and are important for maintaining the structural integrity of the protein, whereas the other residues group together on the protein surface. HP67 contains a single contiguous hydrophobic core, but none the less, the C-terminal helical subdomain can fold independently.

HP36 is composed of three short α-helices (Figure 1) that pack to form the hydrophobic core. The protein contains four acidic and six basic residues, all of which are located on the protein surface. They are not distributed uniformly on the molecular surface but rather form several clusters. The helical subdomain of the headpiece starts at residue 42, which is a Leu, and continues to the C-terminus. This subdomain is commonly denoted HP35. Many experimental studies of the subdomain use a construct that contains an additional N-terminal Met resulting from the expression system. This protein is usually referred to as HP36, and we have adopted this notation here. The first residue of HP36 is denoted as Met-41 to conform to the notation used in earlier studies.

Small proteins have become useful model systems for experimental, theoretical and computational investigations of protein folding (*15*−*27*). HP36 in particular has been widely studied by a diverse set of investigators. Its small size and rapid folding have made it an exceptionally popular model protein for theoretical and computational studies of protein folding and dynamics. At least 17 separate research groups have published theoretical or computational studies of its folding (*26*−*44*). There are fewer experimental studies of its folding, but we and independently the Eaton group have demonstrated that HP36 folds on the microsecond time scale (*21*, *22*). One issue that has limited experimental studies is the modest stability of the domain. Wild type HP36 has a T_m of 73.0 °C, a relatively high value, but its free energy of unfolding at 25 °C is only 3.22 kcal mol^{-1} at pH 5.0 (Table 1). This is typical for small globular proteins. Their small size normally means they have a small $\Delta C_p°$ and a small $\Delta H°$ of unfolding which leads to relatively high T_m values, even though $\Delta G°$ at 25 °C can be quite modest. The small value of $\Delta G°$ has limited mutational studies of the folding of HP36, particularly because these investigations typically involve temperature jump studies that are usually carried out above room temperature, where the protein is even less stable. For example, the hydrophobic core of HP36 contains an unusual triad of phenylalanines, which pack together in an approximately triangular arrangement. Studies of this interesting interaction have made use of Phe to Leu substitutions but have been hindered because some of the mutants are so destabilizing that the domain is not completely folded (*18*, *45*). In this work, we design and characterize a hyperstable mutant of HP36. We used X-ray crystallography to determine the structures of the single mutant K70M and the thermal hyperstable double mutant N68AK70M. Thermodynamic and H−D exchange experiments were also used to characterize the hyperstable mutant.

Table 1: Summary of Equilibrium Stability Measurements for Wild Type HP36 and Its Mutants[a]

protein	T_m (°C)	T_m in 3 M urea(°C)	$\Delta G°$ (kcal mol^{-1})	*m*-value (kcal mol^{-1} M^{-1})
WT HP36	73.0 ± 1.5	54.8	3.22 ± 0.12	0.52 ± 0.01
N68A	76.1	N/A	4.16 ± 0.08[d]	
K70M	82.2	N/A	4.41 ± 0.09[d]	
N68AK70M	90.6[b] 90[c]	73.9	4.94 − 5.06[e]	

[a] The quoted uncertainties represent the standard error to the fit. The uncertainty in T_m as estimated by repeated measurements of wild type is 1.5 °C. The estimated uncertainties in $\Delta G°$ and the *m*-values are on the order of 0.12 kcal mol^{-1} and 0.01 kcal mol^{-1} M^{-1}, respectively, on the basis of repeated measurements of wild type. Thermal unfolding studies were conducted at protein concentrations of 20.4 μM (wild type HP36), 26 μM (HP36 K70M), 27 μM (HP36 N68A), and 28.7 μM (HP36 N68AK70M). [b] Value estimated by directly fitting the thermal unfolding curve. [c] Value estimated by taking the derivative of the curve. [d] Estimated using the wild type *m*-value (0.52 ± 0.01 kcal mol^{-1} M^{-1}) and the measured C_M value. [e] Value estimated by using the wild type *m*-value and C_M determined either by directly fitting the curve or by taking the derivative of the curve.

MATERIALS AND METHODS

Cloning, Expression, and Purification. The plasmid (pET3a-NTL9-FXa-HP36) containing the gene for HP36 was prepared as described (*46*). The proteins with coded amino acid substitutions were expressed and purified as described (*46*). ^{15}N-labeled HP36 and N68AK70M were produced by using M9T minimal media with $^{15}NH_4Cl$. All proteins were more than 95% pure. The identities of the proteins were analyzed by electro-spray ionization mass spectrometry (ESI MS) or matrix-assisted laser desorption and ionization time-of-flight mass spectrometry (MALDI-TOF). The expected and observed molecular weights were as follows: N68A, expected 4146.8, observed 4146.8; N68AK70M, expected 4149.9,

Table 2: X-ray Data

	HP36 K70M	HP36 N68AK70M
wavelength (Å)	1.1	1.0
resolution limits (Å)	50−1.41	50−1.79
number of reflections	5,976	3,020
completeness	0.942 (0.578)	0.942 (0.701)
mean redundancy	3.9 (1.5)	5.9 (3.8)
R_{sym}[a]	0.062 (0.176)	0.038 (0.176)
$\langle I/\sigma I\rangle$[b]	21.1 (6.4)	34.5 (6.6)
space group	*P*41	*P*41
unit cell dimensions *a*, *b*, *c* (Å)	32.2, 32.2, 31.7	32.4, 32.4, 31.7

[a] $R_{sym} = \Sigma_{hkl}\Sigma_i|I_i - \langle I\rangle|/\Sigma_{hkl}\Sigma_i I_i$, where I_i is the i^{th} measurement, and $\langle I\rangle$ is the weighted mean of all measurements of I. [b] $\langle I/\sigma I\rangle$ indicates the average of the intensity divided by its standard deviation. The numbers in parentheses refer to the respective highest resolution data shell in each data set.

observed 4150.2. ^{15}N-labeled HP36, expected 4238.9, observed 4238.5; ^{15}N-labeled N68AK70M, expected 4195.9, observed 4195.6.

Circular Dichroism (CD) Spectroscopy. All CD experiments were performed on an Aviv 62A DS and a 202SF Circular Dichroism spectrophotometer. A 10 mM sodium acetate buffer was used for all CD experiments at pH 5.0. The protein concentrations were determined from absorbance measurements using the method of Pace and co-workers (*47*). Far-UV wavelength scans were performed with five repeats and a signal averaging time of 1 s per point, in a 1 mm quartz cuvette, over the range of 195 to 260 nm. Thermal unfolding and urea denaturation experiments were carried out in a 1 cm quartz cuvette by monitoring the signal at 222 nm. The concentration of urea solutions was measured by measuring the refractive index. The concentration of urea was increased from 0 to 10 M in ~0.25 M steps. Wavelength scans and urea denaturation experiments were performed at 25 °C. Thermal unfolding experiments were performed from 2 to 98 °C with a 2 °C interval. For the thermal unfolding experiment in 3 M urea, the protein samples were dissolved in 10 mM sodium acetate, 150 mM sodium chloride, and 3 M urea at pH 5.0. All the denaturation data were analyzed by nonlinear least squares curve fitting using SigmaPlot, as described (*48*−*50*).

Structural Determination. Both the HP36 K70M and HP36 N68AK70M mutants were lyophilized and dissolved in 10 mM sodium acetate (pH 5.0) and 150 mM NaCl. The final protein concentrations of K70M and N68AK70M were around 1 mM and 500 μM, respectively. Crystals were obtained by incubating the sample at 4 °C. Diffraction data were collected on beamline X26C at the National Synchrotron Light Source (NSLS) at Brookhaven National Laboratory. Data were indexed, integrated, and scaled using HKL. The structures of the both mutants were solved by molecular replacement with MOLREP using 1YRF (HP35-N68H) (*22*) as a search model. Parameters are given in Table 2. All structural Figures were generated using PyMOL, version 0.97 (*51*).

Analytical Ultracentrifugation. Analytical ultracentrifugation of HP36 N68AK70M mutant was performed with a Beckman Optima XL-A analytical ultracentrifuge at 25 °C using rotor speeds of 38,000 rpm (24 h) and 48,000 rpm (24 h). Molecular masses were determined at initial concentrations of 30, 60, 100 μM in 10 mM sodium acetate and 150 mM sodium chloride buffer (pH 5.0). Six channel, 12 mm path length, charcoal-filled Epon cell with quartz windows were used. The wavelength used for this experiment was 280 nm. Ten scans were averaged. The partial specific volume (0.750 mL g^{-1}) and solution density (1.005 g l^{-1}) were calculated from the software program SEDNTERP (*52*). The HeteroAnalysis program from the Analytical Ultracentrifugation Facility at the University of Connecticut was used for data analysis.

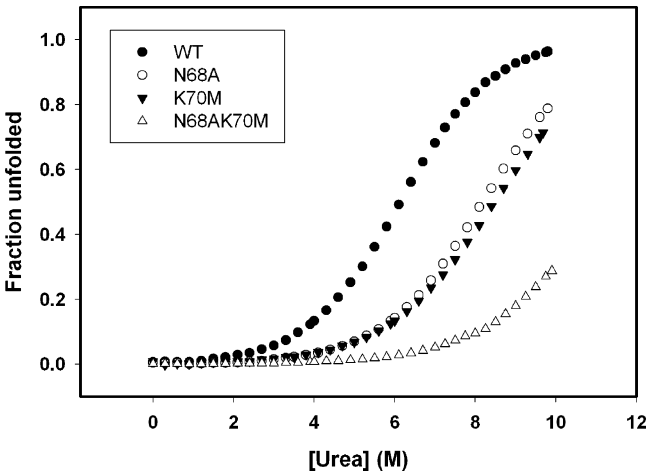

FIGURE 2: Urea denaturation of HP36 (●), N68A (○), K70M (▼), and N68AK70M (△). Measurements were made at 25 °C and pH 5.0 in 10 mM sodium acetate buffer. Data are plotted as fraction unfolded.

NMR Spectroscopy. All NMR experiments were carried out on a Varian Instruments Inova 500 Mhz nuclear magnetic resonance spectrometer at 25 °C. Samples of 1 mM proteins for 1D NMR were prepared in 10% D_2O, 90% 10 mM sodium acetate, and 150 mM sodium chloride at pH 5.0. ^{15}N-labeled HP36 and ^{15}N-labeled N68AK70M for H1/H2 exchanged experiments were dissolved in 10 mM sodium acetate and 150 mM sodium chloride buffer, adjusted to pH 5.0, and then frozen and lyophilized. The dry samples were dissolved in ^2H_2O and then transferred into the NMR tube. The H/H^2 exchange experimental procedures and data analysis were performed as described (*50*).

RESULTS

Characterization of HP36 N68A and HP36 K70M. Early Cys scanning studies of the entire villin headpiece domain showed that a K70C mutation lead to an increase in thermostability (*53*). In that study, thermostability was measured using CD detected thermal melts. The wavelength used, 222 nm, is dominated by the α-helices. The vast majority of the helical structure in the villin headpiece is located in the HP36 region, thus we reasoned that this study was probing the thermal stability of the C-terminal helical subdomain. Inspired by this observation, we sought to test whether the mutation of K70 would indeed stabilize the isolated C-terminal subdomain. We have recently analyzed the contribution of the six positively charged residues to the stability to the helical subdomain of the headpiece (i.e., HP36). We replaced each charged residue by Met, reasoning that this is a more conservative mutation compared to a Lys to Cys or an Arg to Cys mutation because it preserves the hydrophobic portion of the side chain. We observed that mutation of three of the Lysines, K48, K65, and K70 led to an increase in thermal stability of HP36. The HP36 K70M

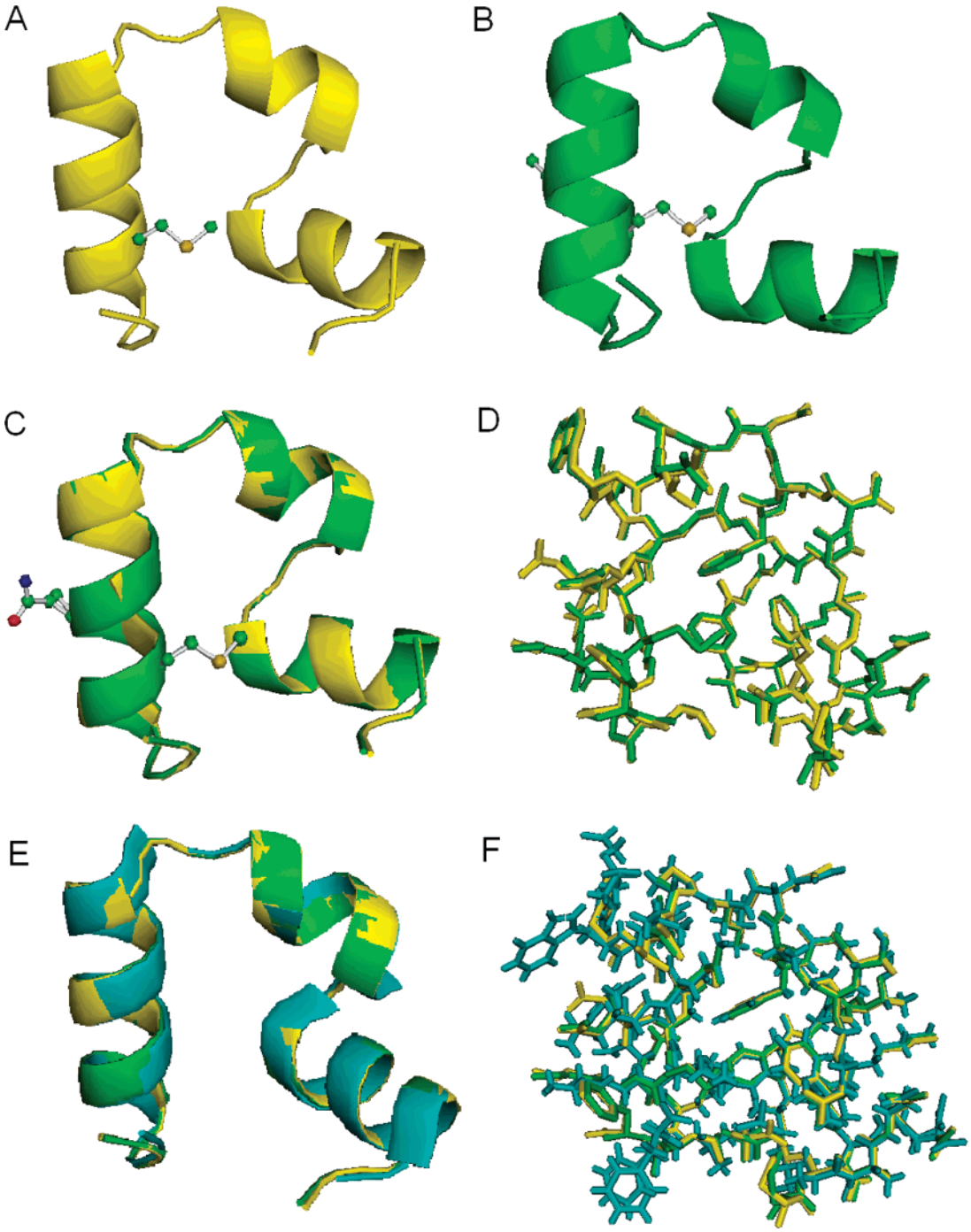

FIGURE 3: (A) Ribbon diagram of the crystal structure of HP36 K70M (yellow) and (B) HP36 N68AK70M (green). The side chains of residue 68 and 70 are shown in ball-and-stick representation. (C) Comparison of the X-ray structures of HP36 K70M (yellow) and HP36 N68AK70M (green) shown in ribbon diagram representation and (D) all bonds representation. (E) Comparison of the X-ray structure of HP35 N68H (1YRF, cyan) and X-ray structures of HP36 K70M (yellow) and HP36 N68AK70M (green) shown in ribbon diagram representation and (F) all bonds representation. The Figures were created using PyMOL, version 0.99 (*51*). The X-ray strycture of HP35 shows two orientations for the Phe76 and Trp64 side chains. Both are shown in the all bonds representation in panel F.

mutant was the most stabilizing among those mutations, and led to an increase in T_m of 9.8 °C. The effect of the other two mutants, K48M and K65M, were smaller, leading to increases in T_m of 5 and 4.2 °C, respectively. The K70M mutant also had the largest effect on $\Delta G°$ of unfolding, increasing it by 1.19 kcal mol^{-1} (Table 1, Figure 2). In the absence of high-resolution structural data on the HP36 K70M mutant, it is not clear what is responsible for the increase in stability. Consequently, we solved the structure of the mutant by X-ray crystallography at 1.41 Å resolution (Table 2, Figure 3). The structure is identical to that of the HP35 N68H (pdb code 1YRF) (*22*) mutant previously studied by the Eaton group. The rmsd for the two structures calculated for residues Leu42 to Leu75 of HP36 K70M with Leu42 to Leu75 of HP35 N68H is 0.32 Å for the backbone. The all atom rmsd, excluding the mutated site, is 0.91 Å. The K70 and M70 side chains in the two proteins adopt the same side chain conformation with a χ_1 value of 124°. The fact that the structures are identical strongly supports our conjecture that the K70M mutant exerts its effect by eliminating unfavorable electrostatic interactions rather than by altering core packing.

In this study, we also examined an additional substitution that did not include a charged residue: Asn-68 to Ala. The choice was again based on the early Cys scanning study of the intact headpiece (*53*). Those studies reported that replacement of Asn-68 by Cys increased the T_m of domain, but no detailed analysis was reported. Asn-68 is a surface residue located in the C-terminal helix of HP36 (Figure 1). Asn has a relatively low helix propensity; thus, we reasoned that its replacement by a residue with a high helical propensity could stabilize the domain (*54*, *55*).

At 25 °C, the far-UV CD spectrum indicates that the N68A mutant has significant α-helix content, and the shape and intensity of the spectrum are very similar to that of HP36 (data not shown). In the 1D NMR spectrum, characteristic resonances appeared at −0.1 ppm (Val-50) and at 5.6 ppm (Phe-47) and were ring-current shifted. These resonances show that the N68A mutant folds to the same native state as that of the wild type. The stability of the mutant was probed using both thermal unfolding experiments and urea denaturation experiments monitored by CD. The HP36 N68A mutant is more thermostable than wild type HP36 with a T_m of 76.1 °C (Table 1). The HP36 N68A mutant does not exhibit a complete urea unfolding curve, and the unfolded baseline is not reached, even at the highest urea concentration (Figure 2). The estimated stability of the mutant from directly fitting the curve was 3.95 kcal mol^{-1}. However, this value is not the most reliable because the complete transition cannot be observed. Alternatively, the stability can be estimated from the midpoint concentration, C_M, which can be accurately determined even from an incomplete unfolding curve, whereas the *m*-value often cannot. However, the stability can still be estimated using the measured C_M value and the wild type *m*-value (0.52 ± 0.01 kcal mol^{-1} M^{-1}). Using this method gives 4.16 ± 0.08 kcal mol^{-1} for the stability of the HP36 N68A mutant (Table 1).

N68AK70M a Thermal Hyperstable Double Mutant. The analysis of the single mutants demonstrates that N68A and K70M significantly stabilize the domain. Examination of the structure of the domain suggests that these two mutations could be simultaneously accommodated by the protein. Consequently, we prepared the N68AK70M double mutant. At 25 °C, the far-UV CD spectrum and 1D NMR spectrum indicate that the double mutant adopts the same fold as that of the wild type (data not shown). We solved the crystal structure of the double mutant to 1.79 Å resolution (Table 2 and Figure 3). The structure is identical to those of the HP36 K70M and HP35 N68H mutants, and the rmsd for residues 42−75 is 0.13 Å relative to HP36 K70M. The all atom rmsd, excluding the mutated residues, is 0.25 Å.

The results of thermal unfolding with the double mutant are impressive. The protein does not completely thermally unfold, and the melting profile does not show a post-transition (Figure 4A). The T_m for the wild type is 73.0 °C, whereas the T_m estimated for the double mutant is at least 90.6 °C (Table 1). We also compared the thermal stability of the wild type and N68AK70M in 3 M urea (Figure 4B). Under these conditions, the T_m of WT HP36 is 54.8 °C, while it is 73.9 °C for the N68AK70M mutant, an increase of 19.1 °C (Table 1). Urea denaturation studies were attempted, but the C_M of the mutant is close to 10 M urea, making it difficult to accurately measure stability (Figure 2). We estimated the stability by using the wild type *m*-value

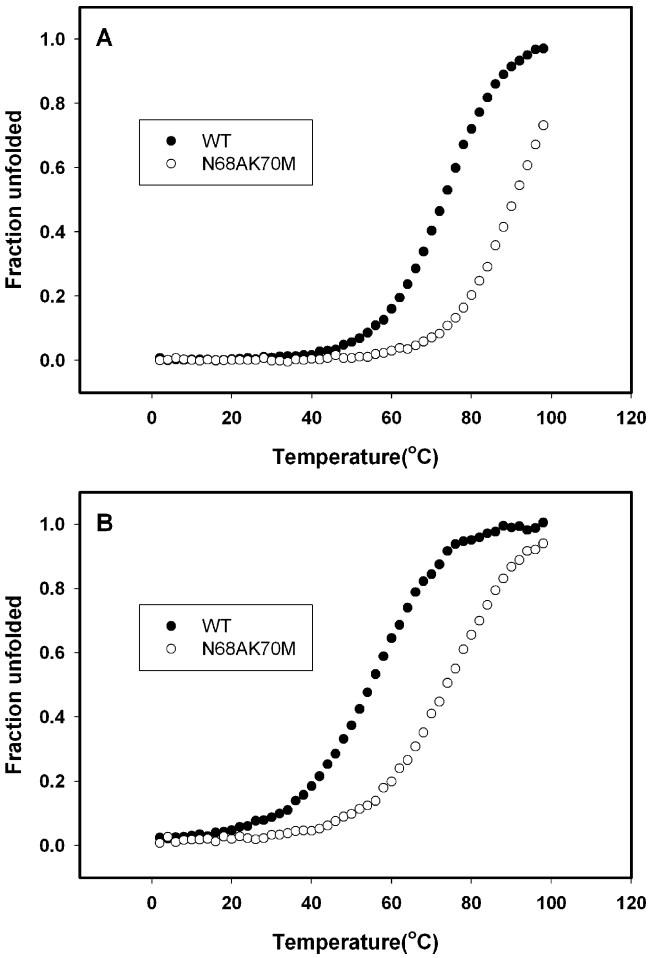

FIGURE 4: Thermal unfolding of HP36 (●) and N68AK70M (○), (A) with 0 M urea and (B) with 3 M urea. All experiments were conducted in 10 mM sodium acetate buffer at pH 5.0. Data are plotted as fraction unfolded.

and C_M for the double mutant calculated by directly fitting the curve or by analysis of the derivative. $\Delta G°$ of unfolding for the double mutant is 1.7−1.8 kcal mol^{-1} larger than the value for the wild type. Guanidine hydrochloride induced unfolding is not a viable option for measuring the stability because electrostatic interactions make a significant contribution to the domain stability, and preliminary studies have shown that guanidine hydrochloride and urea denaturation give different estimates of $\Delta G°$.

We next undertook amide H/^{2}H exchange studies to probe the slow time scale dynamics of the double mutant and to compare it to the wild type. Amide H/^{2}H exchange experiments were performed for WT HP36 and HP36 N68AK70M at pD 5.0 and 25 °C under native conditions. For WT HP36, 21 amide ^{1}H-^{15}N cross-peaks were detected in the first HSQC spectrum, which was taken 27 min after the dry protein was mixed with ^2H_2O. Only two amide ^{1}H-^{15}N cross-peaks (Val 50 and Leu 61) were detected after 6 h. While for HP36 N68AK70M, 23 amide ^{1}H-^{15}N cross-peaks were detected in the first HSQC spectrum, and 7 amide ^{1}H-^{15}N cross-peaks (Phe 58, Ala 59, Leu 61, Lys 65, Asn 66, M70, and K71) were detected even after 17 h. A histogram of the apparent free energies of the opening reaction calculated from the observed exchange rates for wild type and the N68AK70M mutant is shown in Figure 5. The pattern of exchange is very similar for the two proteins,

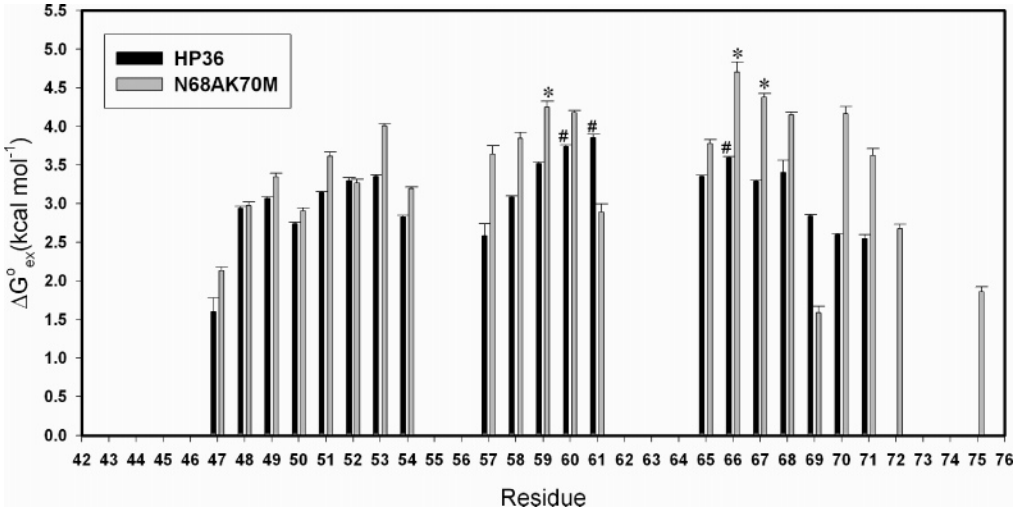

FIGURE 5: Histogram of the apparent free energy change for the opening reaction, ΔG°_{ex} at 25 °C vs residue number for WT HP36 (black bars) and N68AK70M (gray bars). Samples were prepared in 10 mM sodium acetate buffer at pD 5.0 and 25 °C. Error bars are shown. The three residues with the largest protection factor are labeled (#) for HP36 and (*) for HP36 N68AK70M.

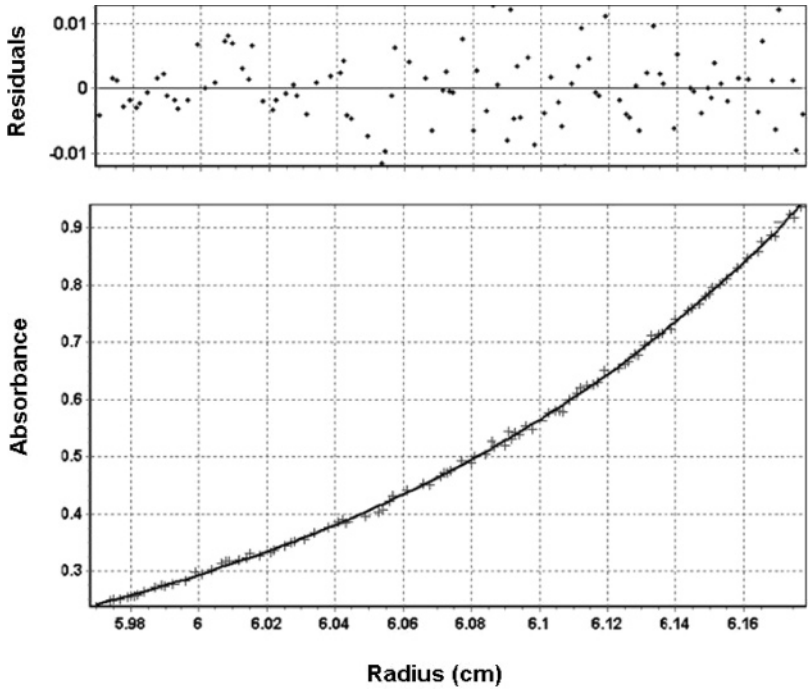

FIGURE 6: Sedimentation equilibrium studies of N68AK70M. The sample contained 100 μM N68AK70M in 10 mM sodium acetate buffer solution at pH 5.0 and 25 °C. The rotor speed was 48,000 rpm. The observed molecular mass was 4278, determined from a single-species fit; the expected mass is 4149.9.

arguing that the mutations have not significantly altered the slow dynamics of the domain. In principle, the free energy of unfolding of a protein can be accurately determined by amide exchange, provided that exchange occurs in the EX2 limit. The analysis requires a known subset of residues that exchange by global unfolding and do not experience local unfolding or, conversely, so-called super protection due to structure in the unfolded state. A priori, it is not always clear which residues exchange strictly by global unfolding, and furthermore, stabilization of a domain can increase the fraction of residues that exchange by subglobal unfolding events. For these reasons, we believe that the amide H/^{2}H exchange studies are best viewed as probes of slow dynamics. Nevertheless, it is interesting to compare the estimated ΔG° of unfolding determined by H/^{2}H exchange for the two proteins. Using the three most protected residues in wild type HP36, ΔG° is 3.73 kcal mol^{-1}, whereas the value calculated for the double mutant is 4.44 kcal mol^{-1}, an increase of 0.7 kcal mol^{-1}. For the reasons outlined above, we believe the ΔG° value determined by the urea denaturation experiment more accurately reflects the increase in stability caused by the double mutations.

The Hyperstable Mutant Is Monomeric. HP36 N68AK70M is much more stable than WT HP36; however, it is necessary to check whether this hyperstable mutant is monomeric or not because self-association can lead to an apparent increase in stability. Analytical ultracentrifugation experiments show that HP36 N68AK70M is monomeric. Figure 6 shows representative sedimentation equilibrium data for HP36 N68AK70M. The data were fit well by an ideal single-species

model with a molecular weight within 5% of the monomer molecular weight. The average experimental molecular mass determined from multiple experiments with HP36 N68AK70M is 4285, and the expected mass is 4149.9.

DISCUSSION

Using two simple surface mutations, we have developed a monomeric hyperstable variant of HP36, whose structure is identical to that of the wild type. The mutations were designed to eliminate unfavorable electrostatic interactions (HP36 K70M) and to increase helical propensity (HP36 N68A). The increase in stability for the N68A mutant is very close to that expected on the basis of helical propensity. Our high-resolution structure of HP36 K70M is consistent with stabilization by modification of surface electrostatics rather than core repacking. The high-resolution structure of the double mutant shows that there is no change in the hydrophobic core and, again, is consistent with the design strategy that targeted surface interactions. It is interesting to compare these results to experiments designed to improve protein stability by altering interactions in the core. In these cases, multiple mutations are often required to achieve increases in stability comparable to what is observed here, suggesting that targeting surface residues may be a particularly efficient strategy for stabilizing proteins (*5*). It is also noteworthy that the significant increase in T_m is associated with a relatively more modest increase in $\Delta G°$. This is, in fact, expected for small proteins. Rees and Robertson (*56*) analyzed a data base of larger proteins and derived a simple relationship between $\Delta\Delta G°$ and ΔT_m, showing that $d\Delta G°/dT_m$ scales with the number of residues, N. Thus the same increase in $\Delta G°$ will lead to a large increase in T_m for small proteins and a small increase in T_m for large proteins. $d\Delta G°/dT_m$ is predicted to be N*0.0023 kcal mol^{-1} K^{-1} at the T_m of the wild type (346 K). Thus a 19° increase in T_m is predicted to be linked to a 1.6 kcal mol^{-1} increase in $\Delta G°$, which is in good agreement with the value estimated from an analysis of the urea denaturation experiment. The observation that such simple mutations can drastically increase protein stability provides direct evidence that the stability of miniature protein domains has not been evolutionarily optimized; instead, like large normal domains, their stability can be increased by optimizing interactions and removing unfavorable interactions. In addition, the fact that the relationship between the increase in T_m and $\Delta\Delta G°$ is well predicted by the properties of large proteins provides further evidence that the thermodynamics of miniature proteins is not unusual and that they do not require specialized interactions to fold.

The modest stability of wild type HP36 has hampered mutational studies (*18*). The use of HP36 N68AK70M as a pseudo-wild type structure will allow more drastic mutations in the hydrophobic core than are tolerated in the normal wild type protein and will facilitate experimental studies of this interesting protein.

Finally, it is interesting to contemplate the consequences of these mutations in the larger HP67 construct. The increase in stability caused by the K70M is not expected to propagate to the N-ternimal subdomain because K70 together with E39 forms a buried salt bridge that links the two subdomains (*13*). Removal of the ϵ-amino group of K70 will leave the buried carboxylate E39 uncompenstated by a salt bridge partner and is thus expected to destabilize the N-terminal subdomain. Sequence analysis is often used to choose potential sites for mutation in protein engineering studies and can be used to design stabilizing substitutions in favorable cases (*57*). However, analysis of aligned sequences of HP36 would fail to predict that K70 is a viable site for mutation because it is rigorously conserved in all HP36 sequences (*12*). The conservation arises because of the need to form the E39 K70 salt bridge and not because K70 is needed to directly stabilize the HP36 fold. It is interesting to note that virtually all characterized miniature proteins are domains or subdomains of large sequences; thus, if the goal of a protein engineering exercise is to stabilize the protein, then sequence alignments may not provide a useful guide because residues may be conserved for functional reasons in the intact protein, even if they make a unfavorable contribution to the stability of the smaller domain. K70 in HP36 provides a striking example of this.

ACKNOWLEDGMENT

We thank Professor Mcknight for his continued interest in this work and for numerous helpful discussions. We also thank Dr. Yuefeng Tang for her help with the NMR experiments and for helpful discussions, Ms. Benben Song for helpful discussions and Ms. Lauren Wickstrom for help with the rmsd calculations.

SUPPORTING INFORMATION AVAILABLE

Amide exchange rates measured for wild type and the N68AK70M double mutant (Table S1), plots of the CD signal versus urea concentration of HP36, HP36 K70M, HP36 N68A, and HP36 N68AK70M (Figure S1), plots of the CD signal versus temperature of HP36 and HP36 N68AK70M without urea (Figure S2) and with urea (Figure S3). This material is available free of charge via the Internet at http://pubs.acs.org.

REFERENCES

1. Chen, J., Lu, Z., Sakon, J., and Stites, W. E. (2000) Increasing the thermostability of staphylococcal nuclease: implications for the origin of protein thermostability, *J. Mol. Biol. 303*, 125−130.
2. Kumar, S., Tsai, C. J., and Nussinov, R. (2000) Factors enhancing protein thermostability, *Protein Eng. 13*, 179−191.
3. Lien, S., and Lowman, H. B. (2003) Therapeutic peptides, *Trends Biotechnol. 21*, 556−562.
4. Kubelka, J., Hofrichter, J., and Eaton, W. A. (2004) The protein folding 'speed limit', *Curr. Opin. Struct. Biol. 14*, 76−88.
5. Malakauskas, S. M., and Mayo, S. L. (1998) Design, structure and stability of a hyperthermophilic protein variant, *Nat. Struct. Biol. 5*, 470−475.
6. Ghosh, T., Garde, S., and Garcia, A. E. (2003) Role of backbone hydration and salt-bridge formation in stability of alpha-helix in solution, *Biophys. J. 85*, 3187−3193.
7. Durrschmidt, P., Mansfeld, J., and Ulbrich-Hofmann, R. (2005) An engineered disulfide bridge mimics the effect of calcium to protect neutral protease against local unfolding, *FEBS. J. 272*, 1523−1534.
8. 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.
9. Strickler, S. S., Gribenko, A. V., Gribenko, A. V., Keiffer, T. R., Tomlinson, J., Reihle, T., Loladze, V. V., and Makhatadze, G. I. (2006) Protein stability and surface electrostatics: a charged relationship, *Biochemistry 45*, 2761−2766.

10. Bretscher, A., and Weber, K. (1979) Villin: the major microfilament-associated protein of the intestinal microvillus, *Proc. Natl. Acad. Sci. U.S.A. 76*, 2321−2325.
11. Arpin, M., Pringault, E., Finidori, J., Garcia, A., Jeltsch, J. M., Vandekerckhove, J., and Louvard, D. (1988) Sequence of human villin: a large duplicated domain homologous with other actin-severing proteins and a unique small carboxy-terminal domain related to villin specificity, *J. Cell Biol. 107*, 1759−1766.
12. Vardar, D., Buckley, D. A., Frank, B. S., and McKnight, C. J. (1999) NMR structure of an F-actin-binding "headpiece" motif from villin, *J. Mol. Biol. 294*, 1299−1310.
13. Meng, J., Vardar, D., Wang, Y., Guo, H. C., Head, J. F., and McKnight, C. J. (2005) High-resolution crystal structures of villin headpiece and mutants with reduced F-actin binding activity, *Biochemistry 44*, 11963−11973.
14. Vardar, D., Chishti, A. H., Frank, B. S., Luna, E. J., Noegel, A. A., Oh, S. W., Schleicher, M., and McKnight, C. J. (2002) Villin-type headpiece domains show a wide range of F-actin-binding affinities, *Cell Motil. Cytoskeleton 52*, 9−21.
15. McKnight, C. J., Doering, D. S., Matsudaira, P. T., and Kim, P. S. (1996) A thermostable 35-residue subdomain within villin headpiece, *J. Mol. Biol. 260*, 126−134.
16. McKnight, C. J., Matsudaira, P. T., and Kim, P. S. (1997) NMR structure of the 35-residue villin headpiece subdomain, *Nat. Struct. Biol. 4*, 180−184.
17. Vugmeyster, L., Trott, O., McKnight, C. J., Raleigh, D. P., and Palmer, A. G., III. (2002) Temperature-dependent dynamics of the villin headpiece helical subdomain, an unusually small thermostable protein, *J. Mol. Biol. 320*, 841−854.
18. Frank, B. S., Vardar, D., Buckley, D. A., and McKnight, C. J. (2002) The role of aromatic residues in the hydrophobic core of the villin headpiece subdomain, *Protein Sci. 11*, 680−687.
19. Tang, Y., Rigotti, D. J., Fairman, R., and Raleigh, D. P. (2004) Peptide models provide evidence for significant structure in the denatured state of a rapidly folding protein: the villin headpiece subdomain, *Biochemistry 43*, 3264−3272.
20. Tang, Y., Goger, M. J., and Raleigh, D. P. (2006) NMR characterization of a peptide model provides evidence for significant structure in the unfolded state of the villin headpiece helical subdomain, *Biochemistry 45*, 6940−6946.
21. Wang, M., Tang, Y., Sato, S., Vugmeyster, L., McKnight, C. J., and Raleigh, D. P. (2003) Dynamic NMR line-shape analysis demonstrates that the villin headpiece subdomain folds on the microsecond time scale, *J. Am. Chem. Soc. 125*, 6032−6033.
22. Chiu, T. K., Kubelka, J., Herbst-Irmer, R., Eaton, W. A., Hofrichter, J., and Davies, D. R. (2005) High-resolution x-ray crystal structures of the villin headpiece subdomain, an ultrafast folding protein, *Proc. Natl. Acad. Sci. U.S.A. 102*, 7517−7522.
23. Duan, Y., Wang, L., and Kollman, P. A. (1998) The early stage of folding of villin headpiece subdomain observed in a 200-nanosecond fully solvated molecular dynamics simulation, *Proc. Natl. Acad. Sci. U.S.A. 95*, 9897−9902.
24. Shen, M. Y., and Freed, K. F. (2002) All-atom fast protein folding simulations: the villin headpiece, *Proteins 49*, 439−445.
25. Ripoll, D. R., Vila, J. A., and Scheraga, H. A. (2004) Folding of the villin headpiece subdomain from random structures. Analysis of the charge distribution as a function of pH, *J. Mol. Biol. 339*, 915−925.
26. Wickstrom, L., Okur, A., Song, K., Hornak, V., Raleigh, D. P., and Simmerling, C. L. (2006) The unfolded state of the villin headpiece helical subdomain: computational studies of the role of locally stabilized structure, *J. Mol. Biol. 360*, 1094−1107.
27. Bandyopadhyay, S., Chakraborty, S., and Bagchi, B. (2005) Secondary structure sensitivity of hydrogen bond lifetime dynamics in the protein hydration layer, *J. Am. Chem. Soc. 127*, 16660−16667.
28. Jayachandran, G., Vishal, V., Garcia, A. E., and Pande, V. S. (2006) Local structure formation in simulations of two small proteins, *J. Struct. Biol.*, in press.
29. Faccioli, P., Sega, M., Pederiva, F., and Orland, H. (2006) Dominant pathways in protein folding, *Phys. Rev. Lett. 97*, 108101.
30. Jang, S., Sreerama, N., Liao, V. H., Lu, S. H., Li, F. Y., Shin, S., Woody, R. W., and Lin, S. H. (2006) Theoretical investigation of the photoinitiated folding of HP-36, *Protein Sci. 15*, 2290−2299.
31. Zagrovic, B., and Pande, V. S. (2006) Simulated unfolded-state ensemble and the experimental NMR structures of villin headpiece yield similar wide-angle solution X-ray scattering profiles, *J. Am. Chem. Soc. 128*, 11742−11743.
32. Trebst, S., Troyer, M., and Hansmann, U. H. (2006) Optimized parallel tempering simulations of proteins, *J. Chem. Phys. 124*, 174903.
33. Jayachandran, G., Vishal, V., and Pande, V. S. (2006) Using massively parallel simulation and Markovian models to study protein folding: examining the dynamics of the villin headpiece, *J. Chem. Phys. 124*, 164902.
34. Guner, U., Arkun, Y., and Erman, B. (2006) Optimum folding pathways of proteins: their determination and properties, *J. Chem. Phys. 124*, 134911.
35. Zagrovic, B., and van Gunsteren, W. F. (2006) Comparing atomistic simulation data with the NMR experiment: how much can NOEs actually tell us? *Proteins 63*, 210−218.
36. Carr, J. M., and Wales, D. J. (2005) Global optimization and folding pathways of selected alpha-helical proteins, *J. Chem. Phys. 123*, 234901.
37. Fogolari, F., Tosatto, S. C., and Colombo, G. (2005) A decoy set for the thermostable subdomain from chicken villin headpiece, comparison of different free energy estimators, *BMC Bioinf. 6*, 301.
38. Massi, F., and Palmer, A. G., III. (2003) Temperature dependence of NMR order parameters and protein dynamics, *J. Am. Chem. Soc. 125*, 11158−11159.
39. De Mori, G. M., Colombo, G., and Micheletti, C. (2005) Study of the villin headpiece folding dynamics by combining coarse-grained Monte Carlo evolution and all-atom molecular dynamics, *Proteins 58*, 459−471.
40. Herges, T., and Wenzel, W. (2004) An all-atom force field for tertiary structure prediction of helical proteins, *Biophys. J. 87*, 3100−3109.
41. Zagrovic, B., and Pande, V. S. (2004) How does averaging affect protein structure comparison on the ensemble level? *Biophys. J. 87*, 2240−2246.
42. Hansmann, U. H. (2004) Simulations of a small protein in a specifically designed generalized ensemble, *Phys. Rev. E: Stat., Nonlinear, Soft Matter Phys. 70*, 012902.
43. Mukherjee, A., and Bagchi, B. (2004) Contact pair dynamics during folding of two small proteins: chicken villin head piece and the Alzheimer protein beta-amyloid, *J. Chem. Phys. 120*, 1602−1612.
44. Srinivasan, R., Fleming, P. J., and Rose, G. D. (2004) Ab initio protein folding using LINUS, *Methods Enzymol. 383*, 48−66.
45. Brewer, S. H., Vu, D. M., Tang, Y., Li, Y., Franzen, S., Raleigh, D. P., and Dyer, R. B. (2005) Effect of modulating unfolded state structure on the folding kinetics of the villin headpiece subdomain, *Proc. Natl. Acad. Sci. U.S.A. 102*, 16662−16667.
46. Bi, Y., Tang, Y., Raleigh, D. P., and Cho, J. H. (2006) Efficient high level expression of peptides and proteins as fusion proteins with the N-terminal domain of L9: application to the villin headpiece helical subdomain, *Protein Expression Purif. 47*, 234−240.
47. Pace, C. N., Vajdos, F., Fee, L., Grimsley, G., and Gray, T. (1995) How to measure and predict the molar absorption coefficient of a protein, *Protein Sci. 4*, 2411−2423.
48. Spector, S., Kuhlman, B., Fairman, R., Wong, E., Boice, J. A., and Raleigh, D. P. (1998) Cooperative folding of a protein mini domain: the peripheral subunit-binding domain of the pyruvate dehydrogenase multienzyme complex, *J. Mol. Biol. 276*, 479−489.
49. Spector, S., Young, P., and Raleigh, D. P. (1999) Nativelike structure and stability in a truncation mutant of a protein minidomain: the peripheral subunit-binding domain, *Biochemistry 38*, 4128−4136.
50. Tang, Y., Grey, M. J., McKnight, J., Palmer, A. G., III, and Raleigh, D. P. (2006) Multistate folding of the villin headpiece domain, *J. Mol. Biol. 355*, 1066−1077.
51. DeLano, W. L. (2002) *PyMOL*, Version 0.97, DeLano Scientific, San Carlos, CA; http://www.pymol.org.

52. Laue, T. M., Shah, B. D., Ridgeway, T. M., and Pelletier, S. M. (1992) *Analytical Ultracentrifugation in Biochemistry and Polymer Science* (Harding, S. E., Rowe, A. J., and Horton, J. C., Eds.), Royal Society of Chemistry, London, England.
53. Doering, D. S., and Matsudaira, P. (1996) Cysteine scanning mutagenesis at 40 of 76 positions in villin headpiece maps the F-actin binding site and structural features of the domain, *Biochemistry 35*, 12677−12685.
54. O'Neil, K. T., and DeGrado, W. F. (1990) A thermodynamic scale for the helix-forming tendencies of the commonly occurring amino acids, *Science 250*, 646−651.
55. Fersht, A. (1998) *Structure and Mechanism in Protein Science: A Guide to Enzyme Catalysis and Protein Folding*, W. H. Freeman, New York.
56. Rees, D. C., and Robertson, A. D. (2001) Some thermodynamic implications for the thermostability of proteins, *Protein Sci. 10*, 1187−1194.
57. Rath, A., and Davidson, A. R. (2000) The design of a hyperstable mutant of the Abp1p SH3 domain by sequence alignment analysis, *Protein Sci. 9*, 2457−2469.

BI6026314