Synthesis and structural studies of peptides containing a mannose-derived furanoid sugar amino acid†‡

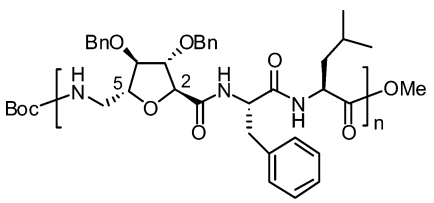
Tushar Kanti Chakraborty,* Dipankar Koley, Rapolu Ravi and Ajit Chand Kunwar*

Received 10th August 2007, Accepted 19th September 2007
First published as an Advance Article on the web 3rd October 2007
DOI: 10.1039/b712365p

A mannose-derived furanoid sugar amino acid (Maa) induced helical turns in peptides having repeat units of Maa(Bn_2)-Phe-Leu, which aggregated into head-to-tail duplexes in the longer oligomers.

Introduction

Since they were first reported as useful peptide building blocks, the pyranoid[1] and furanoid[2] sugar amino acids have been used extensively by many research groups world wide as conformationally constrained scaffolds in peptidomimetic studies.[3] They have also emerged as an important class of synthetic monomers leading to many *de novo* oligomeric libraries with interesting structures and properties.[4] For example, the linear homooligomers of a glucose-derived furanoid sugar amino acid (Gaa) had a well-defined structure in $CDCl_3$ with repeating β-turns, each involving a 10-membered ring structure with intramolecular hydrogen bonds between $NH_i \rightarrow C{=}O_{i-2}$.[4g] A very similar structure was also seen in the corresponding heterooligomers with repeating Gaa-Leu-Val units which displayed well-defined turn structures with repetitive 10/10/9-H-bonding patterns.[4e] In contrast, the octameric chain of *C*-glycosyl α-D-*lyxo*furanose configured tetrahydrofuran amino acid, having the C-2 and C-5 substituents on the tetrahydrofuran ring *trans* to each other, adopted a well-defined helical secondary structure stabilized by 16-membered (i, $i-3$) inter-residue hydrogen bonds, similar to a π-helix, in a nonpolar solvent like $CDCl_3$.[4d] However, homooligomers of a mannose-derived sugar amino acid (Maa), having a similar 2,5-*trans* configuration, failed to exhibit any well-defined structure in polar solvents.[4n] In the present study, we investigate the conformational biases imposed by the Maa residues in the linear oligomeric peptides **1–4** having repeat units of Maa(Bn_2)-Phe-Leu.



1: n = 1, **2**: n = 2, **3**: n = 3, **4**: n = 4

Indian Institute of Chemical Technology, Hyderabad, 500 007, India. E-mail: chakraborty@iict.res.in, kunwar@iict.res.in

† IICT Communication No. 060304

‡ Electronic supplementary information (ESI) available: Characterisation data for **1–4** and MD simulation details for **3** and **4**. See DOI: 10.1039/b712365p

Results and discussion

Compounds **1–4** were synthesized using standard solution phase peptide coupling methods and the final products were purified by silica gel column chromatography and fully characterized by spectroscopic methods before the conformational studies.

Detailed NMR studies of **1–4** were carried out in $CDCl_3$. For **1** and **2**, only one set of peaks was observed, even when the temperature was lowered to 228 K. Amide proton chemical shifts and solvent titration studies support the involvement of the PheNH in hydrogen bonding.[5] The presence of NOE connectivities between PheNH and the preceding MaaC5H support PheNH–CO(Boc) H-bonds in **1** and **2** and PheNH(5)-Maa(4)CO H-bonds in **2**, which result in a 10-membered pseudo β-turn in **1** and two such turns in **2**, like those observed earlier.[4c] Evidence for these sheet-like structures is further provided by the $^3J_{NH-C\alpha H}$ being >8 Hz.

NMR spectra of both **3** and **4** in $CDCl_3$ displayed two sets of peaks. The major to minor populations are in the ratio of about 3 : 2 for **3**, whereas for **4** we could not obtain the exact population of the minor isomer; however it was believed to be <5%. The exchange peaks in the ROESY spectra between these species show that they are in equilibrium with each other. The rate of exchange between the two species was slow in the NMR time scale, being 0.9 Hz and 0.1 Hz for **3** and **4**, respectively.[5,6] The first indications of the major isomer having a different and compact structure in **3** and **4**, compared to **1** and **2**, reside in the very large (~8 ppm) chemical shifts of all the amide protons (δ_{NH}), excluding the BocNH. Solvent titration studies were carried out by adding DMSO-d_6 to the $CDCl_3$ solutions of the peptides **3** and **4**. Insignificant changes were observed in the chemical shifts of the amide protons on addition of up to 6% DMSO-d_6. This indicates that most of the amides of the major species of **3** and **4** were H-bonded. Fig. 1 shows the solvent titration plots for **4**. When more DMSO-d_6 was added to the $CDCl_3$ solution, a gradual growth of the minor species at the expense of the major one was observed. In pure DMSO-d_6 only the peaks of the minor species were seen, that belonged to structures similar to those observed for **1** and **2**.

The NMR data for the major species in **3** and **4** strongly suggest the presence of dimeric structures in $CDCl_3$ solution. The extensive network of H-bonds seen in the NMR spectra probably arises due to the formation of this dimer-like association of the individual strands. These observations, along with the changes in the isomer populations during variable temperature

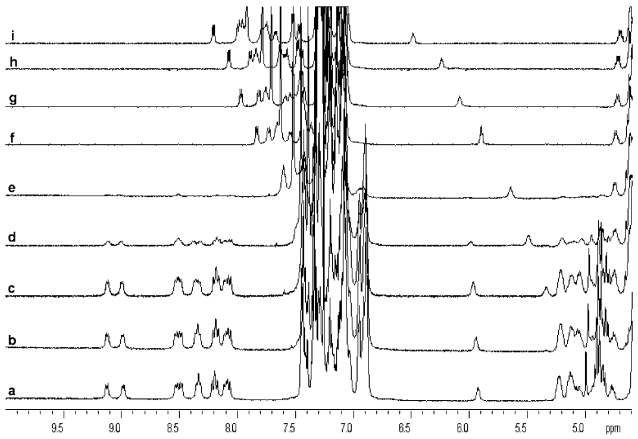


Fig. 1 Solvent titration plots of **4** (500 MHz, 303 K, $CDCl_3$ with 0–33.3% DMSO-d_6): (a) NMR spectrum of **4** in pure $CDCl_3$; (b) to (i) NMR spectra of **4** on addition of 20 μL, 40 μL, 60 μL, 80 μL, 120 μL, 160 μL, 200 μL and 300 μL of DMSO-d_6 to 600 μL $CDCl_3$ solution.

and dilution studies, imply a possible equilibrium between a single- (minor) and a double-stranded (major) species. The dimerisation constants (K_{dim}), for **3** with 40% and **4** with 5% populations of the monomer, were 4.7×10^2 M^{-1} and 3.2×10^4 M^{-1}, respectively.[5,7] ESI mass spectral studies of compounds **3** and **4** showed the presence of doubly-charged dimeric species, in addition to the singly-charged monomer, further supporting the dimerisation of two monomeric units. The experimental peak shapes were matched with the calculated isotopic distribution patterns and the mass differences between the isotopic peaks arising from the doubly-charged dimeric species were 0.5 Da and those from the singly charged monomeric species showed differences of 1.0 Da.[5]

The structure of **4** is discussed here in detail as it existed almost exclusively as a single species. The $^3J_{NH-C\alpha H}$ being >9 Hz for all but the Maa(1) residue strongly supported an antiperiplanar arrangement of the amide and the Cα protons, corresponding to $\phi \sim -120°$ in the β-region of the Ramachandran map, suggesting a basic β-stranded structure. Additional evidence for this structure was provided by the CαH chemical shifts being >5 ppm for both Phe and Leu residues. Since the proteins containing a β-sheet structure have been shown to have the β-strands with a natural right handed twist, it was envisaged that the structure in **4** could also have twisted strands. The $^3J_{NH-C\alpha H}$ of >10.5 Hz for Phe was followed by <9.3 Hz for Leu throughout the strand implying the presence of alternating ϕ magnitudes, a requirement for twisted β-sheets.[8]

The dimer-like association of the individual strands is evident from the large number of inter-residue NOEs shown schematically in Fig. 2. The observed NOEs between the protons at two termini, like Boc–OMe, Maa(1)NH–OMe, Maa(1)NH–Leu(12)CαH and Phe(2)NH–OMe (Fig. 3 shows some of them), provide strong evidence for the proposed anti-parallel dimeric structure. NOE connectivities, Leu(3)CαH–Leu(9)CαH, Leu(3)CαH–Maa(10)NH, Leu(3)NH–Leu(12)NH, Maa(1)C2H–Phe(11)CαH, Maa(4)NH–Leu(9)CαH, Phe(2)CαH–Maa(10)C2H and Leu(6)NH–Leu(9)NH, are consistent with a distinctly complex duplex structure consisting of two anti-parallel twisted β-strands, with all the above-mentioned NOEs being inter-strand in nature.[9] All the NMR observations, *e.g.* $^3J_{NH-C\alpha H}$, H-bonding information and NOEs across the strands, support the formation of a duplex structure from twisted strands.

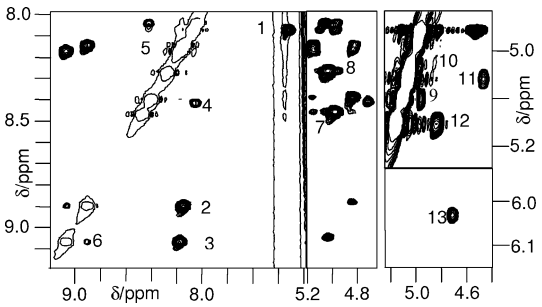


Fig. 2 The characteristic NOEs in the ROESY spectrum of **4**: 1, Phe(2)NH–Leu(3)NH; 2, Phe(5)NH–Leu(6)NH; 3, Phe(8)NH–Leu(9)NH; 4, Phe(11)NH–Leu(12)NH; 5, Leu(3)NH–Leu(12)NH; 6, Leu(6)NH–Leu(9)NH; 7, Maa(4)NH–Leu(9)CαH; 8, Leu(3)CαH–Maa(10)NH; 9, Phe(2)CαH–Maa(10)C2H; 10, Leu(3)CαH–Leu(9)CαH; 11, Maa(1)C2H–Phe(11)CαH; 12, Maa(4)C2H–Phe(8)CαH and Phe(5)CαH–Maa(7)C2H; 13, Maa(1)NH–Leu(12)CαH.

The β-sheets and strands, especially in globular proteins, display right handed twists[10] and the extent of curvature depends on many factors like the entropic considerations, the number and types of residues, the torsional angles of the peptide chains, intra- and

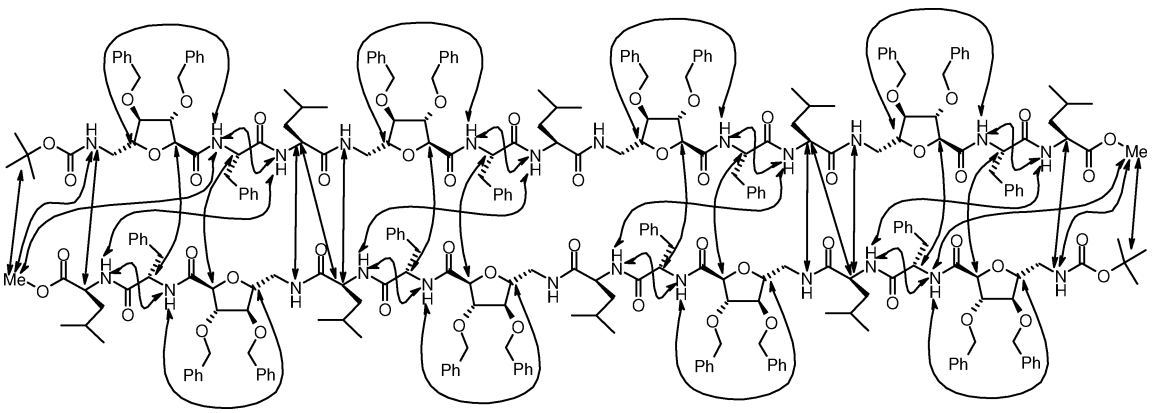


Fig. 2 Schematic representation of the intra- and inter-residue NOEs seen in the ROESY spectrum of **4**.

 This journal is © The Royal Society of Chemistry 2007

inter-strand non-covalent interactions, *etc.*[11–13] Theoretical calculations have shown that periodicity of the twists depends on the nature of the β-sheets.[10,11] For an anti-parallel arrangement, the periodicity is about 15 residues per turn. The mannose-derived furanoid sugar amino acid with 2,5-*trans* geometry permitted significant changes in the direction of the chain propagation, thereby introducing enhanced twists in the strands, which might have resulted in the nucleation of novel structures deduced from the NMR data.

Incorporating all the above information, molecular dynamics calculations were carried out using the distance constraints from the volume integrals from the ROESY correlations and two-spin approximation. Fig. 4 shows the stereo-view of one of the lowest energy MD structures. The structures were revealing. The presence of anti-parallel twisted dimeric β-strands, as envisaged, resulted in an unprecedented right-handed double helical structure. The average values of the dihedral angles, ϕ and ψ, of the backbone from these structures are $-127 \pm 9°$ and $-59 \pm 8°$ for Phe and $-96 \pm 12°$ and $138 \pm 14°$ for Leu, respectively.

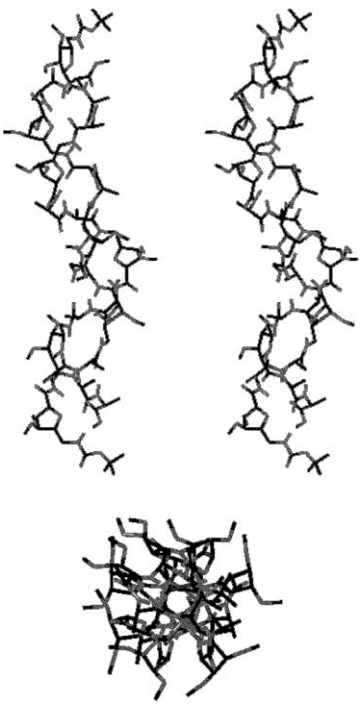


Fig. 3 Stereo-view of one of the energy-minimized MD structures of **4** (top) and its top view (bottom). The benzyl groups and the amino acid side-chains are not shown for clarity.

The MD structures clearly show that the Maa residues impart significant twists in the individual strands with a pitch of ~18 Å and a periodicity of about 6 residues per turn. Two of these strands intertwine in an anti-parallel double helical structure with an inner pore of ~1 Å in diameter (Fig. 4). All the observed NOEs can now be rationalized based on the findings from the MD calculations. We have, for further discussion, differentiated the two strands by marking the residues as natural and primed numbers. Leu carbonyls are involved in three centre H-bonds with Phe and Leu amide protons in the other chain, *e.g.* Leu(3)CO with Phe(11′)NH and Leu(12′)NH, Leu(6)CO with Phe(8′)NH and Leu(9′)NH, Leu(9)CO with Phe(5′)NH and Leu(6′)NH and Leu(12)CO with Phe(2′)NH and Leu(3′)NH. The strong NOE correlations between all the adjacent PheNH and LeuNH are consistent with these observations. The observed H-bonding patterns are also supported by the IR spectrum of **4** in $CHCl_3$, where three amide bands were seen, both in the NH stretch region with values of 3376, 3331 and 3280 cm^{-1} and the C=O stretch region with values of 1676, 1653, and 1639 cm^{-1}. The presence of these bands in the IR spectrum is attributed to all the amides with no H-bonds, single H-bonds as well as three centre H-bonds, based on earlier studies.[14]

All the NMR observations, like $^3J_{NH\text{-}C\alpha H}$, H-bonding information and NOEs across the strands, suggest a similar duplex structure for **3**. The structures for the minor species in both **3** and **4** could not be obtained in detail. However, some of the NMR spectral parameters and NOEs which could be attributed to the minor conformers suggest that their structures were like those proposed for **1** and **2**.

Conclusion

In conclusion, detailed NMR and mass spectral studies carried out here indicate strong aggregation of **3** and **4** into dimeric structures. Furthermore, the inter-strand NOEs seen across the length of these oligomers strongly suggest that the strands are intertwined. The double helical structure described here is unique, very similar to those observed earlier for gramicidins and their structural analogs.[15] It is likely that studies in other solvents and on peptides with different sequences can lead to other interesting structures. Incorporation of charged residues in lieu of Phe and Leu can also provide scope for additional interactions, though at the cost of deviation from a hydrophobic peptide core important for generation of the twisted β-strands.

Experimental section

The peptides **1–4** were synthesized following standard solution phase peptide coupling methods[16] using 1-ethyl-3-[3-(dimethylamino)propyl]-carbodiimide hydrochloride (EDCI) and 1-hydroxybenzotriazole (HOBt) as coupling agents and dry DMF and/or CH_2Cl_2 as solvents. In the racemization free fragment condensation strategy that was followed, the Boc-Maa(Bn_2)-OH was first coupled with the dipeptide H-Phe-Leu-OMe as efficiently as with any normal amino acid using the reagents mentioned above to give the tripeptide Boc-Maa(Bn_2)-Phe-Leu-OMe **1**. While saponification of one-half of **1** using LiOH in THF–MeOH–H_2O gave the acid, Boc-deprotection of the other half using TFA–CH_2Cl_2 freed the *N*-terminus, which was then coupled with the acid to furnish **2**. Coupling of **1** and **2** under similar conditions gave **3** and '**2** + **2**' provided the linear tetramer **4**.

Spectral data of peptide 1

^{1}H NMR ($CDCl_3$, 500 MHz): see Table S1 in the supporting information‡; HRMS (ESI) *m/z* calculated for $[M + H]^+$ $C_{41}H_{54}N_3O_9$: 732.3860, found: 732.3850.

Spectral data of peptide 2

^{1}H NMR ($CDCl_3$, 500 MHz): see Table S2 in the supporting information‡; HRMS (ESI) *m/z* calculated for $[M + NH_4]^+$ $C_{76}H_{98}N_7O_{15}$: 1348.7120, found: 1348.7154.

Spectral data of peptide 3

^{1}H NMR ($CDCl_3$, 600 MHz): see Table S3 in the supporting information‡; HRMS (ESI) *m/z* calculated for $[M + NH_4]^+$ $C_{111}H_{139}N_{10}O_{21}$: 1948.0116, found: 1948.1011.

Spectral data of peptide 4

^{1}H NMR ($CDCl_3$, 750 MHz): see Table S4 in Supporting Information‡; HRMS (ESI) *m/z* calculated for $[M + Na]^+$ $C_{146}H_{176}N_{12}O_{27}Na$: 2552.2665, found: 2552.2766.

Acknowledgements

Authors wish to thank DST, New Delhi for financial support. D.K. and R.R. are thankful to CSIR, New Delhi for research fellowships. A.C.K. thanks Profs. M. Goerlach and R. Ramachandran for 750 MHz NMR spectra and INSA-DFG for supporting the exchange program. Authors thank the Mass Spectrometry Centres at IICT and CCMB for their help.

References

1 E. Graf von Roedern and H. Kessler, *Angew. Chem., Int. Ed. Engl.*, 1994, **33**, 687.
2 T. K. Chakraborty, S. Jayaprakash, P. V. Diwan, R. Nagaraj, S. R. B. Jampani and A. C. Kunwar, *J. Am. Chem. Soc.*, 1998, **120**, 12962.
3 For reviews on sugar amino acids see: (*a*) K. J. Jensen and J. Brask, *Peptide Sci.*, 2005, **80**, 747; (*b*) T. K. Chakraborty, P. Srinivasu, S. Tapadar and B. K. Mohan, *Glycoconjugate J.*, 2005, **22**, 83; (*c*) T. K. Chakraborty, P. Srinivasu, S. Tapadar and B. K. Mohan, *J. Chem. Sci.*, 2004, **116**, 187; (*d*) S. A. W. Gruner, E. Locardi, E. Lohof and H. Kessler, *Chem. Rev.*, 2002, **102**, 491; (*e*) F. Schweizer, *Angew. Chem., Int. Ed.*, 2002, **41**, 230; (*f*) T. K. Chakraborty, S. Ghosh and S. Jayaprakash, *Curr. Med. Chem.*, 2002, **9**, 421; (*g*) T. K. Chakraborty, S. Jayaprakash and S. Ghosh, *Comb. Chem. High Throughput Screening*, 2002, **5**, 373; (*h*) F. Peri, L. Cipolla, E. Forni, B. La Ferla and F. Nicotra, *Chemtracts: Org. Chem.*, 2001, **14**, 481; (*i*) J. Gervay-Hague and T. M. Weathers, Pyranosyl Sugar Amino Acid Conjugates: Their Biological Origins, Synthetic Preparations and Structural Characterisation, in *Glycochemistry: Principles, Synthesis and Applications*, ed. P. G. Wang and C. R. Bertozzi, Dekker, New York, 2001.
4 For some earlier works on oligomers of small linear peptides containing sugar amino acids and linear homooligomers of sugar amino acids see: (*a*) Y. Suhara, M. Kurihara, A. Kittaka and Y. Ichikawa, *Tetrahedron*, 2006, **62**, 8207; (*b*) A. A. Edwards, G. J. Sanjayan, S. Hachisu, G. E. Tranter and G. W. J. Fleet, *Tetrahedron*, 2006, **62**, 7718; (*c*) M. D. P. Risseeuw, G. M. Grotenbreg, M. D. Witte, A. W. Tuin, M. A. Leeuwenburgh, G. A. Van der Marel, H. S. Overkleeft and M. Overhand, *Eur. J. Org. Chem.*, 2006, 3877; (*d*) T. D. W. Claridge, D. D. Long, C. M. Baker, B. Odell, G. H. Grant, A. A. Edwards, G. E. Tranter, G. W. J. Fleet and M. D. Smith, *J. Org. Chem.*, 2005, **70**, 2082; (*e*) T. K. Chakraborty, S. Roy, S. K. Kumar and A. C. Kunwar, *Tetrahedron Lett.*, 2005, **46**, 3065; (*f*) L. F. Bornaghi, B. L. Wilkinson, M. J. Kiefel and S.-A. Poulsen, *Tetrahedron Lett.*, 2004, **45**, 9281; (*g*) T. K. Chakraborty, P. Srinivasu, S. S. Madhavendra, S. K. Kumar and A. C. Kunwar, *Tetrahedron Lett.*, 2004, **45**, 3573; (*h*) T. Q. Gregar and J. Gervay-Hague, *J. Org. Chem.*, 2004, **69**, 1001; (*i*) D. F. A. Hunter and G. W. J. Fleet, *Tetrahedron: Asymmetry*, 2003, **14**, 3831; (*j*) M. D. Smith, T. D. W. Claridge and M. S. P. Sansom, *Org. Biomol. Chem.*, 2003, **1**, 3647; (*k*) A. Vescovi, A. Knoll and U. Koert, *Org. Biomol. Chem.*, 2003, **1**, 2983; (*l*) S. A. W. Gruner, V. Truffault, G. Voll, E. Locardi, M. Stöckle and H. Kessler, *Chem.–Eur. J.*, 2002, **8**, 4365; (*m*) Y. Suhara, Y. Yamaguchi, B. Collins, R. L. Schnaar, M. Yanagishita, J. E. K. Hildreth, I. Shimada and Y. Ichikawa, *Bioorg. Med. Chem.*, 2002, **10**, 1999; (*n*) T. K. Chakraborty, S. Jayaprakash, P. Srinivasu, M. G. Chary, P. V. Diwan, R. Nagaraj, A. R. Sankar and A. C. Kunwar, *Tetrahedron Lett.*, 2000, **41**, 8167; (*o*) D. E. A. Brittain, M. P. Watterson, T. D. W. Claridge, M. D. Smith and G. W. J. Fleet, *J. Chem. Soc., Perkin Trans. 1*, 2000, 3655; (*p*) N. L. H. Hungerford, T. D. W. Claridge, M. P. Watterson, R. T. Aplin, A. Moreno and G. W. J. Fleet, *J. Chem. Soc., Perkin Trans. 1*, 2000, 3666; (*q*) N. Hungerford and G. W. J. Fleet, *J. Chem. Soc., Perkin Trans. 1*, 2000, 3680.
5 See Supporting Information for details‡.
6 B. Reif, V. Wittmann, H. Schwalbe, C. Griesinger, K. Worner, K. Jahn-Hofmann and J. W. Engels, *Helv. Chim. Acta*, 1997, **80**, 1952.
7 C. S. Wilcox, in *Frontiers in Supramolecular Organic Chemistry and Photochemistry*, ed. H.-J. Schneider, H. Durr, VCH, New York, 1991; p. 123.
8 A. G. Cochran, N. J. Skelton and M. A. Starovasnik, *Proc. Natl. Acad. Sci. U. S. A.*, 2001, **98**, 5578.
9 The other likely NOEs, Maa(4)C2H–Phe(8)CαH and Phe(5)CαH–Maa(7)C2H overlapped (cross-peak 12 in Fig. 1) and could not be distinguished. However, in analogy with the data for **3**, we believe that these two long-range NOEs are also present in **4**, although they were not used in the MD calculations.
10 D. Znamenskiy, K. L. Tuan, A. Poupon, J. Chomilier and J.-P. Mornon, *Protein Eng.*, 2000, **13**, 407.
11 L. Wang, T. O'Connell, A. Tropsha and J. Hermans, *J. Mol. Biol.*, 1996, **262**, 283.
12 C. Chothia, *J. Mol. Biol.*, 1973, **75**, 295.
13 A. S. Yang and B. Honig, *J. Mol. Biol.*, 1995, **252**, 366.
14 (*a*) E. Vass, M. Hollossi, F. Besson and R. Buchet, *Chem. Rev.*, 2003, **103**, 1917; (*b*) T. Moriuchi, T. Tamura and T. Hirao, *J. Am. Chem. Soc.*, 2002, **124**, 9356; (*c*) J. Yang, L. A. Christianson and S. H. Gellman, *Org. Lett.*, 1999, **1**, 11; (*d*) T. Heimburg, J. Schunemann, K. Weber and N. Geisler, *Biochemistry*, 1999, **38**, 12727; (*e*) T. Blundell, D. Barlow, N. Borkakoti and J. Thornton, *Nature*, 1983, **306**, 281.
15 (*a*) M. Sastry, C. Brown, G. Wagner and T. D. Clark, *J. Am. Chem. Soc.*, 2006, **128**, 10650; (*b*) H.-D. Arndt, D. Bockelmann, A. Knoll, S. Lamberth, C. Griesinger and U. Koert, *Angew. Chem., Int. Ed.*, 2002, **41**, 4062; (*c*) B. M. Burkhart, R. M. Gassman, D. A. Langs, W. A. Pangborn, W. L. Duax and V. Pletnev, *Biopolymers*, 1999, **51**, 129; (*d*) B. Di Blasio, E. Benedetti, V. Pavone, C. Pedone, C. Gerber and G. P. Lorenzi, *Biopolymers*, 1989, **28**, 203; (*e*) D. A. Langs, *Science*, 1988, **241**, 188; (*f*) A. S. Arseniev, I. L. Barsukov and V. F. Bystrov, *FEBS Lett.*, 1985, **180**, 33; (*g*) G. P. Lorenzi, H. Jackle, L. Tomasic, V. Rizzo and C. Pedone, *J. Am. Chem. Soc.*, 1982, **104**, 1728.
16 (*a*) M. Bodanszky and A. Bodanszky, *The Practices of Peptide Synthesis*, Springer-Verlag, New York, 1984; (*b*) G. A. Grant, *Synthetic Peptides: A User's Guide*, W. H. Freeman, New York, 1992; (*c*) M. Bodanszky, *Peptide Chemistry: A Practical Textbook*, Springer-Verlag, Berlin, 1993.

This journal is © The Royal Society of Chemistry 2007