Current Research in Biotechnology 4 (2022) 350–358

Contents lists available at ScienceDirect

Current Research in Biotechnology

journal homepage: www.elsevier.com/locate/crbiot

Exploration of archaeal nucleotide sugar epimerases unveils a new and highly promiscuous GDP-Gal4E subgroup

Carlos Alvarez Quispe, Matthieu Da Costa, Koen Beerens, Tom Desmet *

Centre for Synthetic Biology (CSB), Faculty of Bioscience Engineering, Ghent University, Coupure Links 653, 9000 Gent, Belgium

ARTICLE INFO

Keywords:
CEP1
UDP-galactose 4-epimerase
L-sugars
Promiscuity
Archaeal enzymes

ABSTRACT

UDP-galactose 4-epimerase (Gal4E, EC 5.1.3.2) catalyses the interconversion of UDP-galactose and UDP-glucose by inverting the configuration of the 4′-hydroxyl group of the sugar moiety. This enzyme is one of the mechanistically best-characterized members of the Short-chain Dehydrogenase/Reductase (SDR) superfamily. Although Gal4E was widely studied throughout all domains of life, ranging from eukaryotes to archaea, its biochemical characterization was often limited to UDP-hexoses, neglecting the possibility that Gal4E might be promiscuous towards other NDP-sugars and derivatives thereof. In this study, we identified a novel Gal4E subgroup displaying an unprecedented specificity on guanosine diphosphate (GDP) sugars. As a proof of concept, a detailed biochemical investigation was performed on Gal4E from *Pyrococcus horikoshii* (phGal4E_1), which revealed that it has, in fact, a clear preference for GDP-sugars. In addition, we confirmed that it accepts a variety of other sugar moieties, including L-sugars like L-galactose and L-fucose.

1. Introduction

Carbohydrate epimerases (CEP) are a group of enzymes that can invert the configuration of a specific hydroxyl group in a sugar ring through a single reaction and without prior activation or protection steps (Tanner, 2002). A large number of these epimerases catalyze conversions at the level of nucleotide diphosphate sugars (NDP-sugars, e.g. UDP-glucose, GDP-mannose, CDP-paratose), which are precursors in essential pathways in all domains of life, including involvement in the biosynthesis of vitamin C in plants (Wolucka et al., 2001), antibiotics (Hoffmeister and Ichinose, 2000) and lipopolysaccharides in bacteria (Pacinelli et al., 2002) or protein N-glycosylation in archaea (Eichler, 2013). The majority of nucleotide sugar epimerases belong to the CEP1 family, which is part of the Short-chain Dehydrogenase/Reductase (SDR) superfamily (Van Overtveldt et al., 2015). The CEP1 family comprises 10 different specificities, including UDP-hexose 4-epimerase (Gal4E) (Beerens et al., 2015), GDP-mannose 3,5-epimerase (GM35E) (Beerens et al., 2022) and CDP-paratose 2-epimerase (CPa2E) (Rapp et al., 2020). These enzymes all contain a Rossmann-fold domain for binding of NAD(P) as cofactor and the motif $[ST]x_nYx_3K$ as catalytic triad (Da Costa et al., 2021).

UDP-galactose 4-epimerase (Gal4E, EC 5.1.3.2) is by far one of the best studied members of the CEP1 family in terms of structural analysis and biochemical characterization (Shin et al., 2015). This is due to its essential role in the Leloir pathway of galactose metabolism in which it interconverts UDP-galactose and UDP-glucose (Daenzer et al., 2012). Gal4E deficiency is responsible for galactosemia, a hereditary disease, highlighting its vital importance (Bang et al., 2009). Gal4E catalysis is based on a transient keto-intermediate mechanism (Fig. 1), which can be divided in three steps: (1) abstraction of a proton from 4′-OH by the catalytic tyrosine (Holden et al., 2003) and transfer of a hydride from C4 to the NAD + cofactor, resulting in a 4-ketopyranose intermediate (Allard et al., 2001) (oxidation step) (2) rotation of the keto-sugar moiety, which allows (3) NADH to transfer the hydride to the opposite site (reduction step), completing the epimerization (Beerens et al., 2015; Van Overtveldt et al., 2015). The proton from the initial abstraction step is re-donated to the oxygen, regenerating the catalytic tyrosine.

Besides the interconversion of UDP-Glc and UDP-Gal, activity on *N*-acetylated substrates like UDP-*N*-acetylgalactosamine (UDP-GalNAc) has also been reported for some Gal4E representatives (Demendi et al., 2005; Shin et al., 2015). Consequently, a subdivision in three groups has been suggested. Group 1 and 3 have a clear preference for non-acetylated and acetylated substrates, respectively, while group 2 is active on both (Ishiyama et al., 2004). An important residue for activity on acetylated substrates is the so-called ‘gatekeeper’, which needs to be a small residue in order to allow rotation of (and thus activity on) the larger acetylated substrates (Fushinobu, 2021).

* Corresponding author.
E-mail addresses: carlosjosue.alvarezquispe@ugent.be, tom.desmet@ugent.be (T. Desmet).

https://doi.org/10.1016/j.crbiot.2022.08.003
Received 31 March 2022; Revised 12 August 2022; Accepted 18 August 2022

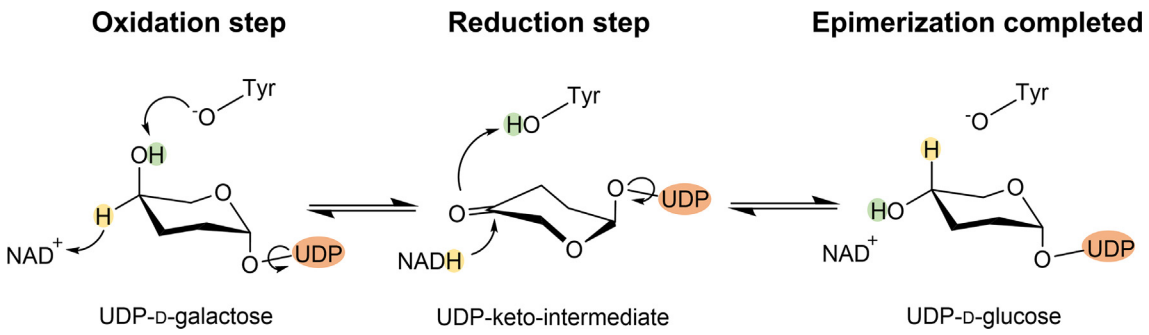


Fig. 1. Representation of the keto-intermediate mechanism in Gal4Es. The epimerization at an asymmetrically substituted carbon (C4) in UDP-D-galactose results in UDP-D-glucose. Proton transfer to and from the NAD + cofactor is depicted in yellow. Proton transfer to and from the catalytic base tyrosine is illustrated in green. Hydroxyl groups on other positions are not shown for clarity reasons. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Although this classification provides a first glimpse into the study of substrate specificity within Gal4Es, it is only limited to UDP-hexoses, neglecting the possibility that Gal4E might be potentially active on other NDP-sugars and derivatives.

The above-mentioned classification would not have been possible without the extensive study of the Gal4E representatives through all domains of life, ranging from eukaryotes (*Homo sapiens* (Thoden et al., 2000), *Trypanosoma brucei* (Friedman et al., 2012)), mesophilic bacteria (*Escherichia coli* (Thoden et al., 1996), *Bifidobacterium longum* (Nam et al., 2019)), thermophilic bacteria (*Thermotoga maritima* (Shin et al., 2015), *Marinithermus hydrothermalis* (Beerens et al., 2013)) to archaea (*Pyrococcus horikoshii* (Chung et al., 2012), *Thermus thermophilus* (Niou et al., 2009), *Pyrobaculum calidifontis* (Sakuraba et al., 2011)). Archaeal Gal4E could be especially appealing for protein evolution studies and enzyme engineering approaches since archaea are considered the earliest known life forms and their enzymes might contain some valuable ancestral features like high thermostability (Bloom et al., 2006) or substrate promiscuity (Tokuriki and Tawfik, 2009). However, identifying promiscuous activities in Gal4E is not a trivial task since most NDP-sugars are expensive and not readily available (Chung et al., 2012). In addition, there is no way to predict the type of NDP-sugar that will be used as a potential substrate because of the limited knowledge on the physiological role of archaeal epimerases (Sato and Atomi, 2011; Verhees et al., 2003) compared to bacterial and eukaryote homologs.

Facing this issue, the recently proposed heptagonal box model (Da Costa et al., 2021) offers a simple but powerful solution that allows to identify sugar- and regio-specificities of the NDP-sugar epimerases (in CEP1) only by pinpointing the model motifs. With this in mind, variation in the composition of Gal4E's heptagonal box motifs might also harbor specificity information, in a similar way as the aforementioned gatekeeper. Therefore, a detailed biochemical study of Gal4E promiscuity in archaeal representatives, all by using the heptagonal box model as a guide, could unveil relevant features to broaden the substrate specificity of Gal4E, so far restricted to UDP-hexoses. In the present study, we identified a novel Gal4E subgroup, namely one with a clear preference for guanosine diphosphate (GDP) sugars. A detailed biochemical investigation of a representative from *P. horikoshii* (phGal4E_1) as well as its potential physiological function confirmed its uniqueness compared to other Gal4E subgroups. In phylogenetic analysis, such enzymes form a separate cluster that is, in fact, more closely related to GM35E, another CEP1 specificity with a GDP-sugar as substrate (i.e., GDP-Man). In other words, phGal4E_1 could potentially be an evolutionary intermediate between traditional Gal4E and GM35E. Interestingly, this novel subgroup also shows activity towards other sugar (e.g. L-galactose and L-fucose) and nucleotide (e.g. ADP and TDP) moieties.

2. Materials and methods

2.1. Materials and chemicals

Nucleotides and sugar nucleotides were obtained from Carbosynth (Compton, Berkshire, UK). *E. coli* BL21(DE3) competent cells were prepared in-house. HPAEC-PAD standards such as galactose, glucose, L-arabinose and 6-deoxy-D-glucose were all purchased from Sigma-Aldrich (St. Louis, MO) or Carbosynth (Compton, Berkshire, UK) and stocks were prepared at 100 mM in ultra pure water. All other chemicals and reagents were of the highest available purity, unless stated otherwise.

2.2. 3DM database and phylogenetic analysis

Protein sequences were classified as Gal4E and GM35E through the signature motifs from the heptagonal box model. Then, they were extracted from the 3DM database (https://www.bio-prodict.com) (Kuipers et al., 2010) and aligned with ClustalO (Sievers and Higgins, 2018) using default parameters. The 3DM database allows the collection of protein sequence data and create a structure-based multiple sequence alignment (3D-MSA). From the alignment, it creates a unified 3D numbering scheme, which offers a useful framework to compare different protein superfamily data such as residue conservation, mutation data, ligand-binding data and other features. The phylogenetic tree was generated with PhyML 3.1 with default parameters (Guindon et al., 2010). Finally, the phylogenetic tree was uploaded to Dendroscope software (Huson and Scornavacca, 2012) for further display and manipulations.

2.3. Genome context analysis

The protein sequences from *E. coli* Gal4E (Uniprot: P09147), *T. maritima* Gal4E (Uniprot: Q9WYX9), *P. horikoshii* Gal4E_1 (Uniprot: O73960) and *P. horikoshii* Gal4E_2 (Uniprot: O59375) were input into the STRING web server (Szklarczyk et al., 2019) to search for their functionally associated genes (neighborhood, fusion, gene co-occurrence).

2.4. Structure modeling and docking

The ColabFold (AlphaFold 2) web server was used for protein structure prediction of *P. horikoshii* Gal4E_1 using default parameters (Mirdita et al., 2021). Crystal structures of *Arabidopsis thaliana* GM35E (PDB: 2C5A), *Archaeoglobus fulgidus* Gal4E (PDB: 3EHE) and *P. calidifontis* Gal4E (PDB: 3AW9) were retrieved from the PDB database. GDP-L-galactose was extracted from the crystal structure of *A.*

thaliana GM35E (PDB: 2C5A) and manually docked in the GDP-Gal4E subgroup structures with PyMOL v2.0.

2.5. Gene cloning and transformation

The Gal4E genes from *E. coli*, *T. maritima*, *P. horikoshii_1*, *P. horikoshii_2*, *A. fulgidus* and *P. calidifontis* were codon optimized for *E.coli*, synthesized and subcloned into the pET21 vector at NdeI and XhoI restriction sites, providing a C-terminal His6-tag, by GeneArt Gene Synthesis (Thermo Fisher Scientific, Waltham, MA, USA) (Table S1). The resulting constructs were transformed in *E. coli* BL21 (DE3) electrocompetent cells. The same procedure was applied for L-fucokinase/GDP-L-fucose pyrophosphorylase (FKP) from *Bacteroides fragilis* (Uniprot: Q58T34) (Liu, T., Ito, H., Chiba, 2011) and GM35E from *Methylacidiphilum fumariolicum* strain SolV (UniProt: I0K0X9) (Gevaert et al., 2020, 2019).

2.6. Protein expression and purification

An overnight preculture of *E. coli* transformed with pET21 expression plasmid was used to inoculate the Lysogeny Broth (LB) growth medium (250 mL) containing ampicillin (100 $\mu g \cdot mL^{-1}$) in a 1L shake flask at 37 °C. The culture was grown until an optical density of 0.6 was reached. Subsequently, 0.1 mM isopropyl-β-D-thiogalactopyrano side was added and the culture was incubated for 16 h at 200 rpm and 20 °C (New Brunswick™ Innova ® 40 - Benchtop Orbital Shaker). Cells were harvested by centrifuging for 20 min at 9000 rpm and 4 °C (Thermo Scientific™ Sorvall™ RC 6 Plus Centrifuge). The obtained pellets were frozen and stored at − 20 °C for at least one day.

For enzyme extraction and purification, each pellet of a 250 mL culture was resuspended in 8 mL of lysis buffer (300 mM NaCl, 10 mM imidazole, 100 μM phenylmethane sulfonyl fluoride (PMSF) and 1 $mg \cdot mL^{-1}$ lysozyme in 50 mM sodium phosphate buffer pH 7.5) and cooled on ice for 30 min. Next, the cells were subjected to 3 times 3 min of sonication (Branson sonifier 250, level 3, 30% duty cycle). Finally, cell debris was removed by centrifugation at 9000 rpm for 1 h. Next, the supernatant was purified by Ni-NTA chromatography, with small variations to the supplier's description (Thermo Fisher Scientific, Waltham, MA, USA). The resulting supernatant was incubated with 1.5 mL of equilibrated HisPur™ Ni-NTA Resin (Thermo Fisher Scientific, Waltham, MA, USA) in a 10 mL gravity chromatography column at 4 °C for 1 h. The flowthrough was discarded, and the resin washed four times with 8 mL Ni-NTA wash buffer (300 mM NaCl, 30 mM imidazole in 50 mM sodium phosphate buffer pH 7.4). Protein was eluted with 8 times 1 mL elution buffer (500 mM NaCl, 250 mM imidazole in 50 mM sodium phosphate buffer pH 7.5). As a final step, buffer was exchanged to 100 mM MOPS pH 7 by using Amicon Ultra-15 centrifugal filter units with 30 kDa cut-off (Merck Millipore Darmstadt, Germany).

The protein concentration was determined by measuring the absorbance at 280 nm with a NanoDrop2000 Spectrophotometer (Thermo Scientific) using extinction coefficients calculated with the ProtParam tool on the ExPASy server (https://web.expasy.org/protparam/). Molecular weight and purity of the protein were verified by sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE; 12% gel) (Fig. S1). The enzyme's electrophoretic behavior corresponded well with its predicted molecular mass of about 37 kDa.

2.7. Substrate specificity and kinetic characterization

2.7.1. NDP-sugar assays to evaluate promiscuity

Specific activities for nucleotide (GDP, ADP, TDP and UDP)-activated sugars were determined in 50 μL reaction volume containing 100 mM MES buffer (pH 6.5), 5 mM of substrate and the purified Gal4E (in a range of 0.003 to 1 $mg \cdot mL^{-1}$ to keep measurements within the linear range of the analysis). The reaction was performed at 60 °C. For the preliminary screening, the reaction mix was incubated overnight. For the specific activity, samples of 5 μL were taken every 2 min for 10 min. Subsequently, the enzyme was acid/heat inactivated and the NDP-sugars were hydrolyzed at the same time. The enzyme acid/heat inactivation together with the NDP-sugar hydrolysis step involved a 20-fold dilution of the sample in 100 mM acetic acid and incubation at 95 °C for 1 h. The released sugar moieties were analyzed by HPAEC-PAD. One enzyme unit (U) equals the production on 1 μmol NDP-sugar per minute under the conditions used.

2.7.2. GDP-L-galactose production and screening

GDP-L-galactose production by mfGM35E was determined in 50 μL reaction volume containing 100 mM MOPS buffer (pH 7), 5 mM of GDP-D-mannose and 1 $mg \cdot mL^{-1}$ of mfGM35E. The reaction was performed at 60 °C for 3 h. Next, 1 $mg \cdot mL^{-1}$ of Gal4E was added and samples of 5 μL were taken after 0, 5, 30 and 60 min of reaction. GDP-L-galactose production by FKP was carried out in 100 μL reaction volume containing 100 mM MOPS buffer (pH 7), 5 mM L-galactose, 3 mM ATP, 3 mM GTP, 5 mM MnSO4 and 1 $mg \cdot mL^{-1}$ FKP. The reaction was performed at 37 °C for 2 h. Subsequently, the temperature was raised to 60 °C, 1 $mg \cdot mL^{-1}$ of Gal4E was added and samples of 5 μL were taken after 0, 5, 30 and 60 min of reaction. The reaction samples were hydrolyzed and analyzed as stated in (2.7.1).

2.7.3. Kinetic parameters of phGal4E_1

Enzyme specific activity incubated with various GDP-D-glucose (2–70 mM) and UDP-D-glucose (2–20 mM) concentrations was evaluated in 100 mM MES pH 6.5 and 3 $\mu g \cdot mL^{-1}$ (for GDP-glc), 65 $\mu g \cdot mL^{-1}$ (for UDP-glc) of phGal4E_1 at 60 °C. The kinetic parameters, including the Michaelis–Menten constant (K_M) and turnover number (k_{cat}), were determined using a Michaelis–Menten plot created in SigmaPlot (Systat Software, San Jose, CA, USA).

2.7.4. HPAEC-PAD

Conversion of substrate to product was evaluated by high-performance anion exchange chromatography-pulsed amperometric detection (HPAEC-PAD) using the Dionex ICS-3000 system (Thermo Fischer Scientific) (CarboPac PA20 column-3 × 150 mm). An isocratic flow (0.5 $mL \cdot min^{-1}$ for 15 min) of 100 mM NaOH (4%) and ultrapure water (96%) was used for phGal4E_1 screening, kinetic characterization and FKP reactions. For the reactions involving mfGM35E, an isocratic flow (0.5 $mL \cdot min^{-1}$ for 20 min) of 100 mM NaOH (15%) and ultrapure water (85%) was used. The NDP-sugar conversion was quantified using standards of the respective monosaccharides (Fig. S2) and recalculated based on peak areas.

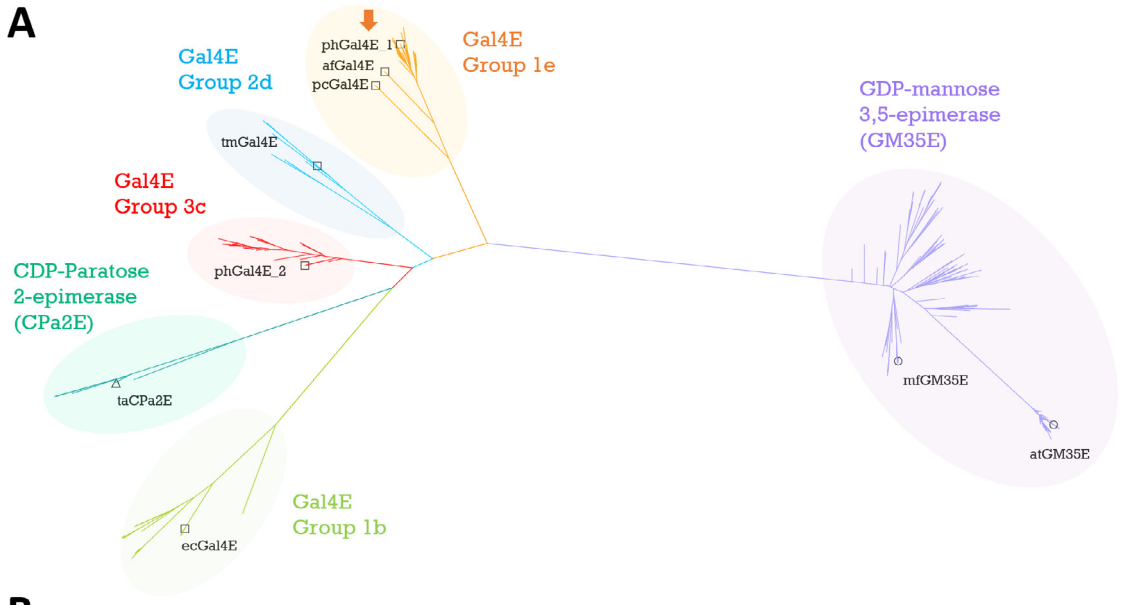
2.8. Statistical analysis

Statistical analysis were performed with R studio using significance cut-offs of $p < 0.05$, $p < 0.01$ and $p < 0.001$. F Test (var.test) was performed to verify the homogeneity of two variances (NDP-sugar vs UDP-glc). Mean values of specific activities on different NDP-sugars were compared to the reference (UDP-glc) using the Student's *t*-test (t.test). Statistics were based on the three replicates of the specific activity per substrate.

3. Results and discussion

3.1. Discovery of the novel Gal4E subgroup

To gain more insight in the diversity and connections between CEP1 members, a phylogenetic tree (Fig. 2A) was constructed for the sequences (~300 in total) classified as Gal4E, CPa2E and GM35E in the 3DM database (Kuipers et al., 2010). The heptagonal box model fingerprints for each representative node of the tree were used for



B

Enzyme	Uniprot/NCBI ID	Red wall	Yellow wall	Cyan wall	Orange wall	Purple wall	Grey wall	Green wall
phGal4E_1	O73960	PEVR	SST	YGAAK	LANI	SNHGV	QRKSY	DVKLM
pcGal4E	A3MUJ4	PEVR	SST	YGAAK	YANV	LRHGV	QRKSY	DVKYM
afGal4E	O29886	PDVR	TST	YGASK	FANV	STHGV	QNKSY	DVPVM
ecGal4E	P09147	KAVG	SAT	YGRSK	YFNP	EDPQGIPNN	GVRDY	DLPAY
tmGal4E	Q9WYX9	ASVA	TggA	YGIAK	YANV	QDPYGEAGV	YVRDY	DVRKS
phGal4E_2	O59375	ISVK	SSA	YGITK	YFNV	QSSAYAGV	QTRDF	DIRRS
mfGM35E	I0K0X9	DMGGIG	SAC	YGWEK	FHNV	GTYKGGREK	QTRSF	VRGR
atGM35E	Q93VR3	DMGGMG	SAC	YGLEK	FHNI	GTWKGGREK	QTRSF	VRGR
taCPa2E	WP_022854415.1	QVAMT	TNK	YGCSK	HSSM	QFATYDQGW	QVRDV	DQKIF

Fig. 2. Clustering of CEP1 members annotated as UDP-galactose 4-epimerases (Gal4E), GDP-mannose 3,5-epimerases (GM35E) and CDP-paratose 2-epimerases (CPa2E). A) Phylogenetic tree of CEP1 members. Putative and characterized (shown) CPa2Es, GM35Es and Gal4Es (including the different subgroups) were grouped according to substrate specificity (confirmed biochemically for at least one member in each group). PhGal4E_1 is classified within the group 1e in orange. B) Heptagonal box motifs for each representative. The different walls of the model were used to verify the classification of the representative's homologs included in the tree. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

sequence selection and classification (Fig. 2B). The phylogenetic tree comprises two major branches, one of these harbors the characterized GDP-Man 3,5-epimerases from *A. thaliana* (atGM35E) (Wolucka et al., 2001) and *Methylacidiphilum fumariolicum* (mfGM35E) (Gevaert et al., 2019). Within this branch, sequences with more similarity to atGM35E (generally from plants) form a narrow group compared to the mfGM35E group (from archaea and bacteria), which is more spread along the branch.

The second major branch contains the Gal4E and CPa2E homologues and shows a broader diversity (mainly within Gal4Es) compared to the GM35E branch. This branch corresponds to the sequences clustering with the characterized CPa2E from *Thermodesulfatator atlanticus* (taCPa2E) (Rapp et al., 2020), Gal4E from *E. coli* (ecGal4E), *T. maritima* (tmGal4E), *P. horikoshii* (phGal4E_1) and *P. calidifontis* (pcGal4E). Sequences clustering with the putative Gal4E from *A. fulgidus* (afGal4E) and, curiously, a potential second Gal4E from *P. horikoshii* (phGal4E_2) were also included. Within this branch, we can identify the CPa2E clade but also four Gal4E clades agreeing with Fushinobu's classification (Fushinobu, 2021), where the green and light-blue clades correspond to Group 1b (solely active on UDP-Glc/Gal) and Group 2d (active on UDP-Glc/Gal and UDP-GlcNAc/GalNAc), respectively. The red clade coincides with group 3c that is supposed to be only active on UDP-GlcNAc/GalNAc. This hypothesis was confirmed experimentally in this work by screening the paralog phGalE_2 on UDP-GlcNAc (Fig. S3). The orange clade contains only archaeal Gal4Es and represents Group 1e, harboring activity on UDP-Glc/Gal. However, the activity on UDP-GlcNAc/GalNAc as well as other NDP-sugars was not measured in previous work. It is important to note that the sequences from the orange and red clades are mostly paralogs, since they are present in the same organism (generally from the genus *Pyrococcus* and *Thermococcus*), except for afGal4E and pcGal4E paralogs that were not found and thus are absent in the red clade.

Of all Gal4E clades, the orange one (Group 1e) is most closely positioned to the GM35E branch. This observation suggests that Group 1e can potentially contain some unknown specificities on NDP-sugars different from UDP-hexoses, with GDP-sugars as the most likely substrates. Indeed, a previous phylogenetic analysis has indicated that enzymes from these thermophilic archaea might have unique characteristics (Fushinobu, 2021).This idea is also supported by the analysis of the heptagonal box motifs within Gal4E groups (Fig. 2B). Group 1e enzymes stand out from the rest by containing a new yellow wall motif (SST/TST), which reinforces the hypothesis of a new specificity. We targeted the yellow wall since it is one of the most conserved parts of the model per specificity (Da Costa et al., 2021). Therefore, this

work will focus on Group 1e for further studies and the characterization of phGal4E_1 will be used as a proof of concept.

3.2. Inspection of genomic context in Gal4E

To shed light on the functional associations of phGal4E_1 we used the webtool STRING, which provides information based on the hypothesis that functionally associated genes are often in physical proximity constituting operons (gene neighborhood) and/or exhibiting the same expression (co-expression) and occurrence pattern (co-occurrence).

This comparative genomic context analysis (Fig. 3) revealed the expected Leloir pathway enzymes as interacting partners with the known representatives ecGal4E and tmGal4E. Interestingly for *Pyrococcus horikoshii* the genetic organization allowed to identify two Gal4E paralogs (phGal4E_1 and phGal4E_2). The former is more likely to participate in the Leloir pathway (even if a galactose mutarotase is absent in the network) than the latter, which only retains a (putative) galactose-1-phosphate uridylyltransferase as associated partner. It is important to note that phGal4E_1 and phGal4E_2 are also interacting partners connected by neighborhood and co-ocurrence STRING channels. This information proposes that phGal4E_1 might be implicated in a more traditional Leloir(-like) pathway or at least in an archaeal variant of it. On the other hand, phGal4E_2 might participate in a different pathway. However, these results can also be explained by the limited number of annotated and characterized proteins in the *P. horikoshii* genome compared to *E. coli*, which reduces the accuracy of the analysis. Consequently, it is possible that galactose mutarotase as well as phGal4E_2 interacting partners are present but located outside the phGal4Es neighborhood. Therefore, more experimental information must be collected to elucidate the main role of phGal4Es.

3.3. Evaluation of phGal4E_1 promiscuity on different NDP-sugars

To further define phGal4E_1′s substrate scope, we evaluated its activity on NDP-sugars with different nucleotide moieties (Fig. 4). The results clearly show that phGal4E_1 exhibits high significant activity towards GDP-D-glucose, which is 10 times higher than the activity on the "native" substrate UDP-D-glucose (at a concentration of 5 mM). In addition, activity towards TDP-D-glucose and ADP-D-glucose was also significantly higher in comparison to the activity on UDP-D-glucose. According to previous studies (Chung et al., 2012), phGal4E_1 has a preference for UDP-D-galactose over UDP-D-glucose, which is consistent with our results. However, the opposite phenomenon was observed with GDP-D-glucose and GDP-D-galactose. Indeed, phGal4E_1 has a remarkably high activity on GDP-D-glucose, suggesting that this enzyme might not participate in the traditional Leloir pathway unlike other subgroups (i.e. 1b and 2d).

Next, the enzyme was also tested on substrates that contain a different sugar moiety. Consequently, activities on GDP-L-fucose (5.84 ± 0.04 U mg^{-1}), UDP-D-xylose (0.06 ± 0.007 U mg^{-1}), GDP-D-mannose (no activity) and UDP-D-GlcNAc (no activity) were measured in this study. Interestingly, it seems that the change/absence of substituents, and thus lack of interactions, at the level of the sugar moiety's C6 (-H for fucose and no C6 for xylose) produces a decrease in activity. Nonetheless, they are still suitable for the C4 epimerization. Differently, a change on the hydroxyl group orientation in the sugar moiety C2 (OH in axial position for D-mannose) or the presence of a bulkier substituent in C2 (amide for GlcNAc rather than OH) totally restricts the activity. Most strikingly, GDP-L-fucose is the second highest activity among all the NDP-sugars tested, highlighting the importance of the GDP-moiety for substrate affinity but also the potential of phGal4E_1 activity on GDP-L-sugars.

To verify the uniqueness of phGal4E_1 we also assayed the activity of ecGal4E and tmGal4E on GDP-D-glucose and GDP-L-fucose but no activity was detected on the latter substrate and only low activity on the former with tmGal4E (18 % conversion after 24 h at pH 6.5 and 60 °C, using 5 mM substrate and 1 mg/ml enzyme). This clearly shows that phGal4E_1 harbors an uncommon specificity towards GDP-sugars within the Gal4E family.

Building on these results, we further assessed phGal4E_1′s activity towards GDP-L-galactose (Fig. 5). Unfortunately, this substrate was difficult to produce from our supplier and thus very expensive. Therefore, we produced GDP-L-galactose in house by two different enzymatic mechanisms. The first employed a GDP-Man 3,5-epimerase that allowed to obtain GDP-L-Gal from GDP-D-Man. For the enzymatic

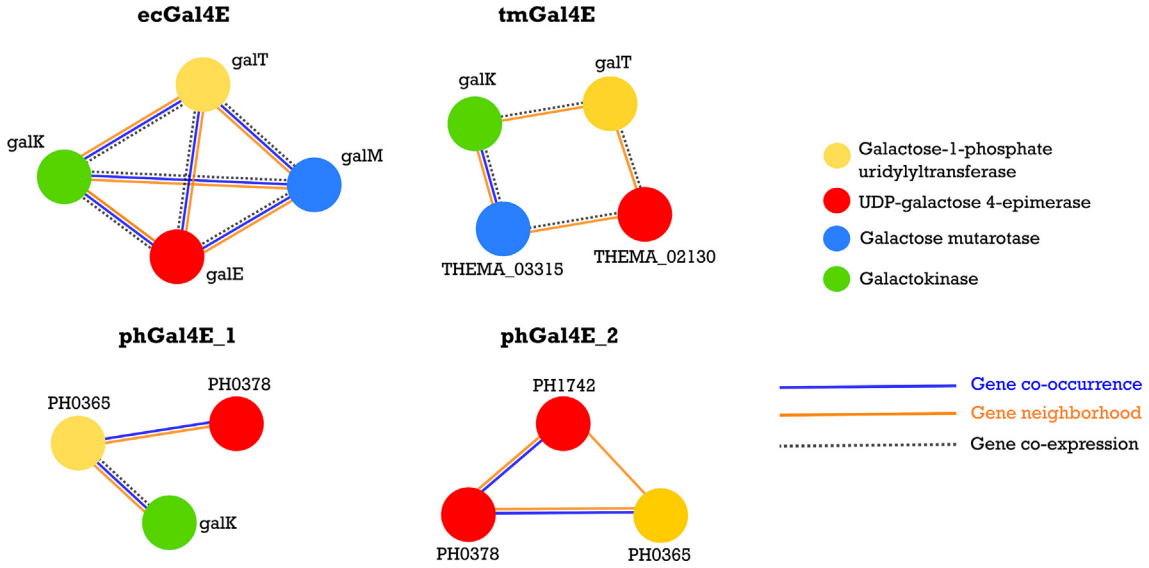


Fig. 3. Association network for the genome context analysis of four Gal4E representatives. The Gal4Es of *E.coli* (ecGal4E), *T. maritima* (tmGal4E) and *P. horikoshii* (phGal4E_1, phGal4E_2) were selected as input. The network contains only Leloir pathway enzymes (galactose-1-phosphate uridylyltransferase, UDP-galactose 4-epimerase, galactose mutarotase and galactokinase) since they showed the highest score in three STRING prediction channels: gene co-occurrence (blue straight line), gene neighborhood (orange straight line) and gene co-expression (black dotted line). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

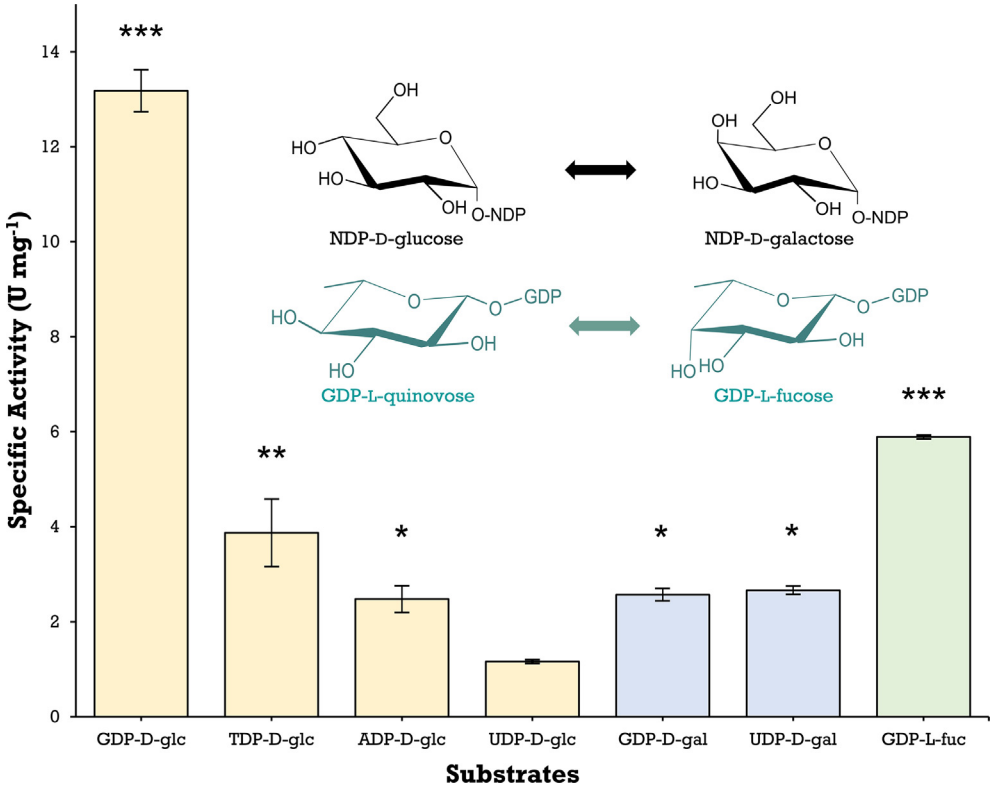


Fig. 4. NDP-sugar screening on phGal4E_1. A substrate concentration of 5 mM was used for this screening (pH 6.5 and 60 °C). The reaction mixtures catalyzed by phGal4E_1 were hydrolyzed prior to HPAEC-PAD analysis. PhGal4E_1 average specific activity was calculated from an experiment with 3 replicates per substrate and the error bars correspond to the standard deviation within the replicates. Statistical significance of the specific activity on other NDP-sugars was compared to UDP-D-glucose and is indicated with asterisks ($p < 0.05$ (*), $p < 0.01$ (**), $p < 0.001$ (***)).

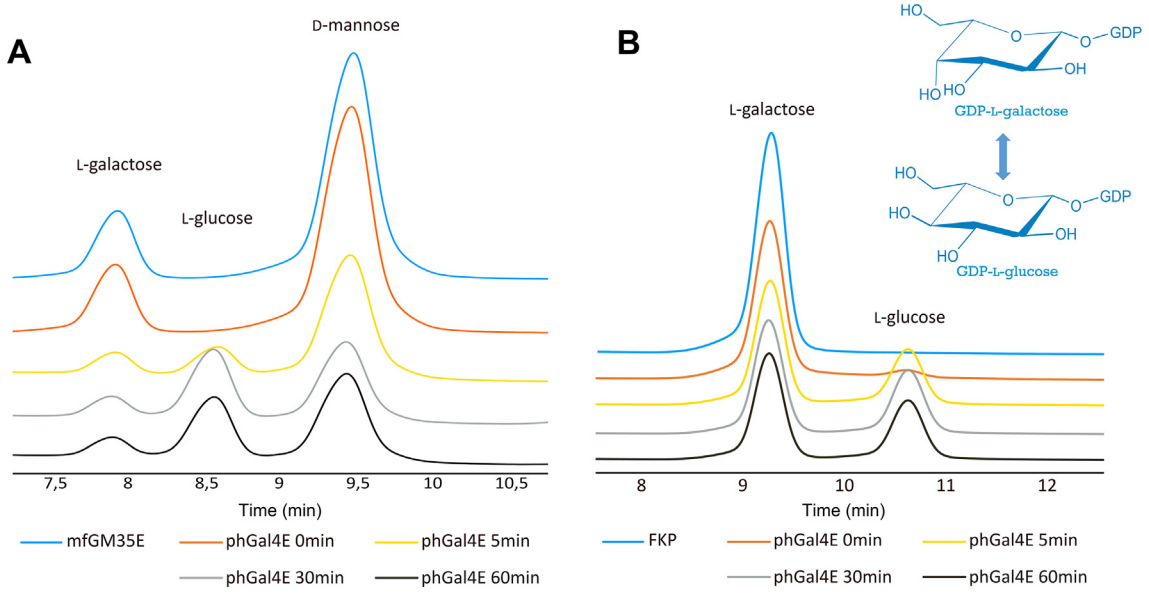


Fig. 5. PhGal4E_1 activity on GDP-L-galactose. A) After GDP-L-galactose production with mfGM35E reaction B) After GDP-L-galactose production with FKP reaction. Peaks shown correspond to the hydrolyzed NDP-sugar (GDP-D-mannose, GDP-L-glucose and GDP-L-galactose) prior to HPAEC-PAD analysis. An isocratic flow (0.5 mL·min^{-1} for 20 min) of 100 mM NaOH (15%) and ultrapure water (85%) was used for the separation in A. An isocratic flow (0.5 mL·min^{-1} for 15 min) of 100 mM NaOH (4%) and ultrapure water (96%) was used for the separation in B.

assemble, mfGM35E was incubated for 3 h at 60 °C in presence of GDP-D-Man and then phGal4E_1 was added and assayed for (0, 5, 30 and 60 min). The results showed that a peak corresponding to L-glucose appeared after 5 min of reaction with phGal4E_1. Equilibrium was reached within 30 min (75% for GDP-L-Glc and 25% for GDP-L-Gal, as seen previously for UDP-D-Glc/Gal). No L-glucose peak was observed in the negative control (same setup without addition of phGal4E_1). We also observed a decrease in the mannose peak that

does not correspond to the increase of L-galactose, which is probably due to the production of GDP-L-gulose and GDP-D-altrose that are also present in mfGM35 reaction (Gevaert et al., 2020) (Fig. S4). Also worth mentioning, L-glucose production was also observed in other representatives from the clade, namely afGal4E and pcGal4E (Fig. S5), which emphasizes the fact that GDP-sugar preference is not an isolated case for phGal4E_1 but a common feature within Gal4Es from Group 1e.

Secondly, GDP-L-galactose was also produced by using the L-fucokinase/GDP-L-fucose pyrophosphorylase (FKP) from *Bacteroides fragilis* (Liu, T., Ito, H., Chiba, 2011). This bifunctional enzyme naturally converts L-fucose towards GDP-L-fucose but was also found to accept L-galactose (Ohashi, 2017). Consequently, the FKP reaction was performed for 2 h at 37 °C for GDP-L-galactose production. Afterwards, phGal4E_1 was added and analyzed for GDP-L-glucose formation. As expected, GDP-L-glucose was quickly produced (already in 5 min reaction sample). However, with the FKP setup we did not reach the same conversion as observed for the mfGM35E setup. In fact, it seems that the reaction stops, or at least the rate decreases after 5 min of reaction. This result could be explained by the fact that GTP and/or ATP might bind the phGal4E_1 active site and thus being competitive inhibitors of GDP-L-galactose. We also observed that phGalE_1 was active on FKP produced GDP-L-fucose but also with a possible inhibition effect (Fig. S6).

In addition, the mfGM35E setup was used to test the activity of ecGal4E and tmGal4E on GDP-L-galactose, but no activity was detected either. These results highlight the distinctiveness of phGal4E_1 subgroup regarding the epimerization of GDP-sugars.

3.4. Comparison of the kinetic parameters in Gal4E

The kinetic parameters of phGal4E_1 were determined on both UDP-D-Glc and GDP-D-Glc (Table 1). The results show that its activity

Table 1
Comparison of the kinetic parameters of phGal4E_1 on UDP-D-glucose and GDP-D-glucose.

Substrate	K_M (mM)	k_{cat} (s^{-1})	k_{cat}/K_M
UDP-D-glucose	2.8 ± 0.8	0.4 ± 0.1	0.1
GDP-D-glucose	17.1 ± 3.1	38.0 ± 2.3	2.2

The enzymatic tests were performed at 60 °C, pH = 6.5 and the reaction mixtures catalyzed by phGal4E_1 were hydrolyzed prior to HPAEC-PAD analysis. Data are presented as means ± standard deviation of the mean.

(k_{cat} = 0.4 s^{-1}) on UDP-Glc is very low compared to that reported for a classical Gal4E, such as the enzyme from *E. coli* (k_{cat} = 18 s^{-1})(Chen et al., 1999; Zhu et al., 2018). Vice versa, its activity on GDP-Glc is remarkably high, whereas ecGal4E shows absolutely no activity on this substrate (see section 3.3). However, the K_M of phGal4E_1 is 6-fold higher on GDP-Glc than on UDP-Glc, and is also higher than that of ecGal4E on its "natural" substrate UDP-Glc (K_M = 1.2 mM)(Chen et al., 1999; Zhu et al., 2018). This could be explained by the fact that the guanosine moiety is bulkier than the uridine, as a result, GDP-Glc might be more difficult to fit on the active site. On the other hand, the GDP-moiety might favorize the reaction distance for sugar C4 epimerization, hence the turnover number increases. Nevertheless, the overall catalytic efficiency of phGalE_1 still is about 20 times higher on GDP-Glc than on UDP-Glc, confirming the results of our preliminary screening (see section 3.3).

3.5. Structural analysis of the novel GDP-Gal4E subgroup

To evaluate the structural features responsible for this GDP-preference in phGal4E_1 we compared the NDP binding site of atGM35E (the only crystallized CEP1 epimerase active on GDP-sugars, namely GDP-Man and GDP-L-Gal) to the NDP binding site of

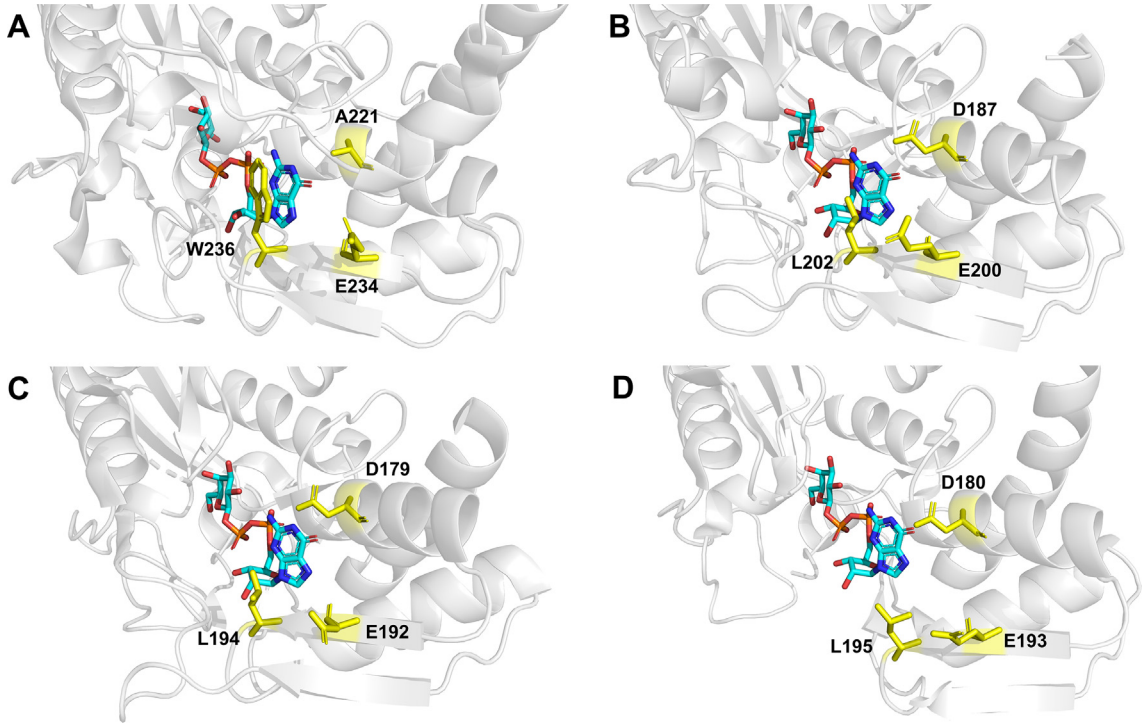


Fig. 6. NDP binding site comparison of the new GDP-Gal4E subgroup with a GM35E representative A) *Arabidopsis thaliana* GM35E (PDB: 2C5A). B) *Pyrococcus horikoshii* Gal4E (Alpha Fold2 model). C) *Pyrobaculum calidifontis* Gal4E (PDB: 3AW9). D) *Archaeoglobus fulgidus* Gal4E (PDB: 3EHE). The GM35E structure was co-crystallized GDP-L-galactose while the GDP-Gal4E representatives were manually docked with the same GDP-L-sugar. Conserved residues are colored in yellow and the nucleotide sugar in cyan. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

the novel Gal4E subgroup representatives, namely phGal4E_1, pcGal4E and afGal4E (Fig. 6). According to a previous study, three residues were identified to be responsible for GDP specificity in atGM35E, namely Ala221, Glu234, Trp236 (Van Overtveldt et al., 2020). Subsequently, these residues were introduced in the Gal4E from *Thermus thermophilus* (which is more active on UDP-Glc). As a result, an increased affinity and activity on GDP-Glc was obtained. In our work, these positions were pinpointed in three representatives from the new GDP-sugar 4-epimerases and we found that the aforementioned Ala, Glu and Trp correspond to the GDP-binding triad Asp, Glu and Leu, respectively. Additionally, the triad was highly conserved within the novel GDP-Gal4E subgroup structures, suggesting that this subgroup might use slightly altered interactions for GDP binding. Interestingly, only Glu conservation was shared between GM35E (Glu234 in atGM35E) and GDP-Gal4E (Glu200, Glu193 and Glu192 in phGal4E_1, pcGal4E and afGal4E, respectively). Although the function of Glu is still not clear, the high conservation of this residue might indicate an evolutionary connection between GM35E and Gal4E, placing the novel Gal4E subgroup as a possible intermediary. Supporting this notion, GDP-L-galactose was extracted from the crystal structure of atGM35 (PDB: 2C59), then manually docked in the active site of the three representatives. As expected, the docking did not show any steric clashes in the GDP-sugar 4-epimerase structures. This correct positioning in the active site was already supported by the activity of these enzymes on GDP-L-galactose (see section 3.3).

4. Conclusion

In this study, we discovered a new subgroup of NDP-sugar C4 epimerases with a high activity for GDP-sugars. Therefore, due to the singularity of phGal4E_1 as well as other representatives of this clade, we propose the name GDP-sugar 4-epimerases (GDP-Gal4E) to refer to this new subgroup. In addition, as a result of the phGal4E_1 analysis, we can infer that enzymes in this subgroup might be highly promiscuous towards a broad range of NDP-sugars, including different NDPs (GDP, TDP, UDP and ADP) as well as sugar moieties (both D- and L-hexoses, deoxysugars, and pentoses). Although the structural basis that allow this extraordinary promiscuity in phGal4E_1 remains to be investigated, our approach has shown that the heptagonal box could be used as a predictive model to discover new subgroups among the NS-SDR enzymes, in this case shown on the group of 4-epimerases. However, the existence of this GDP-Gal4E subgroup suggests that future endeavors to unravel the specificity determinants within Gal4E should also target the NDP-binding site. Indeed, while the heptagonal box model provides insight regarding the interactions in the sugar moiety, the NDP-binding site might contain more valuable information than just its substrate anchoring function (Beerens et al., 2015). Therefore, this region can provide interesting hotspots to target in Gal4E engineering studies.

Our work also reveals the interest of studying archaeal enzymes for glycobiology because they can harbor new specificities and activities. Indeed, an extensive variety of glycosides (including abundant and rare sugars) participate in several pathways in archaea, many of them still undeciphered in their entirety. In this respect, the unprecedented specificity of phGal4E_1 on GDP-L-sugars could offer, in the long term, a myriad of possibilities for both glycodiversification (e.g. oligosaccharides, glycosides) and rare sugar production (e.g. L-sugars, deoxysugars) which find plenty of potential applications in the food, nutraceutical, and pharmaceutical industries.

CRediT authorship contribution statement

Carlos Alvarez Quispe: Conceptualization, Methodology, Investigation, Formal analysis, Data curation, Visualization, Writing – original draft, Writing – review & editing. **Matthieu Da Costa:** Conceptualization, Validation, Writing – review & editing. **Koen Beerens:** Conceptualization, Validation, Supervision, Writing – review & editing. **Tom Desmet:** Conceptualization, Validation, Funding acquisition, Supervision, Writing – review & editing.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

We acknowledge financial support from the Fund for Scientific Research Flanders (FWO-Vlaanderen) via a doctoral scholarship for CA (grant NO. 1105922N) and the DeoxyBioCat project (grant n° G0A7520N).

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.crbiot.2022.08.003.

References

Allard, S.T.M., Giraud, M.F., Naismith, J.H., 2001. Epimerases: Structure, function and mechanism. Cell. Mol. Life Sci. 58, 1650–1665. https://doi.org/10.1007/PL00000803.

Bang, Y.L., Nguyen, T.T.T., Trinh, T.T.B., Kim, Y.J., Song, J., Song, Y.H., 2009. Functional analysis of mutations in UDP-galactose-4-epimerase (GALE) associated with galactosemia in Korean patients using mammalian GALE-null cells. FEBS J. 276, 1952–1961. https://doi.org/10.1111/j.1742-4658.2009.06922.x.

Beerens, K., Gevaert, O., Desmet, T., 2022. GDP-Mannose 3,5-Epimerase: A View on Structure, Mechanism, and Industrial Potential. Front. Mol. Biosci. 8, 1–13. https://doi.org/10.3389/fmolb.2021.784142.

Beerens, K., Soetaert, W., Desmet, T., 2015. UDP-hexose 4-epimerases: a view on structure, mechanism and substrate specificity. Carbohydr. Res. 414, 8–14. https://doi.org/10.1016/j.carres.2015.06.006.

Beerens, K., Soetaert, W., Desmet, T., 2013. Characterization and mutational analysis of the UDP-Glc (NAc) 4-epimerase from Marinithermus hydrothermalis. Appl. Microbiol. Biotechnol. 97, 7733–7740. https://doi.org/10.1007/s00253-012-4635-6.

Bloom, J.D., Labthavikul, S.T., Otey, C.R., Arnold, F.H., 2006. Protein stability promotes evolvability. Proc. Natl. Acad. Sci. U. S. A. 103, 5869–5874. https://doi.org/10.1073/pnas.0510098103.

Chen, X., Kowal, P., Hamad, S., Fan, H., Wang, P.G., 1999. Cloning, expression and characterization of a UDP-galactose 4-epimerase from Escherichia coli. Biotechnol. Lett. 21, 1131–1135. https://doi.org/10.1023/A:1005678225031.

Chung, S.K., Ryu, S.I., Lee, S.B., 2012. Characterization of UDP-glucose 4-epimerase from Pyrococcus horikoshii: Regeneration of UDP to produce UDP-galactose using two-enzyme system with trehalose. Bioresour. Technol. 110, 423–429. https://doi.org/10.1016/j.biortech.2012.01.046.

Da Costa, M., Gevaert, O., Van Overtveldt, S., Lange, J., Joosten, H.-J., Desmet, T., Beerens, K., 2021. Structure-function relationships in NDP-sugar active SDR enzymes: Fingerprints for functional annotation and enzyme engineering. Biotechnol. Adv. 48,. https://doi.org/10.1016/j.biotechadv.2021.107705 107705.

Daenzer, J.M.I., Sanders, R.D., Hang, D., Fridovich-Keil, J.L., Berry, G.T., 2012. UDP-Galactose 4′-epimerase activities toward UDP-Gal and UDP-GalNAc play different roles in the development of Drosophila melanogaster. PLoS Genet. 8 (5), e1002721.

Demendi, M., Ishiyama, N., Lam, J.S., Berghuis, A.M., Creuzenet, C., 2005. Towards a better understanding of the substrate specificity of the UDP-N-acetylglucosamine C4 epimerase WbpP. Biochem. J. 389, 173–180. https://doi.org/10.1042/BJ20050263.

Eichler, J., 2013. Extreme sweetness: Protein glycosylation in archaea. Nat. Rev. Microbiol. 11, 151–156. https://doi.org/10.1038/nrmicro2957.

Friedman, A.J., Durrant, J.D., Pierce, L.C.T., Mccorvie, T.J., Timson, D.J., Mccammon, J. A., 2012. The Molecular Dynamics of Trypanosoma brucei UDP-Galactose 4′-Epimerase: A Drug Target for African Sleeping Sickness. Chem. Biol. Drug Des. 80, 173–181. https://doi.org/10.1111/j.1747-0285.2012.01392.x.

Fushinobu, S., 2021. Molecular evolution and functional divergence of UDP-hexose 4-epimerases. Curr. Opin. Chem. Biol. 61, 53–62. https://doi.org/10.1016/j.cbpa.2020.09.007.

Gevaert, O., Van Overtveldt, S., Beerens, K., Desmet, T., 2019. Characterization of the First Bacterial and Thermostable GDP-Mannose 3,5-Epimerase. Int. J. Mol. Sci. 20 (14), 3530.

Gevaert, O., Van Overtveldt, S., Da Costa, M., Beerens, K., Desmet, T., 2020. GDP-altrose as novel product of GDP-mannose 3,5-epimerase: Revisiting its reaction mechanism.

Int. J. Biol. Macromol. 165, 1862–1868. https://doi.org/10.1016/j.ijbiomac.2020.10.067.

Guindon, S., Dufayard, J.F., Lefort, V., Anisimova, M., Hordijk, W., Gascuel, O., 2010. New algorithms and methods to estimate maximum-likelihood phylogenies: Assessing the performance of PhyML 3.0. Syst. Biol. 59, 307–321. https://doi.org/10.1093/sysbio/syq010.

Hoffmeister, D., Ichinose, K., 2000. The NDP-sugar co-substrate concentration and the enzyme expression level influence the substrate specificity of glycosyltransferases: cloning and characterization of deoxysugar biosynthetic genes of the urdamycin biosynthetic gene cluster. Chem. Biol. 7 (11), 821–831.

Holden, H.M., Rayment, I., Thoden, J.B., 2003. Structure and Function of Enzymes of the Leloir Pathway for Galactose Metabolism. J. Biol. Chem. 278, 43885–43888. https://doi.org/10.1074/jbc.R300025200.

Huson, D.H., Scornavacca, C., 2012. Dendroscope 3: An interactive tool for rooted phylogenetic trees and networks. Syst. Biol. 61, 1061–1067. https://doi.org/10.1093/sysbio/sys062.

Ishiyama, N., Creuzenet, C., Lam, J.S., Berghuis, A.M., 2004. Crystal structure of WbpP, a genuine UDP-N-acetylglucosamine 4-epimerase from Pseudomonas aeruginosa: Substrate specificity in UDP-hexose 4-epimerases. J. Biol. Chem. 279, 22635–22642. https://doi.org/10.1074/jbc.M401642200.

Kuipers, R.K., Joosten, H.J., Van Berkel, W.J.H., Leferink, N.G.H., Rooijen, E., Ittmann, E., Van Zimmeren, F., Jochens, H., Bornscheuer, U., Vriend, G., Martins Dos Santos, V.A.P., Schaap, P.J., 2010. 3DM: Systematic analysis of heterogeneous superfamily data to discover protein functionalities. Proteins Struct. Funct. Bioinforma. 78, 2101–2113. https://doi.org/10.1002/prot.22725.

Liu, T.-W., Ito, H., Chiba, Y., Kubota, T., Sato, T., Narimatsu, H., 2011. Functional expression of L-fucokinase/guanosine 5′-diphosphate-L-fucose pyrophosphorylase from Bacteroides fragilis in Saccharomyces cerevisiae for the production of nucleotide sugars from exogenous monosaccharides. Glycobiology 21 (9), 1228–1236.

Mirdita, M., Ovchinnikov, S., Steinegger, M., 2021. ColabFold - Making protein folding accessible to all. bioRxiv 2021.08.15.456425. https://doi.org/10.1101/2021.08.15.456425.

Nam, Y.W., Nishimoto, M., Arakawa, T., Kitaoka, M., Fushinobu, S., 2019. Structural basis for broad substrate specificity of UDP-glucose 4-epimerase in the human milk oligosaccharide catabolic pathway of Bifidobacterium longum. Sci. Rep. 9, 11081. https://doi.org/10.1038/s41598-019-47591-w.

Niou, Y.K., Wu, W.L., Lin, L.C., Yu, M.S., Shu, H.Y., Yang, H.H., Lin, G.H., 2009. Role of galE on biofilm formation by Thermus spp. Biochem. Biophys. Res. Commun. 390, 313–318. https://doi.org/10.1016/j.bbrc.2009.09.120.

Ohashi, H., 2017. Effective Synthesis of Guanosine 5'-Diphospho-β-L-galactose using Bacterial L-Fucokinase/Guanosine 5'-diphosphate-L-fucose Pyrophosphorylase. Adv. Synth. Catal. 359, 4227–4234. https://doi.org/10.1002/adsc.201700901.

Pacinelli, E., Wang, L., Reeves, P.R., 2002. Relationship of Yersinia pseudotuberculosis O antigens IA, IIA, and IVB: The IIA gene cluster was derived from that of IVB. Infect. Immun. 70, 3271–3276. https://doi.org/10.1128/IAI.70.6.3271-3276.2002.

Rapp, C., van Overtveldt, S., Beerens, K., Weber, H., Desmet, T., Nidetzky, B., Zhou, N.-Y., 2020. Expanding the Enzyme Repertoire for Sugar Nucleotide Epimerization: The CDP-Tyvelose 2-Epimerase from Thermodesulfatator atlanticus for Glucose/Mannose Interconversion. Appl. Environ. Microbiol. 87 (4). https://doi.org/10.1128/AEM.02131-20.

Sakuraba, H., Kawai, T., Yoneda, K., Ohshima, T., 2011. Crystal structure of UDP-galactose 4-epimerase from the hyperthermophilic archaeon Pyrobaculum calidifontis. Arch. Biochem. Biophys. 512, 126–134. https://doi.org/10.1016/j.abb.2011.05.013.

Sato, T., Atomi, H., 2011. Novel metabolic pathways in Archaea. Curr. Opin. Microbiol. 14, 307–314. https://doi.org/10.1016/j.mib.2011.04.014.

Shin, S.M., Choi, J.M., Di Luccio, E., Lee, Y.J., Lee, S.J., Lee, S.J., Lee, S.H., Lee, D.W., 2015. The structural basis of substrate promiscuity in UDP-hexose 4-epimerase from the hyperthermophilic Eubacterium Thermotoga maritima. Arch. Biochem. Biophys. 585, 39–51. https://doi.org/10.1016/j.abb.2015.08.025.

Sievers, F., Higgins, D.G., 2018. Clustal Omega for making accurate alignments of many protein sequences. Protein Sci. 27, 135–145. https://doi.org/10.1002/pro.3290.

Szklarczyk, D., Gable, A.L., Lyon, D., Junge, A., Wyder, S., Huerta-Cepas, J., Simonovic, M., Doncheva, N.T., Morris, J.H., Bork, P., Jensen, L.J., Von Mering, C., 2019. STRING v11: Protein-protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. Nucleic Acids Res. 47, D607–D613. https://doi.org/10.1093/nar/gky1131.

Tanner, M.E., 2002. Understanding nature's strategies for enzyme-catalyzed racemization and epimerization. Acc. Chem. Res. 35, 237–246. https://doi.org/10.1021/ar000056y.

Thoden, J.B., Frey, P.A., Holden, H.M., 1996. Molecular structure of the NADH/UDP-glucose abortive complex of UDP-galactose 4-epimerase from Escherichia coli: Implications for the catalytic mechanism. Biochemistry 35, 5137–5144. https://doi.org/10.1021/bi9601114.

Thoden, J.B., Wohlers, T.M., Fridovich-Keil, J.L., Holden, H.M., 2000. Crystallographic evidence for Tyr 157 functioning as the active site base in human UDP-galactose 4-epimerase. Biochemistry 39, 5691–5701. https://doi.org/10.1021/bi000215l.

Tokuriki, N., Tawfik, D.S., 2009. Protein Dynamism and Evolvability. Science 324, 203–207. https://doi.org/10.1126/science.1169375.

Van Overtveldt, S., Da Costa, M., Gevaert, O., Joosten, H.J., Beerens, K., Desmet, T., 2020. Determinants of the Nucleotide Specificity in the Carbohydrate Epimerase Family 1. Biotechnol. J. 15, 1–8. https://doi.org/10.1002/biot.202000132.

Van Overtveldt, S., Verhaeghe, T., Joosten, H.J., van den Bergh, T., Beerens, K., Desmet, T., 2015. A structural classification of carbohydrate epimerases: From mechanistic insights to practical applications. Biotechnol. Adv. 33, 1814–1828. https://doi.org/10.1016/j.biotechadv.2015.10.010.

Verhees, C.H., Kengen, S.W., Tuininga, J.E., Schut, G.J., Adams, M.W., Vos, D., W.m, Oost, V. Der, J, 2003. The unique features of glycolytic pathways. Archaea. Biochem J 375, 231–246. https://doi.org/10.1042/BJ20021472.

Wolucka, B.A., Persiau, G., Van Doorsselaere, J., Davey, M.W., Demol, H., Vandekerckhove, J., Van Montagu, M., Zabeau, M., Boerjan, W., 2001. Partial purification and identification of GDP-mannose 3″,5″-epimerase of Arabidopsis thaliana, a key enzyme of the plant vitamin C pathway. Proc. Natl. Acad. Sci. U. S. A. 98, 14843–14848. https://doi.org/10.1073/pnas.011578198.

Zhu, H.M., Sun, B., Li, Y.J., Meng, D.H., Zheng, S., Wang, T.T., Wang, F.S., Sheng, J.Z., 2018. KfoA, the UDP-glucose-4-epimerase of Escherichia coli strain O5:K4:H4, shows preference for acetylated substrates. Appl. Microbiol. Biotechnol. 102, 751–761. https://doi.org/10.1007/s00253-017-8639-0.