Published in final edited form as:
J Am Chem Soc. 2008 June 25; 130(25): 7839–7841. doi:10.1021/ja802042c.

Diacid Linkers that Promote Parallel Beta-Sheet Secondary Structure in Water

Felix Freire, **John D. Fisk**, **Aaron J. Peoples**, **Monika Ivancic**, **Ilia A. Guzei**, and **Samuel H. Gellman***
Department of Chemistry, University of Wisconsin, Madison, Wisconsin 53706

Short peptides that adopt specific secondary structures in aqueous solution have proven to be invaluable tools for elucidating the balance of forces that controls the stability of these common protein substructures. Rules for the design of sequences that fold autonomously into the α-helical conformation have been available for nearly two decades,[1] and the analysis of α-helical folding preferences is now a mature field. Guidelines for the design of antiparallel β-sheet-forming sequences have emerged more recently, enabling fundamental studies of the origins of antiparallel β-sheet stability.[2] Parallel β-sheet is common in proteins, but this structural motif has received little attention to date because the rules necessary to design molecules that will adopt this secondary structure in aqueous solution are not yet fully developed.

Parallel β-sheet differs fundamentally from antiparallel β-sheet or α-helix in that the latter two secondary structures can form within relatively short peptides (12-25 residues), but the former cannot. A minimal β-sheet contains two strands; if these strands are to be oriented in parallel fashion, then they must be linked either, via their N-termini or via their C-termini. Thus, creation of a minimal parallel β-sheet requires non-peptide units in the linking segment. A number of linking segments containing diacids (or equivalents) to connect two N-termini or diamines (or equivalents) to connect to C-termini have been explored in organic solvents,[3-5] but only one linking segment has been shown conclusively to support parallel sheet folding between attached peptide strands in aqueous solution (this linker is illustrated in Figure 1a), as evidenced by the detection of characteristic interstrand NOEs.[6] Here we describe simple diacid linkers that promote parallel β-sheet folding between peptide strands attached via their N-termini. These new linkers represent valuable tools for analysis of sequence-stability relationships in parallel β-sheet secondary structure.

Simple molecular mechanics calculations suggested that a diacid formed by allowing glycine to react with *cis*-1,2-cyclohexanedicarboxylic anhydride would be an effective parallel β-sheet promoter (Figure 1b). Our first test of this hypothesis involved preparation of tetrapeptide analogues containing each enantiomer of this linker with L-Leu-N-methylamide attached to the Gly carboxyl and L-Val-N-methylamide attached to the remaining carboxyl on the cyclohexane ring (**1** and **2**). These syntheses relied on the asymmetric reaction of the anhydride with benzyl alcohol, which provides the half-ester with high and complementary enantioselectivity when catalyzed by quinine or quinidine.[7] Diastereomers **1** and **2** both display parallel β-sheet hydrogen bonding patterns in the solid state (Figure 2a). Thus, in both cases, Leu NH forms a 10-membered ring H-bond to one of the carbonyls on the cyclohexane ring (very similar to the H-bonds commonly observed in β-turns), and Leu carbonyl forms a 16-membered ring H-bond to the $NH(CH_3)$ attached to Val. We were intrigued to observe that both configurations of the cyclohexanedicarboxylic acid (CHDA) unit support parallel β-sheet

E-mail: gellman@chem.wisc.edu.

Supporting Information Available: Experimental details, compound characterizations, NMR data and crystallographic data

formation between appended L-amino acid residues. In our previous studies with proline-based linkers for antiparallel β-sheet nucleation and for C-to-C linked parallel β-sheet nucleation, we found that only one proline configuration supported the desired folding pattern.[2a,4c,6] NMR analysis of **1** and **2** in chloroform indicated that in both cases the conformation observed in the solid state is populated in solution: numerous NOEs between a proton on the L-Val residue and a proton on the L-Leu residue were observed for each molecule.[8]

We next examined **3** and **4**, analogues of 1 and 2 in which the Gly residue is replaced by an α-amino-isobutyric acid (Aib) residue. The Aib residue is well-known to promote reverse turn formation in peptides, and Aib-containing turn segments have been used to promote antiparallel β-sheet formation.[9] Similarly, we have found that placing a *gem*-dimethyl pair within a diamine reverse turn segment can promote parallel β-sheet formation.[4c, 6] These precedents led us to examine **3** and **4**, analogues of **1** and **2** in which the Gly residue is replaced by an Aib residue. To our surprise, the Aib-containing molecules adopt not a hairpin but instead a helix-like conformation in the solid state (Figure 2b). In each case one carbonyl from the CHDA unit forms a 10-membered ring with Leu NH, and the other CHDA carbonyl forms a 10-membered ring H-bond with the $NH(CH_3)$ attached to the Leu carbonyl. In light of these results, we did not pursue the Aib-CHDA linker.

The behavior of **1** and **2** in the solid state and in nonpolar solution is promising, but the observation of hairpin-like folding for molecules of this size under these conditions does not guarantee that the linkers will promote parallel β-sheet formation between longer peptide segments in aqueous solution.[10] Folding in water is essential for model systems intended to provide insight on the conformational preferences of proteins. To address this critical issue, we prepared diastereomers **5** and **6**, in which strands containing six L-residues are attached to each carboxyl of enantiomeric Gly-CHDA units (Figure 3a). The design of the peptide strands was based on several considerations. (1) Peptides that form β-sheet secondary structure are often prone to aggregation;[2] therefore, we incorporated several basic residues so that our molecules would bear a net positive charge under the acidic conditions employed for NMR analysis. (2) Aromatic residues were included in order to maximize resonance dispersion in the ^1H NMR spectrum. (3) We placed residues with complementary charges (Lys and Glu) at the terminal strand positions so that electrostatic attraction between the side chains would encourage strand association at the open end of the parallel β-sheet.[11] (4) We selected residues with a high propensity to participate in β-sheet secondary structure,[12] and we designed the strand sequences so as to favor interstrand pairs that appear to be preferred among parallel β-sheets in proteins.[13]

Two-dimensional NMR analysis[14] of 2.5 mM samples of **5** and **6** in 9:1 H_2O:D_2O at pH 3.8 (100 mM sodium acetate buffer) indicated that parallel β-sheet interactions occur between the strand segments in both cases. DOSY measurements[15] carried out with 0.3 mM and 5 mM peptide samples indicated invariant diffusion coefficients, which suggests that little or no peptide aggregation occurs in this concentration range. NOEs between protons from residues that are not adjacent in sequence provided strong evidence for the expected parallel β-sheet interactions between the linked strand segments.[4,16] In parallel β-sheet secondary structure, a pair of aligned residues on adjacent strands has one partner that is H-bonded (HB) to the opposite strand and one partner that is not H-bonded (nHB). For example, in **5** and **6** the target conformations would have a Glu_{HB}-Lys_{nHB} pair at the open end (residues 1 and 14), followed by an Arg_{HB}-Thr_{nHB} pair (residue 2 and 13) (Figure 3b). This interstrand arrangement should give rise to NOEs between NH of the HB partner and $C_\alpha H$ of the nHB partner.[16] For both **5** and **6**, these interstrand NH--$C_\alpha H$ NOEs are observed for five of the six expected residue pairs; only the terminal Glu_{HB}-Lys_{nHB} pair does not give rise to this type of NOE, perhaps because of fraying at the open end of the parallel hairpin. In addition to these backbone NOEs, numerous NOEs are observed between side chains for each of the expected lateral residue pairs, including

the terminal Glu_{HB}-Lys_{nHB} pair. Further interstrand NOEs are observed between side chains juxtaposed in diagonal fashion,[17] Arg2-Tyr11, Lys4-Val9 and Thr5-Phe12, as expected for parallel sheet secondary structure with the commonly observed right-handed twist.[18] A total of 30 NOEs between protons from sequentially non-adjacent residues was observed for **5**, and 42 were observed for **6**. These NOEs were used as constraints for structure calculations with the program CNS.[19] As shown by the backbone overlays of the 10 lowest energy structures (Figure 3c), the set of NOEs is consistent with parallel β-sheet formation in both **5** and **6**. Superimposed NMR structures for **5** and **6** (Figure 3d) indicate that the biggest difference involves the residues that form the turn. Otherwise, the structures are very similar along the β-sheet.

Hutchinson et al. have recently reported an extensive analysis of lateral pairing preferences among parallel β-sheets found in high-resolution protein crystal structures.[13] Their findings suggest that for many residue pairs there is a significant difference between the favorability of the two possible HB-nHB arrangements. For example, the statistical analysis suggests that Arg_{HB}-Thr_{nHB} is strongly favored relative to Thr_{HB}-Arg_{nHB},[13] and Lys_{HB}-Tyr_{nHB} is more favored than Tyr_{HB}-Lys_{nHB}. In the case of Glu_{HB}-Lys_{nHB} and Lys_{HB}-Glu_{nHB} both orientations are strongly favored, with the first one preferred. Asymmetries of this type were rationalized by noting that only the preferred arrangement allows stabilizing interactions between the side chains in the ideal rotameric states. Autonomously folding model systems such as **5** and **6** allow us to ask whether the asymmetric HB-nHB pairing patterns observed in parallel β-sheets of proteins,[13] which are necessarily embedded in specific tertiary structural contexts, represent substantial sources of intrinsic secondary structural stability.

To probe the significance of lateral residue pairings in parallel β-sheet, we prepared **7** and **8**, the sequence isomers of **5** and **6** respectively, in which the attachment points of the strands to the Gly-CHDA linker have been swapped (Figure 4). Of the six lateral residue pairings in the hairpin conformation of **5** or **6**, four are altered in the sequence isomers (two do not change because the residues are identical). In each of these four asymmetric pairings, the orientation in **5** and **6** is predicted to be superior to that in **7** and **8**, based on the protein structure database analysis.[13] The behavior of **7** and **8** is consistent with this prediction: 2D NMR data reveal that neither molecule shows any NOE between protons on sequentially non-adjacent residues. Thus, neither **7** nor **8** appears to form parallel β-sheet secondary structure in aqueous solution. These results show that the Gly-CHDA linkers are not dominant drivers of parallel β-sheet formation; instead, these linkers enable parallel β-sheet interactions between attached strands, but interstrand attractions must contribute to overall conformational stability. Moreover, these results show that intrinsic β-sheet propensities of strand residues are not sufficient to drive folding; instead specific and favorable interactions between side chains on adjacent strands appear to be necessary for parallel β-sheet formation.

We have introduced a new linker, Gly-CHDA, for connecting peptide strands via their N-termini and shown that this unit (in either enantiomeric form) constitutes the first N-terminal linker to support parallel β-sheet folding in aqueous solution. Because β-sheet secondary structure forms without buttressing from a tertiary context, and because folding depends upon intrinsic attractions between the strands themselves, Gly-CHDA linkers will be valuable tools for fundamental study of the factors that influence the stability of parallel β-sheet, a very common structural motif within proteins. We have illustrated this utility by providing a first qualitative test of predictions based on the recent database analysis of lateral residue pairing preferences by Hutchinson et al.[13]

The new Gly-CHDA linkers complement previously described C-terminal linkers that promote parallel β-sheet folding in water.[6] The availability of both N- and C-terminal linkers will be useful for preparation of cyclic peptides, to provide spectroscopic benchmarks for the fully

folded state of parallel hairpin model systems, and these linkers will be essential for generating autonomously folding parallel β-sheets that contain three or more strands.

Acknowledgments

This research was supported by the NIH (GM61238). F. F. is a MEC-Fulbright Postdoctoral Fellow. NMR spectrometers were purchased in part by grants from NIH and NSF; diffraction equipment was supported by grants from NSF. We thank S. H. Choi for growing the crystals of tetrapeptide mimics **1** and **2**.

References

1. Chakrabartty A, Baldwin RL. Adv Protein Chem 1995;46:141. [PubMed: 7771317]
2. (a) Gellman SH. Curr Opin Chem Biol 1998;2:717. [PubMed: 9914187] (b) Searle MS, Ciani B. Curr Opin Struct Biol 2004;14:458. [PubMed: 15313241] (c) Hughes RM, Waters ML. Curr Opin Struc Biol 2006;16:514–524. (d) De Alba E, Rico M, Jiménez MA. Protein Sci 1999;8:2234. [PubMed: 10595526]
3. (a) Levin S, Nowick JS. J Am Chem Soc 2007;129:13043. [PubMed: 17918935]Nowick JS, Smith EM, Noronha G. J Org Chem 1995;60:7386. (b) Nowick JS, Insaf S. J Am Chem Soc 1997;119:10903. (c) Nowick JS. Acc Chem Res 1999;32:287.
4. (a) Wagner G, Feigel M. Tetrahedron 1993;49:10831. (b) Ranganathan D, Haridas V, Kurur S, Thomas A, Madhusudanan KP, Nagaraj R, Kunwar AC, Sarma AVS, Karle IL. J Am Chem Soc 1998;120:8448. (c) Fisk JD, Powell DR, Gellman SH. J Am Chem Soc 2000;122:5443.
5. For analysis of parallel β-sheet model systems in aqueous or mixed aqueous-organic solvents, see: (a) Kemp DS, Blanchard DE, Muendel CC. Peptides-Chemistry and Biology Smith J, Rivier J. ESCOMLeiden1992;319(b)Junquera E, Nowick JS. J Org Chem :1999.64 2527.(c)Chitnumsub P, Fiori WR, Lashuel HA, Diaz H, Kelly JW. Bioorg Med Chem 1999;7:39. [PubMed: 10199655]
6. (a) Fisk JD, Gellman SH. J Am Chem Soc 2001;123:343. [PubMed: 11456526] (b) Fisk JD, Schmitt MA, Gellman SH. J Am Chem Soc 2006;128:7148. [PubMed: 16734453]
7. (a) Bolm C, Atodiresei L, Schiffers I, Kanai M, Shibasaki M. Org Syn 2005;82:120. (b) Bolm C, Schiffers I, Dinter CL, Gerlach A. J Org Chem 2000;65:6984. [PubMed: 11031020]
8. Please see Supporting Information.
9. (a) Setnicka V, Huang R, Thomas CL, Etienne MA, Kubelka J, Hammer RP, Keiderling TA. J Am Chem Soc 2005;127:4992. [PubMed: 15810813] (b) Aravinda S, Shamala N, Rajkishore R, Gopi HN, Balaram P. Angew Chem Int Ed 2002;41:3863–3865.
10. Fisk, JD. PhD Thesis. University of Wisconsin-Madison; 2004.
11. Griffiths-Jones SR, Maynard AJ, Searle MS. J Mol Biol 1999;292:1051–1069. [PubMed: 10512702]
12. (a) Muñoz V, Serrano L. Proteins: Struct Funct Genet 1994;20:301. [PubMed: 7731949] (b) Minor DL, Kim PS. Nature 1994;371:264. [PubMed: 8078589]
13. Fooks HM, Martin ACR, Woolfson DN, Sessions RB, Hutchinson EG. J Mol Biol 2006;356:32. [PubMed: 16337654]
14. COSY: Aue WP, Bartholdi E, Ernst RR. J Chem Phys 1976;64:2229. TOCSY: Bax A, Davis DG. J Magn Reson 1985;65:355. ROESY: Bothner-by AA, Stephens RL, Lee JM, Warren CD, Jeanloz RW. J Am Chem Soc 1984;106:811.
15. (a) Cohen Y, Avram L, Frish L. Angew Chem Int Ed 2005;44:520. (b) Dehner A, Kessler H. Chem Bio Chem 2005;6:1550. (c) Altieri AS, Hinton DP, Byrd RA. J Am Chem Soc 1995;117:7561.
16. Wüthrich, K. NMR of Proteins and Nucleic Acids. Wiley; New York: 1986. (b) Dyson HJ, Wright PE. Annu Rev Biophys Biophys Chem 1991;20:519. [PubMed: 1867725]
17. Syud FA, Stanger HE, Gellman SH. J Am Chem Soc 2001;123:8667. [PubMed: 11535071]
18. Chothia C. J Mol Biol 1973;75:295. [PubMed: 4728692]
19. Brüngerm AT, Adams PD, Clore GM, DeLano WL, Gros P, Grosse-Kunstleve RW, Jiang JS, Kuszewski J, Nilges M, Pannu NS, Read RJ, Rice LM, Simonson T, Warren GL. Acta Cristallogr D Biol Crystallogr 1998;D54:905.

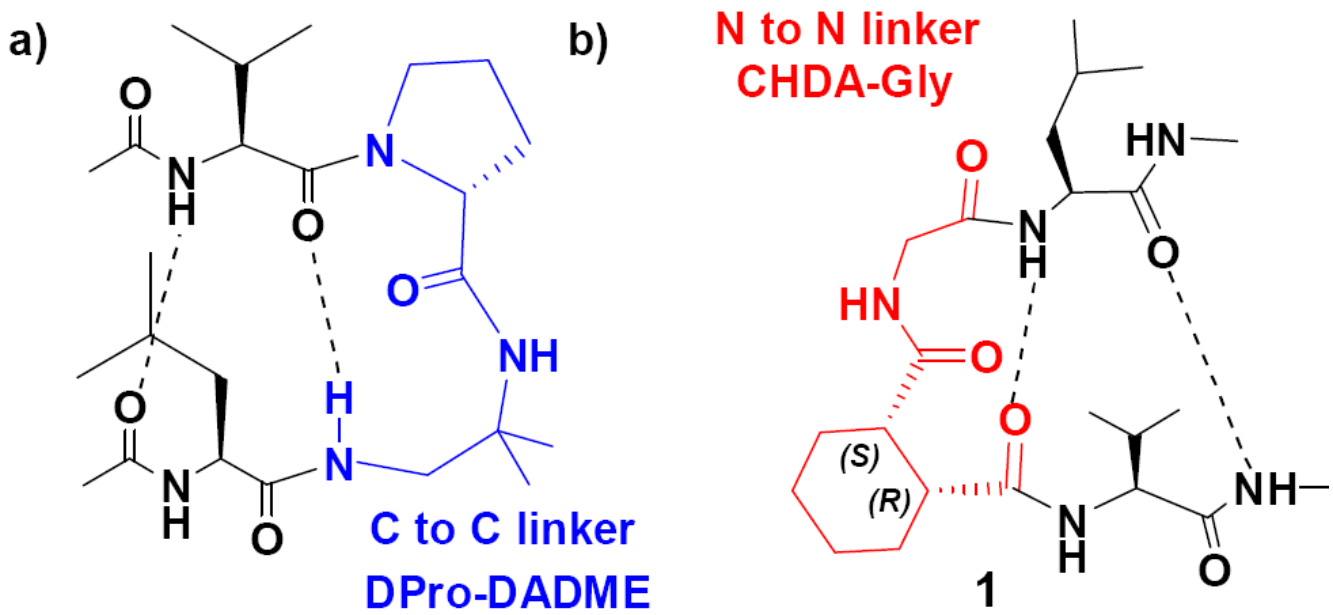

Figure 1.
Tetrapeptide mimics that contain parallel linkers: (a) C-to-C; (b) N-to-N

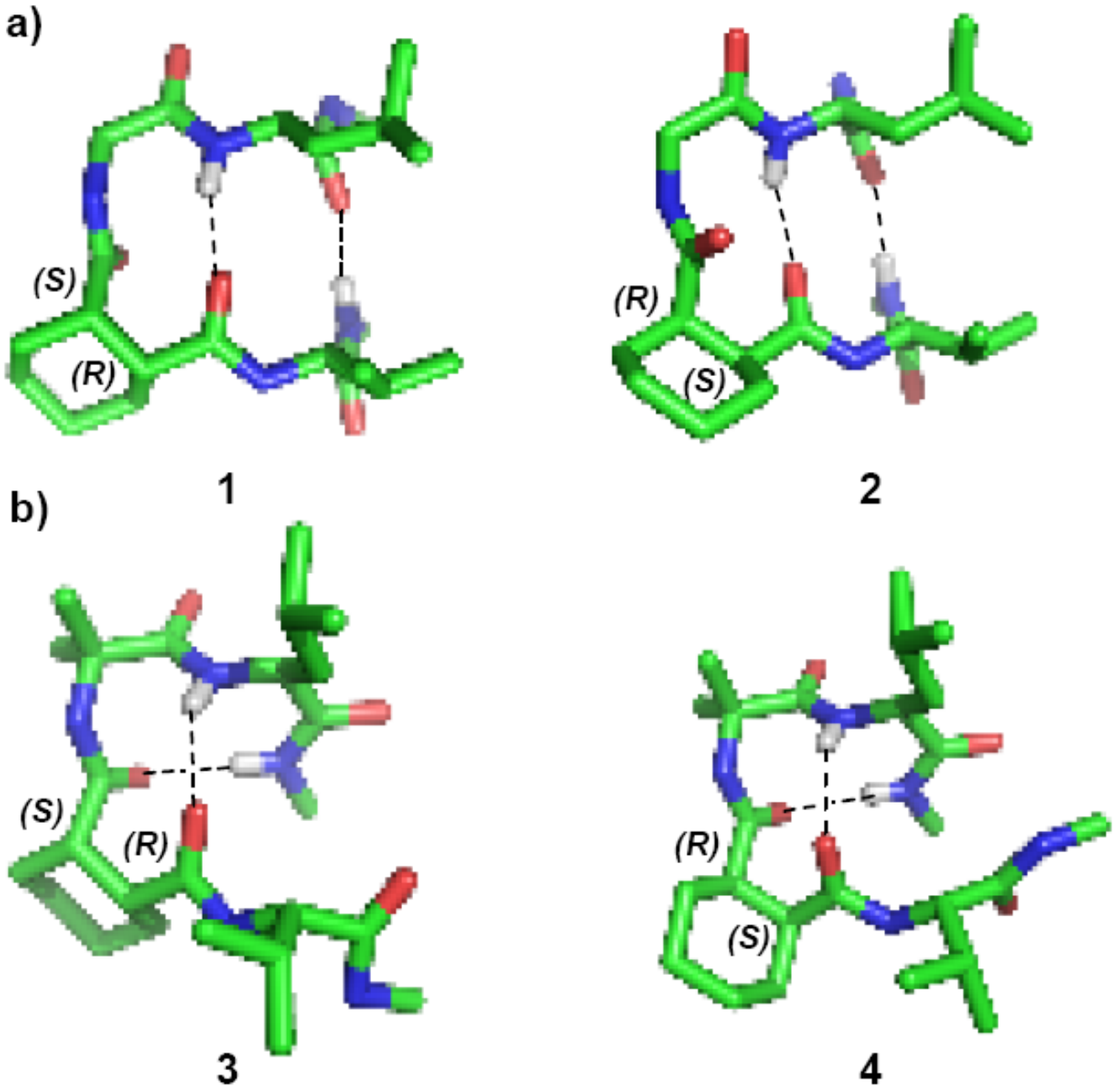

Figure 2.
(a) Solid state conformations of tetrapeptide mimics containing the linkers (S,R)-CHDA-Gly (1) and (R,S)-CHDA-Gly (2). (b) Solid state conformations of tetrapeptide mimics containing the linkers (S,R)-CHDA-Aib (3) and (R,S)-CHDA-Aib (4). In each case, all H atoms other than those on the Val and Leu N atoms are omitted for clarity.

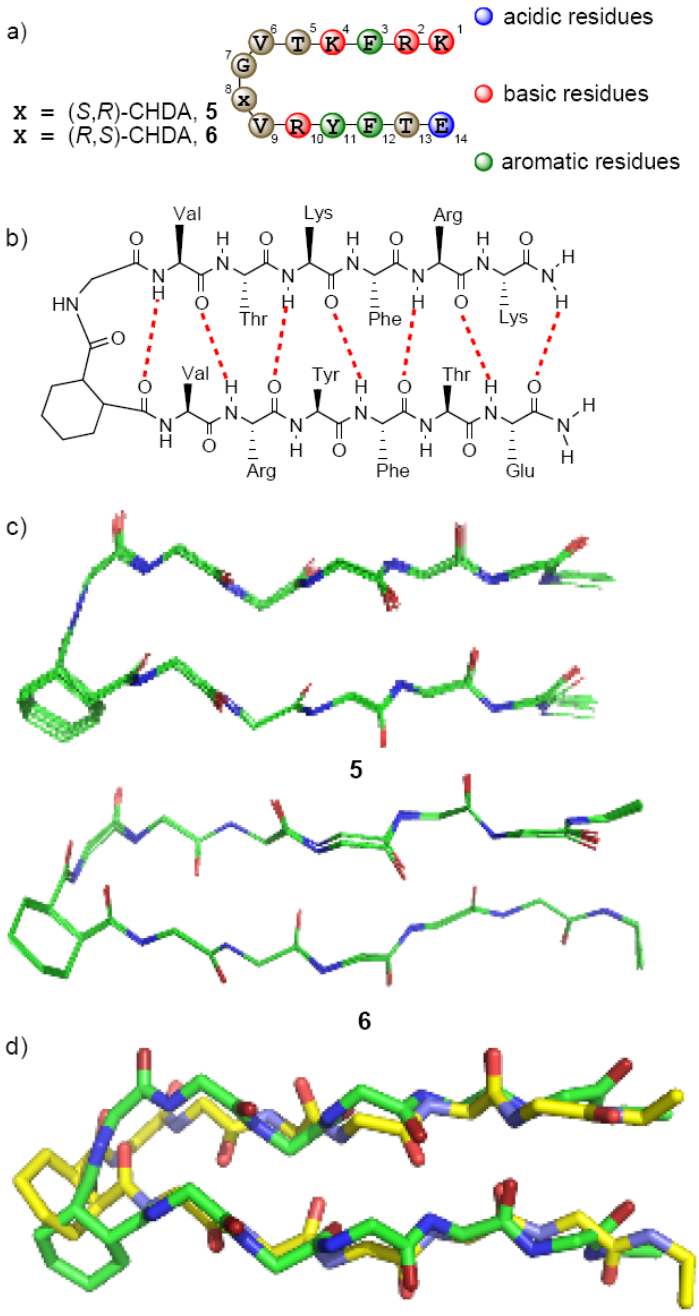

Figure 3.

(a) Sequence for N to N β-hairpins 5 and 6 (red = basic residues, blue = acidic residues, green = aromatic residues). (b) Hydrogen bond pattern in 5 and 6 for the desired parallel β-sheet conformation. (c) Ten best NMR structures obtained by NOE restrains calculations using CNS. The RMSD among backbone heavy atoms for the best structures of 5 is 0.128 ± 0.028Å and of 6 is 0.097± 0.029Å. d) Overlay of the NMR struct ures of peptides 5 (green) and 6 (yellow). RMSD between all the atoms of the backbone 0.326Å.

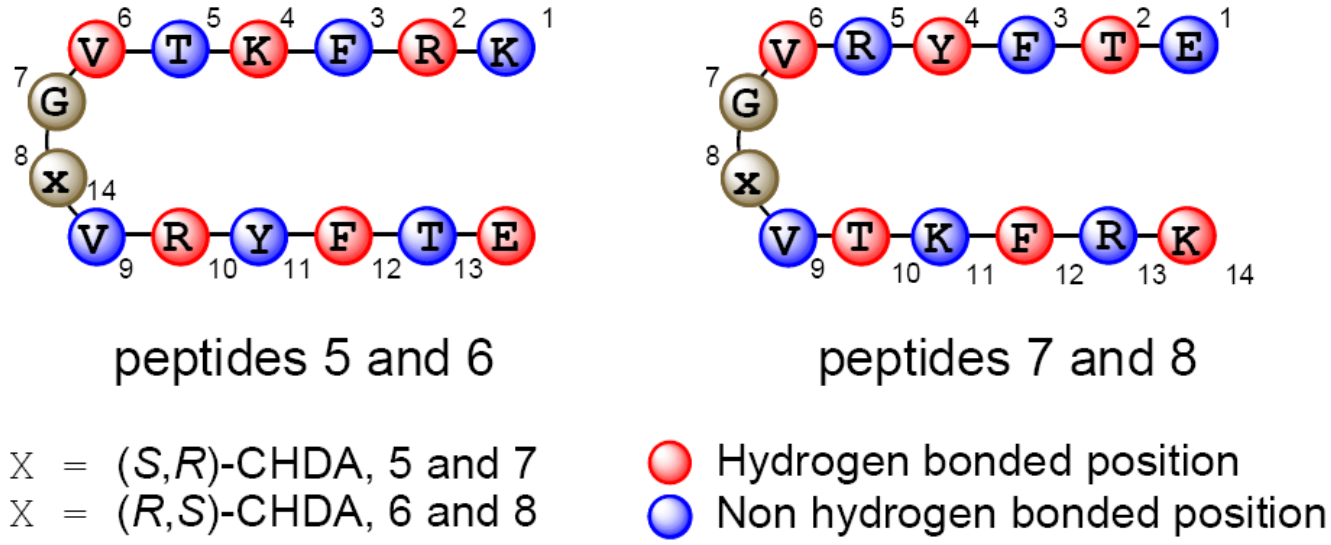

Figure 4.
Sequences of 5-8.