NMR structural analysis of a β-hairpin peptide designed for DNA binding

Allister J. Maynard and Mark S. Searle*

Department of Chemistry, University of Nottingham, University Park, Nottingham, UK NG7 2RD

NMR and circular dichroism (CD) spectroscopy shows that an unconstrained 16 residue linear peptide folds autonomously in water into a β-hairpin: the designed peptide adopts a conformation that mimics the anti-parallel β-sheet DNA binding motif of the *met* repressor protein dimer with key residues for DNA recognition presented in the same positions and orientations principally on one face of the β-hairpin template.

Recent studies have shown that small peptides of 14 to 17 residues, derived from the Bovine Immunodeficiency Virus (BIV) Tat protein[1] and HIV Rev protein,[2] can bind to nucleic acids with high affinity as β-hairpin or α-helical structures inserted into the major groove of RNA. Such model systems open up the possibility of using designed peptides or non-peptidic analogues for sequence specific recognition of nucleic acids in a therapeutically beneficial manner, either through blocking transcription factor binding[3] or by delivering some reactive functional group to the surface of DNA or RNA where single or double strand damage could be inflicted.[4]

The *met* repressor protein dimer represents a novel class of DNA binding proteins that recognise DNA in a sequence specific manner through insertion of a two stranded anti-parallel β-sheet into the major groove.[5] The key residues for binding and recognition by the *met* repressor are shown in Fig. 1(*a*). We have designed a linear 16-residue peptide analogue of this binding motif, that incorporates all of these key residues, by linking two β-strands *via* a two residue turn sequence such that the 16-mer can fold to form a β-hairpin structure [Fig. 1(*b*)] with the correct anti-parallel alignment of β-strands. The most abundant two residue turns in the protein structure data base are of the type I′ variety with the sequence NG particularly favoured.[6,7] To remove some of the sequence degeneracy inherent in the peptide and aid structural analysis by NMR, a tyrosine has been introduced at position 3 on the non-binding face (in place of Ile in the native sequence) to provide a good NMR probe for folding through long range NOEs. Here we describe the characterisation of the 16-mer [Fig. 1(*b*)], which we show by NMR and CD is appreciably folded in aqueous solution, representing one of only a few examples of short peptides that adopt compact folded β-structures in water.[6,8]

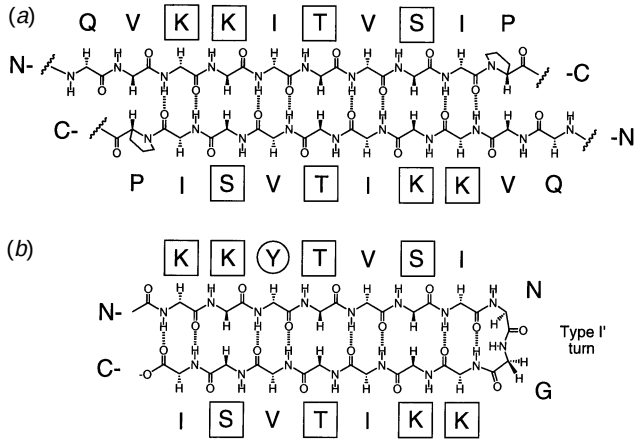

Fig. 1 (*a*) Anti-parallel alignment of the two β-strands of the DNA recognition motif of the *met* repressor protein dimer. Key residues for DNA binding and recognition are highlighted in square boxes. (*b*) Designed 16 residue β-hairpin analogue of (*a*). Peptides were synthesised on a 0.1 mM scale using standard F-moc solid-phase protocols followed by purification by reverse phase HPLC using an Applied Biosystems Aquaphore Octyl column (10 mm × 100 mm). Purity was assessed by mass spectral analysis and ^{1}H NMR spectroscopy.

The *N*-acetylated 16-mer [Fig. 1(*b*)] has been studied by far-UV circular dichroism (CD) spectroscopy in aqueous solution at pH 4.0 and 7.5 μM concentration. A significant negative ellipticity at 216 nm suggests that the peptide forms a high proportion of folded structure in purely aqueous solution with the absorption profile reminiscent of β-structure (both turn and sheet) in equilibrium with some random coil (negative ellipticity below 198 nm)[9] (data not shown). Addition to the peptide of small amounts of methanol (< 20% v/v) produces a significant deepening of the minimum at 216 nm and formation of a maximum at 200 nm indicative of a displacement of the equilibrium already present in aqueous solution further towards the folded state, as has been reported for a number of other model peptide systems.[10]

The most conclusive evidence for folded structure is the observation of long range NOEs between residues in the anti-parallel strands of the β-hairpin. A full resonance assignment has been possible using established procedures,[11] allowing such

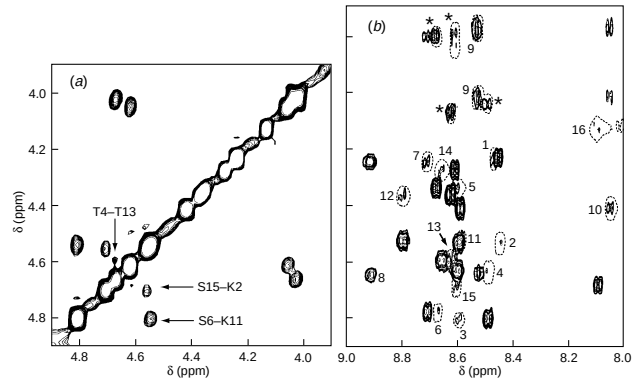

Fig. 2 (*a*) Portion of the NOESY spectrum (200 ms mixing time) of the 16-mer at 278 K illustrating cross-strand αα(i,j) NOEs as highlighted. The intensity of the T4–T13 NOE cross-peak is distorted due to its position close to the diagonal; (*b*) Finger-print regions (NH to Hα) of NOESY (200 ms) and TOCSY spectra (overlayed) of the 16-mer. TOCSY data are plotted as a single broken contour. In general, αN(i,i) NOEs (occuring at the same positions as TOCSY cross-peaks) are very weak for residues in extended β-strands (residues 1–7 and 10–16), while αN(i,i + 1) sequential interresidue NOEs are much more intense. Note: for N8 and G9 NOEs αN(i,i) and αN(i,i + 1) are of comparable intensity reflecting their positions in the β-turn. Asterisks mark the position of NH to Hβ cross-peaks in both NOESY and TOCSY. TOCSY NH to Hα cross-peaks are numbered according to the sequence: N-Ac-K^{1}K^{2}Y^{3}T^{4}V^{5}S^{6}I^{7}N^{8}G^{9}K-^{10}K^{11}I^{12}T^{13}V^{14}S^{15}I^{16}. ^{1}H NMR spectra were recorded at 500 MHz on a Bruker DRX500 spectrometer using standard phase sensitive 2D NMR pulse sequences. Attenuation of the residual solvent peak in NOESY and TOCSY spectra recorded in 90% H_2O solution was achieved by presaturation. The sodium salt of trimethylsilyl propionate was used as an internal chemical shift reference.

long range interactions to be identified. The main chain alignment shown in Fig. 1(*b*) brings a number of Hαs of non-hydrogen bonded residues into close proximity. Accordingly, several intense interstrand αα(i,j)† interactions are readily detected in both NOESY and ROESY spectra recorded in aqueous solution. In Fig. 2(*a*), NOEs are identified between Ser6-Lys11, Thr4-Thr13 and Lys2-Ser15 that are fully consistent with the proposed anti-parallel alignment of the peptide chain. These observations are substantiated by studies in 90% H_2O solution were cross-strand long-range NN(i,j)† interactions between residues involved in hydrogen bonding are also evident (Lys1-Ile16, Val5-Ile12 and Ile7-Lys10). The first of these interactions (Lys1-Ile16) illustrates that even the two terminal residues are held in close proximity for an appreciable proportion of the time, indicative of the peptide adopting a well-defined conformation in aqueous solution. While these NOEs provide evidence for long range order in the peptide chain, the intensity of NOEs between adjacent residues provide additional support for local order between nearest neighbours. The majority of residues give rise to intense αN(i,i + 1) sequential NOEs but very weak (or undetected) NN(i,i + 1) and αN(i,i) interactions [Fig. 2(*b*)], consistent with the involvement of many residues in an extended β-strand conformation.[12] In the random coil these NOEs are expected to be much more similar in intensity due to averaging over a large number of conformational states.[12] Significantly different NOE intensities are observed for residues at the centre of the sequence where, for example, a strong sequential NN(i,i + 1) NOE between G9 and K10, together with similar intensity αN(i,i) and αN(i,i + 1) NOEs for N8 and G9 [Fig. 2(*b*)], are diagnostic of the involvement of these residues in a β-turn.

Chemical shift values for Hα resonances have been widely used as an indicator of folded structure in solution.[6,8,10] Positive deviations from tabulated random coil values[11] generally correlate with residues found in extended β-strand conformations, while negative (upfield) shifts are associated with residues in α-helices or β-turn conformations.[13] In the present case, positive deviations (Δδ) are evident for residues 2–7 and 10–15, while negative (upfield) shifts are identified for residues at the centre of the sequence. The plot of Δδ *versus* amino acid sequence clearly has the characteristic appearance of a β-hairpin, that is, two β-strands separated at the centre of the sequence by a turn. More pronounced effects are apparent in the presence of low concentrations of methanol reflecting the effects of the organic solvent in displacing the equilibrium towards the folded state (Fig. 3). We have established that all

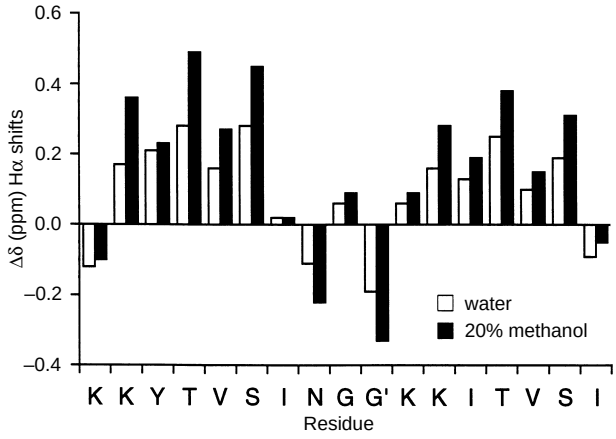

Fig. 3 Deviations of CαH chemical shifts from random coil values in aqueous solution and 20% v/v methanol, pH 4.5 and 278 K

chemical shifts and line widths are independent of sample concentration over a *ca*. 300-fold dilution range (8.0 to 0.03 mM), indicating that the reported effects are a consequence of folding of the monomeric peptide rather than the formation of aggregates.

We have shown unambiguously by CD and NMR experiments that the designed peptide forms a high proportion of folded structure in purely aqueous solution with the desired alignment of the peptide backbone. The 16-mer represents one of only a few examples of short linear peptides that are able to adopt compact folded β-hairpin structures in water, providing a useful model system for probing β-sheet forming interactions relevant to protein folding mechanisms, with the added potential for DNA recognition.

We acknowledge the EPSRC (studentship to A. J. M.), The Royal Society, Nuffield Foundation and the Department of Chemistry at the University of Nottingham for financial support. We thank John Keyte for peptide synthesis and Dr Gary Sharman for helpful discussions.

Footnotes

* E-mail: mark.searle@nottingham.ac.uk

† αα(i,j) interresidue NOE between Hαs of residues i and j; NN(i,j)-NH to NH NOE between residues i and j; αN(i,i + 1)-NOE between Hα of residue i and NH of residue i + 1.

References

1 J. D. Puglisi, L. Chen, S. Blanchard and A. D. Frankel, *Science*, 1995, **270**, 1200; X. Ye, R. A. Kumar and D. J. Patel, *Chem. Biol.*, 1995, **2**, 827.
2 J. L. Battiste, R. A. D. Tan and J. R. Williamson, *J. Biomol. NMR*, 1995, **6**, 375; R. D. Peterson, D. P. Bartel, J. W. Szostak, S. J. Horvath and J. Feigon, *Biochemistry*, 1994, **33**, 5357.
3 J. J. Welch, F. J. Rauscher II and T. A. Beerman, *J. Biol. Chem.*, 1994, **269**, 31 051; D. Henderson and L. H. Hurley, *Nature Med.*, 1995, **1**, 525.
4 P. G. Schultz, J. S. Taylor and P. B. Dervan, *J. Am. Chem. Soc.*, 1982, **104**, 6861; R. P. Herzberg and P. B. Dervan, *J. Am. Chem. Soc.*, 1982, **104**, 313.
5 W. S. Somers and S. E. V. Philips, *Nature*, 1992, **359**, 387.
6 F.J. Blanco, G. Rivas and L. Serrano, *Nat. Struct. Biol.*, 1994, **1**, 584.
7 E. G. Hutchinson and J. M. Thornton, *Protein Sci.*, 1994, **3**, 2207.
8 F. J. Blanco, M. A. Jimenez, J. Herranz, M. Rico, J. Santoro and J. L. Nieto, *J. Am. Chem. Soc.*, 1993, **115**, 5887; F. J. Blanco and L. Serrano, *Eur. J. Biochem.*, 1995, **230**, 634; M. S. Searle, D. H. Williams and L. C. Packman, *Nat. Struct. Biol.*, 1995, **2**, 999; V. Sieber and G. R. Moe, *Biochemistry*, 1996, **35**, 181; E. de Alba, M. A. Jimenez, M. Rico and J. L. Nieto, *Folding and Design*, 1996, **1**, 133.
9 J. T. Yang, C.-S. C. Wu and H. M. Martinez, *Methods Enzymol.*, 1986, **130**, 208; W. C. Johnson, Jr. *Annu. Rev. Biophys. Biophys. Chem.*, 1988, **17**, 145; H. J. Dyson and P. E. Wright, *Annu. Rev. Biophys. Biophys. Chem.*, 1991, **20**, 519.
10 F. J. Blanco, M. A. Jimenez, A. Pineda, M. Rico, J. Santoro and J. L. Nieto, *Biochemistry*, 1994, **33**, 6004; J. P. L. Cox, P. A. Evans, L. C. Packman, D. H. Williams and D. N. Woolfson, *J. Mol. Biol.*, 1993, **234**, 483; M. S. Searle, R. Zerella, D. H. Williams and L. C. Packman, *Protein Eng.*, 1996, **9**, 559.
11 K. Wuthrich, *NMR of Proteins and Nucleic Acids*, Wiley, New York, 1986.
12 L. J. Smith, K. A. Bolin, H. Schwalbe, M. W. MacArthur, J. M. Thornton and C. M. Dobson, *J. Mol. Biol.*, 1996, **255**, 494; L. J. Smith, K. M. Fiebig, H. Schwalbe and C. M. Dobson, *Folding and Design*, 1996, **1**, R95.
13 D. S. Wishart, B. D. Sykes and F. M. Richards, *Biochemistry*, 1992, **31**, 1647.

Received in Glasgow, UK, 16th April 1997; Com. 7/02593I