Synthesis and Structure of a New Type of Polycyclic Peptide

Patrick D. Bailey,* David G. W. Clarke and Gavin A. Crofts

Department of Chemistry, University of York, Heslington, York YO1 5DD, UK

The synthesis of a 'polycyclic peptide' is described, in which the ε-nitrogens of cyclo-(Lys-Lys-Gly)$_2$ are interconnected by four *p*-xylyl moieties; this is the first example of the use of multiple cross-linking to generate highly constrained peptides, and NMR spectroscopic evidence indicates the presence of two unusual β-turns that are predicted from molecular modelling studies.

Peptides, polypeptides and proteins are normally classified according to their size (<15 residues = peptide, 15–50 residues = polypeptide, >50 residues = protein). However, one additional characteristic normally distinguishes between them; most proteins readily fold into well defined tertiary structures, polypeptides often show strong conformational features (secondary structure), whilst peptides usually have considerable conformational freedom. In organisms, peptides and proteins often act as 'guest' and 'host', in which the protein (usually a receptor or enzyme) forces the peptide to adopt a conformation suitable for binding; largely because of this, work towards the use of peptides in medicinal chemistry has been dominated by attempts to prepare conformationally constrained peptide mimics.

These observations highlight two parallel problems in protein and peptide chemistry. For proteins, the molecules are too large for *ab initio* structure prediction to be viable at present, chemical synthesis of modified structures is rarely practicable, and the range of modifications accessible *via* DNA technology is somewhat limited. Peptides, on the other hand, are reasonable synthetic targets that can be modified substantially using standard chemistry, but it is difficult to create peptides that adopt only one (or a few) low energy conformations. We are exploring ways of dramatically limiting the conformational freedom of small cyclic peptides by the incorporation of multiple rings *via* cross-linking of the amino acid side-chains. In this paper, we outline the synthesis of a 'polycyclic peptide' that constitutes the first member of this class of compounds.

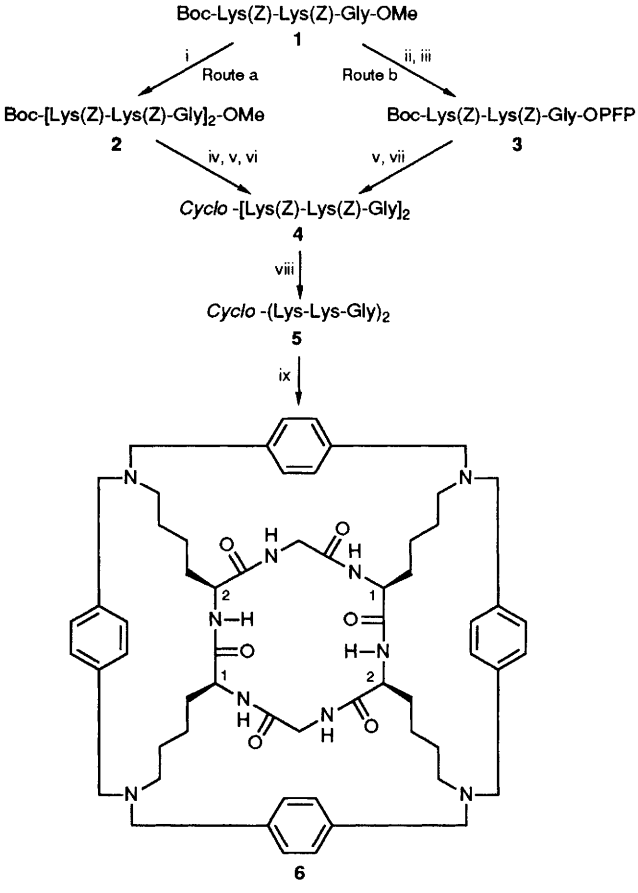
We chose cyclo-(Lys-Lys-Gly)$_2$ as the basic peptide unit, because of the following features: (*i*) the twofold symmetry would allow a convergent synthesis; (*ii*) the glycine residues would allow racemisation-free fragment couplings; (*iii*) the glycine residues might favour two β-turns, aiding cyclisation; and (*iv*) the nucleophilic ε-nitrogens on lysine might facilitate simple, multiple cross-linking reactions.

Standard peptide chemistry gave access to the protected tripeptide **1** as the key building block, and cyclo-[Lys(Z)-Lys(Z)-Gly]$_2$ was initially prepared *via* cyclisation of the linear hexapeptide (Scheme 1, route a); much more efficient was cyclo-dimerisation[1] of H-[Lys(Z)-Lys(Z)-Gly]-OPFP, from which multigram quantities of the cyclic hexapeptide **4** could be obtained (Scheme 1, route b). Removal of the benzyloxycarbonyl protection was efficiently achieved by hydrogenation

over Pearlman's catalyst. With the free cyclic peptide in hand, we started to explore the possibility of multiple intramolecular cross-linking.

The *p*-xylyl cross-linker was selected because the rigid aromatic spacer would preclude cyclisation onto a single ε-nitrogen, whilst also limiting the conformations available to the cross-linked products. Using readily available α,α′-dibromo-*p*-xylene, a series of cross-linking reactions were attempted but, although mass spectral evidence suggested that some of the fully cross-linked product might have been formed, initial work failed to generate sufficient material for isolation and characterisation. Success was finally achieved by the syringe pump addition of cross-linker and then base (Pr^i_2NEt) to a solution of the cyclic peptide **5** in trifluoroethanol at 40–43 °C. Gel filtration (LH20), reversed phase 'flash' chromatography, and recrystallisation from dimethylformamide (DMF) gave the polycyclic peptide **6** as fine white needles [m.p. 212 °C (decomp.)] in 10% yield, for which mass spectral and NMR spectroscopic evidence indicated that the desired cross-links were in place.†

One of our main objectives was that the key structural features of the product could be successfully predicted *ab initio*. Using the BatchMin/MacroModel package of software,[2] a Monte Carlo conformational search was undertaken using the Amber force field.[3] For conformations within 50 kJ mol^{-1} of the global minimum, only two types of structure were generated. In both of them, the cyclic peptide backbone showed (approximate) twofold symmetry, with two unusual β-turns (in which both L-lysyl residues were involved in the loop), and two γ-turns (see Fig. 1). The difference between the two structural types lay in the configuration of the ε-nitrogens (see Fig. 2). For the lower energy structures **6a**, three of the lone pairs were directed inwards, whilst one lone pair was directed outwards. For the higher energy C_2-symmet-



Scheme 1 *Reagents*: i, half of **1** deprotected using KOH–H_2O–MeOH (91% of free acid), and half of **1** deprotected with 90% CF_3CO_2H (aq.) (100% of free amine as $CF_3CO_2^-$salt); these were coupled using DCC–HOBt–DIPEA–DMF (51% yield of **2**); ii, KOH–H_2O–MeOH (91%); iii, PFP-OH–DCC (100%); iv, KOH–H_2O–DMF (80%); v, 90% CF_3CO_2H (aq.) (100% as $CF_3CO_2^-$salt); vi, DPPA (10 equiv.–$NaHCO_3$–DMF (46%); vii, Cs_2CO_3–DMF (42%); viii, H_2–Pd$(OH)_2$ on C–CF_3CH_2OH (100%); ix, *p*-$(BrCH_2)_2C_6H_4$–DIPEA–DMF, 40–43 °C, syringe pump (10%) (DCC = dicyclohexylcarbodiimide; HOBt = 1-hydroxybe zotriazole; DIPEA = diisopropylethylamine; PFP = pentafluorophenyl; DPPA = diphenylphosphoryl azide).

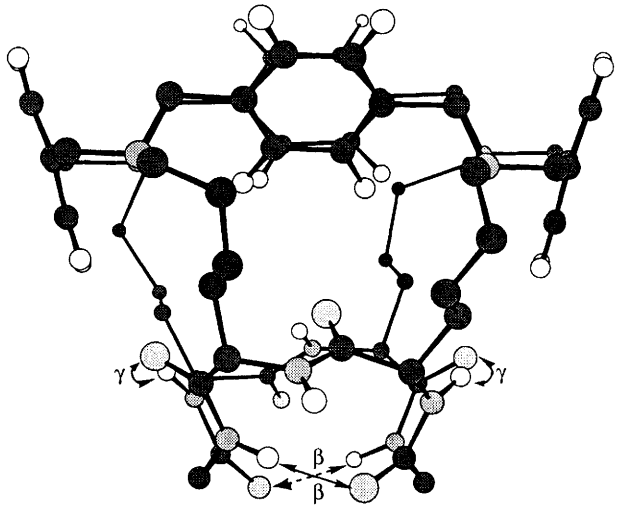


Fig. 1 The lowest energy 'C_2-symmetric' conformation of **6**, from molecular modelling studies. The β- and γ-turns are indicated.

† *Selected data* for **6**: ^{1}H NMR (400 MHz, [$^2H_7$]DMF, 70 °C): δ 1.18–1.48 (8H, m, Lys γ-CH), 1.48–1.68 [8H, m, Lys δ-CH (6H) and Lys^2 β-CH (2H)], 1.82 [4H, m, Lys δ-CH (2H) and Lys^1 β-CH (2H)], 2.02 (2H, m, Lys^1 β-CH), 2.16 [4H, m, Lys^2 ε-CH (2H) and Lys^2 β-CH (2H)], 2.29 [2H, d (br), *J* 11 Hz, Lys^1 ε-CH], 2.49 (2H, m, Lys^2 ε-CH), 2.59 (2H, td, *J* 11 and 3 Hz, Lys^1 ε-CH), 2.75 (2H, d, *J* 12.1 Hz, CH_2Ar), 2.77 (2H, d, *J* 12.1 Hz, CH_2Ar), 2.95 (2H, d, *J* 12.9 Hz, CH_2Ar), 3.01 (2H, d, *J* 12.9 Hz, CH_2Ar), 3.44 (2H, d, *J* 12.3 Hz, CH_2Ar), 3.45 (2H, d, *J* 12.3 Hz, CH_2Ar), 3.49 (2H, dd, *J* 14.2 and 6.4 Hz, Gly α-CH), 4.06 (2H, d, *J* 12.9 Hz, CH_2Ar), 4.100 (2H, m, Lys^1 α-CH), 4.103 (2H, d, *J* 12.9 Hz, CH_2Ar), 4.19 (2H, dd, *J* 14.4 and 4.3 Hz, Gly α-CH), 4.39 (2H, m, Lys^2 α-CH), 7.26 (8H, ABq, *J* 7 Hz, Ar), 7.49 [2H, s (br), Lys^2 NH], 7.63 [2H, s (br), Lys^1 NH], 7.69 (8H, ABq, *J* 7 Hz, Ar), 8.75 [2H, s (br), Gly NH]. The large number of vicinal protons that fail to show coupling is indicative of the conformational constraints imposed on **6**. The different lysyl residues have been labelled with superscripts (Lys^1 and Lys^2); see structure 6 in Scheme 1. ^{1}H–^{1}H COSY and NOESY (400 MHz, [$^2H_7$]DMF, 60 °C) allowed all of the protons to be assigned (except for the exact location of some of the overlapping Lys γ- and δ-protons). Two isomers of **6** are sterically possible, in which pairs of cross-linkers span adjacent or skipped lysyl residues, but NOEs between the aryl and β-lysyl/NH protons would not occur for these compounds (*cf.* Fig. 3). The rectangular nature of the tetraaza ring is characterised by the large difference in the chemical shift for the two sets of (isolated) aromatic protons, and by the presence of 12 (not 16) NOEs between the aryl and benzylic protons; these results are also incompatible with isomers of **6**.

^{13}C NMR (75 MHz, [$^2H_7$]DMF, 60 °C): δ 22.32 (CH_2), 22.55 (CH_2), 26.58 (CH_2), 27.05 (CH_2), 31.57 (CH_2), 31.72 (CH_2), 45.94 (CH_2), 50.61 (CH_2), 51.69 (CH_2), 55.84 (CH), 56.49 (CH), 56.72 (CH_2), 59.30 (CH_2), 59.40 (CH_2), 59.71 (CH_2), *ca.* 129.6 (unresolved CH's), *ca.* 139.9 (unresolved C's), 170.17 (C), 171.360 (C), 173.52 (C).

MS *m/z* [+ve ion FAB (fast atom bombardment), 2-nitrobenzyl alcohol matrix] 1057.6 $[M + Na]^+$ and 1035.6 $[M + H]^+$ (M^+ for $C_{60}H_{78}N_{10}O_6$ calc. as 1034.6); in support of the polycyclic nature of **6**, no other peaks (except for matrix) were observed above those at 104/91/77/65, which were assigned to double fragmentation of the cross-linker.

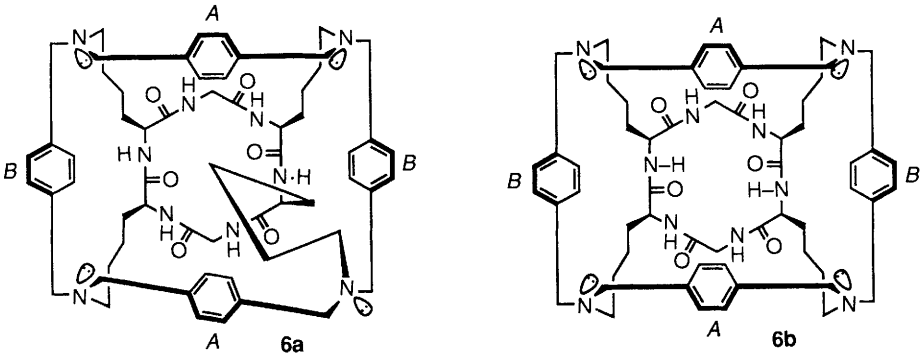


Fig. 2 The two types of low energy conformation for **6** differ in the configuration of the ε-nitrogens (conformation of peptide backbone not depicted here), and molecular modelling studies also predict that the tetraaza-paracyclophane ring should be roughly rectangular, as shown

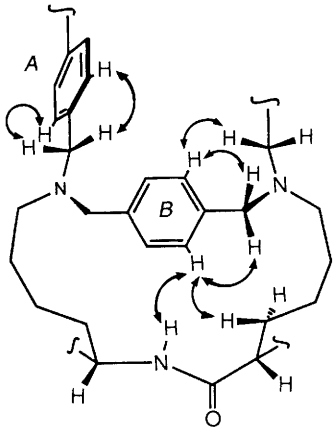


Fig. 3 The main NOE enhancements for the aromatic protons of **6**. Only two protons for each aryl ring are shown for clarity. The δ 7.70 protons were assigned to ring *B*, and δ 7.26 to ring *A*, on the basis of chemical shift and NOE to the Lys^2(NH).

ric structures **6b**, all four N$^\varepsilon$-lone pairs pointed inwards. The lower energy for **6a** presumably results from allowing one of the lysine side-chains to fill the otherwise empty central cavity of the molecule.

To our delight, extensive COSY (correlation spectroscopy) and NOESY (nuclear Overhauser effects spectroscopy) experiments were in full agreement with these structural predictions. At room temperature, the 400 MHz ^{1}H NMR spectrum was quite broad, but at 60–70 °C the C_2-symmetry was apparent on the NMR time scale. Critical NOEs between Gly(NH)-Lys^2(αCH) and between Gly(αCH)-Lys^1(NH) were consistent with the predicted (but unusual) β- and γ-turns. Moreover, two distinct types of aromatic proton were apparent (AB quartets at δ 7.26 and 7.70), and their observed NOEs (indicated on Fig. 3) are those expected for the 'rectangular' tetraaza-macrocycle depicted in structures **6a** and **b** [*i.e.* one pair of phenyl rings (*e.g.* A) are much closer together than the other pair (*e.g.* B)]. It thus appears that the different forms of structure **6a** are rapidly interconverting at 60–70 °C, presumably *via* **6b**.

When this work was started, it was not clear that polycyclic peptides were viable synthetic targets. Our synthesis of **6** has now shown that it is indeed possible to introduce multiple cross-links between adjacent and skipped residues in a cyclic peptide. Moreover, the correlation between the calculated and observed conformational features for **6** suggests that tailor-made polycyclic peptides of predictable 3D structure are accessible. This could be of particular relevance to highly constrained analogues of medicinally important peptides, and to the design of novel peptidic host molecules. For example, **6b** possesses a cavity that should be large enough to accommodate small protected amino acids‡ or, by analogy with the cubic aza-*para*-cyclophane 'kyuphane',[4] substituted benzenes. Studies are currently underway to find guests that might lock **6** in the open conformation **6b**,‡ and to prepare other polycyclic peptides for which the predicted global minimum possesses an open cavity.

P. D. B. wishes to thank Professor W. C. Still for access to his research facilities at Columbia University, and for helpful advice and discussions, particularly concerning the use of MacroModel; the advice and help of Dr J. Simon with computational and host–guest studies, and of Dr C. Turner with NMR studies, is gratefully acknowledged. We also thank Dr G. K. Barlow, Mrs B. Chamberlain, Dr T. A. Dransfield, Mr B. R. Glennie and Dr N. R. McLay for NMR and mass spectra at York, Vinka Parmakovich for mass spectra at Columbia University, the SERC for studentships (to D. G. W. C. and G. A. C.), The Nuffield Foundation and The Royal Society for science travel awards (to P. D. B.), and the Yorkshire Cancer Research Campaign for a career development award (to P. D. B.).

Received, 9th January 1992; Com. 2/00102K

References

1 K. D. Kopple, *J. Pharm. Sci.*, 1972, **61**, 1345.
2 F. Mohamadi, N. G. J. Richards, W. C. Guida, R. Liskamp, M. Lipton, C. Caufield, G. Chang, T. Hendrickson and W. C. Still, *J. Comput. Chem.*, 1990, **11**, 440.
3 G. Chang, W. C. Guida and W. C. Still, *J. Am. Chem. Soc.*, 1989, **111**, 4379.
4 Y. Murakami, J. Kikuchi, T. Ohno, T. Hirayama, Y. Hisaeda, H. Nishimura, J. P. Snyder and K. Steliou, *J. Am. Chem. Soc.*, 1991, **113**, 8229.

‡ A preliminary binding study (NMR titration in $CDCl_3$) of **6** with acetyl-L-alanine indicated the formation of a 1:1 complex with a binding energy of about 12 kJ mol^{-1}; no such complexation was apparent with acetyl-D-alanine, although the titrations are hampered by the low solubility of **6** in suitable solvents, the broadness of the ^{1}H NMR spectrum at room temperature, and by questions concerning the degree of protonation of **6**.