PCCP

Accepted Manuscript

This is an *Accepted Manuscript*, which has been through the Royal Society of Chemistry peer review process and has been accepted for publication.

Accepted Manuscripts are published online shortly after acceptance, before technical editing, formatting and proof reading. Using this free service, authors can make their results available to the community, in citable form, before we publish the edited article. We will replace this *Accepted Manuscript* with the edited and formatted *Advance Article* as soon as it is available.

You can find more information about *Accepted Manuscripts* in the **Information for Authors**.

PCCP

ARTICLE

Nature of aryl-tyrosine interactions contribute to β-hairpin scaffold stability: NMR evidence for alternate ring geometry

Kamlesh Madhusudan Makwana[a] and Radhakrishnan Mahalakshmi[a]*

The specific contribution of the acidic-aromatic β-sheet favouring amino acid tyrosine to the stability of short octapeptide β-hairpin structures is presented here. Solution NMR analysis in near-apolar environments suggests the energetically favourable mode of interaction to be T-shaped face-to-edge (FtE), and that a Trp-Tyr interacting pair is the most stabilizing. Alternate aryl geometries also exist in solution, which readily equilibrate between a preferred $\pi...\pi$ conformation to an aromatic-amide conformation, without any change in the backbone structure. While the phenolic ring is readily accommodated in the 'edge' of FtE aryl interactions, it exhibits an overall lowered contribution to scaffold stability in the 'face' orientation. Such differential tyrosine interactions are key to its dual nature in proteins.

1. Introduction

Aromatic amino acids are crucial for protein stability. It is well established that hydrophobic non-covalent interactions in the protein core cause tight packing among bulky buried aliphatic and aryl ring systems.[1] In such environments, aromatic rings display a multitude of interactions ($\pi...\pi$, N-H...π, C-H...π, O-H...π, hydrogen bonds), which are important for protein folding, as well as chemical and biological recognition.[1a, 1b, 2] The diversity of such non-covalent forces has gathered increasing attention to their role in the formation of ordered amyloid aggregates.[3]

Studies involving phenylalanine and tyrosine residues in short peptide model systems, suggest that "*favourable conformations of aromatic interaction*" can be a rate limiting factor in amyloid formation.[3-4] Indeed, a Phe-Tyr interaction is known to accelerate amyloid formation in the islet amyloid polypeptide.[5] This is not surprising, considering that tyrosine is the least soluble amino acid, promoting its aggregation properties.[6] In contrast, the polar hydroxyl moiety of tyrosine exhibits extensive hydrogen bonding geometries,[6] allowing this amino acid to establish strong intra- and inter-protein interactions. Tyrosine, therefore, displays dual nature (hydrophobic and polar) in proteins.

Of the three aryl residues, tyrosine is of particular interest, due to its preferential localization at the interface,[7] bridging the protein core with the charged exterior. Moreover, the co-existence (chemically) of the acidic hydroxyl moiety and the hydrophobic π-ring system in the phenolic side chain of tyrosine,[8] leads to ~1.4 kcal/mol favourable increase in the system free energy over phenylalanine.[1c, 6, 9] At the polar protein surface, the stacked or parallel displaced arrangement of aryl pairs is preferred over the more popular T-shaped

geometry seen in the hydrophobic core.[7] Hence, precedence for both modes of aromatic interactions (displaced and T-shaped) must be available for tyrosine, and would depend on the local polarity experienced by the phenolic ring in the folded scaffold.

Most studies of aryl interactions in peptides are directed towards Trp-Trp pairs, since it is believed to impart the highest stability to structural scaffolds.[10] However, we have recently demonstrated that the heterologous Trp-aryl pairs are more stabilizing,[11] suggesting that a context-dependent evaluation of aromatic interactions is necessary. The ability of tyrosine to adopt multiple geometries with comparable energy minima *in vacuo*[7] makes it intriguing to examine such interactions experimentally. Moreover, aryl pairs involving tyrosine are most frequent in proteins.[1a, 8] Hence, a detailed analysis that addresses the preferred association modes of the phenolic ring is crucial to understand the dual environment-dependent behaviour of this residue.

Here, we have addressed the involvement of tyrosine in forming homologous and heterologous aryl interactions. As aromatic pairs reinforce β-sheet structures,[1a, 1b, 12] we chose short β-hairpin scaffolds, stapled at one end by the tight type II' turn-forming $^{\text{D}}$Pro-Gly unit,[12c, 13] in our study. Our results provide novel insight on the conformations and energetics of intramolecular interaction geometries involving tyrosine. Studying such interactions is important to further our understanding of amyloid nucleation and its aggregation kinetics, and may help in designing potential inhibitors for these pathological processes.

2. Experimental Methods

2.1 Peptide synthesis and purification. All peptides were synthesized by solid phase synthesis using Fmoc chemistry, as

Table 1. Peptide sequences described in this study.

Peptide Sequence	Code
Ac-L-Y-V-^{D}P-G-L-Y-V-NH$_2$	**YY**
Ac-L-F-V-^{D}P-G-L-Y-V-NH$_2$	**FY**
Ac-L-W-V-^{D}P-G-L-Y-V-NH$_2$	**WY**

reported.[14] Reactions were carried out in dry dimethylformamide and Fmoc removal was achieved using 20% piperidine. Upon completion of the synthesis, peptides were simultaneously deprotected and cleaved from the resin using a TFA: phenol: water: triisopropylsilane cleavage cocktail in the ratio 88:5:5:2, and precipitated using cold ether. Peptides were purified on a C$_{18}$ column using methanol-water gradients. Peptide identities were confirmed by mass spectrometry (Fig. S1, ESI†).

2.2 High resolution NMR measurements. Solution NMR spectra of all peptides were recorded in CD$_3$OH (99.8% D) on a Bruker Avance III 700 MHz NMR spectrometer using a cryoprobe. Homonuclear ^{1}H-^{1}H TOCSY, ROESY and heteronuclear ^{1}H-^{15}N HSQC experiments were obtained using ~3 mM peptide at 303K, using standard pulse sequences available in the Bruker library. Temperature dependence of the proton chemical shifts were measured on a 500 MHz Bruker Avance III spectrometer from 223 K to 323 K at 10 K increments. All spectra were calibrated using TMS (0 ppm) or residual CH$_3$OH resonance at 3.316 ppm. Plots were generated using Topspin v3.0.

2.3 NMR structure calculation. Solution NMR structures were calculated using CYANA v2.1.[15] All observed NOEs were visually classified as strong, medium and weak, based on their intensities, and assigned upper distances limits of 2.5 Å, 3.5 Å, and 5.0 Å, respectively (Table S4, ESI†). The structure was refined further using the hydrogen bonding constraints obtained from the temperature-dependent NMR chemical shifts and the corresponding dδ/dT values as well as the Φ-restraints obtained from the ^{3}J$_{NH-C}{^\alpha}_H$ values derived from the ^{1}H 1D spectrum. The terminal protecting groups (N-terminal acetylation and the C-terminal amidation) were not included in the structure calculation due to limitations in the input library used for the calculations. A total of 100 structures were calculated for each peptide, as described earlier.[11, 14b] Each structure thus calculated had zero violations of van der Waals constraints as well as the distance and angle constraints used in the calculations. Of the calculated 100 structures, the first 35 structures were selected and superimposed using PyMOL.[16] Further, to derive structures (shown in Figs. 5A and 5C) that specifically corresponded to the alternate ring geometries observed in the ROESY spectrum, we segregated the NOEs in Table S4 (ESI†) into two categories, corresponding to either the aryl 2 χ$_1$ of –*gauche* (see Table S3, ESI†) or *trans* (see Table S5, ESI†). A total of 100 structures were calculated for each set and the first structure was rendered using PyMOL in Figs. 5A and 5C.

2.4 Folded population estimation and free energy calculation. Folded fractions (f_F) were calculated using the reported folded (β-sheet) and random coil chemical shifts available in the database.[17] Only strand residues 2, 3 and 6 were considered, as described earlier.[11, 14b] f_F values were used to derive the folding free energy (ΔG^0_F) and equilibrium constant (K_{eq}).[11]

3. Results and discussion

3.1. Aryl-Tyr peptides adopt a β-hairpin conformation.

^{D}Pro-Gly nucleated peptides, with strand-favouring residues such as Val,[12d] have previously been shown to adopt β-hairpin structures in solution. This strategically positions the side chains of residues at the non-hydrogen bonding position in close spatial proximity.[12d, 13b, 13g] Cross-strand aromatic interaction geometry and strength can therefore readily be monitored using such hairpin scaffolds, as demonstrated earlier.[1b, 10a, 11, 13e, 13g, 18] In this study, we positioned Tyr as the 7th residue in designed octapeptides, and examined the effect of the three aromatic residues Phe, Tyr and Trp, placed at position 2, on the backbone scaffold properties and aryl interactions (Table 1).

We chose to study the synthetic peptides in the amphipathic polar solvent methanol, as methanol is reported to destabilize both dispersive hydrophobic (stacked) interactions and T-shaped geometries seen in apolar environments,[7, 19] providing us with an experimental insight on preferred aryl-Tyr interactions. Further, such short peptides are usually disordered in water, and therefore pose difficulties in their structural characterization in this solvent, whereas these sequences are anticipated to be structured in methanol.[10a, 13a, 13e, 14a, 18a, 20] Using high resolution NMR experiments, we probed the presence of signature backbone NOEs that correspond to a folded β-hairpin scaffold for these peptides. The results support the existence of a type II' turn-nucleated β-hairpin conformation in all three peptides (summarized in Fig. 1 and Figs. S2-S4, ESI†).

Temperature coefficients of amide chemical shifts (dδ/dT) are routinely used to probe solvent-exposed *versus* intramolecular hydrogen-bonded amides.[13a, 13e, 14b, 18c, 20] High dδ/dT values for residues 2, 5 and 7 show that these amides are solvent exposed, and display a temperature-dependent chemical shift change, in all three peptides (Fig. 2). Amides of residues 1, 3, 6 and 8 are involved in intra-strand hydrogen bond formation, typically observed in antiparallel β-strands, and therefore show lower dδ/dT values (Fig. 2D). While this interpretation of temperature coefficients is simplistic, this experiment does not distinguish between solvent shielded non-hydrogen bonded amides and their hydrogen bonded counterparts. Furthermore, the data also bears contributions from changes in the folded peptide population with increasing temperature, giving rise to observable non-linearity in some resonances, particularly Val3 (Figs. 2A-C). Nevertheless, correlation coefficients[21] calculated for the three peptides (Fig. S5, ESI†), support our dδ/dT calculations for solvent exposure of amides 2, 5 and 7.

We used the conformational constraints derived from NMR data to calculate the solution structures of these peptides. Structural statistics, along with superposition of the structures calculated in solution, are summarized in Fig. 3; the list of experimental constraints and average backbone torsion angles are provided in Tables S1-S2 (ESI†). Put together, our NMR measurements confirm that all peptides adopt a well-folded β-hairpin structure in methanol. Using this scaffold, we now addressed the interaction preferences of tyrosine with the three aryl groups.
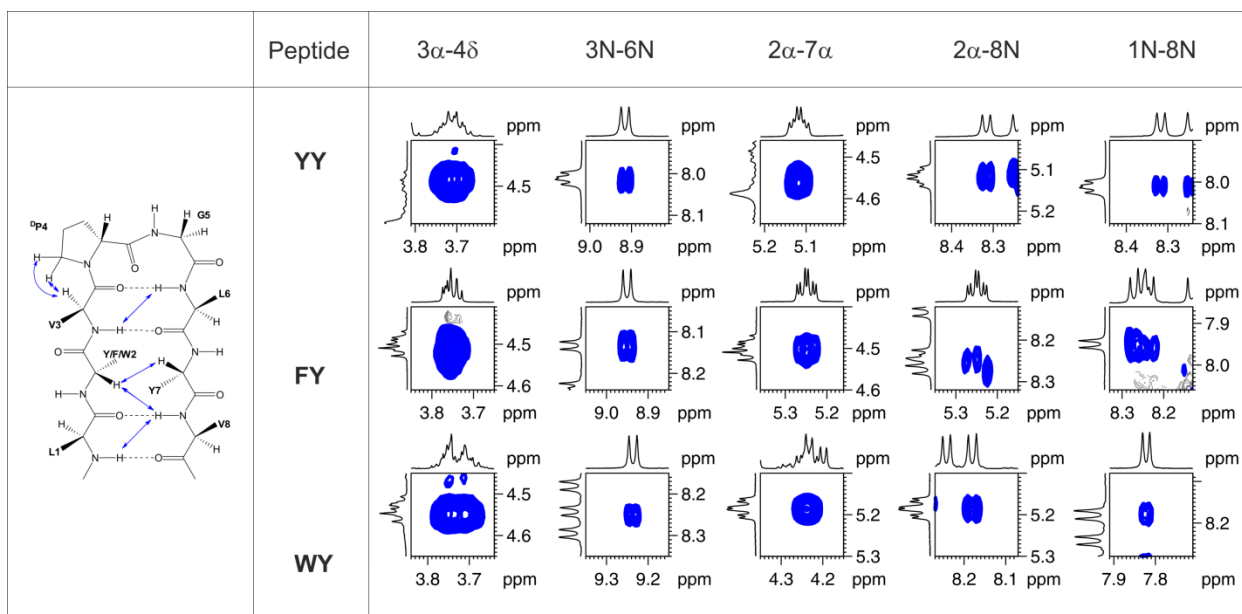
Fig. 1. All peptides adopt a β-hairpin conformation in solution. Diagnostic hairpin NOEs observed in the homonuclear (^{1}H-^{1}H) 2D ROESY spectra recorded in methanol at 303K, across all three peptides. The strong 3α-4δ NOE confirms the presence of a trans Val-^{D}Pro peptide unit necessary to nucleate type II' turn, in a majority of the population. The characteristic 3N-6N and 1N-8N NOEs indicate formation of the 1st and 2nd pair of hydrogen bonds. Along with the 2α-7α NOE, these NOEs are diagnostic to a folded hairpin and the existence of a well-maintained strand registry. The 2α-8N NOE is uncommon in short peptides; observation of this NOE further proves that strand fraying at the termini is minimal. All long-range NOEs are proportionately scaled to the 3α-4δ NOE in each peptide.
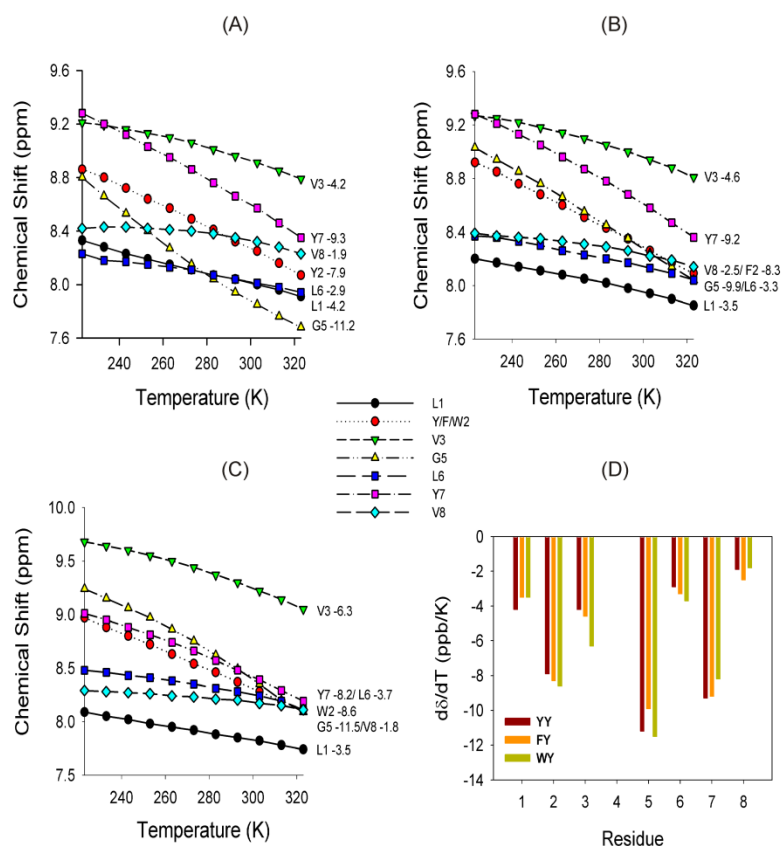

Fig. 2. Temperature-dependent melting curves. Change in backbone amide chemical shift in peptide **YY** (A), **FY** (B) and **WY** (C), with temperature, measured using ^{1}H 1D NMR spectra from 223 K to 323 K at 10 K intervals. (D) Residue-wise plot of temperature coefficients (dδ/dT) indicates higher values for residues 2, 5 and 7 and is a result of solvent-exposed amides. Residues 1, 3, 6 and 8 are involved in intramolecular hydrogen bonds and therefore show lower coefficients.
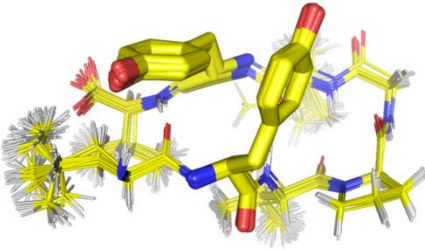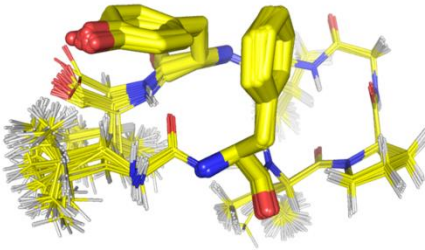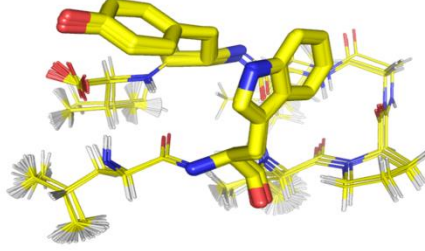
NMR statistics for structure ensemble:	YY	FY	WY
NOE-based distance constraints[1]			
Total	76	78	78
Intraresidue [i = j]	38	38	40
Sequential [|i - j| = 1]	16	15	15
Medium to long range [|i - j| ≥ 3]	22	25	23
Backbone dihedral angle restraints[2]	12	12	12
Hydrogen bond constraints[3]	04	04	04
Structural statistics			
Total # of structures calculated	100	100	100
Violations observed	0	0	0
Global RMSD[4]			
Average backbone RMSD (Å)	0.06 ± 0.04	0.23 ± 0.16	0.02 ± 0.04
Average heavy atom RMSD (Å)	0.54 ± 0.21	0.71 ± 0.26	0.24 ± 0.21

[1] NOE-based distance constraints were obtained from the ROESY spectrum
[2] Angle constraints were obtained from $^3J_{NH-C\alpha H}$ coupling constants from the ^1H 1D spectrum
[3] Hydrogen bond constraints were obtained from the temperature-dependent amide proton chemical shifts
[4] Global RMSDs were calculated for the first 35 structures, using MolMol, and superposed using PyMOL.

Fig. 3. NMR structural statistics and calculated structures. Also see Tables S1-S2 (ESI†).

3.2. Unusual anomalous upfield shifted Y7 C$^\delta$H of WY in FtE interactions.

In isolated peptide β-hairpins, results from our laboratory and others have demonstrated that T-shaped face-to-edge (FtE) or edge-to-face (EtF) tertiary geometries are observed for proximal aryl pairs at the non-hydrogen bonding position.[11, 13e, 14b, 18b, 18c] Due to this T-shaped geometry, the ring protons of the 'edge' aryl group experience shielding by the proximal 'face' electron cloud, resulting in an upfield shift of the 'edge' side chain and backbone resonances. In the ^1H 1D spectra of the peptides described here, we observe a prominent upfield shift in the Y7 C$^{\alpha/\beta}$H and C$^\delta$H resonances, and marginal shift for Y7 C$^\epsilon$H (Figs. 4 and S6-S11, ESI†). This indicates a FtE geometry for the interacting aryl pairs, with Tyr7 occupying the 'edge' and the π2...H-C(π7) interaction established through Y7 C$^\delta$H.

The strength of FtE interactions is determined by the nature and polarity of the interacting aryl pair. Insight into the strength of this aromatic interaction can be obtained deriving the dependence of Y7 C$^\delta$H and C$^\epsilon$H to temperature. While the two resonances are chemically distinguishable, in the presence of a spatially proximal electron cloud, the chemical inequivalence between the Y7 C$^\delta$H and C$^\epsilon$H can be altered. In our peptides, due to the FtE aryl interactions, the aryl 2 electron cloud acts as a shielding agent, causing an unusual upfield shift of the Y7 C$^\delta$H resonance – this is increased further at lower temperatures. At the temperature wherein the C$^\delta$H and C$^\epsilon$H resonances possess the same chemical shift, the chemical inequivalence is abolished, and a singlet is observed.[13e, 22] Fig. 4A and 4B illustrate the temperature at which singlet formation occurs in

YY and FY, respectively (also see Fig. S12, ESI†). Assuming comparable effect of the π-cloud of both Phe and Tyr rings, it may be speculated that the 'singlet' temperature reflects the population of folded β-hairpins in solution, in both peptides. This indicates that a Phe-Tyr pair imparts stronger aromatic contribution as compared to Tyr-Tyr.

Surprisingly, however, singlet formation is not observed in **WY** (Fig. 4C), while the overwhelming NMR evidence points to a well-folded hairpin in this peptide (Figs. 1-3). This anomaly is due to the abnormally upfield shifted Y7 C$^\delta$H resonance, which, even at 323 K, is upfield to the Y7 C$^\epsilon$H (Fig. 4D). Hence, the 'cross-over' singlet formation occurs at temperatures near 350 K (Fig. S13, ESI†), which is higher than methanol boiling point (~338 K), and cannot be experimentally captured. To our knowledge, this is the first instance of an ~0.97 ppm shift (at 303 K) for a Tyr C$^\delta$H resonance in very short peptide hairpins. Similar singlet formation at ~350 K is seen in a dodecapeptide hairpin (peptide name: WYWY) possessing two stabilizing Trp-Tyr pairs, in aqueous buffers.[18c] In both **WY** and the dodecapeptide WYWY,[18c] the T-shaped aryl interaction geometry is retained – that is, the FtE Trp-Tyr interaction involves Trp occupying the 'face' and Tyr as 'edge' in both peptides. Despite differences in solvent contribution to these interactions,[19] our observation of strong T-shaped interactions suggests that **WY** may possess similar stability as longer peptides.

Based on the prominent temperature-dependent upfield shift of the Y7 C$^\delta$H resonance, our studies provide the order Trp>Phe>Tyr, for favoured stabilizing aromatic interactions with the phenolic ring of tyrosine. Further, this also validates
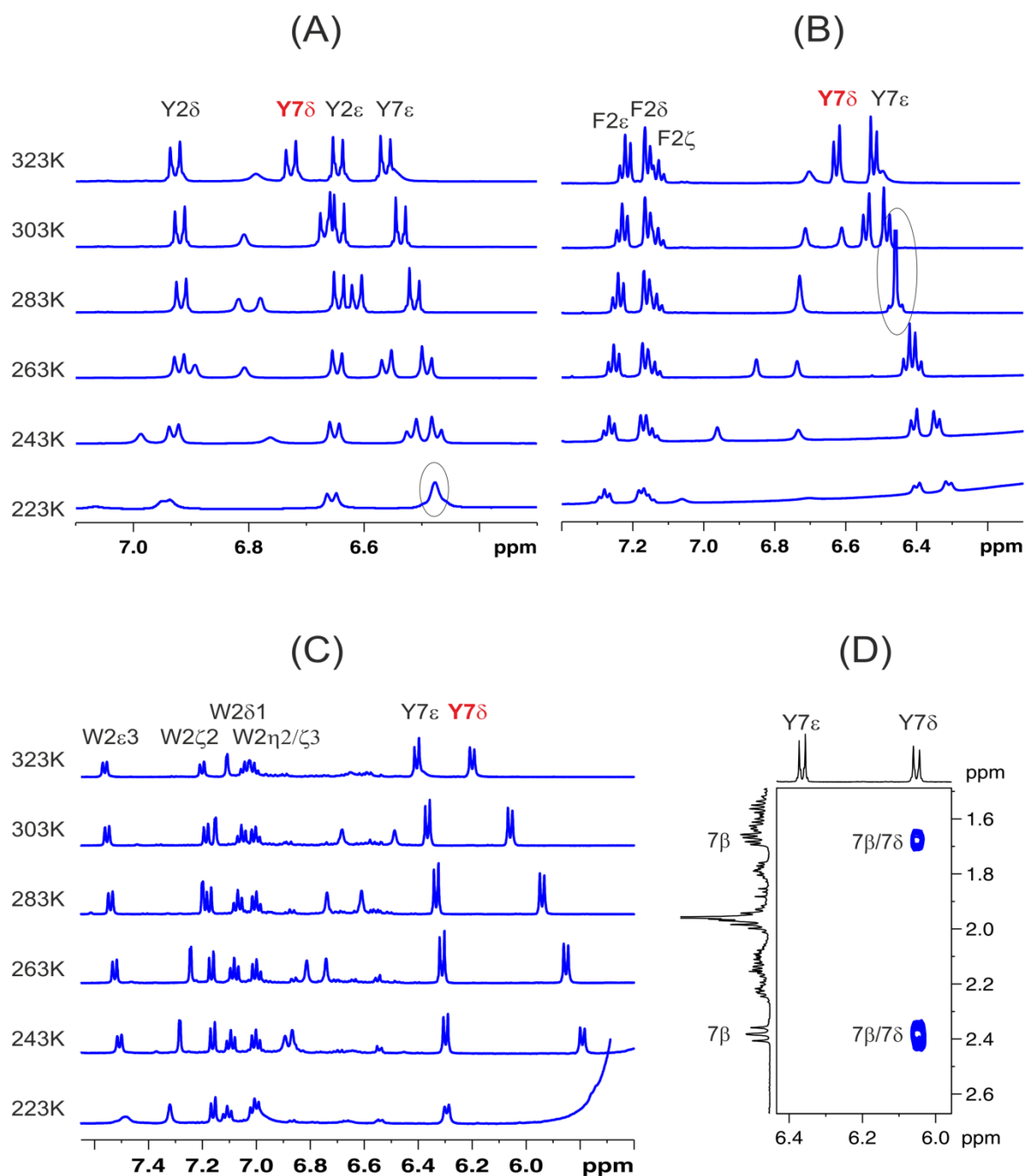
Fig. 4. Ring current effects on Y7δ resonance. Expansion of the aryl resonances, highlighting the temperature dependence of aromatic ring proton chemical shifts in **YY** (A), **FY** (B) and **WY** (C). A face-to-edge geometry between the interacting aryl rings of Y/F/W2 and Y7 results in an upfield temperature-dependent shift of Y7δ. The surprising merging of Y7δ and Y7ε proton chemical shifts arises from degeneracy of the two resonances giving rise to a singlet. This is observed at 223 K for **YY** (A) and at 283 K for **FY** (B), and is circled. In **WY** (C), the resonance degeneracy occurs at much higher temperatures and cannot be captured experimentally. The result is an anomalous Y7δ shift (marked in red) that is *upfield* to the Y7ε even at 323 K (also see Figs. S12-S13, ESI†). Confirmation of this anomalous resonance assignment is provided in the partial expansion of the ROESY spectrum of this peptide recorded at 303 K (D).

the existence of favourable FtE T-shape aromatic interactions even in an organic solvent, where solvent-driven forces are minimal. However, conformational flexibility in short peptides is not uncommon, and is indeed crucial for biological activity of peptides and proteins.[23] We therefore probed for the presence of alternate aryl-Tyr interaction modes.

3.3. Alternate aryl ring geometries stabilized by favourable interactions with the turn

Under ambient temperatures, it is well-recognized that octapeptide hairpins exhibit innumerable, conformationally
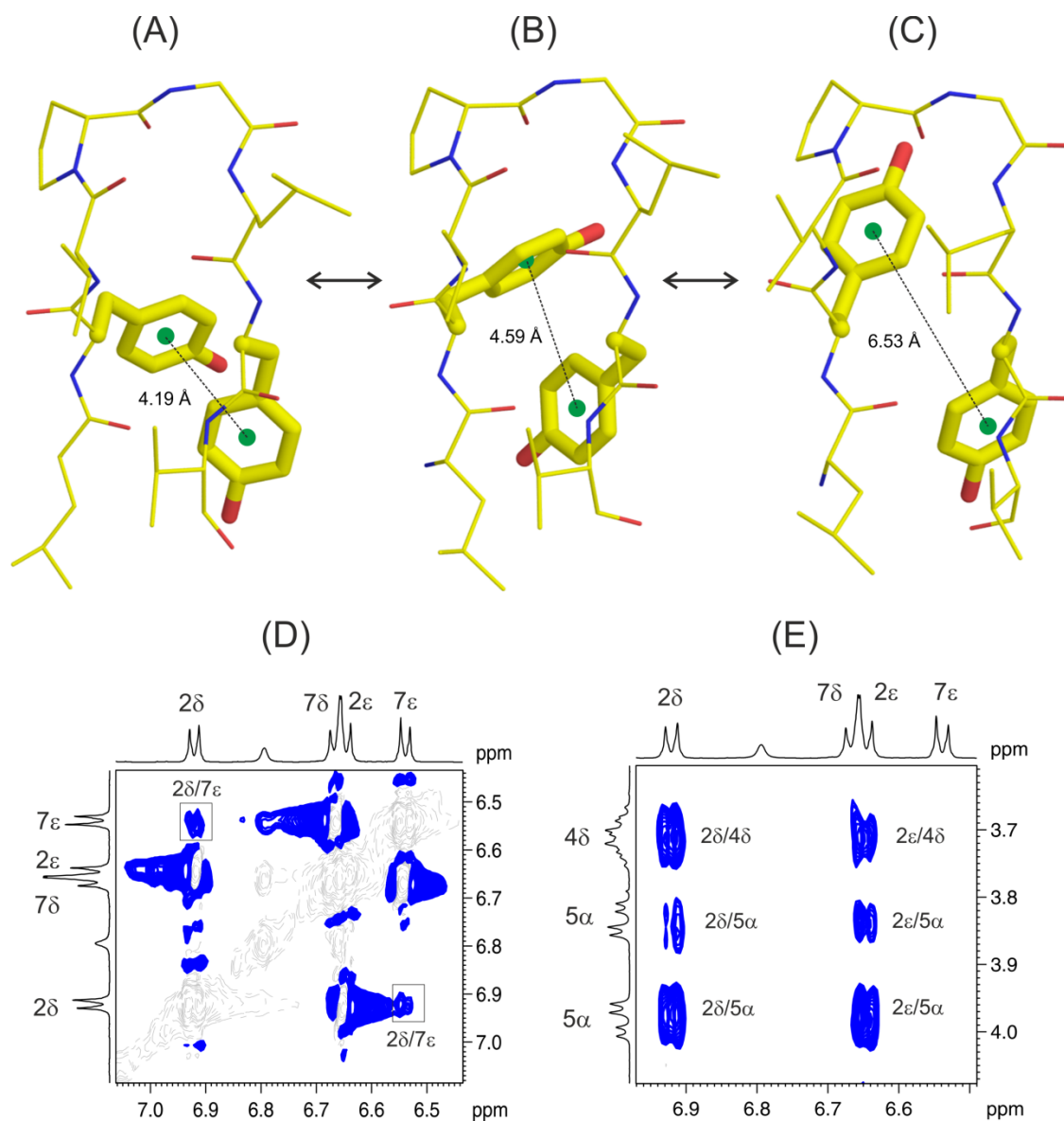
Fig. 5. Ring geometry inter-conversion in solution. The two possible χ_1 geometries seen to occur in solution for ring 2, namely, the ideal $-g$ (A) and *trans* (C) arrangements, are depicted here. NOE evidence for existence of these structures are provided in D (for A) and E (for C). The calculated NMR structure is averaged between the two geometries, and is shown in B. This structure possesses a disallowed χ_1 value for aryl 2. Note that the interconversion between (A) and (C) only requires a change in the orientation of aryl 2, with no significant alteration in aryl 7 orientation. While strong face-to-edge interactions favour structure A, the structure C is stabilized by aromatic-amide interactions involving the peptide unit of residues 2 and 3. The switch from A to C causes an increase in the centroid distances between the aryl groups (marked in the structures), and represent weakening of aromatic interactions from A→ C. The calculated NMR structures of **YY** is shown here as a representative example.

allowed geometries in solution. These populations are under rapid equilibrium in NMR timescales and cannot be readily demarcated. However, when we examined the ring orientations and the corresponding dihedral angles of the two interacting rings, we observed a surprising deviation from the anticipated −*gauche* (−*g*, −60°) χ_1 of aryl 2 (Fig. 5A),[24] to the observed ~ -120° in our calculated structures (Fig. 5B), a value that is conformationally disallowed for an aryl χ_1 (Y/F/W). This clearly suggests that in addition to the detectable peptide population possessing T-shaped aromatic interactions, a second population could exist, which possesses a stereochemically allowed *alternate* geometry for the aryl 2 side chain. Our

spectra therefore possess NOEs arising from at least two stable peptide conformers observable under NMR timescales.

To further explore this anomaly, we systematically examined NOEs between the aryl 2 ring with Y7 and other backbone protons (Fig. 5). All three peptides display NOEs between the aryl resonances (**YY** is illustrated in Fig. 5D), as well as upfield shifted Y7 C$^\delta$H resonance, which indicates close ring proximity. This is feasible only with a χ_1 of −*g*. However, we also obtain NOEs between aryl 2 and the turn residues (Fig. 5E), giving us a χ_1 of 180° (*trans*, *t*) (Fig. 5C). This population does not possess aromatic interactions, but is stabilized by multiple weak interactions of the individual aryl rings with
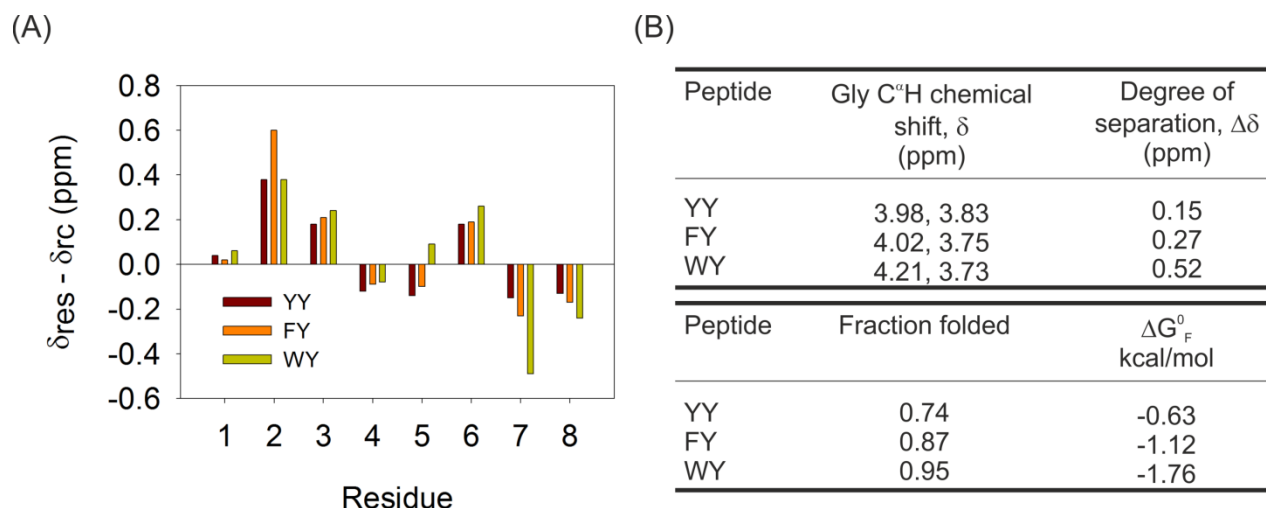
(A)

(B)

Peptide	Gly $C^{\alpha}H$ chemical shift, δ (ppm)	Degree of separation, $\Delta\delta$ (ppm)
YY	3.98, 3.83	0.15
FY	4.02, 3.75	0.27
WY	4.21, 3.73	0.52

Peptide	Fraction folded	ΔG^0_F kcal/mol
YY	0.74	-0.63
FY	0.87	-1.12
WY	0.95	-1.76

Fig. 6. Chemical shift indexing and extent of folding across **YY**, **FY** and **WY**. (A) Plot comparing the residue-wise variation in $C^{\alpha}H$ chemical shift of a given residue (δ_{res}) with respect to the corresponding random coil chemical shift (δ_{rc}). (B) Extent of Gly $C^{\alpha}H$ splitting (top); larger $\Delta\delta$ may represent a higher folded population. Fraction folded values and free energy calculation (bottom) derived from backbone $C^{\alpha}H$ chemical shifts of residues 2, 3 and 6. Calculated free energy values (ΔG^0_F) are similar to results obtained from proteins,[6, 9-10, 13e, 18c] suggesting that identical contributions are made by aryl interactions in different solvent systems and the observed ΔG^0_F values are dependent on the specific interacting aryl moieties.

spatially proximal backbone atoms.

Our data indicates that in aryl-Y interactions, alternative tertiary aryl geometries that do not necessarily involve aromatic interactions are likely. Our calculated NMR structure shown in Fig. 5B is therefore an ensemble of at least two stable and allowed conformations for aryl 2. Co-existence of these two distinct locally stabilized positions for aryl 2 gives rise to our calculated χ_1 with a value intermediate between $-g$ (Fig. 5A) and t (Fig. 5C). Furthermore, the availability of both structures (Figs. 5A and 5C) reveals a simple mechanism of interconversion between the two forms through transient intermediates (such as Fig. 5B) by rotation along χ_1. Another such structural interconversion has previously been documented crystallographically for a Phe-Phe interaction.[20]

Peptides with Tyr in T-shaped interactions are previously known to exhibit complex interactions;[18c] however, our observation of alternate aryl 2 χ_1 can imply the presence of aromatic-amide interactions near the turn, which is believed to be more favourable in a χ_1 of *trans*,[12b, 13e] and can provide local stabilization, especially in β-sheets.[25]

3.4. Folded populations and rank order of peptides offer WY as most stable

Since we observe two alternate geometries in all three peptides, such interactions seem independent of the sequence. However, the strength of either interaction ($\pi...\pi$ *versus* aromatic-amide) would be strongly influenced by the chemical nature of aryl 2, which would, in turn, decide (i) the aryl-Tyr interaction strength and (ii) aromatic-amide interaction energy. We therefore examined the contribution of these interactions on the overall scaffold stability. Comparison of the backbone $C^{\alpha}H$ chemical shifts shows larger positive values for the strand segments for **WY** (Fig. 6A). Similarly, folded fractions (f_F) and corresponding free energy (ΔG^0_F),[11] estimated for the three peptides (Fig. 6B), are concurrent with the rank order **WY**>**FY**>**YY**. We do note that such calculations could also be influenced by anomalous chemical shifts resulting from the

presence of the proximal π-cloud from the interacting aryl rings. Hence, our observed rank order might indeed reflect the strengths of aryl interactions in solution, and not necessarily the folded peptide populations. By and large, assessment of the global NMR parameters for the various resonances including Gly splitting,[18b] indicate that the higher aromatic interaction strengths correlate well with the overall higher f_F in **WY**, when compared with the other two peptide hairpins.

Strong indole interactions with the phenolic group thereby give rise to a well-folded hairpin population in **WY**. Calculated average χ_1 values from the peptide structures follows -122 (**YY**), -115 (**FY**) and -106 (**WY**), suggesting that populations with $\pi...\pi$ interactions are marginally greater in **WY**. This matches the extent of upfield shift of Y7 $C^{\delta}H$ proton (section 3.2), and is in good agreement with the observed rank order. Our observation is also supported by previous database analyses of protein structures, wherein multiple interaction forces ($\pi...\pi$, N-H...O, N...H-O, $\pi...H-O$) are observed between these two aromatic rings.[26] Tryptophan is also known to form strong aromatic-amide interactions,[12b] and it can be envisioned that the combined indole-phenol and indole-amide interaction energies stabilize **WY** in a non-additive manner.

3.5. Assessing aryl interactions and rank order using far-UV CD thermal melts.

Short peptide sequences that possess aromatic interactions are known to display contributions from exciton coupling in the far-UV region of circular dichroism (CD) spectra.[10a, 13b, 13f, 27] As a result, it is exceedingly difficult to derive meaningful information on the secondary structure content of these peptide hairpins using the negative maximum at ~215 nm, usually observed for sheet structures. The magnitude and observed wavelengths of the positive or negative exciton effect depends on the chemical nature of the aryl group, and it is no surprise that the contribution of the indole ring of tryptophan is most predominant among the aromatic amino acids.[27a, 28]

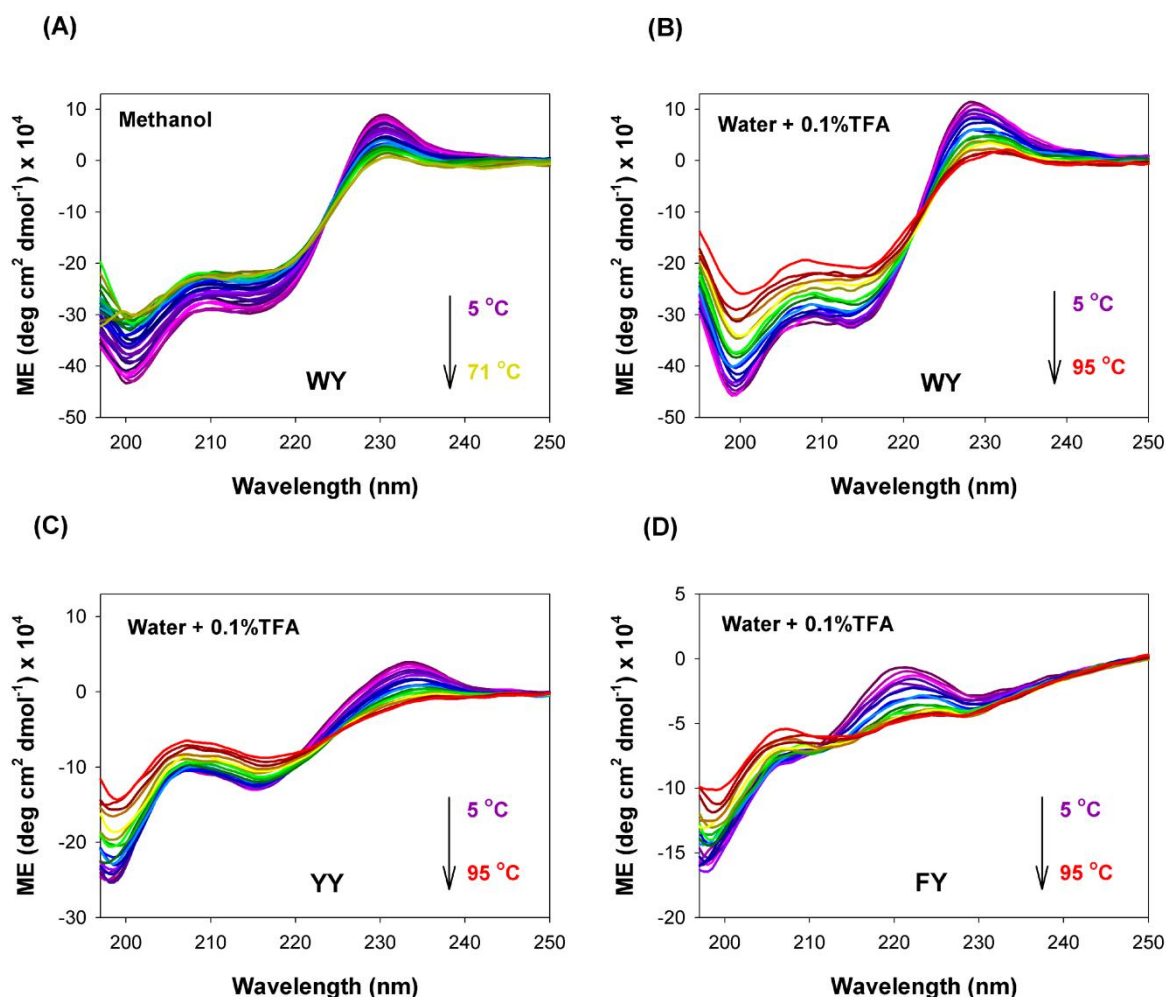CD thermal denaturation measurements that monitor

Fig. 7. Temperature dependence of the far-UV CD spectrum of **WY** in methanol (A) and water containing 0.1% TFA (B). The corresponding spectra for **YY** and **FY** recorded in water (containing 0.1% TFA) are shown in (C) and (D). The characteristic contribution of tryptophan to the CD spectrum, arising from exciton coupling, is observed in (A) and (B), with two negative maxima at ~200 nm and ~212 nm, and a positive couplet at ~228 nm. The extent of change in the CD spectra with temperature for **YY** and **FY** is less than **WY**, suggesting that the contribution of Tyr and Phe to the far-UV CD is considerably lower than Trp. Spectra were recorded from 5 °C to 71 °C in methanol and 5 °C to 95 °C in water/TFA and are colour-coded as purple (low temperature) to red (high temperature). Additional thermal denaturation profiles for **YY** and **FY** in methanol and for all peptides in 80% methanol are presented in Figs. S14-S15 (ESI†). Change in the ME (molar ellipticity) values at 233 nm (**YY**), 221 nm (**FY**) and 228 nm (**WY**) with temperature for data recorded in water (+0.1% TFA) were fitted to a sigmoidal equation to derive the mid-point of the thermal transition, and is presented in Fig. S16 (ESI†).

changes in the indole contribution, to derive ΔG^0_F, have been successful for water-soluble peptides.[2a, 10a, 13d, 13f, 18a-c, 27b] We attempted to measure similar temperature-dependence of the observed CD spectra for our peptides, under select solvent conditions (Figs. S14-S15, ESI†). In **WY**, our measurements reveal that the overall unfolding profile, involving loss of the ~200 nm, ~212 nm and ~228 nm bands observed previously for Trpzip-based sequences,[10a, 18c] is also seen herein (Figs. 7A-B). In our experiments in methanol, we do not observe significant changes in the CD profiles of **YY** and **FY**, which could arise from the poor(er) contributions of Tyr/Phe ring exciton coupling to the CD spectra of these octapeptides (Figs. 7C-D and S15, ESI†).

A two-state dependence to temperature is seen only when the peptides are solubilized in water (containing 0.1% TFA; TFA was added to facilitate peptide solubilization in water). Changes in the ellipticity values to increase in temperature

could be fitted with confidence to a sigmoidal function, in this solvent (Fig. S16, ESI†). The fits provided us with the mid-point of the temperature-dependent transition for aryl contribution to the far-UV CD measurements, as follows: ~41 °C for **YY**, ~37 °C for **FY** and ~49 °C for **WY** (Fig. S16, ESI†). Calculation of this thermal transition provides us with the rank order **WY**>**YY**≈**FY**. While there is an overall agreement of the NMR and CD measurements, since we cannot unambiguously quantify the contributions of the various chromophores to the CD spectrum,[27a] the results must be interpreted with caution.

3.6. C-terminal protecting group influences the observed hairpin population.

We also examined whether differences in the C-terminal modification would influence our observed scaffold

Table 2. Influence of the C-terminal capping group.

Peptide	f_F [#]	ΔG^0_F [#]
Ac-L-Y-V-^{D}P-G-L-Y-V-OCH$_3$ (peptide **1**) [13e]	0.83	-0.95
Ac-L-Y-V-^{D}P-G-L-Y-V-NH$_2$ (**YY**)	0.74	-0.63

[#] f_F = fraction folded; ΔG^0_F = free energy in kcal/mol

Table 3. Positional dependence of Phe↔Tyr interaction from NMR.

Peptide	f_F [#]	ΔG^0_F [#]
Ac-L-Y-V-^{D}P-G-L-F-V-NH$_2$ (YF) [14b]	0.77	-0.72
Ac-L-F-V-^{D}P-G-L-Y-V-NH$_2$ (FY)	0.87	-1.12

[#] f_F = fraction folded; ΔG^0_F = free energy in kcal/mol

characteristics, by comparing the calculated thermodynamic contributions stabilizing **YY** with peptide **1**[13e] (Table 2). Interestingly, we find that a C-terminal methylation significantly enhances peptide stability (~0.3 kcal/mol). This difference in free energy is considerably higher when we take into account the marginal variation in the terminal protecting group. In order to examine the source of this difference, we compared the densitometry analysis of backbone NOE intensities in both peptides. The results are provided in Fig. S18 (ESI†). We observe that the signature 3N↔6N NOE is stronger for **YY**, indicating turn formation and establishment of the first hydrogen bond pair. However, the 2^α↔7^α and 1N↔8N NOE, which correspond to strand propagation and formation of the second hydrogen bond pair between the terminal residues is weakened in this peptide (**YY**). Hence, the differences in free energy values could arise from variations in the hydrogen bonding strengths.

It is likely that methylation sterically occludes solvent molecules more effectively compared to the amide moiety, and therefore shields the Leu1-Val8 hydrogen bonds, resulting in lowered strand fraying at the termini. Alternately, amidation could promote solvation at the termini, and thereby weaken the terminal hydrogen bonds. The influence of the local dipole moment of both terminal groups could also be different in both cases, and could give rise to our observations. Note that in methanol, the C-terminal amide remains polar, yet uncharged, as is evident from the two singlet resonances observed at ~6.5 ppm for the two amide protons (see Figs. S2-S4, ESI†). We believe that methylation would bear a similar influence on the other peptide analogues reported herein. C-terminal therefore acts as a hydrophobic gatekeeper in such peptides. However, amidation is a better mimetic of the peptide unit in proteins, and we believe that the observed thermodynamic values for our peptides are better representatives of aryl interaction energies in proteins.

3.7. Positional effect of aryl interactions involving tyrosine

Finally, we addressed the aryl group positional effect on hairpin characteristics by comparing the NMR-derived properties of **FY** with **YF**[14b] (Table 3), another identical model system studied previously. Earlier reports attempting double-mutant cycles have observed independent contributions of individual residues to strand (and structure) stabilization.[18a] However, comparison of **FY** with **YF** indicates that the Tyr-Phe tertiary interaction is different from Phe-Tyr interaction, although both tend to adopt a T-shaped geometry.[11, 14b] We have reported earlier that subtle variations in the π-cloud interaction geometry affect the biophysical properties of peptide scaffolds,[11] such as the far-UV CD spectra (Fig. S16, ESI†). This has been observed by other groups in different systems.[13b, 18c, 27b, 28a, 29] Similarly, a recent study on amyloid fibrils, utilizing CD, demonstrated the differences in contribution of aromatic side chains to CD, based on their conformation.[28b] Therefore, the relative orientation in three dimensional space demarcates the contribution of either aromatic ring to scaffold free energy, and the additional hydroxyl group differentiates Tyr contributions from that of its homologue Phe.

Differential contributions of the Tyr↔Phe tertiary interactions and evidence for β-hairpin formation across all three sequences described herein suggest that aryl interactions in our peptide systems are independent of peptide folding. This is not surprising, when we consider that ^{D}P-G, used in our peptides, is a strong turn nucleator; hairpin formation is further stabilized by methanol, which is evident from the persistent residual secondary structure in these peptides even at higher temperatures (see CD spectra in Fig. S15, ESI†). Our peptide systems therefore behave in a manner similar to other model peptides studied earlier, wherein aromatic interactions are seemingly a consequence of the β-hairpin structure.[1a, 1b, 8, 30] This is in contrast to observations on Trpzip peptides, wherein cross-strand aromatic interactions are crucial for peptide folding and stability.[10a, 13f, 18a, 18c, 27b, 31] Hence, the extent to which aryl interactions contribute to scaffold stability and protein folding varies contextually.[13d, 32]

4. Conclusions

Our systematic analysis of interaction geometry preferences of an aryl-tyrosine pair at the non-hydrogen bonding position of model peptide β-hairpins indicates that the FtE T-shaped interaction is prevalent even in amphipathic solvents, wherein entropic contributions are minimal compared to aqueous solvents.[13e] The need for such strategically positioned aromatic interactions for the proper folding and stabilization of short secondary structure elements, particularly the β-hairpin scaffold, has been demonstrated using dodecapeptide libraries, both experimentally and *in silico*. These studies together point to a general, largely solvent-independent, T-shaped aromatic interaction geometry in diverse model sequences.

The observation of comparable aryl behaviour of octa- and dodecapeptides (eg., **WY** and WYWY[18c]) indicates that aromatic interaction preferences are preserved in both water and methanol, and their global contribution to the folding and stability of peptide hairpins could bear certain similarities. Furthermore, our study with these peptide model systems in methanol can provide a better picture of aromatic interaction modes in the hydrophobic protein interior. Indeed, a short survey of hairpin structures in proteins from the protein data bank justifies our observation of FtE interaction among aryl-Tyr pairs (Fig. 8).

We also observe, for the first time, a remarkably anomalous upfield shifted resonance for tyrosine in **WY**. While anomalous chemical shifts have been reported earlier for several peptide hairpin models involving tyrosine (examples: [13e, 13g, 18c]), the $C^\delta H$ chemical shift of the Tyr7 ring is upfield to the $C^\epsilon H$ by a significant ~1.0 ppm even at 303 K in our octapeptide hairpin, which is unusual for such short peptides. Observation of deviations in aryl chemical shifts can serve as a signature feature to derive interaction modes and strengths among aryl-Tyr pairs in proteins. Surprisingly, despite the strong aromatic interactions, our observation of alternate aryl side chain geometries that do not involve in aromatic interactions suggests that intramolecular aromatic-amide interactions can also stabilize local aromatic ring conformations and supplement for
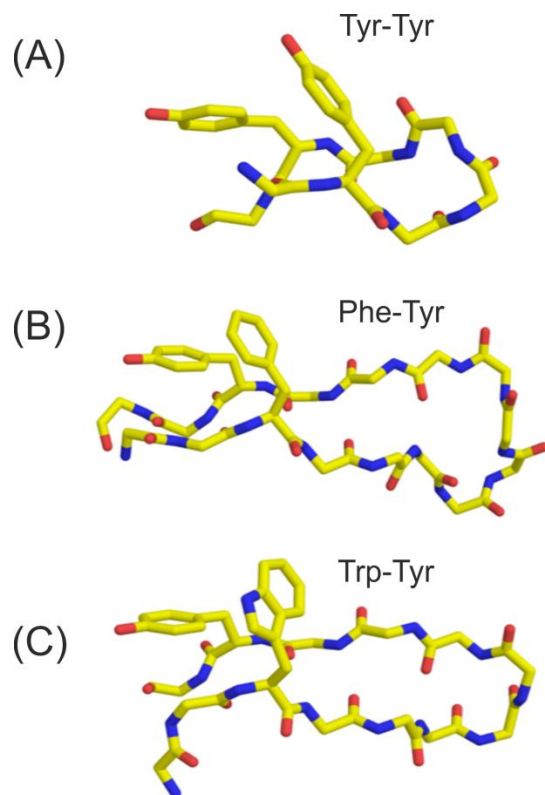
(A)   Tyr-Tyr

(B)   Phe-Tyr

(C)   Trp-Tyr

Fig. 8. Cross-strand FtE aromatic interactions in proteins. Shown here are representative examples of Tyr-Tyr interaction (A) in 1QHO[33], Phe-Tyr interaction (B) and Trp-Tyr interaction (C) in 1H8G[34], at the non-hydrogen bonding position of anti-parallel strands. Coordinates for all the structures were obtained from the Protein Data Bank (www.rcsb.org/pdb).

$\pi...\pi$ interacting forces. Such interactions have a strong positional influence for an interacting aryl pair, supporting previous studies that aromatic interactions stabilize pre-formed scaffolds.[1a, 8, 30a, 35]

Overall, our studies indicate that tyrosine establishes strongest interactions with tryptophan, and exhibits dynamic weaker interactions with phenylalanine and itself. These dynamic interactions may promote the stabilization of alternate geometries for the aryl ring, as seen in our study. Further, the presence of tyrosine promotes alternate polar interactions with the backbone and solvent molecules, which could affect peptide solubilization. It has previously been alluded that such aryl interactions can influence folding pathways of peptide scaffolds or the initial phases of protein folding itself.[13f, 18d, 30c, 30d, 31a, 36] These observations also have great bearing in amyloid nucleation, since studies using model amyloid systems assign diverse roles for aromatics, in fibril assembly.[3-5, 37] Similar conclusions have been made from studies that use Phe interactions (also commonly observed in amyloids) in designed peptides.[4, 5b, 14b, 37b]

In conclusion, our data suggests that Tyr contributes favourably when placed in the 'edge' of an FtE interaction. Our results can shed important insight to our understanding of the role of aromatic interactions occurring on the protein surface or in the hydrophobic core, and in amyloid formation.

Acknowledgements

Notes and references

a Molecular Biophysics Laboratory, Department of Biological Sciences, Indian Institute of Science Education and Research, Bhopal – 462023. India. E-mail: maha@iiserb.ac.in.

† Electronic Supplementary Information (ESI) available: [Supplementary Figures and tables]. See DOI: 10.1039/b000000x/

1. (a) A. Thomas, R. Meurisse, B. Charloteaux and R. Brasseur, *Proteins*, 2002, **48**, 628-634; (b) I. L. Budyak, A. Zhuravleva and L. M. Gierasch, *J. Mol. Biol.*, 2013, **425**, 3522-3535; (c) C. Nick Pace, J. M. Scholtz and G. R. Grimsley, *FEBS Lett.*, 2014, **588**, 2177-2184.

2. (a) M. L. Waters, *Curr. Opin. Chem. Biol.*, 2002, **6**, 736-741; (b) L. M. Espinoza-Fonseca, *Mol. Biosyst.*, 2012, **8**, 237-246.

3. (a) R. Azriel and E. Gazit, *J. Biol. Chem.*, 2001, **276**, 34156-34161; (b) E. Gazit, *FASEB J.*, 2002, **16**, 77-83.

4. J. D. Pham, R. K. Spencer, K. H. Chen and J. S. Nowick, *J. Am. Chem. Soc.*, 2014.

5. (a) P. Marek, A. Abedini, B. Song, M. Kanungo, M. E. Johnson, R. Gupta, W. Zaman, S. S. Wong and D. P. Raleigh, *Biochemistry*, 2007, **46**, 3255-3261; (b) L. H. Tu and D. P. Raleigh, *Biochemistry*, 2012, **52**, 333-342.

6. C. N. Pace, G. Horn, E. J. Hebert, J. Bechert, K. Shaw, L. Urbanikova, J. M. Scholtz and J. Sevcik, *J. Mol. Biol.*, 2001, **312**, 393-404.

7. R. Chelli, F. L. Gervasio, P. Procacci and V. Schettino, *J. Am. Chem. Soc.*, 2002, **124**, 6133-6143.

8. R. Meurisse, R. Brasseur and A. Thomas, *Proteins*, 2004, **54**, 478-490.

9. L. Serrano, M. Bycroft and A. R. Fersht, *J. Mol. Biol.*, 1991, **218**, 465-475.

10. (a) A. G. Cochran, N. J. Skelton and M. A. Starovasnik, *Proc. Natl. Acad. Sci. U. S. A.*, 2001, **98**, 5578-5583; (b) S. M. Butterfield, W. J. Cooper and M. L. Waters, *J. Am. Chem. Soc.*, 2005, **127**, 24-25; (c) C. M. Santiveri and M. A. Jimenez, *Biopolymers*, 2010, **94**, 779-790.

11. K. M. Makwana and R. Mahalakshmi, *ChemBioChem*, 2014, **15**, 2357-2360.

12. (a) P. Y. Chou and G. D. Fasman, *Biochemistry*, 1974, **13**, 211-222; (b) G. Toth, C. R. Watts, R. F. Murphy and S. Lovas, *Proteins*, 2001, **43**, 373-381; (c) C. M. Santiveri, J. Santoro, M. Rico and M. A. Jimenez, *Protein Sci.*, 2004, **13**, 1134-1147; (d) S. T. Phillips, G. Piersanti and P. A. Bartlett, *Proc. Natl. Acad. Sci. U. S. A.*, 2005, **102**, 13737-13742.

13. (a) I. L. Karle, C. Das and P. Balaram, *Proc. Natl. Acad. Sci. U. S. A.*, 2000, **97**, 3034-3037; (b) C. X. Zhao, P. L. Polavarapu, C. Das and P. Balaram, *J. Am. Chem. Soc.*, 2000, **122**, 8228-8231; (c) J. F. Espinosa, F. A. Syud and S. H. Gellman, *Protein Sci.*, 2002, **11**, 1492-1505; (d) R. M. Hughes and M. L. Waters, *Curr. Opin. Struct. Biol.*, 2006, **16**, 514-524; (e) R. Mahalakshmi, S. Raghothama and P. Balaram, *J. Am. Chem. Soc.*, 2006, **128**, 1125-1138; (f) L. Wu, D. McElheny, V. Setnicka, J. Hilario and T. A. Keiderling, *Proteins*,

2012, **80**, 44-60; (g) S. Aravinda, U. S. Raghavender, R. Rai, V. V. Harini, N. Shamala and P. Balaram, *Org. Biomol. Chem.*, 2013, **11**, 4220-4231.

14. (a) K. M. Makwana, S. Raghothama and R. Mahalakshmi, *Phys. Chem. Chem. Phys.*, 2013, **15**, 15321-15324; (b) K. M. Makwana and R. Mahalakshmi, *Org. Biomol. Chem.*, 2014, **12**, 2053-2061.

15. P. Guntert, *Methods Mol. Biol.*, 2004, **278**, 353-378.

16. *The PyMOL Molecular Graphics System, Version 1.2r3pre, Schrödinger, LLC.*

17. E. L. Ulrich, H. Akutsu, J. F. Doreleijers, Y. Harano, Y. E. Ioannidis, J. Lin, M. Livny, S. Mading, D. Maziuk, Z. Miller, E. Nakatani, C. F. Schulte, D. E. Tolmie, R. Kent Wenger, H. Yao and J. L. Markley, *Nucleic Acids Res.*, 2008, **36**, D402-408.

18. (a) S. Russell and A. G. Cochran, *J. Am. Chem. Soc.*, 2000, **122**, 12600-12601; (b) C. D. Tatko and M. L. Waters, *J. Am. Chem. Soc.*, 2002, **124**, 9372-9373; (c) L. Wu, D. McElheny, T. Takekiyo and T. A. Keiderling, *Biochemistry*, 2010, **49**, 4705-4714; (d) M. A. Jimenez, *Methods Mol. Biol.*, 2014, **1216**, 15-52.

19. A. J. Maynard, G. J. Sharman and M. S. Searle, *J. Am. Chem. Soc.*, 1998, **120**, 1996-2007.

20. R. Sonti, R. Rai, S. Ragothama and P. Balaram, *J. Phys. Chem. B*, 2012, **116**, 14207-14215.

21. N. H. Andersen, J. W. Neidigh, S. M. Harris, G. M. Lee, Z. H. Liu and H. Tong, *J. Am. Chem. Soc.*, 1997, **119**, 8547-8561.

22. S. Honda, N. Kobayashi and E. Munekata, *J. Mol. Biol.*, 2000, **295**, 269-278.

23. (a) R. Huber, *Biochem. Soc. Trans.*, 1987, **15**, 1009-1020; (b) Y. Zhang, B. Stec and A. Godzik, *Structure*, 2007, **15**, 1141-1147.

24. A. Rajagopal, S. Aravinda, S. Raghothama, N. Shamala and P. Balaram, *Biopolymers*, 2012, **98**, 185-194.

25. H. Kaur and G. P. Raghava, *FEBS Lett.*, 2004, **564**, 47-57.

26. U. Samanta, D. Pal and P. Chakrabarti, *Acta Crystallogr. D Biol. Crystallogr.*, 1999, **55**, 1421-1427.

27. (a) R. Mahalakshmi, G. Shanmugam, P. L. Polavarapu and P. Balaram, *ChemBioChem*, 2005, **6**, 2152-2158; (b) L. Wu, D. McElheny, R. Huang and T. A. Keiderling, *Biochemistry*, 2009, **48**, 10362-10371.

28. (a) C. D. Andrew, S. Bhattacharjee, N. Kokkoni, J. D. Hirst, G. R. Jones and A. J. Doig, *J. Am. Chem. Soc.*, 2002, **124**, 12706-12714; (b) K. Matsuo, H. Hiramatsu, K. Gekko, H. Namatame, M. Taniguchi and R. W. Woody, *J. Phys. Chem. B*, 2014, **118**, 2785-2795.

29. R. W. Woody, *Biopolymers*, 1978, **17**, 1451-1467.

30. (a) D. M. Chung, Y. Dou, P. Baldi and J. S. Nowick, *J. Am. Chem. Soc.*, 2005, **127**, 9998-9999; (b) S. Enemark and R. Rajagopalan, *Phys. Chem. Chem. Phys.*, 2012, **14**, 12442-12450; (c) S. Enemark, N. A. Kurniawan and R. Rajagopalan, *Sci. Rep.*, 2012, **2**, 649; (d) R. B. Best and J. Mittal, *Proc. Natl. Acad. Sci. U. S. A.*, 2011, **108**, 11087-11092; (e) U. S. Raghavender, *J. Biomol. Struct. Dyn.*, 2013, **31**, 1404-1410.

31. (a) A. Popp, L. Wu, T. A. Keiderling and K. Hauser, *J. Phys. Chem. B*, 2014; (b) Z. Yu, S. Selvam and H. Mao, *Biochemistry*, 2014.

32. M. L. Waters, *Acc. Chem. Res.*, 2013, **46**, 873.

33. Z. Dauter, M. Dauter, A. M. Brzozowski, S. Christensen, T. V. Borchert, L. Beier, K. S. Wilson and G. J. Davies, *Biochemistry*, 1999, **38**, 8385-8392.

34. C. Fernandez-Tornero, R. Lopez, E. Garcia, G. Gimenez-Gallego and A. Romero, *Nat. Struct. Biol.*, 2001, **8**, 1020-1024.

35. C. M. Santiveri, M. Rico and M. A. Jimenez, *Protein Sci.*, 2000, **9**, 2151-2160.

36. (a) Q. Shao, H. Wei and Y. Q. Gao, *J. Mol. Biol.*, 2010, **402**, 595-609; (b) N. A. Kurniawan, S. Enemark and R. Rajagopalan, *J. Am. Chem. Soc.*, 2012, **134**, 10200-10208; (c) C. A. Jimenez-Cruz and A. E. Garcia, *Phys. Chem. Chem. Phys.*, 2014, **16**, 6422-6429.

37. (a) A. H. Armstrong, J. Chen, A. F. McKoy and M. H. Hecht, *Biochemistry*, 2011, **50**, 4058-4067; (b) K. E. Marshall, K. L. Morris, D. Charlton, N. O'Reilly, L. Lewis, H. Walden and L. C. Serpell, *Biochemistry*, 2011, **50**, 2061-2071; (c) J. Adler, H. A. Scheidt, M. Kruger, L. Thomas and D. Huster, *Phys. Chem. Chem. Phys.*, 2014, **16**, 7461-7471.