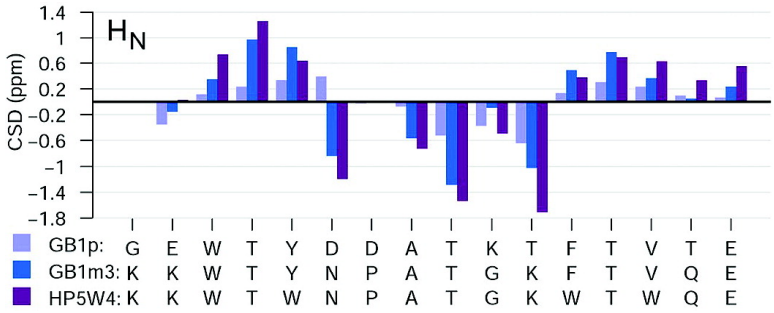
Article

Enhanced Hairpin Stability through Loop Design: The Case of the Protein G B1 Domain Hairpin

R. Matthew Fesinmeyer, F. Michael Hudson, and Niels H. Andersen

J. Am. Chem. Soc., **2004**, 126 (23), 7238-7243• DOI: 10.1021/ja0379520 • Publication Date (Web): 22 May 2004



More About This Article

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

- Supporting Information
- Links to the 7 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 05/22/2004

Enhanced Hairpin Stability through Loop Design: The Case of the Protein G B1 Domain Hairpin

R. Matthew Fesinmeyer, F. Michael Hudson, and Niels H. Andersen*

Contribution from the Department of Chemistry, University of Washington, Seattle, Washington 98195

Received August 15, 2003; E-mail: andersen@chem.washington.edu

Abstract: A mutational study of the peptide corresponding to the second hairpin of the protein G B1 domain (GB1p) provided a series of mutants with significantly increased fold stability. Mutations focused on improvement of the direction-reversing loop and the addition of favorable Coulombic interactions at the sequence termini. The loop optimization was based on a database search for residues that occur with the greatest probability in similar hairpin loops in proteins. This search suggested replacing the native DDATKT sequence with NPATGK, which resulted in a 4.5 kJ/mol stabilization of the hairpin fold. The introduction of positively charged lysines at the N-terminus provided an additional 2.4 kJ/mol of stabilization, affording a GB1p mutant that is 86 ± 3% folded at 25 °C with a melting temperature of 60 ± 2 °C. The trpzip version of this peptide, in which three of the hydrophobic core residues were mutated to tryptophan, yielded a sequence that melted at 85 °C. Throughout, fold populations and melting temperatures were derived from the mutation and temperature dependence of proton chemical shifts and were corroborated by circular dichroism (CD) melts. The study also suggests that the wild-type GB1p sequence is significantly less stable than reported in some other studies: only 30% folded in water at 25 °C.

Introduction

Developing peptide models of protein β-sheets has been an active area of research for the past decade, with numerous successes reported.[1−3] In 1993, Kobayashi[4] observed chemical shifts for the C-terminal hairpin sequence (residues 41−56) of the B1 domain of protein G (GB1p) that suggested hydrophobic interactions such as those in the protein. Blanco et al. established that the sequence formed a hairpin and estimated it to be 42% folded in water at 5 °C.[5,6]

GB1p became a reference standard for the hairpin state and has been the subject of numerous studies: NMR measures of the thermodynamics of the coil/hairpin transition,[8] T-jump studies of fold lifetime,[7] modeling of the folding pathway,[10] and computational folding simulations.[11] The stability of the hairpin state of GB1p is attributed to hydrophobic cluster formation between a valine side chain and the rings of three aromatic residues (W,Y,F). These four residues are arranged in pairs, located in opposing non-H-bonded strand positions (S ± 2 and S ± 4); the strands are linked by a six-residue loop forming a [4:4]- or [4:6]-hairpin (Figure 1), using the hairpin classification scheme of Thornton.[12]

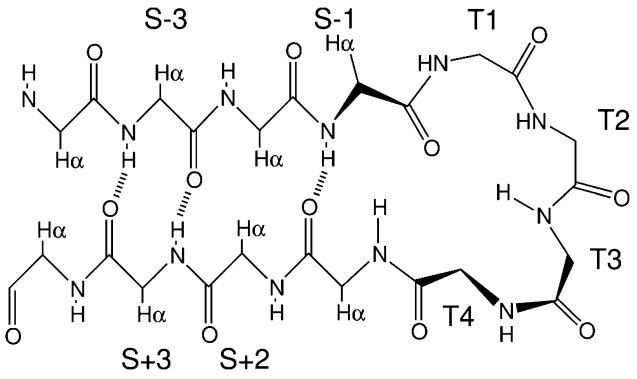


Figure 1. Strand alignment, H-bonding pattern, and nomenclature used for [4:4]- and [4:6]-hairpins: T1−T4 are turn residues, the strand positions (S) are numbered from the turn locus. The connecting loop between the hydrophobic strands consists of residues S − 1 through S + 1 (DDATKT in GB1p).

The hydrophobic strands of GB1p have been connected by shorter and longer loops,[13−15] and a partial Ala scan[16] has been

(1) Gellman, S. H. *Curr. Opin. Chem. Biol.* **1998**, *2*, 717−725.
(2) Syud, F. A.; Espinosa, J. F.; Gellman, S. H. *J. Am. Chem. Soc.* **1999**, *121*, 11577−11578.
(3) Searle, M. S. *J. Chem. Soc., Perkin Trans. 2* **2001**, 1011−1020.
(4) Kobayashi, N.; Endo, S.; Munekata, E. *Pept. Chem.* **1993**, 278−280.
(5) Blanco, F. J.; Rivas, G.; Serrano, L. *Nat. Struct. Biol.* **1994**, *1*, 584−590.
(6) Subsequent estimates of the fold population of GB1p in water (or D_2O) at 5 ± 2 °C vary from 40 to 82%.[5,7−9] The chemical shift methods used in this report yield a fold population estimate of 43 ± 8% for the H_2O data; the estimate from the D_2O data range was 55 ± 5% (vide infra and Supporting Information), with a fold population estimate of 30% at 298 K.
(7) Muñoz, V.; Thompson, P. A.; Hofrichter, J.; Eaton, W. A. *Nature* **1997**, *390*, 196−199.
(8) Honda, S.; Kobayashi, N.; Munekata, E. *J. Mol. Biol.* **2000**, *295*, 269−278.
(9) Cochran, A. G.; Skelton, N. J.; Starovasnik, M. A. *Proc. Natl. Acad. Sci. U.S.A.* **2001**, *98*, 5578−5583.
(10) Muñoz, V.; Henry, E. R.; Hofrichter, J.; Eaton, W. A. *Proc. Natl. Acad. Sci. U.S.A.* **1998**, *95*, 5872−5879.
(11) (a) Dinner, A. R.; Lazaridis, T.; Karplus, M. *Proc. Natl. Acad. Sci. U.S.A.* **1999**, *96*, 9068−9073. (b) Pande, V. S.; Rokhsar, D. S. *Proc. Natl. Acad. Sci. U.S.A.* **1999**, *96*, 9062−9067. (c) Roccatano, D.; Amadei, A.; Di Nola, A.; Berendsen, H. J. C. *Protein Sci.* **1999**, *8*, 2130−2143.
(12) Sibanda, B. L.; Thorton, J. M. *Methods Enzymol.* **1991**, *202*, 59−82.
(13) Espinosa, J. F.; Gellman, S. H. *Angew. Chem., Int. Ed.* **2000**, *39*, 2330−2333.
(14) Espinosa, J. F.; Muñoz, V.; Gellman, S. H. *J. Mol. Biol.* **2001**, *306*, 397−402.

10.1021/ja0379520 CCC: $27.50

reported. The only stabilizing loop mutations noted were the replacement of Asp at T2 with Ala, and the replacement of the entire six-residue loop with a shorter turn-favoring sequence, Val-D-Pro-Gly-Lys[17]; D-Pro is abbreviated as **p** below. For designed [2,2]-hairpins incorporating **p**G or NG at the turn locus, fold estimates have been based on Hα chemical shifts, with the L-Pro-Gly mutant serving as the unfolded control and the 100% folded reference values being those observed when the hairpin is cyclized. Cyclization is accomplished either by linking the strand ends with a D-Pro-Gly unit[2,13,15] or by adding[18] a complete loop sequence (Val-D-Pro-Gly-Orn). It is unclear how this strategy can be extended to [4,4]- or [4,6]-hairpins.

		% folded (4 ± 1 °C)	T_m (°C)
GB1p	GE**W**T**Y**DDATKT**F**T**V**TE	ca. 49%	
	R**W**Q**Y**-V**p**GK-**F**T**V**Q−NH_2	ca. 61%[15]	
	R**W**Q**Y**-VNGK-**F**T**V**Q−NH_2	ca. 33%[15]	
trpzip4	GE**W**T**W**DDATKT**W**T**W**TE−NH_2		70
trpzip2	S**W**T**W**-ENGK-**W**T**W**K−NH_2		72
trpzip3	S**W**T**W**-E**p**GK-**W**T**W**K−NH_2		78.5

Cochran[9] has demonstrated significant hairpin stabilization for the GB1p sequence when all four side chains in the cluster are indole rings, with trpzip4 displaying a melting point (T_m) of 70 °C. Here also, shorter loops, particularly if they bear a **p**G turn, can improve hairpin stability. In the present work, we show that loop sequence optimization and Coulombic interactions of the chain termini can produce GB1p mutants with T_m values as high as 62 °C while retaining the native hydrophobic cluster and loop length. The same loop and terminal mutations raise the melting temperature of trpzip4 to 85 °C.

Methods and Materials

Peptides were synthesized on an Applied Biosystem 433A synthesizer employing standard Fmoc solid-state peptide synthesis methods. Wang resin preloaded with the C-terminal amino acid in the syntheses provided an unprotected C-terminus upon cleaving. Peptides were cleaved using 95% trifluoroacetic acid, with 2.5% triisopropylsilane and 2.5% water. The cleavage product was then purified using reverse-phase HPLC on a Varian C_{18} preparatory scale column using gradients of water and acetonitrile spiked with 0.1% and 0.085% TFA, respectively. Collected fractions were then lyophilized, and their identity and molecular weight were confirmed on a Bruker Esquire ion trap mass spectrometer. Sequence and purity were verified by ^{1}H NMR.

Circular Dichroism Methods. Circular dichroism (CD) stock solutions were prepared by dissolving approximate amounts of peptide in 20 mM aqueous pH 7 phosphate buffer to make a solution of about 200 μM. The concentration of the stock solution was measured by the UV absorption of tyrosine and tryptophan (ϵ = 1280 M^{-1} cm^{-1} and 5580 M^{-1} cm^{-1}, respectively, at 280 nm). CD samples were diluted appropriately to obtain 30 μM solutions of the peptide in buffer. Fluoro alcohols were added as needed.

Spectra were recorded on a Jasco J715 spectropolarimeter using 0.10 cm path length cells. The calibration of the wavelength and degree ellipticity scales has been described previously.[19] Typical spectral accumulation parameters were a scan rate of 100 nm/min and a 0.2 nm step resolution over the range of 190−270 nm with 16 scans averaged for each spectrum. The accumulated average spectra were trimmed at a dynode voltage of 600 V prior to blank subtraction and smoothing, the latter of which used the reverse Fourier transform procedure in the Jasco spectra analysis software. For melting experiments, the temperature was increased from 5 to 95 °C in 10 °C increments. Each step was equilibrated at the target temperature for 5 min before acquiring data. CD data are reported in residue molar ellipticity units (deg cm^2 residue $dmol^{-1}$).

NMR Methods. All NMR experiments were collected on either a Bruker DRX-500 or DMX-750 spectrometer. A combination of TOCSY and NOESY 2D NMR spectra was used to assign all resonances. A 60 ms MLEV-17 spinlock was employed for the TOCSY and a 150 ms mixing time for the NOESY. The samples consisted of ~2 mM peptide in buffered water (phosphate buffer for pH 6−7 and formate buffer for pH 2.5) with 10% D_2O. Some samples used deuterated trifluoroethanol (TFE) or hexafluoro2-propanol (HFIP) as cosolvents, which were added in the appropriate amounts by gastight syringe. Sodium 2,2-dimethyl-2-silapentane-5-sulfonate (DSS) was used as the internal chemical shift reference and set to 0 ppm for all conditions independent of temperature, pH, and cosolvent concentration.

CSD Calculations. Our previously published method[20] relied on a set of literature-derived random coil values, which were then corrected for temperature, solvent, and end charge effects. The only nearest-neighbor effects included were those due to proline insertion. The method we employ here includes updated random coil shifts, temperature corrections for both H_N and Hα protons, and a significantly increased table of nearest-neighbor effects, extending out to the $i \pm 2$ residues. The new random coil values are the average of six data sets: (a) our prior values,[20] (b) a listing provided by J. Nieto and M. Rico (personal communication), (c) Merutka et al.,[21] (d) Wishart et al.,[22] and (e) Schwarzinger et al.,[23] with further modifications introduced based on four series of random coil peptides prepared in our lab (Ac-RKAXAGK, Ac-AKAXAGK, Ac-RKAXAGS, Ac-RKAXAGD) for key residues (X). To have glycine as a perturbation rather than the context, the data from sets c−e were corrected for a Gly-to-Ala substitution using the nearest-neighbor effects reported by Schwarzinger et al.[24] Amide H_N data from set c was corrected for the effect expected for the uncapped N-terminus (the three-effect[20,25]). Because of subsequent refinement, the near-neighbor corrections we employed did not correspond to the large effects observed by Schwarzinger et al.[24] Testing against an in-house set of longer random coil peptides containing aromatic and proline residues resulted in smaller but nonzero near-neighbor corrections for the following residues: Cys, Gly, His, Pro, Phe, Thr, Trp, and Tyr. In shorter peptides, including the new "AXAG" series prepared specifically for refinement of our random coil values, we continue to observe larger near-neighbor effects for X = Tyr (and particularly for Trp), but these have not been included in our CSD calculation algorithm. We find that these do not improve agreement when applied to longer peptides and denatured proteins; the same conclusion has been reported by Schwarzinger et al.[24]

The algorithms for CSD calculations have been automated and are publicly available (http://andersenlab.chem.washington.edu/CSDb). The program calculates and graphs Hα and H_N CSDs across a sequence. For H_N's, alternative temperature gradients are used depending on the degree of solvent sequestration present in the folded state and the degree of unfolding observed over the temperature range employed in a melting study.

(15) Espinosa, J. F.; Syud, F. A.; Gellman, S. H. *Protein Sci.* **2002**, *11*, 1492−1505.
(16) Kobayashi, N.; Honda, S.; Yoshii, H.; Munekata, E. *Biochemistry* **2000**, *39*, 6564−6571.
(17) Haque, T. S.; Gellman, S. *J. Am. Chem. Soc.* **1997**, *119*, 2303−2304.
(18) Kiehna, S. E.; Waters, M. L. *Protein Sci.* **2003**, *12*, 2657−2667.
(19) (a) Andersen, N. H.; Liu, Z.; Prickett, K. S. *FEBS Lett.* **1996**, *399*, 47−52. (b) Andersen, N. H.; Brodsky, Y.; Neidigh, J. W.; Prickett, K. S. *Bioorg. Med. Chem.* **2002**, *10*, 79−85.
(20) Andersen, N. H.; Neidigh, J. W.; Harris, S. M.; Lee, G. M.; Liu, Z.; Tong, H. *J. Am. Chem. Soc.* **1997**, *119*, 8547−8561.
(21) Merutka, G.; Dyson, H. J.; Wright, P. E. *J. Biomol. NMR* **1995**, *5*, 14−24.
(22) Wishart, D. S.; Bigam, C. G.; Holm, A.; Hodges, R. S.; Sykes, B. D. *J. Biomol. NMR* **1995**, *5*, 67−81.
(23) Schwarzinger, S.; Kroon, G. J. A.; Foss, T. R.; Wright, P. E.; Dyson, H. J. *J. Biomol. NMR* **2000**, *18*, 43−48.
(24) Schwarzinger, S.; Kroon, G. J. A.; Foss, T. R.; Chung, J.; Wright, P. E.; Dyson, H. J. *J. Am. Chem. Soc.* **2001**, *123*, 2970−2978.
(25) Andersen, N. H.; Chen, C.; Lee, G. M. *Protein Pept. Lett.* **1995**, *1*, 215−222.

Results

The GB1 hairpin is listed in the SLoops EaaagE class (www-cryst.bioc.cam.ac.uk/~sloop).[26] In GB1p, Lys[10] is at position T4 (with ϕ/ψ values of +52°/+43° in the protein); we reasoned that Gly would be a better residue for this position. The other loop mutations were suggested by the frequency of residues at each position in similar protein loops (see Supporting Information for the detailed statistics). The analysis was done for **E**-E-α_R-α_R-γ_R-α_L-E-**E** ($n = 223$) and the related **E**-b-α_R-α_R-γ_R-α_L-b-**E** ($n = 62$) loops, both of which can support the strand alignment observed for GB1p.

	S − 2	S − 1	T1	T2	T3	T4	S + 1	S + 2
	E-	E	α_R	α_R	γ_R	α_L	E	**E**
		D/N	Pro	Glu	Thr	Gly	Thr	
%-occurrence		**69**	16	15	**22**	17	18	
				Ala	Ser	Asn	Lys	
%-occurrence				13	15	24	13	
	E-	b	α_R	α_R	γ_R	α_L	b	**E**
		D/N	Pro	Ala	Thr	Gly	Glu	
%-occurrence		**52**	**23**	10	**39**	**66**	23	
			Leu	Glu	Ser		Arg	
%-occurrence			16	11	**32**		15	

On the basis of both sets of percent occurrences (and selectivities for residues at any specific loop position) and our preference for positively charged or neutral side chains over negatively charged ones (a solubility consideration), the following loop sequences were selected for examination:

	S − 1	T1	T2	T3	T4	S + 1
1st	Asn	**Pro**	Ala	Thr	**Gly**	Lys
2nd	Asn	**Pro**	Ala	Thr	**Gly**	Thr
3rd	Asp	**Pro**	Ala	Thr	**Gly**	Arg
4th	His	**Pro**	Ala	Thr	**Gly**	Arg
5th	Asn	Ala	Ala	Thr	**Gly**	Arg
6th	Asn	**Pro**	Ala	Thr	**Gly**	Glu
7th	His	Asn	Pro	Ser	**Gly**	Asn
8th	Asp	**Pro**	Glu	Thr	**Gly**	Glu

To date, the first four have yielded stable hairpin peptides (unpublished data). Herein we report only the application of the first loop sequence to GB1p analogues.

In addition to the loop modifications, we examined the effect of including two lysine residues at the N-terminus to take advantage of potential hairpin stabilization ($\Delta\Delta G = 0.5-2.5$ kJ/mol)[3,27,28] due to attractive Coulombic effects at the strand termini. The series of GB1 and trpzip mutants examined and the fold stability measures obtained appear in Table 1. The ^{1}H NMR chemical shift data obtained for these peptides allow us report the fold populations and melting temperatures and to arrive at a better estimate of the fold stability for GB1p. For the more stable species, the melting data have been corroborated and extended to higher temperatures by CD data.

Since β-hairpin lifetimes are short, equal to or less than 20 μs,[7,29] chemical shifts *are* population-weighted averages and thus provide fold populations when the unfolded state approximates random coil norms and the chemical shifts for the β-state can be measured or accurately approximated. In the present case, our more stable constructs display melting curves that indicate we are near the fully folded values at the low-temperature limit; thus, reasonable estimates of the shift parameters for the β-state are available. Prior data[16] indicate that GB1m1 can serve as an alternate unfolded control. A large number of folding probes exist. (1) The upfield ring current shifts of the valine methyls and aryl-H sites in residue S-2 (Tyr[5] in the GB1 mutants)[30] provide measures of folding.[4,16] (2) The threonines (S − 3 and S + 3) separating the hydrophobic residues in the strands are H-bonded residues; on the basis of analogies to other hairpins,[28,31−34] the inwardly directed amide protons (H_N's) at these positions are expected to shift downfield upon folding.[35] (3) Likewise, the inwardly directed α protons (Hα's) at the hydrophobic cluster sites should also display positive chemical shift deviations (CSDs).[31,38] Both inwardly[31] and outwardly directed Hα shifts (in the H-bonded residues)[13,15] have been used to measure hairpin fold populations[15,18,31,38] and NMR melts.[8] H_N CSDs have been used less frequently. In one case,[31] it was noted that H_N CSDs from H-bonded strand sites provide the same population estimates and melting curves as the CSDs of the inwardly directed Hα sites. To our knowledge, other prior applications of H_N structuring shifts have been qualitative rather than quantitative.[28,33] In the present case, H_N chemical shift deviations provide the clearest delineation of fold population trends. Figure 2 shows the H_N CSDs for a series of GB1 mutants (for related graphs of Hα CSDs see Figure 5 and the Supporting Information). It is clear that all of the peptides favor the same turn and strand alignment geometries and that the peptides can be ranked from least to most folded solely on the basis of the magnitude of the H_N shift deviations: (unfolded) **GB1m1** < **HP5A** < GB1p < **GB1m2** < **GB1m3** (most folded).

We turned to melting studies for more precise measures of relative stability. The temperature dependence of the CSDs

Table 1. Hairpin Sequences and Fold Populations

		% folded (298 K)	T_m (°C)
GB1p	GE**W**T**Y**DDATKT**F**T**V**TE	ca. 30%	
GB1m1	GE**W**T**Y**DDATKT**A**T**V**TE	6 ± 6%	
GB1m2	GE**W**T**Y**NPATGK**F**T**V**TE	74 ± 5%	47 ± 2
GB1m3	KK**W**T**Y**NPATGK**F**T**V**QE	86 ± 3%	60 ± 2
trpzip4	GE**W**T**W**DDATKT**W**T**W**TE		70
HP5W4	KK**W**T**W**NPATGK**W**T**W**QE	>96%	85
HP5W	KK**Y**T**W**NPATGK**W**T**V**QE	92 ± 2%	66
HP5F	KK**Y**T**W**NPATGK**F**T**V**QE	82 ± 4%	53 ± 5
HP5A	KK**Y**T**W**NPATGK**A**T**V**QE	21 ± 10%	

(26) Burke, D. F.; Deane, C. M.; Blundell, T. L. *Bioinformatics* **2000**, *16*, 513−519.
(27) (a) de Alba, E.; Blanco, F. J.; Jiménez, M. A.; Rico, M.; Nieto, J. L. *Eur. J. Biochem.* **1995**, *233*, 283−292. (b) Ramírez-Alvarado, M.; Blanco, F. J.; Serrano, L. *Protein Sci.* **2001**, *10*, 1381−1392.
(28) Griffiths-Jones, S. R.; Maynard, A. J.; Searle, M. S. *J. Mol. Biol.* **1999**, *292*, 1051−1069.
(29) Andersen, N. H.; Dyer, R. B.; Fesinmeyer, R. M.; Gai, F.; Maness, S.; Werner, J. H. In *Peptides 2000*; Martinez, J., Fehrentz, J.-A., Eds.; EDK: Paris, France, 2001; pp 553−554.
(30) These large structuring shifts, −0.98 ppm for Tyr[5]-Hδ in **GB1m3** and −2.1 ppm for Trp[5]-Hϵ3 in **HP5W**, disappear in the Ala[12] mutants (**GB1m1** and **HP5A**).
(31) Andersen, N. H.; Dyer, R. B.; Fesinmeyer, R. M.; Gai, F.; Liu, Z.; Neidigh, J. W.; Tong, H. *J. Am. Chem. Soc.* **1999**, *121*, 9879−9880.
(32) Stanger, H. E.; Syud, F. A.; Espinosa, J. F.; Giriat, I.; Muir, T.; Gellman, S. H. *Proc. Natl. Acad. Sci. U.S.A.* **2001**, *98*, 12015−12020.
(33) Tatko, C. D.; Waters, M. L. *Protein Sci.* **2003**, *12*, 2443−2452.
(34) Tatko, C. D.; Waters, M. J. *J. Am. Chem. Soc.* **2004**, *126*, 2028−2034.
(35) The upfield shift of the S + 1 H_N has previously been noted in a three-stranded sheet,[36] in designed hairpins,[28,33,34] and in short β-turn peptides[37] that cannot form hairpin structures.
(36) Sharman, G. J.; Searle, M. S. *J. Am. Chem. Soc.* **1998**, *120*, 5291−5300.
(37) (a) Dyson, H. J.; Rance, M.; Houghten, R. A.; Lerner, R. A.; Wright, P. E. *J. Mol. Biol.* **1988**, *201*, 161−200. (b) Wright, P. E.; Dyson, H. J.; Lerner, R. A. *Biochemistry* **1988**, *27*, 7167−7175. (c) Campbell, A. P.; Sykes, B. D.; Norrby, E.; Assa-Munt, N.; Dyson, H. J. *Folding Des.* **1996**, *1*, 157−165.
(38) Sharman, G. J.; Griffiths-Jones, S. R.; Jourdan, M.; Searle, M. S. *J. Am. Chem. Soc.* **2001**, *123*, 12318−12324.

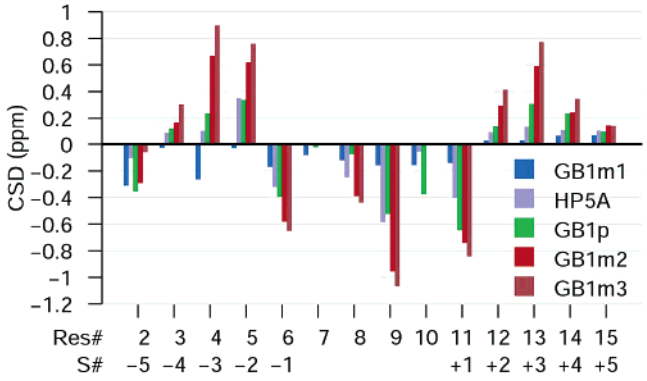


Figure 2. Backbone amide H_N chemical shift deviations for a series of peptides appearing in the following order: **GB1m1**, **HP5A**, GB1p, **GB1m2**, **GB1m3**. All data are at 27 °C in pH 6 aqueous buffer except for GB1p, which uses literature values at 5 °C. CSD calculations employed an updated version of our prior coil state shift and shift gradient method[20] with added nearest-neighbor effects. The H_N CSDs at positions T3 and S + 1 (residues 9 and 11) are large and diagnostic of this hairpin fold.[35] The strand positions (S) are assigned on the basis of the scheme in Figure 1.

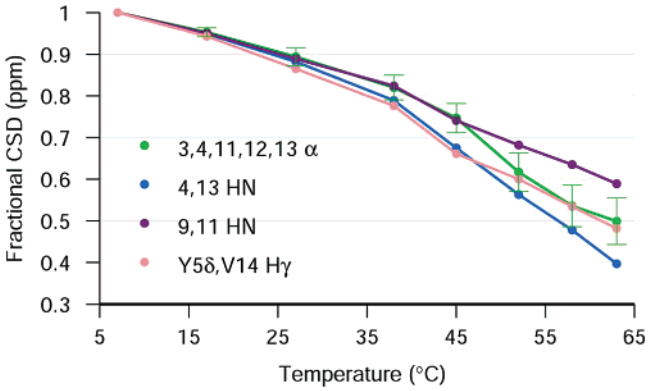


Figure 3. Chemical shift measures of the melting behavior of peptide **GB1m3**, fractional CSDs (with the 280 K value set to 1.00) versus temperature. Four different measures are shown: (1) the average (and SE) fractional CSD for five strand Hα resonances (residues 3, 4, 11, 12, and 13), (2) the average fractional CSD of two strand H_N's (residues 4 and 13), (3) two H_N's (residues 9 and 11) from the loop region, and (4) two ring current shifted side chain sites. The 9,11-H_N values presumably measure disordering of the loop region while the ring current shifted side chain sites reflect only the loss of the specific cross-strand hydrophobic interactions.

reflects changes in the folded population and can thus provide T_m estimates. For the hairpins reported herein, all resonances with a |CSD| greater than 0.4 ppm display monotonic behavior in melts and give the same T_m (within ±5 °C or less) independent of whether the CSD is due to secondary structuring, ring currents, or a combination of these effects.[39] This behavior is expected only for cooperative unfolding transitions. Complete melting data for all of the peptides are included in the Supporting Information, but here we focus on **GB1m3** and **HP5W4**, the most stable construct in each series. Figure 3 shows representative chemical shift melts for **GB1m3** in a format that serves to illustrate the evidence for cooperativity. More classical melting curves appear in Figure 4.

We observe excellent agreement in the chemical shift melts for hydrogens that are attached to carbon rather than nitrogen. Both the strand Hα's and the ring current shifted side chain resonances indicate a melting temperature of 60 ± 2 °C

(39) There are some exceptions to the CSDs predicted by Figure 1; these are attributed to ring current shifts that are in the opposite direction of the diamagnetic anisotropy effects due to backbone amide units.

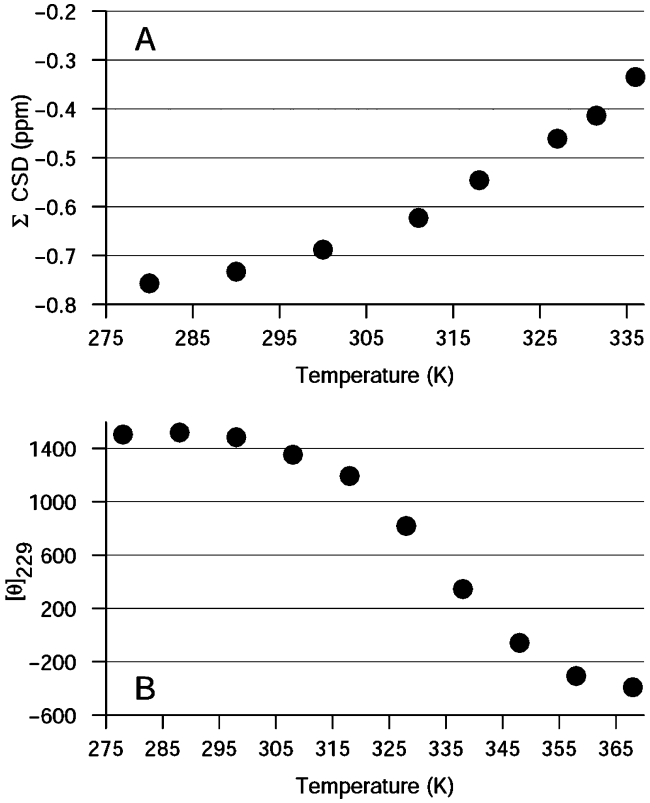


Figure 4. Representative NMR and CD melting curves for **GB1m3**. Panel A tracks melting as the change in the sum of the CSDs for two upfield-shifted Hα sites (Tyr[5] and Pro[7]). Panel B follows the ellipticity of a 229 nm CD band resulting from the ordering of either a Trp or Tyr aromatic side chain in the folded state.

(assuming a folded fraction of 0.96 at the lowest temperature recorded). The data based on H_N shifts have an additional source of error: assumptions must be made regarding the intrinsic H_N temperature shift gradients for the unfolded and folded states. This is particularly problematic for the loop sites (9,11-H_N); for the present figure, we have assumed fully sequestered H_N's, which yields the fractional CSD value of 0.59 at 63 °C. If instead full exposure is assumed, the value would be as small as 0.43.

A typical melting curve for **GB1m3**, based on the Hα CSDs of residues 5 and 7 (which are the most upfield shifted Hα's and were not included in Figure 3), appears in panel A of Figure 4. Alternative presentations of this and additional NMR melts using Kobayashi's δ vs T^{-1} derivative method[8] for deriving T_m appear in the Supporting Information.

Although none of the GB1 mutants display the classic β-structure CD signal, CD did corroborate the NMR melting data and extended the study to higher temperatures and lower concentration, where aggregation cannot compete with refolding. In the case of the GB1 mutants, structuring is indicated by a modest intensity positive band at 228 to 229 nm that is absent in **GB1m1**. For **GB1m3**, a plot of $[\theta]_{229}$ versus T indicates a melting temperature of 62 °C (panel B of Figure 4), while the NMR data in panel A indicate 60 °C. A similar CD study of **GB1m2** afforded a T_m estimate of 49 °C.

Estimating the Fold Population of WT GB1p from Literature NMR Data. Four sets of NMR data for WT GB1p have been reported in sufficient detail to extract chemical shifts: Blanco et al. (5 mM phosphate pH 6.3, 9:1 H_2O/D_2O),[5] the same medium with added TFE (30 vol %) or 6 M urea,[40]

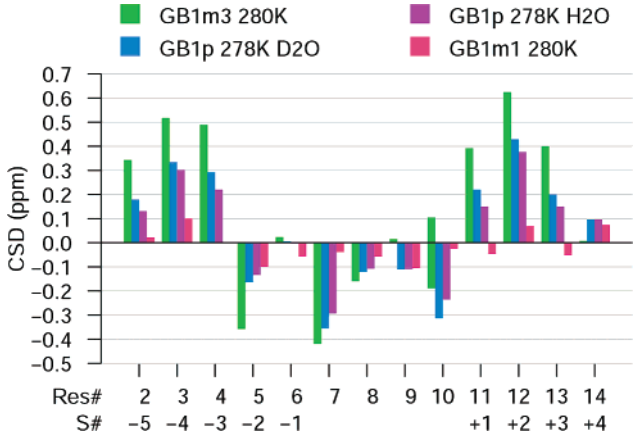


Figure 5. Residue by residue comparisons (residues 2−14) of the Hα CSDs of GB1p and mutants (in the order appearing) **GB1m3** (280 K, most folded example), GB1p (D_2O, 278 K), GB1p (H_2O, 278 K), and **GB1m1** (unfolded control, 280 K). The strand positions (S) are assigned on the basis of the scheme in Figure 1.

and Honda et al. (5 mM phosphate pH 7 D_2O).[8] In addition, limited data are available for a number of single-site mutants, including both a more stable (D7A) and two essentially unfolded mutants (D6A and T9A).[16] The Kobayashi and Honda reports included a range of temperatures, with the highest data points representing the unfolded state. In Figure 5, the Hα shifts from residues 2−14 for WT GB1p are compared to those of **GB1m3** and **GB1m1**.

We can use **GB1m1** as an unfolded control and attribute the shifts of **GB1m3** to a 96% folded population. Using the shift differences at each shifted site (2,3,4,5,11,12α), we calculated the folded fraction for GB1p at 278 K to be 0.438 ± 0.062 (in H_2O) and 0.556 ± 0.052 (in D_2O). For the H_2O data, eight backbone NHs can be used similarly and afford the same calculated folded fraction, 0.428 ± 0.089.

Blanco and Serrano also reported data for GB1p in 30% TFE, and we have examined all of our GB1 mutants in this medium. In the case of GB1p, there were nine sites displaying TFE-induced structuring shifts greater than 0.4 ppm; the CSD increases upon TFE addition were 92 ± 69%. There are 15 sites in **GB1m3** that display >0.4 ppm structuring shifts versus **GB1m1**. The effect of TFE addition at these sites was determined: the increase in these CSDs was 4 ± 15%. The lack of further structuring upon TFE addition, in the case of **GB1m3**, serves as evidence of the stability of the hairpin in water. The percent-folded estimation method used above can be applied, although with less confidence, to GB1p in 30% TFE, with the following result: 74 ± 10% folded at 5 °C.

Kobayashi and Honda[8,16] have noted that ring current shifts are a particularly good measure of folding in these systems. The two largest structuring shifts (both upfield) are at S − 2 (Tyr^5-Hδ) and S + 4 (Val^{14}-Hγ). The latter is extremely remote from the loop mutation positions, and thus the same shifts should be observed at this site for the hairpin state with the stable hydrophobic cluster configuration. The two panels of Figure 6 collect all of the available literature data for WT GB1p, the D6A and D7A mutants,[16] and a comparison with the data for **GB1m2** (T_m = 45.5 °C, by backbone site NMR shift melting curves, 49 °C by CD) and **GB1m3** (T_m = 60 ± 2 °C by CD and backbone site NMR shifts).

(40) Blanco, F. J.; Serrano, L. *Eur. J. Biochem.* **1995**, *230*, 634−649.

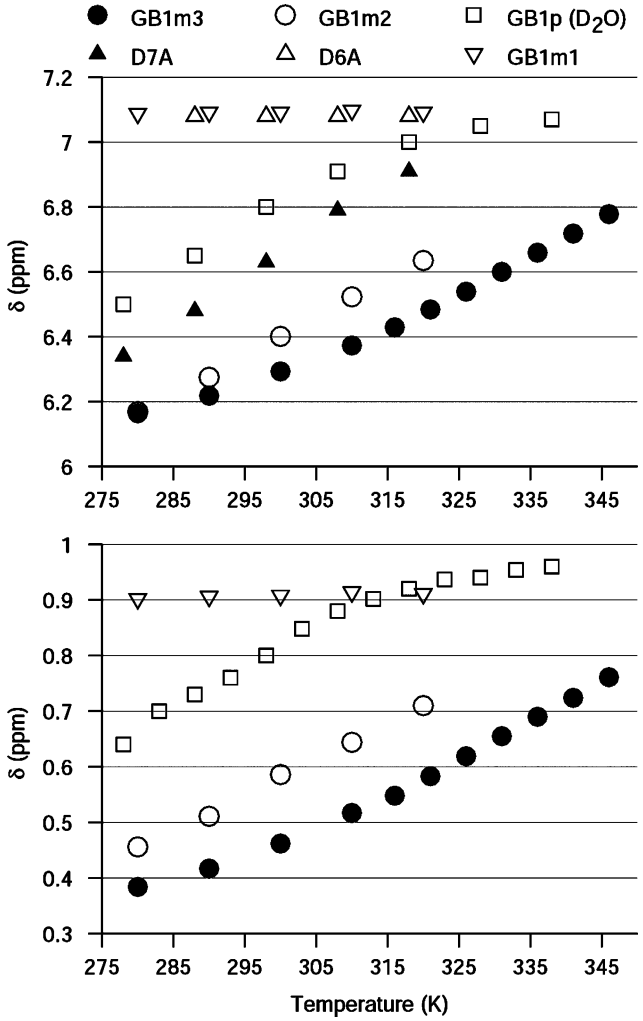


Figure 6. A comparison of melting curves based on the chemical shifts of Tyr^5-Hδ (Panel A) and the upfield Val^{14}-Hγ (Panel B) for WT GB1p (data from ref 8) and a number of analogues including **GB1m3**, the most stable construct to date. The data are shown as δ versus *T*. We have assumed, without evidence, that the same diastereotopic Val-Me is upfield in **GB1m1**.

It is immediately apparent that all of these curves appear to have essentially the same values for the folded and unfolded shifts. GB1p is simply less stable with a T_m that is near the low-temperature limit of the plot. The estimates of folded fraction for GB1p in D_2O that result at 25 °C are 0.29 (Figure 6, panel A) and 0.27 (panel B). Using our CSD criterion for T_m, we find that both plots give a value of 59 °C for **GB1m3**.

When the same shift data are plotted versus 1/*T*, the slopes of the melting rate in the nearly linear portion of the curves are only very slightly less steep for **GB1m3**, suggesting little change in the cooperativity of the transition. For the Tyr^5-Hδ data, the slope is still increasing over the highest temperature increment (68−73 °C); for the Val^{14}-Hγ data, this may also be the case, but the slope changes over the 58−73 °C range are not statistically significant. All of this (and previously presented) data supports our conclusion that **GB1m3** has a T_m of 59 °C (or higher), while the wild-type GB1p data suggests a T_m of 7−12 °C in D_2O.[41]

(41) Because studies of the folding kinetics of GB1p[7] assumed a significantly larger fold population, we suggest that the conclusions may need to be reevaluated in light of the present study. In addition, GB1p should no longer be viewed as the model for folding simulations of β hairpins as there are many better-folded β-hairpin structures available.

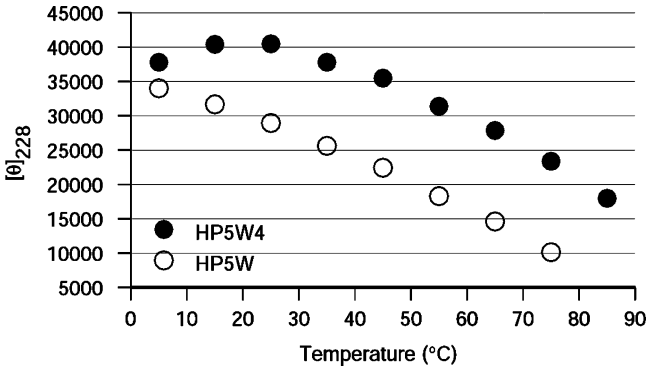


Figure 7. CD melt for HP5W and HP5W4: the temperature dependence of $[\theta]_{228}$. The values for **HP5W** were adjusted to account for the presence of unfolded Trp-containing impurities.

The introduction of the favorable loop and termini mutations used for **GB1m3** into the trpzip4 system yielded **HP5W4**, which displays only modest (<8%) structure loss up to 37 °C by NMR. Some melting is observed at 47 and 62 °C; extrapolation of this trend suggests a melting temperature greater than 80 °C (see Supporting Information). We turned to CD to probe higher temperatures. In analogy to a trpzip peptide,[9] **HP5W4** and **HP5W** (but not **HP5F** and **HP5A**) display an aromatic chromophore exciton couplet ($[\theta]_{max}$ at 228 nm, $[\theta]_{min}$ at 213 nm) superimposed on a β-secondary structure signature. CD melts monitoring either the amplitude of the couplet or the ellipticity at the maximum (228 nm) have been examined. The latter has some advantage since there is essentially no backbone secondary structure signal at this wavelength. The unfolded reference value is on the order of −1000 to −2500° per Trp residue. The temperature dependence of $[\theta]_{228nm}$ indicates T_m = 85 °C (**HP5W4**) and T_m = 66 °C (for **HP5W**) (Figure 7).

The melting transitions are quite broad, with **HP5W4** displaying some cold denaturation as well, which also appears in the NMR melts. Unfolding is far from complete even at the highest temperatures examined. For both peptides the slope ($\Delta[\theta]/\Delta T$) is still increasing over the highest ΔT increment recorded, 65−75° (**HP5W**) and 75−85° (**HP5W4**), placing T_m in or higher than these temperatures.

Mutants **HP5F** and **HP5A** were examined to provide further data on the stabilizing effects of cross-strand aromatic side chain interactions in hairpins. **HP5F** retains a hairpin conformation with a melting temperature of 53 ± 5 °C based on the temperature dependence of chemical shift measures of structuring (Supporting Information). **HP5A** displayed little or no preference for a hairpin fold.

Discussion

This study extends the use of occurrence statistics in proteins as a guide to designing more stable peptide folds and establishes H_N chemical shifts as particularly valuable parameters for the diagnosis and quantitation of β-structure in hairpin and sheet models. The advantages of H_N shifts include diagnostic CSDs for turn and loop positions and for the S + 1 position, as well as the strand H_N's involved in cross-strand H-bonding. Melts based on H_N shifts are, however, complicated by the intrinsic temperature dependence of H_N shifts in both the folded and unfolded states which depend on the hydrogen-bonding status, including the strength of the H bond.[20,42] Nonetheless, chemical shift melts likely represent the most reliable measure of fold stability for peptides that display rapid folding/unfolding rates and are certainly the best test of the cooperativity of the state transition.

The present series of peptides confirms the stabilizing effects of cross-strand (S − 2/S + 2) and diagonal (S − 4/S + 2) aryl side chain interactions, with cross-strand W/W interactions being the most stabilizing: introducing a cross-strand W/W interaction, **HP5A → HP5W**, provided at least 7.4 kJ/mol of stabilization at 27 °C. Peptide **HP5W4**, with the best interstrand interactions, displayed a T_m on the order of 85 °C. This is 15 °C higher than that reported for trpzip4 and, to our knowledge, the greatest fold stability observed for any non-cross-linked polypeptide of less than 36-residue length. The notable stabilization achieved in this series by both loop optimization ($\Delta\Delta G$ = 4.5 kJ/mol)[43] and Coulombic interactions of the chain termini ($\Delta\Delta G$ = 2.4 kJ/mol)[44] are either unprecedented or larger than previous examples.

It should be noted, however, that the substantial hairpin stabilization associated with the NPATGK loop is not a reflection of a rigid, persistent, turn geometry in the absence of hydrophobically driven hairpin alignment. The much-reduced upfield shifts of Thr^9-H_N and Lys^{11}-H_N in **HP5A** (ΔCSD = 0.6 and 0.9 ppm for the F12A mutation) indicate that the interactions of the aryl side chains at positions S − 2 and S + 2 are essential not only for hairpin formation but also for loop structuring. Hydrophobically driven folding is also indicated by the cold denaturation observed for some of the species. While the present melting data are not, in our opinion, of sufficient precision to derive accurate thermodynamic parameters, in all cases a positive ΔC_p for unfolding was required to fit the curvature of the melting curves. Our main reasons for not detailing the thermodynamic fits to melting curves are the minor differences observed using different probes (CD versus strand residue shifts versus loop residue shifts); whether these represent some non-two-state folding behavior cannot be established at this point.

For proteins, mutations that increase stability typically produce a sharper melting transition. The increased hairpin stability obtained by turn mutations in the present case likely represents a decrease in the chain configuration component of ΔS_U and, as a result, does not produce sharper melting transitions. Further mutational studies and minimization of the HP5 hairpin series will be reported in due course.

Acknowledgment. This work was supported by the U.S. Public Health Service (GM59658). R.M.F. was a recipient of an NIH Molecular Biophysics Training Grant Biophysics Program Training Grant (GM08268).

Supporting Information Available: A complete account of the residue occurrence statistics in hairpin loops, four figures, and complete NMR assignments for seven peptides (PDF). This material is available free of charge via the Internet at http://pubs.acs.org.

JA0379520

(42) Cierpicki, T.; Zhukov, I.; Byrd, R. A.; Otlewski, J. *J. Magn. Reson.* **2002**, *157*, 178−180.
(43) Derived from the fraction folded values for GB1p and **GB1m2** at 25 °C.
(44) Derived from the fraction folded values for **GB1m2** and **GB1m3** at 37 °C; the higher temperature was selected since the calculated folded fraction of **GB1m3** is less influenced by the choice of the 100% folded shift values at this temperature.