Novel hydrogen bonding ring motifs in a model peptide: crystal and molecular conformation

Ashwani K. Thakur, Ashish and Raghuvansh Kishore*

Institute of Microbial Technology, Sector 39-A, Chandigarh - 160 036, India. E-mail: kishore@imtech.ernet.in

Received (in Cambridge, UK) 14th June 1999, Accepted 16th July 1999

The crystal molecular structure of a model peptide, Boc-Ile-Thr-NH_2, unexpectedly revealed two unusual intramolecularly hydrogen bonded ring motifs, *i.e.* an intraresidue main-chain to main-chain interaction (N_i–H···O=C_i) and an interresidue 'newly discovered' side-chain to main-chain interaction (N_i–H···O_i^{γ}Thr), across a polar proteinogenic residue.

The design, synthesis and identification of peptide backbone secondary structural features, stabilised by intramolecularly hydrogen bonding (H-bonding) ring motifs, are of paramount significance since they are targets of *de novo* design, and they also serve as powerful molecular tools for probing the problems of structure–function relationships, protein folding and thermodynamic stabilities.[1] For example, α-bend, β-bend and γ-bend secondary structures, which are frequently observed in proteins and polypeptides and are stabilised by single intramolecular H-bonding interactions, constitute 13-membered (C_{13}-form), 10-membered (C_{10}-form) and 7-membered (C_7-form) ring motifs respectively, and have been the subject of intense conformational investigations over three decades.[2]

The fully extended *intraresidue* H-bonded secondary structure, *i.e.* N_i–H···O=C_i interaction, often referred to as a C_5-structure, constitutes the smallest possible intramolecularly H-bonded conformation across an α-amino acid. Ideally, this five-membered ring is characterised by the backbone torsion angles: $\phi \approx \psi \approx \pm 180$ (±20°) in the Ramachandran map.[3] The conformational characteristics available for this structural motif across proteinogenic residues are extremely rare and poorly understood, presumably due to the non-availability of suitable conformational models. As an attempt to investigate systematically the conformational characteristics of the secondary structure element,[4] we describe here the first unambiguous X-ray diffraction analysis and molecular conformation of this fully extended intramolecularly H-bonded structure, across a chiral proteinogenic residue in a simple model peptide, Boc-Ile-Thr-NH_2 **1**.† In addition to the C_5-interaction, the analysis also revealed the existence of a newly discovered six-membered ring motif stabilised by a main-chain to side-chain H-bond (*i.e.* N_{i+1}–H···O_i^{γ}) interaction, and provided further insight into the folding behaviour of the two ring motifs.[5]

A perspective view of the molecular conformation observed for **1** in the solid state is shown in Fig. 1 and the relevant torsional angles are listed in Table 1. The most surprising feature of the molecular conformation is the observation of an *intramolecular* H-bond between the Thr N2–H and C15=O5 groups. This H-bond results in the formation of a pentagonal ring motif, the 'C_5-structure'. This structure is characterised by the significantly extended backbone torsion angles (ϕ = −163.5°, ψ = +175.6°) and H-bond geometric parameters [bond distance d(N···O) = 2.566 Å, d(H···O) = 2.166 Å, bond angle (N–H···O) = 108.01° and the planarity of the hydrogen bond torsion angle (N_i–H···O=C_i) = 12.4°] (Fig. 2a). The apparent unusual degree of conformational rigidity, imposed by steric effects arising from the exposed amphiphilic C^{β} of the Thr side-chain, appears to be the primary cause of the strong conformational preferences, influencing the location of the Thr side-chain and the secondary-structure-dependent preferred rotation about the C^{α}–C^{β} bond in favour of a C_5-structure.

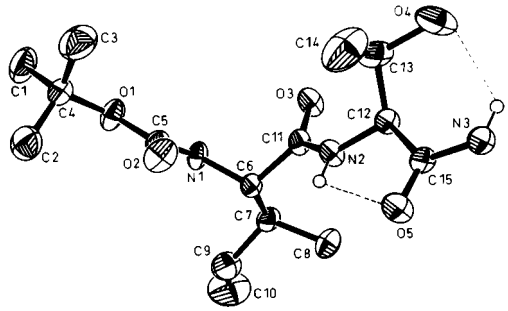

Fig. 1 An ORTEP representation of the molecular structure of Boc-Ile-Thr-NH_2, in the solid state. The thermal ellipsoids are shown at the 40% probability level. Intramolecular H-bonds are indicated by dashed lines.

Table 1 Selected torsion angles (°) for Boc-Ile-Thr-NH_2 **1**

Torsion angle	Ile	Thr
ϕ	−89.7	−163.4
ψ	129.4	175.6
χ_1	170.3,[a] −67.2	−167.8,[b] 69.3
χ_2	163.6	—
ω[c]	179.4	171.3

[a] Torsion angles: N1–C6–C7–C8 = 170.3°; [b] N2–C12–C13–O4 = −167.8°. [c] The C-terminal amide bond, ω_{i+1}, torsion angle C12–C15–N3–H3D = 179.9°.

Another important structural feature of **1** is the influence of the C_5-structural parameters on the unusual energetically unfavourable rotameric distribution of the Thr side-chain.[6] Interestingly, a shift of the Thr C^{α}–C^{β} torsion angle towards a rarely observed *trans* conformation ($\chi_1 \approx -167.8°$) clearly favours the formation of a novel six-membered H-bonded ring motif between the main-chain amide N3–H group and the $O4^{\gamma}$ atom of the Thr side-chain, *i.e.* N_{i+1}–H···O_i^{γ} Thr, interaction. The geometric parameters [d(H···O) = 2.288 Å, d(N···O) = 2.854 Å, bond angle N–H···O = 123.51°, H-bond torsion angle (N_{i+1}–H···O_i^{γ}–C^{β}) = −17.7°] characterise the newly discovered ring motif[5] across a chiral proteinogenic residue (Fig. 2b). In **1**, besides the planarity of the C-terminal amide bond, *i.e.* ω_{i+1}, the set of torsional angles (ψ_i = +175.6° and χ_1 = 167.8°)

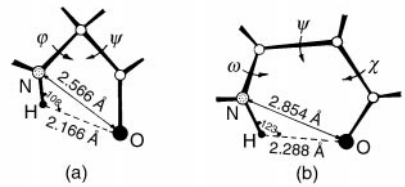

Fig. 2 The geometric characteristics of the two H-bonded ring motifs across the Thr residue: (a) the C_5-structure, N_i–H···O=C_i, main-chain to main-chain interaction; (b) the six-membered ring, N_{i+1}–H···O_i^{γ}Thr, side-chain to main-chain interaction. Relevant torsion angles are marked. Intramolecular H-bonds are indicated by dashed lines. Interacting oxygen and hydrogen atoms are indicated by large and small black circles, respectively.

of the preceding i^{th} residue appear to be the key determinants for the conformational preferences. The topological considerations of the significantly extended torsion angles, $\psi \approx \chi_1 \approx \pm 180$ (±20°), of the Thr residue may not allow the N–H group to approach the O^{γ} atom, precluding the formation of a normal five-membered H-bonding (*i.e.* N_i–H···$O_i{}^{\gamma}$ Thr) interaction. Therefore, an extended backbone ψ torsion angle, in addition to the *trans* χ_1 torsion angle of the i^{th} residue, appears to be the stringent requirement for the juxtaposition of the interacting atoms (N_{i+1}–H···$O_i{}^{\gamma}$) in order to form a six-membered ring across a chiral proteinogenic Thr or Ser residue.[5] In the crystal packing each molecule participates, as a donor or as an acceptor, in the formation of a complex network of intermolecular hydrogen bonding interactions, particularly those involving the Thr and the *trans*-carboxamide NHs. All these interactions probably contribute to the stability of the crystal molecular conformation.

The observation of a rare C_5-interaction across a chiral proteinogenic residue is indeed very interesting since, to date, extensively investigated derivatives and homopeptides incorporating achiral non-proteinogenic $C^{\beta\beta'}$ symmetrically disubstituted Gly residues alone have indicated strong preferences for this structure.[7] Although in these cases the reported average geometric parameters for the C_5-structure are in excellent agreement with those observed for **1**, the conformationally informative τ(N–C^{α}–C′) bond angle of the Thr residue (τ = 106.89°), contrary to our expectation, is not markedly reduced to an average value ≈103°. These results are distinctly inconsistent with the previous results and clearly suggest that the significant narrowing of the τ value may not be mandatory for an effective C_5-interaction, at least across chiral proteinogenic residue(s).

One- and two-dimensional ^{1}H NMR spectroscopic results provided unequivocal evidence, largely consistent with the crystal structure analysis, that in the compact state peptide **1** adopts an intramolecularly H-bonded C_5-structure in solution.[4] However, because of exchange effects (rotation about the C–N bond) the moderately low temperature coefficient ($\Delta\delta/\Delta T \approx -3.66 \times 10^{-3}$ ppm K^{-1}) value obtained for one of the C-terminal primary amide N–H groups was not interpreted as indicative of its involvement in an intramolecular H-bonding interaction stabilising the six-membered ring motif.[4,8] The thermodynamic characteristics of the two stable ring motifs, in a poorly interacting hydrophobic environment, were investigated to gain further insight into their folding–unfolding behaviour. Assuming a two-state process (an equilibrium between intramolecularly H-bonded and non-H-bonded states), the thermodynamic parameters deduced from van't Hoff plots of the variable temperature ^{1}H NMR data indicate that although both the H-bonded ring motifs are enthalpically favoured and entropically disfavoured, among both the ring sizes the more strained five-membered ring is relatively less preferred enthalpically while more favoured entropically (unpublished data).[4,8]

The results of the conformational analysis of **1** indicate that the C_5-structural motif, across an α-amino acid having oxygen atoms(s) in the γ-positions, may facilitate the formation of another six-membered H-bonded ring motif, *i.e.* an N_{i+1}–H···$O_i{}^{\gamma}$ interaction. The variations in the torsional angles ($\phi_i \approx \psi_i$ and $\psi_i \approx \chi_i \approx \omega_{i+1}$) of the five- and six-membered ring motifs, respectively, may characterise the geometric and thermodynamic parameters. It is worth stressing that the combined effect of the two specific ring motifs originating from the chiral moiety, on the molecular conformation appears to be the significant planar arrangement. The marked influence of the main-chain conformational preferences on the side-chain geometry (or *vice versa*) may indicate that χ_1 constrained amino acid derivatives can be exploited as structural tools for *de novo* design of relatively more rigid ring motifs for investigating unexplored structural and functional properties of bioactive molecules. From these observations we are also inclined to suggest that unlike reverse turn structures, the occurrence of such 'flat ring motifs' in proteins and polypeptides may not necessarily be responsible for the chain reversal.

In conclusion, this communication establishes the surprising existence of two fascinating ring motifs, stabilised by intramolecular H-bonding interactions across a chiral proteinogenic residue, and may argue for their intrinsic stabilities. The analysis of conformation-directing effects indicates that unexplored unique local, short range interactions (similar to the Asx-turn motif) might be accessible to small linear peptides incorporating residue(s) bearing a C^{β}-sterogenic centre with short polar side-chain(s). Unambiguous experimental characterisation, in solution as well as in the crystalline state, of the newly described H-bonded topologies are of fundamental importance and may provide an opportunity to examine their occurrence in the highly refined protein data bank; this may further enhance our understanding of polypeptide/protein conformations. Finally, such well defined H-bonded peptide templates/motifs may promise a 'bright future in biochemical and materials science applications.'[9]

This work was supported by a grant from the DBT Government of India. The authors wish to acknowledge the use of the National Diffractometer Facility (DST), New Delhi, for data collection. R. K. is grateful to the Director for his consistent encouragement. This is IMTech communication number 02/99.

Notes and references

† Peptide **1** was synthesised by employing solution phase procedures. A single crystal suitable for an X-ray diffraction study was grown from a methanol solution. The crystal data were collected at 293 K on CAD4 Enraf-Nonius diffractometer with graphite monochromated Cu-Kα radiation (λ = 1.5418 Å) using ω–2θ scan mode. *Crystal data for* **1**: $C_{15}H_{28}N_3O_5$, M = 332.42, orthorhombic, $P2_12_12_1$, colourless needles measuring 1.1 × 0.7 × 0.25 mm^3, a = 8.040(4), b = 12.582(7), c = 18.128(7) Å, $\alpha = \beta = \gamma$ = 90°, V = 1839.7(8) Å^3, Z = 4. The structure was solved by direct methods using SHELXS-97.[10] The structure was refined by full matrix least-squares on F^2 by using SHELXL-97.[11] Final R = 0.0548 [$I > 2\sigma(I)$], R_w = 0.1605. All non hydrogen atoms were refined anisotropically. All hydrogen atoms were placed in idealised positions with assigned isotropic parameters. CCDC 182/1338. See http://www.rsc.org/suppdata/cc/1999/1643/ for crystallographic files.

1 E. N. Baker and R. E. Hubbard, *Prog. Biophys. Mol. Biol.*, 1984, **44**, 97; T. E. Creighton, *Proteins: Structures and Molecular Principles*, 2nd edn., Freeman, New York, 1993; *Peptides: Synthesis Structure and Applications*, ed. B. Gutte, Academic Press, New York, 1995; V. Munõz, P. A. Thompson, J. Hofrichter and W. A. Eaton, *Nature*, 1997, **390**, 196; S. H. Gellman, *Acc. Chem. Res.*, 1998, **31**, 173.

2 C. M. Vankatachalam, *Biopolymers*, 1968, **6**, 1425; *Tetrahedron Symposia-in-Print No. 50*, 1993, **49**, 3433; B. V. V. Prasad and P. Balaram, *CRC Crit. Rev. Biochem.*, 1984, **16**, 307; G. D. Rose, L. M. Gierasch and J. A. Smith, *Adv. Protein Chem.*, 1985, **37**, 1; D. J. Tobias and C. L. Brooks, *Biochemistry*, 1991, **30**, 6059; J. Rizo and L. M. Gierasch, *Annu. Rev. Biochem.*, 1992, **61**, 387.

3 G. N. Ramachandran and V. Sasisekharan, *Adv. Protein Chem.*, 1968, **23**, 283.

4 Ashish and R. Kishore, *Tetrahedron Lett.*, 1997, **38**, 2767; Ashish and R. Kishore, *FEBS Lett.*, 1997, **417**, 97.

5 W. M. Wolf, M. Stasiak, M. T. Leplawy, A. Bianco, F. Formaggio, M. Crisma and C. Toniolo, *J. Am. Chem. Soc.*, 1998, **120**, 11 558.

6 J. Janin, S. Wodak, M. Levitt and B. Maigret, *J. Mol. Biol.*, 1978, **125**, 357; E. Benedetti, G. Morelli, G. Nemethy and H. A. Scheraga, *Int. J. Peptide Protein Res.*, 1983, **22**, 1; J. W. Ponder and F. M. Richards, *J. Mol. Biol.*, 1987, **193**, 775; R. L. Dunbrack Jr. and M. Karplus, *J. Mol. Biol.*, 1993, **230**, 543; P. Chakrabarti and D. Pal. *Protein Eng.*, 1998, **8**, 631.

7 C. Toniolo and E. Benedetti, in *Molecular Conformation and Biological Interactions: G. N. Ramachandran Festschift*, ed. P. Balaram and S. Ramaseshan, Indian Academy of Sciences, Bangalore, India, 1991, p. 511; C. Toniolo, *Janssen Chim. Acta*, 1993, **11**, 10.

8 K. Wüthrich, *NMR in Biological Peptides and Proteins*, North Holland, Amsterdam, 1967; R. Kishore and P. Balaram, *Indian J. Chem.*, 1984, **23B**, 1137.

9 M. Crisma, F. Formaggio, C. Toniolo, T. Yoshikawa and T. Wakamiya, *J. Am. Chem. Soc.*, 1998, **121**, 3272.

10 G. M. Sheldrick, SHELXS-97, A program for crystal structure solution, *Acta Crystallogr., Sect. A*, 1990, **46**, 467.

11 G. M. Sheldrick, SHELXL-97, University of Göttingen, 1997.

Communication 9/04711E