J|A|C|S COMMUNICATIONS

Published on Web 08/26/2010

Pyranose Ring Transition State Is Derived from Cellobiohydrolase I Induced Conformational Stability and Glycosidic Bond Polarization

Christopher B. Barnett, Karl A. Wilkinson, and Kevin J. Naidoo*

Scientific Computing Research Unit and Department of Chemistry, University of Cape Town, Rondebosch 7701, South Africa

Received May 12, 2010; E-mail: kevin.naidoo@uct.ac.za

Abstract: Understanding carbohydrate ring pucker is critical to rational design in materials and pharmaceuticals. Recently we have generalized our adaptive reaction coordinate force biasing method to perform calculations on multidimensional reaction coordinates. We termed this the *Free Energies from Adaptive Reaction Coordinate Forces* (FEARCF) method. Using FEARCF in SCC-DFTB QM/MM non-Boltzmann simulations, we are able to calculate multidimensional ring pucker free energies of conformation. Here we apply this to the six-membered glucopyranose ring located in an eight-membered β 1−4 linked octaose oligosaccharide (cellooctaose). The cellooctaose was built following the conformation of the saccharides bound to cellobiohydrolase I (CBHI) of *Trichoderma reesei* as reported in the 7CEL crystal structure obtained from the PDB. We calculate the free energy of ring puckering of the glucopyranose ring at the −1 position in vacuum, in water, and bound to the protein. We find that the protein induces ^{4}E and $^4H_3$ conformations that are much more stable than the usually preferred $^4C_1$ conformer. Furthermore, for the $^4H_3$ conformation in the catalytic binding domain, there is significant electronic rearrangement that drives the structure toward the transition state of the glycosylation reaction.

Cellulases catalyze the hydrolysis of cellulose, a linear polymer of glucose found in all plants, in a reaction that is a key step in the production of the commercial biofuel, ethanol. In retaining glycosidases, hydrolysis occurs via the glycosylation and deglycosylation of oligosaccharides where in each of these reactions the glucose ring puckers away from the commonly occurring $^4C_1$ solution chair conformation to take on strained conformers in the transition state. However, the lignocellulosic biomass has evolved to resist enzymatic assisted breakdown[1] making enzymatic cellulose conversion, despite years of research, a slow process compared with the breakdown of amylose. Answering the questions of how and why the sugar rings are puckered in hydrolysis reactions may be pivotal in understanding the relative inefficiency of the cellulases.

The enzymes secreted by the high cellulase producing filamentous fungus *Trichoderma reesei* (*T. reesei*) are known to be highly effective against native, crystalline cellulose. *T. reesei* produces cellobiohydrolase I and II (CBHI and CBHII) which hydrolyze cellulose chains from their reducing and nonreducing ends respectively, with CBHI able to act on the inaccessible, crystalline regions of cellulose.[2,3] It cleaves off cellobiose from the chain ends in a processive manner.[4−6] Here, we investigate the nature of glucose ring puckering at the site in cellulose where cellobiose is being cleaved by CBHI. This puckering is considered at the stage preceding glycosylation.

The IUPAC convention defines all possible puckers of a six-membered ring into chair (C), boat (B), skew-boat (S), envelope (E), and half-chair (H) conformers.[7] Analysis and classification of these canonical conformers is most commonly done in terms of Cremer and Pople coordinates.[8] The Cremer−Pople pucker coordinates (q_i) are often transformed into spherical polar coordinates to more easily interpret the pucker conformation. More recently, Hill and Reilly proposed an intuitive description of ring puckering.[9] Similar to the classic Cremer−Pople definition, a puckering conformation is described as deviations from a mean plane. The Hill−Reilly definitions are derived from the triangular decomposition of an *n*-membered monocyclic ring into rotatable planes about a reference plane. The pucker coordinates are specified by the angle that the rotatable planes make with respect to the reference plane. For $n = 6$ there are three puckering coordinates θ_0, θ_1, $\theta_2 \in [-\pi/2,\pi/2]$. The angle of puckering is calculated from

$$\theta_i = \pi/2 - \cos^{-1}[(q_i \cdot n) \cdot (||q_i|| \cdot ||n||)^{-1}] \qquad (1)$$

where $\boldsymbol{q_i}$ is a vector normal to the rotatable plane i and $\boldsymbol{n}$ is the vector normal to the reference plane. Positive and negative angles indicate displacement of the rotatable plane above and below the reference plane respectively. We define a three-dimensional reaction coordinate (ξ) as the combination of these pucker coordinates, i.e., $\xi = (\theta_0, \theta_1, \theta_2)$. In the case of glucose the reference plane is made up of ring atoms C2, C4, and O5, while the rotatable planes are O5−C1−C2 (θ_0), C2−C3−C4 (θ_1), and C4−C5−O5 (θ_2). When plotting ξ the canonical conformers produce a spheroidal volume where the 2 chair conformations ($^4C_1$ and $^1C_4$) are located at the poles, the 6 boats and 6 skew-boats are at the equator, and the 12 half-chairs and 12 envelopes are at the "tropics" (Figure 1a).

We use a generalized free energy approach termed Free Energies from Adaptive Reaction Coordinate Forces (*FEARCF*). This method has been shown to be versatile with preliminary versions having, for example, been applied to glycosidic linkage rotation,[10] ion pairing,[11,12] and chemical reactions.[13] Here we calculate the potential of mean force (PMF) of pyranose ring pucker using the Hill−Reilly definitions as introduced earlier.[14] We follow a similar procedure as we had done before for a *FEARCF* produced three-dimensional ring pucker PMF of the β-D-glucose pyranose ring.[15]

This method delivers comprehensive sampling where, on analysis of the final free energy volume, it was shown that the sampling ratio of the most probable (lowest energy $^4C_1$) to the most improbable (highest energy planar ring marked by a black spot in the center of the sphere in Figure 1a) conformer is 1:1.17.[15] Using the SCCDFTB-D level of theory[16,17] as implemented in CHARMM[18] we calculated the ring pucker free energy volumes as a function of the reaction coordinate ξ for glucose with our *FEARCF* module.[19] We have compared several semiempirical methods to DFT RB3LYP/6-311++G(d,p) optimized and single-point calculations. We found that SCCDFTB-D is the most reliable for prediction of pyranose conformations.[20] We performed this free

 10.1021/ja103766w

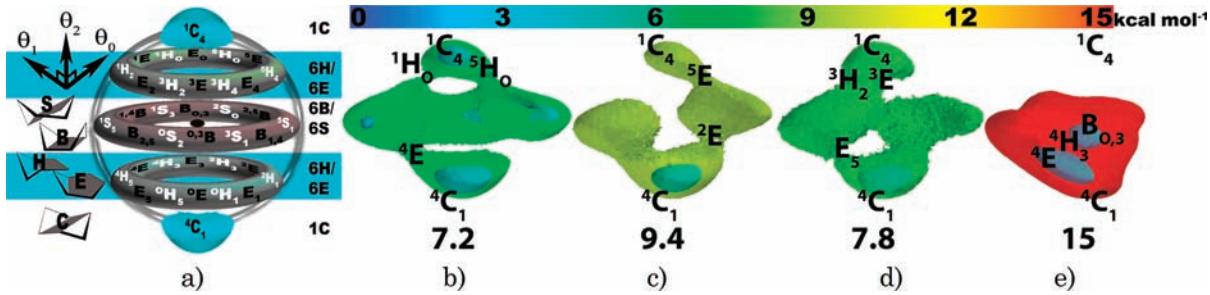

Figure 1. (a) Pucker definition shown on a reaction coordinate sphere where the planar conformer is at the center (θ_0, θ_1, θ_2) = (0,0,0). The ring pucker free energy volumes for (b) the glucose monomer in vacuum and cellooctaose in (c) vacuum, (d) water, and (e) CBHI.

energy analysis for a glucose monomer in vacuum and an internal glucose of a short cellulose strand comprising eight glucose monomers (cellooctaose) in vacuum, in water, and at the catalytic domain (CD) of the CBHI from *T. reesei*. The free energy volumes for these systems are shown in Figure 1b, c, d, and e respectively.

At room temperature (298.15 K), pucker conformations separated by low energy barriers (3−5*kT*) can be interchangeably accessed. The relationship between these conformers can be observed from the free energy iso-surface at 3 kcal/mol (the inner iso energy contours shown in Figure 1b−e). It is immediately obvious for all but one of the pucker volumes (the one determined for glucose in the CD of CBHI) that the $^4C_1$ chair and conformers topologically closely related to it are the most favored. This has been known for some time for glucose in water. Boat and skew-boat conformers along with the less favored $^1C_4$ can be found for glucose *in vacuo* (inner iso energy contours Figure 1b). However, when a glucose monomer is glycosidically linked to other monosaccharides within an oligosaccharide, it appears unable to pucker out of the preferred $^4C_1$ chair conformer. This is the case whether the oligosaccharide is in vacuum (Figure 1c) or in water (Figure 1d). To pucker the glucose ring from the favored chair form to the less favored $^1C_4$ requires forces considerably greater than those to be had from the thermal energy available at room temperature. Therefore, the minimum free energy connecting the $^4C_1$ and $^1C_4$ chair conformers is observed at much higher energy levels (outer contours in Figure 1). For glucose in vacuum (Figure 1b) this is at 7.2 kcal/mol, while cellooctaose in vacuum (Figure 1c) requires a free energy of 9.4 kcal/mol to flip the ring and cellooctaose in water (Figure 1d) requires a free energy of 7.8 kcal/mol.

The free energy volumes mapped out at thermally accessible pucker volumes for cellooctaose bound to CBHI reveals that the sugar ring is puckered out of the favored chair conformer and is restricted to a region of pucker space where half-chair and envelope conformers are preferred. Even when contouring at a higher energy level, i.e. 15 kcal/mol (red), the sugar ring at the −1 position of the active site is unable to access the vast majority of the 38 canonical conformers defined by IUPAC.

The glucose monomer is able to traverse several pathways between the $^4C_1$ and $^1C_4$ conformers (Figure 1b) where for example $^4C_1 \rightarrow {}^4H_5/{}^4E \rightarrow {}^{1,4}B \rightarrow {}^1S_3 \rightarrow {}^1H_O \rightarrow {}^1C_4$ and $^4C_1 \rightarrow {}^4H_5/{}^4E \rightarrow {}^{1,4}B \rightarrow {}^1S_3 \rightarrow B_{O,3} \rightarrow {}^2S_O \rightarrow {}^5H_O \rightarrow {}^1C_4$ are two of the lowest energy pathways that have a similar energy profile but follow different conformational paths. We are not elaborating on this here as the ring puckering of pyranose and furanose monomers are described in detail elsewhere.[20] When exploring the pucker free energy volume of the cellooctaose glucose in vacuum, it is immediately apparent that the path linking the $^4C_1$ and $^1C_4$ conformers ($^4C_1 \rightarrow {}^2E \rightarrow {}^2S_O/{}^{2,5}B \rightarrow {}^5E \rightarrow {}^1C_4$) becomes much more restricted (Figure 1c) compared with that of the single glucose monomer (Figure 1b). The glycosidic bonds to neighboring sugar monomers tether the puckering glucose. This limits the ring flexibility compared with the monomer case. Although the $^3H_2$ conformation is not on the minimum energy path for the cellooctaose vacuum PMF volume, it is an interesting conformer since significant strain on the glycosidic bond (1.56 Å) is observed. The electron density (ρ) in the bond decreases from 0.2540 au in the $^4C_1$ conformer to 0.1922 au in the $^3H_2$ half-chair form.

When that same cellooctaose strand is hydrated the resulting glucose ring pucker free energy volume is very different (Figure 1d), as is the minimum energy path which is now $^4C_1 \rightarrow E_5 \rightarrow {}^OS_2 \rightarrow B_{2,5} \rightarrow {}^3E/{}^3H_2 \rightarrow {}^1C_4$. Of the conformers in the minimum energy pathway just indicated, the $B_{2,5}$ conformer shows the most strained glycosidic bond (1.54 Å) yet the electron density (ρ) changes only slightly (0.2104 au) compared with the equilibrium $^4C_1$ case.

In the presence of the CBHI enzyme, cleavage of the glycosidic bond was observed and it was surmised that the glucose would assume a high energy boat conformation.[5] In general for glycosidases an oxocarbenium ion forms in the Michaelis complex that is accompanied by the distortion of the saccharide ring. The nature of the distortion is complex and poorly understood. This is mainly because enzyme substrates often contain saccharide rings, which exhibit several forms of puckering in the transition state. Using X-ray crystallography Davies et al. commented that there is good evidence for enzymes to utilize four possible transition-state conformations $^4H_3$, $^3H_4$, $^{2,5}B$, and $B_{2,5}$.[21] The puckering conformation of saccharides in the transition state of various glycosidases results in selective reaction pathways and specific reaction products. In the case of pyranoses competing ring puckers have been observed for several members of the glycoside hydrolase family.[22]

Our complete exploration of the puckering free energy landscape of the glucose at catalytic site −1 of the CBHI enzyme reveals that there is a dramatic restriction of pyranose ring flexibility (Figure 1e). In the enzyme both pyranose chair conformations are highly unfavorable. The $^4C_1$ conformer, that was previously the most favored in all other environments, is now 5.97 kcal/mol higher in energy than the most favored conformation (^4E) on the CBHI pucker free energy volume (Table 1). The ring puckering conformational space for the glucose at the CBHI catalytic domain is very restricted with most puckers being unachievable. The simplest path through the free energy pucker volume that traverses from the $^4C_1$ conformer (5.97 kcal/mol) to a stable boat conformer (1.56 kcal/mol) is $^4C_1 \rightarrow {}^4E \rightarrow {}^4H_3 \rightarrow B_{O,3}$ where ^4E is the most stable pucker conformation overall. When the glucose in the oligosaccharide chain is puckered in the CD an envelope conformation appears that is 11.96 and 9.42 kcal/mol lower in energy compared with the $^4C_1$ conformer in the vacuum and aqueous environments respectively. The solvent stabilizes the ^4E conformer by 2.5 kcal/mol more than the vacuum case. However, the protein makes ^4E the most preferred pyranose ring pucker, stabilizing it by more than 15 kcal/mol compared with the hydrated oligosaccharide. The $^4H_3$ and ^4E conformers are closely

Table 1. Bond Lengths (Å) and Corresponding Electron Densities (ρ in au) in the Breaking Glycosidic Bond Observed for an Internal Glucose of an Octaose *in Vacuo* (v), in Water (w), and at the −1 Position of the CBHI, i.e., the CBHI Catalytic Site (e)[a]

	Cellooctaose (v)			Cellooctaose (w)			Cellooctaose (e)		
Conformer	C1---O4	ρ	ΔG	C1---O4	ρ	ΔG	C1---O4	ρ	ΔG
$^4C_1$	1.45	0.2540	0	1.40	0.2480	0	1.48	0.2213	5.97
^2E	1.43	0.2808	9.6	1.42	0.2619	7.2	—	—	—
^2S_O	1.44	0.2527	6.5	1.51	0.2145	3.1	—	—	—
$^{2,5}B$	1.47	0.2391	6.5	1.42	0.2645	5.5	—	—	—
^5E	1.49	0.2230	8.9	1.44	0.2473	9.6	—	—	—
E_5	1.41	0.2635	8.25	1.46	0.2467	7.82	—	—	—
$^OS_2$	1.43	0.2531	7.0	1.44	0.2254	3.38	—	—	—
$B_{2,5}$	1.51	0.2200	7.0	1.54	0.2104	3.91	—	—	—
^3E	1.41	0.2156	12.48	1.50	0.2183	7.3	—	—	—
$^3H_2$	1.56	0.1922	12.48	1.41	0.2621	7.3	—	—	—
^4E	1.45	0.2503	11.96	1.48	0.2332	9.42	1.48	0.2208	0
$^4H_3$	1.49	0.2328	11.16	1.44	0.2498	8.83	1.61	0.1657	3.4
$B_{O,3}$	1.50	0.2199	11.42	1.43	0.2509	6.5	1.56	0.1900	1.56
$^1C_4$	1.49	0.2241	6.34	1.46	0.2397	12.79	—	—	—

[a] Free energies are reported in kcal/mol.

Table 2. Selected Atomic Charges for Cellooctaose in Water (w) and in the CBHI Catalytic Site (e)

	Cellooctaose (w)		Cellooctaose (e)	
Atom	$^4C_1$	$^4H_3$	$^4C_1$	$^4H_3$
Ring: O5	−0.51	−0.40	−0.41	−0.26
Anomeric: C1	0.51	−0.16	0.51	0.44
Glycosidic: O4	−0.39	−0.25	−0.42	−0.55

related to each other by the movement of C3. The $^4H_3$ conformer can be accessed via ^4E by distorting the C-3 atom below the ring plane from that found in the envelope conformation.

In these calculations only pucker forces were applied and no deliberate forces were used to break or form any bonds especially not those participating in the hydrolysis reaction. The difference in bond lengths and electron densities (Table 1) of the glycosidic bond (C1—O4) in the $^4C_1$ (both for water and enzyme) and $^4H_3$ conformations in the enzyme environment is therefore remarkable. CBHI stabilizes the ^4E conformer leading to a $^4H_3$ transition state conformer where the anomeric carbon takes on an sp^2 character making it susceptible to attack from the glutamate 212 (GLU212) nucleophile in the glycosylation step proposed for the CBHI catalyzed reaction. Here, the $^4H_3$ conformation is just 3.4 kcal/mol above the ^4E global minimum and 2.57 kcal/mol more stable than the normally preferred $^4C_1$ chair conformer. Compared with the chair, the C1—O4 bond lengthens from 1.40 to 1.61 Å. This is accompanied by a decrease in bond electron density from 0.2480 to 0.1657 au.

To determine the extent to which the enzyme has induced an electronic change on the $^4H_3$ conformer the Merz—Singh—Kollman (MK) scheme[23] was used to calculate the charges on the O5, C1, and O4 atoms (Table 2). When comparing the $^4H_3$ half-chair in the enzyme with that of the $^4C_1$ chair conformer in water, it is clear that the glycosidic oxygen and anomeric carbon in the former case have increased in negative and positive charge respectively.

The free energy volumes (Figure 1) show that CBHI induces conformational deformations in the pyranose ring at position −1 of the catalytic site. This, in part, results from a bend in the binding channel at the −1 position. The bent channel causes the cellulosic substrate (Figure 2a) to kink away from the $^4C_1$ topology as it passes through the enzyme.

Stabilization of the envelope and half-chair conformers occurs through hydrogen bonding to a glutamate (GLU212). The C2 ring carbon is levered via its hydroxyl hydrogen bonding with GLU212, which is located on the underside of the −1 sugar (Figure 2b). Through a dynamic "push—pull" mechanism it librates the ring between the ^4E envelope and $^4H_3$ half-chair. It does this by sustained hydrogen bonding interactions (between its oxygens, OE1 and OE2, and the hydrogen of O2, Figure 2b) which pull the C2 carbon into a plane with respect to the other pyranose ring atoms except C4 resulting in an ^4E conformation. A time series analysis (see Supporting Information p S2) indicates that only 2.8% of the distorted pyranose conformers (as seen in Figure 1e) are not stabilized by this hydrogen bond interaction. The electron sharing character of the hydrogen bond (specifically GLU212 OE1 to H of O2) is seen from the high electron densities calculated for it in the ^4E and $^4H_3$ conformers, which are 0.0456 au and 0.0329 au respectively (see Supporting Information p S2). This movement "opens up" the bottom face of the anomeric carbon assisting attack

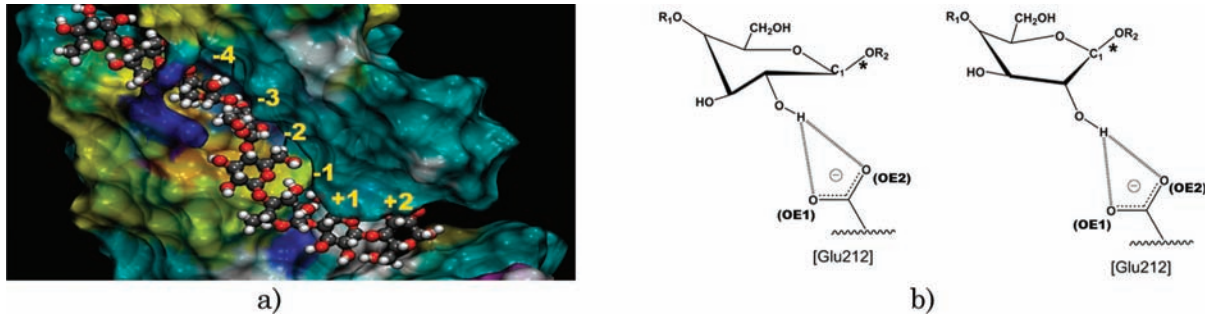

Figure 2. (a) A surface cut-away of CBHI of *T. Reesei* reveals the bent shape of the long catalytic binding tunnel. Cellooctaose is drawn as a CPK model and the catalytic −1 location along with others in CD are shown. (b) At −1 the half-chair (left) and envelope (right) conformers stabilized by the glutamate residue (GLU212) forming hydrogen bonds with the hydroxyl at the C2 ring carbon. R_1 is the −8 to −2 part of the cellulosic strand; R_2 is the +1,+2 part of the cellulosic strand where * denotes the glycosidic bond that is catalytically cleaved.

of the nucleophilic GLU212. Furthermore the distortion of the ring compromises the tetrahedrality at C1, weakening the glycosidic bond and so lowering the reaction barrier.

These glucose ring pucker free energy volumes reveal, for the first time on a natural substrate, that the way in which CBHI catalyzes the hydrolysis of the C1−O4 glycosidic bond is through the stabilization of usually strained conformers and by facilitating an electronic rearrangement in the ligand that is susceptible to nucleophilic and acid attack. We observe from these studies that the CBHI binding pocket limits the glucose at the CD to the ^4E, $^4H_3$, and $B_{O,3}$ conformations. In the presence of the enzyme there is an electronically significant change in the $^4H_3$ structure such that the glycosidic bond is elongated, weakened, and polarized in the period immediately before the formation of the transition state.

Acknowledgment. This work is based upon research supported by the South African Research Chairs Initiative (SARChI) of the Department of Science and Technology and National Research Foundation to K.J.N. C.B. and K.A.W. thank SARChI for doctoral and postdoctoral fellowship support respectively.

Supporting Information Available: Details of the calculations performed and presented here are given in p S1 along with time series and electron density data in p S2. Also complete ref 18 is provided. This material is available free of charge via the Internet at http://pubs.acs.org.

References

(1) Himmel, M. E.; Ding, S.-Y.; Johnson, D. K.; Adney, W. S.; Nimlos, M. R.; Brady, J. W.; Foust, T. D. *Science* **2007**, *315*, 804–807.
(2) Vrsanska, M.; Biely, P. *Carbohydr. Res.* **1992**, *227*, 19–27.
(3) Barr, B. K.; Hsieh, Y. L.; Ganem, B.; Wilson, D. B. *Biochemistry* **1996**, *35*, 586–592.
(4) Divne, C.; Stahlberg, J.; Reinikainen, T.; Ruohonen, L.; Pettersson, G.; Knowles, J.; Teeri, T.; Jones, T. *Science* **1994**, *265*, 524–528.
(5) Divne, C.; Stahlberg, J.; Teeri, T. T.; Jones, T. A. *J. Mol. Biol.* **1998**, *275*, 309–325.
(6) Beckham, G. T.; Matthews, J. F.; Bomble, Y. J.; Bu, L. T.; Adney, W. S.; Himmel, M. E.; Nimlos, M. R.; Crowley, M. F. *J. Phys. Chem. B* **2010**, *114*, 1447–1453.
(7) McNaught, A. D. *Pure Appl. Chem.* **1996**, *68*, 1919–2008.
(8) Cremer, D.; Pople, J. A. *J. Am. Chem. Soc.* **1975**, *96*, 1354–1358.
(9) Hill, A. D.; Reilly, P. J. *J. Chem. Inf. Model.* **2007**, *47*, 1031–1035.
(10) Naidoo, K. J.; Brady, J. W. *J. Am. Chem. Soc.* **1999**, *121*, 2244–2252.
(11) Naidoo, K. J.; Lopis, A.; Westra, A. N.; Robinson, D. J.; Koch, K. R. *J. Am. Chem. Soc.* **2003**, *125*, 13330–13331.
(12) Matthews, R. P.; Naidoo, K. J. *J. Phys. Chem. B* **2010**, *114*, 7286–7293.
(13) Rajamani, R.; Naidoo, K. J.; Gao, J. *J. Comput. Chem.* **2003**, *24*, 1775–1781.
(14) Barnett, C. B.; Naidoo, K. J. *J. Phys. Chem. B* **2008**, *112*, 15450–15459.
(15) Barnett, C. B.; Naidoo, K. J. *Mol. Phys.* **2009**, *107*, 1243–1250.
(16) Elstner, M.; Porezag, D.; Jungnickel, G.; Elsner, J.; Haugk, M.; Frauenheim, T.; Suhai, S.; Seifert, G. *Phys. Rev. B.* **1998**, *58*, 7260–7268.
(17) Cui, Q.; Elstner, M.; Kaxiras, E.; Frauenheim, T.; Karplus, M. *J. Phys. Chem. B* **2001**, *105*, 569–585.
(18) Brooks, B. R.; et al. *J. Comput. Chem.* **2009**, *30*, 1545–1614.
(19) Strümpfer, J.; Naidoo, K. J. *J. Comput. Chem.* **2010**, *31*, 308–316.
(20) Barnett, C. B.; Naidoo, K. J. *J. Phys. Chem. B*, in review.
(21) Davies, G. J.; Ducros, V. M. A.; Varrot, A.; Zechel, D. L. *Biochem. Soc. Trans.* **2003**, *31*, 523–527.
(22) Fushinobu, S.; Mertz, B.; Hill, A. D.; Hidaka, M.; Kitaokac, M.; Reilly, P. J. *Carbohydr. Res.* **2008**, *343*, 1023–1033.
(23) Singh, U. C.; Kollman, P. A. *J. Comput. Chem.* **1984**, *5*, 129–145.

JA103766W