ARTICLE

Structural impact of thioamide incorporation into a β-hairpin

Kristen E. Fiore,[a] Martijn J. Patist,[a] Sam Giannakoulias,[a] Cheng-Hsin Huang,[b] Hitesh Verma,[c] Bhavesh Khatri,[c] Richard P. Cheng,[b] Jayanta Chatterjee,[c] and E. James Petersson[a]*

Received 00th January 20xx,
Accepted 00th January 20xx

DOI: 10.1039/x0xx00000x

The thioamide is a naturally-occurring single atom substitution of the canonical amide bond. The exchange of oxygen to sulfur alters the amide's physical and chemical characteristics, thereby expanding its functionality. Incorporation of thioamides in prevalent secondary structures has demonstrated that it can either have stabilizing, destabilizing, or neutral effects. We performed a systematic investigation of the structural impact of thioamide incorporation in a β-hairpin scaffold with nuclear magnetic resonance (NMR). Thioamides as hydrogen bond donors did not increase the foldedness of the more stable "YKL" variant of this scaffold. In the less stable "HPT" variant of the scaffold, the thioamide could be stabilizing as a hydrogen bond donor and destabilizing as a hydrogen bond acceptor, but the extent of the perturbation depended upon the position of incorporation. To better understand these effects we performed structural modelling of the macrocyclic Folded HPT variants. Finally, we compare the thioamide effects that we observe to previous studies of both side-chain and backbone perturbations to this β-hairpin scaffold to provide context for our observations.

Introduction

The thioamide is an intriguing isostere of the canonical amide bond. Although it differs by only a single atom, the thioamide has unique chemical and physical properties that have been employed by chemists and biophysicists. For example, the thioamide has a lower oxidation potential.[1] Consequently, thioamides can quench fluorescence in a distance dependent manner: Förster resonance energy transfer (FRET)-based quenching of UV fluorophores and photoinduced electron transfer (PeT)-based quenching of visible fluorophores.[2] Therefore, fluorophore/thioamide pairs can be utilized as minimal biophysical probes to study protein folding or dynamics. This has been done to monitor peptide-protein binding,[3-5] to monitor protease activity in real time,[6] and to monitor protein conformational changes during refolding,[3] unfolding,[7, 8] or misfolding.[5]

Another unique property of the thioamide is that is has a red-shifted π-to-π^* absorption,[9] giving it a unique circular dichroism (CD) signature.[10] This red-shifted absorption also lowers the excitation energy required for photoisomerization.[11] Therefore, upon irradiation the thioamide can selectivity photoisomerize from *trans*-to-*cis*, enabling its use as a photoswitch in peptides[12-15].

Nature installs thioamides in ribosomally synthesized and post-translationally modified peptides (RiPPs),[16] as well as in at least two proteins: methyl coenzyme M reductase (MCR)[17, 18] and the uL16 protein of the *E. coli* 70S ribosome[19]. Although the method of installation is well-studied, the effect of thioamidation on protein function is relatively unknown.

To develop a systematic understanding of how the thioamide can affect biological activity, as well as to promote its utility as a biophysical probe, a more fundamental understanding of the structural impact of thioamide incorporation is needed. Previously, small molecule studies have suggested that the lower electronegativity of the thioamide sulfur results in it being a weaker hydrogen bond acceptor.[20-22] Conversely, the lower N-H pK_a (12 versus 14)[23] should result in the thioamide being a stronger hydrogen bond donor (**Fig. 1**).[24, 25] Additionally the larger van der Waals radius of sulfur[26] results in a longer C=S double bond[27, 28] which could be perturbative depending upon the environment. To determine the impact on protein thermodynamic stability, we incorporated thioamides into native proteins of different secondary structures: calmodulin (α-helical), the B1 domain of protein G (GB1, β-sheet), and collagen (PPII triple-helix).[29] In some cases, reasonable explanations could be made as to why incorporation at some positions was more destabilizing than others based on the existing structures of the native proteins. However, many seemingly similar thioamide substitutions resulted in very different effects on protein stability. We were particularly intrigued by the effects on GB1, where substitution in the same β-strand had destabilizing effects differing by 2 kcal/mol. The same variation was observed in another β-sheet structure, the Pin1 WW domain (three β-strands), where the thermostability ($\Delta T_{M(Thio\text{-}Oxo)}$) varied from -0.9 to 14.8˚C depending upon the microenvironment of the position.[30] To further investigate the effects of thioamides in β-sheets, we turned to model peptide systems, which have proven to be valuable for rigorous investigation of protein modifications.

[a] *Department of Chemistry, University of Pennsylvania, 231 S. 34th Street, Philadelphia, 19104, USA*
[b] *Department of Chemistry, National Taiwan University, No 1, Sec 4, Roosevelt Road, Taipei 10617, Taiwan*
[c] *Molecular Biophysics Unit, Indian Institute of Science, Bangalore 560012, India*
Electronic Supplementary Information (ESI) available: peptide synthesis and characterization data, and details of NMR-based computational modeling. Links to coordinates for structures are provided. See DOI: 10.1039/x0xx00000x

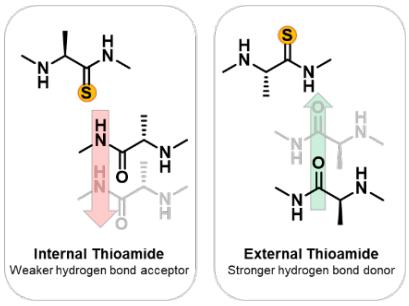

Fig. 1. Expected thioamide effects on hydrogen bonding. Based on previous small molecule studies, it is expected that an internal thioamide will be disruptive to β-hairpin structure due to the thioamide being a weaker hydrogen bond acceptor and having a larger van der Waals radius. An external thioamide will be stabilizing since it is a stronger hydrogen bond donor than an amide.

Thioamides have been previously investigated in α-helical[31, 32] and poly-proline II (PPII) helical model systems[33]. The structural impacts were scaffold and position specific. In contrast, there has been very limited study of thioamides in β-sheet model systems with only four examples to our knowledge. In two of the studies, thioamide substitution occurred in the turn, and therefore is not informative on how the chemical and physical properties of the thioamide affect cross-strand β-sheet interactions.[34, 35] In a well-studied β-hairpin, the tryptophan zipper (Trpzip), thioamides were incorporated as hydrogen bond acceptors and the thermodynamic stability measured using CD spectroscopy.[36] Overall, the thioamides within the strands were destabilizing by about 1 kcal/mol, with the terminal position being the least perturbative. Thioamides were also incorporated into a TrpZip with an azobenzene derivative at the β-turn.[37] This allowed for control of the folding state of the peptide by photo-initiated *cis*/*trans* isomerization (*cis* = folded, *trans* = unfolded). Folding was observed with time-resolved IR spectroscopy and CD. For these TrpZips, two thioamides were incorporated to observe site-specific coupling spectroscopically. Thioamides on the same strand serving as hydrogen bond donors were minimally perturbative with unfolding rates similar to the all-amide reference. Thioamides on opposite strands serving as hydrogen bond donors stabilized the β-hairpin relative to the reference. Thioamides on opposite strands as hydrogen bond acceptors strongly destabilized the β-hairpin.

In these studies, the impact of thioamides on β-sheets are largely interpreted in terms of differing hydrogen bonding properties. However, such a simple interpretation is inconsistent with our observations in GB1, where local structure significantly altered the impact of thioamide substitution. Moreover, although these two studies include elegant kinetic and thermodynamic studies, they lack any direct structural information. Although TrpZips have been well-characterized, we were interested in studying a less-folded β-hairpin scaffold that might be more sensitive to effects of the thioamide and one for which an extensive body of literature on other non-covalent interactions is available for comparison. Therefore, we have designed a systematic investigation of thioamide incorporation using a model β-hairpin system that meets these requirements and performed structural analysis with NMR. This experimental data is supplemented with structural modelling of the macrocyclic folded variants.

Results and Discussion

Scaffold Design

In choosing a host scaffold for our thioamide guests, we analyzed β-hairpins that are water soluble, monomeric, and have significant β-sheet character. A well-established construct that meets these characteristics is the Gellman "YKL" β-hairpin.[38, 39] This scaffold has been utilized to study the β-hairpin stabilization of cross-strand interactions[38] and strand length[40], as well as the β-sheet propensity of charged derivatives of β-branched-amino acids[41] and aza-amino acids[42]. This extensively studied scaffold is also a good starting point for structural characterization as several previous studies have reported structural models based on NMR data.[38-40, 43] This β-hairpin is enforced with a stabilized two-residue β-turn, proGly (where pro is D-proline). Examination of 2 residue β-hairpin loops in proteins determined that there is a preference for type I' and type II' β-turns[44] and pro promotes the right-handed twist needed for this biologically relevant β-turn.[45] Conversely, ProGly incorporation results in a left-handed turn that eradicates the β-sheet structure, which can be utilized for synthesis of an Unfolded control peptide.[39] This 12 residue anti-parallel "YKL" β-sheet is stabilized by a diagonal cation-π interaction (Tyr_2 and Lys_9). To avoid electrostatic associations from the termini, the N-terminus is acetylated, and the C-terminus is a carboxamide (**Fig. 2**). We hypothesized that the β-hairpin would be stabilized when the thioamide is positioned as a hydrogen bond donor.

YKL Scaffold CSD Analysis

We initially chose to incorporate thioamides at the following hydrogen bond donor positions: Glu_4, Lys_9 and Leu_{11}. These constructs are denoted YKL-Glu^S_4, YKL-Lys^S_9, and YKL-Leu^S_{11}-OH, using the superscript S convention for naming thioamides from Mahanta *et al.*[16] The Leu^S_{11} peptide was synthesized with a C-terminal carboxylate (indicated as -OH) due to the propensity for thioamides at the penultimate position to cause hydrolysis and epimerization at C-terminal amides (ESI, **Fig. S4**).[3] For each thioamide position of interest as well as the all-amide reference peptide denoted "YKL", two peptides were synthesized, the Test peptide (with proGly at the turn) and the corresponding

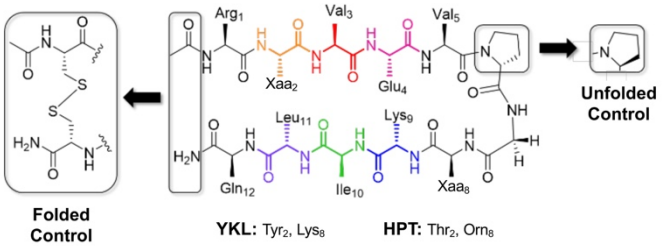

Fig. 2. YKL and HPT test β-hairpin scaffolds. The Unfolded control has a ProGly β-turn. The Folded control has terminal cysteines that are oxidized to form a disulfide linked cyclic peptide. The Folded control was only synthesized for the HPT scaffold.

Unfolded control peptide. The Unfolded control has the same sequence as the Test peptide but has a ProGly β-turn so it does not have any β-sheet secondary structure.

Except for Glu$^{S}_{4}$, the Test peptides displayed significant anti-parallel β-sheet character (minimum at 215 nm) in the far-UV region of their CD spectra (ESI, **Fig. S6A**). This indicated that the level of folding in the thioamide peptides was comparable to the YKL reference peptide. For YKL-Glu$^{S}_{4}$, the π-to-π* absorbance of the thioamide has a greater intensity, which complicates the spectra and could explain the lack of a minimum at 215 nm. Since the contribution of a thioamide residue to the CD spectrum is not well-defined, we sought other measurements to quantify the relative stabilities of the thioamide variants. CD thermal melts measured at the 215 nm signature were linear, and therefore could not be fit to two-state models to derive folding energetics (ESI, **Fig. S6B**). Therefore, we turned to 2D NMR measurements instead. Peptides were dissolved in sodium deuteroacetate buffer (ESI, **Table S3**)‡ and TOCSY and ROESY were collected at 10 ˚C. Lateral NOEs between Tyr$_2$ and Leu$_{11}$, as well as diagonal NOEs between Tyr$_2$ and Lys$_9$ were observed for all Test β-hairpins (ESI, **Fig. S7**). This suggests that all thioamide-containing variants are in a β-hairpin conformation. Although the connectivities of the observed NOEs differ slightly, the similarity of the overall patterns is enough to deem the structures alike.

As expected, the Unfolded controls lack cross-strand NOEs (ESI, **Fig. S8**). Inclusion of the ProGly β-hairpin allows calculation of the chemical shift deviation (CSD or Δδ = Test δ – random coil δ) with the ProGly Unfolded control (Δδ = Test δ – Unfolded control δ). β-hairpins are dynamic structures, and CSD analysis is reflective of the global average, providing a quantitative measure indicative of secondary structure. A $\Delta\delta_{H\alpha}$ of greater than 0.1 ppm for three consecutive residues is considered to be evidence of a β-sheet.[46] The $\Delta\delta_{H\alpha}$ data suggest that all tested thioamide-peptides have β-sheet character comparable to the YKL reference peptide (**Fig. 3**). At each thioamide position, the $\Delta\delta_{H\alpha}$ for the n+1 residue is significantly different in comparison to other peptides. However, this is merely indicative of a local perturbation of the electronic environment since $\Delta\delta_{H\alpha}$ is not increased for other residues. Although YKL-Glu$^{S}_{4}$ has β-sheet character according to $\Delta\delta_{H\alpha}$, the values are lower than the other positions tested. This could be due to additional conformational rigidity from the nearby β-turn, which results in a less favorable β-hairpin structure with a thioamide at Glu$_4$. Thus, we conclude from these data as well as variable temperature NMR experiments that the thioamide hydrogen bond donor substitutions do not significantly increase the stability of the β-hairpin (ESI, **Table S4**).

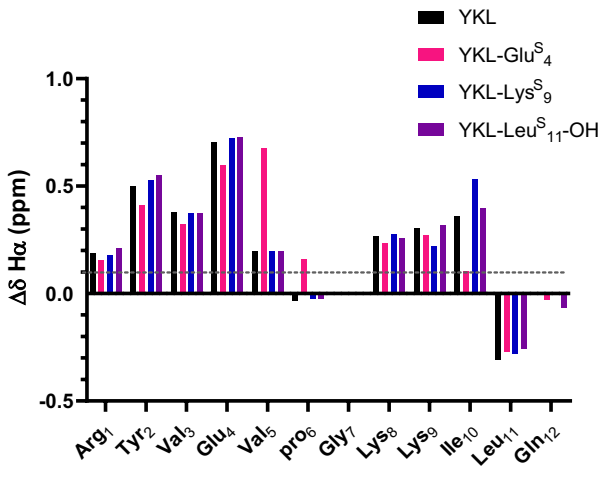

Fig. 3. $\Delta\delta_{H\alpha}$ (Test $\delta_{H\alpha}$ – Unfolded control $\delta_{H\alpha}$) for YKL-Glu$^{S}_{4}$, YKL-Lys$^{S}_{9}$, and YKL-Leu$^{S}_{11}$-OH in comparison to YKL. Three consecutive $\Delta\delta_{H\alpha}$ values of greater than 0.1 ppm (shown in Fig.) is indicative of β-sheet structure. Besides slight variations around where the thioamide is incorporated, the $\Delta\delta_{H\alpha}$ values are like YKL for all positions tested suggesting that the thioamide is not increasing foldedness. $\Delta\delta_{NH}$ analysis also demonstrates the same trends (**Fig. S9**).

HPT Scaffold CSD Analysis

In light of the unexpected failure of the thioamide hydrogen bond donor to increase foldedness, we questioned whether the YKL scaffold was too stable to observe potential perturbations to structure due to thioamide incorporation. Removal of the cation-π interaction would decrease the stability of the scaffold, making it more sensitive to perturbation. Indeed, the Cheng lab previously replaced Tyr$_2$ with Thr to achieve this purpose and used Orn$_8$ instead of Lys$_8$ to help with chemical shift assignment.[47-49] Also, the internal Tyr side-chain causes ring-currents that result in upfield shifts of other internal protons. Therefore, removal of Tyr additionally leaves the internal chemical shifts unaffected by the ring-currents and allows for more accurate determination of the effect of incorporation of an internal thioamide. This variant of the YKL scaffold is referred to as HPT (**H**air**P**ins with **T**hr at position 2) (**Fig. 2**). The hypothesis remained that incorporation of a thioamide as a hydrogen bond donor would increase foldedness, whereas a thioamide as a hydrogen bond acceptor would decrease foldedness.

Thioamides were incorporated at Thr$_2$, Val$_3$, Ile$_{10}$, and Leu$_{11}$ in both Test (proGly) and Unfolded (ProGly) forms: HPT-Thr$^{S}_{2}$, HPT-Val$^{S}_{3}$, HPT-Ile$^{S}_{10}$, and HPT-Leu$^{S}_{11}$-OH. Again, a carboxylate was included at the C-terminus of the Leu$^{S}_{11}$ peptide, in this case due to hydrolysis of the amide form (ESI, **Figs. S4-5, Tables S6-7**). The Test β-hairpins were first examined by CD and have a significant thioamide *π*-to-*π** absorption band (ESI, **Fig. S10**). However, the less stable Test HPT peptides do not have as strong a β-sheet signature at 218 nm, which is further complicated by the strong thioamide absorbance. Therefore, the effect of thioamide incorporation on global secondary structure could not be determined using CD. As a result, we again relied on NMR for information.

HPT, HPT-Thr$^{S}_{2}$, HPT-Val$^{S}_{3}$, HPT-Ile$^{S}_{10}$, and HPT-Leu$^{S}_{11}$-OH were dissolved in sodium deuteroacetate or phosphate buffer to 1-10 mM concentration (ESI, **Table S3**)‡. TOCSY and ROESY spectra were collected at 25˚C. The HPT NMR data was obtained from a previous publication.[49]

For the Test peptide, the HPT β-hairpin has both lateral NOEs between Thr$_2$ and Leu$_{11}$, as well as diagonal NOEs between Thr$_2$ and Lys$_9$. All thioamide Test β-hairpins have

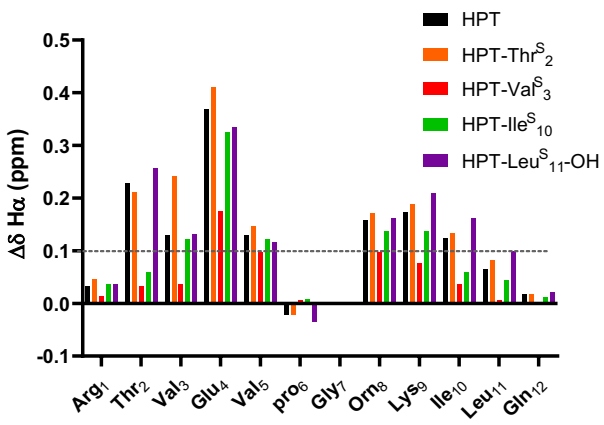

Fig. 4. $\Delta\delta_{H\alpha}$ (Test $\delta_{H\alpha}$ – Unfolded control $\delta_{H\alpha}$) for HPT-Thr$^S_2$, HPT-Val$^S_3$, HPT-Ile$^S_{10}$, and HPT-Leu$^S_{11}$-OH in comparison to HPT. Three consecutive $\Delta\delta_{H\alpha}$ values of greater than 0.1 ppm (shown in Fig.) is indicative of β-sheet structure. The $\Delta\delta_{H\alpha}$ data demonstrates that HPT-Thr$^S_2$, HPT-Ile$^S_{10}$, and HPT-Leu$^S_{11}$-OH have β-sheet character, whereas HPT-Val$^S_3$ does not. Discussion of $\Delta\delta_{NH}$ is in the SI (**Fig. S13**).

mainly lateral NOEs between Glu_4 and Lys_9, with HPT-Thr$^S_2$ having a second cross-strand NOE between Thr_2 and Leu_{11} (ESI, **Fig. S11**). The presence of cross-strand NOEs for all positions tested is suggestive of β-hairpin conformation. As expected, all Unfolded controls lack cross-strand NOEs (ESI, **Fig. S12**). For a more quantitative measure of β-hairpin structure, we again used CSD analysis.

The $\Delta\delta_{H\alpha}$ values for HPT are generally lower than those for YKL, indicating that it has less β-sheet structure, in agreement with the CD data and expectations based on prior literature.[49] The $\Delta\delta_{H\alpha}$ values for the HPT-Thr$^S_2$, HPT-Ile$^S_{10}$ and HPT-Leu$^S_{11}$-OH peptides demonstrate that all have β-sheet structure, whereas HPT-Val$^S_3$ does not (**Fig. 4**). Besides HPT-Ile$^S_{10}$, these observations agree with our YKL work that a thioamide as a hydrogen bond donor is neutral, and additionally agrees with our hypothesis that an internal thioamide is destabilizing. We did not expect that an internal thioamide (Ile$^S_{10}$) could be minimally perturbing. The difference observed for both internal thioamide positions (Val$^S_3$ and Ile$^S_{10}$) could be due to the right-handed twist of the β-hairpin[50], which would better accept the additional steric bulk at Ile_{10}, but not Val_3.

HPT Scaffold ΔΔG Analysis

To quantitatively compare the effect of thioamide incorporation, Folded controls were synthesized to allow for calculation of fraction folded (%) and $\Delta\Delta G_{Folding}$. For the Folded control, terminal cysteines were added and oxidized to produce a disulfide-linked cyclic peptide following a strategy previously employed by the Cheng laboratory (**Fig. 2**).[43, 47-49] However, this analysis is only valid if the Folded controls display sufficient chemical shift dispersion. As the β-hairpin is more folded, both $\delta_{H\alpha}$ and δ_{NH} should be shifted downfield. This was observed when the thioamide is externally facing (i.e., positioned as a hydrogen bond donor). However, the Folded controls with an internal thioamide did not display enough dispersion to be a true folded control (**Fig. 5**). This could be due to puckering as the internal electron-rich thioamide is forced towards the opposing strand due to steric constraints. Since we observed abundant cross-strand NOEs for the HPT-Val$^S_3$ and HPT-Ile$^S_{10}$ Folded controls (ESI, **Fig. S14**), we performed fraction folded analysis despite the lack of chemical shift dispersion. We note that in **Fig. 5** $\Delta\delta_{H\alpha}$ is dramatically affected for the residue following the thioamide, consistent with previous studies of the steric and electronic impact of thioamide incorporation on neighboring residues.[51, 52] The $\Delta\delta_{H\alpha}$ value is increased relative to the HPT value for more stable peptides with externally facing thioamides and decreased for less stable peptides with internally facing thioamides.

Fraction folded was calculated for each residue of a Test β-hairpin using **Equation 1**. The final fraction folded value reported is an average of position 3 and 10. These positions were chosen because Val_3 and Ile_{10} are in the middle of the β-strands and therefore not affected by the flexible termini or the constrained β-turn. Secondly, for Val_3 and Ile_{10}, Hα is externally facing and therefore $\delta_{H\alpha}$ is not affected by the internal micro-environment of the β-hairpin which may be perturbed by the thioamide. Thus, $\delta_{H\alpha}$ for these positions is a direct indicator of foldedness. Similar to $\Delta\delta_{H\alpha}$ analysis, HPT-Thr$^S_2$ has the same

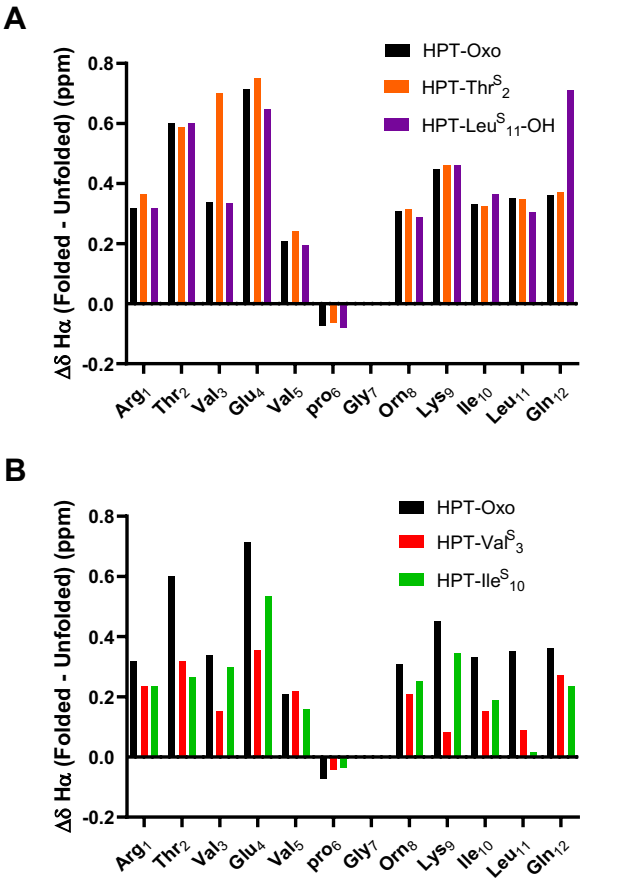

Fig. 5. $\Delta\delta_{H\alpha}$ (Folded $\delta_{H\alpha}$ – Unfolded control $\delta_{H\alpha}$) for external thioamides (**A**, HPT-Thr$^S_2$ and HPT-Leu$^S_{11}$-OH) and internal thioamides (**B**, HPT-Val$^S_3$ and HPT-Ile$^S_{10}$) in comparison to HPT. Folded controls with internal thioamides result in a decreased chemical shift dispersion. The dispersion can be quantified by comparing the $\Delta\delta_{H\alpha}$ between the Folded and Unfolded controls. For the external thioamides (**A**), there is a similar dispersion observed for all the peptides. This is not observed for the internal thioamides (**B**), where the values greatly differ in comparison to HPT.

Table 1. Fraction Folded and $\Delta\Delta G_{Folding}$ for HPT β-Hairpins

Peptide	Fraction Folded (%)	$\Delta\Delta G_{Folding}$ (kcal/mol)
HPT	38 ± 1%	----
HPT-Thr$^{S}_{2}$	38 ± 3%	0.01 ± 0.1
HPT-Val$^{S}_{3}$	25 ± 1%	0.38 ± 0.01
HPT-Ile$^{S}_{10}$	37 ± 5%	0.04 ± 0.1
HPT-Leu$^{S}_{11}$-OH	42 ± 3%	-0.09 ± 0.1

Fraction folded calculated using Equation 1, reported as an average of positions 3 and 10. Equation 2 and 3 were used to calculate $\Delta\Delta G_{Folding}$. Both methods of analysis are suggestive that HPT-Thr$^{S}_{2}$ and HPT-Ile$^{S}_{10}$ have a similar energy of folding and fraction folded profile to HPT. HPT-Leu$^{S}_{11}$-OH is more folded and has a more favorable free energy of folding than HPT. Thioamide incorporation at Val$^{S}_{3}$ is disruptive to β-hairpin structure. See **Table S8** for all values used in calculating fraction folded and ΔΔG.

fraction folded value as HPT (**Table 1**). HPT-Leu$^{S}_{11}$-OH has a higher fraction folded, HPT-Ile$^{S}_{10}$ is slightly less folded, and HPT-Val$^{S}_{3}$ is the least folded. $\Delta\Delta G_{Folding}$ was calculated with **Equations 2** and **3**. The $\Delta\Delta G_{Folding}$ provide a general metric that can be compared to other β-hairpin studies to place the magnitude of the observed effects in context.

$$Fraction\ Folded = \frac{\delta H\alpha_{Test} - \delta H\alpha_{Unfolded}}{\delta H\alpha_{Folded} - \delta H\alpha_{Unfolded}} \quad \textbf{Eq. 1}$$

$$\Delta G = -RT * \ln(K_{eq}) = -RT * \ln\left(\frac{\delta H\alpha_{Test} - \delta H\alpha_{Unfolded}}{\delta H\alpha_{Folded} - \delta H\alpha_{Test}}\right) \quad \textbf{Eq. 2}$$

$$\Delta\Delta G = \Delta G_{Thio} - \Delta G_{oxo} \quad \textbf{Eq. 3}$$

These data clearly show that even in a short β-hairpin the context of the thioamide substitution is very important. While the thioamide as hydrogen bond acceptor can indeed be destabilizing (HPT-Val$^{S}_{3}$), it can also be a neutral modification (HPT-Ile$^{S}_{10}$). Likewise, while the thioamide as hydrogen bond donor can be stabilizing in the more sensitive HPT scaffold (HPT-Leu$^{S}_{11}$-OH), it too can be neutral (HPT-Thr$^{S}_{2}$), depending on local context.

Structural Modelling

To elucidate the mechanistic basis of the thioamide effects we observed with NMR, we utilized structural modelling with PyRosetta.[53] Since the Folded HPT peptides exhibited significantly stronger NOEs (ESI, **Figs. S11** and **S14**) due to their macrocyclic constraint, we took advantage of this stability to avoid difficulty in modelling and analysis because of the flexibility in the Test peptides. As a starting point, we used the average NMR structure (PDB ID 1jy9) previously solved for a YKL derivative with four additional Thr residues at the termini (see Steric Interactions sub-section below).[40] The structure was modified to convert it to the HPT Folded control by removing the two terminal Thr residues, converting the penultimate Thr residues to Cys, forming the disulfide bond, acetylating the N-terminus, converting the C-terminus to a carboxamide, and converting Tyr_2 to Thr and Lys_8 to Orn. Following generation in PyRosetta, the initial HPT structure is similar to the starting 1jy9 PDB structure (ESI, **Fig. S16**). Next, we performed a constrained relax in Rosetta using the NOE-derived distances for the HPT Folded control to generate our final HPT structure (ESI, **Fig. S17-S18).** Previously, the Petersson laboratory developed Rosetta patches for the thioamide based on *ab initio* calculations and experimental data for thioamides in protease substrates.[54-56] With the thioamide patches and the corresponding distance constraints, the thioamide Folded control peptides: HPT-Thr$^{S}_{2}$, HPT-Val$^{S}_{3}$, HPT-Ile$^{S}_{10}$, and HPT-Leu$^{S}_{10}$ were simulated in PyRosetta. On average, 20 constraints were used per structure and only three distance pairs per β-hairpin had a deviation of greater than 0.15 Å between the experimentally-derived and computed distances (ESI, **Table S9**).

The HPT-Val$^{S}_{3}$ Folded control structure deviates greatly from the HPT Folded control with a backbone root mean squared deviation (RMSD) of 2.15 Å (**Fig. 6** and ESI, **Fig. S20** and **Table S10**). Although the structure near the turn overlays well with the HPT Folded control, accommodation of the internal

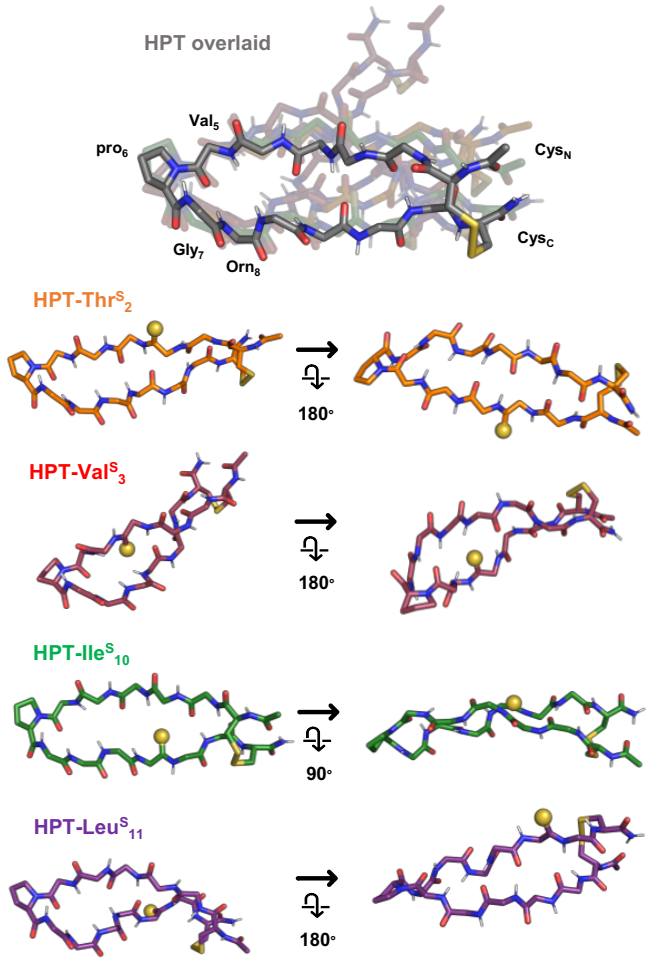

Fig. 6. Structural models of the HPT Folded control peptides. Only the backbone is displayed, except for pro_6 and the terminal Cys residues. In HPT overlaid, all of the thioamide HPT peptides are aligned to HPT (grey, other peptides in indicated colors) based on the coordinates of Val_5, pro_6, Gly_7, and Orn_8 to enable comparison (RMSDs in ESI, **Table S10**). Each thioamide peptide structure from the overlay is displayed individually from two angles, with the thioamide shown as a sphere. HPT-Thr$^{S}_{2}$ and HPT-Ile$^{S}_{10}$ are similar in structure to HPT. As a result of the right-hand twist of the β-hairpin, the thioamide for HPT-Ile$^{S}_{10}$ is more solvent-exposed. HPT-Val$^{S}_{3}$ has a dramatic twist, with differences in backbone arrangement around the internal thioamide. HPT-Leu$^{S}_{11}$ has a more pronounced twist at the terminus near the thioamide. Additional views and discussion of structures are in ESI, **Fig. S18** – **S22**.

thioamide at Val_3 results in a dramatic twist in the hairpin at Thr_2/Leu_{11}. For the non-perturbing HPT-Thr^S_2 and HPT-Ile^S_{10} Folded macrocycles, the structures of both are more like the HPT Folded peptides (**Fig. 6** and ESI, **Figs. S19**, **S21** and **Table S10**). The C-terminal strand for HPT-Thr^S_2 is closer to the N-terminal strand (hydrogen bonds among the four terminal residues are 0.5 Å shorter), potentially a result of the thioamide acting as a stronger hydrogen bond donor to the carbonyl of Ile_{10} as well as a flip of the Thr_2 side-chain due to breaking of a hydrogen bond with the Thr_2 carbonyl (ESI, **Fig. S19**). In agreement with our proposed hypothesis, the internal thioamide at Ile_{10} is better accepted due to the right-handed twist of this scaffold, allowing for a longer hydrogen bond with the Val_3 N-H without altering the hairpin shape (RMSD = 0.63 Å). For the slightly stabilized HPT-Leu^S_{11}, the overall structure is like the HPT Folded control (RMSD = 1.23 Å), with a nearly identical structure near the hairpin turn, but with an additional twist at the terminus near the thioamide (**Fig. 6**and ESI, **Fig. S22** and **Table S10**). Although the Leu_{11}-N-H/Arg_1C=O hydrogen bond distance does not change, the angle does, leading to a rotation of the Leu_{11} side-chain and movement relative to the Cys_C/Cys_N disulfide bond. Experimentally, this can be seen from an increase in the Leu_{11} and Cys_C/Cys_N NOE distances (ESI, **Table S10**).

To further analyze these macrocyclic peptide structures, we used a Backrub[57] protocol to generate ensembles. With slightly higher average deviations from the experimentally-derived distances (ESI, **Table S11**), these ensembles (except for HPT-Ile^S_{10}) demonstrate low backbone RMSDs (< 1 Å) for the 10 lowest energy structures (ESI, **Fig. S23-26**, links to coordinate files in pdb format are also provided), and align well with the constrained relax structures. For the HPT-Ile^S_{10} Folded macrocycle, there is increased rotation at the β-turn and the strand opposite the thioamide. This provides a potential mechanism for how the corresponding Test peptide can accommodate the internal thioamide (ESI, **Fig. S25**).

We note that these mechanistic explanations must be taken with some caution as a relatively small number of NOEs were available for modelling constraints. Additionally, the chemical shift dispersion is very small for HPT-Val^S_3, where the Folded control structure deviates significantly from HPT (backbone RMSD of 2.15 Å), raising some concern over whether it is a true Folded control. While these structures provide snapshots of the Folded control structures, simulation of the Unfolded and Test peptides would be required for direct comparison of energetics to the experimental results. However, by simulating these structures we were able to provide plausible mechanistic explanations for how Ile^S_{10} is well-accepted because of the right-hand twist, whereas Val^S_3 is destabilizing and results in a different configuration than HPT.

Discussion

As noted, we chose the YKL/HPT scaffold because multiple studies have used it as a host system to investigate the β-sheet propensity of various amino acids and their derivatives. These studies have included a variety of non-covalent interactions such as ion-pairing, π-system interactions, and steric constraints. A review of these findings which utilize the amino acid derivatives shown in **Fig. 7** can be found in the ESI. To enable accurate comparisons, all stability measurements are reported as $\Delta\Delta G_{Folding}$ with the parent peptide as a reference.

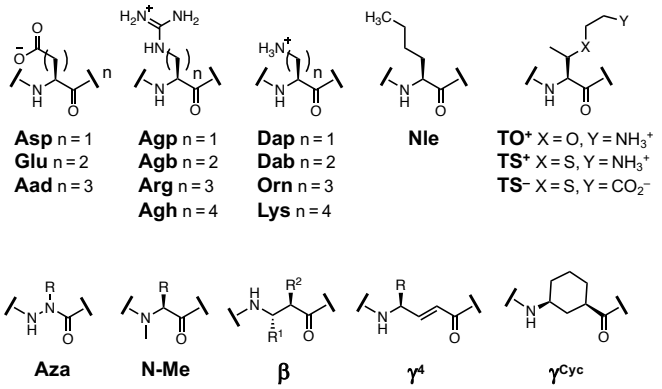

Fig. 7. Structures of amino acids derivatives previously studied in β-hairpin scaffolds.

To place the thioamide modification in the context of the field, we found the following interactions important to mention. Strengthening cation-π interactions by methylation of Lys_9 or Arg_9 side-chains stabilized the hairpin by about -0.2 kcal/mol per methylation.[58-61] Sulfur-arene interactions[62] increased the stability by -0.3 to -0.5 kcal/mol.[63] Side-chain phosphorylation demonstrated that anion-π interactions were destabilizing by ~ 1 kcal/mol.[64, 65] Introduction of charged β-branched derivatives (TS^-_4 and TS^+_9) was highly stabilizing (-0.5 kcal/mol and -0.6 kcal/mol), whereas TO^+ was slightly destabilizing (+0.1 kcal/mol).[41] Addition of β-branched residues to the termini increased stability by -0.3 kcal/mol.[40] For backbone derivatives, Aza-Val_3 incorporation was disruptive to foldedness (1.26 kcal/mol), whereas aza-Gly_3 was better accepted (0.75 kcal/mol), but still less stable than YKL.[42] β-amino acid or linear (E)-vinylogous γ^4-residues substitution was moderately destabilizing (0.5-0.6 kcal/mol).[66, 67] Whereas a cyclically constrained γ-residue was stabilizing (-0.3 to -0.6 kcal/mol).[67, 68] These perturbation studies are summarized in **Table 2**.

Thioamide effects are comparable in scale to these previous modifications. HPT-Thr^S_2, HPT-Ile^S_{10}, and HPT-Leu^S_{11}-OH demonstrate a similar energy of folding to HPT (-0.09 to +0.04 kcal/mol) where thioamide incorporation is less stabilizing than a cation-π or ion-pairing interaction. The internal thioamide at HPT-Val^S_3 is the most disruptive (0.38 kcal/mol), but is still not as disruptive as β-amino acid incorporation (0.5-0.6 kcal/mol) or phosphorylation (1 kcal/mol).

There are differing opinions as to the relative importance of backbone hydrogen bonds, side-chain electrostatic and/or hydrophobic interactions on β-hairpin stability, and their importance can change depending on the β-hairpin construct.[69-71] For the scaffolds we have discussed, it appears as though hydrophobic interactions such as aromatic stacking or addition of the β-branched derivatives are more stabilizing than electrostatic interactions (ion-pairing, cation-π). The favorable

Table 2. Summary of YKL/HPT β-hairpin Perturbation Studies

Interaction	Perturbation[Reference]	Effect on $\Delta\Delta G_{Folding}$
Cation-π	Methylation of Lys_9 across from Trp_2[58]	-0.2 kcal/mol per methylation
	Replacement of Lys_9 with Arg across from Trp_2[61]	-0.3 kcal/mol
	Methylation of Arg_9 across from Trp_2[59]	-0.6 kcal/mol for first methylation
Sulfur-arene	Replacement of Lys_9 with Met across from Trp_2 or Phe_2[63]	-0.3 to -0.5 kcal/mol
Anion-π	Phosphorylated Ser_9, Thr_9, or Tyr_9 across from Trp_2[64, 65]	+~1 kcal/mol
Steric	2 Thr added to each terminus[40]	-0.3 kcal/mol
	TS_4^- substitution[41]	-0.5 kcal/mol
	TS_9^+ substitution[41]	-0.6 kcal/mol
	TO_9^+ substitution[41]	+0.1 kcal/mol
Backbone	γ^{Cyc} substitution[67, 68]	-0.3 to -0.6 kcal/mol
	γ^4 substitution[67]	+0.5 kcal/mol
	β-amino acid replacement of two α-amino acids[66]	+0.5 to 0.6 kcal/mol per αα substitution
	Aza-Gly_3 substitution[42]	+0.8 kcal/mol
	Aza-Val_3 substitution[42]	+1.3 kcal/mol

hydrophobic interaction suggests that desolvation could play a major role in stability for the β-hairpin. Since the thioamide is less polar than the canonical amide bond, an internal thioamide could further stabilize a hydrophobic interaction. Indeed, a recent study by Chatterjee indicates that the altered desolvation of the thioamide contributes to stability in the Pin1 WW β-sheet system.[30] Stabilization by desolvation of the thioamide would be most prominent for internal thioamides. However, in the structural modelling of the HPT Folded controls with internal thioamides, the thioamides are within hydrogen bonding distance of the opposing strand and do not appear to be engaging with a hydrophobic pocket (ESI, **Fig. S20-21**). It is important to note that the previously discussed hydrophobic interactions are between side-chains (or thioamide and a side-chain), therefore the same might not be true for the backbone. Also, these β-hairpins, particularly the less stable HPT scaffold, are flexible substrates lacking tertiary structure so the ability to stabilize via backbone desolvation is very limited.

Conformational rigidity and hydrogen bonding are both backbone properties that can influence the stability of this β-hairpin scaffold. The combination of these properties, as well as differences in the micro-environment of each residue result in a complex system that does not behave as predicted based on small molecule studies. Our results show that an external thioamide can be slightly stabilizing (HPT-Leu^S_{11}-OH), whereas an internal thioamide can be destabilizing (HPT-Val^S_3), as predicted. However, they also show that an internal thioamide can be neutral (HPT-Ile^S_{10}) without significantly altering the peptide structure. The fact that these trends do not match simple interpretations of the hydrogen bonding properties of the thioamide demonstrates that the effect of incorporation is position specific.

The results also reflect the importance of certain interactions at a position in the β-hairpin. The increased stability of HPT-Leu^S_{11}-OH where the thioamide is positioned as a hydrogen bond donor suggests that backbone hydrogen bonding is important at the terminus. The lack of change in stability for HPT-Thr^S_2 and all thioamide-containing YKL β-hairpins suggests that backbone hydrogen bonds are less impactful at this position in HPT and in the YKL scaffold. However, interactions cannot always be neatly parsed into backbone and side-chain effects. For example, in our model of HPT-Thr^S_2 we observe that breaking of a backbone C=O side-chain OH hydrogen bond upon thionation leads to a twist that contributes to overall stability (ESI, **Fig. S19**). Observations such as these highlight the importance of structural data attained using the macrocyclized Folded peptides in understanding the impact of thioamide modification in model peptides and the growing number of thioamide-containing natural products.

Previous experimental work has suggested that introduction of a thioamide reduces conformational flexibility,[72, 73] and theoretical studies demonstrate increased steric constraints for the n+1 residue[51, 52, 74, 75]. Thioamide incorporation at a residue closer to the β-turn (Glu^S_4) in the YKL scaffold, has β-sheet character based on $\Delta\delta_{H\alpha}$ analysis, but it is less prevalent than the other positions tested. The decreased stability of YKL-Glu^S_4 could be due to an inability of Val_5 to accommodate the additional conformational restraint since it is already constrained by pro_6 and the β-turn. This would also explain the dramatic twist observed in our structural modelling work for the HPT-Val^S_3 Folded control. Since this macrocyclic construct is also sterically constrained by the disulfide, the twist occurs to relieve the rigidity imposed by the thioamide.

Even in the macrocyclic Folded peptides, the overall stability derives from an interplay of interactions that vary by position, making it difficult to define a single causative feature for thioamide stability effects. However, the NMR-derived models of these macrocyclic systems enable one to apply more sophisticated electronic structure calculations[51] to help to explain stability effects as well as observations such as the effect on the $\Delta\delta_{H\alpha}$ value for the n+1 residue. We will pursue such computational analysis in conjunction with additional structure determination efforts for constrained systems.

Conclusions

The collection and analysis of ^1H-^1H NMR data for thioamide incorporation into two β-hairpin scaffolds, as well as structural modelling of the macrocyclic Folded controls, suggests structural trends which deviate from expectations based on previous thioamide small molecule studies. For a stable scaffold, the YKL β-hairpin, incorporation of thioamides as hydrogen bond donors did not increase foldedness. Instead, all positions of incorporation demonstrated a similar structure to that of the YKL parent peptide. In a less stable scaffold, the HPT β-hairpin, thioamide incorporation had different structural impacts depending on position. Incorporation of a thioamide as a hydrogen bond donor was either minimally stabilizing (HPT-Leu^S_{11}-OH) or neutral (HPT-Thr^S_2). Conversely, incorporation as a hydrogen bond acceptor was either destabilizing (HPT-Val^S_3) or neutral (HPT-Ile^S_{10}). To elucidate why these two positions

were different we performed structural modelling of the Folded controls. The conformation of HPT-Val$^{S}_{3}$ is highly unlike the others as a result of structural alterations to accommodate the destabilizing internal thioamide. Conversely, the Folded control of HPT-Ile$^{S}_{10}$ is similar in structure to HPT. In this position, the internal thioamide is more solvent exposed due to the right-handed twist of the β-hairpin, and therefore the thioamide steric bulk is better accommodated. This deviation from expectation based on the environment of the thioamide residue follows our previous observations with protein secondary structures.[29]

Our results reinforce the idea that it is difficult to develop simple rules regarding how thioamide modifications will impact β-sheet structure since specific details such as twists, conformational rigidity or the relative importance of those hydrogen bonding interactions will play a major role. Currently, for utilization of the thioamide as a non-perturbing biophysical probe in fluorescence quenching or CD experiments, we recommend consulting the wild-type protein structure and incorporating the thioamide at β-sheet locations where the residue does not engage as a hydrogen bond acceptor. Although we have observed here that hydrogen bond acceptor positions can be tolerated, they are best avoided until criteria for identifying tolerated positions are determined. The increases in stability observed to date for incorporation as a hydrogen bond donor are minimal and should not significantly alter protein folding. To realize computational models that are predictive of the structural impact of thioamide incorporation at a position in a protein, we will use macrocyclic peptides like those shown to be useful in structure determination here, as well as host-guest studies of peptide/protein complexes, to gather sufficient data for machine learning models similar to those that have been successful in our protease studies.[54] Ultimately, we hope to be able to rationally design peptides containing single or even multiple thioamide substitutions, as well as full-sized proteins synthesized through SPPS and/or native chemical ligation.[3, 5, 30, 33, 76-80]

Author Contributions

K.E.F. synthesized and purified the HPT-Thr$^{S}_{2}$ and HPT-Val$^{S}_{3}$ peptides, collected CD data for all HPT peptides, and acquired NMR data for the HPT-Ile$^{S}_{10}$ and HPT-Leu$^{S}_{11}$-OH peptides. M.P. synthesized and purified the HPT-Ile$^{S}_{10}$, HPT-Leu$^{S}_{10}$-OH, and HPT peptides. C.-H.H. acquired the NMR data for the HPT-Thr$^{S}_{2}$ and HPT-Val$^{s}_{3}$ peptide and initially made assignments. H.V. and B.K. synthesized all YKL peptides, performed CD analysis, collected NMR data (including variable temperature studies), and made initial assignments. K.E.F. processed, assigned, and analyzed all NMR data (using the initial assignments for assistance). S.G.G. performed the structural modelling. K.E.F. and E.J.P. wrote the manuscript with input from all authors.

Conflicts of interest

There are no conflicts to declare.

Acknowledgements

This work was supported by the University of Pennsylvania (Penn) and the National Science Foundation (NSF CHE-1708759 to E.J.P.). K.E.F. thanks the National Institutes of Health (NIH) for funding through the Structural Biology & Molecular Biophysics Training Program (T32-GM-008275). S.G thanks the NSF for funding through the NSF Graduate Research Fellowship Program (DGE-1845298). The University of Pennsylvania Bruker AVANCE NEO 600 MHz NMR spectrometer was supported by NIH supplement awards 3R01GM118510-03S1 and 3R01GM087605-06S1, and the Vagelos Institute for Energy Science and Technology. At Penn, the 600 MHz spectrometer is managed by Dr. Jun Gu and Dr. Chad Lawrence. J.C. acknowledges IISc (MHRD), DST-FIST, UGC-CAS, and DBT-IISc partnership program for funding and infrastructural support. R.C. acknowledges support from the Ministry of Science and Technology in Taiwan (MOST-110-2113-M-002-010). Bruker AVIII 800 MHz spectrometer in the Instrumentation Center at National Taiwan University was supported by the Ministry of Science and Technology in Taiwan (MOST-110-2731-M-002-001). At NTU, the 800 MHz spectrometer is managed by Dr. Shing-Jong Huang. K.E.F. acknowledges Dr. Chad Lawrence for his assistance with acquiring the NMR data at Penn.

Notes and references

‡ For NMR experiments, the YKL peptides were dissolved in 100 mM sodium deuteroacetate buffer pH 3.8 (9:1 v/v H_2O/ D_2O). The HPT peptides were dissolved in 50 mM sodium deuteroacetate pH 5.5 (9:1 v/v H_2O/ D_2O) or 50 mM NaH_2PO_4 pH 5.5 (9:1 v/v H_2O/D_2O). Variability in solubility based on construct and experimental time required for data collection is the reason different buffers and pH values were used. Since the buffer remained consistent for the Test, Unfolded control, and Folded control peptides of each HPT thioamide position, the difference in salt should have minimal to no effect on fraction folded and ΔΔG analysis. Discussion of NMR collection across the different universities can be found in ESI, **Table S3** and **Fig. S15**.

1. F. G. Bordwell, D. J. Algrim and J. A. Harrelson, *J. Am. Chem. Soc.*, 1988, **110**, 5903-5904.
2. E. J. Petersson, J. M. Goldberg and R. F. Wissner, *Phys. Chem. Chem. Phys.*, 2014, **16**, 6827-6837.
3. R. F. Wissner, S. Batjargal, C. M. Fadzen and E. J. Petersson, *J. Am. Chem. Soc.*, 2013, **135**, 6529-6540.
4. J. M. Goldberg, R. F. Wissner, A. M. Klein and E. J. Petersson, *Chem. Commun.*, 2012, **48**, 1550-1552.
5. S. Batjargal, Y. J. Wang, J. M. Goldberg, R. F. Wissner and E. J. Petersson, *J. Am. Chem. Soc.*, 2012, **134**, 9172-9182.
6. J. M. Goldberg, X. Chen, N. Meinhardt, D. C. Greenbaum and E. J. Petersson, *J. Am. Chem. Soc.*, 2014, **136**, 2086-2093.
7. J. M. Goldberg, S. Batjargal and E. J. Petersson, *J. Am. Chem. Soc.*, 2010, **132**, 14718-14720.
8. J. M. Goldberg, L. C. Speight, M. W. Fegley and E. J. Petersson, *J. Am. Chem. Soc.*, 2012, **134**, 6088-6091.
9. M. Hollósi, E. Kollát, J. Kajtár, M. Kajtár and G. D. Fasman, *Biopolymers*, 1990, **30**, 1061-1072.
10. T. Sifferlen, M. Rueping, K. Gademann, B. Jaun and D. Seebach, *Helv. Chim. Acta*, 1999, **82**, 2067-2093.

11. J. Zhao, D. Wildemann, M. Jakob, C. Vargas and C. Schiene-Fischer, *Chem. Commun.*, 2003, **22**, 2810-2811.
12. J. Helbing, H. Bregy, J. Bredenbeck, R. Pfister, P. Hamm, R. Huber, J. Wachtveitl, L. De Vico and M. Olivucci, *J. Am. Chem. Soc.*, 2004, **126**, 8823-8834.
13. H. Satzger, C. Root, P. Gilch, W. Zinth, D. Wildemann and G. Fischer, *J. Phys. Chem. B*, 2005, **109**, 4770-4775.
14. H. Bregy, H. Heimgartner and J. Helbing, *J. Phys. Chem. B*, 2009, **113**, 1756-1762.
15. Y. Huang, Z. Cong, L. Yang and S. Dong, *J. Pept. Sci.*, 2008, **14**, 1062-1068.
16. N. Mahanta, D. M. Szantai-Kis, E. J. Petersson and D. A. Mitchell, *ACS Chem. Biol.*, 2019, **14**, 142-163.
17. W. Grabarse, F. Mahlert, S. Shima, R. K. Thauer and U. Ermler, *J. Mol. Biol.*, 2000, **303**, 329-344.
18. D. D. Nayak, N. Mahanta, D. A. Mitchell and W. W. Metcalf, *eLife*, 2017, **6**, 1-18.
19. Z. L. Watson, F. R. Ward, R. Méheust, O. Ad, A. Schepartz, J. F. Banfield and J. H. D. Cate, *eLife*, 2020, **9**, e60482.
20. C. Alemán, *J. Phys. Chem. A*, 2001, **105**, 6717-6723.
21. H.-J. Lee, Y.-S. Choi, K.-B. Lee, J. Park and C.-J. Yoon, *J. Phys. Chem. A*, 2002, **106**, 7010-7017.
22. M. Hollósi, M. Zewdu, E. Kollát, Z. Majer, M. Kajtár, G. Batta, K. Kövér and P. Sándor, *Int. J. Pept. Protein Res.*, 1990, **36**, 173-181.
23. K. A. Jensen, *Arch Pharm Chem Sci Ed*, 1981, **9**, 93-116.
24. E. P. Dudek and G. O. Dudek, *J. Org. Chem.*, 1967, **32**, 823-824.
25. M. Hollósi, Z. Majer, M. Zewdu, F. Ruff, M. Kajtár and K. E. Kövér, *Tetrahedron*, 1988, **44**, 195-202.
26. A. Bondi, *J. Phys. Chem.*, 1964, **68**, 441-451.
27. M. R. Truter, *J. Chem. Soc.*, 1960, **0**, 997-1007.
28. R. Bardi, A. M. Piazzesi, C. Toniolo, O. E. Jensen, R. S. Omar and A. Senning, *Biopolymers*, 1988, **27**, 747-761.
29. C. R. Walters, D. M. Szantai-Kis, Y. Zhang, Z. E. Reinert, W. S. Horne, D. M. Chenoweth and E. J. Petersson, *Chem. Sci.*, 2017, **8**, 2868-2877.
30. B. Khatri, S. Raghunathan, S. Chakraborti, R. Rahisuddin, S. Kumaran, R. Tadala, P. Wagh, U. D. Priyakumar and J. Chatterjee, *Angew. Chem., Int. Ed.*, 2021, **60**, 24870-24874.
31. A. Reiner, D. Wildemann, G. Fischer and T. Kiefhaber, *J. Am. Chem. Soc.*, 2008, **130**, 8079-8084.
32. J. H. Miwa, L. Pallivathucal, S. Gowda and K. E. Lee, *Org. Lett.*, 2002, **4**, 4655-4657.
33. R. W. Newberry, B. VanVeller and R. T. Raines, *Chem. Commun.*, 2015, **51**, 9624-9627.
34. J. H. Miwa, A. K. Patel, N. Vivatrat, S. M. Popek and A. M. Meyer, *Org. Lett.*, 2001, **3**, 3373-3375.
35. B. Khatri, P. Majumder, J. Nagesh, A. Penmatsa and J. Chatterjee, *Chem. Sci.*, 2020, **11**, 9480-9487.
36. R. M. Culik, H. Jo, W. F. Degrado and F. Gai, *J. Am. Chem. Soc.*, 2012, **134**, 8026-8029.
37. J. Spekowius, R. Pfister and J. Helbing, *J. Phys. Chem. B*, 2021, DOI: 10.1021/acs.jpcb.1c03327.
38. F. A. Syud, H. E. Stanger and S. H. Gellman, *J. Am. Chem. Soc.*, 2001, **123**, 8667-8677.
39. H. E. Stanger and S. H. Gellman, *J. Am. Chem. Soc.*, 1998, **120**, 4236-4237.
40. H. E. Stanger, F. A. Syud, J. F. Espinosa, I. Giriat, T. Muir and S. H. Gellman, *Proc. Natl. Acad. Sci. U. S. A.*, 2001, **98**, 12015.
41. S. J. Maynard, A. M. Almeida, Y. Yoshimi and S. H. Gellman, *J. Am. Chem. Soc.*, 2014, **136**, 16683-16688.
42. M. A. McMechen, E. L. Willis, P. C. Gourville and C. Proulx, *Molecules*, 2019, **24**.
43. F. A. Syud, J. F. Espinosa and S. H. Gellman, *J. Am. Chem. Soc.*, 1999, **121**, 11577-11578.
44. B. L. Sibanda and J. M. Thornton, *Nature*, 1985, **316**, 170-174.
45. C. Chothia, *J. Mol. Bio.*, 1973, **75**, 295-302.
46. D. S. Wishart, B. D. Sykes and F. M. Richards, *Biochemistry*, 1992, **31**, 1647-1651.
47. L.-H. Kuo, J.-H. Li, H.-T. Kuo, C.-Y. Hung, H.-Y. Tsai, W.-C. Chiu, C.-H. Wu, W.-R. Wang, P.-A. Yang, Y.-C. Yao, T. W. Wong, S.-J. Huang, S.-L. Huang and R. P. Cheng, *Biochemistry*, 2013, **52**, 7785-7797.
48. H.-T. Kuo, S.-L. Liu, W.-C. Chiu, C.-J. Fang, H.-C. Chang, W.-R. Wang, P.-A. Yang, J.-H. Li, S.-J. Huang, S.-L. Huang and R. P. Cheng, *Amino Acids*, 2015, **47**, 885-898.
49. H.-T. Kuo, C.-J. Fang, H.-Y. Tsai, M.-F. Yang, H.-C. Chang, S.-L. Liu, L.-H. Kuo, W.-R. Wang, P.-A. Yang, S.-J. Huang, S.-L. Huang and R. P. Cheng, *Biochemistry*, 2013, **52**, 9212-9222.
50. C.-H. Huang, T. W. Wong, C.-H. Yu, J.-Y. Chang, S.-J. Huang, S.-L. Huang and R. P. Cheng, *Molecules*, 2021, **26**.
51. B. J. Lampkin and B. VanVeller, *J. Org. Chem.*, 2021, **86**, 18287-18291.
52. D. R. Artis and M. A. Lipton, *J. Am. Chem. Soc.*, 1998, **120**, 12200-12206.
53. S. Chaudhury, S. Lyskov and J. J. Gray, *Bioinformatics*, 2010, **26**, 689-691.
54. S. Giannakoulias, S. R. Shringari, C. Liu, H. A. T. Phan, T. M. Barrett, J. J. Ferrie and E. J. Petersson, *J. Phys. Chem. B.*, 2020, **124**, 8032-8041.
55. J. Yoon, J. J. Ferrie and E. J. Petersson, *J. Phys. Chem. B.*, 2020, **124**, 10653-10662.
56. H. A. T. Phan, S. G. Giannakoulias, T. M. Barrett, C. Liu and E. J. Petersson, *Chem. Sci.*, 2021, **12**, 10825-10835.
57. C. A. Smith and T. Kortemme, *J. Mol. Biol.*, 2008, **380**, 742-756.
58. R. M. Hughes, M. L. Benshoff and M. L. Waters, *Chem. – Eur. J.*, 2007, **13**, 5753-5764.
59. R. M. Hughes and M. L. Waters, *J. Am. Chem. Soc.*, 2006, **128**, 12735-12742.
60. J. C. Ma and D. A. Dougherty, *Chem. Rev.*, 1997, **97**, 1303-1324.
61. C. D. Tatko and M. L. Waters, *Protein Sci.*, 2003, **12**, 2443-2452.
62. L. M. Salonen, M. Ellermann and F. Diederich, *Angew. Chem., Int. Ed.*, 2011, **50**, 4808-4842.
63. C. D. Tatko and M. L. Waters, *Protein Sci.*, 2004, **13**, 2515-2522.
64. A. J. Riemen and M. L. Waters, *J. Am. Chem. Soc.*, 2009, **131**, 14081-14087.
65. A. J. Riemen and M. L. Waters, *Org. Biomol. Chem.*, 2010, **8**, 5411-5417.
66. G. A. Lengyel and W. S. Horne, *J. Am. Chem. Soc.*, 2012, **134**, 15906-15913.
67. G. A. Lengyel, Z. E. Reinert, B. D. Griffith and W. S. Horne, *Org. Biomol. Chem.*, 2014, **12**, 5375-5381.
68. G. A. Lengyel, G. A. Eddinger and W. S. Horne, *Org. Lett.*, 2013, **15**, 944-947.
69. A. J. Maynard, G. J. Sharman and M. S. Searle, *J. Am. Chem. Soc.*, 1998, **120**, 1996-2007.

70. C. S. Colley, S. R. Griffiths-Jones, M. W. George and M. S. Searle, *Chem. Commun.*, 2000, **7**, 593-594.
71. S. R. Griffiths-Jones, A. J. Maynard and M. S. Searle, *J. Mol. Biol.*, 1999, **292**, 1051-1069.
72. Tran, J. Zeng, H. Treutlein and A. W. Burgess, *J. Am. Chem. Soc.*, 2002, **124**, 5222-5230.
73. H. Verma, B. Khatri, S. Chakraborti and J. Chatterjee, *Chem. Sci.*, 2018, **9**, 2443-2451.
74. T. T. Tran, A. W. Burgess, H. Treutlein and J. Zeng, *J. Mol. Graphics Modell.*, 2001, **20**, 245-256.
75. T. T. Tran, H. Treutlein and A. W. Burgess, *J. Comput. Chem.*, 2001, **22**, 1026-1037.
76. Y. Huang, J. J. Ferrie, X. Chen, Y. Zhang, D. M. Szantai-Kis, D. M. Chenoweth and E. J. Petersson, *Chem. Commun.*, 2016, **52**, 7798-7801.
77. C. R. Walters, J. J. Ferrie and E. J. Petersson, *Chem. Commun.*, 2018, **54**, 1766-1769.
78. S. Batjargal, Y. Huang, Y. J. Wang and E. J. Petersson, *J. Pept. Sci.*, 2014, **20**, 87-91.
79. Y. J. Wang, D. M. Szantai-Kis and E. J. Petersson, *Org. Biomol. Chem.*, 2015, **13**, 5074-5081.
80. B. Khatri, P. Bhat and J. Chatterjee, *J. Peptide Sci.*, 2020, **26**, e3248.