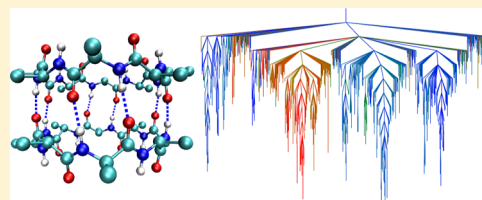


# Energy Landscapes and Global Optimization of Self-Assembling Cyclic Peptides

Mark T. Oakley and Roy L. Johnston\*

School of Chemistry, University of Birmingham, Edgbaston, Birmingham, B15 2TT, United Kingdom

**ABSTRACT:** Self-assembled cyclic peptide nanotubes have attracted much attention because of their antimicrobial properties. Here, we present calculations on the formation of cyclic peptide dimers using basin-hopping and discrete path sampling. We present an analysis of the basin-hopping move sets that most efficiently explore the conformations of cyclic peptides. Group rotation moves, in which sections of the ring are rotated as a rigid body, are the most effective for cyclic peptides containing up to 20 residues. For cyclic peptide dimers, we find that a combination of group rotation intramolecular moves and rigid body intermolecular moves performs well. Discrete path sampling calculations on the cyclic peptide dimers show significant differences in the dimerization of hexa- and octapeptides.



## 1. INTRODUCTION

Self-assembly is the process by which small molecules aggregate to form large, ordered structures. Some head-to-tail cyclic peptides are known to self-assemble to form nanotubes.<sup>1–5</sup> Self-assembling cyclic peptide nanotubes have attracted attention for their possible biomedical uses, including activity against bacteria.<sup>6,7</sup> Cyclic peptide nanotubes have been observed to form across lipid bilayers.<sup>3,8,9</sup> The hydrogen-bonding arrangement of cyclic peptide nanotubes is similar to that seen in amyloid fibrils, and it has been shown that cyclic peptides can disrupt amyloid formation.<sup>10</sup> Cyclic peptide–polymer conjugates have potential uses as self-assembling materials.<sup>11,12</sup>

Self-assembling cyclic peptides have sequences comprising alternating D- and L-peptide residues because this arrangement places the side-chains equatorial to the peptide rings with the peptide groups aligned axially to hydrogen bond to other cyclic peptides. Cyclic octapeptides are the most widely studied because they have the highest tendency to form nanotubes.<sup>1–4</sup> However, dimerization of cyclic hexapeptides has been observed,<sup>13</sup> and nanotubes of larger cyclic peptides are known.<sup>9,14</sup> Nanotubes with antiparallel hydrogen bonding arrangements are more stable than those with parallel hydrogen bonds.<sup>15</sup>

Several molecular dynamics (MD) investigations have been undertaken to model the properties of cyclic peptide nanotubes. The formation of dimers has been modeled with MD.<sup>16</sup> The composition of cyclic peptides has a substantial effect on the stability and flexibility of their nanotubes.<sup>17</sup> Longer nanotubes are more stable than shorter nanotubes.<sup>18,19</sup> MD has been used to study the motion of water<sup>20</sup> and other small molecules<sup>21</sup> through cyclic peptide nanotubes. Cyclic peptide–polymer conjugates have also been studied with MD.<sup>12</sup> All-atom MD is very computationally demanding, and modeling the formation of nanotubes is prohibitively expensive. For this reason, coarse-grained models have been used to study the aggregation of cyclic peptides in lipid bilayers, and

nanotubes have been shown to form across<sup>22,23</sup> or along<sup>24</sup> the membrane, depending on the peptide sequence.

Here, we take a different approach and consider only the stationary points on the energy landscape. The solvent is treated implicitly, which substantially reduces the computational resources required. Locating minima on the energy landscape then becomes a global optimization problem that is split into two parts: searching the conformations of individual cyclic peptides and assembly of cyclic peptide aggregates.

Several methods are available to search the conformations of small flexible molecules. Those that have been used on cyclic peptides include distance geometry methods,<sup>25</sup> Monte Carlo,<sup>26</sup> LowMode molecular dynamics,<sup>27</sup> iterative stochastic search,<sup>28</sup> dihedral angle sampling<sup>29</sup> and diffusion equation evolutionary programming simulated annealing.<sup>30</sup> Optimization of clusters of flexible molecules is a more difficult problem. An evolutionary algorithm has been used to optimize the dimer of kanamycin A.<sup>31,32</sup> Basin-hopping has been used to optimize the small aggregates of the amyloid- $\beta$  peptide.<sup>33,34</sup> Here, we use the basin-hopping algorithm<sup>35</sup> to locate the minima on the cyclic peptide potential energy surfaces.

To understand the conformations of molecules fully we need to know both the minima and the transition states linking them. Discrete path sampling<sup>36–38</sup> is an efficient method for exploring energy landscapes and we have previously used it to study the conformations of peptides.<sup>39,40</sup> Here, we present energy landscapes for the peptide dimers from discrete path sampling calculations.

## 2. METHODS

The energies of all structures were evaluated using the AMBER FF03 force field.<sup>41,42</sup> The effect of solvent was included with the Generalized Born implicit solvation model.<sup>43</sup> Calculations

Received: January 3, 2014

Published: February 25, 2014

were performed in water ( $\epsilon_r = 80$ ) or in a low-dielectric solvent ( $\epsilon_r = 2$ ) to model a membrane environment.

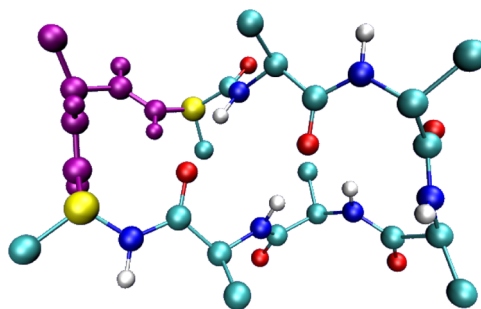
Global optimization of single cyclic peptides and dimers was performed using the basin-hopping algorithm<sup>35</sup> as implemented in GMIN.<sup>44</sup> In this algorithm, structural deformations are followed by a gradient-driven quench to a local minimum before a Monte Carlo acceptance test. Thus, all attempted moves are between local minima on the potential energy surface. For efficient searching, the moves must be sufficiently large to escape the basin of attraction of the initial local minimum. However, moves that enter basins that are not physically accessible should be avoided. Inversion of the chirality of the  $C_\alpha$  atoms is not possible in experiment and has a barrier of about 70 kcal mol<sup>-1</sup> in the Amber force field.<sup>45</sup> Therefore, any moves that lead to structures where the chirality of one or more  $C_\alpha$  atoms are inverted are rejected.

We consider three types of intramolecular move: atom displacements, MD moves, and group rotations. A series of basin-hopping optimizations were performed to benchmark each move class. For cyclic peptides comprising 10 or fewer peptide residues, the global minimum structures are found by all searches after fairly short basin-hopping runs. In these cases, the performance was measured in terms of the mean time to the first encounter of the global minimum from 100 independent searches. For larger cyclic peptides, the conformational spaces are too large for us to be confident that the lowest minima we find are the global minima. In these cases, basin-hopping runs were run for 10 000 Monte Carlo steps and the performance was measured in terms of the lowest energies found at the end of each run. The initial structures for all of these searches were taken from a molecular dynamics simulation at 1600 K with snapshots taken every 10 ps.

The parameters used for all move classes were optimized by performing sets of 10 searches with different values of the parameters for peptides containing up to 10 residues. These initial 10 searches are included in the set of 100 searches used to calculate the mean first encounter time. Unless noted, the same search parameters are used for all of the cyclic peptides studied here.

The atom displacement moves (STEP keyword) involve displacement of all the Cartesian coordinates of all atoms by a up to 1.5 Å. In the MD moves (AMBERMDSTEPS keyword), short MD runs of 1000 steps are performed to escape the basins of attraction of local minima. These runs are performed at high temperature of 1600 K to allow the move to rapidly surmount large barriers, such as those between cis and trans isomers. In the group rotation moves (GROUPROTATION keyword), two atoms in the peptide backbone are selected and all atoms between them are rotated around the axis defined by those two atoms (Figure 1).<sup>46</sup> Four types of group rotation are used (Table 1): cis-trans, peptide, dipeptide, and tripeptide moves. Cis-trans and peptide moves are attempted for all cyclic peptides. Dipeptide moves are attempted for cyclic peptides with six or more residues. The performance of tripeptide moves is evaluated for cyclic peptides comprising 12 or more residues. The magnitudes of all group rotation moves are in the range  $\pm 180^\circ$ . The probability of selecting a group for group rotation is defined so that each basin-hopping step comprises on average  $1/3$  cis-trans moves, three peptide moves and, where appropriate, one of each of the larger moves.

A combined set of intramolecular and intermolecular moves is needed to search the structures of dimers, and we consider two classes of move. MD moves use a short, high-temperature



**Figure 1.** Atoms involved in a dipeptide group rotation move. The axis is defined by two  $C_\alpha$  atoms (yellow). All of the intermediate atoms (purple) are rotated around this axis. Some hydrogen atoms are omitted for clarity.

**Table 1.** Atoms Involved in the Each Type of Group Rotation Move

move	ends	
cis-trans	$C_{\alpha,i}$	$N_{i+1}$
peptide	$C_{\alpha,i}$	$C_{\alpha,i+1}$
dipeptide	$C_{\alpha,i}$	$C_{\alpha,i+2}$
tripeptide	$C_{\alpha,i}$	$C_{\alpha,i+3}$

MD run to perform both the intra- and intermolecular moves. We also use a combination of group rotation intramolecular moves and rigid-body<sup>47</sup> intermolecular moves. For both types of move, a hybrid minimization scheme is used (Figure 2). The second step is particularly important when using MD moves because the two cyclic peptides can drift apart during the MD run.

(1) Deform structure with MD or rigid body moves.
(2) Apply a compressive force and minimise to loose convergence treating each ring as a rigid body
(3) When distances between rigid bodies' centres of masses are below a threshold, turn off the compressive force and continue relaxation.
(4) Once the RMS force has reached the loose convergence criterion, minimise fully atomistically.

**Figure 2.** Basin-hopping scheme with hybrid minimization.

We evaluate the move sets on two types of cyclic peptides. Cyclic oligoglycines do not have any side chains and allow us to study the effect of the moves on the peptide backbones. We have not studied the dimerization of these peptides because they are not known to form nanotubes. Cyclic peptides with alternating D-Ala and L-Ala residues were selected as representative cyclic D-, L-peptides. Other peptide sequences form more stable nanotubes,<sup>17</sup> but alanine was selected to reduce the effort of modeling the peptide side chains.

The energy landscapes of some of the cyclic peptides and dimers were explored with discrete path sampling,<sup>36–38</sup> as implemented in PATHSAMPLE.<sup>48</sup> The databases of stationary points were seeded with minima from the basin-hopping calculations. Pathways between minima were found using OPTIM,<sup>49</sup> with candidate transition states found by doubly nudged elastic band calculations<sup>50</sup> and refined with hybrid eigenvector following.<sup>51–53</sup> Pairs of minima were initially selected for connection by the missing connections algorithm<sup>54</sup> and later using the UNTRAP method to remove artificial frustration.<sup>55</sup> The resulting databases of stationary points are visualized as disconnectivity graphs.<sup>56,57</sup> Stationary points that are permutation-inversion isomers are grouped together in

these disconnectivity graphs. All disconnectivity graphs show the free energy surface at 298 K. The free energies of all stationary points are calculated using the harmonic superposition approximation.<sup>58,59</sup>

We use the pattern of hydrogen bonds as a metric to measure the parallel or antiparallel alignment of these dimers. Hydrogen bonds in the dimers were counted using the definition from the Dictionary of Secondary Structure Prediction.<sup>60</sup> Intermolecular hydrogen bonds were then assigned as parallel or antiparallel based on the angles between the two vectors joining the  $C_\alpha$  atoms on either side of the peptide groups involved. The alignment is defined as  $n_{\text{par}} - n_{\text{antipar}}$ .

### 3. RESULTS

**3.1. Intramolecular Moves.** We have tested the performance of basin-hopping moves on a series of cyclic oligoglycines. For consistency with previous studies,<sup>39,40</sup> all basin-hopping searches on single cyclic peptides were performed using water as a solvent. We have previously shown that the global minimum for cyclo-(Gly<sub>4</sub>) has all four peptide groups aligned axially, with no intramolecular hydrogen bonds (Figure 3a).<sup>39</sup> The global minima for all of the larger cyclic oligoglycines exhibit several intramolecular hydrogen bonds (Figure 3).

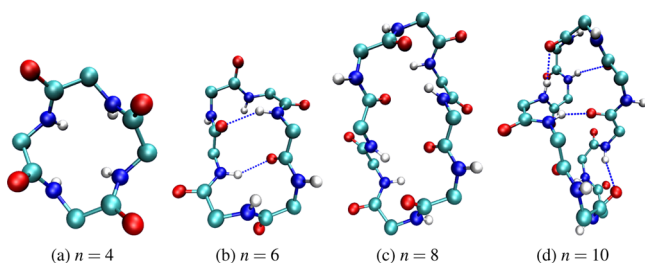


Figure 3. Global minima of cyclo-(Gly<sub>n</sub>) in water.

For all four of the cyclic oligoglycines, the group rotation moves give the most efficient optimization in terms of the CPU time and the number of energy evaluations (Table 2). For  $n \leq 8$ , the MD moves require a similar number of quenches to the group rotation moves but need more CPU time because of the overhead of performing the short MD run. When  $n = 10$ , molecular dynamics moves perform poorly and visit the

Table 2. Mean Times to the First Encounter of the Global Minimum for 100 Basin-Hopping Runs on Cyclic Oligoglycines

molecule	move	mean first encounter time		
		energy evaluations	minimizations	CPU time (s)
cyclo-(Gly <sub>4</sub> )	displacement	$1.8 \times 10^4$	99	2.6
	MD	$3.7 \times 10^4$	32	4.7
	GR	$3.2 \times 10^3$	18	0.45
cyclo-(Gly <sub>6</sub> )	displacement	$2.6 \times 10^4$	90	7.9
	MD	$2.7 \times 10^4$	21	6.8
	GR	$1.3 \times 10^4$	51	3.8
cyclo-(Gly <sub>8</sub> )	displacement	$3.0 \times 10^5$	680	140
	MD	$2.9 \times 10^5$	190	120
	GR	$8.8 \times 10^4$	260	40
cyclo-(Gly <sub>10</sub> )	displacement	$5.2 \times 10^5$	1000	370
	MD	$3.0 \times 10^6$	1700	1900
	GR	$1.8 \times 10^5$	420	120

second-lowest structure several times before locating the global minimum. The displacement moves require a similar amount of CPU time to MD moves when  $n \leq 8$ .

We have also tested the performance of basin-hopping on a series of cyclic peptides with alternating D-Ala and L-Ala residues. We have previously shown that the global minimum of cyclo-((D-Ala-L-Ala)<sub>2</sub>) includes one peptide in a cis conformation (Figure 4a).<sup>39</sup> The global minima for all of the larger cyclo-((D-Ala-L-Ala)<sub>n</sub>) have all of their peptide bonds in trans conformations (Figure 4).

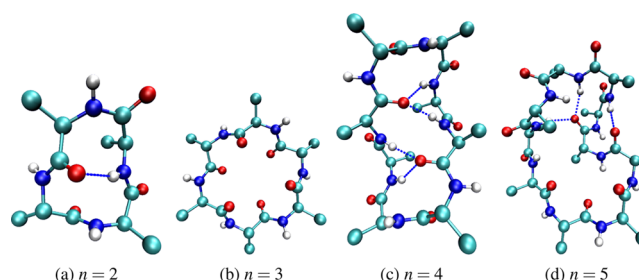


Figure 4. Global minima of cyclo-((D-Ala-L-Ala)<sub>n</sub>) in water.

The performance of the displacement moves is poor for these molecules. Cartesian displacements that are sufficiently large to step between basins on the potential energy surface also lead to the inversion of at least one of the chiral centers in about 45% of the attempted moves. For comparison, MD moves never lead to inversion of chiral centers and group rotation moves invert in 5–10% of attempts. In most cases, the group rotation moves give the most efficient optimization. The MD moves locate the global minimum of cyclo-((D-Ala-L-Ala)<sub>4</sub>) (Figure 4c) very rapidly (Table 3). However, this seems to be a particularly easy global minimum to locate because the mean first encounter times with all move types are faster than the corresponding times for cyclo-((D-Ala-L-Ala)<sub>3</sub>).

Basin-hopping optimization has also been performed on several larger cyclic peptides comprising up to 20 residues (table 4). For these molecules, we only consider the MD and group rotation moves because of the poor performance of the displacement moves in the previous calculations. The performance was assessed in terms of the mean energy of the lowest structure found in each basin-hopping run and the lowest energy found across all basin-hopping runs. For all of the cyclic peptides, the lowest minimum found using group rotation moves up to and including the dipeptide move is more stable than the lowest found with molecular dynamics moves. By both measures, group rotation with dipeptide moves out-performs MD moves for all cyclic oligoglycines. The inclusion of tripeptide moves makes only a small difference to the lowest energies for the smaller rings but gives a substantial improvement for cyclo-(Gly<sub>20</sub>). The CPU time taken for each Monte Carlo step is also substantially longer for the MD moves.

The energy landscape of cyclo-((D-Ala-L-Ala)<sub>3</sub>) is funnelled, with small downhill barriers separating all local minima from the global minimum (Figure 5). The conformation with the lowest free energy in water is similar to that seen in cyclic peptide nanotubes. In nonpolar solvent, the lowest free energy structure exhibits some intramolecular hydrogen bonding. In cyclo-((D-Ala-L-Ala)<sub>3</sub>), conformers containing cis peptides are particularly unstable and are at least 8.8 kcal mol<sup>−1</sup> above the global minimum in water and 10.3 kcal mol<sup>−1</sup> higher in

Table 3. Mean Times to the First Encounter of the Global Minimum in 100 Basin-Hopping Runs on Cyclic Oligo-D,L-alanines

molecule	move	mean first encounter time		
		energy evaluations	minimizations	CPU time (s)
cyclo-((D-Ala-L-Ala) <sub>2</sub> )	displacement	$4.5 \times 10^4$	180	11
	MD	$1.6 \times 10^4$	13	3.7
	GR	$1.1 \times 10^4$	44	2.7
cyclo-((D-Ala-L-Ala) <sub>3</sub> )	displacement	$1.9 \times 10^6$	5200	1000
	MD	$1.5 \times 10^6$	1100	670
	GR	$1.2 \times 10^5$	300	67
cyclo-((D-Ala-L-Ala) <sub>4</sub> )	displacement	$1.2 \times 10^6$	2600	1100
	MD	$6.2 \times 10^4$	39	48
	GR	$6.9 \times 10^4$	130	63
cyclo-((D-Ala-L-Ala) <sub>5</sub> )	displacement	$6.4 \times 10^6$	12000	8100
	MD	$6.6 \times 10^6$	3900	7700
	GR	$3.3 \times 10^5$	530	430

Table 4. Lowest Energies Found in 20 × 10 000 Step Basin-Hopping Runs on Cyclic Peptides with 12 or More Residues

molecule	move	mean lowest energy	lowest energy	mean CPU time (s)
cyclo-(Gly <sub>12</sub> )	MD	−10.076	−10.370	$2.1 \times 10^4$
	GR (dipeptide)	−10.514	−10.514	$5.8 \times 10^3$
	GR (tripeptide)	−10.436	−10.514	$7.3 \times 10^3$
cyclo-(Gly <sub>14</sub> )	MD	−15.066	−15.983	$3.1 \times 10^4$
	GR (dipeptide)	−15.770	−16.540	$8.0 \times 10^3$
	GR (tripeptide)	−15.795	−16.540	$1.0 \times 10^4$
cyclo-(Gly <sub>16</sub> )	MD	−17.997	−19.140	$4.7 \times 10^4$
	GR (dipeptide)	−19.552	−21.067	$1.1 \times 10^4$
	GR (tripeptide)	−19.007	−21.174	$1.5 \times 10^4$
cyclo-(Gly <sub>18</sub> )	MD	−22.833	−24.282	$5.8 \times 10^4$
	GR (dipeptide)	−21.548	−26.437	$1.9 \times 10^4$
	GR (tripeptide)	−23.480	−24.674	$1.9 \times 10^4$
cyclo-(Gly <sub>20</sub> )	MD	−25.779	−27.779	$6.8 \times 10^4$
	GR (dipeptide)	−25.788	−29.236	$2.5 \times 10^4$
	GR (tripeptide)	−27.465	−29.471	$2.5 \times 10^4$
cyclo-((D-Ala-L-Ala) <sub>6</sub> )	MD	−12.207	−13.252	$4.3 \times 10^4$
	GR (dipeptide)	−13.243	−13.252	$1.3 \times 10^4$
	GR (tripeptide)	−13.252	−13.252	$1.6 \times 10^4$
cyclo-((D-Ala-L-Ala) <sub>7</sub> )	MD	−17.690	−18.637	$5.9 \times 10^4$
	GR (dipeptide)	−18.380	−18.637	$1.9 \times 10^4$
	GR (tripeptide)	−18.291	−18.637	$2.5 \times 10^4$
cyclo-((D-Ala-L-Ala) <sub>8</sub> )	MD	−25.577	−26.480	$8.5 \times 10^4$
	GR (dipeptide)	−26.638	−26.814	$2.6 \times 10^4$
	GR (tripeptide)	−26.438	−26.814	$3.4 \times 10^4$
cyclo-((D-Ala-L-Ala) <sub>9</sub> )	MD	−29.764	−31.668	$1.1 \times 10^5$
	GR (dipeptide)	−30.590	−31.668	$3.3 \times 10^4$
	GR (tripeptide)	−30.540	−31.668	$4.3 \times 10^3$
cyclo-((D-Ala-L-Ala) <sub>10</sub> )	MD	−35.782	−37.079	$1.2 \times 10^5$
	GR (dipeptide)	−35.578	−38.493	$4.3 \times 10^4$
	GR (tripeptide)	−35.470	−38.434	$5.6 \times 10^4$

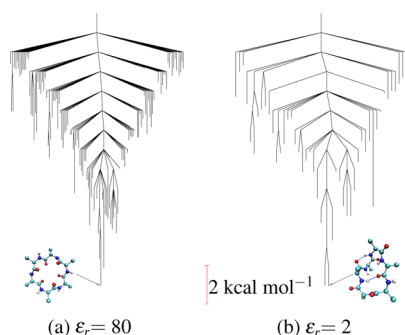
nonpolar solvent. This represents a substantial destabilization of the cis isomers, which are about 4 kcal mol<sup>−1</sup> above the trans isomers in acyclic peptides and unstrained cyclic peptides.<sup>40</sup>

The free energy landscape of cyclo-((D-Ala-L-Ala)<sub>4</sub>) is also funnelled (Figure 6). The lowest free energy structures in both solvents have some intramolecular hydrogen bonds. Cis isomers are destabilized to a lesser extent than in cyclo-((D-Ala-L-Ala)<sub>3</sub>) but are still at least 6.3 kcal mol<sup>−1</sup> above the lowest all-trans structures.

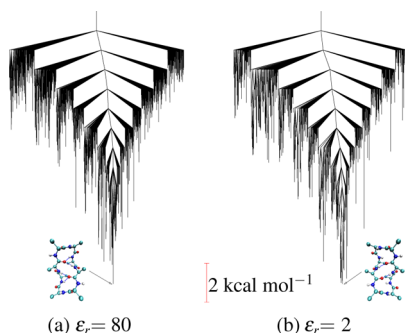
**3.2. Self-assembly.** Cyclic peptides with repeating D- and L-residues show a tendency for self-assembly, and we focus on these in this section. The global potential energy minimum for

the dimer of cyclo-((D-Ala-L-Ala)<sub>3</sub>) is a parallel nanotube with six intermolecular hydrogen bonds (Figure 7a). The antiparallel nanotube structure (Figure 7b) lies 9 kcal mol<sup>−1</sup> above this. The parallel structure is also the global free energy minimum, with the antiparallel structure less stable by 2.8 kcal mol<sup>−1</sup>. In both cases, the peptide groups are tilted away from the optimal alignment for nanotube growth, which is consistent with the dimer being the largest aggregate that has been observed in experiments on cyclic hexapeptides.<sup>2,13</sup> The free energy landscape of the cyclo-((D-Ala-L-Ala)<sub>3</sub>) dimer shows two funnels that correspond to parallel and antiparallel structures

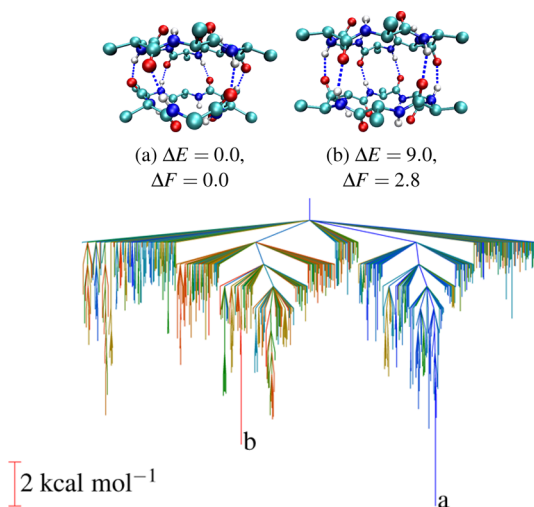




**Figure 5.** Disconnectivity graphs showing the free energy landscape of cyclo-((D-Ala-L-Ala)<sub>3</sub>) at 298 K in water (left) and nonpolar solvent (right).



**Figure 6.** Disconnectivity graphs showing the free energy landscape of cyclo-((D-Ala-L-Ala)<sub>4</sub>) at 298 K in water (left) and nonpolar solvent (right).



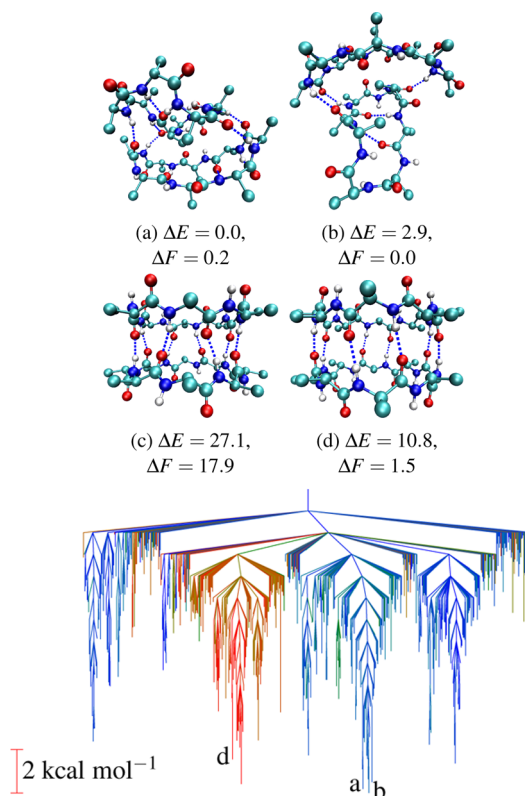
**Figure 7.** Low-lying minima of cyclo-((D-Ala-L-Ala)<sub>3</sub>) dimers in nonpolar solvent. Energies are in kcal mol<sup>−1</sup> and are relative to the global potential or free energy minimum. Also shown is a disconnectivity graph of the free energy landscape at 298 K. The 542 minima and 810 transition states accessible by transition states lower than 13 kcal mol<sup>−1</sup> from the global minimum are shown. Coloring is according to hydrogen bonding pattern from parallel (blue) to antiparallel (red).

(Figure 7). The barrier for the parallel to antiparallel conversion is 12.8 kcal mol<sup>−1</sup>.

For cyclo-((D-Ala-L-Ala)<sub>4</sub>), the global potential energy minimum does not resemble a nanotube and has both intermolecular and intramolecular hydrogen bonds (Figure

7a). Perfect antiparallel and parallel structures are 10.8 and 27.1 kcal mol<sup>−1</sup> less stable, respectively.

The disconnectivity graph of the cyclo-((D-Ala-L-Ala)<sub>4</sub>) dimer (Figure 8) comprises three main funnels that become



**Figure 8.** Low-lying minima of the cyclo-((D-Ala-L-Ala)<sub>4</sub>) dimer. Energies are in kcal mol<sup>−1</sup> and are relative to the global potential or free energy minimum. Also shown is a disconnectivity graph of the free energy landscape at 298 K. The 610 minima and 832 transition states accessible by transition states lower than 13 kcal mol<sup>−1</sup> from the global minimum are shown. Coloring is according to hydrogen bonding pattern from parallel (blue) to antiparallel (red).

mutually accessible about 12 kcal mol<sup>−1</sup> above the global minimum. The funnel containing the global minimum contains several structures with both inter- and intramolecular hydrogen bonds. The next deepest funnel contains the antiparallel nanotube and is competitive with the global minimum, with the lowest structures 0.4 kcal mol<sup>−1</sup> above the global free energy minimum. The third funnel contains parallel structures and has minima at least 1.4 kcal mol<sup>−1</sup> higher than the global free energy minimum. Note that several distorted structures are lower in energy than the perfect nanotubes in both of these funnels.

We have tested the performance of MD and group rotation moves by measuring the mean time to the first encounter of the global potential energy minimum over 50 basin-hopping runs (Table 5). For the cyclo-((D-Ala-L-Ala)<sub>3</sub>) dimer, both move sets require a similar number of quenches to locate the global minimum. However, for cyclo-((D-Ala-L-Ala)<sub>4</sub>), MD moves require substantially more quenches to locate the global minimum. For both cyclic peptide dimers, the computational overheads of the MD moves lead to each attempted move taking 2–3 times as much CPU time as the group rotation moves.

**Table 5. Mean Times to the First Encounter of the Global Minimum in 50 Basin-Hopping Runs on Cyclic Peptide Dimers**

molecule	move	mean first encounter time		
		energy evaluations	minimizations	CPU time (s)
cyclo-((D-Ala-L-Ala) <sub>3</sub> )	MD	$1.6 \times 10^6$	880	2700
	GR	$5.8 \times 10^5$	850	1000
cyclo-((D-Ala-L-Ala) <sub>4</sub> )	MD	$8.9 \times 10^6$	4100	27000
	GR	$1.9 \times 10^6$	2200	6200

#### 4. CONCLUSIONS

We have tested the performance of several move classes in basin-hopping optimization of cyclic peptides and their dimers. Simple Cartesian displacement moves perform poorly because of their tendency to invert chiral centers. MD moves are better, but the computational cost of a single basin-hopping step is significantly higher. The best performance, in terms of CPU time and the effectiveness of each basin-hopping step, is obtained with group rotation moves.

When searching the conformations of cyclic peptide dimers, a hybrid minimization strategy, where quenches are performed treating each molecule as a rigid body before an all-atom quench, is very effective. The combination of group rotation and rigid body moves is more efficient than MD moves, both in terms of the number of basin-hopping steps required to locate the global minimum and the computational cost of each step.

Analysis of the energy landscape of the cyclo-((D-Ala-L-Ala)<sub>3</sub>) dimer shows that the parallel structure is the global minimum, but also that this structure is not well-aligned to seed the growth of a nanotube. For the cyclo-((D-Ala-L-Ala)<sub>4</sub>) dimer there are several competing low-energy structures, and some of these could proceed to form nanotubes.

In the future, we will consider the formation of larger cyclic peptide nanotubes. From some preliminary calculations, we find that basin hopping searches using all atom models are practical for systems up to at least four cyclic peptides. For nanotubes that are too large to study with all-atom models, we will use a course-grained potential, such as the Martini model.<sup>61</sup> We have used an implicit solvation model in this study. Cyclic peptide nanotubes are large enough to contain small molecules, and it is possible that explicit solvent models will lead to different hydrogen bonding patterns. Basin-hopping calculations including a small number of explicit solvent molecules could be a computationally feasible way of capturing this effect.

#### AUTHOR INFORMATION

##### Corresponding Author

\*E-mail: r.l.johnston@bham.ac.uk.

##### Notes

The authors declare no competing financial interest.

#### ACKNOWLEDGMENTS

The authors thank Prof. David Wales for helpful discussions and Dr. Chris Whittleston for implementing some of the Monte Carlo move classes in GMIN. We acknowledge the Engineering and Physical Sciences Research Council, U.K. (EPSRC) for funding under Programme Grant EP/I001352/1. The computations described in this paper were performed using the University of Birmingham's BlueBEAR HPC service, which provides a High Performance Computing service to the

University's research community. See <http://www.birmingham.ac.uk/bear> for more details.

#### REFERENCES

- (1) Ghadiri, M. R.; Granja, J. R.; Milligan, R. A.; McRee, D. E.; Khazanovich, N. *Nature* **1993**, 366, 324–327.
- (2) Hartgerink, J. D.; Granja, J. R.; Milligan, R. A.; Ghadiri, M. R. *J. Am. Chem. Soc.* **1996**, 118, 43–50.
- (3) Kim, H. S.; Hartgerink, J. D.; Ghadiri, M. R. *J. Am. Chem. Soc.* **1998**, 120, 4417–4424.
- (4) Zhu, J.; Cheng, J.; Liao, Z.; Lai, Z.; Liu, B. *J. Comput. Aided Mol. Des.* **2008**, 22, 773–781.
- (5) Chapman, R.; Danial, M.; Koh, M. L.; Jolliffe, K. A.; Perrier, S. *Chem. Soc. Rev.* **2012**, 41, 6023–6041.
- (6) Fernandez-Lopez, S.; Kim, H.-S.; Choi, E. C.; Delgado, M.; Granja, J. R.; Khasanov, A.; Kraehenbuehl, K.; Long, G.; Weinberger, D. A.; Wilcoxon, K. M.; Ghadiri, M. R. *Nature* **2001**, 412, 452–455.
- (7) Dartois, V.; Sanchez-Quesada, J.; Cabezas, E.; Chi, E.; Dubbelde, C.; Dunn, C.; Granja, J.; Gritzen, C.; Weinberger, D.; Ghadiri, M. R.; Parr, T. R. *Antimicrob. Agents Chemother.* **2005**, 49, 3302–3310.
- (8) Ghadiri, M. R.; Granja, J. R.; Buehler, L. K. *Nature* **1994**, 369, 301–304.
- (9) Granja, J. R.; Ghadiri, M. R. *J. Am. Chem. Soc.* **1994**, 116, 10785–10786.
- (10) Richman, M.; Wilk, S.; Chemerovski, M.; Wärmländer, S. K. T. S.; Wahlström, A.; Gräslund, A.; Rahimpour, S. *J. Am. Chem. Soc.* **2013**, 135, 3474–3484.
- (11) ten Cate, M. G. J.; Severin, N.; Börner, H. G. *Macromolecules* **2006**, 39, 7831–7838.
- (12) Bertran, O.; Curcó, D.; Zanuy, D.; Alemán, C. *Faraday Discuss.* **2013**, 166, 59.
- (13) Sun, X.; Lorenzi, G. P. *Helv. Chim. Acta* **1994**, 77, 1520–1526.
- (14) Khazanovich, N.; Granja, J. R.; McRee, D. E.; Milligan, R. A.; Ghadiri, M. R. *J. Am. Chem. Soc.* **1994**, 116, 6011–6012.
- (15) Kobayashi, K.; Granja, J. R.; Ghadiri, M. R. *Angew. Chem., Int. Ed. Engl.* **1995**, 34, 95–98.
- (16) Khurana, E.; Nielsen, S. O.; Ensing, B.; Klein, M. L. *J. Phys. Chem. B* **2006**, 110, 18965–18972.
- (17) Vijayaraj, R.; Van Damme, S.; Bultinck, P.; Subramanian, V. *Phys. Chem. Chem. Phys.* **2012**, 14, 15135.
- (18) Vijayaraj, R.; Sundar Raman, S.; Mahesh Kumar, R.; Subramanian, V. *J. Phys. Chem. B* **2010**, 114, 16574–16583.
- (19) Vijayaraj, R.; Van Damme, S.; Bultinck, P.; Subramanian, V. *J. Phys. Chem. B* **2012**, 116, 9922–9933.
- (20) Comer, J.; Dehez, F.; Cai, W.; Chipot, C. *J. Phys. Chem. C* **2013**, 117, 26797–26803.
- (21) Vijayaraj, R.; Van Damme, S.; Bultinck, P.; Subramanian, V. *Phys. Chem. Chem. Phys.* **2013**, 15, 1260.
- (22) Khurana, E.; DeVane, R. H.; Kohlmeyer, A.; Klein, M. L. *Nano Lett.* **2008**, 8, 3626–3630.
- (23) Khalfa, A.; Treptow, W.; Maigret, B.; Tarek, M. *Chem. Phys.* **2009**, 358, 161–170.
- (24) Khalfa, A.; Tarek, M. *J. Phys. Chem. B* **2010**, 114, 2676–2684.
- (25) Bonnet, P.; Agrafiotis, D. K.; Zhu, F.; Martin, E. *J. Chem. Inf. Model.* **2009**, 49, 2242–2259.
- (26) Büttner, F.; Norgren, A. S.; Zhang, S.; Prabpai, S.; Kongsaree, P.; Arvidsson, P. I. *Chem.—Eur. J.* **2005**, 11, 6145–6158.
- (27) Labute, P. *J. Chem. Inf. Model.* **2010**, 50, 792–800.
- (28) Rayan, A.; Senderowitz, H.; Goldblum, A. *J. Mol. Graphics Modell.* **2004**, 22, 319–333.
- (29) Rezai, T.; Bock, J. E.; Zhou, M. V.; Kalyanaraman, C.; Lokey, R. S.; Jacobson, M. P. *J. Am. Chem. Soc.* **2006**, 128, 14073–14080.
- (30) Goldtzvik, Y.; Goldstein, M.; Benny Gerber, R. *Chem. Phys.* **2013**, 415, 168–172.
- (31) Dieterich, J. M.; Hartke, B. *Mol. Phys.* **2009**, 108, 279–291.
- (32) Dieterich, J. M.; Gerstel, U.; Schröder, J.-M.; Hartke, B. *J. Mol. Model.* **2011**, 17, 3195–3207.
- (33) Strodel, B.; Wales, D. J. *J. Chem. Theory Comput.* **2008**, 4, 657–672.

- (34) Strodel, B.; Lee, J. W. L.; Whittleston, C. S.; Wales, D. J. *J. Am. Chem. Soc.* **2010**, *132*, 13300–13312.
- (35) Wales, D. J.; Doye, J. P. K. *J. Phys. Chem. A* **1997**, *101*, 5111–5116.
- (36) Wales, D. J. *Mol. Phys.* **2002**, *100*, 3285–3305.
- (37) Wales, D. J. In *Energy Landscapes*; Wales, D. J., Ed.; Cambridge University Press: Cambridge, U.K., 2003; pp 397–409.
- (38) Wales, D. J. *Mol. Phys.* **2004**, *102*, 891–908.
- (39) Oakley, M. T.; Johnston, R. L. *J. Chem. Theory Comput.* **2013**, *9*, 650–657.
- (40) Oakley, M. T.; Oheix, E.; Peacock, A. F. A.; Johnston, R. L. *J. Phys. Chem. B* **2013**, *117*, 8122–8134.
- (41) Duan, Y.; Wu, C.; Chowdhury, S.; Lee, M. C.; Xiong, G.; Zhang, W.; Yang, R.; Cieplak, P.; Luo, R.; Lee, T.; Caldwell, J.; Wang, J.; Kollman, P. J. *Comput. Chem.* **2003**, *24*, 1999–2012.
- (42) Lee, M. C.; Duan, Y. *Proteins* **2004**, *55*, 620–634.
- (43) Onufriev, A.; Bashford, D.; Case, D. A. *Proteins* **2004**, *55*, 383–394.
- (44) Wales, D. J. GMIN: A program for finding global minima and calculating thermodynamic properties from basin-sampling. <http://www-wales.ch.cam.ac.uk/GMIN/> (accessed Oct. 11, 2013).
- (45) Somani, S.; Wales, D. J. *J. Chem. Phys.* **2013**, *139*, 121909.
- (46) Mochizuki, K.; Whittleston, C. S.; Somani, S.; Kusumaatmaja, H.; Wales, D. J. *Phys. Chem. Chem. Phys.* **2014**, *16*, 2842–2853.
- (47) Kusumaatmaja, H.; Whittleston, C. S.; Wales, D. J. *J. Chem. Theory Comput.* **2012**, *8*, S159–S165.
- (48) Wales, D. J. PATHSAMPLE: A program for refining and analysing kinetic transition networks. <http://www-wales.ch.cam.ac.uk/PATHSAMPLE/> (accessed Oct. 11, 2013).
- (49) Wales, D. J. OPTIM: A program for characterizing stationary points and reaction pathways. <http://www-wales.ch.cam.ac.uk/OPTIM/> (accessed Oct. 11, 2013).
- (50) Trygubenko, S. A.; Wales, D. J. *J. Chem. Phys.* **2004**, *120*, 2082–2094.
- (51) Henkelman, G.; Jónsson, H. *J. Chem. Phys.* **1999**, *111*, 7010–7022.
- (52) Munro, L. J.; Wales, D. J. *Phys. Rev. B* **1999**, *59*, 3969–3980.
- (53) Kumeda, Y.; Munro, L. J.; Wales, D. J. *Chem. Phys. Lett.* **2001**, *341*, 185–194.
- (54) Carr, J. M.; Trygubenko, S. A.; Wales, D. J. *J. Chem. Phys.* **2005**, *122*, 234903.
- (55) Strodel, B.; Whittleston, C. S.; Wales, D. J. *J. Am. Chem. Soc.* **2007**, *129*, 16005–16014.
- (56) Becker, O. M.; Karplus, M. *J. Chem. Phys.* **1997**, *106*, 1495–1517.
- (57) Wales, D. J.; Miller, M. A.; Walsh, T. R. *Nature* **1998**, *394*, 758–760.
- (58) Krivov, S. V.; Karplus, M. *J. Chem. Phys.* **2002**, *117*, 10894–10903.
- (59) Evans, D. A.; Wales, D. J. *J. Chem. Phys.* **2003**, *118*, 3891–3897.
- (60) Kabsch, W.; Sander, C. *Biopolymers* **1983**, *22*, 2577–2637.
- (61) Marrink, S. J.; Tieleman, D. P. *Chem. Soc. Rev.* **2013**, *42*, 6801–6822.