J. Dairy Sci. 103:7730–7741
https://doi.org/10.3168/jds.2020-18153
© 2020 American Dairy Science Association®. Published by Elsevier Inc. and Fass Inc. All rights reserved.

Computer-aided search for a cold-active cellobiose 2-epimerase

Qiuming Chen,[1] Yaqin Xiao,[1] Wenli Zhang,[1] Timo Stressler,[2] Lutz Fischer,[2] Bo Jiang,[1,3] and Wanmeng Mu[1,3]*
[1]State Key Laboratory of Food Science and Technology, Jiangnan University, Wuxi, Jiangsu, 214122, China
[2]University of Hohenheim, Institute of Food Science and Biotechnology, Department of Biotechnology and Enzyme Science, 70599 Stuttgart, Germany
[3]International Joint Laboratory on Food Safety, Jiangnan University, Wuxi 214122, China

ABSTRACT

Cellobiose 2-epimerase (CE) is a promising industrial enzyme that can catalyze bioconversion of lactose to its high-value derivatives, namely epilactose and lactulose. A need exists in the dairy industry to catalyze lactose bioconversions at low temperatures to avoid microbial growth. We focused on the discovery of cold-active CE in this study. A genome mining method based on computational prediction was used to screen the potential genes encoding cold-active enzymes. The CE-encoding gene from *Roseburia intestinalis*, with a predicted high structural flexibility, was expressed heterologously in *Escherichia coli*. The catalytic property of the recombinant enzyme was extensively studied. The optimum temperature and pH of the enzyme were 45°C and 7.0, respectively. The specific activity of this enzyme to catalyze conversion of lactose to epilactose was measured to be 77.3 ± 1.6 U/mg. The kinetic parameters, including turnover number (k_{cat}), Michaelis constant (K_m), and catalytic efficiency (k_{cat}/K_m) using lactose as a substrate were 117.0 ± 7.7 s^{-1}, 429.9 ± 57.3 mM, and 0.27 m$M^{-1}s^{-1}$, respectively. In situ production of epilactose was carried out at 8°C: 20.9% of 68.4 g/L lactose was converted into epilactose in 4 h using 0.02 mg/mL (1.5 U/mL, measured at 45°C) of recombinant enzyme. The enzyme discovered by this in silico method is suitable for low-temperature applications.
Key words: cellobiose 2-epimerase, MD simulation, rational genome mining, cold-active enzyme, epilactose

INTRODUCTION

In the context of global needs for sustainable manufacturing technologies, biocatalysts are more attractive than traditional catalysts for chemical transformations. A serious obstacle for the application of biocatalysts is that they usually cannot work under harsh environments. Many industrially relevant enzymes originate from microorganisms; those from extremophiles are a particularly interesting source for the discovery of hot- or cold-active biocatalysts (Uchiyama and Miyazaki, 2009). The enzymes with high catalytic activity at temperatures below 25°C are known as cold-active enzymes (Santiago et al., 2016). Cold-active enzymes have potential applications in broad ranges of the food industry (Adapa et al., 2014), such as the dairy industry. Their ability to catalyze reactions at low temperatures makes them cost-efficient biocatalysts that can work without energy-consuming heating processes. Moreover, manufacturing at low temperatures (less than or equal to 8°C) using cold-active enzymes can prevent microbial growth and the undesirable reactions that occur during heat processes, such as Maillard reactions, which will improve the taste and quality of the products.

Permanently cold environments (less than 5°C) cover approximately 80% of the Earth's biosphere. Thus, a large portion of organisms live in cold environments. Most living organisms inhabiting permanently cold environments are psychrophilic microorganisms. The routine approach to discover novel cold-active enzymes is to isolate the enzyme of interest from cultures of psychrophilic microorganisms (Adapa et al., 2014; Santiago et al., 2016). Another approach is to extract the genes of interest from metagenome libraries using the metagenomics analysis method (Temperton and Giovannoni, 2012). Computational genomics is an attractive culture-independent technique for laboratorial novel enzyme discovery (Gong et al., 2013). The uncharacterized genes mined from a genome database can be expressed in heterologous hosts and identified under laboratory conditions. Genome mining is a promising method to access a wealth of untapped sequence diversity. However, the present approach of genome mining is of some degree of empiricism. There is a certain degree of probability of obtaining enzymes with undesired properties. For example, 7 unknown cellobiose 2-epimerase (**CE**) from various aerobes were characterized by

Received January 3, 2020.
Accepted May 4, 2020.
*Corresponding author: wmmu@jiangnan.edu.cn

Ojima et al., and 4 of them showed significantly lower catalytic efficiencies and were considered unsuitable for producing epilactose (Ojima et al., 2013).

A relatively higher enzymatic activity at low temperatures requires an increase in enzyme flexibility. First, high plasticity of the active site leads to a more accessible binding site, so that the substrates can fit into the enzyme more easily. Second, higher flexibility of an enzyme contributes to the enthalpic change caused by a reduced number of enzyme-substrate interactions, which can increase the turnover number (k_{cat}; Feller and Gerday, 2003). In summary, an important strategy to discover cold-active enzymes is to identify the enzymes with potentially higher flexibilities.

Recently Xiao et al. (2019) developed a genome mining method combined with molecular dynamics (**MD**) simulations to discover novel enzymes with desired stabilities. Simulation of MD is a potent tool to study the behavior of the biomacromolecules. Some easy-to-obtain data from MD simulation can be used to predict protein flexibilities. A sequence of what is currently the most thermostable CE (EC 5.1.3.11) from *Dictyoglomus thermophilum* (Dith-CE) was successfully discovered from 11 uncharacterized CE sequences by this method.

Cellobiose 2-epimerase are well-studied industrial enzymes that can catalyze the bioconversion of lactose to lactulose or epilactose. Lactulose and epilactose are isomeric and epimeric derivatives of lactose, respectively, which both have beneficial effects on health (Watanabe et al., 2008; Aït-Aissa and Aïder, 2014). The epimerization activity of CE is much higher than its isomerization activity (Kuschel et al., 2017). Therefore, the isomeric product will hardly be produced using CE as the biocatalyst in short-time and low-temperature conditions. Furthermore, epilactose will be produced more slowly at low temperatures because of the slow molecular motions. Thus, it is necessary to discover cold-active CE that can efficiently convert lactose into its epimeric product at industrially relevant low temperatures and times. More than 20 genes encoding CE have been heterologously expressed and characterized to date. Most cold-adapted CE were isolated from mesophilic microorganisms (Table 1). The optimum temperatures of the proteins in this table do not correlate well with the growth temperatures of the corresponding microorganisms. For example, the optimum temperatures of Spli-CE and Rual-CE are 45°C and 30°C, respectively, but the growth temperature of *Spirosoma linguale* is lower than that of *Ruminococcus albus*. Therefore, empirical screening of the enzymes derived from cold-adapted microorganisms is a less-efficient and more time-consuming way to discover cold-active enzymes.

In this study, the rational genome mining method was used to discover novel flexible CE, which are supposed to be cold-active. The gene encoding the putative flexible CE originating from *Roseburia intestinalis* (Roin-CE) was expressed in *Escherichia coli*. Afterward, the enzymatic property of this CE was characterized and used for the production of epilactose from lactose in whey at low temperature.

MATERIALS AND METHODS

Materials and Chemicals

Isopropyl β-D-1-thiogalactopyranoside and ampicillin sodium salt were purchased from Sangon Biotech (Shanghai, China). Electrophoresis reagents were purchased from Bio-Rad (Hercules, CA). The competent cells of *E. coli* BL21 (DE3) used for the cloning and expression of the proteins were purchased from Promega (Madison, WI). Chelating Sepharose Fast Flow resin was supplied by GE Healthcare (Uppsala, Sweden). Cheese whey powder was obtained from Apple Foods Tech (Shanghai, China). Epilactose and lactulose were of analytical grade and were obtained from Sigma-Aldrich (St. Louis, MO). All other materials and reagents were purchased from Sinopharm Chemical Reagent (Shanghai, China).

Sequence and 3-Dimensional Structures of CE

Fifteen putative uncharacterized CE sequences (Table 2) were randomly selected from the Universal Protein Resource Protein knowledgebase (UniProtKB; UniProt Consortium, 2018). The structures of uncharacterized CE were homologically modeled by the SWISSMODEL online server (Waterhouse et al., 2018) using Protein Data Bank (Berman et al., 2000) deposited structures as templates.

MD Simulation

Molecular dynamics simulations were performed using the GROMACS package (Version 2018.1; Abraham et al., 2015). Each structure of CE was solvated in a periodic boundary conditions box that left a space of 1.0 nm around the solute, with extended simple point-charge water model and 150 m*M* NaCl. Each system was subjected to GROMOS96 54a7 all-atom force field (Schmid et al., 2011) and then relaxed through energy minimization, to avoid bad molecular contacts. A 100-ps isochoric isothermal (NVT) simulation followed by a 100-ps isobaric isothermal (NPT) simulation was performed for each system. These equilibrium simulations were conducted at 353 K and 100 kPa pressure with the position restraining force on the heavy atoms of the protein. Velocity rescaling thermostat and Berend-

Table 1. Temperature properties of the characterized cellobiose 2-epimerases (CE) and their microbial sources

Enzyme	Enzyme optimum temperature (°C)	Microorganism source	Growth temperature[1] (°C)	Optimum temperature[1] (°C)	Temperature range	Reference
Bafr-CE	45	*Bacteroides fragilis*	37	—	Mesophilic	(Senoura et al., 2009)
Caob-CE	70	*Caldicellulosiruptor obsidiansis*	55–85	78	Hyperthermophilic	(Chen et al., 2017)
Casa-CE	75	*Caldicellulosiruptor saccharolyticus*	65	—	Hyperthermophilic	(Park et al., 2011)
Cele-CE	40	*Cellulosilyticum lentocellum*	30	—	Mesophilic	(Krewinkel et al., 2015)
Cevu-CE	40	*Cellvibrio vulgaris*	20–30	—	Psychrophilic/mesophilic	(Saburi et al., 2015)
Dith-CE	75	*Dictyoglomus thermophilum*	70	73–78	Hyperthermophilic	(Xiao et al., 2019)
Ditu-CE	70	*Dictyoglomus turgidum*	75–80	72	Hyperthermophilic	(Kim et al., 2012)
Dyfe-CE	50	*Dyadobacter fermentans*	15–37	—	Mesophilic	(Ojima et al., 2013)
Dyga-CE	40	*Dysgonomonas gadei*	37	—	Mesophilic	(Krewinkel et al., 2015)
Euce-CE	35	*Eubacterium cellulosolvens*	39	—	Mesophilic	(Taguchi et al., 2008)
Fljo-CE	35	*Flavobacterium johnsoniae*	25–30	—	Mesophilic	(Ojima et al., 2013)
Heau-CE	45	*Herpetosiphon aurantiacus*	30	—	Mesophilic	(Ojima et al., 2013)
Pehe-CE	35	*Pedobacter heparinus*	25–30	—	Mesophilic	(Ojima et al., 2013)
Rhma-CE	80	*Rhodothermus marinus*	65	—	Hyperthermophilic	(Ojima et al., 2011)
Rual-CE	30	*Ruminococcus albus*	37	—	Mesophilic	(Ito et al., 2007)
Sade-CE	35	*Saccharophagus degradans*	28–30	30	Mesophilic	(Ojima et al., 2013)
Spli-CE	45	*Spirosoma linguale*	26	—	Mesophilic	(Ojima et al., 2013)
Spth-CE	60	*Spirochaeta thermophila*	65	—	Hyperthermophilic	(Park et al., 2013)
Tetu-CE	35	*Teredinibacter turnerae*	23–30	30–35	Mesophilic	(Ojima et al., 2013)
Thsa-CE	60	*Thermoanaerobacterium saccharolyticum*	60	—	Hyperthermophilic	(Chen et al., 2015)
Thth-CE	60	*Thermoanaerobacterium thermosaccharolyticum*	60	—	Hyperthermophilic	(Xiao et al., 2019)

[1]Data from BacDive database (Reimer et al., 2019). Optimum temperature = optimum temperature for the microorganism to grow. Dashes indicate that data were not provided from the database.

sen barostat were used for temperature and pressure controls, respectively. The bond interactions were calculated using the LINCS algorithm. The cutoff scheme of buffered Verlet lists was exerted to calculate the nonbonded interactions. Coulomb potential was calculated by particle-mesh Ewald algorithm accelerated by graphical processing units with a cutoff distance of 1.2 nm. Lennard-Jones potential was smoothly cut off with a switching function between 1.0 nm and 1.2 nm. The final outputs of the well-equilibrated simulations were then subjected to 100-ns MD simulations without position restraints. Nosé-Hoover thermostat and Rahman-Parrinello barostat were implemented in the production simulations. Three replications were done for each simulation with different initial velocities. The hydrogen bonds and root mean square deviation (**RMSD**) of backbone atom positions in each trajectory (20–100 ns) were calculated using g_hbond and g_rms (GROMACS analysis tools), respectively.

Recombinant Cell Construction

The pET-22b (+) plasmid containing the gene of the CE from *R. intestinalis* was synthesized by Sangon Biotech. The Roin-CE gene sequence was obtained from *R. intestinalis* L1–82 chromosome sequence (NCBI no. NZ_LR027880.1, locus tag RIL182_RS05565). The gene was cloned between the sites of *Nde*I and *Xho*I. An in-frame 6 × His-tag sequence was fused after the 3′-terminal sequence of the open reading frame. The plasmid pET-22b(+)-Roin-CE was transformed into *E. coli* BL21(DE3) cells. The recombinant cells for the expressions of the Dith-CE, Ditu-CE, Casa-CE, Caob-CE, and Rhma-CE in this study were constructed using the same method.

Expression and Identification of the Recombinant Protein

Each single colony of the reconstructed strains was inoculated into the LB medium containing 100 μg/mL ampicillin and cultivated at 37°C under agitation (200 rpm) until optical density at 600 nm reached 0.6. Isopropyl β-D-1-thiogalactopyranoside was added to a final concentration of 0.5 m*M* to induce target protein expression. The cultures were then incubated at 28°C for 8 h. The recombinant cells were harvested by centrifugation and washed with 0.9% NaCl solution. The supernatant was applied to a His-Trap affinity chromatography column that was equilibrated with 50 m*M* phosphate buffer (pH 7.0) containing 300 m*M* NaCl, and the bound protein was eluted with a linear gradient from 50 to 500 m*M* imidazole at 1 mL/min. The eluted purified protein was detected at 280 nm during chromatography. The subunit molecular weight of purified Roin-CE was examined by SDS-PAGE (12% wt/vol acrylamide). The gel was dyed with 0.1% Coomassie Brilliant Blue staining and was destained with a mixture of ethanol, acetic acid, and water (5:10:85, vol/vol/vol) for 24 h. The protein concentration was determined according to the method described by Bradford (Bradford, 1976).

Table 2. Molecular dynamics simulation results of the uncharacterized cellobiose 2-epimerases (CE)

Originating microorganism	Uniprot entry no.[1]	AA sequence length	Growth temperature[2] (°C)	RMSD[3] (nm)	Hbond/AA[4]
Bacteroides helcogenes	E6STC4	395	37	0.443	0.821
Roseburia intestinalis	A0A173VYD4	409	37	0.373	0.78
Geobacillus spp.	D3EK23	399	50	0.363	0.838
Eubacterium siraeum	A0A174ZCD1	386	37	0.361	0.806
Phycisphaerae bacterium	A0A1U9NLQ6	427	—	0.356	0.79
Bacteroides cellulosilyticus	A0A0P0G5I9	395	30–39	0.348	0.833
Opitutus terrae	B1ZW09	403	28	0.337	0.781
Ruminococcus champanellensis	D4LBQ1	397	37–39	0.334	0.789
Paenibacillus macerans	A0A091A1I0	396	30–37	0.334	0.817
Mucilaginibacter gotjawali	A0A0X8X5I1	404	28	0.333	0.803
Roseburia faecis	A0A173ULA0	391	37	0.325	0.848
Niastella koreensis	G8TH40	406	10–37	0.308	0.797
Paenibacillus riograndensis	A0A0E4H975	406	30	0.304	0.767
Fusicatenibacter saccharivorans	A0A174TW20	392	37	0.285	0.837
Caldicellulosiruptor bescii	B9MP66	390	42–90	0.259	0.84

[1]Universal Protein Resource Protein knowledgebase (UniProtKB; UniProt Consortium, 2018)

[2]Data from BacDive database (Reimer et al., 2019). Dash indicates data not provided from the database.

[3]RMSD = root mean square deviation of protein backbone atom positions.

[4]Hbond/AA = hydrogen bond divided by length of the AA sequence.

Enzyme Assay

The activity of Roin-CE was determined by the rate of epimerization reaction catalyzed by the purified enzyme. The measurement was performed in a 1-mL reaction system containing 200 m*M* lactose, 50 m*M* sodium phosphate buffer (pH 7.0), and 0.02 mg of purified enzyme. The reaction at 45°C was stopped after 15 min by addition of HCl to the reaction system, to HCl final concentration of 200 m*M*. Each assay was performed in triplicate. One unit of epimerization activity was defined as the mount of enzyme catalyzing the formation of 1 μmol of epilactose/min.

Analytical Methods

The reaction product was filtered through a 0.22-m membrane after dilution to a proper concentration and was analyzed subsequently using an HPLC system (Waters 2695, Waters Corporation, Milford, MA), equipped with a refractive index detector (Waters Alliance 2695) and a Shodex VG-50 4E column (4.6-mm internal diameter × 250 mm, 5 μm; particle size, Shodex, Tokyo, Japan). The working temperature of the column was set at 40°C. The column was eluted with a mixture of methanol, acetonitrile, and water (75:20:5, vol/vol/vol) at a flow rate of 1 mL/min. The peak times for lactulose, epilactose, and lactose were 13.14, 14.38, and 16.51 min, respectively.

Biochemical Characterization

The characterization of the recombinant CE was carried out using lactose as a substrate. The effect of pH on epimerization activity of CE was measured at 45°C in NaOAc-HOAc buffer (50 m*M*, pH 5.0 to 6.0), sodium phosphate buffer (50 m*M*, pH 6.0 to 7.5) and Britton-Robinson buffer (50 m*M*, pH 7.5 to 8.5). The effect of temperature was investigated at 25 to 60°C in 50 m*M* sodium phosphate buffer (pH 7.0). The thermostability of the purified protein was investigated by measuring the residual enzyme activities using the standard assay at 45°C after preincubation at various temperatures (35 to 55°C). The relative activity of the enzyme without preincubation was regarded as 100%. The kinetic parameters of epimerization activity were calculated via the nonlinear regression (MyCurveFit online tool, 2019; https://mycurvefit.com/) to the Michaelis-Menten equation, using 50 to 800 m*M* substrates and 0.02 mg of purified enzyme. The reaction was conducted at 45°C in 50 m*M* sodium phosphate buffer (pH 7.0).

The melting temperature of CE was determined using a Q2000 differential scanning calorimeter (Nano DSC, TA Instruments, New Castle, DE). The enzyme was re-dialyzed using sodium phosphate buffer (50 m*M*, pH 7.0). The samples and reference buffer were loaded into 2 separate cells after being degassed under vacuum (635 mmHg) for 10 min. The cells were heated from 25 to 100°C at a scan rate of 1°C/min. The melting temperature was calculated using NanoAnalyze software (TA Instruments).

Bioconversion of Epilactose from Whey Lactose

The whey powder was dissolved in 50 m*M* sodium phosphate buffer (pH 7.0). The solution was then centrifuged (8,000 × *g*) at 4°C for 20 min. The supernatant was filtered before production. The lactose accounted for about 76% of the net weight of the whey powder according to the HPLC analysis. We used 90 g/L whey powder solution containing 68.4 g/L lactose for epilactose production. No epilactose was detected in the whey solution before production. The production was carried out at 8°C using 0.02 g/L Roin-CE and Dith-CE as biocatalyst.

RESULTS AND DISCUSSION

Selection of Putative Flexible CE

Fifteen out of 2,730 uncharacterized CE sequences were randomly selected from the UniProt database (UniProt Consortium, 2018; Table 2). Fourteen of the uncharacterized CE sequences originate from mesophilic bacteria, and one of them is from a hyperthermophile. Most of the mesophilic bacteria have a reported optimum growth temperature near 37°C. It is difficult to tell the cold adaptability of the encoded proteins according to the growth temperature of their microbial sources. It is tedious and costly to measure the enzymatic properties of such a high number of enzymes experimentally. Thus, theoretical screening was carried out by investigating the structural flexibility of the putative CE to select a sequence with the desired property.

Because of lack of selective pressure for the proteins to be stable, organisms living in a cold environment adjust their enzymes to be more mobile or flexible to adapt to slow molecular motion at low temperatures (D'Amico et al., 2003). The dynamic quenching of tryptophan fluorescence experimentally proved that psychrophilic enzymes possess higher flexibility than do thermophilic enzymes (Cipolla et al., 2012). According to this adaptive strategy, the flexibility of the encoded proteins was predicted using MD simulation to look for cold-adapted enzymes; MD simulation can assess

the dynamic features of enzymes without experimental expressions. The average RMSD value of Cα atom positions reflects the flexibility of the protein (Niu et al., 2015; Khan et al., 2017; Dong et al., 2018). The number of the formed hydrogen bonds inside the protein also represents the structural stability of the corresponding protein (Srivastava and Sinha, 2014). The average backbone RMSD and the hydrogen bond divided by the length of the amino acid sequence (**Hbond/AA**) were obtained from the MD simulation. The enzyme encoded by the CE gene from *R. intestinalis* had a relatively high RMSD value and a relatively low Hbond/AA value, and thus was considered to be a flexible protein (Figure 1). Conversely, the CE from *Bacteroides helcogenes* and *Paenibacillus riograndensis* had either high Hbond/AA value or low RMSD value and thus were not considered in this study. It is worth mentioning that the CE from hyperthermophile *C. orbescii* having a high Hbond/AA value and low RMSD value indicates that it is a thermostable enzyme, which agrees with what one might expect for enzymes for a microorganism adapted to that environment. The potentially flexible CE sequence from *R. intestinalis* was then selected for further study. The amino acid sequence and enzymatic properties of Roin-CE were analyzed.

Amino Acid Sequence Analysis

The similarity of the sequences in this study was investigated by multiple sequence alignments and phylogenetic tree analysis, constructed using Clustal Omega (Sievers et al., 2011) and Mega software (Kumar et al., 2016), respectively. Most of the sequence identities between the characterized and uncharacterized CE sequences range from 30 to 60%. The CE sequence from *R. intestinalis* shows a low-level evolutionary relationship with the currently characterized CE sequences (Figure 2). Although it shares the highest identity with the CE sequence from *Roseburia faecis*, the MD simulation results of these 2 were quite different.

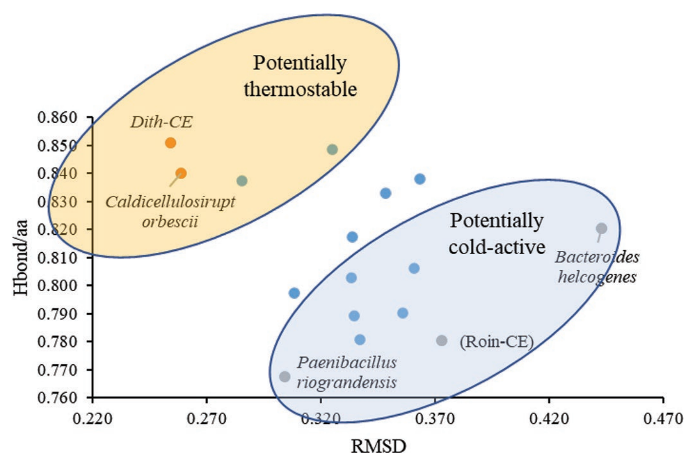

Figure 1. Average backbone root mean square deviation (RMSD) values and hydrogen bond divided by length of AA sequence (Hbond/AA) values of uncharacterized cellobiose 2-epimerases (CE). Data for Dith-CE was obtained from Xiao et al. (2019).

Expression and Activity Investigation of Roin-CE

The recombinant plasmids containing the Roin-CE gene were expressed by *E. coli.* Purified Roin-CE, with a purification yield of 9.8%, exhibited a single strong band of approximately 42 kDa on SDS-PAGE (Supplemental Figure S1, https://doi.org/10.3168/jds.2020-18153; theoretical mass, based on amino acid sequence, is 47.5 kDa). The purified CE was used for further investigations. Quantitative HPLC analysis showed that in such a short reaction time (15 min), Roin-CE can only catalyze the epimerization reactions toward lactose and cellobiose, and no isomerization products were found under these conditions. However, the result did not contradict the inference proposed by Kuschel et al. (2017), that all CE are likely able to catalyze both epimerization and isomerization. This is true, because they carried out the reaction over a long period, with high CE amounts, and only a small amount of isomeric product was produced by CE from the mesophilic microorganism.

Specific Activity and Kinetic Parameters of Roin-CE

The specific epimerization activity of Roin-CE converting lactose to epilactose was measured to be 77.3 ± 1.6 U/mg. For comparison purposes, the epimerization activity of the characterized CE toward lactose are listed in Table 3. Comparing with those of the other thermostable CE, the specific activities of Ditu-CE (Kim et al., 2012) and Spth-CE (Park et al., 2013), measured with low substrate concentrations, were significantly low. One possible reason might be that the corresponding experiments were conducted in the nonlinear portion of the course, in which the enzymatic reactions reached the plateau. The activities of Caob-CE, Dith-CE, and Rhma-CE were higher than that of Roin-CE. However, these enzymes originate from hyperthermophiles and were investigated at their optimum temperatures, which are 70, 75, and 80°C, respectively. The Roin-CE showed the highest epimerization activity toward lactose among the reported mesophilic CE. It is well accepted that the flexibility of the active site is important for enzyme activity (Tsou, 1993; Broos et al., 1995). Thus, the high activity of Roin-CE suggested that this enzyme is relatively flexible, as the MD simulation predicted.

The kinetic parameters of Roin-CE were investigated via lactose and cellobiose epimerization reactions (Supplemental Figure S2, https://doi.org/10.3168/jds.2020-18153). The Roin-CE had a higher catalytic efficiency toward cellobiose than toward lactose. The catalytic efficiency [k_{cat}/Michaelis constant (K_m) value for cellobiose was 0.75 mM^{-1}s^{-1}, and 0.27 mM^{-1}s^{-1} for lactose. The k_{cat} values of Roin-CE toward cellobiose and lactose were 98.5 ± 2.9 and 117.0 ± 7.7/s, respectively. The K_m values of Roin-CE toward cellobiose and lactose were 131.0 ± 11.6 mM and 429.9 ± 57.3 mM, respectively, which were both lower than the corresponding values of Dith-CE, investigated by us previously (Xiao et al., 2019). Thus, Roin-CE had a higher affinity for cellobiose and epimerase then did the more stable Dith-CE. This is probably because of the high flexibility of Roin-CE, as this enzyme can position the substrate better in its active site when it binds.

Effects of pH and Temperature on Activity

The enzymatic properties of Roin-CE were characterized by measuring epimerization activity using lactose as substrate. A mesophilic enzyme, Roin-CE displayed its maximal activity at pH 7.0 and 45°C (Figure 3A and 3B). The activity of Roin-CE dropped dramatically when the working temperature increased from 45 to 60°C.

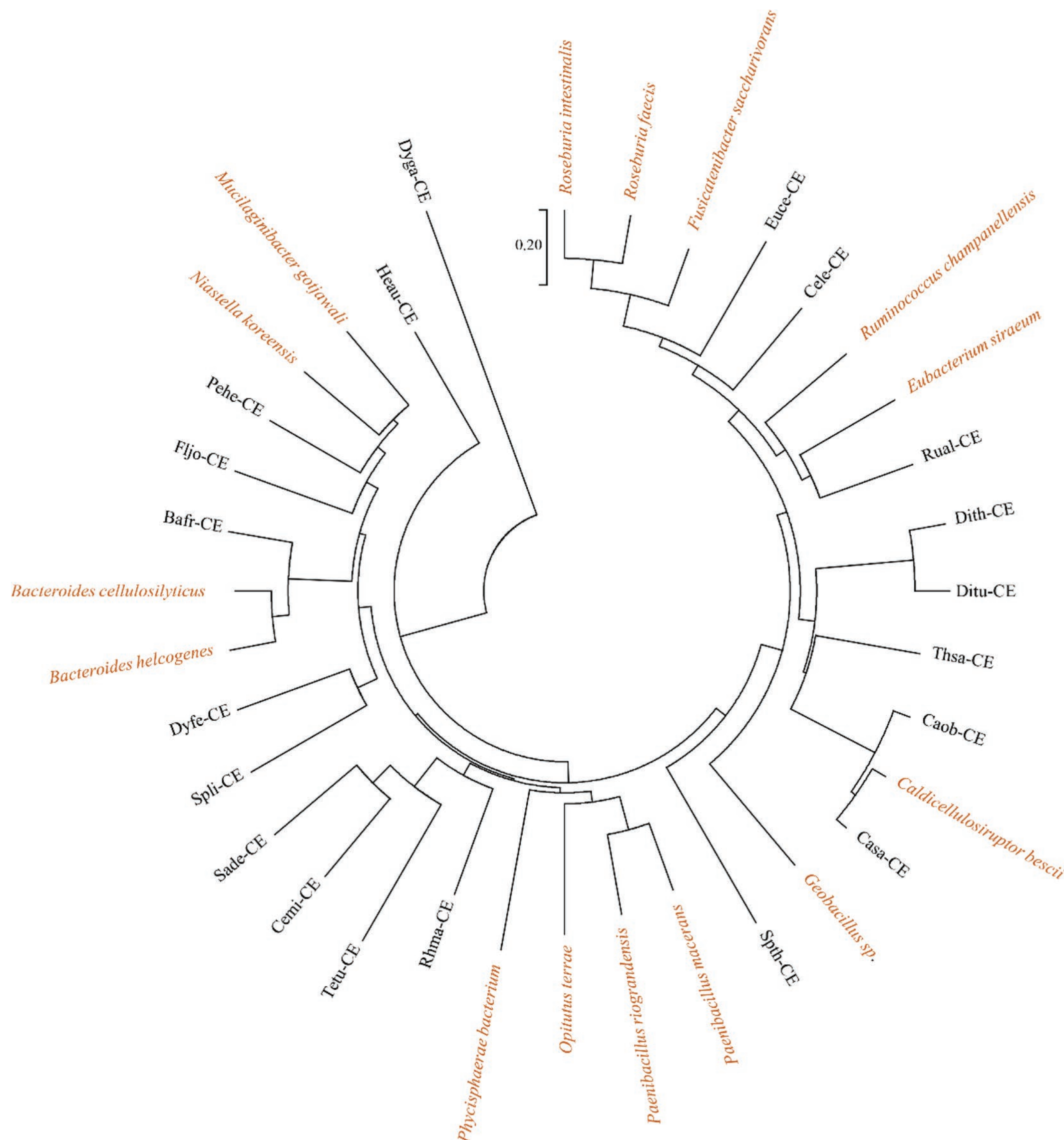

Figure 2. Phylogenetic tree of the characterized and uncharacterized cellobiose 2-epimerases (CE) in this study. The tree was built using the neighbor-joining method in Mega software (Kumar et al., 2016). For CE abbreviations, see Table 1.

Table 3. Epimerization activities of different cellobiose 2-epimerases (CE) toward lactose

Enzyme	Epimerization activity (U/mg)	Optimum temperature (°C)	Buffer condition[1]	Substrate concentration	Reference
Bafr-CE	76.8	45	10 m*M* PIPES buffer, pH 7.5	150 m*M*	Kuschel et al., 2017
Caob-CE	118.1	70	50 m*M* PB, pH 7.5	200 m*M*	Chen et al., 2017
Cele-CE	0.7	40	20 m*M* Tris-HCl buffer, pH 8.0	50 g/L	Krewinkel et al., 2015
Ditu-CE	9.7	70	50 m*M* PIPES buffer, pH 7.0	10 m*M*	Kim et al., 2012
Dyfe-CE	46.9	50	10 m*M* PIPES buffer, pH 7.7	150 m*M*	Kuschel et al., 2017
Dyga-CE	6.5	40	20 m*M* Tris-HCl buffer, pH 7.5	50 g/L	Krewinkel et al., 2015
Fljo-CE	48.5	35	10 m*M* PIPES buffer, pH 8.4	150 m*M*	Kuschel et al., 2017
Rhma-CE	87.5	80	10 m*M* PIPES buffer, pH 6.3	150 m*M*	Kuschel et al., 2017
Rual-CE	38.8	30	10 m*M* PIPES buffer, pH 7.5	150 m*M*	Kuschel et al., 2017
Spth-CE	7.8	60	50 m*M* PIPES buffer, pH 7.0	10 m*M*	Park et al., 2013
Thsa-CE	13.5	60	50 m*M* PB, pH 7.0	100 m*M*	Chen et al., 2015
Dith-CE	160.1	75	50 m*M* PB, pH 7.0	200 m*M*	Xiao et al., 2019
Roin-CE	77.3 ± 1.6	45	50 m*M* PB, pH 7.0	200 m*M*	This study

[1]PIPES = piperazine-*N*,*N*′-bis(2-ethanesulfonic acid); PB = phosphate buffer.

Most of the optimum temperatures of the cold-active enzymes are lower than 40°C. Further, the activities of the mesophilic enzymes are not expected to be high at low temperatures (Santiago et al., 2016). However, because of its high flexibility, Roin-CE maintained more than 60% of its activity at 18°C, 27 degrees below its optimum temperature of 45°C. For comparison, Rual-CE from mesophile *R. albus* NE1 maintained about 55% of its activity at 20 degrees below its optimum temperature (Ito et al., 2008). Therefore, Roin-CE is suitable for the low-temperature production.

Not all of the cold-active enzymes are from psychrophilic or mesophilic organisms. A β-galactosidase from hyperthermophilic archaeon *Pyrococcus furiosus* was still active at low temperatures. This thermophilic enzyme had optimal activity at 90°C but retained 8% of its maximal activity at 0°C and was capable of hydrolysis of lactose in milk processing (Dong et al., 2014). Thus, the structural flexibility of the enzyme might be a more suitable measurement standard for cold-active enzymes than the optimum temperature of the enzymatic reaction.

Thermostability of the Recombinant Enzyme

Because the efficient reactions catalyzed by enzymes at low temperatures require increases in the flexibilities of proteins, the stability of cold-active enzymes is usually not high. Often, the melting temperature is used as an indicator for enzyme thermostability (Syed et al., 2014). The differential scanning calorimetry analysis was carried out to detect the structural stability of the recombinant enzyme. The melting temperature of Roin-CE was measured by differential scanning calorimetry to be 57.3°C, which is much lower than for thermostable enzymes. The melting temperatures of Dith-CE (Xiao et al., 2019) and Caob-CE (Chen et al., 2017) were 92.2°C and 86.7°C, respectively.

The stability of Roin-CE was investigated at 8, 35, 40, 45, 50, and 55°C. The half-life values of Roin-CE at 35, 40, 45, 50, and 55°C were calculated to be 151.0, 137.8, 23.4, 4.2, and 0.5 h, respectively (Figure 3C). This enzyme is relatively stable and can maintain more than 90% of its initial activity after 1-d production below 40°C. The activity of Roin-CE at 8°C remained the same during our investigation. Although it is not as stable as enzymes from thermophiles, such as Dith-CE (Xiao et al., 2019) and Casa-CE (Park et al., 2011), Roin-CE is stable enough to work at low temperatures.

Production of Epilactose from Whey at Low Temperature

Production of epilactose was conducted at 8°C, a suitable process temperature for the dairy industry (Krewinkel et al., 2015; Rentschler et al., 2015). Production was carried out in lactose solution from whey, to take into consideration the potential industrial application of Roin-CE. Epilactose production abilities using Roin-CE, Casa-CE, Caob-CE, Ditu-CE, Rhma-CE, and Dith-CE were investigated under the same conditions for comparison. Time-course analysis showed that, at a low temperature, Roin-CE is more efficient for epilactose production than the other thermostable CE, which exhibited higher specific activities at their optimum temperatures. Using Roin-CE as biocatalyst, 20.9% of 68.4 g/L lactose was converted into epilactose in 4 h, whereas only 2.35% of lactose was converted by Dith-CE in the same time (Figure 4). The epimerization specific activity of Roin-CE was 27.40 ± 0.51 U/mg at 8°C, retaining more than 35% of its maximum specific activity, whereas the specific activity of Dith-CE went

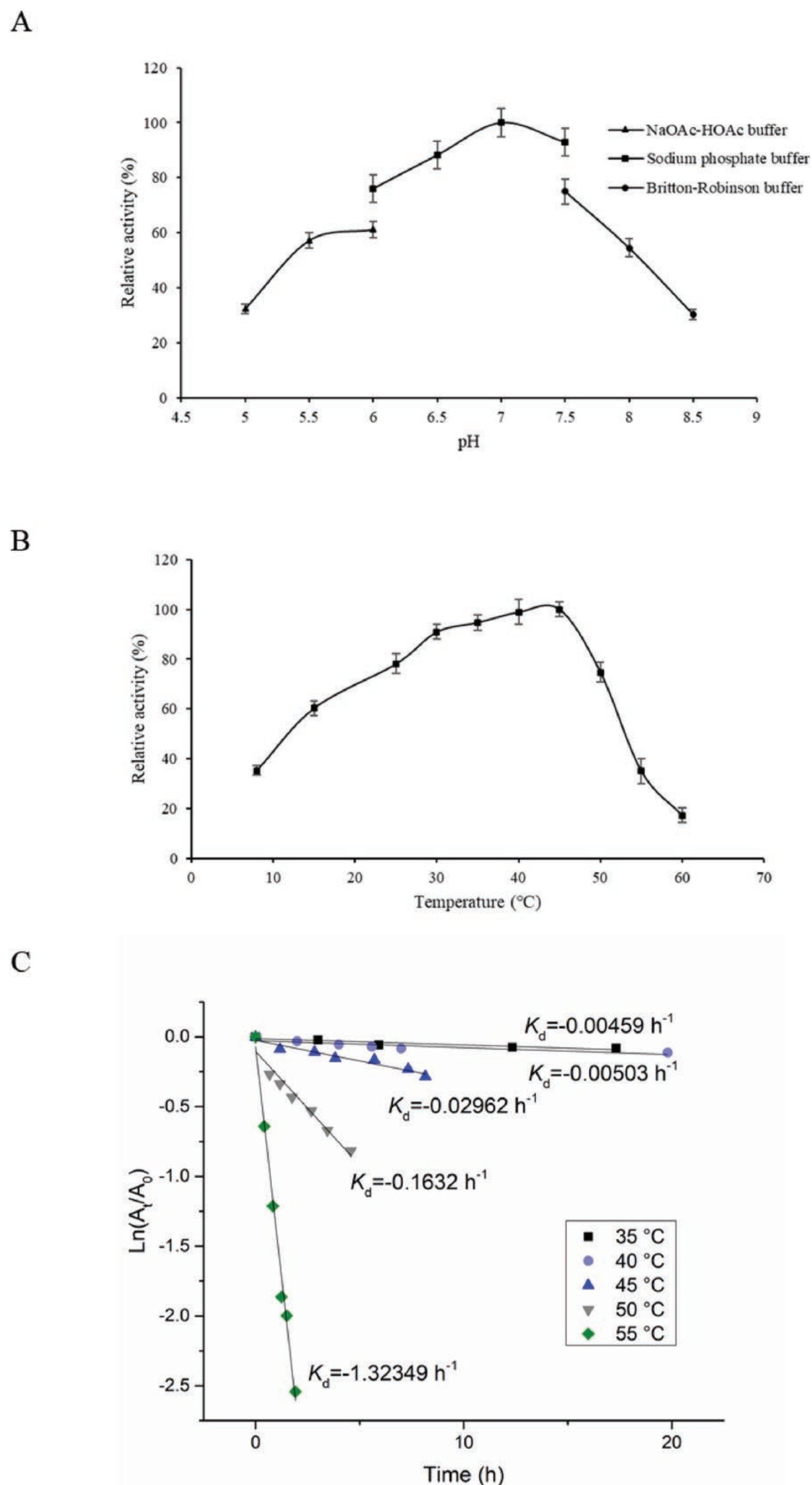

Figure 3. pH and temperature profiles of cellobiose 2-epimerase Roin-CE: (A) effect of pH on the activity of Roin-CE; (B) effect of temperature on the activity of Roin-CE; (C) effect of temperature on the stability of Roin-CE. The vertical axis of panel C represents $\log(A_t/A_0)$, where A_t is the residual activity of Roin-CE after heat treatments at certain temperatures and A_0 represents the initial activity. The decay constant K_d is defined as follows: $A_t = A_0 \exp(-K_d t)$, where t is the time elapsed during reaction. Values are means of 3 replications ± SD.

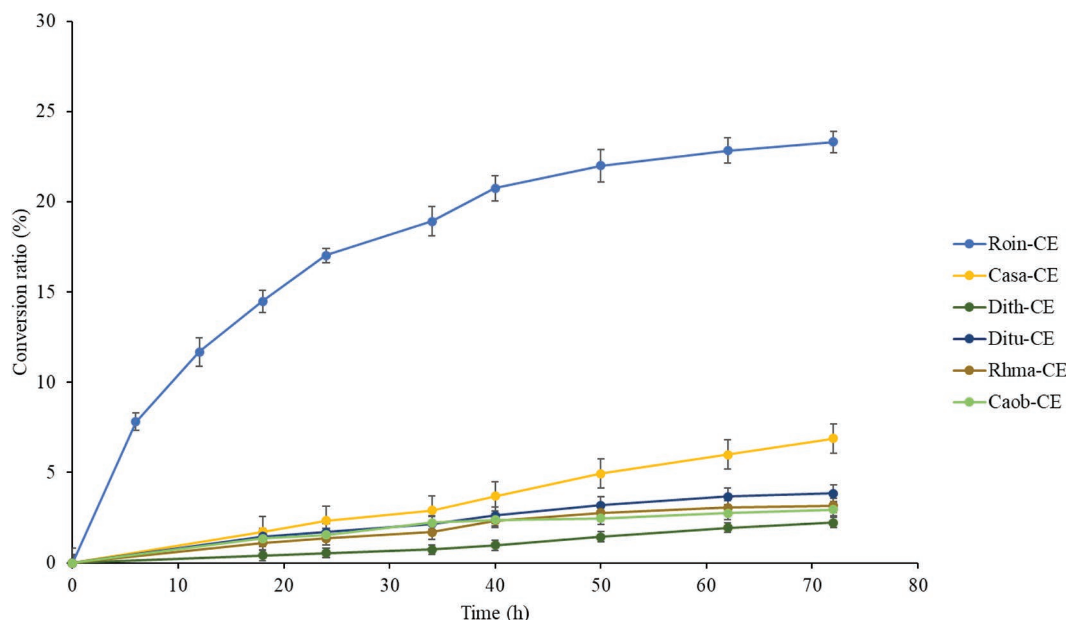

Figure 4. Epilactose production by Roin-CE and other cellobiose 2-epimerases (CE) at 8°C. Conversion ratio is the relative proportion of the produced epilactose to the initial lactose concentration. Values are means of 3 replications ± SD.

down to 1.48 ± 0.01 U/mg, which was 18.5-fold lower than that of Roin-CE. The activity of Roin-CE was not significantly influenced by the other ingredients in whey, compared with production using lactose solution (data not shown).

The biotransformation catalyzed by Fljo-CE and Pehe-CE at 8°C was reported in an early study (Krewinkel et al., 2014). Fljo-CE and Pehe-CE retained 5.5 and 15.9% of their maximum activity, respectively, at this low temperature. Krewinkel et al. (2015) characterized 2 CE from mesophilic bacteria to discover enzymes suitable for low-temperature applications. The strategy to discover novel cold-active enzymes from the enzymes originating from mesophilic bacteria is feasible but less reliable. With lower growth temperature, Cele-CE is less efficient than Dyga-CE to catalyze lactose epimerization. However, 16% of maximal activity was detected using Cele-CE as biocatalyst in lactose buffer at 8°C, whereas the counterpart for Dyga-CE is 2%.

CONCLUSIONS

The genome mining method, combined with MD simulations, was used to discover and characterize a novel cold-active enzyme: Roin-CE, with high predicted flexibility. Our study showed that this enzyme had a relatively high activity in cold environment and could be suitable for low-temperature epilactose production. This genome mining method is a powerful tool to discover novel enzymes with desirable stabilities.

ACKNOWLEDGMENTS

The simulations in this paper were run on the Odyssey cluster, supported by the FAS Division of Science, Research Computing Group at Harvard University, Cambridge, MA. This work was supported by the National Natural Science Foundation of China (No. 31801583 and No. 31922073; Beijing), the Natural Science Foundation of Jiangsu Province (No. BK20180607; Nanjing, China), the Zhangjiagang International Cooperation Project (ZKH1905; Zhangjiagang, China), and the National First-Class Discipline Program of Food Science and Technology (No. JUFSTR20180203; Wuxi, China). The authors have not stated any conflicts of interest.

REFERENCES

Abraham, M. J., T. Murtola, R. Schulz, S. Páll, J. C. Smith, B. Hess, and E. Lindahl. 2015. GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. SoftwareX 1–2:19–25. https://doi.org/10.1016/j.softx.2015.06.001.

Adapa, V., L. Ramya, K. Pulicherla, and K. S. Rao. 2014. Cold active pectinases: advancing the food industry to the next generation. Appl. Biochem. Biotechnol. 172:2324–2337.

Aït-Aissa, A., and M. Aïder. 2014. Lactulose: Production and use in functional food, medical and pharmaceutical applications. Practical and critical review. Int. J. Food Sci. Technol. 49:1245–1253. https://doi.org/10.1111/ijfs.12465.

Berman, H. M., J. Westbrook, Z. Feng, G. Gilliland, T. N. Bhat, H. Weissig, I. N. Shindyalov, and P. E. Bourne. 2000. The Protein Data Bank. Nucleic Acids Res. 28:235–242. https://doi.org/10.1093/nar/28.1.235.

Bradford, M. M. 1976. 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. https://doi.org/10.1016/0003-2697(76)90527-3.

Broos, J., A. J. Visser, J. F. Engbersen, W. Verboom, A. van Hoek, and D. N. Reinhoudt. 1995. Flexibility of enzymes suspended in organic solvents probed by time-resolved fluorescence anisotropy. Evidence that enzyme activity and enantioselectivity are directly related to enzyme flexibility. J. Am. Chem. Soc. 117:12657–12663. https://doi.org/10.1021/ja00156a001.

Chen, Q., R. Levin, W. Zhang, T. Zhang, B. Jiang, T. Stressler, L. Fischer, and W. Mu. 2017. Characterisation of a novel cellobiose 2-epimerase from thermophilic *Caldicellulosiruptor obsidiansis* for lactulose production. J. Sci. Food Agric. 97:3095–3105. https://doi.org/10.1002/jsfa.8148.

Chen, Q., W. Zhang, T. Zhang, B. Jiang, and W. Mu. 2015. Characterization of an epilactose-producing cellobiose 2-epimerase from *Thermoanaerobacterium saccharolyticum*. J. Mol. Catal., B Enzym. 116:39–44. https://doi.org/10.1016/j.molcatb.2015.03.005.

Cipolla, A., F. Delbrassine, J. L. Da Lage, and G. Feller. 2012. Temperature adaptations in psychrophilic, mesophilic and thermophilic chloride-dependent alpha-amylases. Biochimie 94:1943–1950. https://doi.org/10.1016/j.biochi.2012.05.013.

D'Amico, S., J. C. Marx, C. Gerday, and G. Feller. 2003. Activity-stability relationships in extremophilic enzymes. J. Biol. Chem. 278:7891–7896. https://doi.org/10.1074/jbc.M212508200.

Dong, Q., X. Yan, M. Zheng, and Z. Yang. 2014. Characterization of an extremely thermostable but cold-adaptive β-galactosidase from the hyperthermophilic archaeon *Pyrococcus furiosus* for use as a recombinant aggregation for batch lactose degradation at high temperature. J. Biosci. Bioeng. 117:706–710. https://doi.org/10.1016/j.jbiosc.2013.12.002.

Dong, Y. W., M. L. Liao, X. L. Meng, and G. N. Somero. 2018. Structural flexibility and protein adaptation to temperature: Molecular dynamics analysis of malate dehydrogenases of marine molluscs. Proc. Natl. Acad. Sci. USA 115:1274–1279. https://doi.org/10.1073/pnas.1718910115.

Feller, G., and C. Gerday. 2003. Psychrophilic enzymes: Hot topics in cold adaptation. Nat. Rev. Microbiol. 1:200–208. https://doi.org/10.1038/nrmicro773.

Gong, J. S., Z. M. Lu, H. Li, Z. M. Zhou, J. S. Shi, and Z. H. Xu. 2013. Metagenomic technology and genome mining: Emerging areas for exploring novel nitrilases. Appl. Microbiol. Biotechnol. 97:6603–6611. https://doi.org/10.1007/s00253-013-4932-8.

Ito, S., S. Hamada, K. Yamaguchi, S. Umene, H. Ito, H. Matsui, T. Ozawa, H. Taguchi, J. Watanabe, J. Wasaki, and S. Ito. 2007. Cloning and sequencing of the cellobiose 2-epimerase gene from an obligatory anaerobe, *Ruminococcus albus*. Biochem. Biophys. Res. Commun. 360:640–645. https://doi.org/10.1016/j.bbrc.2007.06.091.

Ito, S., H. Taguchi, S. Hamada, S. Kawauchi, H. Ito, T. Senoura, J. Watanabe, M. Nishimukai, S. Ito, and H. Matsui. 2008. Enzymatic properties of cellobiose 2-epimerase from *Ruminococcus albus* and the synthesis of rare oligosaccharides by the enzyme. Appl. Microbiol. Biotechnol. 79:433–441. https://doi.org/10.1007/s00253-008-1449-7.

Khan, F. I., B. Nizami, R. Anwer, K. R. Gu, K. Bisetty, M. I. Hassan, and D. Q. Wei. 2017. Structure prediction and functional analyses of a thermostable lipase obtained from *Shewanella putrefaciens*. J. Biomol. Struct. Dyn. 35:2123–2135. https://doi.org/10.1080/07391102.2016.1206837.

Kim, J. E., Y. S. Kim, L. W. Kang, and D. K. Oh. 2012. Characterization of a recombinant cellobiose 2-epimerase from *Dictyoglomus turgidum* that epimerizes and isomerizes beta-1,4- and alpha-1,4-gluco-oligosaccharides. Biotechnol. Lett. 34:2061–2068. https://doi.org/10.1007/s10529-012-0999-z.

Krewinkel, M., M. Gosch, E. Rentschler, and L. Fischer. 2014. Epilactose production by 2 cellobiose 2-epimerases in natural milk. J. Dairy Sci. 97:155–161. https://doi.org/10.3168/jds.2013-7389.

Krewinkel, M., J. Kaiser, M. Merz, E. Rentschler, B. Kuschel, J. Hinrichs, and L. Fischer. 2015. Novel cellobiose 2-epimerases for the production of epilactose from milk ultrafiltrate containing lactose. J. Dairy Sci. 98:3665–3678. https://doi.org/10.3168/jds.2015-9411.

Kumar, S., G. Stecher, and K. Tamura. 2016. MEGA7: Molecular Evolutionary Genetics Analysis Version 7.0 for Bigger Datasets. Mol. Biol. Evol. 33:1870–1874. https://doi.org/10.1093/molbev/msw054.

Kuschel, B., I. Seitl, C. Gluck, W. Mu, B. Jiang, T. Stressler, and L. Fischer. 2017. Hidden reaction: Mesophilic cellobiose 2—Epimerases produce lactulose. J. Agric. Food Chem. 65:2530–2539. https://doi.org/10.1021/acs.jafc.6b05599.

Niu, C., L. Zhu, P. Zhu, and Q. Li. 2015. Lysine-based site-directed mutagenesis increased rigid beta-sheet structure and thermostability of mesophilic 1,3–1,4-beta-glucanase. J. Agric. Food Chem. 63:5249–5256. https://doi.org/10.1021/acs.jafc.5b00480.

Ojima, T., W. Saburi, H. Sato, T. Yamamoto, H. Mori, and H. Matsui. 2011. Biochemical characterization of a thermophilic cellobiose 2-epimerase from a thermohalophilic bacterium, *Rhodothermus marinus* JCM9785. Biosci. Biotechnol. Biochem. 75:2162–2168. https://doi.org/10.1271/bbb.110456.

Ojima, T., W. Saburi, T. Yamamoto, H. Mori, and H. Matsui. 2013. Identification and characterization of cellobiose 2-epimerases from various aerobes. Biosci. Biotechnol. Biochem. 77:189–193. https://doi.org/10.1271/bbb.120742.

Park, C. S., J. E. Kim, J. G. Choi, and D. K. Oh. 2011. Characterization of a recombinant cellobiose 2-epimerase from *Caldicellulosiruptor saccharolyticus* and its application in the production of mannose from glucose. Appl. Microbiol. Biotechnol. 92:1187–1196. https://doi.org/10.1007/s00253-011-3403-3.

Park, C. S., J. E. Kim, S. H. Lee, Y. S. Kim, L. W. Kang, and D. K. Oh. 2013. Characterization of a recombinant mannobiose 2-epimerase from *Spirochaeta thermophila* that is suggested to be a cellobiose 2-epimerase. Biotechnol. Lett. 35:1873–1880. https://doi.org/10.1007/s10529-013-1267-6.

Reimer, L. C., A. Vetcininova, J. S. Carbasse, C. Söhngen, D. Gleim, C. Ebeling, and J. Overmann. 2019. Bac Dive in 2019: Bacterial phenotypic data for High-throughput biodiversity analysis. Nucleic Acids Res. 47:D631–D636. https://doi.org/10.1093/nar/gky879.

Rentschler, E., K. Schuh, M. Krewinkel, C. Baur, W. Claassen, S. Meyer, B. Kuschel, T. Stressler, and L. Fischer. 2015. Enzymatic production of lactulose and epilactose in milk. J. Dairy Sci. 98:6767–6775. https://doi.org/10.3168/jds.2015-9900.

Saburi, W., Y. Tanaka, H. Muto, S. Inoue, R. Odaka, M. Nishimoto, M. Kitaoka, and H. Mori. 2015. Functional reassignment of *Cellvibrio vulgaris* EpiA to cellobiose 2-epimerase and an evaluation of the biochemical functions of the 4-O-β-D-mannosyl-D-glucose phosphorylase-like protein, UnkA. Biosci. Biotechnol. Biochem. 79:969–977. https://doi.org/10.1080/09168451.2015.1012146.

Santiago, M., C. A. Ramirez-Sarmiento, R. A. Zamora, and L. P. Parra. 2016. Discovery, molecular mechanisms, and industrial applications of cold-active enzymes. Front. Microbiol. 7:1408. https://doi.org/10.3389/fmicb.2016.01408.

Schmid, N., A. P. Eichenberger, A. Choutko, S. Riniker, M. Winger, A. E. Mark, and W. F. van Gunsteren. 2011. Definition and testing of the GROMOS force-field versions 54A7 and 54B7. Eur. Biophys. J. 40:843–856. https://doi.org/10.1007/s00249-011-0700-9.

Senoura, T., H. Taguchi, S. Ito, S. Hamada, H. Matsui, S. Fukiya, A. Yokota, J. Watanabe, J. Wasaki, and S. Ito. 2009. Identification of the cellobiose 2-epimerase gene in the genome of *Bacteroides fragilis* NCTC 9343. Biosci. Biotechnol. Biochem. 73:400–406. https://doi.org/10.1271/bbb.80691.

Sievers, F., A. Wilm, D. Dineen, T. J. Gibson, K. Karplus, W. Li, R. Lopez, H. McWilliam, M. Remmert, J. Söding, J. D. Thompson, and D. G. Higgins. 2011. Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. Mol. Syst. Biol. 7:539. https://doi.org/10.1038/msb.2011.75.

Srivastava, A., and S. Sinha. 2014. Thermostability of in vitro evolved *Bacillus subtilis* lipase A: A network and dynamics perspective. PLoS One 9:e102856. https://doi.org/10.1371/journal.pone.0102856.

Syed, K., K. Shale, K. N. H. Nazir, N. Krasevec, S. S. Mashele, and N. S. Pagadala. 2014. Genome-wide identification, annotation and

characterization of novel thermostable cytochrome P450 monooxygenases from the thermophilic biomass-degrading fungi *Thielavia terrestris* and *Myceliophthora thermophila*. Genes Genomics 36:321–333. https://doi.org/10.1007/s13258-013-0170-9.

Taguchi, H., T. Senoura, S. Hamada, H. Matsui, Y. Kobayashi, J. Watanabe, J. Wasaki, and S. Ito. 2008. Cloning and sequencing of the gene for cellobiose 2-epimerase from a ruminal strain of *Eubacterium cellulosolvens*. FEMS Microbiol. Lett. 287:34–40. https://doi.org/10.1111/j.1574-6968.2008.01281.x.

Temperton, B., and S. J. Giovannoni. 2012. Metagenomics: Microbial diversity through a scratched lens. Curr. Opin. Microbiol. 15:605–612. https://doi.org/10.1016/j.mib.2012.07.001.

Tsou, C. L. 1993. Conformational flexibility of enzyme active sites. Science 262:380–381. https://doi.org/10.1126/science.8211158.

Uchiyama, T., and K. Miyazaki. 2009. Functional metagenomics for enzyme discovery: Challenges to efficient screening. Curr. Opin. Biotechnol. 20:616–622. https://doi.org/10.1016/j.copbio.2009.09.010.

UniProt Consortium. 2018. UniProt: The universal protein knowledgebase. Nucleic Acids Res. 46:2699. https://doi.org/10.1093/nar/gky092.

Watanabe, J., M. Nishimukai, H. Taguchi, T. Senoura, S. Hamada, H. Matsui, T. Yamamoto, J. Wasaki, H. Hara, and S. Ito. 2008. Prebiotic properties of epilactose. J. Dairy Sci. 91:4518–4526. https://doi.org/10.3168/jds.2008-1367.

Waterhouse, A., M. Bertoni, S. Bienert, G. Studer, G. Tauriello, R. Gumienny, F. T. Heer, T. A. P. de Beer, C. Rempfer, L. Bordoli, R. Lepore, and T. Schwede. 2018. SWISS-MODEL: Homology modelling of protein structures and complexes. Nucleic Acids Res. 46(W1):W296–W303. https://doi.org/10.1093/nar/gky427.

Xiao, Y., Q. Chen, E. I. Shakhnovich, W. Zhang, and W. Mu. 2019. Simulation-guided enzyme discovery: A new microbial source of cellobiose 2-epimerase. Int. J. Biol. Macromol. 139:1002–1008. https://doi.org/10.1016/j.ijbiomac.2019.08.075.

ORCIDS

Qiuming Chen https://orcid.org/0000-0002-0780-1199

Timo Stressler https://orcid.org/0000-0001-9392-4031

Wanmeng Mu https://orcid.org/0000-0001-6597-527X