

Article

Chemical shifts provide fold populations and register of β hairpins and β sheets

R. Matthew. Fesinmeyer^a, F. Michael Hudson^a, Katherine A. Olsen^a,
George W. N. White^{a,b}, Anna Euser^a & Niels H. Andersen^{a,*}

^aDepartment of Chemistry, University of Washington, Seattle, WA, 98195, USA; ^bRotation student in the University of Washington Biomolecular Structure & Design program.

Received 22 June 2005; Accepted 29 September 2005

Key words: chemical shift deviations, CSD, hairpin fold population, hairpin register, turn signatures, two-state folding

Abstract

A detailed analysis of peptide backbone amide (H_N) and $H\alpha$ chemical shifts reveals a consistent pattern for β hairpins and three-stranded β sheets. The $H\alpha$'s at non-hydrogen-bonded strand positions are inwardly directed and shifted downfield ~ 1 ppm due largely to an anisotropy contribution from the cross-strand amide function. The secondary structure associated $H\alpha$ shift deviations for the H-bonded strand positions are also positive but much smaller (0.1–0.3 ppm) and the turn residues display negative $H\alpha$ chemical shift deviations (CSDs). The pattern of H_N shift deviations is an even better indicator of both hairpin formation and register, with the cross-strand H-bonded sites shifted downfield (also by ~ 1 ppm) and with diagnostic values for the first turn residue and the first strand position following the turn. These empirical observations, initially made for [2:2]/[2:4]-type-I' and -II' hairpins, are rationalized and can be extended to the analysis of other turns, hairpin classes ([3:5], [4:4]/[4:6]), and three-stranded peptide β -sheet models. The $H\alpha$'s at non-hydrogen-bonded sites and H_N 's in the intervening H-bonded sites provide the largest and most dependable measures of hairpin structuring and can be used for melting studies; however the intrinsic temperature dependence of H_N shifts deviations needs to reflect the extent of solvent sequestration in the folded state. Several observations made in the course of this study provide insights into β -sheet folding mechanisms: (1) The magnitude of the H_N shifts suggests that the cross-strand H-bonds in peptide hairpins are as short as those in protein β sheets. (2) Even L-Pro-Gly turns, which are frequently used in unfolded controls for hairpin peptides, can support hairpin populations in aqueous fluoroalcohol media. (3) The good correlation between hairpin population estimates from cross-strand H-bonded H_N shift deviations, $H\alpha$ shift deviations, and structuring shifts at the turn locus implies that hairpin folding transitions approximate two-state behavior.

Introduction

Peptide models of the two most common protein secondary structures, α helices and β sheets, have

been objects of interest for some time. The rules for the *a priori* design of helical peptides are well developed (Doig and Baldwin, 1995). Peptide models of β sheets have not attained the same level of understanding, even with the simplest systems: β hairpins. While numerous protein-derived helical peptides fold to their native structure, though to a

* To whom correspondence should be addressed. E-mail: andersen@chem.washington.edu

lesser degree (first observed in 1971 (Brown and Klee, 1971)), only a few β -hairpin sequences displaying similar structuring propensity outside of the protein context have been found, e.g. the ubiquitin (1–17) sequence (Cox et al., 1993), and the second hairpin (residues 41–56) of the B1 domain of protein G, GB1p (Kobayashi et al., 1993; Blanco et al., 1994). Many hairpin peptides (Blanco et al., 1994; Chen et al., 2001; Gibbs et al., 2002; Fesinmeyer et al., 2004) fail to display the classic CD signature for an antiparallel β sheet observed for both proteins and peptide β oligomers (Cort et al., 1994). This has, with a few exceptions (Ramírez-Alvarado et al., 1996; Maynard et al., 1998; Andersen et al., 1999; Griffiths-Jones and Searle, 2000; Blandl et al., 2003; Ciani et al., 2003; Fesinmeyer et al., 2004), precluded the use of CD for estimating hairpin fold populations. NMR methods have been employed with varying success; for example, GB1p in water at 5 °C has been reported to be 42–82% folded based on NMR criteria (Blanco et al., 1994; Honda et al., 2000; Cochran et al., 2001; Fesinmeyer et al., 2004). There have been reports of conservative single-site loop mutations resulting, based on NOESY connectivities, in dramatic changes in hairpin register—for example a [2:2]- to [3:5]-hairpin interconversion (Searle et al., 1995; Chen et al., 2001). In other cases, designed peptide hairpins have been reported to populate two different hairpin folds (de Alba et al., 1996, 1999). Clearly, dependable methods for determining both hairpin register and fold population are needed. The formalism used to discuss hairpin geometry appears in Figure 1.

For peptides (in some contrast to proteins), hairpin formation requires residues at the turn locus that are particularly favorable for chain direction reversal, notably Asn-Gly, (Sibanda and Thornton, 1991; Ramírez-Alvarado et al., 1996) D-Pro-Gly (**pG**, throughout D-configured amino acids are designated by the bold lower case one letter amino acid symbols), (Karle et al., 1996; Syud et al., 1999) and Pro-Asp-Gly, (Blanco et al., 1993; Searle et al., 1995) but measurable hairpin populations in water are never achieved in the absence of cross-strand hydrophobic clustering *and* the inclusion of β -branched residues in the strands. The same set of favorable turns have been used for constructing 3- and 4-stranded sheet models (Schenck and Gellman, 1998; Sharman

and Searle, 1998; Santiveri et al., 2003; Syud et al., 2003), of which the Schenck-Gellman model, which is re-examined in the present study, serves as an excellent example (Figure 1).

Cross-strand α/α NOEs have been employed to ascertain strand register and their relative intensity was among the first measures of fold population (Searle et al., 1995; de Alba et al., 1996; Ramírez-Alvarado et al., 1996). The latter is a questionable method as NOE intensities are not simple population-weighted averages: the effective correlation time for the folded state cross-strand interactions is unlikely to be the same as that for the distance references that are averaged over the entire ensemble. However, local NOE ratios do not suffer from this problem and can be used for determining fractional helicity (Bradley et al., 1990; Lee et al., 1994) and the extent of β -strand formation (Sharman and Searle, 1998). Since β -hairpin fold lifetimes are short, equal to or less than 20 μ s, (Muñoz et al., 1997; Xu et al., 2003; Dyer et al., 2004, 2005) chemical shifts *are* population-weighted averages and 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. Chemical shift deviations (CSDs) have been used to estimate hairpin populations since the earliest observations (Blanco et al., 1994; Searle et al., 1995). However, the common occurrence of aromatic residues in designed β hairpins has, to some extent, obscured the chemical shift effects of the backbone conformation by superimposing large ring current effects. Gellman and coworkers (Espinosa and Gellman,

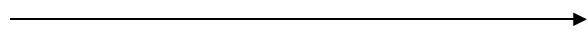
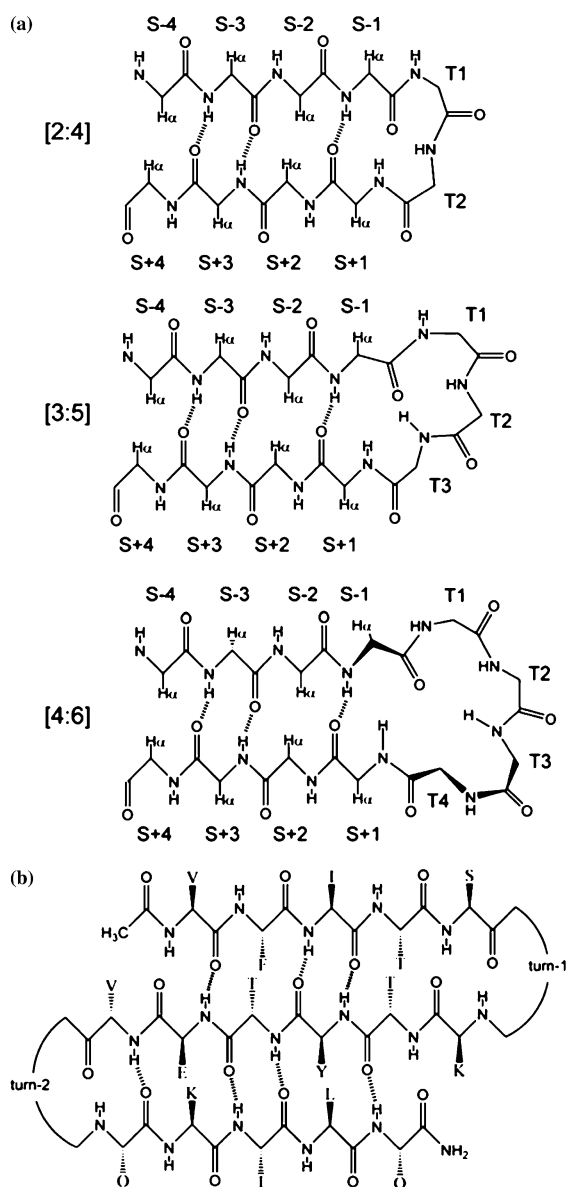


Figure 1. β -Hairpin (a) and three-stranded sheet (b) nomenclature. In (a), *T* indicates turn or loop positions, *S* indicates strand positions numbered from the turn locus. The $S \pm 2$ positions have their $C\alpha$ atoms directed inward and display a short $H\alpha/H\alpha$ distance. The odd-numbered strand positions have their $H\alpha$'s outwardly directed with the sidechains displayed on the top surface of the hairpin as depicted here. The odd-numbered strand positions are also designated as "H-bonded pairs". We employ Thornton's hairpin classification scheme (Sibanda and Thornton, 1991) throughout. The distinction between [2:2]- vs. [2:4]-hairpins (and between [4:4]- and [4:6]-hairpins) is whether or not there is a S-1 C=O to S+1 NH H-bond; the [2:4] and [4:6] H-bonding patterns are illustrated. Henceforth, we refer to these hairpins as [2:2]- and [4:4]-hairpins. In (b), the sequence and strand alignment of the Schenck-Gellman three-strand sheet model are shown. In the original sequence (Schenck and Gellman, 1998) both turns consisted of **pG** (D-Pro-Gly) sequences.



2000; Syud et al., 2001, 2003) employ the H_{α} shifts in the H-bonded residues ($S \pm 1$, $S \pm 3$, etc.) to avoid complications attributed to cross-strand ring current effects. A survey of protein β structures has established that the non-H-bonded residue positions in β strands ($S \pm 2$, $S \pm 4$, etc.) produce larger H_{α} structuring shifts (Sharman et al., 2001). Most studies using 1H CSDs to quantitate fold population have made one or more assumptions: (1) that 30–40% trifluoroethanol (TFE) gives the fully folded species (Ramírez-Alvarado et al., 1996, 1997; Syud et al., 2001),

(2) that the average H_{α} CSDs in β strands should be +0.40 ppm (Santiveri et al., 2001), or (3) that the shifts observed for the D-Pro-Gly turn species, particularly cyclic ones, (e.g. Espinosa et al., 2003; Tatko and Waters, 2004) correspond to 100% folded.

During the course of our efforts to determine the thermodynamics and kinetics of hairpin folding (Andersen et al., 1999), we observed diagnostic patterns in both H_{α} and backbone amide NH (H_N) CSDs that provided independent and unambiguous assignment of hairpin register. We further found the H_N CSD magnitude to be useful for providing population estimates even though H_N shift trends have rarely figured in the assignment of secondary structure and it has generally been observed that it is more difficult to predict or calculate NH shifts (both for the 1H and ^{15}N) based on polypeptide structure (Osapay and Case, 1991; Xu and Case, 2001). Here we demonstrate that consideration of both H_{α} and H_N shift deviations improves structural definition and we derive expectation values for diagnostic CSDs in the fully folded state.

Methods and materials

Peptide synthesis

Representatives of the Met-repressor-inspired (MrH) (Maynard et al., 1998), BH (Ramírez-Alvarado et al., 1996), and ubiquitin/tendamistat-inspired (UTH) (Searle et al., 1995; de Alba et al., 1996; de Alba et al., 1999) hairpins were prepared using fast Fmoc chemistry on an ABI 433A peptide synthesizer and purified by reverse phase HPLC (C_{18} column) with a water (0.1% TFA)/acetonitrile (0.085% TFA) gradient. The Schenck-Gellman (SG) 3-stranded sheet model (Schenck and Gellman, 1998) and its mutants were similarly prepared and purified. Molecular mass was confirmed on a Bruker Esquire ion-trap mass spectrometer, while sequence and purity were verified by 1H NMR. The sequences that were prepared (O = Orn, p = D-Pro) are listed here together with the abbreviations employed; the turn residues are shown in bold.

Additional turn mutants of the SG model were also prepared, the residues present at the first and second turn loci are indicated: SG(PG-pG) and SG(pG-PG) (Schenck and Gellman, 1998),

SG(DG-pG) (Chen et al., 2001), SG(pG-GG), SG(GG-pG), SG(GP-pG), and SG(NG-pG).

NMR data collection

All experiments were performed on a Bruker DRX operating at 500 MHz. All peptide H_N and H_α resonances could be assigned through a combination of 2D TOCSY and NOESY experiments with WATERGATE (Piotto et al., 1992) solvent suppression. The former employed a 60 ms MLEV-17 spinlock and the latter a 150 ms mixing time. Samples contained 1.5–2.5 mM peptide in buffered H_2O with 10% D_2O . Sodium 2,2-dimethyl-2-silapentane-5-sulfonate (DSS) was used as an internal chemical shift reference and was set to 0 ppm for all temperature and pH conditions. The MrH series of peptides was observed at pH 2.5 (50 mM formate), 4.8 (10 mM acetate) and 6.0 (20 mM phosphate). The SG 3-stranded sheet models were examined predominantly using a pH 3 buffer (20 mM formate); detailed studies with SG(pG-pG) and SG(NG-NG) established that there were minor changes in structuring shifts or fold populations on going from pH 3 to pH 6. The SG(NG-NG) mutant displayed more favorable folding about the second hairpin at pH 6. Deuterated TFE and hexafluoroisopropanol (HFIP) were added to the vol-% noted; the aqueous portion was added by pipette with the volatile fluoroalcohols delivered by glass microliter syringes.

Literature NMR data

In their study of H_α chemical shifts, Sharman et al. (Sharman et al., 2001) mined a 100-structure database of PDB structures for which NMR data, via the BioMagResBank (<http://www.bmrb.wisc.edu/>), was also available. We employed the same set of data to study H_N chemical shifts. Seven of the 100 structures were found to be unsuitable for the study as the structure either lacked identifiable strands or the NMR shift data available was either minimal or devoid of H_N chemical shifts. For the remaining 93 structures, the STRIDE application (Frishman and Argos, 1995) was used to identify residues in β strands. NMR chemical shift data was pulled from the corresponding BMRB files, yielding 2137 H_N chemical shifts. The H_N envi-

BH8	RG-ITVNG KTY- GR
MrH1	Ac-KKYTVSINGKKITVSI
MrH3a	KKYTVSINGKKITVSA
MrH3b	KKYTVSIpGKKITVSA
MrH3c	KKYTVSIGGKKITVSA
MrH3d	KKYTVSIPGKKITVSA
MrH4a	KKLTVSINGKKITVSA
MrH4b	KKLTVSIpGKKITVSA
MrH5a	KKYTVSDPATGRKITVSA
MrH5b	KKYTVSHPATGRKITVSA
MrH6a	KKYTVSNPDGTKITVSA
UTH3	SEIYSNPDGTWIVTE
SG(pG-pG)	Ac-VFITSpGKTYTEVpGOKILQ-NH ₂
E12G	Ac-VFITSpGKTYTGvpGOKILQ-NH ₂
SG(NG-NG)	Ac-VFITSNGKTYTEVNGOKILQ-NH ₂

ronment for a given residue was then determined based on: (1) the H-bond state of the NH and carbonyl of that residue, (2) the preceding residue's backbone conformation, and (3) its carbonyl's H-bond state. Residues in which the H_N and carbonyl were involved in H-bonds to the same residue were considered to be the inwardly directed backbone amides of a sheet. *This definition is specific to antiparallel strand alignments as present in hairpins.* Outwardly directed sheet H_N sites were defined as those not involved in a hydrogen bond, but whose preceding residue was assigned as having a β -strand conformation and whose carbonyl was in an H-bond. Of the 2137 H_N 's initially identified, 1042 were classified as inwardly directed, while 482 were classified as outwardly directed. When the definition of inwardly directed was modified to exclude S + 1 sites by requiring that the preceding residue have an extended conformation, 855 inwardly directed H_N sites remained. Chemical shift deviations were calculated by first converting the observed chemical shifts to a 298 K expectation value: for outward-directed sites [$\delta^{298} = \delta^{obs} + (T^{exp} - 298) * 0.0076$], for inwardly directed sites [$\delta^{298} = \delta^{obs} + (T^{exp} - 298) * 0.0034$]. The random-coil shift for the residue was then subtracted from the corrected chemical shift to determine the CSD; neighboring residue corrections were not included.

In addition, we have examined the complete assignments of all hairpins reported out of the laboratories of M. Searle, S. Gellman, L. Serrano, M. Waters, the Madrid group (M. Rico, E. deAlba, C. Santiveri and coworkers), and cyclic peptides that

are forced to assume a hairpin conformation (Trabi et al., 2001; Gibbs et al., 2002) calculating CSDs as detailed in the next section.

CSD calculation

An updated version (Fesinmeyer et al., 2004) of our previously published (Andersen et al., 1997) method for determining random coil values is used throughout. The CSD is defined as $\delta_{\text{obs}} - \delta_{\text{ref}}$, but there are a number of alternative tabulations of reference shift values available (Wishart and Sykes, 1994; Merutka et al., 1995; Wishart et al., 1995; Andersen et al., 1997; Schwarzingler et al., 2000) and, in some cases, there is evidence that reference values need to be corrected for co-solvent and temperature. For our algorithm, we assume an alanine context (AAXAA) with near-neighbor corrections for Cys, Gly, His, Pro, Phe, Thr, Trp, and Tyr and additional corrections for temperature and solvent. A web-based program incorporating our algorithm is available for public use (<http://andersenlab.chem.washington.edu/CSDb>) and is able to both calculate and graph $H\alpha$ 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.

NOE connectivities

NOESY spectra were analyzed for each peptide, but we have not used the NOEs to derive NMR structure ensembles. In all cases we observed $i/i+2$ and/or $i/i+3$ connectivities at the turn locus. As a test of hairpin register, we obtained NOESY spectra in D_2O medium to improve the ease of observation of cross-strand $H\alpha/H\alpha$ connectivities and these confirmed the register indicated by CSD profiles.

CD spectroscopy

Far UV circular dichroic spectra were also recorded for most of the peptides in this account using methods previously described (Andersen et al., 2002). Fluoroalcohols were added in the appropriate amounts as needed using gas-tight mi-

cro-syringes. The ellipticity difference, $[\theta]_{198} - [\theta]_{216}$, in residue-molar ellipticity units ($\text{deg cm}^2 \text{ residue-dmol}^{-1}$) (Andersen et al., 1999) was used to measure β -hairpin character; some of the data has already been published (Dyer et al., 2005).

Results

We employ chemical shift deviations calculated vs. random-coil expectation values as our primary source of fold population estimates; these are calculated using an automated CSD calculation protocol (Fesinmeyer et al., 2004), see Methods. Based on a large body of data for unstructured peptides, random-coil $H\alpha$ shifts appear to be predicted to within 0.04 ppm and H_N values to within 0.08 ppm. Given that hairpin formation CSDs are on the order of 1.0 ppm (*vide infra*), our fold population estimates are likely accurate to $\pm 5\%$. The use of CSDs calculated vs. δ_{ref} -values from a control peptide bearing the same local sequence, but modified (or truncated) so as to preclude folding (Schenck and Gellman, 1998; Espinosa et al., 2001; Tatko and Waters, 2004), is also examined.

Chemical shift data for [2:2] hairpins

Our initial work with hairpins centered on MrH1, a sequence developed by Searle (Maynard et al., 1998) from the dimer interface of Met-repressor protein. We were attracted to this system by three features: (1) the sequence has only one aromatic residue (thus limiting ring current shifts), (2) the authors observed cold denaturation in water, and (3) the hairpin population increased upon alcohol addition. We confirmed these observations and found enhanced cold denaturation in 8% HFIP (Andersen et al., 1999). This allowed us to determine the thermodynamic parameters for hairpin unfolding from the temperature dependence of the chemical shifts of the inwardly directed (even-numbered S residues in Figure 1) $H\alpha$'s. Similar thermodynamic data could also be derived from either the H_N shifts of odd-numbered strand positions or the CD melts.

Issues with solubility and aggregate formation by MrH1 led us to replace the hydrophobic Ile¹⁶ with alanine and restore a charged N-terminus.

The resulting peptide (MrH3a) displayed an increased fold population and did not aggregate at NMR concentrations. Three turn mutants of MrH3a were prepared: 3b (N8**p**), 3c (N8G), 3d (N8P). In addition, 4a (Y3L), and 4b (Y3L/N8**p**) were synthesized to ascertain the extent to which ring current effects might influence backbone CSDs of the tyrosine-containing sequence. The peptides were examined in water and fluoroalcohol-containing media.

The CSDs of MrH3a and the mutants displayed a distinctive pattern around the hairpin structure. Strand alignment was clearly observable from the backbone protons as a series of positive deviations with alternating magnitude (Figure 2). The similarity between the MrH3 and MrH4 peptides indicates the pattern is not the result of ring-current effects associated with Tyr³. Fluoroalcohol addition, which has been previously shown to enhance β -hairpin structuring, (Blanco and Serrano, 1995; Ramírez-Alvarado et al., 1996; Andersen et al., 1999), or the replacement of the NG at the turn locus with **pG**, creates a more pronounced pattern of alternating CSD intensities. The D-Pro substitution should enforce a type II' turn geometry, as has been confirmed by crystallography (Karle et al., 1996). Figure 2 also includes a mutant with an L-Pro-Gly 'turn' (MrH3d), a common substitution for generating non-folding control sequences (Schenck and Gellman, 1998; Espinosa et al., 2001). Whether the turn is type-I' or -II', the inwardly directed H_N and H α protons appear further downfield. Residues in and near the turn display a distinctive pattern of deviations that may be useful as a fingerprint for turn type. Of particular interest is the negative H_N CSD observed at the S+1 position, an inwardly directed amide proton in the strand region. Upfield shifts of the S+1 H_N's have previously been noted in a three-stranded sheet (Sharman and Searle, 1998) and in designed hairpins (Griffiths-Jones et al., 1999; Tatko and Waters 2003). This shift observation can be rationalized in two ways: (1) as the result of shielding by the diamagnetic anisotropy of the T1/T2 amide bond or (2) that the amide proton is solvent sequestered (unable to form an inter-molecular hydrogen bond) but forms only a weak intramolecular bond (Andersen et al., 1997) due to local rigidity; the latter corresponds to a [2:4]- rather than [2:2]-hairpin motif.

The CSDs characteristic of hairpin formation were enhanced upon fluoroalcohol addition, indicating increased structuring. Independent of the turn sequence (**pG**, GG or NG) and the Y3L mutation, inwardly directed H_N and H α protons shift downfield, yielding larger CSD values (see Figure 2). Solvent-exposed protons were less consistent: H α CSDs generally displayed a more modest increase in magnitude, while some H_N CSDs decreased in value even though other shift measures of folding increased. This was observed with increased folding due to mutation and/or fluoroalcohol addition. As a result, we view the CSD values for inwardly directed H_N sites as the better measure of %-structuring when making comparisons, particularly across a series of solvents. The positive T1 H_N CSD appears to monitor the hairpin fold population; the CSD is comparable to the S-1 H_N (about +0.7 ppm for well-folded Ile⁷-Asn⁸ sequences). The same shift trends are observed in the shortest known hairpin sequence, BH8 (Ramírez-Alvarado et al., 1996). BH8 is hardly folded in water, but ~65% folded in 10% HFIP at 280 K (confirmed by CD, data not shown). The BH8 shift histograms (Supplementary materials) display, in exact analogy to the MrH hairpins, an upfield shift for the S+1 H_N and downfield shifts for the S-1, T1, and S+3 H_N's as well as the S \pm 2 H α 's.

In order to use CSDs as a quantitative measure of %-folded, it is necessary to establish the changes in chemical shift expected for a 0 to 100% folding transition. The 0% folded limit can be equated with established 'random-coil chemical shifts' (CSD=0) or derived from a 'control peptide' with a turn sequence that disfavors the ϕ/ψ values required for hairpin strand alignment (the L-Pro-Gly sequence in the present case). As previously noted, [2:2]-hairpins with a D-Pro-Gly turn sequence have often been viewed as 'fully-folded', particularly when the other end of the hairpin has a covalent closure constraint (Syud et al., 1999; Tatko and Waters, 2004). Indeed, the CSDs observed for **pG** turn species in water have been consistently larger in magnitude than those for corresponding NG turn species (MrH3b vs. 3a and 4b vs. 4a). However, the present **pG** turn species display larger structuring shifts upon addition of HFIP (CD spectra also indicate a larger hairpin population, data not shown). At 300 K, H α and

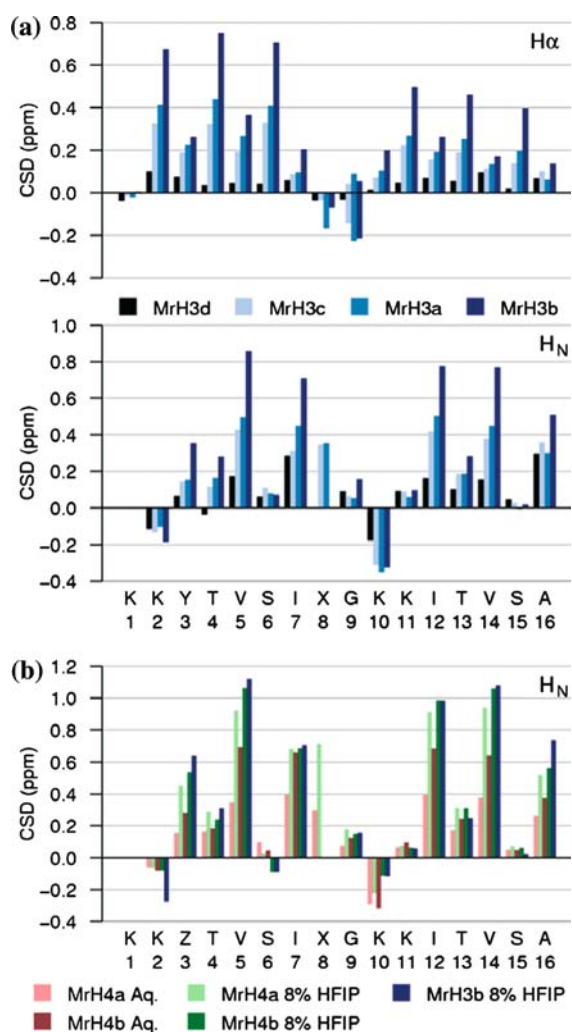


Figure 2. CSD histograms for MrH peptides at 300 K in pH 6 buffer. H_{α} and H_N CSDs for MrH3d (X=P), -3c (G), -3a (N) and -3b (p) show increasing hairpin population in that order (a). The H_{α} CSDs at the non-H-bonded sites in the C-terminal strand are consistently smaller than those in the N-terminal strand. The effect of HFIP addition to 8 vol-% upon the H_N CSDs of MrH4a (X=N, Z=L), MrH4b (X=p, Z=L) and MrH3b (X=N, Z=Y) is also illustrated (b).

H_N sites at hydrogen-bonded strand positions of MrH3b increase by $38 \pm 22\%$ over the values observed in the absence of fluoroalcohol, reaching values as high as $+0.94$ and $+1.14$ ppm, respectively, see Figures 2 and 3. Similar increases are observed for the MrH4 series and when TFE is used in place of HFIP (data not shown). Mutants that are less stable in water, including ones with an NG turn locus and/or Y3L mutation, show very

similar CSD values in the 8% HFIP medium; the Y3L mutation decreases the large CSDs of cross-strand H-bonded H_N sites by only $5.6 \pm 6.6\%$ in 8% HFIP and by $15 \pm 8\%$ in water. The ‘saturation’ of the hairpin population increases associated with HFIP addition for both NG and pG turn species with and without the Y3L mutation, Figure 2b, can only be rationalized if both pG turn species are nearly fully folded in 8% HFIP. As a result, we view CSDs on the order of 1 ppm as essentially 100%-folded values.

While MrH3d, bearing an L-Pro at position T2, served as an unfolded control above, it does display some non-zero CSDs, including some hairpin-like periodicity in the H_N CSDs (Figure 3). At low temperatures in 30% TFE, the H_{α} and H_N sites clearly show the same hairpin register observed for the stable type I'/II' turn sequences. The $\sim 55\%$ hairpin population for peptide MrH3d at 280 K in 30% TFE melts out rapidly on warming.

Application of CSD methods to hairpins with longer loops

Two longer loops are quite common in protein hairpins. Gly-bulge [3:5]-hairpins are the second most common hairpin class in protein structures (Sibanda and Thornton, 1991) and a number of peptide models of this motif have been reported using PDG as the 3-residue turn sequence. The [4:4]-hairpin class is exemplified by the second hairpin of the B1 domain of protein G (GB1p), which has been the subject of extensive study (Blanco et al., 1994; Muñoz et al., 1997; Honda et al., 2000; Fesinmeyer et al., 2004). These hairpins typically have aromatic residues at the non-H-bonded strand positions ($S \pm 4$, $S \pm 2$) and the resulting ring current shifts can obscure the ‘hairpin pattern’ of H_{α} and H_N shifts in the strands. The MrH3 strands provided a good template for turn sequence studies as they lack significant ring current effects. Incorporating the longer turns required replacement of INGK with NPDGT (de Alba et al., 1999) to give MrH6a and XPATGR to give MrH5a (X=D) and MrH5b, (X=H). The CSDs observed for strand sites indicate that both NPDGT and XPATGR loops support hairpin formation, particularly in HFIP-containing media. The alternating pattern of strand H_N CSDs (Figure 4a) establishes that these longer loops afford hairpins with the same cross-strand H-bonding register as MrH3a.

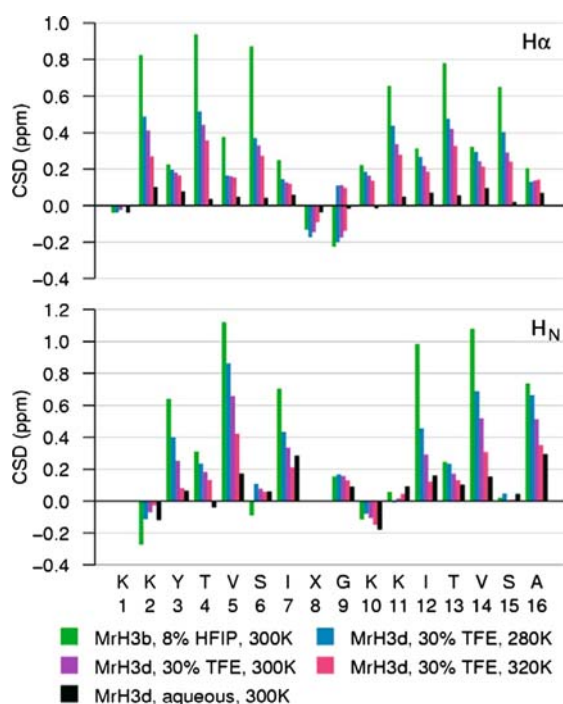


Figure 3. Comparisons of the H_{α} and H_N CSD histograms for MrH3d in water and TFE with the folded control (MrH3b, in 8% HFIP).

The magnitude of both H_N and H_{α} strand CSDs suggests that the aqueous MrH6a equilibrium ensemble contains $\sim 60\%$ of the fold population attained by MrH3a; in 8% HFIP, the constructs are almost equally folded. Based on the same criteria, [4:4]-hairpin MrH5a is present at approximately 55 and 70% in water and 8% HFIP, respectively. Peptide MrH5b was only 12–23% folded in water at 290 K, but $\sim 40\%$ folded in 8% HFIP.

The longer loops display large H_N CSDs, presumably these are diagnostic for turn type (Figure 4b). In the case of the Gly-bulge [3:5]-hairpins, the largest H_N structuring shift in the loop is the upfield shift at T2; the upfield shift at S+1 observed in [2:2] loops is retained. The generality of this observation was supported by studies of UTH3 (de Alba et al., 1999); an extended analysis appears in the Supplementary material. For six-residue loops, such as those in MrH5a and other GB1-like systems (Fesinmeyer et al., 2004), large H_N upfield shifts are observed at positions T3 and S+1. Three additional examples appear in Figure 4b. These include instances of

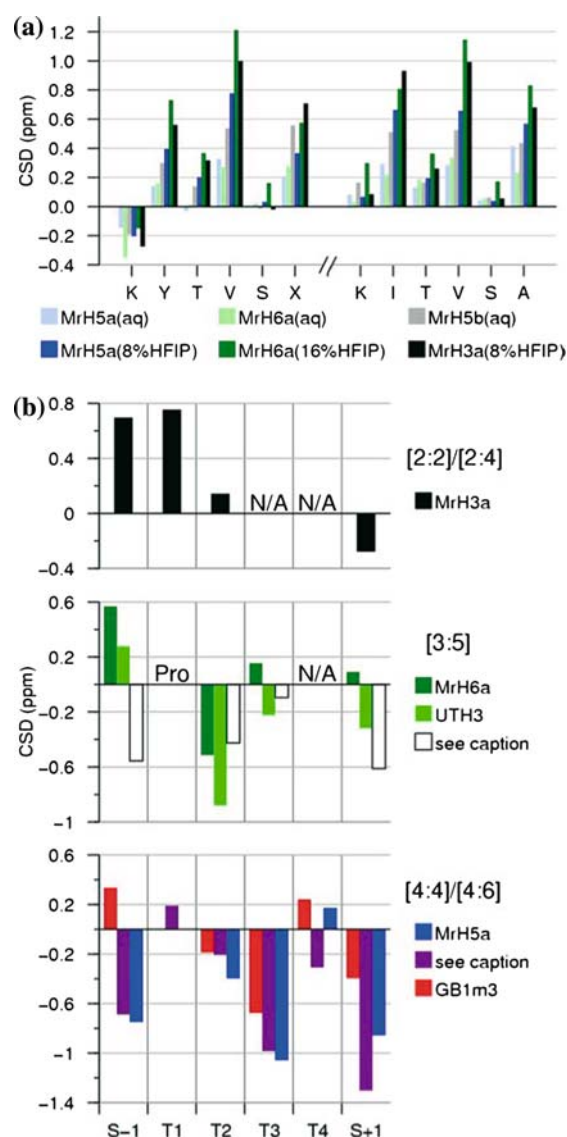


Figure 4. The H_N CSD histograms for the common strand positions of hairpin peptides with different loop lengths (a). Representative H_{α} CSDs from studies of MrH5a, 5b and 6a appear in the Supplementary Materials. The H_N shift patterns observed over the loop residues of [2:2]-, [3:5]- and [4:4]-hairpins in fluoroalcohol-containing media are also illustrated (b). The [3:5] turns are MrH6a (16% HFIP), UTH3 (30% TFE), and KTWEPDGKWTE (10% HFIP). The [4:4]/[4:6] turns are from MrH5a (in 8% HFIP), KTWNAATGKWTE (in 10% HFIP), and GB1m3 (Fesinmeyer et al., 2004) (KKWTYNPATGKFTVQE, 30% TFE). The KTWNAATGKWTE sequence is the only hairpin of that turn type to have an H_N at T1.

exaggerated upfield shifts at S-1 sites, which are observed only when the S-2 residue is an aromatic amino acid.

Application of chemical shift data to 3-stranded sheets

The Schenck-Gellman three-stranded sheet model (Figure 1) displays significant increases in β structure for NG \rightarrow pG mutations at the turn loci and further increases in β structure with the addition of HFIP are observed (Figure 5). The SG(pG-pG) construct is, in water, significantly better folded than SG(NG-NG) in 8% HFIP, but still shows an increase in folded population upon addition of HFIP. All of the significant H_N and H_α CSDs are those that would be expected for [2:2] hairpins: the T1 H_N 's (See also Figure 6) are downfield, the S+1 H_N 's (Lys⁸ and Orn¹⁶) are shifted upfield, and all of the cross-strand directed backbone protons, whether H_α or H_N , display CSD increases for the NG \rightarrow pG mutation.

Our series of mutants included SG(DG-pG) which has been reported to display a [3:5]-hairpin preference due to the mutation in the first turn (Chen et al., 2001). Like others (Schenck and Gellman, 1998; Chen et al., 2001; Syud et al., 2003) we observed the expected cross-strand

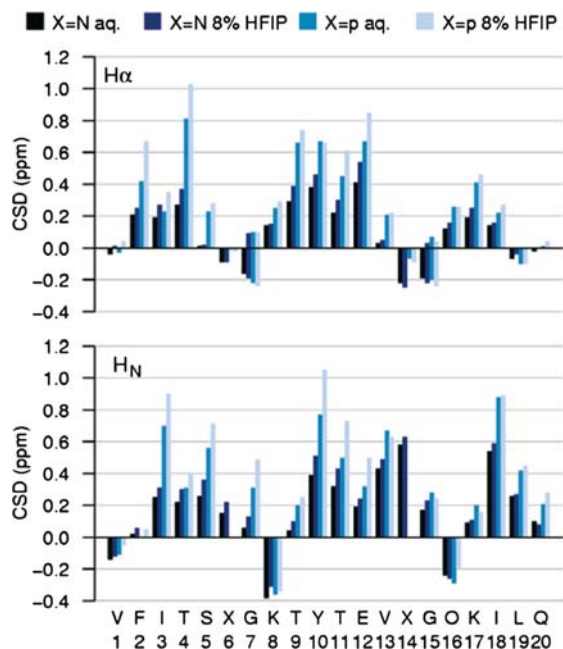


Figure 5. H_α and H_N CSDs observed for the NG-NG turn mutant and the original SG hairpin in water and 8% HFIP at 290 K. The upfield shifts at the turn glycines (G7 and G15) correlate with strand CSD measures of hairpin population.

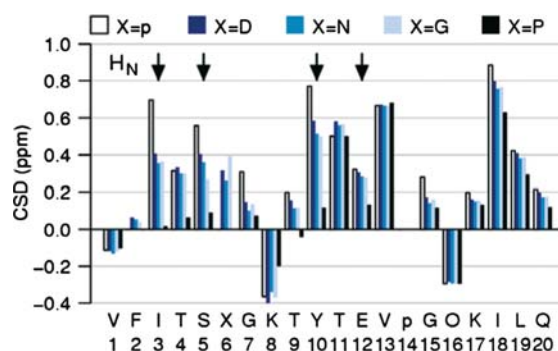


Figure 6. The effect of mutations (D-Pro to Asp, Asn, Gly, and L-Pro) at the first turn locus on the H_N CSDs of the Schenck-Gellman three strand sheet model. All data are for aqueous buffer at 290 K. The arrows indicate H_N sites that are downfield due to H-bonding interactions between the N-terminal and middle strands.

NOEs ($2\alpha/11\alpha$, $3N/11\alpha$, $3N/10N$, $4\alpha/10N$) for SG(pG-pG), see Figure 1. In both the (NG-pG)- and (DG-pG)-mutants, these NOEs were evident to the extent allowed by chemical shift dispersion, which varies in different analogs, but of reduced intensity relative to comparable cross-strand NOEs for the second hairpin (e.g. $10\alpha/19\alpha$ and $12\alpha/17\alpha$). However, the $1\alpha/11\alpha$, $2N/11\alpha$, $2N/10N$, $3\alpha/10N$, and $3\alpha/9\alpha$ connectivities expected for the strand alignment in a [3:5]-hairpin with an SXG turn sequence were not evident in our spectra. The pattern of CSDs provides additional evidence for a common turn through the series. As can be seen in Figure 4b, the central Asp at [3:5]-hairpin turn loci displays a large upfield H_N shift (-0.6 to -1.0 ppm); the turn Asp of SG(DG-pG) displays a positive H_N CSD as was observed for all of the [2:2]-hairpins in this study. Thus, all evidence supports a common [2:2]-hairpin topology for both hairpins independent of the identity of X in these XG turns.

Determining the extent of folding about the two turn loci

Even under the most fold-stabilizing conditions, signals from α protons in non-hydrogen bonded positions along the 3rd strand are distinctly less shifted than those in the central strand or the comparable positions in the 1st strand, a trend observed in other three-stranded sheet models (López de la Paz et al., 2001). In contrast, the CSDs for the cross-strand directed amide protons display similar downfield shifts for strands 1 and 3.

Table 1. Hairpin populations and thermal stability of Schenk-Gellman β -sheet model mutants. All CSD measures of fold population are from data at 290 K with the values for the **pG-pG** form in 8% HFIP arbitrarily set to 100% for this tabulation^a

cmpd	solvent	H1 %	H1melt ^b	H2 %	H2melt ^b	G7 $\Delta\alpha$	G15 $\Delta\alpha$	X6H _N	X14H _N
pG-pG	8% HFIP	100	0.04	100	0.06	0.331	0.272	<i>n.a.</i>	<i>n.a.</i>
NG-pG	8% HFIP	65 (8)	0.03	92 (4)	0.11	0.353	0.243	+0.409	<i>n.a.</i>
GG-pG	8% HFIP	59 (7)	0.10	91 (2)	0.12	0.267	0.234	+0.538	<i>n.a.</i>
pG-GG	8% HFIP	89 (2)	0.19	76 (5)	0.29	0.300	0.278	<i>n.a.</i>	+0.691
NG-NG	8% HFIP	43 (8)	0.32	62 (6)	0.38	0.272	0.245	+0.216	+0.634
pG-pG	Water	77 (9)	0.34	87 (14)	0.19	0.320	0.265	<i>n.a.</i>	<i>n.a.</i>
E12G ^c	Water	51 (5)	0.25	38 (11)	0.31	0.157	< 0.03	<i>n.a.</i>	<i>n.a.</i>
DG-pG	Water	53 (7)	0.25	84 (5)	0.14	0.275	0.239	+0.313	<i>n.a.</i>
NG-pG	Water	46 (7)	0.24	80 (5)	0.13	0.212	0.195	+0.261	<i>n.a.</i>
GG-pG	Water	44 (6)	0.37	80 (5)	0.20	0.185	0.185	+0.391	<i>n.a.</i>
GP-pG	Water	7 (8)	<i>n.a.</i>	67 (4)	0.12	<i>n.a.</i>	0.187	+0.048	<i>n.a.</i>
PG-pG	Water	10 (7)	<i>n.a.</i>	68 (3)	0.13	<0.03	0.213	<i>n.a.</i>	<i>n.a.</i>
pG-GG	Water	77 (3)	0.21	66 (4)	0.27	0.240	0.222	<i>n.a.</i>	+0.635
pG-PG	Water	64 (3)	0.25	24 (19)	<i>n.a.</i>	0.218	0.037	<i>n.a.</i>	<i>n.a.</i>
NG-NG	Water	33 (6)	0.29	50 (8)	0.24	0.197	0.201	+0.148	+0.578
NG-pG	30% TFE	63 (12)	0.45	93 (6)	0.18	0.373	0.222	0.393	<i>n.a.</i>
GG-pG	30% TFE	33 (14)	0.57	79 (11)	0.25	0.224	0.256	<i>n.o.</i>	<i>n.a.</i>
GP-pG	30% TFE	<11	<i>n.a.</i>	79 (7)	0.23	<i>n.a.</i>	0.250	0.063	<i>n.a.</i>
PG-pG	30% TFE	<10	<i>n.a.</i>	81 (10)	0.18	<0.03	0.233	<i>n.a.</i>	<i>n.a.</i>

^a The percent folding of hairpin 1 is based on the fractional CSDs for F2 α , T4 α , T9 α , I3N, S5N, Y10N. The measure for hairpin 2 is based on Y10 α , E12 α , K17 α , T11N, and I18N. G7 $\Delta\alpha$ and G15 $\Delta\alpha$ are the diastereotopic chemical shift differences for the glycines at T2 positions. The Gly $\Delta\alpha$ data is differentiated based on the identity of the turn residues – with **pG** turns shown in bold.

^b Hairpin melting tendencies (H1 and H2melt) are calculated as [CSD(280 K) – CSD(310 K)] / CSD(280 K) and employ only the H α CSDs.

^c E12G mutation was performed on SG(**pG-pG**).

Though there are notable differences in the magnitude of the structuring CSDs, fold populations for the 1st and 2nd hairpins based on the appropriate H α and H_N CSDs are in excellent agreement. The individual hairpin (N- and C-terminal) populations derived for a series of SG turn mutants appear in Table 1. The primary measures of hairpin fold populations are based strictly on cross-strand directed H α and H_N backbone sites; the fold populations were calculated for each probe site. The small standard errors indicate that individual H α and H_N probes provide quite similar population estimates.

From Table 1, it is apparent that replacing **pG** with **PG** or **GP** at either turn locus reduces CSD measures of strand alignment about the altered turn to near zero. Other turn replacements (NG, DG and GG) produce an intermediate level of hairpin destabilization. When the same turn sequence is used at both loci and assuming the CSDs observed in the third strand, including the smaller H α CSDs, are proportional to fold populations,

the second hairpin is the better folded and displays less unfolding upon warming. To convert the changes in fold populations to $\Delta\Delta G$ values we assume that SG(**pG-pG**) in 8% HFIP is 92% folded, rather than 100%, at 290 K. The effects of the E12G mutation provide an example: it has a more dramatic destabilizing effect on the second ($\Delta\Delta G_F = 4.7$) than first ($\Delta\Delta G_F = 2.1$ kJ/mol) hairpin. This may reflect the disruption of a potentially attractive, cross-strand E12/K17 Coulombic interaction.¹ The added backbone flexibility of a glycine in the middle strand would be expected to reduce the stability of both hairpins. We cannot, however,

¹ We observe very little effect of sidechain carboxyl protonation (pH 6 vs. 3) on the folded fraction about turn 2 when there is a **pG** turn locus: this likely reflects the high propensity to form a stable hairpin alignment about a **pG** locus in turn 2. In the NG-NG mutant, ionization of the E12 sidechain increases folding about hairpin 2 ($\Delta\Delta G_U = 2.2$ kJ/mol) with very little effect (0.40 ± 0.24 kJ/mol) on the folding measures for hairpin 1.

rule out the possibility that the greater effect on the second hairpin reflects the location of the glycine relative to the two turns: residue 12 is immediately next to turn 2 but can be viewed as being in the frayable end of the N-terminal hairpin.

Data from Table 1 indicates more facile thermal melting of hairpin alignments about GG turns. Of the four analogs studied in 30% TFE, all displayed enhanced melting in 30% TFE, attributed to a more negative ΔC_{pU} from diminution of the hydrophobic effect (Andersen et al., 1996, 1999). Unlike the other analogs, GG-pG does not display TFE-induced fold increases in either hairpin. This is likely a reflection of the increased ΔS_U associated with the greater backbone flexibility at glycines. The combination of an increased ΔS_U and decreased ΔC_{pU} eliminates fluoroalcohol-induced structuring.

Since H_N CSDs are larger, and do not show the 3rd strand intensity anomaly, we use these for presenting mutational studies. The H_N CSDs for a series of 1st turn locus changes (pG, DG, NG, GG, PG) appear in Figure 6. The PG mutation serves as a control for essentially complete disruption of the first hairpin. The other three mutations show an intermediate degree of hairpin destabilization evident by similar reductions in the H_N CSDs for I3, S5 and Y10, which reflect cross-strand H-bonding between the strands that align about turn 1. Comparable changes occur at the turn locus (G7- H_N); smaller, but similarly trending changes are observed in the 3rd strand.

Discussion

The present study leaves no doubt that the backbone α and amide protons of β strands display an $i/i+2$ periodicity of downfield shifts when the strands are aligned to form a hairpin or are located at the edge of a sheet. Cross-strand H-bonding has been previously suggested as the cause of residue-specific downfield H_N shifts in β strands (Griffiths-Jones et al., 1999; Tatko and Waters, 2003, 2004) and it also appears in chemical shift calculations (Xu and Case, 2002). With regard to the $H\alpha$ CSD pattern, the greater downfield shifts of the cross-strand directed $H\alpha$'s at the non-H-bonded sites (+0.85, vs. +0.4 ppm for the H-bonded sites) have been predicted and a number of examples where this pattern is observed in protein sheets have been

noted (Osapay and Case, 1994). Searle and coworkers (Sharman and Searle, 1998; Sharman et al., 2001) have also noted this pattern in designed hairpin peptides. A survey of chemical shifts in protein sheets (Sharman et al., 2001) revealed a smaller difference, 0.3 ppm, between the average $H\alpha$ CSDs for non-hydrogen-bonded (0.51 ± 0.53) and hydrogen-bonded (0.21 ± 0.38) sites.

Our data indicates that the periodicity in strand $H\alpha$ shifts is more pronounced in designed peptide hairpins than that predicted from protein sheet analogies and increases with hairpin fold stability and, presumably, faithfulness of register. Restricting our data set to the S-5 through S-2 sites of hairpin peptides for which an extrapolation to 100% folded is possible, the expectation values are: +0.25, +1.05, +0.3, and +0.95, respectively. The same periodicity is observed in the C-terminal strand, but in all cases the maximum downfield shifts are smaller, possibly reflecting sheet twist. In the case of three-stranded sheets, the rather small $H\alpha$ CSDs observed in the terminal strand (see Figure 5a) have led other investigators to conclude that the C-terminal portion of the structure is less well formed at equilibrium; see for example an account from the Serrano lab (López de la Paz et al., 2001). The recent extension of the Gellman systems to 4-stranded sheet structures (Syud et al., 2003) indicates that this reduction in CSDs for C-terminal strands continues to be the case.

The present study reveals a richer diversity, and more diagnostic patterns, of strand and turn H_N structuring shifts for peptide sheet models with a variety of reversing turns. The strand sites that are presumed to be H-bonding ($S \pm 5$, $S \pm 3$, and $S-1$) display positive CSDs (typically 0.8–1.2 ppm) with no loss in magnitude in the C-terminal strand, while the $S+1$ H_N site is shifted upfield. The upfield shift at the $S+1$ H_N is viewed as a turn effect and its magnitude varies with turn type; the H_N patterns for a series of turn types appear in Figure 4b. Type I' [2:2]-hairpins display $S+1$ CSDs ≈ -0.4 ppm with larger downfield shifts at T1. Our observations are confirmed by literature data reported for other hairpin peptides (e.g. Sharman and Searle, 1998; Griffiths-Jones et al., 1999; Trabi et al., 2001; Gibbs et al., 2002). For [3:5]-hairpins, the T2 sites provide the largest diagnostic CSD, ≈ -1 ppm. The pattern observed for [4:4]-loops, upfield shifts at T3 and $S+1$, has

now been validated for a wide variety of loop sequences (15 to date), including HPATGR, Asx-PAAGT, Asx-XXTXX, and NGGTGK, where X can be numerous other residues (data not shown). The upfield shift at S+1 approaches 2 ppm when the S+2 site is an aromatic residue.

Can these H_N CSD observations be extended to antiparallel strand alignments in proteins? We employed the same set of protein structure/chemical shift data employed in the $H\alpha$ survey of Sharman et al. (Sharman et al., 2001). H_N CSDs were calculated to be $+0.63 \pm 0.67$ ppm for hydrogen-bonded positions and $+0.43 \pm 0.66$ ppm ($n = 1042$) for non-hydrogen-bonded positions. When the definition of H-bonded sites excluded those not preceded by a residue in an extended conformation, the CSD increases to $+0.75 \pm 0.59$ ppm ($n = 855$). The CSD for the excluded positions was $+0.10 \pm 0.77$; this new category should include S+1 sites in protein sheets. The CSD difference between the two primary types of strand H_N 's is 0.3 ppm, the same as was reported for the $H\alpha$ strand types, rather than the larger value (0.5–0.85 ppm) we observe for isolated hairpins. We view the hairpin data as the better guide for expectation values; the smaller $\Delta\delta$ values (and large standard deviations) from protein data libraries likely arise from diamagnetic anisotropy effects (for example ring currents, which would be more common in proteins) and H_N -type classification errors.

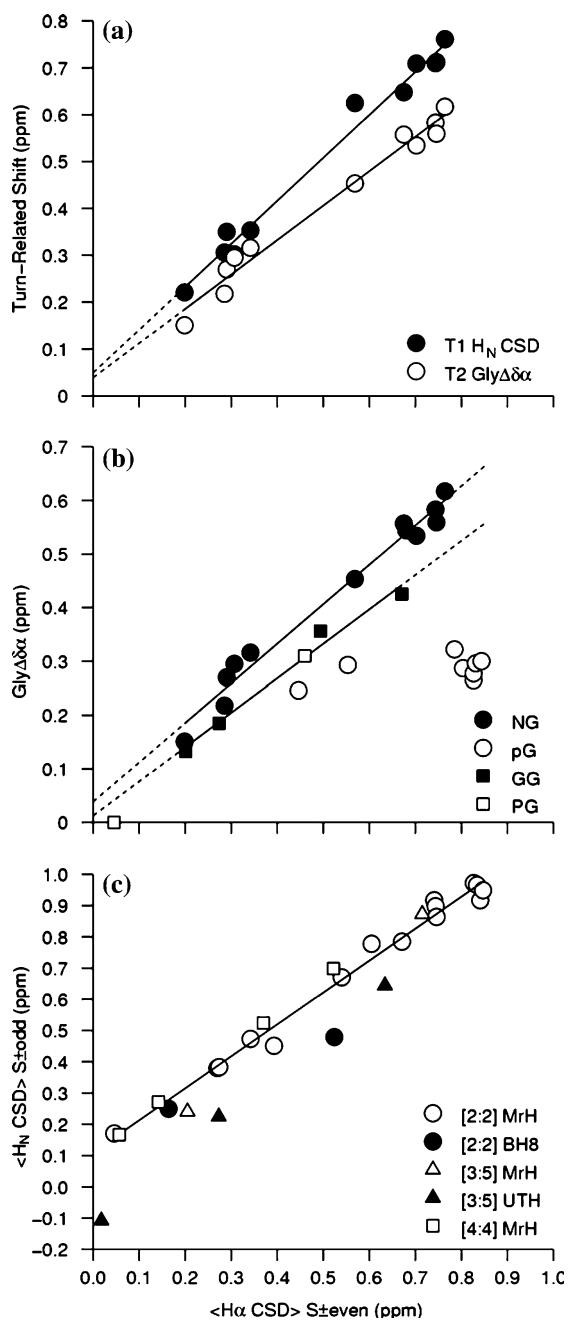
Further examination of shift/topology correlations in protein hairpins was limited to spot checks of hairpins and sheet edges in specific small proteins (<90 residues). In those examined, the shift periodicity more nearly approximates that observed in the peptide models. When the H_N CSD values are binned, the S±3 sites fall predominantly in the +0.8 to +1.2 category. The negative CSD at S+1 is also observed, particularly with [4:4]-hairpins. These also display the upfield T3 shift of the peptide models; the T3 and S+1 H_N 's often shift >1 ppm upfield. Excluding data from [4:4]- and [3:5]-hairpins, the S+1 H_N CSDs range from -0.7 to +0.1 ppm with a typical value of -0.4 ppm. With the evidence provided in the present study, there is a sound basis for the use of H_N CSDs for determining the register and typology of β hairpins.

The correlation of increased CSDs (downfield shifts for amide NHs) with tight H-bonding has been established for nearly two decades (Wagner

et al., 1983; Wishart et al., 1991). A literature account suggests that CSDs on the order of +0.9–1.2 in β -sheets imply $H_N \cdots O=C$ distances of 1.9–1.75 Å (Cierpicki and Otlewski, 2001). Thus it appears that well-populated β hairpins have cross-strand H-bond lengths that are as short as those observed in protein β sheets.

Correlations between, and the relative merits of, alternative chemical shift measures of peptide hairpin populations

A number of $H\alpha$ chemical shift measures of hairpin fold populations have been suggested in the literature. In our studies (Andersen et al., 1999; Fesinmeyer et al., 2004), we have urged the use of the larger positive CSDs observed for inwardly directed $H\alpha$ sites in the strands (notably at the S±2 and S±4 residues) and these have also been used by the Searle group (Griffiths-Jones et al., 1999). Two other measures have been more widely used: the chemical shift difference ($Gly\Delta\delta\alpha$) for the two $H\alpha$'s of a glycine at position T2 in [2:2]-hairpins (e.g. Griffiths-Jones et al., 1999; Griffiths-Jones and Searle, 2000; Kiehna and Waters, 2003; Santiveri et al., 2004; Hughes and Waters, 2005), and the smaller $H\alpha$ CSDs for H-bonded strand sites, such as positions S±3 and S±1 (Espinosa and Gellman, 2000; Syud et al., 2001; Syud et al., 2003). To these, we can now add the H_N CSDs at H-bonded strand sites and possibly H_N CSDs at specific turn positions. When very stable hairpin models are available, we have established that all CSD measures, whether due to secondary structure, ring currents, or a combination of these effects, give comparable estimates for less stable analogs. The different measures also display identical melting behavior, as would be expected for a strict two-state folding transition (Fesinmeyer et al., 2004). Most designed hairpins are significantly less stable; what chemical shift measure of folding should be used for such systems? Do the different measures suggested give similar results? To this end we have examined the cross correlation of CSD-based folding measures over the set of hairpins prepared for this study (Figures 7 and 8). We also address an additional question: whether coil shift-reference values or specifically designed unfolded controls provide more reliable reference values for the calculation of CSDs.



It could be argued that $\text{Gly}\Delta\delta\alpha$ and the $S+1$ H_N CSD might reflect the stability of the turn rather than the overall stability of the hairpin. Several groups, (e.g. Griffiths-Jones et al., 1999; Santiveri et al., 2002), have suggested that turn regions of hairpins retain structure in truncated forms and upon partial melting of strand alignment. The Waters laboratory, which uses only the

Figure 7. The correlation between alternative CSD structuring measures and the average CSD for inwardly directed strand $H\alpha\alpha$ sites. The common axis is labeled “ $\langle H\alpha \text{ CSD} \rangle_{S \pm \text{even}}$ ”, the average non-H-bonded $H\alpha$ CSD. In panel C, less than four sites are averaged: $H\alpha$ at $S+2$ and $S\pm 4$ for the [4:4]-hairpin and UTH3, only $H\alpha$ $S-2$ for the short BH8 sequence. The illustrated correlations are: (a) turn-associated shift measures ($\text{Gly}\Delta\delta\alpha$ and the T1 H_N CSD) vs. the average non-H-bonded $H\alpha$ for MrH3a and MrH4a under a variety of conditions, (b) turn residue dependence of $\text{Gly}\Delta\delta\alpha$ for [2:2]-hairpins with different residues at the turn locus (residues T1 and T2 are listed); the linear regression lines are for the NG and GG turn species, and (c) H-bonded H_N sites in the β strands of hairpins. The sites included in the H_N CSD averages are: for [2:2]- and [3:5]-hairpins with the MrH strands (H_N at $S-1$, $S\pm 3$, and $S+5$), for the [4:4]-hairpin and UTH3 (H_N at $S\pm 3$, and $S+5$), and for BH8 (H_N $S-1$ and $S+3$). The regression line is for the [2:2]-hairpins of the MrH series.

$\text{Gly}\Delta\delta\alpha$ -measure for deriving melting curves and folding transition thermodynamic parameters, has noted that this measure “consistently overestimates hairpin populations in peptides ... [particularly for] the D-Pro-Gly turn” (Kiehna and Waters, 2003). Over the MrH series of peptides (for the NG- and GG-turn species), both $\text{Gly}\Delta\delta\alpha$ and the upfield shift of the T1 H_N correlate with the CSDs for the strand $H\alpha$ sites that display the largest structuring shifts (Figure 7a). Since both correlation lines pass very close to the origin, there appears to be no residual structuring of the turn in the absence of cross-strand interactions. Given the

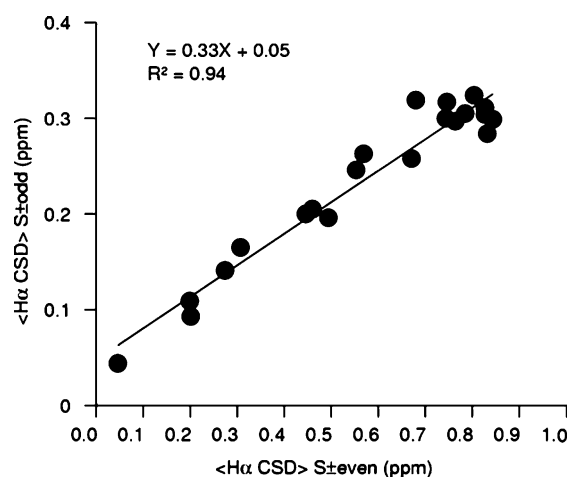


Figure 8. The correlation of “structuring” $H\alpha\alpha$ CSDs at H-bonded residue sites ($S\pm 3$ and $S+1$) with the larger $H\alpha\alpha$ CSDs observed at the non-H-bonded sites ($S\pm 2$ and $S\pm 4$) that measure strand alignment.

excellent correlation of these strand $H\alpha$ sites with the shifts of the H_N sites that form cross-strand H-bonds (Figure 7c), turn formation in the absence of full hairpin-like features appears highly unlikely in these hairpins. An additional conclusion is that T1 H_N CSDs for complete hairpin formation are comparable to the $H\alpha$ CSDs at the non-H-bonded sites in the strands for type-I β turns. When the Gly $\Delta\delta\alpha$ correlations for NG, GG, PG, **pG** turns are examined separately (Figure 7b), it appears that the maximal value depends on turn residue identity. In the case of D-Pro-Gly turns, it is clearly smaller and residual turn structuring may be present in the absence of strand alignment.

In the case of correlations between inward $H\alpha$ CSDs and inward (H-bonded) H_N sites in the strands, we have extended the dataset beyond the MrH series of hairpins (Figure 7c). Our analysis includes: (1) [2:2] hairpins of the MrH series, (2) the shorter [2:2] hairpin (BH8), (3) two [3:5]-hairpins (UTH3, MrH6a), and (4) two [4:4] hairpins (MrH5a and -5b). We have included data in media with and without fluoroalcohol (HFIP or TFE) cosolvents. In order to avoid potential problems associated with differences in the intrinsic temperature gradients of H_N shifts, all of the observations were for data collected at 290–300 K, closely bracketing the temperature for which H_N random-coil reference shifts have been calibrated. Independent of the turn type, the same linear correlation ($R^2=0.985$, slope=1.04) was obtained when the strands were constant. Systems with different strand residues deviate somewhat from this line but still display a clear correlation between H_N and $H\alpha$ CSDs ($R^2=0.93$, slope=1.08). The non-zero intercept in Figure 7c suggests that H_N sites in these strand sequences, which favor extended β conformations even in the absence of hairpin formation, display a small downfield shift associated with this local conformational preference. Given that even the most-folded systems are likely to have unfolded populations on the order of 0.1–0.15 at these temperatures, Figure 7c suggests that 100% folded values for inwardly directed $H\alpha$ and H_N sites in hairpins are on the order of 1.0 and 1.15 ppm, respectively. Turn site H_N 's also achieve structuring shifts of this magnitude in all fully validated hairpin structures. We suggest that these shifts can serve as 100%-folded expectation val-

ues for other hairpins as well and that any report of a designed hairpin that is said to be well structured but does not display at least some $H\alpha$ and H_N CSDs approaching 1 ppm should be viewed as suspect.

In analogy to β hairpins, well-folded 3-stranded sheet models also display 1 ppm downfield shifts for inwardly directed $H\alpha$ and H_N sites and smaller $H\alpha$ CSDs in the C-terminal strand. The extensive series of SG 3-stranded sheet analogs prepared for the present study presents additional tests of the relative value (and correlation between) different chemical shift measures of hairpin fold populations (Table 1). If we once again assume that SG(**pG-pG**) shifts in 8% HFIP represent 92% folding, the tabulated relative folding estimates can provide expectation CSDs for 100% folding at other sites. Using this method, the 100%-folded T1 H_N CSDs are +0.96–1.04 for GG turn loci and vary considerably more for DG and NG but always exceed +0.61 ppm. The 100%-folded value for the Gly $\Delta\delta\alpha$ measure at XG turns was dependent on the identity of X: NG (0.57 ± 0.10 , $n=7$), DG (0.57), GG (0.49 ± 0.14 , $n=5$), and **pG** (0.31 ± 0.04 , $n=15$). These should be compared to, for example, the largest values observed in the MrH series at 17–27 °C (Figure 7b): NG (0.62) and GG (0.50), and **pG** (0.30 ppm). As a further comparison we note that the Waters laboratory uses a Gly $\Delta\delta\alpha$ -value of 0.45 ppm as the 100% folded reference value for NG turns (Tatko and Waters, 2004) and a smaller value, ~ 0.25 ppm (Kiehn and Waters, 2003), for a series of **pG** turn species. The smaller Gly $\Delta\delta\alpha$ value at 100% folded found for D-Pro-Gly turns in all of these studies may represent either the switch from a type-I' to -II' turn or a more specific effect of the neighboring proline; this remains to be determined.

The Gellman group (Espinosa and Gellman, 2000; Syud et al., 2001, 2003; Espinosa et al., 2002) uses the CSDs of the outwardly directed $H\alpha$'s of the H-bonded strand sites (particularly the S-3, S+1 and S+3 positions) to measure both hairpin and β sheet populations. Because of the smaller CSDs of outwardly directed $H\alpha$'s, the accuracy of the reference shifts becomes a significant issue; as a result, more specific reference sequences, including L-Pro 'turn mutants', may be required for shift reference values. In hairpins that bear little resemblance to the Gellman motifs (the MrH series), we find that the correlation between

the $H\alpha$ CSDs at $S\pm 3$ and $S+1$ sites and those at the $S\pm 2$ and $S\pm 4$ sites to be quite good (Figure 8), though not as good as that between inwardly directed $H\alpha$ and H_N shift data.

The non-zero y -intercept in Figure 8 corresponds closely to the average value of these sites in the L-Pro-Gly mutants of the species included in the graph. This suggests that there is a basis for using more specific control peptides as the reference compounds for $H\alpha$ shifts at H-bonded strand sites. A more detailed analysis (Supplementary material) shows that this is certainly the case for mutants of the Schenk-Gellman three-stranded sheet model. In the three-stranded sheet models, we find that the $S+1$ and $S\pm 3$ sites, but not the $S-1$ sites, track the estimates based on the larger CSDs at non-H-bonding sites *and* that use of the L-Pro substituted references does improve the agreement.

L-Pro substitutions at turn loci provide reference values for both MrH β hairpins and the SG 3-stranded sheet models in water and 8% HFIP. However, for both systems the data in 30% TFE suggests that these are not useful random-coil references. Some helicity appears to be induced in some 3-stranded sheet models and data for peptide MrH3d (with an IPGK loop) clearly shows hairpin populations approaching 0.55 in 30% aqueous TFE.

Further analysis of 3-stranded sheet mutants

For this analysis, we maintain an analogy to Gellman's method, assuming that $pG \rightarrow PG$ mutations at each turn should set the fold populations to null about the mutated turn locus, but we rely on the folding measures in Table 1, which do not include $H\alpha$ sites in H-bonded residues, and the prior calibration (SG(pG - pG) in 8% HFIP equals 92% folded) for calculating ΔG 's. The ranking of turn propensities was $pG \gg DG > NG \approx GG \gg PG$ and, to the extent examined, it applies to both turns. The destabilization associated with both NG and GG units at turn 1 are 2.8 and 3.06 kJ/mol in water with a smaller effect (1.6 kJ/mol) noted for the $pG \rightarrow GG$ mutation at turn 2. In 8% HFIP the differences are accentuated: 4.95, 5.45 and 3.86 kJ/mol, listing them in the same order, and the NG-NG double mutant shows greater hairpin alignment loss in hairpin 1 than 2 (6.9 vs.

5.2 kJ/mol). By all criteria available, hairpin 2 was found to be more robust, displaying less thermal fraying and somewhat smaller destabilization effects with mutations at that turn locus.

In the absence of HFIP, the data for the NG-NG mutant indicates a smaller difference in hairpin fold loss, 4.14 and 3.72 kJ/mol for the N- and C-terminal hairpins, respectively. The observation that each hairpin is comparably destabilized in water should not be viewed as evidence for a strict two-state, rather than four-state, folding equilibrium between a coil and a three-stranded sheet (Searle, 2001; Syud et al., 2003). In the original report (Schenck and Gellman, 1998) on the three-stranded sheet model, the authors concluded that there was a significant degree of cooperativity in the formation of the 3-stranded sheet. The key observations were changes in CSDs in the PG- pG and pG -PG mutants vs. the pG - pG form. Based on a number of strand $H\alpha$ sites (8 for the first hairpin, 7 for the second hairpin, including both H-bonded and non-H-bonded sites), the disruption of either hairpin by the D-Pro to L-Pro mutation significantly decreased the fold population of the hairpin formed about the unaltered turn. In subsequent reports, using $H\alpha$ sites in pairs of H-bonded residues for fold quantitation, the increase in %-fold associated with a first-turn $PG \rightarrow pG$ mutation corresponded to a 1.7 kJ/mol stability increase; a similar increase of 1.55 ± 0.05 kJ/mol was observed for the same mutation in a four-stranded antiparallel sheet model (Syud et al., 2003).

The SG(XG- pG) series examined herein provides some evidence for cooperativity. The fold population of hairpins 1 and 2 decrease congruently as the turn propensity of the first turn is reduced (Figure 6), but the effects are greatly diminished at the C-terminal hairpin. For example, the $pG \rightarrow PG$ mutation at turn locus 1 (a turn propensity difference of 2.8 kJ/mol) produces a muted 0.9 kJ/mol destabilization of the 2nd hairpin. Complete disruption of the 1st hairpin by $pG \rightarrow PG$ mutation results in a 2.1 kJ/mol loss in stability at the 2nd hairpin. Additional data is required for a four-state analysis that defines the extent of cooperativity, but the current data clearly requires an equilibrium ensemble containing significant population of a state in which only the C-terminal hairpin is formed. This, and the greater robustness of the 2nd hairpin noted above, are in agreement with a recent replica-exchange MD study of SG(pG - pG) (Roe et al., 2005).

Turn/loop sequence effects on hairpin fold stability depend on the locations of the hydrophobic residues in the β strands

Based on relative stabilities of trpzip (Cochran et al., 2001; Blandl et al., 2003) and GB1 hairpin analogs (Fesinmeyer et al., 2004) the following order of turn preferences emerges: EpNK & AsxPATGK > ENGK \geq DDATAKT > EGNK > EPDNGK > EPDNK. These hairpin systems bear aryl and hydrophobic residues at the $S \pm 2$ and $S \pm 4$ positions. The present study includes an extensive series of loops connecting identical strands (KKYTVS-loop-KITVSA) in which the aryl and hydrophobic residues are at H-bonded sites ($S \pm 3$ and $S \pm 5$). In distinct contrast to the $S \pm 2/S \pm 4$ hydrophobic cluster series, the relative hairpin stabilities in water were: IpGK > XNGK (X = S or I) \geq IGGK > NPDGT > DPATGR > HPATGR > IPGK. The hairpin populations at 290 K range from >90% to \sim 10%, $\Delta\Delta G_F = 11$ kJ/mol for this series. Even the worst two loops of this series can, however, support hairpin populations >40% upon addition of fluoroalcohols. Apparently, the optimization of loop geometry presents different requisites when the hydrophobic residues are at the H-bonded strand sites (as in the MrH series) rather than non-H-bonded sites (the GB1 and trpzip hairpins).

Melting behavior of MrH series hairpins

Monotonic loss of structure is seen in the CSDs of MrH3d and BH8. This is observed for all of the hairpins and three-stranded sheets reported in this paper when the experiments are carried out in 20% HFIP or 30% TFE. In the absence of co-solvent addition and in 8% HFIP, the CSDs reach their maximal values at intermediate temperatures rather than at the lowest temperature examined. Cold denaturation of a hairpin was first noted in the CD spectra of MrH1 recorded in both aqueous methanol and HFIP (Maynard et al., 1998; Andersen et al., 1999) and has since been observed using NMR folding measures for other hairpins (Searle, 2001; Kiehna and Waters, 2003; Fesinmeyer et al., 2004; Tatko and Waters, 2004). Examples of CSD profiles reflecting cold denaturation appear in Figure 9: cold denaturation is observed in both water and

in 8% HFIP. Contrasting data obtained in 20% HFIP and 30% TFE, reflecting monotonic melting, is shown for comparison.

The structural features responsible for cold denaturation remain unknown. MrH4a and 4b in water display an *increase* in fold population with temperature throughout the 280–320 K range (Figure 9). The largest decreases in fold population with cooling are observed for MrH3c and MrH4a in 8% HFIP. The MrH4 examples, with a Y3L mutation, establish that aryl sidechains are not required in the β strands for the observation of cold denaturation: aliphatic hydrophobic effects may even exceed those of aryl groups.

Modeling H_N shift changes during melting transitions

H_α CSDs were used to monitor melting in Figure 9, the use of H_N CSDs for melting studies is complicated by the fact that H_N chemical shifts display large temperature gradients even when there is no change in structure. These gradients

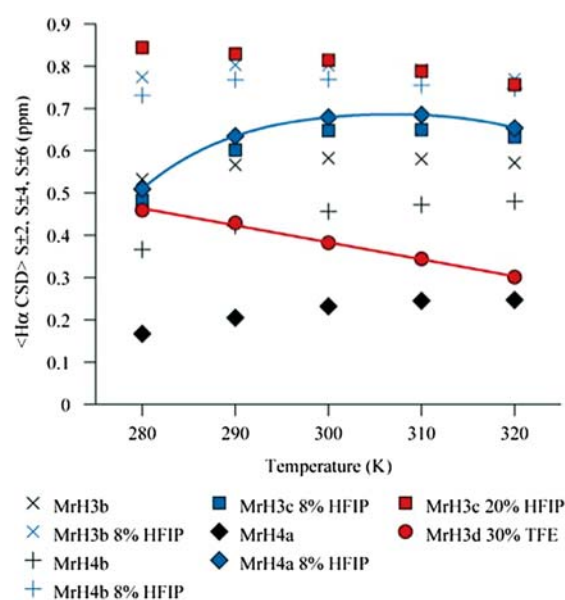


Figure 9. Examples of hairpin cold denaturation evidenced by H_α CSD data: data for aqueous buffer conditions are shown in black, 8% HFIP data in blue. This melting data is contrasted to that observed in 30% TFE and 20% HFIP (red symbols) for the same systems where there is no evidence of cold denaturation. For each system (labeled on the figure), the average H_α CSDs (over the six cross-strand directed sites: $2H_\alpha$, $4H_\alpha$, $6H_\alpha$, $11H_\alpha$, $13H_\alpha$, and $15H_\alpha$ at 280 through 320 K are shown.

$(\partial\delta_{\text{NH}}/\partial T)$ depend on whether the NH is available for, or sequestered from (typically by intramolecular H-bonding), H-bonding interactions with water (Andersen et al., 1997). In the original CSDb program, H_N CSDs were calculated by subtracting a temperature-corrected random-coil shift from the observed shift. The temperature correction applied was $\partial\delta_{\text{NH}}/\partial T = -7.6$ ppb/ $^{\circ}\text{C}$, assuming that the reference value applies at 25 $^{\circ}\text{C}$. For a peptide involved in a fold/coil transition with a temperature-dependent ΔG_F , this method gives CSDs that are directly proportional to fraction folded only for H_N 's that have the same $\partial\delta_{\text{NH}}/\partial T$ in the folded state. When we applied this CSD calculation method to hairpin peptides that are significantly folded and melt incompletely over the temperature range observed, the calculated CSDs for cross-strand, H-bonded H_N 's would frequently display no apparent melting even though the $H\alpha$ CSDs gave a clear melting curve. As a result, we created two versions of the CSD calculation program (v2003S and v2003X, for sequestered and exposed

sites, respectively). The version for sequestered sites employs a smaller correction, $(\partial\delta_{\text{NH}}/\partial T)_{\text{ref}} = -3.6$ ppb/ $^{\circ}\text{C}$. By examining which program affords, for each H_N , a CSD melt concordant with the $H\alpha$ CSD melt, one can determine whether an H_N site is sequestered or exposed in the folded state.

The differences in CSD melts that result are illustrated (Figure 10, top panel) for MrH4a in 8% HFIP, a system that displays cold denaturation based on $H\alpha$ CSDs (Figure 9) and CD (Dyer et al., 2005). Both measures indicate maximal folding at 300–310 K. When CSDb.v2003S is employed, the H-bonded $S_{\pm\text{odd}}$ H_N sites replicate this temperature profile.

A similar reading of the CSDs at the $S_{\pm\text{even}}$ sites “suggests” that only melting occurs, with some sites “melting” past zero. With CSD.v2003X, the $S_{\pm 4}$ sites gave the known dependence of χ_F with temperature and the instances of CSDs passing through zero nearly disappear.

If two-state folding can be assumed, it is possible to simulate the temperature dependence of

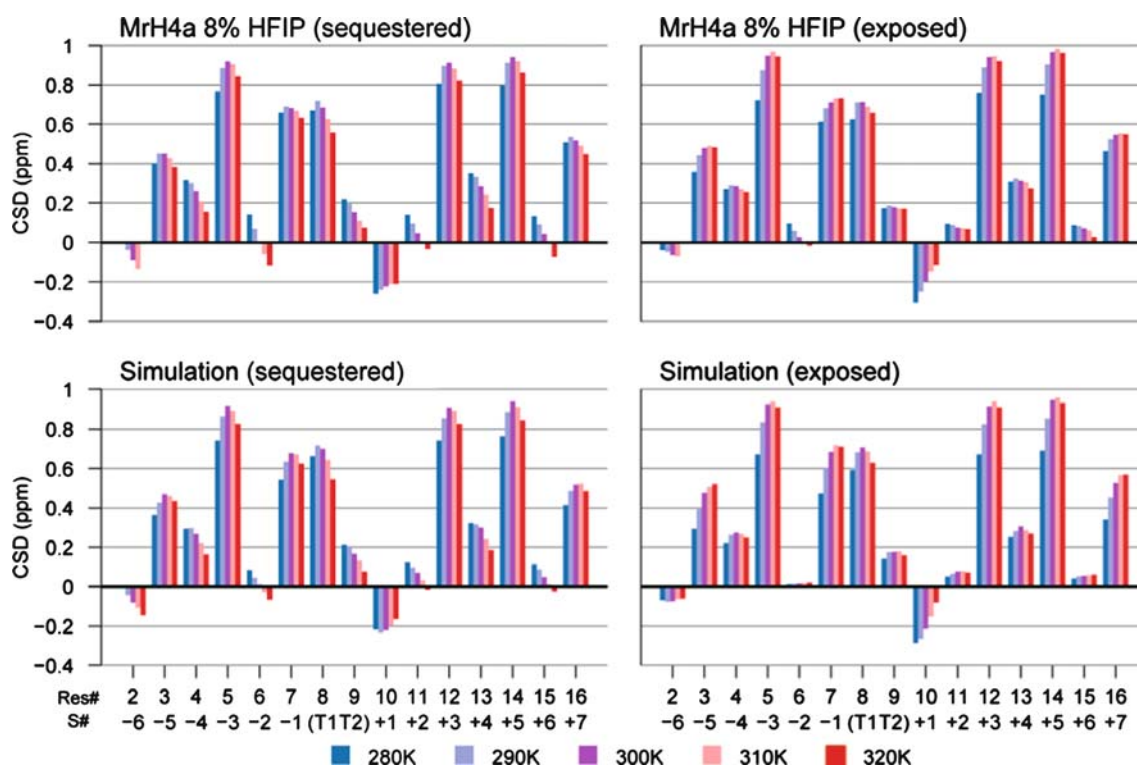


Figure 10. Comparison of CSDs derived from experimental and simulated H_N shift data for MrH4a in 8% HFIP. The upper traces show the ‘experimental’ temperature profile of CSDs at each site as displayed by the ‘sequestered’ and ‘exposed’ site CSDb algorithms. The lower panels show the comparable CSDb outputs for simulated H_N shifts based on a two-state folding assumption with the $S_{\pm\text{odd}}$ H_N sites assigned folded-state temperature gradients of -1.5 ppb/ $^{\circ}\text{C}$.

H_N shifts. For two-state folding, all chemical shifts are population-weighted averages,

$$\delta_{\text{obs}} = \chi_F \delta_F + (1 - \chi_F) \delta_{\text{r.c.}} \quad \text{and}$$

$$\delta_F = \text{CSD}^{298} + (\partial\delta_{\text{NH}}/\partial T)_F (T - 298);$$

with temperature correction required for both δ_F and $\delta_{\text{r.c.}}$. We used this method to calculate a set of simulated H_N chemical shifts for MrH4a in 8% HFIP. For this simulation, we assumed that $(\partial\delta_{\text{NH}}/\partial T)_F = -1.5$ ppb/ $^{\circ}\text{C}$ for all $S \pm \text{odd}$ H_N sites and that $(\partial\delta_{\text{NH}}/\partial T) = -7.8$ ppb/ $^{\circ}\text{C}$ applied to the other sites in the folded state and to *all* sites in the unfolded state. We employed a χ_F vs. T profile consistent with the known extent of cold denaturation and fitted the values of CSD^{298} for each H_N to match the observed CSDs at 300 K. We then employed the two equations above to calculate H_N shifts at each temperature. The calculated H_N shifts were analyzed using the two versions of CSDb with the results displayed in the lower panel of Figure 10. The shifts, simulated using a strict two-state assumption and with the cross-strand H-bonded sites accorded appropriate temperature gradients, reproduce the experimental data to a remarkable degree. We suggest that this represents support for two-state folding of peptide hairpins. We further suggest this supports the use of our CSD programs for determining whether specific sites are sequestered or exposed in the folded states of peptides and miniproteins in rapid equilibrium with unfolded ensembles approximating the random-coil state.

Concluding remarks

CSD profiles and their temperature dependence can provide a surprising wealth of structural and thermodynamic data. In our opinion, for hairpins and other small peptide structural models, NOE-based structure ensemble generation is only warranted when the folded fraction is greater than 0.75. When the folded populations are less than 0.70, the observed NOEs (1) could reflect partially folded states, (2) will certainly contain contributions from the unfolded ensemble that will yield distance ratios incompatible with the folded state, and (3) could, in combination with the diluted long-range connectivities from the folded state, produce false convergence. In the absence of cyc-

lizing cross-links and/or cross-strand ($S \pm \text{odd}$) aromatic/aromatic interactions, the most stable isolated hairpins constructed from Gly and L-AAAs are only 20–50% folded in water. The inclusion of D-Pro-Gly turns can increase hairpin fold populations in water (but only up to about 75%). HFIP (or TFE) addition typically produces further improvements in hairpin populations, to fold populations over 90% at 298 K for some pG turn species. CSD analysis can provide both register and turn type even when the fold populations are quite small. Independent of precise turn type it is the inwardly directed H_N and $H\alpha$ sites that display the largest downfield shifts. We recommend these, and specific turn H_N sites (Figure 4b), as the best probes for assessing the effects of mutations, media changes, and temperature on folding equilibria. The excellent correlations between alternative chemical shift probes of hairpin population is consistent with hairpins being two-state folding species. These correlations, which include sites that are in the turn region, suggest that residual structuring of the turn region is not significant either under partial melting conditions or in mutants with less favorable cross-strand interactions. We expect chemical shift analysis to continue to be a powerful tool for understanding peptide structuring phenomena that may replicate the nucleating events in protein folding.

Electronic supplementary material is available at <http://dx.doi.org/10.1007/s10858-005-3731-7>

Acknowledgments

This research was predominantly supported by NSF Grant CHE0315361 which is hereby acknowledged. Some of the peptides were first prepared in the course of NIH-supported research efforts (GM059658).

References

- Andersen, N.H., Brodsky, Y., Neidigh, J.W. and Prickett, K.S. (2002) *Bioorg. Med. Chem.*, **10**, 79–85.
- Andersen, N.H., Cort, J.R., Liu, Z., Sjoberg, S.J. and Tong, H. (1996) *J. Am. Chem. Soc.*, **118**, 10309–10310.
- Andersen, N.H., Dyer, R.B., Fesinmeyer, R.M., Gai, F., Liu, Z., Neidigh, J.W. and Tong, H. (1999) *J. Am. Chem. Soc.*, **121**, 9879–9880.
- Andersen, N.H., Neidigh, J.W., Harris, S.M., Lee, G.M., Liu, Z. and Tong, H. (1997) *J. Am. Chem. Soc.*, **119**, 8547–8561.

- Blanco, F.J., Jiménez, M.A., Herranz, J., Rico, M., Santoro, J. and Nieto, J.L. (1993) *J. Am. Chem. Soc.*, **115**, 5887–5888.
- Blanco, F.J., Rivas, G. and Serrano, L. (1994) *Nat. Struct. Biol.* **1**, 584–590.
- Blanco, F.J. and Serrano, L. (1995) *Eur. J. Biochem.* **230**, 634–649.
- Blandl, T., Cochran, A.G. and Skelton, N.J. (2003) *Protein Sci.*, **12**, 237–247.
- Bradley, E.K., Thomason, J.F., Cohen, F.E., Kosen, P.A. and Kuntz, I.D. (1990) *J. Mol. Biol.*, **215**, 607–622.
- Brown, J.E. and Klee, W.A. (1971) *Biochemistry*, **10**, 470–476.
- Chen, P.-Y., Lin, C.-K., Lee, C.-T., Jan, H. and Chan, S.I. (2001) *Protein Sci.*, **10**, 1794–1800.
- Ciani, B., Jourdan, M. and Searle, M.S. (2003) *J. Am. Chem. Soc.*, **125**, 9038–9047.
- Cierpicki, T. and Otlewski, J. (2001) *J. Biomol. NMR*, **21**, 249–261.
- Cochran, A.G., Skelton, N.J. and Starovasnik, M.A. (2001) *Proc. Natl. Acad. Sci. USA*, **98**, 5578–5583.
- Cort, J.R., Liu, Z., Lee, G.M., Harris, S.M., Prickett, K.S., Gaeta, L.S.L. and Andersen, N.H. (1994) *Biochem. Biophys. Res. Commun.*, **204**, 1088–1095.
- Cox, J.P.H., Evans, P.A., Packman, L.C., Williams, D.H. and Woolfson, D.N. (1993) *J. Mol. Biol.*, **234**, 483–492.
- de Alba, E., Jiménez, M.A., Rico, M. and Nieto, J.L. (1996) *Fold. Des.*, **1**, 133–144.
- de Alba, E., Rico, M. and Jiménez, M.A. (1999) *Protein Sci.*, **8**, 2234–2244.
- Doig, A.J. and Baldwin, R.L. (1995) *Protein Sci.*, **4**, 1325–1336.
- Dyer, R.B., Manas, E.S., Peterson, E.S., Franzen, S., Fesinmeyer, R.M. and Andersen, N.H. (2004) *Biochemistry*, **43**, 11560–11566.
- Dyer, R.B., Maness, S.J., Franzen, S., Fesinmeyer, R.M., Olsen, K.A. and Andersen, N.H. (2005) *Biochemistry*, **44**, 10406–10415.
- Espinosa, J.F. and Gellman, S.H. (2000) *Angew. Chem. Int. Ed.*, **39**, 2330–2333.
- Espinosa, J.F., Muñoz, V. and Gellman, S.H. (2001) *J. Mol. Biol.*, **306**, 397–402.
- Espinosa, J.F., Syud, F.A. and Gellman, S.H. (2002) *Protein Sci.*, **11**, 1492–1505.
- Fesinmeyer, R.M., Hudson, F.M. and Andersen, N.H. (2004) *J. Am. Chem. Soc.*, **126**, 7238–7243.
- Frishman, D. and Argos, P. (1995) *Proteins*, **23**, 566–579.
- Gibbs, A.C., Bjorndahl, T.C., Hodges, R.S. and Wishart, D.S. (2002) *J. Am. Chem. Soc.*, **124**, 1203–1213.
- Griffiths-Jones, S.R., Maynard, A.J. and Searle, M.S. (1999) *J. Mol. Biol.*, **292**, 1051–1069.
- Griffiths-Jones, S.R. and Searle, M.S. (2000) *J. Am. Chem. Soc.*, **122**, 8350–8356.
- Honda, S., Kobayashi, N. and Munekata, E. (2000) *J. Mol. Biol.*, **295**, 269–278.
- Hughes, R.M. and Waters, M.L. (2005) *J. Am. Chem. Soc.*, **127**, 6518–6519.
- Karle, I.L., Awasthi, S.K. and Balaram, P. (1996) *Proc. Natl. Acad. Sci. USA*, **93**, 8189–8193.
- Kiehna, S.E. and Waters, M.L. (2003) *Protein Sci.*, **12**, 2657–2667.
- Kobayashi, N., Endo, S. and Munekata, E. (1993) *Peptide Chem.*, 278–280.
- Lee, G.M., Chen, C., Marschner, T.M. and Andersen, N.H. (1994) *FEBS Lett.*, **355**, 140–146.
- López de la Paz, M., Lacroix, E., Ramírez-Alvarado, M. and Serrano, L. (2001) *J. Mol. Biol.*, **312**, 229–246.
- Maynard, A.J., Sharman, G.J. and Searle, M.S. (1998) *J. Am. Chem. Soc.*, **120**, 1996–2007.
- Merutka, G., Dyson, H.J. and Wright, P.E. (1995) *J. Biomol. NMR*, **5**, 14–24.
- Muñoz, V., Thompson, P.A., Hofrichter, J. and Eaton, W.A. (1997) *Nature*, **390**, 196–199.
- Osapay, K. and Case, D.A. (1991) *J. Am. Chem. Soc.*, **113**, 9436–9444.
- Osapay, K. and Case, D.A. (1994) *J. Biomol. NMR*, **4**, 215–230.
- Piotto, M., Saudek, V. and Sklenar, V. (1992) *J. Biomol. NMR*, **2**, 661–665.
- Ramírez-Alvarado, M., Blanco, F.J., Niemann, H. and Serrano, L. (1997) *J. Mol. Biol.*, **273**, 898–912.
- Ramírez-Alvarado, M., Blanco, F.J. and Serrano, L. (1996) *Nat. Struct. Biol.*, **3**, 604–612.
- Roe, D.R., Hornack, V. and Simmerling, C. (2005) *J. Mol. Biol.*, **352**, 370–381.
- Santiveri, C.M., Rico, M. and Jiménez, M.A. (2001) *J. Biomol. NMR*, **19**, 331–345.
- Santiveri, C.M., Rico, M., Jiménez, M.A., Pastor, M.T. and Pérez-Payá, E. (2003) *J. Pept. Res.*, **61**, 177–188.
- Santiveri, C.M., Santoro, J., Rico, M. and Jiménez, M.A. (2002) *J. Am. Chem. Soc.*, **124**, 14903–14909.
- Santiveri, C.M., Santoro, J., Rico, M. and Jiménez, M.A. (2004) *Protein Sci.*, **13**, 1134–1147.
- Schenck, H. and Gellman, S. (1998) *J. Am. Chem. Soc.*, **120**, 4869–4870.
- Schwarzinger, S., Kroon, G.J.A., Foss, T.R., Wright, P.E. and Dyson, H.J. (2000) *J. Biomol. NMR*, **18**, 43–48.
- Searle, M.S. (2001) *J. Chem. Soc. Perkin Trans.*, **2**, 1011–1020.
- Searle, M.S., Williams, D.H. and Packman, L.C. (1995) *Nat. Struct. Biol.*, **2**, 999–1006.
- Sharman, G.J., Griffiths-Jones, S.R., Jourdan, M. and Searle, M.S. (2001) *J. Am. Chem. Soc.*, **123**, 12318–12324.
- Sharman, G.J. and Searle, M.S. (1998) *J. Am. Chem. Soc.*, **120**, 5291–5300.
- Sibanda, B.L. and Thornton, J.M. (1991) *Methods Enzymol.*, **202**, 59–82.
- Syud, F.A., Espinosa, J.F. and Gellman, S.H. (1999) *J. Am. Chem. Soc.*, **121**, 11577–11578.
- Syud, F.A., Stanger, H.E. and Gellman, S.H. (2001) *J. Am. Chem. Soc.*, **123**, 8667–8677.
- Syud, F.A., Stanger, H.E., Mortell, H.S., Espinosa, J.F., Fisk, J.D., Fry, C.G. and Gellman, S.H. (2003) *J. Mol. Biol.*, **326**, 553–568.
- Tatko, C.D. and Waters, M.L. (2003) *Protein Sci.*, **12**, 2443–2452.
- Tatko, C.D. and Waters, M.L. (2004) *Protein Sci.*, **13**, 2515–2522.
- Trabi, M., Schirra, H.J. and Craik, D.J. (2001) *Biochemistry*, **40**, 4211–4221.
- Wagner, G., Pardi, A. and Wuthrich, K. (1983) *J. Am. Chem. Soc.*, **105**, 5948–5949.
- Wishart, D.S., Bigam, C.G., Holm, A., Hodges, R.S. and Sykes, B.D. (1995) *J. Biomol. NMR*, **5**, 67–81.
- Wishart, D.S. and Sykes, B.D. (1994) *Methods Enzymol.*, **239**, 363–392.
- Wishart, D.S., Sykes, B.D. and Richards, F.M. (1991) *J. Mol. Biol.*, **222**, 311–333.
- Xu, X.P. and Case, D.A. (2001) *J. Biomol. NMR*, **21**, 321–333.
- Xu, X.P. and Case, D.A. (2002) *Biopolymers*, **65**, 408–423.
- Xu, Y., Oyola, R. and Gai, F. (2003) *J. Am. Chem. Soc.*, **125**, 15388–15394.