Stability and growth mechanism of self-assembling putative antifreeze cyclic peptides

Brotzakis, Z.F.; Gehre, M.; Voets, I.K.; Bolhuis, P.G.

DOI
10.1039/c7cp02465g

Publication date
2017

Document Version
Final published version

Published in
Physical Chemistry Chemical Physics

PCCP

PAPER

View Article Online
View Journal | View Issue

Cite this: *Phys. Chem. Chem. Phys.*, 2017, **19**, 19032

Stability and growth mechanism of self-assembling putative antifreeze cyclic peptides†

Z. Faidon Brotzakis, ‡[a] Mascha Gehre,[a] Ilja K. Voets [b] and Peter G. Bolhuis *[a]

Received 16th April 2017,
Accepted 14th June 2017

DOI: 10.1039/c7cp02465g

rsc.li/pccp

Cyclic peptides (CPs) that self-assemble in nanotubes can be candidates for use as antifreeze proteins. Based on the cyclic peptide sequence cyclo-[(L-LYS-D-ALA-L-LEU-D-ALA)$_2$], which can stack into nanotubes, we propose a putative antifreeze cyclic peptide (AFCP) sequence, cyclo-[(L-LYS-D-ALA)$_2$-(L-THR-D-ALA)$_2$], containing THR-ALA-THR ice binding motifs. Using molecular dynamics simulations we investigate the stability of these cyclic peptides and their growth mechanism. Both nanotube sequences get more stable as a function of size. The relative stability of the AFCP sequence CPNT increases at sizes greater than a dimer by forming intermolecular THR side chain H-bonds. We find that, like the naturally occurring AF protein from spruce budworm (*Choristoneura fumiferana*), the THR distances of the AFCP's ice binding motif match the ice prism plane O–O distances, thus making the AFCP a suitable AF candidate. 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 association, involving different metastable intermediates, including a state in which the peptide was misfolded inside the hydrophobic core of the tube. Finally, we 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 insights into the mechanisms of peptide assembly, and might lead to an improved design of self-assembling antifreeze proteins.

1 Introduction

The binding of antifreeze proteins onto ice lowers the freezing point below the melting point generating a hysteresis gap of a few centigrades in which ice crystals remain small and further crystal growth is halted.[1] The ability of antifreeze proteins (AFPs) to block ice recrystalization is also interesting from an applied perspective, since it causes tissue damage in cryopreservation[2] and deteriorates the texture of frozen foods.[3]

Although all AFPs share antifreeze activity, they show a significant diversity in this activity, in molecular weight, amino-acid sequence and structure. Many hyperactive AFPs have a β-helical rich structure with repetitive sequence motifs.[4–6] For example the spruce budworm AFP shows a β-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 *Cf*AFP-501), Zhou *et al.*[7] showed that the THR O–O distances in the TXT motif of the spruce budworm antifreeze protein as well as the THR O–O distances of neighbouring coils are similar to the ice prism plane O–O distances.

Hyperactive antifreeze proteins that have a β-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 (CPs) containing the TXT ice binding motif, as shown in Fig. 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 β-sheet hollow structure, stabilized by backbone hydrogen bonds between adjacent CPs and by side chain interactions[11] (see Fig. 1 bottom). Vijayraj *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 sequence in such cyclic peptides

[a] *Van't Hoff Institute for Molecular Sciences, Universiteit van Amsterdam, Science Park 904, 1098 XH Amsterdam, The Netherlands. E-mail: p.g.bolhuis@uva.nl*
[b] *Laboratory of Macromolecular and Organic Chemistry, Laboratory of Physical Chemistry, and Institute for Complex Molecular Systems, Eindhoven University of Technology, Post Office Box 513, 5600 MB, Eindhoven, The Netherlands*

† Electronic supplementary information (ESI) available: Molecular dynamics simulation details. See DOI: 10.1039/c7cp02465g

‡ Present address: Department of Chemistry and Applied Bioscience, ETH Zürich and Facoltà di Informatica, Istituto di Scienze Computazionali, Universita della Svizzera Italiana, *Via* Buffi 13 Lugano, CH-6900, Lugano, Switzerland.

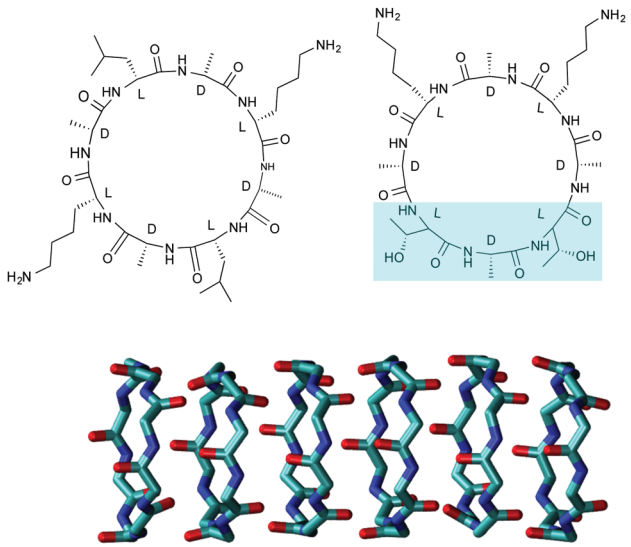

Fig. 1 Top: Original cyclic peptide sequence (left), and AFCP sequence (right). In cyan is highlighted the ice-binding motif TXT. Bottom: Backbone atoms of stacks of cyclic peptide (AFCP hexamer nanotube).

could lead to self-assembled nanotubes with great potential as antifreeze agents. It has been shown[12] that cyclic antifreeze glycopeptides (AFGPs) exhibit high antifreeze activity, however no study to our knowledge has focused on the antifreeze activity and the assembly mechanism of non-glycosylated antifreeze cyclic peptides. Here, we employ molecular simulations to investigate the stability and fluctuations of putative antifreeze 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 the original CP nanotube). After we have established that both the (AF)CP and the original CP nanotube are stable in solution, we continue with the question of how much does the AFCP match the ice-lattice distances, as most AF proteins do. We proceed with the formation mechanism. The formation mechanism of the AFCP nanotube in 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 into 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.

2 Methods

2.1 System setup

2.1.1 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.[13] 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.[14] The D-conformational orientation of the D-Alanine amino acids was obtained by flipping the L-Alanine amino acids in PyMOL[15] in order to obtain the original and AFCP sequences (see also Tables S5 and S6 of the ESI†). The geometry of each single CP unit was energy minimized using steepest descent while the Ramachandran dihedral angles (Φ,Ψ) of all amino acids were restrained to their average anti-parallel β-sheet values as given in Table S1 of the ESI,† 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 10 ns *NPT* simulations.

The final geometry of the CPs was used to build various nanotubes of different sizes (*eg.* 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. 1 bottom and detailed in Tables S5 and S6 of the ESI.†

2.2 Molecular dynamics

All energy minimization and molecular dynamics simulations were performed with the Gromacs 4.5.4 package.[14] Molecular dynamics simulations by Vijayaraj and Khurana[10,16] suggest that using the amberf99sb[17] stabilizes the cyclic peptide nanotubes by the formation of intramolecular hydrogen bonds. Therefore we choose here the same protein force-field, coupled with TIP3P water.[18] 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 lists were updated every 10 fs with a cutoff of 0.8 nm and the time step was 2 fs.[17] The leap-frog algorithm was used for integrating Newton's equations of motion. In the *NPT* simulations the v-rescale thermostat[19] with a coupling time constant of 0.2 ps controlled the temperature, while the Parrinello–Rahman barostat[20] 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[18] water molecules extending 1.7 nm away from the solute atoms. The energy minimization, equilibration, and production runs of various nanotubes (CPNTs) 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_α 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 S2 of the ESI† 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.

2.3 Transition path sampling

2.3.1 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 trajectory forward and backward by integrating the time reversible Newtonian equations of motion. For the backward integration the velocities at the selected time frame are reversed. Acceptance or rejection of the trial trajectory is done using a detailed balance obeying the Metropolis rule.[21,22] For the standard two way shooting move with a fixed path length the Metropolis acceptance rule just checks if the trial path connects the two stable states. If not the trial path is rejected.

In this work we use the more efficient one way flexible shooting algorithm.[22,23] First, a time frame t_{sel} is 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 τ_{part} is generated by molecular dynamics until the initial state (in the case of backward shot) or the final state (in the case of a forward shot) is reached. Because of the use of the stochastic v-rescale thermostat[19] 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, $N_n = (N_o - \tau_{sel}) + \tau_{part}$. 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 the computation time in very long paths connecting the two states, the maximum allowed path length $N_{max} = N_{(o)}/\xi$ is computed in advance, where ξ is a random number in the interval [0,1]. This TPS algorithm has been previously used in other protein systems.[25,26]

2.3.2 Mechanistic analysis *via* the path density. For gaining better insight into the mechanism of the transitions we project the TPS ensemble on two dimensional path density plots. We construct the path density histograms by choosing two order parameters (OPs), and creating a 2D grid initialized to zero. Examples of such pairs are the minimum distance between the CP6 peptide and the centre of mass of the backbone of CP5 (d_{min}(top1$_{sidechain}$, top2$_{cm}$)) as a function of the RMSD of CP6 with respect to the correctly bound nanotube configuration (rmsd(top1,top$_B$)), or as a function of the minimum distance between CP6 and the nanotube (d_{min}(top1, rest)). 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 show 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.

3 Results and discussion

3.1 Nanotube equilibrium properties

3.1.1 Stability. We performed a series of 100 ns MD simulations for each of the systems denoted in Table S2 of the ESI.† We define a system as being stable if it maintains an anti-parallel cyclic β-sheet tube-like structure throughout the course of 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 simulations. 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 Tables S3 and S4 of the ESI†) shows that most of the CPNT systems maintain their anti-parallel β-sheet tube-like structure throughout the simulations (see Tables S7 and S8 of the ESI†). 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. This can also be shown in Table S8 of the ESI,† by the deviation of the average ϕ and ψ from the average anti-parallel β-sheet value shown in Table S1 of the ESI.†

Fig. 2 shows the root mean square fluctuations (RMSFs) of backbone atoms within all CPNT systems of both sequences. The RMSF for CPNT2 is higher than those for 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.

Fig. 2a and 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 (Fig. 3a),

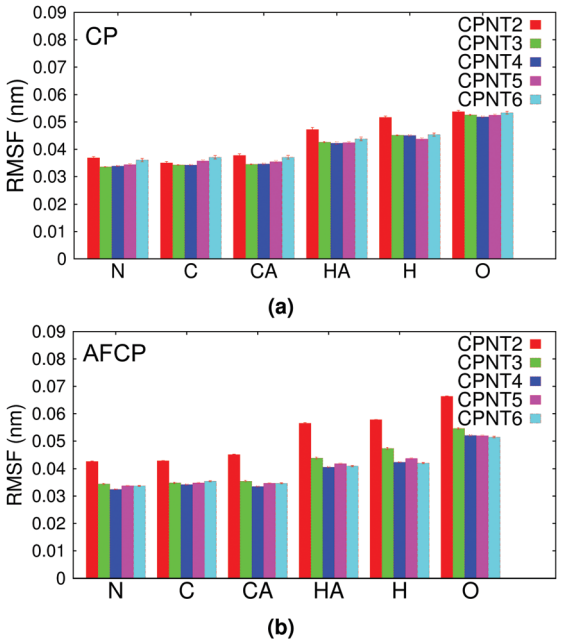

Fig. 2 (a) Root mean square fluctuation (RMSF) of backbone atoms for (a) the original CP sequence and (b) for the AFCP sequence.

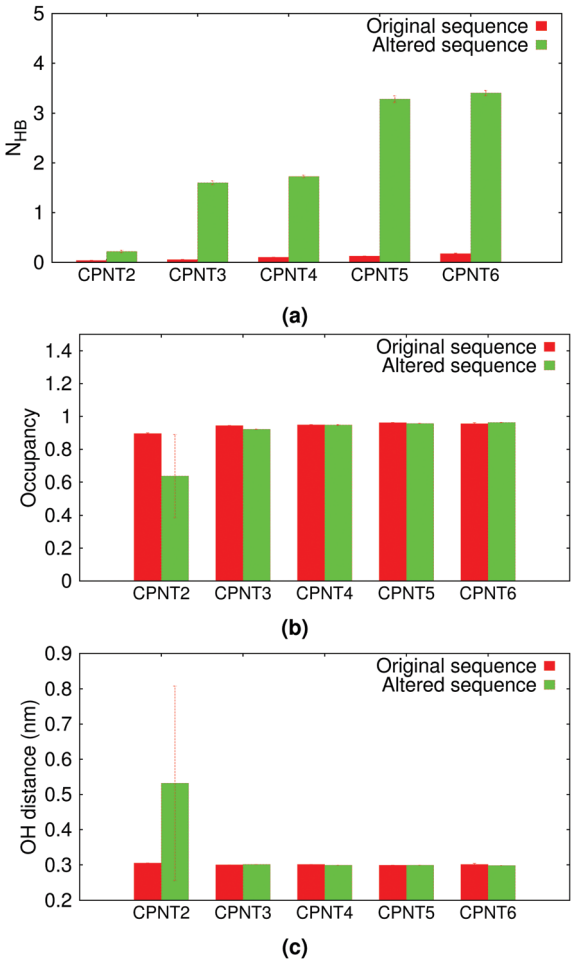

Fig. 3 (a) Number of side chain hydrogen bonds (N_{HB}) between adjacent peptides, (b) backbone hydrogen bond occupancy and (c) O–H distances, for the AFCP sequence (altered) and the original CP sequence and for each nanotube size.

which is higher for the AFCP sequence than for the original CP sequence as well as by the higher occupancy of backbone hydrogen bonds the trimer has compared to the dimer of the AFCP sequence (see Fig. 3b). These stabilization interactions explain why the peptides of the trimer CPNT of the AFCP sequence do not dissociate, compared with the dimer ones (see Table S4 of ESI†), which also reflects their lower backbone OH distance compared to the very much fluctuating ones of the dimer (see Fig. 3c). 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 S3 of the ESI†), suggesting that 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]

3.1.2 Ice lattice matching of AFCP nanotubes. In a computational study, Zhou *et al.*[7] correlated the high interaction energy between the *Cf*AFP-501 isoform and ice with its antifreeze activity and with the matching of the intercoil and intracoil side chain O–O distances of THR in the TXT motif of the same and adjacent coil respectively with the ice lattice. The distances found were (5.90–7.46 Å and 3.95–4.96 Å) respectively. These distances are similar to the values reported by Li *et al.*[27] for the same protein and for different isoforms (respectively, 6.75 Å and 4.57 Å). The above mentioned distances match the ice Ih unit cell dimensions c = 7.356 Å and a = 4.518 Å.[28]

It is conceivable that in order to engineer an antifreeze sequence which contains the TXT motif and has a large affinity for ice, the THR side chain O–O distances should have a very good match with the ice crystal cell dimensions. Fig. 4 shows (a) the intramolecular and (b) the inter-molecular C_α–C_α distances of THR residues for the AFCP nanotubes of different sizes. For the stable CPNTs (>CPNT2) the average value of the intramolecular and the intermolecular THR C_α–C_α distances is, respectively, 7.04 Å and 4.88Å. Therefore, we expect that the AFCP sequence binds ice and depresses the freezing point as the *Cf*AFP-501 spruce budworm isoform.

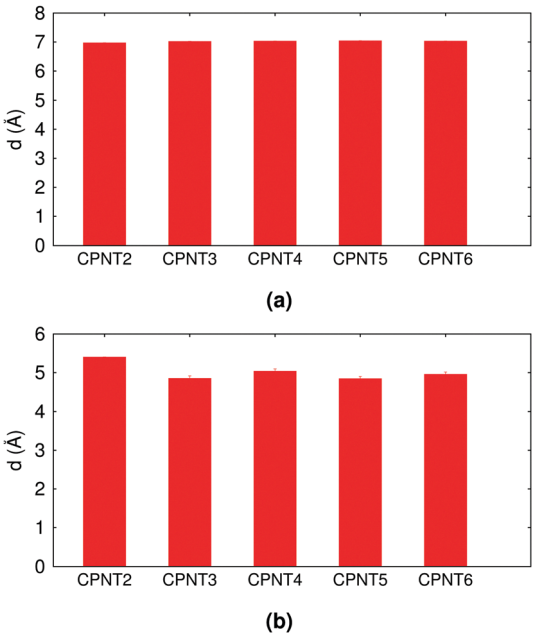

Fig. 4 C_α–C_α distance d of THR residues belonging to (a) the same peptide and (b) to adjacent peptides for the AFCP sequence.

Many hyperactive proteins have flat ice-binding surfaces which increase their affinity for planar ice surfaces, as well as form strong hydrogen bonds with ice *via* the hydroxyl group of the ice binding site threonine.[29,30] Fig. 5 shows that the AFCP sequence has indeed a flatter THR-ALA-THR motif surface and a less perturbed H-bond network compared to LYS-ALA-LEU of the original CP sequence hexamer. Moreover, we find that the water radial distribution function around the THR side-chain oxygen is reminiscent of that of a hydrophobic site, since the side-chain hydroxyl is surrounded by the THR methyl group and the adjacent ALA sidechain methyl group (see Fig. 1 of the ESI†). However, the rdf also shows a very distinct first hydration layer (minimum) at a distance of 0.35 nm, which can be explained by the hydrogen bond the side-chain hydroxyl donates to water. The occupancy of that H-bond is 1.35_1, indicating the presence of more than one water accepting an H-bond with the THR hydroxyl group on average.

3.2 Mechanism of self-assembly by transition path sampling

3.2.1 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 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 CPNT5 and a 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_α RMSD of the outermost peptide CP6 from its natively bound state is less than 0.03 nm, and

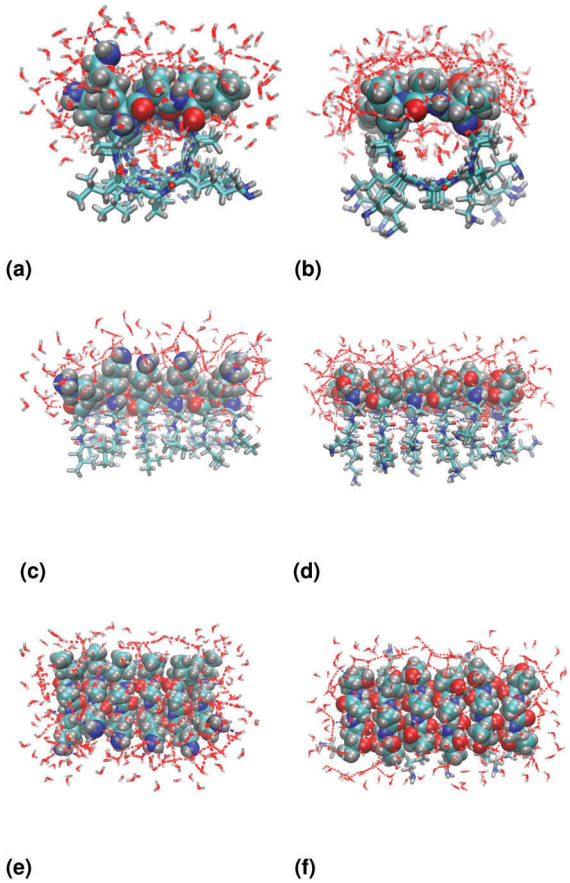

Fig. 5 Representative snapshots from MD simulations of the CPNT6 system viewed from three different angles. (a, c and e) The original CP sequence, (b, d and f) the AFCP sequence.

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 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^{-1}, 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 S4 of the ESI†), 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

Table 1 Stable state definitions as a function of number of backbone hydrogen bonds, C_α-RMSD of CP6 and minimum distance d_{min}

	RMSD (nm)		d_{min} (nm)		H-bonds	
State	Min	Max	Min	Max	Min	Max
B	0	0.03	—	—	8	∞
U	—	—	1	∞	0	0

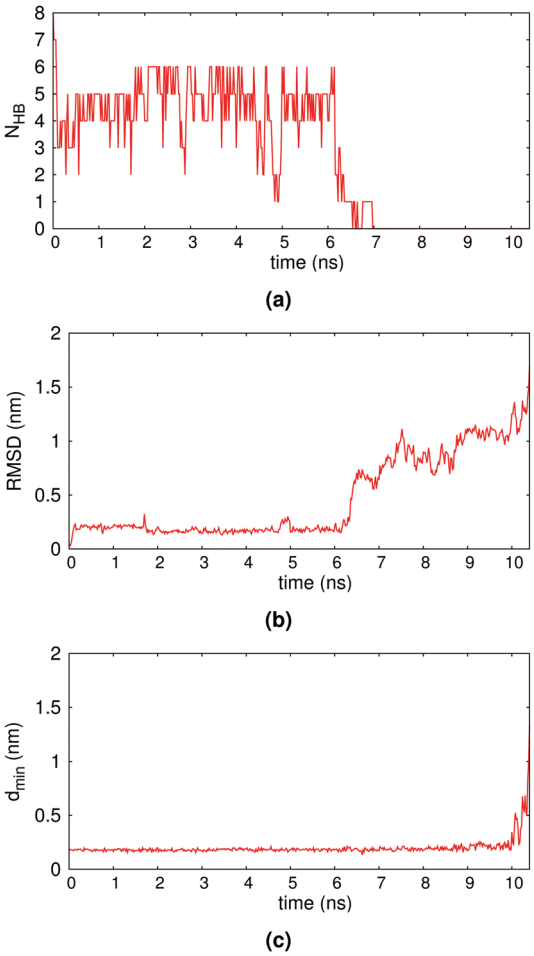

Fig. 6 The initial pathway for the TPS, as a function of (a) backbone hydrogen bonds between CP6 and CP5 (N_{HB}), (b) CP6 RMSD, and (c) minimum distance (d_{min}) between peptide CP6 and the remainder of the tube.

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 same frame, and ending up in different states. The initial path is graphically illustrated as a function of time in Fig. 6.

3.2.2 Transition path sampling. 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 Fig. 7, 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 simulations we obtained 20 completely decorrelated paths. Inspection of the tree reveals a distribution of different path lengths with shorter and longer reactive pathways. Fig. 8 shows this reactive path length distribution. The path length distribution 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 transition.

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 Fig. 9). 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 with the side of the tube. From the distribution of the path length (Fig. 8) 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) has an average length of 3.28 ns (green curve). The dissociation transition I visiting on-pathway intermediate iB (BiAiBU) takes on average 19 ns (blue curve). In addition, we identified several paths mixing these mechanisms such as BiAiBiCU and BiAiCiBU (see Fig. 8).

To further understand the association/dissociation transition, we analyze the mechanisms of three selected pathways 3, 35, and 23, respectively, corresponding to transitions I, II and III. All dissociation/association pathways visit the docked intermediate state iA. Fig. 9 shows that in intermediate iA CP6 has lost its planar β sheet conformation and has instead a V-shape, with only a fraction of the backbone hydrogen bonds to CP5 being intact.

Fig. 10a 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 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 metastable 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. 11c 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. Moreover, Fig. 11d displays that

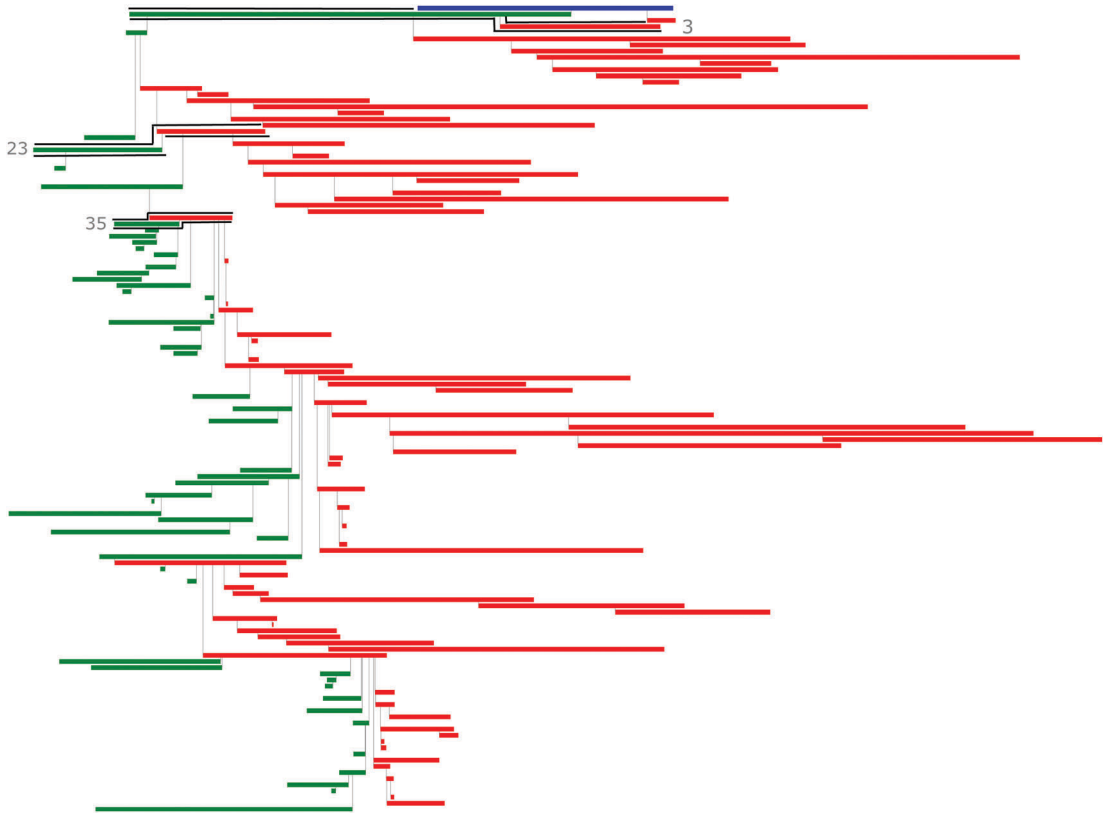

Fig. 7 Transition path ensemble tree representation. In green are depicted the backward partial paths and in red the forward partial paths. Note that the forward paths are on average longer than the backward paths. Paths 3, 23, 35 are delimited by black solid lines.

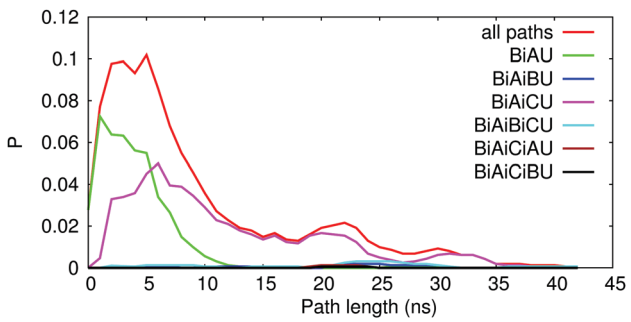

Fig. 8 Length distribution of the transition path ensemble trajectories. The total distribution is highlighted in red, and is split in the underlying distributions according to their transition mechanism: BiAU (green), BiAiBU (blue), BiAiCU (purple), BiAiBiCU (cyan), BiAiCiAU (brown), and BiAiCiBU (black).

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 the CP5 backbone (Fig. 10b). 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. 11c) or water molecules at the top of the tube (Fig. 11d). As shown in Fig. 8, 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. 11e and f 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 CP6 and the tube are the backbone ones. Indeed, the red and black lines of Fig. 11e coincide.

Mechanism of transition III (path 23). Mechanism III consists of transitions B ⇆ iA ⇆ iC ⇆ U. Metastable intermediate iC does not resemble intermediates iA and iB. Fig. 11g 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, in contrast, 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 suggest that CP6 may slide along the tube. Indeed the RMSD shown in Fig. 11h increases, whereas the minimum distance remains small, indicating that the peptide remains associated with the nanotube.

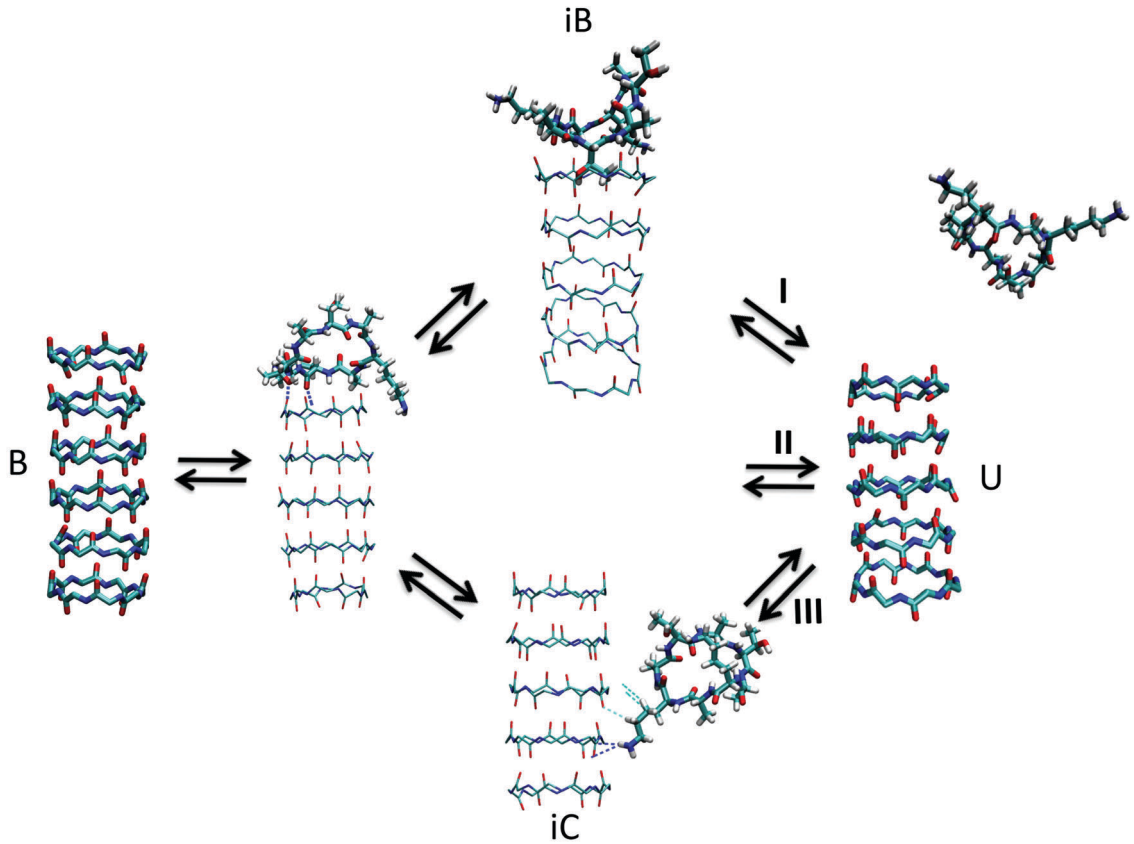

Fig. 9 Mechanisms found in the transition path ensemble. The three corresponding mechanisms are: mechanism I (B ⇄ iA ⇄ iB ⇄ U), mechanism II (B ⇄ iA ⇄ U), mechanism III (B ⇄ iA → iC → U).

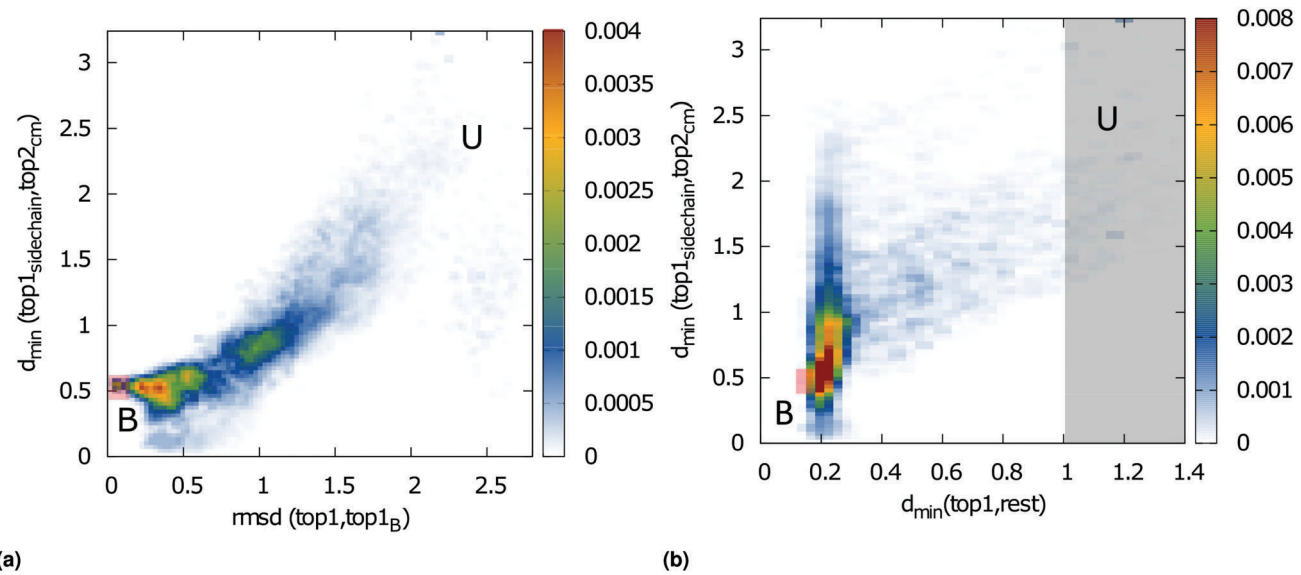

Fig. 10 Path density plots of (a) the RMSD of CP6 with respect to the correctly bound nanotube configuration *vs.* the minimum distance (d_{min}($top1_{sidechain}$, $top2_{cm}$)) between the CP6 and the center of mass of CP5 and (b) the minimum distance (d_{min}(top1, rest)) of CP6 to the nanotube, *vs.* the minimum distance (d_{min}($top1_{sidechain}$, $top2_{cm}$)) between the CP6 and the center of mass of CP5. States B and U are labeled. State U is highlighted by a gray rectangle in (b). Note that, since the unbound state is defined solely with a minimum distance criterion, state U can have different values of RMSD in (a).

Fig. 11 shows that 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

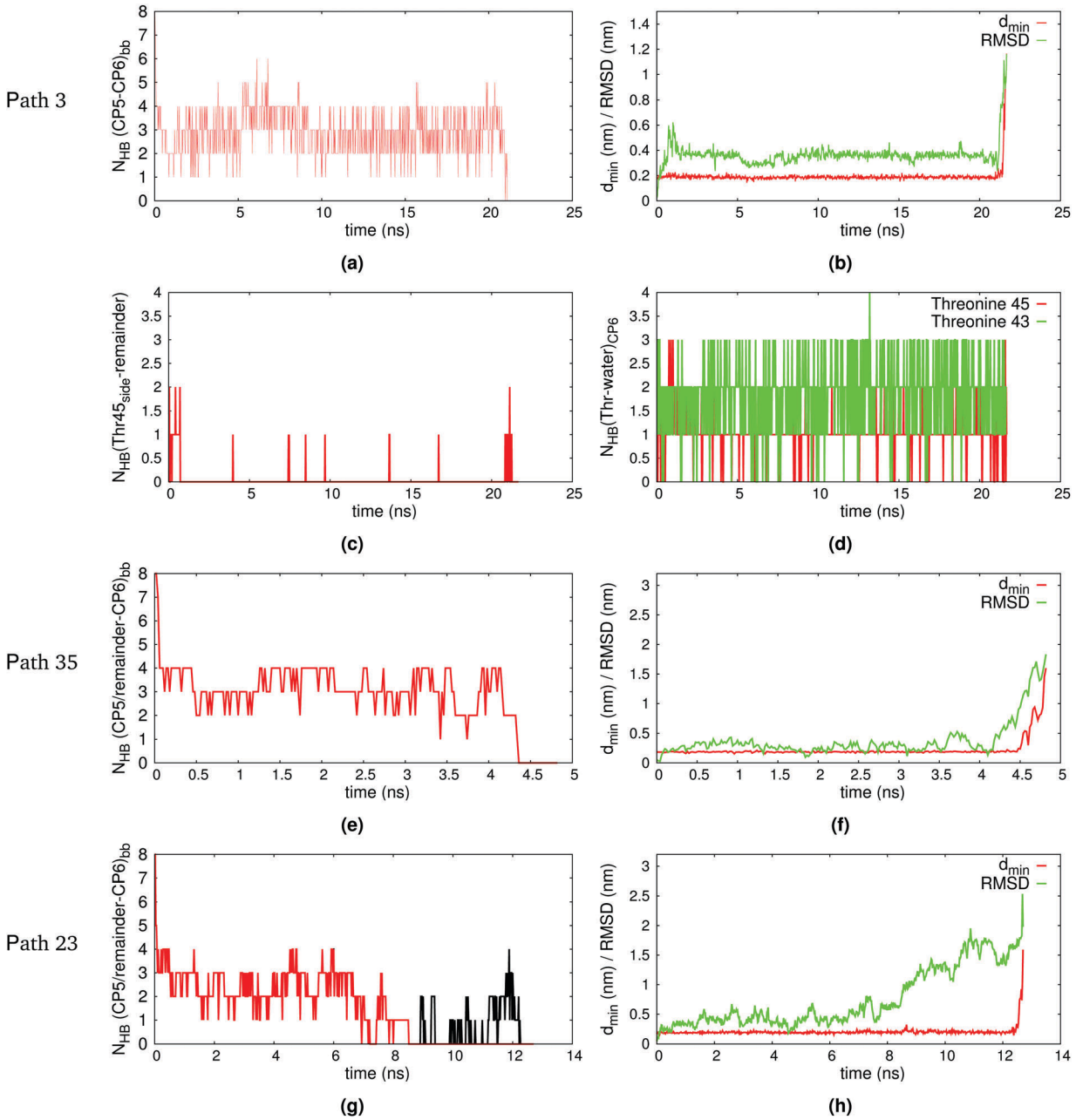

Fig. 11 Time evolution of the order parameters. Path 3: (a) number of H-bonds (N_{HB}(CP5–CP6)$_{bb}$) between backbone atoms of CP5 and CP6. (b) RMSD in nm of CP6 (green) and minimum distance (d_{min}) in nm between CP6 and the remainder of the tube (red). (c) Number of H-bonds (N_{HB}(Thr-45$_{side}$–remainder)) between the side chain of THR-45 of CP6 and the remainder of the tube. (d) Number of H-bonds (N_{HB}(Thr–water)$_{CP6}$) between THR side-chains and water molecules. The side chain of THR-45 is displayed in red and the side-chain of THR-43 in green. Path 35: (e) Number (N_{HB}(CP5/remainder–CP6)$_{bb}$) of H-bonds between backbone atoms of CP6 and CP5 (red) and H-bonds between CP6 and the remainder of the tube (black). (f) RMSD in nm of CP6 (green) and minimum distance (d_{min}) in nm between CP6 and the remainder of the tube (red). Path 23: (g) number of H-bonds between backbone atoms of CP6 and CP5 (red line) and H-bonds between CP6 and the remainder of the tube (black line). (h) RMSD in nm of CP6 (green) and the minimum distance (d_{min}) in nm between CP6 and the remainder of the tube (red).

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 metastable 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 al.*[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 with 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.[31–34]

Intermediate iC is characterized by the presence of H-bonds between the dissociating CP6 and the tube. This might have consequences for the solubility of the nanotubes. The original cyclic peptide sequence forms nanotubes without aggregation. Replacing LEU with THR in the AFCP sequence introduces additional sidechains capable of forming H-bonds. Whereas backbone H-bond interactions are required for 1D fiber formation,[35] additional H-bond capable sidechains at the outside of the tube can enhance 3D aggregation, and hence lower the solubility. Indeed, pathways corresponding to mechanism III (see Fig. 11g) show that during the sliding transition H-bonds form. However, since THR residues are important for the putative antifreeze character of the nanotube, we speculate that in future research this problem might be alleviated by reducing the H-bond capability of the charged residues, by replacing *e.g.* lysine with histidine residues.

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 to the 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 σ

$$k_{\mathrm{on}} = \frac{4\pi D\sigma}{V}, \tag{1}$$

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 $\tau_{\mathrm{on}} = {k_{\mathrm{on}}}^{-1}$. The association rate constant of a CP associating with a nanotube of length L is instead given by[36]

$$k_{\mathrm{on}}^{\mathrm{tube}} = \frac{4\pi DL}{V\ln(2L/\sigma)} \tag{2}$$

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

$$\tau_1 = \frac{L^2}{3D_1}. \tag{3}$$

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

$${\tau_{\mathrm{on}}}' = \frac{V\ln(2L/\sigma)}{4\pi DL} + \frac{L^2}{3D_1} \tag{4}$$

This timescale can always be made smaller than the timescale of the direct mechanism, τ_{on} because the first term is dominant at a low concentration, and can be made smaller by considering longer L.

To identify the crossover, we equate the two timescales ${\tau_{\mathrm{on}}}' = \tau_{\mathrm{on}}$

$$\frac{V\ln(2L/\sigma)}{4\pi DL} + \frac{L^2}{3D_1} = \frac{V}{4\pi D\sigma} \tag{5}$$

$$\frac{\sigma\ln(2L/\sigma)}{L} + \frac{L^2 4\pi D\sigma}{3D_1 V} = 1 \tag{6}$$

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

4 Conclusion

Inspired by the depression of the freezing point caused by TXT rich AF proteins, we are motivated towards designing and simulating self-assembling putative antifreeze cyclic peptides (CPs). By performing Molecular Dynamics simulations, we have compared the stability of the experimentally found nanotube structure made out of CPs with sequence cyclo-[(L-LYS-D-ALA-L-LEU-D-ALA)$_2$] and a proposed nanotube made out of a putative antifreeze CP sequence cyclo-[(L-LYS-D-ALA)$_2$-(L-THR-D-ALA)$_2$]. In order to elucidate the mechanism of growth of the putative AFCP sequence, we performed enhanced sampling Transition Path Sampling simulations. MD simulations indicate that both nanotubes become more stable as a function of size, with the AFCP sequence CPNT being relatively stabilized even more by forming intermolecular THR side chain H-bonds at sizes greater than a dimer. Our hypothesis that the AFCP sequence has a potential ice affinity and antifreeze activity is corroborated by our finding that the AFCP sequence shows a very good matching of intercoil and intracoil THR O–O side chain distances with the ice-crystal lattice distances, similar to natural TXT-rich proteins like *Cf*AFP-501.

We elucidated the growth mechanism of a single CP to a CPNT5 by Transition Path Sampling. TPS simulations indicated that the association process of a single CP to a stable nanotube is initiated by the formation of H-bonds and hydrophobic interactions between CP6 and the remainder of the nanotube, followed by the formation of native backbone intermolecular β-sheet H-bonds. The bottleneck in the association of the CP is the formation of all backbone–backbone H-bonds as reported by Vijayaraj *et al.* In the case of the sliding pathways the bottleneck can be the formation of all side-chain and backbone interactions between CP6 and the remainder of the tube. The observation 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.[31–34]

We found evidence for at least three metastable 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, with H-bonds, which could lead to low solubility.

Using a simple rate theory expression for the association timescale, we conclude that at a low concentration of the AFCPs, association should occur *via* a mechanism involving intermediate iC. Our simulations give further insight into the design principles to use for synthesizing AFCPs.

Acknowledgements

This work is supported by NanoNextNL, a micro and nanotechnology consortium of the Government of the Netherlands and 130 partners. We acknowledge support from the Nederlandse Organisatie voor Wetenschappelijk Onderzoek (NWO) for the use of supercomputer facilities. I. K. Voets acknowledges the European Union (ERC-2014-StG Contract No. 635928) and the Dutch Science Foundation (NWO ECHO Grant No. 712.016.002) for funding.

References

1 C. P. Garnham, R. L. Campbell and P. L. Davies, *Proc. Natl. Acad. Sci. U. S. A.*, 2011, **108**, 7363.
2 O. B. Usta, Y. Kim, S. Ozer, B. G. Bruinsma, J. Lee, E. Demir, T. a. Berendsen, C. F. Puts, M.-L. Izamis, K. Uygun, B. E. Uygun and M. L. Yarmush, *PLoS One*, 2013, **8**, e69334.
3 C. Lu, H. Dai and S. Guo, *Shipin Yu Fajiao Gongye*, 2001, **27**, 44.
4 S. P. Graether, M. J. Kuiper, V. K. Walker, Z. Jia, B. D. Sykes and P. L. Davies, *Nature*, 2000, **133**, 325.
5 S. P. S. B. D. Graether, *Eur. J. Biochem.*, 2004, **271**, 3285.
6 K. Meister, S. Ebbinghaus, Y. Xu, J. G. Duman, A. DeVries, M. Gruebele, D. M. Leitner and M. Havenith, *Proc. Natl. Acad. Sci. U. S. A.*, 2013, **110**, 1617.
7 Y. Zhou, H. Tan, Z. Yang, Z. Jia, R. Liu and G. Chen, *Sci. China, Ser. B: Chem.*, 2007, **50**, 266.
8 M. R. Gadiri, J. R. Granja, R. A. Milligan, D. E. McRee and N. Khazanovich, *Nature*, 1993, **366**, 324.
9 R. Hourani, C. Zhang, R. van der Weegen, L. Ruiz, C. Li, S. Keten, B. A. Helms and T. Xu, *J. Am. Chem. Soc.*, 2011, **133**, 15296.
10 R. Vijayaraj, *J. Phys. Chem. B*, 2010, **114**, 16574.
11 R. Vijayaraj, S. Van Damme, P. Bultinck and V. Subramanian, *Phys. Chem. Chem. Phys.*, 2012, **14**, 15135.
12 M. Hachisu, H. Hinou, M. Takamichi, S. Tsuda, S. Koshida and S.-I. Nishimura, *Chem. Commun.*, 2009, 1641.
13 D. A. Case, T. Darden, T. E. C. Iii, C. Simmerling, S. Brook, A. Roitberg, J. Wang, U. T. Southwestern, R. E. Duke, U.-c. Hill, R. Luo, U. C. Irvine, D. R. Roe, R. C. Walker, S. Legrand, J. Swails, D. Cerutti, J. Kaus, R. Betz, R. M. Wolf, K. M. Merz, M. State, G. Seabra, P. Janowski, F. Paesani, J. Liu, X. Wu, T. Steinbrecher, H. Gohlke, N. Homeyer, Q. Cai, W. Smith, D. Mathews, R. Salomon-ferrer, C. Sagui, N. C. State, V. Babin, T. Luchko, S. Gusarov, A. Kovalenko, J. Berryman, P. A. Kollman and U. C. S. Francisco, *Amber*, Univerisity of California, San Fransisco, 2013.
14 S. Pronk, S. Páll, R. Schulz, P. Larsson, P. Bjelkmar, R. Apostolov, M. R. Shirts, J. C. Smith, P. M. Kasson, D. van der Spoel, B. Hess and E. Lindahl, *Bioinformatics*, 2013, **29**, 845.
15 The PyMOL Molecular Graphics System, Version 1.8 Schrödinger, LLC.
16 E. Khurana, S. O. Nielsen, B. Ensing and M. L. Klein, *J. Phys. Chem. B*, 2006, **110**, 18965.
17 V. Hornak, R. Abel, A. Okur, B. Strockbine, A. Roitberg and C. Simmerling, *Proteins: Struct., Funct., Bioinf.*, 2006, **65**, 712–725.
18 W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey and M. L. Klein, *J. Chem. Phys.*, 1983, **79**, 926.
19 G. Bussi, D. Donadio and M. Parrinello, *J. Chem. Phys.*, 2007, **126**, 014101.
20 M. Parrinello and A. Rahman, *J. Appl. Phys.*, 1981, **52**, 7182.
21 C. Dellago, P. G. Bolhuis, F. S. Csajka and D. Chandler, *J. Chem. Phys.*, 1998, **108**, 1964.
22 C. Dellago, P. Bolhuis and P. Geissler, *Adv. Chem. Phys.*, 2002, **123**, 1.
23 P. Bolhuis, *J. Phys.: Condens. Matter*, 2003, **113**, S113.
24 T. van Erp, D. Moroni and P. Bolhuis, *J. Chem. Phys.*, 2003, **118**, 7762–7774.
25 J. Vreede, J. Juraszek and P. G. Bolhuis, *Proc. Natl. Acad. Sci. U. S. A.*, 2010, **107**, 2397.
26 J. Juraszek and P. G. Bolhuis, *Biophys. J.*, 2008, **95**, 4246–4257.
27 C. Li, X. Guo, Z. Jia, B. Xia and C. Jin, *J. Biomol. NMR*, 2005, **32**, 251–256.
28 K. Röttger, A. Endriss, J. Ihringer, S. Doyle and W. F. Kuhs, *Acta Crystallogr., Sect. A: Found. Crystallogr.*, 1994, **B50**, 644.
29 U. S. Midya and S. Bandyopadhyay, *J. Phys. Chem. B*, 2014, **118**, 4743.
30 K. Meister, S. Lotze, L. L. C. Olijve, A. L. DeVries, J. G. Duman, I. K. Voets and H. J. Bakker, *J. Phys. Chem. Lett.*, 2015, **6**, 1167.
31 W. P. Esler, E. R. Stimson, J. M. Jennings, H. V. Vinters, J. R. Ghilardi, J. P. Lee, P. W. Mantyh and J. E. Maggio, *Biochemistry*, 2000, **39**, 6288.
32 S. R. Collins, A. Douglass, R. D. Vale and J. S. Weissman, *PLoS Biol.*, 2004, **2**, 1582.
33 P. H. Nguyen, M. S. Li, G. Stock, J. E. Straub and D. Thirumalai, *Proc. Natl. Acad. Sci. U. S. A.*, 2007, **104**, 111.
34 D. O'Brien, P. Edward, Y. Okamoto, J. E. Straub, B. R. Brooks and D. Thirumalai, *J. Phys. Chem. B*, 2014, **113**, 14421.
35 M. Garzoni, M. B. Baker, C. M. A. Leenders, I. K. Voets, L. Albertazzi, A. R. A. Palmans, E. W. Meijer and G. M. Pavan, *J. Am. Chem. Soc.*, 2016, **138**, 13985.
36 P. H. Richter and M. Eigen, *Biophys. Chem.*, 1974, **2**, 255.

