ELSEVIER

Journal of Molecular Graphics and Modelling 21 (2003) 341–355

Journal of Molecular Graphics and Modelling

www.elsevier.com/locate/JMGM

Designing supramolecular structures from models of cyclic peptide scaffolds with heterocyclic constraints

A.J. Lucke, J.D.A. Tyndall, Y. Singh, D.P. Fairlie*

Centre for Drug Design and Development, Institute for Molecular Bioscience, University of Queensland, Brisbane, Qld 4072, Australia

Abstract

Cyclic peptides containing oxazole and thiazole heterocycles have been examined for their capacity to be used as scaffolds in larger, more complex, protein-like structures. Both the macrocyclic scaffolds and the supramolecular structures derived therefrom have been visualised by molecular modelling techniques. These molecules are too symmetrical to examine structurally by NMR spectroscopy. The cyclic hexapeptide ($[Aaa–Thz]_3$, $[Aaa–Oxz]_3$) and cyclic octapeptide ($[Aaa–Thz]_4$, $[Aaa–Oxz]_4$) analogues are composed of dipeptide surrogates (Aaa: amino acid, Thz: thiazole, Oxz: oxazole) derived from intramolecular condensation of cysteine or serine/threonine side chains in dipeptides like Aaa–Cys, Aaa–Ser and Aaa–Thr. The five-membered heterocyclic rings, like thiazole, oxazole and reduced analogues like thiazoline, thiazolidine and oxazoline have profound influences on the structures and bioactivities of cyclic peptides derived therefrom. This work suggests that such constrained cyclic peptides can be used as scaffolds to create a range of novel protein-like supramolecular structures (e.g. cylinders, troughs, cones, multi-loop structures, helix bundles) that are comparable in size, shape and composition to bioactive surfaces of proteins. They may therefore represent interesting starting points for the design of novel artificial proteins and artificial enzymes.

Keywords: Macrocycle; Cyclic peptide; Thiazole; Oxazole; Peptidomimetic; Protein; Peptide; Artificial enzymes

1. Introduction

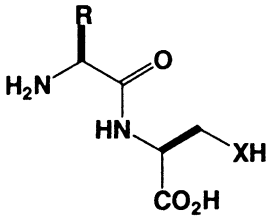
Nature commonly uses hierarchical systems to build complex structures from simple constraints [1]. For example, simple amino acids are “information rich” building blocks that combine to form peptides featuring *trans*-amide bond constraints. These interact with one another through constraints like hydrogen bonds and side chain packing to create uniquely folded helical/turn/loop/sheet secondary structures, which in turn fold into unique three-dimensional tertiary structures that constitute proteins. We now describe a novel hierarchical system for the construction of complex “protein-like” unnatural structures, beginning with heterocyclic five-membered ring constraints incorporated by Nature into amino acids and then into cyclic peptides [2–11]. Theoretically these constrained cyclic peptides could then, in turn, be used as building blocks for elaboration to larger and more complex unnatural structures that resemble proteins.

Intramolecular condensation of cysteine or serine/threonine side chains in dipeptides like Aaa–Cys, Aaa–Ser and Aaa–Thr (**1**) results in dipeptide surrogates (**2**) incorporating five-membered heterocyclic rings like thiazole (X=S), oxazole (X=O) and reduced analogues (thiazoline, thiazolidine, oxazoline). Nature frequently uses heterocycles as structural restraints, particularly in cyclic peptides (e.g. **3**, **4**), where they induce β-turns that dictate structure and biological activities [11]. Cyclisation brings an additional level of restraint to peptide structures, and also confers hydrolytic and proteolytic stability required for pharmaceutical applications [11,12]. Numerous cyclic peptides are now known to be potent enzyme inhibitors and receptor antagonists, with a wide range of potential therapeutic applications [12–16].

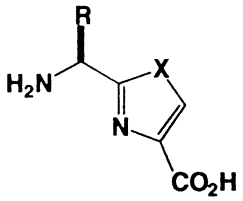
Families of macrocycles like **3** and **4**, recently synthesised via cyclo-oligomerisation [17–20] of units like **2**, have very high symmetry that limits structure determination in solution by NMR spectroscopy. Also a surprising number of such molecules do not produce crystals that diffract well enough to determine high resolution X-ray crystal structures. We therefore examine their putative structures here using molecular modelling techniques and compare the derived cyclic peptide structures with those examples for which X-ray crystal or NMR solution structures are available. We also attempt to use modelling to predict and visualise the sizes, shapes and utility of a range of supramolecular

* Corresponding author. Fax: +61-7-33651990.
E-mail address: d.fairlie@imb.uq.edu.au (D.P. Fairlie).

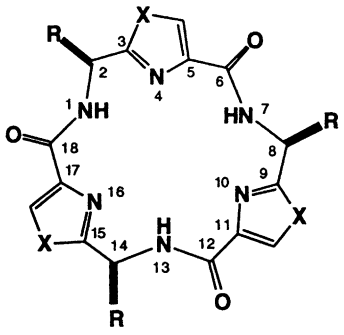
1093-3263/02/$ – see front matter
PII: S1093-3263(02)00181-X



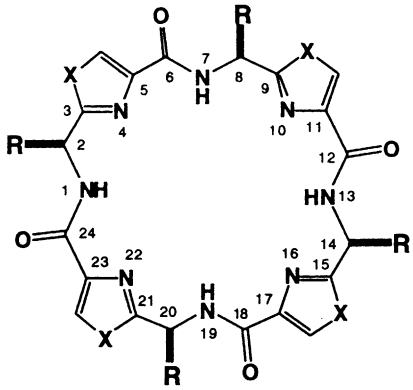
1



2



3



4

3a X = S, R = Me; **3b** X = O, R = Me

structures that can theoretically be built from such conformationally constrained cyclic peptide templates.

2. Methodology

Simulations were performed with InsightII software [21] running on a Silicon Graphics R10000 Octane workstation. InsightII modules used were Builder, Discover, Discover 3, Analysis and Search and Compare. All molecules were initially built and optimised to an acceptable conformation within the Builder module. These structures were then used as initial starting points for further energy minimisations and molecular dynamics simulations. An initial simulation using a quenched molecular dynamics protocol using explicit solvent with periodic boundary conditions was found to take about 2 weeks, so instead all simulations were performed in vacuo.

2.1. Quenched molecular dynamics

Cyclic peptide analogues **3** and **4** along with the cone **5** and cylindrical molecules **6** and **7** were subjected to unrestrained quenched molecular dynamics simulations using the CFF91 force field in Discover. These were performed using a protocol with a dielectric constant of 1.0 and a temperature of 1000 K unless specified otherwise in the text. Initially the molecules were subjected to minimisation using steepest decent (100 iterations) and VA09A algorithms (5000 iterations) until the maximum derivative was less than 1.0 kcal/Å. The molecule was then subjected to molecular dynamics, undergoing heating and equilibration for 5000 fs using steps of 1 fs. Conformations of the resulting structures were then saved every 100 fs for 1000 iterations. These conformations were minimised using steepest descent (500 iterations or a derivative <10 kcal/Å) and VA09A algorithms (5000 iterations) until the maximum derivative was less than 0.001 kcal/Å. The minimised structures were analysed to determine low energy conformers using the Analysis and Search and Compare modules. We have used two measures to examine simulation convergence. Firstly, the $D_{\min}$ value [22] is defined as the smallest number of times a conformer is found (we have used a cut-off within 3 kcal/mol of lowest energy conformer), values greater than 5 give a reasonable indication that a search is complete. Secondly, a simple measure of the probability that a search is complete is given by the equation $1-(1-(1/N))^{M}$, where N is the total number of conformers and M is the number of search steps [23]. A value close to 1 indicates good probability that the search is complete.

Comparisons between trough **8**, loop structure **9**, helix bundle **10** and the corresponding protein surfaces in Figs. **8**, **9** and **10**, respectively, were made by positioning the simple energy minimised modelled structures onto the respective structures of the active site of a protease from dengue virus, [24] an antibody, [25] and the four-helix bundle of leukaemia inhibitory factor [26]. Template forcing was used

to model the loop structure **9** using the complementarity determining region (CDR) of the NC41 antibody crystal structure as a template. Template forcing minimises the RMSD of a specified number of carbon and nitrogen atoms of the peptide backbone constituting the template, using a specific minimisation algorithm that allows rotation around single bonds as well as translation and rotation in space [27].

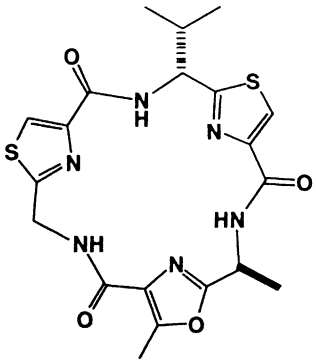
2.2. Energy minimisation

All structures examined were treated in the same manner as described here for the cyclic hexapeptide analogue **3**. Initially, a single heterocycle dipeptide analogue **2** (R = methyl) was constructed in the Builder module and its structure optimised using default parameters. This dipeptide analogue was then copied twice and the resulting three units combined through *trans*-amide coupling, producing the macrocycle **3** with all R groups on the same face of the ring as would be expected from a coupling reaction. Force field CFF91 was selected and the potentials, partial charges and formal charges of the model fixed. Discover 3 was then used to perform energy minimisation simulations using default parameters and gradient methods. The energy minimisation was stopped once the final convergence value reached 0.001.

The only exception to this procedure was for the cone **5** that contained a Zn^{2+} ion co-ordinated to the TACN, in which the ESFF force field was used in the energy minimisation.

3. Results

3.1. Cyclic hexapeptide analogues



Nostocyclamide

Nostocyclamide (X-ray crystal structure shown in Fig. 1B) is a naturally occurring macrocycle with anti-cyanobacterial and anti-algal activity [28]. It is an 18-membered ring containing one oxazole and two thiazole heterocycles linked together by amide bonds. This ring is planar resulting from the large barrier to rotation of the *trans*-amide bonds, the sterically favourable orientation of the aromatic imine nitrogens towards the centre of the macrocycle, and the subsequent secondary interactions of the intramolecular array of alternating hydrogen bond donors and acceptors [29,30]. This makes it an excellent reference molecule for comparison with synthetic analogues. To validate the protocols used in this research, quenched molecular dynamics were carried out upon the crystal structure of nostocyclamide. The resulting low energy conformations (40% of the total calculated) gave RMSD values of 0.16 Å over 18 pairs of macrocyclic atoms and closely resembles the X-ray crystal structure.

Quenched molecular dynamics simulations were performed on unrestrained models of the cyclic hexapeptide analogue, **3**, heated to 1000 K with a series of 1000 frames collected over time. Individual frames were then minimised and analysed to identify low energy conformers. The simulation of **3a** (D_{min} 830, probability value 1) resulted in several similar low energy conformations with the lowest energy conformation shown in Fig. 1A. It is highly symmetric with a C_3 fold axis of symmetry. The macrocyclic structure is essentially flat with each heterocycle lying in the same plane as the macrocycle. Side chain methyl groups of the linking amino acid residues all project from the same face of the macrocycle (*syn*, *syn*, *syn*). All amide carbonyl oxygens were found to point outside the ring with corresponding *trans*-amide nitrogen–hydrogen bonds pointing into the ring as seen in the X-ray structure of nostocyclamide. This structure was observed for 83% of the total conformations resulting from the simulation, the next lowest energy conformer being 5.7 kcal higher in total energy. The corresponding simulation of **3b** showed similar statistics with the lowest energy conformation seen 57% of the time and the next conformation (27%) being 4.4 kcal higher in energy. The lowest energy conformation found is also relatively planar and possesses the alternating imine nitrogen:amide NH functionality. This very high proportion of one single conformer confirms that molecule **3** is highly constrained, due in part to the planar region formed by the heterocycles conjugated to amide bonds. Higher energy conformations, if not planar, indicate rotation around two of the sp^3 αC atoms thus tilting the constrained aromatic ring out of the plane and increasing steric clashes with substituents on the αC atoms as well as disrupting the secondary hydrogen bonding interactions.

The average dihedral angles of the low energy conformations of the linking amide group NHαCH were 139 and 137° for **3a** and **3b**, respectively which agrees well with those seen in the L-amino acid derivatives of nostocyclamide (127 and 131° for the alanine and glycine analogues, respectively) [28]. Macrocycle distances were measured between the atom pairs in **3a**: N_4–N_{10} 4.68, N_1–N_7 5.10, C6–C12 6.60 and C2–C8 6.67 Å. The hole created by the macrocycle is too small to fit even a small guest molecule like methane without significant VDW overlap. In fact a Connolly surface created using a 1.4 Å probe indicated that the centre of the macrocycle was a continuous surface with no hole. Thus, the

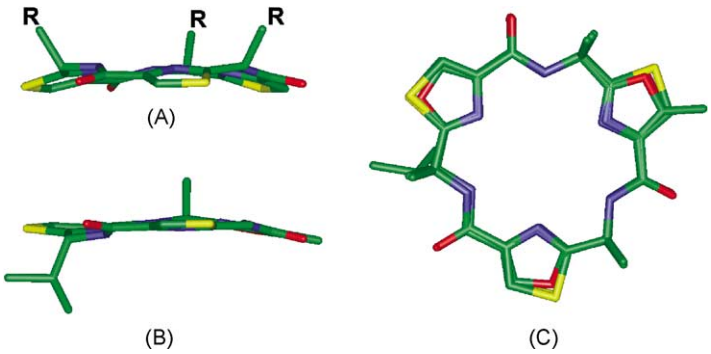


Fig. 1. Quenched molecular dynamics simulation at 1000 K, heavy atoms only displayed. (A) Lowest unrestrained energy minimised conformer of cyclic hexapeptide analogue **3** viewed parallel to the plane of the macrocycle, (B) comparative view of nostocyclamide crystal structure, (C) superimposition of lowest unrestricted energy minimised conformations of **3** (X=O and S) and nostocyclamide viewed perpendicular to plane of macrocycles.

cycle can be considered as a "filled" planar building block with side chain functionality projecting from the same face.

A comparison was made between the modelled tris-oxazole and tris-thiazole macrocycles with the crystal structure of nostocyclamide [28] by superimposing **3a** and nostocyclamide (Fig. 1C), for which there was an RMSD of 0.33 Å over the 27 non-hydrogen atoms not including R substituents. Clearly, in this case modelling has predicted a structure for **3** that very closely mimics a natural cyclic peptide with three heterocyclic constraints. Conformer **3a** is a relatively rigid and planar scaffold with three functional groups projecting from the same face of the scaffold. This suggests the possibility of using various such cyclic templates with predetermined tethering points for engineering larger, more elaborate molecules with designed shapes and functions.

3.2. Cyclic octapeptide analogues

Ascidiacyclamide [31–33] is a naturally occurring 24-membered ring isolated from sea squirts, also known as tunicates or ascidians. This macrocycle contains two planar thiazoles and two oxazolines linked by amide bonds. This

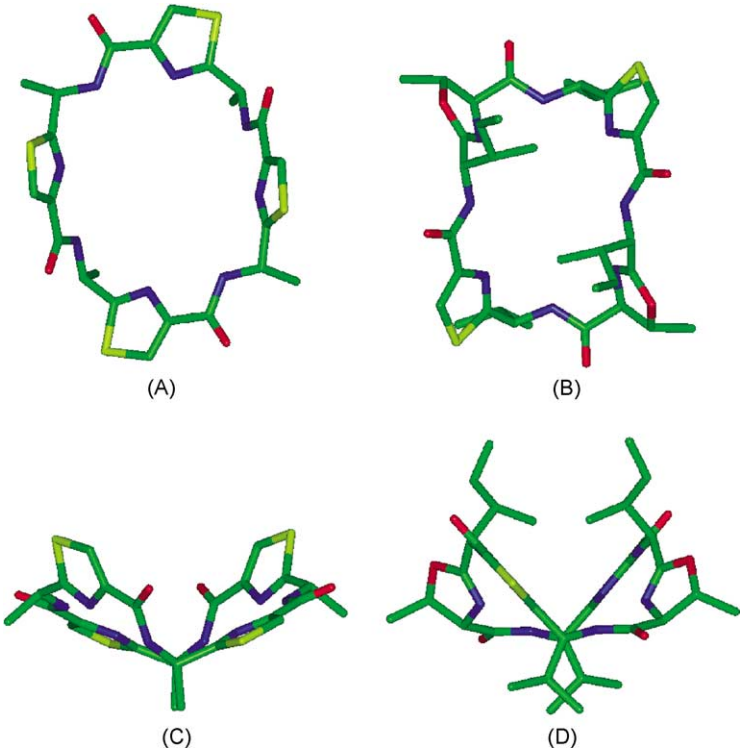
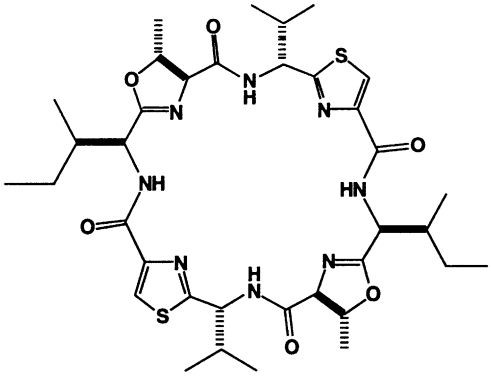


Fig. 2. Molecular dynamics simulation quenched at 1000 K, heavy atoms only displayed. (A) Lowest unrestrained energy minimised conformer of cyclic octapeptide analogue **4** viewed from above and the side (C) of the macrocycle respectively, (B) and (D) are the comparative views of the crystal structure of ascidiacyclamide [31].

makes ascidiacyclamide an excellent template molecule to use as a reference point from which to build the cyclic octapeptide analogue **4**.



Ascidiacyclamide

A model of **4** containing four oxazole heterocycles was constructed so the macrocycle was square in shape and all amide NH protons pointed to the interior of the ring similar to those of ascidiacyclamide and nostocyclamide. Unrestricted quenched molecular dynamic simulations of **4** were then performed to explore available conformational space producing 1000 structures (D_{min} **4**, probability value 0.9993). This produced 140 unique conformations, the lowest energy conformation was found 165 times and is shown in Fig. 2A.

The lowest energy conformation found (Fig. 2A) has a symmetrical rectangular shape with a C_2 fold axis of symmetry. This 24-membered ring was not flat but formed a slightly twisted or saddle-shaped fold. All of the side chain R substituents from the amino acid residues that separate the five-membered ring heterocycles projected from the same face of the macrocycle. All amide carbonyl oxygens were found to point outside the ring with corresponding amide hydrogens pointing into the ring. The minimised conformations of this simulation show that the atoms between the αC's line in the plane. This translates into essentially four rotation points around each αC giving a less constrained scaffold than **3**. The lowest energy minimised structure found (Fig. 2A) contained two distinct NHαCH dihedral angles −126 and 138°). These were considered to be comparable to the reported experimental value (153°) calculated from $^3J_{NH\alpha CH}$ derived from NMR spectra for an analogue of **4** in DMSO [20]. Dimensions of the "square-like" cyclic octapeptide analogue shown in Fig. 2A were measured between the following opposing atoms, N_1–N_{13} 7.90, $C\alpha_2$–$C\alpha_{14}$ 9.27, N_4–N_{16} 6.89, C_6–C_{18} 8.08, N_7–N_{19} 7.02, $C\alpha_8$–Ca_{20} 8.39, N_{10}–N_{22} 7.44 and C_{12}–C_{24} 9.82 Å.

Simple energy minimisation of the same starting conformer used for the quenched molecular dynamics described above produced a structure similar to that shown in Fig. 2A. Superimposition of the energy minimised structure with the lowest energy conformation from quenched molecular dynamics resulted in an RMSD value of 0.53 Å over all 80 non-hydrogen atoms. This RMSD value is higher than for **3** as a result of the larger size and greater flexibility of **4**. Total energies for conformers determined by energy minimised, and lowest energy quenched molecular dynamics methods were 142.49 and 143.21 kcal/mol, respectively.

Ascidiacyclamide [31–33] has a crystal structure slightly rectangular in shape with distances between opposing αC positions of 8.51 and 6.32 Å. The modelled structure of the cyclic octapeptide analogue **4** (Fig. 2A), containing thiazole heterocycles, was compared with the crystal structure of ascidiacyclamide (Fig. 2B) by molecular superimposition (Fig. 2C). There was good agreement between the structures with an RMSD of 1.1 Å over 24 non-hydrogen atoms of the macrocyclic ring. This indicates a high degree of similarity between the structures of the two molecules, despite greater flexibility in the two oxazoline rings of ascidiacyclamide versus the two more rigid planar oxazole heterocycles in **4**.

Molecular dynamic simulations of Patellamide B and C [34] which are structurally analogous to **4** and ascidiacyclamide, though not symmetrical, have shown that they form "figure eight-like" conformations. However, symmetrical compounds such as ascidiacyclamide, form "square-like" conformations as our results have also shown.

Compared to the hexapeptide analogue **3**, the larger macrocycle **4** was observed to be more flexible during the quenched molecular dynamics simulations and gave a larger number of possible energy minimised structures. To highlight this, the next four lowest energy conformers found during the quenched molecular dynamics simulation are shown in Fig. 3 for which A, B, C and D increase progressively in total energy. These conformers were within 2.5 kcal/mol of the lowest energy conformer (Fig. 2A). It is important to note that in each of these structures the side chain R substituents all remained on the same face of the macrocycle, suggesting the possibility of regiofacial functionalisation.

A Connolly surface of the conformer in Fig. 2A was generated and indicated that the centre of the macrocycle was not filled, leaving a hole for potential capture of a guest molecule. This contrasts with **3** in which the Connolly surface filled the macrocycle. The size and flexibility of the cavity formed by **4** was investigated by simple energy minimisation of host–guest assemblies using a variety of simple guest molecules. It was found that guests such as benzene and anthracene could be inserted perpendicularly into the hole, or like while glucose fitted into the saddle-shaped cavity of **4** similar to that observed in the crystal structure of ascidiacyclamide [32] containing a guest benzene molecule.

The presence and nature of the four heterocycles is crucial to the three-dimensional structure of the cyclic molecules **4** and ascidiacyclamide. For the latter, we have previously found experimentally that replacement of the two oxazoline rings with two amino acids allows the macrocycle to adopt a chair-like conformation [35]. Further replacement of the two thiazoles by common amino acids generates an extremely flexible macrocycle [35]. These differently

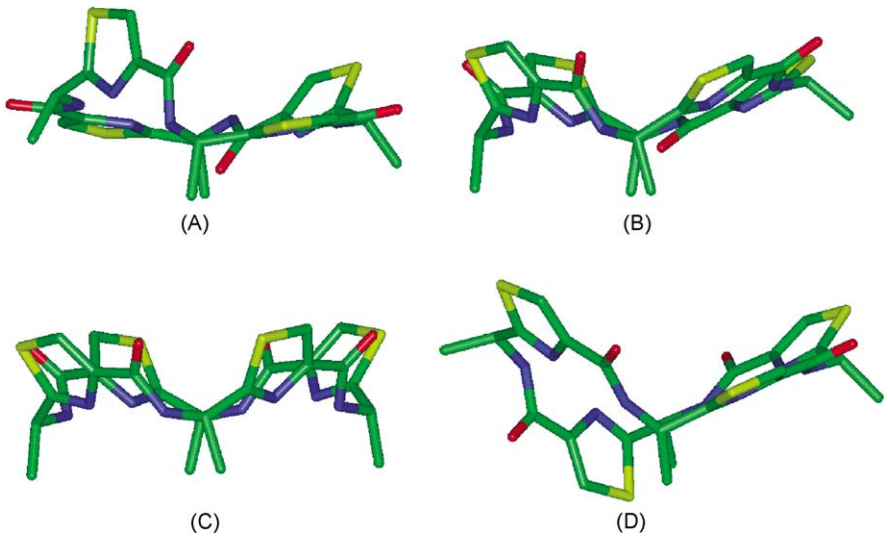


Fig. 3. Molecular dynamics simulation quenched at 1000 K, heavy atoms only displayed. (A, B, C and D) side views of the next four lowest unrestrained energy minimised conformers, respectively for cyclic octapeptide analogue **4**.

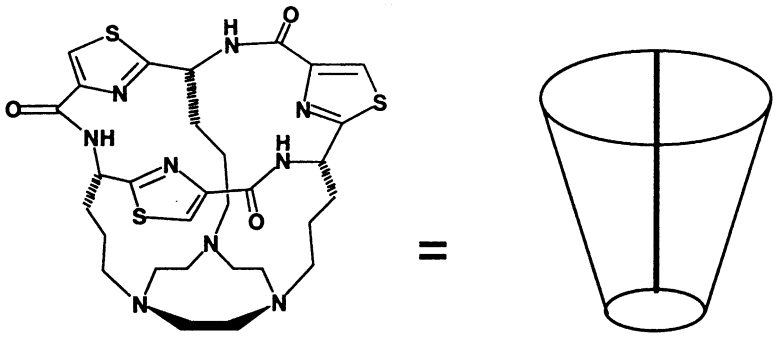
constrained macrocycles have very different affinities for various metal ions [35–40].

3.3. Cones derived from cyclic peptides

Towards the goal of using constrained macrocyclic peptides to design novel, synthesisable, protein-like macromolecules with defined three-dimensional shapes and sizes, we have modelled a simple cone, **5**, composed of cyclic hexapeptide analogue **3** linked by three trimethylene units to 1,4,7-tetraazacyclononane (TACN). TACN can be considered as a reduced analogue of a cyclic tripeptide with a rigid structure and all nitrogen lone pairs pointing into the ring due to the high thermal barrier to inversion of configuration about the nitrogen atoms [41].

The average RMSD for these structures superimposed over the lowest energy conformer was found to be 0.44 Å over all 90 non-hydrogen atoms, indicating a highly constrained molecule. This rigidity is consistent with a molecule containing **5** macrocycles and **3** aromatic heterocycles.

The shape and orientation of the capping tris-thiazole macrocycle **3** did not change significantly when incorporated into **5**. The dihedral angle (NHαCH 149°) of the linking amide group matched more closely the value (160°) derived from the measured coupling constant ($^3J_{HH\alpha CH}$) from NMR spectra for DMSO solutions [20]. The TACN ring at the base of the cone did not undergo any inversion of configuration at the amine nitrogens during modelling, no doubt due to the large energy barrier to inversion [41]. The size and shape of the conical structure **5** is defined by a



5

Simulation of the tris-thiazole cone, **5**, was performed using an unrestrained quenched molecular dynamics simulation as described above. Superimposition and analysis was performed using the 10 lowest energy conformers (Figs. 3 and 4, $\Delta E_{total} = 1.7$ kcal/mol) which constituted 25% of the total number of frames (D_{min} 9, probability value 0.9617). diameter for the tris-thiazole cycle of 5.6 Å, a diameter for the TACN ring of ∼4 Å, and a length for the cone of ∼5.1 Å.

A simple energy minimisation of the same initial starting conformer used for the quenched molecular dynamics simulation produced a structure that was nearly identical to the lowest energy conformer determined by quenched

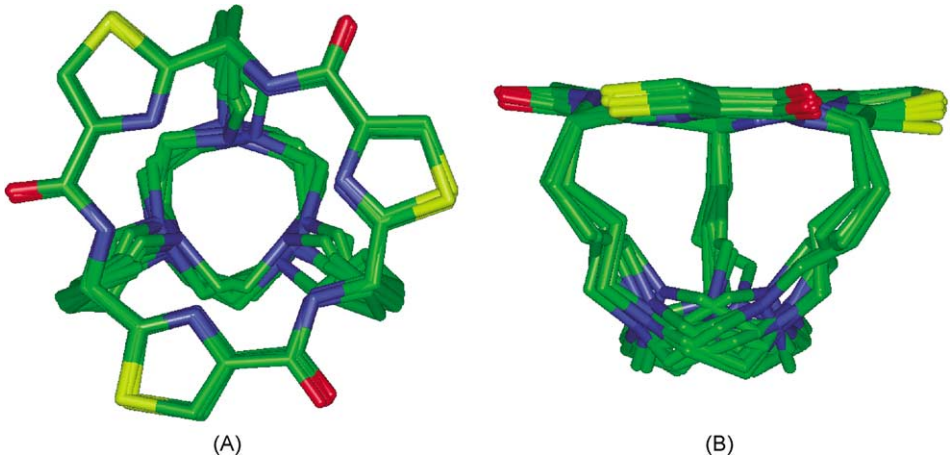


Fig. 4. Molecular dynamics simulation quenched at 1000 K, heavy atoms only displayed. Superimposition of the 10 lowest unrestrained energy minimised conformations of tris-thiazole cone **5** viewed from above (A) and side on (B).

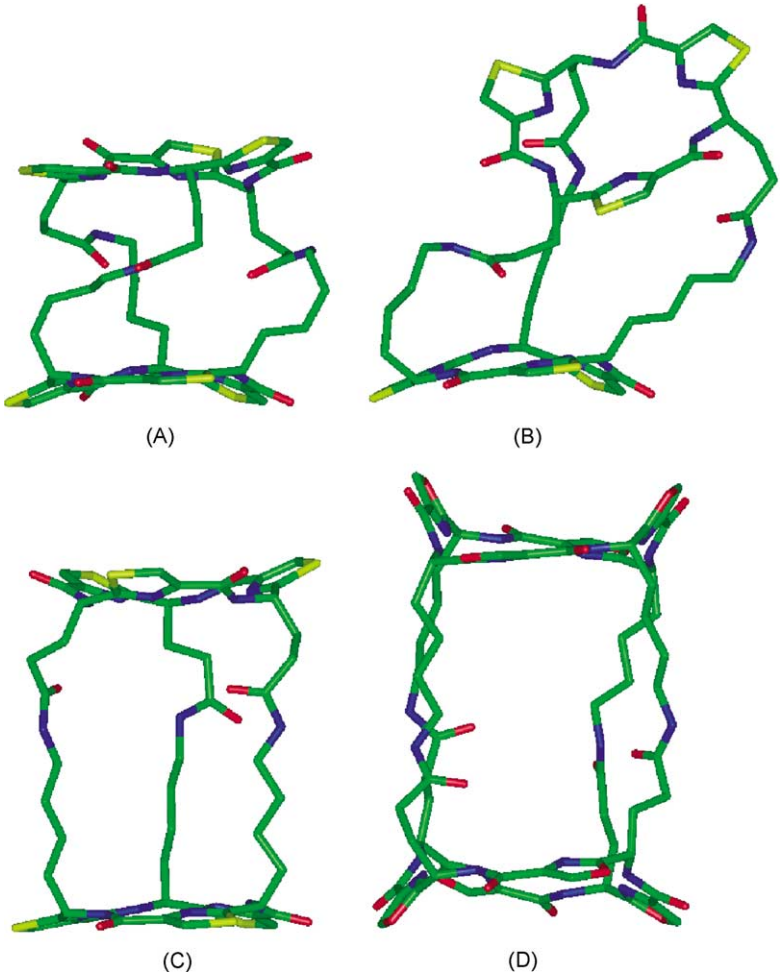


Fig. 5. Molecular dynamics simulation quenched at 1000 K, heavy atoms only displayed. (A) lowest unrestrained energy minimised conformation of cylindrical analogue **6**. (B) lowest unrestrained energy minimised conformation of **6** from quenched molecular dynamics at 300 K. (C) minimised conformation of cylindrical analogue **6**, (D) minimised conformation of cylindrical analogue **7**.

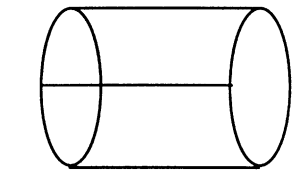
molecular dynamics. Superimposition of the simple energy minimised structure with the lowest energy conformation from quenched molecular dynamics resulted in an RMSD value of 0.05 Å over all 90 heavy atoms. This indicated very good agreement between energy minimisation and quenched molecular dynamics of cone **5**. Total energies for final conformers determined by energy minimised and lowest energy quenched molecular dynamics methods were 180.7 and 159.9 kcal/mol, respectively.

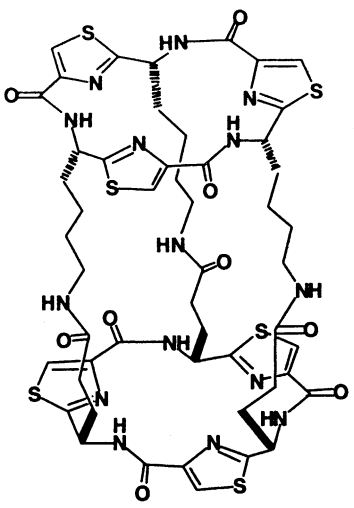
The size and shape of the lowest unrestrained energy minimised conformation of the cone created a small interior cavity. To ascertain whether this was large enough to potentially capture small metal ions, a simple energy minimisation was performed by including a Zn^{2+} ion bound to the TACN unit within the cone. This resulted in a structure in which the planarity of the peptide macrocycle amide bonds was heavily distorted causing the N–H bonds to point up from the plane of the ring. This suggests that there was too much strain to permit metal capture, however, this strain could be relieved by extending the polymethylene chains by two or more methylene units. The cavity within the cone **5** was too small to accept even a small organic guest like methane.

3.4. Cylinders derived from cyclic peptides

Modelled cylinders (**6**, **7**) were built from two tris-thiazole macrocycles **3** or two tetraoxazole macrocycles **4** by condensing side chains of lysine and glutamate amino acids from different macrocycles. In this way macrocycles like **3** or **4** could be linked through 3 or 4 simple *trans*-amide bonds, respectively to create cylindrical molecular shapes. The cylinders were constructed in this manner for ease of potential synthesis.

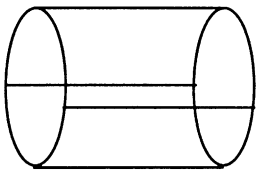
Putative shape of cylinder **6** (ex **3**)

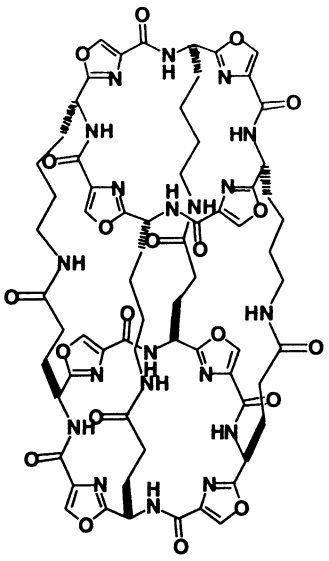




6

Putative shape of cylinder **7** (ex **4**)





7

Simple minimisation of these model cylinders produced structures (Fig. 5C and D) consisting of capping macrocycles **3** and **4**, respectively, linked through Lys and Glu amino acid side chains. Approximate dimensions of cylinders **6** and **7** were found to be ∼5.7 and 7.2 Å diameters, respectively with identical lengths of ∼11 Å. Tris-thiazole cylinder **6** was subjected to quenched molecular dynamics as described earlier and the lowest energy conformations for simulations at 1000 K and 300 K are shown as Fig. 5A and B, respectively. The quenched molecular dynamics simulations of **6** did not reach convergence, although the lowest energy conformers found are described here. Total energies of structures in Fig. 5A, B and C were 128, 133 and 162 kcal/mol, respectively. Cylindrical conformer Fig. 5A when viewed down the axis of the cylinder contains arms that twist in a helical nature forming an angle between the corresponding α-carbons of the upper and lower macrocycles of ∼65°. This extreme twisting causes the height of the cylinder to be reduced by ∼4.5 Å (ca. Fig. 5C). Interestingly, a similar twisted cylinder has been reported for other cavitands called carciplexes [42].

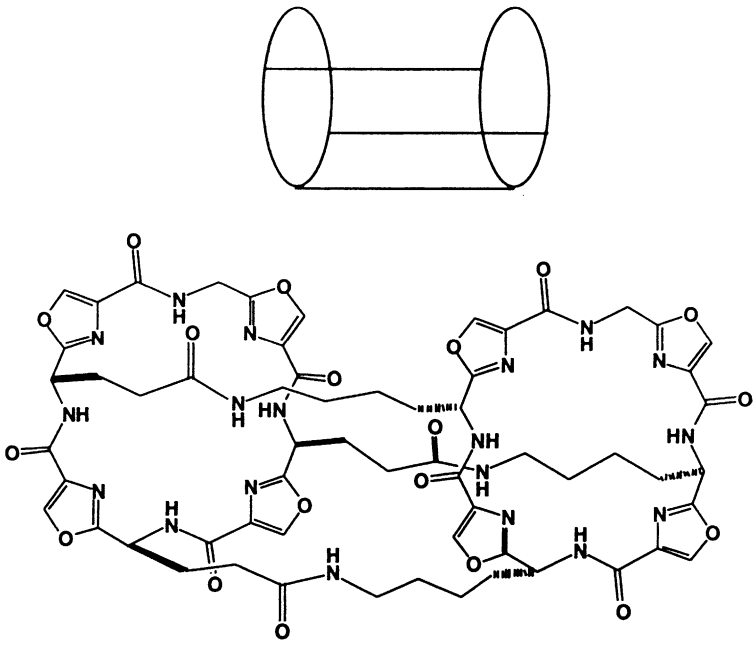
Although Fig. 5A has the lowest energy of the conformers shown we do not think it accurately represents the solution structure for **6**. Structure **6** has been synthesised [20] and NOE experiments have shown no NOE correlations between the amide NH in the macrocycle **3** to the linking arms, as required by the model in Fig. 5A. This is despite the fact that nearly all of the methylene groups in the linking arms of Fig. 5A are within 5 Å of the macrocycle **3** amide NH. The lack of explicit solvent during the modelling of the cylinder **6** is most likely to be the cause of the effect. If explicit solvent had been present, it is reasonable to expect it to fill the space created within the cylinder **6** and provide internal solvent pressure thus reducing the structures ability to collapse during modelling. These models could conceivably be elaborated with longer linkers between the cycles. If the linkers consisted of peptide chains, they might create pseudo-cylindrical cavity structures not unlike protein crevices or grooves.

A Connolly surface of cylinder **6** (Fig. 5C) indicated a narrow interior cavity that was not filled, thus leaving room for guest molecules. The size of the cavity formed by **6** was investigated by simple energy minimisation of some host–guest assemblies. It was found that guests such as hexane and glycerol, fitted easily whereas heptane caused a slight disturbance of one of the macrocycles and glucose caused disturbances in the linking arms. This indicates that small linear molecules should fit within the cavity of cylinder **6**.

A Connolly surface of cylinder **7** (Fig. 5D) also indicated that the interior of the cylinder was not filled, thus leaving room for guest molecules. The size and flexibility of the cavity formed by **7** was investigated by simple energy minimisation of host–guest assemblies using a variety of simple guest molecules. It was found that guests such as glycerol, glucose, ribose and extended dipeptide Ala–Ala fitted easily into the interior of the cylinder making a number of hydrogen bonds with the capping macrocycles and the linking arms. Extended tripeptide Ala–Ala–Ala was almost exactly the same length as the cylinder and formed H bonds with both macrocycles as well as with linking side arms. The dipeptide Phe–Phe was too large to be contained within the cylinder, causing the benzylic side chains to protrude between the linking arms of the cylinder.

3.5. Troughs derived from cyclic peptides

We are currently synthesising simple trough-shaped molecules like **8** (Fig. 6A) based on cylinders like **7**. Simple modelling of these structures is a valuable aid to visualising the size and shape of desired troughs. Such troughs can be viewed as potential models of crevices like binding grooves in proteins, DNA or RNA. Modelling allows us to select appropriately sized scaffolds (macrocycles) upon which to base our synthetic efforts towards such model structures.



8

For the purposes of modelling a simple trough-like structure, we simply removed one of the linking arms of cylinder **7** and then minimised the resulting structure to give Fig. 6A. This model has the following approximate dimensions, length ~11 Å, diameter ~7.2 Å, and depth of groove ~3.5 Å. Though this is only a small trough, the potential of larger trough-like structures with longer linkers between cyclic scaffolds can readily be envisaged and one such example is discussed ahead (Fig. 8).

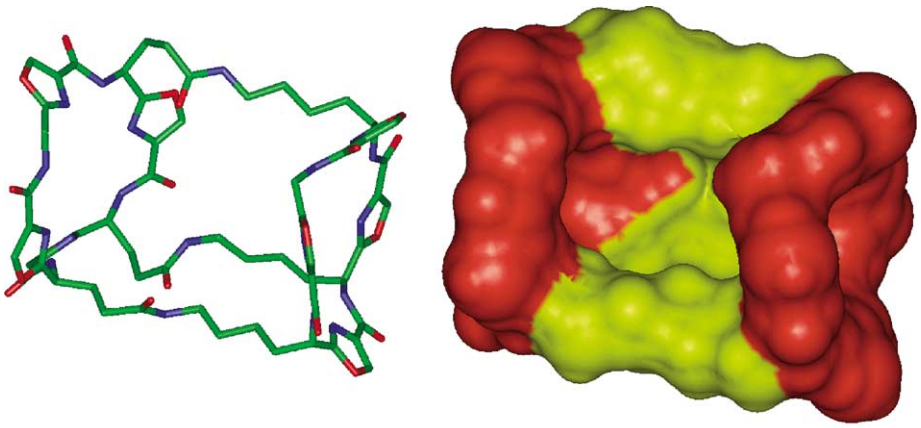
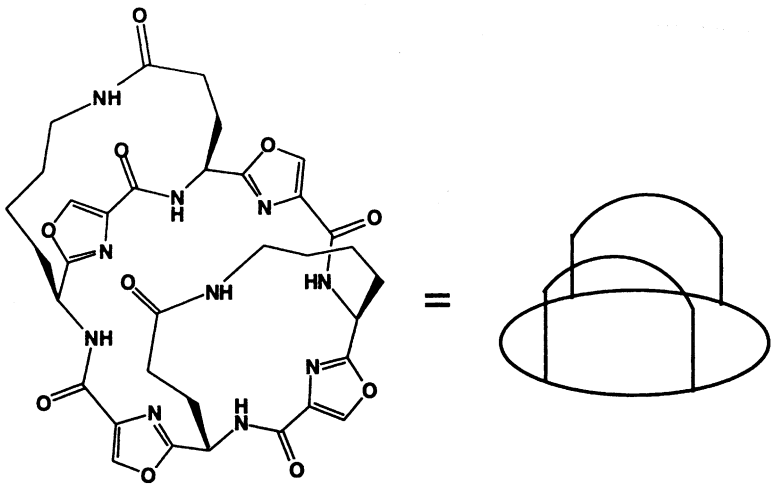


Fig. 6. Trough **8**. (A) Heavy atoms only, minimised structure viewed from above. (B) Connolly surface showing capping macrocycles in red and linkers between cycles in yellow.

3.6. Discontinuous loops

Protein folding tends to bring non-contiguous or discontinuous protein surfaces close together in space even though their component amino acids can be well separated in sequence. A major difficulty lies in mimicking such discontinuous surfaces. The constrained nature of the macrocycles examined makes them ideal scaffolds on which to mount peptidic loops for the purpose of mimicking selected protein surfaces.



9

For example, structure **9** (Fig. 7A) is a prototype for the construction of such multi-loop mimetics, consisting of a tetraoxazole macrocyclic scaffold **4** attached to two loops formed by condensing pairs of lysine and glutamic acid side chains (R) of **4**. However, two isomeric forms of **9** are possible (Fig. 7A, D) depending upon how the two loops are created. When joined above the macrocyclic N–H bond derived from glutamic acid (Fig. 7A), the N–H bond derived from lysine is left outside the loop whereas when two loops are joined above the macrocyclic N–H bond derived from lysine (Fig. 7D), the N–H bond derived from glutamic acid is left outside the loop. These isomers would be expected to show nearly identical NMR spectra making isomer identification difficult.

Fig. 7 shows the energy minimised structures of the possible loop isomers. Fig. 7A and B show the top and side views, respectively of isomer **9**, whereas Fig. 7C and D are the corresponding views of the second possible isomer. For both isomers the macrocyclic scaffold has retained the structure observed earlier for macrocycle **4** and each of the attached loops remains on the same face of the scaffold in a pseudo-symmetric fashion. However, the isomers contain distinct NHαCH dihedral angles (Fig. 7A: 150.3, −142.3; Fig. 7D: 140.6, 178.6) which reflect different folding influences in the macrocyclic scaffold. Differences such as these identified by modelling experiments can greatly assist in confirming the identity of synthetic products.

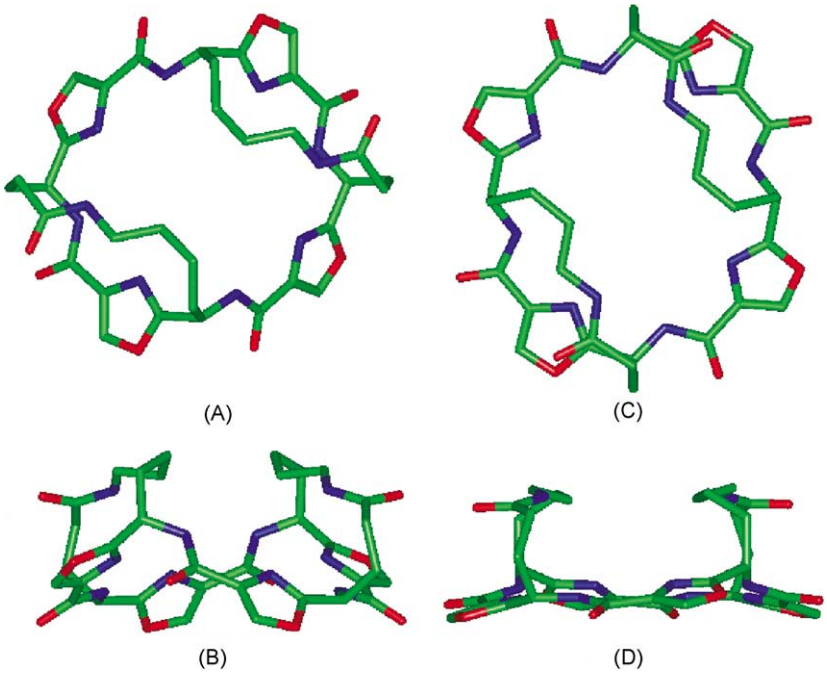


Fig. 7. Minimised conformation of loop structure **9** viewed from above (A) and side (B), and its isomer viewed from above (C) and sides (D).

Recently, in our laboratory an analogue of **9** was synthesised where every second oxazole was replaced by a thiazole [49]. NMR studies suggest that it has high symmetry and dihedral angles (via coupling constants) predict that its solution conformation is analogous to Fig. 7A and B.

4. Discussion

The intention of this article was to visualise some simple and highly symmetrical cyclic hexa- and octapeptide analogues (**3**, **4**), synthesised by us or others, [17–20,30,43–45] as potential scaffolds for creating novel supramolecular structures that begin to resemble protein surfaces and enzyme active sites. Using simple molecular dynamics simulations and energy minimisations the molecules were built in silico and the resulting predicted structures were compared with the limited structural data available from crystal structures, solution NMR structures, or dihedral angles derived from $^{3}J_{\mathrm{NH}\alpha\mathrm{CH}}$ coupling constants in NMR spectra. This allowed us to qualitatively estimate how accurately such simple modelling and molecular graphics representations were in predicting likely sizes and shapes of larger and

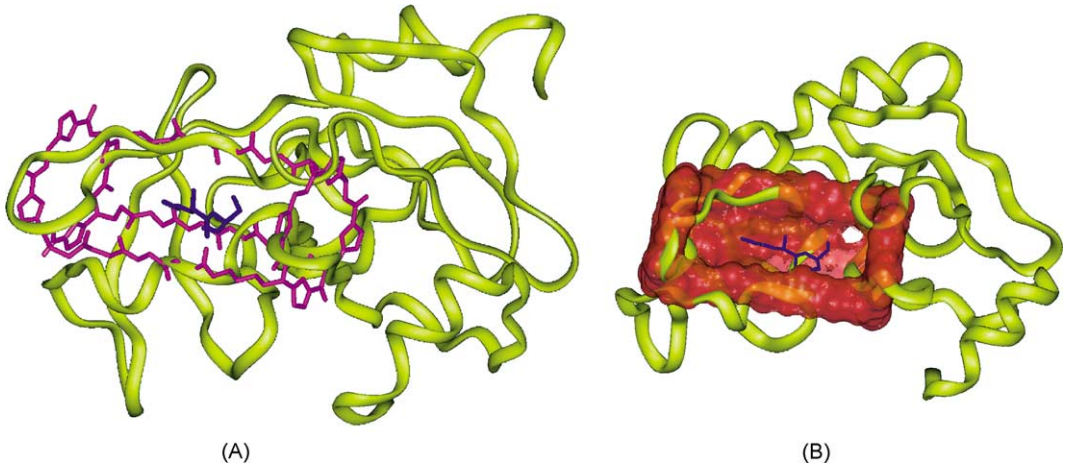


Fig. 8. (A) Simple energy minimised trough analogue (heavy atoms in purple) designed using two cyclic scaffolds **4** linked by amino acids to mimic approximate size and shape of the active site of dengue virus NS3 Protease (yellow ribbon). (B) Connolly surface of trough superimposed on the active site of dengue protease.

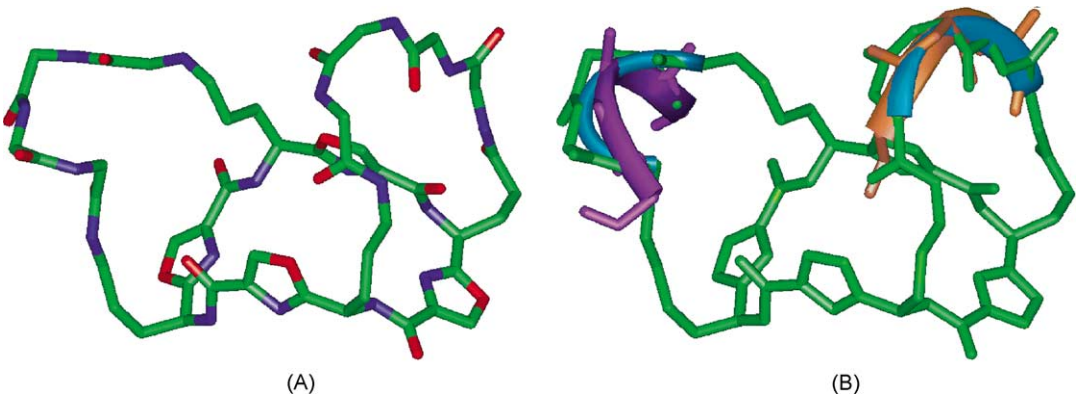


Fig. 9. Backbone heavy atoms only displayed. (A) Model of the tetrapeptide loops from L3 and H3 loops of antibody–antigen complex NC41 and N9 sialidase [25] grafted onto cyclic scaffold **4** with template forcing to mimic corresponding eight residues of the crystal structure. (B) Superimposition of minimised structure onto the template (H3 and L3 loops). The simulated model is shown in green with a ribbon representation of the loop chains shown in pale blue. The corresponding loops H3 and L3 loops are shown in orange and purple, respectively.

more elaborate molecules composed of such macrocyclic templates.

The results were that for the constrained systems such as **3**–**5**, both simple energy minimisation and quenched molecular dynamics produced conformations that were consistent with observed experimental data. However, it was also observed that quenched molecular dynamics also found conformers slightly higher in energy (e.g. **4**), even though the overall size and shape of the molecule remained essentially the same. Increasing the degree of flexibility in the molecules examined (e.g. **6** or **7**) showed the importance of solvation during modelling. The lack of explicit solvent during the

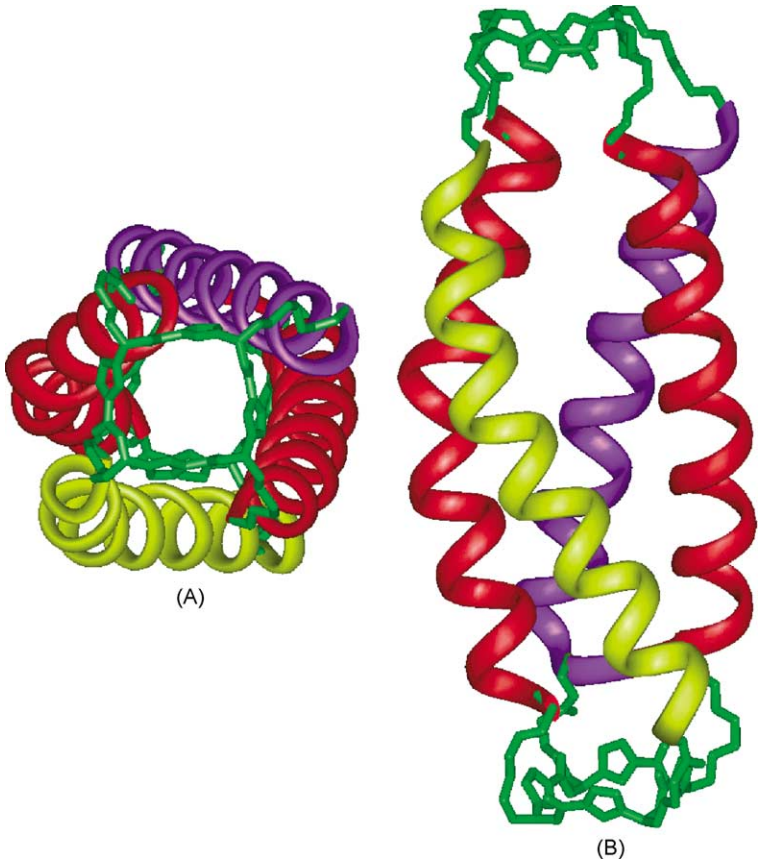


Fig. 10. Minimised structure for two macrocycles **4** capping the ends of a four-helix bundle derived from the four centre helices in the crystal structure of leukaemia inhibitory factor (A) end on, (B) side on [26].

modelling of the cylinders **7** and **8** produced low energy conformers that did not agree with physical data obtained from experimental measurements.

Our graphical and modelling work has suggested that macrocycle **3** is too small to accommodate either first row transition metal ions or even a methane molecule within its 'hole', and so this cyclic template can effectively be considered as a 'filled' cyclic template within supramolecular structures. The models of the larger cyclic octapeptide analogue **4** suggest the possibility of trapping small molecules or metal ions within the saddle-shaped structure. Indeed, ascidiacyclamide has been reported to trap solvents [31,33] like benzene, methanol, water as well as metal ions like Cu^{2+} [36]. When the number or oxidation state of heterocyclic rings within **4** is decreased, the resulting cyclic octapeptides and analogues do bind quite avidly to K^+, Ca^{2+}, Cu^{2+} and Zn^{2+} [34–39]. By analogy, derivatives of cones and cylinders composed of **3** and **4** might trap metal ions and other small molecules. The TACN unit in the cone would be expected to co-ordinate metal ions [41,46] if the dimensions of the cone are increased through longer alkyl linkers between the cycles. The cylinder is only likely to bind metal ions within the macrocyclic templates if the larger cyclic octapeptide template is used.

In terms of the more elaborate supramolecular structures created from these macrocycles, an important question was how well do they compare in terms of sizes and shapes with bioactive protein motifs. In Figs. 8–10 we show examples of mimics of the surfaces of an active site of a protease, two antigen-binding loops of an antibody, and a four-helix bundle of a protein, represented by trough **8**, double loop **9**, and four-helix bundle **10**, respectively.

Fig. 8A shows an energy minimised trough-like structure **8**, derived from two cyclic octapeptide analogues **4**, used to prospectively mimic the approximate size and shape of the active site of a proteolytic enzyme (dengue virus NS3 protease shown in yellow ribbon) pdb 1d9 [24]. The blue residues correspond to residues 520 and 521 of the Mung-bean Bowman–Birk inhibitor. Fig. 8A shows the trough consisting of two tetraoxazole rings which attempt to mimic two loops of the protease, and three straps which connect the two ends and mimic to some extent the strands surrounding the active site. Fig. 8B shows a Connolly surface of the "trough" placed on the active site of the dengue protease.

Fig. 9A shows a model of two loops of an antibody surface mounted upon the cyclic octapeptide analogue **4** used as a template. Two tetrapeptides were attached to the linkers of the cyclic template to construct the loops. Each loop corresponds to a loop from the crystal structure of the antibody–antigen complex NC41 and N9 sialidase [25]. The loop on the left corresponds to the L3 loop (His91–Ty2–Se3–Pro94) while the loop on the right corresponds to the H3 loop (Glu96–Asp97–Asn98–Phe99). This was modelled using template forcing with the corresponding eight residues of the crystal structure (pdb lnca) used as the template for superimposition. Compared with the initial energy of the loops (108 kcal, no restraints RMSD 3.37 Å for αC–βC, 16 atom pairs), the energy after template forcing became 128.6 kcal (restraints, RMSD 0.62 Å) versus 112.09 kcal (RMSD 1.74 Å) when energy minimised without restraints. This shows that the modelled loops mimic fairly well the known interaction conformation of the crystal structure of NC41.

Fig. 9B shows a superimposition of the final minimised structure onto the template (H3 and L3 loops). The loop model is shown in green with a ribbon representation of the αC chain shown in pale blue. The corresponding loops H3 and L3 are shown in orange and purple, respectively. The model uses only lysine and glutamate side chains to form amide bonds from the cyclic peptide scaffolds to the tetrapeptide loops. Clearly, derivatives with longer linkers would be more appropriate for this system.

Fig. 10 illustrates how two macrocycles **4** can be used to cap the ends of a four helix bundle. The four centre helices from leukaemia inhibitory factor (LIF, pdblLKI) [26] were cropped from the crystal structure, capped with the two cycles and minimised to show proof of concept. Whether such a four-helix bundle can be created in such a way remains to be seen. However, the diameter of **4** (∼8 Å) seems appropriate for it to be considered a possible template for mounting four peptides to generate a Template Assembled Synthetic Protein like that in Fig. 10. Indeed we have previously reported that a similar sized cyclic octapeptide, containing four *meta*-aminobenzoic acids alternating with four lysine amino acids, does cause formation of a four-helix bundle when attached to 4 equivalents of a prohelical 14-residue test peptide [47]. We concluded in that work that templates of 6–13 Å can be used to support four attached peptides and still permit four-helix bundle formation so long as the linkers between the template and peptide chains are long enough to allow the peptides to interact. Proteins with 4α-helix bundles have interhelix distances between axes that range between 6 and 16 Å, but commonly 10 ± 3 Å, [48] depending upon the packing requirements for their different sized hydrophobic peptide side chains.

There is obviously a great deal of scope to modify all of the cyclic scaffolds and derived supramolecular structures described herein. The resulting cavities can in principle be made longer, larger in diameter, more hydrophobic or hydrophilic, and diverse functionality can be incorporated into the linkers between cycles for possible applications as catalysts, models of protein active sites, artificial enzymes, and for fundamental host–guest investigations. It is clear to us from this study, that this novel system of supramolecular structures offers some particularly promising possibilities for developing protein-like molecules with protein-like functions and we are consequently in the process of developing synthetic routes to such novel molecules.

Acknowledgements

We thank the Australian Research Council for partial financial support of this work.

References

[1] Y. Singh, N. Sokolenko, G. Abbenante, D. Fairlie, The use of thiazole amino acids for the formation of conformationally constrained scaffolds or cavitands, in: Proceedings of the 15th National Conference of the Medicine and Agriculture Division, 11th National Convention of the Royal Australian Chemical Institute, Canberra, Feb. 6–11, 2000, pp. P102.
[2] D.J. Faulkner, Marine natural products, Nat. Prod. Rep. 15 (1998) 113–158.
[3] D.J. Faulkner, Marine natural products, Nat. Prod. Rep. 14 (1997) 259–302.
[4] D.J. Faulkner, Marine natural products, Nat. Prod. Rep. 13 (1996) 75–125.
[5] D.J. Faulkner, Marine natural products, Nat. Prod. Rep. 12 (1995) 233–269.
[6] D.J. Faulkner, Marine natural products, Nat. Prod. Rep. 11 (1994) 355–394.
[7] R.P. McGeary, D.P. Fairlie, Macrocyclic peptidomimetics: potential for drug development, Curr. Opin. Drug Discovery Dev. 1 (1998) 208–217.
[8] G.D. Fate, C.P. Benner, S.H. Grode, T.J. Gilbertson, The biosynthesis of sulfomycin elucidated by isotopic labeling studies, J. Am. Chem. Soc. 118 (1996) 11363–11368.
[9] P. Wipf, Synthetic studies of biologically active marine cyclopeptides, Chem. Rev. 95 (1995) 2115–2134.
[10] J.P. Michael, G. Pattenden, Marine metabolites and the complexation of metal ions: facts and hypotheses, Angew. Chem. Int. Ed. Engl. 31 (1993) 1–23.
[11] D.P. Fairlie, G. Abbenante, D.R. March, Macrocyclic Peptidomimetics: forcing peptides into bioactive conformations, Curr. Med. Chem. 95 (1995) 654–686.
[12] D.P. Fairlie, M.W. West, A.K. Wong, Towards protein surface mimetics, Curr. Med. Chem. 5 (1998) 29–62.
[13] T.K. Sawyer, V.J. Hruby, P.S. Darman, M.E. Hadely, [half-Cys4, half-Cys10]-α-Melanocyte-stimulating hormone: a cyclic α-melanotropin exhibiting superagonist biological activity, Proc. Natl. Acad. Sci. U.S.A. 79 (1982) 1751.
[14] B. Pernow, Substance P, Pharmacol. Rev. 35 (1983) 85–141.
[15] L.T. Tan, R.T. Williamson, W.H. Gerwick, K.S. Watts, K. McGough, R. Jacobs, *cis*,*cis*- and *trans*, *trans*-Ceratospongamide, new bioactive cyclic heptapeptides from the Indonesian red alga ceratodictyon spongiosum and symbiotic sponge sigmadocia symbiotica, J. Org. Chem. 65 (2000) 419–425.
[16] Z. Xia, C.D. Smith, Total synthesis of dendroamide A, a novel cyclic peptide that reverses multiple drug resistance, J. Org. Chem. 66 (2001) 3459–3466.
[17] N. Sokolenko, G. Abbenante, M.J. Scanlon, A. Jones, L.R. Gahan, G.R. Hanson, D.P. Fairlie, Cyclooligomerization of thiazole-containing tetrapeptides symmetrical macrocycles with up to 76 amino acids, J. Am. Chem. Soc. 121 (1999) 2603–2604.
[18] P. Wipf, C.P. Miller, C.M. Grant, Synthesis of cyclopeptide alkaloids by cyclooligomerization of dipeptidyl oxazolines, Tetrahedron 56 (2000) 9143–9150.
[19] A. Bertram, J.S. Hannam, K.A. Jolliffe, F. Gonzalez-Lopez de Turiso, G. Pattenden, The synthesis of novel thiazole containing cyclic peptides via cyclooligomerization reactions, Synlett 11 (1999) 1723–1726.
[20] Y. Singh, N. Sokolenko, M.J. Kelso, L.R. Gahan, G. Abbenante, D.P. Fairlie, Novel cylindrical, conical and macrocyclic peptides from the cyclooligomerization of functionalized thiazole amino acids, J. Am. Chem. Soc. 123 (2001) 333–334.
[21] InsightII Modeling Environment, Release 2000, Accelrys Inc., San Diego, 2001.
[22] G. Chang, W.C. Guida, W.C. Still, An internal co-ordinate Monte Carlo method for searching conformational space, J. Am. Chem. Soc. 111 (1989) 4379–4386.
[23] M. Saunders, Stochastic search for the conformations of bicyclic hydrocarbons, J. Comput. Chem. 10 (1989) 203–208.
[24] H.M. Krishna Murthy, K. Judge, L. DeLucas, R. Padmanabhan, Crystal structure of dengue virus NS3 protease in complex with a Bowman–Birk inhibitor: implications for flaviviral polyprotein processing and drug design, J. Mol. Biol. 301 (2000) 759–767.
[25] W.R. Tulip, J.N. Varghese, W.G. Laver, R.G. Webster, P.M. Colman, Refined crystal structure of the influenza virus N9 neuraminidase-NC41 Fab complex, J. Mol. Biol. 227 (1992) 122–148.
[26] R.C. Robinson, L.M. Grey, D. Staunton, A.B. Vankelecom, A.B. Vernallis, J.F. Moreau, D.I. Stuart, J.K. Heath, E.Y. Jones, The crystal structure and biological function of leukemia inhibitory factor: Implications for receptor binding, Cell 77 (1994) 1101–1116.
[27] A.R. Leach, Molecular Modelling: Principles and Applications, Longman, Harlow, England, 1996, ISBN 0-582-23933-8.
[28] A.K. Todorova, F. Juttner, A. Linden, T. Pluss, W. von Philipsborn, Nostocyclamide: a new macrocyclic, thiazole-containing allelochemical from Nostoc sp. 31 (cyanobacteria), J. Org. Chem. 60 (1995) 7891–7895.
[29] W.L. Jorgensen, J. Pranata, Importance of secondary interactions in triply hydrogen bonded complexes: guanine–cytosine vs. uracil-2,6, diaminopyridine, J. Am. Chem. Soc. 112 (1990) 2008–2010.
[30] D. Mink, S. Mecozzi, J. Rebek Jr., Natural products analogs as scaffolds for supramolecular and combinatorial chemistry, Tetrahedron Lett. 39 (1998) 5709–5712.
[31] T. Ishida, M. Inoue, Y. Hamada, S. Kato, T. Shioiri, X-ray crystal structure of ascidiacylamide, a cytotoxic cyclic peptide from ascidian, J. Chem. Soc., Chem. Commun. (1987) 370–371.
[32] T. Ishida, M. Tanaka, M. Nabae, M. Inoue, S. Kato, Y. Hamada, T. Shioiri, Solution and solid-state conformations of ascidiacyclamide, a cytotoxic cyclic peptide from ascidian, J. Org. Chem. 53 (1988) 107–112.
[33] T. Ishida, Y. In, M. Doi, M. Inoue, Y. Hamada, T. Shioiri, Molecular conformation of ascidiacyclamide, a cytotoxic cyclic peptide from ascidian: X-ray analyses of its free form and solvate crystals, Biopolymers 32 (1992) 131–134.
[34] T. Ishida, Y. In, F. Shinozaki, M. Doi, D. Yamamoto, Y. Hamada, T. Shioiri, M. Kamigauchi, M. Sugiura, Solution conformations of patellamides B and C, cytotoxic cyclic hexapeptides from marine tunicate, determined by NMR spectroscopy and molecular dynamics, J. Org. Chem. 60 (1995) 3944–3952.
[35] G. Abbenante, D.P. Fairlie, L.R. Gahan, G.R. Hanson, G. Pierens, A.L. van den Brenk, Conformational control by thiazole and oxazoline rings in cyclic octapeptides of marine origin, novel macrocyclic chair and boat conformations, J. Am. Chem. Soc. 118 (1996) 10384–10388.
[36] R.M. Cusack, L. Grøndahl, D.P. Fairlie, L.R. Gahan, G.R. Hanson, Cyclic octapeptides containing thiazole. Effect of stereochemistry and degree of flexibility on calcium binding properties, Perkin Trans. 2 (2002) 556–563.
[37] A.L. Van den Brenk, K.A. Byriel, D.P. Fairlie, L.R. Gahan, G.R. Hanson, T. Hambley, C.J. Hawkins, C.H.L. Kennard, B.J. Moubaraki, K.S. Murray, Crystal structure and electrospray ionization mass spectrometry, electron paramagnetic resonance, and magnetic susceptibility study of [Cu2(ascidH2)(1,2-μ-CO3)(H2O)2]·2H2O, the bis(copper(II)) complex of ascidiacyclamide (ascidH4), a cyclic

peptide isolated from the ascidian lissoclinum patella, Inorg. Chem. 33 (1994) 3549–3557.
[38] A. van den Brenk, D.P. Fairlie, L.R. Gahan, G.R. Hanson, T.W. Hambley, A novel potassium-binding hydrolysis product of ascidiacyclamide, a cyclic octapeptide isolated from the ascidian lissoclinum patella, Inorg. Chem. 35 (1996) 1095–1100.
[39] L. Grøndahl, N. Sokolenko, G. Abbenante, D.P. Fairlie, G.R. Hanson, L.R. Gahan, Interaction of Zn(II) with the cyclooctapeptides cyclo-[Ile-Oxn-D-Val(Thz)]2 and ascidiacyclamide, a cyclic peptide from Lissolclinum patella, J. Chem. Soc., Dalton Trans. 8 (1999) 1227–1234.
[40] R.M. Cusack, L. Grøndahl, G. Abbenante, D.P. Fairlie, L.R. Gahan, G.R. Hanson, T.W. Hambley, Conformations of cyclic octapeptides and the influence of heterocyclic ring constraints upon calcium binding, J. Chem. Soc., Perkin Trans. 2 (2000) 323–331.
[41] P. Chaudhuri, K. Wieghardt, The chemistry of 1,4,7-triazacyclononane and related tridentate macrocyclic compounds, Prog. Inorg. Chem. 35 (1987) 329–436.
[42] R.G. Chapman, J.C. Sherman, Elucidation of "twistomers" in container compounds, J. Am. Chem. Soc. 121 (1999) 1962–1963.
[43] C. Boss, P.H. Rasmussen, A.R. Wartini, S.R. Waldvogel, Functionalized platforms based on marine cyclopeptides: different pathways to the hexapeptide, Tetrahedron Lett. 41 (2000) 6327–6331.
[44] G. Haberhauer, L. Somogyi, J. Rebek Jr., Synthesis of a second-generation pseudopeptide platform, Tetrahedron Lett. 41 (2000) 5013–5016.
[45] L. Somogyi, G. Haberhauer, J. Rebek Jr., Improved synthesis of functionalized molecular platforms related to marine cyclopeptides, Tetrahedron Lett. 57 (2001) 1699–1708.
[46] A.A. Watson, A.C. Willis, D.P. Fairlie, Organization of amino acids using a metallotriazacyclononane template, Inorg. Chem. 36 (1997) 752–753.
[47] A.K. Wong, M.P. Jacobsen, D.J. Winzor, D.P. Fairlie, Template assembled synthetic proteins (TASPS). Are template size, shape and directionality important in formation of four-helix bundles? J. Am. Chem. Soc. 120 (1998) 3836–3841.
[48] B.V.B. Reddy, T.L. Blundell, Packing of secondary structural elements in proteins. Analysis and prediction of inter-helix distances, J. Mol. Biol. 233 (1993) 464–479.
[49] Y. Singh et al., Org. Letts. (2002), in press.