J|A|C|S
COMMUNICATIONS

Published on Web 02/18/2006

Design of a Hyperstable Protein by Rational Consideration of Unfolded State Interactions

Burcu Anil,[†] Rebecca Craig-Schapiro,[‡] and Daniel P. Raleigh*[,†,§]

Department of Chemistry, State University of New York at Stony Brook, Stony Brook, New York 11794-3400, and Graduate Program in Biophysics and Graduate Program in Biochemistry and Structural Biology, State University of New York at Stony Brook, Stony Brook, New York 11794

Received November 20, 2005; E-mail: draleigh@notes.cc.sunysb.edu

A common objective in protein design is the enhancement of protein stability. Increased protein stability is an important area of interest, particularly for applications in biotechnology. The performance of protein pharmaceuticals and proteins used as biocatalysts can often be improved by stabilization, resulting in more effective exploitation of these products.[1] A wide variety of methodologies have been applied, including both rational approaches and selection-based strategies.[2] To date, all rational design approaches, both experimental and computational, have focused on native state interactions.[3] However, protein stability arises due to the balance between unfolded state and folded state energetics. Denatured state interactions can play an important role in the stability of proteins, and mutation of residues involved in such interactions can alter the stability.[4] Here we demonstrate that a hyperstable protein can be generated by rationally targeting unfolded state interactions. Using the N-terminal domain of the ribosomal protein L9 (NTL9) as a model system, we show that two simple substitutions that target the unfolded state generate a hyperthermostable protein.

NTL9 is an α-β protein composed of a three-stranded β-sheet sandwiched between two α-helices (Figure 1) and represents a simple example of the common ABCαD motif found in many proteins.[5] The protein is known to fold via two-state kinetics under a wide range of conditions.[6] We chose NTL9 as our model for this study because it provides a useful system to test the effects of targeting the unfolded state. Previous work has demonstrated that there are significant electrostatic interactions in the unfolded state ensemble of this protein.[7] This appears to be a general phenomenon in globular proteins.[4b,8] Our strategy is two-fold. First, we target these interactions by mutagenesis, taking advantage of prior studies which have shown that lysine-12 plays a key role in the unfolded state.[9] Mutation of this residue to methionine eliminates the unfolded state interactions and stabilizes the protein by 1.9 kcal mol^{-1}.[9] The second part of our approach is more general and can be applied to a range of proteins; it involves targeting unfolded state entropy by a Gly to D-Ala mutation. Glycines experience the least amount of local restrictions on their conformation, thus imposing an entropic penalty on the protein backbone for folding. The largest effect should be observed for glycines which are found in well-ordered conformations in the native state. Such sites often require that the glycine adopt unusual ϕ, ψ angles with $\phi > 0$. Substitutions of glycines have been used to stabilize proteins by reducing unfolded state entropy.[10] However, substitution of a glycine by an L-amino acid can introduce significant strain into the structure as glycine often adopts conformations with positive ϕ angles.[11]

We have shown that replacement of such a glycine by a D-amino acid can be used to stabilize proteins while avoiding strain in the native state. We substituted G34 with D-Ala in NTL9 to reduce the unfolded state entropy. Modeling indicates that the new methyl group does not introduce any strain to the backbone and is solvent-exposed. The G34 to D-Ala mutation was previously shown to increase the stability of NTL9 by 1.87 kcal mol^{-1}.[13]

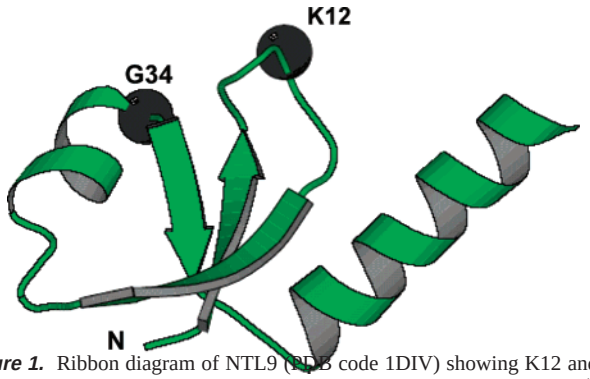

Figure 1. Ribbon diagram of NTL9 (PDB code 1DIV) showing K12 and G34, the N-terminus is labeled. The diagram was created using Molscript.[12]

K12 is located on the surface loop of NTL9 that connects the first and second β-strand, and G34 is located on the loop that connects the central helix and the third β-strand (Figure 1). The double mutant, K12M_G34D-A NTL9, was prepared by FMOC solid-phase peptide synthesis and purified using reverse-phased high performance liquid chromatography (HPLC). The protein was verified to be monomeric at a concentration of 400 μM using analytical ultracentrifuge (Supporting Information). The apparent molecular weight obtained from the fit to the centrifuge data is 6165 Da, within 1.1% of the known actual mass, 6239 Da. The far-UV CD spectrum of K12M_G34D-A shows no changes in the structure of the protein when compared to wild-type NTL9 spectrum (data not shown). The structure of the variant was further analyzed by 2D NMR experiments. CαH chemical shifts are sensitive to secondary structure, and the CαH chemical shifts of wild-type NTL9 show significant deviations from random coil values (Figure 2A). The CαH chemical shifts of wild-type NTL9 and K12M_G34D-A are on average within 0.03 ppm of each other, and the maximum variation between CαH chemical shifts is 0.1 ppm, showing that there is no change in the structure (Figure 2B).

Moreover, the distinctive set of ring-current-shifted methyl resonances observed for wild-type NTL9 and the set of CαH resonances downfield of water found in wild-type NTL9 are also found in the spectrum of K12M_G34D-A, confirming further that the native fold is conserved.

The mutations enhance the stability of NTL9 significantly. K12M_G34D-A does not show a complete melt under native conditions, and no post-transition baseline can be detected. To observe the complete thermal unfolding of the variant, experiments were conducted in the presence of 6 M urea monitored by CD at

[†] Department of Chemistry.
[‡] Present address: School of Medicine, Washington University of St. Louis.
[§] Graduate Program in Biophysics and Graduate Program in Biochemistry and Structural Biology.

10.1021/ja057874b CCC: $33.50

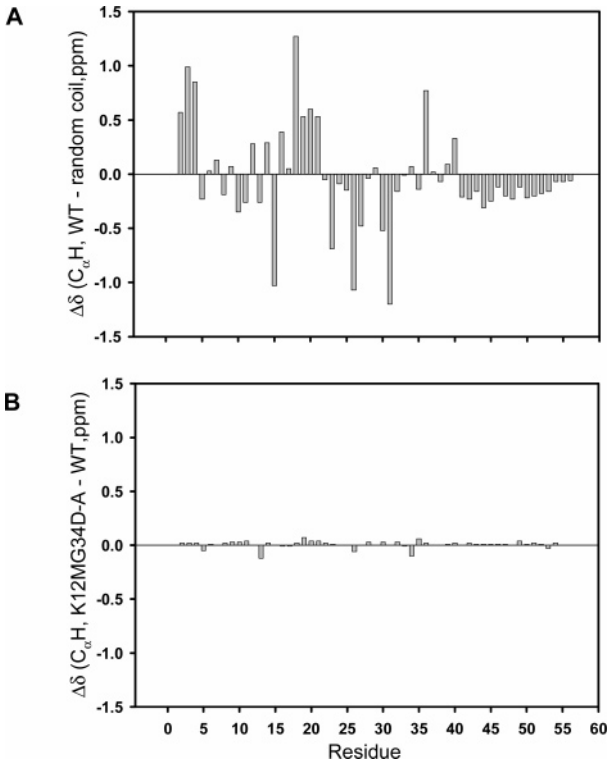

Figure 2. Plots of the deviation between measured chemical Cα proton shifts for (A) wild-type NTL9 and random coil, (B) K12M_G34D-A and wild-type NTL9. The measurements were made at pH 5.4, 25 °C. The random coil chemical shift values are from Wuthrich.[14]

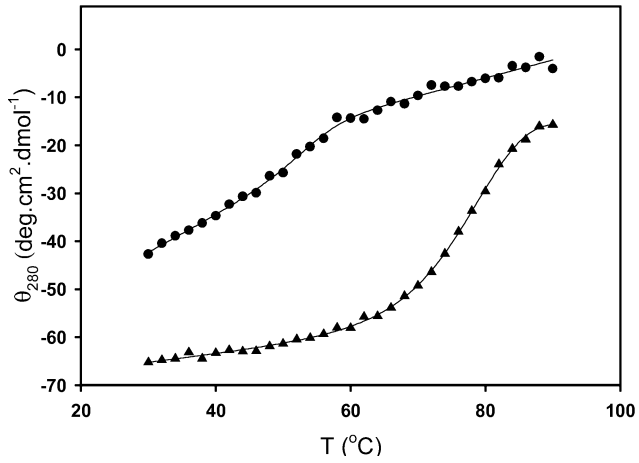

Figure 3. Temperature denaturation curves for NTL9 variants (▲) K12M_G34D-A and (●) wild-type monitored by the near-UV CD signal at 280 nm in 20 mM sodium acetate, 100 mM NaCl, 6 M urea, pH 5.4. The solid lines represent the best fit to the data.

280 nm with 300 μM protein samples. Wild-type NTL9 is already significantly unfolded under these conditions at low temperature and does not show a cooperative melt in 6 M urea (Figure 3). In contrast, K12M_G34D-A is fully folded at temperatures up to at least 60 °C under these conditions. The observed T_m for the double mutant in 6 M urea is 83.1 °C (Figure 3).

We conclude that deletion of interactions that stabilize the unfolded state of proteins and substitutions of D-Ala for glycine residues with positive ϕ angles can be a very useful strategy for rationally stabilizing proteins. Such glycines are common in proteins, and methodology advances have made the incorporation of D-amino acids into proteins more straightforward. Hence, this strategy should offer a general method for stabilizing proteins, provided that the substitutions do not introduce any steric clashes in the native state or introduce additional favorable interactions in the unfolded state.

Acknowledgment. This work was supported by NIH Grant GM70941 to D.P.R.; R.C.S. was supported by NSF-REU fellowship CHE-0139256. We thank Jaehyun Cho for helpful discussions and characterization of the K12M mutant.

Supporting Information Available: Plot of absorbance versus radius obtained from analytical ultracentrifugation experiment. This material is available free of charge via the Internet at http://pubs.acs.org.

References

(1) (a) Chen, J.; Lu, Z.; Sakon, J.; Stites, W. E. *J. Mol. Biol.* **2000**, *303*, 125−130 and references therein. (b) Kumar, S.; Tsai, C. J.; Nussinov, R. *Protein Eng.* **2000**, *13*, 179−191. (c) Lien, S.; Lowman, H. B. *Trends Biotechnol.* **2003**, *21*, 556−562.
(2) (a) Schildbach, J. F.; Milla, M. E.; Jeffrey, P. D.; Raumann, B. E.; Sauer, R. T. *Biochemistry* **1995**, *34*, 1405−1412. (b) Malakauskas, S. M.; Mayo, S. L. *Nat. Struct. Biol.* **1998**, *5*, 470−475. (c) Lehmann, M.; Wyss, M. *Curr. Opin. Biotechnol.* **2001**, *12*, 371−375. (d) van den Burg, B.; Eijsink, V. G. *Curr. Opin. Biotechnol.* **2002**, *13*, 333−337.
(3) (a) Lee, B.; Vasmatzis, G. *Curr. Opin. Biotechnol.* **1997**, *4*, 423−428. (b) Pokkuluri, P. R.; Raffen, R.; Dieckman, L.; Boogaard, C.; Stevens, F. J.; Schiffer, M. *Biophys. J.* **2002**, *82*, 391−398.
(4) (a) Shortle, D. *FASEB J.* **1996**, *10*, 27−34. (b) Pace, C. N.; Alston, R. W.; Shaw, K. L. *Protein Sci.* **2000**, *9*, 1395−1398. (c) Cho, J. H.; Raleigh, D. P. *J. Mol. Biol.* **2005**, *353*, 174−185.
(5) Efimov, A. V. *FEBS Lett.* **1994**, *355*, 213−219.
(6) (a) Kuhlman, B.; Boice, J. A.; Fairman, R.; Raleigh, D. P. *Biochemistry* **1998**, *37*, 1025−1032. (b) Kuhlman, B.; Luisi, D. L.; Evans, P. A.; Raleigh, D. P. *J. Mol. Biol.* **1998**, *284*, 1661−1670. (c) Luisi, D. L.; Raleigh, D. P. *J. Mol. Biol.* **2000**, *299*, 1091−1100.
(7) (a) Kuhlman, B.; Luisi, D. L.; Young, P.; Raleigh, D. P. *Biochemistry* **1999**, *38*, 4896−4903. (b) Luisi, D. L.; Wu, W. J.; Raleigh, D. P. *J. Mol. Biol.* **1999**, *287*, 395−407.
(8) (a) Whitten, S. T.; Garcia-Moreno, E. B. *Biochemistry* **2000**, *39*, 14292−14304. (b) Paci, E.; Vendruscolo, M.; Karplus, M. *Proteins: Struct., Funct., Genet.* **2002**, *47*, 379−392.
(9) Cho, J. H.; Sato, S.; Raleigh, D. P. *J. Mol. Biol.* **2004**, *338*, 827−837.
(10) (a) Matthews, B. W.; Nicholson, H.; Becktel, W. J. *Proc. Natl. Acad. Sci. U.S.A.* **1987**, *84*, 6663−6667. (b) Ganter, C.; Plückthun, A. *Biochemistry* **1990**, *29*, 9395−9402. (c) Stites, W. E.; Pranata, J. *Proteins* **1995**, *22*, 132−140.
(11) (a) Stites, W. E.; Meeker, A. K.; Shortle, D. *J. Mol. Biol.* **1994**, *235*, 27−32. (b) Takano, K.; Yamagata, Y.; Yutani, K. *Proteins* **2001**, *45*, 274−280. (c) Kim, J.; Brych, S. R.; Lee, J.; Logan, T. M.; Blaber, M. *J. Mol. Biol.* **2003**, *328*, 951−961.
(12) Anil, B.; Song, B.; Tang, Y.; Raleigh, D. P. *J. Am. Chem. Soc.* **2004**, *126*, 13194−13195.
(13) Kraulis, P. J. *J. Appl. Crystallogr.* **1991**, *24*, 946−950.
(14) Wuthrich, K. *NMR of Proteins and Nucleic Acids*; John Wiley and Sons: New York, 1986.

JA057874B