

Structural and biochemical characterization of a multidomain alginate lyase reveals a novel role of CBM32 in CAZymes

Qianqian Lyu^{a,b}, Keke Zhang^a, Qiaoyun Zhu^a, Zhijian Li^a, Yujie Liu^a, Elisabeth Fitzek^c, Tanner Yohe^c, Liming Zhao^d, Weihua Li^e, Tao Liu^a, Yanbin Yin^c, Weizhi Liu^{a,b,*}

^a MOE Key Laboratory of Marine Genetics and Breeding, College of Marine Life Sciences, Ocean University of China, Qingdao 266003, China

^b Laboratory for Marine Biology and Biotechnology, Qingdao National Laboratory for Marine Science and Technology, Qingdao 266235, China

^c Department of Biological Sciences, Northern Illinois University, DeKalb, IL 60115, USA

^d School of Biotechnology, State Key Laboratory of Bioreactor Engineering, R&D Center of Separation and Extraction Technology in Fermentation Industry, Shanghai 200237, China

^e Shanghai Key Laboratory of New Drug Design, School of Pharmacy, East China University of Science and Technology, Shanghai 200237, China

ARTICLE INFO

Keywords:

CBM32
Alginate lyase
Alpha helix linker
X-ray crystallography
Molecular docking

ABSTRACT

Noncatalytic carbohydrate binding modules (CBMs) have been demonstrated to play various roles with cognate catalytic domains. However, for polysaccharide lyases (PLs), the roles of CBMs remain mostly unknown. AlyB is a multidomain alginate lyase that contains CBM32 and a PL7 catalytic domain. The AlyB structure determined herein reveals a noncanonical alpha helix linker between CBM32 and the catalytic domain. More interestingly, CBM32 and the linker does not significantly enhance the catalytic activity but rather specifies that trisaccharides are predominant in the degradation products. Detailed mutagenesis, biochemical and cocrystallization analyses show “weak but important” CBM32 interactions with alginate oligosaccharides. In combination with molecular modeling, we propose that the CBM32 domain serves as a “pivot point” during the trisaccharide release process. Collectively, this work demonstrates a novel role of CBMs in the activity of the appended PL domain and provides a new avenue for the well-defined generation of alginate oligosaccharides by taking advantage of associated CBMs.

1. Introduction

Carbohydrate binding modules (CBMs) are generally defined as accessory modules of carbohydrate active enzymes (CAZymes), which fold independently into discrete domains. By binding carbohydrates, CBMs perform diverse functions in CAZyme biology, including substrate recognition, insoluble polysaccharide structure disruption, general substrate adherence, and structure-function contributions to the catalytic site [1, 2]. The CBMs are recognized to play different important roles in enzyme function and currently a total of 83 families of CBMs has been included in the CAZy database. The accumulation of structurally characterized CBMs demonstrates that most CBMs fold into β -sandwich-like scaffolds, which can accommodate two distinct potential binding sites located in the variable loops (VLS) and concave face (CFS) [1, 3, 4]. Although many CBMs share a well-conserved structural scaffold, their ligand-binding specificities and functions are diverse, especially those of CBMs belonging to family 32 [3].

Knowledge of CBM roles in their cognate CAZymes is not only fundamental to biology but also facilitates their biotechnological

applications. Therefore, many efforts have been made to explore CBM functions. For example, phylogenetic predictions were used to help identify potential CBM binding specificity and key residues. In addition, an increasing number of solved three-dimensional CBM structures provides more direct information for the prediction of CBM structure-function relationships, especially regarding CBM binding specificity. For example, the VLS (A–E) region models of CBM6 and CBM35 as well as the CFS (A–C) aromatic cradle models of α -linked glucan-binding CBMs were established based on the integrated analysis of sequences and three-dimensional structures [1]. These models will also facilitate the characterization of future CBM families.

However, accurately predicting CBM functions in the context of entire CAZymes remains difficult due to the spatially coordinated relationship between CBMs and their associated domains. For example, the linkers connecting CBMs with their associated domains influence the roles of CBMs in the whole enzymes. Previous studies showed that flexible linkers were commonly observed between the CBMs and their cognate domains to provide more conformational flexibility and that both the linker length and flexibility were associated with catalytic

* Corresponding author at: MOE Key Laboratory of Marine Genetics and Breeding, College of Marine Life Sciences, Ocean University of China, Qingdao 266003, China.
E-mail address: liuweizhi@ouc.edu.cn (W. Liu).

activity [5, 6]. Additionally, a more extended rigid linker domain (97 residues) was observed, suggesting the codependent relationship between the CBMs and the associated domains [7]. Therefore, to better understand the diverse roles of CBMs in the entire enzymes, more full-length structures from different CAZymes and biochemical characterization are required.

Polysaccharide lyases (PLs) are the crucial enzymes in metabolic polysaccharide degradation, especially that of pectin and alginates [8, 9]. While CBMs are commonly discovered in PLs, their role in PL function is poorly understood [10]. Alginates are linear polymers of two uronates, β -D-mannuronate (M) and α -L-guluronate (G), linked by the 1,4-glycosidic bond yielding various sequences. They can be degraded by alginate lyases via β -elimination reactions, which are divergent in several PL groups (i.e., PL 5, 6, 7, 14, 15, 17 and 18) [11]. Among them, PL7 alginate lyases have attracted the most attention, and several CBM32-containing PL7 alginate lyases have been characterized. However, the role of CBM32 on the PL domain remains unclear.

Recently, we found the CBM32-PL7 alginate lyase AlyB to be the most important enzyme for alginate utilization in *Vibrio spendidus* OU02 (unpublished data). The AlyB structure determined herein reveals that CBM32 and the PL7 catalytic domain are connected by an unexpected alpha helix linker, which represents a new type of linker between CBMs and the catalytic domain. More importantly, comprehensive approaches were applied to determine the novel function of the CBM32 domain. Together with the unique alpha helix linker, CBM32 specifies the AlyB products to be predominantly trisaccharides.

2. Material and methods

2.1. Enzymatic activity assays

Alginate was purchased from Shanghai Yuanye Co., Ltd. The alginate viscosity is 4500 cp at the 2% (w/v) concentration at 25 °C, and the M/G ratio of alginate is 0.6, as determined by NMR analysis. PolyM and polyG were prepared according to the method described by Haug et al. with some modifications [12]. Briefly, HCl was added to the alginate solution (2% (w/v)) to a final concentration of 1 M. The mixture was incubated at 90 °C for 6 h. The polyG and polyM fractions were then separated by adjusting the pH value. The polyG fraction precipitates at pH 2.85, while the polyM fraction precipitates at pH 1.0.

The alginate lyase activity was determined by measuring an increase in absorbance at 235 nm due to the formation of unsaturated double bonds at the nonreducing end of the sugar [13]. The activity assays were carried out in Tris-HCl buffer (50 mM, pH 7.5) with 200 mM NaCl containing 0.2% (w/v) alginate. The 1-ml reaction mixture was incubated for 5 min at 25 °C. One unit of enzyme activity was defined as an A_{235} increase of 1 per min.

For kinetic analysis, 1.58 μ M enzymes were used, and the substrate concentrations varied from 0.15 mg/ml to 8 mg/ml. The reactions (200 μ l for each assay) were carried out in Tris-HCl buffer (50 mM, pH 7.5) with 200 mM NaCl. After incubation at 25 °C for 5 min, 200 μ l of dinitrosalicylic acid was added to the reaction mixture, followed by heating at 100 °C for 5 min and centrifugation. The absorbance of the resulting supernatant was measured at 520 nm. One unit of alginate lyase was defined as the amount of enzyme required to liberate 1 μ mol of reducing sugar per min.

2.2. Gene cloning, mutagenesis, expression and protein purification

The gene encoding AlyB was amplified from genomic *Vibrio spendidus* OU02 DNA (unpublished data) and inserted into the pET32a vector to obtain a recombinant construct. The resulting recombinant plasmid was transformed into the *E. coli* BL21(DE3) expression strain, which was cultured at 37 °C until an OD_{600} of 0.6–0.8. Then, protein expression was induced with 0.2 mM isopropyl- β -D-thiogalactopyranoside (IPTG), and the culture was further incubated overnight at 16 °C.

After harvesting by centrifugation (1500 g, 30 min, 4 °C), the cell pellets were resuspended in Tris-HCl buffer (50 mM, pH 7.5) with 500 mM NaCl and disrupted by sonication, and the lysates were clarified by centrifugation (20,000 g, 30 min, 4 °C). The protein was purified from the resulting crude extract using affinity chromatography (Ni-Agarose Resin, CWBio), and protease (prepared in our lab) was used to cleave the 6 $\times$ His tag from the fusion protein. For crystallographic studies, after affinity purification, AlyB and CBM32 were further purified on a Sephacryl-S-100 column previously equilibrated with 20 mM Tris-HCl (pH 7.5) and 200 mM NaCl. The protein concentrations were then determined based on their absorption at 280 nm using molar extinction coefficients calculated by the online ExPASy server. Three truncation versions of AlyB (CBM32, CD1 and CD2) were prepared using the same method.

Mutants (e.g., W129A) were generated using the QuikChange® site-directed mutagenesis method (Agilent Technologies) and verified by DNA sequencing. Then, these mutants were expressed and purified similarly to AlyB.

To generate the AlyB^{dis} construct, a disordered linker comprising 38 amino acids from a glycoside hydrolase (ABP56033.1) was selected to replace the alpha helix linker of the AlyB (residues 152–180). The nucleotide sequence of the disordered linker was synthesized by Shanghai Sangon Biotech Co., Ltd. Then, three DNA fragments of CBM32, the disordered linker and CD2 were ligated to make one unit (AlyB^{dis}) using the overlap extension PCR. After verification by DNA sequencing, AlyB^{dis} was expressed and purified as described above.

2.3. Crystallization, data collection and structure refinement

Crystals were screened using Hampton screen kits (Screen I, II) and JCSG+ (Qiagen) by the hanging drop vapor diffusion method. AlyB was concentrated to 10 mg/ml before screening. The crystals grew for approximately 2–3 days, and after optimization, the crystals used for diffraction were grown at 15 °C in 0.2 M NaAC, 0.1 M Tris-HCl (pH 8.0) and 25% PEG4000. To explore the interactions between CBM32^{AlyB} and the substrate, we attempted to cocrystallize CBM32^{AlyB} with oligomeric alginate. The oligomers were prepared by alginate degradation using AlyB as described in the following section. The concentrated CBM32^{AlyB} (25 mg/ml) was incubated with the oligomers at 4 °C for 1 h before screening with the kits described above. The crystals were acquired in buffer containing 2.4 M sodium malonate (pH 7.0). The crystals were picked up and dipped for a few seconds in cryoprotectant (mother solution containing 10% glycerol) prior to flash-freezing in liquid nitrogen.

Diffraction data were collected at the Shanghai Synchrotron Radiation Facility (SSRF) on the BL17/19 U beamline and processed using HKL2000 [14]. For AlyB, the initial structure solution was obtained by molecular replacement using both the alginate lyase from *Klebsiella pneumoniae* (PDB: 4OZX) and CBM32 (PDB: 2JD9) as search templates with a phaser [15]. For CBM32^{AlyB}, the initial structure solution was obtained by molecular replacement using the determined CBM32 domain (residues 1–151) in AlyB. The overall structure was then iteratively refined using REFMAC5 [16] in combination with manual model building in Coot [17]. The final structural figures were generated using CCP4mg [18], and the coordinates of the final complex will be deposited in the Protein Data Bank (PDB, Table 1).

2.4. Products analysis

All enzymatic degradation reactions were conducted in 1 ml of buffer containing 50 mM Tris-HCl (pH 7.5) and 200 mM NaCl at 25 °C using 0.2% (w/v) alginate as the substrate, and the enzyme concentrations were 1.6 μ M. These reactions were terminated by heating at 100 °C for 5 min followed by centrifugation.

Table 1
Protein structure data processing and refinement statistics.

Data collection	AlyB	CBM32 ^{AlyB}
Space group	<i>P</i> 2 ₁ 2 ₁ 2 ₁	<i>P</i> 2 ₁ 2 ₁ 2 ₁
Cell dimensions, <i>a</i> , <i>b</i> , <i>c</i> (Å)	65.3, 78.4, 112.3	43.8, 60.7, 66.9
Cell dimensions, α , β , γ (°)	90.0, 90.0, 90.0	90.0, 90.0, 90.0
X-ray source	SSRF BL 19U1	SSRF BL 19U1
Wavelength (Å)	1	1
Resolution (Å) ^a	50–1.60 (1.66–1.60)	50–1.40 (1.45–1.40)
Unique reflections ^a	76,839	35,712
Completeness (%)	99.8 (99.9)	100 (100)
<i>R</i> _{sym} (%) ^a	12.1(74.0)	12.4(46.7)
<i>Mn</i> (I/σ) ^a	22.2 (4.3)	21.7(6.7)
Redundancy ^a	13.7(13.4)	12.5(11.8)
Refinement statistics		
Resolution (Å) ^a	50.0–1.60(1.64–1.60)	44.9–1.4(1.44–1.40)
<i>R</i> _{factor} (%)	19.7(24.6)	18.9(24.0)
Free <i>R</i> _{factor} (%)	23.6(25.6)	21.1(25.0)
Residues built (range) total	Chain A: 1–493	Chain A: 1–152
Free <i>R</i> reflections (%)	5.0	5.0
Free <i>R</i> reflections no.	3846	1786
No. nonhydrogen atoms	4436	1405
No. water molecules	589	214
Model quality		
RMSD bond length (Å)	0.02	0.02
RMSD bond angles (°)	1.81	1.92
Mean B-factors		
Overall (Å ²)	16.4	14.1
Protein atoms (Å ²) (A)	14.9	14.1
Water (Å ²) (Z)	25.7	26.0
Solvent (Å ²) (B/C)	18.1/11.3	12.1
Ramachandran plot (%) favored/allowed/disallowed	97.8/2.2/0	99.0/1/0

^a Parentheses indicate the highest resolution.

2.4.1. TLC assay

Degradation products were analyzed by thin layer chromatography (TLC) on Silica gel 60F254* (Merck) plates using n-butyl alcohol: formic acid: water (4:6:1, v/v/v) as the solvent. The oligomeric products were visualized by a reagent comprising 1 ml of HCl, 2 ml of aniline, 10 ml of H₃PO₃, 100 ml of acetone and 2 g of diphenylamine.

2.4.2. SEC assay

The product distribution patterns of AlyB, CD2 and AlyB^{dis} were also analyzed by size exclusion chromatography (SEC). Products were loaded onto a Sephadex G15 column connected to an ÄKTA purifier system monitored at 235 nm. The column was eluted with 0.2 M ammonium bicarbonate at a flow rate of 0.3 ml/min.

2.4.3. MS assay

Negative ion electrospray ionization mass spectrometry (ESI-MS) was used to determine the molecular masses of some products.

2.5. Affinity electrophoresis

The alginate binding activities of CBM32 and Y466A were evaluated by affinity electrophoresis [3, 19]. BSA was used as a noninteracting control. The stacking gels contained 5% w/v polyacrylamide, and the separating gels contained 13% w/v polyacrylamide with or without 0.2% alginate. Each protein (5 μg) was loaded onto gels, and the gels were electrophoresed simultaneously for 2 h at 4 °C using a constant voltage of 120 V. Then, the gels were stained with Coomassie Blue for protein visualization.

2.6. Fluorescence polarization (FP) assay

As described above, the purified trisaccharides were prepared by AlyB alginate degradation, followed by size exclusion chromatography. The trisaccharides were labeled with fluorescein 1-aminionpyrene-

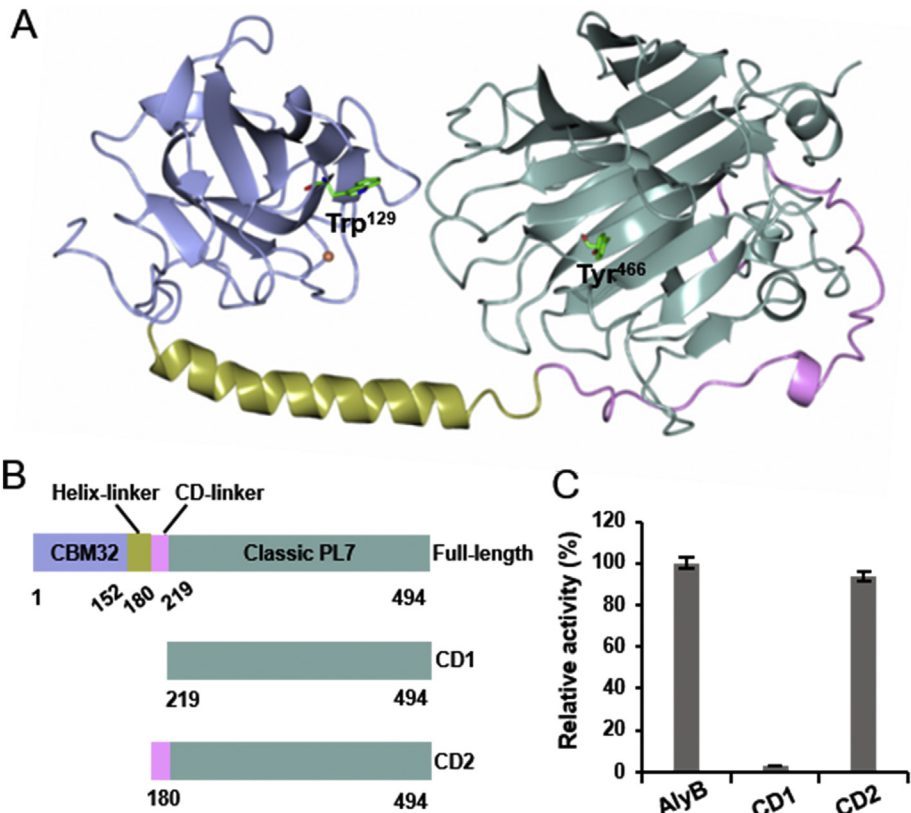


Fig. 1. Full-length crystal structure of AlyB revealing new features of the modular domain organization. (A): The full-length AlyB crystal structure (PDB: 5ZU5) is presented as four colors corresponding to the four regions highlighted in (B). Two key residues, Trp¹²⁹ (located in CBM32 and critical for alginate binding) and Tyr⁴⁶⁶ (located in CD and critical for catalysis), are shown as stick models. (B): AlyB and its two truncations CD1 and CD2 are presented with different modules. The CBM32 domain is in ice blue, the catalytic domain is in sea green, the alpha helix linker is in yellow, and the CD linker is in pink. (C): Enzymatic activities measured for the full-length AlyB, CD1 and CD2.

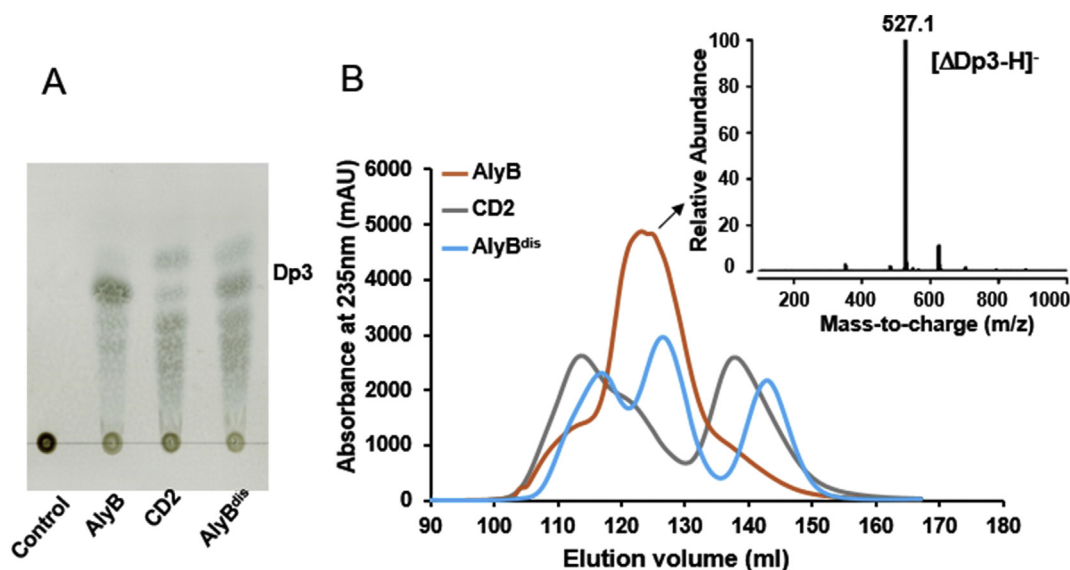


Fig. 2. Importance of CBM32 and the rigid alpha helix linker in specifying that the products of AlyB are primarily trisaccharides.

(A): TLC separation of reaction products for the control (no enzymes), AlyB (full-length), CD2 (no CBM32 and helix linker), and AlyB^{dis} (AlyB with CBM32 and a modified disordered linker in place of the alpha helix linker).

(B): Similar to (A), but size exclusion chromatography was used to analyze the products of AlyB, CD2, and AlyB^{dis}. Together with (A), this figure shows that AlyB produces predominantly trisaccharides (also confirmed with mass spectrometry analysis [arrow pointing to]. Δ : denotes an unsaturated residue, 4-deoxy-L-erythrohex-4-enopyranosyluronic, at the nonreducing end of the products; Dp: degree of polymerization). However, if CBM32 is removed or the alpha helix linker is replaced, the product size specificity is lost.

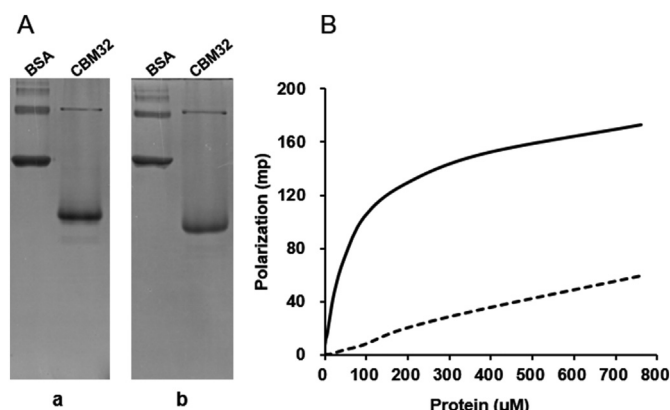


Fig. 3. Substrate binding profiles of CBM32.

(A): Affinity gel electrophoresis demonstrates that CBM32 is incapable of binding alginate. (a): Native polyacrylamide gel without alginate. (b): Native polyacrylamide gel with alginate (0.2% w/v). BSA was used as the non-interacting control.

(B): Fluorescence polarization (FP) assay. The observed increase in the FP signal for the samples containing CBM32 and trisaccharides indicates the CBM32 can bind oligo-alginate. However, such a curve was not observed when Trp¹²⁹ was substituted with Ala.

3,6,8-trisulfonate (APTS; $\lambda_{\text{exc}} = 455 \text{ nm}$, $\lambda_{\text{em}} = 512 \text{ nm}$) via a reductive reaction [20]. The trisaccharide solution (1 M; 5 μl) was mixed with 10 μl of 0.5 M APTS in 10% acetic acid and 0.5 M sodium cyanoborohydride and incubated at 45 °C for 4 h. After incubation, the solute molecules were precipitated twice with a 10-fold excess of ethanol and stored at –20 °C before analysis [21].

The APTS-labeled trisaccharides and proteins (wild-type CBM32 and the W129A mutant) were dissolved in buffer containing 20 mM Tris-HCl (pH 7.5) and 500 mM NaCl. Next, 100 μl aliquots of the APTS-labeled trisaccharide solutions (250 nM) were transferred to each well of a black 96-well microplate. Then, 100 μl of the wild-type CBM32 or W129A mutant (concentrations ranging from 0.37 μM to 760 μM) was added to the total sample volume of 200 μl . Control wells contained

100 μl of the APTS-labeled trisaccharide solution and 100 μl of buffer. After 1 h of incubation, the fluorescence polarization signals were measured using the Synergy™ H1 multimode microplate reader (BioTek). All experiments were performed in triplicate. The FP value of each protein concentration was calculated by subtracting the measurements of the control wells [22].

2.7. Molecular docking

Docking studies were performed using the Schrödinger suite (version 2016) (Schrödinger, LLC, New York, NY, 2016). The crystal structure of AlyB was prepared using the “Protein Preparation Wizard” module in Schrodinger Maestro (version 11.0) with the default protein parameters. Hydrogen atoms were added, hydrogen bonds were optimized, partial charges for all atoms were assigned, and hydrogen atoms were energy-minimized using the OPLS-2005 force field. Finally, all crystal water was removed. To set up the receptor grid with Schrodinger Glide, the W129 and Y466 center of mass was defined as the grid center; the size of the site was set to 36 Å, which is large enough to accommodate the hexamer polysaccharide; the hydrogen bond constraint was applied to Y466.OH; and positional constraints were applied to the Y466.OH and W129 center of mass, respectively. The hexamer polyM was built using CarBuilder (version 2.2.18) [23], the protonation state was calculated using the Schrodinger Epik, and 20 conformations of the hexamer were generated using Schrodinger ConfGen. All conformers were docked to the prepared protein structure using Glide (docking parameters: standard precision, dock flexibly, all constraints in grid were activated).

3. Results and discussion

3.1. The first full-length AlyB protein structure is determined

Sequence alignment showed that AlyB-OU02 is highly identical (> 70%) to the AlyBs from *Vibrio splendidus* 12B01 and *Vibrio alginovor* SA2 (Fig. S1) [9, 24], which indicates that AlyB proteins (CBM32-PL7) are highly conserved and may play important roles in the *Vibrio* species.

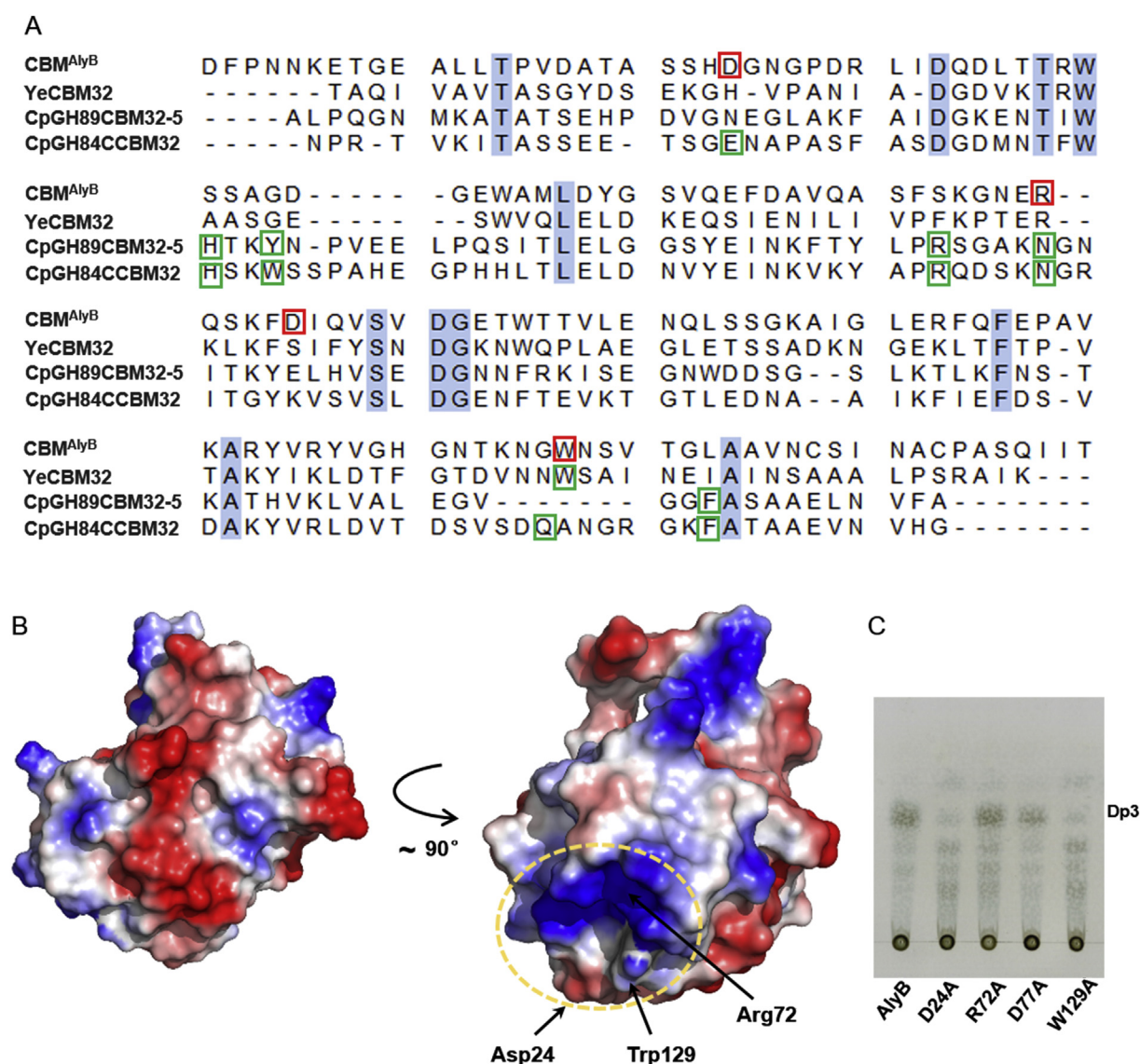


Fig. 4. Substrate binding mechanism of CBM32.

(A): Sequence alignment of CBM32^{AlyB} with three characterized CBM32 modules. Identical residues are shaded in light blue. The characterized key residues for substrate binding are highlighted in green boxes. Based on the alignment, four residues in CBM32^{AlyB}, highlighted in red boxes, are potential substrate binding sites and were selected for mutagenesis analysis.

(B): Electrostatic surface analysis of the CBM32 structure. Three potential substrate binding residues (Asp²⁴, Arg⁷² and Trp¹²⁹) are located in a positively charged pocket, labeled with a yellow circle.

(C): Mutagenesis analysis showing that changing Asp²⁴ or Trp¹²⁹ to Ala yields products significantly different from those of wild-type AlyB. These two residues must be critical for substrate anchoring.

The AlyB full-length protein structure (PDB: 5ZU5) was solved at a resolution of 1.6 Å (Table 1) for the first time. As shown in Fig. 1A and B, the full-length protein contains an N-terminal CBM32 domain (ice blue) and a C-terminal PL7 domain (sea green). The CBM32 domain (Asp¹-Thr¹⁵²) contains eight β-strands and a small fraction of α-helices between the first and second strand, which represents a typical CBM fold [25]. The CBM32 domain has the lowest r.m.s.d of 1.6 Å with YeCBM32 from *Yersinia enterocolitica* (PDB: 2JDA) based on an online Dali server search [26]. The C-terminal PL7 catalytic domain (Asn²¹⁹-Lys⁴⁹⁴) shows a typical β-sandwich jelly roll fold structure, which is composed of two layers of antiparallel β-strands stacked against each other (Fig. 1A). This domain displays the lowest r.m.s.d of 1.2 Å with an alginate lyase from *Klebsiella pneumoniae* (PDB: 4OZX).

3.2. AlyB has a longer PL7 catalytic domain, and its CBM32 domain does not enhance its enzymatic activity

According to sequence alignment with the published PL7 structures, the AlyB PL7 domain boundary was initially defined to be 219–494. However, our newly determined full-length structure (PDB: 5ZU5) suggested that the Arg¹⁸¹-Asn²¹⁸ region might also be a part of the catalytic domain (pink in Fig. 1A and B). To dissect the role of this short region, two catalytic domain (CD) truncation versions were developed (Fig. 1B). The first CD truncation, termed CD1, contained residues Asn²¹⁹ to Lys⁴⁹⁴, while the second CD truncation, termed CD2, was longer, containing residues Arg¹⁸¹ to Lys⁴⁹⁴.

Enzymatic assays (Fig. 1C) showed that without CBM32 and the short region (residues Arg¹⁸¹-Asn²¹⁹), named herein CD linker, CD1 has almost no catalytic activity on alginates. Surprisingly, and more interestingly, with this CD linker but without CBM32, CD2 has catalytic

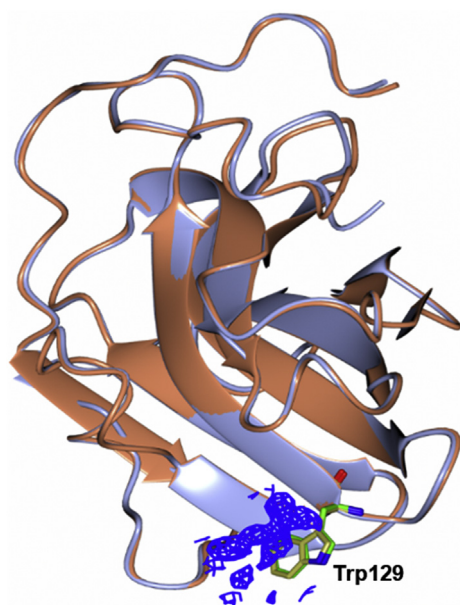


Fig. 5. Electron density map around Trp¹²⁹. The electron density map around Trp¹²⁹. The CBM32^{AlyB}-substrate complex structure was solved (PDB: 5Y1C). The electron density map (fo–fc map) around the potential substrate binding residue Trp¹²⁹ is shown in blue mesh at the 2.5 σ level. Superposing the CBM32^{AlyB} domain structure (PDB: 5Y1C) and the AlyB full-length structure (PDB: 5ZU5) induces no significant change upon substrate binding even in the substrate binding pocket around Trp¹²⁹.

activity very similar to that of the full-length AlyB. Structural superimposition of the CD2^{AlyB} structure over the published Aly structures (top PDB structures based on a search using the Dali server [26]) demonstrated that CD2^{AlyB} is the first determined Aly structure with this unique but critical region (Fig. S2). As shown in Fig. 1A, this region (Arg¹⁸¹-Asn²¹⁹) forms extensive contacts with the PL7 domain within a 146 Å² buried area, which probably stabilizes the intact substrate binding pocket for the catalytic domain, and is therefore critical for the catalytic efficiency. In addition, the enzymatic kinetics assay shows that both the full-length AlyB and CD2 have similar *K_m* and *k_{cat}* values (0.97 mg/ml and 3900 min^{−1} for AlyB; 1.21 mg/ml and 4000 min^{−1} for CD2), indicating that the presence of CBM32 does not significantly increase the enzymatic activity of AlyB. These observations indicated that (i) the CBM32 domain has little effect on AlyB's enzymatic activity, and (ii) the short region (Arg¹⁸¹-Asn²¹⁹) is essential for the catalytic activity of AlyB and should be a part of its PL7 CD domain (Fig. 1A).

3.3. A unique alpha helix linker is discovered between CD2 and CBM32

Another very interesting finding from the new full-length AlyB structure was a unique alpha helix linker (Ser¹⁵³-Leu¹⁸⁰) between CD2 and CBM32 (yellow in Fig. 1A and B). AlyQ from *Persicobacter* sp. CCB-QB2 is also a multidomain alginate lyase that contains two CBMs (CBM16 and CBM32) and a PL7 catalytic domain. The most recently determined truncated AlyQ structure demonstrated a typical disordered linker between CBM32 and the catalytic domain [27]. Such a disordered linker is common to the previously reported linker regions between CBMs and their associated modules, e.g., the linker in *Clostridium thermocellum* (ct) CBM22_GH10 structure (PDB: 2W5F) and an unknown rumen bacterium GH29_0940 structure (PDB: 5K9H). These disordered linker regions provide a necessary flexibility between these domains, which is essential for their binding and catalytic activities [28, 29]. Alternatively, a relatively rigid linker domain containing 97 residues from SpuA has been reported [7]. By contrast, our newly determined alpha helix linker is significantly different from all previously reported linker regions between CBMs and catalytic domains.

Intuitively, the alpha helix linker, named herein helix linker, will result in more constraints in the AlyB structure. Due to the rigid helix linker, the CBM32 does not intimately interact with CD2 (Fig. 1A), which has a small buried area of 46 Å². This is also different from all previously determined CBM catalytic domain structures. For example, the ctCBM22-GH10 complex structure has a putative xylan binding pocket that forms between the CBM22 and GH10 domains with a buried area of 900 Å² [28]. This apparent difference indicates that this unique helix linker may play a different role.

3.4. CBM32 and helix linker are responsible for the preferred trisaccharide products

Knowing that the CBM32 domain and the helix linker do not contribute to the enzymatic activity of full-length AlyB using alginate as a substrate (Fig. 1C), we assessed whether CBM32 contributes to the substrate cleavage specificity. Therefore, both polyM and polyG were used as substrates to investigate whether CBM32 alters the substrate cleavage specificity of AlyB. However, the enzymatic assay showed that AlyB and CD2 still have similar enzymatic activities using polyM or polyG as substrates (Fig. S3).

Since the CBM32 domain and the helix linker are not essential for the enzymatic activity and substrate specificity of AlyB, we further analyzed the product profiles of AlyB, CD2 and AlyB^{dis} to study the function of CBM32 and the helix linker. AlyB^{dis} is an AlyB construct in which the helix linker region is replaced with a disordered loop (see Material and Methods). Both TLC (Fig. 2A) and size exclusion chromatography (Fig. 2B) analyses showed that AlyB produces primarily trisaccharides (see also the MS result in Fig. 2B pointed by arrows). By contrast, both CD2 and AlyB^{dis} generate oligomers of various lengths with no preference for trisaccharides (Fig. 2A and B). Therefore, we conclude CBM32 in combination with the helix linker in AlyB functions to predominantly produce trisaccharides rather than to increase the enzymatic activity. Obviously, this function of CBM32 in AlyB of OU02 is very unusual compared to those of all previously reported CBMs [1, 25].

3.5. Biochemical features of CBM32 that are relevant to its unique role

As described above, CBM32 and the helix linker do not significantly enhance the catalytic activity but rather specifies that trisaccharides are predominant in the degradation products. Thus, whether and how CBM32 interacts with the substrate must be determined. First, affinity electrophoresis was performed to monitor the alginate binding profiles of CBM32. However, alginate binding by CBM32 was not detected (Fig. 3A). This result appeared to be consistent with the observation that the full-length AlyB and CD2 have similar *K_m* values, but it contradicted the fact that the presence of CBM32 in AlyB influenced the degradation product distribution (Fig. 2A, B), which suggested that CBM32 should interact with the substrate. Therefore, the fluorescence polarization (FP) assay was utilized to monitor whether CBM32 can bind oligo-alginate. In this study, the alginate oligomers (trisaccharides) were fluorescently labeled with APTS at their reducing ends. As shown in Fig. 3B, the observed FP signal indicated that an interaction between CBM32 and trisaccharides did in fact occur.

To elucidate the binding mechanism of CBM32 to the oligo-alginate, both sequence alignment and cocrystallization assays were carried out. Several potential substrate binding sites, i.e., Asp²⁴, Arg⁷², Asp⁷⁷ and Trp¹²⁹, were identified according to the sequence alignment of CBM32^{AlyB} against three previously determined CBM32 structures (PDB: 2JD9, 2J1A, 4A45) (Fig. 4A). Furthermore, electrostatic surface analysis showed that Asp²⁴, Arg⁷² and Trp¹²⁹ were located in a positively charged surface area (Fig. 4B), which might form a binding pocket for the negatively charged alginate substrates. Furthermore, four amino acid residues were mutated in AlyB (Asp²⁴/Ala, Arg⁷²/Ala, Asp⁷⁷/Ala and Trp¹²⁹/Ala) to evaluate their roles in the full-length

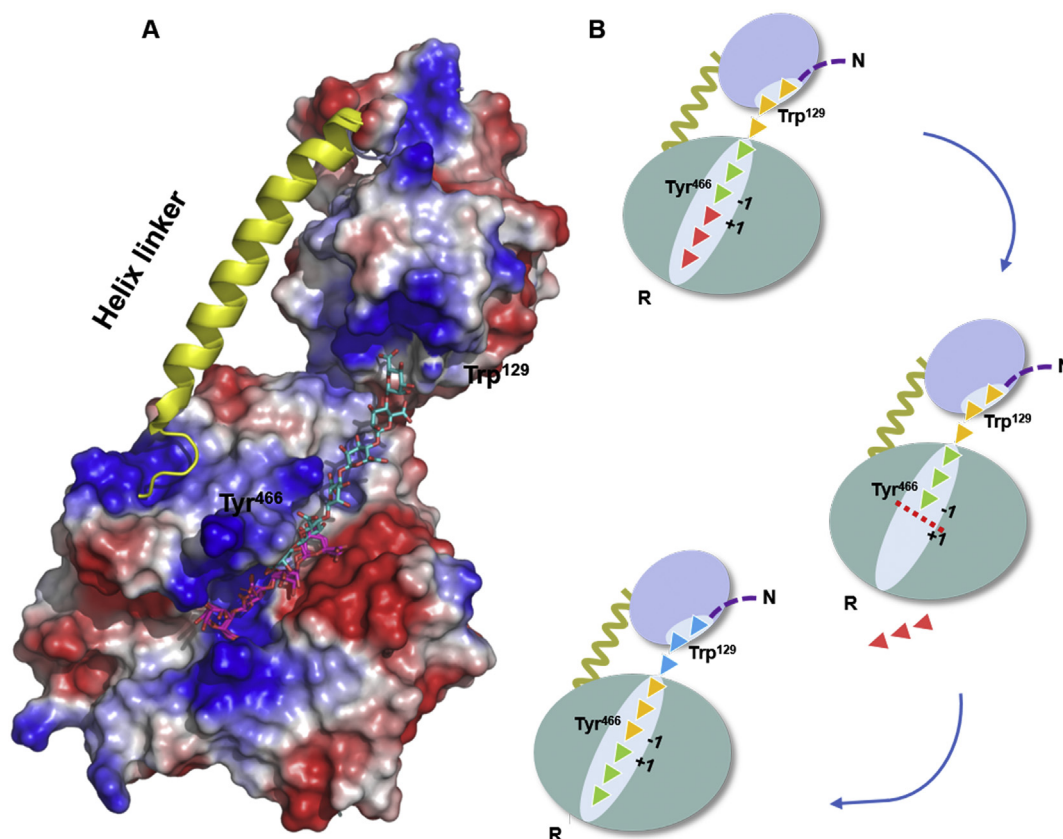


Fig. 6. Molecular docking of alginate onto the AlyB full-length structure helped to construct a trisaccharide release model.

(A): Molecular docking of alginate (hexameric polyM) onto the AlyB structure. The Schrödinger suite (2016) was used for docking. Superposition of the PDB protein-substrate complex PL7-GGMG (PDB: 2ZAA) helped to determine the best substrate binding conformation. PolyM is shown as cyan sticks, and GGMG is shown as pink sticks.

(B): The proposed model for trisaccharide release. The monomer unit of alginate is shown as a triangle. The red dashed line indicates that the cleavage happens between the -1 and $+1$ sites. The sliding direction of the substrate relative to the enzyme is from the nonreducing end (N) to the reducing end (R).

protein. The D24A and W129A mutants produced product profiles that were substantially differed from that of AlyB (but similar to that of CD2) (Fig. 4C), suggesting the importance of Asp²⁴ and Trp¹²⁹ in determining the preference of AlyB to produce trisaccharides.

We also attempted to cocrystallize AlyB with alginates but failed. However, we were able to cocrystallize the CBM32 domain with an alginate trisaccharide and obtained a high-resolution (1.4 Å) structure of the CBM32^{AlyB} + substrate complex (PDB: 5Y1C, Table 1). In this complex structure, an additional electron density around Trp¹²⁹ was observed (Fig. 5). Although the poorly resolved density did not allow us to define the sugar unit perfectly (Fig. 5), the result suggests that the alginate trisaccharide is highly flexible and forms only weak interactions with CBM32. In addition, the interaction between Trp¹²⁹ and the substrate was further confirmed by mutagenesis. As shown in Fig. 3B, the loss of fluorescence polarization change for W129A toward the trisaccharide substrate indicates a significant role of Trp¹²⁹ in substrate binding. Additionally, the TLC analysis described above demonstrates that the weak interactions between CBM32 and the substrate are essential for product distribution (Fig. 4C).

3.6. Molecular docking for AlyB with an alginate hexamer

This new structure reveals two surprising findings: (i) the CBM32 domain does not contribute to the catalytic activity of the PL7 domain, and (ii) a rigid alpha helix linker instead of a flexible disordered loop linker exists between the CBM32 domain and the PL7 domain. Further biochemical data underscore the importance of CBM32 and the helix linker.

To shed some light on the catalytic features of AlyB, molecular docking was carried out for an alginate hexamer docked to the AlyB structure. According to a structural comparison between our AlyB structure (PDB: 5ZU5) and a published structure (PDB: 2ZAA), Tyr⁴⁶⁶ in AlyB was predicted to be the catalytic residue involved in the cleavage of the alginate glycosidic linkage. Indeed, mutagenesis showed that the Y466A mutation negated the catalytic activity but did not affect the substrate binding ability (Fig. S4). With constraints from the two key amino acid sites, Trp¹²⁹ in CBM32 and Tyr⁴⁶⁶ in the catalytic domain, we docked an alginate hexamer polyM (poly-β-D-mannuronate) to the AlyB structure (Fig. 6A). A superposition of the modeled AlyB-polyM structure and the published PL7-GGMG complex structure (PDB: 2ZAA) showed a reasonable consistency at the catalytic active center (Fig. 6A). According to this modeled AlyB-polyM structure, the rigid alpha helix linker holds two substrate binding pockets, one in CBM32 and the other in the catalytic domain, in an appropriate orientation. The modeled structure shows that the CD domain of AlyB alone forms a well-defined substrate binding cleft (Fig. 6A), which suggests that AlyB likely adopts a processive degradation model [30].

3.7. A product release model for AlyB

Taken together, the oligomer release model is depicted in Fig. 6B, where the sliding direction is from the nonreducing end to the reducing end. The substrate sites $+1$, $+2$ and $+3$ at the reducing end are located in the catalytic domain, serving as the product releasing sites. After the cleavage between -1 and $+1$, a trisaccharide is released into solution. Then, the substrate slides into the binding pocket of the catalytic

domain to position the next trisaccharide for enzymatic cleavage (Fig. 6B). During this processive movement, the rigid alpha helix linker keeps the CBM32 and CD2 at an appropriately mutual orientation and distance between the two binding pockets to ensure that the alginate substrate enters the catalytic cavity properly. Therefore, CBM32 and the alpha helix linker together serve as the “pivot point” influencing product distribution, resulting in consistent trisaccharide release.

This process should be strictly controlled with precise cooperation between CBM32, the catalytic domain, and the helix linker and is essential for the production of specific trisaccharides. However, traces of oligomers of sizes other than three also exist. This might be because alginate substrates shorter than a certain length interact with only the catalytic domain and are no longer constrained by the CBM32 domain. This is supported by the observation that at least 9-mers are needed to interact with both CBM32 and CD (Fig. 6A).

4. Conclusions

In summary, the determined AlyB structure helped clarify the boundary of the catalytic domain and revealed an alpha helix linker connecting CBM32 and the catalytic domain. This unique alpha helix linker was discovered for the first time, representing a new linker type. Further biochemical analysis determined that CBM32 and the helix linker do not contribute to the enzymatic activity of AlyB but rather assist in the predominant production of trisaccharides. CBM32 and the helix linker may serve as a “pivot point” to influence product distribution. This unique role enriches our understanding of the function of CBMs in the entire enzymes, and this study will facilitate the applications of CBMs in biotechnology areas.

Conflicts of interests

The authors have no conflicts of interest to declare.

Authors' contributions

W.L., Y.Y., Q.L., L.Z., T.L., designed the study; Q.L., K.Z., Q.Z., Z.L., E.F., W.L., and Y.L. performed the research; Q.L., W.L., and Y.Y. analyzed the data and wrote the paper.

Acknowledgements

We thank Prof. Liang Zhang of Shanghai Jiao Tong University and Prof. Qingjun Ma of the Institute of Oceanology (Chinese Academy of Sciences) for their kind assistance in crystallization data collection. The crystallization data were collected at the Shanghai Synchrotron Radiation Facility (SSRF) on beamlines BL17U and 19U. We also thank Dr. Zhenting Gao for his discussions on molecular docking. This work was supported by the National Natural Science Foundation of China [grant numbers 41706151, 31371725, 41376143, 31728013], the National Key Research and Development Program of China [grant number 2017YFB0309302], the Shandong Provincial Natural Science Foundation [grant number ZR2016DB17], the China Postdoctoral Science Foundation [grant number 861505020040], the China-ASEAN Maritime Cooperation Fund National Science Foundation [grant number DBI-1652164], and the US National Institutes of Health [grant number 1R15GM114706].

Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.bbagen.2018.05.024>.

References

- [1] D.W. Abbott, A.L. van Bueren, Using structure to inform carbohydrate binding

- module function, *Curr. Opin. Struct. Biol.* 28 (2014) 32.
- [2] D. Guillén, S. Sánchez, R. Rodríguez-Sanoja, Carbohydrate-binding domains: multiplicity of biological roles, *Appl. Microbiol. Biotechnol.* 85 (2010) 1241–1249.
- [3] D.W. Abbott, S. Hrynuk, A.B. Boraston, Identification and characterization of a novel periplasmic polygalacturonic acid binding protein from *Yersinia enterocolitica*, *J. Mol. Biol.* 367 (2007) 1023–1033.
- [4] H.J. Gilbert, J.P. Knox, A.B. Boraston, Advances in understanding the molecular basis of plant cell wall polysaccharide recognition by carbohydrate-binding modules, *Curr. Opin. Struct. Biol.* 23 (2013) 669–677.
- [5] I.V. Ossowski, J.T. Eaton, M. Czjzek, S.J. Perkins, T.P. Frandsen, M. Schülein, P. Panine, B. Henrissat, V. Receveurbréchet, Protein Disorder: conformational distribution of the flexible linker in a chimeric double cellulase, *Biophys. J.* 88 (2005) 2823–2832.
- [6] D.W. Sammond, C.M. Payne, R. Brunecky, M.E. Himmel, M.F. Crowley, G.T. Beckham, Cellulase linkers are optimized based on domain type and function: insights from sequence analysis, biophysical measurements, and molecular simulation, *PLoS One* 7 (2012) e48615.
- [7] A.V.B. Lammerts, E. Fickoblean, B. Pluvina, J.H. Hehemann, M.A. Higgins, L. Deng, A.D. Ogunniyi, U.H. Stroeder, N.W. El, R.D. Burke, The conformation and function of a multimodular glycogen-degrading pneumococcal virulence factor, *Structure* 19 (2011) 640–651.
- [8] D. Ndeh, A. Rogowski, A. Cartmell, A.S. Luis, A. Baslé, J. Gray, I. Venditto, J. Briggs, X. Zhang, A. Labourel, Complex pectin metabolism by gut bacteria reveals novel catalytic functions, *Nature* 544 (2017) 65.
- [9] A.J. Wargacki, E. Leonard, M.N. Win, D.D. Regitsky, C.N. Santos, P.B. Kim, S.R. Cooper, R.M. Raiser, A. Herman, A.B. Sivitz, et al., An engineered microbial platform for direct biofuel production from brown macroalgae, *Science* 335 (2012) 308–313.
- [10] M.L. Garron, M. Cygler, Uronic polysaccharide degrading enzymes, *Curr. Opin. Struct. Biol.* 28 (2014) 87–95.
- [11] H.M. Qin, T. Miyakawa, A. Inoue, R. Nishiyama, A. Nakamura, A. Asano, T. Ojima, M. Tanokura, Structural basis for controlling the enzymatic properties of polymannuronate preferred alginate lyase FIAlyA from the PL-7 family, *Chem. Commun.* 54 (2018) 555–558.
- [12] A. Haug, B. Larsen, O. Smidsrød, G. Eriksson, R. Blinc, S. Paušak, L. Ehrenberg, J. Dumanović, Studies on the sequence of uronic acid residues in alginic acid, *Acta Chem. Scand.* 21 (1967) 691–704.
- [13] F. Thomas, L.C.E. Lundqvist, M. Jam, A. Jeudy, T. Barbeyron, C. Sandström, G. Michel, M. Czjzek, Comparative characterization of two marine alginate lyases from *Zobellia galactanivorans* reveals distinct modes of action and exquisite adaptation to their natural substrate, *J. Biol. Chem.* 288 (2013) 23021–23037.
- [14] Z. Otwinowski, W. Minor, Processing of X-ray diffraction data collected in oscillation mode, *Methods Enzymol.* 276 (1997) 307–326.
- [15] A.J. McCoy, R.W. Grosse-Kunstleve, L.C. Storoni, R.J. Read, Likelihood-enhanced fast translation functions, *Acta Crystallogr. D Biol. Crystallogr.* 61 (2005) 458–464.
- [16] G.N. Murshudov, P. Skubák, A.A. Lebedev, N.S. Pannu, R.A. Steiner, R.A. Nicholls, M.D. Winn, F. Long, A.A. Vagin, REFMAC5 for the refinement of macromolecular crystal structures, *Acta Crystallogr. D Biol. Crystallogr.* 67 (2011) 355–367.
- [17] P. Emsley, K. Cowtan, Coot: model-building tools for molecular graphics, *Acta Crystallogr. D Biol. Crystallogr.* 60 (2004) 2126–2132.
- [18] E. Potterton, P. Briggs, M. Turkenburg, E. Dodson, A graphical user interface to the CCP4 program suite, *Acta Crystallogr. D Biol. Crystallogr.* 59 (2003) 1131.
- [19] P. Tömmé, A. Boraston, J.M. Kormos, R.A. Warren, D.G. Kilburn, Affinity electrophoresis for the identification and characterization of soluble sugar binding by carbohydrate-binding modules, *Enzym. Microb. Technol.* 27 (2000) 453–458.
- [20] F.N. Lamari, R. Kuhn, N.K. Karamanos, Derivatization of carbohydrates for chromatographic, electrophoretic and mass spectrometric structure analysis, *J. Chromatogr. B Analyt. Technol. Biomed. Life Sci.* 793 (2003) 15.
- [21] M. Stefansson, Assessment of alginic acid molecular weight and chemical composition through capillary electrophoresis, *Anal. Chem.* 71 (1999) 2373–2378.
- [22] Y. Du, Fluorescence polarization assay to quantify protein-protein interactions in an HTS format, in: Cheryl L. Meyerkord, Haian Fu (Eds.), *Methods in Molecular Biology*, Springer Science and Business Media, New York, 2015, pp. 529–543.
- [23] M.M. Kuttel, J. Stähle, G. Widmalm, CarBuilder: software for building molecular models of complex oligo- and polysaccharide structures, *J. Comput. Chem.* 37 (2016) 2098–2105.
- [24] H. Doi, Y. Tokura, Y. Mori, K. Mori, Y. Asakura, Y. Usuda, H. Fukuda, A. Chinen, Identification of enzymes responsible for extracellular alginate depolymerization and alginate metabolism in *Vibrio alginivorans*, *Appl. Microbiol. Biotechnol.* 101 (2017) 1581–1592.
- [25] A.B. Boraston, D.N. Bolam, H.J. Gilbert, G.J. Davies, Carbohydrate-binding modules: fine-tuning polysaccharide recognition, *Biochem. J.* 382 (2004) 769–781.
- [26] L. Holm, P. Rosenstrom, Dali server: conservation mapping in 3D, *Nucleic Acids Res.* 38 (2010) W545–W549.
- [27] P.F. Sim, G. Furusawa, A.H. Teh, Functional and structural studies of a multidomain alginate lyase from *Persicobacter* sp. CCB-QB2, *Sci. Rep.* 7 (2017).
- [28] S. Najmudin, B.A. Pinheiro, J.A. Prates, H.J. Gilbert, M.J. Romão, C.M. Fontes, Putting an N-terminal end to the Clostridium thermocellum xylanase Xyn10B story: crystal structure of the CBM22-1-GH10 modules complexed with xylohexaose, *J. Struct. Biol.* 172 (2010) 353–362.
- [29] E.L. Summers, C.D. Moon, R. Atua, V.L. Arcus, The structure of a glycoside hydrolase 29 family member from a rumen bacterium reveals unique, dual carbohydrate-binding domains, *Acta Crystallogr. F Struct. Biol. Commun.* 72 (2016) 750–761.
- [30] W.A. Breyer, B.W. Matthews, A structural basis for processivity, *Protein Sci.* 10 (2001) 1699.