# Deciphering Cellodextrin and Glucose Uptake in *Clostridium thermocellum*

Fei Yan,[a,b,c,d,e] Sheng Dong,[a,b,c,d,e] Ya-Jun Liu,[a,b,c,d,e] Xingzhe Yao,[a,b,c,d,e*] Chao Chen,[a,b,c,d,e] Yan Xiao,[a,b,c] [iD] Edward A. Bayer,[f,g] Yuval Shoham,[h] Chun You,[i] Qiu Cui,[a,b,c,d,e] [iD] Yingang Feng[a,b,c,d,e]

aCAS Key Laboratory of Biofuels, Shandong Provincial Key Laboratory of Synthetic Biology, Qingdao Institute of Bioenergy and Bioprocess Technology, Chinese Academy of Sciences, Qingdao, China
bShandong Energy Institute, Shandong, Qingdao, China
cQingdao New Energy Shandong Laboratory, Shandong, Qingdao, China
dDalian National Laboratory for Clean Energy, Shandong, Qingdao, China
eUniversity of Chinese Academy of Sciences, Beijing, China
fDepartment of Biomolecular Sciences, The Weizmann Institute of Science, Rehovot, Israel
gDepartment of Life Sciences and the National Institute for Biotechnology in the Negev, Ben-Gurion University of the Negev, Beer-Sheva, Israel
hDepartment of Biotechnology and Food Engineering, Technion-Israel Institute of Technology, Haifa, Israel
iTianjin Institute of Industrial Biotechnology, Chinese Academy of Sciences, Tianjin, China

Fei Yan, Sheng Dong, and Ya-Jun Liu contributed equally to this work. Author order was determined in order of increasing seniority.

**ABSTRACT** Sugar uptake is of great significance in industrially relevant microorganisms. *Clostridium thermocellum* has extensive potential in lignocellulose biorefineries as an environmentally prominent, thermophilic, cellulolytic bacterium. The bacterium employs five putative ATP-binding cassette transporters which purportedly take up cellulose hydrolysates. Here, we first applied combined genetic manipulations and biophysical titration experiments to decipher the key glucose and cellodextrin transporters. *In vivo* gene inactivation of each transporter and *in vitro* calorimetric and nuclear magnetic resonance (NMR) titration of each putative sugar-binding protein with various saccharides supported the conclusion that only transporters A and B play the roles of glucose and cellodextrin transport, respectively. To gain insight into the structural mechanism of the transporter specificities, 11 crystal structures, both alone and in complex with appropriate saccharides, were solved for all 5 putative sugar-binding proteins, thus providing detailed specific interactions between the proteins and the corresponding saccharides. Considering the importance of transporter B as the major cellodextrin transporter, we further identified its cryptic, hitherto unknown ATPase-encoding gene as *clo1313_2554,* which is located outside the transporter B gene cluster. The crystal structure of the ATPase was solved, showing that it represents a typical nucleotide-binding domain of the ATP-binding cassette (ABC) transporter. Moreover, we determined that the inducing effect of cellobiose (G2) and cellulose on cellulosome production could be eliminated by deletion of transporter B genes, suggesting the coupling of sugar transport and regulation of cellulosome components. This study provides key basic information on the sugar uptake mechanism of *C. thermocellum* and will promote rational engineering of the bacterium for industrial application.

**IMPORTANCE** Highly efficient sugar uptake is important to microbial cell factories, and sugar transporters are therefore of great interest in the study of industrially relevant microorganisms. *Clostridium thermocellum* is a lignocellulolytic bacterium known for its multienzyme complex, the cellulosome, which is of great potential value in lignocellulose biorefinery. In this study, we clarify the function and mechanism of substrate specificity of the five reported putative sugar transporters using genetic, biophysical, and structural methods. Intriguingly, the results showed that only one of them, transporter B, is the major cellodextrin transporter, whereas another, transporter A, represents the major glucose

transporter. Considering the importance of transporter B, we further identified the missing ATPase gene of transporter B and revealed the correlation between transporter B and cellulosome production. Revealing the mechanism by which *C. thermocellum* utilizes cellodextrins will help pave the way for engineering the strain for industrial applications.

**KEYWORDS** ABC transporter, genetic inactivation, sugar-binding protein, protein structure, cellulosome

Lignocellulose biorefineries provide new pathways for the eco-friendly and sustainable production of energy and chemicals (1–4). One of the major bottlenecks for cost-effective production of biofuels and bio-based chemicals is the recalcitrance of lignocellulose, and thus, studies on lignocellulolytic microorganisms are of great interest (5–8).

*Clostridium thermocellum* (also named *Ruminiclostridium thermocellum*, *Hungateiclostridium thermocellum*, and *Acetivibrio thermocellus*) is a thermophilic, lignocellulolytic, Gram-positive bacterium with great potential to be integrated into various strategies of lignocellulose biorefinery, particularly in consolidated bioprocessing and consolidated bio-saccharification (9–12). *C. thermocellum* produces a multienzyme complex, the cellulosome, which efficiently degrades cellulose into oligosaccharides (13, 14). The intracellular metabolism of carbohydrates in *C. thermocellum* has been extensively investigated and engineered for the production of high-value chemicals (15). However, the process of sugar uptake by *C. thermocellum* has not been comprehensively studied, except for the identification of several putative sugar ATP-binding cassette (ABC) transporters that have been proposed to be responsible for cellodextrin transport (16, 17).

Highly efficient uptake of hydrolyzed substrates is important to cell factories in lignocellulose biorefineries, and sugar transporters are therefore of great interest in the study of lignocellulolytic microorganisms (18). Unlike most cellulolytic fungi and bacteria, which prefer glucose as a carbon source, *C. thermocellum* prefers to utilize oligosaccharides and exhibits poor growth on monosaccharides (19–22). Previously, five putative sugar transporters, denoted A to D and L and belonging to the ABC transport systems group, were identified in the genome of *C. thermocellum*, and the solute-binding lipoproteins (SBP) of these ABC transporters were shown to bind different dextrins by isothermal titration calorimetry (ITC) assays (17). Four SBPs, denoted cellodextrin-binding proteins A to D (CbpA to CbpD), were shown to bind to glucose and/or different lengths of cellodextrins; and one SBP, denoted laminaribiose-binding protein (Lbp), was shown to bind laminaribiose. However, the function(s) of these putative sugar transporters has not been verified *in vivo*.

In this study, we first investigate the function of the five putative sugar transporters by genetic inactivation of each transporter. Surprisingly, the growth phenotypes of these mutants indicated that transporter B is the sole cellodextrin transporter, whereas transporter A is the sole glucose transporter. These results were further supported by both biophysical and structural evidence, which provided the mechanism of substrate specificity of the sugar-binding proteins in these transporters. Considering the importance of transporter B as the sole cellodextrin transporter, we identified the ATPase gene which is missing in the transporter B gene cluster. The correlation between transporter B and cellulosome production was determined by characterizing the transporter B deletion mutant.

## RESULTS

**Genetic identification of key ABC transporters for glucose and cellodextrin.** To further verify the function of the putative sugar ABC transporters *in vivo*, we inactivated each transporter gene by using thermotargetrons, which insert a group II intron RNA into the first gene of the operon (Fig. 1). Inactivation mutants of transporters A, C, D, and L (denoted Δ*transporterA*, Δ*transporterC*, Δ*transporterD*, and Δ*transporterL*) were obtained with cellobiose (G2) as the carbon source in the GS-2 medium, while the transporter B mutant (denoted Δ*transporterB*) was obtained with glucose as the carbon source.

We then monitored the growth curves of the mutant and wild-type strains in GS-2 media with glucose, cellobiose, or Avicel (microcrystalline cellulose) as the sole carbon source.
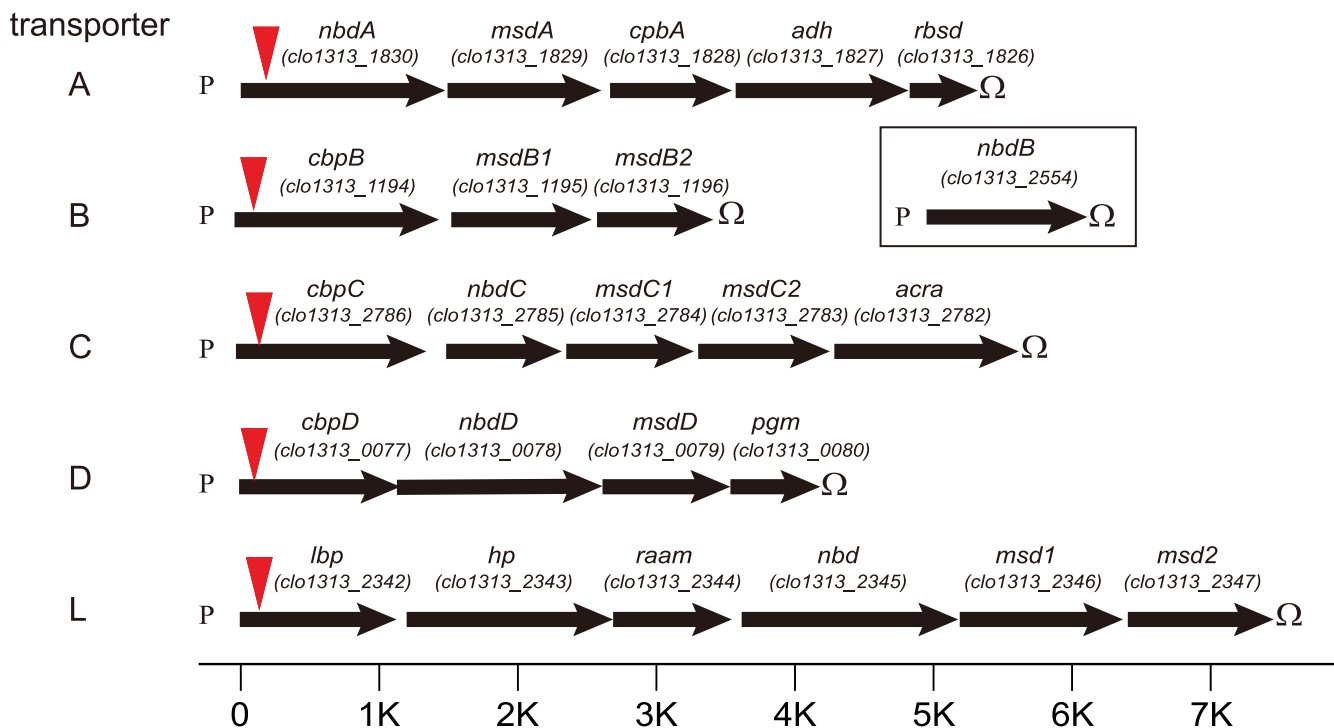
**FIG 1** Schematic representation of the genetic organization of the sugar ATP-binding cassette (ABC) transporters in *Clostridium thermocellum*. The letters P and Ω symbolize potential promoters and rho-independent terminators, respectively. Red triangle indicates intron insertion site. Gene names: *nbd*, nucleotide-binding domain; *msd*, membrane-spanning domain; *cbp*, cellodextrin-binding protein; *adh*, alcohol dehydrogenase; *rsbd*, RbsD/FucU transport protein family; *acra*, multidrug efflux pump subunit AcrA; *lbp*, laminaribiose-binding protein; *pgm*, phosphoglycerate mutase; *hp*, hypothetical protein; *rsam*, radical *S*-adenosylmethionine.

Although a previous study (17) showed that Cbp A to D can all bind to cellodextrin and only CbpC binds to glucose, we found in this study that the Δ*transporterA* strain lost the ability to grow on glucose and the Δ*transporterB* strain was unable to grow on cellobiose or Avicel (Fig. 2). Unexpectedly, the Δ*transporterC*, Δ*transporterD*, and Δ*transporterL* strains grew normally on all three substrates (Fig. 2). Considering that thermotargetrons potentially have additional off-target insertions (23, 24), we additionally constructed transporter A and transporter B deletion mutants using seamless genome editing (25). As shown in Fig. S1 in the supplemental material, the growth phenotype of the mutant strains was the same as that of the mutants obtained by thermotargetron inactivation, and the respective pheno-type can be restored by using plasmid-based expression of the transporter genes. These results indicated that transporter A is the major glucose transporter and transporter B is the major cellodextrin transporter in *C. thermocellum*.

**Substrate specificity of the Cbps.** The results of our genetic experiments indicated that transporters A and B are essential for supporting growth on glucose and cellodextrin, respectively. However, previous ITC experiments using SBPs and the appropriate sugars showed that transporters A to D can all bind to cellodextrins, whereas only CbpC can bind to glucose (17). To further address the discrepancy between our genetic data and previous ITC data in the literature, we performed ITC and NMR experiments to determine the inter-actions between Cbps and various sugars, including cellodextrins (G2 to G5), glucose (G1), and xylose. The ITC experiments showed that among all of the tested substrates, only CbpA binds to glucose, with a dissociation constant ($K_D$) of 240 $\mu$M (Fig. 3, Table 1). In agreement with the ITC data, NMR titrations showed that CbpA binds glucose but not cel-lobiose or cellotriose (Fig. S2). CbpB showed high affinities for all of the tested cellodex-trins (G2 to G5), with the highest affinity for cellotetraose ($K_D$ = 0.562 $\mu$M) (Fig. 3, Table 1). Interestingly, CbpB also appeared to interact with glucose, but the data could not be fitted to obtain reasonable N and $K_D$ values. Furthermore, NMR titration showed that CbpB can bind to cellobiose but not to glucose (Fig. S2). For CbpC and CbpD, no interaction was detected for all of the tested carbohydrate substrates in both ITC and NMR experiments
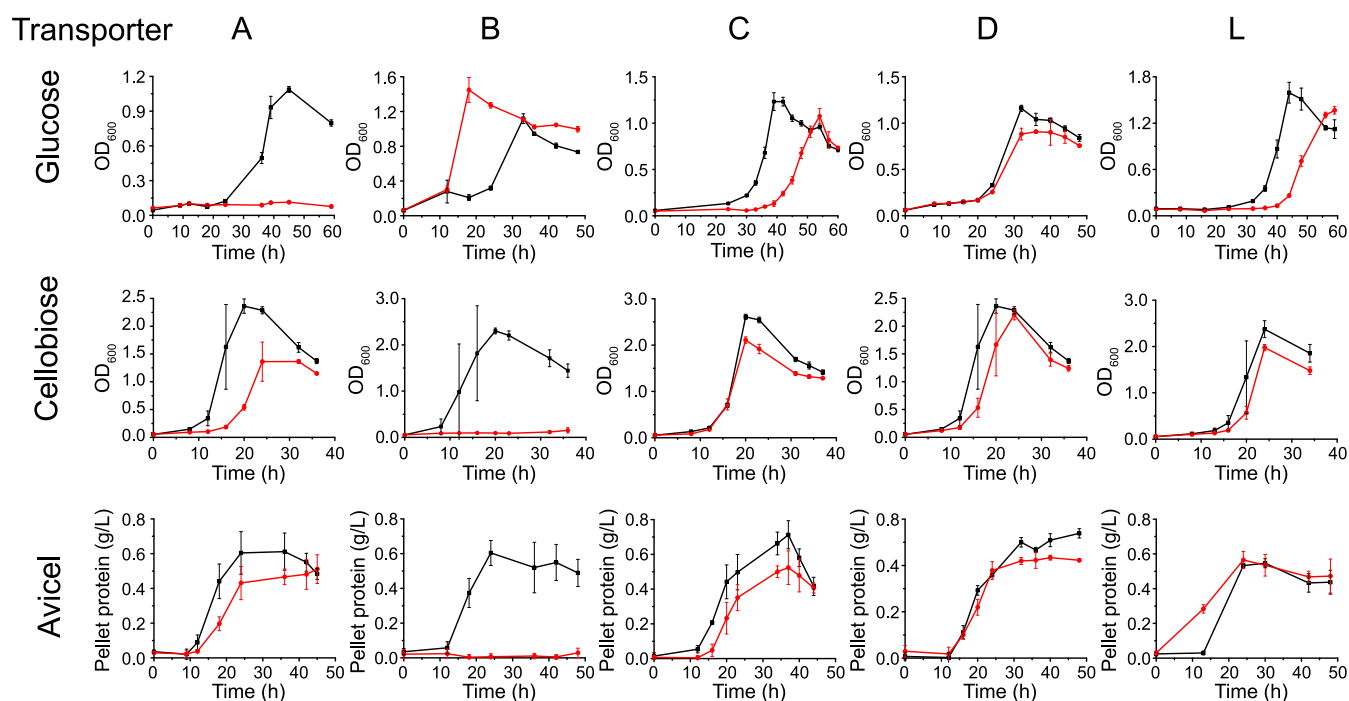
**FIG 2** Growth curves of *C. thermocellum* transporter inactivation mutants (red) and wild type (WT) (black) on glucose, cellobiose, and Avicel. Optical density at 600 nm ($OD_{600}$) and the pellet protein were measured on soluble (glucose and cellobiose) and insoluble (Avicel) carbon sources, respectively. All data represent the mean of triplicate cultures and bars indicate ± standard deviation (SD).

(Fig. S2), suggesting that transporters C and D are not responsible for glucose or cellodextrin transport. Thus, the ITC and NMR results are consistent with the genetic experiments, confirming the essential roles of transporters A and B for glucose and cellodextrin transportation, respectively.

**Overall structures of the Cbps.** To identify the carbohydrate ligand-binding mechanism of the different Cbps, we successfully solved the crystal structures of unliganded CbpA, CbpB, CbpC, and CbpD, as well as the complex structures of CbpA with glucose and CbpB with cellobiose, cellotriose, cellotetraose, and cellopentaose (Table S1). Similar to the typical SBP fold, all of the Cbps adopted two-domain structures connected by three hinge regions, with the buried ligand-binding site between the two domains. Although each Cbp was crystallized in the presence of various carbohydrates, including glucose, xylose, cellobiose, cellotriose, cellotetraose, and cellopentaose, only those Cbps and ligands showing interactions in the ITC experiments were successfully crystallized in the liganded forms.

**Structure of CbpA.** The crystal structures of both unliganded and glucose-liganded CbpA were determined to 2.10 and 1.85 Å resolution, respectively (Table S1). Unambiguous electron density for $\beta$-D-glucose was observed in the CbpA/Glc complex, indicating that the ligand molecule was bound in a single chair conformation. According to the "Venus-flytrap" mechanism (26, 27), i.e., ligand binding causes a dramatic conformational change of two-domain SBPs from an open to a closed state. Hence, the structure of unliganded CbpA assumes an open conformation (Fig. 4A), while CbpA in complex with glucose adopts a closed-conformation, resulting in approximately 37° rigid-body rotation with respect to the unliganded CbpA structure (Fig. 4B). The glucose is bound via hydrogen-bonding interactions and CH-Pi interactions with CbpA (Fig. 4C). Furthermore, the binding pocket of CbpA is distinctly small for the accommodation of a monosaccharide, and no additional space appears available for the binding of larger cellodextrins in CbpA, thus explaining the functional specificity of transporter A.

We searched similar structures in the Protein Data Bank (PDB) using the BLAST (28) and Dali servers (29) with the amino acid sequence and structure of CbpA, respectively. The top hits of both searches were monosaccharide-binding proteins of ABC transporters for the transport of glucose, xylose, arabinose, ribose, etc. Structural comparisons with some of them (30–34) showed that they could be superimposed well in the overall structure (Fig. 4D).
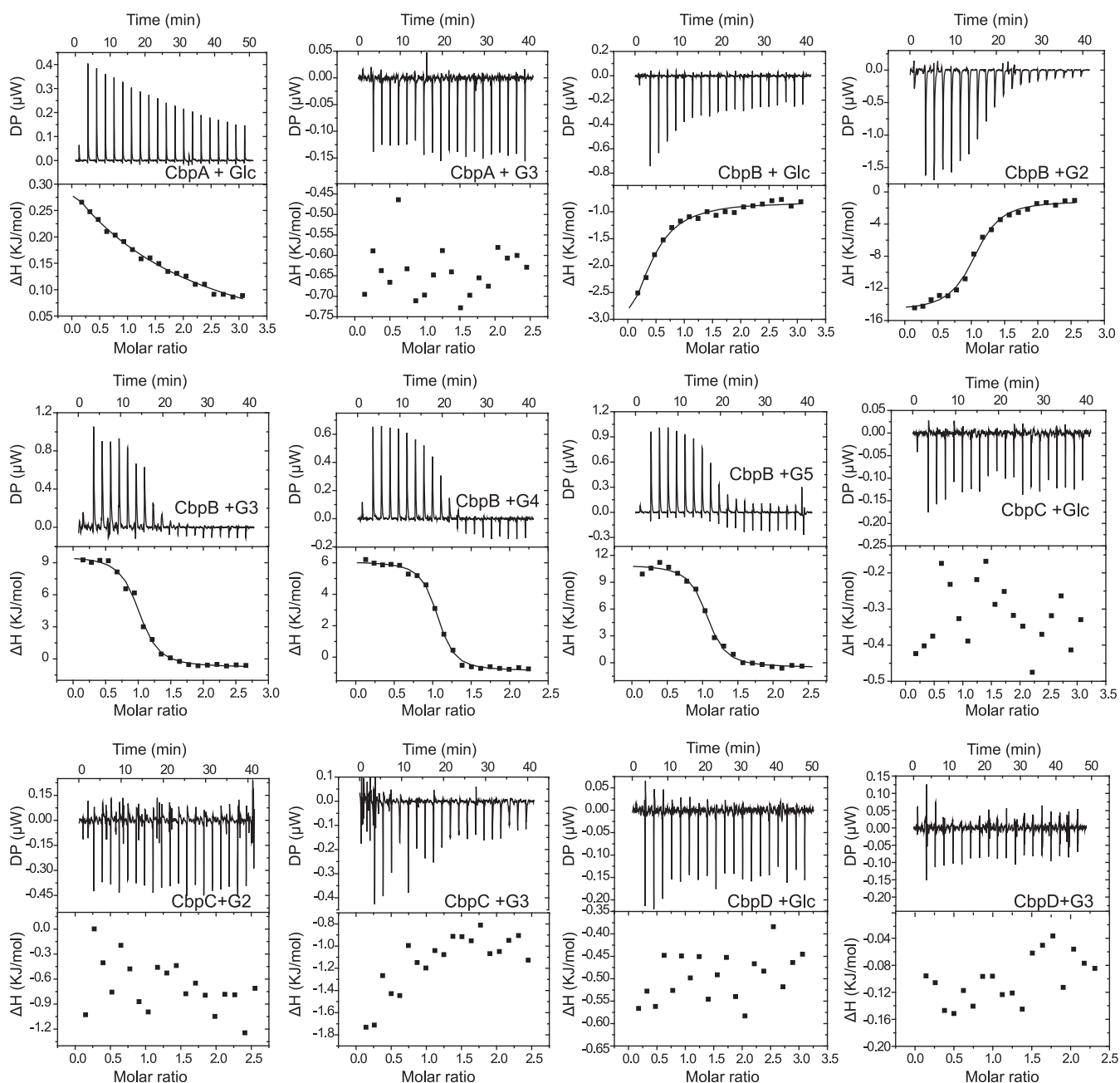
**FIG 3** Representative isothermal titration calorimetry (ITC) data for titrations of various carbohydrate ligands with the cellodextrin-binding proteins (Cbps). ITC traces shown on top, integrated binding isotherms shown on the bottom. The Cbp and substrate used are shown in each ITC trace. Glc, glucose; G2, cellobiose; G3, cellotriose; G4, cellotetraose; G5, cellopentaose.

However, some hydrogen bonds between residues with the ligand were missing in CbpA compared to the reference structures. For example, in XylFII from *Clostridium beijerinkii* and tmGBP from *Thermotoga maritima*, H22 and K9 form hydrogen bonds with O4 of the ligand, respectively. However, in *C. thermocellum* CbpA, a phenylalanine residue (F47) occupies the corresponding location and thus cannot form a similar hydrogen bonding interaction with the substrate. In addition, W54 in CbpA adopted a conformation that is also inconsistent with hydrogen bond formation (Fig. 4E). Furthermore, Y190 in tmGBP forms aromatic ring stacking with glucose, as opposed to D229 which occupies the same position in CbpA (Fig. 4F). The reduced interactions between sugar and protein in CbpA are consistent with the fact that the affinity between glucose and CbpA ($K_D$ of 240 $\mu$M) is much lower than that of the other binding proteins ($K_D$ of 1 $\mu$M or less).

**TABLE 1** Thermodynamic parameters determined by ITC measurements[a]

| Sbp | Substrate | N | $K_D$ ($\mu$M) | $K_a$ ($\times 10^5 \cdot M^{-1}$) | $\Delta$H (kJ/mol) | $\Delta$G (kJ/mol) | T$\Delta$S (kJ/mol) |
|---|---|---|---|---|---|---|---|
| CbpA | Glucose | 1.06 ± 0.07 | 240 ± 95 | 0.042 ± 0.016 | 1.71 ± 0.36 | −21.0 | 22.7 |
| CbpB | Cellobiose | 1.03 ± 0.02 | 1.89 ± 0.36 | 5.29 ± 0.99 | −13.7 ± 0.4 | −33.2 | 19.5 |
| | Cellotriose | 0.97 ± 0.02 | 1.17 ± 0.23 | 8.55 ± 1.69 | 10.4 ± 0.3 | −34.5 | 44.9 |
| | Cellotetraose | 1.02 ± 0.08 | 0.56 ± 0.08 | 17.8 ± 2.5 | 6.95 ± 0.11 | −36.3 | 43.3 |
| | Cellopentaose | 1.02 ± 0.01 | 0.67 ± 0.15 | 14.9 ± 3.3 | 11.5 ± 0.3 | −35.9 | 47.3 |
| | Laminaribiose | 0.98 ± 0.04 | 10.7 ± 2.8 | 0.93 ± 0.25 | 4.57 ± 0.42 | −28.9 | 33.4 |

[a]ITC, isothermal titration calorimetry; N, number of binding site; $K_D$, dissociation constant; $K_a$, absorption rate constant; $\Delta$H, change in enthalpy; Cbp, cellodextrin-binding protein.

**Structure of CbpB.** The unliganded CbpB structure and the complex structures of CbpB with cellodextrins (G2 to G5) were determined to 1.70 to 2.00 Å resolution (Table S1). A hinge motion was observed between the unliganded CbpB and its complexed structures, and the rigid body rotation angle was ~17° (Fig. 5A). The complex structures of CbpB with G2 to G5 could be superimposed well (Fig. 5B), with a root mean square deviation (RMSD) of 0.15 to 0.51 Å. CbpB exhibited a long substrate-binding pocket between the two domains, and the glucose residues at subsites +1 to +3 showed unambiguous electron densities (Fig. 5C), indicating tight ligand fixations at these subsites. The electron densities for the glucose residue at subsite +4, however, were somewhat ambiguous, and those at subsite +5 were hardly observed, indicating the flexibility of the glucose residues at these subsites. The fifth glucose residue in the CbpB-G5 complex indeed showed a problematic conformation in validation, resulting from insufficient electron density. A large number of residues at subsites +1 to +4 are involved in the extensive hydrogen-bond network formed with the cellodextrin hydroxyl groups in CbpB (Fig. 5D). Furthermore, W90 and W296 lay parallel with the glucose residues at subsites +2 and +3, respectively, and formed aromatic stacking interactions with the ligand. The fifth glucose residue extended outside the protein and had no significant direct interaction with the protein.

After we determined the structure of CbpB, we found that structures of a $\beta$-glucosides-binding protein ($\beta$GlyBP) from *Thermus thermophilus* HB8 with 39% sequence identity had been reported in 2020 (35). $\beta$GlyBP can bind to different types of oligosaccharides with various glycosidic linkages and 2 to 5 glucosyl units (35). Structural comparison of these two proteins showed that the overall structures could be superimposed well (Fig. 5E). The residues at subsites +1 to +3 were conserved and showed similar conformations, whereas some variations existed at subsites +4 and +5 for the two proteins (Fig. 5F). The structural similarity confirms that CbpB is a conserved cellodextrin-binding protein. The structures of the CbpB-cellodextrin complex revealed that the longer cellodextrins had more interactions with protein until DP 4, which is consistent with its highest binding affinity with cellotetraose (Table 1).

Considering that $\beta$GlyBP from *Thermus thermophilus* can bind to laminaribiose (35), we suspected that CbpB also serves as a laminaribiose-binding protein. ITC experiments indicated that CbpB can indeed bind to laminaribiose, albeit with a lower affinity compared to cellobiose (Fig. S3A, Table 1). We also successfully determined the crystal structure of the CbpB-laminaribiose complex (Fig. S3B, Table S1). Laminaribiose was bound in the same binding site as cellobiose with the same hydrogen-bonding pattern at subsite +1 but a different pattern at subsite +2 (Fig. S3C). These data suggest that transporter B can also serve as a laminaribiose transporter. However, the mutant $\Delta$*transporterB* strain can still grow well with laminaribiose as the sole carbon source (Fig. S3D), suggesting that transporter B is not the sole laminaribiose transporter in *C. thermocellum*.

**Structure of other potential sugar-binding proteins.** The unliganded CbpC and CbpD structures were determined to 2.00 and 1.50 Å resolution, respectively (Table S1) (Fig. S4A and C). A Dali search using the CbpC structure revealed similarities with several sugar- and other solute-binding proteins of ABC transporters (Fig. S4B). However, the expected substrate-binding pocket between domain I and domain II in CbpC was blocked by an additional loop that was absent in the other proteins (Fig. S4B), suggesting that CbpC might not be a functional sugar-binding protein. A Dali search using the CbpD structure revealed similarity with several monosaccharide-binding proteins in the closed conformation (Fig. S4D). A nearly
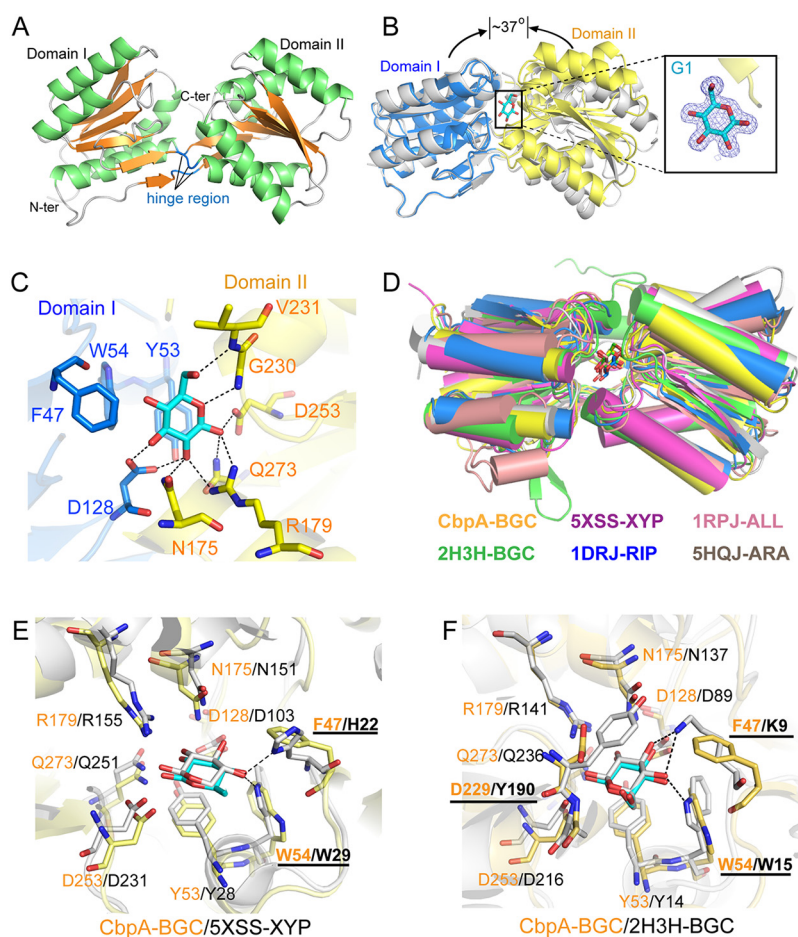
**FIG 4** Structure of cellodextrin-binding protein A (CbpA). (A) Unliganded structure. $\alpha$-Helices and $\beta$-strands are colored in green and orange, respectively. The hinge region includes three loops (colored blue) which connect the two domains. (B) Conformational change of CbpA upon binding of glucose. The CbpA-glucose complex structure (domain I in blue and domain II in yellow) is overlaid with the unliganded CbpA (gray) by the superimposition of domain I. In the inset, the glucose (G1) in the complex structure is shown in cyan sticks, and the 2mFo-DFc densities for the glucose are contoured in blue at 1.0 $\sigma$. (C) Interactions between glucose and surrounding residues of CbpA. Potential hydrogen bonds are shown as dashed lines. (D) Structural comparison of CbpA and other monosaccharide-binding proteins. The Protein Data Bank (PDB) numbers and monosaccharides (BGC, glucose; XYP, xylose; RIP, ribose; ALL, allose; ARA, arabinose according to identifiers from the PDB Chemical Component Dictionary) are indicated with the same color of proteins. (E and F) Detailed comparison of the interactions between the protein and monosaccharide for CbpA (yellow with orange labels) and two other monosaccharide-binding proteins (xylose-binding protein XylFII in panel E and glucose-binding protein tmGBP in panel F, gray with black labels). Additional hydrogen bonds in the structures of XylFII-xylose (PDB ID 5XSS) and tmGBP (PDB ID 2H3H) are shown as dashed lines. Labels of non-conserved binding residues are underlined.

closed substrate-binding pocket between domains I and II was found in the unliganded CbpD crystal structure, which is suitable for accommodating a monosaccharide (Fig. S4D). The small substrate-binding pocket suggests that CbpD cannot bind to cellodextrin, in agreement with our titration and genetic experimental results. CbpD has been proposed to be the transporter of xylose (36), but we could not detect any xylose-binding ability in either ITC or NMR experiments.

The structure of Lbp was determined to 1.50 Å resolution (Table S1, Fig. S5A). Although no sugar was added during crystallization, a clear electron density of a small molecular substrate was observed in the structure. However, neither laminaribiose nor cellobiose can fit into the density. Considering that Lbp has about 35% sequence identity with nucleoside-binding proteins (37, 38) as shown by a BLAST search, we tried to fit different nucleotides and found that guanosine fits well into the density (Fig. S5A). The guanosine has reasonable interactions with the surrounding residues (Fig. S5B), resembling the interactions of the nucleotide-binding
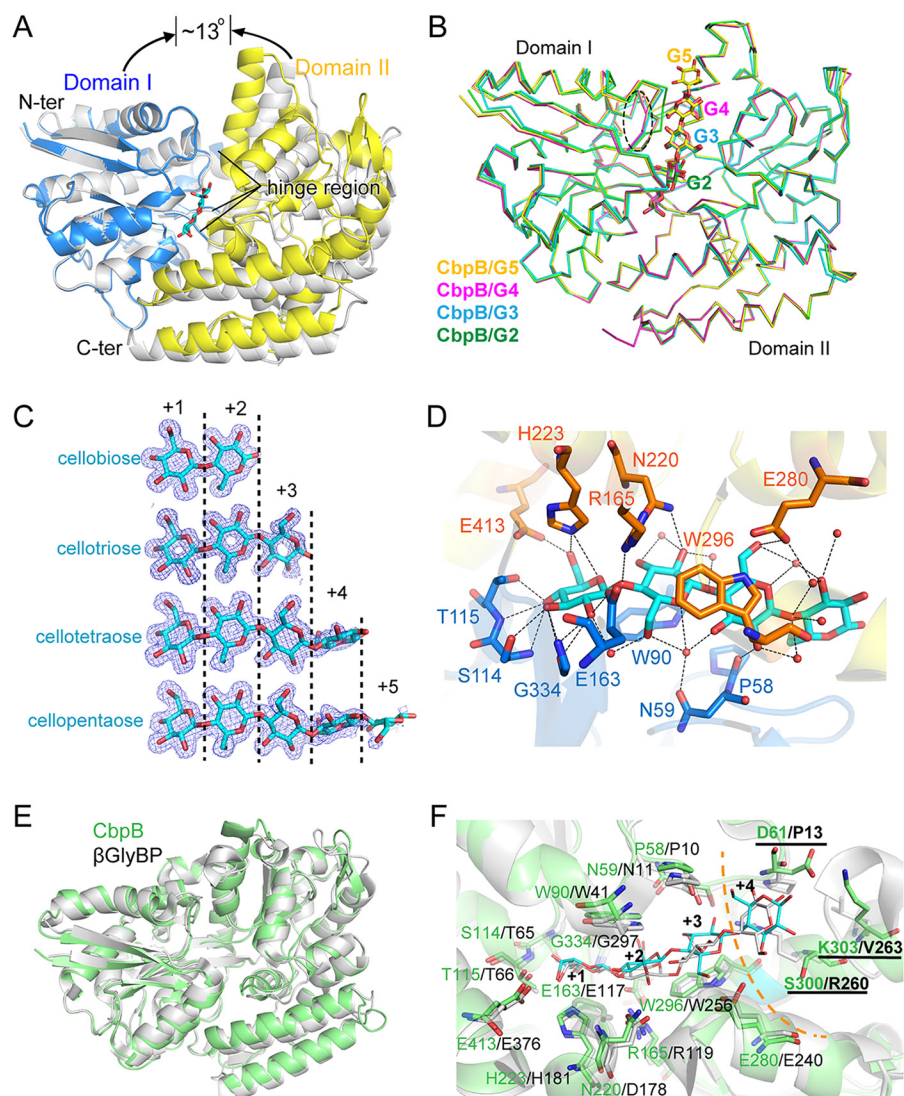
**FIG 5** Structure of CbpB. (A) Conformational change of CbpB upon binding of cellobiose (G2). The CbpB-G2 complex structure (domain I in blue and domain II in yellow) was overlaid with the unliganded CbpB (gray) by the superimposition of domain I. (B) Superimposed backbone traces of CbpB-cellodextrin complexes. CbpB-G2, green; CbpB-G3, cyan; CbpB-G4, magenta; CbpB-G5, yellow. (C) Conformations of cellodextrins in the CbpB-cellodextrin structures. The 2mFo-DFc densities for the cellodextrins are contoured in blue at 1.0 $\sigma$. (D) Residues involved in cellotetraose binding of CbpB. Cellotetraose is shown as cyan sticks and residues from domains I and II are shown as orange and blue sticks, respectively. Water molecules are shown as red spheres. Potential hydrogen bonds between cellotetraose and CbpB are shown as dashed lines. (E and F) Comparison between the structures (E) and sugar-binding sites (F) of CbpB (green) and the $\beta$-glucosides-binding protein ($\beta$GlyBP) from *Thermus thermophilus* HB8 (gray). The cellotetraose molecules complexed with CbpB and $\beta$GlyBP are shown as cyan and gray sticks, respectively. The dashed orange curve separates conserved +1 to +3 subsites (residues with plain labels) and the non-conserved +4 and +5 subsites (residues with underlined labels).

protein from *Treponema pallidum* (PDB ID 2FQX) (Fig. S5C) which can bind various nucleotides (37). This suggests that Lbp may be a nucleotide-binding protein. Furthermore, the Δ*transporterL* strain grew well in the medium with laminaribiose as the sole carbon source (Fig. S5D). Further studies, however, are still required to elucidate the interaction of Lbp with laminaribiose and whether it can function in laminaribiose transport.

**Identification of the gene encoding the ATPase subunit of transporter B.** As shown in a previous study (17), the gene cluster of transporter B in *C. thermocellum* lacks the gene encoding the nucleotide-binding domain NBD subunit (i.e., the ATPase subunit). Because *C. thermocellum* degrades cellulose into cellodextrins as the main form for uptake,
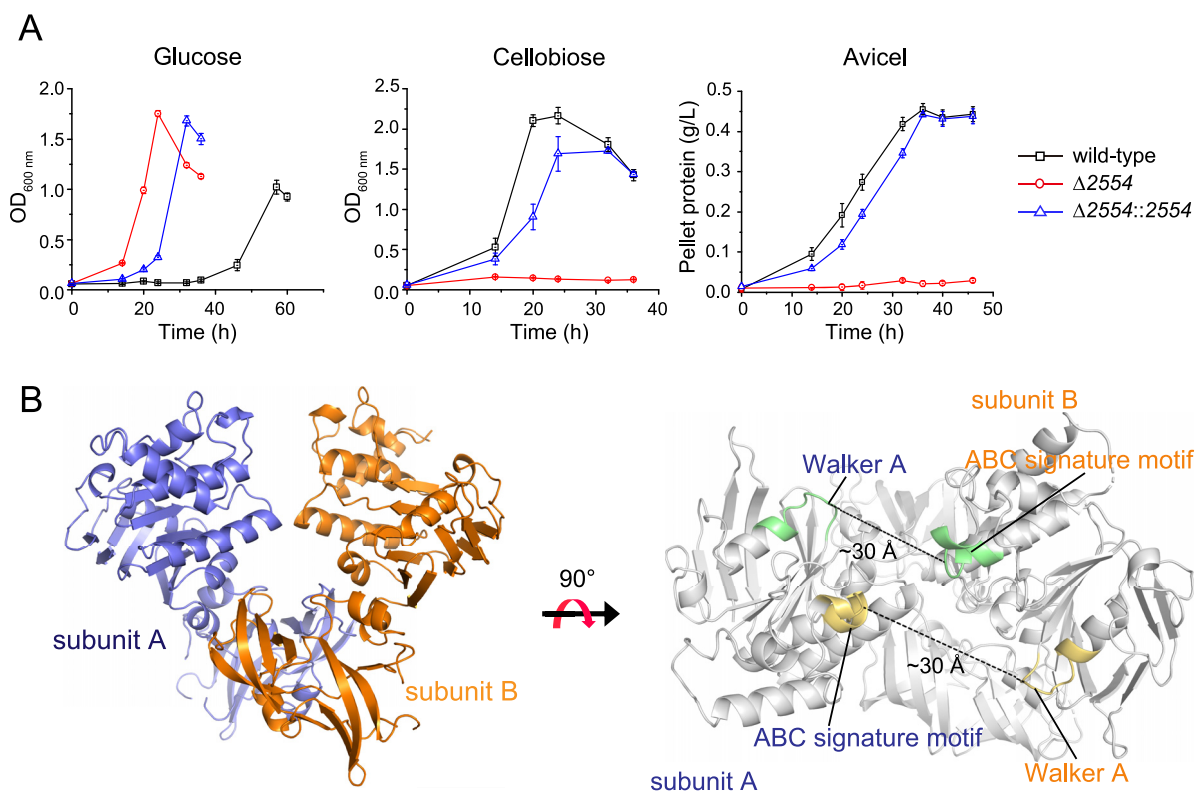
**FIG 6** Function of Clo1313_2554 as the ATPase subunit of transporter B. (A) Growth curves of wild-type (black), *clo1313_2554*-deletion strain (red), and plasmid-based complementation strain (blue) on the designated carbon sources. $OD_{600}$ and the pellet protein were measured on soluble (glucose and cellobiose) and insoluble (Avicel) carbon sources, respectively. All data represent the mean of triplicate cultures and the bars show $\pm$ SD. (B) Structural analysis of Clo1313_2554. Left panel shows the dimeric ATPase structure of Clo1313_2554; the two subunits are shown in blue and orange, respectively. Right panel shows that the two pairs of ATP binding motifs, i.e., the ABC signature motif and Walker A motif (green and orange, respectively), are separated by a distance of $\sim$30 Å, indicating an open conformation.

the major cellodextrin transporter B is of major interest, and identification of its NBD gene is essential. Based on previously reported transcriptomic and proteomic studies (39–44), we noticed that one potential sugar-transporter ATPase gene, *clo1313_2554* in *C. thermocellum* DSM 1313 (*cthe_1862* in *C. thermocellum* ATCC 27405), was highly expressed together with transporter A and B genes. We therefore suspected that Clo1313_2554 represents the ATPase subunit of transporter B. Therefore, we deleted *clo1313_2554* using homologous recombination and obtained the Δ*2554* mutant. The Δ*2554* mutant lost the ability to grow on cellobiose and Avicel (Fig. 6A), in a manner similar to that of the Δ*transporterB* strain (Fig. 2). The growth phenotype was complemented successfully by expression of the *clo1313_2554* gene on a plasmid. In the genome, *clo1313_2554* is predicted to be a standalone gene, i.e., not part of any gene cluster (45). Therefore, the same phenotypes of the Δ*2554* and Δ*transporterB* mutants imply that Clo1313_2554 is the missing ATPase subunit of transporter B, which is the sole cellodextrin transporter in *C. thermocellum*.

To better confirm the function of Clo1313_2554, the crystal structure of Clo1313_2554 was determined at 2.8 Å resolution (Table S1). As shown in Fig. 6B, Clo1313_2554 presents the general $\alpha/\beta$-type ATPase domain fold in homodimer form, similar to that of other NBD crystal structures (46–48). The ATP-binding Walker A motif from subunit A and the ABC signature motif from subunit B were approximately 30 Å apart in the structure, indicating that the nucleotide-free homodimer was in an open conformation (Fig. 6B).

**Inactivation of transporter B eliminated the inducer effect of cellobiose and Avicel for cellulosome expression.** Previous studies have shown that cellodextrin and Avicel can induce the expression of cellulosomes in *C. thermocellum* (49). Since transporter B is the sole cellodextrin transporter, we suspected that it may play a role in the induction of cellulosome expression. The cellulase activities of wild-type, Δ*transporterB*, and Δ*2554* cells were therefore compared in the glucose and glucose+cellobiose media. Wild-type cells
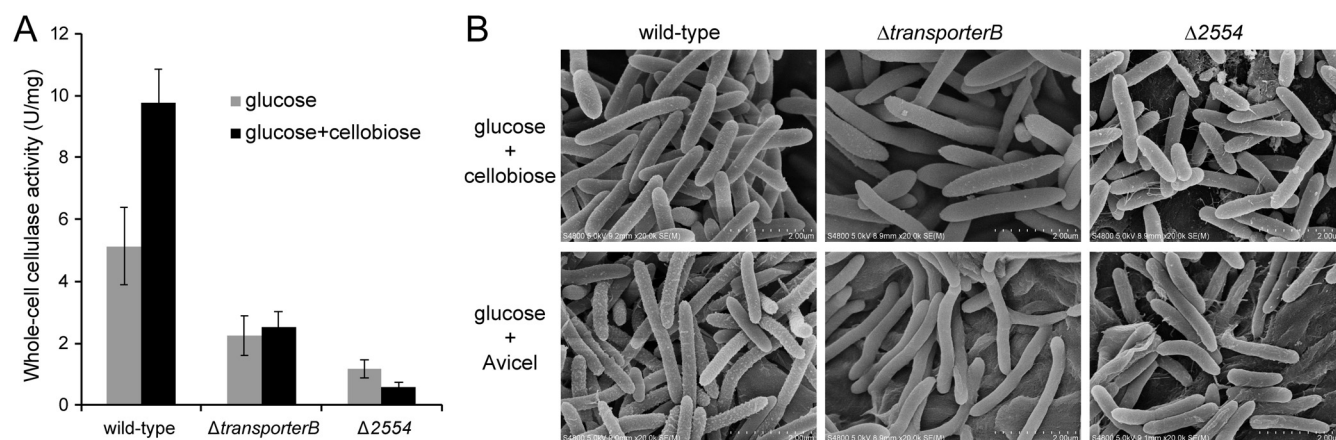
**FIG 7** Cellulase activity (A) and scanning electron microscopic visualization (B) of wild-type, transporter B gene cluster deletion, and *clo1313_2554*-deletion strains of *C. thermocellum* grown on glucose, glucose supplemented with 5 g/L cellobiose, or glucose supplemented with 5 g/L Avicel.

exhibited a significant induction effect of cellobiose for cellulosome expression, while no induction effect was observed in Δ*transporterB* and Δ*2554* cells (Fig. 7A). In scanning electron microscope images of the cells, protuberance structures, indicative of cell surface-borne cellulosomes, can be clearly observed on wild-type cells cultured in glucose+cellobiose- or glucose+Avicel-containing media (Fig. 7B), indicating that the expression of poly cellulosomes is induced by cellobiose/Avicel (50). In contrast, very few protuberance-like structures were observed on the cell surfaces of Δ*transporterB* and Δ*2554* cells in media containing glucose supplemented with cellobiose or Avicel (Fig. 7B). These results indicate that cellodextrin transport and/or subsequent cellodextrin metabolism plays a key role in cellulosome production by the bacterium.

## DISCUSSION

Combined genetic, biophysical, and structural studies have provided solid evidence that transporter A and transporter B are the major glucose and cellodextrin transporters, respectively, among the five potential sugar transporters in *C. thermocellum*. These findings are quite unexpected, because many microorganisms are known to have a multiplicity of redundant sugar transporters for their major carbon source (51, 52). *C. thermocellum* takes up cellodextrins derived from cellulose as its major carbon source, while monosaccharides, such as glucose and fructose, can also be utilized by the bacterium after a lengthy adaptation period (21, 22). Although the physiological and evolutionary benefits of these restrictive transporters are still unknown, these findings suggest that engineering sugar transporters in *C. thermocellum* could be easier than in species with multiple redundant transporters.

The ATPase subunit of transporter B was identified as the encoded product of *clo1313_2554*, which is located outside the gene cluster of transporter B. Interestingly, this feature is found in many other bacteria. For example, the cellodextrin transporter gene cluster of *Clostridium cellulolyticum* H10 also lacks the ATPase gene, and Ccel_2909, which exhibits the highest homology to Clo1313_2554, was proposed to be its ATPase (53). This ATPase was demonstrated to be shared by several sugar transporters, and this phenomenon has been observed in several Gram-positive bacteria (53, 54). We noticed that both gene clusters of the β-glucosides transporter from the Gram-negative bacterium *T. thermophilus* HB8 and the Gram-positive bacterium *Listeria innocua*, which include a CbpB homologous gene, also lack a gene for the ATPase subunit (35). Therefore, the non-cluster location of *clo1313_2554* appears to be a conserved feature in some transporters in both Gram-negative and Gram-positive bacteria. Although this feature was proposed to be the basis for shared ATPase by several sugar transporters (53, 54), the potential benefit of this arrangement in transporter B of *C. thermocellum* remains to be elucidated.

Contrary to previous reports (17), transporter A in *C. thermocellum* was confirmed to be the sole glucose transporter in our study, instead of transporter C. This finding is indeed

consistent with those of two recent studies involving the adaptive evolution of *C. thermocellum* on glucose and fructose, which showed that transporter A genes contained the most frequent mutations in the adaptive strains (21, 22). Unlike other monosaccharide-binding proteins with strong affinities to the substrate ($K_D$ in the $\mu$M range) (51), the binding of CbpA to glucose is remarkably weak as shown by ITC experiments. This weak binding might account for the impaired growth on glucose displayed by *C. thermocellum*. The weak ability of glucose uptake is consistent with the sugar metabolism of *C. thermocellum* in which cellodextrins metabolized through a phosphorolytic mechanism have more bioenergetic benefits than glucose (55).

At this stage, we have no satisfactory explanation for the contradictory results in the binding properties of CbpC and CbpD compared to the previous study by Nataf et al. (17). However, regardless of the exact specificity of CbpC and CbpD, the knockout mutants of these associated transporters clearly indicate that they do not play a significant role in the utilization of cellulose by *C. thermocellum*.

The data in this study showed that transporter B is related to cellulosome expression induced by cellodextrins and Avicel. Although in lignocellulolytic fungi, various sugar transporters are closely related to cellulase induction and production (52, 56), the relationship in bacterial systems is currently unclear. Since bacterial sugar transporters and the cellulosome system components differ from fungal transporters and cellulases, future studies are needed to address the regulation mechanism of transporter B inactivation on cellulosome expression in *C. thermocellum*. Recent studies have shown that some solute-binding proteins of bacterial ABC transporters, including the maltose- and xylose-binding proteins, can activate bacterial receptors for signal transduction with diverse mechanisms (57). Previous studies have shown that several distinct sigma/anti-sigma factors (SigI/RsgI) are responsible for substrate sensing (i.e., RsgI is the receptor) and cellulosome expression (58–61). Therefore, future studies should focus on whether and how transporter B is coupled to SigI/RsgI factors in the regulation of cellulosome expression. These studies will provide both mechanistic insights into cellulosome production and knowledge for effective rational engineering of *C. thermocellum* for lignocellulose biorefineries.

## MATERIALS AND METHODS

**Bacterial strains and cultivation.** The bacterial strains used in this study are listed in Table S2 in the supplemental material. *C. thermocellum* cells were grown anaerobically at 60℃ in GS-2 medium with 5 g/L glucose, cellobiose, laminaribiose, or Avicel as the sole carbon source (62). When necessary, the media were supplemented with 3 $\mu$g/mL thiamphenicol (Tm), 10 $\mu$g/mL 5-fluoro-2-deoxyuridine (FUDR), and 500 $\mu$g/mL 5-fluoroorotic acid (FOA). Cell growth of the *C. thermocellum* strains on glucose, cellobiose, or laminaribiose was determined by monitoring the optical density at 600 nm ($OD_{600}$) every 4 to 8 h. To determine cell growth on Avicel, the cell pellet proteins were quantified by the Bradford method as previously described (24).

**Plasmid construction.** All plasmids for gene expression, inactivation, or deletion were constructed using the *E. coli* Top10 strain and verified by colony PCR and sequencing using the primers listed in Table S3. The thermotargetrons for gene inactivation in *C. thermocellum* were constructed with pHK-TT1A as the template as previously described (23). The resulting targetrons are denoted by a number which corresponds to the 5′ nucleotide residue of the targetron insertion site within the target gene, followed by an "a" or "s" indicating the antisense/bottom or sense/top strand of the insertion site (Table S3). To construct plasmids for gene deletion via homologous recombination in *C. thermocellum*, upstream and downstream homology regions of ～1.2 Kb were obtained by PCR with the genomic DNA of *C. thermocellum* DSM1313 as the template and corresponding primers (Table S3) and were ligated to plasmid pHK-HR as previously described (25). For gene complementation in *C. thermocellum*, the target genes were successively cloned into pHK-$P_{2638}$-BGL to replace the *bgl* gene using a seamless cloning kit (Vazyme Biotech, Beijing, China). Gene expression would be driven by the promoter of gene *clo1313_2638* as previously reported (63). All obtained plasmids for gene inactivation, deletion, and complementation were verified using the primer sets TT1A-F/R, HR-F/R, and pHK-F/R, respectively.

**Electrotransformation and screening of *C. thermocellum*.** The plasmids constructed in *E. coli* Top10 were transformed into *E. coli* BL21(DE3) to remove Dcm methylation. Plasmid transformation to *C. thermocellum* DSM1313 was then performed via electroporation using 0.1-cm electroporation cuvettes and a homemade electroporator with a pulse amplitude of 1.5 kV as previously described (24). After electroporation, cells were recovered for 12 h at 51℃ in 4 mL of GS-2 medium before screening on a solid medium containing Tm and 5 g/L cellobiose or glucose as the sole carbon source. The obtained colonies were screened by colony PCR and sequencing as the target transformants.

Transformants containing the thermotargetron plasmid were analyzed by colony PCR using the primer set TT1A-F/R. The precise targetron insertion in each of the target genes in the genome DNA was then determined using the corresponding sequencing primers (Table S3). After gene targeting, the plasmids were cured by continuous inoculation and growth of cells in a Tm-free fresh medium (23). Transformants containing homologous

recombination plasmid were screened following a previously described two-step protocol (25). The seamless deletion of the operons of transporters A and B and the gene *clo1313_2554* was subsequently verified by colony PCR and sequencing using the primer sets ΔA-F/R, ΔB-F/R, and Δ2554-F/R, respectively.

**Protein expression and purification.** To facilitate heterologous expression and crystallization, the genes encoding CbpA to CbpD, Lbp, and Clo1313_2554 proteins without signal peptide were amplified with primers shown in Table S3 using the genome of *C. thermocellum* DSM1313 as a template. The DNA fragments were digested by the BamHI and XhoI restriction enzymes and then ligated into the pET28a-SMT3 vector. The encoded target proteins in these expression vectors contain an N-terminal His$_6$-SMT3 tag, which could be removed by ULP1 protease treatment in the purification procedures. After sequencing, the expression plasmids were transformed into *E. coli* BL21(DE3) to produce the target proteins.

The proteins were purified by affinity chromatography, ULP1 protease treatment, and gel filtration as previously described (64), except that the final buffer for gel filtration was 10 mM Tris-HCl and 100 mM NaCl (pH 8.0). Protein concentration was determined by UV absorption at 280 nm using the theoretical molar extinction coefficient.

Selenomethionine derivatives of CbpB and CbpC were expressed in an M9 minimal medium supplemented with selenomethionine, lysine, valine, threonine, leucine, isoleucine, and phenylalanine. $^{15}$N-labeled proteins for NMR experiments were obtained by cell cultivation using M9 minimal medium containing $^{15}$N-NH$_4$Cl as the sole nitrogen source. The labeled proteins were purified using the same procedures as for the unlabeled proteins.

**Isothermal titration calorimetry measurements.** Isothermal titration calorimetry measurements were performed at 30°C using a MicroCal PEAQ-ITC (Malvern Panalytical Ltd., Malvern, United Kingdom). Protein samples were dialyzed against buffer containing 50 mM Tris-HCl (pH 8.0) and 100 mM NaCl. Carbohydrate-ligand solutions of cellodextrins, glucose, and xylose were prepared using the same buffer as for protein dialysis. The sample cell was loaded with 280 $\mu$L of protein sample ($\sim$50 $\mu$M), and the reference cell contained distilled water. The syringe was filled with 50 $\mu$L of ligand (500 to 750 $\mu$M). Titrations were carried out by adding 0.8 $\mu$L of ligand for the first injection and 2 $\mu$L for the subsequent 19 injections, with stirring at 750 rpm. Binding parameters were determined by fitting the experimental binding isotherms using a single-site model.

**NMR titration.** NMR titration experiments were performed at 298 K on a Bruker Avance III 600 MHz NMR spectrometer equipped with a cryoprobe. The $^{15}$N-labeled protein samples contained 0.1 or 0.2 mM sugar-binding protein in 20 mM Bis-Tris (pH 6.5), 100 mM KCl, 0.02% (wt/vol) dextran sulfate sodium, and 90% H$_2$O/10% D$_2$O. The concentrations of carbohydrate substrates were 10, 20, or 100 mM in the same buffer. For the titration, the substrates were gradually added to the labeled proteins until the ratio of their concentrations reached 1:20. A series of two-dimensional $^1$H-$^{15}$N HSQC spectra were recorded at different protein-substrate ratios during the titration.

**Crystallization, data collection, structure determination, and refinement.** The purified proteins for crystallization were concentrated to approximately 20 mg/mL in 10 mM Tris-HCl (pH 8.0) and 100 mM NaCl. Crystals were obtained using sitting-drop vapor diffusion for screening and hanging-drop vapor diffusion for optimization at 18°C. To obtain crystals of the Sbp/carbohydrate complex, CbpA protein was mixed with glucose at a 1:50 molar ratio, while CbpB was mixed with each cellodextrin (G2 to G5) at a 1:5 molar ratio. High-quality crystals were obtained under the conditions shown in Table S1. All of the crystals used for data collection were cryoprotected by soaking in well solution supplemented with 20% (vol/vol) glycerol for 10 s, and then flash-cooled in liquid nitrogen. X-ray diffraction data were collected on the BL17U1 or BL19U1 beamline at the Shanghai Synchrotron Radiation Facility (65–67).

The diffraction data indexing, integration, and scaling were conducted using XDS (68). The crystal structures of CbpA, CbpD, Lbp, and Clo1313_2554 were determined by molecular replacement using the protein tmGBP (PDB ID 2H3H) (30), XylFII (PDB ID 5XSS) (31), PnrA (PDB ID 2FQX) (37), and an NBD from *Pyrococcus horikoshii* (PDB ID 1V43) (69), respectively, as search models in the PHENIX program Phaser-MR (70, 71). The structures of the CbpB/cellotriose complex and CbpC Se-derivative were determined by single-wavelength anomalous dispersion phasing using CCP4 CRANK2 (72, 73). One monomer from the resulting structure was subsequently used as the search model for molecular replacement to determine the structures of ligand-free CbpB, CbpC, and the CbpB-cellodextrin complexes. Refinements of the structures were performed using Coot (74) and PHENIX (71). Carbohydrates in all structures were further validated using Privateer in CCP4 (72, 75). All structure figures were made using PyMOL (Schrödinger LLC).

**Scanning electron microscopy.** Scanning electron microscopy was performed with *C. thermocellum* cells according to a previously reported method using a field emission scanning electron microscope S-4800 (Hitachi, Tokyo, Japan) (24).

**Cellulosome activity measurement.** *C. thermocellum* cells were cultivated to the early exponential stage with various carbon sources. Cell suspensions (1.0 mL) were sampled and centrifuged at 15,000 × *g* for 10 min to separate supernatant from the pellet. Cell-associated cellulosome activity was measured in a 1.0-mL reaction volume containing the pellet and 15 mg Avicel as the substrate, using GS-2 medium as the reaction buffer. The concentration of reducing sugars was determined by the 3,5-dinitrosalicylic acid method after incubation at 55°C for 24 h unless otherwise stated. One unit of enzyme activity is defined as the amount of enzyme that releases 1 nmol reducing sugar (glucose equivalent) per min.

**Data availability.** Atomic coordinates and structure factors have been deposited in the PDB. The accession numbers are listed in Table S1.

## SUPPLEMENTAL MATERIAL

Supplemental material is available online only.

**FIG S1**, PDF file, 0.6 MB.

**FIG S2**, PDF file, 1.1 MB.

**FIG S3**, PDF file, 2.1 MB.
**FIG S4**, PDF file, 0.5 MB.
**FIG S5**, PDF file, 2 MB.
**TABLE S1**, PDF file, 0.2 MB.
**TABLE S2**, PDF file, 0.1 MB.
**TABLE S3**, PDF file, 0.1 MB.

## REFERENCES

1. Demain AL. 2009. Biosolutions to the energy problem. J Ind Microbiol Biotechnol 36:319–332. https://doi.org/10.1007/s10295-008-0521-8.
2. Usmani Z, Sharma M, Awasthi AK, Lukk T, Tuohy MG, Gong L, Nguyen-Tri P, Goddard AD, Bill RM, Nayak SC, Gupta VK. 2021. Lignocellulosic biorefineries: the current state of challenges and strategies for efficient commercialization. Renew Sust Energ Rev 148:111258. https://doi.org/10.1016/j.rser.2021.111258.
3. Liu YS, Tang YH, Gao HY, Zhang WM, Jiang YJ, Xin FX, Jiang M. 2021. Challenges and future perspectives of promising biotechnologies for lignocellulosic biorefinery. Molecules 26:5411. https://doi.org/10.3390/molecules26175411.
4. Liu YZ, Cruz-Morales P, Zargar A, Belcher MS, Pang B, Englund E, Dan QY, Yin K, Keasling JD. 2021. Biofuels for a sustainable future. Cell 184:1636–1647. https://doi.org/10.1016/j.cell.2021.01.052.
5. Hasunuma T, Okazaki F, Okai N, Hara KY, Ishii J, Kondo A. 2013. A review of enzymes and microbes for lignocellulosic biorefinery and the possibility of their application to consolidated bioprocessing technology. Bioresour Technol 135:513–522. https://doi.org/10.1016/j.biortech.2012.10.047.
6. Sethupathy S, Morales GM, Li Y, Wang Y, Jiang J, Sun J, Zhu D. 2021. Harnessing microbial wealth for lignocellulose biomass valorization through secretomics: a review. Biotechnol Biofuels 14:154. https://doi.org/10.1186/s13068-021-02006-9.
7. Mankar AR, Pandey A, Modak A, Pant KK. 2021. Pretreatment of lignocellulosic biomass: a review on recent advances. Bioresour Technol 334:125235. https://doi.org/10.1016/j.biortech.2021.125235.
8. Shen X, Sun R. 2021. Recent advances in lignocellulose prior-fractionation for biomaterials, biochemicals, and bioenergy. Carbohydr Polym 261:117884. https://doi.org/10.1016/j.carbpol.2021.117884.
9. Lynd LR, van Zyl WH, McBride JE, Laser M. 2005. Consolidated bioprocessing of cellulosic biomass: an update. Curr Opin Biotechnol 16:577–583. https://doi.org/10.1016/j.copbio.2005.08.009.
10. Olson DG, McBride JE, Shaw AJ, Lynd LR. 2012. Recent progress in consolidated bioprocessing. Curr Opin Biotechnol 23:396–405. https://doi.org/10.1016/j.copbio.2011.11.026.
11. Liu YJ, Li B, Feng Y, Cui Q. 2020. Consolidated bio-saccharification: leading lignocellulose bioconversion into the real world. Biotechnol Adv 40:107535. https://doi.org/10.1016/j.biotechadv.2020.107535.
12. Banner A, Toogood HS, Scrutton NS. 2021. Consolidated bioprocessing: synthetic biology routes to fuels and fine chemicals. Microorganisms 9:1079. https://doi.org/10.3390/microorganisms9051079.
13. Artzi L, Bayer EA, Morais S. 2017. Cellulosomes: bacterial nanomachines for dismantling plant polysaccharides. Nat Rev Microbiol 15:83–95. https://doi.org/10.1038/nrmicro.2016.164.
14. Alves VD, Fontes C, Bule P. 2021. Cellulosomes: highly efficient cellulolytic complexes. Subcell Biochem 96:323–354. https://doi.org/10.1007/978-3-030-58971-4_9.
15. Mazzoli R, Olson DG. 2020. *Clostridium thermocellum*: a microbial platform for high-value chemical production from lignocellulose. Adv Appl Microbiol 113:111–161. https://doi.org/10.1016/bs.aambs.2020.07.004.
16. Strobel HJ, Caldwell FC, Dawson KA. 1995. Carbohydrate transport by the anaerobic thermophile *Clostridium thermocellum* LQRI. Appl Environ Microbiol 61:4012–4015. https://doi.org/10.1128/aem.61.11.4012-4015.1995.
17. Nataf Y, Yaron S, Stahl F, Lamed R, Bayer EA, Scheper TH, Sonenshein AL, Shoham Y. 2009. Cellodextrin and laminaribiose ABC transporters in *Clostridium thermocellum*. J Bacteriol 191:203–209. https://doi.org/10.1128/JB.01190-08.
18. Podolsky IA, Schauer EE, Seppala S, O'Malley MA. 2022. Identification of novel membrane proteins for improved lignocellulose conversion. Curr Opin Biotechnol 73:198–204. https://doi.org/10.1016/j.copbio.2021.08.010.
19. Nochur SV, Roberts MF, Demain AL. 1990. Mutation of *Clostridium thermocellum* in the presence of certain carbon-sources. FEMS Microbiol Lett 71:199–204. https://doi.org/10.1111/j.1574-6968.1990.tb03822.x.
20. Nochur SV, Demain AL, Roberts MF. 1992. Carbohydrate utilization by *Clostridium thermocellum*: importance of internal pH in regulating growth. Enzyme Microb Tech 14:338–349. https://doi.org/10.1016/0141-0229(92)90001-5.
21. Ha-Tran DM, Nguyen TTM, Lo SC, Huang CC. 2021. Utilization of monosaccharides by *Hungateiclostridium thermocellum* ATCC 27405 through adaptive evolution. Microorganisms 9:1445. https://doi.org/10.3390/microorganisms9071445.
22. Yayo J, Kuil T, Olson DG, Lynd LR, Holwerda EK, van Maris AJA. 2021. Laboratory evolution and reverse engineering of *Clostridium thermocellum* for growth on glucose and fructose. Appl Environ Microbiol 87:e03017-20. https://doi.org/10.1128/AEM.03017-20.
23. Mohr G, Hong W, Zhang J, Cui GZ, Yang Y, Cui Q, Liu YJ, Lambowitz AM. 2013. A targetron system for gene targeting in thermophiles and its application in *Clostridium thermocellum*. PLoS One 8:e69032. https://doi.org/10.1371/journal.pone.0069032.
24. Hong W, Zhang J, Feng Y, Mohr G, Lambowitz AM, Cui GZ, Liu YJ, Cui Q. 2014. The contribution of cellulosomal scaffoldins to cellulose hydrolysis by *Clostridium thermocellum* analyzed by using thermotargetrons. Biotechnol Biofuels 7:80. https://doi.org/10.1186/1754-6834-7-80.

25. Zhang J, Liu S, Li R, Hong W, Xiao Y, Feng Y, Cui Q, Liu Y-J. 2017. Efficient whole-cell-catalyzing cellulose saccharification using engineered *Clostridium thermocellum*. Biotechnol Biofuels 10:124. https://doi.org/10.1186/s13068-017-0796-y.

26. Mao B, Pear MR, McCammon JA, Quiocho FA. 1982. Hinge-bending in L-arabinose-binding protein. The "Venus's-flytrap" model. J Biol Chem 257:1131–1133. https://doi.org/10.1016/S0021-9258(19)68161-5.

27. Scheepers GH, Lycklama ANJA, Poolman B. 2016. An updated structural classification of substrate-binding proteins. FEBS Lett 590:4393–4401. https://doi.org/10.1002/1873-3468.12445.

28. Johnson M, Zaretskaya I, Raytselis Y, Merezhuk Y, McGinnis S, Madden TL. 2008. NCBI BLAST: a better web interface. Nucleic Acids Res 36:W5–W9. https://doi.org/10.1093/nar/gkn201.

29. Holm L. 2020. Using Dali for protein structure comparison. Methods Mol Biol 2112:29–42. https://doi.org/10.1007/978-1-0716-0270-6_3.

30. Tian Y, Cuneo MJ, Changela A, Hocker B, Beese LS, Hellinga HW. 2007. Structure-based design of robust glucose biosensors using a *Thermotoga maritima* periplasmic glucose-binding protein. Protein Sci 16:2240–2250. https://doi.org/10.1110/ps.072969407.

31. Li J, Wang C, Yang G, Sun Z, Guo H, Shao K, Gu Y, Jiang W, Zhang P. 2017. Molecular mechanism of environmental D-xylose perception by a XylFII-LytS complex in bacteria. Proc Natl Acad Sci U S A 114:8235–8240. https://doi.org/10.1073/pnas.1620183114.

32. Vyas NK, Vyas MN, Quiocho FA. 1988. Sugar and signal-transducer binding sites of the *Escherichia coli* galactose chemoreceptor protein. Science 242:1290–1295. https://doi.org/10.1126/science.3057628.

33. Chaudhuri BN, Ko J, Park C, Jones TA, Mowbray SL. 1999. Structure of D-allose binding protein from *Escherichia coli* bound to D-allose at 1.8 Å resolution. J Mol Biol 286:1519–1531. https://doi.org/10.1006/jmbi.1999.2571.

34. Bjorkman AJ, Binnie RA, Zhang H, Cole LB, Hermodson MA, Mowbray SL. 1994. Probing protein-protein interactions: the ribose-binding protein in bacterial transport and chemotaxis. J Biol Chem 269:30206–30211. https://doi.org/10.1016/S0021-9258(18)43798-2.

35. Chandravanshi M, Samanta R, Kanaujia SP. 2020. Conformational trapping of a β-glucosides-binding protein unveils the selective two-step ligand-binding mechanism of ABC importers. J Mol Biol 432:5711–5734. https://doi.org/10.1016/j.jmb.2020.08.021.

36. Verbeke TJ, Giannone RJ, Klingeman DM, Engle NL, Rydzak T, Guss AM, Tschaplinski TJ, Brown SD, Hettich RL, Elkins JG. 2017. Pentose sugars inhibit metabolism and increase expression of an AgrD-type cyclic pentapeptide in *Clostridium thermocellum*. Sci Rep 7:43355. https://doi.org/10.1038/srep43355.

37. Deka RK, Brautigam CA, Yang XF, Blevins JS, Machius M, Tomchick DR, Norgard MV. 2006. The PnrA (Tp0319; TmpC) lipoprotein represents a new family of bacterial purine nucleoside receptor encoded within an ATP-binding cassette (ABC)-like operon in *Treponema pallidum*. J Biol Chem 281:8072–8081. https://doi.org/10.1074/jbc.M511405200.

38. Abdullah MR, Batuecas MT, Jennert F, Voß F, Westhoff P, Kohler TP, Molina R, Hirschmann S, Lalk M, Hermoso JA, Hammerschmidt S. 2021. Crystal structure and pathophysiological role of the Pneumococcal nucleoside-binding protein PnrA. J Mol Biol 433:166723. https://doi.org/10.1016/j.jmb.2020.11.022.

39. Raman B, McKeown CK, Rodriguez M, Jr., Brown SD, Mielenz JR. 2011. Transcriptomic analysis of *Clostridium thermocellum* ATCC 27405 cellulose fermentation. BMC Microbiol 11:134. https://doi.org/10.1186/1471-2180-11-134.

40. Rydzak T, McQueen PD, Krokhin OV, Spicer V, Ezzati P, Dwivedi RC, Shamshurin D, Levin DB, Wilkins JA, Sparling R. 2012. Proteomic analysis of *Clostridium thermocellum* core metabolism: relative protein expression profiles and growth phase-dependent changes in protein expression. BMC Microbiol 12:214. https://doi.org/10.1186/1471-2180-12-214.

41. Burton E, Martin VJJ. 2012. Proteomic analysis of *Clostridium thermocellum* ATCC 27405 reveals the upregulation of an alternative transhydrogenase-malate pathway and nitrogen assimilation in cells grown on cellulose. Can J Microbiol 58:1378–1388. https://doi.org/10.1139/cjm-2012-0412.

42. Wei H, Fu Y, Magnusson L, Baker JO, Maness PC, Xu Q, Yang S, Bowersox A, Bogorad I, Wang W, Tucker MP, Himmel ME, Ding SY. 2014. Comparison of transcriptional profiles of *Clostridium thermocellum* grown on cellobiose and pretreated yellow poplar using RNA-Seq. Front Microbiol 5:142. https://doi.org/10.3389/fmicb.2014.00142.

43. Rydzak T, Garcia D, Stevenson DM, Sladek M, Klingeman DM, Holwerda EK, Amador-Noguez D, Brown SD, Guss AM. 2017. Deletion of Type I glutamine synthetase deregulates nitrogen metabolism and increases ethanol production in *Clostridium thermocellum*. Metab Eng 41:182–191. https://doi.org/10.1016/j.ymben.2017.04.002.

44. Dumitrache A, Klingeman DM, Natzke J, Rodriguez M, Jr., Giannone RJ, Hettich RL, Davison BH, Brown SD. 2017. Specialized activities and expression differences for *Clostridium thermocellum* biofilm and planktonic cells. Sci Rep 7:43583. https://doi.org/10.1038/srep43583.

45. Mao X, Ma Q, Zhou C, Chen X, Zhang H, Yang J, Mao F, Lai W, Xu Y. 2014. DOOR 2.0: presenting operons and their functions through dynamic and integrated views. Nucleic Acids Res 42:D654–D659. https://doi.org/10.1093/nar/gkt1048.

46. Beis K. 2015. Structural basis for the mechanism of ABC transporters. Biochem Soc Trans 43:889–893. https://doi.org/10.1042/BST20150047.

47. Diederichs K, Diez J, Greller G, Muller C, Breed J, Schnell C, Vonrhein C, Boos W, Welte W. 2000. Crystal structure of MalK, the ATPase subunit of the trehalose/maltose ABC transporter of the archaeon *Thermococcus litoralis*. EMBO J 19:5951–5961. https://doi.org/10.1093/emboj/19.22.5951.

48. Chen J, Lu G, Lin J, Davidson AL, Quiocho FA. 2003. A tweezers-like motion of the ATP-binding cassette dimer in an ABC transport cycle. Mol Cell 12:651–661. https://doi.org/10.1016/j.molcel.2003.08.004.

49. Li R, Feng Y, Liu S, Qi K, Cui Q, Liu YJ. 2018. Inducing effects of cellulosic hydrolysate components of lignocellulose on cellulosome synthesis in *Clostridium thermocellum*. Microb Biotechnol 11:905–916. https://doi.org/10.1111/1751-7915.13293.

50. Bayer EA, Lamed R. 1986. Ultrastructure of the cell surface cellulosome of *Clostridium thermocellum* and its interaction with cellulose. J Bacteriol 167:828–836. https://doi.org/10.1128/jb.167.3.828-836.1986.

51. Jeckelmann JM, Erni B. 2020. Transporters of glucose and other carbohydrates in bacteria. Pflugers Arch 472:1129–1153. https://doi.org/10.1007/s00424-020-02379-0.

52. Nogueira KMV, Mendes V, Carraro CB, Taveira IC, Oshiquiri LH, Gupta VK, Silva RN. 2020. Sugar transporters from industrial fungi: key to improving second-generation ethanol production. Renew Sust Energ Rev 131:109991. https://doi.org/10.1016/j.rser.2020.109991.

53. Fosses A, Mate M, Franche N, Liu N, Denis Y, Borne R, de Philip P, Fierobe HP, Perret S. 2017. A seven-gene cluster in *Ruminiclostridium cellulolyticum* is essential for signalization, uptake and catabolism of the degradation products of cellulose hydrolysis. Biotechnol Biofuels 10:250. https://doi.org/10.1186/s13068-017-0933-7.

54. Ferreira MJ, Sa-Nogueira I. 2010. A multitask ATPase serving different ABC-type sugar importers in *Bacillus subtilis*. J Bacteriol 192:5312–5318. https://doi.org/10.1128/JB.00832-10.

55. Zhang YH, Lynd LR. 2005. Cellulose utilization by *Clostridium thermocellum*: bioenergetics and hydrolysis product assimilation. Proc Natl Acad Sci U S A 102:7321–7325. https://doi.org/10.1073/pnas.0408734102.

56. Zhang F, Bai F, Zhao X. 2016. Induction and regulation of cellulase expression in filamentous fungi: a review. Sheng Wu Gong Cheng Xue Bao 32:1481–1495. https://doi.org/10.13345/j.cjb.160105.

57. Matilla MA, Ortega A, Krell T. 2021. The role of solute binding proteins in signal transduction. Comput Struct Biotechnol J 19:1786–1805. https://doi.org/10.1016/j.csbj.2021.03.029.

58. Kahel-Raifer H, Jindou S, Bahari L, Nataf Y, Shoham Y, Bayer EA, Borovok I, Lamed R. 2010. The unique set of putative membrane-associated anti-sigma factors in *Clostridium thermocellum* suggests a novel extracellular carbohydrate-sensing mechanism involved in gene regulation. FEMS Microbiol Lett 308:84–93. https://doi.org/10.1111/j.1574-6968.2010.01997.x.

59. Nataf Y, Bahari L, Kahel-Raifer H, Borovok I, Lamed R, Bayer EA, Sonenshein AL, Shoham Y. 2010. *Clostridium thermocellum* cellulosomal genes are regulated by extracytoplasmic polysaccharides via alternative sigma factors. Proc Natl Acad Sci U S A 107:18646–18651. https://doi.org/10.1073/pnas.1012175107.

60. Ortiz de Ora L, Lamed R, Liu YJ, Xu J, Cui Q, Feng Y, Shoham Y, Bayer EA, Munoz-Gutierrez I. 2018. Regulation of biomass degradation by alternative σ factors in cellulolytic clostridia. Sci Rep 8:11036. https://doi.org/10.1038/s41598-018-29245-5.

61. Wei Z, Chen C, Liu YJ, Dong S, Li J, Qi K, Liu S, Ding X, Ortiz de Ora L, Munoz-Gutierrez I, Li Y, Yao H, Lamed R, Bayer EA, Cui Q, Feng Y. 2019. Alternative σ^I/anti-σ^I factors represent a unique form of bacterial σ/anti-σ complex. Nucleic Acids Res 47:5988–5997. https://doi.org/10.1093/nar/gkz355.

62. Cui GZ, Hong W, Zhang J, Li WL, Feng Y, Liu YJ, Cui Q. 2012. Targeted gene engineering in *Clostridium cellulolyticum* H10 without methylation. J Microbiol Methods 89:201–208. https://doi.org/10.1016/j.mimet.2012.02.015.

63. Qi K, Chen C, Yan F, Feng Y, Bayer EA, Kosugi A, Cui Q, Liu YJ. 2021. Coordinated β-glucosidase activity with the cellulosome is effective for enhanced lignocellulose saccharification. Bioresour Technol 337:125441. https://doi.org/10.1016/j.biortech.2021.125441.

64. Yao X, Chen C, Wang Y, Dong S, Liu YJ, Li Y, Cui Z, Gong W, Perrett S, Yao L, Lamed R, Bayer EA, Cui Q, Feng Y. 2020. Discovery and mechanism of a pH-dependent dual-binding-site switch in the interaction of a pair of protein modules. Sci Adv 6:eabd7182. https://doi.org/10.1126/sciadv.abd7182.

65. Wang QS, Yu F, Huang S, Sun B, Zhang KH, Liu K, Wang ZJ, Xu CY, Wang SS, Yang LF, Pan QY, Li L, Zhou H, Cui Y, Xu Q, Earnest T, He JH. 2015. The macromolecular crystallography beamline of SSRF. Nucl Sci Tech 26:12–17. https://doi.org/10.13538/j.1001-8042/nst.26.010102.

66. Wang QS, Zhang KH, Cui Y, Wang ZJ, Pan QY, Liu K, Sun B, Zhou H, Li MJ, Xu Q, Xu CY, Yu F, He JH. 2018. Upgrade of macromolecular crystallography beamline BL17U1 at SSRF. Nucl Sci Tech 29:68. https://doi.org/10.1007/s41365-018-0398-9.

67. Zhang WZ, Tang JC, Wang SS, Wang ZJ, Qin WM, He JH. 2019. The protein complex crystallography beamline (BL19U1) at the Shanghai Synchrotron Radiation Facility. Nucl Sci Tech 30:170. https://doi.org/10.1007/s41365-019-0683-2.

68. Kabsch W. 2010. XDS. Acta Crystallogr D Biol Crystallogr 66:125–132. https://doi.org/10.1107/S0907444909047337.

69. Ose T, Fujie T, Yao M, Watanabe N, Tanaka I. 2004. Crystal structure of the ATP-binding cassette of multisugar transporter from *Pyrococcus horikoshii* OT3. Proteins 57:635–638. https://doi.org/10.1002/prot.20206.

70. McCoy AJ, Grosse-Kunstleve RW, Adams PD, Winn MD, Storoni LC, Read RJ. 2007. Phaser crystallographic software. J Appl Crystallogr 40:658–674. https://doi.org/10.1107/S0021889807021206.

71. Adams PD, Afonine PV, Bunkoczi G, Chen VB, Davis IW, Echols N, Headd JJ, Hung LW, Kapral GJ, Grosse-Kunstleve RW, McCoy AJ, Moriarty NW, Oeffner R, Read RJ, Richardson DC, Richardson JS, Terwilliger TC, Zwart PH. 2010. PHENIX: a comprehensive Python-based system for macromolecular structure solution. Acta Crystallogr D Biol Crystallogr 66:213–221. https://doi.org/10.1107/S0907444909052925.

72. Collaborative Computational Project N. 1994. The CCP4 suite: programs for protein crystallography. Acta Crystallogr D Biol Crystallogr 50:760–763. https://doi.org/10.1107/S0907444994003112.

73. Pannu NS, Waterreus WJ, Skubak P, Sikharulidze I, Abrahams JP, de Graaff RA. 2011. Recent advances in the CRANK software suite for experimental phasing. Acta Crystallogr D Biol Crystallogr 67:331–337. https://doi.org/10.1107/S0907444910052224.

74. Emsley P, Lohkamp B, Scott WG, Cowtan K. 2010. Features and development of Coot. Acta Crystallogr D Biol Crystallogr 66:486–501. https://doi.org/10.1107/S0907444910007493.

75. Agirre J, Iglesias-Fernandez J, Rovira C, Davies GJ, Wilson KS, Cowtan KD. 2015. Privateer: software for the conformational validation of carbohydrate structures. Nat Struct Mol Biol 22:833–834. https://doi.org/10.1038/nsmb.3115.