Designing Stable β-Hairpins: Energetic Contributions from Cross-Strand Residues

Stephen J. Russell and Andrea G. Cochran*

Department of Protein Engineering, Genentech, Inc.
1 DNA Way, South San Francisco, California 94080

Received June 12, 2000

Short peptides that adopt β-hairpin conformations in water have been the object of much recent interest.[1] These peptides are potentially useful models for protein β-sheets and have been used as starting points for design of small three-stranded sheets[2] and structured, metal-free zinc fingers.[3] However, most small β-hairpins are marginally stable, despite the use of very strong β-turns.[1,4] Although it is generally appreciated that favorable cross-strand residue pairing can improve hairpin stability,[1,5] there is no clear procedure for choosing such residue pairs. Favorable residue pairing has been demonstrated to significantly stabilize β-sheet proteins,[6] but it is not yet known whether this would apply directly to hairpins.

We have established a disulfide-cyclized 10-residue peptide as a system in which to compare hairpin stabilities,[7] using changes in the thiol−disulfide equilibrium constant as a probe.[7,8] We compared a series of 19 substitutions in peptide **1**,[7] and our initial efforts revealed tryptophan to be quite stabilizing in the non-hydrogen-bonded (NHB) strand site **X** when paired with a cross-strand leucine. The tryptophan peptide from this series (bhpW) was highly structured in water, adopting the intended hairpin conformation (Figure 1).[7]

Here, we investigate the relationship between the two NHB cross-strand residues. Remarkably, we find that residue preferences for the two structurally inequivalent sites are the same and that specific pair interactions produce only minor deviations from the single site contributions. Accordingly, a tryptophan−tryptophan cross-strand pair is highly stabilizing and appears to be the optimal NHB pair for β-hairpins.

Ac-CTXEGNKLTC-NH_2	**1**	
Ac-CTLEGNKXTC-NH_2	**2**	X = W, Y, F, L, M, I, V, A
Ac-CTXEGNKWTC-NH_2	**3**	
Ac-CTWEGNKXTC-NH_2	**4**	

Our initial observation of a stabilizing contribution from tryptophan[7] prompted us to question how general the effect might be. Because of the twist of the β-strands, the tryptophan side chain

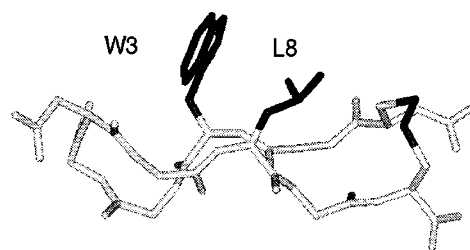

Figure 1. NMR structure (minimized mean) of disulfide-cyclized hairpin bhpW[7] (peptide **1**, **X** = Trp). Side chains W3 and L8 and the disulfide are shown in black. Side chains for the hydrogen-bonded residues (T2, E4, K7, T9) have been omitted for clarity. This figure was made using the program INSIGHT97.0 (Molecular Simulations, Inc.)

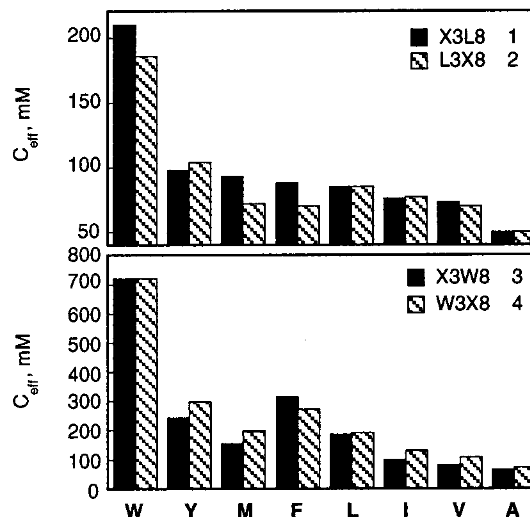

Figure 2. Effective concentration (C_{eff}) values for peptides **1**−**4**. Values for substitutions paired with a cross-strand leucine are shown at top; those for tryptophan pairs are shown below. Values are averages for three or more measurements. Standard deviations were $\leq$5% of the measured C_{eff} (equivalent to $\Delta\Delta G \approx 30$ cal mol^{-1}).

in peptide bhpW (Figure 1) is sandwiched between the side chain of the cross-strand leucine and the type-II′ turn, while the leucine side chain is closer to the disulfide. Therefore, it seems possible that the stabilizing effect of tryptophan might be unique to this structural context. To investigate this, we reversed the hydrophobic pairs (peptide **2**), varying the amino acid at position 8 (nearest the disulfide, Figure 1) with leucine fixed at position 3. Effective concentrations (C_{eff}) of the cysteine thiols were determined as in our previous studies.[7,9]

We find that tryptophan at position 8 is the most stabilizing of those residues tested (Figure 2, top). Significantly, the C_{eff} values are quite close for the Trp−Leu and Leu−Trp pairs, indicating that the two arrangements are about equivalent energetically. This result appears to hold for other residue pairs with leucine: the rank order and numeric values of C_{eff} are similar, but not exact, in the two series (Figure 2, top).

To test whether the equivalence of the reversed hydrophobic pairs might be more general, we examined peptide series **3** and

* Corresponding author. E-mail: andrea@gene.com.

(1) (a) Gellman, S. H. *Curr. Opin. Chem. Biol.* **1998**, *2*, 717. (b) Ramirez-Alvarado, M.; Kortemme, T.; Blanco, F. J.; Serrano, L. *Bioorg. Med. Chem.* **1999**, *7*, 93. (c) Smith, C. K.; Regan, L. *Acc. Chem. Res.* **1997**, *30*, 153.
(2) (a) Kortemme, T.; Ramirez-Alvarado, M.; Serrano, L. *Science* **1998**, *281*, 253. (b) Schenck, H. L.; Gellman, S. H. *J. Am. Chem. Soc.* **1998**, *120*, 4869. (c) Sharman, G. J.; Searle, M. S. *J. Am. Chem. Soc.* **1998**, *120*, 529. (d) de Alba, E.; Santoro, J.; Rico, M.; Jimenez, M. A. *Protein Sci.* **1999**, *8*, 854.
(3) Struthers, M. D.; Cheng, R. P.; Imperiali, B. *Science* **1996**, *271*, 342.
(4) Syud, F. A.; Espinosa, J. F.; Gellman, S. H. *J. Am. Chem. Soc.* **1999**, *121*, 11577.
(5) Griffiths-Jones, S. R.; Maynard, A. J.; Searle, M. S. *J. Mol. Biol.* **1999**, *292*, 1051.
(6) (a) Smith, C. K.; Regan, L. *Science* **1995**, *270*, 980. (b) Merkel, J. S.; Regan, L. *Folding Des.* **1998**, *3*, 44.
(7) Cochran, A. G.; Tong, R. T.; Starovasnik, M. A.; Park, E. J.; McDowell, R. S.; Theaker, J. E.; Skelton, N. J. *J. Am. Chem. Soc.*, in press.
(8) (a) Stroup, A. N.; Gierasch, L. M. *Biochemistry* **1990**, *29*, 9765. (b) Milburn, J.; Konishi, Y.; Meinwald, Y. C.; Scheraga, H. A. *J. Am. Chem. Soc.* **1987**, *109*, 4486. (c) Milburn, P. J.; Meinwald, Y. C.; Takahashi, S.; Ooi, T.; Scheraga, H. A. *Int. J. Pept. Protein Res.* **1988**, *31*, 311. (d) Falcomer, C. M.; Meinwald, Y. C.; Choudhary, I.; Talluri, S.; Milburn, P. J.; Clardy, J.; Scheraga, H. A. *J. Am. Chem. Soc.* **1992**, *114*, 4036.

(9) Briefly, equilibrium constants were measured relative to the reference thiol glutathione (GSH) at pH 8.1 and 20 °C. Concentrations of reduced and oxidized species were determined from HPLC peak areas, and effective concentrations (C_{eff}) were obtained from the following relationship: $C_{eff} = ([GSH]^2 \cdot [peptide_{ox}])/([GSSG] \cdot [peptide_{red}])$. The C_{eff} values reported here for peptide series **2**, **3**, and **4** are averages from three or more HPLC analyses; values for peptide series **1** are from ref 7. Selected peptides were reassayed at lower concentrations to confirm that C_{eff} was concentration independent; no evidence of peptide aggregation was found. A detailed protocol may be found in ref 7.

10.1021/ja002085+ CCC: $19.00 © 2000 American Chemical Society
Published on Web 12/01/2000

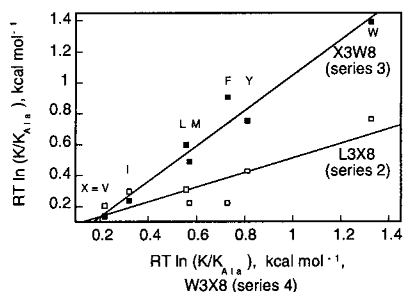

Figure 3. Comparison of substitution free energy differences.

Table 1. Slopes (p) of Free Energy Correlations for Peptide Series **1**−**4**

x-axis data set	p, X3 vs X8	p, Leu vs Trp
W3X8 (**4**)[a]	1.15 (1.11)[b]	0.47 (0.43)[b]
L3X8 (**2**)	0.98 (0.86)[b]	−
X3W8 (**3**)	−	0.43 (0.32)[b]

[a] Plots vs W3X8 (**4**) data are shown in Figure 3. [b] Values in parentheses were obtained using the Trp peptides (**X** = W in series **1**−**4**) as internal reference instead of the Ala peptides (see text and Figure 3).

4, in which residues are instead paired with a cross-strand tryptophan (Figure 2, bottom). As with Leu pairs, a close correspondence is seen between the two Trp series, both in rank order and value of C_{eff}. We conclude that the two cross-strand sites are essentially equivalent and, therefore, that neither residue makes specific packing interactions with the disulfide or with residues in the turn. Because the 3,8-cross-strand pair is, in effect, "isolated", our cyclic peptide is a suitable model system for quantifying side chain pair interactions between β-strands.

The two leucine series (**1** and **2**) may be compared to the tryptophan series (**3** and **4**). The trends in the two data sets are remarkably similar (Figure 2, top vs bottom), suggesting that the cross-strand residues contribute to stability in a roughly independent manner. To explore this idea, we calculated free energy differences for substitutions within each of the peptide series relative to the alanine peptide from that series ($\Delta\Delta G = -RT \ln \{C_{eff,\mathbf{X}}/C_{eff,ala}\}$). Representative comparisons are plotted in Figure 3.

Linear free energy relationships exist among the four data sets. This is seen not only in comparisons of particular cross-strand pairs switched between NHB sites 3 and 8, but also for comparisons of Trp pairs with Leu pairs in the same orientation (Figure 3). (There is more scatter in the latter plots.) Slopes (p) are given in Table 1; the p values do not depend on the choice of reference peptide (**X** = A or W, Table 1).

Consistent with the idea that positions 3 and 8 are equivalent, p is near 1 for plots comparing these data. In contrast, when Leu pairs are compared to Trp pairs, p is about 0.4. This means that for a given pair of residues **X**, the expected difference in hairpin stability is ~2.5-fold larger with Trp as the cross-strand partner than with Leu. Given these simple relationships, $\Delta\Delta G$ could be calculated for *any* cross-strand pair relative to a reference pair by multiplying a substitution energy s_X by p for the cross-strand partner.[10]

The implication of the above analysis is seen by calculation of double-mutant coupling energies. Typically, in proteins, the difference between the effect of a double substitution ($\Delta\Delta G_{fold}$) and the sum of the $\Delta\Delta G$ of the single substitutions is interpreted as an interaction, or coupling, energy.[11] These energies are zero (by definition) if the substitutions produce independent effects but can be substantial if the residues are in contact.[11] Here, the C_{eff} ratios yield $\Delta\Delta G$ for the single or double substitutions. In the example shown in Figure 4, the coupling energy would be −136 cal mol^{-1} for the Trp−Tyr pair relative to a Leu−Leu reference state. If, instead, the single substitution energies are calculated sequentially, scaling by p in the second step, the coupling energy is only +41 cal mol^{-1} (within the error of these experiments, see Figure 2). That is, for those pairs on the line of the appropriate plot, the coupling energy is 0. Therefore, we conclude from the observed linear free energy relationships that single site preferences (s and p) are most important in predicting hairpin stability.[12]

(10) By analogy to the Hammett equation.

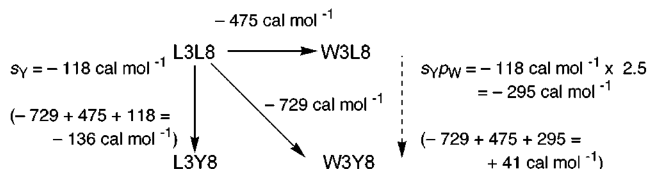

Figure 4. Double mutant analysis[11] of the stability of W3Y8 relative to L3L8. Energies given are $\Delta\Delta G_{mut} = -RT \ln(C_{eff,X3X8}/C_{eff,L3L8})$.

In contrast, statistical analyses of HB and NHB cross-strand pairs in β-sheet proteins find many residue pairs to be positively or negatively correlated with high confidence.[13−15] Largely in accord with the statistical preferences, protein mutagenesis studies have identified coupling energies as large as 1 kcal mol^{-1} between HB pairs.[6] Our data suggest that such large coupling energies do not necessarily indicate *unique* interactions between pairs of amino acid side chains, but instead may reflect differential sensitivity to *all* residue substitutions opposite a given cross-strand partner. Alternatively, they may reflect more stringent packing requirements in an extended β-sheet.

It has been proposed that the use of cross-strand pairs statistically preferred in proteins might improve stability or fix strand register in β-hairpins.[16] Among the most highly correlated NHB pairs are Thr−Thr and Val−Val, and stereochemical arguments have been used to rationalize these preferences.[13,14] We have introduced these pairs into our hairpin at NHB sites 3 and 8 and find the stabilities to be very close to that of the Leu−Leu analogue ($\Delta\Delta G$ = 0 and 50 cal mol^{-1}, for TT and VV respectively; not shown). Notably, the Trp−Trp pair yields a much more stable structure ($\Delta\Delta G$ = −1250 cal mol^{-1} vs LL), and we conclude that, despite its rarity in proteins,[15] it is the optimal NHB pair for isolated β-hairpins.[17]

JA002085+

(11) Wells, J. A. *Biochemistry* **1990**, *29*, 8509.
(12) However, for the Phe−Trp pair (compare Figure 2, top and bottom), a similar analysis yields $\Delta\Delta\Delta G$ = −253 cal mol^{-1} *after* scaling by p. This interaction energy is significant. The value is small when compared to the total range of energies seen for single-site substitutions (Figure 3): we believe that a few small but real interaction energies (and experimental error) are responsible for the *scatter* in our correlations.
(13) Wouters, M. A.; Curmi, P. M. G. *Proteins* **1995**, *22*, 119.
(14) Hutchinson, E. G.; Sessions, R. B.; Thornton, J. M.; Woolfson, D. N. *Protein Sci.* **1998**, *7*, 2287.
(15) Several positively correlated NHB pairs (confidence ≥95%) combine residues used in this study. These are FL, LL, VI, and VV (ref 11); and FA, FL, LW, VV, and WY (ref 12). Both analyses identify WW as positively correlated, but because of its rarity, confidence is low.
(16) (a) de Alba, E.; Rico, M.; Jimenez, A. A. *Protein Sci.* **1997**, *6*, 2548. (b) Ramirez-Alvarado, M.; Blanco, F. J.; Serrano, L. *Nat. Struct. Biol.* **1996**, *3*, 604.
(17) For an example of a structured cyclic peptide with a Trp−Trp NHB pair, see Favre, M.; Moehle, K.; Jiang, L.; Pfeiffer, B.; Robinson, J. A. *J. Am. Chem. Soc.* **1999**, *121*, 2679.