

Cite this: *Chem. Sci.*, 2016, 7, 5212

## Engineering $\beta$ -sheets employing *N*-methylated heterochiral amino acids†

Dipan Ghosh,‡ Priyanka Lahiri,‡ Hitesh Verma, Somnath Mukherjee and Jayanta Chatterjee\*

There is a lack of functional group diversity in the reverse turn motifs nucleating a  $\beta$ -sheet conformation in designed peptides, proteins and foldamers. The majority of these sequences consist of D-Pro–L-Pro, D-Pro–Gly or Asn–Gly as the turn inducing motif restricting their biological application and physicochemical modulation. In this report, for the first time we elucidate that *N*-methylation of heterochiral amino acids in linear peptides nucleates  $\beta$ -sheet conformation without the necessity of having a ring or covalent constraint at the reverse turn. Our results show that D-Pro can be conveniently substituted by any other *N*-methylated D-amino acid followed by an *N*-methylated L-amino acid or sarcosine to adopt a  $\beta$ II' turn inducing the  $\beta$ -sheet folding. Furthermore, we reveal that a single amino acid either at the  $i + 1$  or  $i + 2$  site of the reverse turn can modulate the right-handed twist, which eventually dictates the extent of the foldedness of the  $\beta$ -hairpin.

Received 2nd February 2016  
Accepted 19th April 2016

DOI: 10.1039/c6sc00518g

www.rsc.org/chemicalscience

### Introduction

The design of  $\beta$ -hairpin forming peptides is of great importance in developing chemical tools that can perturb protein–protein/protein–peptide interactions in a chemical biology context<sup>1</sup> and in developing metabolically stable, well-folded synthetic proteins from a protein engineering perspective.<sup>2</sup> Extensive efforts to understand the design principles underlying the formation of reverse turns leading to the formation of  $\beta$ -hairpin peptides and  $\beta$ -sheets have established mainly two dipeptide motifs, D-Pro–Xaa (Xaa is any L-amino acid, although the preferred ones are L-Pro and Gly)<sup>3</sup> and Asn–Gly, as the most favored  $\beta$ -hairpin nucleators in linear peptides.<sup>4</sup> However, this severely limits the biological applications of  $\beta$ -hairpins, as the functionalization of these turn-inducing sequences either compromises their nucleating efficacy<sup>5</sup> or involves tedious chemical synthesis,<sup>6</sup> resulting in gross neglect towards the development of  $\beta$ -hairpins as chemical tools in contrast to  $\alpha$ -helices.<sup>1b</sup> Thus, to expand the repertoire of reverse turn-inducing motifs we focused on employing *N*-methylated amino acids despite their inherent conformational flexibility.<sup>7</sup>

*N*-Methylation has been thoroughly investigated in cyclic peptides but its conformational impact on linear peptides is not well understood. We were thus keen to study the influence of *N*-

methylation on the turn-inducing residues in linear  $\beta$ -hairpin peptides as there is no such report and the few known facts about the *N*-methylation of the turn residues are conflicting.<sup>8</sup> In the present study, we have thoroughly investigated the influence of *N*-methylation on a varied selection of heterochiral residues in linear peptides using circular dichroism and NMR based structure calculations. We show that *N*-methylation indeed nucleates  $\beta$ -hairpin conformation irrespective of the amino acids present at the  $i + 1$  and  $i + 2$  positions of the  $\beta$ -turn. Furthermore, most of these hairpins are conformationally homogeneous<sup>9</sup> in the NMR time scale in apolar and polar environments in spite of having consecutive *N*-methylated heterochiral amino acids. This indicates the wide applicability of the newly identified turn inducing motif towards peptide/protein engineering studies in lipophilic (e.g. membrane) and hydrophilic (e.g. cytosol) environments.<sup>10</sup>

### Results and discussion

To initially study the influence of *N*-methylation on  $\beta$ -hairpin induction in apolar conditions, we selected the model octapeptide **1**,<sup>11</sup> and substituted the D-Pro ( $i + 1$ ) and L-Pro ( $i + 2$ ) at the reverse turn with various *N*-methylated amino acids (Fig. 1a). We chose to *N*-methylate both the  $i + 1$  and  $i + 2$  residues to minimize the conformational entropy about the reverse turn.<sup>12</sup>

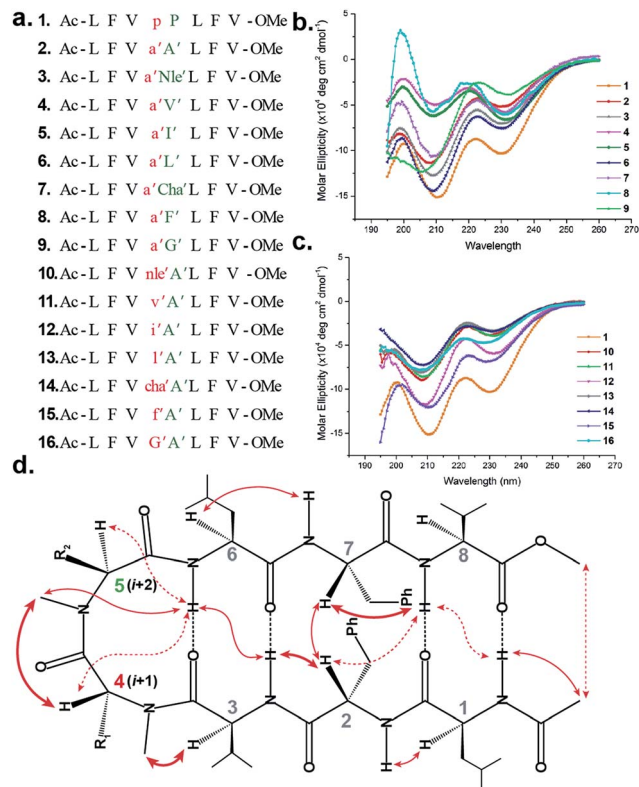
The hydrophobic amino acids: Ala, Val, Ile, Leu and Phe were chosen as substituents since they display the least propensity to occur at the  $\beta$ -turn region, according to Chou–Fasman's analysis of  $\beta$ -turns in protein structures.<sup>13</sup> Besides, these residues also have a minimal stabilizing contribution from the side

Molecular Biophysics Unit, Indian Institute of Science, Bangalore 560012, India.  
E-mail: jayanta@mbu.iisc.ernet.in

† Electronic supplementary information (ESI) available: Detailed experimental procedures, HPLC and MALDI traces, NMR spectra, DMSO- $d_6$  titrations, coupling constants, chemical shifts, NMR structure calculations and overlay of structures. See DOI: 10.1039/c6sc00518g

‡ Equal contribution.





**Fig. 1** (a) Sequences of the model peptides synthesized in this study. The letters in lower case denote D-amino acid and the ' denotes N-methylation. (b) CD spectra of the (i + 2) variants (2–9) and (c) (i + 1) variants (10–16) along with 1 in methanol. (d) The characteristic interresidue backbone NOEs observed in the synthesized peptides. Thick arrows denote distances between 1.8 and 2.2 Å, thin arrows denote 2.3–2.6 Å and dotted arrows denote 2.7–3.5 Å.

chain functionality on the turn conformation in the synthesized peptides. We also chose two unnatural amino acids: norleucine (Nle) which acts as a methionine isostere (methionine also has a low propensity to occur at the turn region) and a ring constrained leucine analog, cyclohexylalanine (Cha). All of the compounds were synthesized on a solid support and the N-methylation of the amino acids was performed on-resin using an optimized protocol.<sup>14</sup> In order to determine whether any of our peptides undergo self-association under the conditions used for spectroscopic evaluation, we examined proton NMR chemical shift data for each peptide as a function of concentration. For the hydrophobic peptides, 1–16, the chemical shifts measured at the concentration employed for 2D NMR studies (1–3 mM) were indistinguishable from the chemical shifts measured after 100-fold dilution. For the hydrophilic peptides, 17–21, the chemical shifts measured at the concentration employed for 2D NMR studies were indistinguishable from the chemical shifts measured after ~20-fold dilution.

The initial qualitative assessment of their potential to form  $\beta$ -hairpins was done using far UV (190–260 nm) CD measurements in methanol. The CD spectra of all of the (i + 2) and (i + 1) (Fig. 1) variants resembled the CD spectrum of 1, showing the characteristic minima at 210 and 232 nm due to the anomalous

behavior arising from the Phe2–Phe7 stacking interaction.<sup>4b,15</sup> This clearly suggested that the replacement of the i + 1 and i + 2 residues in 1 with N-methylated amino acids was not detrimental to the overall topology of the molecule. However, as several compounds showed varying CD intensities with a single amino acid substitution, to understand the underlying cause we calculated their average solution conformation using restraints derived from ROESY. In the following discussion, the D-residues will be denoted in lower case and N-methylation with a prime symbol (').

The end caps in 1–16 were carefully designed, so that if these were well folded, besides the characteristic NOEs (Fig. 1d) one would also expect the NOE between the end caps (NHAc and OMe) in the absence of strand fraying. The <sup>1</sup>H NMR of 1 (in CDCl<sub>3</sub>) showed well dispersed signals in the H<sup>α</sup> and H<sup>N</sup> region with <sup>3</sup>J<sub>H<sup>N</sup>–H<sup>α</sup></sub> > 8 Hz suggestive of a well-defined and extended conformation (Table S1†). The characteristic long-range NOEs Phe2H<sup>α</sup>–Phe7H<sup>α</sup> and NHAc–OMe; strand NOEs Leu1H<sup>α</sup>–Phe2H<sup>N</sup>, Phe2H<sup>α</sup>–Val3H<sup>N</sup>, Leu6H<sup>α</sup>–Phe7H<sup>N</sup> and Phe7H<sup>α</sup>–Val8H<sup>N</sup> along with the turn region NOEs proH<sup>α</sup>–Pro5H<sup>δ1/2</sup> and Pro5H<sup>δ1</sup>–Leu6H<sup>N</sup> with the solvent shielded Leu1, Val3, Leu6 and Val8H<sup>N</sup> (as determined by DMSO titration) (Table S2†) suggest the formation of a  $\beta$ -hairpin conformation. The solution structure of 1 (Fig. 2) revealed the formation of a  $\beta$ -hairpin with a centrally located  $\beta$ II' turn (Table 1).

In 2, the pP motif is substituted by the a'A' motif and has a relatively less restricted  $\Phi_{(i+1)}$  and  $\Psi_{(i+2)}$  at the reverse turn due to the lack of any ring constraint. Furthermore, the tertiary amide bonds have a low rotational barrier and thus 2 can co-exist in *cis/trans* form.<sup>16</sup> However, we were surprised to note that 2 displayed a single conformation in CDCl<sub>3</sub> with the absence of any (i<sup>th</sup>)H<sup>α</sup>–(i + 1<sup>st</sup>)H<sup>α</sup> NOE confirming the presence of an all-trans conformer.<sup>17,18</sup> It was interesting to note that most of the <sup>3</sup>J<sub>H<sup>N</sup>–H<sup>α</sup></sub> values in 2 were higher than the corresponding values in 1. The NOEs ala'H<sup>α</sup>–Ala'5NMe, ala'H<sup>α</sup>–Leu6H<sup>N</sup> and Ala'5NMe–Leu6H<sup>N</sup> along with solvent shielded Leu6H<sup>N</sup> clearly define a  $\beta$ II' turn.<sup>18</sup> The long-range NOEs Leu1H<sup>N</sup>–Val8H<sup>N</sup>, Phe2H<sup>α</sup>–Phe7H<sup>α</sup>, Val3H<sup>N</sup>–Leu6H<sup>N</sup> and NHAc–OMe additionally suggest the anti-parallel strand registry in the molecule. The conformation of 2 (Fig. 2) revealed that the  $\Phi_{(i+1)}$  and  $\Psi_{(i+1)}$  are quite close to the ideal  $\beta$ II' turn (Table 1). Whereas, peptide 2a with an A'A' motif showed the presence of a *cis*-peptide bond between the 4<sup>th</sup> and 5<sup>th</sup> residue<sup>18</sup> and the absence of any characteristic long-range NOEs defining the  $\beta$ -hairpin.

Compound 3, with the long un-branched amino acid norleucine, displays all of the characteristic NOEs (Fig. 1d) with a comparable solvent accessibility of the HNs as observed in 2. The conformation of 3 is quite close to 2 (backbone RMSD of 0.20 Å),<sup>18</sup> with some variation in the dihedrals about the turn-motif. Compounds 4 and 5 both have an N-methylated  $\beta$ -branched amino acid at the i + 2 position and display a similar CD profile although with strikingly low intensity. However, we could assign the characteristic NOEs in both of the compounds, and the high <sup>3</sup>J<sub>H<sup>N</sup>–H<sup>α</sup></sub> suggested the presence of a hairpin conformation. It was interesting to note that in these compounds the Leu1H<sup>N</sup> and Leu6H<sup>N</sup> are comparatively less solvent shielded than in 1, 2 and 3 (Table S2†). In these



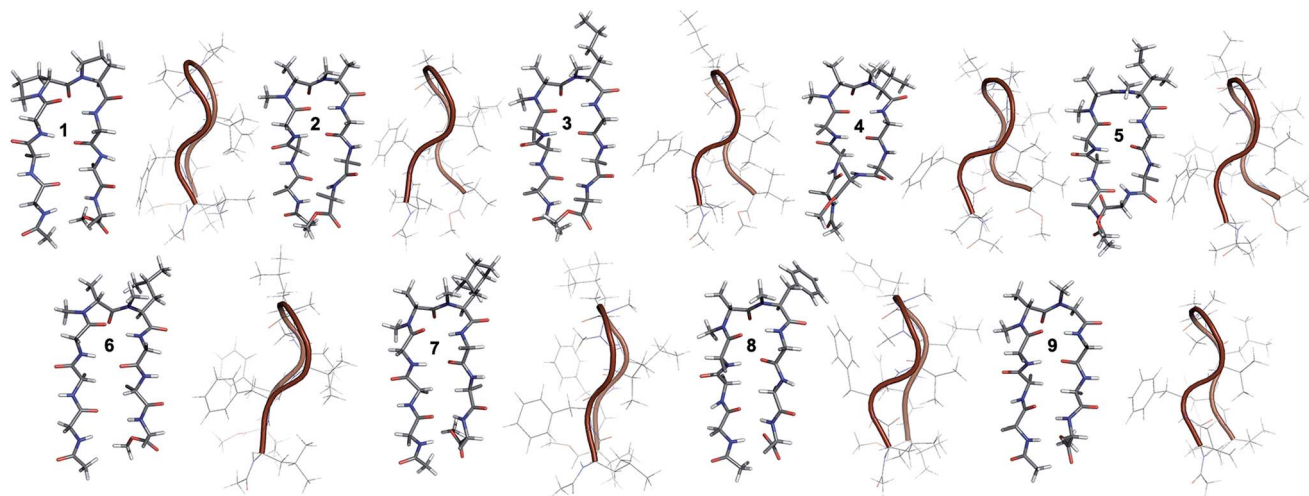


Fig. 2 Solution structure of **1** and all of the  $i + 2$  variants (**2**–**9**) in  $\text{CDCl}_3$ . The front view of the peptide's backbone along with the side chains of the reverse turn motif are shown in a stick representation; all other side chains are restricted to  $\text{C}^\beta$  for the sake of clarity. The edge view is shown in a cartoon representation to depict the right-handed twist.

Table 1 Dihedral angles about the reverse turn in  $i + 2$  variants obtained from the conformations generated by the restrained molecular dynamics simulation

	$i + 1$		$i + 2$	
	$\Phi$ ( $^\circ$ )	$\Psi$ ( $^\circ$ )	$\Phi$ ( $^\circ$ )	$\Psi$ ( $^\circ$ )
1	$50 \pm 14$	$-125 \pm 13$	$-77 \pm 8$	$43 \pm 24$
2	$51 \pm 11$	$-122 \pm 15$	$-106 \pm 14$	$25 \pm 12$
3	$63 \pm 7$	$-139 \pm 6$	$-91 \pm 8$	$23 \pm 12$
4	$63 \pm 8$	$-133 \pm 8$	$-122 \pm 9$	$48 \pm 16$
5	$74 \pm 9$	$-126 \pm 15$	$-113 \pm 12$	$37 \pm 13$
6	$56 \pm 11$	$-131 \pm 9$	$-95 \pm 8$	$14 \pm 10$
7	$66 \pm 11$	$-137 \pm 10$	$-99 \pm 10$	$35 \pm 19$
8	$74 \pm 8$	$-129 \pm 7$	$-99 \pm 9$	$20 \pm 10$
9	$65 \pm 7$	$-127 \pm 6$	$-89 \pm 10$	$15 \pm 8$

$\beta$ -branched analogs, the  $i + 1/i + 2$  and  $i + 2/i + 3$  amide planes are forced away from each other due to the steric repulsion between Val<sup>5</sup>/Ile<sup>5</sup>  $\text{C}^\gamma$  and their respective  $N$ -methyl group resulting in higher values for  $\Phi_{(i+2)}$  and  $\Psi_{(i+2)}$  (Table 1). The conformations of **4** and **5** are quite similar with a backbone RMSD of 0.45 Å and a stark right handed twist (Fig. 2).

Introduction of a  $\gamma$ -branched residue at the  $i + 2$  position (**6**), results in less deviation from the optimal  $\Phi$  and  $\Psi$  values (Table 1) in the turn region in comparison to the  $\beta$ -branched analogs. This is also evident from the CD spectrum of **6** showing a higher intensity than **4** and **5**. The ring constrained leucine isostere cyclohexylalanine substituted analog **7** and its aromatic congener **8**, with an  $N$ -methylated phenylalanine (although **8** shows an anomalous behavior in CD), display a similar solution conformation to **6**, suggesting the compatibility of the appended side chains at the  $i + 2$  site.

Finally, we introduced an  $N$ -methylated glycine (sarcosine) (as it is achiral and lacks side chain) at the  $i + 2$  position to assess its suitability in nucleating the  $\beta$ -hairpin. In spite of the

absence of any side chain constraint, **9** exhibited conformational homogeneity in  $\text{CDCl}_3$  with all of the characteristic NOEs. To our surprise, the dihedral angles about the turn region of **9** (Table 1) in its solution conformation are quite similar to a designed  $\beta$ -hairpin peptide with a central pro-Gly turn motif in its crystal form.<sup>19</sup> However, the solvent-shielded nature of the amide protons in **9** followed the trend that was observed for **5**, suggesting a flexible nature for both of the compounds. Compound **2** with the heterochiral  $N$ -methylated alanine in the turn region can be classified into both the  $i + 1$  and  $i + 2$  libraries. This gave us the clue that an  $N$ -methylated  $D$ -amino acid could potentially replace proline to form a stable  $\beta$ -hairpin structure. However, to assess the suitability of various branched and un-branched amino acids at the  $i + 1$  position we chose to substitute this site with all of the aforementioned amino acids but with reversed chirality.

**10** shows the entire signature NOEs of the reverse turn ( $\text{nle}'\text{H}^\alpha\text{--Ala}'5\text{NMe}$ ,  $\text{nle}'\text{H}^\alpha\text{--Leu}6\text{H}^\text{N}$  and  $\text{Ala}'5\text{NMe--Leu}6\text{H}^\text{N}$ ) and displays the presence of a well-folded conformation ( $\text{Leu}1\text{H}^\text{N}\text{--Val}8\text{H}^\text{N}$ ,  $\text{Phe}2\text{H}^\alpha\text{--Phe}7\text{H}^\alpha$ ,  $\text{Val}3\text{H}^\text{N}\text{--Leu}6\text{H}^\text{N}$  and  $\text{NHAc--OMe}$ ). The solvent-shielded  $\text{Leu}1\text{H}^\text{N}$ ,  $\text{Val}3\text{H}^\text{N}$ ,  $\text{Leu}6\text{H}^\text{N}$  and  $\text{Val}8\text{H}^\text{N}$  along with the  $^3J_{\text{H}^\text{N}\text{--H}^\alpha} > 8$  Hz implies a putative  $\beta$ -hairpin structure. The structure of **10** (Fig. 3) is comparable to **2** and the norleucine swapped analog **3** with a backbone RMSD of 0.56 Å and 0.50 Å respectively.<sup>18</sup>

The  $\beta$ -branched  $i + 1$  analog **11** showed a CD spectrum comparable to **10** (Fig. 1c) suggesting a structural similarity between these two analogs unlike in the corresponding  $i + 2$  analogs. The characteristic NOEs and the high  $^3J_{\text{H}^\text{N}\text{--H}^\alpha}$  revealed the structural integrity of the molecule. The  $\Phi_{(i+1)}$  in **11** (Table 2) showed a significant increase that could be attributed to the steric interaction between the  $\text{val}'\text{C}^\gamma$  and  $\text{val}'\text{NMe}$ , which subsequently results in the reduced right handed twist of the hairpin unlike in **4**. A fairly good agreement between the solution structures of **11** and **10** (backbone RMSD of 0.39 Å) indicates the differential behavior of valine at the  $i + 1$  and  $i + 2$  sites.





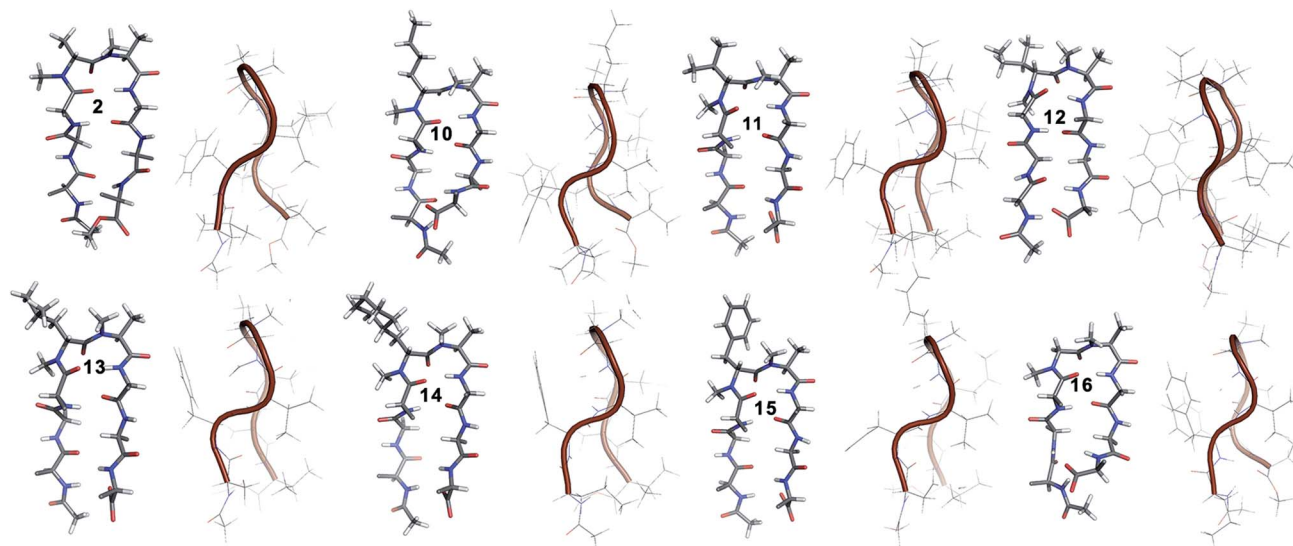


Fig. 3 Solution structure of all of the  $i + 1$  variants in  $\text{CDCl}_3$ . The front view of the peptide's backbone along with the side chains of the reverse turn motif are shown in a stick representation; all other side chains are restricted to  $\text{C}^\beta$  for the sake of clarity. The edge view is shown in a cartoon representation to depict the right-handed twist.

Table 2 Dihedral angles about the reverse turn in  $i + 1$  variants obtained from the conformations generated by the restrained molecular dynamics simulation

	$i + 1$		$i + 2$	
	$\Phi$ ( $^\circ$ )	$\Psi$ ( $^\circ$ )	$\Phi$ ( $^\circ$ )	$\Psi$ ( $^\circ$ )
10	$73 \pm 7$	$-121 \pm 8$	$-105 \pm 10$	$17 \pm 10$
11	$86 \pm 10$	$-118 \pm 7$	$-106 \pm 8$	$26 \pm 9$
12	$93 \pm 12$	$-112 \pm 13$	$-107 \pm 11$	$40 \pm 17$
13	$67 \pm 7$	$-131 \pm 6$	$-90 \pm 9$	$14 \pm 7$
14	$69 \pm 7$	$-135 \pm 6$	$-88 \pm 7$	$14 \pm 7$
15	$65 \pm 8$	$-126 \pm 6$	$-93 \pm 9$	$15 \pm 8$
16	$75 \pm 10$	$-104 \pm 7$	$-128 \pm 8$	$6 \pm 5$

Unlike 5, compound 12 with the Ile' shows a CD spectrum with a high intensity. The explanation was beautifully provided by the solution structure of 12, where the  $\Phi_{(i+1)}$  and  $\Psi_{(i+2)}$  deviate the most from the ideal geometry amongst all of the  $i + 1$  analogs. The increased  $\Phi_{(i+1)}$  is a result of the strong steric interaction between the Ile' $\text{C}^{\gamma 1}$  and Ile' $\text{NMe}$  (Fig. 3) that eventually leads to a heavily restricted conformation about Ile' as suggested by the multiple intra- and interresidue NOEs. This conformational restriction creates further steric clash between Ile' $\text{C}^{\gamma 2}$  and Ala'5NMe resulting in a higher  $\Psi_{(i+2)}$  value. This restriction is analogous to the ring constraint of Pro in 1. A combination of these effects result in a flattened conformation in 12 that is comparable to 1 (backbone RMSD of 0.64 Å).

The CD spectrum of the  $\gamma$ -branched analog 13 is less intense than the CD spectrum of the corresponding  $i + 2$  analog 6. This observation is strikingly opposite to the trend observed for the  $\beta$ -branched analogs 5 and 12. The solution structures of these compounds revealed the presence of a considerable right-handed twist in 13 as compared to 6, validating the CD

spectrum. The basis of this twist in 13 is the steric interaction between the isopropyl and the  $\text{N}$ -methyl group of leu' resulting in a slightly higher value of  $\Phi_{(i+1)}$  in 13 than in 6. On the contrary, the steric repulsion between the isopropyl and the  $\text{N}$ -methyl group of leu' in 6 is considerably less due to the greater torsion angle of  $\text{C}^{\text{(NMe)}}\text{-N-C}^\alpha\text{-C}^\beta_{(i+2)}$  than  $\text{C}^{\text{(NMe)}}\text{-N-C}^\alpha\text{-C}^\beta_{(i+1)}$  (due to the opposite stereochemistry at the  $i + 1$  and  $i + 2$  site) resulting in the flattened hairpin in 6.

Surprisingly, the CD spectrum of 14 showed a lower intensity in comparison to all other  $i + 1$  analogs. Nevertheless, we could identify the characteristic NOEs, and the solution conformation of 14 is almost identical to that of 13 (backbone RMSD of 0.24 Å) displaying the right handed twist. On the other hand the aromatic  $i + 1$   $\gamma$ -branched analog 15 shows the most intense CD spectrum amongst all of the  $i + 1$  congeners indicating the occurrence of a flatter hairpin. However, the conformation of 15 and the solvent exposure of amide protons show a striking similarity with the corresponding  $i + 2$  analog 8 (backbone RMSD of 0.26 Å), which shows a less intense CD spectrum. This re-emphasizes the anomalous behavior of the electronic CD for peptides with aromatic residues.<sup>20</sup> Furthermore, we did not observe any aggregation at the concentrations used for NMR and CD measurements,<sup>18</sup> therefore the extent of twist in these peptides arises mainly due to the substitution pattern at the reverse turn.

To determine the importance of the  $i + 1$  side chain on the induction of the  $\beta\text{II}'$  turn and the subsequent folding of the hairpin, we synthesized compound 16 with sarcosine at the  $i + 1$  position. Although the CD spectrum of 16 is comparable to other  $i + 1$  analogs, it has considerably less intensity than the CD spectrum of 9. This observation directly correlates with the flexibility of 16 in  $\text{CDCl}_3$ ,<sup>18</sup> although the major conformer displays all of the characteristic signatures (NOEs,  $^3J_{\text{H}^{\text{N}}\text{-H}^\alpha}$  and solvent shielded  $\text{H}^{\text{N}}$ ) of a  $\beta$ -hairpin. It is interesting to note that



in spite of the absence of any side chains at the  $i + 1$  site, there is a clear indication of a  $\beta$ II' turn as suggested by the following NOEs: Gly $^i$ H $^{\alpha 1}$ –Ala $^{i+1}$ 5NMe, Gly $^i$ H $^{\alpha 1}$ –Leu $^{i+2}$ 6H $^N$ , Ala $^{i+1}$ 5NMe–Leu $^{i+2}$ 6H $^N$  and a shorter distance for Ala $^{i+1}$ 5NMe–Ala $^{i+2}$ 5H $^{\beta}$  than Ala $^{i+1}$ 5NMe–Ala $^{i+2}$ 5H $^{\alpha}$ . However, there is a substantial increase in the  $\Phi_{(i+1)}$  and  $\Phi_{(i+2)}$  values. A close assessment of the structure revealed the separation of  $i + 1/i + 2$  and  $i + 2/i + 3$  amide planes to a greater extent in **16** than all of the other analogs including the most sterically demanding  $i + 2$   $\beta$ -branched analogs, **4** and **5**. This is mainly due to the absence of  $i + 1$  C $^{\beta}$  in **16**, which restricts the inward rotation of the  $i + 1/i + 2$  amide plane in the other analogs due to the van der Waals repulsion with the  $i + 1$  CO group (Fig. 4a) (note the greater torsion angle C $^{\beta}$ –C $^{\alpha}$ –C–O $_{(i+1)}$  in **2** and **4** in comparison to H $^{\alpha(Pro-R)}$ –C $^{\alpha}$ –C–O $_{(i+1)}$  in **16**).

The tolerance of various *N*-methylated chiral or achiral amino acids at the turn region clearly indicates that the steric interactions in the reverse turn are crucial in nucleating the  $\beta$ -hairpin that is subsequently stabilized by the intramolecular

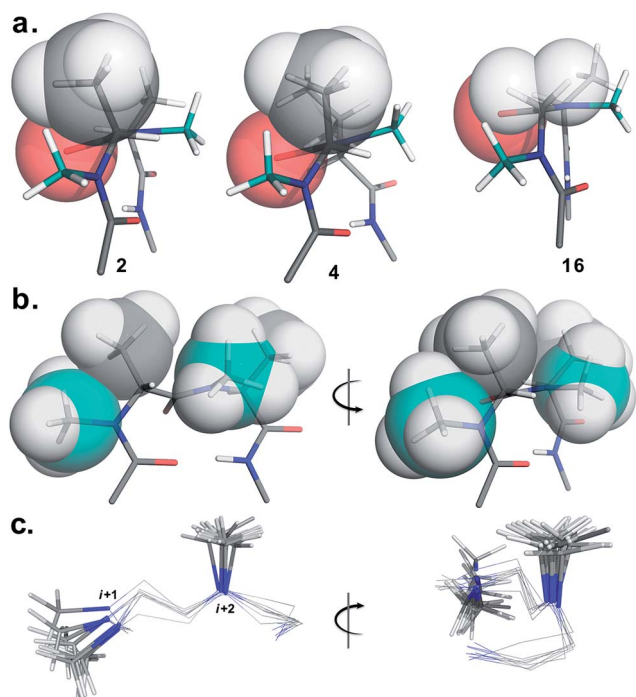
hydrogen bonding.<sup>21</sup> The relative orientations of the  $i + 1$  *N*-methyl,  $i + 1$  C $^{\beta}$ ,  $i + 2$  *N*-methyl and  $i + 2$  C $^{\beta}$  along with the  $i + 1$  CO are critical in dictating the stability of the  $\beta$ -turn (Fig. 4b). Any relaxation in the van der Waals repulsion about these residues destabilizes the global conformation of the hairpin. This is clearly evident from **16**, where the absence of  $i + 1$  C $^{\beta}$  destabilizes the global conformation. The torsion angles C $^{(NMe)}_{(i+1)}$ –N–C $^{\alpha}$ –C $^{\beta}_{(i+1)}$  and C $^{(NMe)}_{(i+2)}$ –N–C $^{\alpha}$ –C $^{\beta}_{(i+2)}$  also play a critical role in stabilizing the reverse turn as observed in the  $\gamma$ -branched analogs **6** and **13**. Finally, the conformational flexibility of the *N*-methyl group at the  $i + 1$  position (Fig. 4c) allows for a better tolerance of bulkier substituents (e.g.  $\beta$ -branched amino acids) at the  $i + 1$  site than at  $i + 2$ . This leads to the differential behavior of certain substituents at these two sites (e.g. **4** and **11**).

Next, to assess the suitability of the designed reverse turn motif in the context of protein engineering, we studied several hydrophilic peptides with a common strand sequence<sup>22</sup> but with varied turn inducing motifs. We were keen to understand the behavior of our turn inducing motif in aqueous conditions, as only a few reverse turn motifs have found utility in protein engineering.<sup>23</sup> We initially synthesized **17**, with D-Ala–L-Ala as the turn inducing motif, since in small cyclic peptides heterochirality is enough to induce the formation of a  $\beta$ II' turn.<sup>14</sup>

The CD spectrum of **17** (Fig. 5a) showed the signature of a random coil, suggesting that heterochirality in a linear peptide is not sufficient to induce a hairpin formation. To emphasize the role of *N*-methylation in inducing the  $\beta$ -hairpin formation, we synthesized **18**, with *N*-methylation at D-Ala and L-Ala. The CD spectrum of **18** showed a broad minima centred around 215 nm that is characteristic of a  $\beta$ -sheet structure (Fig. 5a). Further, to show the compatibility of different amino acid side chains at the reverse turn of the  $\beta$ -hairpin, we synthesized **19**, **20** and **21**. It was gratifying to note that the CD spectrum for these analogs followed a similar pattern as observed in the hydrophobic peptides, with the  $i + 2$   $\gamma$ -branched amino acid (**20**) showing the maximum and the  $\beta$ -branched analog (**19**) showing the minimum intensity. It was also encouraging to observe the differential behavior of the  $i + 2$  and  $i + 1$   $\beta$ -branched analogs **19** and **21**, respectively, that followed the trend as noted for the hydrophobic peptides.

To obtain additional insights into the folding behavior of these compounds, we performed NMR spectroscopy in acetate buffer (pH 3.8). The  $^1$ H spectrum of **17** showed clear indications of an unfolded structure, with the absence of any upfield shifted methyl groups<sup>24</sup> and overlapping resonance for the  $\beta$ -methyl groups ( $\sim 1.4$  ppm) of alanine (Fig. 5b). On the contrary, all of the *N*-methylated analogs show the presence of two upfield shifted methyl groups that progressively increase in the order **19** < **21** < **18** < **20**. This order also correlated with the enhanced dispersion of chemical shifts in the amide region of these compounds.

The secondary chemical shifts of these compounds (obtained from the respective unfolded controls, where the *N*-methylated D-residue was substituted with an *N*-methylated L-residue)<sup>18</sup> also follows the order **19** < **21**  $\approx$  **18** < **20**, suggesting the increased foldedness of the  $\beta$ -sheet from left to right.<sup>25a</sup> The



**Fig. 4** Critical factors determining the reverse turn stability. (a) The relative orientation of the  $i + 1$  C $^{\beta}$  and  $i + 1$  CO is depicted in three different substitution patterns. In spite of the strong steric repulsion at the  $i + 2$  site in **4** the relative orientation of the  $i + 1$  C $^{\beta}$  and  $i + 1$  CO does not alter. However, it effectively alters the orientation of the  $i + 2$  C $^{\beta}$  resulting in a twist in the structure (note the parallel orientation of the strands in **2** and **16**). However, the absence of  $i + 1$  C $^{\beta}$  in **16** relaxes the steric repulsion between  $i + 1$  H $^{\alpha}$  and  $i + 1$  CO resulting in enhanced conformational flexibility. (b) The relative orientation of the  $i + 1$  *N*-methyl,  $i + 1$  C $^{\beta}$ ,  $i + 2$  *N*-methyl and  $i + 2$  C $^{\beta}$  groups that are responsible for the induction and the stability of the  $\beta$ -hairpin conformation (e.g. front and side view of the reverse turn in **2**). (c) Overlay of the  $\beta$ II' turn of **1–9** depicting the spatial distribution of the  $i + 1$  and  $i + 2$  *N*-methyl groups (front and side view). The conformational space spanned by the  $i + 1$  *N*-methyl group is relatively larger than the  $i + 2$  *N*-methyl group. Only the *N*-methyl groups in the turn motif are represented as sticks for the sake of clarity.



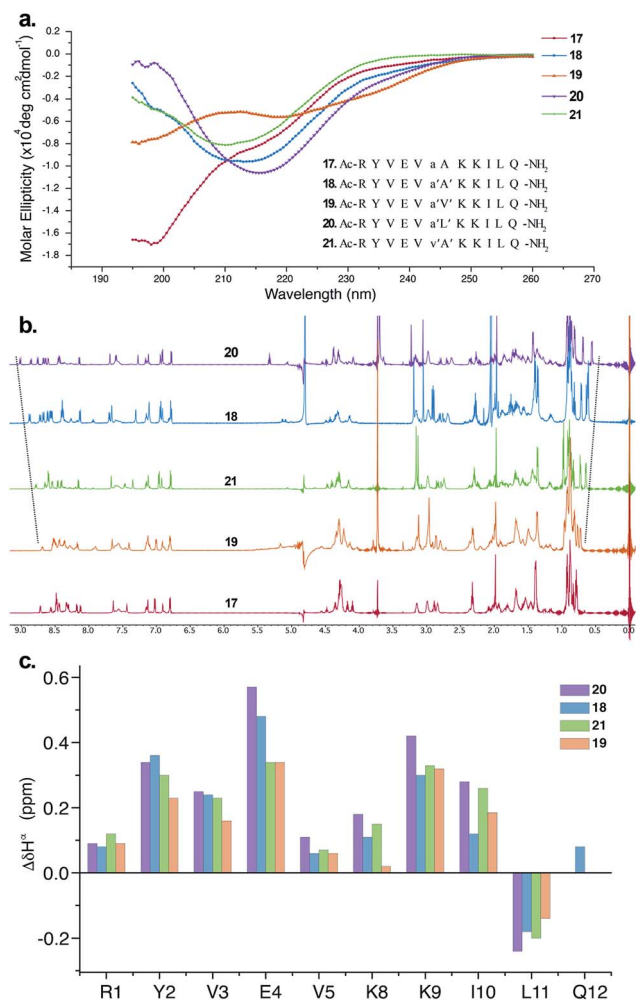


Fig. 5 (a) The CD spectra of the hydrophilic peptides in acetate buffer (pH 3.8) and the synthesized sequences 17–21 (inset), which are color-coded. (b) <sup>1</sup>H NMR spectra of 17–21 showing the gradual upfield shift of Leu11CH<sub>3</sub><sup>γ</sup> as well as the gradual downfield shift of the Ile10H<sup>N</sup>. (c) Comparison of  $\Delta\delta\text{H}^\alpha$  chemical shifts of the strand residues in 18–21 (obtained from the respective unfolded control peptides), showing the modulation of the folded conformation by a single amino acid substitution (assignment of Gln12H<sup>α</sup> in 19, 20 and 21 was not possible due to resonance overlap).

results from the NMR analyses corroborated the different intensities observed in the CD spectrum. Together these results strongly indicate the subtle modulation of the  $\beta$ -sheet folded conformation *via* a single substitution at the  $i + 1$  or  $i + 2$  site, which would find enormous utility in protein engineering.

Finally, to validate the broad scope of our design strategy in foldamer design, we calculated the solution structure of 18 at 25 °C. The conformation of 18 (Fig. 6a) was well-defined by several inter- and intraresidue NOEs. The reverse turn was specified by ala'H<sup>β</sup>-ala'NMe, ala'H<sup>α</sup>-Ala'7NMe, ala'H<sup>α</sup>-Lys8H<sup>N</sup>, and Ala'7NMe-Lys8H<sup>N</sup> NOEs, while Tyr2H<sup>α</sup>-Leu11H<sup>α</sup>, Tyr2H<sup>α</sup>-Gln12H<sup>N</sup>, and Val5H<sup>N</sup>-Lys8H<sup>N</sup> NOEs confirmed the strand registry.<sup>18</sup> The structure of 18 was calculated utilizing the distances derived from the NOEs, which showed the presence of a  $\beta$ -hairpin conformation with a central  $\beta\text{II}'$  turn. It was

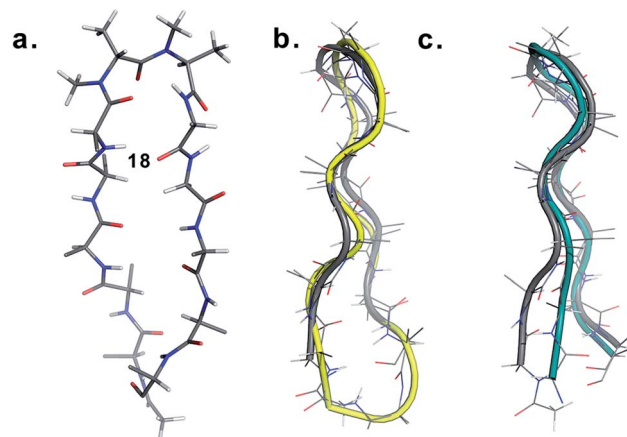


Fig. 6 (a) Structure of 18 in aqueous solution and backbone overlay (in grey) with (b) a cyclic  $\beta$ -sheet peptide TS+ (yellow) (RMSD 1.0 Å) and (c) an engineered GB1 peptide (teal) (RMSD 0.75 Å).

gratifying to see that 18 displayed significant structural similarity with two different  $\beta$ -sheet peptides<sup>25</sup> in aqueous conditions (Fig. 6b and c).

## Conclusions

In conclusion, we show that the *N*-methyl groups in conjunction with the C<sup>β</sup> at the  $i + 1$  and  $i + 2$  amino acid side chains provide sufficient steric locking to fold a linear peptide into a  $\beta$ -hairpin. The introduction of heterochirality in linear peptides alone is not enough to induce the formation of  $\beta$ -sheets, unlike in cyclic peptides. Our engineering strategy is quite modular in terms of decorating the reverse turn with different functional groups to alter the physicochemical properties of the designed  $\beta$ -sheets and attach various probes for biophysical studies. Furthermore, branched amino acids at the reverse turn add another dimension to its modularity by introducing a varying extent of twist that could be utilized to probe the structure activity relationship of designed  $\beta$ -sheets and the modulate folding of proteins. *N*-methylation has found tremendous utility in cyclic peptides, however, its use in linear peptides was limited. With this report we strongly believe that this simple strategy will not only find enormous utility in foldamer design<sup>26</sup> and protein engineering but also in the development of novel materials<sup>27</sup> and peptide based catalysts.<sup>28</sup>

## Acknowledgements

The authors gratefully acknowledge the Science and Engineering Research Board (SERB) through the project SB/YS/LS-220/2013 to JC and the Department of Biotechnology, Ministry of Science and Technology through the DBT-IISc partnership program for providing funding for this work. We also thank the DBT & DST funded NMR facility and the DBT funded Proteomics facility at IISc. We also wish to thank Prof. Siddhartha P. Sarma for assistance with the NMR experiments and Prof. Debnath Pal for the computing facility.





## References

- 1 (a) M. Pelay-Gimeno, A. Glas, O. Koch and T. N. Grossmann, *Angew. Chem., Int. Ed.*, 2015, **54**, 8896; (b) A. M. Watkins and P. S. Arora, *ACS Chem. Biol.*, 2014, **9**, 1747.
- 2 (a) M. D. Struthers, R. P. Cheng and B. Imperiali, *Science*, 1996, **271**, 342; (b) Z. E. Reinert, G. A. Lengyel and S. W. Horne, *J. Am. Chem. Soc.*, 2013, **135**, 12528.
- 3 (a) C. M. Nair, M. Vijayan, Y. V. Venkatachalapathi and P. Balaram, *J. Chem. Soc., Chem. Commun.*, 1979, 1183; (b) T. S. Haque, J. C. Little and S. H. Gellman, *J. Am. Chem. Soc.*, 1996, **118**, 6975; (c) J. A. Robinson, *Acc. Chem. Res.*, 2008, **41**, 1278.
- 4 (a) M. Ramírez-Alvarado, F. J. Blanco and L. Serrano, *Nat. Struct. Biol.*, 1996, **3**, 604; (b) C. D. Tatko and M. L. Waters, *J. Am. Chem. Soc.*, 2002, **124**, 9372; (c) S. R. Griffiths-Jones, A. J. Maynard and M. S. Searle, *J. Mol. Biol.*, 1999, **292**, 1051; (d) A. G. Cochran, N. J. Skelton and M. A. Starovasnik, *Proc. Natl. Acad. Sci. U. S. A.*, 2001, **98**, 5578.
- 5 W. J. Cooper and M. L. Waters, *Org. Lett.*, 2005, **7**, 3825.
- 6 (a) J. Xiao, B. Weisblum and P. Wipf, *Org. Lett.*, 2006, **8**, 4731; (b) G.-B. Liang and S. H. Gellman, *J. Am. Chem. Soc.*, 1999, **121**, 1806; (c) Y. J. Chung, B. R. Huck, L. A. Christianson, H. E. Stanger, S. Krauthäuser, D. R. Powell and S. H. Gellman, *J. Am. Chem. Soc.*, 2000, **122**, 3995; (d) G. Jeannotte and W. D. Lubell, *J. Am. Chem. Soc.*, 2004, **126**, 14334; (e) R. V. Nair, S. B. Baravkar, T. S. Ingole and G. J. Sanjayan, *Chem. Commun.*, 2014, **50**, 13874; (f) S. Chowdhury, G. Schatte and H. B. Kraatz, *Angew. Chem., Int. Ed.*, 2008, **47**, 7056; (g) J. S. Nowick and J. O. Brower, *J. Am. Chem. Soc.*, 2003, **125**, 876; (h) J. S. Nowick, K. S. Lam, T. V. Khasanova, W. E. Kemnitzer, S. Maitra, H. T. Mee and R. Liu, *J. Am. Chem. Soc.*, 2002, **124**, 4972.
- 7 G. A. Lengyel, Z. E. Reinert, B. D. Griffith and S. W. Horne, *Org. Biomol. Chem.*, 2014, **12**, 5375.
- 8 (a) D. K. Chalmers and G. R. Marshall, *J. Am. Chem. Soc.*, 1995, **117**, 5927; (b) Y. Takeuchi and G. R. Marshall, *J. Am. Chem. Soc.*, 1998, **120**, 5363; (c) A. C. Gibbs, T. C. Bjorndahl, R. S. Hodges and D. S. Wishart, *J. Am. Chem. Soc.*, 2002, **124**, 1203.
- 9 (a) J. Chatterjee, D. Mierke and H. Kessler, *J. Am. Chem. Soc.*, 2006, **128**, 15164; (b) J. G. Beck, J. Chatterjee, B. Laufer, M. U. Kiran, A. O. Frank, S. Neubauer, O. Ovadia, S. Greenberg, C. Gilon, A. Hoffman and H. Kessler, *J. Am. Chem. Soc.*, 2012, **134**, 12125.
- 10 N. El Tayar, A. E. Mark, P. Vallat, R. M. Brunne, B. Testa and W. F. van Gunsteren, *J. Med. Chem.*, 1993, **36**, 3757.
- 11 R. Rai, S. Raghothama and P. Balaram, *J. Am. Chem. Soc.*, 2006, **128**, 2675.
- 12 A. E. Wakefield, W. M. Wuest and V. A. Voelz, *J. Chem. Inf. Model.*, 2015, **55**, 806.
- 13 P. Y. Chou and G. D. Fasman, *J. Mol. Biol.*, 1977, **115**, 135.
- 14 J. Chatterjee, B. Laufer and H. Kessler, *Nat. Protoc.*, 2012, **7**, 432.
- 15 R. Mahalakshmi, S. Raghothama and P. Balaram, *J. Am. Chem. Soc.*, 2006, **128**, 1125.
- 16 H. Kessler, *Angew. Chem., Int. Ed. Engl.*, 1970, **9**, 219.
- 17 H. Kessler, U. Anders and M. Schudok, *J. Am. Chem. Soc.*, 1990, **112**, 5908.
- 18 See ESI for details†
- 19 I. L. Karle, S. K. Awasthi and P. Balaram, *Proc. Natl. Acad. Sci. U. S. A.*, 1996, **93**, 8189.
- 20 C. Zhao, P. L. Polavarapu, C. Das and P. Balaram, *J. Am. Chem. Soc.*, 2000, **122**, 8228.
- 21 R. M. Culik, H. Jo, W. F. DeGrado and F. Gai, *J. Am. Chem. Soc.*, 2012, **134**, 8026.
- 22 F. A. Syud, J. F. Espinosa and S. H. Gellman, *J. Am. Chem. Soc.*, 1999, **121**, 11577.
- 23 M. Körling and A. Geyer, *Eur. J. Org. Chem.*, 2015, 2382.
- 24 B. Eckhardt, W. Grosse, L. O. Essen and A. Geyer, *Proc. Natl. Acad. Sci. U. S. A.*, 2010, **107**, 18336.
- 25 (a) S. J. Maynard, A. M. Almeida, Y. Yoshimi and S. H. Gellman, *J. Am. Chem. Soc.*, 2014, **136**, 16683; (b) G. A. Lengyel, R. C. Frank and W. S. Horne, *J. Am. Chem. Soc.*, 2011, **133**, 4246.
- 26 (a) C. M. Goodman, S. Choi, S. Shandler and W. F. DeGrado, *Nat. Chem. Biol.*, 2007, **3**, 252; (b) S. W. Horne and S. H. Gellman, *Acc. Chem. Res.*, 2008, **41**, 1399.
- 27 C. M. Micklitsch, P. J. Knerr, M. C. Branco, R. Nagarkar, D. J. Pochan and J. P. Schneider, *Angew. Chem., Int. Ed.*, 2011, **50**, 1577.
- 28 (a) J. T. Blank, D. J. Guerin and S. J. Miller, *Org. Lett.*, 2000, **2**, 1247; (b) M. E. Diener, A. J. Metrano, S. Kusano and S. J. Miller, *J. Am. Chem. Soc.*, 2015, **137**, 12369.

