

PCCP

Accepted Manuscript



This is an *Accepted Manuscript*, which has been through the Royal Society of Chemistry peer review process and has been accepted for publication.

Accepted Manuscripts are published online shortly after acceptance, before technical editing, formatting and proof reading. Using this free service, authors can make their results available to the community, in citable form, before we publish the edited article. We will replace this *Accepted Manuscript* with the edited and formatted *Advance Article* as soon as it is available.

You can find more information about *Accepted Manuscripts* in the [Information for Authors](#).

Please note that technical editing may introduce minor changes to the text and/or graphics, which may alter content. The journal's standard [Terms & Conditions](#) and the [Ethical guidelines](#) still apply. In no event shall the Royal Society of Chemistry be held responsible for any errors or omissions in this *Accepted Manuscript* or any consequences arising from the use of any information it contains.



PCCP

PAPER

Molecular dynamics of the *Bacillus subtilis* expansin EXLX1: interaction with substrates and structural basis of the lack of activity of mutants

Received 00th January 2015,
Accepted 00th January 2015

DOI: 10.1039/x0xx00000x

www.rsc.org/pccp

Rodrigo L. Silveira^a and Munir S. Skaf*^a

Expansins are non-hydrolytic proteins that loosen growing plant cell walls and can enhance the enzymatic hydrolysis of cellulose. The canonical expansin structure consists of one domain responsible for substrate binding (D2) and another domain (D1) of unknown function, but essential for activity. Although the effects of expansins on cell walls and cellulose fibrils are known, the molecular mechanism underlying their biophysical function is poorly understood. Here, we use molecular dynamics simulations to gain insights into the mechanism of action of the *Bacillus subtilis* expansin BsEXLX1. We show that BsEXLX1 can slide on the hydrophobic surface of crystalline cellulose via the flat aromatic surface of its binding domain D2, comprised mainly of residues Trp125 and Trp126. Also, we observe that BsEXLX1 can hydrogen bond a free glucan chain in a twisted conformation and that the twisting is chiefly induced by means of residue Asp82 located on D1, which has been shown to be essential for expansin activity. These results suggest that BsEXLX1 could walk on the surface of cellulose and disrupt hydrogen bonds by twisting glucan chains. Simulations of the inactive BsEXLX1 mutants Asp82Asn and Tyr73Ala indicate structural alterations around the twisting center in the domain D1, which suggest a molecular basis for the lack of activity of these mutants and corroborate the idea that BsEXLX1 works by inducing twists on glucan chains. Moreover, simulations of the double mutant Asp82Asn/Tyr73Leu predict the recovery of the lost activity of BsEXLX1-Asp82Asn. Our results provide a dynamical view of the expansin-substrate interactions at the molecular scale and help shed light on the expansin mechanism.

1 Introduction

Plant cell walls are composed of cellulose microfibrils surrounded by a non-crystalline polymeric matrix containing hemicellulose, pectin, lignin, and proteins.¹ Such a polymeric matrix is arranged in a network that tethers the cellulose microfibrils together by noncovalent interactions.² This cell wall architecture provides the mechanical strength that enables plant cells to resist the turgor pressures originated in the cytoplasm. During the growth period, however, increases in the cellular volume require the concomitant extension of the cell walls.³ Depending on the plant species and type of tissue, cells can undergo a 30,000-fold volume enlargement before reaching maturity.^{4,5} As such, the cell wall architecture must somehow be modified in order to allow for plant growth.⁶

Expansins are loosening proteins that regulate cell wall extension in growing plants.^{4,5,7,8} According to genetic and functional studies, expansins induce extension and stress relaxation in plant cell walls, which are essential processes for plant growth.^{9,12} Expansins are also involved in other processes that require cell wall loosening, such as root hair initiation,¹³ seed germination,¹⁴

fruit softening^{15,16} and abscission.¹⁷ While growing tissues naturally express expansin genes, plant growth decreases upon expansin gene silencing.^{7,18,19} Moreover, addition of exogenous expansin to normal plants stimulates further growth, indicating that endogenous expansins may limit plant development.^{5,11,20}

Four types of expansins (α -expansin, β -expansin, expansin-like A, expansin-like B) occur in plants.²¹ Similar proteins are found in other organisms, such as pathogenic bacteria that colonize plant roots²²⁻²⁴ and fungi that degrade lignocellulosic biomass.²⁵⁻²⁷ These expansin-related proteins induce modifications in plant cell walls without detectable lytic activity and have been generally referred to as non-hydrolytic disruptive proteins.²⁸ Because of their biological role, these disruptive proteins enhance biomass degradation for biofuel production.^{29,30} Plant cell walls are recalcitrant to enzymatic degradation, in part because the cell wall nanoarchitecture limits the enzyme accessibility to the cellulosic substrate.^{1,31,32} When mixed with hydrolases, disruptive proteins enhance the enzymatic activity by increasing the substrate accessibility to the enzymes.^{23,26,27,28,29,33} It has been proposed that such proteins act by disrupting the hydrogen bond network within the cell wall structure,²⁸ but no direct evidence has been found so far.

The expansin-like protein BsEXLX1 from the soil bacteria *Bacillus subtilis* has received special attention recently.³⁰ BsEXLX1 has been employed in several studies aiming at unveiling the structure-function relationship of expansins, especially due to the difficulties in obtaining heterologous expression of plant expansins in bacteria.

^a Institute of Chemistry, University of Campinas, Campinas, Cx. P. 6154, 13084-862, SP, Brazil. E-mail: skaf@iqm.unicamp.br

† Electronic Supplementary Information (ESI) available. See DOI: 10.1039/x0xx00000x

BsEXLX1 exhibits plant cell wall extension qualitatively similar to β -expansins,²² has cellulose-weakening activity^{34,35} and enhances the cellulose degradation by cellulases.³⁴ Accordingly, BsEXLX1 exhibits the same two-domain structure of plant expansins, in which the domains D1 and D2 contact each other and are linked by a short linker.^{22,36} The D1 domain presents some structural similarity to family 45 glycosyl hydrolases,³⁷ but lacks catalytic activity. Its putative carbohydrate binding surface is populated by polar/charged residues important for activity on cell walls. The D2 domain, an immunoglobulin-like domain, has been recognized as a type-A carbohydrate binding module of the family 63 (CBM63).³⁵ As such, D2 displays a flat aromatic surface, suitable for the recognizing crystalline cellulose.^{22,35,36,38,39}

Functional aspects of each of the expansin domains have been addressed by measuring the activities displayed by the intact BsEXLX1 expansin, the individual domains D1 and D2 separately, and D1 and D2 simultaneously free in the solution. Results conclusively show that only the intact expansin is active, indicating that both expansin domains must be connected to yield expansinic activity.³⁵ Moreover, the D2 domain and the intact expansin display the same binding capacity to cellulose and to whole cell walls, demonstrating that substrate binding is D2's primary function.³⁵

In spite of the known biophysical effects of expansins on cellulose and whole plant cell walls, the molecular mechanism underlying their function remains elusive, especially because of the difficulties of addressing biophysical changes in complex systems at the nanoscale. In such contexts, molecular simulations have been successfully employed to obtain molecular-level insights into protein mechanisms, particularly of those involved in cell wall modification and deconstruction.⁴⁰ Here, we present molecular dynamics (MD) simulations of the bacterial expansin BsEXLX1 to investigate: (i) the interactions between BsEXLX1 and a cellulose elementary fibril; (ii) BsEXLX1 interactions with a single glucan chain; and (iii) the molecular basis underlying the loss of expansinic activity caused by the mutations Asp82Asn and Tyr73Ala, previously reported by Georgelis *et al.*³⁵ Our simulation results indicate that the D1 domain interacts with a free glucan chain in a twisted conformation while domain D2 can slide on the crystalline cellulose surface. The mutations Asp82Asn and Tyr73Ala cause structural perturbations that impair the twisting capability of the D1 domain. The results provide new insights into the loosening mechanism of expansins at the molecular level.

2 Computational Details

All simulations were performed with NAMD⁴¹ using CHARMM22 force field with the CMAP correction for the protein^{42,43} and the CHARMM36⁴⁴ for the carbohydrates. The TIP3P model was used for water molecules.^{42,45} The chemical bonds involving hydrogen atoms were constrained at their equilibrium values and a timestep of 2 fs was employed to numerically solve the equations of motion. The temperature was kept constant at 300 K by means of a Langevin thermostat and the pressure was kept at 1 atm with a Langevin piston, except in the BsEXLX1-cellulose fibril simulations in which the volume, instead of the pressure, was kept fixed. Analyses were carried out with in-house programs and the figures generated with VMD.⁴⁶

The initial coordinates of the BsEXLX1 expansin were taken from the crystallographic structure obtained by Kerff *et al.* (PDB code: 3D30).²² Hydrogen atoms were added according to the estimated pK_a values at pH=7.0 using the server H++.^{47,48} The neutral pH was chosen considering that BsEXLX1 is active at pH values ranging from

5.5 to 9.5.³⁵ At pH 7.0, residues His7 and His203 were considered as protonated and all other titratable residues were considered at the standard protonation state. The structure was then surrounded by a rectangular box of water molecules containing 0.15 M NaCl with excess counter-ions to render the system electrically neutral. The simulation box was built with Packmol⁴⁹ such that the water layer surrounding the protein was at least 16 Å thick. Periodic boundary conditions were imposed during simulations. Long-range interactions were treated using Particle Mesh Ewald⁵⁰ and short-range interactions were truncated at the cutoff radius of 12 Å. The structure was then submitted to the following steps: (i) with all the protein atoms fixed, 1000 steps of energy minimization via the conjugated gradients (CG) algorithm followed by 200 ps of MD simulations; (ii) with all the α -carbon fixed, 1000 steps of CG followed by 200 ps of MD simulation; (iii) 600 ps of MD with all atoms free. After these preparation steps, a 150 ns long MD simulation was performed for analysis.

Three configurations from this 150-ns-long MD simulation were randomly taken (as the configurations do not differ substantially) to set up the substrate-free systems for the mutants Asp82Asn, Tyr73Ala and Asp82Asn/Tyr73Leu. For the BsEXLX1-cellulose and BsEXLX1-glucan complexes, one of these configurations was used. The coordinates of the mutated residues were inserted according to the CHARMM22 internal coordinates. One sodium ion was removed from the Asp82Asn and Asp82Asn/Tyr73Leu systems to keep the electroneutrality. The systems were then submitted to 100 steps of CG followed by 10 ns of MD simulations. Afterwards, the trajectories were propagated by 150 ns, over which the analyses were performed.

For the MD simulations involving the cellulose microfibril, the 36-chain hexagonal model was employed.⁵¹ A cellulose fibril with 24-glucose-long chains was generated with Cellulose-Builder⁵² according to the crystallographic structure of cellulose I β .⁵³ For studying the adsorption process, the BsEXLX1 was placed close to the hydrophobic surface of the cellulose microfibril in an orientation in which the planar aromatic surface of the domain D2 was approximately parallel to the cellulose surface and 10 Å apart. A water layer of at least 20 Å from the protein or cellulose was created with Packmol.⁴⁹ Such configuration allowed the water molecules to occupy the interstitial region between the expansin and the cellulose surface. Before a 50-ns-long production run, the system was submitted to the following steps: (1) 1000 steps of CG and 500 ps of MD simulation with the expansin and cellulose fixed; (2) 1000 steps of CG and 2 ns of MD with the whole system free. Two adsorption simulations were carried out starting from step (1). To study the behavior of the expansin adsorbed on the cellulose surface, another system was built in which the expansin was initially in contact with the cellulose. This allowed the use of a thinner water layer to solvate the system in comparison with the previous system. A simulation lasting 700 ns was performed with this system after the same preparation steps (1) and (2) described for the adsorption simulations.

The initial structure of the BsEXLX1-glucan complex was obtained by taking the expansin coordinates and the glucan chain in contact with it in the BsEXLX1-cellulose crystallographic complex. After removing all water molecules, the glucan chain was submitted to 10,000 steps of CG and 200 ps of MD with the BsEXLX1 fixed. This allowed the glucan chain to adjust to both D1 and D2 expansin domains. The final configuration of such BsEXLX1-glucan complex was then solvated by a 16 Å-thick electrically neutral water layer containing 0.15 M NaCl. The same preparation procedure described for the substrate-free expansin was adopted before the production

runs. Initially, one simulation lasting 200 ns was carried out. In this simulation, the glucan chain remained bound to both domains for 100 ns, after which it started exploring conformations unbound to the domain D1 (but still bound to D2). The dissociation from D1 was expected due to its weak interactions with the substrate, as previously shown by binding experiments.³⁵ In order to increase the statistics of the bound state, three random configurations were taken from the BsEXLX1-substrate complex along the first 100 ns of the trajectory. New simulations starting from these configurations, with different initial velocities, were then performed. In these simulations, the glucan chain remained bound during the first 60, 38, and 32 ns. Analyses were performed over these three time windows.

3 Results

3.1 BsEXLX1 on crystalline cellulose: adsorption and translation

MD simulations were performed for a system comprised of a cellulose fibril and the intact BsEXLX1 dispersed in aqueous solution. The expansin was initially placed close (~ 10 Å) to the cellulose hydrophobic face and the system was allowed to evolve. Within the first 30 ns of simulation (after equilibration), the BsEXLX1 becomes fully adsorbed onto the cellulose surface. Fig. 1a shows the shortest distances between the cellulose face and the domains D1 and D2, as well as the system's configuration at times $t=0$, $t=10$ ns and $t>30$ ns. The adsorption starts with D2 approaching the cellulose surface, led by residue Trp125 (10 ns). As the other aligned aromatic residues of domain D2 (Trp125 and Tyr157) approach the cellulose surface, the D1 domain also starts contacting the cellulose due to the geometrical constraints imposed on D1 and D2 by the linker and by the D1–D2 interface. During this process, water is driven out of the interfacial region between the planar aromatic surface on D2 and cellulose (Fig. 1b). These results indicate that the interactions between the BsEXLX1 aromatic surface and the cellulose hydrophobic surface – followed by water expelling from the mutual interacting surfaces – constitute the driving force for BsEXLX1 adsorption onto cellulose.

The interaction energy between the cellulose fibril and the BsEXLX1 provides further information on the nature of the BsEXLX1–cellulose interactions. The BsEXLX1–cellulose total interaction energy decreases during the adsorption process due to a strongly favourable van der Waals attraction; meanwhile, the electrostatic component fluctuates around positive values (Fig. 1c, left panel). The main source of such van der Waals interactions comes from the D2 domain, rather than domain D1, which interacts with the cellulose mostly via electrostatic forces (Fig. 1c, center and right panels). Also, we observe that the total interaction energy between D2 and cellulose converges to vanishingly small values in the final stages of the simulation. This suggests the predominance of entropic stabilization factors associated with water expelling from the mutual interacting surfaces as the BsEXLX1 binds to cellulose. Together, these results reflect the hydrophobic nature of the BsEXLX1–cellulose interactions and the entropic nature of the adsorption driving force.

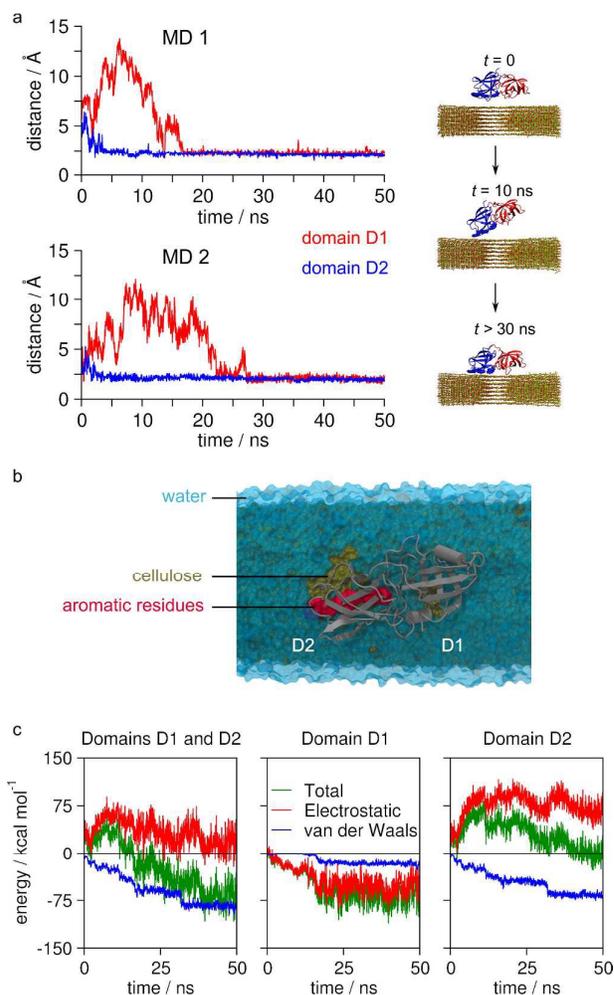


Fig. 1 BsEXLX1 adsorption onto the cellulose hydrophobic surface. (a) Minimum distance, computed from two independent simulations (MD1 and MD2), between BsEXLX1 domains D1 and D2 from the cellulose microfibril (in brown) surface during the adsorption process, with snapshots from the simulation. Interactions of the domain D2 with the cellulose surface mediate the BsEXLX1 adsorption. (b) Molecular surface of water molecules around the cellulose surface, showing that binding of BsEXLX1 causes the dehydration of the cellulose in contact with the domain D2. (c) Interaction energies between cellulose and the intact BsEXLX1 expansin (both domains D1 and D2; left), the domain D1 (middle) and domain D2 (right) during adsorption process in the MD1 run. The same analysis is provided for MD2 in Fig. S1a (ESI[†]). Decrease of the van der Waals interactions between the cellulose elementary fibril and the domain D2 follows the adsorption process.

After further extending one of the adsorption simulations to 700 ns, we observed that the BsEXLX1–cellulose interaction energies fluctuates around approximately constant values, indicating that the BsEXLX1 has reached an state close to equilibrium on the cellulose surface within the first 50 ns of the adsorption process (Fig. S1b, ESI[†]). During the 700-ns simulation, we observed that the BsEXLX1 can translate longitudinally back and forth over the cellulose surface (Fig. 2a). The glucan chains of the cellulose surface serve as tracks on which the aromatic surface of D2 can slide. This

suggests that the domain D2 plays roles beyond simply binding to cellulose and that the expansin slippage could be essential to the mechanism of cellulose loosening. Translation along the hydrophobic surfaces of cellulose has also been observed for another type-A carbohydrate binding module.⁵⁴⁻⁵⁶ Structural fluctuations of the BsEXLX1 (computed after structural alignment of the cellulose fibril) over the time frame between 100 and 200 ns of the trajectory (in which no translation event is observed) show that domain D1 interacts with cellulose less tightly than domain D2 (Fig. 2b), reinforcing the view that D2 is responsible for anchoring the BsEXLX1 onto the substrate.

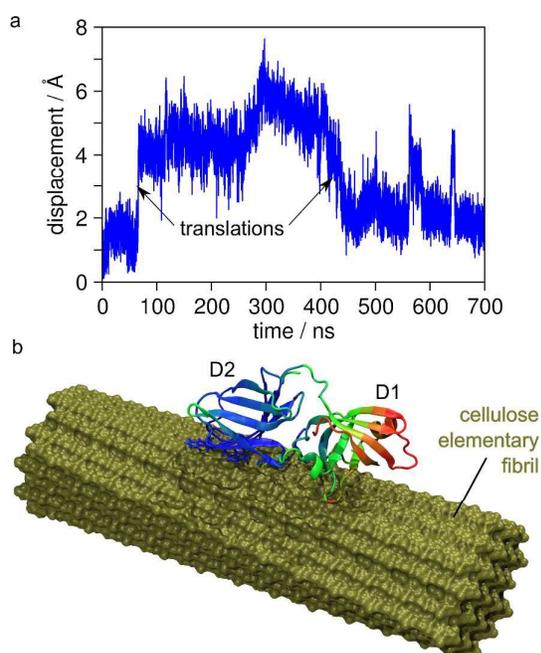


Fig. 2 BsEXLX1 on the cellulose hydrophobic surface. (a) Displacement of the α -carbon of residue Trp126 (located at the aromatic surface of the domain D2) from the beginning of the simulation. The time evolution of such displacement illustrates back and forth translations of the BsEXLX1 along the cellulose surface. (b) Average structure of BsEXLX1 on the cellulose surface over the time frame from 100 ns to 200 ns of simulation, during which the protein does not translate, showing that the domain D1 is less tightly bound to cellulose than the domain D2. The color scale represents average fluctuation ranging from values lesser than 1 Å (blue) to values greater than 3 Å (red). The protein fluctuations were computed after removing the translational and rotational motions of the cellulose fibril.

BsEXLX1 did not promote any significant changes in the cellulose structure within the time scale of these simulations. Similar result was obtained by Wang *et al.* in a 40-ns-long MD simulation of the BsEXLX1 expansin on cellulose.⁵⁷ This is not entirely unexpected, as the time scale of the expansin-induced effects on cell walls is of the order of seconds to hours.²²

3.2 Interactions between BsEXLX1 and a single glucan chain

The current hypothesis for the expansin mechanism is that it disrupts the hydrogen bonds within the cell wall structure.⁴ Since

the D1 domain possesses a number of polar/charged residues important for activity, it is reasonable to suppose that D1 bears the expansin disruption function. This view is supported by the experimental observation that loosenin, another non-hydrolytic expansin-like protein, lacks the domain D2 and yet exhibits disrupting activity.²⁷ To further comprehend the role of the domain D1, we simulated the BsEXLX1 expansin complexed to a free glucan chain. As the free glucan chain is not packed in the crystalline cellulose, the model gives access to the BsEXLX1-substrate interactions that would be established once the BsEXLX1 has reached a region of lower crystallinity in an actual cellulose fibril or overcome the interchain interactions within the crystalline cellulose structure. The simulations show that the glucan chain, while tightly bound to D2, binds the D1 domain weakly, as the chain is seen dissociating rather easily from D1 (Fig. 3a). This can be explained by the gain of conformational entropy when the chain leaves the domain D1 associated with the absence of enthalpic interactions strong enough to overcome the entropic penalty upon binding. On a solid matrix – such as a partially crystallized glucan chain on the cellulose surface – the entropic penalty would be reduced, and more structurally stable D1-substrate interactions could be established. We observe that the domain D2 holds the glucan chain in a planar conformation (Fig. 3b), in contrast with the slight twisted conformation around Tyr157 observed in the crystallographic structure of celohexaose-BsEXLX1 complex.⁵⁸ In the MD simulations, the thermal fluctuations can overcome the influence of the twisted Tyr157 on the glucan chain conformation, so that the chain can assume a non-twisted conformation along the domain D2 aromatic surface. These results are consistent with experimental results that show the cellulose binding function of D2³⁵ and corroborate the dependence of the D1 domain on D2 for its activity.

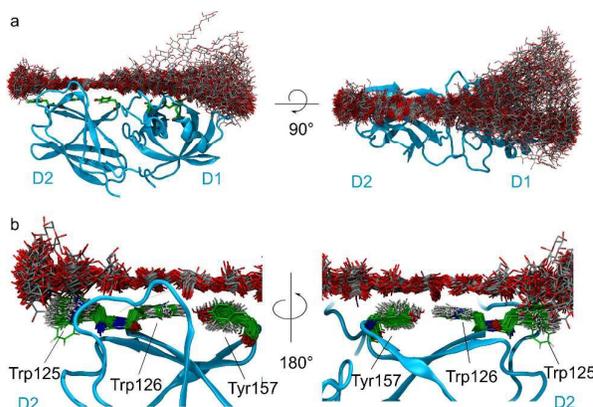


Fig. 3 (a) Glucan chain dynamics, showing that the domain D2 is able to hold the substrate even when it dissociates from the domain D1. The dissociation from the domain D1 reflects the weak thermodynamic force that this domain exerts on the substrate and illustrates the binding function of D2. (b) Glucan chain conformation along the domain D2 aromatic surface, with superposed configurations taken from the MD simulations. Even though the residue Tyr157 assumes a twisted conformation in relation to the planar surface of the residues Trp125 and Trp126, thermal fluctuations prevents the substrate from being twisted in the domain D2.

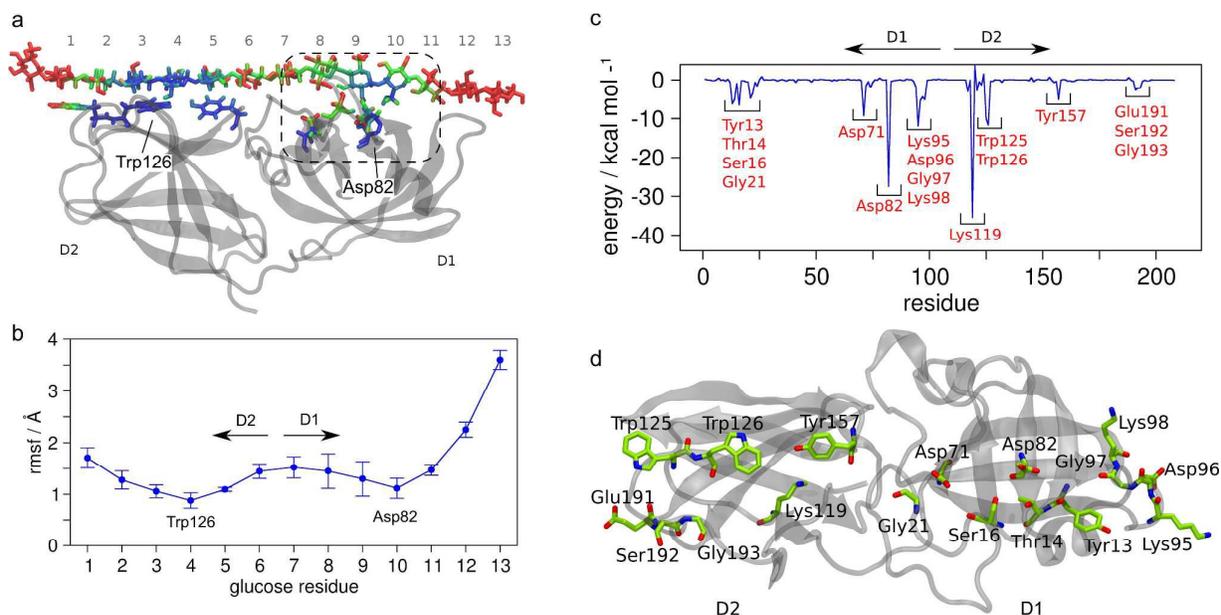


Fig. 4 Glucan chain bound to both D1 and D2 BsEXLX1 domains. (a) BsEXLX1 bound to a glucan chain colored according to its atomic mobilities obtained during the course of the simulations. From blue to red, the color scale represents average fluctuations from <1.5 Å to >3.0 Å. While the domain D2 holds the glucan chain in a planar conformation, the domain D1 induces a 90° twist in the chain (indicated inside the dashed box). (b) Average fluctuation (rmsf: root mean squared fluctuation) of the glucose residues along the glucan chain. Error bars indicate standard deviation from different simulations. The numbers of the glucose residues are indicated in the panel (a). Both the domain D1 and D2 bear residues able to stabilize the conformation of the glucan chain. (c) Average interaction energy between the glucan chain and each residue of BsEXLX1. (d) BsEXLX1 structure showing the residues that interact with the glucan chain, indicated in red in (c).

Analyses were performed for the configurations in which the glucan chain is bound to D1 (summing up a total of 230 ns of simulations), in order to examine the interactions between the substrate and the D1 domain. Fig. 4a shows a representative snapshot of the BsEXLX1 structure complexed to the glucan chain. From blue to red, the colors represent a scale of conformational fluctuation of the glucan chain measured as the atomic root mean squared fluctuations (rmsf), ranging from 1.5 to 3.0 Å. The chain mobility varies along the BsEXLX1 binding sites and reaches minimum values in two specific regions (Fig. 4a,b): around the Trp126 residue located in domain D2 and around the Asp82 residue in domain D1. Whereas the aromatic residues in D2 hold the chain in a planar conformation (Fig. 3b and Fig. 4a), several polar residues around Asp82 in domain D1 stabilize a twisted conformation of the chain.

Analysis of the interaction energies between each of the BsEXLX1 residues and the glucan chain shows the residues involved with the glucan chain complexation along the BsEXLX1 binding sites (Fig. 4c). In the domain D2, the aromatic residues Trp125, Trp126, and Tyr157 interact with the glucan chain via CH- π stacking, while residue Lys119 helps stabilizing such interactions with an additional hydrogen bond to the glucan chain. Consistently, it has been shown that mutation of these four residues by alanine results in impaired BsEXLX1 binding to crystalline cellulose.³⁵ D2 residues Glu191, Ser192, and Gly193 interact only weakly with the glucan chain because these residues are far from the aromatic binding sites where the glucan chain binds to (Fig. 4d). With the exception of Asp82, the residues in D1 involved in binding the glucan chain do so by means of hydrogen bonding and the average interaction energies are around -10 kcal/mol. Only Tyr13 does not hydrogen bond the glucan chain. Differently from these D1 residues, Asp82 interacts with the glucan chain with an average energy of -30

kcal/mol and, hence, seems to be pivotal to the D1 interaction with the glucan chain. In fact, mutation of this residue by alanine or asparagine completely inactivates BsEXLX1.³⁵

A closer look at the interactions between the glucan chain and the D1 domain indicates that residues Thr14, Ser16, Asp71, and Asp82 establish hydrogen bonds with the twisted glucan chain (Fig. 5a). The hydrogen bonding frequency between D1 and the glucan chain indicates that residues Thr14 and Asp82 play the most important role in the BsEXLX1–glucan interactions, as they are more persistently involved in hydrogen bonds than residues Ser16 and Asp71 (Fig. 5b). We also notice that Asp82 can interact with the glucan chain via two simultaneous hydrogen bonds. This is consistent with the mobility profile of the glucose units 8, 9, and 10 (Fig. 4b), since glucose units that hydrogen bond less persistently to D1 residues tend to exhibit higher mobility. Since residue Asp82 interacts more efficiently with the glucan chain and induce the highest structural stabilization in the twisted glucose unit, the simulation results suggest the existence of a twisting center located around the residue Asp82. The conformational stability of Asp82 is maintained with the assistance of a hydrogen bond with residue Thr12 (Fig. 5a). We advocate that such a D1-induced twisting of the glucan chain may underlie the mechanism by which BsEXLX1 weakens cellulose and loosens the primary cell wall.

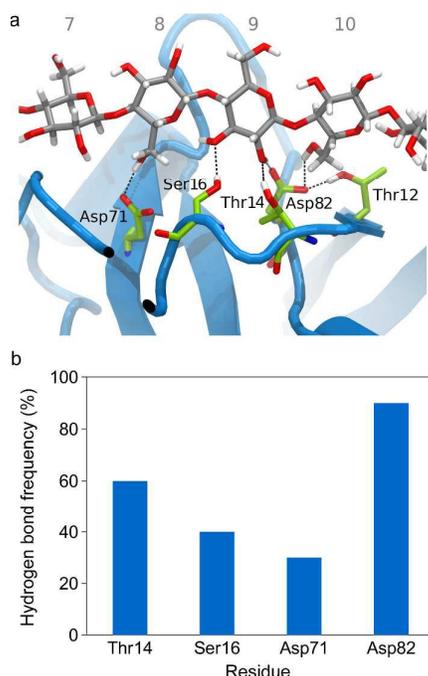


Fig. 5 Hydrogen bonding between the domain D1 and the glucan chain. (a) Close view of the interactions of the glucan chain with the residues in domain D1 responsible for the twist. The numbering of the glucose residues corresponds to the ones shown in the Fig. 4a. (b) Frequency of the hydrogen bonds between the glucan chain and the BsEXLX1 expansin. The residue Asp82 makes the most persistent hydrogen bond with the glucan chain.

3.3 The Asp82Asn mutant

According to the simulations, the residue Asp82 appears to be essential for the BsEXLX1 function by inducing a conformational twist in the glucan chain. It has been demonstrated that replacement of Asp82 by alanine causes a complete inactivation of the BsEXLX1 creep activity.³⁵ This result is likely to be extendable to the BsEXLX1 cellulose weakening activity as well, as the mutation Asp82Ala in the *Pectobacterium carotovorum* Ex1 expansin, whose domain D1 binding surface is nearly-identical to the BsEXLX1's (Fig. S2, ESI[†]), eliminates its cellulose disrupting activity.⁵⁹ The effect of replacing Asp82 by other polar residues capable of hydrogen bonding would not, in principle, affect BsEXLX1 activity. However, this is not the case; the mutant Asp82Asn does not exhibit any creep activity on cell walls.³⁵ At first, that seems counterintuitive, as the asparagine residue is also able to hydrogen bond glucose residues via its amide moiety. To understand the molecular basis for the lack of activity of the Asp82Asn mutant and further comprehend the role played by Asp82 in the BsEXLX1 activity, we carried out MD simulations with the substrate-free Asp82Asn BsEXLX1 mutant. Only minor changes in the α -carbon fluctuation

were detected for the mutant compared to the native BsEXLX1 [$\Delta(\text{rmsf}) < 0.6 \text{ \AA}$], indicating that the Asp82Asn mutation does not destabilize the BsEXLX1 tertiary structure (Fig. S3, ESI[†]).

Unlike residue Asp82 in the wild type BsEXLX1, the Asn82 residue in the mutant becomes unavailable for binding the substrate. By means of a mechanism that involves the neighbouring residue Tyr73, Asn82 regularly moves to a solvent-inaccessible region of the substrate-free BsEXLX1, thus preventing occasional contact with the substrate. Residues Asn82 and Tyr73 are seen to engage in a mutual equilibrium between two conformational states, hereafter denoted active and inactive states (Fig. 6a). In the active conformation, Asn82 and all other residues in its immediate vicinity, assume a configuration that closely resembles the one exhibited by the wild type BsEXLX1. In the inactive state, instead, residue Asn82 interacts with Tyr73, causing conformational changes that result in Asn82 ducking underneath the aromatic ring of Tyr73. Residue Thr14 seems to facilitate this process by making a hydrogen bond with Tyr73 (Fig. 6a). As a result, residue Asn82 in the inactive state becomes unavailable for binding the substrate, and residue Tyr73 obstructs the D1 binding site. Interactions between the amide and aromatic groups, like the Asn82-Tyr73 interaction in the inactive state, are common in proteins and have both dispersion and electrostatic origin.⁶⁰

The time evolution of the Asn82 structural fluctuation shows that the lifetime of the active and inactive states is of the order of several nanoseconds (Fig. 6b) and that the inactive state is slightly more likely than the active state by an approximately 60/40 population ratio (Fig. 6c). Thus, the conformational states interchange many times over the course of the simulation, in a timescale much lower than the reported timescale for expansin's action on the cell wall. Although the active state represents a considerable fraction of the accessible conformational states of key residues in D1, the simulations suggest it does not persist long enough to allow for expansin's activity. These results further corroborate the central role of the Asp82 residue in the expansin mechanism.

Replacing the Tyr73 residue by a bulky, non-aromatic residue that does not interact strongly with Asn82 could, in principle, prevent the BsEXLX1-Asp82Asn mutant from reaching the conformationally inactive state. To verify this possibility, we replaced the residue Tyr73 by Leu73 in the mutant Asp82Asn and performed simulations for such system. The results show that the accessibility of the residue Asn82 to an eventual binding to the substrate in the double mutant Asp82Asn/Tyr73Leu is similar to the accessibility of the residue Asp82 in the wild type BsEXLX1, as illustrated in Fig. 7. Moreover, except by increased local fluctuations around residue Asn82 relative to the native Asp82 (also observed for the single mutant), no major conformational changes were detected for the double mutant Asp82Asn/Tyr73Leu, as shown by the comparative analysis of the protein's rmsf (Fig. S3, ESI[†]). These results, thus, suggest that the loss of activity of the BsEXLX1-Asp82Asn mutant may be at least in part recovered by introducing the additional Tyr73Leu mutation.

PCCP

PAPER

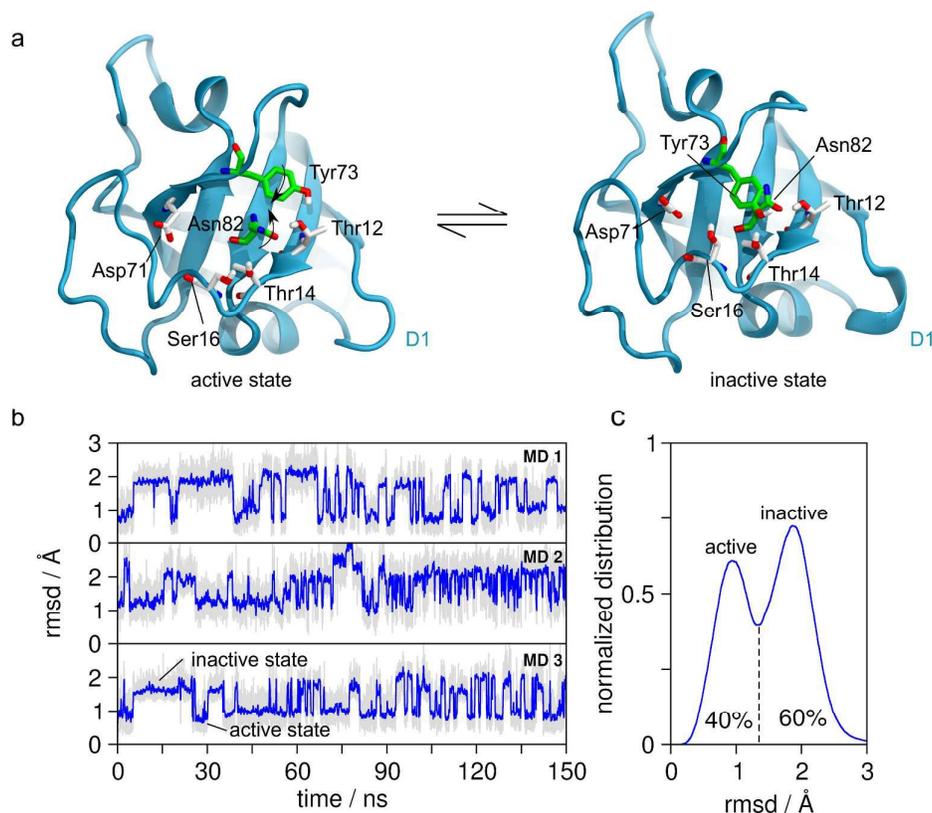


Fig. 6 Twisting center conformations in the Asp82Asn mutant. (a) Conformational equilibrium between the active and inactive states, in which the pair Asn82/Tyr73 can assume two distinct conformations in the domain D1. (b) Asn82 rmsd (root mean squared deviation) from the initial coordinates (after relaxation) in three independent simulations. The rmsd fluctuates around two values, corresponding to the active (rmsd ~ 1 Å) and inactive states (rmsd ~ 2 Å) of the mutant Asp82Asn. The blue curves represent running averages of the raw rmsd data (gray). (c) Normalized distribution of rmsd values for residue Asn82, showing the probabilities (integrated area) of the active (40%) and inactive states (60%).

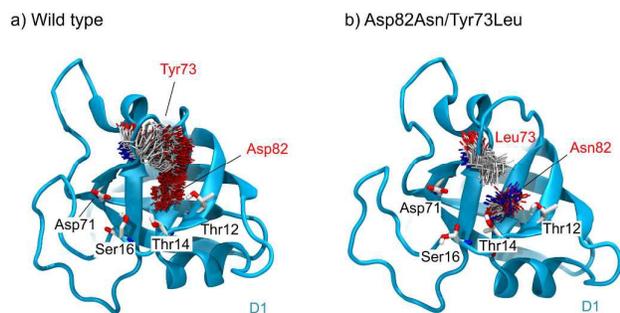


Fig. 7 Dynamic picture of the twisting center of the (a) native and (b) Asp82Asn/Tyr73Ala mutant BsEXLX1, showing superposed configurations of the residues 73 and 82 taken from three

independent simulations. In both cases, the residues Tyr73 and Leu73 do not bury the residues Asp82 and Asn82, respectively, as in the single mutant Asp82Asn (Fig. 6). Thus, the twisting center remains intact for interacting with the substrate in the double mutant Asp82Asn/Tyr73Leu, suggesting that the mutation Tyr73Leu may restore the activity of the Asp82Asn mutant.

3.4 The Tyr73Ala mutant

Site-directed mutagenesis shows that replacing the Tyr73 residue by an alanine yields an essentially inactive expansion, whereas replacement of Tyr73 by phenylalanine results in no change in activity.³⁵ Such mutagenesis assays reveal that Tyr73 plays a role in the BsEXLX1 activity via the tyrosine aromatic ring rather than involving its hydrogen bonding capability. In light of the idea of a

glucan chain twisting center pivoted by the Asp82 residue, it is unclear how the aromatic Tyr73 residue influences the BsEXLX1 function. In fact, no significant interaction between the free glucan chain and residue Tyr73 was observed during the course of our simulations (Fig. 4c). To understand the function of the Tyr73 residue and its relationship to the twisting mechanism previously described, we performed MD simulations of the Tyr73Ala mutant.

Simulation results reveal that the mutant BsEXLX1 lacks the structural stability around the twisting center necessary for effective protein-substrate interactions. Comparison of the dynamics of the wild type and Tyr73Ala mutant BsEXLX1 shows that the Tyr73Ala mutation perturbs the structural fluctuations of the D1 domain by increasing the protein mobility in several regions of this domain, as depicted in Fig. 8a. The mutation perturbs particularly the loop region near residue position 73, as shown by the color scale representation of residue mobility directly on the BsEXLX1 structure (Fig. 8b). Other regions of the D1 domain, distant from Ala73, which include functional residues such as Thr14 and Ser16, are also affected, indicating that the Tyr73Ala mutation induces non-local structural perturbations on D1.

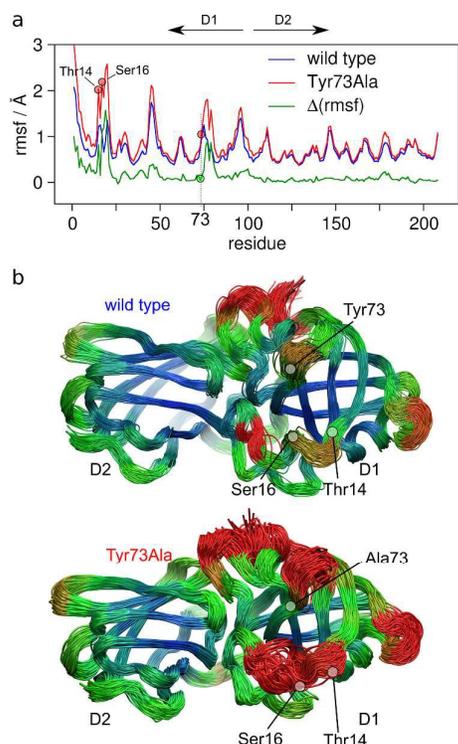


Fig. 8 Structural fluctuations of the expansin mutant Tyr73Ala. (a) Residue rmsf for wild type and mutant expansins, shown as the average over three independent simulations. (b) Expansin structure with the rmsf mapped in a color scale ranging from blue (rmsf < 0.3 Å) to red (rmsf > 1.5 Å). The mutation enhances structural fluctuations around the twisting center.

We also observe that the hydrogen bonding between residues Thr12 and Asp82 – strictly conserved residues among structurally related proteins, including plant expansins, GH45 endoglucanases and lytic transglycosylases²² – is regularly broken in the mutant BsEXLX1, but not in the wild type. As shown in Fig. 9, the Thr12-Asp82 distance distribution is predominantly peaked at the hydrogen bonding length of 1.8 Å in the wild type BsEXLX1 and

bimodal in the Tyr73Ala mutant, with peaks around 1.8 and 2.6 Å. These results indicate that residue Tyr73 participates only indirectly in the BsEXLX1 mechanism, as the role of Tyr73 seems essentially to preserve the twisting center in its active conformation. The results further support the idea that the twisting center is at the core of BsEXLX1 activity.

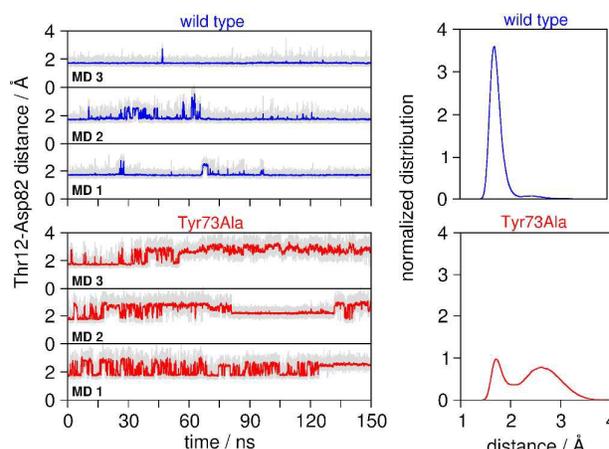


Fig. 9 Distances between residues Asp82 and Thr12 in three independent simulations (left panels) and their distributions (right panels) in the wild type and Tyr73Ala mutant BsEXLX1. The mutation causes the conserved hydrogen bond between these residues to break regularly.

The reason why residue 73 has to be either tyrosine or phenylalanine seems to be that these aromatic residues have bulky side chains that sterically prevent the D1 domain from accessing inactive states. The void created by replacing the tyrosine by an alanine gives room to protein motions of larger amplitude, allowing non-active states to be visited regularly. To illustrate this point, we have performed a principal component analysis⁶¹ of the wild type and Tyr73Ala BsEXLX1 dynamics. The structural fluctuations of both the wild type and mutant BsEXLX1 were decomposed into principal components and the results, depicted in Fig. 10, show that the first two modes capture the major differences in the fluctuations. This indicates that the effect of the mutation appears on the high amplitude motions of BsEXLX1 (Fig. 10a). Qualitative free energy landscape along these two principal components reveal that whereas the native BsEXLX1 explores a locally confined conformational space (Fig. 10b), the Tyr73Ala mutant is capable of exploring several other minima within the same time scale, indicating deviations from the wild-type conformations (Fig. 10c). Further analyses show that there is little superposition between the principal components from the wild type and Tyr73Ala BsEXLX1 simulations (Fig. S4 and Table S1, ESI[†]).

4 Discussion

Here, we report a set of MD simulations of the *Bacillus subtilis* expansin (BsEXLX1) to understand structural and dynamical aspects underlying the molecular basis of its activity. The results compare favorably with the experimental knowledge available to date and provide further insights into the mechanism of action of expansins. We observed that CH- π interactions between cellulose and the aromatic residues located in the D2 domain stabilize the BsEXLX1 adsorption onto the hydrophobic face of crystalline cellulose. The

binding involves the expulsion of water molecules from the contact interface to the bulk. This is consistent with calorimetric studies, which have shown that, similarly to another type-A carbohydrate binding module,⁶² binding of BsEXLX1 to cellulose is entropically driven,⁵⁸ likely owing to an increase of rotational and translational freedom of water molecules upon expelling from hydrophobic surfaces. Analysis of the interaction energy further corroborates the hydrophobic nature of the BsEXLX1-cellulose interactions, in which only the van der Waals component of the interaction energy arising from the D2 domain decreases during the adsorption process. The primary role of the van der Waals forces was also recently observed for interactions of cellulose and the type-A CBM1 of *Trichoderma reesei* cellobiohydrolase Cel7A.⁵⁶ The interactions between the D1 domain and cellulose are mainly electrostatic, suggesting that water competition could be the reason for D1's inability to bind cellulose by itself, as previously observed.³⁵

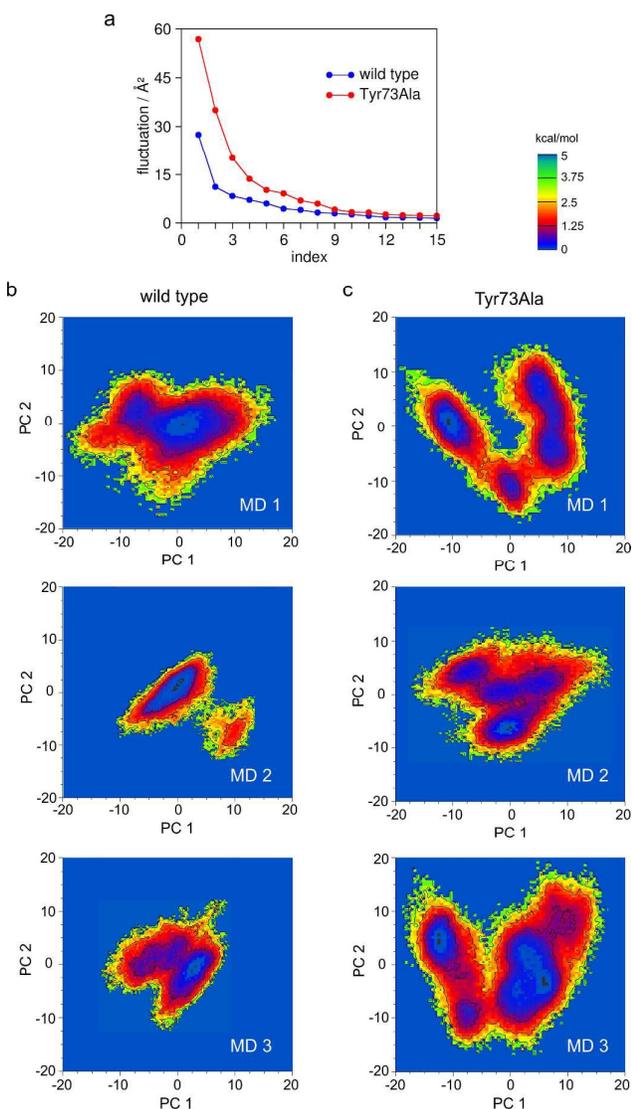


Fig. 10 Principal component analysis of BsEXLX1 fluctuations. (a) Squared fluctuations, shown as average over three independent simulations, along the first 15 principal modes and qualitative free energy landscape along the first (PC1) and second (PC2) principal components for three simulations (labeled MD1, MD2 and MD3) of

(b) wild type and (c) Tyr73Ala mutant. Most of the structural changes induced by the Tyr73Ala mutation are manifested on the high-amplitude motions of the BsEXLX1. Unlike the wild type, the mutant explores multiple substates around different free energy basins, away from the active substates explored by the native BsEXLX1.

While we observed no BsEXLX1-induced structural changes in the cellulose fibril, solid-state NMR measurements suggest that cellulose chains in close contact with BsEXLX1 exhibit different conformational properties in comparison with bulk cellulose,⁵⁷ indicating that BsEXLX1 induces changes in the cellulose structure in a timescale longer than that typically accessed by MD simulations. So, to further study the BsEXLX1-cellulose interactions, we simulated BsEXLX1 bound to a free glucan chain. Strikingly, we found that the D2 domain holds the glucan chain in planar conformation, whereas D1 induces a 90-degree twist in the chain. We advocate that such a twist is essential to the mechanism of BsEXLX1 activity on the cellulose surface. BsEXLX1 could induce local twists in single chains on the cellulose surface and, hence, disrupt cellulose interchain hydrogen bonds. That could explain the synergy observed between BsEXLX1 and cellulases acting on cellulose.³⁴ The ability of breaking such hydrogen bonds has long been proposed as the molecular function of expansins,⁹ but no experimental or computational evidence has been presented hitherto.

The observation that the BsEXLX1 is able to slide on the cellulose surface via the D2 domain combined with D1's capacity of twisting glucan chains, provides a compelling physical picture for the expansin mechanism by means of which the expansin would translate along cellulose causing local, unzipping twists on the polymer chains, thereby weakening the cell wall architecture. Similar ideas were previously suggested by Yennawar *et al.*³⁶ Thus, the D2 domain seems to have not only the binding function, but is also involved in activity. In fact, Georgelis *et al.* replaced the three aromatic residues of D2 involved in binding by alanine residues and detected no expansin activity on cell walls, even at high expansin concentrations that were used in order to compensate for its low substrate affinity.³⁵

The twisting function of the D1 domain is chiefly mediated by four residues: Thr14, Ser16, Asp71, and Asp82. Replacing any of them by alanine reduces the expansin activity in cell walls.³⁵ Among these residues, Asp82 shows highest capacity of hydrogen bonding to the substrate, which hinges on its ability to hold the twisted chain more effectively. Thus, residue Asp82 acts as the BsEXLX1's glucan chain twisting center and seems to be key to BsEXLX1's molecular activity. Interestingly, replacing residues Thr14, Ser16, or Asp71 by alanine results in a significant decrease in the BsEXLX1 activity, but some residual activity remains, whereas substituting residue Asp82 by either alanine or asparagine completely inactivates the BsEXLX1.³⁵ These observations correlate with the stability of the hydrogen bonds these residues make with the glucan chain: Asp82 engages in one or two hydrogen bonds with the substrate in approximately 90% of the simulation time, whereas the other three residues are not that effective at hydrogen bonding to the substrate. Our results support the conclusion that residues Thr14, Ser16, and Asp71 are important to the BsEXLX1 activity, but only residue Asp82 is essential. Based on our simulation results for the Asp82Asn mutant, we advocate that the twisting of the glucan chain would not occur in the absence of the residue Asp82.

If residue Asp82 plays a key role in effectively making hydrogen bonds with the substrate, as the simulations suggest, then it is

trivial to understand why its replacement by alanine inactivates the BsEXLX1. Less obvious is the fact that the Asp82Asn substitution also inactivates the expansin, since both residues bear hydrogen bonding capabilities. During the course of the simulations we observe a stacking of the aromatic ring of Tyr73 over Asn82 planar amide group, which suggests that the lack of activity of the Asp82Asn mutant is due to conformational changes that bury residue Asn82 under the side chain of Tyr73, rendering Asn82 unavailable to the substrate. This hypothesis was addressed by coupling the Asp82Asn mutation with the bulky, non-aromatic mutation Tyr73Leu. MD simulation of the double Asp82Asn/Tyr73Leu mutant showed that the absence of the aromatic residue close to Asn82 prevents such residue from being hidden by stacking interactions, suggesting that the BsEXLX1 activity could be at least partially recovered by incorporating the Tyr73Leu mutation into the Asp82Asn mutant. These predictions could be verified experimentally and, if confirmed, it would imply the non-catalytic role of residue Asp82 in BsEXLX1 and its importance in breaking hydrogen bonds within cell walls.

According to our simulations of BsEXLX1-glucan complex, residue Tyr73 does not participate directly in the interactions between the sugar and the D1 domain. Nevertheless, the replacement of such residue by alanine drastically impairs the BsEXLX1 activity on cell walls.³⁵ At first, it is unclear why Tyr73 is needed for the BsEXLX1 activity, since this residue is not associated with the glucan twisting. The simulations of the BsEXLX1 Tyr73Ala mutant suggest that the lack of activity is related to perturbations of the dynamics of functional residues in domain D1. The mobility of the region around the twisting center increases in the Tyr73Ala mutant, suggesting that the stability of the interactions between D1 residues and the substrate would be markedly affected. The reason for this structural perturbation is the extra space freed by inserting the alanine in place of the bulky Tyr73, which allows the protein to span inactive conformational states more frequently. This steric role of Tyr73 is supported by the fact that the mutant Tyr73Phe works as effectively as the native BsEXLX1.³⁵ Therefore, the structural integrity of the loops of the D1 domain seems to be essential for the BsEXLX1 activity.

5 Conclusions

Expansins stimulate cell wall extension and weaken cellulose structure by means of a mechanism that is not fully understood. Using molecular dynamics simulations, we have analyzed the interactions between the *Bacillus subtilis* expansin BsEXLX1 and cellulose in detail and investigated the dynamics of the system. While the BsEXLX1 was found to slip on the cellulose surface via the aromatic planar surface formed by residues Trp124 and Trp125 on domain D2, interactions of the protein with a single glucan chain indicated an BsEXLX1-induced twisting of the glucan chain carried out by domain D1. These results suggest that the cellulose-weakening effects could arise from expansin-induced chain conformational perturbations across the cellulose surface, which in turn break cellulose-cellulose hydrogen bonds. The results also showed that the twisting function is primarily mediated by residue Asp82 on D1. The experimentally observed lack of activity for the mutants Asp82Asn and Tyr73Ala were here associated with structural perturbations around the twisting center of BsEXLX1. Our simulations also predict that the double Asp82Asn/Tyr73Leu BsEXLX1 mutant would exhibit expansin activity. Overall, such results provide molecular-level insights into the BsEXLX1 mechanism with implications for the comprehension of the cell wall

fundamentals as well as the synergistic effects observed when expansins are mixed with cellulases. Fully unveiling the expansin mechanism would allow the incorporation of cellulose-disrupting functions into engineered cellulases.

Acknowledgements

This work was supported by the São Paulo Research Foundation (grants 2014/10448-1 and 2013/08293-7).

References

- 1 S. P. S. Chundawat, G. T. Beckman, M. E. Himmel and B. E. Dale, *Annu. Rev. Chem. Biomol. Eng.*, 2011, **2**, 121–145.
- 2 K. Keegstra, *Plant Physiol.*, 2010, **154**, 483–486.
- 3 S. Wolf, K. Hématy and H. Höfte, *Annu. Rev. Plant Biol.*, 2012, **63**, 381–407.
- 4 D. J. Cosgrove, *Nature*, 2000, **407**, 321–326.
- 5 D. J. Cosgrove, *Nat. Rev. Mol. Cell Biol.*, 2005, **6**, 850–861.
- 6 F. Marga, M. Grandbois, D. J. Cosgrove and T. I. Baskin, *Plant J.*, 2005, **43**, 181–190.
- 7 D. Choi, Y. Lee, H. T. Cho and H. Kende, *Plant Cell*, 2003, **15**, 1386–1398.
- 8 D. J. Cosgrove, *Curr. Opin. Plant Biol.*, 2015, **25**, 162–172.
- 9 S. J. McQueen-Mason and D. J. Cosgrove, *Proc. Natl. Acad. Sci. USA*, 1994, **91**, 6574–6578.
- 10 S. J. McQueen-Mason and D. J. Cosgrove, *Plant Physiol.*, 1995, **107**, 87–100.
- 11 D. J. Cosgrove, *Plant Cell*, 1997, **9**, 1031–1041.
- 12 S. E. C. Whitney, M. J. Gidley and S. J. McQueen-Mason, *Plant J.*, 2000, **22**, 327–334.
- 13 H. T. Cho and D. J. Cosgrove, *Plant Cell*, 2000, **14**, 3227–3253.
- 14 F. Chen, P. Dahal and K. J. Bradford, *Plant Physiol.*, 2001, **127**, 928–936.
- 15 D. A. Brummell, M. H. Harpster, P. M. Civallo, J. M. Palys, A. B. Bennett and P. Dunsmuir, *Plant Cell*, 1999, **11**, 2203–2216.
- 16 P. M. Civallo, A. L. T. Powell, A. Sabehat and A. B. Bennett, *Plant Physiol.*, 1999, **121**, 1273–1279.
- 17 E. J. Belfield, B. Ruperti, J. A. Roberts and S. J. McQueen-Mason, *J. Exp. Bot.*, 2005, **56**, 817–823.
- 18 H. T. Cho and D. J. Cosgrove, *Proc. Natl. Acad. Sci. USA*, 2000, **97**, 9783–9788.
- 19 S. Pien, J. Wyrzykowska, S. J. McQueen-Mason, C. Smart and A. Fleming, *Proc. Natl. Acad. Sci. USA*, 2001, **98**, 11812–11817.
- 20 A. J. Fleming, S. J. McQueen-Mason, T. Mandel and C. Kuhlemeier, *Science*, 1997, **276**, 1415–1418.
- 21 J. Sampedro and D. J. Cosgrove, *Genome Biol.*, 2005, **6**, 242.
- 22 F. Kerff, A. Amoroso, R. Herman, E. Sauvage, S. Petrella, P. Filée, P. Charlier, B. Joris, A. Tabuchi, N. Nikolaidis and D. J. Cosgrove, *Proc. Natl. Acad. Sci. USA*, 2008, **105**, 16876–16881.
- 23 H. J. Lee, S. Lee, H. Ko, K. H. Kim and I. Choi, *Mol. Cells*, 2010, **29**, 379–385.
- 24 N. Georgelis N. Nikolaidis and D. J. Cosgrove, *Carb. Polymers*, 2014, **100**, 17–23.
- 25 M. Saloheimo, M. Paloheimo, S. Hakola, J. Pere, B. Swanson, E. Nyyssonen, A. Bhatia, M. Ward and M. Penttilä, *Eur. J. Biochem.*, 2002, **269**, 4202–4211.
- 26 G. Jäger, M. Girfoglio, F. Dollo, R. Rinaldi, H. Bongard, U. Commandeur, R. Fischer, A. C. Spiess and J. Büchs, *Biotechnol. Biofuels*, 2011, **4**, 33.
- 27 R. E. Quiroz-Castañeda, C. Martínez-Anaya, L. I. Cuervo-Soto, L. Segovia and J. L. Folch-Mallol, *Microb. Cell Fact.*, 2011, **10**, 8.

- 28 K. Gourlay, V. Arantes, J. N. Saddler, *Biotechnol. Biofuels*, 2012, **5**, 51.
- 29 V. Arantes and J. N. Saddler, *Biotechnol. Biofuels*, 2010, **3**, 4.
- 30 N. Georgelis, N. Nikolaidis and D. J. Cosgrove, *Appl. Microbiol. Biotechnol.*, 2015, **99**, 3807–3823.
- 31 M. E. Himmel, S. -Y. Ding, D. K. Johnson, W. S. Adney, M. R. Nimlos, J. W. Brady and T. D. Foust, *Science*, 2007, **315**, 804–807.
- 32 S. -Y. Ding, Y. -S Liu, Y. Zeng, M. E. Himmel, J. O. Baker and E. A. Bayer, *Science*, 2012, **338**, 1055–1060.
- 33 K. Gourlay, J. Hu, V. Arantes, M. Andberg, M. Saloheimo, M. Penttilä and J. Saddler, *Biores. Technol.*, 2013, **142**, 498–503.
- 34 E. S. Kim, H. J. Lee, W. -G. Bang, I. -G. Choi and K. H. Kim, *Biotechnol. Bioeng.*, 2009, **102**, 1342–1353.
- 35 N. Georgelis, A. Tabuchi, N. Nikolaidis and D. J. Cosgrove, *J. Biol. Chem.*, 2011, **286**, 16814–16823.
- 36 N. H. Yennawar, L. -C. Li, D. M. Dudzinski, A. Tabuchi and D. J. Cosgrove, *Proc. Natl. Acad. Sci. USA*, 2006, **103**, 14664–14671.
- 37 G. J. Davies, G. G. Dodson, R. E. Hubbard, S. P. Tolley, Z. Dauter, K. S. Wilson, C. Hjort, J. M. Mikkelsen, G. Rasmussen and M. Schuelein *Nature*, 1993, **365**, 362–364.
- 38 A. B. Boraston, D. N. Bolam, H. J. Gilbert and G. J. Davies, *Biochem. J.*, 2004, **382**, 769–781.
- 39 I. J. Kim, H. J. Ko, T. W. Kim, I. G. Choi and K. H. Kim, *Biotechnol. Bioeng.*, 2013, **110**, 401–407.
- 40 G. T. Beckham, Y. J. Bomble, E. A. Bayer, M. E. Himmel and M. F. Crowley, *Curr. Opin. Biotechnol.*, 2011, **22**, 231–238.
- 41 J. C. Phillips, R. Braun, W. Wang, J. Gumbart, E. Tajkhorshid, E. Villa, C. Chipot, R. D. Skeel, L. Kalé and K. Schulten, *J. Comput. Chem.*, 2005, **26**, 1781–1802.
- 42 A. D. MacKerell Jr., D. Bashford, M. Bellott, R. L. Dunbrack Jr., J. D. Evanseck, M. J. Field, S. Fischer, J. Gao, H. Guo, S. Ha, D. Joseph-McCarthy, L. Kuchnir, K. Kuczera, F. T. K. Lau, C. Mattos, S. Michnick, T. Ngo, D. T. Nguyen, B. Prodhom, W. E. Reiher III, B. Roux, M. Schlenkrich, J. C. Smith, R. Stote, J. Straub, M. Watanabe, J. Wiorcikiewicz-Kuczera, D. Yin and M. Karplus, *J. Phys. Chem. B*, 1998, **102**, 3586–3616.
- 43 A. D. MacKerell Jr., M. Feig and C. L. Brooks III, *J. Comput. Chem.*, 2004, **25**, 1400–1415.
- 44 O. Guvench, E. Hatcher, R. M. Venable, R. W. Pastor and A. D. MacKerell Jr., *J. Chem. Theory Comput.*, 2009, **5**, 2353–2370.
- 45 W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey and M. L. Klein, *J. Chem. Phys.*, 1983, **79**, 926–935.
- 46 W. Humphrey, A. Dalke and K. Schulten, *J. Mol. Graph.*, 1996, **14**, 33–38.
- 47 J. C. Gordon, J. B. Myers, T. Folta, V. Shoja, L. S. Heath and A. Onufriev, *Nucleic Acids Res.*, 2005, **33**, W368–W371.
- 48 J. Myers, G. Grothaus, S. Narayanan and A. Onufriev, *Proteins*, 2006, **63**, 928–938.
- 49 L. Martínez, R. Andrade, E. G. Birgin and J. M. Martínez, *J. Comput. Chem.*, 2009, **30**, 2157–2164.
- 50 P. Darden, D. York and L. Pedersen *J. Chem. Phys.*, 1993, **98**, 10089–10092.
- 51 S. -Y. Ding and M. E. Himmel, *J. Agric. Food Chem.*, 2006, **54**, 597–606.
- 52 T. C. F. Gomes and M. S. Skaf, *J. Comput. Chem.*, 2012, **33**, 1338–1346.
- 53 Y. Nishiyama, P. Langan and H. Chanzy, *J. Am. Chem. Soc.*, 2002, **124**, 9074–9082.
- 54 J. Wohler and L. A. Berglund, *J. Chem. Theory Comput.*, 2011, **7**, 753–760.
- 55 M. R. Nimlos, G. T. Beckham, J. F. Matthews, L. Bu, M. E. Himmel and M. F. Crowley, *J. Biol. Chem.*, 2012, **287**, 20603–20612.
- 56 E. M. Alekozai, P. L. GhattyVenkataKrishna, E. C. Uberbacher, M. F. Crowley, J. C. Smith and X. Cheng, *Cellulose*, 2014, **21**, 951–971.
- 57 T. Wang, Y. B. Park, M. A. Caporini, M. Rosay, L. Zhong, D. J. Cosgrove and M. Hong, *Proc. Natl. Acad. Sci. USA*, 2013, **110**, 16444–16449.
- 58 N. Georgelis, N. H. Yennawar and D. J. Cosgrove, *Proc. Natl. Acad. Sci. USA*, 2012, **109**, 14830–14935.
- 59 M. Olarte-Lozano, M. A. Nuñez-Mendoza, N. Pastor, L. Segovia, J. Folch-Mallol and C. Martínez-Anaya, *Plos One*, 2014, **9**, e95638.
- 60 G. Duan, V. H. Smith Jr. and D. F. Weaver, *J. Phys. Chem. A*, 2000, **104**, 4521–4532.
- 61 A. Amadei, A. B. M. Linssen and H. J. C. Berendsen, *Proteins*, 1993, **17**, 412–425.
- 62 A. L. Creagh, E. Ong, E. Jervis, D. G. Kilburn and C. A. Haynes, *Proc. Natl. Acad. Sci. USA*, 1996, **93**, 12229–12234.