The structural basis of substrate promiscuity in UDP-hexose 4-epimerase from the hyperthermophilic Eubacterium Thermotoga maritima

Sun-Mi Shin, Jin Myung Choi, Eric di Luccio, Yong-Jik Lee, Sang-Jae Lee, Sang Jun Lee, Sung Haeng Lee, Dong-Woo Lee

A R T I C L E   I N F O

Article history:
Received 2 July 2015
Received in revised form
18 August 2015
Accepted 31 August 2015
Available online 3 September 2015

Keywords:
Crystal structure
UDP-galactose 4-epimerase
Substrate specificity
Hyperthermophiles
Evolution

A B S T R A C T

UDP-galactose 4-epimerase (GalE) catalyzes the interconversion of UDP-glucose (UDP-Glc) and UDP-galactose (UDP-Gal), which is a pivotal step in the Leloir pathway for D-galactose metabolism. Although GalE is widely distributed in prokaryotes and eukaryotes, little information is available regarding hyperthermophilic GalE. We overexpressed the TM0509 gene, encoding a putative GalE from Thermotoga maritima (TMGalE), in Escherichia coli and characterized the encoded protein. To further investigate the molecular basis of this enzyme’s catalytic function, we determined the crystal structures of TMGalE and TMGalE bound to UDP-Glc at resolutions of 1.9 Å and 2.0 Å, respectively. The enzyme was determined to be a homodimer with a molecular mass of 70 kDa. The enzyme could reversibly catalyze the epimerization of UDP-GalNAc/UDP-GlcNAc as well as UDP-Gal/UDP-Glc at elevated temperatures, with an apparent optimal temperature and pH of 80°C/14°C and 7.0, respectively. The enzyme was determined to be a homodimer with a molecular mass of 70 kDa. The enzyme could reversibly catalyze the epimerization of UDP-GalNAc/UDP-GlcNAc as well as UDP-Gal/UDP-Glc at elevated temperatures, with an apparent optimal temperature and pH of 80°C/7.0, respectively. Our data showed that TM0509 is a UDP-galactosugar 4-epimerase involved in D-galactose metabolism; consequently, this study provides the first detailed characterization of a hyperthermophilic GalE. Moreover, the promiscuous substrate specificity of TMGalE, which is more similar to human GalE than E. coli GalE, supports the notion that TMGalE might exhibit the earliest form of sugar-epimerizing enzymes in the evolution of galactose metabolism.

1. Introduction

Carbohydrates are essential for all forms of life. Many types of carbohydrate structures play vital roles in biological systems as source of energy, structural elements, molecular recognition markers, and precursors for the biosynthesis of other molecules. In nature, many sugar-decorating and -modifying enzymes (e.g., epimerases, isomerases, glycohydrolases, glycosyltransferases, oxidoreductases, and dehydrogenase) are involved in carbohydrate metabolism. Epimerases, which are widespread among animals, plants, and microorganisms, can catalyze the inversion of the configuration of asymmetrically substituted carbon in linear or cyclic molecules such as carbohydrates. These enzymes are responsible for removing hydrogen from one face of a central carbon and returning it to the opposite face. Although simple to draw, the chemistry behind such a transformation is complex because carbohydrates are extremely stable, and epimerization does not occur spontaneously except mutarotation. During the past decades, a variety of carbohydrate epimerases, such as UDP-α-D-glucose 4-epimerase (GalE) [1], D-ribulose-5-phosphate 3-epimerase [2], UDP-N-acetylgalactosamine 2-epimerase [3], L-ribulose 4-phosphate 4-epimerase [4], and D-galactose mutarotase [5], have been intensively studied to unravel their epimerization reactions.
Epimerization reactions are categorized according to the reaction mechanism [6]: transient keto intermediate, proton abstraction/addition, nucleotide elimination/readdition, carbon–carbon bond cleavage, and mutarotation. These studies suggested that a variety of cofactors and/or coenzymes are involved in the sugar epimerases.

In living organisms (animals, plants, and bacteria), galactosyl groups can be utilized for the biosynthesis of complex carbohydrates such as glycoproteins and lipopolysaccharides [7,8], as well as for catabolism of D-galactose (Gal) through the Leloir pathway [9]. UDP-α-D-galactose 4-epimerase (EC 5.1.3.2, GalE), which catalyzes the interconversion of UDP-α-D-glucose (UDP-Glc) and UDP-α-D-galactose (UDP-Gal), is essential for de novo biosynthesis of UDP-Gal in the Leloir pathway, by which β-D-galactose is converted to α-D-glucose 1-phosphate (Glc-1-P) to allow catabolism of β-galactose [9–11]. First, β-D-galactose is produced from β-D-galactose by galactokinase (GalK). The third enzyme, Gal-1-P uridyltransferase (GalT), catalyzes the transfer of an uridyl group from UDP-Glc to Gal-1-P. Accordingly, these enzymes are required for the interconversion of α-galactose moiety to β-D-glucose moiety in cells. In the absence of GalT or GalE, the accumulation of intermediary metabolites such as Gal-1-P and UDP-Gal [13], or the depletion of pyrimidine nucleotides [14] could cause the dys-regulation of robust metabolism in bacteria. In humans, impaired galactose metabolism, ascribed to missense mutations in either GalT or GalE, results in galactosemia, a serious metabolic disease [10,16]. Moreover, GalE deficiency exacerbates β-D-galactose toxicity in plants [17] and yeast [18]. In light of this, GalE is of great importance, not only for understanding carbohydrate metabolism in various organisms, but also for its physiological relevance to diseases and cellular phenotypes.

During the last decade, genome sequences of hyperthermophiles have revealed that multiple carbohydrate metabolic pathways are also available to hyperthermophilic bacteria and archaea [19]. These genomic data shed light on metabolism under extreme environments, indicating that sugar catabolism in hyperthermophiles is different relative to those of other domains of life. In this regard, functional annotation and characterization of gene products derived from hyperthermophiles are essential for investigation and understanding of early forms of life and the molecular evolution of metabolic enzymes [20]. The order Thermotogales, consisting mainly of anaerobically fermenting extremophilic bacteria, can grow on a range of both simple and complex carbohydrates, including glucose, starch, cellobiose, xylan, and pectin [21–23]. The genome sequence of Thermotoga maritima, which has an optimal growth temperature of 80 °C [24], revealed that many genes in this organism (10–15%) are involved in carbohydrate metabolism [25–27]. In addition, about half of the protein coding genes of T. maritima are highly conserved relative to bacterial genes, whereas a quarter of genes are more similar to archaeal genes, providing evidence for lateral gene transfer [28]. Moreover, sugar-nucleotide C4 epimerase is of special interest because it is a key enzyme in broad-spectrum microbial carbohydrate utilization. Sugar-nucleotide C4 epimerases can be classified into three groups based on their substrate specificity [29,30]: group 1 (e.g., Trypanosoma brucei GalE) is highly specific for non-acetylated sugars; group 2 (e.g., human GalE) can epimerize acetylated and non-acetylated sugars equally well; and group 3 (e.g., WbpP from Pseudomonas aeruginosa) exhibits a strong preference for acetylated sugars. Intriguingly, sequence alignments of the TM0509 gene encoding a putative sugar-nucleotide C4 epimerase from T. maritima with galE genes derived from a wide range of prokaryotes and eukaryotes demonstrated that although the sequence of the TM0509 gene is more similar to those of prokaryotic homologues, its substrate specificity is more similar to those of its eukaryotic homologues [31]. Although GalEs are widely distributed across bacteria and eukaryotes, little information is available regarding hyperthermophilic GalEs.

In this study, we overexpressed the TM0509 gene encoding the T. maritima GalE (TmGalE) in Escherichia coli, characterized the enzyme, and determined its crystal structure. In addition, we discuss the molecular evolution of GalE in the context of β-galactose metabolism under extreme environments, as well as the functional annotation of the TM0509 gene with respect to substrate specificity.

2. Materials and methods

2.1. Materials

Restriction enzymes, PrimeSTAR DNA polymerase, deoxy-nucleotide triphosphates, and chemicals for PCR were obtained from TaKaRa Biomedicals (Takara Co., Shiga, Japan). The pTOP Blunt V2 vector for cloning was obtained from Enzymonics (Daejeon, Korea). The pET-15b- and pET-22b expression vectors and the His-Bind Resin kit were obtained from Novagen (Madison, WI, USA). The pBAD-hisA expression vector was obtained from Invitrogen (Carlsbad, CA, USA). Genomic-tip, gel extraction, PCR purification, and plasmid miniprep kit were obtained from Qiagen (Hilden, Germany). Electrophoresis reagents were obtained from Bio-rad (Hercules, CA, USA). All chemicals used for enzyme assays and characterization (e.g., UDP-glucose, UDP-galactose, UDP-N-acetyl-glucosamine, UDP-N-acetylgalactosamine, and nictinamide adenine dinucleotide) were purchased from Sigma (St.Louis, MO, USA).

2.2. Bacterial strains and plasmids

E. coli strains, plasmids, and primers used in this study are listed in Table S1. Cultures were grown in Luria-Bertani (LB) and M9 media at 37 °C with appropriate concentrations of antibiotics (kanamycin at 50 μg/ml and ampicillin at 100 μg/ml, respectively). E. coli strains DH5α and BL21 (DE3) were used as bacterial hosts for recombinant plasmids. For complementation experiments, E. coli BW25113 ΔgalE Keio strain was used as a host and grown in minimal M9 medium, supplemented with 0.5% β-D-galactose as the sole carbon source and 0.2% L-arabinose as an inducer. Growth was monitored by determining the absorbance at 600 nm with an Ultratrace spectrophotometer (GE Healthcare). The pTOP Blunt V2 vector was used for cloning and sequencing, and the pET-15b(+), pET-22b(+), and pBAD-hisA were used for expression.

2.3. Cloning and expression of the TM0509 gene

A search of the microbial genome sequences in GenBank revealed a galE homologue (TM0509) in T. maritima encoding a putative UDP-glucose 4-epimerase (GalE). The galE gene was amplified by PCR by using genomic DNA as the template. The PCR mixture (total volume, 50 μl) contained 20 ng of genomic DNA, 10 pmol of primer TM G4E-NdeI-F, 10 pmol of primer TM G4E-XholR (Table S1), 1× PCR buffer, 0.2 mM dNTP mix, and 2.5 U of PfuTurbo polymerase (Takara). The PCR product was cloned into pTOP Blunt V2, and the resultant construct was transformed into E. coli DH5α competent cells. Transformants containing pTOP Blunt V2 harboring the gene encoding T. maritima GalE (TM0509) were
selected on LB medium-ampicillin plates.

Plasmid DNA was isolated from the transformants with inserts and digested with NdeI and Xhol. The digested DNA was purified and ligated into the Ndel and Xhol sites of pET-15b, yielding pET-15b-TM0509. For expression of the recombinant enzyme, *E. coli* BL21 (DE3) cells transformed with pET-15b-TM0509 were grown in LB medium (1.0 L) containing 100 μg/ml of ampicillin at 37 °C to an optical density at 600 nm of 0.5–0.6. After induction by 1 mM isopropyl-β-D-thiogalactoside (IPTG, IPTG), the cells were grown for an additional 6 h and harvested by centrifugation (10,000 × g, 20 min, 4 °C). Cells expressed were analyzed by sodium dodecyl sulfate (SDS)-12% polyacrylamide gel electrophoresis (PAGE) and visualized by staining with Coomassie blue. The harvested cells were stored at −80 °C prior to use.

2.4. Purification and characterization of TMGalE

The centrifuged cells were resuspended in 50 ml of 1× Histag binding buffer (20 mM Tris–HCl, 500 mM NaCl, and 5 mM imidazole; pH 7.9) and disrupted by sonication. The lysate was centrifuged at 10,000 × g for 20 min to remove cell debris, and the supernatant was heated at 70 °C for 20 min. Next, the suspension was centrifuged at 10,000 × g for 20 min to remove denatured *E. coli* protein, and the soluble fraction was filtered (pore size, 0.45 μm). The filtrate was loaded on a His-Bind Resin column (10 ml, Nova- gen) equilibrated with the same buffer. The column was washed with ten column volumes of the washing buffer (20 mM Tris–HCl, 500 mM NaCl, 60 mM imidazole; pH 7.9), and then 250 mM imidazole was applied to elute the recombinant protein. The eluent was then subjected to thrombin digestion to remove the N-terminal His tag. Prior to protease digestion, the buffer was exchanged by dialysis against 20 mM Tris–HCl (pH 7.5) containing 150 mM NaCl. Thrombin (Novagen, San Diego, CA, USA) digestion was performed by incubating ten units of thrombin with 10 mg of fusion protein and 2.5 mM CaCl₂ at 23 °C, resulting in full cleavage within 24 h. Further purification was carried out on a HiLoad 16/600 Superdex S200 column (GE Healthcare) in 20 mM Tris–HCl (pH 7.0) containing 150 mM NaCl. The fractions containing enzyme were pooled and dialyzed against 20 mM Tris–HCl buffer (pH 7.0), and the dialyzed enzyme preparation was stored at 4 °C prior to use. Protein concentrations were determined by the bicinchoninic acid (BCA) method [32], with bovine serum albumin as a standard. Enzyme fractions were analyzed by SDS-PAGE and visualized with Coomassie blue.

2.5. Enzyme assay

TMGalE activity was determined in a reaction mixture (200 μl) containing 50 mM potassium phosphate buffer (pH 7.0), 1 mM NAD⁺, 1 mM UDP-sugar and enzyme. After incubation at 80 °C for 10 min, the reaction was terminated by chilling on ice for 5 min. The mixture was centrifuged at 16,000 × g for 5 min, and a 20 μl aliquot was analyzed on a Waters Alliance HPLC system (model 2695; Waters, Milford, MA), equipped with a model 2487 Dual λ Absorbance Detector (Waters) and C18 column (Waters, SunFire™, 250 × 4.6 mm, 5 μm particle size). The column temperature was 37 °C. The mobile phase consisted of two buffers: (1) buffer A: 100 mM KH₂PO₄ containing 8 mM tetrabutylammonium hydrogen sulfate (adjusted to pH 5.3 with 2 M KOH); (2) buffer B: 70% buffer A and 30% methanol (pH 5.9). The column was eluted at 1.0 ml/min by a 0–20% (v/v) gradient of buffer B in buffer A under the following conditions: 0–15 min, linear gradient of 0–20% buffer B concomitant with 100-80% buffer A; 15–17 min, linear gradient of 20-0% buffer B; 17–30 min 100% buffer A. The reaction products were monitored by absorbance at 270 nm, and the amount of UDP-sugar produced was estimated from the peak area based on the curve of UDP-Glc (UDP-GlcNAc) or UDP-Gal (UDP-GalNAc) as the calibration standard. One unit of enzyme activity was defined as the amount of enzyme that converts the sugar moieties of 1 μmol of UDP-sugars into its 4-epimer per min at 80 °C. When various UDP-sugars were examined for their efficiency as substrates for the enzyme, each UDP-sugar produced was confirmed by comparison of its elution time, as follows: UDP-Glc, 20.3 min; UDP-Gal, 19.2 min; UDP-GlcNAc, 35.2 min; and UDP-GalNAc 34.0 min.

To confirm the UDP-Gal 4-epimerase activity of TMGalE, we also performed the UDP-Glc dehydrogenase (UDP-GlcDH) activity assay by coupling the 4-epimerization of UDP-Gal with the oxidation of UDP-Glc to UDP-glucuronate (UDP-GlcA) by UDP-GlcDH and measured the generated NADH absorbance at 340 nm. To obtain UDP-GlcDH, the tauD gene, encoding UDP-GlcDH, was cloned and expressed from Bacillus subtilis 168. Briefly, the gene was amplified from B. subtilis 168 genomic DNA using primers tauD-Ncol-F and tauD-Xhol-R (Table S1). The PCR products were digested with Ncol and Xhol, and then ligated into the same sites of pET22b, yielding pET22b-tauD. After overexpression in *E. coli* BL21(DE3), the recombinant enzyme was purified to homogeneity by Ni²⁺ affinity chromatography. Enzymatic reaction mixtures contained 0.2 ml of the resulting supernatants as substrates, 150 mM glycine/NaOH (pH 9.0), 5 mM NAD⁺, and 25 μl of UDP-GlcDH enzyme solution (~1 mg/ml) and were incubated at 37 °C for 20 min. The increase in A₃₄₀ due to NADH-formation was used to calculate the UDP-Glc content of the aliquot, using 6.22 × 10⁻³ M⁻¹ · cm⁻¹ as the extinction coefficient for NADH.

2.6. Biophysical and biochemical characteristics

To determine the effect of temperature on TMGalE activity, the enzyme (2 μg/ml) was incubated in 50 mM potassium phosphate buffer (pH 7.0) containing 1 mM NAD⁺ and 1 mM substrate (UDP-Gal) for 10 min at different temperatures ranging from 30 °C to 100 °C. For a thermostability test, the enzyme was incubated at various temperatures in the presence of 50 mM potassium phosphate (pH 7.0) for 3 h. Residual enzyme activity was measured under the standard assay conditions.

For pH studies, TMGalE activity was measured with UDP-Gal as a substrate as described above, except that potassium phosphate buffer was replaced by 50 mM sodium acetate-acetic acid (pH 4.0 to 6.0), 50 mM potassium phosphate (pH 6.0 to 8.0), and 50 mM glycine-NaOH (pH 8.0 to 10.0). All pH values were adjusted at room temperature, and the ΔpK₅/ΔTₕ for each buffer was taken into account when the results were analyzed. The residual enzyme activity was measured under the above assay conditions at 80 °C.

To investigate the effect of various metal ions on TMGalE activity, we first prepared apo-TMGalE by treatment with 10 mM EDTA at 60 °C for 2 h, followed by overnight dialysis against 20 mM Tris–HCl buffer (pH 7.0) at 4 °C with several changes of buffer. Removal of metals was confirmed by inductively coupled plasma mass spectrometry (ICP-MS). The effects of various metal ions were assessed by adding 1 mM CoCl₂, MnCl₂, MgCl₂, ZnCl₂, CaCl₂, CuCl₂, FeCl₃, or NiSO₄ to the dialyzed enzyme and assaying TMGalE activity under the standard assay conditions. The dependence of TMGalE activity on NAD⁺ concentration was determined by measuring enzyme activity in the presence of 0–10 mM NAD⁺ under the standard assay conditions. The substrate conversion was estimated using the HPLC analysis as described above. To determine the kinetic parameters, all enzymatic assays were performed in 50 mM potassium phosphate buffer (pH 7.0) containing 1 mM NAD⁺ and 0–5 mM substrate (UDP-Glc, UDP-Gal, UDP-GlcNAc or UDP-GalNAc) at 80 °C for 1 min. Kinetic parameters were determined by fitting data to the Michaelis–Menten equations using
2.7. Crystallization of TMGalE

Purified TMGalE in 20 mM Tris–HCl (pH 7.5) containing 50 mM NaCl was concentrated to 12 mg/ml by using a Vivaspin MWCO 10K (Sartorius) and used for crystallization trials using a 24-well crystallization plate (Hampton VDX) by the hanging drop vapor-diffusion method. Crystallization setups in each well (0.5 ml) consisted of a 1:1 (v/v) mixture of protein and crystallization solutions (1 μl–1 μl) on the Hampton siliconized cover slide. Initial crystals were obtained within 7 days from the Hampton Research Screen solution PEG/ION2 #12, containing 8% (v/v) Tacsimate (pH 5.0) and 20% polyethylene glycol 3350 (PEG3350). Crystal quality was improved by slightly increasing the concentration of PEG3350 (to 20% PEG3350, and 30% glycerol as a cryoprotectant. The UDP-Glc complex of TMGalE was prepared by soaking TMGalE crystals in the cryo-solution containing 8% (v/v) Tacsimate (pH 5.0) and 25% UDP-Glc have been deposited in the Protein Data Bank under ID codes 4ZRN and 4ZRM, respectively.

2.8. Data collection and structure determination

A native dataset was collected from TMGalE crystals (Table 1) using beamline 5C at the Pohang Light Source (PLS) (Pohang, Korea). The enzyme without bound sugar was flash-frozen in a cryo-solution containing 8% (v/v) Tacsimate (pH 5.0), 25% PEG3350, and 30% glycerol as a cryoprotectant. The UDP-Glc complex of TMGalE was prepared by soaking TMGalE crystals in the well solution containing 10 mM UDP-Glc for 1 h or overnight at 20 °C, and then frozen in the cryo-solution.

The coordinates for TMGalE structures with and without UDP-Glc have been deposited in the Protein Data Bank under ID codes 4ZRN and 4ZRM, respectively.

### Table 1

<table>
<thead>
<tr>
<th>Data collection</th>
<th>TMGalE</th>
<th>TMGalE with UDP-Glc</th>
</tr>
</thead>
<tbody>
<tr>
<td>Beam line</td>
<td>PAL-5C</td>
<td>PAL-5C</td>
</tr>
<tr>
<td>Space group</td>
<td>P2_12</td>
<td>P2_12</td>
</tr>
<tr>
<td>Wavelength (Å)</td>
<td>0.97949</td>
<td>0.97952</td>
</tr>
<tr>
<td>Unit-cell parameters (Å)</td>
<td>a = 82.54, b = 150.26, c = 60.61</td>
<td>a = 82.11, b = 149.94, c = 62.12</td>
</tr>
<tr>
<td>Resolution (Å)</td>
<td>50–1.90 (1.93–1.90)</td>
<td>50–2.00 (2.03–2.00)</td>
</tr>
<tr>
<td>Total reflections</td>
<td>797,474</td>
<td>398,270</td>
</tr>
<tr>
<td>Unique reflection</td>
<td>59,945</td>
<td>50,135</td>
</tr>
<tr>
<td>Completeness (%)</td>
<td>99 (96.3)</td>
<td>100 (100)</td>
</tr>
<tr>
<td>Redundancy</td>
<td>13.2 (13.2)</td>
<td>7.7 (7.7)</td>
</tr>
<tr>
<td>R-merge (%)</td>
<td>8.6 (54.5)</td>
<td>12.8 (50.4)</td>
</tr>
<tr>
<td>Average I/σ(I)</td>
<td>29.9 (8.3)</td>
<td>15.6 (4.9)</td>
</tr>
<tr>
<td>Solvent content (%)</td>
<td>54.34</td>
<td>54.39</td>
</tr>
<tr>
<td>Matthews coefficient (Å³ Da⁻¹)</td>
<td>2.69</td>
<td>2.7</td>
</tr>
</tbody>
</table>

**Refinement**

| Rwork/Rfree (%) | 15.77/18.47 | 16.43/19.65 |
| Protein residues | 615 | 617 |
| Waters | 272 | 299 |
| NAD | 2 | 2 |
| UDP-glucose | 2 | 2 |
| RMSD Angle (°) | 1.149 | 1.366 |
| Length (Å) | 0.007 | 0.008 |
| Average B-factors (Å²) | 16.3 | 28.6 |
| Ramachandran plot | | |
| Most favored regions (%) | 98.53 | 98.04 |
| Allowed regions (%) | 1.47 | 1.31 |
| Outliers | 0 | 0.65 |

Values in parentheses correspond to highest resolution shell.

* R-merge = Σhkl|<I(hkl)>−|Σhkl|I(hkl)|/Σhkl|I(hkl)|, where |I(hkl)| is the observed intensity and |<I(hkl)>| is the average intensity of symmetry-related observations.
T. maritima genomic DNA and cloned into the pTOP Blunt V2 cloning vector for sequencing. For expression in E. coli, as well as to facilitate subsequent purification, the TM0509 gene was cloned into the pET-15b expression vector. DNA sequence analysis of galE revealed an open reading frame of 930 bp, encoding a 309 amino acid protein with a calculated molecular weight of 34,899 and a theoretical pI of 5.72. The recombinant protein was successfully expressed as an N-terminal hexahistidine (6× His)-tagged fusion protein in E. coli BL21 (DE3) upon induction with IPTG. After sonication, E. coli lysates were heated at 70 °C to remove the majority of endogenous proteins while TMGalE remained soluble and active. The expressed TMGalE was purified by a Ni²⁺ chelate affinity chromatography step, followed by cleavage of the 6× His tag with thrombin. Thereafter, the sample was subjected to Superdex S200 gel filtration chromatography. The apparent Mr of the intact protein was estimated to be 35 kDa by SDS-PAGE, consistent with the value of Mr (34,899 Da) calculated from the presumptive amino acid sequence. Size-exclusion chromatography suggested that the native enzyme was a homodimer with a molecular mass of 70 kDa (Fig. 1A).

### 3.2. Effects of temperature and pH on TMGalE activity

The temperature dependence of the recombinant enzyme was determined in the presence of 1 mM NAD⁺ with 1 mM UDP-Gal as a substrate at various temperatures. TMGalE had the apparent optimum temperature of 80 °C (Fig. 1B). Remarkably, the catalytic activity at 30 °C was 53% of the maximal activity at 80 °C. As shown in Fig. 1C, no significant decrease in enzyme activity was observed over a 3 h incubation at 70 °C. Even after 3 h at 80 °C, TMGalE retained more than 85% of its original activity under standard assay conditions. Moreover, the enzyme had almost 50% of its original activity after 40 min of incubation at 90 °C, but its catalytic activity was lost completely after 2 h of incubation. These results demonstrated that TMGalE exhibits substantial thermostability, and has catalytic activity over a broad temperature range. As shown in Fig. 1D, the optimum pH at 80 °C was 7.0.

### 3.3. Effects of metal ions and cofactor NAD⁺ on TMGalE activity

Several sugar 4-epimerases require divalent metal ions as cofactors for catalytic activity [4,38–40]. To investigate the effect of divalent metal ions on TMGalE activity, we prepared apo-TMGalE and determined its activity in the presence of divalent metal ions such as Co²⁺, Mn²⁺, Mg²⁺, Zn²⁺, Ca²⁺, Cu²⁺, Fe²⁺, or Ni²⁺ under standard assay conditions. The apo form of TMGalE exhibited considerable activity (353.1 U/mg) that was not significantly increased by addition of any metal ions, demonstrating that the enzyme is not specifically dependent on divalent metal ions for its catalytic activity (Fig. S1A). Previously, uridine diphosphogalactose 4-epimerase was shown to be highly dependent on NAD⁺ as a cofactor in the epimerization reaction [41]. Hence, to investigate the effect of NAD⁺ on TMGalE activity, we performed enzyme assays under standard conditions in the presence of 1–10 mM NAD⁺. As shown in Fig. S1B, NAD⁺ did not appear to be a strong activator of enzyme activity because there was no significant difference in enzyme activity regardless of the presence and absence of NAD⁺.

Nevertheless, although purified TMGalE exhibited no absolute...
requirement for metals, we could not exclude the possibility that NAD⁺ remained tightly bound to the purified TMGalE, because no spontaneous exchange with free NAD⁺ took place, as in the case of partially purified UDP-Gal 4-epimerase [41].

3.4. In vivo functional and complementation analyses

To assess whether TMGalE can complement a strain lacking the epimerization activity of UDP-Gal, we transformed the pBAD vector harboring the TMGalE gene into E. coli BW25113 ΔgaLE (the control strain) [42], yielding TMGalE-E. coli ΔgaLE cells. We measured the absorbance at 600 nm of TMGalE-E. coli ΔgaLE cells grown on M9 minimal media supplemented with 5 g/l of D-galactose as the sole carbon source (Fig. 2A). As expected, BW25113 (the wild-type strain) grew in D-galactose-containing medium, whereas M9ΔgaLE mutant strains at 37°C exhibited significant growth on D-galactose-containing medium after a long lag phase, probably due to the lower activity of TMGalE at 37°C relative to its optimal temperature. This was not surprising, however; as noted above, the enzyme activity at 30°C was 53% of the maximum activity at 80°C, sufficient to support the growth of the ΔgaLE cells. In addition, we sequenced the plasmid to verify that it contained no specific mutation that altered the properties of the TMGalE expressed in E. coli. Thus, this result strongly suggested that TMGalE might possess UDP-Gal epimerization activity sufficient to support the growth of E. coli cells lacking GaLE responsible for bacterial viability in β-galactose metabolism.

3.5. Substrate specificity of TMGalE toward UDP-sugars

To study the substrate specificity of TMGalE, we used HPLC to measure its activity toward UDP-Glc, UDP-Gal, UDP-GlcNAc, and UDP-GalNac under standard assay conditions. As shown in Fig. 2C, the HPLC data revealed that TMGalE could catalyze the epimerization of UDP-Gal and UDP-Glc, as expected. Of the two substrates, TMGalE had higher activity toward UDP-Gal with the 1:3 conversion ratio of UDP-Gal to UDP-Glc at equilibrium (Table 2). Moreover, TMGalE also had an unusually high activity for epimerization of UDP-GalNAc to UDP-GlcNAc. Next, to determine whether UDP-Glc was produced from UDP-Gal by TMGalE, we carried out a 4-epimerization reaction with 1 mM UDP-Gal as the substrate for TMGalE, followed by the oxidation of UDP-Glc to UDP-glucuronate (UDP-GlcA) by UDP-GlcDH derived from B. subtilis 168. These reactions were run in the presence of NAD⁺ as a cofactor, as described in Ref. [41].

Furthermore, in order to further characterize the substrate specificity of TMGalE, we obtained the kinetic parameters for UDP-Gal/Glc and UDP-GalNAc/GlcNAc as substrates at 80°C and pH 7 using Michaelis–Menten kinetics (Table 3 and Fig. 2C). The catalytic efficiency (kcat/Km) for UDP-Gal (1,406 mM⁻¹ s⁻¹) was approximately 2.1 times higher than that for UDP-Glc (1134 mM⁻¹ s⁻¹), indicating that this enzyme might have a preference for UDP-Gal over UDP-Glc. Notably, the catalytic efficiencies of TMGalE for UDP-GalNAc and UDP-GlcNAc were approximately 25-fold and 10-fold higher than those for UDP-Gal and UDP-Glc, respectively, indicating that TMGalE might have broader (i.e., more promiscuous) substrate specificity than E. coli GaLE (ECGaLE) [43]. Therefore, although a direct comparison is difficult, TMGalE appears to be

![Fig. 2. Comparison of the growth of E. coli BW25113 as the wild-type strain and ΔgaLE mutant strains at 37°C in minimal M9 liquid (A) and solid (B) media. (A and B) M9ΔgaLE cells did not grow at all. However, TMGalE-E. coli ΔgaLE cells exhibited significant growth on D-galactose-containing media after a long lag phase, probably due to the lower activity of TMGalE at 37°C relative to its optimal temperature. (C) The reaction mixtures (200 μl) contained 0.2 μg/ml of purified enzyme, 1 mM UDP-sugars (UDP-Glc, UDP-Gal, UDP-GlcNAc, and UDP-GalNAc), 1 mM NAD⁺, and 50 mM potassium phosphate (pH 7.0). Aliquots were withdrawn after 60 min of incubation at 80°C. Samples (20 μl) were analyzed by HPLC analysis with a C18 column (Waters, SunFire™).

<table>
<thead>
<tr>
<th>Substrate</th>
<th>Enzyme activity* (U/mg of protein)</th>
<th>Equilibrium ratio⁴</th>
<th>Proportion (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>UDP-Gal</td>
<td>192.4 ± 2.1</td>
<td>100</td>
<td>26.2 ± 0.1</td>
</tr>
<tr>
<td>UDP-Glc</td>
<td>69.7 ± 2.1</td>
<td>36.2</td>
<td>74.5 ± 0.1</td>
</tr>
<tr>
<td>UDP-GalNAc</td>
<td>93.7 ± 3.3</td>
<td>48.5</td>
<td>30.1 ± 0.1</td>
</tr>
<tr>
<td>UDP-GlcNAc</td>
<td>48.7 ± 0.1</td>
<td>25.3</td>
<td>72.9 ± 0.1</td>
</tr>
</tbody>
</table>

* Activities of recombinant enzymes toward various UDP-sugars were determined under standard conditions. Activities are compared with activity toward UDP-Gal and expressed as percentages.

Table 3

<table>
<thead>
<tr>
<th>Substrate</th>
<th>Vmax (U/mg)</th>
<th>Km (mM)</th>
<th>kcat (s⁻¹)⁵</th>
<th>kcat/Km (mM⁻¹ s⁻¹)</th>
</tr>
</thead>
<tbody>
<tr>
<td>UDP-Glc</td>
<td>25,234 ± 1400 *</td>
<td>12.9 ± 0.1</td>
<td>14,677</td>
<td>1134</td>
</tr>
<tr>
<td>UDP-Gal</td>
<td>13,337 ± 500</td>
<td>5.53 ± 0.3</td>
<td>7757</td>
<td>1406</td>
</tr>
<tr>
<td>UDP-GalNAc</td>
<td>47,574 ± 4000</td>
<td>2.40 ± 0.1</td>
<td>27,671</td>
<td>11,465</td>
</tr>
<tr>
<td>UDP-GlcNAc</td>
<td>65,839 ± 5770</td>
<td>1.06 ± 0.1</td>
<td>38,295</td>
<td>36,120</td>
</tr>
</tbody>
</table>

Data are means ± standard deviations from duplicate independent experiments.

Kinetic parameters were obtained by fitting the experimental data to the Michaelis–Menten equation, as follows: v = Vmax[S]/(Km + [S]).

The data are means ± standard deviations.

kcat is the number of substrate molecules reacted per active site per sec.
closer to Human GalE (HUGalE) than bacterial ECGalE with respect to substrate specificity.

3.6. Overall structure of TMGalE

Previously, the crystal structure of ECGalE with bound UDP-benzene revealed a Rossmann fold as a specific regulatory site for pyrimidine nucleotide binding distinct from the catalytic site [44]. Analogously, we determined the X-ray crystal structures of TMGalE with and without bound UDP-Glc as a substrate at 2.0 Å and 1.9 Å, respectively, and its NAD cofactor was also well-defined. TMGalE, the asymmetric unit is formed by two molecules consistent with the form of the fully active dimer in solution (Fig. 3). The interface analysis by PDBePISA [45] indicated that TMGalE formed a dimer in the asymmetric unit with total buried surface areas of 1283.5 Å² for TMGalE and 1286.8 Å² for TMGalE bound to UDP-Glc, respectively.

When compared to ECGalE and HUGalE, TMGalE superimposed with these mesophilic counterparts with the RMSD values of 1.385 Å and 1.719 Å, respectively (Fig. S3A). As shown in Fig. 3, each subunit of TMGalE consists of an N-terminal domain with NAD⁺ bound and a C-terminal domain resembling that of ECGalE [46]. The N-terminal domain (Met1 to Tyr174) bound to NAD⁺ consists of seven strands of parallel β-pleated sheet flanked on either side by a total of seven α-helices arranged in a modified “Rossmann-type” fold, which is generally found in dinucleotide-binding enzymes. Moreover, the characteristic YXXXX motif (Tyr143 to Lys147) of the superfamily was also found in the N-terminal domain of TMGalE. In addition, the characteristic GGXXGX motif of the short-chain dehydrogenase/reductase (SDR) family members was also found in the β1-loop-α1 stretch of TMGalE. This cofactor-binding site is highly conserved in the prokaryotic and eukaryotic homologues as shown in Fig. S3. According to the domain organization described above, TMGalE seems to belong to the SDR superfamily. Meanwhile, the C-terminal region (Gly175 to Glu309) of the subunit, consisting of four β-strands and five α-helices, is associated with UDP-sugar binding (Fig. 3). Functional dimeric TMGalE contains one tightly

![Fig. 3. Overall structure of TMGalE.](image-url)
bound NAD\(^+\) per subunit, and the active site of each subunit is formed at the interface in between the two domains, which coincides with ECGalE [47] and HUGalE [31].

3.7. Cofactor- and substrate-binding sites of TMGalE

The electron density map corresponding to the NAD\(^+\) and UDP-Glc is clearly shown in Fig. 4. The NAD cofactor is coordinated with eleven residues (Phe11, Ile12, Asp31, Ser34, Ser35, Glu53, Leu73, Thr92, Tyr143, Lys147, and Val173) in the N-terminal domain, and superimposes closely on the corresponding structures of ECGalE and HUGalE (Fig. S3). NAD\(^+\) is firmly anchored via twelve hydrogen (H) bonds (Fig. 4A). Phe11, Ile12, Leu73, and Val173 are involved in sugar binding via Cα-carbonyl or amide group-mediated H-bonding. As shown in Fig. S4, there are, however, a number of differences in the NAD\(^+\)-binding cleft of TMGalE when compared to that of ECGalE; Phe11, Ser35, Glu53, Leu73, Ala77, Thr117, Ala171, and Val173. Among those residues, both Glu53 and Val173, corresponding to Arg60 (Leu68) and Pro180 (Pro188) in ECGalE (HUGalE), respectively, are H-bonded to the NAD\(^+\) cofactor only in TMGalE. As in HUGalE and ECGalE, the nicotine amide ring of NADH adopts the syn conformation [48], indicating a proper hydride transfer between C4 of sugar and the dinucleotide cofactor.

A close-up view of the active site bound to UDP-Glc reveals that the side chains of Thr117, Asn172, Arg210, Arg267, and Asp270 are involved in substrate binding via H-bonding (Fig. 4B). Additionally, six water molecules at H-bonding distance (~3.1 Å) interact with the sugar substrate. In the case of Val186, His201, Phe203, Arg267, Asp270, and Arg210, significant Cα-carbonyl or amide group-mediated bonding was involved in sugar binding. During the entry of substrate into the active site, four water molecules were replaced with substrate, and another four water molecules filled the space around the substrate. Notably, the substrate binding site-separator residue Glu183 was pushed ~7.96 Å away (between α-carbons) and replaced with the sugar moiety of the substrate. Unlike the positioning of ligand from HUGalE [31], the sugar positioning was not further stabilized via H-bonding between Tyr143 and C3 of the sugar ring. Additionally, in contrast to other GalEs including HUGalE and ECGalE, the sugar moiety in TMGalE positioned rotating to 180° around a C1-UDP bond, which C3 of sugar superpose C4 of sugar in HUGalE and ECGalE structures. This keep the distance by 4.15 Å between C4 of sugar and C4 nicotinamide ring.

3.8. Conformational changes in the active site of TMGalE

Structural changes occur between the C-terminal regions of TMGalE, HUGalE, and ECGalE with and without bound sugar, even though the ligand-bound conformations of those three structures are similar, as stated above (Fig. 5A). The RMSDs between forms of proteins with and without bound sugar were 0.308 Å for TMGalE and 0.722 Å for HUGalE. Remarkably, superimposition of those structures showed that the marginal extent of deviation seemed to be mainly due to the substrate-binding domain. A detailed inspection of TMGalE structures with and without bound sugar revealed more distinct features of the dinucleotide-binding pocket in the active site than in other GalEs (Fig. 5A). Intriguingly, the distance between Tyr157 and Asn206 in HUGalE without bound sugar is ~5.99 Å [49], whereas the corresponding residues (i.e., Tyr143 and Glu183) in TMGalE forms H-bond (~2.82 Å) between

![Fig. 4. Stereoviews of the ligand binding sites of TMGalE in complex with NAD\(^+\) (A) and UDP-Glc (B). The omit difference maps (mFo-DFc) around ligands contoured at 2.0 \(\sigma\) are depicted as gray lines, and water molecules are illustrated as limon spheres. Side chains of conserved (white) and variable (purple) residues that are involved in ligand coordination are represented as sticks. Ligands and side chains are colored by atom type (oxygen in red, nitrogen in blue, phosphorous in orange, and carbons of NAD\(^+\) and UDP-Glc in white and yellow, respectively). Hydrogen bonds are shown as dashed lines in black.](#)
them, indicating that a clear partitioning two ligand binding sites for NAD\(^+\) and UDP-Glc can be observed in the TMGalE. In addition, Tyr143 and Glu183 in TMGalE without bound sugar are both H-bonded (\(\sim 2.82\) Å) to the carboxamide group of NAD\(^+\). However, Glu183 in the loop between \(\beta 7\) and \(\alpha 8\) of TMGalE is displaced \(\sim 13.91\) Å away from the carboxamide group on substrate (UDP-Glc) binding, suggesting that such a large displacement of the loop is likely to increase the cavity size of TMGalE. This may contribute to its broad substrate specificity. Nevertheless, no significant conformational changes of secondary structures around the active site occurred in response to substrate binding, except in the loop (Tyr174 to Val186) containing Glu183 (Fig. 5A). This might indicate the less flexible nature of chemical binding pocket in the case of hyperthermophilic GalEs. Comparison of the conformational changes related to UDP-Glc binding revealed distinctive differences in the C-terminal region of the two proteins: major changes occurred in multiple regions in HUGalE and in only one region in TMGalE (Fig. 5B). Notably, two regions in HUGalE underwent significant conformational changes when ligand bound [31]. First, the \(\beta 8\)-strand (Ile217 to Phe226) moved toward the substrate-binding pocket by \(\sim 3.88\) Å as a distance of Phe226. Secondly, the loop (Val297 to Ala306) also dislocated from the active site by \(\sim 4.97\) Å. On the other hand, when UDP-Glc bound in TMGalE, Arg267 in the loop (Lys264 to Lys273) is closer to the active site bound to ligand. This observation strongly suggests that the least conformational change in the active site of TMGalE is commensurate with the larger changes occurring in those of mesophilic counterparts.

### 3.9. Thermostability

Hyperthermophilic enzymes require proper maintenance of its overall architecture at elevated temperatures during catalysis. Such thermal stability can be ascribed to several factors such as amino acid composition, surface ion pairs, hydrophobic core residues, water-mediated H-bond, and so on [50,51]. One may speculate that two factors including ion pairs and H-bonds may contribute to thermal stability amongst these enzymes. Indeed, comparison of the amino acid composition of TMGalE with those of other homologues revealed that charged amino acid content in hyperthermophilic GalEs is higher than their mesophilic counterparts (Table S3). Unlike mesophilic GalEs, hyperthermophilic GalEs exhibited several H-bonds and salt bridges by charged amino acids located inside or peripheral region of the active site, which may contribute to thermostability at elevated temperatures. Although GalE from the hyperthermophilic archaean *Pyrobaculum calidifontis* [52] appears to be structurally more similar to ECGalE and HUGalE rather than TMGalE in terms of RMSD values, such additional interactions may contribute to the integrity of the substrate binding site of hyperthermophilic GaE at elevated temperatures.

In addition, the residues that interact directly (or indirectly via water molecules) with substrate binding ligands are well-correlated among those structures (Fig. 6). Inspection of TMGalE (PDB 4ZRN) and HUGalE (PDB 1EK6) bound to UDP-Glc revealed that there is a notable difference in the number and location of water molecules in the active sites of GalEs (Figs. 6 and S5). In
TMGalE with and without bound UDP-Glc, totally six water molecules were placed in their substrate binding pockets (Fig. 6A and B). However, binding of UDP-Glc retained only two water molecules (W1 and W3), whereas four new water molecules are coordinated to interact with the substrate in the active site, resulting in the stabilization of overall active site architecture (Figs. 6 and S5). By contrast, binding of UDP-Glc to the active site of HUGalE significantly altered the coordination of water molecules as well as the number of water molecules in the active site (Fig. 6C and D). In particular, two additional water molecules seems to be involved in coordination of substrate binding with pre-existing water ones in maintaining the volume of the active site constant between with and without bound sugar, implying the role of peripheral water molecules associated with the integrity of the substrate binding site of hyperthermophilic GalEs.

Further, detailed structural comparison between TMGalE and HUGalE shed light on the residues potentially involved in conformational stability. In TMGalE without bound sugar, Arg210 tightly interacts with Asp270, and both residues are also connected via a water molecule (W1). In addition, Asp270 is H-bonded to W2, which, in turn, interacts with W3 and the backbone carbonyl oxygen of Phe203 (Fig. 6). Similarly, GalE from *T. thermophilus* HB8 (2PSY) also exhibits a salt bridge (Arg215-Asp274) and water-mediated H-bonds. However, such serial bonding connections are not observed in HUGalE in the absence of sugar, indicating that the active site of TMGalE contains three more H-bonds and one more salt bridge than HUGalE without bound sugar in the C-terminal region. In terms of thermodynamics, an individual H-bond and salt bridge contribute an average of 5.5 kJ/mol (−1.3 kcal/mol) and 22 kJ/mol (−5.2 kcal/mol) to stabilization, respectively [53]. Even though TMGalE without bound sugar contains three more water molecules than HUGalE (Fig. 6), it is likely that three water molecules (W1, 2, and 3) participating in the formation of serial H-bonds are properly positioned to stabilize the conformation, as well as prohibit drastic changes in active-site conformation at elevated temperatures, by tightening the loop (between β8 and α7) and the β8-strand (Fig. 5A). Indeed, the corresponding scaffold of HUGalE underwent significant changes due to the lack of waters in the positions (Figs. 6 and S5).

Therefore, structural comparison of TMGalE with mesophilic counterparts suggested that higher contents of charged residues responsible for salt-bridges and water molecules for H-bonds in the active site of TMGalE might play a significant role in its thermostability at elevated temperatures.

### 4. Discussion

Analysis of the genome of hyperthermophilic *T. maritima* revealed that this heterotrophic strain has extensive and highly diversified machineries for carbohydrate utilization [28,54]. This feature was also analyzed by an integrated approach for functional mapping of diverse sugar kinomes, indicating that a large variety of sugar catabolic pathways exist in *T. maritima*. However, to identify the biochemical functions of unknown proteins and improve the quality of reconstructed pathways, both bioinformatics analysis and biochemical characterization are still essential for functional gene annotation. Among the annotated genes responsible for α-galactose metabolism, the TM0509 gene was predicted to encode...
UDP-glucose 4-epimerase (GalE).

First, to characterize this putative GalE protein, we expressed the TM0509 gene and purified the encoded protein for further analysis. The recombinant protein is a homodimer with maximal catalytic activity at 80 °C and pH 7 under our assay conditions (Fig. 1). Notably, although TMGalE originated from the hyperthermophilic bacterium, it exhibited pronounced activity, corresponding to about 50% of its maximal epimerization activity, at suboptimal temperature conditions, as is also the case for other proteins from *T. maritima* [54]. This characteristic allowed us to carry out complementation experiments using mesophilic *E. coli* cells. Indeed, TMGalE expressed in *E. coli* cells lacking the galE gene could complement the mutant cells to grow on M9 minimal media supplemented with 0.5% D-galactose as the sole carbon source (Fig. 2A and B). Taken together, our data strongly suggest that the TM0509 gene may encode a sugar-nucleotide C4 epimerase specific for the D-galactose sugar moiety.

Secondly, we investigated the substrate specificity of TMGalE and compared its amino acid sequence with those of other GalE homologues. In comparison to others, the TM0509 gene has relatively low levels of amino acid sequence identity (similarity): 30.1% (44.2%) relative to ECGalE [46], 26.4% (40.7%) to HUGalE [31], 26.9% (39.2%) to *T. brucei* GalE (TBGalE) [55], and 31.6% (47.5%) to *P. aeruginosa* GalE (PAGalE) [30] (Table S2), indicating that the amino acid sequences of these enzymes are not highly conserved across all forms of life. Indeed, the substrate specificities of these proteins are relatively variable. Notably, unlike the bacterial enzyme, HUGalE appears to be a bifunctional enzyme specific for UDP-GlcNAc/UDP-GalNAc, as well as UDP-Glc/UDP-Gal. Such bifunctional epimerases were described in a variety of other organisms [49,56–58]. As described in Tables 2 and 3, TM0509 mainly catalyzed the interconversion of UDP-Gal and UDP-Glc, with a preference for UDP-Gal over UDP-Glc. However, it also exhibited significant activity toward bulkier substrates (i.e., UDP-GalNAc and UDP-GlcNAc), indicating that TM0509 encodes UDP-Gal 4-epimerase with broad substrate specificity more similar to that of HUGalE than other microbial GalEs. This distinct feature is supported by comparing the cavity size of GalE homologues (Fig. S4).

Third, to study the evolutionary relationship between TMGalEs and other orthologues derived from various organisms, we constructed a phylogenetic tree based on their amino acid sequences (Fig. 7). Sugar-nucleotide C4 epimerases can be classified into three

---

**Fig. 7.** Dendrogram for the UDP-hexose 4-epimerase family. An unrooted radial phylogram was drawn from a multiple sequence alignment of 41 GalE homologues including TMGalE. Alignment was performed using ClustalW and the radial-style tree was drawn using MEGA6. Predicted substrate spectrum profiles were superimposed on this tree and color-coded as follows: group 1 (black), group 2 (gray), and group 3 (light gray). The scale bar represents 0.1 substitutions per amino acid position.
groups based on their substrate specificities [29,30]: group 1, including ECGalE derived from E. coli K12, prefers non-acetylated UDP-sugar substrates [44]; group 2, represented by HUGalE, can epimerize both acetylated and non-acetylated substrates [49]; and group 3, including PAGalE, exhibits a strong preference for acetylated substrates [29,56,59]. In this regard, TMGalE appears to be most closely related to group 2 GalE with respect to substrate specificity. However, as shown in Fig. 7, this phylogenetic analysis indicated that TMGalE belongs to group 3, which preferentially converts between UDP-GlcNAc and UDP-GalNAc. This finding is not consistent with the previous conclusion described in Tables 2 and 3, and seemingly indicates that TMGalE should be placed on a major divergence between groups 2 and 3. This discrepancy led us to question why TMGalE exhibits such unusual promiscuity for substrate preference relative to other bacterial enzymes. Prior to answering this question, we first attempted to determine the 3D structure of TMGalE in order to compare it with those of HUGalE (group 2) and PAGalE (group 3).

Previous work showed that the size and shape of the GalE active site varies across species, allowing for variable substrate specificity [30,60]. As shown in Fig. 54, three critical binding sites for the saccharide moiety are significantly modified in TMGalE relative to other homologues (Figs. S6–S8). Previously, it was proposed that Gly102 in PAGalE (corresponding to Ala77 in TMGalE) might play an important role in the binding of bulkier substrates by enlarging the sugar binding cavity to reside in Ref. [30]. This might be also the case for Ala77 in TMGalE that showed significant activity for UDP-Glc/Gal-Nac. Indeed, the corresponding residue in group 3 GalE is either Ala or Gly, rather than a bulkier amino acid residue. In addition, Ala184 in TMGalE (Ala209 in PAGalE), which can bind bulkier sugars, is replaced by either His in Group 1 (e.g., TBGalE) or Asn in Group 2 (e.g., HUGalE), consistent with the results described in Ref. [30]. Lastly, the Leu residue (e.g., Leu342 in TBGalE) that is highly conserved in group 1 is replaced by the Ser274 in TMGalE (Cys307 in HUGalE and Tyr299 in ECGalE), which functions as a gatekeeper to control the size of the active cleft involved in sugar binding [30,60]. Therefore, overall comparison of these structures clearly indicated that TMGalE might possess a larger binding niche for a bulkier saccharide moiety, as in the case of PAGalE in group 3. However, the biochemical and kinetic data regarding substrate specificity strongly suggested that TMGalE might be closely related to GalE homologues in group 2 (Tables 2 and 3).

Fourth, to understand the molecular basis of promiscuity of substrate specificity for TMGalE, we compared the crystal structures of TMGalE with and without bound sugar to those of other GalE’s (Figs. S3 and S4). The crystal structures of HUGalE bound to UDP-Glc and UDP-GlcNAc demonstrate that the difference in the volume of the active site in HUGalE may be ascribed primarily to Asn207 and Cys307. Indeed, simple rotation of the side chain of Asn207 could increase the volume of one side of the pocket up to ~15% larger than that of ECGalE, reflecting the ability of HUGalE to accommodate the N-acetyl group at the C2 position of glucose [49]. Moreover, the Y299C mutation in ECGalE altered its substrate preference towards the bulkier UDP-GalNAc [43], indicating that Cys307 in HUGalE (Tyr299 in ECGalE) accounts for a 15% increase in the volume of the active-site cleft in HUGalE relative to ECGalE [60]. Remarkably, Ala184 and Ser274 of TMGalE (Asn207 and Cys307 in HUGalE, respectively) seem to function similarly to the analogous residues in HUGalE and ECGalE (Fig. S4D). Ala184 is located at the C-terminus of the extended loop (Gly175 to Val186), a little further away from the active site than the same region in HUGalE and ECGalE. This expanded space might facilitate the binding of bulkier sugar substrates, as supported by previous findings that the UDP group is bound strongly (as the binding anchor), whereas the hexopyranosyl group is bound more weakly [61–63]. The size of the

5. Conclusions

In this study, we overexpressed the TM0509 gene encoding the T. maritima GalE (TMGalE) in E. coli, characterized the enzyme, and determined its crystal structure. We confirmed that the TM0509 gene of T. maritima, which is annotated as a putative UDP-glucose 4-epimerase, exhibited UDP-galactose 4-epimerase activity with a high affinity for UDP-N-acetylgalactosamine (UDP-GalNAc), similar to human GalE. To further investigate the molecular basis of this enzyme’s catalytic function, we determined the crystal structures of TMGalE and TMGalE bound to UDP-Glc at resolutions of 1.9 Å and 2.0 Å, respectively. The recombinant enzyme was determined to be a homodimer with a molecular mass of 70 kDa. The enzyme could reversibly catalyze the epimerization of UDP-GalNAc/UDP-GlcNAc as well as UDP-Gal/UDP-Glc in the presence of NAD⁺ at elevated temperatures, with an apparent optimal temperature and pH of 80°C and 7.0, respectively.

Recently, the crystal structure of an archaeal GalE from P. calidivintos was determined [52]. In contrast to TMGalE, this hyperthermophilic archaeal enzyme exhibited catalytic activity only toward UDP-Gal/UDP-Glc. As described above, it is evident that the three domains of life underwent their own molecular evolution and/or adaptation after they diverged from the last universal common ancestor (LUCA); this is clearly reflected by the fact that homologous enzymes from the three domains exhibit a wide range of substrate specificities in different environments. Based on the genome context of T. maritima, which contains evidence of lateral gene transfer events [28], we speculate that TMGalE encoded by the TM0509 gene might have gained the ability to utilize a broad range of substrates via a trade-off involving a minimal change in active-site conformation, in order to maintain its integrity during substrate binding at elevated temperatures.

Acknowledgments

We thank the Beamline (PLS-5C and 7A) staff at the Pohang Light Source for technical support during data collection.

This work was supported by the National Research Foundation of Korea (NRF) grants (2014R1A2A2A01006765 & 2013R1A1A2057465) to Dr. Dong-Woo Lee and to Dr. Sung Haeng Lee, respectively, funded by the Korean government (MSIP), and by a Grant 311042-05-1-HD120 from the Korea Institute of Planning & Evaluation for Technology (iPET) funded by the Ministry for Food, Agriculture, Forestry and Fisheries to Dr. Dong-Woo Lee.

Appendix A. Supplementary data

Supplementary data related to this article can be found at http:// dx.doi.org/10.1016/j.abb.2015.08.025.

References
