

SCIENTIFIC REPORTS

OPEN

Biochemical properties of GH94 cellodextrin phosphorylase THA_1941 from a thermophilic eubacterium *Thermosipho africanus* TCF52B with cellobiose phosphorylase activity

Yuanyuan Wu¹, Guotao Mao¹, Haiyan Fan¹, Andong Song¹, Yi-Heng Percival Zhang^{1,2} & Hongge Chen¹

A hypothetic gene (THA_1941) encoding a putative cellobiose phosphorylase (CBP) from *Thermosipho africanus* TCF52B has very low amino acid identities (less than 12%) to all known GH94 enzymes. This gene was cloned and over-expressed in *Escherichia coli* BL21(DE3). The recombinant protein was hypothesized to be a CBP enzyme and it showed an optimum temperature of 75 °C and an optimum pH of 7.5. Beyond its CBP activity, this enzyme can use cellobiose and long-chain cellodextrins with a degree of polymerization of greater than two as a glucose acceptor, releasing phosphate from glucose 1-phosphate. The catalytic efficiencies (k_{cat}/K_m) indicated that cellotetraose and cellopentaose were the best substrates for the phosphorolytic and reverse synthetic reactions, respectively. These results suggested that this enzyme was the first enzyme having both cellodextrin and cellobiose phosphorylase activities. Because it preferred cellobiose and cellodextrins to glucose in the synthetic direction, it was categorized as a cellodextrin phosphorylase (CDP). Due to its unique ability of the reverse synthetic reaction, this enzyme could be a potential catalyst for the synthesis of various oligosaccharides. The speculative function of this CDP in the carbohydrate metabolism of *T. africanus* TCF52B was also discussed.

The β -D-glycoside phosphorylases, which act on the β -linked substrates such as cellobiose, cellodextrin, laminaribiose, N,N'-diacetylchitobiose, and cellobionic acid, are mainly classified into glycoside hydrolase family (GH) 94¹. Among GH94 enzymes, cellobiose phosphorylase (CBP, EC 2.4.1.20) is the biggest member catalyzing the reversible phosphorolysis of cellobiose into α -D-glucose-1-phosphate (G-1-P) and D-glucose (Eq. 1). This enzyme exists in anaerobic bacteria, e.g., *Clostridium thermocellum*², *Ruminococcus albus*³, *Cellvibrio gilvus*⁴, *Clostridium stercorarium*⁵, *Ruminococcus flavefaciens*⁶, *Cellulomonas uda*⁷, *Formis annosus*⁸, *Thermotoga neapolitana*⁹, and *Thermotoga maritima*¹⁰. Cellobiose, an intermediate of enzymatic cellulose degradation, can be assimilated across the cellular membrane followed by intracellular phosphorolysis with CBP or by hydrolysis with β -glucosidase (β G). Bioenergetic studies demonstrated that intracellular phosphorolytic cleavage of cellobiose (or cellodextrins) was preferred to hydrolytic cleavage in cellulolytic anaerobic bacteria^{11,12} because one ATP is saved per beta-1,4-glycosidic bond cleavage of glucose units of cellobiose and cellodextrins.



where G-1-P denotes glucose-1-phosphate, and P_i denotes inorganic phosphate.

¹College of Life Sciences, Henan Agricultural University, 95 Wenhua Road, Zhengzhou, 450002, China. ²Biological Systems Engineering Department, Virginia Tech, 304 Seitz Hall, Blacksburg, Virginia, 24061, USA. Correspondence and requests for materials should be addressed to H.C. (email: honggeyz@163.com)

Cellodextrin phosphorylase (CDP, EC 2.4.1.49) also belongs to the GH94 family, catalyzes the phosphorolysis of long-chain cellodextrins with a degree of polymerization (DP) of more than two and the synthesis of cellodextrin from short cellodextrins with a DP of at least two (Eq. 2).



where G_n denotes a β -glucan oligomer of length n ($n \geq 2$), and G_{n+1} denotes a β -glucan oligomer of length $n + 1$. Although CBP and CDP belong to the same glycoside hydrolase family, they have different substrate specificities^{13,14}.

Laminaribiose phosphorylase (LBP, EC 2.4.1.31), N,N'-diacetylchitobiose phosphorylase (briefly noted chitobiose phosphorylase, ChBP, EC 2.4.1.280), and cellobionic acid phosphorylase (CBAP, EC 2.4.1.321) are relatively new members in GH94 family compared to CBP and CDP, catalyzing reversible phosphorolysis of their corresponding β -glucosidic substrates.

GH94 phosphorylases are involved in the intracellular catabolism of β -D-glycosides, enabling direct production of phosphorylated glucose without ATP consumption. Due to this energetic benefit, these enzymes could find broad uses in metabolic engineering, biocatalysis and *in vitro* synthetic biology. For example, CBPs have been introduced into non-cellulolytic ethanol-producing microbes, such as *Saccharomyces cerevisiae* and *Escherichia coli*, for intracellular phosphorolysis of cellobiose for enhanced bioenergetic levels^{15–19}. Compared to cells employing a β G for intracellular cellobiose hydrolysis, Ha *et al.*¹⁶ demonstrate that CBP-containing *S. cerevisiae* had more biomass synthesis and higher ethanol yields under strict anaerobic conditions and under acetate stress conditions. Shin *et al.*¹⁷ further shows that CBP-containing *E. coli* cells tolerated acetate more than BG-containing cells under both anaerobic and aerobic conditions. Another promising application of CBP is *in vitro* cascade biocatalysis. By utilization of its phosphorolytic ability, this enzyme can produce high-yield hydrogen from cellobiose and cellodextrin²⁰ and from oligoxylosaccharides²¹. Also, by combining *C. thermocellum* CBP along with potato alpha-glucan phosphorylase, a fraction of cellulose can be converted to artificial starch²². By utilization of their reverse synthetic ability, CBP and CDP can be used to synthesize diverse disaccharides, oligosaccharides, and glycolipids^{23–25}. Great potentials of CBP and CDP motivate us to discover their new species with more desirable properties, for example, high thermostability, high activity, broad substrate specificity and so on.

In this study, the open reading frame (THA_1941) encoding a putative CBP from *Thermosiphon africanus* TCF52B was cloned and overexpressed in *E. coli* BL21(DE3). Its basic biochemical properties were characterized. This enzyme was the first enzyme with both cellodextrin and cellobiose phosphorylase activities.

Results

Discovery for a putative CBP from *T. africanus* TCF52B. By considering potential industrial applications and easy thermophilic protein purification expressed in a mesophilic host *E. coli*, putative thermostable CBP genes were searched from exploring genomic DNA database of thermophilic microorganisms. *T. africanus* TCF52B, which was isolated from a high-temperature oil reservoir in the North Sea with an optimal growth temperature of 70 °C, was sequenced and annotated in 2009²⁶. Although the locus THA_1941 (protein_id ACJ76363.1) was annotated as a hypothetical protein in both KEGG and NCBI, there existed a region COG3459 in the sequence annotated as “cellobiose phosphorylase [carbohydrate transport and metabolism]”. So we speculated THA_1941 as a putative CBP. Also, signal peptide analysis predicted that the protein without a signal peptide was located intracellularly, in consistent with its intracellular function. However, sequence alignment with ClustalW showed that this putative 1,019 amino acid enzyme had a very low sequence identity to all characterized CBPs, as well as to characterized CDPs, LBPs, ChBPs, and CBAPs. Among them, it has the highest identity of 11.9% with *C. thermocellum* CBP (GenBank No.: ABN51514.1).

Structure basis for the putative CBP's enzymatic function. A phylogenetic analysis was conducted to examine the genetic relationship of THA_1941, the putative CBP with characterized GH94 phosphorylases, including CBPs, CDPs, LBPs, ChBPs, and CBAPs. The phylogenetic tree was generated by Neighbour-Joining Method based on amino acid sequences (Fig. 1). It was clear that all CBPs fell into one cluster, and ChBPs, CBAPs, and LBPs were all in their own cluster, while CDPs could go into the same cluster at the larger genetic distance with CBPs, ChBPs and CBAPs. Yet, THA_1941 separately located on an earliest diverging clade, indicating that the putative CBP had the greatest genetic distance with all others and therefore could not be categorized into any of the known GH94 enzymes.

In order to find the conserved domains responsible for the putative CBP's possible catalytic functions, those structure-solved enzymes in Fig. 1 were picked out for amino acid sequence alignment (i.e., *C. gilvus* CBP (CgCBP)²⁷, *C. uda* CBP (CuCBP)²⁸, *C. thermocellum* CBP (CtCBP)²⁹, *V. proteolyticus* ChBP (VpChBP)³⁰, and *S. degradans* CBAP (SdCBAP)³¹) (Fig. 2). The results showed that the secondary structural components of THA_1941 were almost the same as all the five enzymes despite low amino acid sequence identities. It can be divided into four distinct domains: N-terminal domain (1–287), a helical linker (288–319), (α/α)₆-barrel domain (328–931), and C-terminal domain (320–327 and 932–1,019). The N-terminal β -sandwich domain was designated as *Glyco_trans_36* (PF06165) in Pfam database (later reclassified into GH94 family). In this domain THA_1941 had the conserved residues, such as Asn140, Phe230, Pro262 and Gly283, corresponding to Asn136, Phe217, Pro245 and Gly268 in CgCBP, respectively. The (α/α)₆-barrel domain of THA_1941 was larger than that of the modular enzymes, owing to some insertions. Although low overall sequence identity in the (α/α)₆-barrel domains, the potential catalytic residue Asp625 for THA_1941 (corresponding to Asp490 for CgCBP) near the center of (α/α)₆-barrel was found based on the conserved sites in the modular enzymes. The potential phosphate binding sites in THA_1941 such as Arg373, His847, Ser913 and Gly914 (corresponding to Arg351, His666, Thr731 and Gly732 for CgCBP, respectively) and the potential sugar binding sites such as Arg393, Asp394,

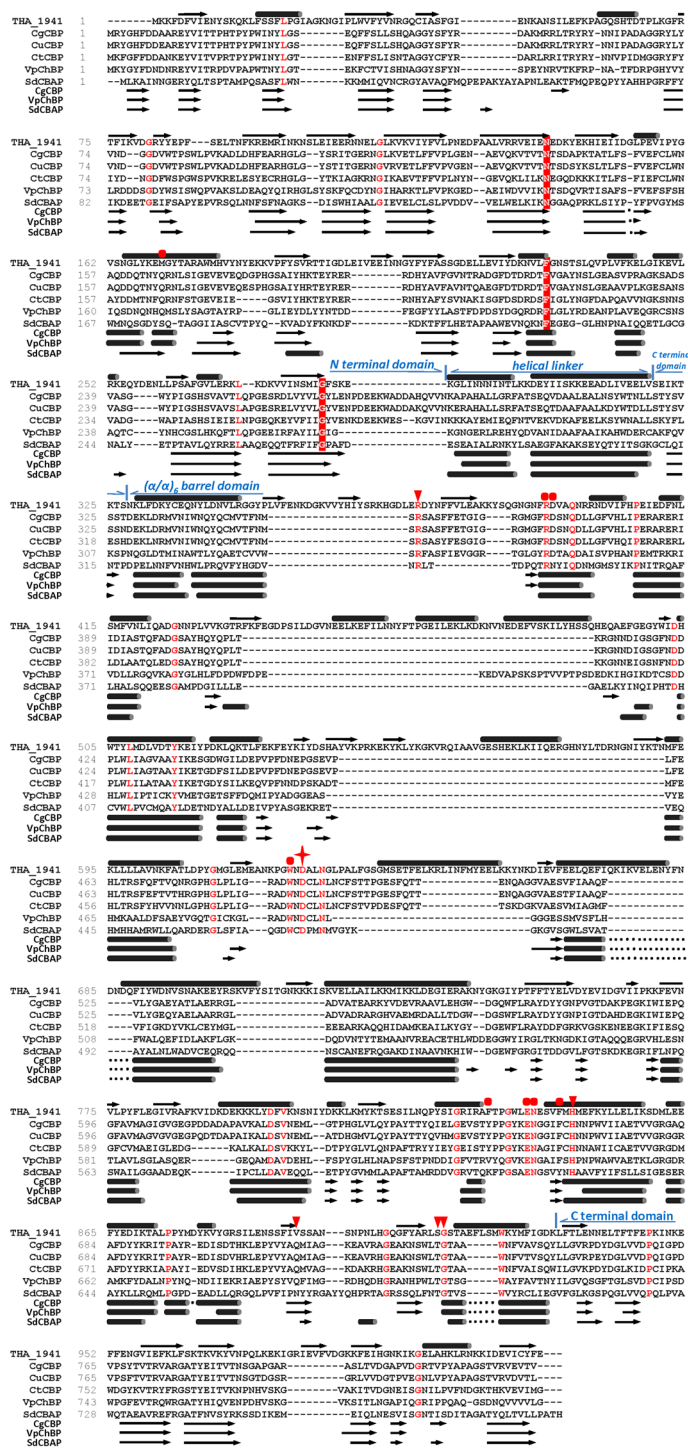


Figure 2. Amino acid sequence alignment of THA_1941 with structure-solved GH94 enzymes. Sequence alignment of THA_1941 with structure-solved CBPs from *Cellvibrio gilvus* (CgCBP, BAA28631.1), *Cellulomonas uda* (CuCBP, AAQ20920.1), and *Clostridium thermocellum* (CtCBP, AAL67138.1), ChBP from *Vibrio proteolyticus* (VpChBP, BAC87867.1), and CBAP from *Saccharophagus degradans* (SdCBAP, ABD80168.1) was performed using the program ClustalX²⁴ and formatted with BioEdit. Secondary structure was predicted with PSIPRED sever⁴³ and the secondary structural elements are marked on the top. Secondary structure elements from CgCBP, VpChBP and SdCBAP are shown below the sequence alignment (the secondary structure of CgCBP represents that of CuCBP and CtCBP because of their high similarity). Arrows and columns represent the β strands and α helices, respectively. Conserved residues in the N terminal sandwich domain (*Glyco_trans_36*, PF06165, reclassified into GH94) are highlighted with red rectangles. The catalytic residues are marked with red stars, while the phosphate binding sites and the sugar binding sites are marked with triangles and dots, respectively.

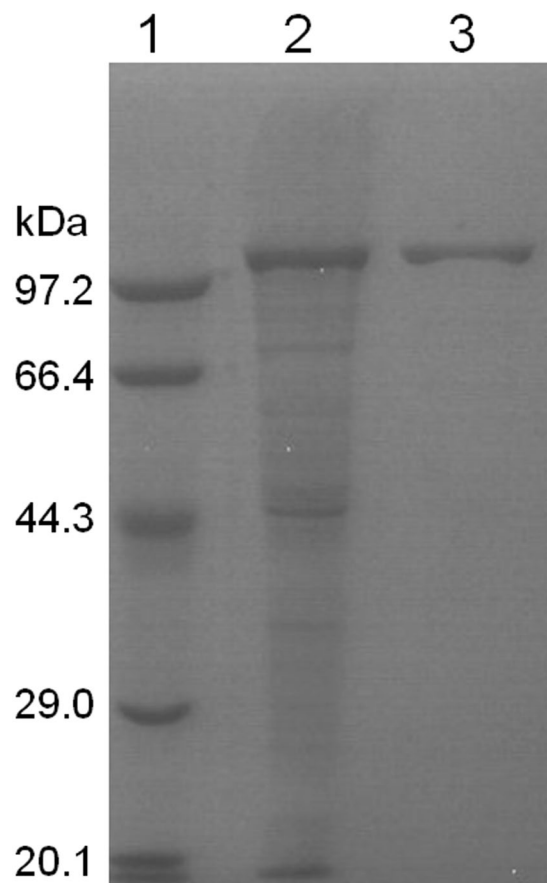


Figure 3. SDS-PAGE analysis of recombinant putative CBP. Lane 1, Molecular mass markers; Lane 2, Crude extract of recombinant *E. coli*; Lane 3, Purified THA_1941 protein.

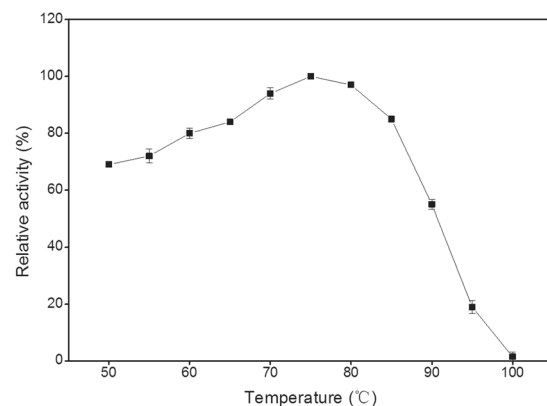


Figure 4. Effect of temperature on the activity of TaCBP.

Kinetic parameters. In the synthetic reaction, initial reaction rates were measured on varying concentrations of D-glucose, D-xylose, G-1-P, and cellodextrins and the kinetic parameters are summarized in Table 3. TaCBP showed a much higher catalytic efficiency on glucose (30.3 s^{-1}) than on D-xylose (1.37 s^{-1}). The K_m values for cellodextrins decreased with an increase in substrate chain length, suggesting TaCBP's higher affinity on longer-chain substrates. The k_{cat} for cellotetraose was the highest among the cellodextrins tested, yet cellopentaose was the best substrate in terms of its k_{cat}/K_m .

The kinetic parameters to a series of cellodextrins for the phosphorylytic reaction were determined in the presence of 100 mM inorganic phosphate (Table 4). Similar with those for the synthetic reaction, the K_m values for cellodextrins decreased with an increase in substrate chain length. Cellotetraose was the best substrate for

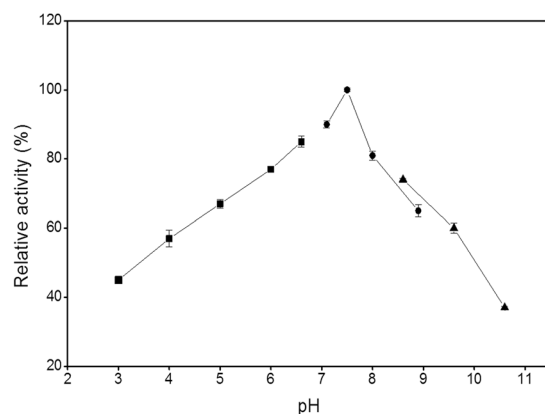


Figure 5. Effect of pH on the activity of TaCBP.

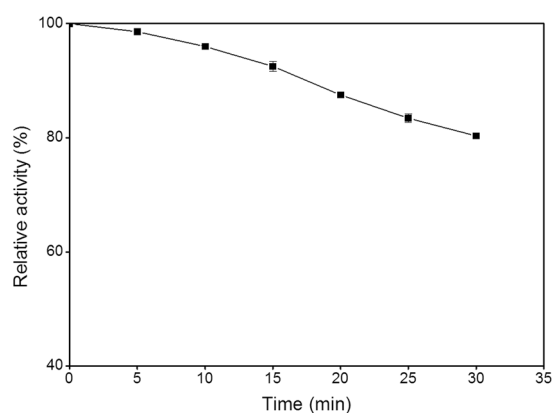


Figure 6. The stability of TaCBP incubated at 75 °C.

Substrate	Rate ($\mu\text{mol}/\text{min}/\text{mg}$)
D-Glucose	1.86 ± 0.06
D-Xylose	0.35 ± 0.01
D-Galactose	0.27 ± 0.02
D-Mannose	0.38 ± 0.01
D-Arabinose	0.22 ± 0.01
N-Acetyl-D-glucosamine	0.22 ± 0.01
D-Fructose	0.22 ± 0.01
D-Glucosamine	1.08 ± 0.04
2-Deoxy-D-glucose	0.85 ± 0.01
D-Maltose	0.16 ± 0.01
D-Cellobiose	58.61 ± 0.35
D-Gentiobiose	49.98 ± 0.28

Table 1. Substrate specificities of TaCBP in the synthetic reaction. Concentration of each substrate used was 10 mM.

phosphorolysis in terms of the value of k_{cat}/K_m , consistent with the fact that TaCBP had the highest phosphorolytic activity towards cellotetraose.

Discussion

We validated that the hypothetical protein THA_1941 from *T. africanus* TCF52B was an enzyme having both CDP and CBP activities. Comparing its two activities, though in phosphorolytic reaction they were in the similar level, in the synthetic reaction it showed remarkably higher specificity to cellodextrins than to D-glucose and D-xylose, indicating its higher CDP activity than CBP activity. Therefore this enzyme was categorized as cellodextrin phosphorylase (EC 2.4.1.49) and designated as TaCDP despite of the initial name TaCBP.

Substrate	Synthetic direction		Phosphorolytic direction	
	Rate ($\mu\text{mol}/\text{min}/\text{mg}$)	Relative rate (%)	Rate ($\mu\text{mol}/\text{min}/\text{mg}$)	Relative rate (%)
Cellobiose	42.2 \pm 0.20	38.6	0.069 \pm 0.01	71.1
Cellotriose	98.1 \pm 0.60	89.8	0.071 \pm 0.01	73.2
Cellotetraose	109.3 \pm 0.73	100.0	0.097 \pm 0.01	100.0
Cellopentaose	69.2 \pm 0.50	63.3	0.085 \pm 0.02	87.6

Table 2. Substrate specificities of TaCBP on cellodextrins in the synthetic and phosphorolytic reactions. Concentration of each cellodextrin used was 5 mM.

Substrate	K_m (mM)	k_{cat} (s^{-1})	k_{cat}/K_m ($\text{s}^{-1}\text{mM}^{-1}$)
D-Glucose ^a	14.3 \pm 1.53	30.3 \pm 2.20	2.12
D-Xylose ^b	25.7 \pm 1.30	1.37 \pm 0.13	0.05
G-1-P ^a	1.57 \pm 0.20	7.61 \pm 0.66	4.85
Cellobiose ^c	7.39 \pm 0.75	361 \pm 9.80	48.87
Cellotriose ^c	4.05 \pm 0.31	603 \pm 21.2	148.90
Cellotetraose ^c	3.7 \pm 0.25	612 \pm 18.8	165.46
Cellopentaose ^c	2.59 \pm 0.22	527 \pm 15.5	203.50

Table 3. Kinetic parameters for the synthetic reaction catalyzed by TaCBP. Concentrations used: ^a0.5–10 mM; ^b5–100 mM; ^c0.5–5 mM.

Substrate	K_m (mM)	k_{cat} (s^{-1})	k_{cat}/K_m ($\text{s}^{-1}\text{mM}^{-1}$)
Cellobiose	0.120 \pm 0.03	0.23 \pm 0.06	1.92
Cellotriose	0.094 \pm 0.01	0.28 \pm 0.05	2.98
Cellotetraose	0.062 \pm 0.015	0.49 \pm 0.10	7.91
Cellopentaose	0.060 \pm 0.02	0.34 \pm 0.11	5.67

Table 4. Kinetic parameters for the phosphorolytic reaction of TaCBP. Concentrations used: 0.02–1.0 mM.

Compared with properties of the known CBPs and CDPs (Table 5), TaCDP had the highest optimal temperature (75 °C) among known CDPs, representing its best thermostability. As for substrate specificity, no known CDP was reported to phosphorolyze cellobiose, neither to use monosaccharides such as D-glucose, D-xylose and D-glucose-derivatives as glucosyl acceptors in its reverse synthetic reaction. Different from known CDPs, TaCDP showed a wide range of substrate specificity in both reaction directions, making it a unique bifunctional enzyme with both CDP and CBP activities. It was noteworthy that TaCDP had much lower catalytic efficiency in the phosphorolytic reaction than in the reverse synthetic reaction, indicating that it had much greater preference for the synthetic reaction than for the phosphorolytic reaction. Moreover TaCDP's catalytic efficiency in the synthetic reaction was shown to be the highest among the known CDPs as comparing their highest k_{cat} values for cellodextrins in synthetic reaction: the highest k_{cat} (612 s^{-1} , 60 °C) of TaCDP was 13-fold and 38-fold higher than that of RaCDP (47.1 s^{-1} , 37 °C)³³ and CtCDP (16.2 s^{-1} , 37 °C)¹³, respectively.

Though TaCDP's enzymatic functions were characterized *in vitro*, it was difficult to investigate its biological function in *T. africanus* TCF52B without genetic modification tools. As suggested by Taylor II *et al.*³⁴, both CBPs and CDPs were included in the class of "cellulase accessory enzymes", which acted on cellulose oligosaccharides, the products of cellulases *i.e.* β -1,4-endoglucanases and/or cellobiohydrolases. By searching the genomic sequence of *T. africanus* TCF52B, only 2 ORFs (THA_83 and THA_328) were found to be possible endoglucanases related to cellulose degradation, yet both of them have no putative signal peptides. Hence, we presumed that *T. africanus* TCF52B was unlikely to be a lignocellulose-degrading bacterium. Considering the fact that TaCDP's phosphorolysis efficiency for cellodextrins was much lower than its synthesis efficiency, we then speculated that TaCDP might be responsible for the cellodextrins formation other than for cellodextrins degradation, where cellodextrins formed by TaCDP might be intracellular energy reserve material, like poly- β -hydroxybutyrate accumulated in many bacteria³⁵ and glycogen accumulated by *C. cellulolyticum*³⁶. A possible carbohydrate metabolism pathway was constructed to illustrate the role of involved TaCDP (Fig. 7). Here, when *T. africanus* TCF52B grows in rich nutrition conditions, cytoplasmic G-6-P will be converted to G-1-P by phosphoglucomutase (locus THA_RS09865 and THA_1027) and G-1-P then transfer its glucose unit to D-glucose or other monosaccharides by TaCDP, forming gradually longer cellodextrin. When this bacterium grows in poor nutrition conditions, accumulated cellodextrin will be hydrolyzed by intracellular β -glucosidase (locus THA_1926 and THA_1942). This scheme for synthesizing energy reserves uses ATP-derived energy more efficiently than that of synthesizing

Strains	Enzyme	MW (kDa)	pH optima	Temperature optima	Thermostability	Substrate specificity						Reference
						Synthetic reaction			Phosphorolytic reaction			
						Substrate	K_m (mM)	k_{cat} (s^{-1})	Substrate	K_m (mM)	K_{cat} (s^{-1})	
<i>Clostridium thermocellum</i> ATCC 27405	CBP	150	6.0–8.0	60 °C	<60 °C	α -D-Glucose-1-phosphate/D-Mannose/	8/170/	—/—/				32
						D-Xylose/D-Glucosamine/	40/10/	—/—/				
						2-Deoxy-D-glucose	250	—				
<i>Ruminococcus albus</i> NE1	CBP	187	6.2	50 °C	<40 °C	D-Glucose/D-Mannose/	1.46/23.5/	92.3/2.83/	Cellobiose	1.27	110	3
						2-Deoxy-D-glucose/D-Glucosamine/	60.2/13.3/	21.5/9.56/	Pi	0.427	—	
						D-Xylose/6-Deoxy-D-glucose/	25.5/9.8/	16.5/124/				
						1,5-Anhydro-D-glucitol/Gentiobiose	9.5/10.8	2.79/4.01				
<i>Clostridium stercorarium</i>	CBP	93	6.0–7.0	65 °C	50% residual activity after 20 h at 65 °C			Cellobiose	6.2	1.1	4	
<i>Thermotoga neapolitana</i>	CBP	93.5	5.0	85 °C	remained active for 2 h at 85 °C			Cellobiose	1.42	26.3	9	
<i>Clostridium thermocellum</i>	CBP	—	6.5(D-Xylose as acceptor)	—	—	6-Deoxy-glucose/D-Xylose/	9.2/35/	—/—/	Cellobiose	7.3	—	2
						2-Deoxy-D-glucose/D-Glucosamine/	73/9.5/	—/—/	Pi	2.9	—	
						D-Mannose/D-Arabinose/	85/240/	—/—/	D-Glucose	—	—	
						L-Fucose/ α -D-Glucose-1-P	160/2.1	—/—				
<i>Thermotoga maritima</i> MSB8	CBP	150	6.2	80 °C	<70 °C	D-Glucose/D-Mannose/	0.69/67/	8/4.4/	Cellobiose	0.29	5.4	10
						D-Glucosamine/2-Deoxy-D-glucose/	5.7/47/	5.2/16/	Pi	0.15	—	
						6-Deoxy-D-glucose/D-Xylose/	4.1/14/	17/40/				
						Methyl- β -D-glucoside	135	6.5				
<i>Cellulomonas uda</i>	CBP	—	—	—	—	D-Glucose/ β -D-Glucose/	2.26/4.7/	44.2/47.6/	D-Cellobiose	4.3	29.1	7
						2-Deoxy-D-glucose/	24.3/	11.6/	Pi	1.2	—	
						2-Deoxy-2-fluoro-D-glucose/	17.2/	8/				
						D-Mannose/D-Glucosamine/	27.3/10.8/	6.6/5.6/				
						3-Deoxy-D-glucose/	4/	0.1/				
						3-Deoxy-3-fluoro-D-glucose/	ND/	ND/				
						6-Deoxy-D-glucose/	10.2/	55.3/				
						6-Deoxy-6-fluoro-D-glucose/	7.4/	44.6/				
<i>Cellvibrio gilvus</i>	CBP	280	7.6	—	complete inactivation after 10 min at 60 °C			Cellobiose	1.25	—	4	
								Pi	0.77	—		
<i>Prevotella ruminicola</i> (crude extracts)	CBP	—	—	—	—			Cellobiose	0.17	—	41	
<i>Clostridium stercorarium</i>	CDP	91	6.0–7.0	65 °C	$t_{1/2(60^\circ C)} = 30$ h			Cellotriose	0.04	2.7	5	
								Cellotetraose	0.05	6.9		
								Cellopentaose	0.17	6.6		
								Cellohexaose	0.19	6.1		
<i>Ruminococcus albus</i>	CDP	272	6.0	50 °C	<40 °C	Cellobiose/ Cellotriose/	13.2/5.01/	47.1/43.7/	Cellobiose	ND	ND	33

Continued

Strains	Enzyme	MW (kDa)	pH optima	Temperature optima	Thermostability	Substrate specificity						Reference
						Synthetic reaction			Phosphorolytic reaction			
						Substrate	K_m (mM)	k_{cat} (s^{-1})	Substrate	K_m (mM)	k_{cat} (s^{-1})	
						Cellopentaose/ Cellohexaose/	2.73/3.22/	28.9/18.2/	Cellotriose	6.04	76.2	
						Sophorose/ Laminaribiose/	343/119/	14.2/33.3/	Cellotetraose	4.16	92.8	
						Xylobiose/ Mannobiose/	50.9/65/	15.4/41.3/	Cellopentaose	2.41	83.8	
						Cellobiitol// Cellotetraose	73.7//3.97	3.26//37.8	Cellohexaose	1.04	55.8	
<i>Clostridium thermocellum</i> YM4	CDP	210	7.5	60 °C	<60 °C	Glucose/Cellobiose/	ND/0.89/	ND/10.1/	Cellobiose	ND	ND	13
						Cellotriose/ Cellotetraose	1.75/2.65	13.9/16.2	Cellotriose	0.81	4	
									Cellotetraose	0.82	3.2	
<i>Thermosiphon africanus</i> strain TCF52B	CDP/ CBP	120	7.5	75 °C	80% residual activity after 30 min at 75 °C	D-Glucose/D-Xylose/	14.3/25.7/	30.3/1.37/	Cellobiose	0.120	0.23	This work
						α -D-Glucose-1-P/ Cellobiose/	1.57/7.39/	7.61/361/	Cellotriose	0.094	0.28	
						Cellotriose/	4.05/	603/	Cellotetraose	0.062	0.49	
						Cellotetraose/	3.7/	612/	Cellopentaose	0.060	0.34	
						Cellopentaose	2.59	527				

Table 5. Properties of characterized CBPs and CDPs from different strains. “–”: not tested; ND: not detected.

glycogen, as only one ATP is consumed per glucose unit addition while two ATPs in glycogen synthesis. Actually, we did not find any putative UDP/ADP-glucose pyrophosphorylases in *T. africanus* TCF52B genome, enzymes responsible for activated glucosyl donor formation for the glycogen synthesis, which may increase the possibility of our assumption on TaCDP's function.

Phosphorolytic enzymes like CBPs and CDPs had advantages in oligosaccharides synthesis over chemical catalysts because they have both stereoselectivity and regioselectivity^{37,38}. TaCDP had high synthetic activities towards monosaccharides, disaccharides, and long-chain oligosaccharides, making it a valuable biocatalyst for cost-effective enzymatic synthesis of various oligosaccharides.

Methods

Cloning and expression of *T. africanus* THA_1941. The DNA sequence of THA_1941 can be found in KEGG database. The gene was amplified using the polymerase chain reaction (PCR) from the *T. africanus* TCF52B genomic DNA using 5'-CCTAG CTAGC ATGAA AAAAT TTGAC TTTGT G-3' and 5'-CCGCT CGAGT TCAAA ATAAC ATATA ACTTC GTC-3' as the forward and reverse primers, respectively (*Nhe*I and *Xho*I restriction sites underlined, respectively). The PCR product was digested with *Nhe*I and *Xho*I prior to insertion into pET21a(+). The ligation product was transformed into *E. coli* DH5 α competent cells and the plasmid was verified by DNA sequencing service provided at Genewiz Inc. China. The plasmid was then transformed into *E. coli* BL21(DE3) competent cells. A single colony was picked to grow in Luria-Bertani (LB) medium supplemented with 50 μ g/ml ampicillin. The culture was grown at 37 °C and 220 rpm until an absorbency of 600 nm reached 0.6–0.8. Expression of the protein was induced with the addition of isopropyl β -D-1-thiogalactopyranoside (IPTG) to a final concentration of 1 mM and the culture temperature was decreased to 25 °C for six hours. The cells were harvested by centrifugation at 3,800 \times g for 10 min at 4 °C.

Enzyme purification. The cell pellets were re-suspended in 20 mM Tris-HCl buffer (pH 7.0) with 500 mM NaCl (pH8.0) followed by sonication. The cell lysate was centrifuged at 10,000 \times g at 4 °C for 20 min and the supernatant was applied to a nickel-charged resin column (Takara, Shiga, Japan). The column was washed with gradual washing containing 20–50–200–500 mM imidazole. The fraction eluted with 200 mM imidazole was collected and dialyzed against 20 mM Tris-HCl buffer (pH 7.4). After centrifugation the enzyme solution was loaded onto a pre-equilibrated HiTrap Q HP column (GE Healthcare) and the elution was performed with a linear gradient of NaCl from 100 to 500 mM. Under these conditions, most target protein was in the fractions of 200 mM NaCl eluate. The purity of each fraction was assessed by SDS-PAGE³⁹. Only those fractions showing a single band were pooled and the concentration of the resulting purified sample was determined to be 1.40 mg/ml using the method of Bradford with bovine serum albumin as a standard⁴⁰.

Enzyme assays. In the synthetic reaction, CBP activity was assayed by measuring the amount of P_i liberated from G-1-P³². D-Glucose was routinely used as an acceptor of glucosyl, unless otherwise noted. A reaction mixture of 200 μ l contained 50 μ l appropriately diluted enzyme, 40 mM G-1-P, 1 mM MgCl₂, 10 mM dithiothreitol (DTT), 20 mM D-glucose, and 50 mM Tris-HCl buffer (pH 7.5). The mixture was incubated for 15 min at 60 °C and then the reaction was terminated by the addition of 2 ml of the molybdate reagent containing 15 mM ammonium molybdate, 100 mM zinc acetate [pH 5.0], and 500 μ l of ascorbic acid reagent (10% [wt/vol], pH 5.0) was

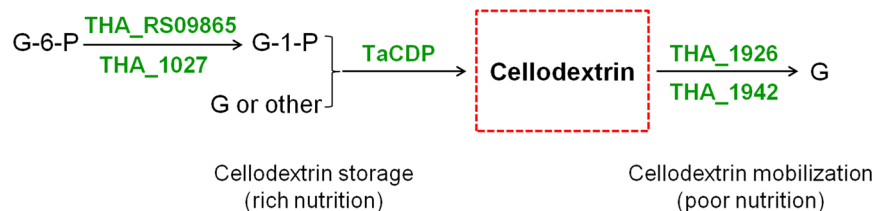


Figure 7. Presumed carbohydrate metabolism pathway with TaCDP involved in *Thermosiphon africanus* TCF52B.

then added to the mixture. This mixture was incubated at 30 °C for 15 min, and the absorbance was measured at 850 nm. One unit of CBP in synthetic reaction was defined as the amount of enzyme that produced 1 μ mole of phosphate per min.

The phosphorolytic activity of TaCBP was assayed by measuring the formation of G-1-P from cellobiose (or cellobioses when stated). A reaction mixture of 40 μ l containing 5 mM cellobiose or cellobioses, 100 mM sodium phosphate buffer (pH 7.5), and appropriate amount of enzyme was incubated for 15 min at 60 °C. The reaction was stopped by boiling for 10 min, and the amount of G-1-P produced was determined by a coupled enzyme assay measuring the appearance of NADPH at 340 nm. The assay mixture contained phosphoglucomutase (4.0 U/ml), glucose-6-phosphate dehydrogenase (2.0 U/ml), where the two enzymes were purchased from Sigma, 3 mM NADP⁺, and 5 μ M glucose 1,6-bisphosphate (Sigma) in 80 mM triethanolamine buffer (pH 7.5) containing 4.4 mM MgCl₂. One unit of phosphorolytic activity was defined as the release of one μ mole of G-1-P per min.

Optimum pH, optimum temperature, and thermal stability. Determinations in this section were all based on the synthetic reaction. The optimum pH was investigated by measuring the enzyme activity as described above at various pH values (citric acid-sodium citrate buffers for pH 3.0–6.6, Tris-HCl buffers for pH 7.1–8.9, and glycine-NaOH buffers for pH 8.6–10.6). The optimum temperature was measured over a temperature range of 50–100 °C at pH 7.5. To evaluate thermostability, the enzyme (0.092 mg/ml) was incubated at 75 °C and pH 7.5 for different time periods (5–30 min) and the residual enzyme activity was measured. All assays were performed in triplicates.

Substrate specificity. The substrate specificity for the synthetic direction was investigated by measuring initial velocities (μ mol/min/mg) for 10 mM various sugars with 40 mM G-1-P at 60 °C. The final concentration of enzyme used for each sugar was adjusted properly to make sure the P_i amount released was suitable for accurate determination. Thus the final enzyme concentrations were 0.67 μ g/ml and 2 μ g/ml for D-cellobiose and D-gentiobiose, respectively, while 20 μ g/ml of enzyme was used for D-glucose, D-maltose, and 2-deoxy-D-glucose, and 100 μ g/ml for all other sugars.

The substrate specificity for the phosphorolytic direction was investigated by measuring the initial velocities for 5 mM different cellobioses at 60 °C with 87.5 μ g/ml of the enzyme. The other conditions were the same as for the phosphorolytic activity assay mentioned earlier.

Kinetic analyses. To determine the kinetic parameters for the synthetic reaction, 200 μ l of reaction system was used and the final concentrations of the enzyme used for D-glucose, D-xylose, and cellobioses were 20, 100, and 0.67 μ g/ml, respectively. The initial reaction rates at varying concentrations of D-glucose (0.5–10 mM), D-xylose (5–100 mM), or cellobioses (0.5–5 mM) with 40 mM of a fixed concentration of G-1-P were determined. While determining the kinetic parameters to G-1-P, the reaction mixture contained 20 mM of a fixed concentration of D-glucose and G-1-P (0.5–10 mM) with 20 μ g/ml of the enzyme. To determine the kinetic parameters for the phosphorolytic reaction to cellobioses, 40 μ l of reaction system and 100 μ g/ml of the enzyme concentration were used. The initial reaction rates at varying concentrations of cellobioses with 100 mM sodium phosphate buffer (pH 7.5) were determined. Both the synthetic reaction and the phosphorolytic reaction were incubated at 60 °C for 10 min. Each result was an average of at least three repetitions. K_m and k_{cat} values were calculated based on the Hanes-Woolf plots.

References

- Nakai, H., Kitaoka, M., Svensson, B. & Ohtsubo, K. Recent development of phosphorylases possessing large potential for oligosaccharide synthesis. *Curr. Opin. Chem. Biol.* **17**, 301–309 (2013).
- Alexander, J. K. Purification and specificity of cellobiose phosphorylase from *Clostridium thermocellum*. *J. Biol. Chem.* **243**, 2899–2904 (1968).
- Hamura, K. *et al.* Enzymatic characteristics of cellobiose phosphorylase from *Ruminococcus albus* NE1 and kinetic mechanism of unusual substrate inhibition in reverse phosphorolysis. *Biosci. Biotechnol. Biochem.* **76**, 812–818 (2012).
- Sasaki, T., Tanaka, T., Nakagawa, S. & Kainuma, K. Purification and properties of *Cellvibrio gilvus* cellobiose phosphorylase. *Biochem. J.* **209**, 803–807 (1983).
- Reichenbecher, M., Lottspeich, F. & Bronnenmeier, K. Purification and properties of a cellobiose phosphorylase (CepA) and a cellobiosin phosphorylase (CepB) from the cellulolytic thermophile. *Clostridium Stercorarium*. *Eur. J. Biochem.* **247**, 262–267 (1997).
- Ayers, W. A. Phosphorylation of cellobiose and glucose by *Ruminococcus flavefaciens*. *J. Bacteriol.* **76**, 515–517 (1958).

7. Nidetzky, B., Eis, C. & Albert, M. Role of non-covalent enzyme-substrate interactions in the reaction catalysed by cellobiose phosphorylase from *Cellulomonas uda*. *Biochem. J.* **351**, 649–659 (2000).
8. Hüttermann, A. & Volger, C. Cellobiose phosphorylase in *Fomes annosus*. *Nat. New Biol.* **245**, 64 (1973).
9. Yernool, D. A., McCarthy, J. K., Eveleigh, D. E. & Bok, J. D. Cloning and characterization of the glucooligosaccharide catabolic pathway β -glucan glucohydrolase and cellobiose phosphorylase in the marine hyperthermophile *Thermotoga neapolitana*. *J. Bacteriol.* **182**, 5172–5179 (2000).
10. Rajashekhar, E., Kitaoka, M., Kim, Y. K. & Hayashi, K. Characterization of a cellobiose phosphorylase from a hyperthermophilic eubacterium, *Thermotoga maritima* MSB8. *Biosci. Biotechnol. Biochem.* **66**, 2578–2586 (2002).
11. Lou, J., Dawson, K. A. & Strobel, H. J. Cellobiose and cellodextrin metabolism by the ruminal bacterium *Ruminococcus albus*. *Curr. Microbiol.* **35**, 221–227 (1997).
12. Zhang, Y. H. P. & Lynd, L. R. Kinetics and relative importance of phosphorolytic and hydrolytic cleavage of cellodextrins and cellobiose in cell extracts of *Clostridium thermocellum*. *Appl. Environ. Microbiol.* **70**, 1563–1569 (2004).
13. Krishnareddy, M., Kim, Y. K., Kitaoka, M., Mori, Y. & Hayashi, K. Cellodextrin phosphorylase from *Clostridium thermocellum* YM4 strain expressed in *Escherichia coli*. *J. Appl. Glycosci* **49**, 1–8 (2002).
14. De, G. M. R. *et al.* Construction of cellobiose phosphorylase variants with broadened acceptor specificity towards anomeric substituted glucosides. *Biotechnol. Bioeng.* **107**, 413–420 (2010).
15. Sadie, C. J., Rose, S. H., den Haan, R. & van Zyl, W. H. Co-expression of a cellobiose phosphorylase and lactose permease enables intracellular cellobiose utilization by *Saccharomyces cerevisiae*. *Appl. Microbiol. Biotechnol.* **90**, 1373–1380 (2011).
16. Ha, S. J. *et al.* Energetic benefits and rapid cellobiose fermentation by *Saccharomyces cerevisiae* expressing cellobiose phosphorylase and mutant cellodextrin transporters. *Metab. Eng.* **15**, 134–143 (2013).
17. Shin, H. D., Wu, J. & Chen, R. Comparative engineering of *Escherichia coli* for cellobiose utilization: Hydrolysis versus phosphorolysis. *Metab. Eng.* **24**, 9–17 (2014).
18. Sekar, R., Shin, H. D. & Chen, R. Engineering *Escherichia coli* cells for cellobiose assimilation through a phosphorolytic mechanism. *Appl. Environ. Microbiol.* **78**, 1611–1614 (2012).
19. Chomvong, K. *et al.* Overcoming inefficient cellobiose fermentation by cellobiose phosphorylase in the presence of xylose. *Biotechnol. Biofuels* **7**, 85 (2014).
20. Ye, X. *et al.* Spontaneous high-yield production of hydrogen from cellulosic materials and water catalyzed by enzyme cocktails. *ChemSusChem* **2**, 149–152 (2009).
21. Moustafa, H. M. A. *et al.* Water splitting for high-yield hydrogen production energized by biomass xylooligosaccharides catalyzed by an enzyme cocktail. *ChemCatChem* **8**, 2898–2902 (2016).
22. You, C. *et al.* Enzymatic transformation of nonfood biomass to starch. *Proc. Natl. Acad. Sci. USA* **110**, 7182–7187 (2013).
23. Ajsaka, K. & Yamamoto, Y. Control of regioselectivity in the enzymatic syntheses of oligosaccharides using glycosides. *Trends Glycosci. Glyc.* **14**, 1–11 (2002).
24. Nakai, H. *et al.* Efficient chemoenzymatic oligosaccharide synthesis by reverse phosphorolysis using cellobiose phosphorylase and cellodextrin phosphorylase from *Clostridium thermocellum*. *Biochimie* **92**, 1818–1826 (2010).
25. Tran, H. G. *et al.* Biocatalytic production of novel glycolipids with cellodextrin phosphorylase. *Bioresour. Technol.* **115**, 84–87 (2012).
26. Nesbo, C. L. *et al.* The genome of *Thermosiphon africanus* TCF52B: lateral genetic connections to the firmicutes and archaea. *J. Bacteriol.* **191**, 1974–1978 (2009).
27. Hidaka, M. *et al.* Structural dissection of the reaction mechanism of cellobiose phosphorylase. *Biochem. J.* **398**, 37–43 (2006).
28. Van Hoorebeke, A. *et al.* Crystallization and X-ray diffraction studies of cellobiose phosphorylase from *Cellulomonas uda*. *Acta Crystallogr. Sect. F Struct. Biol. Cryst. Commun.* **66**, 346–351 (2010).
29. Bianchetti, C. M., Elsen, N. L., Fox, B. G. & Phillips, G. N. Structure of cellobiose phosphorylase from *Clostridium thermocellum* in complex with phosphate. *Acta Crystallogr. Sect. F Struct. Biol. Cryst. Commun.* **67**, 1345–1349 (2011).
30. Hidaka, M. *et al.* Chitobiose phosphorylase from *Vibrio proteolyticus*, a member of glycosyl transferase family 36, has a clan GH-L-like (α/α)₆ barrel fold. *Structure* **12**, 937–47 (2004).
31. Nam, Y. W. *et al.* Crystal structure and substrate recognition of cellobionic acid phosphorylase, which plays a key role in oxidative cellulose degradation by microbes. *J. Biol. Chem.* **290**, 18281–18292 (2015).
32. Tanaka, T., Kawaguchi, T., Imada, Y., Ooi, T. & Arai, M. Purification and properties of cellobiose phosphorylase from *Clostridium thermocellum*. *J. Ferment. Bioeng.* **79**, 212–216 (1995).
33. Sawano, T., Saburi, W., Hamura, K., Matsui, H. & Mori, H. Characterization of *Ruminococcus albus* cellodextrin phosphorylase and identification of a key phenylalanine residue for acceptor specificity and affinity to the phosphate group. *FEBS J.* **280**, 4463–4473 (2013).
34. Taylor, L. E. II *et al.* Complete Cellulase System in the Marine Bacterium *Saccharophagus degradans* Strain 2-40(T). *J. Bacteriol.* **188**, 3849–3861 (2006).
35. Sharma, P. & Bajaj, B. K. Production of poly- β -hydroxybutyrate by *Bacillus cereus* PS 10 using biphasic-acid-pretreated rice straw. *Int. J. Biol. Macromol.* **79**, 704–710 (2015).
36. Desvaux, M. *Clostridium cellulolyticum*: model organism of mesophilic cellulolytic clostridia. *FEMS Microbiol. Rev.* **29**, 741–764 (2005).
37. Shintate, K., Kitaoka, M., Kim, Y.-K. & Hayashi, K. Enzymatic synthesis of a library of β -(1 \rightarrow 4) hetero- d-glucose and d-xylose-based oligosaccharides employing cellodextrin phosphorylase. *Carbohydr. Res.* **338**, 1981–1990 (2003).
38. Wohlgemuth, R., Liese, A. & Streit, W. Biocatalytic phosphorylations of metabolites: past, present, and future. *Trends Biotechnol.* doi:10.1016/j.tibtech.2017.1001.1005 (2017).
39. Laemmli, U. K. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* **227**, 680–685 (1970).
40. Bradford, M. M. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal. Biochem.* **72**, 248–254 (1976).
41. Lou, J., Dawson, K. A. & Strobel, H. J. Role of phosphorolytic cleavage in cellobiose and cellodextrin metabolism by the ruminal bacterium *Prevotella ruminicola*. *Appl. Environ. Microbiol.* **62**, 1770–1773 (1996).
42. Larkin, M. A. *et al.* Clustal W and Clustal X version 2.0. *Bioinformatics* **23**, 2947–2948 (2007).
43. Buchan, D. W., Minneci, F., Nugent, T. C., Bryson, K. & Jones, D. T. Scalable web services for the PSIPRED protein analysis workbench. *Nucleic Acids Res* **41**, W349–W357 (2013).

Acknowledgements

The authors thank Dr. Chun You in Tianjin Institute of Industrial Biotechnology, Chinese Academy of Sciences for his help in the verification of enzyme properties. This work was supported by the National Natural Science Foundation of China (No. 31571775) to H.C.

Author Contributions

Y.W., G.M. and H.F. performed the experiments and analyzed the data. A.S., Y.Z. and H.C. designed the study and wrote the manuscript. All authors contributed to data interpretation and manuscript preparation.

Additional Information

Competing Interests: The authors declare that they have no competing interests.

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2017