

MM-PBSA free energy analysis of *endo*-1,4-xylanase II (XynII)–substrate complexes: binding of the reactive sugar in a skew boat and chair conformation †

Tuomo Laitinen,^a Juha Rouvinen^b and Mikael Peräkylä^{*a}

^a Department of Chemistry, University of Kuopio, PO Box 1627, FIN-70211 Kuopio, Finland

^b Department of Chemistry, University of Joensuu, PO Box 111, FIN-80101 Joensuu, Finland

Received 30th June 2003, Accepted 5th September 2003

First published as an Advance Article on the web 23rd September 2003

The binding of xylotetraose in different conformations to the active site of *endo*-1,4- β -xylanase II (XynII) from *Trichoderma reesei* was studied using molecular dynamics (MD) simulations and free energy analyses employing the MM-PBSA (Molecular Mechanics–Poisson–Boltzmann Surface Area) method. MD simulations of 1 ns were done for the substrate xylotetraose having the reactive sugar, which is bound in the –1 subsite of XynII in the 4C_1 (chair) and 2S_0 (skew boat) ground state conformations, and for the transition state of the XynII catalysed hydrolysis of the β -glycosidic linkage. According to the simulations and free energy analysis, XynII binds the substrate with the –1 sugar in the 2S_0 conformation 59.8 kJ mol⁻¹ tighter than the substrate with the sugar in the 4C_1 conformation. The reactive 2S_0 conformation resembles closely the reaction transition state and has the breaking glycosidic bond in a *pseudo*-axial orientation ready for facile bond cleavage. The transition state was calculated to be bound 77.1 kJ mol⁻¹ tighter than the 4C_1 ground state conformation. The molecular mechanical interaction energy between the enzyme and the reactive pyranoside unit at the –1 subsite was 75.7 kJ mol⁻¹ more favorable for the binding of the 2S_0 conformation than the 4C_1 conformation, explaining the clearly tighter binding of the reactive structure. The results of this study indicate that in the Michaelis complex XynII, a member of the family 11 xylanase, the substrate is bound in a skew boat conformation and in the catalytic reaction, the –1 sugar proceeds from the 4C_1 conformation through 2S_0 to the transition state with the –1 sugar in the ${}^{2,5}B$ conformation.

Introduction

Glycosyl hydrolases are enzymes for which an increasing amount of structural and biological data has become available. They can be divided into superfamilies on the basis of their reaction mechanisms, 3-dimensional structures, and amino acid sequence similarities. Xylanases of the families 10 and 11 are involved in the hydrolysis of nature's most abundant hemicellulose, xylan. These xylanases follow a reaction mechanism (Fig. 1), which yields overall retention of the anomeric configuration of the substrate.^{1–3} The retaining mechanism involves two functional groups in the active site, one carboxyl and one carboxylate, which display an average separation of 5 Å between the two catalytic residues. The transition states to and from the covalent intermediate have considerable oxocarbenium ion character.^{1,4} Several X-ray structures of protein–inhibitor complexes have provided atomic resolution data on the structure of the reaction intermediate of glycosidases. In the case of the family 10 xylanase⁵ and the family 12 cellulase⁶, the 2-deoxy-2-fluoro-xylose residue was observed to bind to the –1 binding subsite in a 4C_1 conformation. In contrast, the 2-deoxy-

2-fluoro-xylobiosyl binds in a ${}^{2,5}B$ conformation to the –1 binding subsite of the family 11 xylanases.^{7,8} The unusual ${}^{2,5}B$ conformation is a result of lateral *syn*-protonation architecture of the active site of the family 11 xylanases.⁷ However, it is less clear in which conformation the reactive pyranoside of the substrate is bound in the Michaelis complex. This is mainly due to the lack of suitable substrates, which are similar to natural substrates, not hydrolyzed and also occupy the reducing end of the active site. Distortion of the reactive pyranoside ring has been observed in the X-ray structures of several β -glycosidases complexed with uncleaved substrates.^{6,7,9}

In the present work we have computationally investigated the catalytic properties of *endo*-1,4- β -xylanase II (XynII) from *Trichoderma reesei*, a member of the retaining family 11 xylanase. Molecular dynamics (MD) simulations and free energy analyses using the MM-PBSA (Molecular Mechanics–Poisson–Boltzmann Surface Area) method¹⁰ were applied to study the binding of xylotetraose to the active site of XynII. The MM-PBSA method is a quite recently introduced approach to calculate free energies of macromolecules,^{11–13} which can also be applied to estimate the free energies for ligand binding to proteins and for carrying out computational mutagenesis.^{14–17} MD simulations of 1 ns were carried out for the ground state (Michaelis complex) and transition state-

† Electronic supplementary information (ESI) available: xylotetraose charges. See <http://www.rsc.org/suppdata/ob/b3/b307335a/>

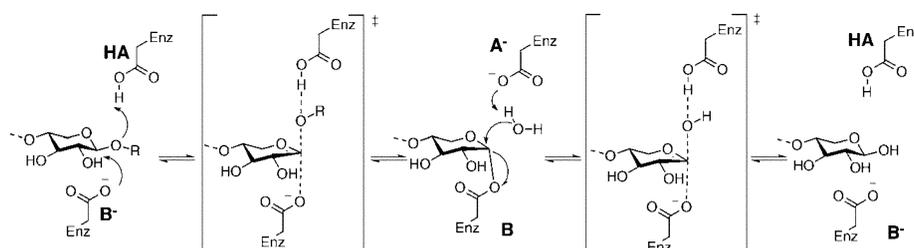


Fig. 1 Reaction scheme for the retaining mechanism of glycoside hydrolases.

mimicking (oxocarbenium ion) structures of the XynII–xylotetraose complexes. The free energies of the simulated structures were analysed using the MM-PBSA method. In addition, the roles of individual binding-site residues in xyloxytetraose binding were further probed by computational alanine scanning with the MM-PBSA method. On the basis of the structural and energetic analyses, we describe how XynII binds enzyme reaction ground and transition state structures. Specifically, we show that XynII binds substrate preferentially in a ground state conformation which has the reactive –1 sugar in the 2S_0 (skew boat) conformation.

Results and discussion

Structures of ground state complexes

The *endo*-1,4- β -xyloxytetraose (XynII) from *Trichoderma reesei* belongs to family 11 of the xyloxytetraoses and it has an active site geometry in which the proton donor interacts with the scissile glycosidic bond from the side of the carbohydrate plane allowing *syn* protonation.¹⁸ Xyloxytetraose (X4) was docked to subsites –2 and –1 using the complex structure of *Bacillus circulans* xyloxytetraose mutant E172C as a reference,¹⁹ whereas several orientations were considered for xylose residues occupying subsites +1 and +2. After preliminary equilibration dynamics, two xyloxytetraose conformations were selected for further simulations: one with the –1 sugar in the 2S_0 conformation (marked as X4(sb)) and the rest in the 4C_1 conformation, and another with all sugars in the 4C_1 conformation (marked as X4(c)) (see Fig. 2). These structures had no significant steric overlap with XynII and they fulfilled the geometrical requirements for the lateral protonation of the scissile glycosidic bond in the enzymatic reaction.¹⁸ The binding orientations of the +2 and +1 sugars of the simulation structures are close to those modeled by Gruber *et al.* for the family 11 xyloxytetraose from *Thermomyces lanuginosus*.²⁰

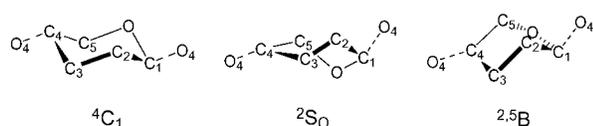


Fig. 2 Xylopyranoside conformations 4C_1 , 2S_0 and ${}^{2.5}B$.

The two XynII–xyloxytetraose complexes were first equilibrated for 300 ps after which a production simulation of 1 ns was started. Both simulations equilibrated in 300 ps and they were stable during the production runs according to the *Ca* rms deviation (Fig. 3) and energies (data not shown). However, in the XynII–X4(sb) simulation the binding site cleft opened after 700 ps and, consequently, resulted in a decrease in the binding affinity (Fig. 4). Similar protein motion, believed to be important for reactant binding and product release in the catalytic

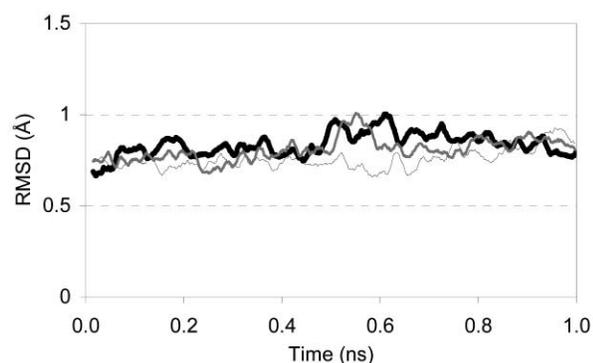


Fig. 3 *Ca* rms deviation of molecular dynamics simulations of X4(c) (thin line), X4(sb) (thick line) and transition state (gray line) complexed with XynII.

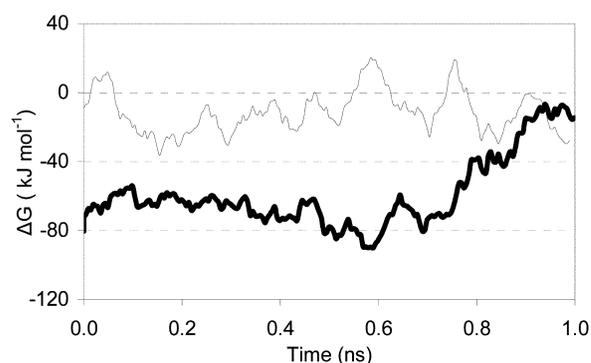


Fig. 4 Binding free energies (kJ mol^{-1}) as a function of time for X4(c) (thin line) and X4(sb) (thick line).

cycle of the enzyme, has been observed both experimentally and theoretically.^{21,22} Therefore, the structures collected during the first 700 ps of production simulation were used in the free energy analysis.

The average ϕ ($\text{O}5_{i-1}\text{C}1_{i-1}\text{O}4_{i+1}\text{C}4_{i+1}$) and ψ ($\text{C}1_{i-1}\text{O}4_{i+1}\text{C}4_{i+1}\text{C}3_{i+1}$) torsion angles between xylose subunits –2/–1, –1/+1 and +1/+2 were $-74^\circ/-72^\circ/-75^\circ$ and $+158^\circ/+163^\circ/+154^\circ$ in a free water simulation of X4(c). In the enzyme simulation the torsion angles between the –2/–1 and +1/+2 subunit pairs were $-98^\circ/-74^\circ$ for ϕ and $148^\circ/161^\circ$ for ψ torsion angle of X4(c). In the enzyme simulation of X4(sb) the torsion angles between the –2/–1 and +1/+2 pyranosides were $-75^\circ/-84^\circ$ for ϕ and $+134^\circ/+162^\circ$ for ψ torsion angle. These numbers are close to those observed in the free water simulations of xyloxytetraose in all-chair (4C_1) conformation. However, binding to the enzyme resulted in a kink in the polysaccharide chain between the –1/+1 pyranosides. The torsion angles of the –1/+1 linkage were -170° (ϕ) and $+90^\circ$ (ψ) for X4(sb) and -177° (ϕ) and $+82^\circ$ (ψ) for X4(c). That binding to an enzyme changes the torsion angle of the reactive glycosidic bond has been observed, for example, in the X-ray structure of the family 5 endocellulase E1 from *Acidothermus cellulolyticus*.²³

The average structures of the XynII–X4(c) and XynII–X4(sb) simulations (Figs. 5a and b, averages between the 0.2–0.7 ps of production simulations) show that the 2S_0 conformation of the reactive sugar forms more favourable interactions with the enzyme than the 4C_1 conformation. The hydrogen bond distances between the C2–OH of the reactive sugar and E86 are 2.7 Å for X4(sb) and 4.0 Å for X4(c). In addition, the distance from the carboxylic oxygen of the catalytic nucleophile (E86) to the anomeric C1 is 3.4 Å for X4(sb), whereas it is 4.4 Å for X4(c). In this subsite (–1) X4(sb) forms two hydrogen bonds with the NH and NH_2 groups of R122 (2.9 Å), whereas X4(c) forms only one hydrogen bond with the NH_2 group of R122 (2.9 Å). The C3–OH of the –1 pyranoside forms a hydrogen bond with the carbonyl oxygen of P126 (O–O distances are 2.6 Å for X4(sb) and 2.7 Å for X4(c)). The glycosidic bond that is cleaved in the enzyme reaction is close to the acid/base catalyst with a distance of 3.7 Å between the glycosidic oxygen and carboxylate oxygen of E177 in both substrate conformations. In the subsites –1 and –2 both substrate conformations form a hydrogen bonding network with Y77. In the subsite +1, Y73 forms a hydrogen bond with C2–OH. This distance is 2.8 Å for X4(c) and 2.9 Å for X4(sb) conformation. In the subsite +2, proper hydrogen bond partners are not present but Y96 and Y179 are able to provide stacking interactions with the substrate.

Structure of the transition state complex

To study the properties of the reaction transition state (intermediate) MD simulations were performed for the oxocarbenium ion intermediate (Fig. 5c). The simulated structure

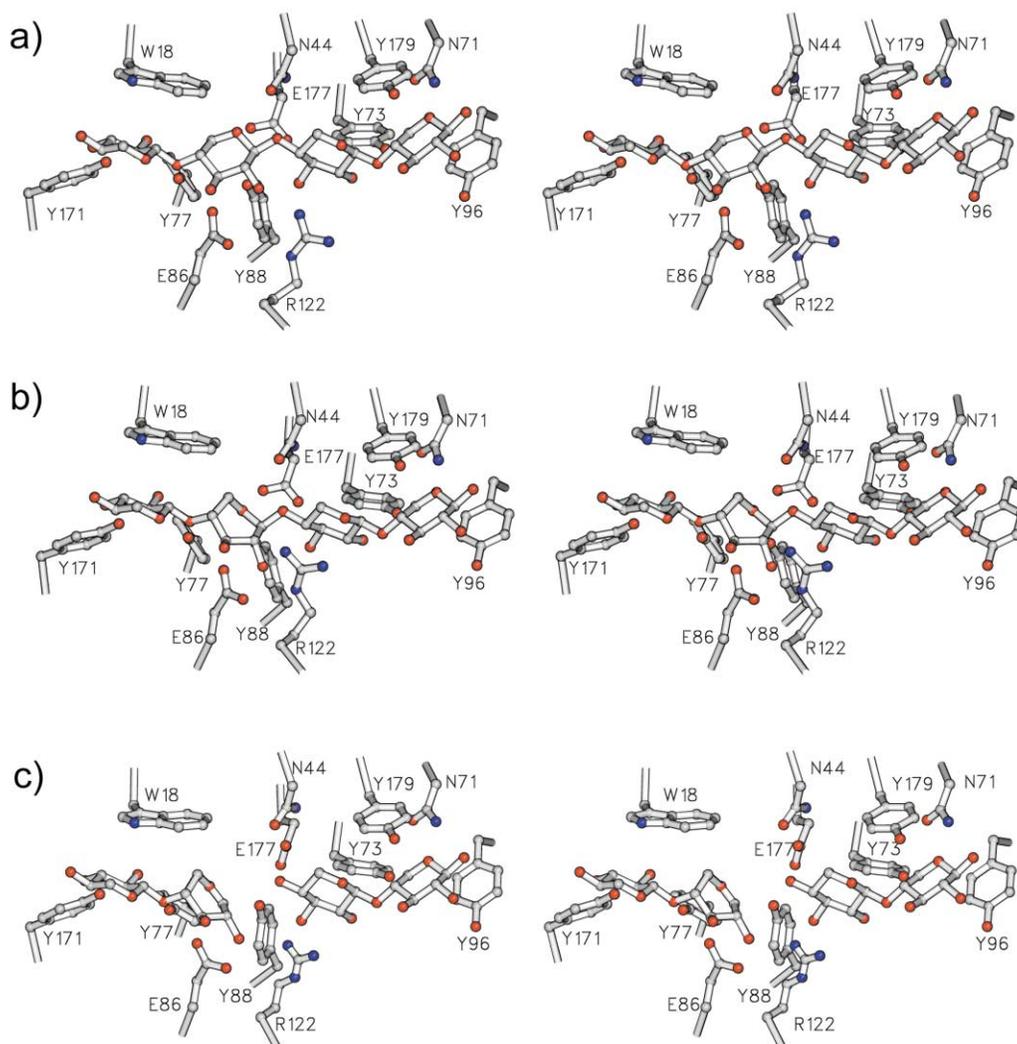


Fig. 5 Stereo presentation of average structures of a) XynII–X4(c) b) XynII–X4(sb) and c) XynII–transition state complexes. Xylotetraose occupies the subsites –2, –1, +1 and +2 (from left).

represents a state where the glycosidic bond is cleaved, the glycosidic oxygen is protonated, the C1 carbon carries a positive charge and both catalytic glutamates are charged. This structure is thought to be close to the transition state of the retaining glycosyl hydrolases and therefore likely shows interactions important for transition state stabilization.^{24,25} In the simulated structure the glycosidic bond between the xylobioside parts of tetraose is cleaved and the distance between the C1 of the oxocarbenium ion and the glycosidic oxygen was set at 4.0 Å by a restraint of 21 kJ mol⁻¹. With this restraint the newly formed OH is still in the close vicinity of the acid/base catalyst but not too close to the positively charged C1. A modified set of parameters for the intermediate was created to keep the carbocation part in a half chair conformation, which was tested in a free water simulation. The planar starting structure was assembled at the subsites –1 and –2 according to the three-dimensional structures of the family 11 xylanase–2-fluoro-xylobiosyl complexes (Fig. 5c).^{7,8} The xylobioside of the reducing end was positioned at the subsites +1 and +2 following the guidelines provided by the ground state simulations (Figs. 5a and b). It was observed that interactions with the enzyme caused the oxocarbenium ion to adopt the ^{2,5}B conformation after 200 ps of the equilibration run (Figs. 2 and 5c). The ^{2,5}B structure formed was stable along the rest of the equilibration and production simulations. E86 plays a major role in the formation and stabilization of this structure, forming a hydrogen bond with C2–OH and by coordinating with the positively charged anomeric C1 carbon. The average distance between the catalytic E86 and C1 was 2.7 Å and the distance

between E86 and C2–OH was 2.5 Å. Furthermore, the structure formed during the simulation is almost identical to the X-ray structure of the *Bacillus circulans* xylanase–deoxy-2-fluoro-xylobioside complex (1BCX).⁸ The xylobioside part of the +1 and +2 binding sites in the simulated oxocarbenium ion represents a structure with a newly formed hydroxyl of the leaving disaccharide that is still hydrogen bonded to the catalytic E177 (2.6 Å).

Energy analyses of the complexes

The binding free energies for the two xylotetraose conformations (X4(c) and X4(sb)) and the transition state were estimated (Table 1) using the MM-PBSA method.¹⁰ The calculated binding free energies were –10.1 kJ mol⁻¹ for X4(c), –69.9 kJ mol⁻¹ for X4(sb) and –87.2 kJ mol⁻¹ for the transition state. The analysis of the free energy components shows that the –73.3 kJ mol⁻¹ more favorable electrostatic XynII–substrate interactions are the main reason for the higher affinity of X4(sb) as compared to X4(c). On the other hand, the van der Waals interactions and the solvation free energy are more favorable for X4(c) conformation by 2.0 kJ mol⁻¹ and 11.5 kJ mol⁻¹ respectively. For the transition state the binding is due to strong electrostatic interactions between the positively charged carbocation and the negatively charged nucleophile. This effect is partly compensated by the unfavorable solvation energy term.

In the solution the total free energy of the xylotetraose in the skew boat (X4(sb)) conformation is 515.2 kJ mol⁻¹ and in the chair (X4(c)) conformation 490.5 kJ mol⁻¹. In the enzyme the

Table 1 Calculated average binding free energies (kJ mol^{-1}) for substrates in different conformations (standard deviations in parentheses)^a

	⁴ C ₁	² S ₀	TS
$\langle \Delta E_{\text{elec}} \rangle$	-236.6 (36.0)	-309.9 (30.7)	-685.3 (37.5)
$\langle \Delta E_{\text{vdw}} \rangle$	-259.5 (12.7)	-257.5 (16.1)	-222.2 (18.9)
$\langle \Delta E_{\text{MM}} \rangle$	-496.1 (37.3)	-567.4 (25.9)	-907.5 (34.9)
$\langle \Delta \Delta G_{\text{np}} \rangle$	-18.8 (0.5)	-18.9 (0.6)	-18.5 (0.7)
$\langle \Delta \Delta G_{\text{PB}} \rangle$	504.8 (28.8)	516.4 (22.9)	838.9 (35.5)
$\langle \Delta \Delta G_{\text{solv}} \rangle$	486.0 (28.7)	497.5 (22.7)	820.3 (35.4)
$\langle \Delta \Delta G_{\text{PB,elec}} \rangle$	268.2 (24.8)	206.5 (26.2)	153.6 (26.2)
$\langle \Delta G_{\text{bind}} \rangle$	-10.1 (25.7)	-69.9 (19.7)	-87.2 (19.7)
$\langle \Delta \Delta G_{\text{bind}} \rangle$	0.0	-59.8	-77.1

^a Definition of energy contributions: $\langle \Delta E_{\text{elec}} \rangle$ = electrostatic molecular mechanical energy, $\langle \Delta E_{\text{vdw}} \rangle$ = van der Waals molecular mechanical energy, $\langle \Delta E_{\text{MM}} \rangle = \langle \Delta E_{\text{elec}} \rangle + \langle \Delta E_{\text{vdw}} \rangle$, $\langle \Delta \Delta G_{\text{np}} \rangle$ = non-polar solvation energy, $\langle \Delta \Delta G_{\text{PB}} \rangle$ = electrostatic solvation energy, $\langle \Delta \Delta G_{\text{solv}} \rangle = \langle \Delta \Delta G_{\text{np}} \rangle + \langle \Delta \Delta G_{\text{PB}} \rangle$, $\langle \Delta \Delta G_{\text{PB,elec}} \rangle = \langle \Delta E_{\text{elec}} \rangle + \langle \Delta \Delta G_{\text{PB}} \rangle$, $\langle \Delta G_{\text{bind}} \rangle$ = calculated binding energy, $\langle \Delta \Delta G_{\text{bind}} \rangle$ = relative binding energy.

total free energy of xylootetraose is $521.2 \text{ kJ mol}^{-1}$ when the -1 sugar is in the ²S₀ conformation and $511.7 \text{ kJ mol}^{-1}$ when it is in the ⁴C₁ conformation. Because these calculated values are well converged, we may conclude that upon binding to the enzyme the energy difference between the two X4 conformations becomes considerably smaller (24.7 kJ mol^{-1} (water) vs. 9.5 kJ mol^{-1} (enzyme)).

The molecular mechanical energies of individual pyranoside subunits were examined using a strategy in which the substrate coordinates were taken from the molecular dynamics trajectories, the extracted structures were rebuilt as xylose monomers and the average internal molecular mechanical energies of the monomers and their interaction energies with the protein were calculated (Figs. 6 and 7). In the enzyme the X4(c) conformation has internal energies of $144.0 \text{ kJ mol}^{-1}$ and $143.4 \text{ kJ mol}^{-1}$ in subsites -2 and -1, and $136.3 \text{ kJ mol}^{-1}$ and $135.2 \text{ kJ mol}^{-1}$

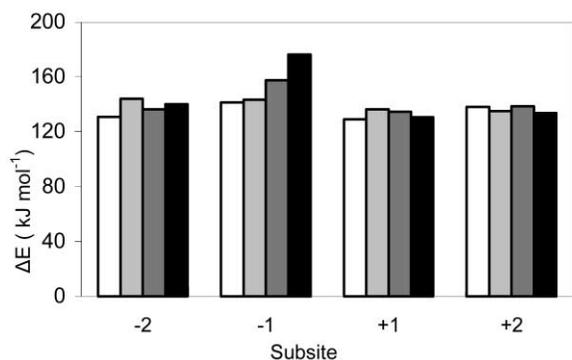


Fig. 6 Internal molecular mechanical energies (kJ mol^{-1}) of xylose monomers in water and within the enzyme. Energies of X4(c) are shown with white (water) and light gray (enzyme) columns and energies of X4(sb) are shown with dark gray (water) and black (enzyme) columns.

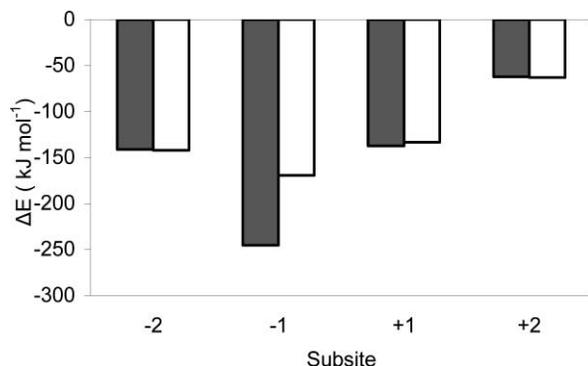


Fig. 7 Molecular mechanical interaction energies (kJ mol^{-1}) between xylopyranoside monomers and XynII. Energies of X4(c) are shown with white and energies of X4(sb) are shown with gray columns.

in subsites +1 and +2 (Fig. 6). In solution the energy of the X4(c) conformation is about 9 kJ mol^{-1} higher in subsites -1 and +2 than in subsites -2 and +1. In the case of X4(sb) the internal energy of the pyranoside in the skew boat conformation is $157.6 \text{ kJ mol}^{-1}$ in solution and $176.3 \text{ kJ mol}^{-1}$ when bound to the enzyme. Internal energies of the pyranoside units in the ⁴C₁ conformation are within 6.3 kJ mol^{-1} both in solution and within the enzyme.

The interaction energy of the reactive pyranoside unit in the subsite -1 is 75.7 kJ mol^{-1} more favourable for the binding of the skew boat conformation than the chair conformation (Fig. 7). In both cases the protein-monomer interaction energies are within 8.8 kJ mol^{-1} in the subsites -2, +1 and +2. The subsite +2 is the least tightly bound of the four binding subsites with 62.8 kJ mol^{-1} lower interaction energies than in the -2 and +1 subsites. According to the XynII-X4 interaction energies, the most important subsite for the substrate binding is subsite -1. The subsites -2 and +1 are also important for the tight binding and presumably for the correct positioning of the reactive sugar unit. That the sugar moiety of the +2 subsite is bound the least tightly of the four moieties may be linked to the product release step of the reaction cycle: when the glycosidic bond is cleaved, the sugar moieties of the +1 and +2 subsites need to leave the substrate-binding site fast enough in order not to hinder the binding of the substrate of the next reaction cycle.

Computational mutagenesis of active site residues

Computational alanine scanning of selected amino acid residues of the active site groove was carried out to elucidate the role of individual residues on the binding of the substrate with the -1 sugar in the ⁴C₁ and ²S₀ conformations. The residues mutated are shown in Figs. 5a and b and the results of the mutagenesis are presented in Fig. 8. It must be emphasized that the results of the computational mutagenesis of this study, done with a single trajectory method, are not meant to reproduce the corresponding energies measured experimentally, but to provide an estimate of the interactions between the sugar and the residue mutated in the original xylanase-X4 complex.

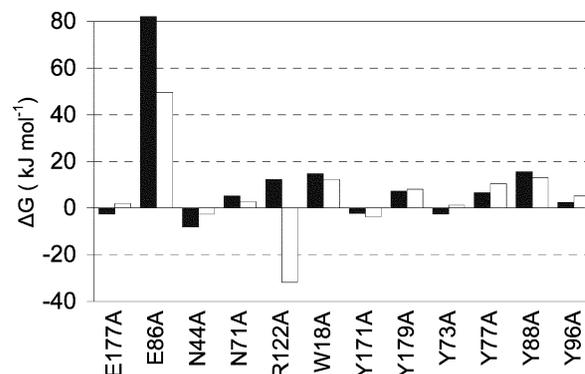


Fig. 8 Computational alanine mutants (kJ mol^{-1}) calculated for the XynII-X4(c) (white columns) and XynII-X4(sb) (black columns) complexes.

The E86A mutation lowers the affinity of X4(sb) by $125.5 \text{ kJ mol}^{-1}$ and that of X4(c) by 49.6 kJ mol^{-1} . Other main components of the binding are aromatic residues Y179, Y77, Y88, and W18, which all have effects of $6\text{--}16 \text{ kJ mol}^{-1}$ when mutated to alanine. The mutation R122A lowers the binding energy by 12.3 kJ mol^{-1} for X4(sb) but increases it by 31.8 kJ mol^{-1} for X4(c) although in both cases R122 is at the hydrogen bonding distance from the substrate (Figs. 5a and b). The difference above is partly due to the fact that in the case of X4(sb) R122 forms two hydrogen bonds with the C2-OH of the -1 sugar whereas only one hydrogen bond is detected in the case of X4(c). In addition, in the case of X4(sb), R122 forms tighter interactions with the +1 sugar because X4(sb) is

positioned slightly deeper in the substrate binding cavity than X4(sb). In both cases the mutation of the neutral acid/base catalyst E177A has only a minor effect on the affinities.

Implications for glycosidase catalysis

The common principles of molecular recognition, *i.e.* hydrogen bonds, dispersion forces and hydrophobic stacking, are the main forces in substrate binding to glycosidases. Recently, it has been demonstrated that enzymes tend to bind their substrates in a reactive ground state conformation and in this way catalyse chemical reactions considerably.^{26,27} In such a reactive ground state conformation, the substrate and the catalytic groups of the enzyme are precisely placed at correct positions ready to meet the reaction transition state. In the case of family 11 glycosidases, it is well established that in the high-energy reaction intermediate the anomeric carbon of the substrate is covalently bound to the catalytic nucleophile.⁸ In this computational investigation of XynII-xylo-tetraose complex structures we found that when xylo-tetraose is bound in an all-chair (⁴C₁) conformation the catalytic groups are not positioned in a reactive ground state conformation. However, the xylo-tetraose that has its -1 sugar in a ²S₀ conformation has the shape that is needed for short reactive distances and to have catalytic groups at the correct positions. In the simulations of the reaction transition state structure, the -1 sugar adopted the ^{2,5}B conformation (see Figs. 2 and 5c) during the equilibrium run, which is similar to the structures observed when fluoro inhibitors are covalently bound to family 11 xylanases. Structurally the ²S₀ conformation is located between the ⁴C₁ and ^{2,5}B conformations and, therefore, likely represents a structure formed along the reaction pathway.

Experimental

MD simulations

The initial coordinates of the *endo*-1,4-β-xylanase (XynII) from *Trichoderma reesei* were obtained from the X-ray crystal structure of the protein complexed with 2,3-epoxypropyl-β-D-xyloside determined at 1.8 Å resolution.¹⁹ In the ground state simulations, the acid/base catalyst was protonated and the nucleophile was charged and in the intermediate simulation both catalytic acids were charged.²⁸ The MD simulations were performed with the AMBER7 program,²⁹ using the Cornell *et al.* force field.³⁰ The modified parameters of the GLYCAM_93 set³¹ were used for the xylopyranoside parts, and their atomic point charges were calculated using the two-stage RESP method³² at the HF/6-31G* level using geometries optimized at the same level (xylo-tetraose charges are available as electronic supplementary information †). An additional set of parameters was created to keep the carbocation unit in a half-chair conformation. The *leap* program³³ was used to set up the simulation systems. The simulations were done in an explicit solvent (average 5502 TIP3P waters) under periodic boundary conditions (truncated octahedron 63*63*63 Å) using the particle-mesh Ewald method (PME) for electrostatic interactions.³⁴ The simulations were carried out at a temperature of 300 K and a pressure of 1 atm. The Van der Waals interactions were truncated by using a cutoff value of 8 Å. The SHAKE algorithm³⁵ was applied to constrain bonds including hydrogen atoms within their equilibrium values. In the simulation of the intermediate complex a restraint of 21 kJ mol⁻¹ was used to keep the distance between the two-disaccharide parts at 4 Å. At the beginning of each simulation the solvent box was equilibrated for 50 ps, keeping the protein fixed. After that the whole simulation system was energy minimized (1000 steps), heated up to the simulation temperature of 300 K in 5 ps and equilibrated with an additional 300 ps simulation. We used a time step of 1.5 fs and the structures were saved every 30 fs for further analyses. The rms deviation from the initial starting

structure was calculated using *ptraj*, the trajectory analysis suite of AMBER7. Snapshots of the trajectories were visually examined with the assistance of the MOIL-View program³⁶ and the illustrations were generated using the SETOR program.³⁷

Energy analyses

The MM-PBSA method used here involves the calculation of energies for the snapshot structures taken from the MD trajectories, followed by calculation of the averages of the energy values. In this work the snapshot structures for the energy calculations of the protein-ligand complex and separated protein and ligand were taken from the MD trajectory of the protein-ligand complex and the reported binding free energies are averages from the 50 snapshots. It has been observed that this single-trajectory method provides fairly good estimates for the relative binding energies.^{10,13,15,16} In addition, it was observed in our previous study that 50 snapshots are sufficient for converged results.¹⁷ In the MM-PBSA method the average binding free energy (ΔG_{bind}) is calculated from the average molecular mechanical gas-phase energies (ΔE_{MM}), solvation free energies ($\Delta \Delta G_{\text{solv}}$) and entropy contributions ($-T\Delta S$) of the binding reaction:

$$\Delta G_{\text{bind}} = \Delta E_{\text{MM}} + \Delta \Delta G_{\text{solv}} - T\Delta S \quad (1)$$

The molecular mechanical (E_{MM}) energy of each snapshot was calculated using the *sander* program of AMBER7 with all pair-wise interactions included using a dielectric constant (ϵ) of 1. The solvation free energy (ΔG_{solv}) was estimated as the sum of electrostatic solvation free energy, calculated by the finite-difference solution of the Poisson-Boltzmann equation (ΔG_{PB}) as implemented in the *Delphi* program³⁸ and non-polar solvation free energy (ΔG_{np}), calculated from the solvent-accessible surface area (SASA):

$$\Delta G_{\text{solv}} = \Delta G_{\text{PB}} + \Delta G_{\text{np}} \quad (2)$$

We used $\epsilon = 1$ for the solute and $\epsilon = 80$ for the solvent in the electrostatic solvation free energy (ΔG_{PB}) calculations. A probe radius of 1.4 Å and atomic radii of the PARSE parameter set³⁹ were used to determine the molecular surface. Atomic charges of the Cornell *et al.* force field³⁰ were used for amino acid residues and RESP charges were calculated at the HF/6-31G* level for the carbohydrates.³² An 80% boxfill cubic lattice and a grid resolution of 0.5 Å/grid point were used in the *Delphi* calculations. The *molsurf* program was used to calculate the solvent-accessible surface area (SASA)⁴⁰ for the estimation of the non-polar solvation free energy (ΔG_{np}) using eqn. (3) and $\gamma = 0.0227 \text{ kJ mol}^{-1} \text{ \AA}^{-2}$ and $\beta = 3.85 \text{ kJ mol}^{-1}$.⁴¹

$$\Delta G_{\text{np}} = \gamma \times \text{SASA} + \beta \quad (3)$$

In this work, the energy contribution from entropy changes upon ligand binding was not included. This was justified by the fact that it is likely that entropy does not contribute much to the relative binding free energies of the carbohydrates to the same protein. In addition, there is no straightforward way to quantitatively calculate entropy contribution to binding. The normal mode analysis, which is often used to estimate entropy changes, only provides qualitative estimates.¹⁵

The energy contributions of individual sugar moieties were examined by extracting the coordinates of the sugar monomers from the molecular dynamics trajectories using the *carnal* module of AMBER7. The missing valences of the extracted xylopyranoside monomers were filled with hydrogen atoms with the *leap* program of AMBER7. The charges of the added hydrogens were adjusted to produce sugar monomers with a net zero charge. Finally, the internal molecular mechanical energies of xylopyranoside units in different subsites and molecular

mechanical interaction energies between XynII and individual xylopyranoside units were calculated using the *anal* program of AMBER7.

Mutagenesis

We examined the role of selected active site residues in the substrate binding by calculating selected point mutations with the MM-PBSA method for XynII–xylootetraose complexes. In the mutagenesis, the modified snapshot structures were used to calculate the binding free energies for the mutant proteins. The xylanase structures were modified by truncating the selected side chains to alanines and by building the missing hydrogen atoms of the alanine C β atoms in a standard geometry. 50 snapshots were taken from the MD trajectories at even intervals for the binding energy analyses.¹⁷

Conclusions

In the simulations of this work we found two ground state conformations, which have a glycosidic bond suitably positioned for the enzyme reaction. In these structures the reactive sugar bound in the –1 subsite is in the ⁴C₁ (chair, X4(c)) and ²S₀ (skew boat, X4(sb)) conformation and the acid/base catalyst of XynII ready to protonate the substrate in the plane of the pyranoside with the scissile glycosidic bond positioned for the *syn*-protonation trajectory. These observations are in line with the suggestion put forward by Heightman and Vasella¹⁸ with regard to family 11 of the glycosidases. It is assumed that, in the reactive ground state structure the catalytic groups and the substrate are precisely at the correct positions for the reaction.^{26,27} In the XynII-catalyzed reaction the scissile glycosidic oxygen should point towards the proton donor E177 and the catalytic nucleophile E86 should be positioned above the α -face of the pyranoside ring close to the anomeric carbon. The criteria of the reactive ground state structure are fulfilled in the case of the X4(sb) conformation. The X4(sb) conformation is bound considerably tighter than the X4(c) and it has its glycosidic linkage in the *pseudo*-axial orientation ready for facile cleavage of the glycosidic bond. In the ²S₀ conformation the –1 sugar is geometrically close to the reaction transition state structure.

Acknowledgements

This work was supported by the Academy of Finland (grant 74097 to M. P.). We also thank CSC – Scientific Computing Ltd. (Espoo, Finland) for providing computational resources.

References

- 1 D. E. Koshland, *Biol. Rev.*, 1953, **28**, 314–436.
- 2 G. Davies and B. Henrissat, *Structure*, 1995, **3**, 853–859.
- 3 J. McCarter and S. G. Withers, *Curr. Opin. Struct. Biol.*, 1994, **4**, 885–892.
- 4 D. L. Zechel and S. G. Withers, *Acc. Chem. Res.*, 2000, **33**, 11–18.
- 5 G. J. Davies, L. Mackenzie, A. Varrot, M. Dauter, A. M. Brzozowski, M. Schulein and S. G. Withers, *Biochemistry*, 1998, **37**, 11707–11713.
- 6 G. Sulzenbacher, L. F. Mackenzie, K. S. Wilson, S. G. Withers, C. Dupont and G. J. Davies, *Biochemistry*, 1999, **38**, 4826–4833.
- 7 A. Sabini, G. Sulzenbacher, M. Dauter, Z. Dauter, P. L. Jorgensen, M. Schulein, C. Dupont, G. J. Davies and K. S. Wilson, *Chem. Biol.*, 1999, **6**, 483–492.

- 8 G. Sidhu, S. G. Withers, N. T. Nguyen, L. P. McIntosh, L. Ziser and G. D. Brayer, *Biochemistry*, 1999, **38**, 5346–5354.
- 9 I. Tews, A. Perrakis, A. Oppenheim, M. Dauter, K. S. Wilson and C. E. Vorgias, *Nat. Struct. Biol.*, 1996, **3**, 638–648.
- 10 I. Massova and P. A. Kollman, *J. Am. Chem. Soc.*, 1999, **121**, 8133–8143.
- 11 B. Jayaram, D. Sprous, M. A. Young and D. L. Beveridge, *J. Am. Chem. Soc.*, 1998, **120**, 10629–10633.
- 12 Y. N. Vorobjev, J. C. Almagro and J. Hermans, *Proteins*, 1998, **32**, 399–413.
- 13 P. A. Kollman, I. Massova, C. Reyes, B. Kuhn, S. Huo, L. Chong, M. Lee, T. Lee, Y. Duan, W. Wang, O. Donini, P. Cieplak, J. Srinivasan, D. A. Case and T. E. III. Cheatham, *Acc. Chem. Res.*, 2000, **33**, 889–897.
- 14 B. Kuhn and P. A. Kollman, *J. Am. Chem. Soc.*, 2000, **122**, 3909–3916.
- 15 B. Kuhn and P. A. Kollman, *J. Med. Chem.*, 2000, **43**, 3786–3791.
- 16 M. Peräkylä and N. Nordman, *Protein Eng.*, 2001, **14**, 753–758.
- 17 N. Nordman, J. Valjakka and M. Peräkylä, *Proteins*, 2003, **50**, 135–143.
- 18 T. D. Heightman and A. T. Vasella, *Angew. Chem., Int. Ed.*, 1999, **38**, 750–770.
- 19 W. W. Wakarchuk, R. L. Campbell, W. L. Sung, J. Davoodi and M. Yaguchi, *Protein Sci.*, 1994, **3**, 467–475.
- 20 K. Gruber, G. Klintschar, M. Hayn, A. Schlacher, W. Steiner and C. Kratky, *Biochemistry*, 1998, **37**, 13475–13485.
- 21 R. Havukainen, A. Törrönen, T. Laitinen and J. Rouvinen, *Biochemistry*, 1996, **35**, 9617–9624.
- 22 J. Muilu, A. Törrönen, M. Peräkylä and J. Rouvinen, *Proteins*, 1998, **31**, 434–444.
- 23 J. Sakon, W. S. Adney, M. E. Himmel, S. R. Thomas and P. A. Karplus, *Biochemistry*, 1996, **35**, 10648–10660.
- 24 A. Vasella, G. Davies and M. Böhm, *Curr. Opin. Chem. Biol.*, 2002, **6**, 619–629.
- 25 G. J. Davies, V. M.-A. Ducros, A. Varrot and D. L. Zechel, *Biochem. Soc. Trans.*, 2003, **31**, 523–527.
- 26 T. C. Bruice and F. C. Lightstone, *Acc. Chem. Res.*, 1999, **32**, 127–136.
- 27 R. A. Torres and T. A. Bruice, *Proc. Natl. Acad. Sci. U. S. A.*, 1998, **95**, 11077–11082.
- 28 A. Varrot, C. A. Tarling, J. M. Macdonald, R. V. Stick, D. L. Zechel, S. G. Withers and G. J. Davies, *J. Am. Chem. Soc.*, 2003, **125**, 7496–7497.
- 29 D. A. Case, D. A. Pearlman, J. W. Caldwell, T. E. Cheatham III, J. Wang, W. S. Ross, C. L. Simmerling, T. A. Darden, K. M. Merz, R. V. Stanton, A. L. Cheng, J. J. Vincent, M. Crowley, V. Tsui, H. Gohlke, R. J. Radmer, Y. Duan, J. Pitera, I. Massova, G. L. Seibel, U. C. Singh, P. K. Weiner and P. A. Kollman, AMBER 7, University of California, San Francisco, 2002.
- 30 W. D. Cornell, P. Cieplak, C. I. Bayly, I. R. Gould, K. M. Merz, D. M. Ferguson, D. C. Spellmeyer, T. Fox, J. W. Caldwell and P. A. Kollman, *J. Am. Chem. Soc.*, 1995, **117**, 5179–5197.
- 31 R. J. Woods, R. A. Dvek, D. J. Edge and B. Fraser-Reid, *J. Phys. Chem.*, 1995, **99**, 3822–3846.
- 32 C. I. Bayly, P. Cieplak, W. D. Cornell and P. A. Kollman, *J. Phys. Chem.*, 1993, **97**, 10269–10280.
- 33 C. E. A. F. Schafmeister, W. S. Ross and V. Romanovski, LEAP, University of California, San Francisco, 1995.
- 34 T. A. Darden, D. M. York and L. Pedersen, *J. Chem. Phys.*, 1993, **100**, 10089–10092.
- 35 J. P. Ryckaert, G. Ciccotti and H. J. C. Berendsen, *J. Comput. Phys.*, 1977, **23**, 327–341.
- 36 C. Simmerling, R. Elber and J. Zhang, *Modelling of Biomolecular Structure and Mechanisms*, ed. A. Pullman, Kluwer, Netherlands, 1995, 241–265.
- 37 S. V. Evans, *J. Mol. Graphics*, 1993, **11**, 134–138.
- 38 A. Nicholls, K. A. Sharp and B. Honig, Delphi, Department of Biochemistry and Molecular Biophysics, Columbia University, NY, 1990.
- 39 D. Sitkoff, K. A. Sharp and B. Honig, *J. Phys. Chem.*, 1994, **98**, 1978–1988.
- 40 M. L. Connolly, *J. Appl. Crystallogr.*, 1983, **16**, 548–558.
- 41 T. E. Cheatham III, J. Srinivasan, D. A. Case and P. A. Kollman, *J. Biomol. Struct. Dyn.*, 1998, **16**, 265–280.