

# Extended Ensemble Molecular Dynamics Study of Cellulose I–Ethylenediamine Complex Crystal Models: Atomistic Picture of Desorption Behaviors of Ethylenediamine

Toshifumi Yui (✉ [tyui@cc.miyazaki-u.ac.jp](mailto:tyui@cc.miyazaki-u.ac.jp))

University of Miyazaki <https://orcid.org/0000-0002-6939-7750>

Takuya Uto

University of Miyazaki

---

## Research Article

**Keywords:** Cellulose I-ethylenediamine complex crystal, Crystal model, EDA desorption, Accelerated MD, Adaptive steered MD, Umbrella sampling method

**Posted Date:** June 17th, 2021

**DOI:** <https://doi.org/10.21203/rs.3.rs-587409/v1>

**License:** © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

---

**Version of Record:** A version of this preprint was published at Cellulose on August 23rd, 2021. See the published version at <https://doi.org/10.1007/s10570-021-04136-7>.

# Abstract

Cellulose I crystals swell on exposure to ethylenediamine (EDA) molecules to form a cellulose I–EDA complex, and successive extraction of EDA molecules converts the complex crystalline phase to either original cellulose I or cellulose III<sub>I</sub>, depending on the treatment procedure. The present study reports the extended ensemble molecular dynamics (MD) simulation of the cellulose I–EDA complex models. An accelerated MD simulation allows most of the EDA molecules to desorb from the crystal model through a hydrophilic channel between the piles of cellulose chains, one at a time. Migration of a single EDA molecule along the channel is simulated by the adopted steered MD method combined with the umbrella sampling method to evaluate the potential of mean force (PMF) or free energy change on its movement. The PMF continues to increase during the migration of an EDA molecule to give a final PMF value of more than 30 kcal/mol. The PMF profiles are largely lowered by the removal of EDA molecules in the neighboring channels and by the widening of the channel. The former suggests that the EDA desorption cooperates with that in the neighboring channels, and, in the latter case, an EDA migration is efficiently promoted by solvation with water molecules in the expanded channel. We conclude that the atomistic picture of the EDA desorption behaviors observed in the crystal models is applicable to the real crystalline phase.

## Introduction

Among the versatile properties of cellulosic materials, crystal polymorphism has been attracting much attention in terms of both scientific and technological interests. In particular, cellulose forms complex phases with various amines and liquid ammonia as an intermediate state in the crystalline conversion between the cellulose I<sub>β</sub> and III<sub>I</sub> forms. (Chanzy et al. 1987; Wada 2001; Wada et al. 2001; Wada et al. 2004; Wada et al. 2006; Wada et al. 2008; Wada et al. 2009b) In comparison with the native cellulose allomorphs, I<sub>α</sub> and I<sub>β</sub>, cellulose III<sub>I</sub> is more susceptible to certain chemical treatments (Clark and Parker 1937; Davis et al. 1943; Pentoney 1966; Perez et al. 2003) and enzymatic degradation (Igarashi et al. 2007; Chundawat et al. 2011; Igarashi et al. 2011). Pretreatment of native cellulosic materials with liquid ammonia and amines, followed by the chemical and enzymatic methods, has been expected to be an efficient method for biomass conversion to produce biofuel. In particular, ammonia fiber expansion (AFEX) pretreatment effectively decreases the crystallinity of the original native cellulose material and reduces the amounts of hemicellulose and lignin (Lau and Dale 2009; Zheng et al. 2009). Information on the underlying molecular events of the crystal conversions, such as rearrangement of cellulose chains, penetration and desorption of guest amines, and exchange of the hydrogen bonding network, will improve the amine treatments for an increased effective biomass conversion. The high-resolution crystal structures of the relevant cellulose allomorphs have exclusively been reported by groups in Tokyo, Grenoble, Los Alamos, and Oak Ridge, using synchrotron fiber diffraction analysis, sometimes combined with nuclear magnetic resonance (NMR) measurement and theoretical calculations (Nishiyama et al. 2002; Wada et al. 2004; Wada et al. 2009a; Nishiyama et al. 2010; Wada et al. 2011; Sawada et al. 2013). These allomorphs include cellulose I<sub>α</sub> and I<sub>β</sub>, cellulose III<sub>I</sub>, and cellulose I–ethylenediamine (EDA) and

ammonia complexes. Among them, the crystal structure of the cellulose I–EDA complex includes four cellulose chains in parallel polarity and eight EDA molecules, which are packed in a monoclinic unit cell ( $a = 4.546 \text{ \AA}$ ,  $b = 11.330 \text{ \AA}$ ,  $c = 10.368 \text{ \AA}$ , and  $\gamma = 94.017^\circ$  in the space group  $P2_1$ ) (Wada et al. 2009b). The  $ab$  projection structure reveals that cellulose chains are piled by their pyranose faces upon each other along the  $a$ -axis and that EDA molecules align along the hydrophilic channel composed by the hydrophilic edges of the piled cellulose chains (see Fig. 1). Intermolecular hydrogen bonds are formed between cellulose chains and EDA molecules (Wada et al. 2009b), which have been exchanged for the direct hydrogen bonds between cellulose chains in the parent allomorphs, cellulose I $_{\beta}$  and III $_I$  (Nishiyama et al. 2002; Wada et al. 2004). A time-resolved X-ray microscopy diffraction measurement detects that no intermediate phase appears during the crystalline conversion, which suggests the successive crystalline conversion from cellulose I $_{\beta}$  to the I–EDA complex (Nishiyama et al. 2010).

Theoretical calculations using molecular dynamics (MD) and density functional theory (DFT) have often been applied to examine the cellulose crystal structures to define the positions of the hydroxyl hydrogen atoms and, for the complex crystals, to refine the ligand structures. The former method has also revealed the dynamic features of the crystal structures such as exchange in the hydrogen bonding network coupled with rotation of hydroxyl groups. For both methods, periodic systems are mostly adopted; the crystalline phase continues beyond the boundary of a unit cell, usually consisting of several crystal unit cells, so that molecular chains of infinite length can be established. In contrast to the infinite system, French proposed a molecular modeling methodology based on the miniature crystals, or mini-crystals, for carbohydrates, which deals with isolated small crystal models with finite dimensions (French and Dowd 1993; French et al. 1993). Intensive hydrogen bonding networks, which are characteristic of most carbohydrate crystals, enable the mini-crystal models to substantially retain the original crystal structures after energy minimization. One can work with the optimized models to define the hydrogen bonding network and thermodynamics properties. We first combined mini-crystal modeling with the MD method to examine the structure dynamics of the  $\beta$ -chitin structure (Yui et al. 2007). The combination was found to be more valid for an extended application to simulate a large-scale structural change, which includes a fiber twist of the native cellulose crystal (Matthews et al. 2006; Yui et al. 2006), the decrystallization of crystalline fibers (Beckham and Crowley 2011; Beckham et al. 2011), and the conversion of cellulose III $_I$  (Yui and Hayashi 2008; Yui et al. 2010; Uto et al. 2013). In a similar fashion, the finite cellulose chain sheet models, which were extracted from the corresponding cellulose allomorphs, were studied by energy minimization using the DFT method to examine the deformation of the initial sheet forms (Uto et al. 2014; Uto and Yui 2018).

We recently reported an MD study of the cellulose I–EDA crystal models in a solution state (Uto et al. 2019). Although our primary objective of the MD calculations was to observe the diffusion and desorption of EDA molecules, it was observed that desorption of EDA molecules was limited to the first and second EDA layers at the early simulation time and that it apparently ceased for the remainder of the 100-ns simulation time. Despite the small sizes of the crystal models, which consist of 48 cellulose chains with a degree of polymerization (DP) of 40, compared with the real crystalline phase, EDA

desorption would be an excessively slow event under the ordinary simulation time. Similarly, an MD study aimed at the complexation of a native cellulose crystal with ammonia molecules was carried out only to observe their penetration into the surface cellulose layers (Bellesia et al. 2011). In the present study, MD simulations of the cellulose I–EDA complex models have been carried out, this time by adopting the extended ensemble theories. We have successfully achieved almost complete desorption of the EDA molecules and estimated the free energy change on diffusion of an EDA molecule in the crystal model.

## Computational Methods

### Accelerated molecular dynamics

Accelerated molecular dynamics (aMD) is an enhanced sampling technique that flattens the potential energy surface by introducing a non-negative boost potential, by which one can observe the long-time behaviors of simulation systems with limited computational time (Hamelberg et al. 2004; Hamelberg et al. 2007). Figure 1(a) shows the *ab* base planes of the two crystal models, which consist of 24 and 48 cellulose chains with DP = 40, denoted as the 40mer × 24 and 40mer × 48 models, respectively. Both crystal models share the (010) lattice planes of the same length, whereas the length of the (100) lattice plane of the former is doubled in the latter. The 40mer × 24 was solvated in 50,000 TIP3P water molecules (Jorgensen 1981; Jorgensen et al. 1983) containing 546 EDA molecules, and in 90,000 water molecules and 1,365 EDA molecules for the 40mer × 48 model.

The present aMD simulations adopted two acceleration levels, termed as a dual boost, which applied a boost potential to both dihedral and total potentials. The aMD parameters, the threshold boost energy,  $E$ , and the acceleration factor,  $a$ , are defined for each potential according to the Amber Tutorial (Pierce et al. 2012; Salomon-Ferrer et al. 2013b; Miao et al. 2014), following:

$$\begin{aligned} E_{dihedral} &= V_{dihedral\_avrg} + 4.0N_{residues} & \alpha_{dihedral} &= 0.8N_{residues} \\ E_{total} &= V_{total\_avrg} + 0.16N_{atoms} & \alpha_{total} &= 0.16N_{atoms} \end{aligned} \quad (1)$$

where  $N_{residues}$  is the number of solute residues including EDA,  $N_{atoms}$  is the total number of atoms, and  $V_{dihedral\_avrg}$  and  $V_{total\_avrg}$  are the average dihedral and total potential energies from conventional MD calculations, respectively.

As described in the next section, many EDA molecules are released from the crystal models in the early stage of the aMD calculations, which fills a solvent box with desorbed EDA molecules that should be removed from the estimation of the aMD parameters. The present aMD simulations were carried out by the multi-cycle procedure involving a solvent displacement step accompanied by re-estimation of the aMD parameters, following:

1. Equilibration step. The simulation system, the cellulose I–EDA complex crystal model with water molecules, is equilibrated by the standard procedure described below. The final part of the procedure involves the conventional MD run for 1 ns for estimation of the aMD parameter.
2. aMD calculations. The equilibrated system is subject to aMD calculations for 10 ns with the aMD parameters determined based on the conventional MD trajectory performed in the previous cycle, using the NPT ensemble at a constant temperature of 300 K or 370 K and pressure of 1 bar. The number of EDA molecules remaining in the crystal model is counted for each 1-ns aMD run.
3. Displacement of the solvent box. A cellulose crystal accommodating the interior EDA molecules and water molecules, and the water molecules attached to the crystal surface are extracted from the final aMD structure. A new solvate box system is built for the crystal model extracted, where the number of water molecules is adjusted to be same as the original number. The new simulation system is then moved to the equilibration step of the next cycle.

### Adaptive steered molecular dynamics

Steered molecular dynamics (SMD) uses a pseudo particle that applies a steering force to analyze the physical movement of atoms and molecules along a reaction coordinate at a particular velocity (Park et al. 2003; Park and Schulten 2004). SMD, in combination with the Jarzynski's non-equilibrium work relation (Jarzynski 1997), predicts the potential of mean force (PMF), which is derived from the exponential average of numbers of non-equilibrium works. Adaptive steered molecular dynamics (ASMD) is an improved method of SMD for effective sampling, where the reaction coordinate is divided into smaller segments in which the work distribution becomes narrower (Ozer et al. 2010; Ozer et al. 2012). The 16 cello-octamers (8mer × 16 model) model was adopted in the ASMD shown in Fig. 1(b). From the crystal surface, four EDA molecules were removed from the central hydrophilic channel to produce a hole for an EDA migration path. The model was termed as the one-hole model and, in a similar fashion, the two- and three-hole models were composed by digging an additional hole on one side or both sides of the original hole, respectively. The one-hole model was additionally modified by expanding the space of the migration path, such that both crystal model halves (the 8mer × 8 units) were slid outward along the *a*-axis for a given distance. The model was termed as the gap model. All the crystal models were solvated in 11,000 to 15,748 TIP3P water molecules (Jorgensen 1981; Jorgensen et al. 1983) and the one-, two-, and three-hole models contained 45, 40, and 35 EDA molecules, respectively.

The EDA molecule at the bottom of the hole, which occupied the fifth EDA layer from the crystal surface, was assigned to be the target molecule to be pulled along the hydrophilic channel. The reaction coordinate was defined to connect the center of the target EDA with that of the opening of the hole at the crystal surface.

A harmonic potential with a spring constant of 71.5 kcal/(mol•Å<sup>2</sup>) was imposed on the center of the target EDA, which traveled at the constant pulling velocity of 0.5 Å/ns for 4 Å, which corresponded to the spacing of the cellulose layers. PMF, or the free energy difference, was determined from the ensemble average of the work, *W*, done by 50 SMD runs by the following Jarzynski's equality (Jarzynski 1997):

$$PMF = -\frac{1}{\beta} \ln \langle e^{-\beta W} \rangle \quad (2)$$

where  $\beta = 1/(k_B T)$ ,  $k_B$  is Boltzmann's constant, and  $T$  is the temperature. At the switching point of the SMD stages, a single trajectory was selected according to that whose work value was the closest to the PMF of the current SMD stage, and the structure at the final frame of the selected trajectory was used as the starting structure of the next SMD stage. The four SMD stages covered the overall distance of the migration path of 16 Å in the hydrophilic channel; meanwhile, the target EDA traveled through the four cellulose layers until it reached the crystal surface. The SMD calculations were performed using the NVT ensemble at a constant temperature of 300 K. Throughout the ASMD calculations, the positions of all backbone atoms of a cellulose chain, C1, C2, C3, C4, C5, and O5, and of the two carbon atoms of the EDA molecules, except those of the target EDA, were weakly constrained with a force constant of 0.1 kcal/(mol•Å<sup>2</sup>) to prevent the crystal model from irreversible distortion resulting from EDA movement.

### **Umbrella sampling method**

In contrast to the SMD and ASMD methods with a non-equilibrium basis, an umbrella sampling (US) simulation is a classified equilibrium molecular dynamics method (Torrie and Valleau 1974; Torrie and Valleau 1977), although it is also used to calculate PMF or a free energy change along a given reaction path, where biased potentials are applied to an atom or molecule at certain closely spaced intervals along the path. In the present US simulations, the PMF trajectory obtained by the SMD calculations was adopted as a reaction coordinate of which the overall length, 16 Å, was divided into 0.01 Å intervals, which gave 1,600 umbrella windows. At each interval, a harmonic potential with a force constant of 100 kcal/(mol•Å<sup>2</sup>) was applied to the target EDA molecule and the NPT production run was performed for 1 ns at a constant temperature of 300 K and pressure of 1 bar. As was performed in the present ASMD, all of the backbone atoms of cellulose and the carbon atoms of all EDA molecules, except the target EDA, were positionally constrained with a force constant of 0.1 kcal/(mol•Å<sup>2</sup>). The biased EDA distributions obtained from the production runs were unbiased and combined using the weighted histogram analysis method (WHAM) to obtain a PMF profile (Kumar et al. 1992; Kumar et al. 1995; Roux 1995).

### **General MD procedures and parameters**

All energy minimizations and MD calculations of the solvated cellulose I–EDA complex systems were carried out using the PMEMD and PMEMD.CUDA modules (Le Grand et al. 2013; Salomon-Ferrer et al. 2013a) of the Amber 18 program package (Case et al. 2018) with an NVIDIA Kepler GPU system. The equilibration procedure of the system was composed by two-step energy minimizations: optimization of water configurations and full optimization of the whole system, and the gradual heating of the system from 100 to 300 K for 500 ps using the NVT ensemble. The density of the system was then equilibrated by the NPT run at a constant temperature of 300 K and pressure of 1 bar for, typically, 1 ns. In the first energy minimization and the heating runs, the positions of the solute molecules, cellulose chains, and

EDA molecules were constrained with a force constant of 50 kcal/(mol·Å<sup>2</sup>). In the ASMD calculations, the density equilibrating NPT runs were extended to 50 ns until the holes in the hydrophilic channels had filled with waters with the positional constraint. The temperature and pressure of the system were regulated by a Langevin thermostat (Hoover et al. 1982; Evans 1983) with a collision frequency of 5 ps<sup>-1</sup> and a Berendsen barostat (Berendsen et al. 1984), respectively. A time step of 2.0 fs, enabled by the SHAKE option (Ryckaert et al. 1977), was employed to ensure rigid hydrogen atoms. The particle mesh Ewald (PME) method (Essmann et al. 1995) was adopted for long-range interactions and the van der Waals interactions were cut off at 10 Å. The three extended ensemble MD options for aMD, SMD, and US, implemented by the AMBER package, were used.

Cellulose molecules were described by the GLYCAM06 force field (Kirschner et al. 2008) and EDA molecules were modeled using the revised GAFF parameters (Wang et al. 2004) developed by us (Uto et al. 2019).

The MD trajectories were analyzed using the CPPTRAJ module (Roe and Cheatham 2013) of the AmberTools 19 and, for a visual analysis, VMD 1.9.3 (Humphrey et al. 1996). US–PMF profiles were calculated using the WHAM code ver. 2.0.10 written by Grossfield Lab. (Grossfield 2018). The molecular graphics software, PyMOL ver. 2.5.0 (Schrödinger LLC. 2021), was also used for visualizing the system.

## Results And Discussion

### Desorption behaviors of EDA molecules

Figure 2(a) shows changes in the amounts of remaining EDA molecules in percent with respect to simulation time. Note that a 1 ns trajectory of every 11 ns trajectories corresponds to the conventional MD for sampling thermodynamics values, as described in the Computational methods section. Both the 40mer × 24 and 40mer × 48 models finally released more than 80% of the initial EDA molecules at 300 K. Both PMF profiles almost coincided with each other having a similar desorption rate of EDA. At the elevated temperature of 370 K, the desorption rate remarkably increased in the 40mer × 24 crystal model and the PMF finally reached a constant value of approximately 10%. This was consistent with the observation that EDA molecules in the complex crystal became mobile at a temperature larger than 70°C (Wada et al. 2008). Figure 2(b) and (c) depicts the cross-views of the sliced structures of the 40mer × 48 model and its overall structure obtained from the final simulation time, respectively. The respective EDA desorption behaviors are given in Movies S1 and S2. As one can clearly see in Movie S1, EDA molecules diffused along the hydrophilic channel between the cellulose chain piles, one at a time, and then moved away from the crystal model, rather than simultaneous diffusion involving more than one EDA molecule. The vacant space, once occupied by an EDA molecule, was quickly filled by water molecules that diffused from the bulk. As the hydrophilic channel lost most of intercalated EDA molecules, the cellulose chain piles collapsed and enclosed the remaining EDA molecules in the disintegrated crystal model. This was

obviously the cause of the deceleration of EDA desorption. As shown in Fig. 2(a), another ASMD simulation for the 40mer × 24 crystal model with the constrained cellulose chain structure exhibited almost complete desorption, because the hydrophilic channels were retained throughout the simulation. In the real crystalline phase, such a noticeable cellulose chain collapse is unlikely, owing to less mobile cellulose chains with a large DP, so that the crystalline phase can successfully convert to another cellulose crystal form after complete EDA desorption. In the present crystal model, EDA molecules most quickly disappeared in the central channel (see Movie S2), which led to dividing the crystal model into halves along the longitudinal (fiber) direction (see Movie S1). As a result, the overall desorption behavior can be represented by those of the crystal model halves, as also suggested by the similar desorption curves between the 40mer × 24 and 40mer × 48 crystal models. We have reported in a previous MD study that the cellulose I–EDA crystal model developed a right-handed twist as it loses the guest EDA molecules (Uto et al. 2019). Similarly, the present halves of the crystal model twisted and moderately intertwined with each other. The crystalline deformation involving fiber twisting and entangling likely takes place in the real crystal fibers, but on a much larger scale, which results in the degradation of crystallinity and, consequently, a decrease in the dimensions of the single-crystalline phase through the crystal conversion. In comparison with a native cellulose crystal, the enhanced accessibility and chemical reactivity of the cellulose III<sub>1</sub> crystal have been attributed to not only the physical nature of the latter form but also crystal degradation resulting from amine treatments. Salminen et al. reported the solid-state polymorphic transition of cellulose nanocrystals immobilized on a chemically modified silicon surface from cellulose I to III<sub>1</sub> through its EDA complex (Salminen et al. 2017). The width of the cellulose III<sub>1</sub> nanocrystals decreased to half the original size of the cellulose I nanocrystals, whereas their length remained unchanged. It should be noted that they also reported no change in crystal dimensions of the cellulose nanocrystals when treated in a bulk dispersion. Nishiyama et al. studied the crystalline conversion from the cellulose I to cellulose I–EDA complex phases by time-resolved X-ray diffraction measurements, which suggested that the lateral width dimensions were limited to half of the cellulose I phase as a result of penetration of EDA molecules (Nishiyama et al. 2010). On the basis of the present aMD results, we wish to propose that crystal thinning can also take place on removal of EDA molecules even in the dispersion state.

## Diffusion behaviors of EDA molecule in the crystal models

The ASMD simulations were intended to simulate the diffusion of a single EDA molecule in the hydrophilic channel of the complex crystal, which is based on the discontinuous desorption of EDA molecules observed in the preceding aMD simulations. The time-course changes in PMF, with respect to EDA movement, for the one-, two-, and three-hole models are shown in Fig. 3. In the one-hole model, the PMF continued to increase with the movement of the target EDA molecule and reached an unexpected, large amount of endothermic free energy of more than 30 kcal/mol after the EDA molecule passed through the four cellulose chain layers. The PMF profile exhibited the four peaks being barely appreciable, where each peak corresponded to the position of a cellulose chain layer. Obviously, spontaneous EDA desorption would be impossible with the endothermic free energy change. Therefore, the PMF of the one-

hole model was reevaluated by applying the US method to the PMF trajectory derived from the ASMD simulation as a reaction coordinate. Figure 3 shows that both the ASMD–PMF and US–PMF profiles, initially overlapping, deviated to each other with a difference of approximately 5 kcal/mol near a 1.2 nm movement, while the latter exhibited a smoother curve. For further validation of the PMF of the one-hole model, ASMD simulations with a slower pulling speed of 0.2 Å/ns, combined with the US method, were also carried out only to reproduce similar PMF profiles (Figure S1).

The first US–MD trajectory was examined to analyze the spatial directions and conformations of the target EDA molecule by monitoring the changes in direction of the C-C bond vector and its rotation (Figures S2 and S3). The C-C rotations, defined by the N-C-C-N atom sequence, mostly preferred to be around the  $\pm$  *gauche* conformations. Therefore, the apparent rotational motions of the target EDA molecule observed in the visual inspection of the trajectory may reflect rotations of the molecule as a whole, rather than the C-C bond rotation. The directional change in the C-C bond vector with respect to the initial direction of the vector in the crystal structure somewhat diverged at the EDA positions from 0.5 nm to 1.0 nm and, otherwise, they were restricted to certain angles.

In contrast to the PMF profiles of the one-hole model, those of the two- and three-hole models, shown in Fig. 3, displayed moderate increases in the baselines with final values of a few kcal/mol, accompanied by four rounded peaks. The free energy changes indicated that EDA molecules can surmount the energy landscapes by the thermal motions at an ambient temperature. On the basis of these findings, we propose a plausible desorption scheme applicable to the real crystalline phase as follows. Once an initial EDA hole appears on the crystal surface after releasing several EDA molecules, those EDA molecules located in the adjacent hydrophilic channels can cooperatively migrate and diffuse into the bulk water. An alternative scheme would be that EDA desorption is initiated from the interfacial regions with amorphous or para-crystalline phases, or the terminals of crystal fibers where EDA molecules are loosely bound. In fact, EDA desorption apparently proceeds from both the terminals along the fiber axis in the aMD simulation shown in Movie S2.

We performed supplemental ASMD calculations for the one-hole model with the modified potential functions to confirm the above discussions. Figure S4 compares the PMF profiles obtained by the ASMD calculations adopting the modified potential functions with no torsional potential about the C-C and N-C bonds of an EDA molecule and by those implemented with the various dielectric constant values of 1.5, 2.0, and 4.0, along with the original PMF curve of the one-hole model given in Fig. 3. The PMF with no torsional potential of an EDA molecule was lowered from the original one, although both exhibited similar appearances in their profiles. The amount of decrease was moderately limited despite the conformation of the EDA molecule being allowed to change freely. More notable changes were observed as the dielectric constant value was increased, which resulted in not only effectively suppressing the PMF level but also changing the shapes of the PMF profile to be similar to those of the two- and three-hole model, as represented by the rounded peaks with moderate or gentle slopes. In the original crystal structure, EDA molecules participate in the intermolecular hydrogen bonding network with O2, O3, and O6 hydroxyl groups of the cellulose chains along the *a*-axis direction. It is unlikely that, in the one-hole model, the

unfavorable endothermic PMF is caused by breaking the intermolecular hydrogen bonds involving the target EDA, because it can form another hydrogen bond in the next cellulose chain layer when it moves. A high frequency of breaking and formation of hydrogen bonds caused by thermal motions of the hydroxyl groups was suggested in the real cellulose I–EDA complex (Sawada et al. 2013). Instead, a certain dragging force against the movement of the target EDA may arise in the one-hole channel, such that EDA migration successively generates an endothermic free energy, although we do not have a straightforward interpretation for the origin of the dragging force at present. The present ASMD simulation constrained the positions of the pyranose backbones and carbon atoms of EDA molecules, except the target EDA molecule, while allowing the exocyclic and amino groups to rotate. The functional groups can more frequently rotate as they become liberated from the intermolecular hydrogen bonds, which is introduced either by removal of adjacent EDA molecules in the two- and three-hole models or attenuation of electrostatic interactions introduced by larger dielectric values. Consequently, the target EDA molecule can find a lower energy path with a higher probability, especially as it moves through the cellulose chain layers. This may be applicable to lowering the PMF in the simulation with no torsional potential of an EDA molecule, although the effect is relatively restrictive compared with that adapting larger dielectric values.

While the introduction of the multi-hole models was to modify the crystal models along the fiber axis, the gap models were examined for the effect of expanding the crystal models along the *b*-axis (see Fig. 1). Figure 4 compares the PMF profiles of the five gap models with different gap spaces and that of the original model shown in Fig. 3, where the ASMD calculations were performed only for the 1st step with a pulling distance of 0.4 nm. The PMF curves drastically decreased as the gap space increased from 0.4 nm to 0.6 nm. The final PMF value of the latter model indicated that spontaneous diffusion of EDA molecules can occur from the interior of the crystal model. Slight exothermic free energy changes were detected in the PMF curves near a pulling distance of 0.03 nm for the three gap models with a gap space of 0.1, 0.2, or 0.3 nm, where positional relaxation of the target EDA molecule occurs as a result of the developing attractive interactions with the surrounding cellulose chains. However, the target EDA molecules in the 0.4 and 0.6 nm gap models, which exhibited no relaxations, were solvated with water during the preceding equilibration process. The results suggest that water solvation of an EDA molecule plays an important role in its spontaneous departure from the original, crystalline position, which is consistent with the discontinuous diffusion observed in the aMD simulations. It is likely that diffusion of EDA molecules is also promoted by their solvation in the real crystalline phase.

## Conclusions

We have reported desorption behaviors of guest EDA molecules from the cellulose I–EDA complex crystal models studied by adopting a combination of three extended ensemble MD methods. EDA molecules migrate along the hydrophilic channel between the cellulose chain piles on the (010) lattice plane. The same EDA migration path is proposed experimentally on complexation of native cellulose fibers with EDA (Nishiyama et al. 2010). The diffusion behavior of EDA molecules appears to be discontinuous; they migrate in the channel one at a time. A vacancy in the channel emerging after EDA diffusion is quickly

filled by water molecules. This obviously suggests that a departure of an EDA molecule from the crystal position is, at least partly, triggered by its solvation. Following complete desolvation and rearrangement of cellulose chains, crystalline conversion is allowed. The scheme may reasonably be consistent with the observation that the destiny of the crystalline phase depends on the solvent used to rinse the cellulose I–EDA complex specimen; cellulose I by hot-water treatment and cellulose III, by methanol or ethanol treatment (Loeb and Segal 1955; Lokhande et al. 1977; Roche and Chanzy 1981; Sugiyama et al. 1987). It has been also suggested by the final structure of the aMD crystal model that drastic desorption of EDA may lead to crystal degradation and fiber thinning in the real complex crystals.

The present MD simulations have also suggested that while a large endothermic free energy is required for EDA diffusion under the regular crystalline environment, EDA molecules readily cooperatively migrate in the hydrophilic channel once the adjacent channels have discharged their own EDA molecules along with the loss of intermolecular hydrogen bonds. Diffusion of EDA molecules may be promoted by the increased rotational frequencies of the exocyclic groups of cellulose chains, which are free from restrictions of intermolecular hydrogen bonds. Naturally, enlarging the space of the hydrophilic channel would be a direct solution to accelerate EDA diffusion and desorption. However, the space needs to be expanded so that it is large enough to afford solvation of EDA molecules for spontaneous diffusion of them from the crystal interior.

One may be critical of applying the present simulation results to describe EDA desorption of the real crystalline phase, especially those obtained from the ASMD simulations where the crystal models essentially retain the regular crystalline framework, in contrast to the real crystalline phase that possibly involves disordered features on EDA desorption. However, disordered or para-crystal modeling would be a much more difficult issue; there is neither applicable theory nor experimental evidence to describe a representative configuration of a disordered cellulose crystal with atomistic detail. In the present MD study, the various crystal models have been assessed by combination of the three extended ensemble approaches. The results are reasonably correlated to each other and lead us to the specific conclusions described above, some of which are in accordance with the experimental results. The present simulation methodology can be applicable to similar complex systems, such as a cellulose I–ammonia complex crystal, which is currently underway.

## **Declarations**

### **Acknowledgements**

MD calculations were partly performed by using the computer cluster systems at Research Center for Computational Science (RCCS), Okazaki Research Facilities, and National Institutes of Natural Sciences (NINS), Japan. We thank Edanz (<https://www.jp.edanz.com/ac>) for editing a draft of this manuscript.

### **Ethics declarations**

### **Conflict of interest**

The authors declare that they have no conflict of interest.

## Ethical approval

The authors have no statement about ethical approval as this research does not involve humans and animals.

## Informed consent

The authors have no statements about informed consent as this research does not involve human participants.

## Funding

The authors did not receive support from any organization for the submitted work.

## References

1. Beckham GT, Crowley MF (2011) Examination of the  $\alpha$ -chitin structure and decrystallization thermodynamics at the nanoscale. *J Phys Chem B* 115:4516–4522. <https://doi.org/10.1021/jp200912q>
2. Beckham GT, Matthews JF, Peters B, Bomble YJ, Himmel ME, Crowley MF (2011) Molecular-level origins of biomass recalcitrance: decrystallization free energies for four common cellulose polymorphs. *J Phys Chem B* 115:4118–4127. <https://doi.org/10.1021/jp1106394>
3. Bellesia G, Chundawat SP, Langan P, Dale BE, Gnanakaran S (2011) Probing the early events associated with liquid ammonia pretreatment of native crystalline cellulose. *J Phys Chem B* 115:9782–9788. <https://doi.org/10.1021/jp2048844>
4. Berendsen HJC, Postma JPM, van Gunsteren WF, DiNola A, Haak JR (1984) Molecular dynamics with coupling to an external bath. *J Chem Phys* 81:3684–3690. <https://doi.org/10.1063/1.448118>
5. Case DA, Ben-Shalom IY, Brozell SR, Cerutti DS, Cheatham III TE, Cruzeiro VWD, Darden TA, Duke RE, Ghoreishi D, Gilson MK, Gohlke H, Goetz AW, Greene D, Harris R, Homeyer N, Huang Y, Izadi S, Kovalenko A, Kurtzman T, Lee TS, LeGrand S, Li P, Lin C, Liu J, Luchko T, Luo R, Mermelstein DJ, Merz KM, Miao Y, Monard G, Nguyen C, Nguyen H, Omelyan I, Onufriev A, Pan F, Qi R, Roe DR, Roitberg A, Sagui C, Schott-Verdugo S, Shen J, Simmerling CL, Smith J, Salomon-Ferrer R, Swails J, Walker RC, Wang J, Wei H, Wolf RM, Wu X, Xiao L, York DM, Kollman PA (2018) Amber 18, University of California, San Francisco, CA <https://ambermd.org>
6. Chanzy H, Henrissat B, Vincendon M, Tanner SF, Belton PS (1987) Solid-state  $^{13}\text{C}$ -N.M.R. and electron microscopy study on the reversible cellulose I $\rightarrow$ cellulose III $_1$  transformation in *Valonia*. *Carbohydr Res* 160:1–11. [https://doi.org/10.1016/0008-6215\(87\)80299-9](https://doi.org/10.1016/0008-6215(87)80299-9)

7. Chundawat SP, Bellesia G, Uppugundla N, da Costa Sousa L, Gao D, Cheh AM, Agarwal UP, Bianchetti CM, Phillips GN, Jr., Langan P, Balan V, Gnanakaran S, Dale BE (2011) Restructuring the crystalline cellulose hydrogen bond network enhances its depolymerization rate. *J Am Chem Soc* 133:11163–11174.  
<https://doi.org/10.1021/ja2011115>
8. Clark GL, Parker EA (1937) An X-ray diffraction study of the action of liquid ammonia on cellulose and its derivatives. *J Phys Chem* 41:777–786. <https://doi.org/10.1021/j150384a001>
9. Davis WE, Barry AJ, Peterson FC, King AJ (1943) X-ray studies of reactions of cellulose in non-aqueous systems. II. Interaction of cellulose and primary amines. *J Am Chem Soc* 65:1294–1299. <https://doi.org/10.1021/ja01247a012>
10. Essmann U, Perera L, Berkowitz ML, Darden T, Lee H, Pedersen LG (1995) A smooth particle mesh Ewald method. *J Chem Phys* 103:8577–8593. <https://doi.org/10.1063/1.470117>
11. Evans DJ (1983) Computer "experiment" for nonlinear thermodynamics of Couette flow. *J Chem Phys* 78:3297–3302. <https://doi.org/10.1063/1.445195>
12. French AD, Dowd MK (1993) Exploration of disaccharide conformations by molecular mechanics. *J Mol Struct (Theochem)* 286:183–201. [https://doi.org/10.1016/0166-1280\(93\)87162-7](https://doi.org/10.1016/0166-1280(93)87162-7)
13. French AD, Miller DP, Aabloo A (1993) Miniature crystal models of cellulose polymorphs and other carbohydrates. *Int J Biol Macromol* 15:30–36.  
[https://doi.org/10.1016/s0141-8130\(05\)80085-6](https://doi.org/10.1016/s0141-8130(05)80085-6)
14. Grossfield A (2018) WHAM: an implementation of the weighted histogram analysis method, ver. 2.0.10, <http://membrane.urmc.rochester.edu/content/wham>: University of Rochester Medical Center: Rochester, NY
15. Hamelberg D, Mongan J, McCammon JA (2004) Accelerated molecular dynamics: A promising and efficient simulation method for biomolecules. *J Chem Phys* 120:11919–11929.  
<https://doi.org/10.1063/1.1755656>
16. Hamelberg D, de Oliveira CA, McCammon JA (2007) Sampling of slow diffusive conformational transitions with accelerated molecular dynamics. *J Chem Phys* 127:155102.  
<https://doi.org/10.1063/1.2789432>
17. Hoover WG, Ladd AJC, Moran B (1982) High-strain-rate plastic flow studied via nonequilibrium molecular dynamics. *Phys Rev Lett* 48:1818–1820.  
<https://doi.org/10.1103/PhysRevLett.48.1818>
18. Humphrey W, Dalke A, Schulten K (1996) VMD: Visual molecular dynamics. *J Mol Graph* 14:33–38.  
[https://doi.org/10.1016/0263-7855\(96\)00018-5](https://doi.org/10.1016/0263-7855(96)00018-5)
19. Igarashi K, Wada M, Samejima M (2007) Activation of crystalline cellulose to cellulose III<sub>I</sub> results in efficient hydrolysis by cellobiohydrolase. *FEBS J* 274:1785–1792.  
<https://doi.org/10.1111/j.1742-4658.2007.05727.x>
20. Igarashi K, Uchihashi T, Koivula A, Wada M, Kimura S, Okamoto T, Penttila M, Ando T, Samejima M (2011) Traffic jams reduce hydrolytic efficiency of cellulase on cellulose surface. *Science* 333:1279–

1282. <https://doi.org/10.1126/science.1208386>
21. Jarzynski C (1997) Nonequilibrium equality for free energy differences. *Phys Rev Lett* 78:2690–2693. <https://doi.org/10.1103/PhysRevLett.78.2690>
22. Jorgensen WL (1981) Quantum and statistical mechanical studies of liquids. 10. Transferable intermolecular potential functions for water, alcohols, and ethers. Application to liquid water. *J Am Chem Soc* 103:335–340. <https://doi.org/10.1021/ja00392a016>
23. Jorgensen WL, Chandrasekhar J, Madura JD, Impey RW, Klein ML (1983) Comparison of simple potential functions for simulating liquid water. *J Chem Phys* 79:926–935. <https://doi.org/10.1063/1.445869>
24. Kirschner KN, Yongye AB, Tschampel SM, Gonzalez-Outeirino J, Daniels CR, Foley BL, Woods RJ (2008) GLYCAM06: A generalizable biomolecular force field. *Carbohydrates*. *J Comput Chem* 29:622–655. <https://doi.org/10.1002/jcc.20820>
25. Kumar S, Rosenberg JM, Bouzida D, Swendsen RH, Kollman PA (1992) THE weighted histogram analysis method for free-energy calculations on biomolecules. I. The method. *J Comput Chem* 13:1011–1021. <https://doi.org/10.1002/jcc.540130812>
26. Kumar S, Rosenberg JM, Bouzida D, Swendsen RH, Kollman PA (1995) Multidimensional free-energy calculations using the weighted histogram analysis method. *J Comput Chem* 16:1339–1350. <https://doi.org/10.1002/jcc.540161104>
27. Lau MW, Dale BE (2009) Cellulosic ethanol production from AFEX-treated corn stover using *Saccharomyces cerevisiae* 424A(LNH-ST). *Proc Natl Acad Sci USA* 106:1368–1373. <https://doi.org/10.1073/pnas.0812364106>
28. Le Grand S, Götz AW, Walker RC (2013) SPFP: Speed without compromise—A mixed precision model for GPU accelerated molecular dynamics simulations. *Comput Phys Commun* 184:374–380. <https://doi.org/10.1016/j.cpc.2012.09.022>
29. Loeb L, Segal L (1955) Studies of the ethylenediamine-cellulose complex. I. Decomposition of the complex by solvents. *J Polym Sci* 15:343–354. <https://doi.org/10.1002/pol.1955.120158003>
30. Lokhande HT, Shukla SR, Chidambareswaran PK, Patil NB (1977) Ethylenediamine-induced conversion of cellulose I to cellulose III. *Journal of Polymer Science: Polymer Letters Edition* 15:97–99. <https://doi.org/10.1002/pol.1977.130150207>
31. Matthews JF, Skopec CE, Mason PE, Zuccato P, Torget RW, Sugiyama J, Himmel ME, Brady JW (2006) Computer simulation studies of microcrystalline cellulose I $\beta$ . *Carbohydr Res* 341:138–152. <https://doi.org/10.1016/j.carres.2005.09.028>
32. Miao Y, Sinko W, Pierce L, Bucher D, Walker RC, McCammon JA (2014) Improved reweighting of accelerated molecular dynamics simulations for free energy calculation. *J Chem Theory Comput* 10:2677–2689. <https://doi.org/10.1021/ct500090q>
33. Nishiyama Y, Langan P, Chanzy H (2002) Crystal structure and hydrogen-bonding system in cellulose I $\beta$  from synchrotron X-ray and neutron fiber diffraction. *J Am Chem Soc* 124:9074–9082.

<https://doi.org/10.1021/ja0257319>

34. Nishiyama Y, Wada M, Hanson BL, Langan P (2010) Time-resolved X-ray diffraction microprobe studies of the conversion of cellulose I to ethylenediamine-cellulose I. *Cellulose* 17:735–745. <https://doi.org/10.1007/s10570-010-9415-9>
35. Ozer G, Valeev EF, Quirk S, Hernandez R (2010) Adaptive steered molecular dynamics of the long-distance unfolding of neuropeptide Y. *J Chem Theory Comput* 6:3026–3038. <https://doi.org/10.1021/ct100320g>
36. Ozer G, Quirk S, Hernandez R (2012) Adaptive steered molecular dynamics: validation of the selection criterion and benchmarking energetics in vacuum. *J Chem Phys* 136:215104. <https://doi.org/10.1063/1.4725183>
37. Park S, Khalili-Araghi F, Tajkhorshid E, Schulten K (2003) Free energy calculation from steered molecular dynamics simulations using Jarzynski's equality. *J Chem Phys* 119:3559–3566. <https://doi.org/10.1063/1.1590311>
38. Park S, Schulten K (2004) Calculating potentials of mean force from steered molecular dynamics simulations. *J Chem Phys* 120:5946–5961. <https://doi.org/10.1063/1.1651473>
39. Pentoney RE (1966) Liquid ammonia-solvent combinations in wood plasticization. Properties of treated wood. *Ind Eng Chem Prod Res Dev* 5:105–110. <https://doi.org/10.1021/i360018a003>
40. Perez DDS, Montanari S, Vignon MR (2003) TEMPO-mediated oxidation of cellulose III. *Biomacromolecules* 4:1417–1425. <https://doi.org/10.1021/bm034144s>
41. Pierce LCT, Salomon-Ferrer R, de Oliveira CAF, McCammon JA, Walker RC (2012) Routine access to millisecond time scale events with accelerated molecular dynamics. *J Chem Theory Comput* 8:2997–3002. <https://doi.org/10.1021/ct300284c>
42. Roche E, Chanzy H (1981) Electron microscopy study of the transformation of cellulose I into cellulose III<sub>I</sub> in *Valonia*. *Int J Biol Macromol* 3:201–206. [https://doi.org/10.1016/0141-8130\(81\)90064-7](https://doi.org/10.1016/0141-8130(81)90064-7)
43. Roe DR, Cheatham TE, III (2013) PTRAJ and CPPTRAJ: Software for processing and analysis of molecular dynamics trajectory data. *J Chem Theory Comput* 9:3084–3095. <https://doi.org/10.1021/ct400341p>
44. Roux B (1995) The calculation of the potential of mean force using computer simulations. *Comput Phys Commun* 91:275–282. [https://doi.org/10.1016/0010-4655\(95\)00053-i](https://doi.org/10.1016/0010-4655(95)00053-i)
45. Ryckaert J-P, Ciccotti G, Berendsen HJC (1977) Numerical integration of the cartesian equations of motion of a system with constraints: Molecular dynamics of n-alkanes. *J Comput Phys* 23:327–341. [https://doi.org/10.1016/0021-9991\(77\)90098-5](https://doi.org/10.1016/0021-9991(77)90098-5)
46. Salminen R, Baccile N, Reza M, Kontturi E (2017) Surface-induced frustration in solid state polymorphic transition of native cellulose nanocrystals. *Biomacromolecules* 18:1975–1982. <https://doi.org/10.1021/acs.biomac.7b00463>

47. Salomon-Ferrer R, Gotz AW, Poole D, Le Grand S, Walker RC (2013a) Routine microsecond molecular dynamics simulations with AMBER on GPUs. 2. Explicit solvent particle mesh ewald. *J Chem Theory Comput* 9:3878–3888. <https://doi.org/10.1021/ct400314y>
48. Salomon-Ferrer R, Pierce L, Walker R (2013b) Accelerated molecular dynamics in AMBER example analysis of all-atom enhanced sampling method accelerated molecular dynamics (aMD) to investigate conformational changes in proteins that typically occur on the millisecond time scale. <https://ambermd.org/tutorials/advanced/tutorial22>. Accessed June 1, 2021
49. Sawada D, Nishiyama Y, Petridis L, Parthasarathi R, Gnanakaran S, Forsyth VT, Wada M, Langan P (2013) Structure and dynamics of a complex of cellulose with EDA: insights into the action of amines on cellulose. *Cellulose* 20:1563–1571. <https://doi.org/10.1007/s10570-013-9974-7>
50. Schrödinger LLC. (2021) The PyMOL molecular graphics system, ver. 2.5.0, <https://pymol.org>: New York, NY
51. Sugiyama J, Harada H, Saiki H (1987) Crystalline morphology of *Valonia macrophysa* cellulose III<sub>I</sub> revealed by direct lattice imaging. *Int J Biol Macromol* 9:122–130. [https://doi.org/10.1016/0141-8130\(87\)90039-0](https://doi.org/10.1016/0141-8130(87)90039-0)
52. Torrie GM, Valleau JP (1974) Monte Carlo free energy estimates using non-Boltzmann sampling: Application to the sub-critical Lennard-Jones fluid. *Chem Phys Lett* 28:578–581. [https://doi.org/10.1016/0009-2614\(74\)80109-0](https://doi.org/10.1016/0009-2614(74)80109-0)
53. Torrie GM, Valleau JP (1977) Nonphysical sampling distributions in Monte Carlo free-energy estimation: Umbrella sampling. *J Comput Phys* 23:187–199. [https://doi.org/10.1016/0021-9991\(77\)90121-8](https://doi.org/10.1016/0021-9991(77)90121-8)
54. Uto T, Hosoya T, Hayashi S, Yui T (2013) Partial crystalline transformation of solvated cellulose III<sub>I</sub> crystals, reproduced by theoretical calculations. *Cellulose* 20:605–612. <https://doi.org/10.1007/s10570-013-9861-2>
55. Uto T, Mawatari S, Yui T (2014) Theoretical study of the structural stability of molecular chain sheet models of cellulose crystal allomorphs. *J Phys Chem B* 118:9313–9321. <https://doi.org/10.1021/jp503535d>
56. Uto T, Yui T (2018) DFT optimization of isolated molecular chain sheet models constituting native cellulose crystal structures. *ACS Omega* 3:8050–8058. <https://doi.org/10.1021/acsomega.8b00834>
57. Uto T, Minamizaki M, Yui T (2019) Molecular dynamics simulation of cellulose I-ethylenediamine complex crystal models. *J Phys Chem B* 124:134–143. <https://doi.org/10.1021/acs.jpcc.9b08153>
58. Wada M (2001) In situ observation of the crystalline transformation from cellulose III<sub>I</sub> to I $\beta$ . *Macromolecules* 34:3271–3275. <https://doi.org/10.1021/ma0013354>
59. Wada M, Heux L, Isogai A, Nishiyama Y, Chanzy H, Sugiyama J (2001) Improved structural data of cellulose III<sub>I</sub> prepared in supercritical ammonia. *Macromolecules* 34:1237–1243.

<https://doi.org/10.1021/ma001406z>

60. Wada M, Chanzy H, Nishiyama Y, Langan P (2004) Cellulose III<sub>I</sub> crystal structure and hydrogen bonding by synchrotron X-ray and neutron fiber diffraction. *Macromolecules* 37:8548–8555. <https://doi.org/10.1021/ma0485585>
61. Wada M, Nishiyama Y, Langan P (2006) X-ray structure of ammonia–cellulose I: New insights into the conversion of cellulose I to cellulose III<sub>I</sub>. *Macromolecules* 39:2947–2952. <https://doi.org/10.1021/ma060228s>
62. Wada M, Kwon GJ, Nishiyama Y (2008) Structure and thermal behavior of a cellulose I–ethylenediamine complex. *Biomacromolecules* 9:2898–2904. <https://doi.org/10.1021/bm8006709>
63. Wada M, Heux L, Nishiyama Y, Langan P (2009a) X-ray crystallographic, scanning microprobe X-ray diffraction, and cross-polarized/magic angle spinning <sup>13</sup>C NMR studies of the structure of cellulose III<sub>II</sub>. *Biomacromolecules* 10:302–309. <https://doi.org/10.1021/bm8010227>
64. Wada M, Heux L, Nishiyama Y, Langan P (2009b) The structure of the complex of cellulose I with ethylenediamine by X-ray crystallography and cross-polarization/magic angle spinning <sup>13</sup>C nuclear magnetic resonance. *Cellulose* 16:943–957. <https://doi.org/10.1007/s10570-009-9338-5>
65. Wada M, Nishiyama Y, Bellesia G, Forsyth T, Gnanakaran S, Langan P (2011) Neutron crystallographic and molecular dynamics studies of the structure of ammonia–cellulose I: rearrangement of hydrogen bonding during the treatment of cellulose with ammonia. *Cellulose* 18:191–206. <https://doi.org/10.1007/s10570-010-9488-5>
66. Wang J, Wolf RM, Caldwell JW, Kollman PA, Case DA (2004) Development and testing of a general Amber force field. *J Comput Chem* 25:1157–1174. <https://doi.org/10.1002/jcc.20035>
67. Yui T, Nishimura S, Akiba S, Hayashi S (2006) Swelling behavior of the cellulose Iβ crystal models by molecular dynamics. *Carbohydr Res* 341:2521–2530. <https://doi.org/10.1016/j.carres.2006.04.051>
68. Yui T, Taki N, Sugiyama J, Hayashi S (2007) Exhaustive crystal structure search and crystal modeling of β-chitin. *Int J Biol Macromol* 40:336–344. <https://doi.org/10.1016/j.ijbiomac.2006.08.017>
69. Yui T, Hayashi S (2008) Structural stability of the solvated cellulose III<sub>I</sub> crystal models: a molecular dynamics study. *Cellulose* 16:151–165. <https://doi.org/10.1007/s10570-008-9265-x>
70. Yui T, Okayama N, Hayashi S (2010) Structure conversions of cellulose III<sub>I</sub> crystal models in solution state: A molecular dynamics study. *Cellulose* 17:679–691. <https://doi.org/10.1007/s10570-010-9422-x>
71. Zheng Y, Pan Z, Zhang R (2009) Overview of biomass pretreatment for cellulosic ethanol production. *Int J Agric & Biol Eng* 2:51–68.

## Figures

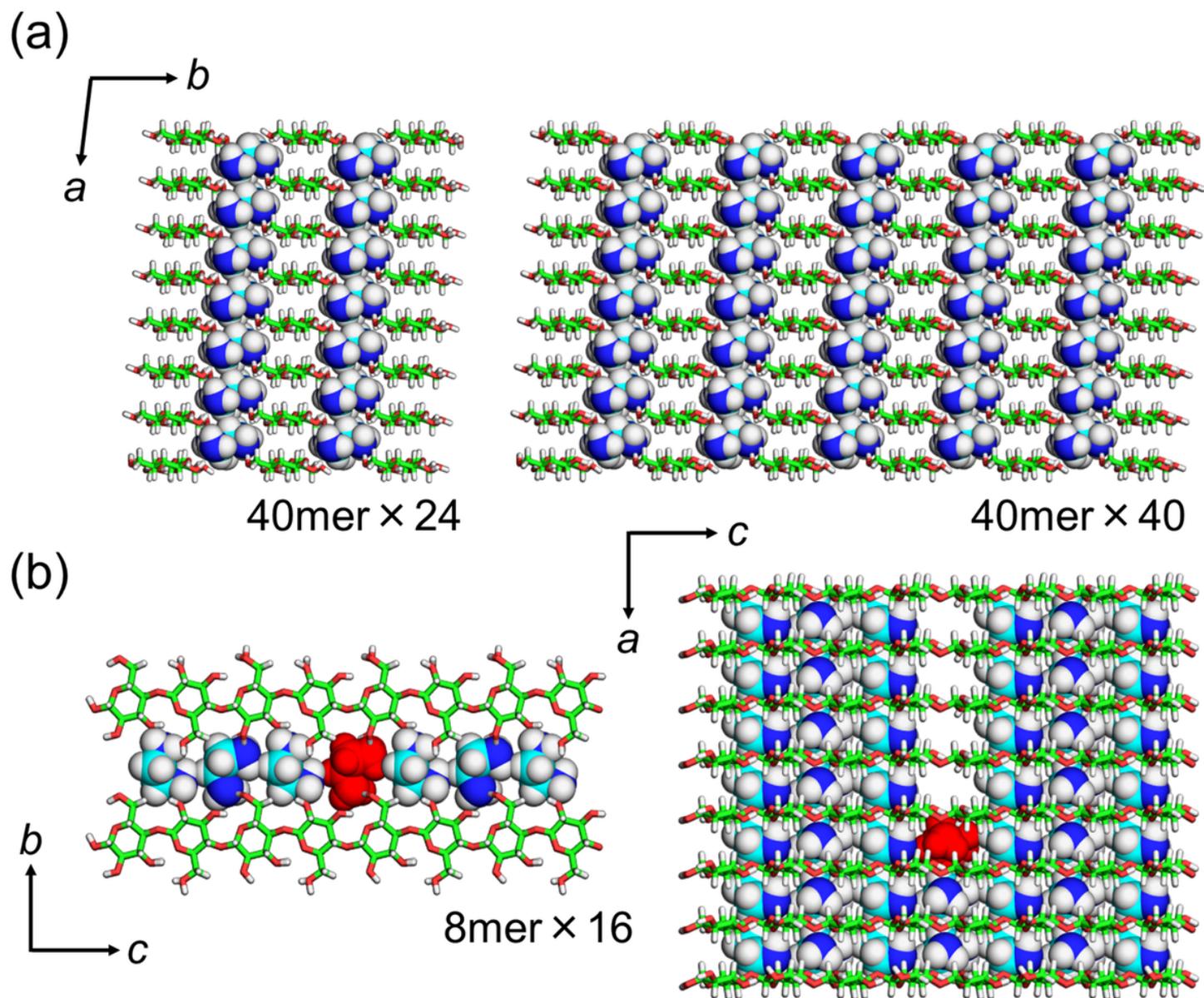


Figure 1

(a) Projections of the 40mer  $\times$  24 and 40mer  $\times$  48 models adopted in the aMD simulation, and (b) the 8mer  $\times$  16 model in the ASMD simulation. Cellulose chains are drawn in stick representation and EDA molecules in sphere representation. The target EDA molecule for ASMD is shown in red in the 8mer  $\times$  16 model.

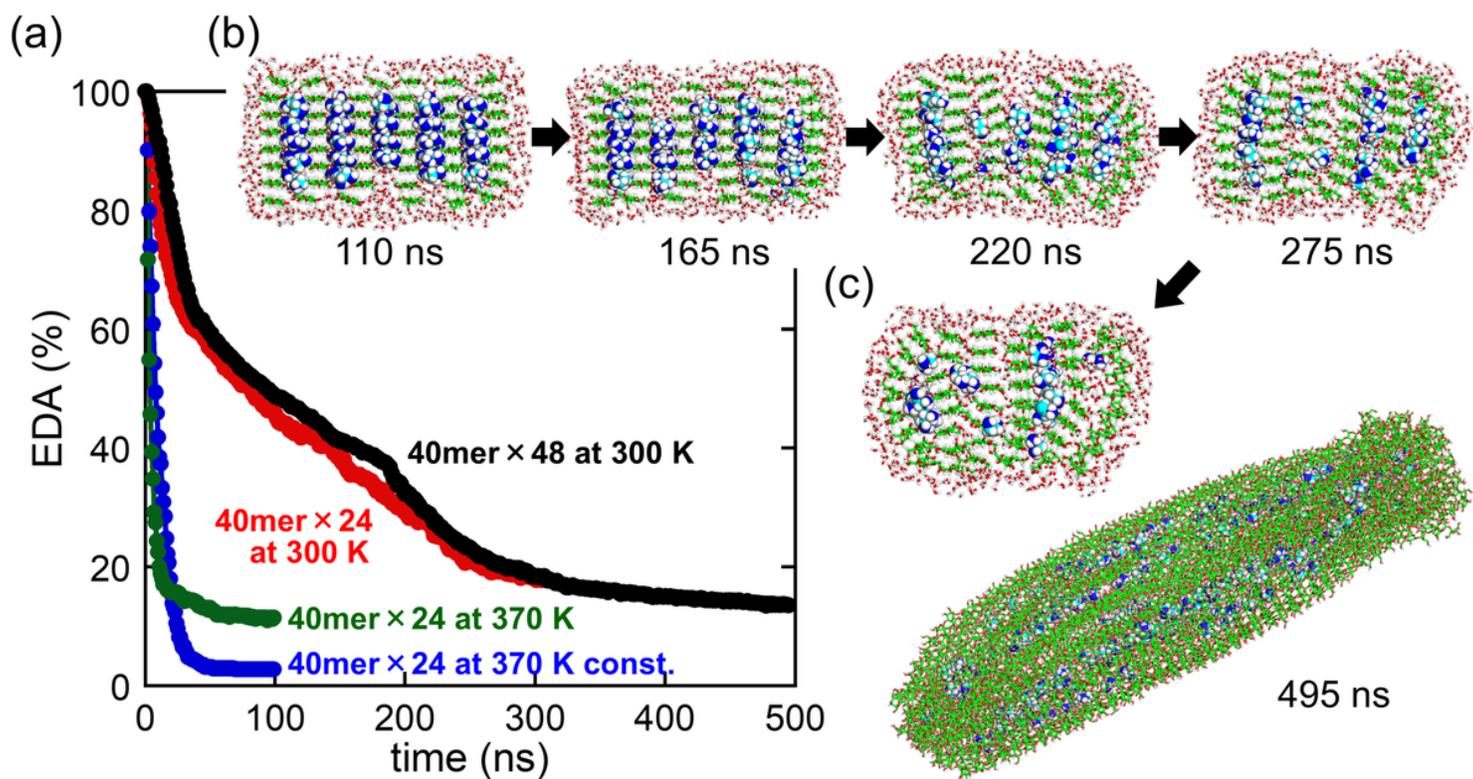


Figure 2

(a) Change in amount of residual EDA in the crystal model with respect to simulation time. (b) Time course of EDA desorption viewed from the sliced cross-sections of the 40mer x 48 model. The sliced cross-section view and overall structure of the 40mer x 48 models at the final time frame of 495 ns.

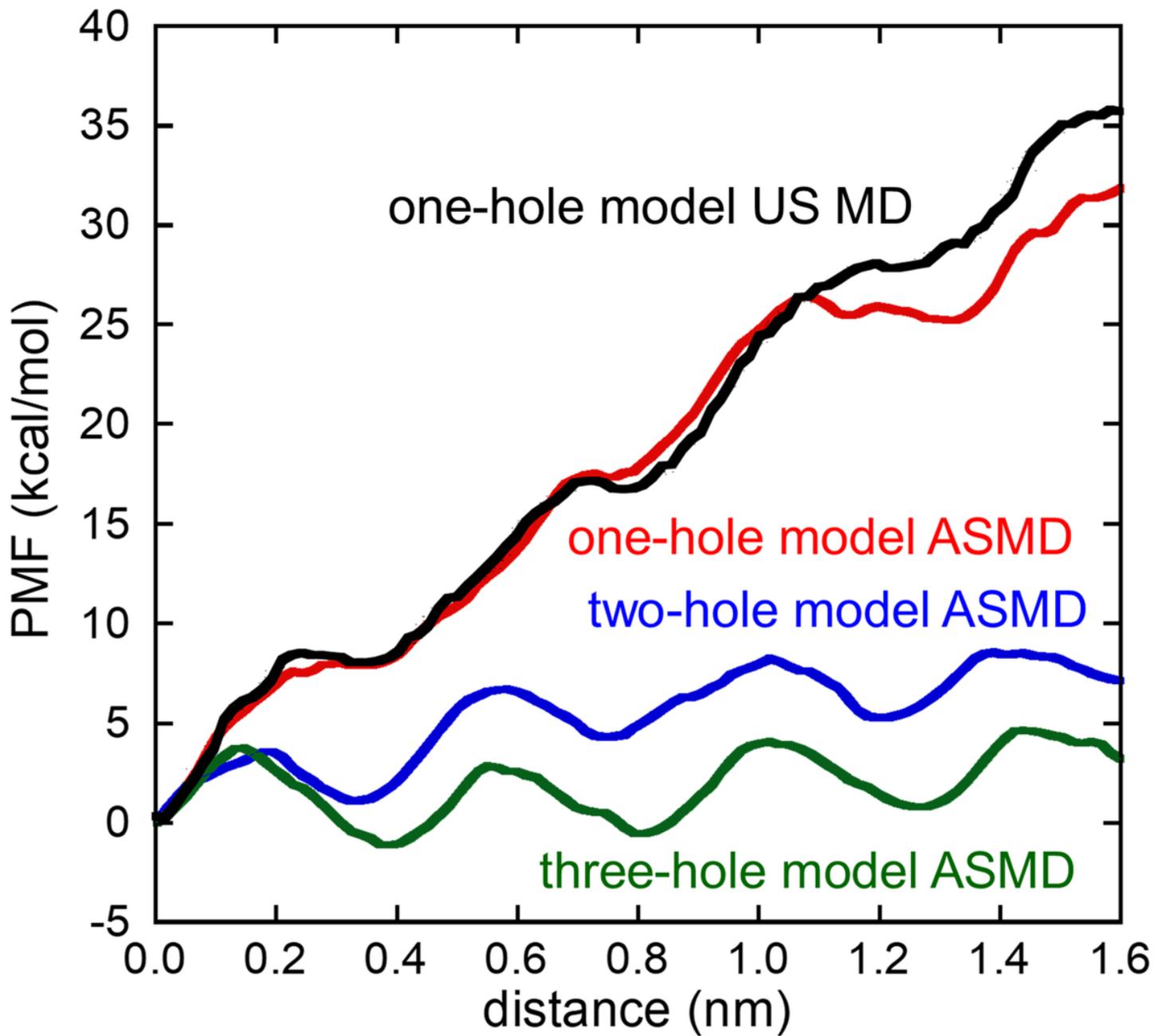


Figure 3

PMF profiles with respect to the pulling distance of the target EDA molecule obtained from ASMD simulations for the one-, two-, and three-hole models and combined MD simulation using the US method for the one-hole model.

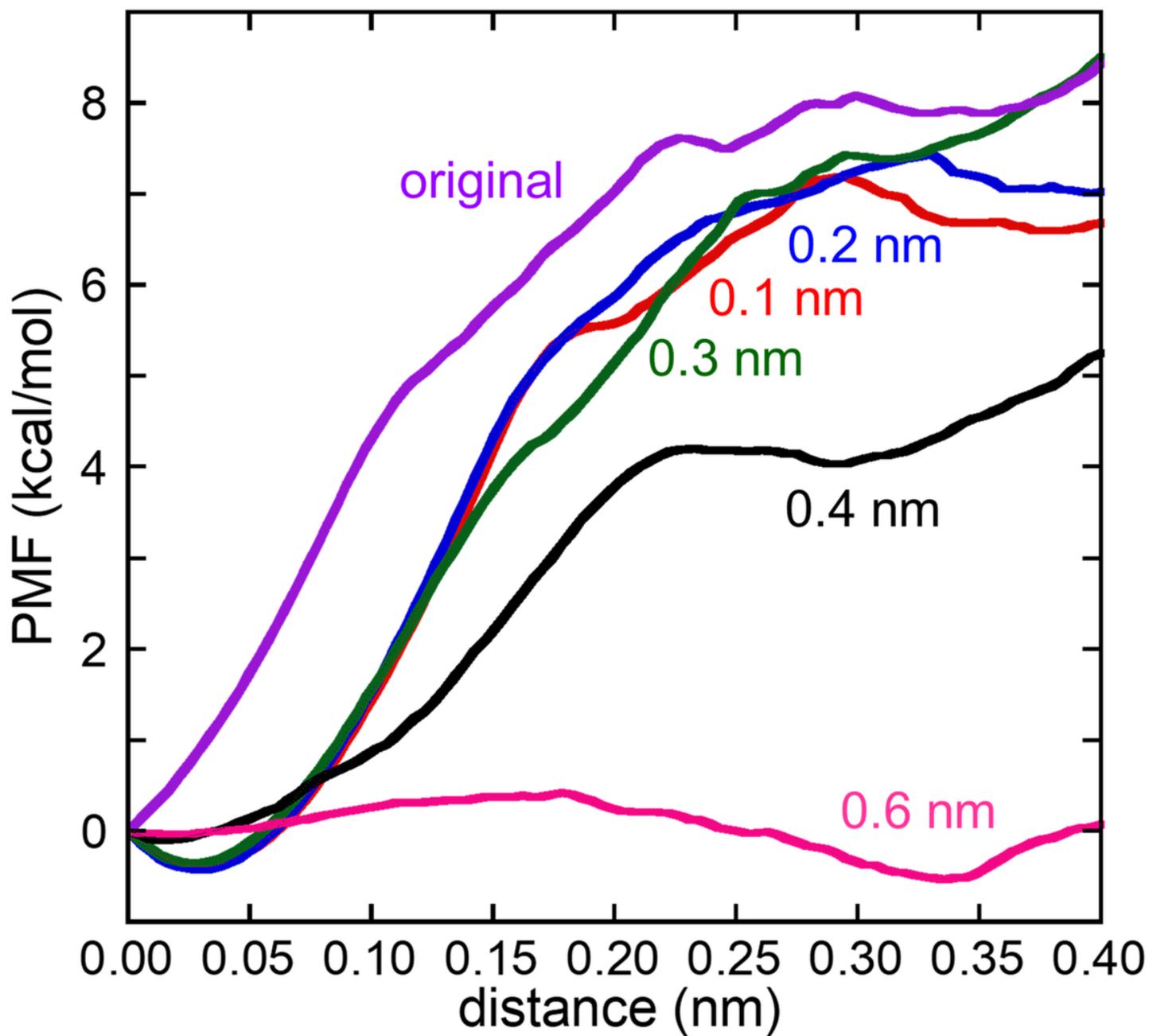


Figure 4

PMF profiles with respect to the pulling distance of the target EDA molecule obtained from ASMD simulations for the one-hole gap models with different gap spaces as noted. Original: the PMF profile of the original one-hole model.

## Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [graphicalabstract.png](#)

- S1a.png
- S2.png
- S3.png
- S4a.png
- movies1.mp4
- movies2.mp4
- supplementaryinformationCelluloseyui.docx