Green Chemistry

Accepted Manuscript

ROYAL SOCIETY OF CHEMISTRY

www.rsc.org/greenchem

Ionic Liquid Induced Inactivation of Cellobiohydrolase I from *Trichoderma Reesei*

Weifeng Li[1,2], Lushan Wang[3], Ruhong Zhou[1,4,5]* and Yuguang Mu[2]*

1. *Institute of Quantitative Biology and Medicine, School for Radiological and Interdisciplinary Sciences (RAD-X) & Collaborative Innovation Center of Radiation Medicine of Jiangsu Higher Education Institutions, Soochow University, Suzhou 215123, China*
2. *School of Biological Sciences, Nanyang Technological University, 637551, Singapore*
3. *The State Key Laboratory of Microbial Technology, Shandong University, Jinan, 250100, China*
4. *Computational Biological Center, IBM Thomas J. Watson Research Center, Yorktown Heights, NY 10598, USA.*
5. *Department of Chemistry, Columbia University, New York, NY 10027, USA*

E-mail: ygmu@ntu.edu.sg; ruhong@us.ibm.com

Abstract

Ionic liquids (ILs) exhibit unique capabilities in dissolving cellulose, the most abundant bioorganic material on earth, thus providing another promising source for biofuels. In this study, structural stability and interactions of cellulase cellobiohydrolase I (CBHI) in ionic liquid (1-butyl-3-methylimidazolium chloride, [Bmim]Cl) solution were investigated through computer simulations. At a moderate low concentration (0.8 M) of [Bmim]Cl, the cellulase CBHI showed both thermo-stability and IL-resistance at room temperature as well as a higher temperature of 350 K. However, several $[Bmim]^+$ cations were found to be able to intrude into the interior of cellulase, especially near the cellulose binding tunnel and the active site for hydrolysis. These $[Bmim]^+$ can compete with the binding of cellulose substrate and directly block the hydrolysis pathway, which are responsible for the cellulase inactivation found in experiment. Three important residues, H228, W376 and R251, are identified to potentially improve the IL-resistance

on cellulase. Motivated by the experiment, *in silico* mutagenesis studies of H228R have been conducted to examine the IL binding at the active site, in which the binding affinity of $[Bmim]^+$ in the mutant is significantly reduced, by 20%, as compared to the wild type due to the repulsive interactions from the new arginine residue. Our findings have uncovered the inactivation mechanism of cellulase in the IL environment and pointed out the potential solutions for industrial-use cellulase enzymes in IL through protein engineering.

1. Introduction

The ever growing consumption of fossil fuels has led to the energy crisis as a result of the rapid depletion of fossil fuel reserves (the consequences of peak oil)[1-5], thereby stunting world economic growth and our daily life. The optimal solution may be a combination of renewables, nuclear power, and other environmental-friendly energy resources. For the renewables, in addition to solar and wind energy, lignocellulosic biomass (e.g., wood, corn stalk), the most abundant organic material on earth, has been proposed as a promising source for fuel ethanol production.[6, 7] Cellulose can be hydrolyzed into glucose and other constituent sugars and subsequently converted to ethanol through fermentation.

However, lignocellulosic biomass is, by nature, highly crystalline with polymeric cellulose chains (1,4-linked β-D-glucose units) held together by hydrogen bonds and van der Waal's forces.[8] As a consequence, cellulose pretreatment is essential in order to reduce the degree of polymerization and to make it amorphous to facilitate substrate access for efficient enzymatic hydrolysis. Traditional preprocessing methods include mechanical reduction of cellulose particle size, followed by acidic or alkaline solution treatment.[7] However, these methods are usually considered not environmentally friendly due to the severe reaction conditions, such as high temperature, high pressure and/or extreme pH values. Thus, a green and large scale industrial production of biofuel from cellulose is of critical importance to reconcile the urgent need for fossil fuel replacement worldwide.

Ionic liquids (ILs) exhibit several unique aspects such as powerful solvents for both organic and inorganic substances, low vapor pressure, high thermal stability and wide electrochemical window. In 2002, a pioneering work conducted by Rogers and coworkers has demonstrated that some ILs are able to dissolve cellulose.[9] Subsequently, more ILs were discovered to possess superiority in dissolving cellulose materials.[7, 10] Among these ILs, the 1-butyl-3-methylimidazolium chloride, [Bmim]Cl, is shown to be one of the most prominent ILs to efficiently dissolve cellulose.[9, 11] The solvated cellulose is then transferred from IL solution to water solution for enzymatic hydrolysis, while the ILs can be recycled for the next round of cellulose dissolution. ILs pretreatment could remarkably accelerate the hydrolysis and improve the productivity. For example, Zhao *et al.* reported that the D-glucose yield from switchgrass by ILs pretreatment could reach 96% after enzymatic hydrolysis for 24 h.[12] This method provides new opportunities for the green processing of biomass for fuel ethanol production and has become a hot research area in recent years. However, there is a critical issue that needs to be

resolved, where the hydrolytic enzyme, namely the cellulase, is found to be inactivated due to the residual IL molecules in the water solvent.[11, 13-15] A thorough removal of the residual ILs after cellulose pretreatment would help to overcome this obstacle,[11] although this could be quite costly and may not be applicable for industrial scale.

Protein engineering to tailor cellulase for functioning in unusual environments can be a promising attempt.[16, 17] For example, Schwaneberg and co-workers have found that a single substitution of H288R in CelA2 (the position 288 is located nearby the active site) could maintain its residual activity in 3-fold concentrated seawater, while the wild-type CelA2 loses >50% of its activity.[16] Finding working systems containing both enzyme-compatible ILs and IL-resistant enzymes through protein engineering would be a green and sustainable strategy.[18] However, limited by the experimental techniques, it is usually hard to monitor the protein structure and functional changes at the atomic level. Hence, the cellulase inactivation process and the underline mechanism in ILs have been poorly documented. Some fundamental questions need to be answered before trial attempts for large scale bioengineering. For example, what is the cause of the cellulase inactivation at the molecular level? Is it the cation or the anion of ILs that is mostly responsible for the inactivation? What are the key residues involved in the protein inactivation? It is evident that a detailed explanation of the inactivation mechanism of cellulase with IL is necessary and desirable in order to provide guidance for future improvements.

Here we present molecular dynamics (MD) simulation studies of the cellulase in [Bmim]Cl solvent to explore the molecular mechanism of this IL-induced cellulase inactivation with atomic details. We have considered the cellulase in 0.8 M concentration of [Bmim]Cl which mimics the experimental conditions. The protein stability was described by root mean square deviation (RMSD) with respect to the crystal structure and radius of gyration (Rg). Different from protein denaturation in urea solvent where protein unfolds, the cellulase maintains a stable secondary and tertiary structure in [Bmim]Cl. Through analyzing the specific interactions between the cellulase and water/IL molecules, we identified that the intrusions of $[Bmim]^+$ cations into the cellulose binding tunnel, especially the active site, are responsible for the inactivation of cellulase. Regarding protein engineering to improve the IL-resistance of cellulase, residues H228, R251 and W376 were revealed as key positions in defending against the $[Bmim]^+$ intrusion. *In silico* mutagenesis study showed that H228R can efficiently weaken the binding affinity of $[Bmim]^+$ at the active site.

2. System and Method

For cellulase, the cellobiohydrolase I (CBHI) from fungus *Trichoderma reesei* was adopted, because the crystal structure, hydrolysis process and inactivation of CBHI in [Bmim]Cl have been well characterized in experiments.[13, 19] The atomic coordinates were taken from the crystal structure (PDB accession code 8CEL).[19] Inspired by the mutagenesis study from experiment[16], a single substitution mutant H228R was also investigated in our study.

Three systems were setup for our study: (*i*) the wild type cellulase at 300 K, (*ii*) the wild type cellulase at 350 K, and (*iii*) the H228R mutant at 300 K. The simulation conducted at 350 K is for the consideration that enzymatic hydrolysis is usually conducted at high temperature, for example, 50 ℃.[13] Moreover, simulation at high temperature can help to accelerate the protein unfolding. The system was first equilibrated for 5 ns where position restraints were applied to the protein heavy atoms, followed by productive simulations of 200 ns. For each case, three independent trajectories were generated for data collection.

Amber03 force field[20] was used to model the cellulase. For [Bmim]Cl, a refined force field by Liu and coworkers was used.[21] Following similar protocols as in our previous studies,[22-28] the current simulation box contains 1 cellulase and 15,856 TIP3P water.[29] To neutralize the 22 positive charges of cellulase, 300 $[Bmim]^+$ and 278 Cl^- were added to the water box, which resulted in a simulation box of ~8.5 × 8.5 × 8.5 nm^3 with the molar concentration of the IL around 0.8M (corresponds to 20.5% v/v and 13.6% w/w).

All simulations were carried out using GROMACS 4.5.4 package[30]. SHAKE constraints were applied to all bonds involving hydrogen atoms.[31] The long-range electrostatic interactions were treated with the Particle Mesh Ewald method,[32] and a typical distance cutoff of 12 Å was used for the van der Waals interactions. The non-bonded interaction pair-list was updated every 10 fs. Constant-pressure (1 atm), constant-temperature (300 K and 350 K, respectively) molecular dynamics with an integration step of 2 fs is used for each system.

3. Results and Discussion

3.1 Structural Evolution of the Wild Type CBHI

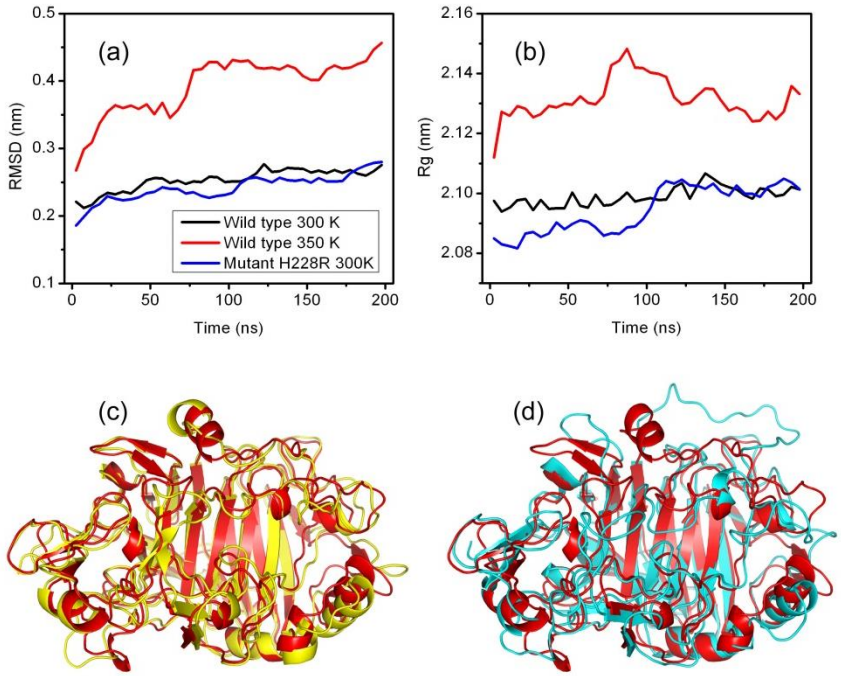


Figure 1. (a) Root mean square deviation (RMSD) with respect to the crystal structure; (b) Radius of gyration (Rg) of protein heavy atoms. The values are averaged between three independent simulations of each case; (c) The final structure of CBHI after simulation at 300 K and (d) The final structure of CBHI after simulation at 350 K. The crystal structure is also shown in red color for comparison.

First, the structural stability of the wild type CBHI is evaluated by calculating the RMSD with respect to the crystal structure and Rg of the protein heavy atoms (see Figure 1a-b). At room temperature (300 K), the values of RMSD and Rg reach to 0.26 and and 2.10 nm, respectively, with little fluctuations during the entire simulation. At higher temperature (350 K), slightly larger structural changes can be observed with the RMSD reaching 0.45 nm. Meanwhile, the Rg fluctuates around 2.13 nm, with slightly larger fluctuations than that observed at 300 K.

The relatively small values of RMSD indicate that CBHI is structurally stable in 0.8M (20.5% v/v and 13.6% w/w) [Bmim]Cl solvent, even at 350 K. The two final structures after 200 ns simulation at 300 and 350 K are illustrated in Figure 1c-d, respectively, where the crystal structure is also shown in red color for comparison. It is clear that CBHI maintains a well ordered secondary and tertiary structure. The only significant changes at 350 K are the opening/twisting of some loops on the CBHI surface and unfolding of a helix (residue 404-410) as shown in Figure 1d. These changes correspond to the increase of RMSD and Rg as shown in Figure 1a-b. The core part (the β-sheet) remains intact. Overall, the cellulase structure is very stable, showing both thermo-stability and IL-resistance. This is well consistent with experimental results where cellulase was found to be in well-folded conformation in comparable concentrations of [Bmim]Cl.[13] Our findings also demonstrate that $[Bmim]^+$ and Cl^- are essentially "salt" instead of chaotropic protein denaturing agent like urea.[33] The inactivation of cellulase with ILs is not

caused by the structural unfolding of cellulase, which commonly happens in chemical denaturation, such as in urea and GdmCl solutions.[34-36]

3.2 ILs and Water around CBHI

After excluding the structure unfolding as a potential factor for the cellulase inactivity, we then expect some specific interactions between solvent (including [Bmim]Cl and water) and CBHI might be responsible for the inactivation.

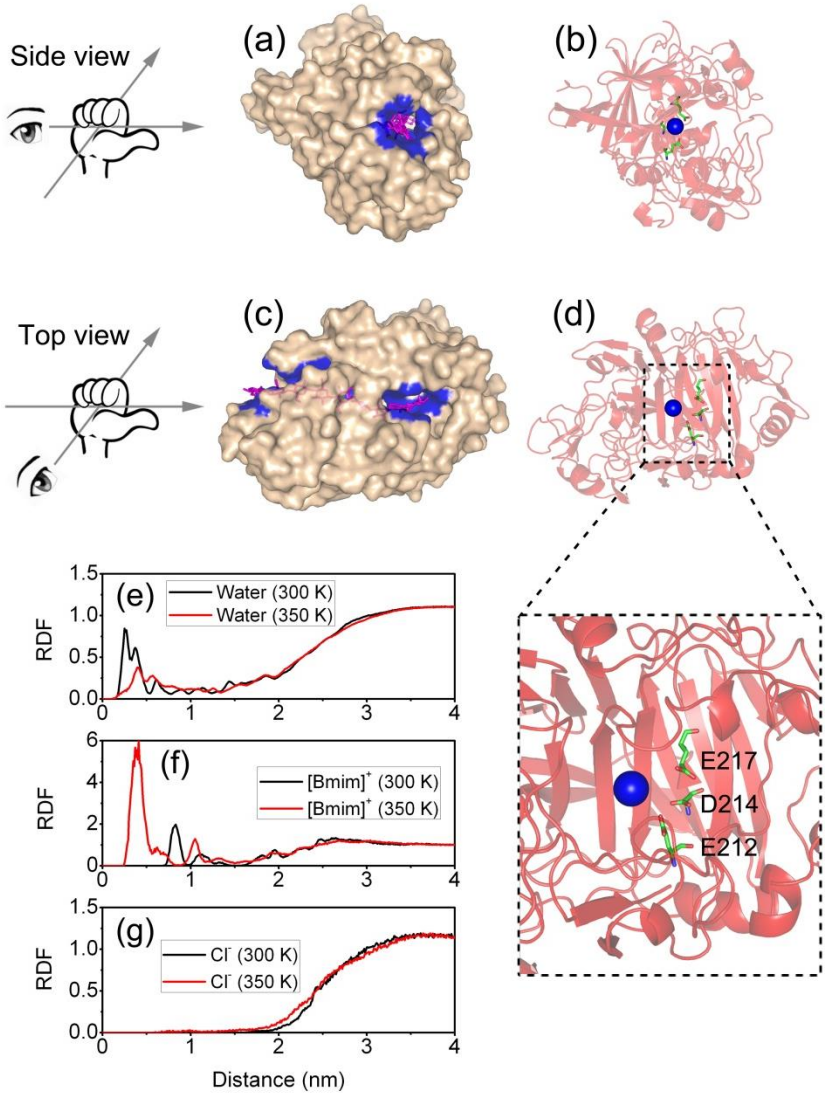


Figure 2. (a, b) Side view of CBHI crystal structure shown as surface and cartoon, respectively; (c, d) Top view of CBHI shown as surface and cartoon, respectively; (e-g) Radial distribution functions (RDF) of water, $[Bmim]^+$ and Cl^- around the cellulase center of mass (COM). In (a) and (c), the cellulose substrate is shown as magenta colored sticks. The CBHI

atoms which are within 5 Å to cellulose are shown as blue colored surface, demonstrating the binding tunnel. In (b) and (d), the blue spheres represent the COM of CBHI. The three key residues, E212, D214 and E217, constituting the active site of CBHI are highlighted which can be found in the insert.

As illustrated in Figure 2a-d, the tertiary structure of CBHI is like a right hand with the fingers curled. The cellulose substrate locates in the "palm of the hand" with its movement direction along the thumb during hydrolysis. The CBHI atoms which are within 5 Å to the cellulose atoms are determined as the "binding tunnel", which are highlighted in blue in Figure 2a and c.

The center of mass (COM) of CBHI locates almost at the middle of the cellulose binding tunnel as shown in Figure 2b and d. Moreover, the active site (E212, D214 and E217) is also at the middle of the protein, thus near the COM. Therefore, for simplicity, we calculated the radial distribution functions (RDF) of water, $[Bmim]^+$ and Cl^- around the cellulase COM. As shown in Figure 2e, some water molecules are found to locate in the interior of cellulase, near the COM. For example, at 300 K, the cellulase COM is surrounded by a few water molecules, corresponding to the peaks at 0.25 and 0.4 nm. The existence of water in the interior of cellulase is expected, because in our simulation we have not considered the cellooligomer which leaves some space for water intrusion. However, it is interesting to find that, some $[Bmim]^+$ cations also appear in the interior of the cellulase, corresponds to the peak at 0.8 nm in Figure 2f (the black curve). At higher temperature of 350 K, these $[Bmim]^+$ cations further intrude into the core region successfully (near the COM), which is correlated with the peak at 0.4 nm (the red curve in Figure 2f). Accordingly, the height of the peak of water at 0.25 nm (red curve in Figure 2e) is suppressed because some water molecules are repelled away by $[Bmim]^+$ cations. At 2.5 nm, the RDF of $[Bmim]^+$ almost reaches to the bulk value of 1, indicating the edge of cellulase. On the contrary, the interaction between Cl^- and cellulase behaves quite differently. As shown in Figure 2g, the RDF for Cl^- starts to increase at distance 2 nm, and reaches to a value of 1 around 3.5 nm without any observable peak. This phenomenon reflects the fact that Cl^- ions do not have specific binding interactions with cellulase, so that they diffuse freely in the solution.

3.3 Binding Statistics of $[Bmim]^+$ with Cellulase

The existence of $[Bmim]^+$ in the interior of CBHI from the RDF analysis suggests that these $[Bmim]^+$ might be responsible for the inactivation of CBHI. We then analyzed the detailed binding statistics of $[Bmim]^+$ with cellulase. The minimum distances between cellulase and each $[Bmim]^+$ were calculated (for the last 100 ns trajectory). Overall, we have found three types of representative binding patterns as shown in Figure 3a. The type 1 (black curve in Figure 3a) is almost freely-moving $[Bmim]^+$ during the simulation, with numbers of instant contacting with cellulase. Type 2 $[Bmim]^+$ (blue curve) could bind to cellulase and last for tens of nanosecond, and then dissolves, after which the behavior is similar to that of type 1. Type 3 $[Bmim]^+$ (red curve) forms strong binding with the cellulase, with an average distance of 0.24 nm.

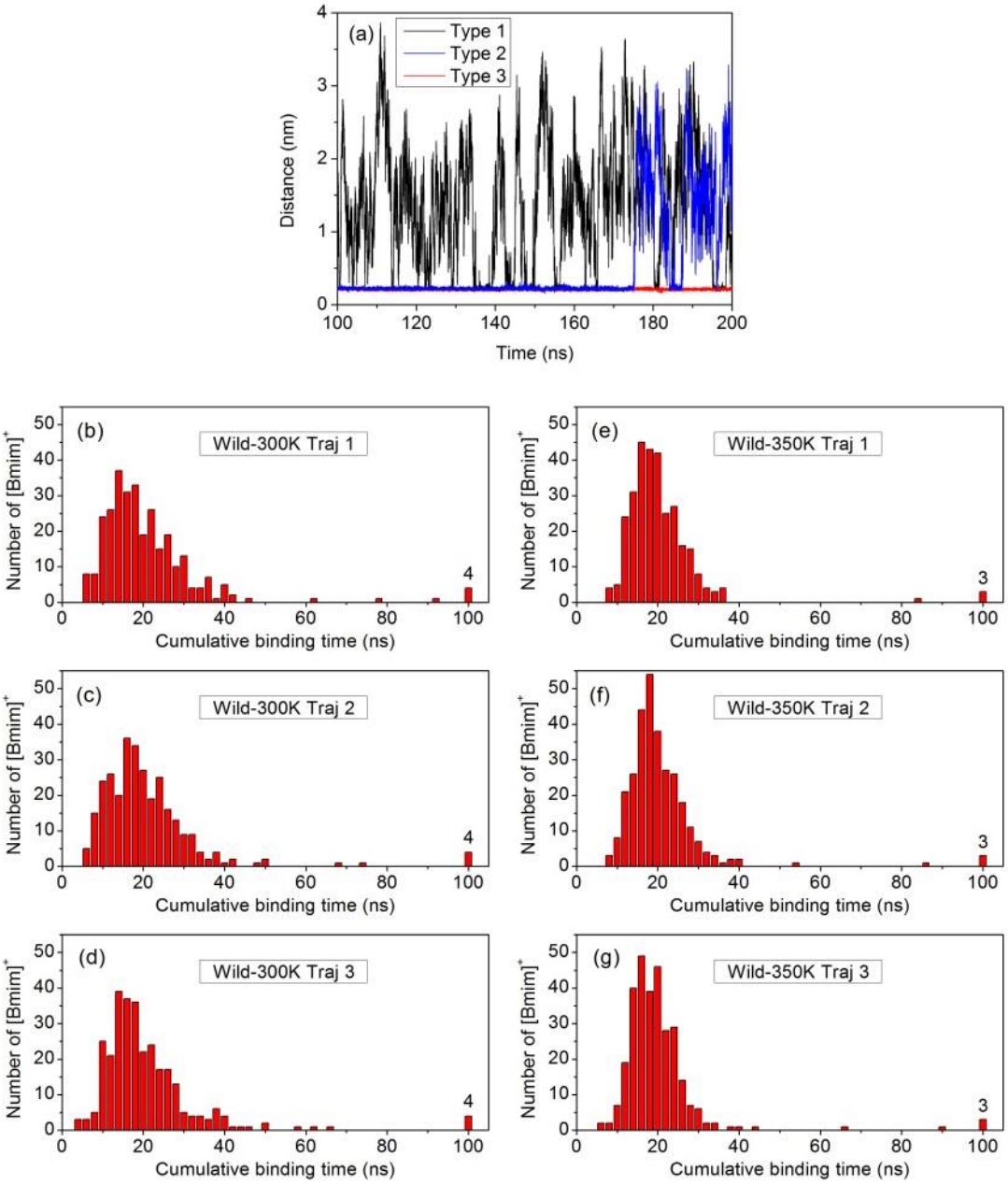


Figure 3. (a) Three types of representative binding patterns of $[Bmim]^+$ with cellulase, (b-d) distributions of cumulative binding time of each $[Bmim]^+$ with cellulase at 300 K, (e-g) distributions of cumulative binding time of each $[Bmim]^+$ with cellulase at 350 K.

In order to quantitatively describe the binding statistics of $[Bmim]^+$ with cellulase, we calculate the cumulative binding time of each $[Bmim]^+$ in the last 100 ns trajectories at 300 K and 350K, respectively. A distance cutoff of 0.3 nm is chosen as the criterion for $[Bmim]^+$-cellulase effective binding. The numbers of $[Bmim]^+$ with respect to the values of cumulative binding time are shown in Figure 3b-g. It is clear that most of the $[Bmim]^+$ have a cumulative binding time of 10-30 ns, corresponding to the free-diffusive $[Bmim]^+$ (type 1) in Figure 3a. In general, more

$[Bmim]^+$ are free-diffusive at 350 K than that at 300 K because of higher thermol kinetic energy. Some $[Bmim]^+$ bind to cellulase in a moderate manner: there are 6-8 $[Bmim]^+$ in the case of 300 K and 1~3 $[Bmim]^+$ in the case of 350 K, which could bind with the cellulase for a longer time (40 ns < t < 100 ns) . More importantly, we have found 4 and 3 $[Bmim]^+$, in the case of simulations at 300 K and 350 K respectively, bound firmly to cellulase during the entire simulation time (t = 100 ns).

The positions of these firmly-bound $[Bmim]^+$ at 300 K from each trajectories are shown in Figure 4. For clarity, the crystal structure (8CEL) with cellooligomer, the natural substrate of the cellulase is also shown. Taking the structure from traj 1 for example, one $[Bmim]^+$ cation, labeled as 1 in Figure 4a, is located at the active site. One $[Bmim]^+$, labeled as 3, is near the product release region. There is also one $[Bmim]^+$, labeled as 4, located at the mouth of the binding tunnel of cellulase. Summarizing all the three simulations, we found that most of the $[Bmim]^+$ coincide with the native cellooligomer binding sites, especially site 1, the active site of the hydrolysis.

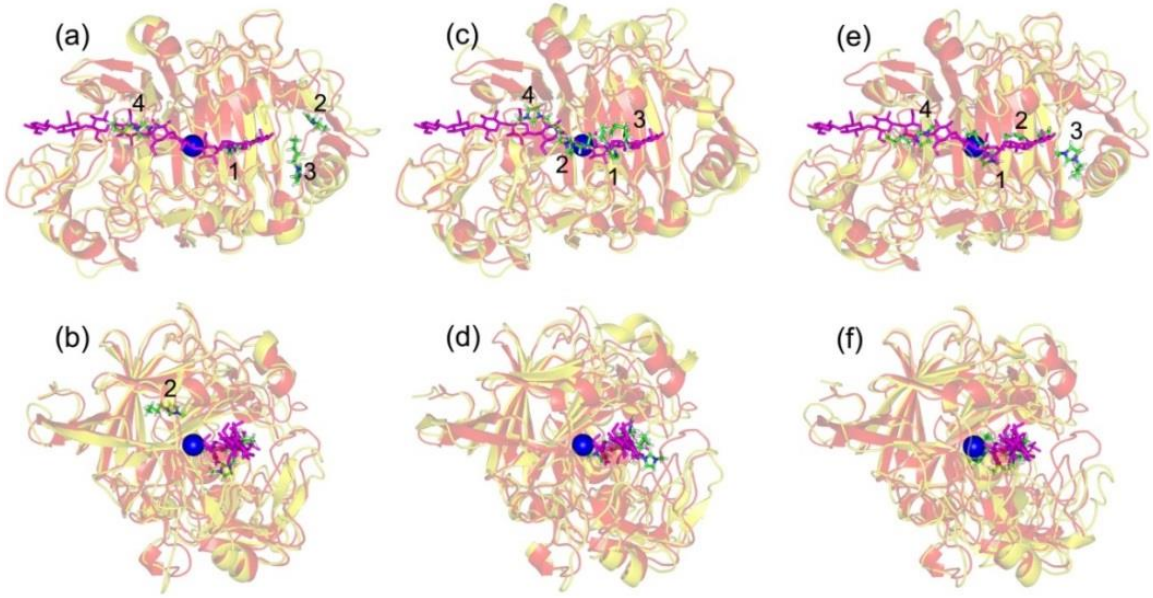


Figure 4. The positions of the four firmly-binded $[Bmim]^+$ at 300 K: top view (a) and side view (b) from traj 1, top view (c) and side view (d) from traj 2, top view (e) and side view (f) from traj 3. The crystal structure is shown in red color with cellooligomer in the binding tunnel (magenta-colored). Blue colored sphere indicates the position of protein center of mass.

From the generalized Born (GB) model, the binding affinities of these $[Bmim]^+$ with cellulase can be evaluated and the values are summarized in Table 1. Among all the firmly-bound $[Bmim]^+$, the one at the active site (labeled as $[Bmim]^+$ 1) has the strongest binding with cellulase, with the binding affinity reaches to a value around -120 kJ/mol. Moreover, other three $[Bmim]^+$ in the cellulose also have considerable binding affinities of -54.89 to -76.91 kJ/mol. This means that the cellulose binding tunnel, especially the active site, is a strong "attractor" for the IL cations. These tightly-binding $[Bmim]^+$ cations would block the binding of cellooligomer, followed by the direct inhibition of the hydrolysis at the active site, hence induce the cellulase inactivation.

Table 1. Binding affinities (kJ/mol) of the $[Bmim]^+$ with cellulase from the GB model. The values in the brackets are error estimations using block averaging where the full trajectory data was divided into 5 blocks.

		$[Bmim]^+$ 1	$[Bmim]^+$ 2	$[Bmim]^+$ 3	$[Bmim]^+$ 4
Wild type	Traj 1	-123.42 (7.06)	-76.91 (7.45)	-67.38 (7.01)	-66.15 (5.96)
	Traj 2	-122.25 (6.46)	-74.40 (6.57)	-59.90 (8.31)	-54.89 (6.57)
	Traj 3	-119.02 (6.91)	-76.18 (8.79)	-74.51 (6.22)	-57.83 (6.81)
Mutant	Traj 1	-98.19 (6.16)	-69.75 (5.01)	-65.93 (6.08)	-58.26 (5.61)
	Traj 2	-96.17 (8.65)	-67.54 (6.62)	-	-
	Traj 3	-91.94 (5.07)	-77.78 (7.82)	-	-

The simulation trajectories also disclose how the $[Bmim]^+$ cations enter the active site and a typical process is shown in Figure 5a-d. In detail, the intrusion process can be divided into four steps according to the nature of driving forces:

Step 1, the $[Bmim]^+$ is free diffusing in the solution, forming random contacts with cellulase;

Step 2, the $[Bmim]^+$ has loaded on the edge of the cellulase and begins to intrude into the binding tunnel;

Step 3, the $[Bmim]^+$ has successfully intruded into the cellulase, with its hydrophobic alkyl tail ponting toward the active site (see Figure 5c);

Step 4, the $[Bmim]^+$ has formed strong Columbic interactions with the active site through its methylimidazolium head.

The intrusions of $[Bmim]^+$ into the active site always happen in a time scale of several ns (see Figure 5f). It is also interesting to note that a quick flip of the $[Bmim]^+$ orientation is found from step 3 to step 4 as shown in Figure 5c-d. This phenomenon can be further quantified through two principle variables: the distance between $[Bmim]^+$ and E212, and one angle determined by two vectors as shown in Figure 5e, *vector 1(v1)*: the center of the methylimidazolium head to the sidechain of E212, and *vector 2 (v2)*: the center of the methylimidazolium head to the hydrophobic alkyl tail. According to this definition, an angle near $0°$ stands for an ideal tail driven intrusion (hydrophobic interaction driven), and $180°$ stands for an ideal head driven intrusion (Columbic attraction driven). The values are illustrated in Figure 5f. It is clear that, when the distance decreases dramatically from 4.0 nm to 1.2 nm (where $[Bmim]^+$ loads to the cellulase surface), the angle decreases from ~$90°$ to ~$40°$. Then the $[Bmim]^+$ undergoes a quick flip from $40°$ to $120°$ before it forms direct contact with E212. This can be attributed to two factors: (1) the binding tunnel of cellulase is relatively hydrophobic; and (2) the head group, which is positively charged, prefers the aqueous environment. As a consequence, the intrusion process is driven by the hydrophobic alkyl tail of $[Bmim]^+$. However, when $[Bmim]^+$ approaches the active site which is constituted by negative charged residues (E212, D214, E217), the positive charged head of $[Bmim]^+$ tends to form strong Columbic attraction with them, and thus the flip.

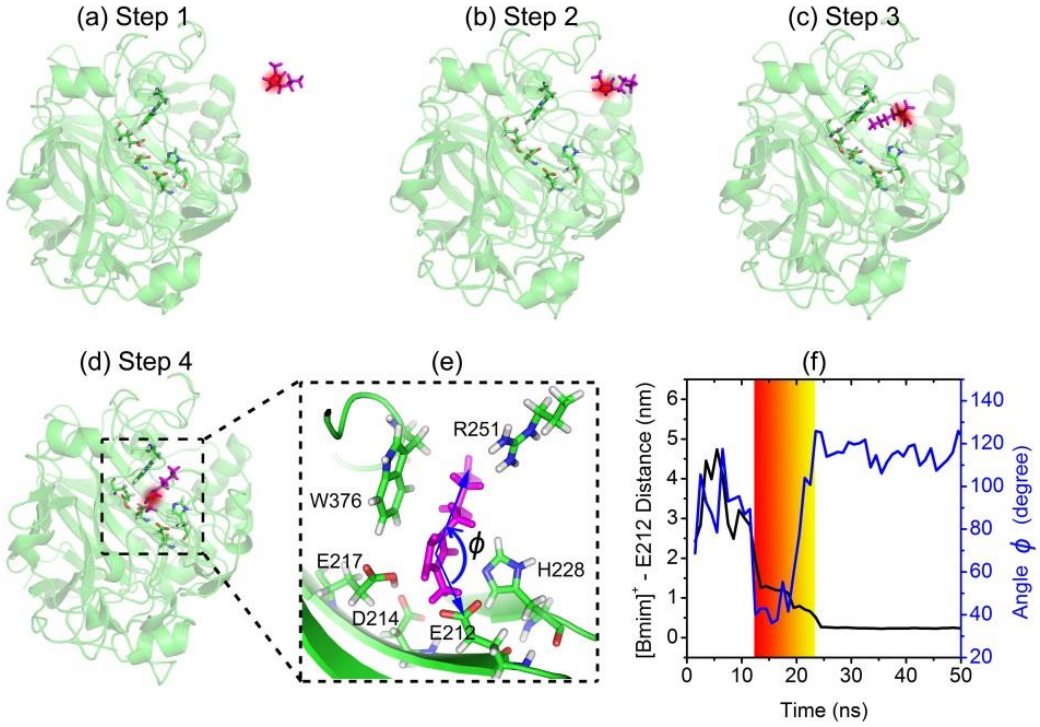


Figure 5. (a-d) The intrusion process of a $[Bmim]^+$ into the active site at 300 K. The positive charge of the methylimidazolium head is colored in red for easy view of the $[Bmim]^+$ orientation; (e) The binding pattern and surrounding key residues around $[Bmim]^+$; (f) Time evolutions of the distance between [Bmim]+ and E212, and the angle determined by two vectors as shown in (e), *v1*: the center of the methylimidazolium head to sidechain of E212, and *v2*: the center of the methylimidazolium head to the hydrophobic alkyl tail (see the main text for more details). The red-to-yellow gradient color bar indicates the $[Bmim]^+$ orientation flip time window.

As illustrated in Figure 5e, there are six residues from CBHI constituting the binding pocket of $[Bmim]^+$: E212, D214, E217, H228, R251 and W376. These key residues are classified into two groups according to their interactions with the $[Bmim]^+$. One group consists of E212, D214 and E217 which are negatively-charged acidic residues. In addition, the ε2 nitrogen of H228 has a negative partial charge of -0.6e. These four residues constitute a locally-negative charged environment as a group. The head group of $[Bmim]^+$ binds to these residues at the active site through the Coulomb attractions. The other group, W376 and R251, on the other hand, binds to the $[Bmim]^+$ alkyl tail mainly through hydrophobic interactions.

3.4 *In silico* mutagenesis study: the H228R mutant

Taking into consideration that E121, D214 and E217 are conserved residues for hydrolysis, the other three residues, H228, R251 and W376, are potential candidates for mutagenesis studies in order to improve the IL-resistance of CBHI. For example, a single mutation of H228 to amino acids with positive-charged residues, like Lysine or Arginine, could effectively weaken the Columbic attraction of the $[Bmim]^+$ head at the active site.

The identification of the importance of H228 in the capture of a $[Bmim]^+$ is consistent with a recent experimental reported by Schwaneberg and coworkers.[16] A single substitution of H288 with an Arginine in cellulase CelA2 resulted in a lower specific activity than the wild type in potassium phosphate buffer (enzyme-friendly environment). On the contrary, in 30% choline chloride: glycerol and in concentrated seawater (deep eutectic solvents), the H288R mutant showed a higher specific activity than that of the wild type. Structurally, the locations of H288 in CelA2 and H228 in CBHI are both near the active site and in close proximity to the substrate, suggesting that the Histidine is detrimental for substrate binding.[16] Inspired by this experimental study, the H228R mutant is also investigated here to show whether the salt resistance mechanism is applicable to [Bmim]Cl as well.

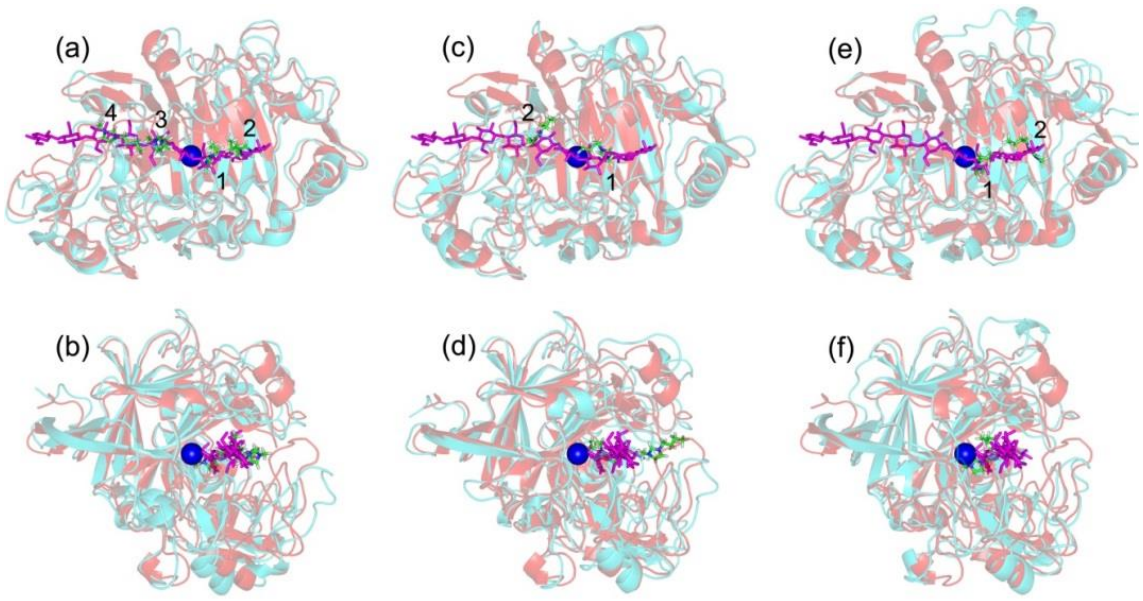


Figure 6. The positions of the firmly-bound $[Bmim]^+$ in the H228R mutant: top view (a) and side view (b) from traj 1, top view (c) and side view (d) from traj 2, top view (e) and side view (f) from traj 3. The crystal structure is shown in red color with cellooligomer in the binding tunnel (magenta-colored). Blue colored sphere indicates the position of protein center of mass.

Three independent simulations of the H228R mutant were conducted under exactly the same conditions at 300 K as that for the wild type. The RMSD and Rg of H228R mutant are shown in Figure 1. It is clear that the values are comparable to that for the wide type at 300 K. Hence, the structure of the mutant is also stable because no structural unfolding or distortion is found. The final structures after 200 ns simulations can also be found in Figure 6, where the secondary and tertiary structures are found to be well conserved.

As observed in the wild type simulations, the intrusions of $[Bmim]^+$ into the cellulose binding tunnel, especially at the active site also happen in the H228R mutant, as shown in Figure 6. The $[Bmim]^+$ at the active site (labeled with 1) is always observed, although the total numbers of $[Bmim]^+$ that intrude into the cellulase are less than that found in the wild type simulations. However, because of the existence of R228 which is a distinct positive residue, the binding affinities of the $[Bmim]^+$ at the active site has been greatly reduced. As shown in Table 1, the binding affinities are around -91.94 to -98.19 kJ/mol. Compared to the wild type, the H228R mutation has effectively reduced the binding affinity by ~20%. Thus, the H228R mutant is less affected by the IL cation intrusion as compared to the wile type.

Finally, it should be noted that there might be one drawback for the H228R mutation, that is the position of H228 is too close to the active site, as shown in Figure 5e. Hence, the R228 substitution results in a strong direct interaction with E212 and D214 through Coulombic attractions. As a consequence, by blocking the binding of cellulose substrate to a certain degree, the intimate interaction between R228 and E212 and D214 may be responsible for the lowered overall activity of cellulase found in experiment. Therefore, the H228R substitution might be a two-edged sword: on one hand, cellulase benefits from the existence of a positive charged residue when the ionic strength of the solution is increased due to residual IL; on the other hand, it lowers the overall substrate binding strength due to the internal attractions between R228 and acidic residues at the active site. Our current simulation data provide atomic-level explanations for these experimental findings.

4. Conclusion

Our current molecular dynamics simulations show that the cellulase CBHI displays good stability in 0.8 M (20.5% v/v and 13.6% w/w) [Bmim]Cl, both at 300 K and 350 K. However, the $[Bmim]^+$ cations can intrude into the binding tunnel and compete with the cellulose substrate. Especially, $[Bmim]^+$ can bind at the active site for hydrolysis which may inactivate cellulase by blocking the hydrolysis pathway. Six key amino acids near the active site were highlighted for the $[Bmim]^+$ binding. Three acidic residues (E212, D214 and E217) and H228 constitute a negative-charged binding pocket which will attract the positive-charged methylimidazolium head of $[Bmim]^+$ through Coulomb interactions. The sidechains of W376 and R251 bind to the alkyl hydrophobic tail of $[Bmim]^+$ through hydrophobic interactions. Considering the sequence conservation of the three active site residues, mutations of H228, W376 and R251, may help to improve the IL-resisting capability of cellulase. *In silico* mutagenesis studies of H228R have been conducted to examine the IL binding at the active site. The binding affinity of $[Bmim]^+$ in the mutant is found to be significantly reduced, by 20%, as compared to the wild type due to the repulsive interactions from the new arginine residue.

Interestingly, the intrusion of $[Bmim]^+$ into cellulase is initialized through its hydrophobic alkyl tail. As a consequence, it is expected that mutations with reduction in the hydrophobic interactions, such as W376/R251 to amino acids with smaller sidechains, could also inhibit the

initial intrusion process of $[Bmim]^+$ to the binding tunnel. Moreover, optimization of the length of alkyl tail of IL cations to weaken the attraction of IL to the surface of cellulase may also help to increase cellulase IL-resistance.

Acknowledgement

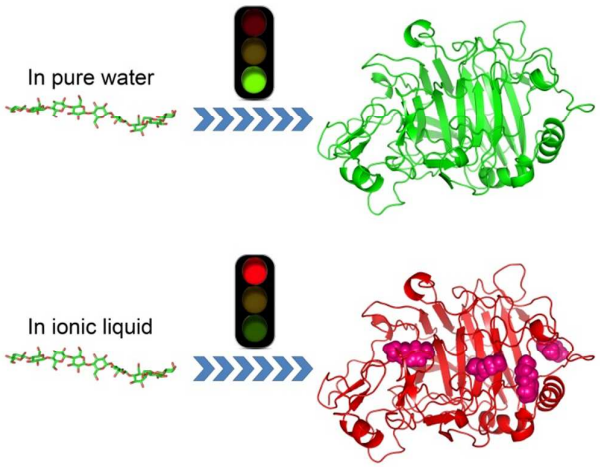
This work was partially supported by NTU Tier 1 Grant RG 23/11, IBM Blue Gene Science Program, the National Natural Science Foundation of China (grant no. 11304214) and the Priority Academic Program Development of Jiangsu Higher Education Institutions (PAPD).

Reference

1. M. K. Hubbert, *Nuclear energy and the fossil fuels*, Shell Development Company, Exploration and Production Research Division Houston, TX, 1956.
2. R. L. Hirsch, R. Bezdek and R. Wendling, 2005.
3. P. Newman, T. Beatley and H. Boyer, *Resilient cities: responding to peak oil and climate change*, Island Press, 2009.
4. U. Bardi, *Energy*, 2009, 34, 323-326.
5. N. A. Owen, O. R. Inderwildi and D. A. King, *Energy Policy*, 2010, 38, 4743-4749.
6. Q. Zhang, S. Zhang and Y. Deng, *Green Chem.*, 2011, 13, 2619-2637.
7. H. Wang, G. Gurau and R. D. Rogers, *Chem. Soc. Rev.*, 2012, 41, 1519-1537.
8. A. Pizzi and N. Eaton, *Journal of Macromolecular Science—Chemistry*, 1985, 22, 105-137.
9. R. P. Swatloski, S. K. Spear, J. D. Holbrey and R. D. Rogers, *J. Am. Chem. Soc.*, 2002, 124, 4974-4975.
10. W. Xiao, W. Yin, S. Xia and P. Ma, *Carbohydr. Polym.*, 2012, 87, 2019-2023.
11. H. Zhao, C. L. Jones, G. A. Baker, S. Xia, O. Olubajo and V. N. Person, *J. Biotechnol.*, 2009, 139, 47-54.
12. H. Zhao, G. A. Baker and J. V. Cowins, *Biotechnol. Prog.*, 2010, 26, 127-133.
13. M. B. Turner, S. K. Spear, J. G. Huddleston, J. D. Holbrey and R. D. Rogers, *Green Chem.*, 2003, 5, 443-447.
14. Â. C. Salvador, M. d. C. Santos and J. A. Saraiva, *Green Chem.*, 2010, 12, 632-635.
15. H. Zhao, *J. Chem. Technol. Biotechnol.*, 2010, 85, 891-907.
16. C. Lehmann, F. Sibilla, Z. Maugeri, W. R. Streit, P. Dominguez de Maria, R. Martinez and U. Schwaneberg, *Green Chem.*, 2012, 14, 2719-2726.
17. L. Li, J. Xie, S. Yu, Z. Su, S. Liu, F. Liu, C. Xie, B. Zhang and C. Zhang, *Green Chem.*, 2013, 15, 1624-1630.
18. H. Zhao, G. A. Baker, Z. Song, O. Olubajo, T. Crittle and D. Peters, *Green Chem.*, 2008, 10, 696-705.
19. C. Divne, J. Ståhlberg, T. T. Teeri and T. A. Jones, *J. Mol. Biol.*, 1998, 275, 309-325.
20. T. A. D. D.A. Case, T.E. Cheatham, III, C.L. Simmerling, J. Wang, R.E. Duke, R. Luo, M. Crowley, Ross C. Walker,W. Zhang, K.M. Merz, B.Wang, S. Hayik, A. Roitberg, G. Seabra, I. Kolossváry, K.F.Wong, F. Paesani, J. Vanicek, X.Wu, S.R. Brozell, T. Steinbrecher, H.

Gohlke, L. Yang, C. Tan, J. Mongan, V. Hornak, G. Cui, D.H. Mathews, M.G. Seetin, C. Sagui, V. Babin, and P.A. Kollman, *University of California, San Francisco*, 2008.
21. Z. Liu, S. Huang and W. Wang, *J. Phys. Chem. B*, 2004, 108, 12978-12989.
22. M. Eleftheriou, R. S. Germain, A. K. Royyuru and R. Zhou, *J. Am. Chem. Soc.*, 2006, 128, 13388-13395.
23. S. G. Kang, T. Huynh, Z. Xia, Y. Zhang, H. Fang, G. Wei and R. Zhou, *J. Am. Chem. Soc.*, 2013, 135, 3150-3157.
24. X. Li, J. Li, M. Eleftheriou and R. Zhou, *J. Am. Chem. Soc.*, 2006, 128, 12439-12447.
25. P. Das, D. Kapoor, K. T. Halloran, R. Zhou and C. R. Matthews, *J. Am. Chem. Soc.*, 2013, 135, 1882-1890.
26. Z. Xia, P. Das, E. I. Shakhnovich and R. Zhou, *J. Am. Chem. Soc.*, 2012, 134, 18266-18274.
27. Y. Tu, R. Zhou and H. Fang, *Nanoscale*, 2010, 2, 1976-1983.
28. P. Das, J. A. King and R. Zhou, *Proc. Natl. Acad. Sci. U. S. A.*, 2011, 108, 10514-10519.
29. W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey and M. L. Klein, *J. Chem. Phys.*, 1983, 79, 926-935.
30. S. David Van Der, L. Erik, H. Berk, G. Gerrit, E. M. Alan and J. C. B. Herman, *J. Comput. Chem.*, 2005, 26, 1701-1718.
31. J. P. Ryckaert, G. Ciccotti and H. J. C. Berendsen, *J. Comput. Phys.*, 1977, 23, 327-341.
32. T. Darden, L. Perera, L. Li and L. Pedersen, *Structure*, 1999, 7, R55-R60.
33. H. Zhao, *J. Chem. Technol. Biotechnol.*, 2006, 81, 877-891.
34. W. Li, R. Zhou and Y. Mu, *J. Phys. Chem. B*, 2012, 116, 1446-1451.
35. D. R. Canchi, D. Paschek and A. E. García, *J. Am. Chem. Soc.*, 2010, 132, 2338-2344.
36. L. Hua, R. Zhou, D. Thirumalai and B. J. Berne, *Proc. Natl. Acad. Sci.*, 2008, 105, 16928-16933.

The table of contents entry



Ionic liquid cations can intrude into the cellulase active site and inactivate it by blocking the cellooligomer binding.