Biochimica et Biophysica Acta 1834 (2013) 1492–1500

Contents lists available at SciVerse ScienceDirect

Biochimica et Biophysica Acta

journal homepage: www.elsevier.com/locate/bbapap

Assembling a xylanase–lichenase chimera through all-atom molecular dynamics simulations

Junio Cota [a,d,1], Leandro C. Oliveira [a,b,1], André R.L. Damásio [a], Ana P. Citadini [a], Zaira B. Hoffmam [a], Thabata M. Alvarez [a], Carla A. Codima [a], Vitor B.P. Leite [b], Glaucia Pastore [c], Mario de Oliveira-Neto [d], Mario T. Murakami [e], Roberto Ruller [a], Fabio M. Squina [a,*]

[a] Laboratório Nacional de Ciência e Tecnologia do Bioetanol — CTBE/CNPEM, Campinas, SP, Brazil
[b] Departamento de Física, IBILCE, Universidade Estadual Paulista - UNESP, São José do Rio Preto, SP, Brazil
[c] Faculdade de Engenharia de Alimentos, Universidade Estadual de Campinas, Campinas, SP, Brazil
[d] Departamento de Física e Biofísica, Instituto de Biociências, UNESP, Botucatu, São Paulo, Brazil
[e] Laboratório Nacional de Biociências — LNBio/CNPEM, Campinas, SP, Brazil

ARTICLE INFO

Article history:
Received 8 October 2012
Received in revised form 23 January 2013
Accepted 20 February 2013
Available online 28 February 2013

Keywords:
Multifunctional enzyme
Small-angle X-ray scattering
Molecular dynamics
Computational characterization
Experimental validation

ABSTRACT

Multifunctional enzyme engineering can improve enzyme cocktails for emerging biofuel technology. Molecular dynamics through structure-based models (SB) is an effective tool for assessing the tridimensional arrangement of chimeric enzymes as well as for inferring the functional practicability before experimental validation. This study describes the computational design of a bifunctional xylanase–lichenase chimera (XylLich) using the *xynA* and *bglS* genes from *Bacillus subtilis*. In silico analysis of the average solvent accessible surface area (SAS) and the root mean square fluctuation (RMSF) predicted a fully functional chimera, with minor fluctuations and variations along the polypeptide chains. Afterwards, the chimeric enzyme was built by fusing the *xynA* and *bglS* genes. XylLich was evaluated through small-angle X-ray scattering (SAXS) experiments, resulting in scattering curves with a very accurate fit to the theoretical protein model. The chimera preserved the biochemical characteristics of the parental enzymes, with the exception of a slight variation in the temperature of operation and the catalytic efficiency (k_{cat}/K_m). The absence of substantial shifts in the catalytic mode of operation was also verified. Furthermore, the production of chimeric enzymes could be more profitable than producing a single enzyme separately, based on comparing the recombinant protein production yield and the hydrolytic activity achieved for XylLich with that of the parental enzymes.

1. Introduction

Plant biomass saccharification and biofuel production have been described as promising renewable alternatives to petroleum and natural gas. However, the polysaccharide network in plant cell walls is one of the most complex structures in nature, which jeopardizes the production of biofuels from plant biomass. First, biomass feedstock must go through a recalcitrance-reducing step (pretreatment). Then, enzymatic cocktails are used to precisely breakdown the polysaccharides into simple sugars suitable for several bioprocesses, such as fermentation to ethanol [1].

The enzymatic cocktail for plant biomass saccharification and biofuel production must include cellulolytic hydrolases, such as cellobiohydrolases, endo-glucanases and β-glucosidases. For hemicellulose degradation, synergistic action by hydrolytic enzymes is required at the polysaccharide backbone, side chains and decorating units. For instance, the hydrolysis of feedstock containing arabinoxylan requires several hydrolytic enzymes, such as endo-xylanases, β-xylosidases, arabinofuranosidases, ferulic acid esterases, glucuronidases and other enzymes [2]. β-glucans, polysaccharides with β-1,3 and β-1,4 glucosidic linkages, are also abundant in many plant cell walls, especially in sugarcane [3]. The enzymatic depolymerization of 1,3-1,4-β-glucans is catalyzed by 1,4-β-D-glucan 4-glucanohydrolase (EC 3.2.1.4), 1,3-β-D-glucan 3-glucanohydrolase (EC 3.2.1.39) and 1,3-1,4-β-D-4-glucanohydrolase or lichenase (EC 3.2.1.73) [4].

The engineering of multifunctional proteins with a synergistic catalytic capacity has the potential to streamline biomass conversion strategies [5]. However, the unsupervised construction of enzyme

* Corresponding author at: Laboratório Nacional de Ciência e Tecnologia do Bioetanol — CTBE/CNPEM, Caixa Postal 6170, 13083-970 Campinas, São Paulo, Brazil. Tel.: +55 19 3518 3111; fax: +55 19 3518 3104.
E-mail address: fabio.squina@bioetanol.org.br (F.M. Squina).
[1] These authors contributed equally to this work.

1570-9639
http://dx.doi.org/10.1016/j.bbapap.2013.02.030

fusions can result in nonfunctional chimeras because of misfolding and catalytic restriction [6]. Catalytic modules connected through a linker peptide are widely used because this process allows inter-domain flexibility and usually retains the original wild-type functionality [7]. Numerous linkers have been described for protein fusion, including AAA [8], GGGG [9,10], HHHHHH [11] and (GGGGS)4 [12].

An important issue regarding the construction of a protein chimera is how to determine the structural organization of the domains in a simple and reliable manner. Currently, small-angle X-ray scattering (SAXS) has become a central tool in structural biology for characterizing proteins in solution. Models using flexible regions and studied with normal mode analysis [13], molecular dynamics (MD) [14–16] or Monte Carlo simulations [17] have provided not only successful data validation but also accurate fitting of the scattering profile because of the potential to explore the protein conformation in space. Motivated by low computational costs, high control of the energetic parameters and good agreement with experiments, the models based on the energy landscape theory [18] have been extensively employed in several molecular systems, including protein folding, conformational changes and dynamic molecular machines [18–24]. The combination of simulations and SAXS can also provide information about the dynamic equilibrium of proteins in solution [25], for instance, conformational changes of *holo* and *apo* protein states and their correlation to regulatory mechanisms [26].

The development of computational approaches for predicting chimeric behavior, such as the stability and arrangement of the domains in solution, can contribute to the development of automated searching pipelines for optimal linkers and enzyme modules. In this paper, we described a computational approach based on energy landscape theory for designing a bifunctional enzyme containing the endo-xylanase and lichenase catalytic modules from *Bacillus subtilis*: XynA and BglS, respectively. The properties related to the structure of the chimera, substrate accessibility to the active site and dynamical behavior in solution were first calculated through computational tools. Then, we designed the multifunctional enzyme (XylLich), which was comprehensively evaluated and validated both biochemically and structurally (SAXS). The effectiveness of the chimera in hydrolyzing beechwood xylan and lichenan polysaccharide composites was also evaluated. Furthermore, based on the recombinant protein production yield and hydrolytic activity achieved for XylLich compared with the parental enzymes, producing chimeric enzymes could be more advantageous than producing single enzymes separately.

2. Material and methods

2.1. Chimera construction for simulations

The chimera was constructed using the endo-xylanase (XynA; PDB ID: 1xxn) and endo-β-1,3-1,4-glucanase (BglS; PDB ID: 3o5s) domains connected by a 4-glycine linker. The chosen linker, which has been studied previously [9], is simple and guarantees reasonable separation between the monomers. The linker was modeled using the MOLMOL program [27], where the angles Π (Phi) and ψ (Psi) were defined as −100° and 120°, respectively, which is an allowed region of the Ramachandran plot. The obtained structure was solvated (water TIP-3), and 5000 steps of energy minimization using the default conjugated gradient were carried out using NAMD 2.9 [28] with the CHARMM 2.2 force field [29]. The conformation employed for further simulations was obtained after removing the water.

2.2. Molecular dynamics simulations of flexible models

The SMOG web server [30] was used to generate the force field for an all-atoms model. The total energy (V) of the system for a given conformation Γ is calculated relative to the native state Γ_0 by the following equation:

$$\begin{aligned} V(\Gamma,\Gamma_0) = & \sum_{bond} \varepsilon_r (r-r_0)^2 + \sum_{angle} \varepsilon_\theta (\theta-\theta_0)^2 + \sum_{planar/improper} \varepsilon_\nu (\nu-\nu_0)^2 \\ & + \sum_{backbone} \varepsilon_{BB} \left\{ [1-\cos(\phi-\phi_0)] + \frac{1}{2}[1-\cos(3(\phi-\phi_0))] \right\} \\ & + \sum_{sidechain} \varepsilon_{SC} \left\{ [1-\cos(\phi-\phi_0)] + \frac{1}{2}[1-\cos(3(\phi-\phi_0))] \right\} \\ & + \sum_{contact} \varepsilon_c \left[\left(\frac{\delta_{ij}}{r_{ij}}\right)^{12} - 2\left(\frac{\delta_{ij}}{r_{ij}}\right)^{6} \right] + \sum_{non-contact} \varepsilon_{NC} \left(\frac{\delta_{NC}}{r_{ij}}\right)^{12} \end{aligned} \quad (1)$$

The values for r_0, θ_0, ν_0, Π_0 and δ_{ij} were taken from the conformation obtained in the previous step (see the Chimera construction for simulations section) and corresponded to the distance between the covalent bonds, the angle between three consecutive connected atoms, the improper/planar dihedral angles, the dihedral angles and the pair distance between the pair of atoms i and j, respectively, for all cases in the initial state. r_{ij} is the pair distance in the conformation Γ. The energetic terms were given as a function of the contact energy $\varepsilon_c = 1$ kT, where $\varepsilon_r = 100\varepsilon_c/\text{Å}^2$, $\varepsilon_\theta = 20\varepsilon_c/\text{rad}^2$, $\varepsilon_5 = 10\varepsilon_c/\text{rad}^2$ and $\varepsilon_{NC} = 0.01\varepsilon_c$. The map of interactions between the atoms was calculated using the "Shadow algorithm" [31]. The excluded volume term was defined by non-contact pairs with δ_{NC} defined as 1.25 Å. ε_c, ε_{BB} and ε_{SC} were adjusted as described by Whitford and collaborators [32]. When increasing the conformational sampling, no inter-domain interactions were included, indicating that the simulations were mostly entropically driven.

The simulation steps were integrated using the GROMACS software package 4 [33] at a temperature slightly lower than the folding temperature and coupled by a thermal bath controlled by Langevin dynamics for a total of 2,000,000,000 steps using a time step = 0.0005 for a total of 100 ns.

The free energy profile was calculated using the simple histogram method [34] for the following reaction coordinates: radius of gyration (R_g) and distance from the center of mass of the XynA to the center of mass of the BglS in the chimera (CM distance Xyl-Lich), both of which were calculated using GROMACS analysis tools [33]. The conformations with lower free energy values were considered the computational suggestions for describing the arrangement of the domains in solution.

2.3. Assessment of the enzymatic functionality using simulations

All-atom SB were employed to simulate 5×10^8 steps using the energetic parameters previously defined and without modifications to the topological files generated by the SMOG web server (the contact map included inter-domain interactions). To evaluate possible unexpected features, the root mean square fluctuation (RMSD) and average solvent accessible surface area (SAS) parameters per residue were calculated using GROMACS tools [33].

2.4. Assembly and protein expression of the XynA-BglS chimera (XylLich)

Genomic DNA from *B. subtilis* 168 was used as the template for PCR amplification of the *xynA* (NCBI-GI: 16078944) and *bglS* (NCBI-GI: 16080958) genes. The forward and reverse primers were 5′-tatatagctagcagcacagactactggcaaaa-3′ and 5′-tatataggatcccacactg ttacgttagaac-3′, respectively, for xynA, and the forward and reverse primers were 5′-tatatagctagccaaacaggtggatcgttttt-3′ and 5′-tatata ggatccttattttttgtatagcgca-3′, respectively, for *bglS*. The restriction sites are underlined in the primer sequences. Two different primers were tailor designed to fuse the genes (*xynA* reverse 5′-accaccacc accccacactgttacgttagaac-3′ and *bglS* forward 5′-ggtggtggtggtcaaacagg tggatcgttttt-3′) and included a four-glycine residues linker [9] (underlined nucleotides). The amplified fragments were used as

templates for the overlap extension PCR technique [35] to fuse the two genes in a unique ORF. The chimeric fragment was digested with the restriction enzymes *Nhe*I and *BamH*I and cloned in pET28a(+) (Novagen).

All plasmid constructions, pET-XynA, pET-BglS and pET-XylLich, were produced in BL21 DE3 cells. The protein expression was induced with a final concentration of 0.5 mM IPTG in 0.5 L LB medium for 5 h with shaking at 200 rpm and 37 °C. The culture was harvested at the end of fermentation and resuspended in cell lysis buffer consisting of 20 mM phosphate buffer at pH 7.4 with 5 mM imidazole, 1 mM PMSF and 0.5 mg/mL lysozyme. Then, the cells were disrupted by sonication. Two protein purification steps, including Ni^{2+}-chelating affinity and size-exclusion chromatography, were performed according to Squina et al. [36]. The purified proteins were further analyzed using SDS-PAGE, and the protein concentration was determined from the absorbance at 280 nm.

2.5. SAXS data collection and validation of the predicted conformations in solution

Small angle X-ray scattering (SAXS) data for XylLich were collected on the SAXS2 beam line at the Brazilian Synchrotron Light Laboratory at the concentrations of 1 and 2 mg/mL. The radiation wavelength was set to 1.48 Å, and a 165 mm MarCCD detector was used to record the scattering patterns. The sample-to-detector distance was set to 1022.5 mm to give a range of the scattering vector q from 0.013 to 0.33 $Å^{-1}$, where q is the magnitude of the q-vector defined by $q = 4\pi\, sin\theta/\lambda$ (2θ is the scattering angle). Protein samples were prepared in a buffer with 20 mM phosphate and 50 mM NaCl at pH 7.4. The SAXS patterns were integrated using Fit2D software [37], and the curves were scaled by the protein concentration. The molecular weight was calculated using the procedure implemented in the web tool SAXSmoW [38].

CRYSOL 2.7 [39] was used to generate theoretical scattering curves and to compare the experimental and theoretical data. The parameter for goodness of fitting, χ^2, was defined as follows:

$$\chi^2 = \frac{1}{N_q}\left[\frac{I^{experimental}(q) - I^{theoretical}(q)}{\varsigma(q)}\right]^2 \quad (2)$$

where N_q is the total number of experimental points, $I^{experimental}(q)$ and $I^{theoretical}(q)$ are the experimental and theoretical intensities of the scattering vector q, respectively, and $\varsigma(q)$ is the standard deviation of the experimental intensity values.

2.6. Assessing the enzymatic properties

All the assays were carried out with an automated pipetting system: epMotion® 5075 (Eppendorf). Reducing sugars were determined using the 3,5-dinitrosalicylic acid (DNS) method and monitored colorimetrically at 540 nm [40] using an Infinite® 200 PRO microplate reader (TECAN). One unit of enzyme was defined as the quantity of enzyme necessary to release reducing sugars at a rate of 1 μmol per minute under standard conditions. The standard assay was conducted for 10 min in 40 mM McIlvaine's buffer with glycine at pH 6, 50 °C. The final concentrations were 2.5 mg/mL for the substrates and 1 μM for the enzymes.

The optimal pH and temperature were determined for XynA and BglS using beechwood xylan and lichenan as substrates, respectively, under standard conditions. The enzymatic activity rates of the chimera and independent modules were evaluated with other substrates, including sugar beet, debranched arabinan, linear arabinan, rye arabinoxylan, larch arabinogalactan, galactomannan, xyloglucan, oat spelt xylan, wheat arabinoxylan and xylan beechwood. The polysaccharides were purchased from Sigma Aldrich or Megazyme.

Xylohexaose and the oligosaccharides derived from lichenan (Megazyme) were derivatized with 8-aminopyrene-1,3,6-trisulfonic acid (APTS) by reductive amination [41]. Thus, the enzymatic hydrolysis of these labeled substrates was performed at 50 °C. Capillary zone electrophoresis (CZE) was performed in a P/ACE MQD instrument (Beckman Coulter) equipped with a laser-induced fluorescence detector, as previously described [42]. A fused-silica capillary (TSP 050375, Polymicro Technologies) with an internal diameter of 50 μm and total length of 31 cm was used as a separation column for the oligosaccharides. The electrophoresis conditions were 15 kV/70–100 μA at a controlled temperature of 20 °C using sodium phosphate buffer (40 mM, pH 2.5).

The kinetic parameters were estimated for all enzymes from the initial rates using twelve substrate concentrations in the range of 1–30 mg/mL. The assays were carried out under standard conditions, using beechwood xylan and lichenan as substrates to assess the V_{max}, K_m and k_{cat} of the activity of xylanase and lichenase, respectively.

Far-UV circular dichroism (CD) spectra were taken on a JASCO J-810 spectropolarimeter (Jasco Inc., Tokyo, Japan.) equipped with a Peltier temperature control unit using a wavelength range of 195–250 nm, a 0.1 cm path quartz cuvette. The solvent spectra were subtracted in all experiments to eliminate background effects. The CD spectra were the average of 8 accumulations using a scanning speed of 100 nm min^{-1}, spectral bandwidth of 1 nm, and response time of 0.5 s. The protein concentration was 0.2 mg/mL in 50 mM sodium phosphate buffer at pH 7.4. The thermal denaturation of the xylanase–lichenase chimera was characterized by measuring the ellipticity changes at 218.5 nm induced by a temperature increase from 20 °C to 100 °C at a heating rate of 1 °C min^{-1} [43].

The biotechnological potential of the chimeric enzyme was evaluated under standard conditions using a polysaccharide composite prepared with beechwood xylan and lichenan in a 1:1 ratio (10 mg/mL). The conversion rates were assessed for XynA, BglS, XynA plus BglS mixture and XylLich at the same molar concentration (1 μM). The amount of reducing sugar was measured as previously described.

To compare the hydrolytic activity and recombinant protein production yield of the chimera and the parental enzymes, the enzymes were produced at the same scale as previously described. The total produced cell mass was weighed. Afterwards, the crude enzyme extract was prepared through cell lysis, as previously described. The enzymatic activities were measured under standard conditions using the crude preparations of XynA, BglS and XylLich. The total protein content was also determined for all the samples using a commercial Bradford kit from Bio-Rad®. Finally, to correlate the overall activity and the total recombinant protein production yield, a parameter dubbed activity yield (AY) was generated by dividing the total enzyme units (U) by the total protein content (mg) and cell mass that was produced (g).

3. Results

3.1. Simulations suggested a unique ensemble of structures

Molecular dynamics simulations were performed to characterize the most probable conformation of the chimera in solution. To assess probable arrangements of the chimera in solution, 2,000,000 conformations were generated through simulations using SB models and evaluated. Inter-region contacts (among XynA, BglS and the linker) were removed to allow an extensive search of the configuration space. Fig. 1A illustrates the universe of conformational possibilities derived from these simulations. The free energy profile was calculated as a function of the following reaction coordinates: CM distance, R_g and total energy of the system. An ensemble of candidate structures covered the lowest free-energy region, within which an accurate chimera arrangement was also expected. As shown in Fig. 1A, a large number of possible configurational states were sampled within

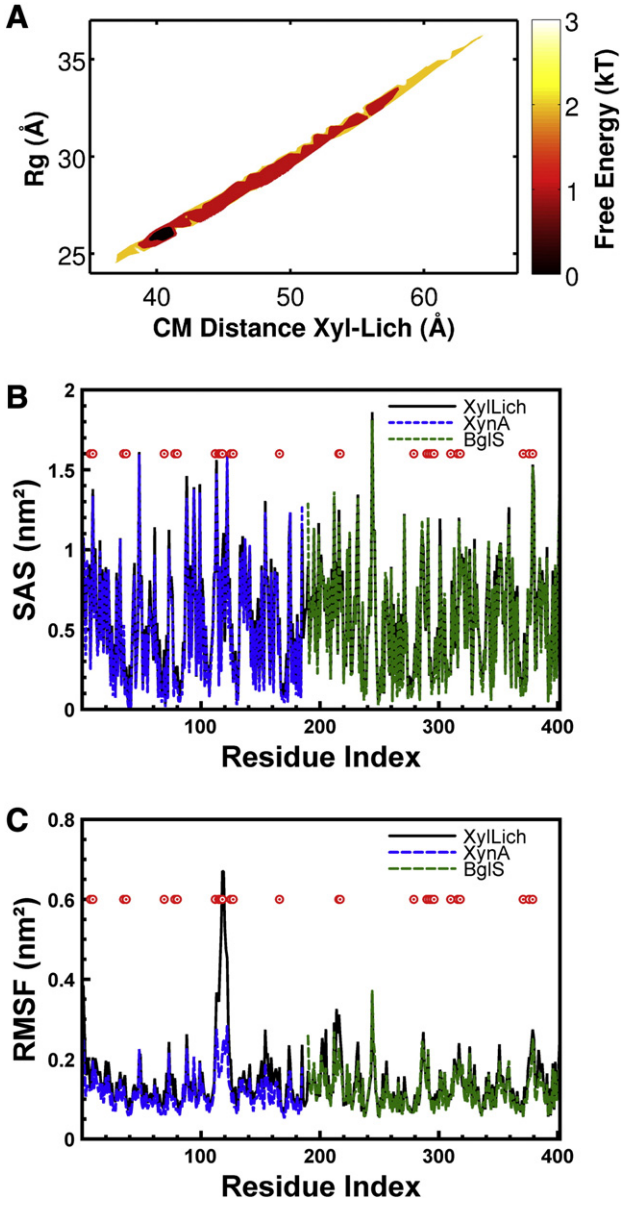

Fig. 1. Characterization of XylLich using structure-based models (SB) through molecular dynamics (MD) simulations. (A) The figure on the top shows the free energy profile as a function of the radius of gyration (R_g) and the distance between the centers of mass of the XynA and BglS domains (CM). Free energy is given in units of kT and presented as a color scale. The figure was obtained using extensive simulation (100 ns) and no intra-domain interactions (see the Material and methods section). (B) Average solvent accessible surface area (SAS) per residue, calculated from structure-based model simulations and (C) root mean square fluctuation (RMSF) per residue are presented. A and B show the chimera (XylLich) in black, XynA only in blue, BglS only in green and the residues related to the catalytic site as red dots.

a distinctive basin of low-energy conformations, which was restricted to the R_g region near 26 Å with a CM close to 42 Å, which represented the accommodated chimera in solution. These results were obtained without any experimental support.

Conformations outside the free-energy basin revealed disagreements (high χ^2 values), which were caused by incorrect protein arrangements (Figs. SI-2). The compacted conformations showed deformations because of the high repulsive force between the residues. The extended conformations presented local unfolding and distortions near the linker. In both cases, the protein accommodation in solution was not quenched.

Simulations using XynA, BglS and the computationally suggested chimera were performed to predict possible modifications in their functional behavior. One conformation from the basin was chosen for a new simulation (R_g = 26.7 Å, *CM* distance = 42.4 Å). Simulations using the chimera and the enzyme domains separately (XynA and BglS) displayed minor variations in the SAS (Fig. 1B), suggesting that the accessibility of the substrate to the pocket was not affected in the designed chimera. RMSF analysis (Fig. 1C) also showed minor variations, except for the XynA residues SER:117 and ILE:118, which were located into the "thumb" region of the protein [44,45].

3.2. Validation of the computationally predicted conformations through SAXS experiments

The genes *xynA* and *bglS* were fused by PCR-mediated overlap extension, and four glycine residues were included as a linker (Fig. 2A). The resulting amplicon, which was cloned into a pET28a vector, was 1209 bp long (Fig. 2A). Restriction analysis and DNA sequencing confirmed the molecular cloning. The chimera and its parental enzymes, XynA (22.8 kDa) and BglS (26.7 kDa), were over expressed in *Escherichia coli* and purified through two chromatographic steps: first, nickel-affinity and then size-exclusion chromatography. The apparent molecular weight of the xylanase–lichenase chimera was 47.2 kDa (Fig. 2B).

Because the candidate conformations were obtained without previous experimental information, SAXS was employed to validate the computationally predicted structures. The superimposed theoretical and experimental scattering curves of the selected conformation and the XylLich model are presented in Fig. 2C. Theoretical scattering curves ($I(q)$ versus q) within the broad basin of low-energy conformations were compared to the experimental scattering curve using CRYSOL. Good agreements were obtained for these conformations, i.e., low χ^2 values, which were equal to 2.9 on average (some examples are presented in Figs. SI-1A and Figs. SI-1B), validating our simulation strategy. The obtained conformations displayed no obstructions or deformations at the catalytic region, indicating that the chimera would be fully functional in vitro.

3.3. The chimera maintained the functional characteristics of the parental enzymes

The optimal temperature and pH of operation did not change drastically between the chimeric enzyme and the parental enzymes. The optimal pH for XynA and BglS was just 0.5 units lower than that of XylLich (Fig. 3A and B). There was no statistically significant shift in the xylanase temperature-dependent activity for the chimera, whereas the optimum temperature for lichenan hydrolysis was slightly greater for XylLich than for the parental enzymes (Fig. 3C and D).

The enzymes were biochemically assessed using a set of 10 natural polysaccharides. The specific activities of the chimera and the parental enzymes on these polysaccharides are summarized in Table 1. XynA and XylLich hydrolyzed birchwood xylan, rye arabinoxylan and beechwood xylan more efficiently than other polysaccharides. Taking into account the standard deviation, there was no statistically significant difference between XynA and XylLich in hydrolyzing birchwood xylan and rye arabinoxylan. XylA and XylLich presented the lowest xylan-degrading activity (29–36%) on wheat arabinoxylan, which was the most complex and insoluble substrate tested. XylLich and BglS hydrolyzed lichenan and β-glucan equally. The relative activities of XylLich and BglS on lichenan were 72 and 75% lower, respectively, on β-glucan.

Fig. 4 describes our attempt to evaluate the mode of action of XylLich, XynA and BglS through CZE analysis of APTS-labeled oligosaccharides. Both parental and chimeric xylanase produced the same degradation pattern for APTS-labeled xylohexaose (X6), which included xylotetraose (X4), xylotriose (X3) and xylobiose (X2) (Fig. 4A). BglS was not able to hydrolyze X6. After hydrolysis of lichenan, XylLich and

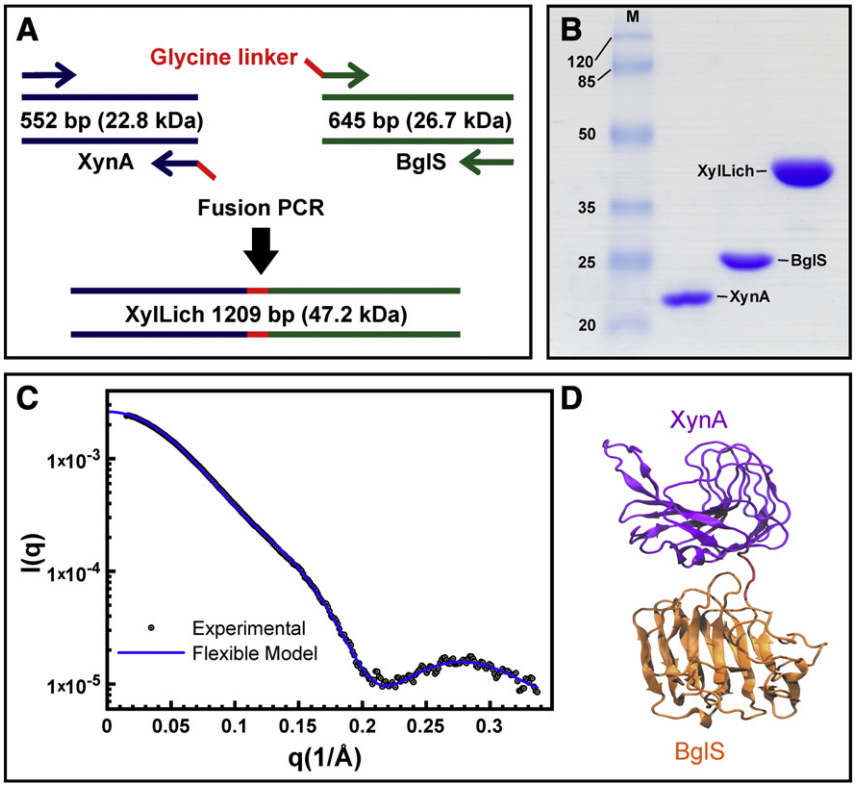

Fig. 2. The chimera construction and experimental validation by SAXS. (A) Flow chart showing the process of fusion PCR for the XylLich construction. (B) SDS-PAGE of the chimera and wild-type proteins, indicating that the fused enzyme xylanase–lichenase had the predicted molecular weight. The protein molecular weight marker is shown in the first lane (M), and the values are displayed in kDa. (C) The small angle X-ray scattering profile for the experimental and theoretical evaluation of XylLich, which was taken from the free energy basin with $\chi^2 = 2.80$, $R_g = 26.0$ Å and *CM* distance = 40.7 Å. The scattering intensity is shown on a logarithm scale as a function of the momentum transfer (q). (D) The XylLich model comprising XynA (purple), the linker (red) and BglS (orange). CRYSOL was employed to generate the theoretical curve and VMD [55] for the denoted cartoon.

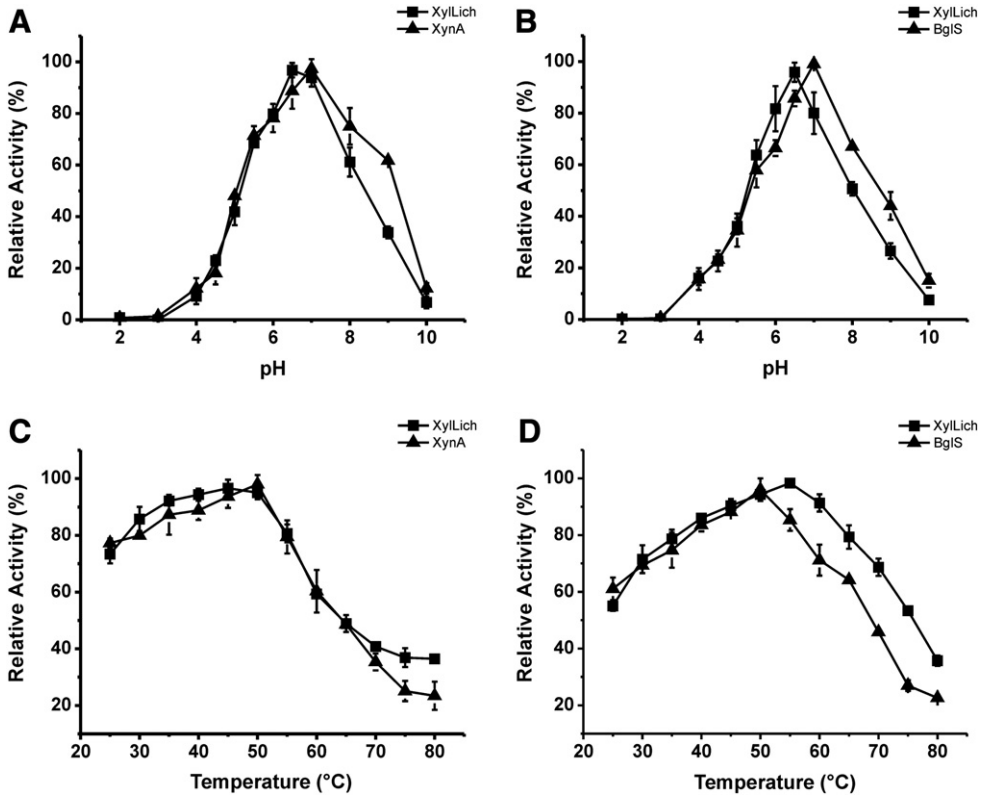

Fig. 3. The effects of pH and temperature on the XylLich catalytic activity. pH (A and B) and temperature (C and D) curves for the chimera (■) and the parental enzymes (▲). The influence of pH on the enzymatic activity of XylLich compared to that of XynA (A) and BglS (B) using beechwood xylan (A) and lichenan (B) as substrates. The effect of temperature on the enzymatic activity of the chimeric enzyme compared to that of xylanase (C) and lichenase (D) using beechwood xylan and lichenan as substrates, respectively, is represented.

Table 1
Specific activities of the parental and chimeric enzymes on different types of substrates.

	Specific activity (U/nmol)		
Substrate	XynA	BglS	XylLich
Birchwood xylan	3.73 ± 0.29	ND	2.71 ± 0.13
Beechwood xylan	3.17 ± 0.07	ND	2.87 ± 0.08
Rye arabinoxylan	3.73 ± 0.14	ND	3.03 ± 0.15
Wheat arabinoxylan	1.36 ± 0.12	ND	0.88 ± 0.07
Oat spelt xylan	3.28 ± 0.27	ND	2.15 ± 0.06
Lichenan	ND	3.65 ± 0.29	3.85 ± 0.16
β-Glucan	ND	5.03 ± 0.20	5.11 ± 0.07

Legend. ND means not determined. There was no activity on laminarin, xyloglucan and glucomannan (konjac) for all enzymes. Values are given by the mean ± S.D. of three independent assays.

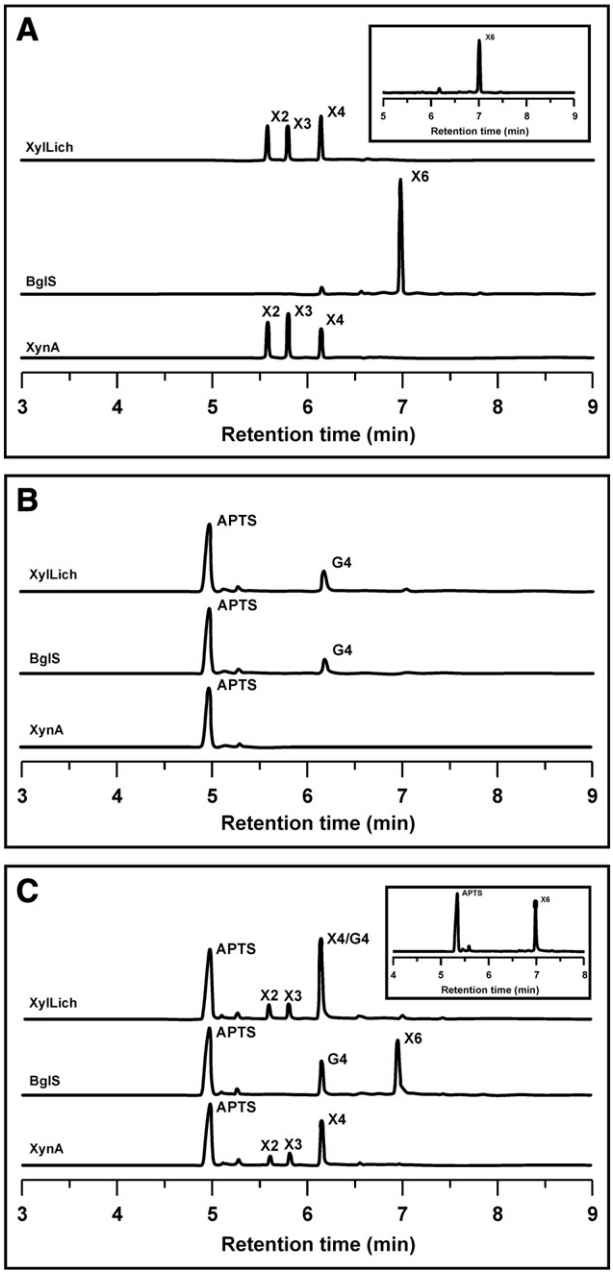

Fig. 4. Capillary zone electrophoresis analysis of the breakdown products released by XylLich, XynA and BglS. The products after enzymatic hydrolysis of APTS-labeled xylohexaose (A), lichenan (B) and xylohexaose plus lichenan together (C) are presented. X2, X3, X4, X6 and G4 indicate the degree of polymerization of the produced xylose and glucose oligomers. The APTS-labeled xylohexaose used in the assays is indicated in the upper right boxes (A and C).

BglS produced glucotetraose (G4), and XynA was not able to break down lichenan (Fig. 4B). As shown in Fig. 4C, the enzymes were also assayed against the two substrates (X6 and lichenan) simultaneously. The results confirmed our previous findings from using the substrates separately (Fig. 4A and B), highlighting that the mode of operation of the chimera was exactly similar to that of the parental enzymes.

Saturation assays were performed using beechwood xylan and lichenan. The kinetic constants of the wild-type and chimeric enzymes are compared in Table 2. The maximum xylan degradation rate (V_{max}) of the chimera was 30% lower than that of the parental enzyme (XynA), but for lichenan hydrolysis, the V_{max} was 33% greater for XylLich than for BglS. Using xylan as a substrate, XylLich had a lower K_m (41%) and turnover number (30%) than XynA. In contrast, the K_m of the chimeric enzyme was approximately 55% higher than that of BglS. In addition, the k_{cat} increased by up to 28% for XylLich using lichenan as substrate. Moreover, the catalytic efficiency (k_{cat}/K_m) of the chimera was greater (18%) for xylan degradation and lower (18%) for lichenan hydrolysis than for the parental enzymes.

The far-UV CD spectra of XylLich showed a negative peak near 218 nm, suggesting that XylLich consists primarily of β-sheets (Fig. 5A). This profile was expected because the wild-type enzymes have β-sheet predominance [46–48]. Another band was detected in the 220–230 nm region of the BglS spectrum, indicating the contribution of side-chain aromatic residues [47]. Thermal denaturation experiments were performed to compare the stability of the chimera and wild-type enzymes. The XylLich melting temperature (Tm) was 43.6 °C, whereas the Tm values of the parental enzymes were 53.3 °C and 42.5 °C for XynA and BglS, respectively (Fig. 5B).

3.4. Biotechnological appeal for producing the chimeric enzyme

To evaluate the biotechnological potential of using the chimeric enzyme in biomass to bio-products applications, hydrolysis assays

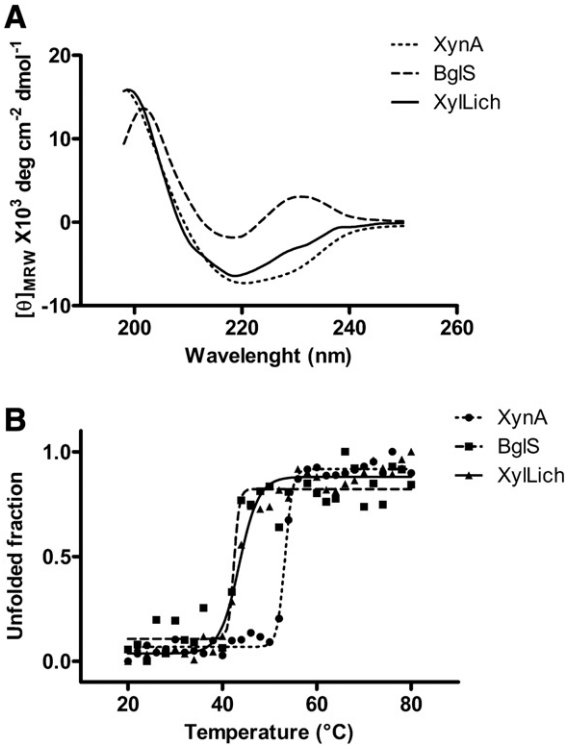

Fig. 5. Secondary structure evaluation by circular dichroism. (A) Far-UV CD spectra of XylLich and individual enzymes at pH 7.4 and 20 °C and (B) thermal denaturation curve at pH 7.4. The circular dichroism spectrum represents an average of eight scans. The thermal denaturation curve was obtained by monitoring at 220 nm for xylanase, 218 nm for lichenase and 218.5 nm for the chimera.

Table 2
Kinetic parameters of the chimeric and parental enzymes.

Substrate	Beechwood xylan				Lichenan			
	V_{max} μmol/min/μmol	K_m mg/mL	k_{cat} s^{-1}	k_{cat}/K_m $mL \cdot mg^{-1} \cdot s^{-1}$	V_{max} μmol/min/μmol	K_m mg/mL	k_{cat} s^{-1}	k_{cat}/K_m $mL \cdot mg^{-1} \cdot s^{-1}$
XynA	10420 ± 497	7.87 ± 0.71	173.7 ± 8.3	22.1 ± 1.1	ND	ND	ND	ND
BglS	ND	ND	ND	ND	11430 ± 407	3.39 ± 0.26	197.2 ± 8.5	58.1 ± 2.5
XylLich	7300 ± 212	4.67 ± 0.26	121.7 ± 3.5	26.1 ± 0.8	15190 ± 1987	5.27 ± 1.13	253.2 ± 33.1	48.1 ± 6.3

Legend. ND means not determined. Values are given by the mean ± S.D. of three independent assays. The V_{max} unit was defined as μmol of reducing sugars released per minute per μmol of enzyme.

were performed using a polysaccharide composite comprising beechwood xylan and lichenan in a 1:1 ratio. The conversion efficiencies of *XynA*, *BglS*, an equimolar mixture of the parental enzymes and the chimeric enzyme on this polysaccharide composite are described in Fig. 6A. The rate of substrate conversion for XylLich was statistically equal to that of the parental enzymes combined (Fig. 6A).

The enzymatic performance was also evaluated using the crude *E. coli* cell lysate as an enzyme source. Based on the total protein content in the cell lysate and the overall mass produced by *E. coli* cells, the amount of enzyme required to release 1 μmol of reducing sugar per minute per protein milligram per gram of cell (U/mg/g) was calculated, which was defined in AY units (see methods). The chimera AY (0.35 U/mg/g) was approximately 6 times higher than XynA AY(0.06 U/mg/g) using xylan as a substrate. Likewise, the AY was 20% greater for XylLich than for BglS (Fig. 6B) using lichenan as a substrate.

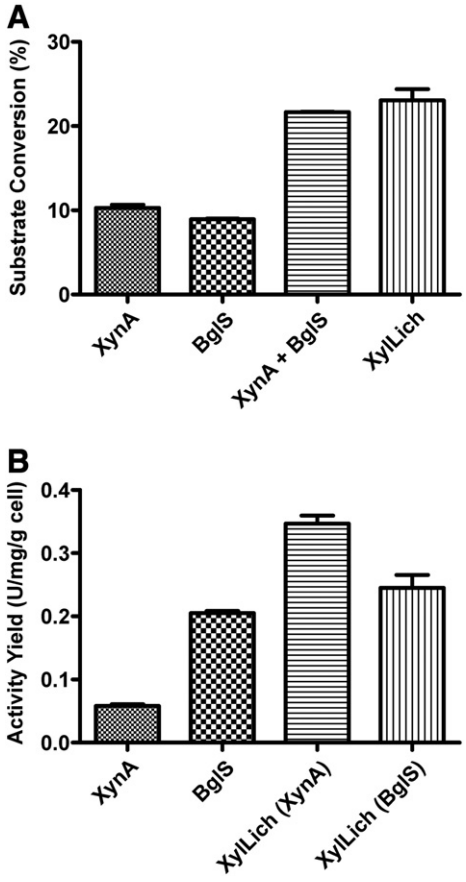

Fig. 6. Biotechnological appealing for producing the chimeric enzyme. (A) Analysis of the conversion efficiency of a composite consisting of beechwood xylan and lichenan by the parental enzymes, the mixture of the parental enzymes (XynA + BglS) and XylLich. (B) Assessment of the activity yield (AY) using the crude *E. coli* cell extract as an enzyme source.

4. Discussion

This study is an initial step toward the development of automated screening pipelines for optimal linkers and modules and functional chimeras before experimental validation. XynA and BglS from *B. subtilis* were chosen because these enzymes have been fully characterized and crystallographic structures are available [44–46,48–50]. Accordingly, the small and simple linker comprising four glycine residues was very well described previously [9]. However, based on the number of conformations available to XylLich (Figs. 1A, SI 1 and SI 2), its predicted organization in solution was not obvious without computational tools or SAXS.

Simulations employing SB models require significantly less computational time than traditional molecular dynamics using explicit water or complex force fields. This key benefit allows the straightforward investigation of several constructions and larger systems. Despite being minimalistic approaches, SB models take into account the atomic restrictions to which real proteins are susceptible and thus corroborate experimental results [22]. The solvent properties implicitly included in these models support the calculation of an SAS, which predicts the change of the substrate accessibility to the enzyme [51] and estimates alterations in the enzymatic functionality. Simulations employing the chimeric model ($\chi^2 = 2.8$) and the parental enzymes (XynA or BglS) presented minor changes in the SAS. These results suggested that variations in the operation mode would not be expected because the binding site was not obstructed by deformations or steric effects between the domains. Supporting our hypothesis, the RMSF analysis (Fig. 1C) showed a similar profile for all cases. The only exception was the residue ILE:118 of the XynA thumb region located near the catalytic binding site, which could influence the enzymatic catalysis [44,45].

To validate the molecular dynamics simulations, comprehensive biochemical characterization was performed with the parental enzymes and the chimera. Collectively, the biochemical analysis confirmed the simulations. The main exception was the slight variation of the XylLich temperature of operation and catalytic efficiency (k_{cat}/K_m). There were no substantial shifts in either the substrate specificities or the mode of operation. In addition, circular dichroism analysis revealed a greater similarity to the thermal denaturation properties of XylLich to BglS than those of XynA.

The minor shifts in the optimal pH and temperature of operation can be explained by an alteration in the microenvironment of the fused enzyme; one enzyme module preferentially catches protons and acidifies the local pH [52]. In addition, protein fusion can disturb the tertiary structure of a chimeric enzyme [52], which can cause displacements in the pH and temperature of operation, phenomena that are very well described in the literature [5,50,52].

The engineering of fused proteins has long been considered attractive for industrial processes because these proteins are more efficient (based on costs and catalytic efficiency) than wild-type proteins [5,49,50,53,54]. Along with the lack of substantial changes in the catalytic performance of the chimera, the parameter AY suggested an advantage in producing the fused protein rather than the separate wild-type ones. Certainly more studies on XylLich properties are required before use in a scaled-up process.

In conclusion, this work presented a novel approach for predicting the arrangement of chimeric domains in solution before experimental validation. The computational strategy proposed herein is fast and robust for characterizing a large number of enzymes in a few weeks, as well as for identifying possible binding site obstructions or large dynamical variations differing from the single domains. The methodology can be extended to multi-domain chimeras and, by the same token, to a wide variety of biological systems. Finally, we expect that our findings will increase the pace of finding novel and cost-effective approaches for converting plant biomass into bio-products. One of the great challenges in biomass saccharification is decreasing the cost of enzyme production. Therefore, the use of multifunctional hydrolases, which act synergistically in plant polysaccharide degradation, is a promising venue for the improvement of enzyme cocktails for second-generation biofuels.

Acknowledgements

This research was supported by grants from FAPESP (2008/58037-9) and CNPq (475022/2011-4 and 310177/2011-1). JC received a scholarship from CNPq (140420/2009-6), and LCO, TMA and ARLD received scholarships from FAPESP (2011/13242-7, 2010/11499-1 and 2011/02169-7, respectively). X-ray scattering data were collected at the Brazilian Synchrotron Light Laboratory (LNLS) that integrates the Center of Research in Energy and Material (CNPEM). The authors would like to thank the LNLS for support. The computational analyses were supported by resources supplied by the Center for Scientific Computing (NCC/Grid UNESP) of São Paulo State University (UNESP) and CENAPAD-SP (Centro Nacional de Processamento de Alto Desempenho em São Paulo), project UNICAMP/FINEP-MCT.

Appendix A. Supplementary data

Supplementary data to this article can be found online at http://dx.doi.org/10.1016/j.bbapap.2013.02.030.

References

[1] M.E. Himmel, S.-Y. Ding, D.K. Johnson, W.S. Adney, M.R. Nimlos, J.W. Brady, T.D. Foust, Biomass recalcitrance: engineering plants and enzymes for biofuels production, Science (New York, N.Y.) 315 (2007) 804–807.
[2] T. Collins, C. Gerday, G. Feller, Xylanases, xylanase families and extremophilic xylanases, FEMS Microbiol. Rev. 29 (2005) 3–23.
[3] D.U. Lima, H.P. Santos, M.a. Tiné, F.R.D. Molle, M.S. Buckeridge, Patterns of expression of cell wall related genes in sugarcane, Genet. Mol. Biol. 24 (2001) 191–198.
[4] A. Planas, Bacterial 1,3-1,4-beta-glucanases: structure, function and protein engineering, Biochim. Biophys. Acta 1543 (2000) 361–382.
[5] Z. Fan, K. Wagschal, C.C. Lee, Q. Kong, K.a. Shen, I.B. Maiti, L. Yuan, The construction and characterization of two xylan-degrading chimeric enzymes, Biotechnol. Bioeng. 102 (2009) 684–692.
[6] S.Y. Hong, J.S. Lee, K.M. Cho, R.K. Math, Y.H. Kim, S.J. Hong, Y.U. Cho, S.J. Cho, H. Kim, H.D. Yun, Construction of the bifunctional enzyme cellulase-beta-glucosidase from the hyperthermophilic bacterium *Thermotoga maritima*, Biotechnol. Lett. 29 (2007) 931–936.
[7] C.J. Crasto, J.A. Feng, LINKER: a program to generate linker sequences for fusion proteins, Protein Eng. 13 (2000) 309–312.
[8] J.L. Casey, A.M. Coley, L.M. Tilley, M. Foley, Green fluorescent antibodies: novel in vitro tools, Protein Eng. 13 (2000) 445–452.
[9] R. Arai, H. Ueda, A. Kitayama, N. Kamiya, T. Nagamune, Design of the linkers which effectively separate domains of a bifunctional fusion protein, Protein Eng. 14 (2001) 529–532.
[10] R.A. Griep, C. van Twisk, J.M. van der Wolf, A. Schots, Fluobodies: green fluorescent single-chain Fv fusion proteins, J. Immunol. Methods 230 (1999) 121–130.
[11] K. Morino, H. Katsumi, Y. Akahori, Y. Iba, M. Shinohara, Y. Ukai, Y. Kohara, Y. Kurosawa, Antibody fusions with fluorescent proteins: a versatile reagent for profiling protein expression, J. Immunol. Methods 257 (2001) 175–184.
[12] M. Peipp, D. Saul, K. Barbin, J. Bruenke, S.J. Zunino, M. Niederweis, G.H. Fey, Efficient eukaryotic expression of fluorescent scFv fusion proteins directed against CD antigens for FACS applications, J. Immunol. Methods 285 (2004) 265–280.
[13] O. Miyashita, C. Gorba, F. Tama, Structure modeling from small angle X-ray scattering data with elastic network normal mode analysis, J. Struct. Biol. 173 (2011) 451–460.
[14] M. Pelikan, G.L. Hura, M. Hammel, Structure and flexibility within proteins as identified through small angle X-ray scattering, Gen. Physiol. Biophys. 28 (2009) 174–189.
[15] S. Yang, L. Blachowicz, L. Makowski, B. Roux, Multidomain assembled states of Hck tyrosine kinase in solution, Proc. Natl. Acad. Sci. U. S. A. 107 (2010) 15757–15762.
[16] S. Yang, S. Park, L. Makowski, B. Roux, A rapid coarse residue-based computational method for X-ray solution scattering characterization of protein folds and multiple conformational states of large protein complexes, Biophys. J. 96 (2009) 4449–4463.
[17] S.J. Kim, C. Dumont, M. Gruebele, Simulation-based fitting of protein–protein interaction potentials to SAXS experiments, Biophys. J. 94 (2008) 4924–4931.
[18] J.N. Onuchic, P.G. Wolynes, Theory of protein folding, Curr. Opin. Struct. Biol. 14 (2004) 70–75.
[19] J.D. Bryngelson, J.N. Onuchic, N.D. Socci, P.G. Wolynes, Funnels, pathways, and the energy landscape of protein folding: a synthesis, Proteins 21 (1995) 167–195.
[20] L.L. Chavez, J.N. Onuchic, C. Clementi, Quantifying the roughness on the free energy landscape: entropic bottlenecks and protein folding rates, J. Am. Chem. Soc. 126 (2004) 8426–8432.
[21] C. Clementi, Coarse-grained models of protein folding: toy models or predictive tools? Curr. Opin. Struct. Biol. 18 (2008) 10–15.
[22] C. Clementi, P.A. Jennings, J.N. Onuchic, Prediction of folding mechanism for circular-permuted proteins, J. Mol. Biol. 311 (2001) 879–890.
[23] J.G. Lyubovitsky, H.B. Gray, J.R. Winkler, Mapping the cytochrome C folding landscape, J. Am. Chem. Soc. 124 (2002) 5481–5485.
[24] L.C. Oliveira, A. Schug, J.N. Onuchic, Geometrical features of the protein folding mechanism are a robust property of the energy landscape: a detailed investigation of several reduced models, J. Phys. Chem. B 112 (2008) 6131–6136.
[25] M.A. Jamros, L.C. Oliveria, P.C. Whitford, J.N. Onuchic, J.A. Adams, D.K. Blumenthal, P.A. Jennings, Proteins at work: a combined SAXS and theoretical determination of the multiple structures involved on the protein kinase functional landscape, J. Biol. Chem. 285 (2010) 36121–36128.
[26] M.A. Jamros, L.C. Oliveira, P.C. Whitford, J.N. Onuchic, J.A. Adams, et al., Substrate-specific reorganization of the conformational ensemble of CSK implicates novel modes of kinase function, PLoS Comput. Biol. 8 (2012) 1–8.
[27] R. Koradi, M. Billeter, K. Wüthrich, MOLMOL: a program for display and analysis of macromolecular structures, J. Mol. Graph. 14 (1996) 51–55, (29–32).
[28] J.C. Phillips, R. Braun, W. Wang, J. Gumbart, E. Tajkhorshid, E. Villa, C. Chipot, R.D. Skeel, L. Kalé, K. Schulten, Scalable molecular dynamics with NAMD, J. Comput. Chem. 26 (2005) 1781–1802.
[29] A.D. MacKerell, D. Bashford, Bellott, R.L. Dunbrack, 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, B. Roux, M. Schlenkrich, J.C. Smith, R. Stote, J. Straub, M. Watanabe, J. Wiórkiewicz-Kuczera, D. Yin, M. Karplus, All-atom empirical potential for molecular modeling and dynamics studies of proteins, J. Phys. Chem. B 102 (1998) 3586–3616.
[30] J.K. Noel, P.C. Whitford, K.Y. Sanbonmatsu, J.N. Onuchic, SMOG@ctbp: simplified deployment of structure-based models in GROMACS, Nucleic Acids Res. 38 (2010) W657–W661, (Suppl.).
[31] J.K. Noel, P.C. Whitford, J.N. Onuchic, The Shadow map: a general contact definition for capturing the dynamics of biomolecular folding and function, J. Phys. Chem. B 116 (2012) 8692–8702.
[32] P.C. Whitford, J.K. Noel, S. Gosavi, A. Schug, K.Y. Sanbonmatsu, J.N. Onuchic, An all-atom structure-based potential for proteins: bridging minimal models with all-atom empirical forcefields, Proteins 75 (2009) 430–441.
[33] B. Hess, C. Kutzner, D.V.D. Spoel, E. Lindahl, GROMACS 4: algorithms for highly efficient, load-balanced, and scalable molecular simulation, J. Chem. Theory Comput. 4 (2008) 435–447.
[34] A.M. Ferrenberg, R.H. Swendsen, New Monte Carlo technique for studying phase transitions, Phys. Rev. Lett. 61 (1988) 2635–2638.
[35] K.L. Heckman, L.R. Pease, Gene splicing and mutagenesis by PCR-driven overlap extension, Nat. Protoc. 2 (2007) 924–932.
[36] F.M. Squina, C.R. Santos, D.a. Ribeiro, J. Cota, R.R. de Oliveira, R. Ruller, A. Mort, M.T. Murakami, R.a. Prade, Substrate cleavage pattern, biophysical characterization and low-resolution structure of a novel hyperthermostable arabinanase from *Thermotoga petrophila*, Biochem. Biophys. Res. Commun. 399 (2010) 505–511.
[37] A.P. Hammersley, ESRF Internal Report, ESRF97HA02T, "FIT2D: An Introduction and Overview", 1997. (Information obtained from: http://www.esrf.eu/computing/scientific/FIT2D/FIT2D_REF/node268.html).
[38] H. Fischer, M. de Oliveira Neto, H.B. Napolitano, I. Polikarpov, a.F. Craievich, Determination of the molecular weight of proteins in solution from a single small-angle X-ray scattering measurement on a relative scale, J. Appl. Crystallogr. 43 (2009) 101–109.
[39] D. Svergun, C. Barberato, M.H.J. Koch, CRYSOL — a program to evaluate X-ray solution scattering of biological macromolecules from atomic coordinates D. Svergun, C. Barberato and M. H. J. Koch, J. Appl. Crystallogr. 28 (1995) 768–773.
[40] G.L. Miller, Use of dinitrosalicylic acid reagent for determination of reducing sugar, Anal. Chem. 31 (1959) 426–428.
[41] R. Naran, M.L. Pierce, A.J. Mort, Detection and identification of rhamnogalacturonan lyase activity in intercellular spaces of expanding cotton cotyledons, Plant J. 50 (2007) 95–107.
[42] C.R. Santos, F.M. Squina, A.M. Navarro, D.P. Oldiges, A.F. Leme, R. Ruller, A.J. Mort, R. Prade, M.T. Murakami, Functional and biophysical characterization of a hyperthermostable GH51 alpha-L-arabinofuranosidase from *Thermotoga petrophila*, Biotechnol. Lett. 33 (2011) 131–137.
[43] J. Cota, T.M. Alvarez, A.P. Citadini, C.R. Santos, M. de Oliveira Neto, R.R. Oliveira, G.M. Pastore, R. Ruller, R.a. Prade, M.T. Murakami, F.M. Squina, Mode of operation and low-resolution structure of a multi-domain and hyperthermophilic endo-β-

1,3-glucanase from *Thermotoga petrophila*, Biochem. Biophys. Res. Commun. 406 (2011) 590–594.
[44] M.T. Murakami, R.K. Arni, D.S. Vieira, L. Degrève, R. Ruller, R.J. Ward, Correlation of temperature induced conformation change with optimum catalytic activity in the recombinant G/11 xylanase A from *Bacillus subtilis* strain 168 (1A1), FEBS Lett. 579 (2005) 6505–6510.
[45] D.S. Vieira, R.J. Ward, Conformation analysis of a surface loop that controls active site access in the GH11 xylanase A from *Bacillus subtilis*, J. Mol. Model. 18 (2012) 1473–1479.
[46] G.P. Furtado, L.F. Ribeiro, C.R. Santos, C.C. Tonoli, A.R. Souza, R.R. Oliveira, M.T. Murakami, R.J. Ward, Biochemical and structural characterization of a β-1,3-1,4-glucanase from *Bacillus subtilis* 168, Process Biochem. 46 (2011) 1202–1206.
[47] I.B. Grishina, R.W. Woody, Contributions of tryptophan side chains to the circular dichroism of globular proteins: exciton couplets and coupled oscillators, Faraday Discuss. 99 (1994) 245–262.
[48] R. Ruller, L. Deliberto, T.L. Ferreira, R.J. Ward, Thermostable variants of the recombinant xylanase A from *Bacillus subtilis* produced by directed evolution show reduced heat capacity changes, Proteins 70 (2008) 1280–1293.
[49] J. Aÿ, F. Götz, R. Borriss, U. Heinemann, Structure and function of the Bacillus hybrid enzyme GluXyn-1: native-like jellyroll fold preserved after insertion of autonomous globular domain, Proc. Natl. Acad. Sci. U. S. A. 95 (1998) 6613–6618.
[50] L.F. Ribeiro, G.P. Furtado, M.R. Lourenzoni, A.J. Costa-Filho, C.R. Santos, S.C. Peixoto Nogueira, J.a. Betini, M.D.L.T.M. Polizeli, M.T. Murakami, R.J. Ward, Engineering bifunctional laccase–xylanase chimerae for improved catalytic performance, J. Biol. Chem. 286 (2011) 40026–40038.
[51] M.L. Connolly, SCIENCE: solvent-accessible surfaces o f proteins and nucleic acids, Science 221 (1983) 709–713.
[52] L. Bülow, Characterization of an artificial bifunctional enzyme, beta-galactosidase/galactokinase, prepared by gene fusion, Eur. J. Biochem. 163 (1987) 443–448.
[53] L. Chang, M. Ding, L. Bao, Y. Chen, J. Zhou, H. Lu, Characterization of a bifunctional xylanase/endoglucanase from yak rumen microorganisms, Appl. Microbiol. Biotechnol. 90 (2011) 1933–1942.
[54] R. Khandeparker, M.T. Numan, Bifunctional xylanases and their potential use in biotechnology, J. Ind. Microbiol. Biotechnol. 35 (2008) 635–644.
[55] W. Humphrey, A. Dalke, K. Schulten, VMD: visual molecular dynamics, J. Mol. Graph. 14 (1996) 33–38.