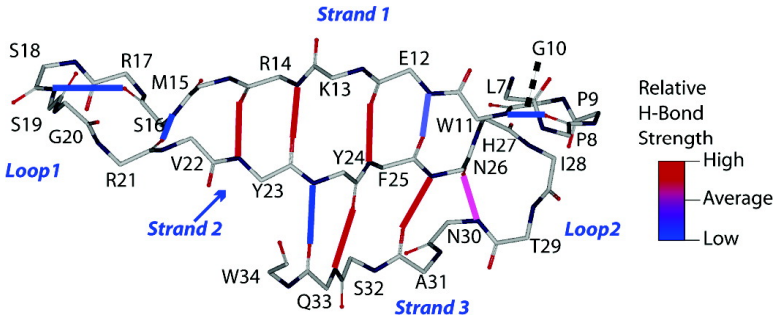
Article

Toward Assessing the Position-Dependent Contributions of Backbone Hydrogen Bonding to β-Sheet Folding Thermodynamics Employing Amide-to-Ester Perturbations

Songpon Deechongkit, Philip E. Dawson, and Jeffery W. Kelly

J. Am. Chem. Soc., **2004**, 126 (51), 16762-16771• DOI: 10.1021/ja045934s • Publication Date (Web): 01 December 2004



More About This Article

Additional resources and features associated with this article are available within the HTML version:

- Supporting Information
- Links to the 2 articles that cite this article, as of the time of this article download
- Access to high resolution figures
- Links to articles and content related to this article
- Copyright permission to reproduce figures and/or text from this article

View the Full Text HTML

J|A|C|S
ARTICLES
Published on Web 12/01/2004

Toward Assessing the Position-Dependent Contributions of Backbone Hydrogen Bonding to β-Sheet Folding Thermodynamics Employing Amide-to-Ester Perturbations

Songpon Deechongkit,[†] Philip E. Dawson,[†,‡] and Jeffery W. Kelly*[,†]

Contribution from the Departments of Chemistry and Cell Biology, The Skaggs Institute of Chemical Biology, The Scripps Research Institute, 10550 North Torrey Pines Road, BCC 506, La Jolla, California 92037

Received July 7, 2004; E-mail: jkelly@scripps.edu

Abstract: An amide-to-ester backbone substitution in a protein is accomplished by replacing an α-amino acid residue with the corresponding α-hydroxy acid, preserving stereochemistry, and conformation of the backbone and the structure of the side chain. This substitution replaces the amide NH (a hydrogen bond donor) with an ester O (which is not a hydrogen bond donor) and the amide carbonyl (a strong hydrogen bond acceptor) with an ester carbonyl (a weaker hydrogen bond acceptor), thus perturbing folding energetics. Amide-to-ester perturbations were used to evaluate the thermodynamic contribution of each hydrogen bond in the PIN WW domain, a three-stranded β-sheet protein. Our results reveal that removing a hydrogen bond donor destabilizes the native state more than weakening a hydrogen bond acceptor and that the degree of destabilization is strongly dependent on the location of the amide bond replaced. Hydrogen bonds near turns or at the ends of β-strands are less influential than hydrogen bonds that are protected within a hydrophobic core. β-Sheet destabilization caused by an amide-to-ester substitution cannot be directly related to hydrogen bond strength because of differences in the solvation and electrostatic interactions of amides and esters. We propose corrections for these differences to obtain approximate hydrogen bond strengths from destabilization energies. These corrections, however, do not alter the trends noted above, indicating that the destabilization energy of an amide-to-ester mutation is a good first-order approximation of the free energy of formation of a backbone amide hydrogen bond.

Introduction

The hydrophobic effect and hydrogen bonding are the two major forces that stabilize a protein's native state relative to an ensemble of unfolded conformations.[1] Numerous studies have focused on their energetic contributions to protein folding.[1−9] Since the hydrophobic effect largely originates from interactions between side chains, it can be perturbed by traditional site-directed mutagenesis employing side chains of variable structure and hydrophobicity.[10−13] Side chain hydrogen bonding interactions can also be altered by traditional site-directed mutagenesis, for example, by using Ser to Ala or Tyr to Phe mutations.[13−18] The experimental ease of traditional mutagenesis has facilitated many studies concerning the role of the hydrophobic effect and side chain hydrogen bonding in protein folding energetics.

In contrast, data addressing specific backbone hydrogen bonding contributions to the thermodynamics and kinetics of protein folding are scarce because ordinary mutagenesis does not alter the backbone. There are backbone amide replacements that can be incorporated into polypeptides using solid phase peptide synthesis approaches that perturb hydrogen bonding. Each alternative (Figure 1) has advantages and disadvantages. Replacement of a secondary amide by a tertiary amide is achieved employing a C-terminal Pro substitution in a dipeptide subsequence, removing an N−H hydrogen bond donor.[19] The

[†] Department of Chemistry.
[‡] Department of Cell Biology.

(1) Dill, K. A. *Biochemistry* **1990**, *29*, 7133−7155.
(2) Grantcharova, V.; Alm, E. J.; Baker, D.; Horwich, A. L. *Curr. Opin. Struct. Biol.* **2001**, *11*, 70−82.
(3) Gruebele, M. *Curr. Opin. Struct. Biol.* **2002**, *12*, 161−168.
(4) Miranker, A. D.; Dobson, C. M. *Curr. Opin. Struct. Biol.* **1996**, *6*, 31−42.
(5) Southall, N. T.; Dill, K. A.; Haymet, A. D. J. *J. Phys. Chem. B* **2002**, *106*, 521−533.
(6) Levitt, M. *Curr. Opin. Struct. Biol.* **1991**, *1*, 224−229.
(7) Rose, G. D. *Annu. Rev. Biophys. Biomol. Struct.* **1993**, *22*, 381−415.
(8) (a) Englander, S. W. *Annu. Rev. Biophys. Biomol. Struct.* **2000**, *29*, 213−238. (b) Englander, S. W.; Mayne, L. *Annu. Rev. Biophys. Biomol. Struct.* **1992**, *21*, 243−265.
(9) (a) Ciani, B.; Jourdan, M.; Searle, M. S. *J. Am. Chem. Soc.* **2003**, *125*, 9038−9047. (b) Sharman, G. J.; Griffiths-Jones, S. R.; Jourdan, M.; Searle, M. S. *J. Am. Chem. Soc.* **2001**, *123*, 12318−12324. (c) Searle, M. S. *J. Chem. Soc., Perkin Trans.* **2001**, *2*, 1011−1020. (d) Griffiths-Jones, S. R.; Searle, M. S. *J. Am. Chem. Soc.* **2000**, *122*, 8350−8356. (e) Colley, C. S.; Griffiths-Jones, S. R.; George, M. W.; Searle, M. S. *Chem. Commun.* **2000**, *7*, 593−594.
(10) Meiering, E. M.; Serrano, L.; Fersht, A. R. *J. Mol. Biol.* **1992**, *225*, 585−589.
(11) Vendruscolo, M.; Paci, E.; Dobson, C. M.; Karplus, M. *Nature* **2001**, *409*, 641−645.
(12) Jager, M.; Nguyen, H.; Crane, J. C.; Kelly, J. W.; Gruebele, M. *J. Mol. Biol.* **2001**, *311*, 373−393.
(13) Chen, Y. W.; Fersht, A. R.; Henrick, K. *J. Mol. Biol.* **1993**, *234*, 1158−1170.
(14) Takano, K.; Yamagata, Y.; Kubota, M.; Funahashi, J.; Fujii, S. *Biochemistry* **1999**, *38*, 6623−6629.
(15) Bulaj, G.; Goldenberg, D. P. *J. Mol. Biol.* **2001**, *313*, 639−656.
(16) Pace, C. N.; Horn, G.; Hebert, E. J.; Bechert, J.; Shaw, K.; Urbanikova, L.; Scholtz, J. M.; Sevcik, J. *J. Mol. Biol.* **2001**, *312*, 393−404.
(17) Pace, C. N. *Methods Enzymol.* **1995**, *259*, 538−554.
(18) Shirley, B. A.; Stanssens, P.; Hahn, U.; Pace, C. N. *Biochemistry* **1992**, *31*, 725−732.

10.1021/ja045934s CCC: $27.50

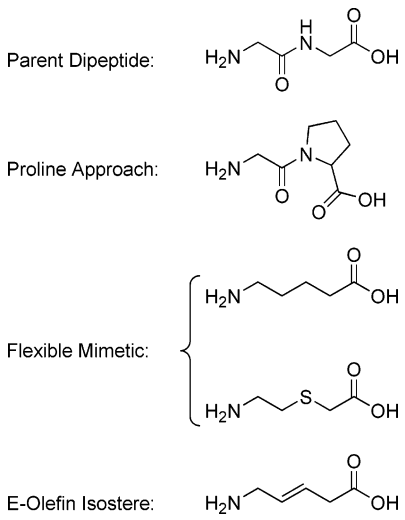


Figure 1. Using a Gly-Gly dipeptide as an example, four substitutions that perturb the ability of the amide backbone to participate in hydrogen bonding are shown: (from top) amide to secondary amide via proline substitution, amide to dimethylene, amide to thiomethylene, and amide to *E*-olefin isostere.

disadvantage of Pro incorporation is that it strongly alters the conformational preferences of the backbone, as proline prefers to be in reverse β-turns or extended polyproline type II conformations.[20,21] Another approach is to replace the amide bond of interest with a dipeptide mimetic lacking an amide bond, such as 5-aminopentanoic acid or (2-aminoethylsulfanyl)acetic acid.[22] These dipeptide mimetics often have altered conformational preferences about what was the amide bond and different side chain substructures owing to the difficulty involved in the synthesis of dipeptide mimetics retaining the original side chains and chirality. Dipeptide isosteres where the amide bond is replaced with a *trans*-alkene functional group, referred to as an "*E*-olefin isostere", are ideal backbone amide replacements as the trans-olefin, which is incapable of hydrogen bonding, replaces the trans amide bond, which is capable of serving as both a hydrogen bond donor and acceptor (Figure 1). However, the synthetic challenge associated with the stereospecific synthesis of *E*-olefin isosteres limits their utility at this time.[23,24]

Currently, the most accessible and attractive method for the perturbation of backbone hydrogen bonding is to replace the amide bond with an ester bond (Figure 2).[25−36] The synthesis

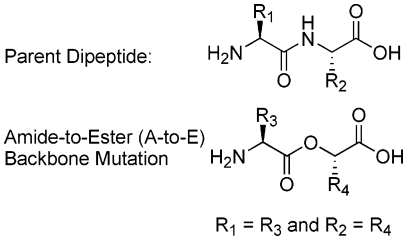


Figure 2. Illustration of the A-to-E backbone mutation strategy, where the amide bond of the parent dipeptide is replaced with an ester bond. Note that the side chains of both amino acids (R_1 and R_2) are preserved in the α-hydroxy acid dipeptide substructure.

of amide-to-ester (A-to-E) variants of a protein is accomplished by replacing the α-amino acid residue of interest by the corresponding chiral α-hydroxy acid residue bearing the same side chain. Incorporating α-hydroxy acids into a protein can be done either by chemical synthesis or by biosynthesis,[35−36] the latter using technology developed by Schultz and co-workers.[30−33] Proteins or peptides that contain an ester linkage instead of an amide linkage are called depsipeptides ("depsi" comes from the Greek word "*depsidi*", which means ester).[34]

The A-to-E substitution eliminates the NH hydrogen bond donor by replacing it with an O (Figure 2).[37] Furthermore, the carbonyl of an ester is a weaker hydrogen bond acceptor than that of an amide. Thus, A-to-E mutations eliminate one of the hydrogen bonds that a given backbone amide could make (by removing the donor) and perturbs the other one (by weakening the acceptor). An A-to-E substitution is a desirable backbone amide replacement because amides and esters have similar conformational preferences. Previous reports, including a recent crystal structure of a depsipeptide and conformational data on polyesters, reveal that depsipeptides and their all-amide counterparts have similar bond angles and lengths, prefer similar ϕ and ψ dihedral angles, and strongly prefer the trans ester conformation, like secondary amides in proteins.[27,38−42]

It is important to maintain the structure of the side chains when using A-to-E mutations to study backbone hydrogen bonding in protein folding, because altering the side chain in an A-to-E mutant contaminates the effect of perturbing backbone hydrogen bonding with the effect of changing the side chain. The above is especially true with β-sheet structures, in which interactions between side chains on different strands are often crucial to the stability of the folded state. As a result, the scope of A-to-E mutation as a method for studying backbone H-bonding has been limited by the lack of commercial availability of some of the α-hydroxy acid analogues of the α-amino acids. To avoid this dilemma, our laboratory has synthesized the α-hydroxy acid equivalents of the α-amino acids not currently commercially available, thus allowing all possible A-to-E

(19) Wales, T. E.; Fitzgerald, M. C. *J. Am. Chem. Soc.* **2001**, *123*, 7709−7710.
(20) Sibanda, B. L.; Thornton, J. M. *J. Mol. Biol.* **1993**, *229*, 428−447.
(21) Hutchinson, E. G.; Thornton, J. M. *Protein Sci.* **1994**, *3*, 2207−2216.
(22) Ferguson, N.; Pires, J. R.; Toepert, F.; Johnson, C. M.; Pan, Y. P.; Volkmer-Engert, R.; Schneider-Mergener, J.; Daggett, V.; Oschkinat, H.; Fersht, A. *Proc. Natl. Acad. Sci. U.S.A.* **2001**, *98*, 13008−13013.
(23) Hart, S. A.; Sabat, M.; Etzkorn, F. A. *J. Org. Chem.* **1998**, *63*, 7580−7581.
(24) Gardner, R. R.; Liang, G.-B.; Gellman, S. H. *J. Am. Chem. Soc.* **1999**, *121*, 1806−1816.
(25) Bramson, H. N.; Thomas, N. E.; Kaiser, E. T. *J. Biol. Chem.* **1985**, *260*, 15452−15457.
(26) Coombs, G. S.; Rao, M. S.; Olson, A. J.; Dawson, P. E.; Madison, E. L. *J. Biol. Chem.* **1999**, *274*, 24074−24079.
(27) Arad, O.; Goodman, M. *Biopolymers* **1990**, *29*, 1652−1668.
(28) Baskakov, I. V.; Kumar, R.; Srinivasan, G.; Ji, Y. S.; Bolen, D. W.; Thompson, E. B. *J. Biol. Chem.* **1999**, *274*, 10693−10696.
(29) Aleman, C.; Puiggali, J. *Chem. Phys.* **1997**, *222*, 9−15.
(30) Thorson, J. S.; Chapman, E.; Schultz, P. G. *J. Am. Chem. Soc.* **1995**, *117*, 9361−9362.
(31) Koh, J. T.; Cornish, V. W.; Schultz, P. G. *Biochemistry* **1997**, *36*, 11314−11322.
(32) Chapman, E.; Thorson, J. S.; Schultz, P. G. *J. Am. Chem. Soc.* **1997**, *119*, 7151−7152.
(33) Shin, I.; Ting, A. Y.; Schultz, P. G. *J. Am. Chem. Soc.* **1997**, *119*, 12667−12668.
(34) Shemyakin, M. M.; Ovchinnikov, Y. A. *Recent Dev. Chem. Nat. Carbon Compd.* **1967**, *2*, 1−46.
(35) (a) Blankenship, J. W.; Balambika, R.; Dawson, P. E. *Biochemistry* **2002**, *41*, 15676−15684. (b) Beligere, G. S.; Dawson, P. E. *J. Am. Chem. Soc.* **2000**, *122*, 12079−12082. (c) Schnolzer, M.; Alewood, P.; Jones, A.; Alewood, D.; Kent, S. B. *Int. J. Pept. Protein Res.* **1992**, *40*, 180−193.
(36) Gallo, E. A.; Gellman, S. H. *J. Am. Chem. Soc.* **1993**, *115*, 9774−9788.
(37) Arnett, E. M.; Mitchell, E. J.; Murty, T. S. S. R. *J. Am. Chem. Soc.* **1974**, *96*, 3875−3891.
(38) Ohyama, T.; Oku, H.; Yoshida, M.; Katakai, R. *Biopolymers* **2001**, *58*, 636−642.
(39) Tseng, C. C.; Bruner, S. D.; Kohli, R. M.; Marahiel, M. A.; Walsh, C. T. *Biochemistry* **2002**, *41*, 13350−13359.
(40) Wiberg, K. B.; Laidig, K. E. *J. Am. Chem. Soc.* **1987**, *109*, 5935−5943.
(41) Ingwall, R. T.; Goodman, M. *Macromolecules* **1974**, *7*, 598−605.
(42) Bunting, J. W.; Kanter, J. P. *J. Am. Chem. Soc.* **1993**, *115*, 11705−11715.

substitutions to be made with retention of side chain structure and α-carbon configuration (Figure 2).[52]

The contribution of each backbone hydrogen bond to the stability of the three-stranded β-sheet structure of the PIN WW domain was probed using twenty A-to-E variants. These variants were required to perturb each of the eleven hydrogen bonds in the PIN WW domain β-sheet. The free energy of destabilization resulting from an A-to-E mutation cannot be directly ascribed to hydrogen bond strength because of possible electrostatic repulsions between the ester oxygen and the amide carbonyl on the opposing strand and energetic differences between ester and amide solvation. We propose electrostatic and solvation corrections to the measured free energies of destabilization caused by the A-to-E mutations that provide a preliminary estimate of hydrogen bond strengths; however, more sophisticated corrections and data from A-to-E scans of other proteins will be needed before accurate position-dependent hydrogen bond strengths can be established. Importantly, both the primary data and the corrected data reveal the same trend: that hydrogen bonds in the low dielectric microenvironment of a hydrophobic cluster contribute substantially more to protein stability than do solvent-exposed hydrogen bonds. This indicates, first, that the measured free energies of destabilization are good approximations of hydrogen bond strengths, and, second, that there are "hot spot" hydrogen bonds in a protein that contribute disproportionately to stability.

Design

The PIN1 protein has two domains and functions in cell cycle regulation by mediating protein−protein interactions.[28,43−47] The WW domain can be excised from the full length PIN1 protein as an independently folded, 34-residue, three-stranded β-sheet miniprotein with two intervening loops.[51b] This folded miniprotein retains its proline-rich phosphoserine/phosphothreonine peptide binding properties.[47] The PIN WW domain is monomeric, small enough to be conveniently synthesized by solid-phase peptide synthesis, and is one of the best studied β-sheet proteins.[12,22,48−50] Moreover, PIN WW domain variants prepared by traditional recombinant methods reveal that this sequence is highly tolerant of conservative side chain mutations, affording folded structures that allow one to probe the role of almost every side chain in protein folding energetics. Kinetic and thermodynamic analyses of these variants reveal that hydrophobic side chain packing is important for thermodynamic stability and that the formation of the 4:6 loop 1 conformation is rate limiting for WW domain folding.[12] Recently, we communicated a summary of the kinetic and thermodynamic effects of A-to-E mutations on PIN WW domain folding, focusing mostly on the role that individual backbone H-bonds play in the transition state structure, and thus in kinetics.[53] In contrast, this manuscript focuses on the effect of A-to-E variants, and therefore of individual backbone H-bonds, on folding thermodynamics. This manuscript also introduces methodology to correct the experimental destabilization free energies of A-to-E mutants for the solvation differences and electrostatic repulsions associated with the A-to-E mutations to obtain estimates of the free energies of formation of individual backbone H-bonds.

The crystal structure of the PIN WW domain, taken from the crystal structure of the entire PIN1 protein, and the solution NMR structure of the isolated PIN WW domain reveal 11 backbone hydrogen bonds (Figure 3).[51] The sequence of the PIN WW domain employed for these studies is shown below.

```
          6                                   39
wt PIN:   KLPPGWEKRMSRSSGRVYYFNHITNASQWERPSG
```

Backbone hydrogen bonding requires one amide backbone carbonyl to serve as the hydrogen bond acceptor, while another backbone amide NH serves as the donor. As a result, two A-to-E backbone mutations can be made to probe each backbone hydrogen bond: one A-to-E mutation eliminates the donor, while the other weakens the acceptor. For example, the hydrogen bond between S16 and S19 in the β-turn of loop 1 can be eliminated by replacing S19 by the α-hydroxy acid equivalent of serine, thus eliminating the hydrogen bond donor. Alternatively, this hydrogen bond can be weakened by substituting R17 with its corresponding α-hydroxy acid to weaken the hydrogen bond acceptor of S16. In all, 20 A-to-E variants of the PIN WW domain were synthesized to probe the 11 hydrogen bonds in the three-stranded β-sheet. These variants, the contributing residues for each hydrogen bond, and their location in the context of the PIN WW domain β-sheet are listed in Table 1. We propose the use of lower case Greek letters to refer to the α-hydroxy acid equivalents of the L-α-amino acids utilized in this study (e.g., α represents lactic acid, the α-hydroxy acid equivalent of alanine, A).

Amino acids:	A	C	D	E	F	G	H	I	K	L	M	N	P	Q	R	S	T	V	W	Y
Hydroxy acids:	α	χ	δ	ε	ϕ	γ	η	ι	κ	λ	μ	ν	n/a	θ	ρ	σ	τ	ϖ	ω	ψ

Results

Synthesis. The synthesis of PIN WW domain A-to-E mutants was carried out using a Boc/benzyl solid-phase peptide synthesis strategy, as described previously.[35] The wild type PIN WW domain (wt PIN) and its phosphopeptide ligand were synthesized as previously described.[48] All α-hydroxy acids that were not commercially available with appropriate side chain protection for a Boc/benzyl synthesis strategy were synthesized as recently described.[52]

Sample Preparation. One of two methods were employed to ready the PIN WW domain A-to-E variants for biophysical studies. In the first method, utilized for the majority of the A-to-E variants, HPLC-purified PIN WW domains were lyophilized, redissolved in deionized water, eluted through a size exclusion column in the presence of high salt, and dialyzed against an appropriate aqueous buffer. In the second method, HPLC-purified PIN WW domains were lyophilized, redissolved

(43) Lu, K. P.; Hanes, S. D.; Hunter, T. *Nature* **1996**, *380*, 544−547.
(44) Verdecia, M. A.; Bowman, M. E.; Lu, K. P.; Hunter, T.; Noel, J. P. *Nat. Struct. Biol.* **2000**, *7*, 639−643.
(45) Sudol, M.; Hunter, T. *Cell* **2000**, *103*, 1001−1004.
(46) Sudol, M. *Exp. Mol. Med.* **1996**, *28*, 65−69.
(47) Staub, O.; Rotin, D. *Structure* **1996**, *4*, 495−499.
(48) Deechongkit, S.; Kelly, J. W. *J. Am. Chem. Soc.* **2002**, *124*, 4980−4986.
(49) (a) Kaul, R.; Angeles, A. R.; Jager, M.; Powers, E. T.; Kelly, J. W. *J. Am. Chem. Soc.* **2001**, *123*, 5206−5212. (b) Kaul, R.; Deechongkit, S.; Kelly, J. W. *J. Am. Chem. Soc.* **2002**, *124*, 11900−11907.
(50) Ferguson, N.; Johnson, C. M.; Macias, M.; Oschkinat, H.; Fersht, A. *Proc. Natl. Acad. Sci. U.S.A.* **2001**, *98*, 13002−13007.
(51) (a) Ranganathan, R.; Lu, K. P.; Hunter, T.; Noel, J. P. *Cell* **1997**, *89*, 875−886. (b) Kowalski, J. A.; Liu, K.; Kelly, J. W. *Biopolymers* **2002**, *63*, 111−121.
(52) Deechongkit, S.; You, S.-L.; Kelly, J. W. *Org. Lett.* **2004**, *6*, 497−500.
(53) Deechongkit, S.; Nguyen, H.; Powers, E. T.; Dawson, P. E.; Gruebele, M.; Kelly, J. W. *Nature* **2004**, *430*, 101−105.

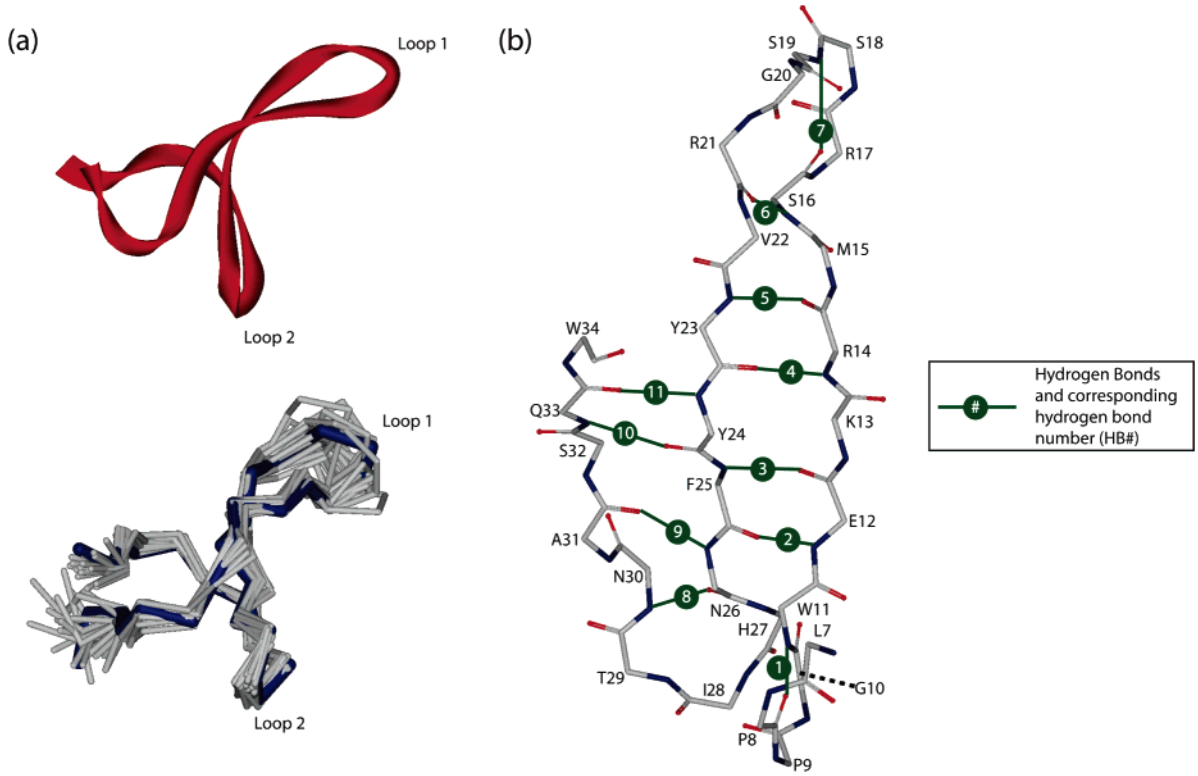


Figure 3. (a) Crystal structure of the PIN WW domain excised from the crystal structure of the full-length two domain PIN1 protein (top) and solution structure of the isolated PIN WW domain (6−39) determined by NMR spectroscopy (bottom). (b) Backbone of the β-sheet region (from the crystal structure) where hydrogen bonds and their corresponding hydrogen bond number (HB#) are indicated by green lines.

Table 1. List of A-to-E Mutants, Their Backbone Hydrogen Bonding Context, Their Pairing Residues, and Their Thermodynamic Characteristics[a]

mutation	affected backbone H-bond	location	pairing residue	T_m (°C)	ΔG_f (kJ/mol)
wt PIN	N/A	N/A	N/A	59 ± 0.1	−13.9 ± 0.4
L7λ	N/A	N/A	N/A	56.7 ± 0.1	−13.9 ± 0.3
W11ω	donor HB1	outer β-turn shown in Figures 3 and 6	N/A	45.2 ± 0.2	−9.2 ± 0.2
E12ϵ	donor HB2	end of sheet	N26	22.2 ± 0.1	−7.6 ± 0.4
K13κ	acceptor HB3	middle of sheet	F25	46.4 ± 0.1	−10.9 ± 0.4
R14ρ	donor HB4	middle of sheet	Y24	n.d.	2.5 ± 0.2
M15μ	acceptor HB5	middle of sheet	Y23	55.3 ± 0.3	−12.2 ± 0.2
S16σ	donor HB6	first loop 1 H-bond	V22	42.2 ± 0.2	−9.2 ± 0.2
R17ρ	acceptor HB7	define β-turn in loop 1	S19	49.1 ± 0.1	−12.6 ± 0.1
S19σ	donor HB7	define β-turn in loop 1	R17	38.4 ± 0.1	−11.3 ± 0.2
V22ϖ	acceptor HB6	first loop 1 H-bond	S16	56.7 ± 0.1	−11.3 ± 0.4
Y23ψ	donor HB5	middle of sheet	M15	21 ± 0.2	−4.6 ± 0.3
Y24ψ	acceptor HB4	end of sheet	W34	25 ± 0.4	−9.3 ± 0.1
	donor HB11	middle of sheet	R14		
F25ϕ	donor HB3	middle of sheet	K13	n.d.	3.7 ± 0.3
	acceptor HB10	middle of sheet	Q33		
N26ν	donor HB9	middle of sheet	S32	n.d.	6.3 ± 0.4
	acceptor HB2	end of sheet	E12		
H27η	acceptor HB8	loop 2	N30	38.7 ± 0.1	−10.5 ± 0.2
N30ν	donor HB8	loop 2	H27	24.8 ± 0.2	−6.3 ± 0.2
A31α	donor to side-chain	backbone to side chain	T29	45.3 ± 0.1	−10.5 ± 0.6
S32σ	acceptor HB9	middle of sheet	N26	41.5 ± 0.3	−9.7 ± 0.4
Q33θ	donor HB10	middle of sheet	F25	n.d.	−0.8 ± 0.1
W34ω	acceptor HB11	end of sheet	Y24	49.5 ± 0.1	−11.8 ± 0.1

[a] Values of ΔG_f (at 2 °C) are derived from guanidine hydrochloride denaturation curves. The A-to-E mutants R14ρ, F25ϕ, N26ν, and Q33θ required the addition of trimethylamine *N*-oxide (TMAO) to determine ΔG_f.

in deionized water to make a concentrated stock solution, and diluted with buffer. Although the former approach was preferred because it further purified the protein, it was not suitable for the R17ρ, S19σ, and H27η variants of the PIN WW domain. The ester bonds of these depsipeptides were hydrolyzed within 12−24 h under high salt conditions, likely because the ester bonds are located in highly solvent accessible loops. Therefore, these A-to-E variants were readied for biophysical studies by the second method.

PIN WW Domain Characterization. A number of approaches were utilized to ensure that the PIN WW domain A-to-E variants adopted a normally folded structure. The A-to-E variants were characterized by far-UV CD and fluorescence spectroscopy. All of the folded variants exhibited the charac-

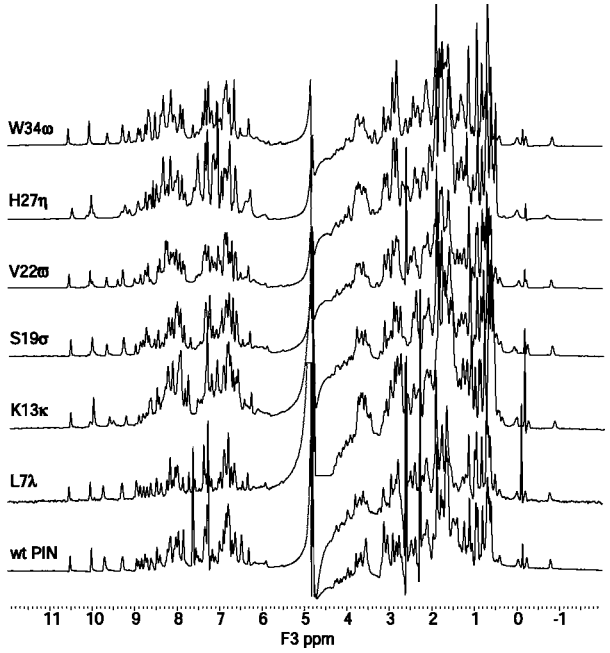


Figure 4. 1D ^{1}H NMR spectra (600 MHz) of wt PIN WW domain and selected A-to-E PIN WW domain mutants (1 mM) in 9:1 H_2O/D_2O (v/v) buffered with 20 mM sodium phosphate (pH 7.0, 2 °C).

teristic far-UV CD maximum at 227 nm, as well as a fluorescence emission spectrum (maximum at 342 nm) similar to that of the wt PIN WW domain. The tertiary structure of a representative set of the PIN WW domain A-to-E variants (L7λ, K13κ, S19σ, V22ϖ, H27η, W34ω) was further scrutinized by 1D ^{1}H NMR spectroscopy employing a Watergate solvent suppression pulse sequence.[49,51b] The variants analyzed displayed well-dispersed resonances in both the downfield aromatic/amide region (δ 6.5−10.0 ppm) and the upfield aliphatic region, which is characteristic of a properly folded WW domain (Figure 4). The similarity between the 1D ^{1}H NMR fingerprints of the A-to-E variants and the wt PIN WW domain is consistent with the expectation that these depsipeptides adopt a normal tertiary structure. To confirm the integrity of the folded state, all PIN WW variants, except for R14ρ, F25ϕ, N26ν, and Q33θ (which were not folded), were subjected to a ligand-binding assay. The phosphorylated peptide YSPTpSPS (CTD-S5) bound to all of the A-to-E variants with a free energy comparable to that of the wt PIN-ligand interaction (20−25 kJ/mol), implying that the variants adopted a native structure.

To apply the two-state assumption to extract thermodynamic data from the denaturation curves of the PIN WW domain A-to-E variants, they must remain monomeric in solution. Their association state was characterized by sedimentation equilibrium analytical ultracentrifugation studies. Under the conditions employed to make biophysical measurements, all of the PIN WW domain A-to-E variants (both folded and unfolded) exhibited radial concentration distribution profiles that fit best to a single species model with a molecular weight corresponding to monomer (~4 kD) and do not fit well to models invoking heterogeneity (data not shown).

Steady-State Thermodynamic Measurements. The aim of this study was to gain insight into the contribution of each of the eleven hydrogen bonds to the thermodynamic stability of the PIN WW domain. Stability parameters that enable comparison of all these variants include the midpoint of thermal denaturation curves (T_m) (Table 1)[54] and the free energy changes based on guanidine hydrochloride-induced unfolding at 2 °C (ΔG_f). Thermal denaturation curves were recorded based on the decrease in ellipticity of the characteristic maximum in the CD spectrum of the PIN WW domain at 227 nm. Chaotrope denaturation curves at 2 °C were recorded based either on the decrease in ellipticiy at 227 nm or on the decrease in the normalized fluorescence signal (the ratio of the fluorescence at 342 and 355 nm) as a function of increasing guanidine hydrochloride concentration. Though chaotrope denaturation experiments were carried out at 2 °C to enable a direct comparison of as many variants as possible, some of the A-to-E variants, namely R14ρ, F25ϕ, N26ν, and Q33θ, were partially unfolded at 2 °C in the absence of denaturant and could not be studied by this approach (a pretransition baseline is necessary for data fitting). Therefore, ΔG_f for these variants was measured by recording their denaturation curves in the presence of varying concentrations of trimethylamine *N*-oxide (TMAO), an osmolyte that can stabilize proteins by up to 16 kJ/mol.[28,55−56] With increasing concentrations of TMAO, these variants became more stable, showing a higher ellipticity at 227 nm and a higher normalized fluorescence signal. The normalized fluorescence signal was related to the fraction of unfolded protein as established from other guanidine denaturation experiments, and plots of the fraction protein unfolded as a function of TMAO concentration were constructed for each of these mutants. Finally, ΔG_f in the absence of TMAO was obtained by extrapolating this plot to 0 M TMAO. The steady-state thermodynamic data and experimental errors are shown in Table 1.[53]

Discussion

Utilization of A-to-E Backbone Mutations. The importance of maintaining side chain integrity when examining backbone hydrogen bonding can be illustrated using A-to-E mutants of W11. Replacement of W11 with lactic acid (α) results in WW domain unfolding, likely because the smaller methyl side chain of the lactic acid disrupts the hydrophobic core (data not shown). Because the effect of altering the side chain of W11 is combined with the effect of perturbing the backbone hydrogen bond, information on the contribution of backbone hydrogen bonding to protein folding energetics at this residue cannot be extracted from the destabilization energy of the W11α mutant. When the tryptophan residue is replaced by ω (W11ω), the resulting WW domain remains folded and the influence of that backbone hydrogen bond can be ascertained, provided that solvation and electrostatic corrections are made (see below). The stability of W11ω is decreased by 4.6 kJ/mol relative to wt PIN, and since no side chain interaction is interrupted, the decrease in stability can be attributed largely to the loss of a backbone hydrogen bond. This hydrogen bond, from the NH of the G10-W11 amide

(54) ΔG values can be calculated from a thermal denaturation curve if the change in ΔC_p upon unfolding ($\Delta C_{p,u}$) is known and the thermal denaturation curve contains both native and unfolded baselines.[48] Since not all A-to-E variants have native baselines, their ΔG values cannot be obtained. Thus, for consistency, the ΔG values discussed herein are those obtained by chaotrope denaturation or osmolyte renaturation.

(55) (a) Baskakov, I.; Bolen, D. W. *J. Biol. Chem.* **1998**, *273*, 4831−4834. (b) Bolen, D. W.; Baskakov, I. V. *J. Mol. Biol.* **2001**, *310*, 955−963. (c) Auton, M.; Bolen, D. W. *Biochemistry* **2004**, *43*, 1329−1342.

(56) (a) Kumar, R.; Lee, J. C.; Bolen, D. W.; Thompson, E. B. *J. Biol. Chem.* **2001**, *276*, 18146−18152. (b) Liu, Y.; Bolen, D. W. *Biochemistry* **1995**, *34*, 12884−12891. (c) Wang, A.; Bolen, D. W. *Biochemistry* **1997**, *36*, 9101−9108.

bond to the carbonyl of the P8−P9 amide bond, defines a reverse turn encompassing residues P8 to W11. The existence of this turn was confirmed by an evaluation of the ϕ and ψ dihedral angles of residues P8, P9, G10, and W11 in the NMR structural ensemble (Figure 3).[20]

L7λ A-to-E Mutant and the Energy of the Unfolded State. The thermodynamic stability of a protein is measured by the difference in free energy between the native and unfolded states. In traditional mutagenesis studies, it is assumed that mutations do not change the free energy of the variant's unfolded state compared to that of the wild type. Thus, the difference in free energies between variant and wild type is assumed to result entirely from changes in the native state.[58] Though the unfolded state of both wt PIN and the A-to-E variants can be assumed to have the same ensemble of unordered structures, the solvation energies of the all-amide backbone of wt PIN and the depsipeptide backbone of the A-to-E variants could be different in the unfolded state. The wt PIN WW domain has 33 amide bonds, while the A-to-E variants have 32 amide bonds and 1 ester bond in the backbone. It is well-known that solvating amides in aqueous solution are more favorable than solvating esters.[37,40,62] Thus, it is possible that an A-to-E mutation affects the unfolded state as well as the native state.

The A-to-E mutant L7λ was used to determine whether changes in the solvation of the unfolded state affect the ΔG_f value of A-to-E mutants, since the ester is fully solvated in both the folded and unfolded states in this mutant. It happens that ΔG_f values for L7λ and the wt PIN WW domain are identical within experimental error (Table 1). Therefore, the solvation differences between the unfolded states of the wt PIN WW domain and its A-to-E mutants were neglected.

Relationship between Measured ΔΔG Values and Hydrogen Bond Strengths. Since A-to-E mutations primarily affect backbone hydrogen bonds, it is tempting to equate differences between the folding free energies of the wt PIN WW domain and its A-to-E mutants ($\Delta\Delta G_f = \Delta G_{f,mut} - \Delta G_{f,wt}$) to hydrogen bond strengths (or strength difference, for cases where an amide CO is replaced by and ester CO). We define the strength of a given hydrogen bond as the free energy of the electrostatic interaction between desolvated amide CO and NH groups in the native state of the PIN WW domain. The value of $\Delta\Delta G_f$ for an A-to-E mutant is only equal to the hydrogen bonding energy when the electrostatic repulsions and solvation differences discussed in the Introduction are negligible. However, this is often not true for the electrostatic repulsions, and almost never true for the solvation differences. We describe a simple, first-order procedure for making such corrections below.

Corrections for solvation energy differences between the wt PIN WW domain and A-to-E mutants require knowledge of the solvation energies of amides and esters and the extent of backbone burial at each residue. The solvation energy of amides and esters in protein folding was estimated from water−octanol partition energies of amide and ester functional groups: 11.3 and 5.5 kJ/mol, respectively.[29,59−65] Amides and esters consist of two fragments that are involved in hydrogen bond formation, i.e., carbonyl and NH (for amide) or carbonyl and O (for ester). While it is debatable how much each fragment contributes to the overall partition energies of amide and ester functional groups, we approximate that each fragment contributes equally to the partition energy of an amide or ester substructure.[66] In other words, partition energies for each fragment of an amide and an ester are 5.7 and 2.8 kJ/mol, respectively. The extent of CO and NH burial at each residue was obtained from the average NMR structure of the PIN WW domain using a 1.4 Å probe radius and is summarized in Table 2. Using this information, $\Delta\Delta G_f$ values were corrected for solvation effects in the folded state by scaling the correction based on percent solvent accessibility. In other words, no correction was applied to fully solvated functional groups, the full correction was applied to fully desolvated functional groups, and for intermediate cases, the percent desolvation was used to derive a fractional correction (Table 2).

Table 2. Summary of the Stability of A-to-E Variants Relative to the wt PIN WW Domain ($\Delta\Delta G_f = \Delta G_{f,mut} - \Delta G_{f,wt}$), Their Amide Solvent Accessibility Based on the NMR Solution Structure, and the Adjusted $\Delta\Delta G_f$ Corrected for Solvation Differences as Described in the Text[e]

					solvation correction (kJ/mol)[c]		
			% solvent accessible				
mutation	hydrogen bond	observed $\Delta\Delta G_f$ (kJ/mol)	–CO–[a]	–NH–[b]	cor from –CO–[c]	cor from –NH–[c]	adjusted $\Delta\Delta G_f$[d]
L7λ	N/A	0	25	29	0	0	0
W11ω	HB1	4.6 ± 0.4	0	18	2.9	2.4	9.9 ± 0.4
E12ε	HB2	6.3 ± 0.5	35	0	1.9	2.9	11.1 ± 0.5
K13κ	HB3	2.9 ± 0.5	0	34	2.9	1.9	7.7 ± 0.5
R14ρ	HB4	16.4 ± 0.4	55	0	1.3	2.9	20.6 ± 0.4
M15μ	HB5	1.7 ± 0.4	0	40	2.9	1.7	6.3 ± 0.4
S16σ	HB6	4.6 ± 0.4	40	0	1.7	2.9	9.2 ± 0.4
R17ρ	HB7	1.3 ± 0.3	10	30	2.6	2.0	5.9 ± 0.3
S19σ	HB7	2.5 ± 0.4	28	20	2.1	2.3	6.9 ± 0.4
V22ϖ	HB6	2.5 ± 0.5	0	30	2.9	2.0	7.4 ± 0.5
Y23ψ	HB5	9.2 ± 0.4	20	0	2.3	2.9	14.4 ± 0.4
Y24ψ	HB4, HB11	4.6 ± 0.3	0	0	2.9	2.9	10.4 ± 0.3
F25φ	HB3, HB10	17.6 ± 0.4	0	0	2.9	2.9	23.4 ± 0.4
N26ν	HB2, HB9	20.2 ± 0.5	0	0	2.9	2.9	26.0 ± 0.5
H27η	HB8	3.4 ± 0.4	0	0	2.9	2.9	9.2 ± 0.4
N30ν	HB8	7.6 ± 0.4	52	0	1.4	2.9	11.9 ± 0.4
A31α	Side-chain	3.4 ± 0.7	40	0	1.7	2.9	8.0 ± 0.7
S32σ	HB9	4.2 ± 0.5	0	44	2.9	1.6	8.7 ± 0.5
Q33θ	HB10	13 ± 0.3	23	0	2.2	2.9	18.1 ± 0.3
W34ω	HB11	2.1 ± 0.3	35	1	1.9	2.9	6.9 ± 0.3

[a] Based on the rationale stated above, the difference between the water/octanol transfer energies of an amide carbonyl (5.7 kJ/mol) and an ester carbonyl (2.8 kJ/mol) is 2.9 kJ/mol. [b] Based on Table 2, the difference between the water/octanol transfer energies of an amide −NH− (5.7 kJ/mol) and an ester −O− (2.8 kJ/mol) is 2.9 kJ/mol. [c] The correction assumes fractional transfer energy. As a result, the correction from −CO− = (100 − −CO− solvent accessible %)(2.9)/100 and the correction from −NH− = (100 − −NH− solvent accessible %)(2.9)/100. [d] Observed $\Delta\Delta G_f$ = adjusted $\Delta\Delta G_f$ − (ΔG correction from −CO− + ΔG correction from −NH−). [e] Values in the table are rounded to the nearest 0.1 kJ/mol. For sums, rounding is performed after summation.

(57) Bergasa-Caceres, F.; Rabitz, H. *J. Phys. Chem. B* **2001**, *105*, 2874−2880.
(58) (a) Shortle, D.; Simons, K. T.; Baker, D. *Proc. Natl. Acad. Sci. U.S.A.* **1998**, *95*, 11158−11162. (b) Wrabl, J. O.; Shortle, D. *Protein Sci.* **1996**, *5*, 2343−2352. (c) Dill, K. A.; Shortle, D. *Annu. Rev. Biochem.* **1991**, *60*, 795−825. (d) Botstein, D.; Shortle, D. *Science* **1985**, *229*, 1193−1201.
(59) Meylan, W. M.; Howard, P. H. *Perspect. Drug Discovery Des.* **2000**, *19*, 67−84.
(60) Abraham, D. J.; Kellogg, G. E.; Holt, J. M.; Ackers, G. K. *J. Mol. Biol.* **1997**, *272*, 613−632.
(61) Chan, H. S.; Dill, K. A. *Annu. Rev. Biophys. Biomol. Struct.* **1997**, *26*, 425−459.
(62) (a) Leo, A. J.; Hoekman, D. *Perspect. Drug Discovery Des.* **2000**, *18*, 19−38. (b)Leo, A.; Hansch, C.; Elkins, D. *Chem. Rev.* **1971**, *71*, 525−616.
(63) Kellogg, G. E.; Abraham, D. J. *Eur. J. Med. Chem.* **2000**, *35*, 651−661.
(64) Mannhold, R.; Rekker, R. F. *Perspect. Drug Discovery Des.* **2000**, *18*, 1−18.
(65) Petrauskas, A. A.; Kolovanov, E. A. *Perspect. Drug Discovery Des.* **2000**, *19*, 99−116.
(66) Honig, B.; Yang, A. S. *Adv. Protein Chem.* **1995**, *46*, 27−58.

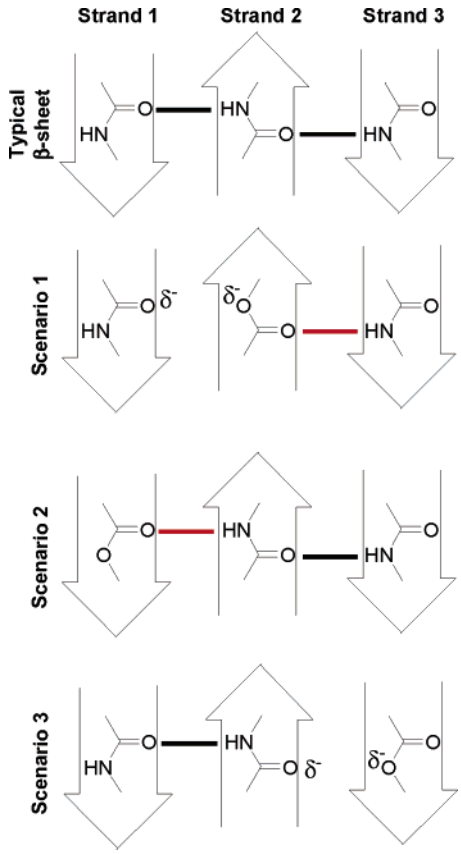


Figure 5. Three scenarios for the introduction of potentially unfavorable electrostatic interactions utilizing an A-to-E mutation. Typical hydrogen bonds are indicated by a black line, while weakened hydrogen bonds are indicated by a red line. Potentially unfavorable electrostatic interactions are formed between the partially charged oxygens of the amide carbonyl and the ester in scenarios 1 and 3.

A-to-E mutations in some contexts can lead to unfavorable electrostatic interactions that should be corrected for to extract hydrogen bond strengths from $\Delta\Delta G_f$ values. There are three possible scenarios where unfavorable electrostatic interactions could complicate A-to-E effects on backbone hydrogen bonding in a β-sheet (Figure 5). First, an A-to-E substitution in the center strand of a three-stranded β-sheet (scenario 1, Figure 5) eliminates the NH hydrogen bond donor and creates an unfavorable electrostatic interaction between the ester O and the amide carbonyl on the exterior strand (strand 1). Of course, minor structural alterations in strand 1 could attenuate these unfavorable electrostatic interactions between the electron-rich centers indicated in Figure 5. Second, an A-to-E mutation in an exterior strand involving an NH that is solvent exposed (scenario 2, Figure 5) weakens the hydrogen bond that the ester carbonyl makes with the amide NH in strand 2 without introducing unfavorable electrostatic interactions. Finally, A-to-E mutations in the exterior strand that eliminate an NH that is hydrogen bonded to an amide carbonyl in strand 2 (scenario 3) introduce an unfavorable electrostatic interaction between the ester oxygen and the amide carbonyl oxygen. This repulsion could also be attenuated by slight rearrangements in structure that increase the separation between the partial negative charges.

The energetic costs of the unfavorable electrostatic interactions between the oxygen atoms in scenarios 1 and 3 can be approximated by Coulomb's law, by accounting for the partial charge on each oxygen atom, the distance between the two atoms, and the dielectric constant of the surrounding medium. While there are many methods for calculating partial charges on ester oxygens,[67] an ab initio calculated set of values for these partial charges calculated by ChemBats3D (CambridgeSoft, Inc.) using a Hückel model are -0.8 and -0.1 for an amide carbonyl oxygen and an ester oxygen, respectively. The distance between these oxygen atoms is approximately 3 Å. In scenario 3, where the functional groups are solvent exposed, the dielectric constant of water (80) is utilized, whereas, in scenario 1, the dielectric constant must be estimated. While the absolute dielectric constant of protein can vary, a value of 16 is chosen here for simplicity.[68] Based on these limits, the electrostatic repulsion between the two oxygen atoms is destabilizing by approximately 0.4 and 2.9 kJ/mol in scenarios 3 and 1, respectively (assuming no motion to separate the oxygens). An experimental approach to approximate this unfavorable electrostatic interaction was carried out by replacing $-O-$ with $-CH_2-$ in the carboxylate ester of vancomycin's ligand and the phosphonate ester of thermolysin's ligand.[69] The differential binding energy assigned to destabilizing lone pair/lone pair interaction was 6.7−11.3 kJ/mol.[69a] These studies likely overestimate the destabilization energy because the solvation/desolvation energies are unaccounted for.

First-order approximations of hydrogen bond strengths, based on $\Delta\Delta G_f$ values adjusted for both electrostatic and solvation contributions, can be found in Table 3 and Figure 6. Clearly, more accurately adjusted $\Delta\Delta G_f$ values could be calculated with

Table 3. Approximation of Individual Hydrogen Bond Strengths[b]

mutation	hydrogen bond	observed $\Delta\Delta G_f$ (kJ/mol)	electrostatic correction (kJ/mol)	adjusted $\Delta\Delta G_f$[a] (kJ/mol)
L7λ	N/A	0	N/A	0
W11ω	HB1	4.6 ± 0.4	0.4	9.5 ± 0.4
E12ϵ	HB2	6.3 ± 0.5	0.4	10.7 ± 0.5
K13κ	HB3	2.9 ± 0.5	N/A	7.7 ± 0.5
R14ρ	HB4	16.4 ± 0.4	2.9	17.7 ± 0.4
M15μ	HB5	1.7 ± 0.4	N/A	6.3 ± 0.4
S16σ	HB6	4.6 ± 0.4	0.4	8.8 ± 0.4
R17ρ	HB7	1.3 ± 0.3	N/A	5.9 ± 0.3
S19σ	HB7	2.5 ± 0.4	0.4	6.5 ± 0.4
V22ϖ	HB6	2.5 ± 0.5	N/A	7.4 ± 0.5
Y23ψ	HB5	9.2 ± 0.4	2.9	11.5 ± 0.4
Y24ψ	HB4, HB11	4.6 ± 0.3	0.4	10.0 ± 0.3
F25ϕ	HB3, HB10	17.6 ± 0.4	2.9	20.5 ± 0.4
N26ν	HB2, HB9	20.2 ± 0.5	2.9	23.1 ± 0.5
H27η	HB8	3.4 ± 0.4	N/A	9.2 ± 0.4
N30ν	HB8	7.6 ± 0.4	0.4	11.5 ± 0.4
A31α	Side-chain	3.4 ± 0.7	0.4	7.6 ± 0.7
S32σ	HB9	4.2 ± 0.5	N/A	8.7 ± 0.5
Q33θ	HB10	13 ± 0.3	2.9	15.2 ± 0.3
W34ω	HB11	2.1 ± 0.3	N/A	6.9 ± 0.3

[a] These values represent a first-order approximation of the hydrogen bond strength utilizing the solvation and electrostatic corrections described in the text. [b] Summary of the stability of A-to-E variants relative to the wt PIN WW domain ($\Delta\Delta G_f$), the electrostatic correction to the $\Delta\Delta G_f$ value using the approach described in the text, and the $\Delta\Delta G_f$ value corrected for both electrostatic and solvation contributions. The fully corrected $\Delta\Delta G_f$ value is a first-order approximation of the hydrogen bond strengths. Values in the table are rounded to the nearest 0.1 kJ/mol. For sums, rounding is performed after summation.

(67) Li, J. B.; Zhu, T. H.; Cramer, C. J.; Truhlar, D. G. *J. Phys. Chem. A* **1998**, *102*, 1820−1831.
(68) (a) Schutz, C. N.; Warshel, A. *Proteins* **2001**, *44*, 400−417. (b) Bergveld, P. *Biosens. Bioelectron.* **1991**, *6*, 55−72.
(69) (a) McComas, C. C.; Crowley, B. M.; Boger, D. L. *J. Am. Chem. Soc.* **2003**, *125*, 9314−9315. (b) Morgan, B. P.; Scholtz, J. M.; Ballinger, M. D.; Zipkin, I. D.; Bartlett, P. A. *J. Am. Chem. Soc.* **1991**, *113*, 297−307.

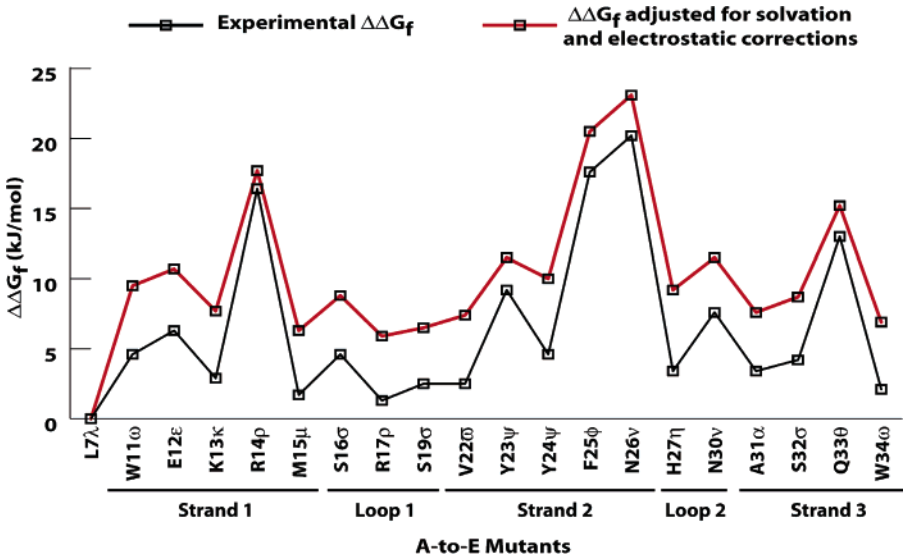


Figure 6. A plot comparing experimental $\Delta\Delta G_f$'s and $\Delta\Delta G_f$'s corrected for solvation and electrostatic contributions. Experimental errors for the data in this plot are reported in Table 3.

Table 4. Statistical Summary of the of A-to-E Mutation Data

	without correction (kJ/mol)		corrected for solvation (kJ/mol)		corrected for both solvation and electrostatic (kJ/mol)	
effects	average	SD (σ)	average	SD (σ)	average	SD (σ)
removal of donor	7.6	4.6	12.2	4.6	11.0	3.6
weakening of acceptor	2.5	0.8	7.5	1.2	7.5	1.2
affecting both donor and acceptor	14.3	8.4	19.9	8.4	17.9	6.9

more knowledge about solvation energies specific for each amide bond and more information regarding the local dielectric constants and structural alterations for the electrostatic corrections discussed above. While the tabulated raw data allows us and others to further refine these corrections, it is clear that the trends in the data are the same whether or not the data are corrected (see below).

Thermodynamic Analysis. Three observations regarding the contributions of backbone hydrogen bonding to PIN WW domain stability are striking. First, the elimination of selected backbone hydrogen bonds shifts the equilibrium toward the unfolded state, demonstrating their important energetic contributions to the stability of the native state. Second, A-to-E mutations where a hydrogen bond donor is removed are more destabilizing than mutants where a hydrogen bond acceptor is weakened. Third, the energetic contributions of backbone hydrogen bonds to the thermodynamic stability of the PIN WW domain are strongly context dependent, as evidenced by standard deviations of the $\Delta\Delta G_f$ values that are much larger (on the order of several kJ) than the experimental error (on the order of tenths of kJ) among mutants where (A) a hydrogen bond donor is removed, (B) a hydrogen bond acceptor is weakened, and (C) both hydrogen bond donor and acceptor are affected (Table 4). These findings emphasize that backbone hydrogen bonding can have highly variable, context dependent effects on the stability of a β-sheet protein, and in extreme cases (R14ρ, F25ϕ, N26ν, and Q33θ), the destabilization resulting from A-to-E substitutions can be sufficient to prevent substantial population of the native state.

The most destabilizing A-to-E variants are those that interrupt a backbone hydrogen bond located in the middle of a β-strand, usually a hydrogen bond buried by hydrophobic core side chains.[70,71] In contrast, less destabilizing A-to-E variants eliminate hydrogen bonds in or near the two loops, at the beginning of strand 1, and at the end of strand 3 (Figure 7). All A-to-E variants exhibiting average to below average $\Delta\Delta G_f$ values were solvent accessible (Figure 7). Hydrogen bonds, being primarily electrostatic interactions, are stronger in a low dielectric medium, explaining why hydrogen bonds enveloped by hydrophobic side chains contribute disproportionately to the thermodynamic stability of the PIN WW domain.

The extent of destabilization of an A-to-E mutation may also relate to the ability of the structure to rearrange. This point can be illustrated by the extreme context dependency of mutations within strand 2. For example, Y23ψ removes a hydrogen bond donor, whereas Y24ψ both removes a hydrogen bond donor and weakens a hydrogen bond acceptor. Yet, Y23ψ is a more destabilizing mutation than Y24ψ. A likely explanation for this finding is that M15, the hydrogen bonding partner of Y23ψ, is in a more constricted location than W34, the hydrogen bonding partner of Y24ψ. The hydrogen bond between M15 and Y23 occurs in a region where there are several other hydrogen bonds. Therefore, conformational changes that might allow M15 to find another hydrogen bonding partner would likely disrupt many interactions. In contrast, W34 is located at the end of strand 3 and likely can form new hydrogen bonds with water without severely distorting the structure of the WW domain.

The points made above regarding the context dependency of the contributions of hydrogen bonds to the stability of the native state are equally valid, whether the uncorrected or corrected data are being considered. This suggests that the destabilization energy due to an A-to-E mutation is an adequate indicator of the strength of a backbone hydrogen bond, even though this energy is not strictly equal to the energy of hydrogen bond formation. The corrections applied for solvation and electrostatics do not change the trends in the raw data, and even though

(70) Loladze, V. V.; Ermolenko, D. N.; Makhatadze, G. I. *J. Mol. Biol.* **2002**, *320*, 343−357.
(71) Gromiha, M. M.; Oobatake, M.; Kono, H.; Uedaira, H.; Sarai, A. *Protein Eng.* **1999**, *12*, 549−555.

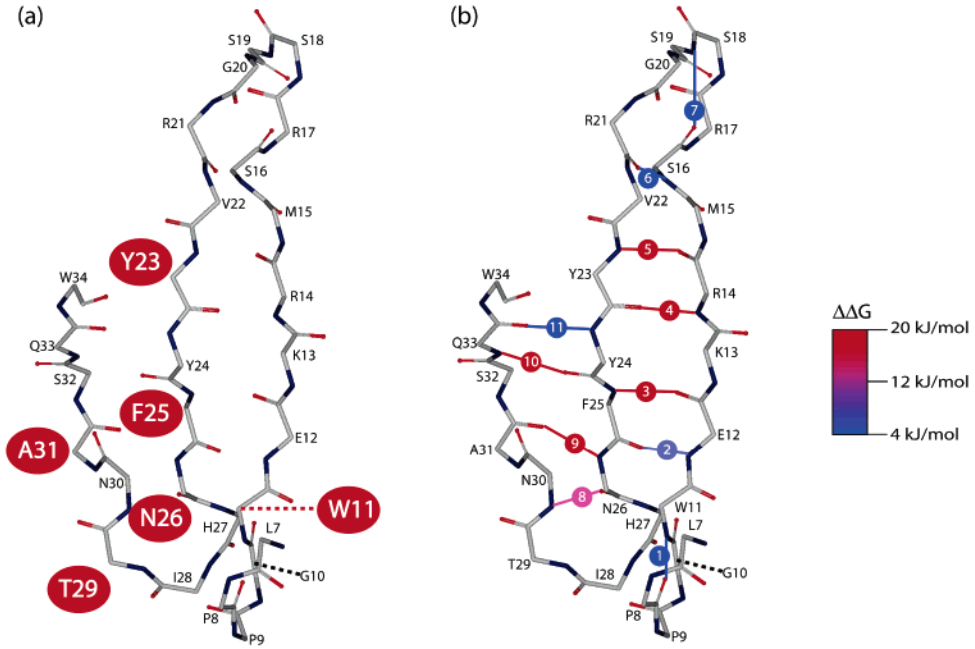


Figure 7. (a) Traditional side chain mutagenesis data reveals the side chains that contribute most to thermodynamic stability as indicated by the single letter codes in red ovals. (b) Relative stability of the A-to-E WW domain variants. The degree of destabilization caused by the A-to-E mutation is indicated by the color of the hydrogen bonds that are perturbed. Red hydrogen bonds are most destabilized by the A-to-E substitution (>15 kJ/mol). The pink hydrogen bond is destabilized by 12−15 kJ/mol. The remainder of the hydrogen bonds depicted in blue are only modestly destabilized (<10 kJ/mol).

they cannot be neglected for determining absolute hydrogen bond energies, they do not dramatically change the reported numbers.

Conclusions

This study used A-to-E mutations to perturb individual backbone hydrogen bonds to discern their contributions to PIN WW domain folding thermodynamics. While the $\Delta\Delta G_f$ values had to be corrected for solvation effects and unfavorable electrostatic interactions to estimate hydrogen bond strengths, such corrections, based on the criteria and approximations presented in this paper, do not substantially change the overall picture suggested by the raw data. The contributions of backbone hydrogen bonds to β-sheet stability are highly context dependent, with a few hydrogen bonds potentially contributing as much as 25 kJ/mol, and others contributing almost nothing. The best predictor of whether a hydrogen bond is highly stabilizing is whether it is buried.[35a] These conclusions suggest that the formation of buried hydrogen bonds is an important driving force for protein folding, a conclusion worthy of further scrutiny utilizing A-to-E scans in structurally diverse proteins where the thermodynamic data are analyzed by correction methods of increasing sophistication.

Experimental Section

Materials and Methods. Boc-amino acids were obtained from Midwest Biotech (Fishers, IN). Boc-Gly-PAM resin was purchased from Applied Biosystems (Foster City, CA). 2-(1*H*-Benzotriazol-1-yl)-1,1,3,3,-tetramethyluronium hexafluorophosphate (HBTU) was obtained from QBiogene (Carlsbad, CA). 1-Hydroxybenzotriazole (HOBt) was purchased from Novabiochem (San Diego, CA). Diisopropylethylamine (DIEA) and *p*-cresol were purchased from Aldrich Chemical Co.. Trifluoroacetic acid was purchased from Halocarbon (River Edge, NJ). All reagent-grade solvents (dichloromethane (DCM), *N*,*N*-dimethylformamide (DMF), ethyl ether, and acetonitrile) were purchased from Fisher. Hydrogen fluoride was purchased from Matheson Gas (Rancho Cucamonga, CA). The following α-hydroxy acids were purchased from Fisher: L-lactic acid (α), L-phenyllactic acid (ϕ), glycolic acid (γ), (*S*)-(+)-2-hydroxy-4-methylpentanoic acid (λ), and (*S*)-(+)-2-hydroxy-2-methylbutyric acid (ϖ). Other α-hydroxy acids with appropriate side chain protecting groups for Boc synthesis were synthesized using previously published methods.[52] Common inorganic compounds and ultrapure guanidine hydrochloride (GuHCl) were purchased from Fisher. Trimethylamine *N*-oxide was purchased from Sigma (St. Louis, MO).

Reversed-phase HPLC (RP-HPLC) was performed on a Waters HPLC system with model 600 pumps and model 486 or 2487 detectors with 214 and 280 nm UV detection using a Vydac C_{18} column. The flow rate for analytical HPLC was 1 mL/min, while the flow rate for preparative HPLC was 10 mL/min. The mobile phase was a gradient from 4.9% acetonitrile, 95% water, 0.1% TFA to 95% acetonitrile, 4.9% water, 0.1% TFA. The purified peptides were identified by ESI-MS performed on a Hewlett-Packard LC-MS (MSD1100).

PIN WW domain variants were subjected to size exclusion chromatography, which was accomplished using a Pharmacia Model UPC900/P-920 FPLC. The stationary phase was Superdex 30. The eluant was 20 mM sodium phosphate buffer (pH 7.0, 100 mM NaCl). The flow rate was 0.5 mL/min. The eluted peptide was collected and dialyzed against 20 mM sodium phosphate (pH 7.0). For some depsipeptide variants with labile ester bonds, samples were prepared by dissolving the lyophilized depsipeptide in deionized water and diluting with 20 mM sodium phosphate buffer (pH 7.0) to an appropriate concentration. The concentrations of all PIN WW domain variants were determined by UV absorbance ($\epsilon_{280} = 13\,940\ M^{-1}\ cm^{-1}$).

Peptide Synthesis. The wt PIN WW domain and the tetramethylrhodamine-labeled YSPTpSPS (CTD-S5) ligand were synthesized by Fmoc-based solid phase peptide synthesis as previously described.[48] Peptides incorporating α-hydroxy acids were synthesized by Boc solid phase peptide synthesis.[35]

CD Studies. CD spectra were recorded using an Aviv model 202SF circular dichroism spectrometer equipped with a cell holder with a Peltier temperature controller and a 0.2 cm path length Suprasil quartz cell (Hellma, Forest Hills, NY). Far-UV CD spectra were recorded from 200 to 250 nm at 2 and 25 °C. The wavelength step size was 0.5 nm, and the signal averaging time was 2 s at each wavelength step. The

peptide sample was dissolved in 20 mM sodium phosphate (pH 7.0). Thermal denaturation was monitored at 227 nm. The temperature range utilized was from 2 °C to 98 °C with a 2 °C step size and a 90 s equilibration time. The signal was averaged for 30 s at each temperature. After the highest temperature was reached, the sample was cooled to 25 °C and another full CD spectrum was measured to ensure that folding was reversible. The fraction of unfolded PIN WW domain (f_u) was determined using the baseline extrapolation method. The T_m value was determined from the fraction unfolding curve (the T_m value was taken to be the temperature at which $f_u = 0.5$) as previously described.[16]

Guanidinium hydrochloride (GuHCl) denaturation was accomplished using the automated titrator accessory on the Aviv model 202SF. Two solutions were prepared. The first being 5 μM protein in 20 mM sodium phosphate buffer at pH 7.0, and the second being 5 μM protein in 7 M GuHCl (20 mM sodium phosphate buffer at pH 7.0). The second solution was added to the first in steps such that the denaturant concentration of the mixture increased by 0.2 M/step while the protein concentration remained fixed (excess solution was removed after each step). The equilibration time was 10 min for each addition with constant stirring. The data were collected at 227 nm with a 30 s averaging time. The denaturation curves were analyzed assuming two-state behavior using a previously described method.[16]

Fluorescence. Fluorescence measurements were carried out on an AVIV model ATF-105 automated titrating differential/ratio spectrofluorometer. Emission spectra were recorded from 310 to 410 nm in 1 nm steps with excitation at 295 nm using a 1 cm × 0.6 cm cuvette. GuHCl denaturation experiments were conducted at 2 °C using 5 μM protein in 20 mM sodium phosphate (pH 7.0) and increasing GuHCl concentration as described above (0.2 M/step). Data were converted to ratios of the fluorescences at 342 and 355 nm (the normalized fluorescence) to eliminate the dependence of the fluorescence on concentration. The data were fit to a two-state model using a previously described method.[48] The fit yielded a relationship between fraction unfolded and the normalized fluorescence, which was used as a standard curve for the TMAO renaturation studies described below.

For A-to-E mutants that were not folded at 2 °C, renaturation was carried out in the presence of the osmolyte trimethylamine *N*-oxide (TMAO). The concentrations of TMAO were determined by a previously described method.[56] Experiments were conducted at 2 °C using 5 μM protein in 20 mM sodium phosphate (pH 7.0) and increasing TMAO concentration (0.2 M/step). The normalized fluorescence signals were compared to the standard curve described above to convert the fluorescence data to the fraction of unfolded protein. The plot of fraction unfolded vs TMAO concentration was fitted by a previously described method to obtain free energies of folding.[28,48,55−56]

^{1}H NMR Studies. The aqueous wt or depsipeptide PIN WW domain samples were analyzed by ^{1}H NMR using data collected on a Bruker AMX 600 MHz spectrometer. The spectra were acquired at 10 °C at a spectral width of 9000 Hz at an operating frequency of 600 MHz using 800 points. Spectra were referenced to the internal standard sodium 2,2-dimethyl-2-silapentane-5-sulfonate (DSS; Cambridge Isotope Laboratories). Sample concentrations of 500 μM in 20 mM sodium phosphate buffer at pH 7.0 (10% D_2O) were employed. Water suppression was achieved using the Watergate pulse sequence.[75] The data were processed using XWINNMR software version 6.0 (Bruker) using a line-broadening parameter of 5 Hz.

Ligand Binding Assay. Fluorescence anisotropies of the samples were measured on an Aviv model ATF105 automated differential/ratio spectrofluorometer. The tetramethylrhodamine-labeled YSPTpSPS (CTD-S5) ligand for the PIN WW domain was prepared as described previously.[48] Samples with varying concentrations of the PIN WW domain or one of its A-to-E mutants (0, 10, 25, 50, and 100 μM) and a constant concentration of phosphopeptide ligand (10 μM) in sodium phosphate (pH 7.0) were prepared. The fluorescence anisotropies at 571 nm of these samples was measured (4 nm bandwidth, excitation at 544 nm with a 1 nm bandwidth) and fit as a function of peptide concentration to the following equation to obtain the dissociation constant, K_d:

$$r = r_f + (r_b - r_f)((K_d + [P] + [L]) - ((K_d + [P] + [L])^2 - 4[P][L]))^{1/2}/2[P] \quad (1)$$

where r = anisotropy, r_f = anisotropy of free ligand, and r_b = anisotropy of bound ligand. [P] = concentration of the PIN WW domain, and [L] = concentration of the ligand

Analytical Ultracentrifugation. The solution molecular weights of the peptides were evaluated by sedimentation equilibrium measurements carried out with a temperature-controlled Beckman XL-I analytical ultracentrifuge equipped with an An-60 Ti rotor and a photoelectric scanner (Beckman Instruments, Palo Alto, CA). Protein samples (100 μM in 20 mM sodium phosphate, pH 7.0) were loaded in a double sector cell equipped with a 12 mm Epon centerpiece and a sapphire optical window. The reference compartment was loaded with the matching buffer. The samples were monitored at 280 nm at a rotor speed of 3000 to 40 000 rpm at 20 °C and analyzed by a nonlinear least squares approach using Origin (Microcal Software Inc., Northampton, MA). The data were fit to the Lamm equation for a single species model

$$A_r = A_0 \exp[\omega^2/2RTM(1 - (\text{v-bar})\rho)](r^2 - r_0^2) \quad (2)$$

where A_r is radial absorbance, A_0 is the baseline absorbance, ω is the rotor speed (s^{-1}), R is the gas constant (J mol^{-1} K^{-1}), T is the temperature (K), v-bar is the partial specific volume (mL g^{-1}), ρ is the density of solvent (g mL^{-1}), r is the variable radius, and r_0 is the meniscus radius.

Acknowledgment. We gratefully acknowledge financial support from the NIH (GM 51105), the Skaggs Institute of Chemical Biology, the Lita Annenberg Hazen Foundation, and a Norton B. Gilula Fellowship (S.D.). P.E.D. was supported by the NIH (GM 59380) and the Alfred P. Sloan Foundation. We thank Professor Evan T. Powers and Dr. Shu-Li You for helpful discussions, Drs. Rema Balambika and John Blankenship for assistance with depsipeptide synthesis, and Drs. John Chung and Gerard Kroon for help with NMR experiments

JA045934S

(72) Dill, K. A.; Bromberg, S.; Yue, K.; Fiebig, K. M.; Yee, D. P. *Protein Sci.* 1995, *4*, 561−602.
(73) Fernandez, A.; Kardos, J.; Goto, Y. *FEBS Lett.* **2003**, *536*, 187−192.
(74) Kentsis, A.; Sosnick, T. R. *Biochemistry* **1998**, *37*, 14613−14622.
(75) Piotto, M.; Saudek, V.; Sklenar, V. *J. Biomol. NMR* **1992**, *2*, 661−665.