Antifreeze glycoproteins (AFGPs) facilitate the survival of various organisms in the polar region by preventing internal ice accumulation via an adsorption-inhibition mechanism. Inhibition of AFGP antifreeze activity by the borate buffers has been widely acknowledged as the direct experimental evidence supporting the hydroxyl, rather than methyl, binding mechanism. On the other hand, perturbation of borate binding on the AFGP configuration, which might have considerable influence on the binding efficiency of not only the hydroxyl but also the methyl groups, has rarely been quantitatively examined. Herein we studied, using molecular dynamics simulations, the perturbation on the configuration of a solvated AFGP8 protein induced by the binding of one single borate anion. Near the freezing point, this binding not only makes the disaccharide groups adjacent to the borate-binding disaccharide close to each other but also affects the entire AFGP8 conformation. The structural changes induced by the binding of borate on different disaccharide sidechains exhibit clear site-specificities and the effect of borate binding on the structural changes is significantly reduced at higher temperatures. Our study is valuable for further understanding the relationship between the structure and antifreeze activity of these antifreeze glycoproteins.

Key words: Antifreeze glycoproteins, Borate anions, Molecular dynamics, Clustering analysis

I. INTRODUCTION

Antifreeze proteins (AFPs) and antifreeze glycoproteins (AFGPs) help inhibit the freezing and associated cell death for diverse life forms, and thereby enable their survival in extremely cold environments [1–6]. These natural proteins are of significant importance for preserving cells, tissues and organs at lower temperature [6–10], cryosurgery [11], enhancing the cold resistance of plants [12, 13], and improving the texture of frozen foods [14, 15]. AFGPs are the most effective ice recrystallization inhibitors which are orders of magnitude more potent than other AFPs [2].

AFGPs are suggested to inhibit the growth of ice via an adsorption-inhibition mechanism, in which the

*Authors to whom correspondence should be addressed. E-mail: wzhuang@fjirsm.ac.cn, klhan@dicp.ac.cn

†Part of Special Issue “John Z.H. Zhang Festschrift for celebrating his 60th birthday”.

DOI:10.1063/1674-0068/cjcp2107120 659 ©2021 Chinese Physical Society
molecules recognize and bind to embryonic ice crystals and prevent macroscopic ice growth [16]. Many aspects of this mechanism remain largely elusive. It is not yet conclusively identified, for instance, whether the ice-binding sites of AFGPs are the hydrophobic methyl groups on the peptide and disaccharides or the hydrophilic hydroxyl groups in the carbohydrate side chains.

Inhibition of AFGP antifreeze activity by the borate buffers has been widely acknowledged as one of the most direct experimental evidences supporting the hydroxyl binding picture, since borate ions can complex with the cis-hydroxyl groups of the β-D-galactopyranosyl group of AFGPs to form a four-coordinate structure [17–21]. The structure of borate-AFGP8 complex in solution is extremely complicated. Since AFGP8 contains four disaccharides of the same composition, when the total boron concentration is high enough, borate can react with the β-D-galactopyranosyl on multiple disaccharides at the same time, and cross-linking occurs between inter or intra AFGP8 molecules to form a gel [22]. Therefore, the influence of borate binding on the AFGP configuration, which might considerably affect the binding efficiency of not only the hydroxyl but also the methyl groups, remains rarely studied at the molecular level. It is yet to be systematically examined, for instance, whether even a single borate binding has a recognizable impact on the peptide conformation, and whether this impact is localized near the hydroxyl or has a larger spatial range. Moreover, the site-specificity and the temperature dependence of this impact remain elusive as well.

In this study, we simulate the impact of borate binding on the configuration distribution of AFGP8 in solution at very low total boron concentration using the replica exchange molecular dynamics (REMD) simulation and clustering analysis, which assume that only one borate ion forms a four coordination structure with the disaccharide on AFGP8. To examine the site-specificity of this impact, we studied four different structures where the borate anion binds to each of the four disaccharide moieties of AFGP8 (noted as AFGP8-B1, AFGP8-B2, AFGP8-B3, and AFGP8-B4, respectively). Our study demonstrates that this binding not only makes the disaccharide groups adjacent to the borate-binding disaccharide close to each other but also affects the entire AFGP8 conformation. The structural changes induced by the binding are related to both the binding sites and temperature.

II. METHODS

A. Atomistic models

In order to study the impact of borate on the structure of antifreeze glycoproteins, we constructed five model systems, native AFGP8 (AFGP8-noB) and four borate-binding AFGP8 (AFGP8-Bi) complexes. For the native AFGP8, the starting structure was obtained by two steps: firstly, building up the initial peptide backbone structure of native AFGP8 with VMD [23] and then adding the disaccharide groups with the help of the website GLYCAM [24]. For the four AFGP8-Bi complexes, a single borate anion was complexed with the cis-hydroxyl groups of carbons 3 and 4 of the galactose in the disaccharide moieties in the first to the fourth Thr residue of AFGP8, respectively (noted as AFGP8-B1, AFGP8-B2, AFGP8-B3 and AFGP8-B4), as shown in Fig. 1. A typical three-dimensional structure of AFGP8-B2 is shown in Fig. 2.

The peptide backbone for all systems was modelled with full atomistic details using AMBER99SB-ILDN [25], the disaccharide moieties were represented by GLYCAM06 [26], and the connection part between the peptide and the sugar using GAFF force field [27]. The partial charge parameters of Thr* and borate binding Thr* residue were computed using the restrained electrostatic potential (RESP) [28] charge-fitting scheme. The force field parameters about borate refer to the articles published by Tafi et al. [29] and Otkidach et al. [30]. The missing angle (O–B–O) parameters adopted the same parameters as the O–B–O angle proposed by Tafi and Otkidach, and the missing dihedral angles used the parameters in the GAFF force field. Using the non-bonding parameters of the OH atom connected to the phosphorus atom of the nucleic acid as the non-bonding parameters of the atom directly connected to borate. Because of the different 1–4 interaction scaling factors of the GLYCAM and Amber family force filed, we provided the LJ-14 parameters for each interaction individually.

The glycoproteins were inserted into a dodecahedron box with a minimal distance of 1.0 nm between the peptide and the box edge, and solvated with 5837 water molecules for AFGP8 and 5839 water molecules for AFGP8-Bi complexes. For AFGP8-Bi complexes, a Na+ ion was added to ensure neutral biological conditions. Water molecules were represented by the
FIG. 1 Chemical structure of the borate-binding AFGP8 complexes. The borate is attached to the disaccharide moieties in the first to the fourth Thr residue of AFGP8, respectively (noted as AFGP8-B1, AFGP8-B2, AFGP8-B3 and AFGP8-B4).

TIP4P/Ice model [31].

B. MD simulations

All molecular dynamics simulations were performed using the GROMACS 2016.4 [32, 33] software with PLUMED 2.4.2 plugin [34–36]. The energy of each system was minimized using the 10000 steepest descent steps followed by 200 ps position restrained equilibration at NVT ensemble and 500 ps position restrained equilibration at NPT ensemble. Periodic boundary conditions were used in all directions. The leap-frog algorithm was used to numerically integrate the equation of motions with a time step of 2 fs. All bonds were constrained using the LINCS algorithm [37]. To prevent energy drift we used a switching function for the non-bonded interactions from 1.0 nm to 1.1 nm. A cut-off of 1.1 nm was used for long-range electrostatic interactions. Long-range electrostatic interactions were calculated by Particle Mesh Ewald summation [38, 39] with a fourth-order interpolation and a grid spacing of 0.16 nm. The production runs were carried out in the NPT ensemble. The solute and solvent were coupled separately to a temperature bath of 300 K using a velocity-rescaling [40] thermostat with a relaxation time of 0.1 ps. The pressure coupling was fixed at 1 bar using the Parrinello-Rahman algorithm [41].

C. REMD simulations

Replica exchange MD (REMD) simulations were performed to enhance the conformational sampling for all the systems in explicit solvent. Starting from the last snapshot of the aforementioned conventional MD simulations, all replicas were performed for 250 ns in the NPT ensemble. The exchanges between neighbouring replicas were attempted every 2 ps. For native AFGP8, 57 replicas temperatures were used ranging from 270 K to 398 K based on the algorithm of van der Spoel [42] and the resulting acceptance probability was about 22%–32% on average. Temperatures ranging from 270 K to 441 K for AFGP8-B1 (64 replicas), AFGP8-B2 (63 replicas), AFGP8-B3 (64 replicas) and AFGP8-B4 (63 replicas) and the average exchange ratio between nearest neighbour temperatures were 19%–32%, 18%–29%, 19%–32% and 18%–28%, respectively. Two sets of independent 250 ns REMD simulations, starting from different initial conformations, were performed for native AFGP8 to check the convergence of the simulation by comparing the radius of gyration ($R_g$) and the number of hydrogen bonds distribution functions (see FIG. S1 in Supplementary materials). The snapshots were saved every 10 ps in order to obtain sufficient data for further analysis.

D. Analysis

The initial 50 ns trajectory was discarded as an equilibration period for further analysis. For REMD simulations, the equilibrium probability distributions of all thermodynamic quantities were computed using the multistate Bennett acceptance ratio (MBAR) method. The radius of gyration ($R_g$) was calculated for the backbone atoms for all the molecules with GROMACS tool gmx gyrate. The GROMACS tool, gmx sasa, was used to determine protein solvent accessible surface area (SASA). The backbone dihedral angles $\phi$ and $\psi$ shown on the Ramachandran plots were computed with gmx rama. The clustering analyses were performed with gmx cluster tool using the gromos method [43].

III. RESULTS AND DISCUSSION

A. Structure distribution of the solvated AFGP8 peptides

Molecular dynamics simulations were carried out to generate the trajectories of solvated AFGP8 (AFGP8-noB) as well as AFGP8-B1, AFGP8-B2, AFGP8-B3 and AFGP8-B4 samples. Replica exchange molecular dynamics (REMD) simulations [44] were used to enhance the conformational sampling, and thermodynamic properties at each temperature were computed using the multistate Bennett acceptance ratio (MBAR)
FIG. 2 (a) A schematic presentation of AFGP8-B2 structure (borate binding in the second glycosylated Thr), as well as the Ramachandran plots of (b) native AFGP8 and (c–f) the four borate-binding AFGP8 complexes at 270 K.

method [45].

FIG. 2 (b–f) presents the Ramachandran angle distributions (ϕ, ψ) of these five samples. In line with the observations of previous circular dichroism (CD), nuclear magnetic resonance (NMR) and linear infrared (IR) experiments as well as the computer simulations [46-51], AFGP8-noB samples the conformations in the PPII, β-sheet, α-helical and left-handed α-helical regions, and PPII helix dominates the distribution (FIG. 2(b)). The Ramachandran distributions of the four AFGP8-Bi structures are very similar to that of the AFGP8-noB (FIG. 2(c–f)). It was consistent with the linear infrared spectra study by Bakker and co-workers which suggested that adding borate hardly changes the secondary structure of AFGP in solution [51]. Single borate binding at the cis-hydroxyl groups of disaccharides, therefore, leads to no significant difference in the dihedral angle distributions.

We further examined the gyration radius (Rg) and the solvent accessible surface area (SASA) for AFGP8-noB and AFGP8-Bi at 270 K. The details of Rg and SASA calculations are provided in the Methods section. FIG. 3 (a) and (b) show the projections of the free energy landscape onto these two dimensions. Rg indicates the compactness of the protein. All five samples explore a wide conformational space and adopt the extended structure. The values of Rg for all the five systems are mainly within a similar range of 8–13 Å. The maximum probability of Rg of AFGP8-noB (11.55 Å) is slightly smaller than the four AFGP8-Bi structures: AFGP8-B1 (12.25 Å), AFGP8-B2 (12.25 Å), AFGP8-B3 (12.15 Å), and AFGP8-B4 (12.05 Å). The distributions of protein SASA of all structures almost overlap, and do not show a significant difference in the range of 23–31 nm² S⁻² N⁻¹. Neither of these two order parameters, therefore, is significantly perturbed by the single borate binding.

B. Clustering analysis of AFGP8 configuration distributions

To further rationalize the subtle conformational perturbation induced by the single borate binding, we carried out a clustering analysis. The Daura algorithm [43] was employed and the root-mean-square deviation
FIG. 3 Projections of the potential of mean force onto the dimension of (a) \(R_g\) and (b) protein SASA of the native AFGP8 as well as the four borate-bound AFGP8 structures at 270 K.

(RMSD) of all protein C\(\alpha\) atoms and all the heavy atoms at the sidechain carbohydrate ring was used for the clustering. A cut-off value of 4.2 Å was used to group the AFGP8-noB configuration ensemble at 270 K into 90 clusters with the top ten clusters accounting for 83.6% of the total population. The population decreases exponentially with the increasing cluster index and becomes less than 2% after the seventh cluster (FIG. 4(a)). In the following, we denoted this ensemble of clusters (states) as \(\{S_i^{\text{whole}}\}\).

We then constructed the population distributions of AFGP8-B\(i\) on the basis of \(\{S_i^{\text{whole}}\}\) by calculating the RMSD between each configuration in the AFGP8-B\(i\) structure ensembles and the central structures of each state in \(\{S_i^{\text{whole}}\}\), and assigned this AFGP8-B\(i\) configuration to the state \(S_k^{\text{whole}}\) with the smallest RMSD.

We first examined the difference of populations between the AFGP8-noB and the AFGP8-B2 complex (FIG. 4 (a) and (b)). Most of the AFGP8-noB populations (\(\sim \)83.0%) stay in the first 6 states (FIG. 4(a)). The overall population change of AFGP8-B2 with respect to AFGP8-noB, defined as the sum of decreased (increased) populations in those states with populations reduced (enhanced) by the borate binding is 49.8%. Compared to that in AFGP8-noB, the state 1 population in AFGP8-B2 has the largest change of \(-27.6\%\) (from 44.5% to 16.9%). We can define the states with “non-negligible population change” as those states with a change of population no smaller than 2.7%, which is 10% of this largest population change. For AFGP8-B2, states 1, 2 and 5 have non-negligible population decrease of 27.6%, 5.4% and 4.9%, respectively, and states 4, 21, 33 and 36 have non-negligible population increase of 13.6%, 3.8%, 6.6% and 3.4%, respectively (FIG. 4(b)). Representative structures of these clusters with significant population changes are shown in FIG. 5. Rest of the states have rather insignificant population changes (<2.7%). For the states with a significant increase in population, their structure shows that the second disaccharide group and the third disaccharide group are close to each other, as shown in the second row of FIG. 5. This is mainly because that the boron atom connected to the second disaccharide group in AFGP8-B2 has a relatively large negative charge and produces a large electrostatic attraction interaction with the positively charged third sugar group. Experiments find that AFGP8 binds to ice through hydroxyl groups on disaccharides [22, 52]. The borate-binding to AFGP8 changed the distance between the hydroxyl groups on the disaccharide groups, which may cause the mismatch between the ice-binding sites of the protein and the ice lattice. This makes it difficult to bind to the ice surface, which in turn leads to the inhibition of antifreeze activity for AFGP8.

To examine whether the influence of single borate binding is localized in vicinity of the binding site, we carried out a similar clustering procedure to construct an ensemble of states \(\{S_i^{\text{noTHG2}}\}\) for AFGP8-noB, with the bound THG2 residue excluded. We then assigned each structure of AFGP8-B2 to \(S_i^{\text{noTHG2}}\) and examined the population difference between AFGP8-noTHG2 and the AFGP8-B2-noTHG2 (FIG. 4(c)). The overall population change of AFGP8-B2-noTHG2 with respect to AFGP8-noTHG2 is 19.6%. States 1, 3, and 4 have non-negligible population decreases of 31.6%, 5.0% and 4.3%, while state 2, 6, and 19 have non-
FIG. 4 (a) Cluster populations of the native AFGP8 and AFGP8-B2 structures with the cluster of AFGP8 as the base set at 270 K and (b) cluster populations with the population difference between AFGP8-B2 and native AFGP8 greater than 2.7%. (c) Clustering analysis of the AFGP8-noTHG2 and AFGP8-B2-noTHG2 with the states of AFGP8-noTHG2 as the basis set at 270 K. Here, THG denotes the glycosylated THR residue. AFGP8-noB with the second THG residue removed is termed AFGP8-noTHG2. AFGP8-B2 removes the borate-binding THG residue called AFGP8-B2-noTHG2. (d) Cluster populations of AFGP8-noTHG2 and AFGP8-B2-noTHG2 with the population difference between AFGP8-B2-noTHG2 and AFGP8-noTHG2 greater than 2.7% are shown here.

FIG. 5 Representative structures of AFGP8 clusters with population changes significantly between AFGP8-B2 and native AFGP8. For AFGP8-B2, the populations of the three clusters in the first row decreased and the populations of the four clusters in the second row increased.

negligible population increases of 9.4%, 4.5% and 4.5% (FIG. 4(d)). The influence of borate binding on the AFGP configurations is therefore not limited locally to the directly bound THG2 residue. It shows that in addition to the short-range interaction between borate and the disaccharides on the protein, there are also long-range interactions, which is caused by van der Waals forces. Note that with borate binding, the indices of states with non-negligible increment of population are different in \( \{S_{\text{whole}}^i\} \) and \( \{S_{\text{noTHG2}}^i\} \), this does not imply that the conformational variations are different with/without considering the bound THG2 residue, since the cluster grouping is different in \( \{S_{\text{whole}}^i\} \) and \( \{S_{\text{noTHG2}}^i\} \).

C. Influence of borate binding at different sites on the structure of AFGP8

We next examined the site-specificity of the borate-binding effect on the AFGP8 configurations by carrying out a similar analysis on the structure of other AFGP8-Bi complexes (FIG. 6). Population decreases by 11.0%, 6.8% and 4.5% for state 1, 2 and 5, and increases by 2.9%, 3.2%, 4.4%, 3.8%, 3.8% for states 4, 21, 27, 28, 36 for AFGP8-B1 (with an overall population change of 34.0%). For AFGP8-B3 (with an overall population change of 35.1%), populations of states 1, 2 and 5 de-
FIG. 6 (a) Cluster populations of the native AFGP8 and the four AFGP8-B complexes with the cluster of AFGP8 as the base set at 270 K, and (b) cluster population of AFGP8-noB and (c) cluster population difference between AFGP8-B structures and AFGP8-noB. Only clusters with a population difference greater than 2.7% are shown here.

crease by 24.6%, 6.6% and 4.5%, while the population of other four states (3, 4, 21, 27) increase by 7.6%, 2.8%, 4.6% and 4.6%. For AFGP8-B4 (with an overall population change of 30.3%), populations of states 1, 2, and 5 decrease by 18.9%, 6.7%, and 4.3%, and that of the 4 states 4, 21, 22, 33 increase by approximately 3.9%, 3.7%, 3.1%, and 3.9%. In FIG. S2 in Supplementary materials, we also present the representative structures of these clusters with remarkable population changes.

Interestingly, at this low temperature of 270 K, the binding of borate at all the four different sites leads to the decrement of populations at the three states 1, 2 and 5. On the other hand, a clear site-specificity is observed for the effect of borate-binding on the AFGP8 structures. When borate binds to the central two disaccharide sites 2 and 3, the two adjacent disaccharide groups in the middle on AFGP8 approach to each other, but when borate binds to the end disaccharide sites 1 and 4, the distance between disaccharide groups do not change significantly. The electrostatic field generated by the negatively charged boron atom induces the orientation polarization of the water molecules in solution, which gives rise to the polarized electric field. This shields part of the electrostatic field generated by the boron atoms. Since the aqueous environment of borate binding to the end disaccharide sites for AFGP8 is more uniform, the Coulomb shielding effect is stronger than that of the borate binding to the central two disaccharides of the protein. The stronger Coulomb shielding effect weakens the effect of borate on the structure of antifreeze protein. This indicates that the binding site of borate on antifreeze glycoproteins has an important impact on the structure and antifreeze activity of antifreeze glycoprotein.

We further studied the locality of borate-binding influence on the structure of AFGP8-B1, AFGP8-B3 and AFGP8-B4 by constructing the ensembles of states \( \{S_{noTHG2}\}_i \), \( \{S_{noTHG3}\}_i \), and \( \{S_{noTHG4}\}_i \) for AFGP8-noB using a similar procedure to that of \( \{S_{noTHG2}\}_i \). The population differences between AFGP8-noTHG and the AFGP8-Bi-noTHG \( i \) were examined in FIG. 7. The cluster populations of all the four AFGP8-noTHG \( i \) structures are very similar (see FIG. S3 in Supplementary materials), however, the cluster populations of all the four AFGP8-Bi-noTHG \( i \) structures are clearly changed compared with AFGP8-noTHG \( i \). There are 6, 6, 7, and 6 states with non-negligible population changes for the four AFGP8-Bi-noTHG \( i \) structures (see Table S1 in Supplementary materials). The influence of borate binding on different THG residues is therefore not limited locally to the vicinity of THG residue but on the entire structure. We also note that these states with non-negligible population changes are different. This indicates that the binding of borate to different THG residues has different effects on the part of AFGP structure beyond the immediate binding site. This effect is not only caused by short-range electrostatic interactions, but also by long-range van der Waals interactions.

D. Influence of borate on the structure of AFGP8 at a higher temperature of 300 K

At the room temperature of 300 K, we again examined the changes of Ramachandran plot, \( R_g \) and SASA by the single borate binding (FIG. S4, FIG. S5 in supplementary materials). Compared with AFGP8-noB, these order parameters demonstrate even less difference comparing to those at 270 K. PPII helix is still dominant and there is no new secondary structure at
FIG. 7 Clustering analysis of the AFGP8-noTHGi and AFGP8-Bi-noTHGi with the cluster of AFGP8-noTHG as the base set at 270 K. AFGP8-noB with the first to fourth THG residue removed are termed AFGP8-noTHG1, AFGP8-noTHG2, AFGP8-noTHG3 and AFGP8-noTHG4, respectively. The AFGP8-Bi remove the borate-binding THG residue called AFGP8-B1-noTHG1, AFGP8-B2-noTHG2, AFGP8-B3-noTHG3, AFGP8-B4-noTHG4, respectively. (a) Cluster populations of AFGP8-noTHG1 and AFGP8-B1-noTHG1, (b) AFGP8-noTHG2 and AFGP8-B2-noTHG2, (c) AFGP8-noTHG3 and AFGP8-B3-noTHG3 and (d) AFGP8-noTHG4 and AFGP8-B4-noTHG4. Only clusters with a population difference greater than 2.7% are shown here.

room temperature, however the conformation distribution becomes broader, indicating more flexible structure at higher temperatures. The free energy landscape a wider range (7–13 Å) for all the $R_g$ sample and almost overlap in 300 K. We also observe that the maximum probability of $R_g$ of the four AFGP8-Bi structures is around 12 Å, which is the same as that at 270 K. The solvent accessible surface areas (SASAs) of all structures have almost no difference.

We next constructed the population distributions of the configuration ensemble of AFGP8-noB and AFGP8-Bi at 300 K on the basis of $\{S_i^\text{whole}\}$ at 270 K using the similar procedure to that in the previous section. FIG. 8 presents the cluster populations of the top nine clusters ranked by the absolute population differences between AFGP8 and AFGP8-Bi ($i=1, 2, 3, 4$). The overall population differences between AFGP8-noB and AFGP8-Bi at 300 K is much less significant compared to that at 270 K, which are 18.7%, 22.0%, 17.1% and 14.1% for AFGP8-Bi, respectively. For AFGP8-B1, state 1 has a non-negligible population decrease of 8.6%. Populations of states 1 and 3 decrease by 3.4% and 3.2%, and those of states 4 and 6 increase by 3.8% and 2.8% for AFGP8-B2. None of the states in AFGP8-B3 and AFGP8-B4 has significant population change higher than 2.7% (FIG. 8 (c) and (d)). Single borate binding therefore leads to a much less significant and less state-specific conformational change at 300 K compared to that at 270 K. At higher temperatures, the vibration amplitudes of the system at the equilibrium position are larger than that at 270 K, and thermal fluctuations reduce the effect of electrostatic interaction on the structure, thus leading to the AFGP8-Bi systems are more similar to the AFGP8-noB. Inhibition of the antifreeze activity of antifreeze glycoproteins by borates is more pronounced at low temperatures.

As a further demonstration, we cast the configuration ensembles of AFGP8-noTHGi and the AFGP8-Bi-noTHGi at 300 K on the basis of $\{S_i^\text{noTHG}\}$ at 270 K to generate their population difference (FIG. 9). The effect of borate binding is much more local at this higher temperature. Cluster populations of the four AFGP8-noTHGi and AFGP8-Bi-noTHGi structure pairs are very similar, but the populations of states between each structure pair indeed show some difference. There are 1, 3, and 1 states in the three AFGP8-Bi-noTHGi ($i=1, 2, 3$) structure pairs with population changed more than 2.7%. Cluster populations changes of states in
 FIG. 8 Cluster populations of the AFGP8 and AFGP8-B*i complexes at 300 K using the states of AFGP8 at 270 K as the base set. (a) Cluster populations of AFGP8 and AFGP8-B1, (b) AFGP8 and AFGP8-B2, (c) AFGP8 and AFGP8-B3 and (d) AFGP8 and AFGP8-B4 complexes. Only the top nine clusters ranked by the absolute population differences between AFGP8 and AFGP8-B*i are shown here.

 FIG. 9 Clustering analysis of the AFGP8-noTHG*i and AFGP8-B*i-noTHG*i at 300 K with the cluster of AFGP8-noTHG*i at 270 K as the base set. (a) Cluster populations of AFGP8-noTHG1 and AFGP8-B1-noTHG1, (b) AFGP8-noTHG2 and AFGP8-B2-noTHG2, (c) AFGP8-noTHG3 and AFGP8-B3-noTHG3 and (d) AFGP8-noTHG4 and AFGP8-B4-noTHG4. Only the top nine clusters ranked by the absolute population differences between AFGP8 and AFGP8-B*i are shown here.

AFGP8-B4-noTHG4 structure pairs all show difference less than 2.7%. The influence of borate on the structure of AFGP8 at 300 K is not just limited to the vicinity of the borate-binding THG residues, but determined by the whole structure. At higher temperatures, the thermal fluctuations compete with the electrostatic interactions and weaken the effect of electrostatic interactions on the structures. The water molecules move faster, and the water molecules around the boron are more disordered. The water molecules have a stronger Coulomb shielding effect on the electrostatic field generated by the boron, which makes the long-range interaction weaker. Overall, the combined effect of thermal fluctuations and Coulomb shielding weakened the long-
range interaction of borate on antifreeze glycoproteins.

IV. CONCLUSION

We carried out a computational analysis of the single borate-binding perturbation on the solvated AFGP8 glycoproteins configuration using REMD simulations. Our research demonstrates that even a single borate binding to disaccharide group for AFGP8 in solution leads to a significant structural change of protein. These changes in the AFGP8 structure caused by borate binding not only make the disaccharide groups adjacent to the borate-binding disaccharide close to each other, but also affect the entire AFGP8 conformation.

The structural changes are mainly because the boron atom attached to the disaccharide group in AFGP8-Bi has a relatively large negative charge and produces a large electrostatic attraction interaction with their adjacent positively charged sugar group. In addition, there are long-range interactions, which are caused by van der Waals forces. The borate-binding to AFGP8 changed the distance between the hydroxyl groups on the disaccharide groups, which may cause the mismatch between the ice-binding sites of the protein and the ice lattice. It is thus difficult for AFGP8 to bind to the ice surface, which in turn leads to the decrement of antifreeze activity. Besides, the binding site of borate on antifreeze glycoprotein and temperature also have an important impact on the structure and antifreeze activity of antifreeze glycoprotein. Therefore, our study on the impact of borate on the structure of antifreeze glycoproteins provides some guidance for understanding the relationship of structure and antifreeze activity of antifreeze glycoproteins.

Supplementary materials: Simulation convergence check is shown in FIG. S1. FIG. S2 presents the representative structures of AFGP8 clusters with population changes significantly between AFGP8-Bi and native AFGP8. FIG. S3 shows cluster populations of the four AFGP8-noTHGi structures at 270 K. Cluster populations with the population changes more than 2.7% between AFGP8-Bi-noTHGi and AFGP8-noTHGi at 270 K in Table S1. Structure distribution of the solvated AFGP8 peptides at 300 K are shown in FIG. S4 and FIG. S5.

V. ACKNOWLEDGMENTS

We thank Professor Lu Zhang for the helpful discussions of analysis. Financial support from the National Natural Science Foundation of China (No.21873101) is acknowledged.

[1] M. M. Harding, P. I. Anderberg, and A. D. J. Haymet, Eur. J. Biochem. 270, 1381 (2003).
[2] M. Bar Dolev, I. Braslavsky, and P. L. Davies, Annu. Rev. Biochem. 85, 515 (2016).
[3] C. I. Biggs, T. L. Bailey, G. Ben, C. Stubbs, A. Fayter, and M. I. Gibson, Nat. Commun. 8, 1546 (2017).
[4] Z. He, K. Liu, and J. Wang, Accounts Chem. Res. 51, 1082 (2018).
[5] L. Weng, S. L. Stott, and M. Toner, Annu. Rev. Biomed. Eng. 21, 1 (2019).
[6] T. Chang and G. Zhao, Adv. Sci. 8, 34 (2021).
[7] S. Liang, B. Yuan, J. W. Kwon, M. Ahn, X. S. Cui, J. K. Bang, and N. H. Kim, Theriogenology 86, 485 (2016).
[8] V. Robles, D. G. Valcarce, and M. E. Riesco, Biomolecules 9, 181 (2019).
[9] M. J. Taylor, B. P. Weegman, S. C. Baicu, and S. E. Giwa, Transfus. Med. Hemoth. 46, 197 (2019).
[10] R. P. Tas, V. Sampaio-Pinto, T. Wennekes, L. W. van Laake, and I. K. Voets, Embo Rep. 22, 52162 (2021).
[11] A. Shitzer, J. Heat Transfer 133, 011005 (2011).
[12] H. K. Khanna and G. E. Daggard, Plant Cell Rep. 25, 1336 (2006).
[13] B. Zhu, A. S. Xiong, R. H. Peng, J. Xu, X. F. Jin, X. R. Meng, and Q. H. Yao, Mol. Biol. Rep. 37, 961 (2010).
[14] W. Boonsupthip and T. C. Lee, J. Food Sci. 68, 1804 (2003).
[15] X. Liu, Y. Pan, F. Liu, Y. He, Q. Zhu, Z. Liu, X. Zhan, and S. Tan, J. Nanomater. 2021, 9990709 (2021).
[16] J. A. Raymond and A. L. Devries, Proc. Natl. Acad. Sci. USA 74, 2589 (1977).
[17] M. Levy, J. Biol. Chem. 84, 763 (1929).
[18] C. A. Zittle, Adv. Enzymol. Relat. Subj. Biochem. 12, 493 (1951).
[19] S. Svensson, S. G. Hammarström, and E. A. Kabat, Immunochemistry 7, 413 (1970).
[20] A. L. Devries. Science 172, 1152 (1971).
[21] T. E. Acree, The Chemistry of Sugars in Boric Acid Solutions, in: H. Carbohydrates in Solution, S. Isbell Ed., Washington, DC: American Chemical Society, 208 (1973).
[22] C. C. M. Groot, K. Meister, A. L. DeVries, and H. J. Bakker, J. Phys. Chem. Lett. 7, 4836 (2016).
[23] W. Humphrey, A. Dalke, and K. Schulten. J. Mol. Graph. 14, 33 (1996).
[24] GLYCAM (http://www.glycam.or).

[25] K. Lindorf-Larsen, S. Piana, K. Palmro, P. Maragakis, J. L. Klepeis, R. O. Dror, and D. E. Shaw, Proteins 78, 1950 (2010).

[26] K. N. Kirschner, A. B. Yongye, S. M. Tschampel, J. González-Outeiriño, C. R. Daniels, B. L. Foley, and R. J. Woods, J. Comput. Chem. 29, 622 (2008).

[27] J. M. Wang, R. M. Wolf, J. W. Caldwell, and P. A. Kollman, J. Comput. Chem. 25, 1157 (2004).

[28] C. I. Bayly, P. Cieplak, W. Cornell, and P. A. Kollman, J. Phys. Chem. 97, 10269 (1993).

[29] M. J. Abraham, T. Murtola, R. Schulz, S. Pall, J. C. Smith, B. Hess, and E. Lindahl, SoftwareX 1-2, 19 (2015).

[30] M. Bonomi, D. Branduardi, G. Bussi, C. Camilloni, D. Provasi, P. Raiteri, D. Donadio, F. Marinelli, F. Pietrucci, R. A. Broglia, and M. Parrinello, Comput. Phys. Commun. 180, 1961 (2009).

[31] G. A. Tribello, M. Bonomi, D. Branduardi, C. Camilloni, and G. Bussi, Comput. Phys. Commun. 185, 604 (2014).

[32] M. Bonomi, G. Bussi, C. Camilloni, G. A. Tribello, P. Banáš, A. Barducci, M. Bernetti, P.G. Bolhuis, S. Bottero, D. Branduardi, R. Capelli, P. Carloni, M. Ceriotti, A. Cesari, H. Chen, W. Chen, F. Colizzi, S. De, M. De La Pierre, D. Donadio, V. Drobot, B. Enssing, A. L. Ferguson, M. Filizola, J. S. Fraser, H. Fu, P. Gasparotto, F. L. Gervasio, F. Giberti, A. Gil-Ley, T. Giorgio, G. T. Heller, G. M. Hocky, M. Iannuzzi, M. Invernizzi, K.E. Jels, A. Jusupow, E. Kirilin, A. Laio, V. Limongelli, K. Lindorf-Larsen, T. Lühr, F. Marinelli, L. Martin-Samos, M. Masetti, R. Meyer, A. Michaelides, C. Molteni, T. Morishita, M. Nava, C. Paissoni, E. Papaleo, M. Parrinello, J. Pfaendtner, P. Piaggi, G. Piccini, A. Pietropaolo, F. Pietrucci, S. Pipolo, D. Provasi, D. Quigley, P. Raiteri, S. Ranilo, J. Rydzewski, M. Salvataggio, G.C. Sosso, V. Spiwok, J. Šponer, D. W. H. Swenson, O. Valsson, M. Vendruscolo, G. A. Voth, and A. White, Nat. Methods 16, 670 (2019).

[33] B. Hess, H. Bekker, H. J. C. Berendsen, and J. G. E. M. Fraaije, J. Comput. Chem. 18, 1463 (1997).

[34] T. Darden, D. York, and L. Pedersen. J. Chem. Phys. 98, 10089 (1993).

[35] U. Essmann, L. Perera, M. L. Berkowitz, T. Darden, H. Lee, and L. G. Pedersen, J. Chem. Phys. 103, 8577 (1995).

[36] M. Parrinello and A. Rahman, J. Appl. Phys. 52, 7182 (1981).

[37] A. Patriksson and D. van der Spoel, Phys. Chem. Chem. Phys. 10, 2073 (2008).

[38] X. Daura, K. Gademann, B. Jaun, D. Seebach, W. F. van Gunsteren, and A. E. Mark, Angew. Chem. Int. Edit. 38, 236 (1999).

[39] Y. Sugita and Y. Okamoto, Chem. Phys. Lett. 314, 141 (1999).

[40] M. R. Shirts and J. D. Chodera, J. Chem. Phys. 129, 124105 (2008).

[41] C. A. Bush and R. E. Feeney, Int. J. Peptide Protein Res. 28, 386 (1986).

[42] A. N. Lane, L. M. Hays, L. M. Crowe, J. H. Crowe, and R. E. Feeney, Protein Sci. 7, 1555 (1998).

[43] A. N. Lane, L. M. Hays, N. Tsvetkova, R. E. Feeney, L. M. Crowe, and J. H. Crowe, Biophys. J. 78, 3195 (2000).

[44] S. S. Mallajosyula, K. Vanommeslaeghe, and A. D. MacKerell Jr., J. Phys. Chem. B 118, 11696 (2014).

[45] K. Mochizuki and V. Molinero, J. Am. Chem. Soc. 140, 4803 (2018).

[46] G. Giubertoni, K. Meister, A. L. DeVries, and H. J. Bakker, J. Phys. Chem. Lett. 10, 352 (2019).

[47] C. A. Knight, E. Driggers, and A. L. DeVries, Biophys. J. 64, 252 (1993).