PCCP

Accepted Manuscript

ROYAL SOCIETY OF CHEMISTRY

www.rsc.org/pccp

Phys Chem Chem Phys

RSCPublishing

ARTICLE

Why Ice-Binding Type I Antifreeze Protein Acts as a Gas Hydrate Crystal Inhibitor

Cite this: DOI: 10.1039/x0xx00000x

S. Alireza Bagherzadeh,[a] Saman Alavi,[a,b] John A. Ripmeester,[a,b] and Peter Englezos[a]

Received ooth October 2014,
Accepted ooth January 2015

DOI: 10.1039/x0xx00000x

www.rsc.org/

Antifreeze proteins (AFPs) prevent ice growth by binding to a specific ice plane. Some AFPs have been found to inhibit the formation of gas hydrates which are a serious safety and operational challenge for the oil and gas industry. Molecular dynamics simulations are used to determine the mechanism of action of the Winter Flounder AFP (*wf*-AFP) in inhibiting methane hydrate growth. The *wf*-AFP adsorbs to the methane hydrate surface via cooperative binding of a set of hydrophobic methyl pendant groups to the empty half-cages at the hydrate/water interface. Each binding set is composed of the methyl side chain of the threonine and two alanine residues, four and seven places further down in the sequence of the protein. Understanding the principles of action of AFPs can lead to the rational design of green hydrate inhibitor molecules with potential superior performance.

Introduction

Gas hydrates are solid solutions composed of water and appropriate hydrate former molecules known as guests. Under favourable thermodynamic conditions, usually met in hydrocarbon transport and processing facilities, undersea sediments, and permafrost deposits, water molecules freeze to form a solid cage-like network, via hydrogen bonding, around guest molecules such as CH_4, C_2H_6, C_3H_8, and CO_2.[1–5]

Inspired by naturally occurring low-concentration antifreeze substances, in recent years efforts in the field of gas hydrate flow assurance are being made to manage the risk of hydrate plug formation (as opposed to the total prevention of hydrate formation in hydrocarbon transport and processing facilities) by injecting chemicals known as kinetic inhibitors (KI's).[6–8] Unlike conventional (thermodynamic) inhibitors such as methanol where concentrations of up to 50 wt.% may be used, the concentrations required for KI's to be effective are significantly lower and fall in the range of 0.1-1.0 wt.%. Although, KI's are experimentally proven to effectively inhibit hydrate growth[9–11] their mechanism of action is still not well understood. Currently, one of the challenges for lab-scale experiments is that it is not possible to predict the performance of a chemical in the field in spite of its excellent antifreeze performance in the laboratory. The elucidation of the physical and chemical interactions between the inhibitors and the gas hydrate crystal is critical to the design of effective inhibitors.

The winter flounder (*Pseudopleuronectes americanus*) can survive sub-zero arctic temperatures by producing an antifreeze protein (*wf*-AFP) which enables it to prevent the growth of ice crystals beyond a size that would lead to cell death. The *wf*-AFP adsorbs to the {20-21} ice plane where protein surface/structural complementarity with the ice surface is believed to be a key factor in the binding and antifreeze action.[12,13] The primary and three-dimensional[14] helix structures of the alanine (ALA or A)-rich 37-amino acid *wf*-AFP are shown in Fig. 1. This protein has four threonine (THR or T) residues which lie on one side (known as the ice binding surface, IBS) of the helix and repeat every eleven residues with an approximate distance of 1.65 nm. Each THR has a hydrophilic hydroxyl (–OH) side chain and a pendant hydrophobic methyl group ($–CH_3$). Chemical structures of the THR and ALA residues are shown in Fig. S1 of the Electronic Supplementary Information (ESI).

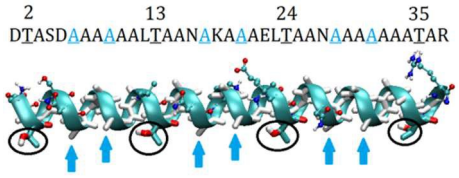

Fig. 1. Primary and three-dimensional structures of *wf*-AFP. Alanine (A): white bars. Threonine (T) residues are circled. The (*i*+4) and (*i*+7)ALA residues with respect to the (*i*)THR residue are shown by arrows. Other amino acid residues are shown with ball-and-stick models. Colour scheme: carbon, cyan; oxygen, red; hydrogen, white; nitrogen, blue

In early models, the hydrogen bonding between the OH group of THR residues and water oxygen at the ice surface was believed to drive the adsorption of *wf*-AFP to the ice surface. However, mutation studies showed that replacing the THR

hydroxyl pendant group with a methyl group does not result in complete loss of anti-freeze activity of the protein.[15] In later studies, the role of hydrophobic ALA residues were investigated and it was concluded that the contribution of hydrogen bonding to the ice surface is secondary to van der Waals interactions of the AFP with the surface.[16,17]

AFPs evolved in living organisms to deal with ice formation but some of them inhibit formation of gas hydrates[9] which have different microstructures from ice. This is an adaptation of a natural product for a different purpose than it was designed for, which could lead to green inhibition practices. For some AFP inhibitors, it was shown that nucleation and growth inhibition on ice and gas hydrates are independent processes and that good performance for ice inhibition does not necessarily lead to good hydrate inhibition activity.[18] To design more effective and environmentally friendly synthetic hydrate inhibitors, it is therefore crucial to determine the mechanism of action of sample AFPs on hydrate surfaces.

The mechanism of action of antifreeze molecules involves interactions that occur on the molecular scale. Molecular simulation techniques are well established and render useful molecular insights on nanometre scales, otherwise very challenging or not possible to capture experimentally.[19–21] Molecular simulation studies of AFPs in water and water/ice systems have been performed. It is reported that the IBS pre-organizes the surrounding solution water molecules, therefore maximizing compatibility with the ice surface and minimizing the entropy penalty of binding.[22] Protein surfaces containing the THR-ALA-ALA sequence are also found to interact more favourably with the ice/water interface.[23]

To our knowledge, this is the first molecular dynamics simulation studying the interactions of *wf*-AFP with the surface of a methane hydrate crystal. The aim is to obtain direct molecular insights on the mechanism of action of *wf*-AFP in surface binding and inhibiting hydrate growth and to determine the responsible structural features on the AFP. Ice has a hexagonal crystal structure and forms many available binding planes for the AFP. The cubic structure I methane hydrate, in contrast, has fewer binding planes available. The (001) plane of the clathrate hydrate phase is used for our AFP binding simulations as it is stable and has the greatest number of complete cage structures on the outer layer, with a relatively small number of "dangling" waters. The (001) face also leaves half cages of water, which may or may not hold a methane molecule, accessible to the solution. These half cages are the site where further growth of the hydrate crystal will proceed, therefore for the AFP to inhibit growth those are the very growth planes that must be blocked. Therefore to look at other possible hydrate planes for the AFP binding, where there is no possibility for hydrate growth is not required. These observations lead us to suspect that the mechanism of binding of *wf*-AFP on the hydrate surface will be substantially different from its binding to the ice surface. Such studies lead to a more comprehensive understanding of the function of AFPs and will be a first step in the rational design of new green and more efficient kinetic hydrate inhibitor molecules with potential superior activity.

Computational Methods

Three sets of simulations at 275 K and 800 bar were performed: APF/Water to equilibrate the protein in water under the simulation conditions (refer to the ESI for the analysis of this simulation) and two production runs where the interactions between the *wf*-AFP and the hydrate surface are closely monitored. A 7×7×3 unit cell replica block of methane hydrate (unit cell taken form the crystallographic data[24]) with all cages occupied by methane is prepared. In the outermost layers in the *z*-direction, methane guests and some water molecules were removed to expose empty half-cages of water to the solution. One (*case I*, Fig. 2) and two (*case II*, Fig. S4 of ESI) AFP molecules are placed ~1 Å away from the hydrate surface with threonine residues (which constitute the ice binding surface of the *wf*-AFP) pointing toward the hydrate surface (see Fig. S5 of ESI for the relative position of protein with respect to the hydrate phase). The relatively stable (001) surface of the hydrate is placed in contact with water on both sides with a thickness of 3.6 nm along the *z*-axis and the system is sandwiched between two reservoirs of methane gas such that the overall composition of water and methane in the simulation cell is kept approximately equal to that of the stoichiometry of a fully occupied sI hydrate (i.e. methane mole fraction = ~0.148). The number of molecules present in each set of simulations is summarized in Table S1 of ESI. The size of the resulting simulation box is 8.42×8.42×15.64 nm.

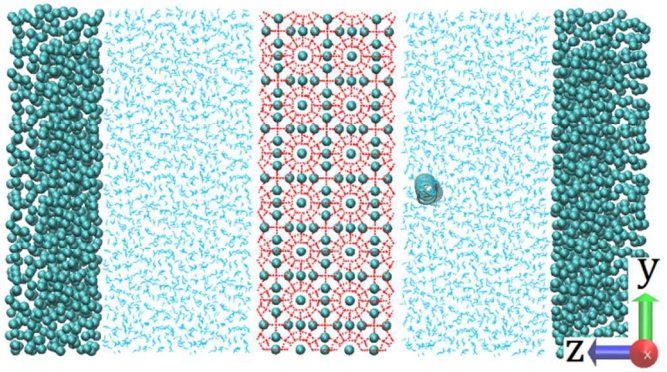

Fig. 2. *yz*-projection of the initial configurations of the *case I* simulation. The AFP is initially aligned with the long axis along the *x*-direction. *wf*-AFP, methane in the gas phase and water in the liquid phase are represented by a cyan ribbon, cyan spheres and blue lines, respectively. Hydrate water and methane are shown as red hydrogen bonds and cyan spheres in the middle of the box. Simulation box sizes along the *x*, *y* and *z*-directions are 8.421, 8.421 and 15.639 nm, respectively.

The GROMACS 4.6.1 MD code[25] was used to run all simulations with the leap-frog algorithm and a time step of 2 fs. Water and protein bonds were constrained using the LINCS algorithm and short range interactions were truncated at 1.5 nm. Coulombic interactions were treated using the particle mesh Ewald (PME) method. The temperature and pressure were

controlled by Nosé-Hoover velocity rescaling with a time constant of 0.1 ps and by Parrinello-Rahman pressure coupling with a time constant of 2 ps. Periodic boundary conditions were applied in three dimensions. Water was modelled by TIP4P-ice,[26] methane by a united atom potential,[27] and the amber99sb-ildn force field[28] was used for the protein.

In each set of simulations, to equilibrate the system, first, an energy minimization of water around the AFP was run. This was followed by a 0.5 ns constant volume–constant temperature (NVT) simulation for temperature equilibration and a subsequent 0.5 ns constant pressure–constant temperature (NPT) run to equilibrate the pressure and temperature to the final desired values. For the *case I* and *case II* simulations, the hydrate phase was kept frozen during the equilibration stages. Finally, the hydrate is unfrozen and hydrate-AFP-water-methane gas system was allowed to naturally evolve in a NPT ensemble at 275 K and 800 bar for 100-200 nanoseconds. This thermodynamic condition is well within the hydrate phase stable region for the TIP4P-ice methane hydrate force field where the equilibrium temperature for a pressure of 400 bar is estimated at 302 K[29] which is reasonably close to the experimental value of 297.23 K.[30] These conditions guarantee that there will be no large scale dissociation of the hydrate phase during simulation. The speed of calculations for the production simulations on the Westgrid computer cluster (jasper) was ~5 ns/day.

Results and discussion

The final configuration of the *case I* and *case II* simulations are shown in Figs. 3(A) and 4(A). In both cases we can see that some dissolved methane and water molecules have formed cages on the surface of the expose hydrate faces in the *z*-direction. In *case I*, at about 55 ns, the head of the AFP became anchored to the hydrate surface and remained so for the rest of simulation. At 200 ns, parts of the *wf*-AFP corresponding to residues 1 through 13 are approximately parallel to the hydrate surface. The protein bends at residue 14 such that it is roughly perpendicular to the hydrate surface between residues 15 and 25. Loops of the helix after residue 26 have unravelled (denatured), and the protein tail has deformed near the gas/water interface. The details of the bending of the protein including snapshots of the protein structure at different times as well the root mean square deviation (RMSD) profile of the protein during the *case I* simulation are given in Fig. S6 of ESI. It should be noted that despite the deformation of the protein, the residues 1 through 13 maintain their strong binding to the hydrate surface.

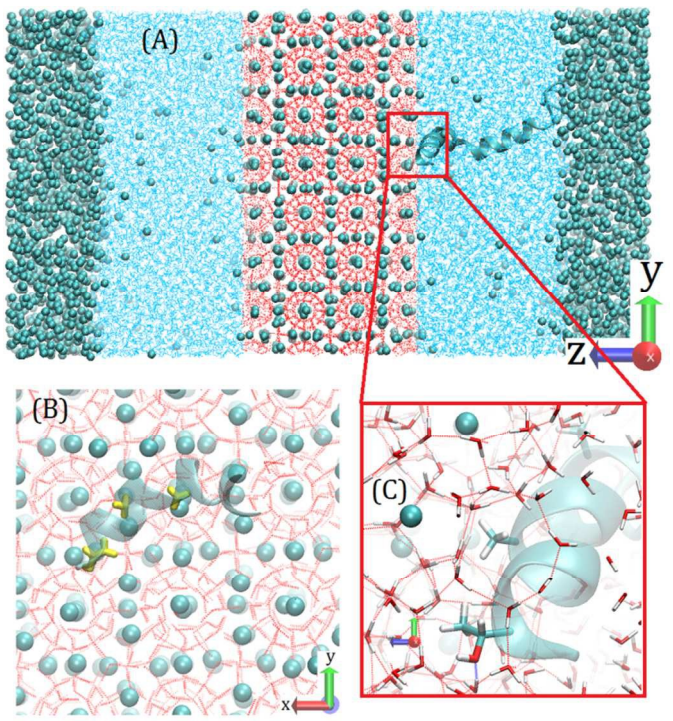

Fig. 3. (A) *yz*- and (B) *xy*-projection and (C) zoomed in yz-view of the *case I* simulation at 200 ns. The binding residues, 2THR, 6ALA and 9ALA, are shown in yellow in panel (B) and as thick bars in panel (C). Liquid water is not shown in panel (B) for clarity. The entrapment of binding residues in the empty half-cages is evident in panel (C). Hydrogen bonds are shown as red dashed lines.

In Fig. 3(B) one can see that the pendant groups of the binding residues (discussed later in the text), shown in yellow, are aligned approximately above empty half-cages of the hydrate surface. The zoomed in view of Fig. 3(C) depicts the capture of the binding 2THR, 6ALA, and 9ALA residues in the half-cages at the hydrate/water interface. A quantitative analysis of the inclusion of these pendant groups in the half-cages is discussed in Figure 5.

In the *case II* simulation, after 100 ns one protein has bound to the methane hydrate surface. Neither protein in this simulation showed bending and both maintained their helical structure throughout the 100 ns simulation (the RMSD of both of the proteins remained well below 0.4 nm). Therefore, the protein distortion behaviour observed in the *case I* simulation is not a general feature of the system.

Figure 4 shows the binding pattern of the *wf*-AFP to the methane hydrate surface in *case II*. Similar to Fig. 3(B), in Fig. 4(B), it is seen that the binding residues of the *wf*-AFP are positioned approximately above adjacent empty half-cages on the surface of the hydrate phase. Panel (C) of Fig. 4 also depicts the entrapment of these residues in the empty half-cages at the hydrate/water interface.

Note that in both *case I* and *case II* simulations, the protein molecules are not aligned in any particular orientation relative to the *xy*-plane of hydrate surface. Any set of neighbouring empty half-cages can act as binding sites for the *wf*-AFP. This observation points to an important difference with the binding

of AFPs on ice surfaces where the protein is aligned in a very specific direction with respect to a particular ice surface.[12,13]

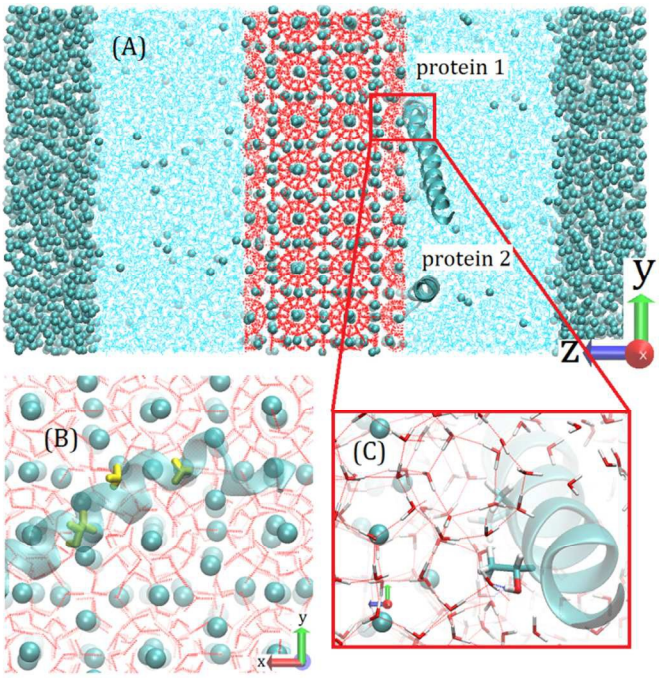

Fig. 4. (A) *yz*- and (B) *xy*-projection and (C) zoomed in yz-view of the protein 1 in the *case II* simulation at 100 ns. The binding residues, 24THR, 28ALA and 31ALA, are shown in yellow in panel (B) and as thick bars in panel (C). Liquid water is not shown in panel (B) for clarity. The entrapment of binding residues in the empty half-cages is evident in panel (C). Hydrogen bonds are shown as red dashed lines.

The F_3 order parameter

To probe the local water environment around the *wf*-AFP residues, the F_3 order parameter around the side chains (pendant groups) of all protein amino acid residues is calculated. F_3 measures the deviation of water network from the ideal tetrahedral arrangement observed in ice and hydrate phases,[31] $F_{3,i} = <[\cos\theta_{jik} \mid \cos\theta_{jik} \mid + \cos^2 104.25]^2>_{j,k}$. The angle θ_{jik} represents the angle formed by a triplet of water oxygen atoms where i is in the centre and j and k indices correspond to neighbouring water oxygen atoms which lie within a spherical shell of radius 0.35 nm around the central oxygen i.

In our simulations, F_3 values close to 0.015 correspond to water molecules arranged in hydrate-like environments, whereas values of 0.09 correspond to a liquid water environment. The F_3 for water molecules near the hydrate/water interface is expected to have the average value of ~0.05 since these water molecules have neighbours in both the ordered hydrate phase and the disordered liquid water phase. In this work the F_3 values were averaged for water molecules within a spherical shell of radius 0.6 nm around each terminal heavy atom of the pendant groups of all 37 amino acid residues of *wf*-AFP (see Table S2 of ESI). To reduce fluctuations F_3 values are averaged over 125 evenly spaced samples for each 5 ns time frame of the simulation (i.e. one sample every 40 ps).

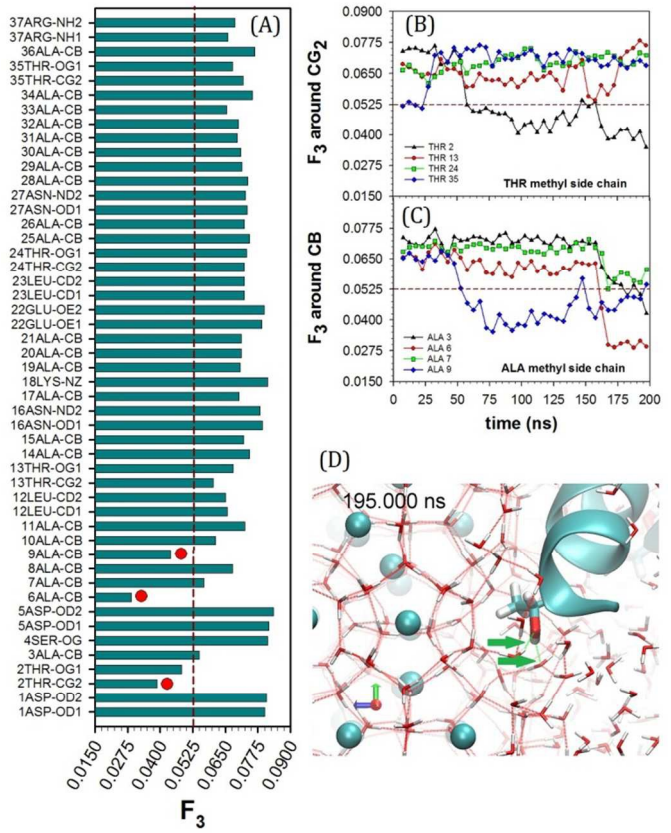

Fig. 5. (A) F_3 values for the heavy atoms of pendant side chains of all residues of *wf*-AFP for *case I* simulation at 172.5 ns. 2THR, 6ALA and 9 ALA have F_3 values below 0.05 suggesting they are trapped within the half-cages at the hydrate/water interface. Refer to Table S2 of ESI for the naming scheme of the atoms. (B) F_3 vs. time profiles of the methyl side chain of the four THR residues and (C) ALA 3, 6,7 and 9 of the *wf*-AFP in *case I* simulation. The adsorption of *wf*-AFP to the hydrate surface is a collaborative action between THR and ALA residues. In this case: 2THR, 6ALA and 9 ALA. (D) A snapshot of the *case I* simulation at 195 ns which shows that the methyl group of 2THR is trapped inside a half-cage at the hydrate surface and its hydroxyl side chain forms two hydrogen bonds, shown as green lines (emphasized by arrows), with local water molecules.

A typical F_3 profile of all residues of *wf*-AFP in *case I* simulation at 172.5 ns is plotted in Fig. 5(A). Interestingly, only 2THR, 6ALA and 9ALA residues display F_3 values below the expected interfacial value of ~0.05, suggesting that these residues are effectively incorporated into the empty half-cages at the hydrate surface.

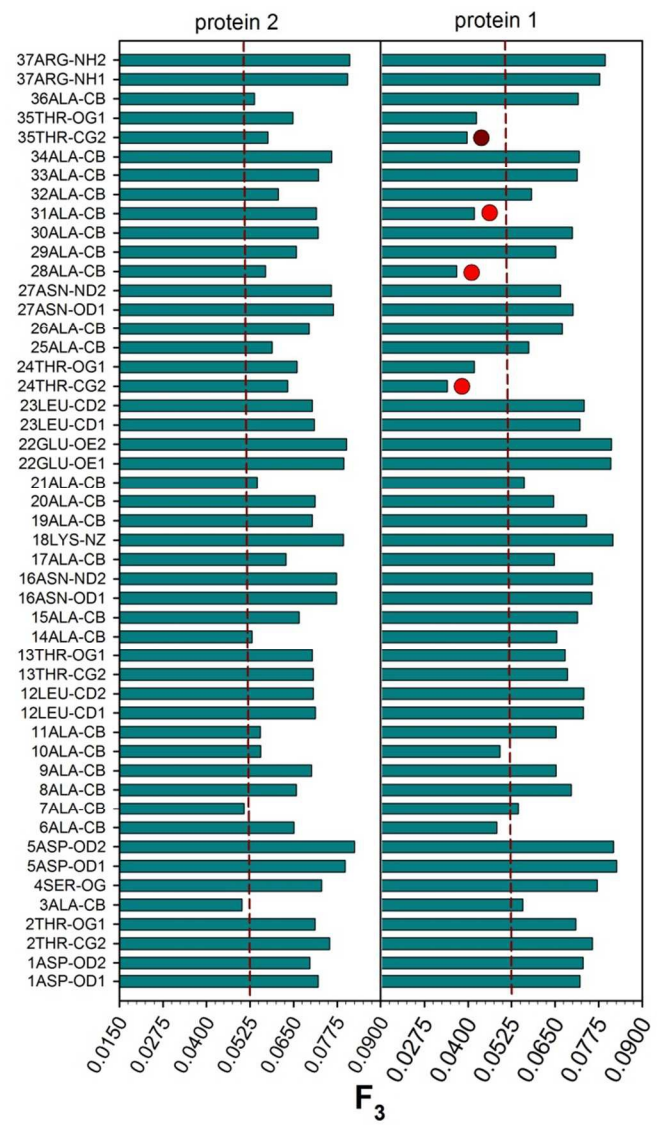

Fig. 6. F_3 values of all side chains of wf-AFP residues for *case II* simulation at 67.5 ns. For protein 1, 24 THR, 28 ALA and 31 ALA (designated by red circles) as well as 35 THR have F_3 values well below 0.0525 (red broken line). This suggests these residues are trapped within the half cages at the hydrate surface. Protein 2 is not effectively anchored to the hydrate surface at this time.

The F_3 profiles for the two protein chains for the *case II* simulation are given in Fig. 6. Only protein 1 was effectively attached to the hydrate surface and 24THR, 28ALA, 31ALA and 35THR displayed F_3 values noticeably lower than the interfacial values. Pendent groups of other *wf*-AFP residues are seen in Fig. 5(A) and Fig. 6 to lead to partial ordering of water molecules around the protein. Even residues which do not bind to the hydrate surface have F_3 values less than the solution value of 0.09.

The time variation of the order parameter, $F_3(t)$, around the methyl side chain of all THR residues and ALA residues number 3, 6, 7 and 9 during the *case I* simulation are plotted in Fig. 5(B) and (C), respectively. The carbon atoms of the THR and ALA methyl group side chains are designated as CG2 and CB. Figure 5(B) illustrates that the F_3 values of THR residues number 13, 24 and 35 remain above the interfacial value of ~0.05. Therefore, these residues are not anchored to the hydrate surface. However, F_3 of 2THR begins to decrease at ~55 ns and stays below the interfacial value afterwards suggesting that 2THR is trapped inside a half-cage at the hydrate/water interface as evident in Fig. 5(D).

The F_3 profile of ALA residues in Fig. 5(C) displays an interesting trend. F_3 of 9ALA starts to decrease at about the same time as 2THR methyl group and ~100 ns later 6ALA shows a sudden drop, which implies its entry into an empty hydrate half-cage as well. The $F_3(t)$ profiles of the two protein chains during the *case II* simulation are given in Fig. 7. For protein 1 as seen in panels (A) and (B), after about 35 ns, THR residues number 24 and 35 as well as ALA residues number 28 and 31 show F_3 values well below the interfacial value. Within the time frame of the simulation, protein 2 did not adsorb to the hydrate surface and its THR methyl F_3 values are all above the interfacial value. We note that at 100 ns, the ALA residues of protein 2 are starting to approach the interfacial values. If this simulation is continued for longer times, protein 2 may also bind to the hydrate surface.

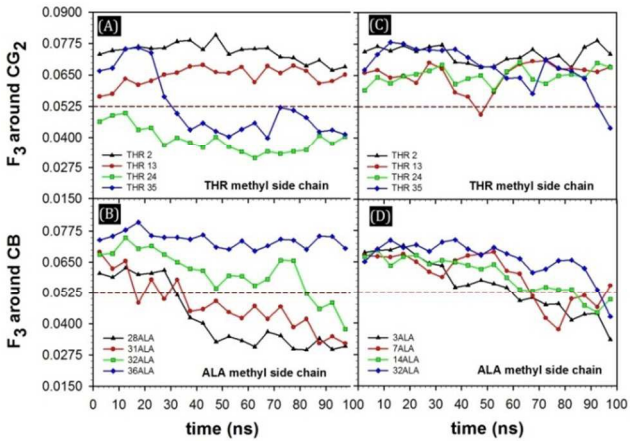

Fig. 7. F_3 vs. time profiles of the methyl side chain of (A) all of the THR residues and (B) ALA 28, 31, 32 and 36 of protein 1 in *case II* simulation. The adsorption of *wf*-AFP to the hydrate surface is a collaborative action between THR residues and ALA residues, in this case 24THR, 28ALA and 31 ALA. F_3(t) profiles of the methyl side chain of (C) all of the THR residues and (D) ALA 3, 7, 14 and 32 of protein 2. During the 100 ns of the simulation, this protein is not effectively bound to the hydrate surface. The protein 2 THR residues do not show F_3 values which are well below the interfacial value of ~0.05.

Based on ^{13}C solid-state NMR spin lattice relaxation time analysis[32] and mutations studies[16] of the AFP – ice binding, it is suggested that the IBS of *wf*-AFP consists of the THR side chains and the (*i*+4)ALA and (*i*+8)ALA residues. However, it is not microscopically determined why this particular residue combination contributes to the antifreeze action. Our results for both the *case I* and *case II* simulations indicate that the *wf*-AFP adsorbs to the methane hydrate surface by the cooperative entrapment of the pendant methyl group of THR residue at location *i* and two pendant methyl groups of (*i*+4)ALA and (*i*+7)ALA residues further down the protein sequence. These residues are in the hydrate-binding surface (HBS) of the *wf*-

AFP (i = 2, 13 or 24). The hydrophobic methyl pendant groups tend to enter into the empty half cages at the methane hydrate interface and effectively act as a guest in the empty hydrate half-cages. This stabilizes the hydrate surface by optimizing the *hydrophobic interaction* of methyl groups with surrounding polar water molecules, similar to what occurs in the formation of gas hydrates themselves from liquid water and methane gas. The size of the methyl group very closely resembles that of CH_4 thus making it an ideal guest to stabilize hydrate cages. An interesting feature observed in Figs. 5(B) and 7(A) is that for *wf*-AFP to adsorb on the hydrate surface, all four THR residues may not be necessary and *wf*-AFP can bind to the hydrate surface via only one or two of its four THR residues. This binding mechanism is different from *wf*-AFP binding on ice which is a collaborative process involving the entire protein molecule.

Hydrogen bonding analysis

The (*i*)THR hydroxyl group can form hydrogen bonds with nearby water molecules at the hydrate surface, leading to further stabilization of the *wf*-AFP binding at the hydrate surface. Fig. 5(D) is a snapshot at 195 ns of the *case I* simulation illustrating how the pendant methyl side chain of 2THR residue is trapped inside one of the half cages at hydrate surface while the THR hydroxyl group forms two hydrogen bonds (green lines) with the local water molecules.

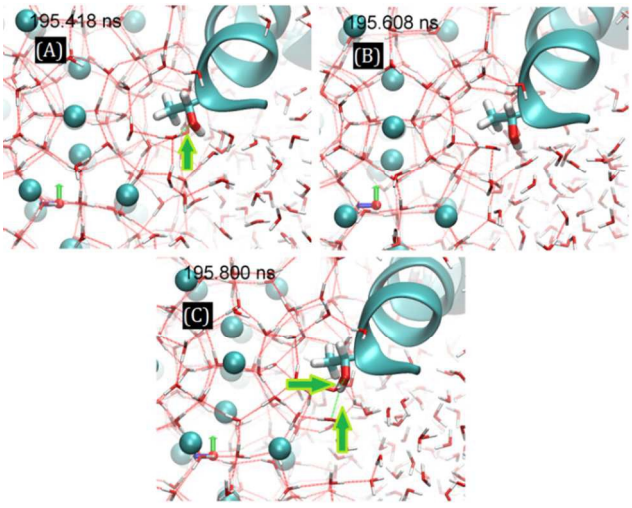

Fig. 8. The *yz*-projection of three snapshots of the *case I* simulation at (A) 195.418 ns, (B) 195.608 and (C) 195.800 ns. The 2THR methyl group is trapped in the canter of the half-cages at the water/hydrate interface while there are 1, 0 and 2 hydrogen bonds (shown in green lines) between the hydroxyl group and nearby water molecules.

Fig. 8 shows three additional snapshots of the 2THR residue in the *case I* simulation at 195.418, 195.608 and 195.800 ns where it can be seen that the methyl group is trapped in the centre of a half-cage at the water/hydrate interface and the hydroxyl group forms 0, 1 or 2 hydrogen bonds with the nearby water molecules, respectively.

The average number of hydrogen bonds between the pendant OH groups of individual THR residues and neighbouring water molecules, given in Table 1, are not significantly different between *wf*-AFP in water and the *case I* where *wf*-AFP interacts with the hydrate surface. Moreover, the hydrogen bonding in the first 5 ns of the *case I* simulation where the *wf*-AFP was not adsorbed to the hydrate surface, and the last 5 ns where the protein was adsorbed to the hydrate surface are similar as well. This observation is consistent with the findings of Dalal et al.[33] who reported that no significant gain of hydrogen bonding is obtained for the *wf*-AFP/ice complex in the interfacial region compared to the solution. The similar hydrogen bonding patterns at the ice/hydrate surface and solution indicate why hydrogen bonding between THR residues and water molecules at the hydrate surface is not the main driving force for the adsorption of *wf*-AFP to the hydrate surface. Hydrogen bonding between the antifreeze molecule and the ice or hydrate surface leads to more efficient inhibition, but is not a sufficient condition for antifreeze effects. A mutant of *wf*-AFP where hydroxyl groups of THR were substituted by pendant methyl groups still showed inhibitory effects, retaining 30% to 50% of the activity of the original AFP.[15,34] The partial loss of inhibition activity is attributed to the absence of hydrogen bonding, but the fact that this mutant is still active for inhibition demonstrates that hydrogen bonding is not the structural feature solely responsible for inhibition activity.

Table 1. Average numbers of hydrogen bonds between the pendant hydroxyl group of THR residues and neighbouring water molecules

System	THR			
	2	13	24	35
AFP/Water (0-5 ns)	1.78	1.48	1.53	1.32
Case I (5-10 ns, not bound to hydrate surface)	1.84	1.57	1.61	1.75
Case I (195-200 ns, bound to hydrate surface)	1.77	1.92	1.54	1.22

Conclusions

We observe that for hydrate inhibition, *wf*-AFP is bound to the hydrate surface by cooperative anchoring of the pendant methyl side chain of (*i*)THR, (*i*+4)ALA and (*i*+7)ALA residues to the half-cages of the hydrate framework at the hydrate/water interface. However, compared to the ice inhibition where AFP adsorption occurs along a very particular direction on the ice surface, in hydrate inhibition the *wf*-AFP has freedom to align on the hydrate surface and position its methyl side chains from THR and ALA residues to optimally bind to the half-cages at the hydrate surface. The average distance of the centres of the neighbour cages of methane hydrate is between 0.602 and 0.672 nm. In *wf*-AFP, (*i*)THR - (*i*+4)ALA residues, and, (*i*+4)ALA - (*i*+7)ALA residues are spaced similarly for i = 2, 13 and 24 at (0.541, 0.558), (0.538, 0.569), and (0.551, 0.600) nm, respectively. The protein structure is flexible and can complement the hydrate surface by matching with the half-cage spacing.

The major consequence of the adsorption of *wf*-AFP on the hydrate surface is that it partially covers the hydrate surface and adds resistance to mass transfer of water and methane to the

hydrate surface, which leads to growth inhibition. The mass transfer barrier of *wf*-AFP with respect to methane is an aspect of hydrate inhibition that does not exist in antifreeze activity of *wf*-AFP on ice, where the AFP-ice interactions are far less specific. In addition, the side chain functional groups on the protein residues facing the solution structure surrounding waters differently than the hydrate, thus adding an additional entropic barrier for the hydrate lattice growth. The water cages of the hydrate surface that accommodate the pendant THR and ALA methyl groups has aspects similar to the semi-hydrate structure observed in the Maxi antifreeze protein.[35] The formation of cages of water molecules which surround hydrophobic groups may be a general aspect of the function of antifreeze proteins on ice as well as hydrates. The presence of hydrate half cages facilitates the AFP binding on gas hydrates, but the nature of the interaction is general.

The simulations identify the *wf*-AFP binding groups and other structural aspects of the protein and hydrate surface which contribute to the binding. In principle, with additional information on the pendant groups of other active AFP molecules, one should be able to design new non-biological molecules that integrate different residues with required spacing between them to attain superior performance. Simulations indicate that the non-binding sites of the *wf*-AFP protein are not essential to the operation of the protein as a hydrate inhibitor, except to give it the overall helix structure and enhancing its water solubility. At present, non-biological additives like polyvinylpyrrolidone (PVP) or polyvinylcaprolactam (PVCap) have been empirically observed to have hydrate inhibition activity. These substances do not have antifreeze activity towards ice. Compared to AFPs, these substances have very simple structures and are perhaps too simple to be very effective hydrate inhibitors. With these points in mind, it is desirable to develop hydrate inhibitor substances without the need for molecules with kilodalton molecular weights. Understanding the binding mechanism of the AFPs to the hydrate surface may lead to a substitution of these substances with more effective and economical alternatives.

Acknowledgements

We thank Compute/Calcul Canada for providing the computational facilities. All calculations of this manuscript were run in Jasper Westgrid cluster. S.A.B. would like to thank the University of British Columbia for providing the financial support through the 4YF scholarship.

Notes and references

[a] Department of Chemical and Biological Engineering, The University of British Columbia, 2360 East Mall, Vancouver BC, Canada V6T1Z3.

[b] National Research Council of Canada, 100 Sussex Dr., Ottawa ON, Canada K1A0R6.

† Electronic Supplementary Information (ESI) available. See DOI: 10.1039/b000000x/

1. D. W. Davidson, in *Water: A Comprehensive Treatise*, ed. F. Franks, Plenum, New York, 1973, vol. 2, pp. 115–234.
2. P. Englezos, *Ind. Eng. Chem. Res.*, 1993, **32**, 1251–1274.
3. J. A. Ripmeester, *Ann. N. Y. Acad. Sci.*, 2000, **912**, 1–16.
4. E. D. Sloan, *Nature*, 2003, **426**, 353–363.
5. P. A. Gale and J. W. Steed, Eds., *Supramolecular Chemistry: From Molecules to Nanomaterials*, 2012.
6. R. L. Christiansen and E. D. Sloan, *Ann. N. Y. Acad. Sci.*, 1994, **715**, 283–305.
7. E. D. Sloan, *Fluid Phase Equilibria*, 2005, **228**, 67–74.
8. M. A. Kelland, *Energy Fuels*, 2006, **20**, 825–847.
9. H. Zeng, L. D. Wilson, V. K. Walker and J. A. Ripmeester, *J. Am. Chem. Soc.*, 2006, **128**, 2844–2850.
10. N. Daraboina, P. Linga, J. Ripmeester, V. K. Walker and P. Englezos, *Energy Fuels*, 2011, **25**, 4384–4391.
11. H. Sharifi, V. K. Walker, J. Ripmeester and P. Englezos, *Cryst. Growth Des.*, 2014, **14**, 2923–2930.
12. C. A. Knight, C. C. Cheng and A. L. DeVries, *Biophys. J.*, 1991, **59**, 409–418.
13. K.-C. Chou, *J. Mol. Biol.*, 1992, **223**, 509–517.
14. F. Sicheri and D. S. Yang, *Nature*, 1995, **375**, 427–431.
15. A. D. J. Haymet, L. G. Ward, M. M. Harding and C. A. Knight, *FEBS Lett.*, 1998, **430**, 301–306.
16. J. Baardsnes, L. H. Kondejewski, R. S. Hodges, H. Chao, C. Kay and P. L. Davies, *FEBS Lett.*, 1999, **463**, 87–91.
17. P. L. Davies, J. Baardsnes, M. J. Kuiper and V. K. Walker, *Philos. Trans. R. Soc. Lond. B. Biol. Sci.*, 2002, **357**, 927–935.
18. H. Ohno, R. Susilo, R. Gordienko, J. Ripmeester and V. K. Walker, *Chem.- Eur. J.*, 2010, **16**, 10409–10417.
19. M. R. Walsh, C. A. Koh, E. D. Sloan, A. K. Sum and D. T. Wu, *Science*, 2009, **326**, 1095.
20. S. A. Bagherzadeh, P. Englezos, S. Alavi and J. A. Ripmeester, *J. Phys. Chem. B*, 2012, **116**, 3188–3197.
21. S. A. Bagherzadeh, S. Alavi, J. A. Ripmeester and P. Englezos, *Fluid Phase Equilibria*, 2013, **358**, 114–120.
22. D. R. Nutt and J. C. Smith, *J. Am. Chem. Soc.*, 2008, **130**, 13066–13073.
23. A. Wierzbicki, P. Dalal, T. E. Cheatham III, J. E. Knickelbein, A. D. J. Haymet and J. D. Madura, *Biophys. J.*, 2007, **93**, 1442–1451.
24. R. K. McMullan and G. A. Jeffrey, *J. Chem. Phys.*, 1965, **42**, 2725–2732.
25. B. Hess, C. Kutzner, D. van der Spoel and E. Lindahl, *J. Chem. Theory Comput.*, 2008, **4**, 435–447.
26. J. L. F. Abascal, E. Sanz, R. G. Fernández and C. Vega, *J. Chem. Phys.*, 2005, **122**, 234511.
27. B. Guillot and Y. Guissani, *J. Chem. Phys.*, 1993, **99**, 8075–8094.
28. K. Lindorff-Larsen, S. Piana, K. Palmo, P. Maragakis, J. L. Klepeis, R. O. Dror and D. E. Shaw, *Proteins Struct. Funct. Bioinform.*, 2010, **78**, 1950–1958.
29. M. Conde and C. Vega, *J. Chem. Phys.*, 2010, **133**, 064507.
30. E. D. Sloan and C. A. Koh, *Clathrate Hydrates of Natural Gases*, CRC Press, Boca Raton, Fl., 3rd edn., 2008.
31. L. A. Baez and P. Clancy, *Ann. N. Y. Acad. Sci.*, 1994, **715**, 177–186.
32. Y. Mao and Y. Ba, *Biophys. J.*, 2006, **91**, 1059–1068.
33. P. Dalal, J. Knickelbein, A. D. Haymet, F. D. Sönnichsen and J. D. Madura, *PhysChemComm*, 2001, **4**, 32–36.
34. W. Zhang and R. A. Laursen, *J. Biol. Chem.*, 1998, **273**, 34806–34812.

35. T. Sun, F.-H. Lin, R. L. Campbell, J. S. Allingham and P. L. Davies, *Science*, 2014, **343**, 795–798.

