

# Computational mutagenesis reveals the role of active-site tyrosine in stabilising a boat conformation for the substrate: QM/MM molecular dynamics studies of wild-type and mutant xylanases

Mahmoud E. S. Soliman,<sup>a,b</sup> Giuseppe D. Ruggiero,<sup>a</sup> J. Javier Ruiz Pernía,<sup>a</sup> Ian R. Greig<sup>a</sup> and Ian H. Williams<sup>\*a</sup>

Received 22nd August 2008, Accepted 24th October 2008

First published as an Advance Article on the web 5th December 2008

DOI: 10.1039/b814695k

Molecular dynamics simulations have been performed for non-covalent complexes of phenyl  $\beta$ -xylobioside with the retaining endo- $\beta$ -1,4-xylanase from *B. circulans* (BCX) and its Tyr69Phe mutant using a hybrid QM/MM methodology. A trajectory initiated for the wild-type enzyme–substrate complex with the proximal xylose ring bound at the –1 subsite (adjacent to the scissile glycosidic bond) in the  ${}^4C_1$  chair conformation shows spontaneous transformation to the  ${}^{2,5}B$  boat conformation, and potential of mean force calculations indicate that the boat is  $\sim 30$  kJ mol $^{-1}$  lower in free energy than the chair. Analogous simulations for the mutant lacking one oxygen atom confirm the key role of Tyr69 in stabilizing the boat in preference to the  ${}^4C_1$  chair conformation, with a relative free energy difference of about 20 kJ mol $^{-1}$ , by donating a hydrogen bond to the endocyclic oxygen of the proximal xylose ring. QM/MM MD simulations for phenyl  $\beta$ -xyloside in water, with and without a propionate/propionic acid pair to mimic the catalytic glutamate/glutamic acid pair of the enzyme, show the  ${}^4C_1$  chair to be stable, although a hydrogen bond between the OH group at C2 of xylose and the propionate moiety seems to provide some stabilization for the  ${}^{2,5}B$  conformation.

## Introduction

Glycosidases<sup>1</sup> (glycoside hydrolases, GHs) are involved in various physiological as well as pathological processes and are thus important therapeutic targets.<sup>2</sup> Much attention has been focussed over the past decade on the design and synthesis of glycosidase inhibitors. Successful examples include the influenza neuraminidase inhibitors, Tamiflu and Relenza,<sup>3</sup> and the anti-diabetes  $\alpha$ -glucosidase inhibitors, Acarbose and Miglitol.<sup>4,5</sup> Other targets under developments are anti-cancer inhibitors for mannosidase<sup>6</sup> and for nucleoside hydrolases/phosphorylases.<sup>7</sup>

The family G/11 endo-1,4- $\beta$ -xylanase (BCX) from *Bacillus circulans* catalyses the hydrolysis of xylan<sup>8</sup> by means of a double-displacement mechanism (Fig. 1), via oxacarbenium-like transition states, with net retention of configuration at the anomeric carbon.<sup>9</sup> Being well characterized experimentally and relatively small, it is an ideal subject for computational modelling. Its 3D fold, its substrate conformation, and its catalytic residues have been identified through a combination of sequence analysis, inhibition experiments and structural determination.<sup>10–15</sup> However, many questions still remain concerning details of the reaction pathway and catalytic mechanism. In this paper we attempt to answer the following question: why does the proximal sugar-ring of the substrate distort from  ${}^4C_1$  chair to  ${}^{2,5}B$  boat conformation?

Substrate distortion appears to be a general feature of  $\beta$ -glycoside hydrolases,<sup>16–19</sup> as seen by X-ray crystallography of enzyme–inhibitor complexes or of complexes with mutated enzymes, but less easily observed with reactive substrates, for

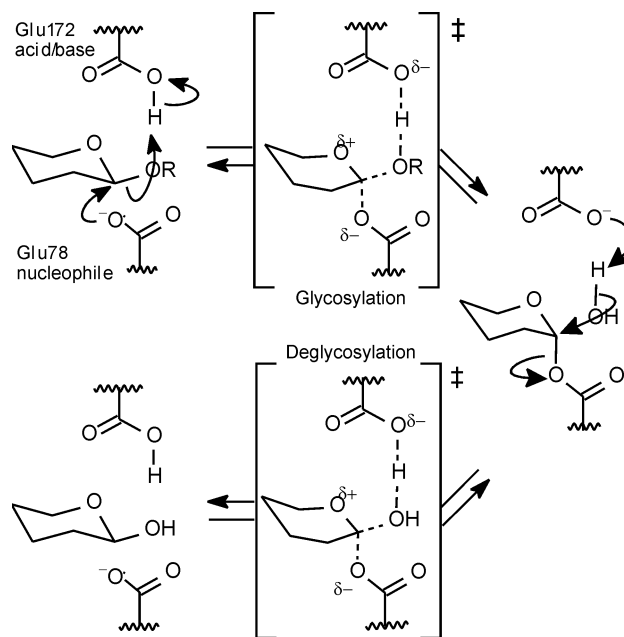


Fig. 1 Mechanism of retaining endo-1,4- $\beta$ -xylanase: catalytic residues are Glu78 and Glu172. (Sugar-ring distortion not shown.)

which computational modeling provides a powerful investigative tool. Many modelling studies have confirmed that substrate ring distortion is a common feature among glycosidases.<sup>16–19</sup> Molecular dynamics (MD) simulations have shown that the boat conformation at the –1 subsite is critical in the mechanism of family 18 chitinases,<sup>16</sup> and other studies have demonstrated that the –1 sugar moiety in cellulase Ce16A from *Trichoderma reesei* adopts a skew-boat conformation.<sup>17</sup> Similarly, modelling

<sup>a</sup>Department of Chemistry, University of Bath, Bath, UK BA2 7AY. E-mail: i.h.williams@bath.ac.uk; Fax: 44 1225 386231; Tel: 44 1225 386625

<sup>b</sup>Department of Pharmaceutical Organic Chemistry, Faculty of Pharmacy, Zagazig University, Zagazig, Egypt

studies of  $\beta$ -galactosidases have provided evidence of substrate distortion.<sup>18,19</sup> However, most studies have relied exclusively on classical approximations to describe the interatomic interactions: electronic factors, which could be crucial for substrate conformational analysis and subsequently for the mechanistic behaviour, have been neglected. To model accurately reactions occurring in enzymes (and other condensed phase systems) it is necessary to consider a very large number of interactions which may influence the mechanism. It is not feasible to use quantum mechanical (QM) methods to treat the dynamics of entire mega-dimensional systems, whilst molecular mechanics (MM) methods cannot describe the electronic polarisability that is probably a key component of substrate distortion, as exemplified by the studies of Rovira and co-workers on glucanase<sup>20</sup> and glucose.<sup>21</sup> In this paper, we apply a hybrid quantum mechanics/molecular mechanics (QM/MM)<sup>22</sup> approach in conjunction with MD simulations.

## Methods

### Isolated substrates

Unconstrained geometry optimization was performed for the substrate, starting from each of the two conformations (<sup>2</sup><sub>5</sub>B and <sup>4</sup><sub>C</sub>), using the B3LYP density functional method<sup>23</sup> and 6-31 + G(d) basis set, as implemented in the Gaussian03 package.<sup>24</sup>

### Enzyme–substrate initial structure

The X-ray crystallographic coordinates (Fig. 2a) of the covalent enzyme–inhibitor complex of wild-type BCX (PDB accession code 1BVV)<sup>25</sup> were modified as follows:

(1)  $pK_a$  values for amino acid residues at the pH optimum 5.7 were estimated using the H<sup>++</sup> programme.<sup>26</sup>

(2) Hydrogen atoms were added in accordance with the predicted protonation state.

(3) The whole system (16475 atoms) was energy minimized to a residual gradient of less than 0.001 kJ mol<sup>-1</sup> Å<sup>-1</sup> with a QM/MM method using the DYNAMO package.<sup>27</sup> The QM region (68 atoms) comprised the covalently-attached substrate—a disaccharide of xylose (XYL) and 2-deoxyfluoroxylucose (DFX)—and the two catalytic residues, Glu78 and Glu172, and was described by AM1 semi-empirical Hamiltonian.<sup>28</sup> The MM region

contained the rest of the enzyme, and was described by the OPLS-AA potential.<sup>29</sup> AM1/OPLS is commonly used in QM/MM studies of enzymic systems; although other semiempirical QM methods might give better results for small molecules *in vacuo*, their use in conjunction with incompatible QM/MM non-bonded interaction parameters could detrimentally affect their performance for condensed-phase applications. QM link atoms were placed along the C<sub>α</sub>C<sub>β</sub> bonds of Glu78 and Glu172 (Fig. 2c). At this stage, no additional water molecules were present beyond the crystallographic waters.

(4) A gas-phase QM/MM MD simulation (NVT, 300 K, 5 ps) was performed to pre-equilibrate the protein and to allow the substrate to accommodate in the binding cavity.

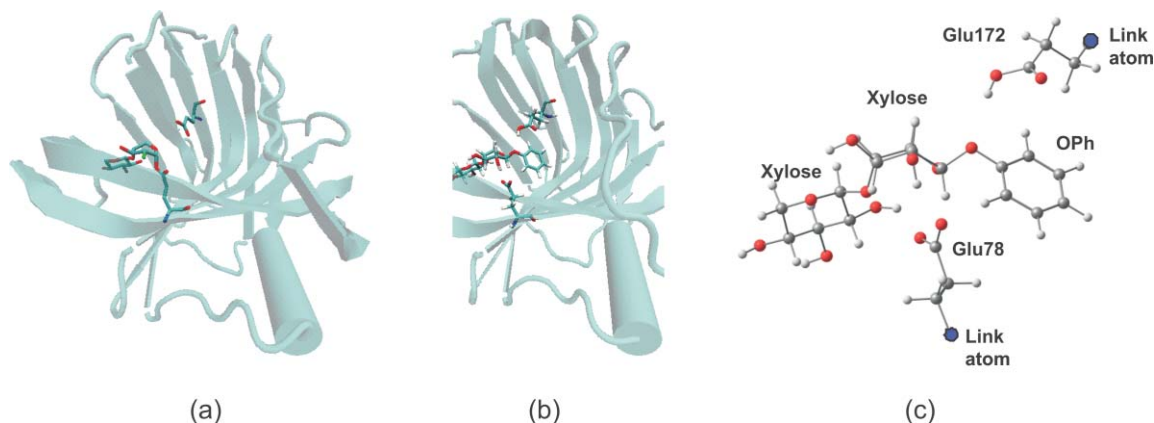
(5) The F atom at position 2 of the proximal sugar moiety of the covalently-attached substrate was changed to OH to simulate the natural substrate, xylose; the system was QM/MM energy-minimized and equilibrated by MD once again.

(6) The covalent bond between the anomeric carbon of the ( $\alpha$ -anomer of the) attached substrate and the oxygen of Glu78 was broken, and a phenoxy (OPh) leaving group was inserted manually to form the  $\beta$ -anomer. This change required care to retain the integrity of the enzyme configuration prior to the whole system being freely optimized again and equilibrated by MD, still without solvating waters.

(7) The whole system was enveloped in a cubic box of TIP3P water<sup>30</sup> of side-length 55.5 Å. First, all water molecules were relaxed with a gradient minimizer, while keeping the protein structure frozen; then the system was subjected to a short MD pre-equilibration using mild constraining forces to maintain the desired interactions between the substrate and the catalytic residues. Next, the whole system was equilibrated by MD for 20 ps at 150 K, using the NVT ensemble, still with the frozen protein. Finally, the whole system was freely minimized without any constraints and subsequently equilibrated for 20 ps at 300 K.

### Sugar ring distortion

In order to understand the nature of ring distortion at the -1 subsite within the environment of the enzyme active site, it is necessary also to explore the conformational behavior of the substrate in solution. To this end, we consider not only the final



**Fig. 2** (a) X-Ray structure for wild-type BCX. (b) Final modelled structure after energy minimization and equilibration; water molecules are not shown for clarity. (c) QM atoms selected in the simulation.

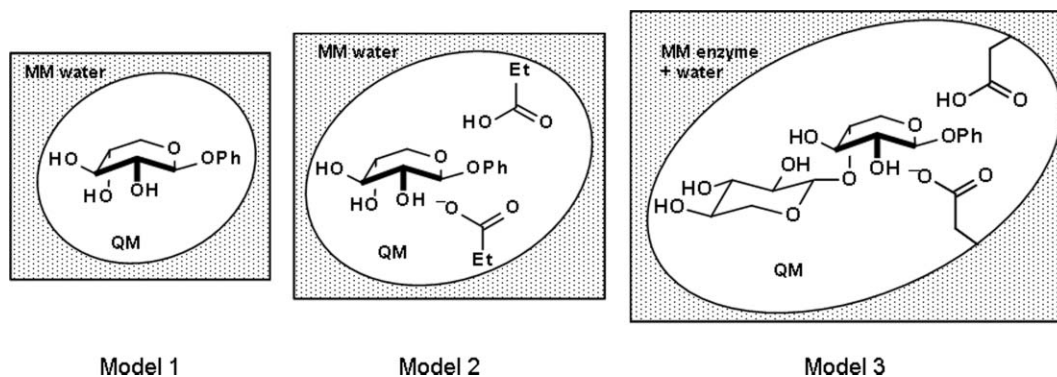


Fig. 3 Schematic representations for models 1, 2 and 3.

full system, as described in step 7 above, as model 3, but also two simpler models (Fig. 3): model 1 includes in the QM region only the proximal sugar ring of substrate in MM water, whereas model 2 additionally includes propionic acid and propionate groups in the QM region, to represent the functionality of Glu78 and Glu172; models 1 and 2 both involve cubic boxes of TIP3P water with side-length 31.4 Å. Subscripts B and C below denote the  $^{2,5}B$  boat and  $^4C_1$  chair conformations, respectively. The QM region of models 1 and 2 does not involve any link atoms.

Models 1 and 2 were allowed to equilibrate for 10 ps, and model 3 for 20 ps (same as above) of MD at 300 K in the NVT ensemble, prior to 30 ps of production MD for each model (1 fs timestep, switched cutoff radius of 16 Å applied to all interactions). No constraints were applied at any stage of the dynamic runs. Atomic charges were monitored along the dynamic simulation. A free-energy pathway for interconversion of  $^{2,5}B$  and  $^4C_1$  conformers was generated using the potential of mean force (PMF) approach and the weighted histogram analysis method (WHAM) as implemented in DYNAMO.<sup>27</sup> Before running PMF calculations, it was necessary to determine the best internal coordinate to describe this conformational change. Ideally the conformations of hexapyranose rings may be described by combinations of distances and angles, such as the puckering coordinates of Cremer and Pople<sup>31</sup> used in a recent study of glucose,<sup>21</sup> but for simplicity we use a single coordinate. Five different dihedral angles were considered and the C3–C4–C5–O5 dihedral was found to be the best for the purpose. An umbrella constraint force of 0.5 kJ mol<sup>-1</sup> Å<sup>-2</sup> was applied in each of the 31 windows; 1 ps of equilibration was performed prior to 10 ps of production MD in each window, starting with a structure perturbed in the distinguished coordinate from the final structure in the previous window. Block averaging within each window suggests that the PMF is satisfactorily converged for its purpose in this study, although this does not mean that equilibrium has been achieved with respect to all hydroxyl group conformations.

Hexapyranose ring conformations were determined from trajectories for each model and were described following the approach of Bércecs *et al.*<sup>32</sup> Briefly, a conformation was specified in terms of a polar angle  $\theta$  ( $0^\circ \leq \theta \leq 180^\circ$ ), an equatorial angle  $\phi$  ( $0^\circ \leq \phi < 360^\circ$ ), and an amplitude  $r$ . Variation in the polar angle specifies a  $^1C_4$  chair ( $\theta = 0^\circ$ ), envelope/half-chair ( $E/H$ ,  $\theta = 45^\circ$ ), boat/skew-boat ( $B/S$ ,  $\theta = 90^\circ$ ), envelope/half-chair ( $E/H$ ,  $\theta = 135^\circ$ ), and  $^4C_1$  chair ( $\theta = 180^\circ$ ) transformation; whereas variation in the equatorial angle specifies a boat/skew-boat (for  $\phi = 90^\circ$ )

or envelope/half-chair (for  $\phi = 45^\circ$  or  $135^\circ$ ) pseudorotational itinerary. Angles  $\theta$  and  $\phi$  are sufficient to specify which of the 38 canonical conformations a molecule is closest to, while the conformational amplitude  $r$  specifies the extent of distortion from the planar cyclohexane structure. A diagrammatic representation of hexopyranose ring conformations in terms of  $\theta$  and  $\phi$  only is shown in Fig. 4.

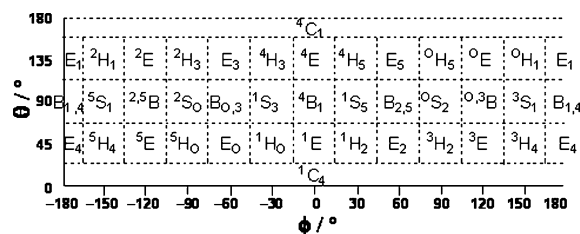


Fig. 4 Division of hexopyranose conformational space into regions corresponding to canonical conformations.

## Results

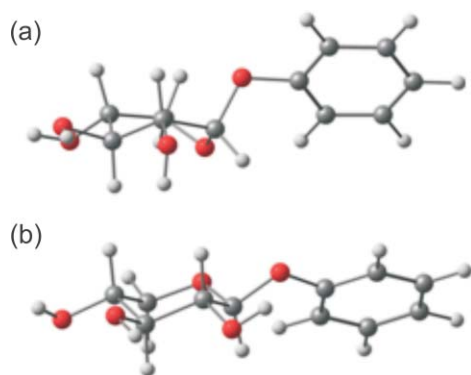
### Isolated substrates: model 1

B3LYP/6-31+G(d) geometry optimization starting from the  $^4C_1$  conformer of the phenyl xyloside substrate in vacuum confirms the chair as a stable minimum-energy species. However, a similar optimization starting from the  $^{2,5}B$  conformer leads to the  $^1S_3$  skew conformation (Fig. 5) as a local minimum, demonstrating that the boat is not a stable conformer in the gas phase.

The QM/MM MD trajectory (Fig. 6a) initiated from the  $^{2,5}B$  conformation of phenyl  $\beta$ -xyloside in water follows an itinerary through  $B_{3,0}$  to  $^2S_0$  and  $^{2,5}B$  over the course of 30 ps, as shown by the plot of  $\phi$  vs. time; these conformations are clustered in the region bounded by  $70^\circ < \theta < 110^\circ$  and  $-120^\circ < \phi < -45^\circ$  (Fig. 6b). The corresponding trajectory initiated from the  $^4C_1$  conformation in water remains stable: the apparent fluctuations in Fig. 6c and the spread of points in Fig. 6d reflect the ill-defined character of the angle  $\phi$  when  $\theta \approx 180^\circ$ .

### Effect of catalytic residues: model 2

The effects of the acid Glu172 and nucleophilic Glu78 residues may be mimicked simply by propionic acid and propionate moieties, respectively. The  $^{2,5}B$  conformation with the additional groups



**Fig. 5** B3LYP/6-31+G(d) optimized structures starting from (a)  $^{2.5}B$  and (b)  $^4C_1$  conformations, respectively.

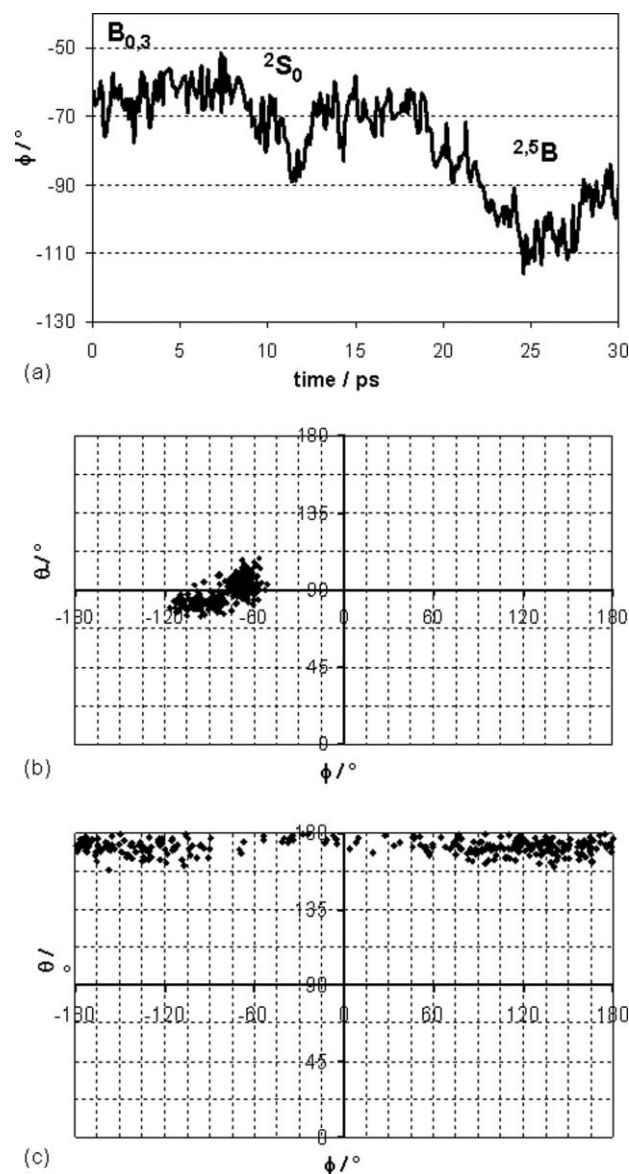
located in positions taken from the structure of the enzyme–substrate complex optimizes to a local  $^{2.5}B$  boat minimum in both gas-phase B3LYP/6-31+G(d) and QM/MM aqueous-phase calculations.

The QM/MM MD trajectory initiated from the  $^4C_1$  conformation in water remains stable and behaves very similarly to that for model 1. The corresponding trajectory initiated from the  $^{2.5}B$  boat in water fluctuates between this conformation and the  $^2S_0$  skew-boat during 30 ps, as shown by the plot of  $\phi$  vs. time (Fig. 7a); these conformations are clustered in the region bounded by  $75^\circ < \theta < 105^\circ$  and  $-135^\circ < \phi < -90^\circ$  (Fig. 7b). In comparison with model 1, the observed relative stability of  $^{2.5}B$  boat may refer to the stabilization by hydrogen bond between the OH group at C2 and the nucleophilic propionate residue (Fig. 8).

### Effect of protein environment: model 3

The QM/MM MD trajectory (Fig. 9a and 9b) initiated from the  $^{2.5}B$  conformer shows this to be a stable species which does not evolve towards a skew conformation. The C2–C1–O5–C5 dihedral angle of the  $^{2.5}B$ -substrate in the enzyme achieves a better coplanarity ( $0$ – $30^\circ$ ) than in model 2.

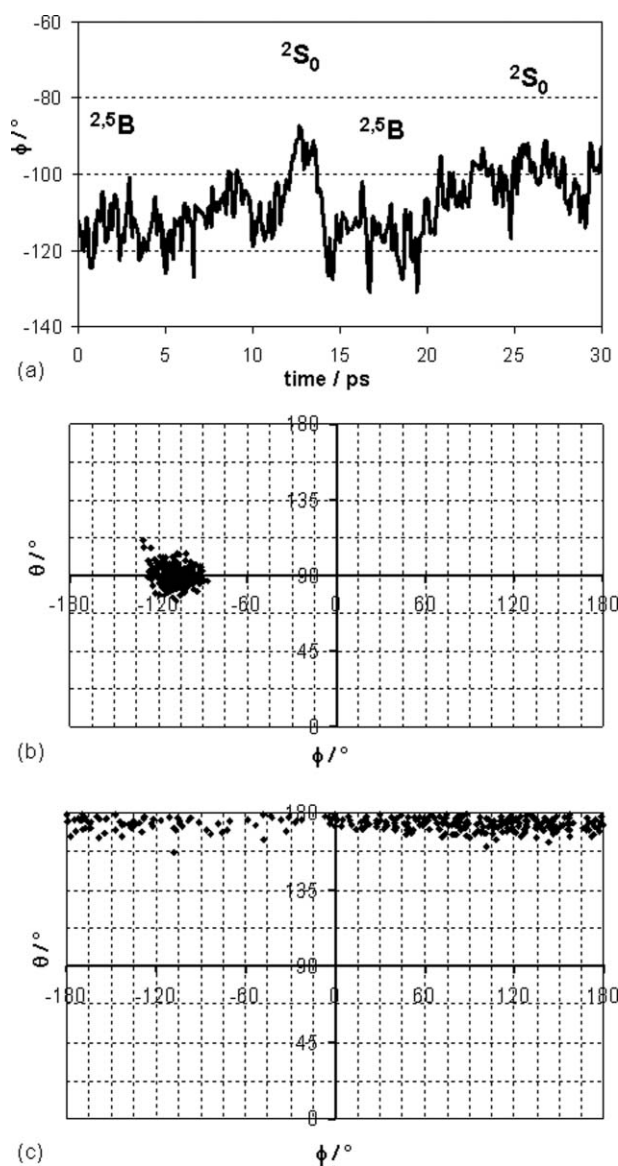
Interestingly, after initial energy minimization, the  $^4C_1$  chair conformation of the proximal sugar ring within the enzyme active was not found as a local minimum but converted to an  $E_1$  envelope and then transformed quickly to  $^{2.5}B$  (via  $^2H_1$  and  $^2E$ ) within 5 ps of MD simulation. (Note that the apparent swings in the value of  $\phi$  between  $+180^\circ$  and  $-180^\circ$  in Fig. 9c are illusory since, as Fig. 9d shows, these points all lie within the same region of conformational space corresponding to  $E_1$ .) The abrupt change in the conformational behavior from  $E_1$  to  $^{2.5}B$  does require an explanation, which we sought by consideration of four nearby residues that interact directly with the xylose ring at the  $-1$  subsite, namely Glu172 and Glu78, Arg112 and Tyr69 (Fig. 10). The acidic group Glu172 donates a hydrogen bond to the glycosidic O1 atom of the substrate, whereas the nucleophilic group Glu78 accepts a hydrogen bond from the OH group at C2 of the sugar ring, and Arg112 forms two hydrogen bonds with the OH group at C3 of the substrate. We have monitored these hydrogen-bond distances along the 30 ps MD trajectories for the chair and the boat conformations of the proximal sugar ring of the substrate, but have found essentially no difference between the two conformers. On the other hand, Tyr69 may form a hydrogen bond with either the endocyclic ring oxygen O5 or with Glu78 (Fig. 11a): when



**Fig. 6** QM/MM MD simulation for Model 1 in water: (a) plot of composite angle  $\phi$  vs. time and (b) conformational space sampled during the QM MD trajectory for xylose ring conformation starting from  $^{2.5}B$  conformation; (c) conformational space sampled starting from  $^4C_1$  conformation.

one hydrogen-bond distance is short, the other is long (Fig. 11b, red and purple lines). In the QM/MM MD simulation started from the  $^4C_1$  chair, the H1...O5 distance between Tyr69 and the substrate is initially  $>3.5 \text{ \AA}$  and distinct from the short distance for the  $^{2.5}B$  boat (Fig. 11c), but after 5 ps the two trajectories behave in the same way (Fig. 11b, blue line).

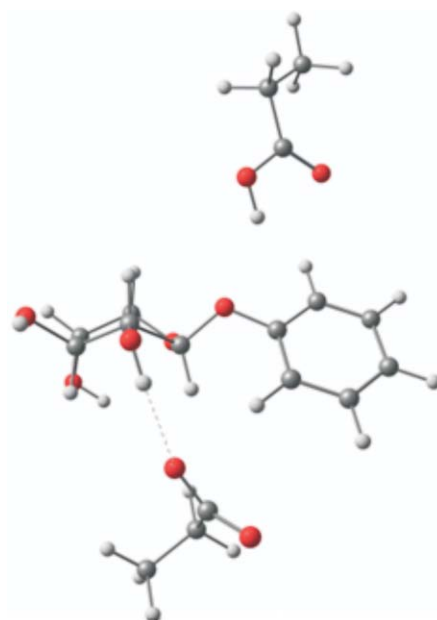
The remarkable change in the hydrogen bond between Tyr69 and the proximal sugar ring suggests that this interaction is required to assist the conformational rearrangement from the  $^4C_1$  chair to the  $^{2.5}B$  boat in the enzyme active site. To test this hypothesis we made the Y69F mutant, replacing tyrosine by phenylalanine, in order to eliminate this hydrogen bonding interaction with the substrate. Using the same computational procedure as before, QM/MM optimization showed the  $^4C_1$



**Fig. 7** QM/MM MD simulation for Model 2: (a) plot of composite angle  $\phi$  vs. time and (b) conformational space sampled during the QM MD trajectory for xylose ring conformation starting from  $^{2.5}B$  conformation; (c) conformational space sampled starting from  $^4C_1$  conformation.

conformation to be a local minimum, and subsequent MD simulation confirmed that it remained as a stable species during the 30 ps trajectory (Fig. 12). This strongly indicates that Tyr plays a vital role in stabilizing the  $^{2.5}B$  conformer.

To quantify the extent of this stabilization, we computed free energy profiles (Fig. 13) for the interconversion of the  $^{2.5}B$  boat and the  $^4C_1$  chair for both the wild-type enzyme and Y69F mutant enzyme–substrate complexes, using the C3–C4–C5–O5 dihedral angle as the reaction coordinate. The (Helmholtz) free changes along this coordinate are evaluated relative to the well-defined boat conformer with C3–C4–C5–O5  $\approx 45^\circ$  for both profiles. Consequently, the energetic stabilization of the boat conformer in the wild-type due to the hydrogen bond between the substrate and Tyr69 appears in Fig. 13 as an apparent stabilization of the chair conformer in the mutant by  $\approx 20$  kJ mol $^{-1}$ . Note that, for the wild-type, the dihedral angle C3–C4–C5–O5  $\approx -60^\circ$  does not



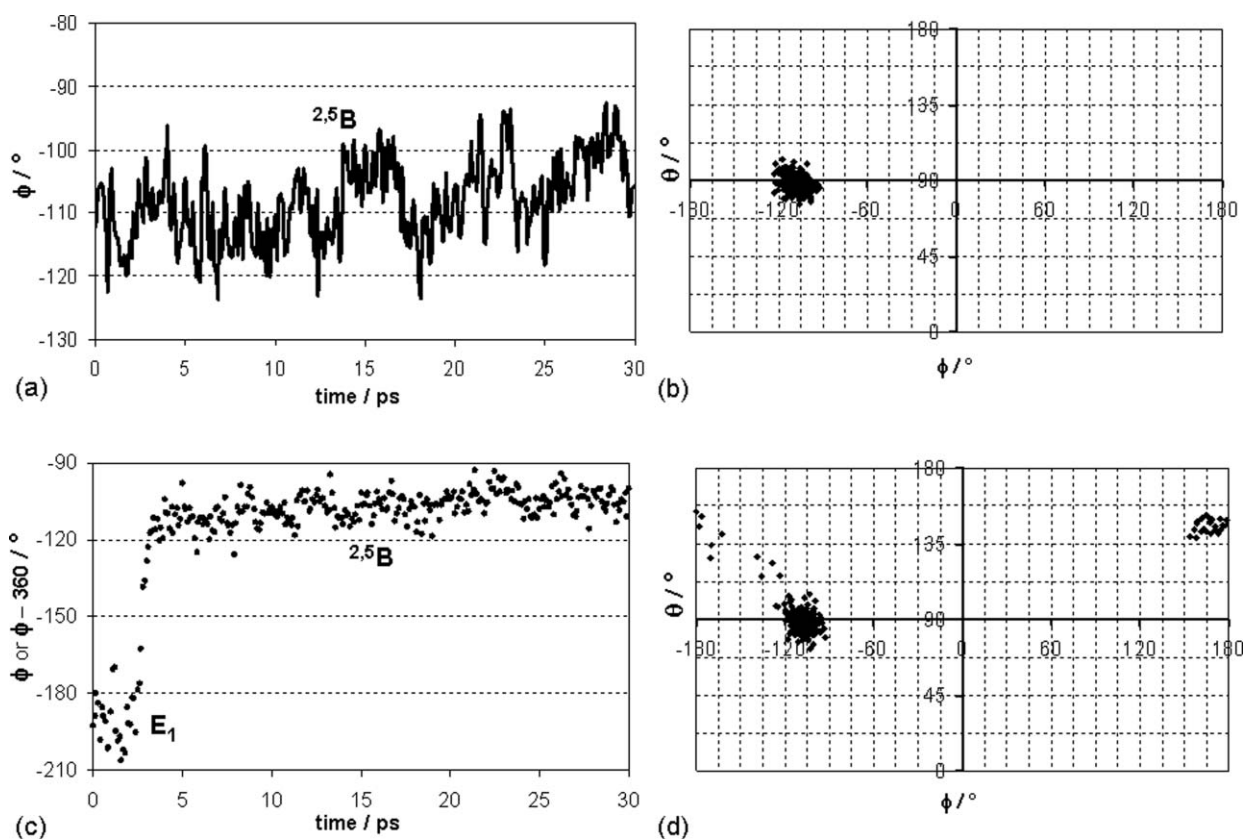
**Fig. 8** Hydrogen bond stabilization between the ring OH at C2 and propionate.

correspond purely to a  $^4C_1$  chair: the population of conformers sampled in the simulation for values of the reaction coordinate in this range includes also envelope and half-chair conformations. It is therefore probably safer to estimate the degree of stabilization by comparison between the local energy minima at  $\approx -60^\circ$  for the mutant and  $\approx -40^\circ$  for the wild-type. Clearly, the barrier for conversion of the  $^4C_1$  chair to the more-stable  $^{2.5}B$  boat in the wild-type enzyme–substrate complex is significantly lower than it is for the Y69F mutant.

The sequence of conformations populated by the wild-type BCX–substrate complex as it is driven along the C3–C4–C5–O5 dihedral angle reaction coordinate in the PMF calculations is presented in Fig. 14. The differently coloured clusters of points represent a selection of the overlapping windows considered during umbrella sampling; for the sake of clarity, not all the windows are plotted in this diagram. Starting from the  $^{2.5}B$  boat, the sequence of MD trajectories sample conformations including  $^2S_0$ ,  $B_{0,3}$ ,  $E_3$ ,  $^2H_3$ ,  $^2E$ ,  $^2H_1$ ,  $E_1$ ,  $^{\circ}H_1$ ,  $^{\circ}E$ ,  $^{\circ}H_5$  and  $E_3$  on the way towards  $^4C_1$ . Note that this sequence of clusters does not define a dynamical trajectory since it is driven by the dihedral angle constraint and is not a function of time. Also note that the “polar”  $^4C_1$  region is distorted by the Mercator projection in regard to area and distance relative to the other sugar-ring conformations.

### Development of oxocarbenium-ion character

Mulliken atomic charges for selected atoms (Table 1) averaged over the 30 ps MD trajectory for each of the three models suggest a gradual increase in the oxocarbenium-ion character in the proximal xylose ring as the substrate interacts with the carboxyl/carboxylate pair in solution and within the wild-type enzyme active site. This is shown by increases in positive charge on C1 and H1 and a decrease of negative charge on O5; these changes are accompanied by increased polarization of the C1–O1 bond as shown by an increase of negative charge on O1. Note that these



**Fig. 9** QM/MM MD simulation for Model 3: (a) plot of composite angle  $\phi$  vs. time and (b) conformational space sampled during the QM MD trajectory for xylose ring conformation starting from  ${}^{2.5}B$  conformation; (c) plot of  $\phi$  vs. time and (d) conformational space sampled starting from  ${}^4C_1$  conformation.

**Table 1** Average atomic charges  $\langle q \rangle / |e|$  on selected atoms for populations of boat and chair conformers

atom	model 1		model 2		model 3		Y69F
	boat	chair	boat	chair	boat	chair	boat
C1	0.11	0.11	0.13	0.11	0.15	0.14	0.16
H1	0.15	0.13	0.18	0.14	0.23	0.21	
O5	-0.28	-0.29	-0.24	-0.29	-0.22	-0.23	-0.27
O1	-0.25	-0.23	-0.27	-0.22	-0.28	-0.27	

**Table 2** Average interatomic distances  $\langle d \rangle / \text{\AA}$  for populations of boat and chair conformers

distance	model 1		model 2		model 3		Y69F
	boat	chair	boat	chair	boat	chair	boat
C1–O5	1.422	1.429	1.419	1.427	1.411	1.412	1.418
C1–O1	1.428	1.421	1.432	1.421	1.439	1.438	1.434
$O_n^- \cdots C1$			4.204	4.311	3.041	3.052	3.089
$H_a \cdots O1$			2.479	2.751	2.619	2.629	2.524
$O_n^- \cdots O_a$			7.596	7.992	7.432	7.426	7.298

charge changes are more pronounced for the boat than for the chair conformer. Moreover, this development of oxocarbenium-ion character is demonstrated by shortening and lengthening, respectively, of the average C1–O5 and C1–O1 bonds (Table 2); curiously, these changes are larger for the chair conformer.

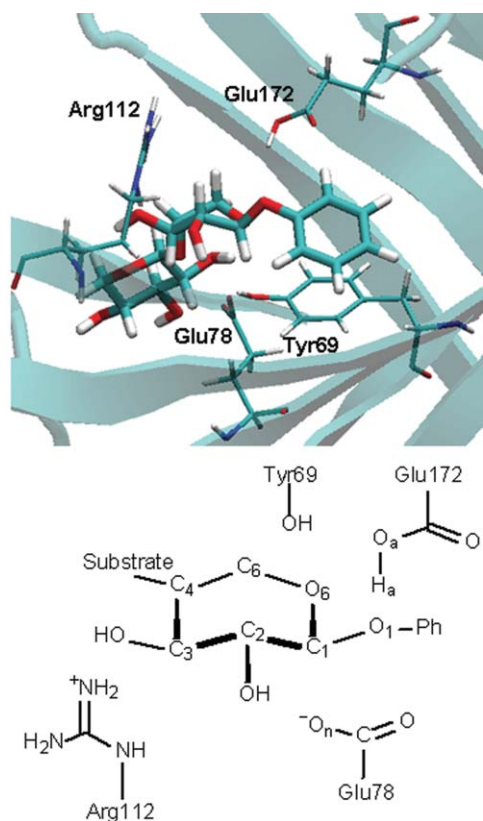
The effect of the Y69F mutation is to reduce the degree of oxocarbenium-ion character in the proximal xylose ring of

the enzyme–substrate complex, as evidenced by the charges and bond lengths shown in Tables 1 and 2. The non-bonded distances  $O_n^- \cdots C1$  and  $H_a \cdots O1$  in Table 2 indicate that greater oxocarbenium-ion character is accompanied by a shorter distance to the nucleophilic carboxylate group but a longer distance to the acidic carboxyl group.

## Discussion

The suggestion that a pyranoside substrate might preferentially adopt a  ${}^{2.5}B$  conformation during enzyme-catalysed glycoside hydrolysis was first made by Hosie and Sinnott<sup>33</sup> on the basis of kinetic isotope effects for yeast  $\alpha$ -glucosidase catalysed hydrolysis of aryl glucosides and glucosyl pyridinium ions. Their data required that breaking of the bond to the aglycone moiety of their substrates was preceded by a kinetically discrete non-covalent transformation of the initial enzyme–substrate complex, which they identified with a change to the  ${}^{2.5}B$  conformation in which the C5–O5–C1–C2 atoms are approximately co-planar. The latter arrangement would facilitate the formation of a transition state with significant oxocarbenium-ion character.

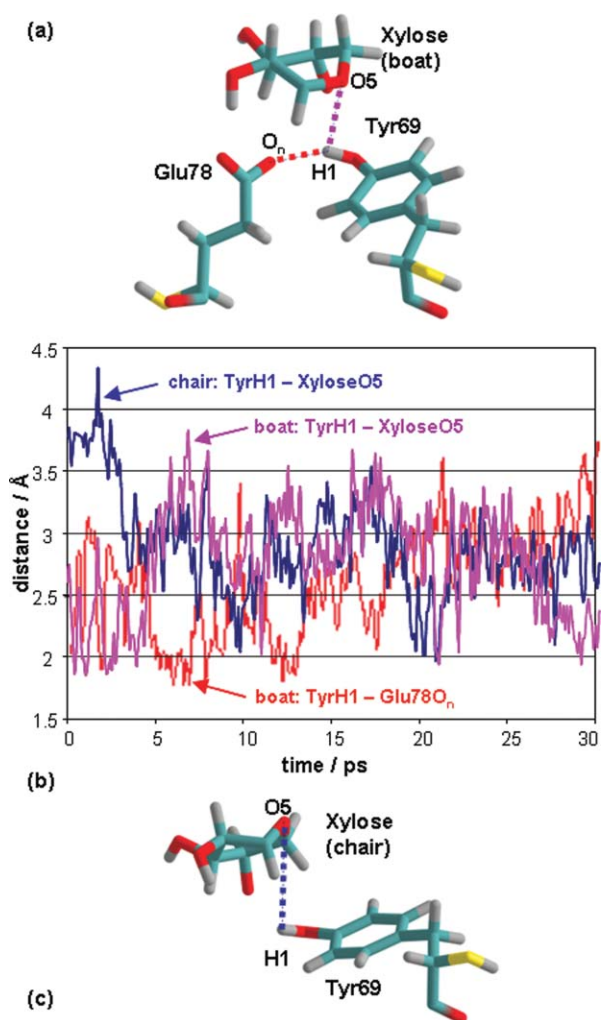
In this paper we have taken one step toward investigating the reason for, as well as the structural implications of, substrate ring distortion in  $\beta$ -1,4-xylanase, by applying hybrid QM/MM molecular dynamics and free energy calculations. How reliable are our results, based as they are upon use of the semiempirical AM1 hamiltonian within the QM/MM method? Momany and co-workers<sup>34</sup> have shown that the B3LYP/6–311++G\*\* level of



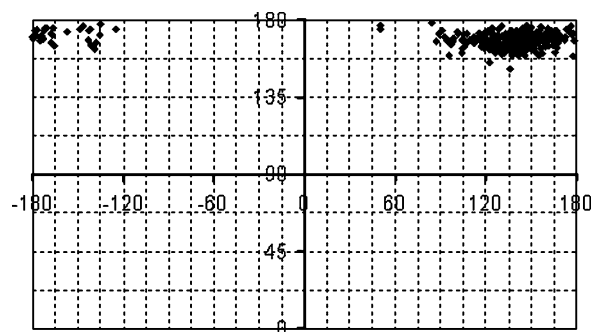
**Fig. 10** Interaction of the proximal sugar ring with the active site residues.

theory gives consistently reliable geometries, conformations, and energies for carbohydrates *in vacuo*, and have noted the importance of hydroxyl group orientations and interactions. Intramolecular interactions between these groups are less important in aqueous solution because of hydrogen bonding with solvent molecules, and in an enzyme active site, interactions with the protein environment may drastically reduce the number of significant hydroxyl group rotameric forms. In regard to the relative energies of the  ${}^4C_1$  and  ${}^{2.5}B$  conformations of the xylose ring in each of our condensed phase models 1, 2 and 3, one may well suppose that there is an underlying systematic error due to AM1 as compared to a large-basis DFT, or other high-level theoretical method, but it is not a straightforward task to estimate its magnitude. Moreover, our concern in this study is not with the energy difference between the boat and chair conformers, but rather with the influence of enzyme environment upon that energy difference. We expect that the change from tyrosine to phenylalanine (both in the MM region) will polarize higher-level QM wavefunctions in a qualitatively similar (though undoubtedly quantitatively different) fashion to AM1. We therefore consider our AM1/OPLS estimate of the free energy of stabilization of the boat conformer due to the hydrogen bond with Tyr69 in BCX to be at least qualitatively reliable. A quantitative assessment of error, as compared to an appropriate high-level QM method, would be meaningless unless it were evaluated for completely converged populations of conformations accessible under condensed-phase conditions, since there are large fluctuations between energy differences determined at arbitrary “snapshot” structures taken from the MD trajectories.

The first structural evidence for sugar-ring distortion to the boat conformer was obtained by Brayer and co-workers<sup>25</sup> who obtained

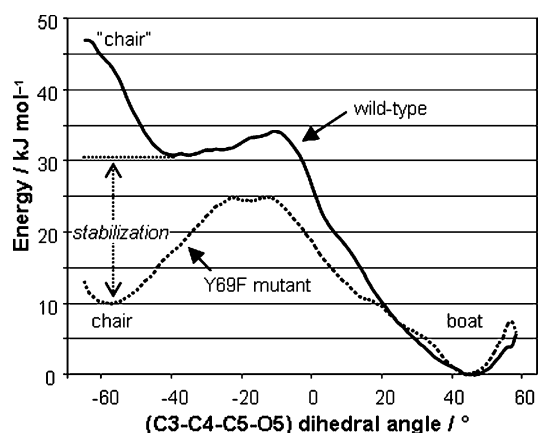


**Fig. 11** (a) Tyr69 forms a hydrogen bond either to the endocyclic ring oxygen O5 of the xylose boat conformer (purple line) or to oxygen O<sub>n</sub> of the Glu78 residue (red line) and (b) switches between the two during a QM/MM MD simulation. (c) The hydrogen bond between Tyr69 and the endocyclic ring oxygen O5 of the xylose chair conformer (blue line) is initially long but after 5 ps behaves in the same way as for the boat conformer.

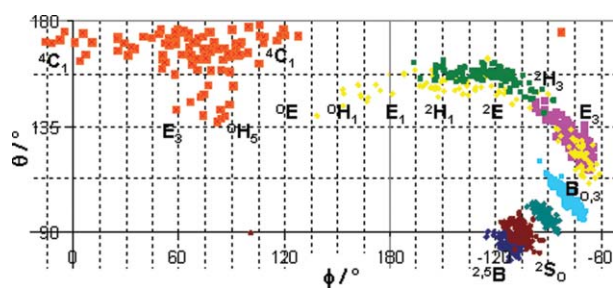


**Fig. 12** Regions of xylose ring conformational space sampled during the MD trajectories for the  ${}^4C_1$  of the Y69F mutant.

the 1.8 Å resolution structure of the covalent glycosyl-enzyme intermediate formed between BCX and a 2-deoxy-2-fluoro- $\beta$ -xylobioside. (This structure was, of course, the starting point for



**Fig. 13** Free energy profiles for interconversion of the  ${}^4C_1$  chair and the  ${}^{2.5}B$  boat conformers of the proximal xylose ring of the substrate in wild-type BCX (solid line) and Y69F mutant (dotted line).



**Fig. 14** Conformations sampled during the PMF calculation for  ${}^{2.5}B \rightarrow {}^4C_1$  interconversion with respect to the C3–C4–C5–O5 dihedral angle in the wild-type enzyme–substrate complex.

the modelling studies described in the present paper.) The proximal xylose ring was found to be heavily distorted from  ${}^4C_1$  to  ${}^{2.5}B$ , thus achieving near coplanarity of the C5–O5–C1–C2 atoms, whereas the distal xylose ring remained in the chair conformation. It was noted that an itinerary for the conformational rearrangement that were to proceed *via*  ${}^4C_1 \rightarrow {}^2H_3 \rightarrow {}^2S_0 \rightarrow {}^{2.5}B$  would involve motion (relatively) of only C5 and O5, and that these atoms do not bear substituents which might add to the energetic cost of the conformational change. These workers also obtained the X-ray crystal structure for the catalytically inactive Y69F mutant BCX, with no substrate in the active site, and reported that it has a very similar structure to that of the covalent intermediate. In particular, the position of Glu78 was almost identical, thus refuting an earlier proposal<sup>35</sup> that Tyr69 might have a crucial catalytic role in correctly positioning the nucleophile. On the other hand, it was noted<sup>25</sup> that the non-covalent enzyme–substrate complex observed by Wakarchuk *et al.*<sup>35</sup> between the catalytically incompetent E172C mutant BCX and a xylotetraose substrate has the proximal xylose ring (occupying the –1 subsite) bound in the  ${}^4C_1$  conformation.

What none of these prior experimental studies has been able to ascertain is the nature of the conformation of a xylose substrate bound non-covalently in the –1 subsite of a catalytically competent BCX, but that is precisely what we have been able to achieve by means of the computational modelling carried out in the present study. Our QM/MM model for the enzyme–substrate complex between wild-type BCX and phenyl  $\beta$ -xylobioside clearly

prefers the proximal ring to adopt the  ${}^{2.5}B$  conformation. An MD trajectory initiated from a structure with this ring in the  ${}^4C_1$  conformation spontaneously transformed into the  ${}^{2.5}B$  boat. The PMF for interconversion of these conformers along the dihedral angle C3–C4–C5–O5 coordinate suggests that the boat is lower in free energy than the chair site by  $\sim 30$  kJ mol $^{-1}$  within the active site of wild-type BCX.

We have demonstrated the key role of hydrogen bond donation from Tyr69 to O5 of the proximal xylose ring in stabilizing the boat over the chair by means of a comparative QM/MM MD study for the Y69F mutant BCX which clearly shows the  ${}^4C_1$  conformation to be a stable local energy minimum in the absence of that hydrogen bond. The PMF for interconversion of boat and chair along the dihedral angle C3–C4–C5–O5 coordinate suggests that the boat is still lower in free energy than the chair site by  $\sim 10$  kJ mol $^{-1}$  within the active site of the Y69F mutant, implying a preferential stabilization of the boat by wild-type BCX by about 20 kJ mol $^{-1}$ .

The  ${}^{2.5}B$  conformation of phenyl  $\beta$ -xyloside is not stable in the gas phase but seems to be metastable in water, being able to fluctuate within at least the  ${}^{2.5}B \rightarrow {}^2S_0 \rightarrow B_{3,0}$  section of the boat/skew-boat pseudorotational itinerary. In the presence of an added carboxylate/carboxyl pair, the boat is a local minimum in the gas-phase and in water it fluctuates between  ${}^{2.5}B$  and  ${}^2S_0$ : a hydrogen bond between the OH group at C2 of xylose and the propionate moiety seems to provide some stabilization for the  ${}^{2.5}B$  conformation. As expected, the  ${}^4C_1$  chair is stable in water both with and without the propionate/propionic acid pair.

The Y69F mutant has been reported<sup>35</sup> to have less than 0.01% of the activity of wild-type BCX at 40 °C, which would imply a free energy difference between the rate-determining transition states of  $>24$  kJ mol $^{-1}$ . It has been hypothesized<sup>25,36</sup> that the catalytic role of Tyr69 is provide a stabilizing dipolar interaction with the partial positive charge on O5 of the oxocarbenium-ion like transition state for the rate-determining glycosylation of BCX by the substrate. We will present the results of QM/MM PMF calculations for the free energies of activation for the glycosylation step in wild-type BCX and the Y69F mutant in a forthcoming paper,<sup>37</sup> and so will not speculate further on this point here. However, it is worth noting an alternative possibility that arises by analogy with the experimental<sup>38</sup> and theoretical<sup>39</sup> studies by Schramm, Schwartz and co-workers on human purine nucleoside phosphorylase (hPNP). The glycosyl transfer reaction catalysed by this enzyme has a transition state with oxocarbenium-ion character and the ribofuranoside substrate possesses a hydroxyl group C5. It is suggested that the neighbouring His257 provides a mechanical push upon O5 towards the endocyclic O4 in a compressive motion with the phosphate nucleophile such that the build-up of electron density stabilizes the oxocarbenium-like transition state and facilitates the reaction.<sup>38,39</sup> Regardless of whether this particular idea is correct, it is striking to note the importance of a suitably located hydroxyl group which exerts an effect upon the endocyclic oxygen of a glycoside by means of its oxygen rather than its proton. Of course, the xylose ring in the –1 subsite of BCX has no hydroxymethyl substituent at C5; indeed, it has been noted that there is no space around C5 to accommodate any substituent.<sup>25</sup> Instead, however, the active site of wild-type BCX presents Tyr69 in close proximity to O5. We suggest that the OH groups of Tyr69 in BCX and of the hydroxymethyl substituent of the ribofuranoside substrate of hPNP may have similar roles,



although we do not wish to speculate here on what exactly those roles may be.

## Conclusion

Molecular dynamics simulations using a hybrid QM/MM potential have demonstrated that wild-type BCX preferentially binds a phenyl  $\beta$ -xylobioside substrate in the  ${}^2,5B$  conformation at the  $-1$  subsite adjacent to the scissile glycosidic bond. This result complements and extends the earlier experimental report that the proximal xylose ring of a covalently-bound 2-deoxy-2-fluoroxxylobioside complex with BCX adopts the same  ${}^2,5B$  boat conformation. Analogous simulations for the Y69F mutant confirm the key role of Tyr69 in stabilizing the boat in preference to the  ${}^4C_1$  chair conformation with a relative free energy difference of about 20 kJ mol $^{-1}$ .

## Acknowledgements

We gratefully acknowledge the award of an Egyptian Government Scholarship to MESS and financial support from the BBSRC (BB/C006011/1), EPSRC (EP/E019455/1) and the Mizutani Foundation for Glycoscience (070034). IRG is a Leverhulme Trust Early Career Fellow.

## Notes and references

- 1 S. Hurlley, R. Service and P. Szuromi, *Science*, 2001, **291**, 2337–2337.
- 2 N. Asano, *Glycobiology*, 2003, **13**, 93r–104r.
- 3 A. Moscona, *New Engl. J. Medicine*, 2005, **353**, 1363–1373.
- 4 H. W. M. Breuer, *Int. J. Clinical Pharmacol. Therapeutics*, 2003, **41**, 421–440.
- 5 L. J. Scott and C. M. Spencer, *Drugs*, 2000, **59**, 521–549.
- 6 P. E. Goss, M. A. Baker, J. P. Carver and J. W. Dennis, *Clin. Cancer Res.*, 1995, **1**, 935–944.
- 7 G. B. Evans, R. H. Furneaux, D. H. Lenz, G. F. Painter, V. L. Schramm, V. Singh and P. C. Tyler, *J. Med. Chem.*, 2005, **48**, 4679–4689.
- 8 B. Henrissat and G. Davies, *Curr. Opin. Struct. Biol.*, 1997, **7**, 637–644.
- 9 J. D. McCarter and S. G. Withers, *Curr. Opin. Struct. Biol.*, 1994, **4**, 885–892.
- 10 J. Gebler, N. R. Gilkes, M. Claeysens, D. B. Wilson, P. Beguin, W. W. Wakarchuk, D. G. Kilburn, R. C. Miller, R. A. J. Warren and S. G. Withers, *J. Biol. Chem.*, 1992, **267**, 12559–12561.
- 11 A. Torronen and J. Rouvinen, *Biochemistry*, 1995, **34**, 847–856.
- 12 U. Krengel and B. W. Dijkstra, *J. Mol. Biol.*, 1996, **263**, 70–78.
- 13 W. W. Wakarchuk, R. L. Campbell, W. L. Sung, J. Davoodi and M. Yaguchi, *Protein Sci.*, 1994, **3**, 467–475.
- 14 S. G. Withers and R. Aebersold, *Protein Sci.*, 1995, **4**, 361–372.
- 15 S. C. Miao, L. Ziser, R. Aebersold and S. G. Withers, *Biochemistry*, 1994, **33**, 7027–7032.
- 16 K. A. Brameld and W. A. Goddard, *J. Am. Chem. Soc.*, 1998, **120**, 3571–3580.
- 17 G. Andre, P. Kanchanawong, R. Palma, H. Cho, X. Deng, D. Irwin, M. E. Himmel, D. B. Wilson and J. W. Brady, *Protein Eng.*, 2003, **16**, 125–134.
- 18 J. F. Espinosa, E. Montero, A. Vian, J. L. Garcia, H. Dietrich, R. R. Schmidt, M. Martin-Lomas, A. Imberty, F. J. Canada and J. Jimenez-Barbero, *J. Am. Chem. Soc.*, 1998, **120**, 1309–1318.
- 19 M. Kankainen, T. Laitinen and M. Perakyla, *Phys. Chem. Chem. Phys.*, 2004, **6**, 5074–5080.
- 20 X. Biarnés, J. Nieto, A. Plana and C. Rovira, *J. Biol. Chem.*, 2006, **281**, 1432–1441.
- 21 X. Biarnés, A. Arévol, A. Plana, C. Rovira, A. Laio and M. Parrinello, *J. Am. Chem. Soc.*, 2007, **129**, 10686–10693.
- 22 A. Warshel and M. Levitt, *J. Mol. Biol.*, 1976, **103**, 227–249.
- 23 A. D. Becke, *Phys. Rev. A*, 1988, **38**, 3098–3100; A. D. Becke, *J. Chem. Phys.*, 1992, **96**, 2155–2160; A. D. Becke, *J. Chem. Phys.*, 1992, **97**, 9173–9177; A. D. Becke, *J. Chem. Phys.*, 1993, **98**, 5648–5652.
- 24 *Gaussian 03*, Revision C.02, M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, J. A. Montgomery, Jr., T. Vreven, K. N. Kudin, J. C. Burant, J. M. Millam, S. S. Iyengar, J. Tomasi, V. Barone, B. Mennucci, M. Cossi, G. Scalmani, N. Rega, G. A. Petersson, H. Nakatsuji, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, M. Klene, X. Li, J. E. Knox, H. P. Hratchian, J. B. Cross, V. Bakken, C. Adamo, J. Jaramillo, R. Gomperts, R. E. Stratmann, O. Yazyev, A. J. Austin, R. Cammi, C. Pomelli, J. W. Ochterski, P. Y. Ayala, K. Morokuma, G. A. Voth, P. Salvador, J. J. Dannenberg, V. G. Zakrzewski, S. Dapprich, A. D. Daniels, M. C. Strain, O. Farkas, D. K. Malick, A. D. Rabuck, K. Raghavachari, J. B. Foresman, J. V. Ortiz, Q. Cui, A. G. Baboul, S. Clifford, J. Cioslowski, B. B. Stefanov, G. Liu, A. Liashenko, P. Piskorz, I. Komaromi, R. L. Martin, D. J. Fox, T. Keith, M. A. Al-Laham, C. Y. Peng, A. Nanayakkara, M. Challacombe, P. M. W. Gill, B. Johnson, W. Chen, M. W. Wong, C. Gonzalez and J. A. Pople, Gaussian, Inc., Wallingford CT, 2004.
- 25 G. Sidhu, S. G. Withers, N. T. Nguyen, L. P. McIntosh, L. Ziser and G. D. Brayer, *Biochemistry*, 1999, **38**, 5346–5354.
- 26 J. C. Gordon, J. B. Myers, T. Folta, V. Shoja, L. S. Heath and A. Onufriev, *Nucleic Acids Res.*, 2005, **33**, W368–W371; J. B. Myers, G. Grothaus, S. Narayanan and A. Onufriev, *Proteins: Struct. Func. Genet.*, 2006, **63**, 928–938; D. Bashford and M. Karplus, *Biochemistry*, 1990, **29**, 10219–10225.
- 27 M. J. Field, M. Albe, C. Bret, F. Proust-De Martin and A. Thomas, *J. Comput. Chem.*, 2000, **21**, 1088–1100.
- 28 M. J. S. Dewar, E. G. Zoebisch, E. F. Healy and J. J. P. Stewart, *J. Am. Chem. Soc.*, 1985, **107**, 3902–3909.
- 29 G. A. Kaminski, R. A. Friesner, J. Tirado-Rives and W. L. Jorgensen, *J. Phys. Chem. B*, 2001, **105**, 6474–6487.
- 30 W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey and M. L. Klein, *J. Chem. Phys.*, 1983, **79**, 926–935.
- 31 D. Cremer and J. A. Pople, *J. Am. Chem. Soc.*, 1975, **97**, 1354–1358.
- 32 A. Bérces, D. M. Whitfield and T. Nukada, *Tetrahedron*, 2001, **57**, 477–491.
- 33 L. Hsieh and M. L. Sinnott, *Biochem. J.*, 1985, **226**, 437–446.
- 34 U. Schnupf, J. L. Willett, W. B. Bosma and F. A. Momany, *Carbohydrate Res.*, 2007, **342**, 196–216.
- 35 W. W. Wakarchuk, R. L. Campbell, W. L. Sung, J. Davoodi and M. Yaguchi, *Protein Sci.*, 1994, **3**, 467–475.
- 36 D. L. Zechel and S. G. Withers, *Acc. Chem. Res.*, 2000, **33**, 11–18.
- 37 M. E. S. Soliman and I. H. Williams, manuscript in preparation.
- 38 A. S. Murkin, M. R. Brinck, A. Rinaldo-Matthis, W. Shi, E. A. Taylor, S. C. Almo and V. L. Schramm, *Biochemistry*, 2007, **46**, 5038–5049.
- 39 S. Núñez, D. Antoniou, V. L. Schramm and S. D. Schwartz, *J. Am. Chem. Soc.*, 2004, **126**, 15720–15729.