Hydration layer dynamics and association mechanisms of food and antifreeze proteins
A Molecular Dynamics and Transition Path Sampling study
Brotzakis, Z.F.

Publication date
2017
Document Version
Other version
License
Other

Citation for published version (APA):

General rights
It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations
If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: https://uba.uva.nl/en/contact, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.
Chapter 5

Stability and growth mechanism of self-assembling anti-freeze cyclic peptides

Cyclic peptides (CPs) that self-assemble in ice-binding nanotubes are great candidates for use as anti-freeze proteins. Based on cyclic peptide sequence, cyclo-[(L-LYS-D-ALA-L-LEU-D-ALA)\textsubscript{2}], which can stack into nanotubes, we propose an anti-freeze cyclic peptide (AFCP) sequence, cyclo-[(L-LYS-D-ALA)\textsubscript{2}-(L-THR-D-ALA)\textsubscript{2}] which contains THR-ALA-THR ice binding motifs. Using molecular dynamics simulations we investigate the stability of cyclic peptides and their growth mechanism. We find that dimers of the AFCP sequence dissociate more frequently and are less stable than dimers of the original CP sequence, while nanotubes consisting of more than two peptides are stable. This sudden increase in stability of nanotubes of the AFCP sequence may be explained by the formation of H-bonds between Threonine side-chains. The Threonine distances in the ice-binding motifs are similar to those in the ant-freeze protein of Christoneura fumiferana, suggesting good ice lattice matching, and a potential for depression of the freezing point. In addition, we investigated the nanotube growth process, i.e. the association/dissociation of a single CP to an existing AFCP nanotube, by Transition Path Sampling. We found a general dock-lock mechanism, in which a single CP first docks loosely before locking into place. Moreover, we identified several qualitatively different mechanisms for dissociation, involving different meta-stable intermediates, including a state in which the peptide was misfolded inside the hydrophobic core of the tube. We also find evidence for a mechanism involving non-specific association followed by 1D diffusion. Under most conditions, this will be the dominant pathway. The results yield insight in the mechanisms of peptide assembly, and might lead to improved design of self-assembling anti-freeze proteins.
5.1 Introduction

Anti-freeze proteins (AFPs) adsorb to the surface of ice crystals and prevent growth [1]. This lowers the freezing point temperature of the solution below its melting point, enabling the survival of many organisms living in subzero temperature environments. AFPs have been used in a variety of applications, such as in super-cooling organ preservation [2] by inhibiting ice-crystal growth and therefore preventing tissue damage, as well as in industry as texture enhancing agents [3].

Although all AFPs share anti-freeze activity, they show a significant diversity in this activity, in molecular weight, amino-acid sequence and structure. Many hyperactive AFPs have a $\beta$-helical rich structure with repetitive sequence motifs [4–6]. For example the spruce budworm AFP shows a $\beta$-sheet region where threonines (THR) are organized in a regular array of THR-$X_{aa}$-THR (TXT) motifs, where $X_{aa}$ can be any amino-acid. Those motifs match both the basal and prism crystal plane of ice and therefore can bind to these planes [1]. In a molecular simulation study of the spruce budworm AFP (isoform CfAFP-501), Zhou et al.[7], showed that the THR O-O distances in the TXT motif of the spruce budworm anti-freeze protein as well as the THR O-O distances of neighbouring coils are similar to the ice prism plane O-O distances.

Hyperactive anti-freeze proteins that have a $\beta$-helical structure and a rich TXT motif can be mimicked using synthetic polypeptides. For instance, one proposed option for such mimicking is a nanotube made of cyclic peptides (CP) containing the TXT ice binding motif, as shown in Fig. 5.2.1 top left. Experiments [8, 9] and simulations [10, 11] indicate that cyclic peptides of alternating L and D amino acids self assemble into nanotubes under proper conditions. These nanotubes consist of an anti-parallel cyclic $\beta$-sheet hollow structure, stabilized by backbone hydrogen bonds between adjacent CPs and by side chain interactions [11] (see Fig. 5.2.1bottom). Vijayaraj et al. [10, 11] investigated the number of required CPs for a stable nanotube, and showed the importance of alanine (ALA) amino acids in the stability and fluctuations of the nanotube. Introducing an ice binding motif in such cyclic peptides would lead to self-assembled nanotubes with great potential as an anti-freeze agent. Here, we employ molecular simulations to investigate the stability and fluctuations of anti-freeze cyclic peptide nanotubes (denoted AFCP nanotubes) comprising stacks of AFCP sequence cyclo-[(L-LYS-D-ALA)$_2$-(L-THR-D-ALA)$_2$] and of the experimentally self assembled nanotube [9] comprising the original CP sequence cyclo-[(L-LYS-D-ALA-L-LEU-D-ALA)$_2$], (denoted original CP nanotube). After we have established that both the AFCP and the original CP nanotube are stable in solution, we continue with the question of the formation mechanism. The formation mechanism of the AFCP nanotube from dilute solution is poorly understood, but is believed to occur via association.
of CPs, nucleation and growth. In the latter stage, growth is dominated by association (and dissociation) events to the end of a growing nanotube. This step can be seen as a rare event, and needs to be addressed with rare event simulation techniques. Here we employ the transition path sampling (TPS) technique, which allows us to harvest an ensemble of unbiased rare transition paths that give valuable information about the association process. Analyzing this ensemble gives insight in the different mechanisms. While we find several mechanisms of association/dissociation, all involve an intermediate docked state. Furthermore, we find evidence for a growth mode involving nonspecific binding to the nanotube, followed by a random walk of the CP along the nanotube until it finds one of the endpoints. This mechanism might have general implications for growing fibril structures in general.

The chapter is organized as follows. In section 5.2 we describe the used methods. In section 5.3 present and discuss the results. We end with concluding remarks.

5.2 Methods

5.2.1 System setup

Construction of the cyclic peptides and nanotubes

We construct two initial CP conformations, one for the original CP cyclo-[(L-LYS-D-ALA-L-LEU-D-ALA)_2], and the other for the AFCP, cyclo-[(L-LYS-D-ALA)_2-(L-THR-D-ALA)_2]. An initial linear structure for constructing these cyclic peptides was created using AmberTools [12]. The linear peptide was turned into a cyclic peptide by constructing a bond between the first and last residue and minimizing the structure using steepest descent in GROMACS[13].The D-conformational orientation of the D-Alanine amino acids was obtained by flipping the L-Alanine amino acids in PyMOL [14] in order to obtain the original and AFCP sequences (see also Table 5.A.3 and Table 5.A.4). The geometry of each single CP unit was energy minimized using steepest descent while the Ramachandran dihedral angles ($\Phi, \Psi$) of all amino acids were restrained to their average anti-parallel $\beta$-sheet values as given in Table 5.2.1, in order to guarantee the planarity of the ring. Since the side chains of the cyclic peptides play a crucial role in the stability of the nanotubes[11], also the side chains of the rings were relaxed in a 10 ns NPT simulations.

The final geometry of the CPs was used to build various nanotubes of different size (e.g CPNT2, CPNT3, for stacks of two or three CP units). The CPs were stacked in an anti-parallel fashion, meaning that two adjacent CPs had opposite chain orientations. During the model building, necessary care was taken to align amino acids (L and D)
in the adjacent chains as shown in Fig 5.2.1 bottom and detailed Table 5.A.3 and Table 5.A.4 of the appendix).

5.2.2 Molecular dynamics

All energy minimization and molecular dynamics simulations were performed with the Gromacs 4.5.4 package[13]. Molecular dynamics simulations by Vijayaraj and Khurana [10, 15], suggest that using the amberf99sb [16] stabilizes the cyclic peptide nanotubes by formation of intramolecular hydrogen bonds. Therefore we choose here the same protein force-field, coupled with TIP3P water [17]. The protonation state of the amino acids corresponds to pH greater than 11 in order to compare to the experiments. All bonds were constrained with the Lincs algorithm. A cutoff of 0.8 nm was used for the non-bonded Lennard-Jones interactions. The Particle Mesh Ewald method was used to calculate the electrostatic interactions with a Fourier spacing of 0.12 nm and a 0.8 nm cutoff for the short range electrostatic interactions. neighbor
Amino acid & $\Phi_{\text{aver}}$ & $\sigma_\Phi$ & $\Psi_{\text{aver}}$ & $\sigma_\Psi$
\hline
L-ALA & -130.2 & 21.4 & 143.8 & 14.6
L-LEU & -115.2 & 15.8 & 131.8 & 13.5
L-LYS & -118.9 & 17.5 & 134.2 & 15.3
L-THR & -123.9 & 14.0 & 138.1 & 14.3
D-ALA & 130.2 & - & -143.8 & -
\hline

Table 5.2.1: Average values for the dihedral angles $\Phi$, $\Psi$ and their estimated standard deviations $\sigma_\Phi, \sigma_\Psi$ for anti-parallel $\beta$-strands [20].

Lists were updated every 10 fs with a cutoff of 0.8 nm and the time step was 2 fs [16]. The leap-frog algorithm was used for integrating Newton’s equations of motion. In the NPT simulations the v-rescale thermostat[18] with a coupling time constant of 0.2 ps controlled the temperature, while the Parrinello-Rahman barostat[19] with a coupling time constant of 1.0 ps kept the pressure constant. After energy minimization the CPs were solvated with a dodecahedron box of TIP3P [17] water molecules extending 1.7 nm away from the solute atoms. The energy minimization, equilibration, and production runs of various nanotubes (CPNT) were carried out in different stages: (i) solvent equilibration for 10 ps by restraining the heavy atoms of the CPNT systems, (ii) a 1000 ps total system equilibration run by restraining the $C_\alpha$ atoms of the CPNTs and (iii) production MD for 100 ns with a 2 fs time step. From the production run, a frame for every 2 ps was collected for trajectory analysis. All the simulations were carried out in the NPT ensemble, except for the high temperature simulations and the pressure was maintained at 1 bar.

Table 5.2.2 gives information on the number of CP units in each CPNT system, total number of residues, water molecules, and initial volume of the periodic box.

5.2.3 Transition Path Sampling

The flexible one-way TPS algorithm

Transition Path Sampling [21, 22] (TPS) allows efficient sampling of infrequent transitions between two predefined stable states by harvesting an ensemble of trajectories that lead over a high free energy barrier, connecting the two stable states. Starting from an initial reactive path connecting the two stable states, TPS performs a Markov Chain Monte Carlo random walk in trajectory space by selecting a time frame of the current trajectory, changing the momenta slightly and shooting off a new trial trajec-
Table 5.2.2: Original CP sequence: Composition of systems containing CPNTs of sequence cyclo-[(L-LYS-D-ALA-L-LEU- D-ALA)\textsubscript{2}]. AFCP sequence: Composition of systems containing nanotubes of AFCP sequence cyclo-[(L-LYS-D-ALA)\textsubscript{2}-(L- THR-D-ALA)\textsubscript{2}].

<table>
<thead>
<tr>
<th>System</th>
<th>#CP</th>
<th>#AA</th>
<th>#H\textsubscript{2}O</th>
<th>#atoms</th>
<th>V (nm\textsuperscript{3})</th>
</tr>
</thead>
<tbody>
<tr>
<td>CPNT2</td>
<td>2</td>
<td>16</td>
<td>4093</td>
<td>12519</td>
<td>127.48</td>
</tr>
<tr>
<td>CPNT3</td>
<td>3</td>
<td>24</td>
<td>4519</td>
<td>13917</td>
<td>139.09</td>
</tr>
<tr>
<td>CPNT4</td>
<td>4</td>
<td>32</td>
<td>4847</td>
<td>15021</td>
<td>151.47</td>
</tr>
<tr>
<td>CPNT5</td>
<td>5</td>
<td>40</td>
<td>5549</td>
<td>17247</td>
<td>174.34</td>
</tr>
<tr>
<td>CPNT6</td>
<td>6</td>
<td>48</td>
<td>6280</td>
<td>19560</td>
<td>197.25</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>System</th>
<th>#CP</th>
<th>#AA</th>
<th>#H\textsubscript{2}O</th>
<th>#atoms</th>
<th>V (nm\textsuperscript{3})</th>
</tr>
</thead>
<tbody>
<tr>
<td>CPNT2</td>
<td>2</td>
<td>16</td>
<td>4035</td>
<td>12325</td>
<td>122.57</td>
</tr>
<tr>
<td>CPNT3</td>
<td>3</td>
<td>24</td>
<td>4100</td>
<td>12630</td>
<td>131.07</td>
</tr>
<tr>
<td>CPNT4</td>
<td>4</td>
<td>32</td>
<td>4662</td>
<td>14426</td>
<td>144.55</td>
</tr>
<tr>
<td>CPNT5</td>
<td>5</td>
<td>40</td>
<td>5417</td>
<td>16801</td>
<td>169.58</td>
</tr>
<tr>
<td>CPNT6</td>
<td>6</td>
<td>48</td>
<td>6267</td>
<td>19317</td>
<td>195.47</td>
</tr>
</tbody>
</table>

In this work we use the more efficient one way flexible shooting algorithm [22, 23]. First, a time frame $t_{sel}$ is uniformly randomly selected from the current (old) path containing $N_o$ frames and the shooting direction (forward or backward) is randomly chosen with equal probability. A new partial trial trajectory of length $\tau_{part}$ is generated by molecular dynamics until the initial state (in case of backward shot) or the final state (in case of a forward shot) is reached. Because of the use of the stochastic v-rescale thermostat[18] the new part of the trial trajectory will diverge from the old path. When the generated path ends in the wrong state, the entire move is rejected. The resulting
new partial path is glued to the complementary part of the previous (old) path to yield the new trial path with a length $N_n$. If performing a forward move, $N_n = \tau_{sel} + \tau_{part}$, while for a backward move, $N_n = (N_0 - \tau_{sel}) + \tau_{part}$. Next, to maintain detailed balance the algorithm accepts the trial path according to $P_{acc} = \min(1, N_{(o)}/N_{(n)})$ [24]. In order to prevent wasting computation time in very long paths connecting the two states, the maximum allowed path length $N_{\text{max}} = N_{(o)}/\xi$ is computed in advance, where $\xi$ is a random number in the interval [0, 1]. This TPS algorithm has been previously used in other protein systems [25, 26].

**Mechanistic analysis via the path density**

For more insight in the mechanism of the transitions we project the TPS ensemble on two dimensional path plots. We construct the path density histograms by choosing two order parameters (OP), and creating a 2D grid initialized to zero. Each path in the ensemble is then projected on that grid. A bin in the 2D histogram is incremented with the weight of that path, if the path visited that bin at least once. The path density plots shows the existence of correlation between particular order parameters in the mechanisms. It is more informative than a configuration projection which is usually overwhelmed by intermediate states.

## 5.3 Results and Discussion

### 5.3.1 Nanotube equilibrium properties

**Stability**

We performed a series of 100 ns MD simulations for each of the systems denoted in Table 5.2.2. We define a system as being stable if it maintains an anti-parallel cyclic $\beta$-sheet tube-like structure throughout the course of a 100 ns MD simulations. An unstable system is, on the other hand, a system that deviates considerably from its tube-like structure during the MD simulation. While this definition is rather heuristic, it serves our purpose here to identify relative nanotube stability as a function of number of CPs. Analysis of the MD trajectories (see also Table 5.A.1 and 5.A.2 of the appendix) shows that most of the CPNT systems maintain their anti-parallel $\beta$-sheet tube-like structure throughout the simulation. The exception is the CPNT2 system of the AFCP sequence, which exhibited strong deviations from a tube-like structure, leading in most cases to dissociation of the two CPs. The CPNT2 system of the AFCP sequence is therefore clearly less stable than the corresponding CPNT2 system of the original CP sequence. Figure 5.2.2 shows the root mean square fluctuations (RMSF) of
Figure 5.2.2: (a) RMSF of backbone atoms for the original CP sequence. (b) RMSF of backbone atoms for the AFCP sequence. (b). Number of side chain hydrogen bonds ($N_{HB}$) between adjacent peptides.

Backbone atoms within all CPNT systems of both sequences. The RMSF for CPNT2 is higher than those of longer CPNTs, as the carbonyl and amide groups of the CPs within the CPNT2 system are exposed to solvent molecules. The RMSF of the AFCP sequence CPNT2 is higher compared to the original CP sequence. The higher RMSF is a consequence of strong interactions of the hydroxyl groups of THR side-chains with solvent molecules. As strongly fluctuating backbone atoms perturb the specific backbone-backbone H-bonds of the AFCP sequence CPNT2, the latter dissociates easier than the original CP sequence CPNT2.

Figure 5.2.2a,b shows that CPNTs containing three or more CPs of either sequence have almost identical RMSFs. The sudden increase in stability of CPNTs of the AFCP sequence between CPNT2 and CPNT3 may be explained by the average number of H-bonds between side chains in the CPNTs (Figure 5.2.2c), which is higher for the AFCP sequence than for the original CP sequence. This can be explained by the formation of H-bonds between hydroxyl groups in adjacent side-chains, as predicted by Vijayaraj et al. [11]. Indeed, when increasing the AFCP nanotube size from CPNT2 to CPNT3, we hypothesize that the hydroxyl groups of threonine side-chains interact less
strongly with water molecules while forming more side chain intramolecular hydrogen bonds, thus decreasing the RMSFs of AFCP sequence CPNTs similar to the original CP sequence RMSF levels.

CPNTs consisting of more than two CPs of both sequences form stable nanotube conformations. Vijayaraj et al. [10] reported, based on RMSD calculations of the whole tube, that larger oligomers (>CPNT3) show less fluctuation and more structural stability. The RMSD of the CPNT3 system should be slightly higher than that of larger nanotubes since two thirds of the system is composed of stronger fluctuating termini. Reanalyzing the data in Ref. [10], the RMSD decreases strongly from CPNT2 to CPNT3 and only slightly between CPNT3 and CPNT4, which suggest that the CPNT3 system, in fact, is stable. Finally, one of the outermost CPs of a CPNT6 system of the original CP sequence left its initial configuration in the course of a 100 ns MD run (Table 5.A.1), suggesting the original CP system might be less stable than the AFCP CPNT6 system, possibly due to additional stabilization of the CPNT by H-bonds between THR side chains, as was predicted by Vijayaraj et al. [10].

**Ice binding properties of AFCP nanotubes**

Figure 5.3.1 shows a) the intramolecular and b) the inter-molecular $C_{\alpha} - C_{\alpha}$ distances of THR residues. For the stable CPNTs (>CPNT2) the average value of the intramolecular and the intermolecular THR $C_{\alpha} - C_{\alpha}$ distances are, respectively, 7.036 and 4.875 Å. These distances are similar to the values reported by Li et al. [27] (respectively, 6.75 and 4.75 Å) and fall within the ranges reported by Zou et al. [7] (5.90-7.46 Å and 3.95-4.96 Å). The distances are very close to the ice Ih unit cell dimensions $a = 4.518$ Å and $c = 7.356$ Å [28] and as a result, lead to good lattice matching. Therefore, we expect that the AFCP sequence nanotubes exhibit similar or the same ice-binding ability as
Table 5.3.1: Stable state definitions as a function of number of backbone hydrogen bonds, $C_\alpha$–RMSD of CP6 and minimum distance $d_{\text{min}}$.

<table>
<thead>
<tr>
<th>State</th>
<th>RMSD (nm) $\text{min}$</th>
<th>$d_{\text{min}}$ (nm) $\text{min}$</th>
<th>H-bonds $\text{min}$</th>
<th>$\text{max}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>B</td>
<td>0</td>
<td>0.03</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>U</td>
<td>-</td>
<td>-</td>
<td>1</td>
<td>$\infty$</td>
</tr>
</tbody>
</table>

CfAFP-501 spruce budworm isoform and are able to depress the freezing point.

### 5.3.2 Mechanism of self-assembly by Transition Path Sampling

#### The initial path

We applied TPS to study the self-assembly process of the AFCP nanotube, in particular, the association step of a single CP to a growing nanotube, here chosen as the CPNT6. However, as the MD trajectories are time reversible we can also instead focus on the dissociation of a single CP from a stable nanotube. This dissociation transition occurs between two stable states, the bound state B and the unbound state U. The unbound state U consists of a CPNT5 and single CP, and can be simply defined by a minimum distance larger than 1 nm. In the bound state B all of the six CPs within the CPNT6 are part of an anti-parallel cyclic β-sheet nanotube structure. In this state the $C_\alpha$ RMSD of the outermost peptide CP6 from its natively bound state is less than 0.03 nm, and eight backbone hydrogen bonds are present between the outermost peptide and the rest of the nanotube. The stable state definitions for U and B are given in Table 5.3.1.

TPS requires an initial path between the two stable states, which for the association process is difficult to obtain with straightforward MD simulations, and might take many microseconds. Indeed the unbinding free energy difference of removing the top CP from a different hexamer sequence (cyclo-[D-ALA-L-ALA]$_4$) was found in the order of $\approx 7$ kcal/mol, giving rise to a microsecond timescale for the association [10]. While we indeed did not observe a spontaneous dissociation in the AFCP CPNT6 system in the stability MD trajectories, (Table 5.A.2), such a dissociation did occur in a high temperature (450 K) simulation. From this we constructed an equilibrated initial path at 300 K by performing a committor analysis, which consists of shooting off several room temperature trajectories for selected frames on the 450 K pathway to find a frame where the probability of returning to the initial state is similar to ending in the final state. The initial path is constructed by gluing two partial trajectories starting from the
Figure 5.3.2: The initial pathway for the TPS, as a function of a) backbone hydrogen bonds between CP6 and CP5 (N\textsubscript{HB}), b) CP6 RMSD, and c) minimum distance (d\textsubscript{min}) between peptide CP6 and the remainder of the tube.

The TPS simulations consisted of 1084 trial one-way flexible shooting moves, which resulted in 325 accepted paths with an overall acceptance ratio of 0.30. The path tree is plotted in figure 5.3.3, and illustrates the decorrelation between the successive accepted shooting moves. Starting from the top each horizontal line indicates an accepted shooting attempt. Red lines indicate forward shots, green lines indicate backward shots. The thin black vertical lines indicate the shooting point location on the previous path. Each accepted new path thus consists of the newly formed green/red partial path, together with the complementary part of the previous path. A measure of decorrelation on the transition path ensemble is the number of decorrelated paths. A path is considered decorrelated when it shares no time slice with the previous decorrelated path. In our simulation we obtained 20 completely decorrelated paths. Inspection of the tree reveals a distribution of different path lengths with shorter and longer reactive pathways. Figure 5.3.4 shows this reactive path length distribution. The path length distribution
Figure 5.3.3: Transition path ensemble tree representation. In green are depicted the backward partial paths and in red the forward partial paths.

is peaked around 2 ns and has an average path length of 8.96 ns. The distribution is roughly Poissonian, with a long tail up to 40 ns. In addition, a second peak visible around 20 ns indicates that there are multiple mechanisms in this transitions.
The distribution is highlighted in red, and is split in the underlying distributions according to their transition mechanism: BiAU (green), BiAiBU (blue), BiAiCU (purple), BiAiBiCU (cyan), BiAiCiAU (grey), and BiAiCiBU (black).

5.3.3 Analysis of the path ensemble

Inspection of the transition path ensemble revealed three different dissociation mechanisms labeled I, II, and III, with three on-pathway intermediate states, denoted iB, iA, and iC (see figure 5.3.5). Intermediate state iA is partly dissociated, but has still a few bound state backbone hydrogen bonds between CP6 and the nanotube intact, and can thus be seen as a ‘docked state’. Intermediate state iB is a misfolded state defined by a THR side chain of CP6 located inside the nanotube. Intermediate state iC is characterized by a CP6 peptide that has all bound state hydrogen bonds broken, but is still associated to the side of the tube. From the distribution of path length (figure 5.3.4) we can deduce that the average path length of each type of transition varies. The paths following mechanism III involving transitions BiAiCU (purple curve) are on average 10.76 ns, and are the origin for the long tail in the distribution. The much faster direct dissociation transition II via intermediate iA (BiAU) is much faster has an average length 3.28 ns (green curve). The dissociation transition I visiting on-pathway intermediate iB (BiAI BU) takes average 19 ns (blue curve). In addition, we identified several paths mixing these mechanisms such as BiAiBiCU and BiAiCiBU (see Figure 5.3.4).

To further understand the association/dissociation transition, we analyze the mechanisms of three selected pathways 3, 35, and 23 respectively corresponding to transition I, II and III. All dissociation/association pathways visit the docked intermediate state iA. Figure 5.3.5 shows that in intermediate iA CP6 has lost its planar $\beta$ sheet conformation and has instead a V-shape, with only a fraction of the backbone hydrogen bonds
Figure 5.3.5: Mechanisms found in the Transition Path ensemble. The three corresponding mechanisms are: mechanism I (B ⇄ iA ⇄ iB ⇄ U), mechanism II (B ⇄ iA ⇄ U), mechanism III (B ⇄ iA → iC → U).

to CP5 being intact.

Fig. 5.3.6a show the path densities plotted as a function of several order parameters, the RMSD of CP6 with respect to the correctly bound nanotube configuration, the minimum distance CP6 to the nanotube, and the minimum distance between the CP6 and the center of mass of CP5. Here, the RMSD values around 0.3 correspond to the docked intermediate state iA. The three main mechanisms are visible as broad channels in the path densities.

Mechanism of transition I (path 3) Mechanism I consists of transitions B ⇄ iA ⇄ iB ⇄ U. The shape of the meta-stable intermediate iB resembles intermediate state iA, but with the side chain of THR-45 no longer H-bonded to the backbone of CP5, but moved towards the interior of the tube. Fig. 5.3.7 bottom left shows that there are almost no hydrogen bonds formed between THR-45 side chains and the remainder
of the nanotube in the course of path 3. Fig. 5.3.7bottom right displays, moreover, that THR-45 has fewer H-bonds with water molecules than THR-43. The threonine side-chain is held in place by a hydrophobic interaction between its methyl group and the interior of the nanotube. Threonine entry into the nanotube is also visible in the path density as a function of the minimum distance of the CP6 side chains and the center of mass of CP5 backbone (Fig. 5.3.6b). Clearly, when the minimum distance is around 0.1 nm, the system, visits the intermediate state iB. In addition THR-45 may be stabilized by the occasional formation of hydrogen bonds between its hydroxyl group and the remainder of the tube (Fig. 5.3.7 bottom left) or water molecules at the top of the tube (Fig. 5.3.7bottom right). As shown in Fig. 5.3.4, transition I is rare compared to transitions II and III.

**Mechanism of transition II (path 35)** Mechanism II consists of transitions B ⇔ iA ⇔ U. Fig. 5.3.8 shows the transition through intermediate iA observed in path 35. Here, several intact backbone-backbone hydrogen bonds are broken at 4.5 ns, while the minimum distance increases and peptide CP6 dissociates without going through another intermediate state. Here the only hydrogen bonds formed between the CP6 and the tube are the backbone ones. Indeed, the red and black line of Fig. 5.3.8a coincide.
Mechanism of transition III (path 23) Mechanism III consists of transitions B ⇔ iA ⇔ iC ⇔ U. Meta-stable intermediate iC does not resemble intermediate iA and iB. Fig. 5.3.9a shows that all of the unique backbone-backbone CP5-CP6 hydrogen bonds are broken in the last part of path 23 where the system is in intermediate iC (red line). The H-bonds between CP6 and the nanotube (black line) are, on the contrary, frequently formed and broken, indicating that intermediate iC is probably not only stabilized by H-bonds, but also by hydrophobic interactions between CP6 and the nanotube. The frequent formation and breaking of these interactions suggests CP6 may slide along the tube. Indeed the RMSD in Fig. 5.3.9b increases, whereas the minimum distance remains small, indicating the peptide remains associated to the nanotube.

Fig. 5.3.7, Fig. 5.3.9 and Fig. 5.3.8 show all dissociation processes are initiated by breaking some of the stable inter-molecular backbone-backbone H-bonds. Once all β-sheet hydrogen bonds between backbone-atoms are broken, H-bonds and hydrophobic interactions between CP6 and the nanotube may exist. Once those are broken CP6...
fully separates from the nanotube. The rate limiting step of the dissociation of CPs is the breaking of all backbone-backbone H-bonds as suggested by Vijayaraj et al. [11]. During this process the system can get trapped in the meta-stable intermediate iB, due to hydrophobic interactions between the methyl group of a THR and the interior of the nanotube. The dissociation is realized by the disappearance of all side-chain and backbone interactions between CP6 and the nanotube. The prediction of Vijayaraj et [11] that the process is followed by the annihilation of various side-chain side-chain interactions is, therefore, too simplified. Furthermore, we found that during dissociation the system may be trapped in intermediate iC, in which CP6 is associated to the side of the tube. From the perspective of association, all TPS pathways are examples of a dock-lock mechanism that has also been identified in protein aggregation studies [29–32].

The existence of intermediate iC also explains why CPNTs align in a linear fashion within their crystal structure. Due to the fact that a single CP interacts strongly with the side of a nanotube, two nanotubes should also interact strongly with one another. As a result solubility of the nanotube is low and crystals should early form as observed in experiment [33]. This is because for nanotube solvation all interactions between the nanotubes have to be broken.

5.3.4 Comparison of the association time scales

From the view point of the association process, a CP might therefore first bind non-specifically to the tube, followed by a random walk along the nanotube before it docks.
Figure 5.3.9: Time evolution of the order parameters of path 23. a) Number \( N_{HB}(CP5/remainder - CP6)_{bb} \) of H-bonds between backbone atoms of CP6 and CP5 (red line) and H-bonds between CP6 and the remainder of the tube (black). b) RMSD in nm of CP6 (green) and minimum distance \( d_{min} \) in nm between CP6 and the remainder of the tube (red).

to end of the nanotube, and locks into place. Under certain conditions this one-dimensional diffusion along the nanotube might be faster than a random search for the end of the growing tube in a 3-dimensional volume. This can be argued as follows. The effective diffusion limited (Smoluchowski) association rate for a particle attaching to a specific point with a contact radius \( \sigma \)

\[
k_{on} = \frac{4\pi D \sigma}{V},
\]  

(5.1)

where \( D \) is the translational diffusion constant, and \( V \) is the volume (determined by the concentration of growing ends). The timescale connected to this rate is simply its inverse \( \tau_{on} = k_{on}^{-1} \). The association rate constant of a CP associating to a nanotube of length \( L \) is instead is given by [34]

\[
k'_{on}^{tube} = \frac{4\pi DL}{V \ln(2L/\sigma)}
\]

(5.2)

The diffusion timescale to diffuse along the nanotube, in a quasi 1D random walk, is of the order of

\[
\tau_1 = \frac{L^2}{3D_1}.
\]

(5.3)

where \( D_1 \) is the translational diffusion constant along the nanotube. This leads to a total association timescale via pathway III of the order of

\[
\tau'_{on} = \frac{V \ln(2L/\sigma)}{4\pi DL} + \frac{L^2}{3D_1}
\]

(5.4)
This timescale can always be made smaller than the timescale of the direct mechanism, \( \tau_{\text{on}} \) because the first term is dominant at low concentration, and can be made smaller by considering longer \( L \).

To identify the crossover, we equate the two times scales \( \tau'_{\text{on}} = \tau_{\text{on}} \):

\[
\frac{V \ln(2L/\sigma)}{4\pi DL} + \frac{L^2}{3D_1} = \frac{V}{4\pi D\sigma}
\]

(5.5)

\[
\frac{\sigma \ln(2L/\sigma)}{L} + \frac{L^2 4\pi D\sigma}{3D_1 V} = 1
\]

(5.6)

Clearly, the left hand side can be made arbitrarily small by increasing the volume \( V \) and/or increasing the length \( L \). For these condition, the indirect mechanism of association is preferred.

### 5.4 Conclusion

In summary, we have investigated the stability of nanotube structures formed by CPs with sequence cyclo-[(L-LYS-D-ALA-L-LEU-D-ALA)\(_2\)] and sequence cyclo-[(L-LYS-D-ALA)\(_2\)-(L-THR-D-ALA)\(_2\)] using molecular dynamics simulations. The AFCP sequence contains TXT ice binding motifs as they are found in anti-freeze protein CfAFP-501 which is the most active anti-freeze protein discovered thus far [7, 27].

MD simulations indicated that the AFCP sequence CPNT2 system is less stable than a original CP sequence CPNT2, because THR side-chains in the AFCP sequence CPNT2 system interact strongly with solvent water molecules, yielding strong fluctuations of backbone atoms and, therefore, perturbation of the inter-molecular H-bonds between the CP units. CPNTs of more than two peptides of both sequences maintained their stable anti-parallel \( \beta \)-sheet tube-like structure. The sudden increase in stability of CPNTs of the AFCP sequence (for sizes greater than two CP units) may be explained by the formation of H-bonds between THR side-chains. As a result, backbone atoms fluctuate less and AFCP CPNTs with three units become equally stable with respect to the original sequence CPNTs of the same size. Nanotubes of more than 3 AFCPs are slightly stabler than those of the original CP sequence due to the formation of H-bonds between side-chains in accordance with the prediction of Vijayaraj [11].

We compared the distances between THR in the constructed nanotube structures to THR distances reported for CfAFP-501 in literature. Comparison of the average distances between THR residues in the constructed nanotube structures of the AFCP sequence and CfAFP-501 showed that the distances are very similar. The distances are close to the O-O distances of ice and may lead to good lattice matching, suggesting
the nanotubes of the AFCP sequence show similar or the same ice-binding ability as CfAFP-501 and are able to depress the freezing point.

We studied the self-assembly and dissociation process of a single CP to a CPNT5 using Transition Path Sampling. The TPS simulations indicated that the dissociation process of a single CP from a stable nanotube is initiated by the breaking of native backbone-backbone H-bonds. Once all native intermolecular $\beta$-sheet H-bonds between backbone-backbone atoms are broken, alternative H-bonds and hydrophobic interactions between CP6 and the remainder of the nanotube tube may be still be present. When the latter also break, CP6 truly separates from the nanotube. The determining step in the dissociation of the CP is the breaking of all backbone-backbone H-bonds as reported by Vijayaraj et al., followed by the loss of all side-chain and backbone interactions between CP6 and the remainder of the tube. The observation of Vijayaraj et al. that during the dissociation various side-chain side-chain interactions are lost is thus too simple [11]. From the perspective of association, all TPS paths are examples of the dock-lock mechanism that has also been identified in protein aggregation studies [29–32].

We found evidence for at least three meta-stable on-pathway intermediates (iA, iB and iC) in the dissociation/associating process. The presence of intermediate iC indicates that CPNTs interact strongly with one another, in accordance with a proposed crystal structure of highly aligned nanotubes. This indicates that synthetic CP crystals should have a low solubility, which is consistent with experimental observation. [33].

Using a simple rate theory expression for the association timescale, we conclude that at low concentration of the AFCPs, association is should occur via a mechanism involving intermediate iC.

Insight in these mechanisms can lead to improved design of the cyclic systems. Current efforts are underway to do so [33].
Appendix

5.A Brute force MD and sequences of different nanotubes

<table>
<thead>
<tr>
<th>System</th>
<th>Run 1</th>
<th>Run 2</th>
<th>Run 3</th>
<th>Run 4</th>
<th>Run 5</th>
</tr>
</thead>
<tbody>
<tr>
<td>CPNT2</td>
<td>B → I</td>
<td>B → I</td>
<td>B → I</td>
<td>B → I</td>
<td>B → I</td>
</tr>
<tr>
<td>CPNT3</td>
<td>B</td>
<td>B</td>
<td>B</td>
<td>B</td>
<td>B</td>
</tr>
<tr>
<td>CPNT4</td>
<td>B</td>
<td>B</td>
<td>B</td>
<td>B</td>
<td>B</td>
</tr>
<tr>
<td>CPNT5</td>
<td>B</td>
<td>B</td>
<td>B</td>
<td>B</td>
<td>B</td>
</tr>
<tr>
<td>CPNT6</td>
<td>B</td>
<td>B</td>
<td>B</td>
<td>B</td>
<td>B → I</td>
</tr>
</tbody>
</table>

Table 5.A.1: Original sequence: Visiting states of different size CPNTs in the course of a 100 ns NPT simulations. $B$ denotes the bound state and $I$ the area of phase space which does not belong neither to $B$ nor $U$.

<table>
<thead>
<tr>
<th>System</th>
<th>Run 1</th>
<th>Run 2</th>
<th>Run 3</th>
<th>Run 4</th>
<th>Run 5</th>
</tr>
</thead>
<tbody>
<tr>
<td>CPNT2</td>
<td>B → I</td>
<td>U → I</td>
<td>U → I</td>
<td>B → I</td>
<td>B → I</td>
</tr>
<tr>
<td>CPNT3</td>
<td>B</td>
<td>B</td>
<td>B</td>
<td>B</td>
<td>B</td>
</tr>
<tr>
<td>CPNT4</td>
<td>B</td>
<td>B</td>
<td>B</td>
<td>B</td>
<td>B</td>
</tr>
<tr>
<td>CPNT5</td>
<td>B</td>
<td>B</td>
<td>B</td>
<td>B</td>
<td>B</td>
</tr>
<tr>
<td>CPNT6</td>
<td>B</td>
<td>B</td>
<td>B</td>
<td>B</td>
<td>B</td>
</tr>
</tbody>
</table>

Table 5.A.2: AFCP sequence: Visiting states of different size CPNTs in the course of a 100 ns NPT simulations. $B$ denotes the bound state and $U$ the unbound state. $I$ denotes the area of phase space which does not belong neither to $B$ nor $U$. 
Table 5.A.3: Original CP sequence: schematic representation of the arrangement and chirality of the amino acids in the CP chains of sequence cyclo-[(L-LYS-D-ALA-L-LEU-D-ALA)\(_2\)] within various CPNTs. The sequential chains are named as CP1 to CP6. The terminals are assumed to be bonded in order to form the cyclic structure.

<table>
<thead>
<tr>
<th>Chain Orientation</th>
<th>Chain Name</th>
</tr>
</thead>
<tbody>
<tr>
<td>N → C</td>
<td>CP1</td>
</tr>
<tr>
<td>C ← N</td>
<td>CP2</td>
</tr>
<tr>
<td>N → C</td>
<td>CP3</td>
</tr>
<tr>
<td>C ← N</td>
<td>CP4</td>
</tr>
<tr>
<td>N → C</td>
<td>CP5</td>
</tr>
<tr>
<td>C ← N</td>
<td>CP6</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Chain Orientation</th>
<th>Chain Name</th>
</tr>
</thead>
<tbody>
<tr>
<td>N → C</td>
<td>CP1</td>
</tr>
<tr>
<td>C ← N</td>
<td>CP2</td>
</tr>
<tr>
<td>N → C</td>
<td>CP3</td>
</tr>
<tr>
<td>C ← N</td>
<td>CP4</td>
</tr>
<tr>
<td>N → C</td>
<td>CP5</td>
</tr>
<tr>
<td>C ← N</td>
<td>CP6</td>
</tr>
</tbody>
</table>

Table 5.A.4: AFCP sequence: schematic representation of the arrangement and chirality of the amino acids in the CP Chains of sequence cyclo-[(L-LYS-D-ALA-L-LEU-D-ALA)\(_2\)] within various CPNTs. The sequential chains are named as CP1 to CP6. The terminals are assumed to be bonded in order to form the cyclic structure.
Bibliography


[33] Personal correspondence with dr. ilja k. voets, macromolecular and organic chemistry & institute for complex systems, eindhoven university of technology, the netherlands.