Evaluating β-turn mimics as β-sheet folding nucleators

Amelia A. Fuller[a], Deguo Du[a,1], Feng Liu[b,1], Jennifer E. Davoren[c], Gira Bhabha[d], Gerard Kroon[d], David A. Case[e], H. Jane Dyson[d], Evan T. Powers[a], Peter Wipf[c], Martin Gruebele[b,2], and Jeffery W. Kelly[a,2]

[a]Departments of Chemistry and Molecular and Experimental Medicine and The Skaggs Institute for Chemical Biology, and [d]Department of Molecular Biology, The Scripps Research Institute, 10550 North Torrey Pines Road, La Jolla, CA 92037; [b]Center for Biophysics and Computational Biology and Departments of Chemistry and Physics, University of Illinois, Urbana, IL 61801; [c]Department of Chemistry, University of Pittsburgh, Pittsburgh, PA 15260; and [e]BioMaPS Institute and Department of Chemistry and Chemical Biology, Rutgers, the State University of New Jersey, 610 Taylor Road, Piscataway, NJ 08854-8087

Edited by Robert L. Baldwin, Stanford University Medical Center, Stanford, CA, and approved May 12, 2009 (received for review December 19, 2008)

β-Turns are common conformations that enable proteins to adopt globular structures, and their formation is often rate limiting for folding. β-Turn mimics, molecules that replace the *i* + 1 and *i* + 2 amino acid residues of a β-turn, are envisioned to act as folding nucleators by preorganizing the pendant polypeptide chains, thereby lowering the activation barrier for β-sheet formation. However, the crucial kinetic experiments to demonstrate that β-turn mimics can act as strong nucleators in the context of a cooperatively folding protein have not been reported. We have incorporated 6 β-turn mimics simulating varied β-turn types in place of 2 residues in an engineered β-turn 1 or β-bulge turn 1 of the Pin 1 WW domain, a three-stranded β-sheet protein. We present 2 lines of kinetic evidence that the inclusion of β-turn mimics alters β-sheet folding rates, enabling us to classify β-turn mimics into 3 categories: those that are weak nucleators but permit Pin WW folding, native-like nucleators, and strong nucleators. Strong nucleators accelerate folding relative to WW domains incorporating all α-amino acid sequences. A solution NMR structure reveals that the native Pin WW β-sheet structure is retained upon incorporating a strong *E*-olefin nucleator. These β-turn mimics can now be used to interrogate protein folding transition state structures and the 2 kinetic analyses presented can be used to assess the nucleation capacity of other β-turn mimics.

beta-sheet nucleator | kinetic assessment of turn mimics | Pin WW domain | protein folding

Loops and turns enable the formation of compact protein structures (1, 2). Numerous reports analyzing structural databases, studying peptide model systems, and perturbing known protein structures have revealed many examples where β-sheet folding is nucleated or envisioned to be nucleated by reverse turn formation (3–10). Specific amino acid sequences have a high propensity for promoting reverse turn formation; they preorganize the flanking polypeptide chains in a favorable geometry for β-sheet formation, thereby lowering the activation barrier for, or nucleating, folding.

The most prevalent turn motif is the four-residue β-turn, in which the *i* and *i* + 3 residues are hydrogen-bonded (5, 11–13). β-Turns are categorized based on the ϕ and ψ dihedral angles of the *i* + 1 and *i* + 2 amino acid residues, with the most common types being I, I′, II, and II′. The signs of the ϕ and ψ angles for type I and type II turns are opposite to those of type I′ and type II′, respectively. Type I′ and type II′ β-turns are compatible with the right-handed twist exhibited by most β-sheets. In the large WW domain family of proteins, the five-residue type I G1 β-bulge turn is the most common reverse turn type. The structure of the β-bulge turn includes an *i* + 3 Gly exhibiting positive ϕ and ψ angles and a hydrogen bond between the *i* and *i* + 4 residues (14).

Numerous mimics of β-turn substructures have been prepared over the past few decades, including those designed to act as folding nucleators (e.g., see Fig. 1) (15–26). For the purpose of this paper, β-turn mimics are defined as molecules that replace the *i* + 1 and *i* + 2 amino acid residues of a β-turn. If β-turn mimics are true folding nucleators, incorporating them in place of the *i* + 1 and *i* + 2 residues of a β-turn should accelerate β-sheet folding. One challenge to validating this hypothesis has been the identification of a cooperatively-folding host protein amenable to the essential kinetic measurements. The WW domain from the human Pin1 protein (Pin WW) (Fig. 2*A*) is well suited (27, 28) to evaluate the nucleation efficacy of β-turn mimics.

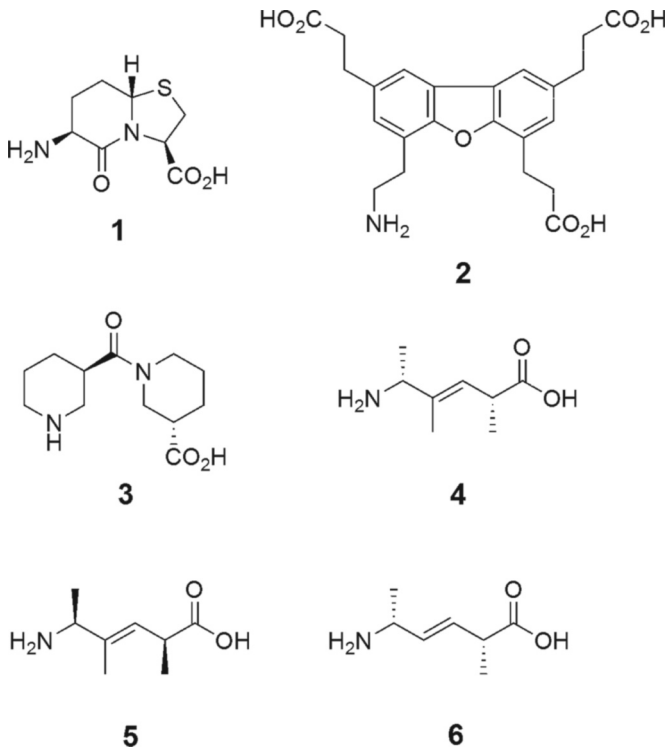

Fig. 1. β-Turn mimics studied.

Pin WW is a three-stranded β-sheet protein whose spontaneous two-state folding is reversible (4, 8, 10, 27–35). Its β-sheet exhibits

Author contributions: A.A.F., P.W., M.G., and J.W.K. designed research; A.A.F., D.D., F.L., J.E.D., and G.B. performed research; A.A.F., D.D., F.L., G.K., D.A.C., H.J.D., and E.T.P. analyzed data; and A.A.F., M.G., and J.W.K. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

Data deposition: The atomic coordinates have been deposited in the Protein Data Bank, www.pdb.org (PDB ID code 2kbu).

[1]D.D. and F.L. contributed equally to this work.

[2]To whom correspondence may be addressed. E-mail: jkelly@scripps.edu or gruebele@scs.uiuc.edu.

This article contains supporting information online at www.pnas.org/cgi/content/full/0813012106/DCSupplemental.

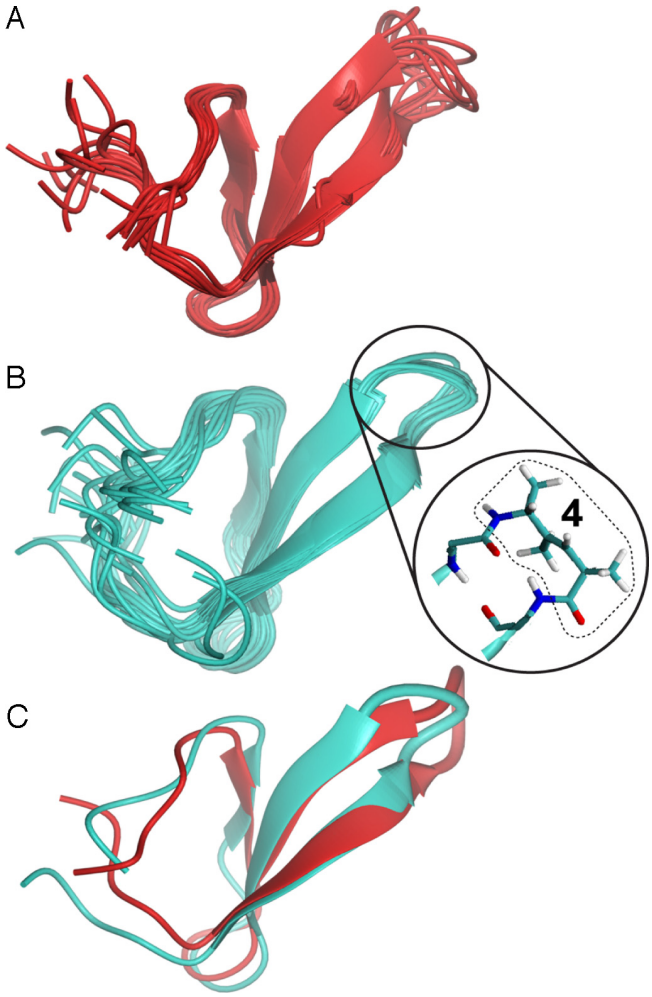

Fig. 2. Comparison of the solution structures of the Pin1 WW domain. (*A*) Superposition of 12 structures of the WT protein (36; PDB ID code 2kcf) (*B*) Superposition of 15 structures of variant **C-4** (PDB ID code 2kbu). (*C*) Superposition of a low-energy structure from each of the ensembles in (*A*) and (*B*). The inset shows the structure of dipeptide replacement **4** in loop 1.

a slight right-handed twist, as revealed by crystal and solution structures (36, 37). Extensive thermodynamic and kinetic analyses of side-chain and backbone mutants demonstrate that the folding of Pin WW is rate-limited by the formation of loop 1 (4, 10, 30, 32). Loop 1 comprises 6 residues (16-SRSSGR-21) stabilized by an internal type II β-turn (16-SRSS-19, inclusive of an i to $i + 3$ H-bond) and a Ser-16 to Arg-21 hydrogen bond. Loop 1 of Pin WW is unusually long relative to the >150 other WW domain family members, but its length is essential for ligand binding (32).

While shortened loop 1 variants lack ligand binding capability, several exhibit superior folding energetics relative to WT Pin WW (32). Remodeling loop 1 to a five-residue type I G1 β-bulge structure (16-S_ADGR-21, characterized structure, PDB ID code 2f21; 16-S_SSGR-21, putative β-bulge), a motif statistically favored in the WW family, dramatically accelerates folding and stabilizes Pin WW (32). Truncation of loop 1 to a four-residue type I′ β-turn (16-S_NGR-21, characterized structure, PDB ID code 1zcn; 16-S_S_GR-21, putative β-turn structure) or a type II′ β-turn (16-S_(D-P)GR-21, putative structure) also accelerates folding, although to a lesser extent than the type I G1 β-bulge turn type. Of the β-turn types assessed, the type II′ version affords the fastest folding WW domain with the type I′ β-turn being very close (8, 32, 38).

We report the folding thermodynamics and kinetics of Pin WW variants incorporating a β-turn mimic in lieu of the $i + 1$ and $i + 2$ β-turn 1 residues or β-bulge 1 residues. We also determined the structure of the Pin WW variant incorporating the best nucleator among the β-turn mimics evaluated. Temperature jump kinetics analyzed by 2 approaches verify that inclusion of certain β-turn mimics can accelerate Pin WW folding.

Results

Design of Loop 1-Modified Pin WW Variants. Six β-turn mimics were synthesized to explore their capacity to nucleate β-sheet formation (Fig. 1). β-Turn mimics **1**, **4**, and **6** were designed to replace the $i + 1$ and $i + 2$ residues of a type II′ β-turn and were anticipated to match the right-handed twist of Pin WW (21, 25). Compound **1** has reduced conformational freedom relative to **4** and **6**. The ϕ and ψ dihedral angle equivalents in **4** are restricted by allylic strain ($A^{1,2}$ and $A^{1,3}$) relative to **6** (39). Compound **5**, designed to replace the $i + 1$ and $i + 2$ residues of a type II β-turn (25), was expected to introduce a Pin WW twist preference mismatch. Although **1** has been incorporated into folded proteins (40, 41), its influence on folding kinetics is unknown. β-Turn mimics **4** and **5** have been included in place of loop 1 residues in Pin WW as part of a study to discern the energetic requirements for the transition from two-state folding to downhill folding (38). However, a detailed examination of the turn mimic features that influence nucleation has not been reported nor has a structure of a protein containing **4** or **5** been characterized. Dibenzofuran-based β-turn mimic **2** was designed to affect β-sheet nucleation by making favorable hydrophobic interactions with the side chains of the flanking i and $i + 3$ α-amino acid residues (18, 19, 24, 34, 42, 43). Compound **2** was previously incorporated into the six-residue loop 1 of Pin WW, but folding kinetics were not explored (43). Heterochiral dinipecotic acid, **3** (17), which does not mimic a specific turn type, was found to stabilize RNase A in a turn context, although kinetic experiments were not performed (44).

Because the five-residue β-bulge and four-residue β-turn sequences are less flexible than the native six-residue loop 1, we reasoned that these structures would offer the best context to evaluate the nucleation capacity of conformationally restricted β-turn mimics **1–6**. All protein variants containing **1–6** were compared to known, structurally characterized Pin WW domain variants **A**, **B**, and **C** (Table 1) (32, 38). Variant **A** comprises the WT Pin WW six-residue loop 1 sequence and harbors a conservative W34F point mutation that obviates a reversible self-association observed in some rapid kinetic traces (32); all of the variants studied herein have this mutation. Variant **B** features a structurally verified five-residue β-bulge reverse turn 1, leading to the fastest folding Pin WW variant studied to date (32, 38). Compounds **1–5** were inserted into the β-bulge turn in place of the $i + 1$ and $i + 2$ residues (variants **B-1**–**B-5**, respectively). Variant **C** adopts a structurally verified four-residue type-I′ β-turn 1 conformation (32, 38). Dipeptide replacements **1-6** were introduced in place of the $i + 1$ and $i + 2$ residues of the β-turn in **C** (variants **C-1**–**C-6**, respectively). Variants **D-3**–**D-5,** in which all of the loop 1 residues of Pin WW were replaced, were also assessed.

Structural Evaluation of the Pin WW Variants. All variants were evaluated by far-UV circular dichroism (CD) spectroscopy and exhibited the spectral maximum at 227 nm diagnostic of folded Pin WW (*SI Appendix*, Fig. S1), although the intensity of this maximum cannot be used to quantitatively compare the extent of folding of non-identical Pin WW sequences (32). The observation of characteristic resonances and well dispersed peaks in the ^{1}H NMR spectra of the **B** and **C** series of variants provides even stronger evidence that incorporation of **1–6** into the Pin WW backbone does not disrupt the global fold of this domain (*SI Appendix*, Fig. S2) (36).

Solution Structure of Pin WW Variant C-4. The structure of **C-4** determined by ^{1}H NMR spectroscopy confirmed that inclusion of **4**, a "strong" nucleator (*vide infra*), in place of the $i + 1$ and $i + 2$

Table 1. Nomenclature and aligned sequences of the WT Pin WW domain and the variants studied here

Variant	Turn mimic	Length of loop 1	Sequence		
			6	16–loop1–21	39
wt	–	6 residues	KLPPGWEKRM	SRSSGR	VYYFNHITNASQWERPSG
A	–	6 residues	KLPPGWEKRM	SRSSGR	VYYFNHITNASQFERPSG
B	–	5 residues	KLPPGWEKRM	S-ADGR	VYYFNHITNASQFERPSG
B-#*	**1–5**	4 residues†	KLPPGWEKRM	S-**XX**GR	VYYFNHITNASQFERPSG
C	–	4 residues	KLPPGWEKRM	S--NGR	VYYFNHITNASQFERPSG
C-#*	**1–6**	3 residues†	KLPPGWEKRM	S--**XX**R	VYYFNHITNASQFERPSG
D-#*	**3–5**	1 residue†	KLPPGWEKRM	--**XX**--	VYYFNHITNASQFERPSG

*"#" represents the compound number (Fig. 1) of the turn mimic used in variants **B-#**, **C-#**, and **D-#** and "**XX**" represents the position of the specified turn mimic in the amino acid sequence.
†Counting the β-turn mimic as 1 residue.

residues of the type I′ β-turn in **C** does not disrupt the native Pin WW fold (Fig. 2*B*; *SI Appendix*, Table S1). The highly dispersed ^{1}H NMR resonances of **C-4** at natural abundance enabled assignment of all protons by examining COSY, TOCSY, and NOESY spectra. The assignments agree well with those obtained from a previous solution structure of Pin WW (36) with the exception of the loop 1 residues, as expected. Assignments of the protons in β-turn mimic **4** in the context of **C-4** were made by examining the connectivities outlined in *SI Appendix*, Fig. S3. The β-sheet structure of **C-4** superimposes with that of Pin WW (Fig. 2*C*). Notably, the β-turn region of **C-4** is less structurally diverse than the corresponding loop 1 region of Pin WW, indicating that the structure of **4** enables formation of the correct protein topology (Fig. 2). The conformation adopted by **4** is closest to a type II′ β-turn structure, with $\phi = 66° \pm 8°$ and $\psi = -113° \pm 6°$ for the equivalent of the $i + 1$ position and $\phi = -79° \pm 5°$ and $\psi = -2° \pm 5°$ for the equivalent of the $i + 2$ position ($\pm$ indicates standard deviation for 15 structures, compare to angles of 60°, −120°, −80°, and 0°, respectively, for the canonical type II′ β-turn).

Stability of the Pin WW Variants. The thermodynamic stability of each Pin WW variant was evaluated by thermal and guanidine hydrochloride (GdnHCl) denaturation studies and compared to variants **A**, **B**, and **C** (*SI Appendix*, Fig. S4) (32). Two different thermodynamic fitting methods yielded consistent midpoints of the thermal transition (T_M) and consistent folding equilibrium constants (K) (*vide infra*, *SI Appendix*, Table S2). The free energy of folding (ΔG_f) was obtained from fitting the GdnHCl denaturation curves to a two-state model (Table 2), applicable under these conditions (2 °C, 20 mM sodium phosphate buffer, pH 7.0). Chaotrope denaturation curves obtained by monitoring changes in either far-UV CD or fluorescence spectra were the same within error. Higher T_M values correlated with more negative chaotrope-derived ΔG_f values (*SI Appendix*, Fig. S5*a*).

Variants **B-1**–**B-5** differ in stability from variant **A** by less than ±0.6 kcal/mol (Table 2, $\Delta\Delta G_f$ column). Notably, the all α-amino acid sequence of **B** exhibits greater stability than any of the β-turn mimic variants (Table 2, ΔG_f column). In contrast, replacement of all 6 loop 1 residues in Pin WW with **3**, **4**, or **5** (variants **D-3**–**D-5**)

Table 2. Summary of thermodynamic and kinetic parameters of loop 1-modified Pin WW domains

Variant	T_M (°C)	T_{max}*	τ_f (T_M),† μs	τ_f (T_{max}),‡ μs	$\Delta G^{\ddagger}_f$,§ kT_0	ΔG_f,¶ kcal/mol	$\Delta\Delta G_f$,‖ kcal/mol
A**	56	43	121	67	3.5	−2.2 ± 0.10	–
B**	75	72	13	11††	2.2	−3.8 ± 0.11	−1.6
B-1	58	55	58	55	3.6	−2.7 ± 0.12	−0.5
B-2	56	41	158	112	4.0	−1.6 ± 0.11	+0.6
B-3	47	32	150	125	3.8	−2.0 ± 0.07	+0.2
B-4**	53	50	122	105	4.2	−2.3 ± 0.16	−0.1
B-5**	54	40	157	113††	4.0	−1.9 ± 0.09	+0.3
C**	66	52	22	16	2.2	−3.4 ± 0.08	−1.2
C-1	68	61	18	16	2.4	−2.8 ± 0.31	−0.6
C-2	61	51	78	61	3.6	−2.1 ± 0.41	+0.1
C-3	56	63	59	51	3.8	−2.1 ± 0.04	+0.1
C-4**	71	71	15	14	2.5	−3.7 ± 0.57	−1.5
C-5**	51	36	192	160††	4.2	−2.2 ± 0.18	0.0
C-6	69	56	26	22	2.6	−3.6 ± 0.15	−1.4
D-3	36	22	219	180	3.8	n.d	n.d
D-4**	21	13	290	221††	3.7	n.d.	n.d
D-5**	37	28	385	212	4.2	n.d.	n.d

*T_{max} is the temperature of the maximum folding rate.
†τ_f (T_M) is the activated folding time at T_M. τ_f (T_M) = k_f (T_M)$^{-1}$.
‡τ_f (T_{max}) is the activated folding time at T_{max}. τ_f (T_{max}) = k_f (T_{max})$^{-1}$.
§$\Delta G^{\ddagger}_f$ are minimum activated free energies for folding, calculated with the τ_f (T_{max}) assuming a prefactor of 2 μs at 59 °C with an inverse bulk solvent viscosity dependence. The reference temperature T_0 = 293 K. The uncertainty of $\Delta G^{\ddagger}_f$ is estimated to be less than 0.2 kT_0.
¶Values at 2 °C as extracted from chaotrope denaturation curves, n.d. is not determined.
‖$\Delta\Delta G_f = \Delta G_f$ (variant i where $i \neq$ **A**) − ΔG_f (variant **A**).
**T_M and kinetic parameters from ref. 38.
††The highest measured activated folding time is tabulated when the folding rate does not have the maximum value within the measured temperature range.

substantially destabilizes the protein relative to **A** (Table 2, compare T_M values).

Replacement of the $i + 1$ and $i + 2$ residues of the type I′ β-turn in **C** affords a better context to compare the influences of **1–6** on Pin WW stability (32). Three variants, **C-1**, **C-4**, and **C-6**, are more stable than **A**, and **C-4** and **C-6** are comparable in stability to **C** (Table 2, ΔG_f column). As **1**, **4**, and **6** emulate the type II′ $i + 1$ and $i + 2$ dihedral angles, their inclusion likely introduces a favorable structural match to the right-handed β-sheet twist in Pin WW, consistent with the solution structure of **C-4** (Fig. 2*B*). Despite the additional conformational freedom of **6** relative to **4**, the stabilities of **C-6** and **C-4** are within error. Inclusion of **5,** a type II β-turn mimic, does not stabilize the protein relative to **A**, consistent with the expected twist mismatch.

Folding Kinetics of the Pin WW Variants. The activated folding and unfolding rates of all Pin WW variants were measured using the ns-resolution laser T-jump relaxation method (32, 38) and are plotted as a function of temperature (Fig. 3*A* and *SI Appendix* Fig. S5*c*). Briefly, the relaxation of each variant after a 5–12 °C temperature jump was recorded over a range of temperatures (*SI Appendix*, Fig. S5*c*). The relaxation at each temperature was normalized and then fitted by least squares to the model

$$S(t) = (1 - A_m)e^{-k_a t} + A_m e^{-(k_m t)^\beta}$$

where $k_m > k_a$. This model accounts for all important scenarios: (*i*) for apparent two-state folding, where $A_m = 0$, and $k_a(T) = k_f(T)\ K(T)/(1+K(T))$ relates the observed activated rate coefficient to the folding rate coefficient and equilibrium constant; (*ii*) for incipient downhill folding, where $A_m \neq 0$, $\beta \leq 1$, and $k_a(T)$ approaches the two-state folding rate when $T < T_M$ (see ref. 45); (*iii*) for rapid three-state folding with an intermediate, where $A_m \neq 0$, $\beta = 1$. Whenever 2 exponential phases were observed, we conservatively used the slow rate coefficient $k_a(T)$ to extract a folding rate $k_f(T)$, using the two-state assumption. As long as the resulting folding rates are extracted at $T<T_M$, this procedure guarantees for all 3 scenarios that we do not overestimate the folding rate, preventing categorization of a nucleator as "strong" when it is not.

The equilibrium constant (K) at T_{max} (where T_{max} is the temperature of maximum folding rate extracted from curve fitting the Arrhenius plot of k_f) was determined by fitting the thermal denaturation data using 2 models. The first model incorporates a temperature dependence based on the folding heat capacity change ΔC_p: $\Delta G_f = \Delta H_m(1-T/T_M) + \Delta C_p((T-T_M)-T\ln(T/T_M))$. The second model was a simple linear model: $\Delta G_f = \Delta G_1(T-T_M)$, wherein ΔC_p was excluded as an explicit variable, providing an estimate for the upper limit of the error in K. The T_M values determined by the 2 approaches differ by less than 4 °C (*SI Appendix*, Table S2). The resulting equilibrium constants at T_{max} lie in a small region near unity (0.5–10.5), and no large errors in K are expected based on the 2 analyses (more than a factor of 2, or 0.35 kcal/mole in free energy). Therefore, we estimate the error in $\tau_f = 1/k_f$ to be less than 30%.

The minimum activated folding times, $\tau_f(T_{max}) = 1/k_f\ (T_{max})$, are of greatest interest because they relate directly to the minimum nucleation barrier. The small error in K discussed above does not have a strong effect on the determination of the folding time $\tau_f(T_{max})$ because T_{max} is never further than 15 °C from T_M. For the 3 variants that do not reach T_{max} in the temperature range covered (Fig. 3*A*), we conservatively report the largest observed k_f, not an extrapolated k_f. Values for $\tau_f(T_{max})$ combined with $1/\eta(T_{max})$ solvent viscosity scaling of the prefactor were used to calculate the folding activation free energies ($\Delta G^{\ddagger}_f$, Table 2). Overall, the τ_f of variants at both T_M and at T_{max} correlates with their thermodynamic stability (T_M) (Table 2, *SI Appendix*, Figs. S6 *a* and *b*, respectively); more stable proteins fold faster, also observed in all α-amino acid remodeled loop 1 Pin WW domains (8, 32, 38).

Two distinct kinetic criteria were applied to classify the nucleation capabilities of **1–6** in each of the Pin WW loop 1 lengths as weak, native-like, or strong. Importantly, both analyses classified nucleators **1–6** identically. In the first approach, the $\Delta G^{\ddagger}_f$ of each loop 1-modified variant was compared to the $\Delta G^{\ddagger}_f$ of **A**, which comprises the native loop 1. Although it could be argued that folding kinetics of β-turn mimic containing variants should instead be compared to the nucleation barriers for **B** or **C**, the best nucleators exhibit a $\Delta G^{\ddagger}_f$ within 0.5 kT_0 of the $\Delta G^{\ddagger}_f$ values of **B** and **C** (Fig. 3*B*), diminishing the significance of such a debate. When inserted in place of the $i + 1$ and $i + 2$ residues within the β-bulge turn, mimics **2**, **4**, and **5** exhibit a $\Delta G^{\ddagger}_f$ at least 0.5 kT_0 greater than that of variant **A** (Fig. 3B); i.e., they are weak nucleators. In the β-bulge context, turn mimics **1** and **3** are native-like nucleators, since the $\Delta G^{\ddagger}_f$ values for variants **B-1** and **B-3** are within 0.5 kT_0 of the $\Delta G^{\ddagger}_f$ of variant **A**. The all α-amino acid β-bulge sequence in **B** is the only sequence that meets the criteria for a strong nucleator; its inclusion reduces the $\Delta G^{\ddagger}_f$ by more than 0.5 kT_0 relative to **A** (Fig. 3*B*).

When **1–6** are incorporated in place of the $i + 1$ and $i + 2$ residues of the type I′ β-turn in **C**, β-turn mimics **1**, **4**, and **6** are classified as strong nucleators; i.e., variants **C-1**, **C-4**, and **C-6** exhibit $\Delta G^{\ddagger}_f$ values at least 0.5 kT_0 less than the $\Delta G^{\ddagger}_f$ of variant **A** (Fig. 3*B*). Notably, these values are within 0.5 kT_0 of the $\Delta G^{\ddagger}_f$ of variants **B** and **C**, 2 of the fastest-folding Pin WW variants comprising all α-amino acid sequences. β-Turn mimics **2** and **3** meet the kinetic criteria for native-like nucleators. Only **5**, a mimic of a type II β-turn, is a weak nucleator, as expected from its twist mismatch. Although the error in k_f could be as much as 30% (*vide supra*), the differences used to categorize the nucleators are much greater than this error.

When all loop 1 residues are replaced with **3**, **4,** or **5** (variants **D-3–D-5**), **3** and **4** are native-like nucleators, while **5** is a weak nucleator (*SI Appendix*, Fig. S6*c*). The unfolding rates, k_u, of **D-3–D-5** are markedly faster than the unfolding rates of the longer loop 1 variants (*SI Appendix*, Fig. S5*b*), implying that the native states of **D-3–D-5** are destabilized with respect to the transition states, probably owing to structural distortion of the connected β-strands.

The second kinetic analysis to categorize nucleators more explicitly accounts for the temperature effects on folding rates. A plot of $\ln k_f(T_{max})$ versus $1{,}000/T_{max}$ for all of the variants enables weak, native-like, and strong nucleators to be grouped (Fig. 3*C*). We analyzed each protein at its fastest fitted folding rate (Fig. 3*A*), rather than at a fixed temperature, as done in a ϕ value analysis, which is an alternative method to analyze the data presented herein (9). The Pin WW variants exhibit a wide range of stabilities, however, making kinetic comparisons at a single temperature for all of the mutants challenging. As discussed above, the temperature of fastest folding generally lies below the melting temperature, allowing us to extract a conservative upper limit on the folding rate, so weak nucleators are not miscategorized as strong.

Native-like nucleators (as defined by the first analysis) afford a straight line (Fig. *3C*, solid black line, R = −0.984). The slope of this line yields an activation barrier of 6.22 kcal/mol. This experimental value is very close to the estimated value of 6.17 kcal/mol reported previously for a WT WW domain (46). The effective diffusion coefficient used in ref. 46 has an activation barrier with 2 contributions. One arises from solvent viscosity (the prefactor assumption used in our first analysis above, dashed orange line in Fig. *3C*), and only accounts for an activation barrier of 3.59 kcal/mol. The other arises from protein internal friction (an additional 2.61 kcal/mol). The sum of these 2 contributions yields the dotted blue line, which is satisfyingly close to the direct linear least squares fit of the data (solid black line). Irrespective of whether a viscosity only or a viscosity plus internal friction model is used, the

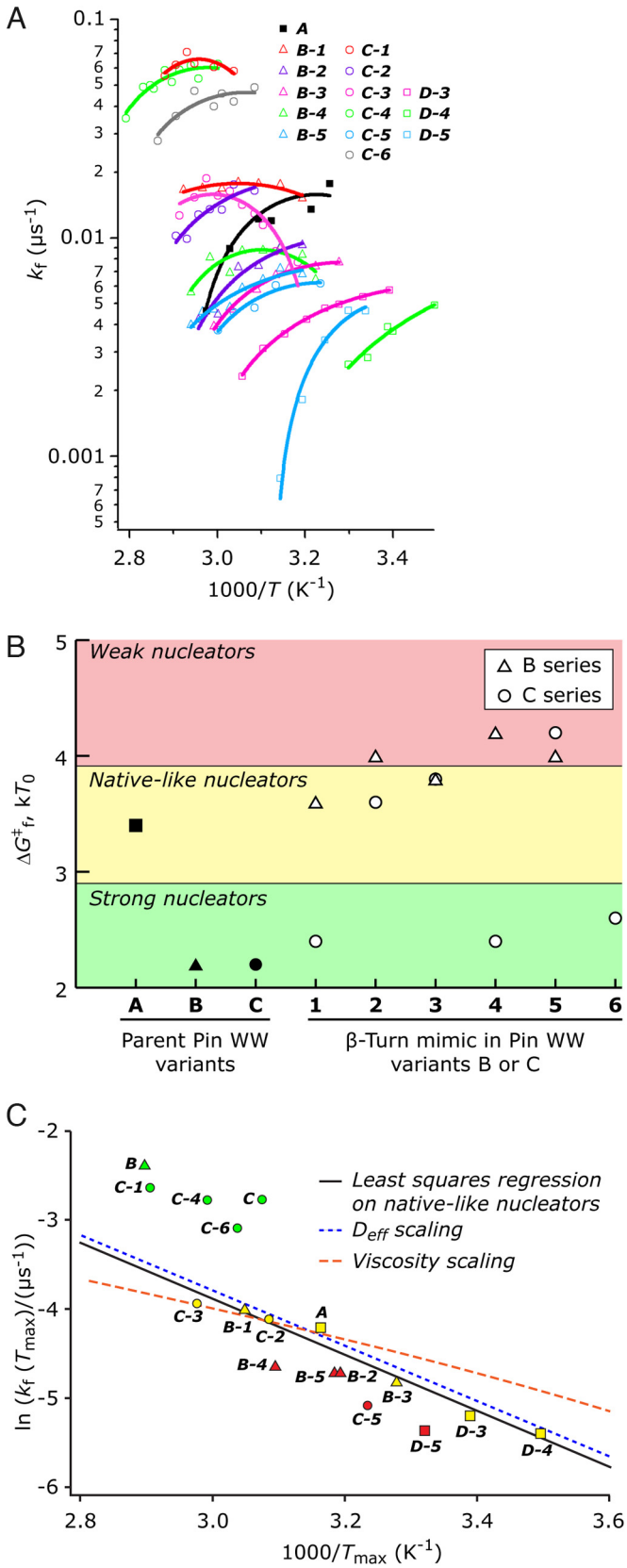

Fig. 3. (*A*) Plot of the folding rate constant (log scale) vs. 1000/*T* for Pin WW domain variants in Table 2. The curves are fits to the data of a Kramers' model for the folding rate with a temperature-dependent energy of activation, as we have reported previously (4). (*B*) Kinetic classification of nucleators **1**–**6** as weak, native-like, or strong in the context of the five-residue β-bulge (**B** series) or the four-residue type I′ β-turn (**C** series) structures based on their free energies of activation for folding (see text). (*C*) Plot of the natural logarithm of the folding rate constant at the temperature at which it is a maximum (T_{max}) vs. 1000/T_{max} for the WW domain variants. The data points are color coded according to the classification of the turn nucleator: red for weak, yellow for native-like, or green for strong. The solid black line represents the linear least squares fit to the data points corresponding to native-like nucleators in the first analysis, with the line slope representing the activation barrier. Strong nucleators exhibit lower activation barriers and are above the line (see text for meaning of orange and blue lines).

variants harboring strong nucleators (Fig. *3C*, green data points) are well above any of the lines afforded by the models or the least squares fit to the data, whereas WW domains comprising weak nucleators (Fig. *3C*, red data points) are well below the lines, so our analysis is robust.

Discussion

More than 2 decades after the first β-turn mimic was described in the literature (21), we have carried out the kinetic experiments necessary to evaluate its nucleation capacity and that of 5 other β-turn mimics in the context of a cooperatively folding host protein (16–19, 24, 25, 38). The re-engineered Pin WW domain **C**, comprising a type I′ β-turn 1, is an optimal host sequence to evaluate β-turn mimics as nucleators (32). Three β-turn mimics are categorized as strong nucleators by 2 kinetic approaches.

Our data demonstrate that merely incorporating a conformationally restricted β-turn mimic into a reverse turn to preorganize the pendant strands is not sufficient to accelerate WW domain folding relative to **A**. The excellent nucleating efficacies of **1**, **4**, and **6** and the poor nucleating capacity of **5** in the context of **C-1**, **C-4**, **C-5**, and **C-6** suggest that matching the preferred ϕ and ψ dihedral angles of the $i + 1$ and $i + 2$ residues of a type II′ β-turn as well as the right-handed twist common to most β-sheets is important. While inclusion of **1–6** in the **B** and **C** contexts affords stable, folded proteins, it is notable that the β-turn mimics influence folding rates more dramatically in the context of a type I′/II′ β-turn than in the context of a β-bulge turn, consistent with their design.

We hypothesize that the inclusion of nucleators **1**, **4**, and **6** in place of the $i + 1$ and $i + 2$ residues of a type I′ β-turn of Pin WW variant **C** (Table 1) destabilizes the unfolded ensemble by virtue of reduced conformational entropy. We also expect that they concomitantly stabilize the transition state by promoting native contacts and twist matching; both of these factors enable accelerated folding. That **C-1** is a fast folder but is less stable than **C**, **C-4,** or **C-6** is consistent with the strain introduced by the conformationally restricted β-turn mimic **1** in the folded state.

This study reveals that *E*-alkene dipeptide isosteres **4** and **6** (25) and the bicyclic turn dipeptide mimic **1** (21) are the best nucleators in the Pin WW variant **C** context, followed by the dibenzofuran-based nucleator (**2**) (43) and the heterochiral dinipecotic acid (**3**) (16, 17). The solution structure of **C-4** reveals a well-defined type II′ β-turn 1 structure, demonstrating that the nucleator matches the native β-sheet twist.

This study and related future studies should expand the utility of β-turn mimics as tools to probe protein folding transition state structures. Mimics **1**, **4,** and **6** can now be incorporated into type I′ or II′ β-turns when their formation is thought to be rate limiting for protein folding without compromising folding rates, whereas incorporation of **5** should slow folding. In addition, these compounds may find other applications, including use in backbone-modified therapeutic proteins.

Materials and Methods

Synthesis of Dipeptide Replacements and Pin WW Domain Variants. Dipeptide replacements **1–6** were synthesized with protecting groups compatible for solid phase peptide synthesis. Compounds **1** and **2** were prepared as the *N*-Fmoc derivatives according to previously described procedures (43, 47). Each enantiomer of commercially available nipecotic acid was protected as the *N*-Fmoc deriv-

ative using standard methods to access **3**. Preparative methods and characterization information for **4**, **5**, and **6** and their synthetic precursors are described in the *SI Appendix*.

Protein variants were prepared by solid phase peptide synthesis, and dipeptide replacements were included into protein backbones as described elsewhere (32, 43, 48). Variants were purified as previously described, and their identities were confirmed by mass spectrometry. See *SI Appendix* for additional details.

NMR Analysis of Pin WW Variants. NMR spectra of Pin WW domain variants were obtained on Bruker Avance 500 MHz and Bruker DRX 600 MHz instruments with TXI cryoprobes and *z* axis gradients at 288 K as solutions in 20 mM sodium phosphate buffer, pH 7.0, 90% H_2O/10% D_2O. For variant **C-4**, 2D ^{1}H TOCSY, NOESY, and COSY spectra were acquired in 100% D_2O as well as 90% H_2O/10% D_2O for backbone and side chain assignment. Solution structures of **C-4** were calculated and refined in AMBER using NOE distance restraints obtained from the ^{1}H NOESY spectrum (49). See *SI Appendix* for detailed NMR methods and *SI Appendix*, Table S1 for refinement statistics.

Equilibrium Unfolding of Pin WW Variants. Thermal stability of the proteins was assessed by monitoring the CD signal at 227 nm of a 40 μM protein solution in 20 mM sodium phosphate buffer, pH 7.0 in 2 °C intervals from 2 – 98 °C. At each temperature, the sample was equilibrated for 2 min, and data were averaged for 30 s. To assess protein stability in chaotrope, a 5–10 μM protein solution in GdnHCl was titrated into a buffered solution of equal protein concentration. The CD maximum at 227 nm and the fluorescence intensity at 341 nm were monitored as a function of [GdnHCl]. Values for T_M and ΔG_f were extracted from thermal and chaotrope denaturation curves, respectively, assuming a two-state model as previously described (50). See *SI Appendix* for additional details.

Relaxation Kinetics of the Pin WW Variants. Relaxation kinetics of Pin WW variants were measured with a laser temperature jump apparatus using jumps between 5 and 12 °C. Protein concentrations were 40–100 μM in 10 mM sodium phosphate buffer, 80% H_2O/20% D_2O, pH 7.0 (without isotope effect corrections). Trp-11 was excited by a 280 nm laser pulse every 14 ns, and fluorescence emission decays were digitized at 500 ps intervals for 500 μs of data per measurement. The relaxation kinetics were obtained from the time evolution of the tryptophan fluorescence decays using our previously reported χ-analysis sensitive to tryptophan fluorescence lifetime changes (51). The normalized curve of relaxation kinetics was fitted to a single exponential decay or to a stretched plus single exponential decay when necessary.

ACKNOWLEDGMENTS. We thank the reviewers for excellent suggestions, including the second method for evaluating candidate nucleators. We thank Colleen Fearns for her careful reading and editing of the manuscript. A.A.F. was supported by a Ruth L. Kirschstein National Research Service Award. P.W. and J.E.D. were supported by the National Institutes of Health (NIH) National Institute of General Medical Sciences Chemical Methodologies and Library Development program (GM067082). M.G. and F.L. were supported by NSF-MCB 0613643. D.A.C. was supported by the NIH (GM045811). H.J.D. was supported by the NIH (GM057374). J.W.K. was supported by the NIH (GM051105), the Skaggs Institute for Chemical Biology and the Lita Annenberg Hazen Foundation.

1. Rose GD, Gierasch LM, Smith JA (1985) Turns in peptides and proteins. *Adv Protein Chem* 37:1–109.
2. Rotondi KS, Gierasch LM (2006) Natural polypeptide scaffolds: β-sheets, β-turns, and β-hairpins. *Biopolymers* 84:13–22.
3. Grantcharova VP, Riddle DS, Santiago JV, Baker D (1998) Important role of hydrogen bonds in the structurally polarized transition state for folding of the src SH3 domain. *Nat Struct Mol Biol* 5:714–720.
4. Jäger M, Nguyen H, Crane JC, Kelly JW, Gruebele M (2001) The folding mechanism of a β-sheet: The WW domain. *J Mol Biol* 311:373–373.
5. Marcelino AMC, Gierasch LM (2008) Roles of β-turns in protein folding: From peptide models to protein engineering. *Biopolymers* 89:380–391.
6. Martinez JC, Pisabarro MT, Serrano L (1998) Obligatory steps in protein folding and the conformational diversity of the transition state. *Nat Struct Biol* 5:721–729.
7. Matouschek A, Kellis JT, Jr, Serrano L, Fersht AR (1989) Mapping the transition state and pathway of protein folding by protein engineering. *Nature* 340:122–126.
8. Nguyen H, Jäger M, Kelly JW, Gruebele M (2005) Engineering β-sheet protein toward the folding speed limit. *J Phys Chem B* 109:15182–15186.
9. Petrovich M, Jonsson AL, Ferguson N, Daggett V, Fersht AR (2006) Φ-analysis at the experimental limits: mechanism of β-hairpin formation. *J Mol Biol* 360:865–881.
10. Sharpe T, Jonsson AL, Rutherford TJ, Daggett V, Fersht AR (2007) The role of the turn in β-hairpin formation during WW domain folding. *Protein Sci* 16:2233–2239.
11. Sibanda BL, Blundell TL, Thornton JM (1989) Conformation of β-hairpins in protein structures. A systematic classification with applications to modelling by homology, electron density fitting and protein engineering. *J Mol Biol* 206:759–777.
12. Sibanda BL, Thornton JM (1985) β-Hairpin families in globular proteins. *Nature* 316:170–174.
13. Venkatachalam CM (1968) Stereochemical criteria for polypeptides and proteins. V. Conformation of a system of three linked peptide units. *Biopolymers* 6:1425–1436.
14. Richardson JS, Getzoff ED, Richardson DC (1978) The β bulge: A common small unit of nonrepetitive protein structure. *Proc Natl Acad Sci USA* 75:2574–2578.
15. Oishi S, et al. (2006) Application of tri- and tetrasubstituted alkene dipeptide mimetics to conformational studies of cyclic RGD peptides. *Tetrahedron* 62:1416–1424.
16. Chung YJ, Christianson LA, Stanger HE, Powell DR, Gellman SH (1998) A β-peptide reverse turn that promotes hairpin formation. *J Am Chem Soc* 120:10555–10556.
17. Chung YJ, et al. (2000) Stereochemical control of hairpin formation in β-peptides containing dinipecotic acid reverse turn segments. *J Am Chem Soc* 122:3995–4004.
18. Diaz H, Espina JR, Kelly JW (1992) A dibenzofuran-based amino acid designed to nucleate antiparallel β-sheet structure: Evidence for intramolecular hydrogen bond formation. *J Am Chem Soc* 114:8316–8318.
19. Diaz H, Tsang KY, Choo D, Espina JR, Kelly JW (1993) Design, synthesis, and partial characterization of water-soluble β-sheets stabilized by a dibenzofuran-based amino acid. *J Am Chem Soc* 115:3790–3791.
20. Gardner RR, Liang GB, Gellman SH (1999) β-Turn and β-hairpin mimicry with tetrasubstituted alkenes. *J Am Chem Soc* 121:1806–1816.
21. Nagai U, Sato K (1985) Synthesis of a bicyclic dipeptide with the shape of β-turn central part. *Tetrahedron Lett* 26:647–650.
22. Schneider JP, Kelly JW (1995) Templates that induce α-helical, β-sheet, and loop conformations. *Chem Rev* 95:2169–2187.
23. Souers AJ, Ellman JA (2001) β-Turn mimetic library synthesis: Scaffolds and applications. *Tetrahedron* 57:7431–7448.
24. Tsang KY, Diaz H, Graciani N, Kelly JW (1994) Hydrophobic cluster formation is necessary for dibenzofuran-based amino acids to function as β-sheet nucleators. *J Am Chem Soc* 116:3988–4005.
25. Wipf P, Henninger TC, Geib SJ (1998) Methyl- and (trifluoromethyl)alkene peptide isosteres: Synthesis and evaluation of their potential as β-turn promoters and peptide mimetics. *J Org Chem* 63:6088–6089.
26. Nesloney CN, Kelly JW (1996) Progress towards understanding β-sheet structure. *Bioorg Med Chem* 4:739–766.
27. Koepf EK, et al. (1999) Characterization of the structure and function of W → F WW domain variants: Identification of a natively unfolded protein that folds upon ligand binding. *Biochemistry* 38:14338–14351.
28. Koepf EK, Petrassi HM, Sudol M, Kelly JW (1999) WW: An isolated three-stranded antiparallel β-sheet domain that unfolds and refolds reversibly; evidence for a structured hydrophobic cluster in urea and GdnHCl and a disordered thermal unfolded state. *Protein Sci* 8:841–853.
29. Crane JC, Koepf EK, Kelly JW, Gruebele M (2000) Mapping the transition state of the WW domain β-sheet. *J Mol Biol* 298:283–292.
30. Deechongkit S, et al. (2004) Context-dependent contributions of backbone hydrogen bonding to β-sheet folding energetics. *Nature* 430:101–105.
31. Jäger M, Nguyen H, Dendle M, Gruebele M, Kelly JW (2007) Influence of hPin1 WW N-terminal domain boundaries on function, protein stability, and folding. *Protein Sci* 16:1495–1501.
32. Jäger M, et al. (2006) Structure-function-folding relationship in a WW domain. *Proc Natl Acad Sci USA* 103:10648–10653.
33. Jiang X, Kowalski J, Kelly JW (2001) Increasing protein stability using a rational approach combining sequence homology and structural alignment: Stabilizing the WW domain. *Protein Sci* 10:1454–1465.
34. Kaul R, Angeles AR, Jäger M, Powers ET, Kelly JW (2001) Incorporating β-turns and a turn mimetic out of context in loop 1 of the WW domain affords cooperatively folded β-sheets. *J Am Chem Soc* 123:5206–5212.
35. Nguyen H, Jäger M, Moretto A, Gruebele M, Kelly JW (2003) Tuning the free-energy landscape of a WW domain by temperature, mutation, and truncation. *Proc Natl Acad Sci USA* 100:3948–3953.
36. Kowalski JA, Liu K, Kelly JW (2002) NMR solution structure of the isolated Apo Pin1 WW domain: comparison to the x-ray crystal structures of Pin1. *Biopolymers* 63:111–121.
37. Ranganathan R, Lu KP, Hunter T, Noel JP (1997) Structural and functional analysis of the mitotic rotamase Pin1 suggests substrate recognition is phosphorylation dependent. *Cell* 89:875–886.
38. Liu F, et al. (2008) An experimental survey of the transition between two-state and downhill protein folding scenarios. *Proc Natl Acad Sci USA* 105:2369–2374.
39. Eliel EL, Wilen SH (1994) *Stereochemistry of Organic Compounds* (Wiley-Interscience, New York).
40. Baca M, Alewood PF, Kent SBH (1993) Structural-Engineering of the HIV-1 Protease Molecule with a β-Turn Mimic of Fixed Geometry. *Protein Sci* 2:1085–1091.
41. Viles JH, et al. (1998) Design, synthesis and structure of a zinc finger with an artificial β-turn. *J Mol Biol* 279:973–986.
42. Choo DW, Schneider JP, Graciani NR, Kelly JW (1996) Nucleated antiparallel β-sheet that folds and undergoes self-assembly: A template promoted folding strategy toward controlled molecular architectures. *Macromolecules* 29:355–366.
43. Kaul R, Deechongkit S, Kelly JW (2002) Synthesis of a negatively charged dibenzofuran-based β-turn mimetic and its incorporation into the WW miniprotein-enhanced solubility without a loss of thermodynamic stability. *J Am Chem Soc* 124:11900–11907.
44. Arnold U, et al. (2002) Protein prosthesis: A semisynthetic enzyme with a β-peptide reverse turn. *J Am Chem Soc* 124:8522–8523.
45. Yang WY, Gruebele M (2004) Folding lambda-repressor at its speed limit. *Biophys J* 87:596–608.
46. Naganathan AN, Doshi U, Munoz V (2007) Protein folding kinetics: Barrier effects in chemical and thermal denaturation experiments. *J Am Chem Soc* 129:5673–5682.
47. Bach AC, II, Markwalder JA, Ripka WC (1991) Synthesis and NMR conformational analysis of a β-turn mimic incorporated into gramicidin S. A general approach to evaluate β-turn peptidomimetics. *Int J Pept Prot Res* 38:314–323.
48. Fu Y, Gao J, Bieschke J, Dendle MA, Kelly JW (2006) Amide-to-E-olefin versus amide-to-ester backbone H-bond perturbations: Evaluating the O-O repulsion for extracting H-bond energies. *J Am Chem Soc* 128:15948–15949.
49. Case DA, et al. (2005) The Amber biomolecular simulation programs. *J Comput Chem* 26:1668–1688.
50. Pace CN, et al. (2001) Tyrosine hydrogen bonds make a large contribution to protein stability. *J Mol Biol* 312:393–404.
51. Ervin J, Sabelko J, Gruebele M (2000) Submicrosecond real-time fluorescence sampling: application to protein folding. *J Photochem Photobio B* 54:1–15.