



The Linker Region Promotes Activity and Binding Efficiency of Modular LPMO towards Polymeric Substrate

Aishwarya Srivastava,^a Pragya Nagar,^a Sumit Rathore,^b  Nidhi Adlakha^a

^aSynthetic Biology and Bioprocessing Laboratory, Regional Centre for Biotechnology, NCR-Biotech Cluster, Faridabad, Haryana, India

^bDepartment of Biotechnology, All India Institute of Medical Sciences, New Delhi, India

ABSTRACT Lytic polysaccharide monoxygenases (LPMOs) mediate oxidative degradation of plant polysaccharides. The genes encoding LPMOs are most commonly arranged with one catalytic domain, while a few are found tethered to additional noncatalytic units, i.e., cellulase linker and carbohydrate-binding module (CBM). The presence of CBM is known to facilitate catalysis by directing the enzymes toward cellulosic polymer, while the role of linkers is poorly understood. Based on limited experimental evidence, linkers are believed to serve merely as flexible spacers between the structured domains. Thus, this study aims to unravel the role of the linker regions present in LPMO sequences. For this, we analyzed the genome of *Botrytis cinerea* and found 9 genes encoding cellulose lytic monoxygenases (AA9 family), of which *BcAA9C* was overexpressed in cellulose-inducible conditions. We designed variants of f^{LPMO} (full-length enzyme) with truncation of either linker or CBM to examine the role of linker in activity, binding, and thermal stability of the associated monoxygenase. Biochemical assays predicted that the deletion of linker does not impact the potential of f^{LPMO} for catalyzing the oxidation of Amplex Red, but that it does have a major influence on the capability of f^{LPMO} to degrade recalcitrant polysaccharide substrate. Langmuir isotherm and SEM analysis demonstrated that linker domain aids in polysaccharide binding during f^{LPMO} -mediated deconstruction of plant cell wall. Interestingly, linker domain was also found to contribute toward the thermostability of f^{LPMO} . Overall, our study reveals that linker is not merely a spacer, but plays a key role in LPMO-mediated biomass fibrillation; these findings are broadly applicable to other polysaccharide-degrading enzymes.

IMPORTANCE The polysaccharide-disintegrating carbohydrate-active enzymes (CAZymes) are often found with multimodular architecture, where the catalytic domain is connected to an accessory CBM domain with the help of a flexible linker region. So far, the linker has been understood merely as a flexible spacer between the two domains. Therefore, the current study is designed to determine the role of linker in polysaccharide fibrillation. To conceive this study, we have selected LPMO as a model enzyme, as it is not only an industrially relevant enzyme but it also harbors a catalytic domain, linker region, and CBM domain. The present study highlighted the crucial and indispensable role of the linker region in mediating polysaccharide disintegration. Considering its role in binding, thermostability, and activity toward polysaccharide substrate, we propose linker as a potential candidate for future CAZyme engineering.

KEYWORDS linker, carbohydrate-binding module, CAZymes, cellulose, *B. cinerea*

With wide-ranging global challenges and environmental concerns, there is a pressing need to transition from a fossil fuel-driven economy to a more sustainable bio-economy (1–4). In this regard, cellulose has emerged as an attractive and renewable source of alternative energy (5, 6). This crystalline cellulose can be depolymerized to fermentable residues with the help of an efficient cocktail of enzymes (7–9).

Editor Ronald P. de Vries, Westerdijk Fungal Biodiversity Institute

Copyright © 2022 Srivastava et al. This is an open-access article distributed under the terms of the [Creative Commons Attribution 4.0 International license](https://creativecommons.org/licenses/by/4.0/).

Address correspondence to Nidhi Adlakha, nidhi.adlakha@rcb.res.in.

The authors declare no conflict of interest.

Received 21 December 2021

Accepted 31 December 2021

Published 26 January 2022

Recent advancement in the understanding of enzyme-based degradation has resulted in the development of efficient enzyme cocktails (10, 11). These hydrolytic consortia are essentially composed of fungal-origin, glycoside hydrolase enzymes, which are archived in a carbohydrate-active enzyme (CAZyme) database and divided into subfamilies based on their activities and active site residues (12, 13). Among these, endo-1,4- β -glucanases (cleave random internal bonds in cellulose chains), exo-1,4- β -glucanases (attack the reducing or nonreducing end of the cellulose polymer), and β -glucosidases (catalyze cellobiose-to-glucose conversion) (14–16) are observed to be major players in catalyzing biomass disintegration. Of late, it has been observed that complete cellulosic degradation requires the inclusion of lytic polysaccharide monoxygenases (LPMOs) in cellulase consortia (17, 18).

Discovered in 2010, lytic polysaccharide monoxygenases (LPMOs) are copper-dependent redox enzymes which require an external supply of electron donors for oxidative cleavage of polymeric substrate (19). This metalloenzyme inserts an oxygen atom at C1 or C4 of an intrachain cellobiosyl moiety, and this random internal cleavage in polymeric structure leads to the depolymerization of complex cellulose chain (20, 21). Based on their substrate specificity, LPMOs are archived in the CAZyme database in the AA (Auxiliary Activity) protein family under subfamilies AA9 to 11 and AA13 to 16: here, the AA9 family of proteins oxidatively cleaves cellulose (22), AA10 is involved in chitin disintegration, and AA13 mediates starch degradation (23).

Many LPMOs exist solely as catalytic modules, while a few are found tethered to additional noncatalytic modules, i.e., cellulase linker and carbohydrate-binding module (CBM). The presence of CBM facilitates catalysis by directing the enzymes toward cellulosic polymer; Chalak et al. (24) elucidated that the release of sugar was drastically reduced in CBM-truncated LPMO enzyme, indicating its role in assisting LPMO-mediated biomass fibrillation. The CBM is generally coupled to the catalytic domain of CAZymes via a linker region. These linkers are intrinsically disordered proteins, widely believed to be flexible spacers between the structured domains (25). However, it has only recently been realized that this linker region might support synergistic action of the catalytic and accessory domains. Linkers typically exist with variable length and little sequence conservation (26) and it is believed that each linker is tailored based on a specific structured domain type (27). There exists one common feature: bias toward specific amino acids such as Ser and Thr (27), which are generally glycosylated, contributing toward protection against proteolysis (28), expansion of the operating range of catalytic domain (29, 30), and/or binding to polymeric substrate (31). Additionally, cellulase linkers have been observed to improve temperature adaptation of cellulases, as seen in the case of Cel5G from *P. haloplanktis* (32). The cellulase linker has also shown to impart specificity to PoCel6A from *Penicillium oxalicum* (33). All these studies conclusively suggest that each cellulase linker has a unique, indispensable functional role in assisting the cellulose depolymerization process.

In spite of these findings, the impact of the linker region on the activity of modular LPMO has not been studied in detail. Courtade et al. (25) elucidated that the linker region acts as a connecting region between LPMO and CBM, which allows these domains to move independently of each other; however, other functional roles of the linker region in LPMO performance have not been revealed. Thus, considering the varied functionality of the linker region, the current study performed biochemical and biophysical analysis to understand its role in the activity and binding of LPMO toward polymeric substrate.

For this, we have mined the genome of phytopathogenic fungus *Botrytis cinerea* (*B. cinerea*) and identified a CBM containing AA9 LPMO (*BcAA9C*). The supplementation of full-length LPMO (f^{LPMO}) to a classical cellulase cocktail led to a 1.6-fold increase in glucose released from biomass disintegration, highlighting its importance in the saccharification of polysaccharide substrate. The enzyme possesses additional linker and CBM domains associated with catalytic domain. To understand the role of the linker and CBM domains associated with f^{LPMO} , we designed and expressed CBM- and linker-truncated

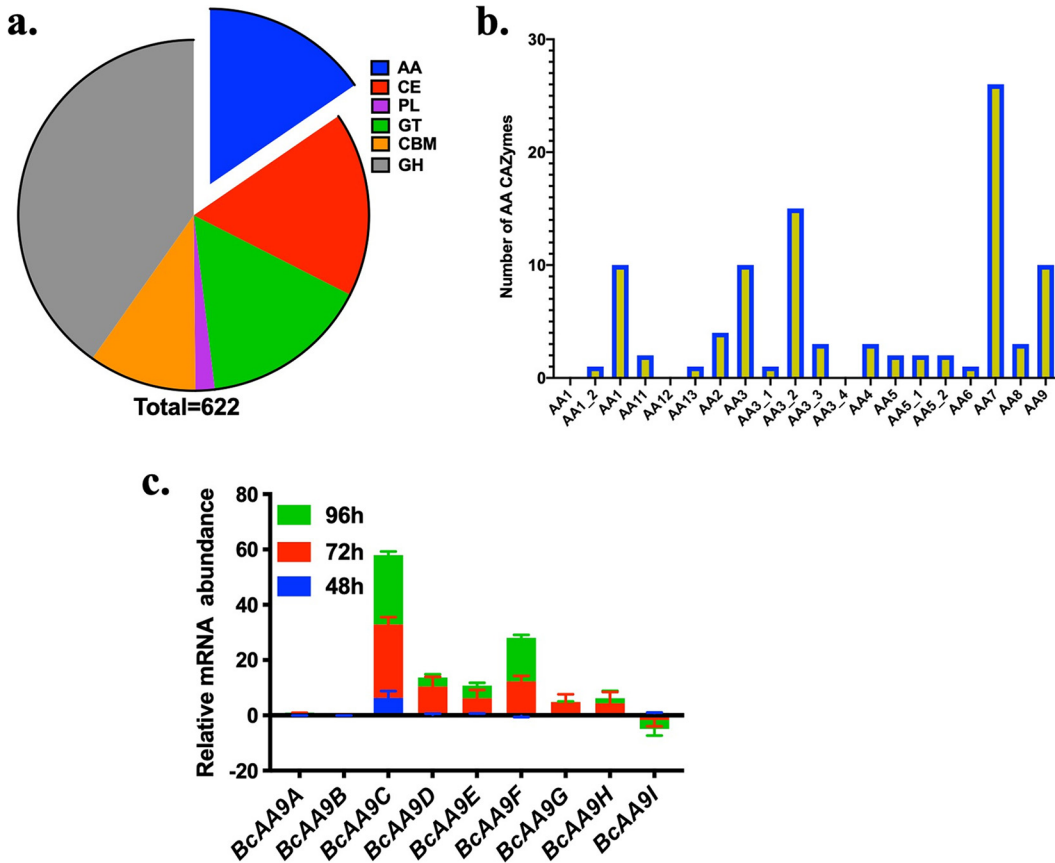


FIG 1 Polysaccharide monoxygenases encoded by *Botrytis cinerea*. (a) The genome sequence of *B. cinerea* 05.10 scrutinized for the presence of carbohydrate active enzymes (CAZymes) using dbCAN2 metaserver, which indicated the presence of 622 CAZymes. AA, Auxiliary Activity; CBM, carbohydrate-binding domain; PL, polysaccharide lyase; GT, glycosyl transferase; GH, glycosyl hydrolase; CE, carbohydrate esterase. (b) Bar graph indicates CAZymes belonging to different subfamilies of Auxiliary Activity, also revealing the presence of nine variants of polysaccharide monoxygenases which belong to the AA9 CAZyme family. (c) Relative mRNA abundance of nine *BcAA9* genes of *Botrytis cinerea* after 48 h, 72 h, and 96 h growth; *actin* was used as a housekeeping gene for RT-PCR. Error bar denotes standard deviation from three biological replicates.

enzymes in *Pichia pastoris*. Biochemical assays and biophysical techniques clearly indicated that the linker region facilitates binding and activity toward polymeric substrate. Furthermore, the presence of linker region was shown to contribute to the thermal stability of *fl*^{PMO}. Collectively, we have elucidated that the linker region present in *fl*^{PMO} is crucial for its activity on polymeric substrate.

RESULTS

***B. cinerea* harbors multiple CAZymes associated with Auxiliary Activity.** The dbCAN analysis of publicly available *B. cinerea* genome indicated the presence of 622 carbohydrate active enzymes (CAZymes), and the those belonging to the Auxiliary Activity (AA) family accounted for almost 15% (Fig. 1a). CAZyme scrutiny, supported with the CAZy database (cazy.org), also suggested the presence of nine variants of cellulose lytic monoxygenases (AA9 family) (Table S2) in contrast to 7, 3, and 2 variants in *A. niger*, *T. reesei*, and *T. cellulolyticus*, respectively (Fig. 1b). Thus, we selected *B. cinerea* as the preferred fungal isolate for isolating enzymes particularly involved in oxidative degradation of cellulose.

To identify outperforming AA9 variants, we analyzed the expression of all nine homologues at the transcript level. qRT-PCR analysis suggested a 30-fold increase in the expression of *BcAA9C*, while *BcAA9D*, *BcAA9E*, and *BcAA9F* showed 4- to 8-fold increases and *BcAA9A*, *BcAA9B*, and *BcAA9I* did not show any expression compared with the *actin* control (Fig. 1c). Overall, *B. cinerea* secreted multiple cellulose lytic

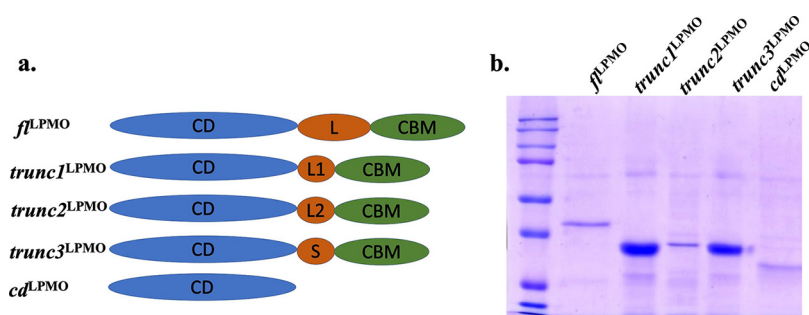


FIG 2 Design and purification of fl^{LPMO} and CD-CBM chimeric proteins. (a) Modular representation of *BcAA9C*-encoded LPMO variants, where fl^{LPMO} represents full-length protein (with catalytic domain, CD; linker, L; carbohydrate-binding module, CBM); $trunc1^{LPMO}$, $trunc2^{LPMO}$, and $trunc3^{LPMO}$ represent CD-CBM chimeric proteins [where either linker is replaced with L1/L2-truncated linker sequence, or spacer sequence (S) is used to separate two domains]; and cd^{LPMO} represents catalytic domain without linker and CBM domains. (b) Purification of fl^{LPMO} , $trunc1^{LPMO}$, $trunc2^{LPMO}$, $trunc3^{LPMO}$, and cd^{LPMO} from *Pichia pastoris* KM71H, where SDS-PAGE gel indicates enzymes purified using Ni-NTA chromatography.

monooxygenases (AA9 family) into the medium (unpublished data), and among all the variants, *BcAA9C* was observed to express maximally. Therefore, we considered *BcAA9C*-encoded protein for further biochemical characterization and supplementation to augment biomass degradation by commercial enzyme cocktail.

Design and *in silico* analysis of linker-truncated construct. *BcAA9C* encodes a 32-kDa protein which harbors a CBM domain and linker sequence in addition to a catalytic domain (Fig. S1 in the supplemental material). Previous studies indicated the importance of the CBM domain in cellulase enzyme activity (34); however, the impact of the linker sequence has not been elucidated in detail. Therefore, to pin down biomass disintegration by monooxygenase protein and evaluate the impact of linker sequence and CBM domain on LPMO-mediated polysaccharide disintegration, we designed the following enzyme variants: full-length (fl^{LPMO}), catalytic domain (cd^{LPMO}), and CD-CBM ($trunc1^{LPMO}$, $trunc2^{LPMO}$, $trunc3^{LPMO}$) (Fig. 2a).

To ensure that the replacement of linker with spacer sequence in the CD-CBM fusion constructs do not distort the stability of either domain, we screened the range of previously reported spacer sequences to ensure proper folding and stability in multi-domain proteins (35). We analyzed the instability index of these chimeric proteins using the ExPASy ProtParam tool. The instability index provides an estimate of the stability of proteins by analyzing the dipeptides, the occurrence of which is significantly different in stable versus unstable proteins (36). A protein with an instability index of greater than 40 is predicted as stable, while a value of less than 40 indicates that the protein may be unstable. Out of all the spacer sequences, CD-CBM fusion constructs with ELNYFQG, (EAAAK)₃, and native linker truncations demonstrated instability indices of less than 40 (Table 1). Furthermore, the tertiary structures of both proteins were threaded using I-TASSER, and the secondary structures of fusion proteins were analyzed using GOR and Pdbsum, which showed that the secondary structures of the chimeric proteins were similar to that of the native protein, with a slight decline in the percentage of extended strands from 34.65% to 31.76%, in contrast to fusion protein with (EAAAK)₃, which demonstrated a reduction in alpha helix with reduced disulfide bonds (Fig. S2). Therefore, we selected ELNYFQG as a suitable spacer sequence and other native linker truncations to construct chimeric CD-CBM fusion constructs.

The tertiary structures of chimeric constructs and fl^{LPMO} were threaded based on PDB templates such as 4e1rA, 6rs6A, 5foh, 6rs6, 4e1r, 5x39A, and 1cbhA, and models with highest confidence scores were selected to estimate structural similarity between two constructs. Structure superimposition using Swiss-PdbViewer indicated high similarity, with a RMSD (root mean square deviation) of -0.297 . I-TASSER also predicted that the ligand binding site was conserved (H85, H1, and Y170) and unperturbed in the chimeric construct. Furthermore, the models were subjected to energy minimization

TABLE 1 The physiochemical properties of f^{LPMO} and chimeric constructs

No.	Spacer sequence	Instability index
1	Native linker	38.38
2	ELNYFQG	39.97
3	(GGGG) ₂	42.81
4	(GGGG) ₃	43.59
5	(Gly) ₆	42.51
6	(Gly) ₈	43.22
7	— ^a	41.99
8	(EAAAK) ₃	38.97
9	TIPGPTPFVCGAAQSTAK ^b	39.87
10	TIPGPTPFVCGAAQSTAKSSSTSTAKPTSTLSTSTVTKTSSS ^c	39.98

^aIndicates no linker.

^bLinker truncation 1 (L1).

^cLinker truncation 2 (L2).

by SPdbV (Swiss-PdbViewer), which displayed that the chimeric constructs (*trunc1*^{LPMO}, -8727.556 KJ/mol; *trunc2*^{LPMO}, -8178.429 KJ/mol; *trunc3*^{LPMO}, -9987.863 KJ/mol) and the full-length construct (-7277.997 KJ/mol) had high structural stability.

Purification and biochemical characterization of f^{LPMO} , *trunc1*^{LPMO}, *trunc2*^{LPMO}, *trunc3*^{LPMO}, and *cd*^{LPMO}. All proteins were expressed in a *Pichia pastoris* system and purified using Ni-NTA affinity chromatography (Fig. 2b). The purified enzymes were assayed against Amplex Red, and the release of H₂O₂ per unit of time was monitored. Interestingly, activity in the truncated constructs was broadly unaffected, which again confirms the integrity of the catalytic domain across all variants (Fig. 3).

However, when assayed on a polysaccharide substrate, PASC (phosphoric acid swollen cellulose), a decline in activity was observed in all *trunc1*^{LPMO}, *trunc2*^{LPMO}, *trunc3*^{LPMO}, and *cd*^{LPMO}. Precisely, 8 g/L PASC was incubated with 10 μ M copper-saturated LPMO enzymes and the oxidized products were hydrolyzed using glucosidase to release gluconic acid, which was monitored using high-performance liquid chromatography (HPLC) (Fig. S3). The experimentation results indicated that truncation of the CBM domain resulted in a 93% decline in cellulose oxidation activity, whereas linker-truncated constructs demonstrated a 34 to 79% reduction in activity, compared to that of the full-length construct (Fig. 3). Precisely, *trunc2*^{LPMO} with a longer stretch of native linker sequence displayed higher activity than the other two constructs, confirming the importance of complete linker sequence in mediating polysaccharide degradation. Overall, the results clearly indicate the essential roles of linker and CBM in enhancing the rate of oxidative degradation of polysaccharide substrate.

Influence of accessory domains on binding to polysaccharide substrate. Here, we aimed at exploring the role of linker sequence in increasing proximity of f^{LPMO} to recalcitrant cellulose. For this, the binding strengths of f^{LPMO} , *trunc1*^{LPMO}, *trunc2*^{LPMO}, *trunc3*^{LPMO}, and *cd*^{LPMO} were tested against polysaccharide substrates. Each construct was evaluated at concentrations ranging from 0 to 150 μ M for binding to pretreated biomass (Fig. 4a). The curve between bound protein (μ mol/g) and amount of free protein (μ M) was used to compute partition coefficients, K_r (B_{max}/K_d). The experimental results indicated that the deletion of linker led to a 1.5- to 7-fold reduction in binding, whereas removal of CBM completely aborted binding to insoluble substrates. On Avicel, the K_r value followed a similar trend as on pretreated biomass (PTB); $K_r(f^{LPMO}) > K_r(trunc2^{LPMO}) > K_r(trunc3^{LPMO}) \sim K_r(trunc1^{LPMO}) > K_r(cd^{LPMO})$ (Table 2).

As detailed in Methods section; the fractions of bound, unbound, and total protein for each construct were run on SDS-PAGE gel; in concordance with depletion isotherm, the binding gel showed that f^{LPMO} exhibited complete binding to PTB, and linker-truncated constructs demonstrated reduced binding to polysaccharide substrate, showing that linker plays a crucial role in binding to polysaccharide substrate (Fig. 4b).

Further, confocal microscopy of biomass confirmed that the binding of f^{LPMO} declines with truncation of the CBM and linker domains (Fig. 4c). We also observed that deletion of the carbohydrate-binding module led to complete loss of the capability of f^{LPMO} to bind

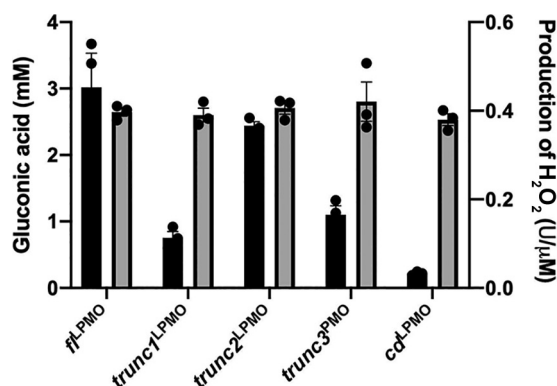


FIG 3 Biochemical characterization of *fl*^{LPMO}, *cd*^{LPMO}, *trunc1*^{LPMO}, *trunc2*^{LPMO}, and *trunc3*^{LPMO}. Specific activity of full-length and truncated monooxygenases on complex and simple substrate. Gray bar indicates Amplex Red assay using 1 μ M protein, where y axis indicates μ mol of H₂O₂ released per unit of time per μ M protein; black bar indicates LPMO-assisted degradation of PASC, where 8 g/L PASC was treated for 48 h at 45°C with 10 μ M copper-saturated LPMO enzymes and 2 mM ascorbate. The hydrolysis of oxidized products was performed at 40°C for 24 h using 5 units of β -glucosidase, and the release of gluconic acid was estimated using HPLC equipped with a photodiode array (PDA) detector. Error bars represent standard deviation of three independent experiments.

recalcitrant biomass (Fig. 4d). Overall, the presence of the linker and CBM domains was found to enhance the binding of LPMO to insoluble and crystalline substrates.

Influence of linker on thermostability of *fl*^{LPMO}. The effect of the linker region on the thermostability of *fl*^{LPMO} was investigated by measuring the residual activities of *fl*^{LPMO}, *cd*^{LPMO}, *trunc1*^{LPMO}, *trunc2*^{LPMO}, and *trunc3*^{LPMO} at various temperatures. The monooxygenase activities of all constructs remained constant for the initial 30 min at a temperature range of 35 to 45°C. At 50°C, an approximately 65% decline in activity was observed within 30 min in the *trunc1*^{LPMO}, *trunc3*^{LPMO}, and *cd*^{LPMO} constructs, whereas activity was still maintained in *fl*^{LPMO} and *trunc2*^{LPMO}. With a 5-degree increase in temperature, complete loss of activity was observed in less than 15 min in the *trunc1*^{LPMO}, *trunc3*^{LPMO}, and *cd*^{LPMO} constructs, whereas the *fl*^{LPMO} and *trunc2*^{LPMO} constructs demonstrated 46% and 65% declines in activity in 25 min, respectively. Furthermore, the presence of linker increased the thermostability of monooxygenase enzyme by 10 min at 60°C (Fig. 5). By and large, linker imparts thermostability to its associated monooxygenase catalytic domain.

Effect of LPMO on degradation of PTB. Recent studies have reported the role of LPMO in accelerating the degradation of recalcitrant biomass (37). However, the importance of the linker and CBM domains in assisting LPMO-mediated disintegration of cellulosic biomass is poorly understood. Here, we have analyzed the effect of LPMO on disintegration of recalcitrant biomass using biophysical and biochemical parameters.

Scanning electron microscopy (SEM) was used to observe surface changes in pretreated biomass post-exposure to 10 μ M *fl*^{LPMO}, *cd*^{LPMO}, or CD-CBM fusion constructs. Figure S6i and iii show the negative and positive controls, respectively; Fig. S6i represents the surface morphology of untreated pretreated biomass, whereas Fig. S6iii represents the morphology of commercial enzyme-treated biomass. SEM analysis of the *cd*^{LPMO} (Fig. S6ii), *trunc1*^{LPMO}, and *trunc3*^{LPMO} (Fig. S6v and S6vii)-treated samples displayed marginal changes in surface architecture, while the *fl*^{LPMO} (Fig. S6iv) and *trunc2*^{LPMO} (Fig. S6vi) microscopy images display increased surface roughness compared to the control, indicating that both the linker and CBM domains are crucial for the binding and hydrolysis of plant biomass structure.

Furthermore, we performed a biochemical assay to measure the increase in glucose release upon LPMO-assisted biomass degradation. As anticipated, the degradation followed the same trend as for SEM: *fl*^{LPMO} > *trunc2*^{LPMO} > *trunc3*^{LPMO} > *trunc1*^{LPMO} > *cd*^{LPMO} (Fig. S4).

Synergistic depolymerization of PASC by LPMOs and classical cellulase cocktail. LPMOs are known to augment the degradation of polysaccharide substrates by classical cellulase cocktail. Therefore, we evaluated the contribution of *fl*^{LPMO} toward depolymerization of crystalline polysaccharide substrate, PASC. Furthermore, we also studied importance of the linker and CBM sequences in *fl*^{LPMO}-assisted cellulose disintegration.

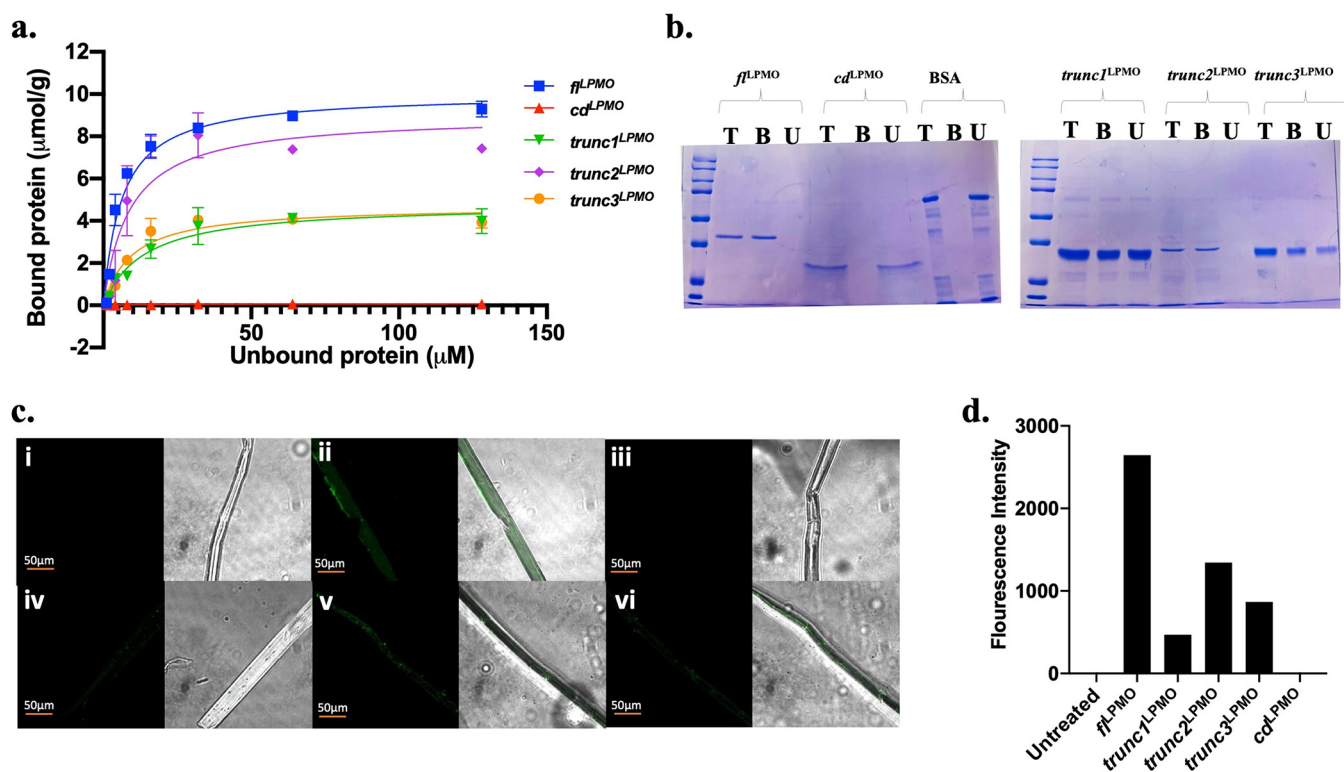


FIG 4 Influence of linker on binding polysaccharide substrate. (a) Binding isotherms for full-length ($f\ell^{\text{LPMO}}$), CD-CBM ($trunc1^{\text{LPMO}}$, $trunc2^{\text{LPMO}}$, $trunc3^{\text{LPMO}}$), and CD (cd^{LPMO}), where specific binding of protein was calculated by the following equation: [total protein-free protein (post binding)]/g biomass. (b) Representation of binding experiment where 120 μM protein was allowed to bind 2 mg biomass for 12 h at 4°C in 250 μL reaction and total protein (T), Bound fraction (B), and Unbound fraction (U) were loaded onto SDS-PAGE gel (14%). BSA was used as a negative control (c) The protein-bound biomass was analyzed for protein binding using confocal microscopy. The protein-bound biomass fraction was probed with anti-His primary antibody and anti-IgG-FITC secondary antibody to visualize specific binding using confocal microscopy, where the binding of different constructs is represented by (i) BSA, (ii) $f\ell^{\text{LPMO}}$, (iii) cd^{LPMO} , (iv) $trunc1^{\text{LPMO}}$, (v) $trunc2^{\text{LPMO}}$, and (vi) $trunc3^{\text{LPMO}}$. (d) FITC fluorescence intensity averaged for nine images indicates quantitative binding of LPMO constructs.

For this, the glucose yield generated by the combined action of cellulase and the $f\ell^{\text{LPMO}}$, cd^{LPMO} , or CD-CBM fusion constructs was analyzed using HPLC equipped with a HiPlex-Ca column (Fig. S5). The addition of $f\ell^{\text{LPMO}}$ led to a 1.6-fold increase in glucose released over time, whereas the addition of $trunc2^{\text{LPMO}}$ and $trunc3^{\text{LPMO}}$ led to mere 1.4- and 1.1-fold increases, respectively; no change was observed in the CBM-truncated or linker-truncated constructs with shorter stretches of linker sequence (Fig. 6). The results also highlighted the importance of the complete linker region in LPMO-assisted biomass disintegration.

DISCUSSION

Cellulose lytic monooxygenases belong to the AA9 family of CAZymes, which catalyze oxidative degradation of cellulose. LPMOs are reported to augment biomass disintegration by classical hydrolytic enzyme cocktail (38), and therefore, persistent efforts are being led to discover newer LPMOs (39–41). These LPMOs mainly interact with polysaccharide substrate via accessory domains, i.e., carbohydrate-binding module (CBM) (24, 25). Generally, these

TABLE 2 The partition coefficients ($K_r = B_{\text{max}}/K_d$) for LPMO variants from the Langmuir isotherm

LPMO variant	K_d	B_{max}	B_{max}/K_d
$f\ell^{\text{LPMO}}$	6.218	10.02	1.61
cd^{LPMO}	18.74	0.07915	4.22×10^{-3}
$trunc1^{\text{LPMO}}$	22.87	5.283	2.31×10^{-1}
$trunc2^{\text{LPMO}}$	8.844	8.997	1.07
$trunc3^{\text{LPMO}}$	20.66	5.081	2.45×10^{-1}

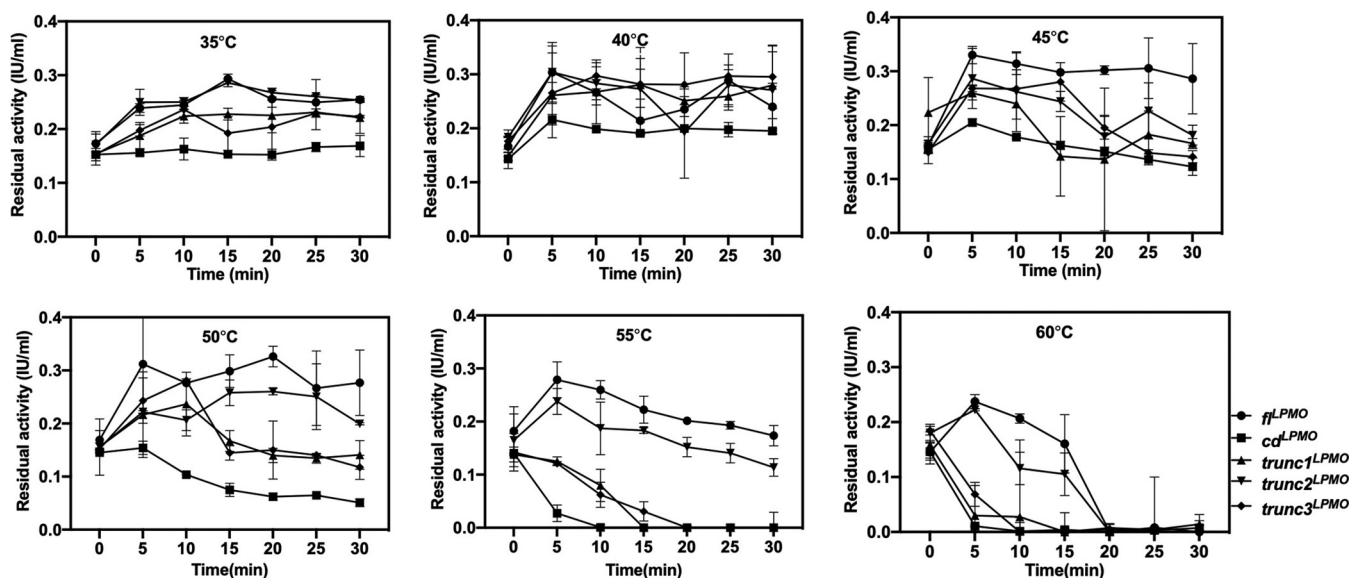


FIG 5 Effect of linker on thermostability of recombinant polysaccharide monooxygenase. LPMO activity of full-length (fl^{LPMO}), CD-CBM ($trunc1^{LPMO}$, $trunc2^{LPMO}$, $trunc3^{LPMO}$), and CD (cd^{LPMO}) construct was tested at a temperature range of 35 to 60°C for 30 min. The enzymes were preincubated at 35°C, 40°C, 45°C, 50°C, 55°C, and 60°C for 5 to 30 min and immediately tested for residual monooxygenase activity using a reference Amplex Red assay, which estimates μmol of H_2O_2 released per unit of time. Error bars represent standard deviation of two independent experiments.

binding modules are tethered to the catalytic domain via linker regions, which are considered to be flexible hinges between two structured domains (25); however, the other functional role of the linker region in LPMO performance has not yet been revealed. Here, we aim to unravel the role of the linker region in the activity, binding, and thermostability of its associated protein.

For this, we have mined the genome of the fungal phytopathogen *B. cinerea* for the presence of cellulose lytic monooxygenases and identified *BcAA9C* as one of the majorly expressed monooxygenases. The *BcAA9C*-encoded, 32-kDa fl^{LPMO} is a protein with a 5.2-kDa linker region and a 3.2-kDa CBM domain. To understand the importance of linker, we designed linker- and CBM-truncated constructs. Briefly, to prune the linker domain, we replaced the linker sequence with a 21-nucleotide helical spacer sequence to allow independent folding of the catalytic and carbohydrate-binding domains.

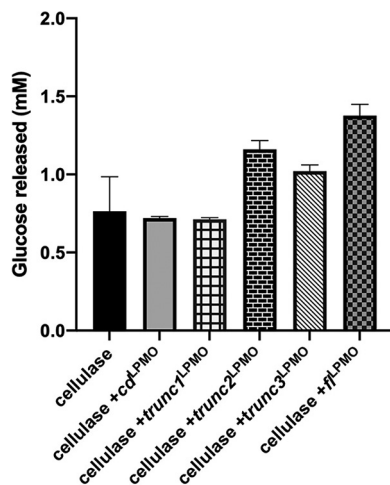


FIG 6 Synergistic action of LPMOs and cellulases. Eight g/L PASC was incubated with 0.5 FPU cellulase (C2730, Sigma), 10 μM copper-saturated fl^{LPMO} , $trunc1^{LPMO}$, $trunc2^{LPMO}$, $trunc3^{LPMO}$, or cd^{LPMO} , and 2 mM ascorbate for 72h at 45°C, and the release of glucose was measured using HPLC equipped with a Hi-Plex Ca column. Error bars indicate standard deviations from three independent experiments.

Furthermore, we designed two additional fusion constructs with truncation of the native linker sequence. We analyzed the physicochemical properties of a CD-CBM chimeric construct to ensure that the switch of linker domain to spacer sequence did not disturb the folding of the protein. Biochemical assays also suggested that the fusion and the native protein displayed similar potential to oxidize Amplex Red (Fig. 3) and that the active center is unperturbed, suggesting that the constructs are ideal to study the role of the linker present in the LPMO sequence.

The oxidative degradation of PASC declined by 93% with CBM deletion, whereas linker truncations resulted in a 34 to 79% reduction, suggesting the importance of the linker and CBM sequences in biomass degradation. Langmuir depletion isotherm also showed that removal of the linker domain resulted in 7-fold reduced binding efficiency to polysaccharide substrate, emphasizing the role of the linker domain in binding to cellulosic chain (Fig. 4a). Using confocal microscopy, we confirmed the role of linker in binding, as a linker-free, chimeric *cd-cbm*^{LPMO} protein demonstrated significantly less binding than *fl*^{LPMO}. Furthermore, it is evident from confocal and binding experiments that a longer stretch of native linker in CD-CBM fusion constructs ensures improved binding, confirming the crucial role of linker in assisting binding to cellulose. This was in concordance with Badino et al. (42) who suggested that moderate modification of linker sequence in cellobiohydrolase could perturb its substrate affinity. Another study by Srisodsuk et al. (43) also demonstrated that, at high enzyme concentration, interdomain linker plays a crucial role in supporting polysaccharide-substrate binding. One possible reason for its role in mediating interaction between substrate and catalytic domain could be its typical glycosylation pattern. Sammond et al. (27) preferentially studied linkers connecting glycosyl hydrolase family 6 or 7 to carbohydrate binding modules 1 or 2, and identified that the linker sequence displays very little sequence conservation and significant amino acid bias toward serine/threonine, which serves as potential hot spots for O-glycosylation. These O-glycosylation sites were shown to expand the operating range of enzymes (44). In addition, glycosylation plays myriad biological roles in providing proteolytic stability to cellulose-degrading enzymes (28). Following these studies, we performed prediction of glycosylation patterns in *fl*^{LPMO} using the Net Ogly 4.0 server (45), which distinctly revealed that in a complete protein sequence, only the linker region is heavily O-glycosylated.

Along with binding, we demonstrated the role of linker region in imparting thermal stability to the catalytic domain. In accordance, Miao et al. (46) has shown that appending the linker region and CBM1 to GH10 xylanase increases its thermostability. The group has also shown that deletion of the linker region led to a marked decline in enzyme stability. Contrary to this, Sonan et al. (32) demonstrated a role of linker in cold adaptation, indicating that each cellulase linker has a unique, indispensable functional role in imparting stability; therefore, it was important to analyze the linker associated with *fl*^{LPMO}. Our study highlighted the importance of linker in increasing the thermostability of associated cellulose monoxygenase (Fig. 5).

The goal of scrutinizing any new CAzyme is to improve existing cellulase cocktail for efficient biomass deconstruction. Therefore, we analyzed the roles of *fl*^{LPMO}, *trunc1*^{LPMO}, *trunc2*^{LPMO}, *trunc3*^{LPMO}, and *cd*^{LPMO} in biomass fibrillation using SEM, which indicated that the linker and CBM domains boost degradation of pretreated biomass. Overall, our experimental results suggest that linker region is as crucial as catalytic domain in dismantling polysaccharide biomass and, therefore, that engineering the linker region may also constitute an efficient method for improving cellulase activity.

MATERIALS AND METHODS

Strains and media. The fungus *Botrytis cinerea* was taken from the Indian Type Culture Collection, India (accession no. –6192). *Botrytis* sp. (1×10^6 spores/mL) was inoculated in complex medium (Soya Peptone 24 g, ammonium sulfate 3.12 g, CaCl₂·2H₂O 0.05 g, KH₂PO₄ 5.9 g, yeast extract 0.05 g per liter of medium) containing 21.4 g/L wheat bran and 24 g/L Avicel and incubated on a shaker for 5 days at 22°C, and samples were withdrawn at different time points for expression studies.

Escherichia coli strain DH5-Alpha was used for cloning genes in pPICZ(alpha)A vector, while protein expression was performed in *Pichia pastoris* KM71H using the standard instructions in an EasySelect *Pichia* Expression kit (Invitrogen).

TABLE 3 Primer sequences for RT-PCR

Gene	Forward primer	Reverse primer
<i>BcAA9A</i>	GCTAGCAATGATCTACGAT	GACCTGATGATAAACCTG
<i>BcAA9B</i>	GTATTGGTTCATGGTTCAAG	TGAACGAAGCAAGTATTAC
<i>BcAA9C</i>	GAATACGTGCGCTGCAATG	ATCATTTCAGGTGTGCTG
<i>BcAA9D</i>	CGTTCCTTACCATTCCATC	CAGACCCAGTTACGCGTAG
<i>BcAA9E</i>	CAGATCCCTACTATGCTACA	GCTGTGAACCAACAATACC
<i>BcAA9F</i>	CACAACCCGATTCTATCGG	GCGTTGGTAGCCTCCTTAG
<i>BcAA9G</i>	CAATCGAACGGTTTCATTGC	CCATTGCAAGTTAACCTCAC
<i>BcAA9H</i>	GTGTGTTATGATTGTGCGC	CCGTCTTAACAATCACAGG
<i>BcAA9I</i>	GCTGGCTGATACCTAAAGAT	CTTTAACAGGAGCCTCGATC
<i>Actin</i>	GAATACGTGCGCTGCAATG	ATCATTTCAGGTGTGCTG

Pretreatment of wheat straw. Sodium hydroxide-treated wheat straw was prepared using the protocol developed by Xu et al. (47). Briefly, wheat straw was treated with 1% NaOH for 12 h at 50°C under continuous agitation. The treated biomass was then washed extensively with distilled water until it reached a neutral pH. Excess water in pretreated biomass (PTB) was removed by keeping it in an oven at 70°C for 5 h. The cellulosic and hemicellulosic content in PTB was analyzed using a standard protocol (48) (Table S1 in the supplemental material).

Functional annotation. Analysis of carbohydrate active enzymes (CAZymes) was performed using proteins downloaded from the *Botrytis cinerea* genome sequence (NCBI GenBank Assembly accession no. GCA_000143535.4). The dbCAN2 metaserver was used to predict the list of CAZymes (49). The Hidden Markov model algorithm (50) was used in HMMER 3.0 (51) to retrieve CAZymes with an E value of 10^{-5} .

RNA extraction and qRT-PCR. *B. cinerea* (1×10^6 spores/mL) was inoculated in Avicel- and wheat bran-containing complex medium and mycelium was recovered after 48 h, 72 h, and 96 h cultivation. The mycelium was crushed in liquid nitrogen using a mortar and pestle, and 100 mg mycelium powder was utilized for RNA extraction using the manufacturer's instructions (Qiagen Plant RNA extraction kit). RNA samples were quantified using NanoDrop (Thermo Fisher) and 2 μ g RNA was used for cDNA synthesis using the manufacturer's instructions (High Capacity cDNA Reverse Transcription Kit, Thermo Fisher Scientific).

For real-time PCR, primers for nine *BcAA9* genes were designed using IDT software considering amplification of 160 bp amplicon and $T_m = 60^\circ\text{C}$ (Table 3), and *actin* was used as a housekeeping gene. Two μ L of cDNA was used as the template and the reaction mix was prepared with 10 μ L SYBR Green Master Mix (Takara Bio), 0.8 μ L Primer1, 0.8 μ L Primer2, 0.4 μ L ROX reference dye, and 7.44 μ L nuclease-free water. After reaction assembly, an Applied Biosystems real-time PCR machine was used to perform qPCR with the following conditions: initial denaturation at 95°C for 2 min; 40 cycles of 95°C for 15 s, 55°C for 15 s, and 72°C for 15 s; and final extension of 72°C for 2 min. Raw data were analyzed using the $\Delta\Delta\text{Ct}$ method and results were plotted in GraphPad Prism 8.0. All experiments were performed in two technical and three biological replicates, and standard deviation in the fold change was calculated accordingly.

Expression of full-length, truncated, and chimeric LPMO constructs. A full-length codon optimized *BcAA9C* gene was ordered from GenScript Co., and the sequence is submitted in NCBI under GenBank ID-732129. The full-length sequence was used as a template to amplify the catalytic domain (1 to 232 aa), whereas the construct without native linker sequence was designed by either truncating the linker sequence or replacing it with a 21-nucleotide spacer sequence; genes were synthesized from Thermo Fisher Scientific. Specifically, the linker-truncated construct has catalytic (1 to 232 aa) and carbohydrate-binding domains (288 to 321 aa), and a spacer sequence to allow free movement of structured domains. This sequence is submitted in NCBI under GenBank ID-732130. All gene fragments were amplified and digested with EcoRI/NotI restriction enzymes and cloned separately in EcoRI/NotI digested pPICZ(alpha)A vector. The cloning in EcoRI and NotI site in pPICZalpha vector inserts the gene in frame with an N-terminal alpha secretion signal and a C-terminal polyhistidine tag. The cloning was performed in DH5-Alpha and correct insertion of the gene was confirmed via colony PCR using standard AOX (alcohol oxidase) primers provided in the EasySelect Pichia Expression Kit, followed by Sanger sequencing. The positive clones were then linearized using a blunt-end restriction enzyme, SacI, transformed in a chemical-competent *P. pastoris* KM71H strain, and plated onto a yeast extract peptone dextrose selection plate (10 g/L yeast extract, 20 g/liter Soya Peptone, 20 g/L dextrose, 1 M sorbitol, and 20 g/L agar) containing 100 μ g/mL Zeocin. Integration in yeast was confirmed by gene-specific PCR, and positive transformants were selected for protein purification.

Protein expression was induced using methanol, with addition of 1% methanol every 24 h, and culture supernatant was harvested after 72 h for protein purification. Briefly, protein secreted in the culture supernatant was precipitated using 60% ammonium sulfate (A4915, Sigma). Following this, precipitate was dialyzed in 20 mM sodium phosphate buffer and 150 mM sodium chloride (pH 8.0), then subjected to Ni-NTA chromatography. The purified enzymes are listed as follows: *fl*^{LPMO} (full-length protein), *cd*^{LPMO} (catalytic domain), and *trunc1*^{LPMO}, *trunc2*^{LPMO}, and *trunc3*^{LPMO} (linker-truncated chimeric protein) were copper saturated by incubating the enzyme with a 10-fold excess concentration of copper sulfate. The copper was allowed to bind at 25°C for 1 h at 200 rpm. Post-binding, the enzyme was dialyzed against a 20 mM sodium phosphate buffer (pH 6.0) to remove unbound copper.

In silico analysis of stability and structure of CD-CBM chimeric construct. The physical and chemical properties of linker-truncated fusion proteins (*trunc1*^{LP_{MO}}, *trunc2*^{LP_{MO}}, *trunc3*^{LP_{MO}}) were compared with those of full-length protein (*fl*^{LP_{MO}}) using ExPASy's ProtParam server (52). Various parameters analyzed include instability index, aliphatic index, grand average hydropathicity, and half-life of protein. The aliphatic index increased along with the number of aliphatic residues, indirectly indicating an increase in the thermostability of the protein; the instability index measures protein stability, and an index of less than 40 indicates that protein is stable.

The secondary structures of proteins were analyzed using the GOR secondary structure prediction method (53) and the phyre2 online server (54), whereas the 3D structures were modeled using the I-TASSER server (55). The FASTA sequences were submitted and threading was performed using I-TASSER, which generated five models; the one with highest C-score was selected for further analysis. C-score values are typically in a range of -5 to 2 where a higher score reflects a model of better quality (56). Tertiary structures of proteins were verified using the Procheck server (57).

After generating the model, energy minimization was determined by analyzing the 3D structural stability of the chimeric protein using Swiss-PdbViewer (58) and the GROMACS 2020 package with the GROMOS 96, 54A7 force field. The steepest descent method was employed for energy minimization with Particle Mesh Ewald (PME) long-range electrostatics for 50,000 steps (59, 60).

Enzyme Assays. An LP_{MO} assay was performed using the protocol described by Ogunmolu et al. (61), with minor modifications. Briefly, the reaction mixture was comprised of 30 μ M Ascorbate (Sigma), 50 μ M Amplex Red (Sigma), and 7.14 units/mL HRP in a citrate phosphate buffer (pH 6.0). The enzyme assay was performed at 40°C, and 20 μ L enzyme was added to 180 μ L reaction mix to initiate the reaction. The kinetics was performed for 10 min with 535 nm and 580 nm excitation and emission wavelengths, respectively. The enzyme was excluded in the control reaction. The concentration of H₂O₂ released in the LP_{MO} reaction was measured using H₂O₂ standard curve (1 to 10 μ M). The activity of LP_{MO} was expressed as μ mol of H₂O₂ released per unit of time.

The cellulose oxidation activities of LP_{MO} enzymes were performed by incubating 10 μ M copper-saturated *fl*^{LP_{MO}}, *cd*^{LP_{MO}}, *trunc1*^{LP_{MO}}, *trunc2*^{LP_{MO}}, and *trunc3*^{LP_{MO}} with 8 g/L PASC (phosphoric acid swollen cellulose) and 2 mM ascorbate at 45°C for 48 h. The control reaction lacked the enzyme component. Oxidized oligosaccharides were subjected to further hydrolysis by 5 units of beta-glucosidase at 40°C for 24 h. The supernatant was harvested and protein was removed using 20 mM H₂SO₄. The release of gluconic acid was estimated using an Agilent 1260 Infinity II HPLC equipped with an Agilent Hi-Plex-H column. The conditions used are as follows: 0.5 mL/min flow rate, column compartment temperature = 30°C, and absorbance 210 nm (62). The concentration of gluconic acid was calculated using a gluconic acid standard curve prepared under similar conditions. The experiment was performed in three biological replicates.

To estimate synergistic depolymerization of complex polysaccharide using LP_{MO} variants and cellulases: we treated 8 g/L PASC with 0.5 FPU commercial enzyme cocktail (C2730; Sigma) and supplemented it separately with 10 μ M LP_{MO} variants; for the negative control, LP_{MO} enzymes were excluded. The reaction mix was incubated at 45°C for 72 h at 200 rpm to ensure proper agitation and mixing of reaction content. The supernatant was harvested and analyzed for increases in the concentration of glucose released from biomass treatment, using an Agilent 1260 Infinity II HPLC equipped with an Agilent Hi-Plex Ca column using the following conditions: column temperature = 80°C, refractive index detector temperature = 30°C. Milli-Q water was used as the mobile phase at a flow rate of 0.6 mL/min. The concentration of glucose was calculated against the standard curve prepared by running glucose standard (1 to 10 g/L).

Binding assays. The binding of *fl*^{LP_{MO}}, *cd*^{LP_{MO}}, *trunc1*^{LP_{MO}}, *trunc2*^{LP_{MO}}, and *trunc3*^{LP_{MO}} enzymes to cellulosic substrates (pretreated biomass and Avicel) was tested by mixing various concentrations of protein (1 to 150 μ M) with a fixed substrate concentration. Briefly, 40 g/L water-saturated PTB (80% hydration) was incubated with different concentrations of enzymes in the presence of 20 mM sodium phosphate buffer (pH 6.0) in a total volume of 250 μ L. The reaction mixture was incubated in a heat block at 4°C for 12 h at 200 rpm. Protein-bound substrate pellet and unbound proteins in the supernatant were recovered by high-speed centrifugation. The protein in the unbound fraction (supernatant) was estimated using a BCA kit (Pierce, Thermo Fisher Scientific), which was then used to calculate the protein in the bound fraction using the following formula: bound protein = total protein - unbound protein, where total protein indicates the amount of initial protein loaded for the binding experiment. GraphPad Prism 8.0 was used to compute dissociation constant (K_d) and the amount of protein bound at saturation (B_{max}). Langmuir binding isotherms were carried out for each protein and bovine serum albumin (BSA) was used as a negative control for the assay. All binding experiments were performed in duplicate.

For SDS-PAGE, 120 μ M copper-saturated *fl*^{LP_{MO}}, *cd*^{LP_{MO}}, *trunc1*^{LP_{MO}}, *trunc2*^{LP_{MO}}, and *trunc3*^{LP_{MO}} were mixed with 40 g/L (80% water-saturated PTB) and incubated for 12 h at 4°C. Samples were centrifuged and supernatant (unbound), pellet (bound), and total protein fraction were mixed with 6 \times gel loading dye and run on SDS-PAGE gel. The gel was stained with Coomassie brilliant blue for visualization.

Thermostability assay. To determine the thermostability of *fl*^{LP_{MO}}, *cd*^{LP_{MO}}, *trunc1*^{LP_{MO}}, *trunc2*^{LP_{MO}}, and *trunc3*^{LP_{MO}}, 150 μ M copper-saturated enzyme was preincubated at a range of temperatures (35 to 60°C) and the sample was analyzed every 5 min for the presence of residual oxidation activity. Briefly, 50 μ L enzyme was incubated for 5 min at gradient temperatures 35, 40, 45, 50, 55, and 60°C in a Bio-Rad thermocycler; postincubation, appropriately diluted LP_{MO} enzymes were analyzed for residual activity by performing an Amplex Red assay. Similarly, five other experimental setups for different time points, i.e., 10, 15, 20, 25, and 30 min, indicated a change in thermostability over time. The release of H₂O₂ per unit of time was measured and compared with that of the control (no preincubation). All experiments were done in duplicate.

Scanning electron microscopy. The topological changes in plant biomass upon exposure to hydrolytic enzymes was analyzed using scanning electron microscopy (SEM). For this, sodium hydroxide-treated biomass was treated with 10 μM copper-saturated β^{LPMO} , cd^{LPMO} , $trunc1^{\text{LPMO}}$, $trunc2^{\text{LPMO}}$, and $trunc3^{\text{LPMO}}$ with 2 μM ascorbate in 20 μM sodium phosphate buffer (pH 6.0) for 48 h at 45°C. Treatment with commercial enzyme cocktail (SAE0020; Sigma) was used as a positive control, while incubation with water was utilized as a negative control. This differentially treated wheat straw was dried on a coverslip under vacuum in a desiccator. The dried biomass was then carbon-coated and viewed via a scanning electron microscope (EVO 40, Carl Zeiss, Germany). The topology of biomass samples post-treatment was observed by the action of an electron beam at 2 KV, and all samples were analyzed at 20 KX magnification.

Confocal assay. Confocal microscopy was used to monitor binding of proteins to PTB. For sample processing, as shown in the binding experiment, 40 g/L (80% water saturated) pretreated biomass was allowed to bind 120 μM β^{LPMO} , cd^{LPMO} , $trunc1^{\text{LPMO}}$, $trunc2^{\text{LPMO}}$, and $trunc3^{\text{LPMO}}$ for 12 h at 4°C under a constant agitation of 200 rpm. For the control, pretreated biomass was incubated with BSA (ab217817, abcam) instead of monoxygenase enzyme. After 12 h, the protein-bound PTB was recovered by high-speed centrifugation. The biomass was washed four times with phosphate-buffered saline (PBS) to remove loosely bound protein. The PTB with bound enzymes (his-tagged) were then incubated with 1:5,000 dilution monoclonal anti-polyhistidine antibody (H1029, Sigma) for 1.5 h at room temperature and washed five times with PBS to remove non-specifically bound antibody. Following this, primary antibody-bound biomass was incubated with 1:15,000 anti-mouse IgG-FITC (fluorescein isothiocyanate) antibody (F0257, Sigma) for 1.5 h in darkness (34). The samples were then extensively washed to remove non-specifically bound antibody. The biomass samples were mounted on the slide with the help of glycerol, and observed on an Olympus BH-2 microscope equipped with epifluorescence irradiation. Confocal microscopy was performed at 492 nm excitation and 520 nm emission maxima, using a Laser Scanning Confocal Microscope (LSM 880, Carl Zeiss). Channel mode visualization was done using 63 \times Plan-Apochromat Oil immersion DIC (NA 1.40). All experiments were done in triplicate and multiple images were captured for quantitative estimation of binding. The fluorescence was analyzed using open source ImageJ software.

SUPPLEMENTAL MATERIAL

Supplemental material is available online only.

SUPPLEMENTAL FILE 1, PDF file, 6.4 MB.

ACKNOWLEDGMENTS

This work was supported by RCB Core grant and DST INSPIRE grant LSPA-30/2015. We also acknowledge Sathish, from the Synthetic Biology and Bioprocessing Lab, RCB, for his help with HPLC experiments.

N.A. and A.S. conceived and designed the research. A.S., S.R. and P.N. conducted experiments. N.A. and A.S. analyzed the data. N.A. wrote the manuscript. All authors read and approved the manuscript.

REFERENCES

1. Scarlat N, Dallemand J-F, Monforti-Ferrario F, Nita V. 2015. The role of biomass and bioenergy in a future bioeconomy: policies and facts. *J Environ Dev* 15:3–34. <https://doi.org/10.1016/j.envdev.2015.03.006>.
2. Berchin II, Silva SA, Bocquillon P, Fornasari VH, Ribeiro LPC, Ribeiro JMP, Guerra J. 2018. Contributions of public policies to greening sugarcane ethanol production in Brazil, pg. 375–393. *In* Filho WL, Pociovalişteanu D, Brito PB, Lima IB (ed), *Towards a sustainable bioeconomy: principles, challenges and perspectives. world sustainability series*. Springer, Cham, Switzerland.
3. Owusu P, Asumadu SS. 2016. A review of renewable energy sources, sustainability issues and climate change mitigation. *Cogent Eng* 3:1167990. <https://doi.org/10.1080/23311916.2016.1167990>.
4. Pfennig A. 2019. Sustainable bio- or CO₂ economy: chances, risks, and systems perspective. *Chem Bio Eng Rev* 6:90–104. <https://doi.org/10.1002/cben.201900006>.
5. Carere CR, Sparling R, Cicek N, Levin DB. 2008. Third generation biofuels via direct cellulose fermentation. *Int J Mol Sci* 9:1342–1360. <https://doi.org/10.3390/ijms9071342>.
6. Ahmed MM, Nasri NS, Hamza DU. 2012. Biomass as a renewable source of chemicals for industrial applications. *Int J Eng Sci and Technol* 4:721–730.
7. Contreras F, Pramanik S, Rozhkova AM, Zorov IN, Korotkova O, Sinityn AP, Schwaneberg U, Davari MD. 2020. Engineering robust cellulases for tailored lignocellulosic degradation cocktails. *Int J Mol Sci* 21:1589. <https://doi.org/10.3390/ijms21051589>.
8. Reyes-Sosa FM, López Morales M, Platero Gómez AI, Valbuena Crespo N, Sánchez Zamorano L, Rocha-Martín J, Molina-Heredia FP, Díez García B. 2017. Management of enzyme diversity in high-performance cellulolytic cocktails. *Biotechnol Biofuels* 10:156. <https://doi.org/10.1186/s13068-017-0845-6>.
9. Knott BC, Erickson E, Allen MD, Gado JE, Graham R, Kearns FL, Pardo I, Topuzlu E, Anderson JJ, Austin HP, Dominick G, Johnson CW, Rorrer NA, Szostkiewicz CJ, Copié V, Payne CM, Woodcock HL, Donohoe BS, Beckham GT, McGeehan JE. 2020. Characterization and engineering of a two-enzyme system for plastics depolymerization. *Proc Natl Acad Sci U S A* 117:25476–25485. <https://doi.org/10.1073/pnas.2006753117>.
10. Lopes A, Ferreira Filho EX, Moreira L. 2018. An update on enzymatic cocktails for lignocellulose breakdown. *J Appl Microbiol* 125:632–645. <https://doi.org/10.1111/jam.13923>.
11. Mohanram S, Amat D, Choudhary J, Arora A, Nain L. 2013. Novel perspectives for evolving enzyme cocktails for lignocellulose hydrolysis in biorefineries. *Sust Chem Process* 1:1–12. <https://doi.org/10.1186/2043-7129-1-15>.
12. Garron ML, Henrissat B. 2019. The continuing expansion of CAZymes and their families. *Curr Opin Chem Biol* 53:82–87. <https://doi.org/10.1016/j.cbpa.2019.08.004>.
13. Kameshwar AKS, Ramos LP, Qin W. 2019. CAZymes-based ranking of fungi (CBRF): an interactive web database for identifying fungi with extrinsic plant biomass degrading abilities. *Bioresour Bioprocess* 6:51. <https://doi.org/10.1186/s40643-019-0286-0>.
14. Payne CM, Knott BC, Mayes HB, Hansson H, Himmel ME, Sandgren M, Stahlberg J, Beckham GT. 2015. Fungal cellulases. *Chem Rev* 115:1308–1448. <https://doi.org/10.1021/cr500351c>.

15. Sreedharan S, Prakasan P, Sasidharan S, Sailas B. 2016. An overview on fungal cellulases with an industrial perspective. *J Nutr Food Sci* 06:461–473. <https://doi.org/10.4172/2155-9600.1000461>.
16. Obeng EM, Adam SNN, Budiman C, Ongkudon CM, Maas R, Jose J. 2017. Lignocellulases: a review of emerging and developing enzymes, systems, and practices. *Bioresour Bioprocess* 4:16. <https://doi.org/10.1186/s40643-017-0146-8>.
17. Hemsworth GR, Déjean G, Davies GJ, Brumer H. 2016. Learning from microbial strategies for polysaccharide degradation. *Biochem Soc Trans* 44: 94–108. <https://doi.org/10.1042/BST20150180>.
18. Johansen KS. 2018. Oxidative copper-enzymes are the key for unlocking recalcitrant biopolymers. Doctoral thesis. University of Copenhagen, Copenhagen, Denmark.
19. Vaaje-Kolstad G, Westereng B, Horn SJ, Liu Z, Zhai H, Sorlie M, Eijsink VGH. 2010. An oxidative enzyme boosting the enzymatic conversion of recalcitrant polysaccharides. *Science* 330:219–222. <https://doi.org/10.1126/science.1192231>.
20. Phillips CM, Beeson WT, Cate JH, Marletta MA. 2011. Cellobiose dehydrogenase and a copper-dependent polysaccharide monooxygenase potentiate cellulose degradation by *Neurospora crassa*. *ACS Chem Biol* 6:1399–1406. <https://doi.org/10.1021/cb200351y>.
21. Quinlan RJ, Sweeney MD, Lo Leggio L, Otten H, Poulsen JC, Johansen KS, Krogh KB, Jørgensen CI, Tovborg M, Anthonson A, Tryfona T, Walter CP, Dupree P, Xu F, Davies GJ, Walton PH. 2011. Insights into the oxidative degradation of cellulose by a copper metalloenzyme that exploits biomass components. *Proc Natl Acad Sci U S A* 108:15079–15084. <https://doi.org/10.1073/pnas.1105776108>.
22. Aachmann FL, Sørlie M, Skjåk-Bræk G, Eijsink VGH, Vaaje-Kolstad G. 2012. NMR structure of a lytic polysaccharide monooxygenase provides insight into copper binding, protein dynamics, and substrate interactions. *Proc Natl Acad Sci U S A* 109:18779–18784. <https://doi.org/10.1073/pnas.1208822109>.
23. Beeson WT, Phillips CM, Cate JH, Marletta MA. 2012. Oxidative cleavage of cellulose by fungal copper-dependent polysaccharide monooxygenases. *J Am Chem Soc* 134:890–892. <https://doi.org/10.1021/ja210657t>.
24. Chalak A, Villares A, Moreau C, Haon M, Grisel S, d'Orlando A, Herpöel-Gimbert I, Labourel A, Cathala B, Berrin J-G. 2019. Influence of the carbohydrate-binding module on the activity of a fungal AA9 lytic polysaccharide monooxygenase on cellulosic substrates. *Biotechnol Biofuels* 12:206. <https://doi.org/10.1186/s13068-019-1548-y>.
25. Courtade G, Forsberg Z, Heggset EB, Eijsink VGH, Aachmann FL. 2018. The carbohydrate-binding module and linker of a modular lytic polysaccharide monooxygenase promote localized cellulose oxidation. *J Biol Chem* 293:13006–13015. <https://doi.org/10.1074/jbc.RA118.004269>.
26. Ruiz D, Turowski V, Murakami M. 2016. Effects of the linker region on the structure and function of modular GH5 cellulases. *Sci Rep* 6:28504. <https://doi.org/10.1038/srep28504>.
27. Sammond DW, Payne CM, Brunecky R, Himmel ME, Crowley MF, Beckham GT. 2012. Cellulase linkers are optimized based on domain type and function: insights from sequence analysis, biophysical measurements, and molecular simulation. *PLoS One* 7:e48615. <https://doi.org/10.1371/journal.pone.0048615>.
28. Amore A, Knott BC, Supekar NT, Shajahan A, Azadi P, Zhao P, Wells L, Linger JG, Hobdley SE, Vander Wall TA, Shollenberger T, Yarbrough J, Tan Z, Crowley MF, Himmel ME, Decker SR, Beckham GT, Taylor LE. 2017. Distinct roles of N- and O-glycans in cellulase activity and stability. *Proc Natl Acad Sci U S A* 114: 13667–13672. <https://doi.org/10.1073/pnas.1714249114>.
29. Beckham GT, Dai Z, Matthews JF, Momany M, Payne CM, Adney WS, Baker SE, Himmel ME. 2012. Harnessing glycosylation to improve cellulase activity. *Curr Opin Biotechnol* 23:338–345. <https://doi.org/10.1016/j.copbio.2011.11.030>.
30. Hansson H, Karkehabadi S, Mikkelsen N, Douglas NR, Kim S, Lam A, Kaper T, Kelemen B, Meier KK, Jones SM, Solomon EI, Sandgren M. 2017. High-resolution structure of a lytic polysaccharide monooxygenase from *Hypocrea jecorina* reveals a predicted linker as an integral part of the catalytic domain. *J Biol Chem* 292:19099–19109. <https://doi.org/10.1074/jbc.M117.799767>.
31. Payne CM, Resch MG, Chen L, Crowley MF, Himmel ME, Taylor LE, 2nd, Sandgren M, Ståhlberg J, Stals I, Tan Z, Beckham GT. 2013. Glycosylated linkers in multimodular lignocellulose-degrading enzymes dynamically bind to cellulose. *Proc Natl Acad Sci U S A* 110:14646–14651. <https://doi.org/10.1073/pnas.1309106110>.
32. Sonan GK, Receveur-Brechot V, Duez C, Aghajari N, Czjzek M, Haser R, Gerday C. 2007. The linker region plays a key role in the adaptation to cold of the cellulase from an Antarctic bacterium. *Biochem J* 407:293–302. <https://doi.org/10.1042/BJ20070640>.
33. Gao L, Wang L, Jiang X, Qu Y. 2015. Linker length and flexibility induces new cellobiohydrolase activity of PoCel6A from *Penicillium oxalicum*. *Biotechnol J* 10:899–904. <https://doi.org/10.1002/biot.201400734>.
34. McCartney L, Blake AW, Flint J, Bolam DN, Boraston AB, Gilbert HJ, Knox JP. 2006. Differential recognition of plant cell walls by microbial xylan-specific carbohydrate-binding modules. *Proc Natl Acad Sci U S A* 103: 4765–4770. <https://doi.org/10.1073/pnas.0508887103>.
35. Chen X, Zaro J, Shen WC. 2013. Fusion protein linkers: property, design and functionality. *Adv Drug Deliv Rev* 65:1357–1369. <https://doi.org/10.1016/j.addr.2012.09.039>.
36. Guruprasad K, Reddy BV, Pandit MW. 1990. Correlation between stability of a protein and its dipeptide composition: a novel approach for predicting *in vivo* stability of a protein from its primary sequence. *Protein Eng* 4: 155–161. <https://doi.org/10.1093/protein/4.2.155>.
37. Eibinger M, Ganner T, Bubner P, Rožker S, Kracher D, Haltrich D, Ludwig R, Plank H, Nidetzky B. 2014. Cellulose surface degradation by a lytic polysaccharide monooxygenase and its effect on cellulase hydrolytic efficiency. *J Biol Chem* 289:35929–35938. <https://doi.org/10.1074/jbc.M114.602227>.
38. Kallioinen A, Puranen T, Siika-Aho M. 2014. Mixtures of thermostable enzymes show high performance in biomass saccharification. *Appl Biochem Biotechnol* 173:1038–1056. <https://doi.org/10.1007/s12010-014-0893-3>.
39. Fowler CA, Sabbadin F, Ciano L, Hemsworth GR, Elias L, Bruce N, McQueen-Mason S, Davies GJ, Walton PH. 2019. Discovery, activity and characterisation of an AA10 lytic polysaccharide oxygenase from the shipworm symbiont *Teredinibacter turnerae*. *Biotechnol Biofuels* 12:232. <https://doi.org/10.1186/s13068-019-1573-x>.
40. Hemsworth GR, Henrissat B, Davies GJ, Walton PH. 2014. Discovery and characterization of a new family of lytic polysaccharide monooxygenases. *Nat Chem Biol* 10:122–126. <https://doi.org/10.1038/nchembio.1417>.
41. Voshol GP, Vijgenboom E, Punt PJ. 2017. The discovery of novel LPMO families with a new Hidden Markov model. *BMC Res Notes* 10:105. <https://doi.org/10.1186/s13104-017-2429-8>.
42. Badino SF, Bathke JK, Sørensen TH, Windahl MS, Jensen K, Peters GHJ, Borch K, Westh P. 2017. The influence of different linker modifications on the catalytic activity and cellulose affinity of cellobiohydrolase Cel7A from *Hypocrea jecorina*. *Protein Eng Des Sel* 30:495–501. <https://doi.org/10.1093/protein/gzx036>.
43. Srisodsuk M, Reinikainen T, Penttilä M, Teeri TT. 1993. Role of the interdomain linker peptide of *Trichoderma reesei* cellobiohydrolase I in its interaction with crystalline cellulose. *J Biol Chem* 268:20756–20761. [https://doi.org/10.1016/S0021-9258\(19\)36847-4](https://doi.org/10.1016/S0021-9258(19)36847-4).
44. Beckham GT, Bomble YJ, Matthews JF, Taylor CB, Resch MG, Yarbrough JM, Decker SR, Bu L, Zhao X, McCabe C, Wohler J, Bergenstråhle M, Brady JW, Adney WS, Himmel ME, Crowley MF. 2010. The O-glycosylated linker from the *Trichoderma reesei* Family 7 cellulase is a flexible, disordered protein. *Biophys J* 99:3773–3781. <https://doi.org/10.1016/j.bpj.2010.10.032>.
45. Steentoft C, Vakhrushev SY, Joshi HJ, Kong Y, Vester-Christensen MB, Schjoldager KT, Lavrsen K, Dabelsteen S, Pedersen NB, Marcos-Silva L, Gupta R, Bennett EP, Mandel U, Brunak S, Wandall HH, Lavery SB, Clausen H. 2013. Precision mapping of the human O-GalNAc glycoproteome through SimpleCell technology. *EMBO J* 32:1478–1488. <https://doi.org/10.1038/emboj.2013.79>.
46. Miao Y, Kong Y, Li P, Li G, Liu D, Shen Q, Zhang R. 2018. Effect of CBM1 and linker region on enzymatic properties of a novel thermostable dimeric GH10 xylanase (Xyn10A) from filamentous fungus *Aspergillus fumigatus* Z5. *AMB Expr* 8:44. <https://doi.org/10.1186/s13568-018-0576-5>.
47. Xu J, Cheng JJ, Sharma-Shivappa RR, Burns JC. 2010. Sodium hydroxide pretreatment of switchgrass for ethanol production. *Energy Fuels* 24: 2113–2119. <https://doi.org/10.1021/ef9014718>.
48. Jin SY, Chen HZ. 2007. Near-infrared analysis of the chemical composition of rice straw. *Ind Crops Prod* 26:207–211. <https://doi.org/10.1016/j.indcrop.2007.03.004>.
49. Yin Y, Mao X, Yang J, Chen X, Mao F, Xu Y. 2012. dbCAN: a web resource for automated carbohydrate-active enzyme annotation. *Nucleic Acids Res* 40:W445–51. <https://doi.org/10.1093/nar/gks479>.
50. Yoon BJ. 2009. Hidden Markov Models and their Applications in Biological Sequence Analysis. *Curr Genomics* 10:402–415. <https://doi.org/10.2174/138920209789177575>.
51. Finn RD, Clements J, Eddy SR. 2011. HMMER web server: interactive sequence similarity searching. *Nucleic Acids Res* 39:W29–37. <https://doi.org/10.1093/nar/gkr367>.

52. Gasteiger E, Hoogland C, Gattiker A, Duvaud S, Wilkins MR, Appel RD, Bairoch A. 2005. protein identification and analysis tools on the ExPASy server. In Walker JM (ed). The proteomics protocols handbook, pg. 571–607. Springer Protocols Handbooks, Humana Press, Clifton, NJ. <https://doi.org/10.1385/1-59259-890-0:571>.
53. Garnier J, Osguthorpe DJ, Robson B. 1978. Analysis of the accuracy and implications of simple methods for predicting the secondary structure of globular proteins. *J Mol Biol* 120:97–120. [https://doi.org/10.1016/0022-2836\(78\)90297-8](https://doi.org/10.1016/0022-2836(78)90297-8).
54. Kelley LA, Mezulis S, Yates CM, Wass MN, Sternberg MJ. 2015. The Phyre2 web portal for protein modeling, prediction and analysis. *Nat Protoc* 10: 845–858. <https://doi.org/10.1038/nprot.2015.053>.
55. Yang J, Zhang Y. 2015. I-TASSER server: new development for protein structure and function predictions. *Nucleic Acids Res* 43:W174–81. <https://doi.org/10.1093/nar/gkv342>.
56. Zhang Y. 2008. I-TASSER server for protein 3D structure prediction. *BMC Bioinformatics* 9:40. <https://doi.org/10.1186/1471-2105-9-40>.
57. Laskowski RA, MacArthur MW, Moss DS, Thornton JM. 1993. PROCHECK: a program to check the stereochemical quality of protein structures. *J Appl Crystallogr* 26:283–291. <https://doi.org/10.1107/S0021889892009944>.
58. Guex N, Peitsch MC. 1997. SWISS-MODEL and the Swiss-PdbViewer: an environment for comparative protein modeling. *Electrophoresis* 18:2714–2723. <https://doi.org/10.1002/elps.1150181505>.
59. Anderson RJ, Weng Z, Campbell RK, Jiang X. 2005. Main-chain conformational tendencies of amino acids. *Proteins* 60:679–689. <https://doi.org/10.1002/prot.20530>.
60. Waterhouse A, Bertoni M, Bienert S, Studer G, Tauriello G, Gumienny R, Heer FT, de Beer TAP, Rempfer C, Bordoli L, Lepore R, Schwede T. 2018. SWISS-MODEL: homology modelling of protein structures and complexes. *Nucleic Acids Res* 46:W296–303. <https://doi.org/10.1093/nar/gky427>.
61. Ogunmolu FE, Kaur I, Gupta M, Bashir Z, Pasari N, Yazdani SS. 2015. Proteomics insights into the biomass hydrolysis potentials of a hypercellulolytic fungus *Penicillium funiculosum*. *J Proteome Res* 14:4342–4358. <https://doi.org/10.1021/acs.jproteome.5b00542>.
62. Cheng C, Haider J, Liu P, Yang J, Tan Z, Huang T, Lin J, Jiang M, Liu H, Zhu L. 2020. Engineered LPMO significantly boosting cellulase-catalyzed depolymerization of cellulose. *J Agric Food Chem* 68:15257–15266. <https://doi.org/10.1021/acs.jafc.0c05979>.