Computational and Structural Biotechnology Journal 27 (2025) 2375–2385

Contents lists available at ScienceDirect

Computational and Structural Biotechnology Journal

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

Research Article

Protein flexibility drives sugar rotation and high substrate promiscuity in a GDP-sugar 4-epimerase

Carlos Josué Alvarez Quispe [a], Koen Beerens [a], Andy-Mark W.H. Thunnissen [b], Xevi Biarnés [c], Antoni Planas [c], Tom Desmet [a,*]

[a] Centre for Synthetic Biology (CSB), Department of Biotechnology, Ghent University, Coupure Links 653, Ghent 9000, Belgium
[b] Groningen Biomolecular Sciences and Biotechnology Institute, University of Groningen, Nijenborgh 4, Groningen 9747 AG, the Netherlands
[c] Laboratory of Biochemistry, Institut Químic de Sarrià, Universitat Ramon Llull, Via Augusta 390, Barcelona 08017, Spain

ARTICLE INFO

Keywords:
Carbohydrate epimerases (CEP1)
GDP-sugar 4-epimerase
UDP-galactose 4-epimerase
Nucleotide-sugars
Heptagonal box model
Sugar ring rotation
Molecular dynamics simulations
L-sugars

ABSTRACT

UDP-galactose 4-epimerases (Gal4Es) catalyze the inversion of the 4-hydroxyl configuration of a sugar moiety from an NDP-sugar through a three-step process: oxidation, rotation and reduction. Despite extensive biochemical and structural studies, the role of protein dynamics on substrate specificity remains poorly understood. The recently identified subgroup of GDP-sugar 4-epimerases, notable for its exceptional substrate promiscuity, provides an intriguing model to investigate the role of dynamics in the Gal4E catalytic mechanism and the unique promiscuity of the subgroup. In this study, we used a multidisciplinary approach to examine the dynamic-function relationships in the *Pyrococcus horikoshii* representative (*Ph*Gal4E_1). First, we determined several crystal structures (WT: 1.9–2.4 Å and Y145F: 3.1 Å), providing structural insights of the *Ph*Gal4E_1 structure bound to GDP-L-fucose in a catalytic conformation. To further explore the enzyme's promiscuity, *in silico* docking studies were conducted with three substrates, namely GDP-L-Fuc, GDP-Glc and UDP-Glc. Molecular dynamics simulations identified a dynamic hydrogen bond network surrounding the sugar moiety and phosphate groups, revealing four key residues: P80, H182, R83 and N174. These residues interact with either the substrate's sugar moiety (H182 and P80 with C2-OH and C3-OH, resp.) or diphosphate backbone (N174 and R83 with β-/α- and α-phosphate, resp.), which facilitates sugar ring positioning. Protein flexibility then initiates disruption of the hydrogen bonds enabling the required rotation of the intermediate. Site directed mutagenesis of these residues was performed to disrupt the interaction network followed by enzyme activity assays on the three substrates, validating their critical role in the epimerization reaction. These results highlight the pivotal role of protein flexibility in *Ph*Gal4E_1 promiscuity and establish a framework for dynamic studies across other Gal4E representatives.

1. Introduction

The diversity of carbohydrates results from their various arrangements occurring at different levels, including the number of carbon atoms of the saccharide ring, stereochemistry of the ring substituents, functional groups (e.g. absence of hydroxyl, as observed in deoxy sugars, or presence of specific substituents such as amino or sulfo groups) and polymerization. Carbohydrate epimerases (CEP) [1] are undoubtedly among the main architects in sugar stereochemistry as they allow the interconversion of abundant sugars (e.g. glucose, mannose) into their rare counterparts (e.g. L-sugars) [2,3], which find a myriad of applications in food, cosmetic, pharmaceutical and nutraceutical industries [2]. Through a single reaction, these CEPs can alter the orientation of a specific hydroxyl group within a sugar ring [4]. This process primarily involves nucleoside diphosphate sugars (NDP-sugars) [5], essential precursors for biosynthesis of polysaccharides [6,7], glycoproteins [8,9], glycolipids [10], vitamins [11] and antibiotics [12].

The CEP1 family belongs to the well-known Short-chain Dehydrogenase/Reductase (SDR) enzyme superfamily [1,13], specifically the group that acts on nucleotide sugars (NS). Within this family, UDP-galactose 4-epimerase (Gal4E, EC 5.1.3.2) is by far the most studied group, regarding biochemical characterization and structural analysis [1,14]. Like all CEP1 and NS-SDR members, Gal4Es consist of a Rossmann-fold domain for binding of the NAD^+ cofactor and contains

* Corresponding author.
E-mail address: Tom.Desmet@UGent.be (T. Desmet).

https://doi.org/10.1016/j.csbj.2025.05.037
Received 28 April 2025; Received in revised form 23 May 2025; Accepted 23 May 2025
Available online 27 May 2025

the catalytic triad [S/T]x_nYx_3K [15,16].

Gal4Es employ a transient keto intermediate mechanism resulting in an epimerization reaction through three steps: oxidation, rotation and reduction (Fig. 1A). Initially, in the oxidation phase, the catalytic tyrosine acts as a base in the form of a tyrosinate anion [17] and abstracts a proton from the hydroxyl group located at the C4 stereocenter of the sugar moiety. Simultaneously, the NAD^+ cofactor's nicotinamide ring acts as a hydride acceptor, producing a keto group at the C4-epimerization site [18,19]. Secondly, the keto intermediate undergoes a 180° rotation within the active site, hereby facing the opposite side of the keto-intermediate towards the tyrosine and NADH [20,21]. Thirdly, in the reduction step, the hydride from NADH is transferred back to the C4 carbon on the opposite side of the sugar plane due to previous rotation. The tyrosine proton, initially removed during oxidation, is simultaneously returned to the oxygen atom, forming a hydroxyl group. This completes the epimerization and regenerates the catalytic base [16]. While the oxidation and reduction steps are extensively documented in the literature [20,22], the rotation step remains insufficiently understood.

Crystallographic analyses of UDP-galactose and UDP-glucose in *Escherichia coli* Gal4E revealed distinct torsional angles at the glycosidic bond and β-phosphate [23]. This observation was crucial to support the rotation mechanism. Subsequent structural analyses reported conformational changes at the NDP-binding site upon substrate binding, [17,24] yet these changes were not linked to rotation. Dynamic studies on certain Gal4E representatives, principally in human and *E. coli* Gal4E, have provided insights into the flexibility of the binding pocket [25–27], but these endeavors remain insufficient to understand its possible impact on substrate specificity. Recent investigations of the UDP-glucuronic acid (UDP-GlcA) 4-epimerases (a CEP1 subgroup very similar to Gal4E) have underscored the significant role of protein dynamics in facilitating sugar ring rotation, as demonstrated through structural analyses, metadynamics and QM/MM simulations. [28,29] Despite these efforts, the role of protein dynamics in shaping Gal4E specificity and sugar rotation remains uncovered.

Research on Gal4Es and related NDP-sugar 4-epimerases has predominantly focused on their substrate specificity, as they can be active in a broad range of substrates [30–34]. A recent and extensive model for specificity is the "heptagonal box model" (HBM) of NS-SDR enzymes [16]. This model describes the substrate binding pocket of the NS-SDR enzymes (which includes the CEP1 members like Gal4E) and conceptualizes it as seven color-coded "walls", each containing conserved amino acid motifs surrounding the substrate's sugar moiety in the active site (Fig. 1B). Developed through sequence conservation analyses within the NS-SDR superfamily, this model facilitated the identification of a novel Gal4E subgroup in *Pyrococcus horikoshii* (*Ph*Gal4E_1, Uniprot: O73960), notable for its preference for GDP-sugars and pronounced promiscuity towards rare NDP-sugars such as GDP-L-fucose and GDP-L-galactose [35]. Yet, despite HBM's role in pinpointing *Ph*Gal4E_1, the specific contributions of residues within these "walls" to catalysis remain elusive, especially given the low sequence similarity among Gal4Es (as low as 20–30 %) [16]. Residues in these walls could play a catalytic role attributed to their structure and dynamics rather than only their sequence. In that account, a thorough structural and dynamic analysis of *Ph*Gal4E_1 can elucidate the importance of underexplored residues in the Gal4E mechanism, such as sugar rotation, while also revealing how dynamic residue interactions drive its promiscuous substrate specificity.

In our present study, we set out to gain a deeper understanding of the substrate specificity and the influence of protein flexibility in the mechanism of the promiscuous *Ph*Gal4E_1. We determined *Ph*Gal4E_1's X-ray crystal structure in the holoenzyme state (with NAD^+ cofactor) and as a Michaelis complex (MC) with NAD^+ cofactor and GDP-L-Fuc substrate. Subsequently, the obtained crystal structures were analyzed with molecular dynamics (MD) simulations to identify flexible regions corresponding to HBM motifs. Through extensive MD simulations, we were able to evaluate the influence of local flexibility on the sugar ring's rotation. Finally, site-directed mutagenesis studies and biochemical characterization of the created mutants were performed to confirm our *in silico* (MD) analyses. Based on this combined approach it was revealed that *Ph*Gal4E_1's dynamics play a crucial role in substrate promiscuity and modulates sugar rotation through a hydrogen bond network that influences the phosphates and sugar moiety.

2. Materials and methods

2.1. Experimental procedures

2.1.1. Materials and chemicals

Nucleotides and nucleotide-sugars were obtained from Biosynth (Compton, Berkshire, UK) and BOC Sciences (New York, USA). *E. coli* BL21(DE3) competent cells were prepared in-house. HPAEC-PAD standards such as galactose, L-glucose and quinovose (6-deoxy-D-glucose) were all purchased from Sigma-Aldrich (St. Louis, MO) and Biosynth

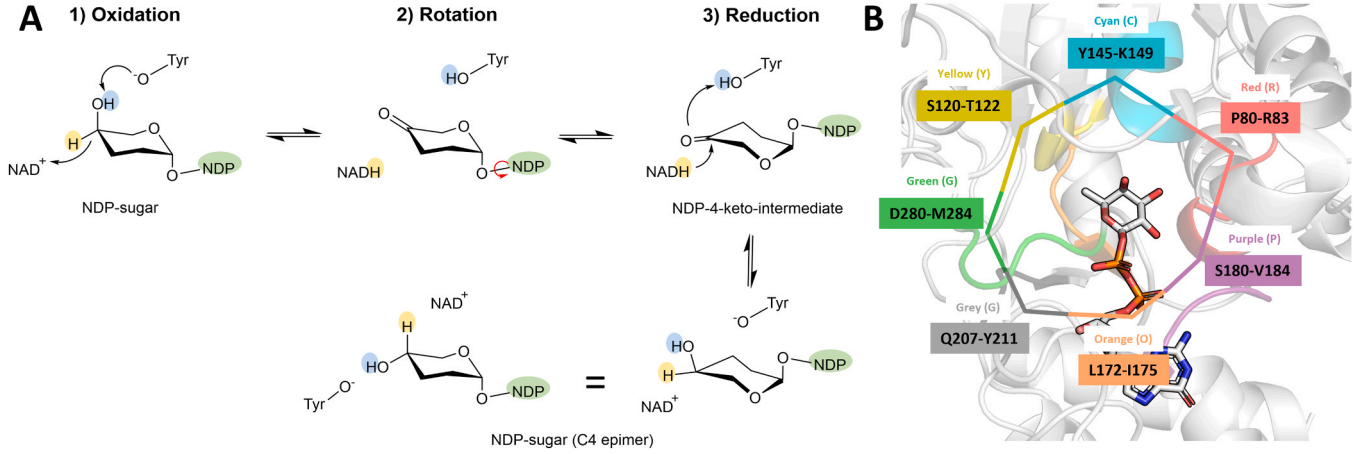

Fig. 1. Overview of the general keto intermediate mechanism for an NDP-sugar in Gal4Es and heptagonal box model (HBM) of *Ph*Gal4E_1. A) The epimerization reaction at the C4 stereocenter results in an equilibrium between the NDP-sugar and its C4 epimer. Hydride transfer to and from the NAD^+ cofactor is illustrated in yellow. Proton transfer to and from the catalytic tyrosinate is depicted in blue. Hydroxyl groups at other positions are omitted for the sake of clarity. B) The heptagonal box model (HBM) is sketched in the catalytic pocket of *Ph*Gal4E_1. The HBM is composed of seven conserved amino acid motifs, also called "walls". The catalytic triad ([S/T]x_nYx_3K) contains residues from the yellow ([S/T]) and the cyan wall (Yx_3K). Besides the yellow and cyan motifs being important for catalysis, additional "HBM walls" could play a significant role in catalysis, such us the required ring rotation.

(Compton, Berkshire, UK), stocks were prepared at 100 mM in ultrapure water. All other chemicals and reagents were of the highest available purity, unless stated otherwise.

2.1.2. Gene cloning, site-directed mutagenesis and transformation

The Gal4E_1 gene of *P. horikoshii* was codon optimized for *E.coli,* synthesized and subcloned into the pET21 vector at *Nde*I and *Xho*I restriction sites, providing a C-terminal His6-tag, by GeneArt Gene Synthesis (Thermo Fisher Scientific, Waltham, MA, USA) [35]. Mutants of *Ph*Gal4E_1 were constructed in the pET21 vector using the Q5 Site-Directed Mutagenesis Kit (NEB) according to manufacturer's instructions. The primers used for mutagenesis are listed in Table S1. *E. coli* BL21 (DE3) electrocompetent cells were transformed with the resulting constructs (WT and mutants) for protein expression.

2.1.3. Protein production

An overnight preculture of *E. coli* transformed with the respective expression plasmid was used to inoculate the Lysogeny Broth (LB) growth medium (250 mL) containing ampicillin (100 $\mu g \cdot mL^{-1}$) in a 1 L shake flask at 37°C. The culture was grown until an optical density OD_{600} of 0.6 was reached. Subsequently, 0.1 mM isopropyl-β-D-thiogalactopyranoside (IPTG) 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.

2.1.4. Protein purification

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 homogenization with glass beads (Sigma-Aldrich 150 – 212 μm) in 7 cycles of 15 sec each (FastPrep-24™high-speed benchtop homogenizer). Finally, cell debris was removed by centrifugation at 9000 rpm for 1 h.

Subsequently, 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 with 100 mM of 2-morpholinoethanesulfonic acid (MES, pH 6.5) by using Amicon Ultra-15 centrifugal filter units with 30 kDa cut-off (Merck Millipore Darmstadt, Germany).

To decrease the size heterogeneity, the protein solutions were further purified by means of size exclusion chromatography (SEC, Superdex 200 increase 10/300 GL), using an Akta pure system. Fractions were collected in a 96-well plate before pooling. For both protein WT and mutant Y145F, fractions were pooled according to their purity. The running buffer was PBS (pH=7.5, 50 mM disodium phosphate with 300 mM NaCl), being the buffer in which the protein samples 1, 3 and 4 were delivered (Figure S1).

The protein concentration was determined by measuring the absorbance at 280 nm with a NanoDrop2000 Spectrophotometer (Thermo Scientific) using extinction coefficient (ε_{280}=38120) which was calculated with ProteinCalculator v.3.4 tool (https://protcalc.sourceforge.net/). Molecular weight and purity of the protein were verified by sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE; 12 % gel) (Figure S2). The enzyme's electrophoretic behavior corresponded well with its predicted molecular mass of about 37 kDa.

2.1.5. Protein crystallography

Aliquots of purified wild-type *Ph*Gal4E_1 and variant Y145F were concentrated to 10 $mg \cdot mL^{-1}$ in 20 mM MOPS, 100 mM NaCl and 0.2 mM TCEP (Tris(2-carboxyethyl)phosphine). Crystallization conditions were identified using the sitting drop vapor diffusion technique, with the help of a Mosquito dispensing robot (SPTLabTech) and employing the PACT Premier, JCSG+ (Molecular Dimensions) and Index (Hampton Research) high-throughput crystallization screens. Drops of 200 nL were dispensed at two protein:reservoir volume ratios (1.25:0.75 and 0.75:1.25) and the plates were incubated at 21°C. Crystals grew within a few days and were used for X-ray data collection without further optimization. Successful crystallization solutions for wild-type *Ph*Gal4E_1 contained 25 % PEG 1500 in 0.1 M PCTP (sodium propionate, sodium cacodylate, Bis-Tris propane) buffer, pH 7.0 (PACT condition C4) or 20 % (w/v) PEG 3350, 0.2 M KSCN in 0.1 M Bis-Tris propane buffer, pH 7.5 (PACT condition G4). Crystals of variant Y145F were obtained with 50 % PEG500, 0.2 M NaCl and 0.1 M Na/K phosphate buffer, pH 6.2 (JCSG+ condition D3). To obtain a substrate-bound complex, wild-type and Y145F crystals were soaked for 1–3 minutes in crystallization solutions supplemented with 10 mM GDP-L-Fuc and flash cooled in liquid nitrogen. As a cryoprotectant, 25 % glycerol was added to the wild-type *Ph*Gal4E_1 crystals grown with PEG 3350.

X-ray diffraction data were collected at 100 K at the ID30A-1 (MASSIF-1) beamline of the European Synchrotron Radiation Facility (ESRF) in Grenoble, France. Initial data processing was performed using the automatic XDSAPP data processing pipeline [53] implemented at the synchrotron. Starting from the unmerged data, merging and reduction to unique structure factors was carried out using the AIMLESS [54] task from the CCP4 software suite (version 8.0). [55] Initial phases and structures were determined by molecular replacement with PHASER [56] using a structure generated by AlphaFold2 [57] as a search model. The structures were subsequently improved via rounds of manual model building using Coot [58], alternated with restrained refinement using Phenix.refine. [59] Waters and bound GDP-L-Fuc and/or NAD^+ were added at the final stages of model building and refinement. Geometry restraints for the co-factor and substrate were copied from the internal monomer library of the Phenix software suite (version 1.20.1–4487). [60] The geometries of the final structures were validated with MolProbity [61] and with the online wwPDB Validation System (https://validate.wwpdb.org). Crystal structures of substrate-free and substrate-bound wild-type *Ph*Gal4E_1 were determined at 1.9 Å and 2.4 Å resolution, respectively, obtained with crystals grown at different conditions (PACT-C4 for substrate-free, PACT-G4 for substrate-bound). Both crystal structures contain a dimer in the asymmetric unit. The electron density maps are well defined and allowed an unambiguous characterization of the substrate binding mode, including the sugar moiety. The substrate-bound crystal structure of variant Y145F was determined at 3.1 Å resolution and contains 2 dimers in the asymmetric unit. Relatively high atomic B-factors hamper the quality of this structure and the binding of GDP-L-Fuc is less well defined by the electron density than in the wild-type structure. A summary of the data collection and refinement statistics is available in Table S2. Coordinates and structure factors have been deposited in the Protein Data Bank.

2.1.6. WT and mutants' specific activity

Specific activities for nucleotide activated sugars (UDP-Glc, GDP-Glc and GDP-L-Fuc) were determined in 50 μL reaction volume containing 100 mM MES buffer (pH 6.5), 2 mM of substrate and the purified *Ph*Gal4E_1 WT or mutants (in a range of 0.005–1 $mg \cdot mL^{-1}$ to keep measurements within the linear range of the analysis). The reaction was performed at 60°C. and 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.

2.1.7. Kinetic parameters of PhGal4E_1

Enzyme specific activity incubated with various GDP-Glc, UDP-Glc and GDP-L-Fuc (0.5 – 2 mM) concentrations was evaluated in 100 mM MES pH 6.5 and 5 $\mu g \cdot mL^{-1}$ (for GDP-Glc), 50 $\mu g \cdot mL^{-1}$ (for GDP-L-Fuc) and 20 $\mu g \cdot mL^{-1}$ (for UDP-Glc) of *Ph*Gal4E_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).

2.1.8. HPAEC-PAD

Conversion of substrate to product was evaluated by high-performance anion exchange chromatography-pulsed amperometric detection (HPAEC-PAD) using the Dionex ICS-6000 system (Thermo Fischer Scientific) (CarboPac PA20 column-3 × 150 mm) (Figure S3). An isocratic flow (0.5 $mL \cdot min^{-1}$ for 15 min) of 10 mM NaOH (10 %) was used for *Ph*Gal4E_1 and mutants specific activity determination and kinetic characterization. The NDP-sugar conversion was quantified using standards of the respective monosaccharides and recalculated based on peak areas.

2.2. In silico procedures

2.2.1. Docking and protein visualization

The strategy consisted in using one subunit of the Michaelis complex crystal of *Ph*Gal4E_1 (PDB ID: 8RDH), removing the substrate GDP-L-fucose and perform the docking calculations in presence of GDP-L-Fuc, GDP-Glc and UDP-Glc. The structure of the ligand GDP-L-Fuc was taken from the *Ph*Gal4E_1 crystal (PDB ID: 8RDH). GDP-Glc and UDP-Glc ligands were directly obtained from the PDB database (PDB ID: 7XPT, crystal structure of NDP-pyranose mutase in complex with GDP-glucose) and (PDB ID: 1A9Y, crystal structure of UDP-galactose 4-epimerase in complex with UDP-glucose). Protein and ligand parametrization was performed with AutoDock Tools4 [62] using the AutoDock 4.2 atom typing. The grid box had custom parameters (center x = -25, center y = 20, center z = -5, size x = 25, size y = 30 and size z = 35). Docking calculations with flexible ligand were performed using AutoDock Vina 1.1.2 [62] with default settings. Reference binding modes were selected based on a dual criterion: low energy binding modes and catalytically competent orientations of the ligand. The results were analyzed and rendered using PyMOL v2.0 (open source; Schrödinger, LLC).

2.2.2. Molecular dynamics simulations

Molecular dynamics (MD) simulations were performed in three systems: (i) the holoenzyme (PDB ID: 8RDG): holo (enzyme + NAD^+), (ii) the Michaelis complex (PDB ID: 8RDH) with the substrate GDP-L-Fuc and NAD^+ (MC_S) and (iii) the Michaelis complex with the intermediate GDP-4-keto-L-Fuc and NADH (MC_I). Stability analysis of the three different systems was also performed in three replicates each (Figure S4). The protein was characterized using the Amber ff14SB force field [63], while the GLYCAM06 [64] force field described the L-fucose unit. The pyrophosphates and guanidine from the GDP moiety were defined by using the force field from Imberty [65]. NAD and NADH were parametrized with the force field from Ryde [66]. The force field parameters for the 4-keto-L-fucose moiety were obtained by following a reported procedure for 4-keto-hexuronic acid [29]. Bond and angle parameters were taken from GLYCAM06 and GAFF2. Dihedral angle parameters, not explicitly covered by GLYCAM06, were determined through relaxed scan calculations at the MP2/6–31 * level of theory with Gaussian09 [67]. The scan results for each dihedral angle were fitted to a molecular mechanics energy profile using the VMD force field toolkit plugin [68].

GROMACS v4.5.3 was used to apply a long MD simulation to each of the systems. The simulations were performed within a triclinic box (x: 7.29710 nm, y: 11.68141 nm, z: 10.49733 nm), with solvent explicit treatment (TIP3P) and neutral charge by adding convenient amount of counterions (Na^+ and Cl^-). The systems were equilibrated following a previous protocol [69] and simulations were conducted under the NVT ensemble. The MD simulation was extended to 500 ns for each of the systems. Root-mean-squared deviation analysis was employed to evaluate the system stability (Figure S4). MD analyses were conducted employing tools from Gromacs and complemented by VMD tools. Data analysis and plotting was performed in a Python Jupyter Notebook. Visualization and rendering of molecular structures were carried out by using PyMOL and VMD.

3. Results

3.1. Crystal structure determination and structural inspection of PhGal4E_1 unveil important flexible regions

Enzyme promiscuity is associated with an active site that exhibits flexibility, allowing the accommodation of diverse substrates [36,37]. It was therefore hypothesized that the catalytic pocket of *Ph*Gal4E_1 may also display an increased degree of flexibility compared to other Gal4Es. To deepen our understanding of the structural features in *Ph*Gal4E_1, its crystal structure was determined. The wild-type (WT) *Ph*Gal4E_1 and its inactive Y145F mutant (at the catalytic base) were recombinantly produced and His-Tag purified, followed by size exclusion chromatography before crystallization (Figure S1, S2). Considering the unique ability of *Ph*Gal4E_1 (and in general the GDP-Gal4E subgroup) to bind GDP-sugars with L-configuration in an active conformation (Fig. 2A, 2B), the enzyme was co-crystallized with GDP-L-Fuc. Notably, prior Gal4E structures predominantly featured UDP-sugars with D-configurations (e.g., UDP-Glc, UDP-Gal, UDP-GlcNAc).

Finally, three crystal structures were obtained: the holoenzyme with the NAD^+ cofactor of the WT (PDB ID: 8RDG, 1.9 Å resolution) as well as the Michaelis complex with the NAD^+ cofactor and GDP-L-Fuc bound in the active site of the WT (PDB ID: 8RDH, 2.4 Å resolution). A third crystal structure of the Michaelis complex of the inactivated Y145F mutant (disabled catalytic triad, PDB ID: 8RDI, 3.1 Å resolution) was obtained. The comparison of the WT and mutant Michaelis complexes revealed that they were structurally almost identical (backbone RMSD of 0.36) and no differences in the ligand conformations were observed. Although the WT structure was anticipated to also display the product (GDP-L-Quinovose), the short soaking duration may explain why no product formation was observed. To the best of our knowledge, these are the first reported Gal4E structures in complex with a GDP-sugar as well as an NDP-L-sugar, more specifically GDP-L-Fuc.

Similar to other Gal4E representatives, *Ph*Gal4E_1 is a homodimer in solution, each subunit comprises a N-terminal Rossman-fold domain (RFD) harboring a tightly bound NAD^+ cofactor, and a C-terminal NDP-sugar-binding domain (NBD) in which GDP-L-Fuc is bound. The epimerization reaction occurs at the interface of the two binding-domains (Fig. 2A). The unique preference of *Ph*Gal4E_1, and other homologs from the archaeal subgroup, for GDP-sugars is likely due to the increased number of hydrogen bonds formed with the GDP moiety compared to UDP (Figure S5). In contrast to other GDP-sugar epimerases, such as the GDP-mannose 3,5-epimerase from *Arabidopsis thaliana* (*At*GM35E) [38], *Ph*Gal4E_1 and other homologs such as *Pyrobaculum calidifontis* Gal4E display a T-shaped π-stacking between a tryptophan residue and the GDP moiety (Figure S5). This arrangement differs from the parallel π-stacking observed between a tryptophan and the GDP moiety of GDP-mannose in *At*GM35E [39].

The sugar moiety of the active site bound GDP-L-Fuc displayed a favorable binding mode for C4-epimerization. The L-fucose pyranose ring adopted an undistorted $^1C_4$ chair conformation, which is typical for this L-sugar [40]. However, it is distinct from other reported Gal4E substrates that predominantly exhibit D-sugars in the $^4C_1$ chair

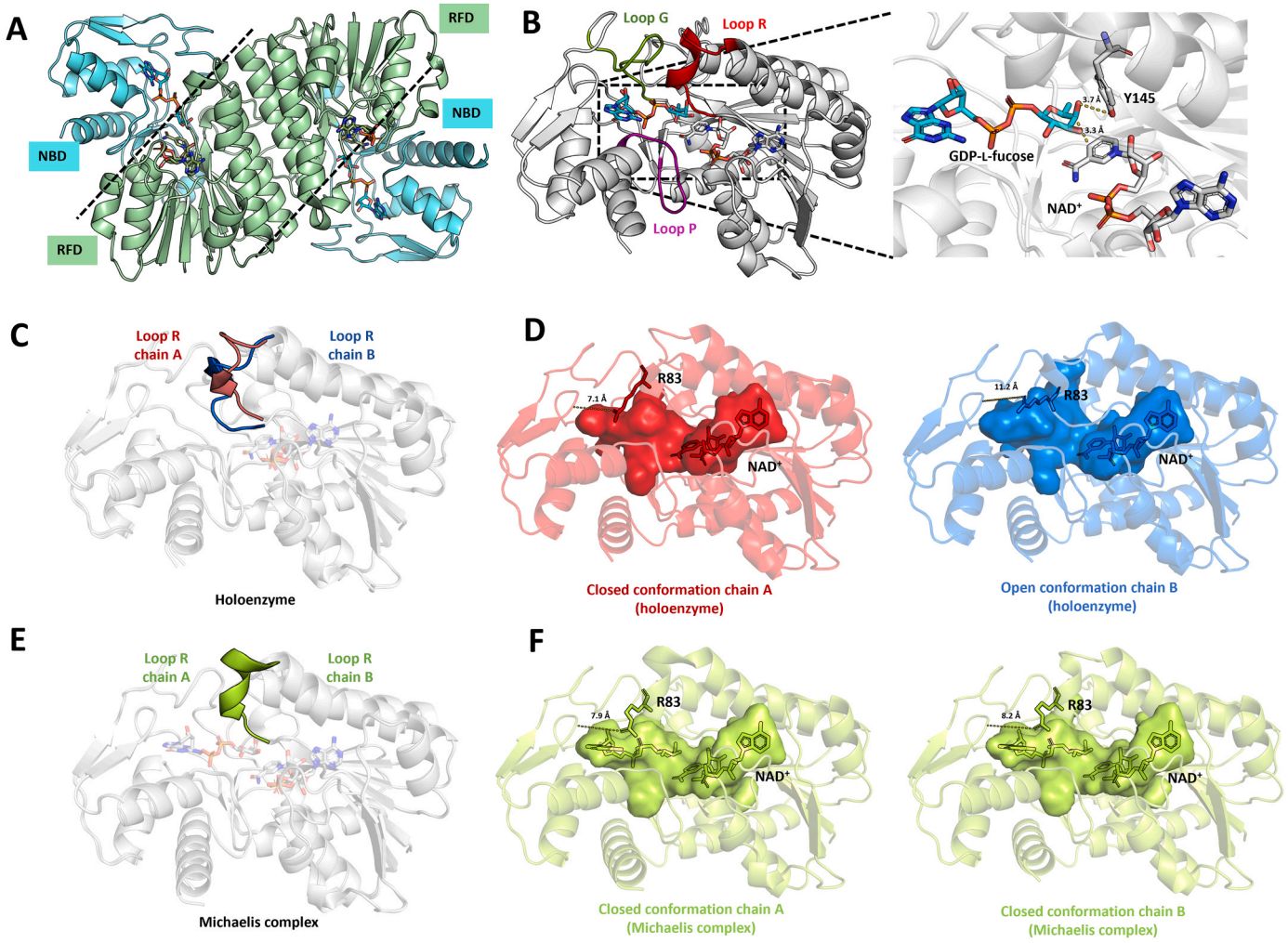

Fig. 2. Structural analysis of *Ph*Gal4E_1. A) Dimer representation of *Ph*Gal4E_1 complexed with GDP-L-Fuc. The NDP-binding domain (NBD) is represented in cyan and the Rossmann-fold domain (RFD) is represented in green. B) Active site of *Ph*Gal4E_1 surrounded by three loops (R (red), G (green) and P (purple)) and catalytic positioning of GDP-L-Fuc (in cyan) with Y145 and NAD^+ (in grey). C) Loop R comparison in chain A (in red) and chain B (in blue) of the holoenzyme. D) Closed and open conformations of the holoenzyme chain A (in red) and B (in blue), respectively, with the binding pocket delimited by R83. E) Loop R comparison in chain A and chain B (in green) of the Michaelis complex (with GDP-L-Fuc bound). F) Closed conformations of the Michaelis complex in chain A and B (in green), with the binding pocket delimited by R83. The width of the different pockets is indicated in Å.

conformation [23,24]. Nonetheless, despite the unusual conformation, the catalytic orientation of the substrate is conserved. The C4-OH is appropriately angled towards the catalytic tyrosinate (Y145), while the C4 hydride is facing the NAD^+ cofactor. The correct positioning of the L-sugar moiety in the active site emphasizes the adaptability of *Ph*Gal4E_1 in catalyzing C4-epimerization of unusual substrates.

Three important loops are present in the active site (Fig. 2B) and they are associated with three of the previously described HBM walls: P80–P89, G274–L285, and I176–V184, corresponding to the red (R), green (G) and purple (P) walls, respectively. Since loops are often responsible for dynamic properties in enzymes [41], due to their flexibility, these secondary structure elements might participate in conformational changes in the active site upon NDP sugar binding, as previously described for other Gal4Es [22]. The three aforementioned loops showed a high degree of flexibility, as evidenced by the higher B-factors in the holoenzyme crystal and the Michaelis complex (Figure S6, S7). Interestingly, the loop R adopted two different conformations in the dimer of the holoenzyme state (Fig. 2C), namely a closed conformation in chain A and an open conformation in chain B given the exposure of bulky R83 (Fig. 2D). This suggests that *Ph*Gal4E_1 already presented conformational changes even when the substrate is not bound. On the other hand, the Michaelis complex only displayed one closed conformation on both chains (Figs. 2E, 2F). These findings imply that *Ph*Gal4E_1's subunits undergo transitions between open and closed conformations, with the closed conformation being favored upon NDP-sugar binding. This suggests that conformational selection precedes the binding event, followed by an induced-fit mechanism. Similar binding mechanisms have been suggested for other nucleotide-sugar binding enzymes, such as glycosyltransferases [42]. However, these interpretations remain hypothetical, and further studies beyond the scope of this work will be required.

High flexibility of the loop R was also found in crystal structures from other representatives of the archaeal (and promiscuous) subgroup of GDP-sugar 4-epimerases such as *Archaeoglobus fulgidus* Gal4E (PDB ID: 3EHE) and *Pyrobaculum calidifontis* Gal4E (PDB ID: 3AW9) (Figure S8). Loop R plasticity was also observed in UDP-GlcA 4-epimerase [28], in which this structural element was suggested to have an intrinsic connection to the sugar ring rotation step and substrate specificity.

3.2. *Docking studies shed light on ring rotation mechanism*

To obtain molecular insights into the distinct substrate specificities of *Ph*Gal4E_1, GDP-Glc and UDP-Glc, as well as GDP-L-Fuc were docked into the crystallographic structure of the WT Michaelis complex after

removing the bound substrate (PDB ID: 8RDH). This structure was selected for flexible ligand docking because the Michaelis complex provides a more accurate side-chain configuration of the amino acids at the active site, facilitating a correct catalytic binding mode for the substrate. GDP-L-Fuc was docked and compared with the bound ligand in the crystal (Figure S9) as a control for the correct substrate positioning after docking. The binding modes of different substrates to which *Ph*Gal4E_1 has shown activity were investigated (Fig. 3). The selected substrates were found to be docked in an active conformation for initial oxidation (catalytic state). Some differences in the binding mode of the diphosphate linker of the docked substrates and the crystal structure of GDP-L-Fuc were observed (Figure S9), notably at the level of the torsion angles of the α- and β-phosphates. Nevertheless, both ligands displayed a correct catalytic positioning. This suggests the active site can accommodate distinct phosphate conformations while still enabling catalysis.

Interestingly, alternative binding modes of the three substrates were also identified, these being associated with sugar ring rotations in the catalytic pocket. For GDP-Glc and UDP-Glc, a torsion in the phosphoanhydride bond between the α- and β-phosphates was observed, leading to a shift in their positioning (Fig. 3). This reorientation could facilitate the rotation of the sugar moiety around the glycosidic bond between the glucose's anomeric carbon (C1) and the β-phosphate. For GDP-L-Fuc only rotation around the glycosidic bond was observed. However, torsional flexibility in both α and β is not discarded, as observed in the comparison of the docked and crystal structures of this ligand (Figure S9). This provides further insight into the substrate rotation along the catalytic itinerary of *Ph*Gal4E_1, which appears to involve an initial reorientation of the phosphates followed by rotation of the sugar moiety at the level of the glycosidic bond. It is important to emphasize that this rotated conformation is not catalytically active but reflects an equivalent state that the keto-intermediate is expected to adopt after the rotation.

The rotated conformations observed after docking slightly differ from the proposed mechanism for *E. coli* Gal4E (*Ec*Gal4E) in which the α- and β-phosphates are fixed by two arginines (Figure S10) and the sugar ring rotation occurs between the β-phosphate and glycosidic bond [22, 23]. This limited ligand mobility in the active site of *Ec*Gal4E could explain the strict substrate specificity of this enzyme. Docking data also implies that *Ph*Gal4E_1 can accommodate different binding modes of the ligand (catalytic and rotated) independently of their stereochemistry (L- or D-sugar). It should be noted that the catalytic parameters of these three substrates were determined (Table S3), confirming that the enzyme is active either on NDP-D-sugars or NDP-L-sugars despite the opposite chair form of the sugar moieties ($^4C_1$ vs $^1C_4$, resp.)

Taken altogether, we hypothesize that the active site of *Ph*Gal4E_1 imposes less constrains to the diphosphate backbone of the substrate compared to *Ec*Gal4E, allowing the accommodation of alternative ligand configurations, which could explain the higher substrate promiscuity of the enzyme (on the sugar moiety).

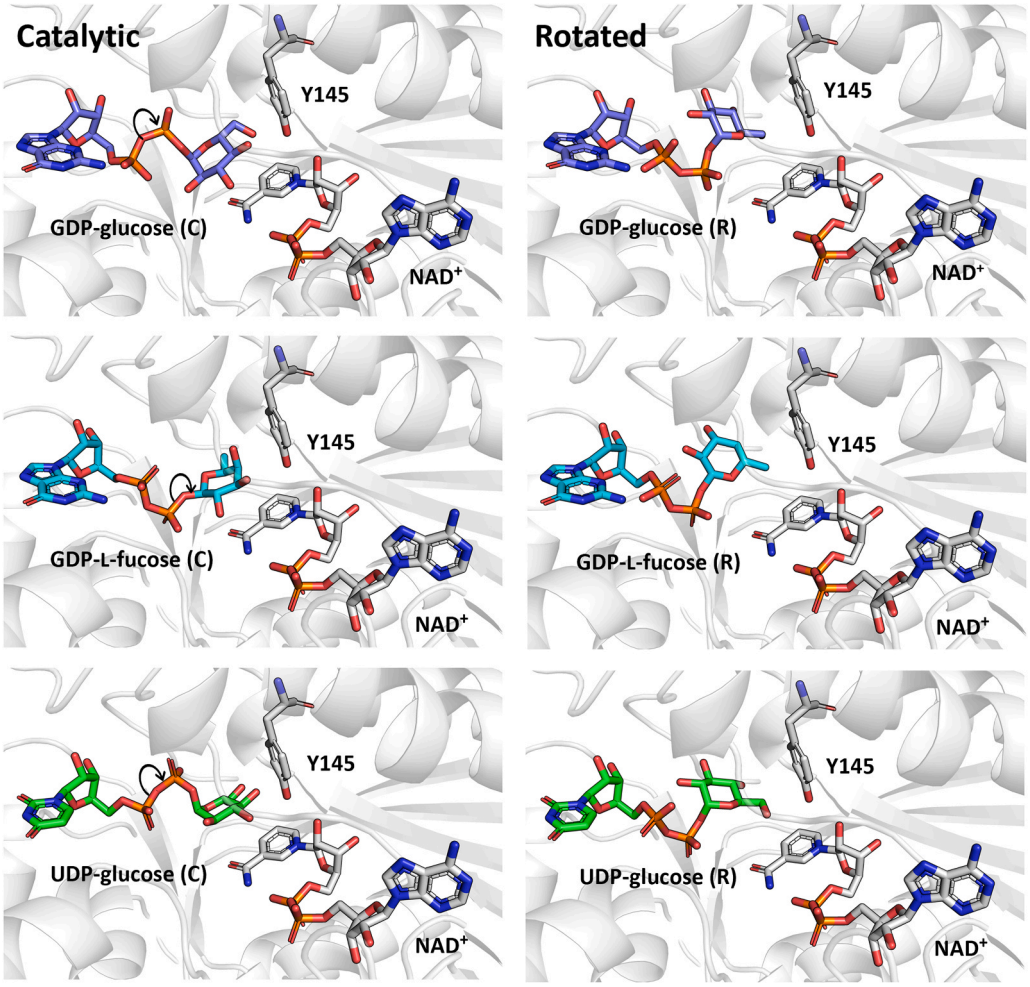

Fig. 3. Docking study of different NDP-sugars. Left side: docked binding modes for GDP-Glc (in purple), GDP-L-Fuc (in cyan) and UDP-Glc (in green) in a catalytic orientation (C). Right side: docked binding modes for GDP-Glc (in purple), GDP-L-Fuc (in cyan) and UDP-Glc (in green) in a rotated orientation (R). The catalytic Y145 and NAD^+ cofactor are depicted in grey. Direction of rotation is indicated by the black arrows.

3.3. The analysis of PhGal4E_1's plasticity through MD simulations provides insights into loop flexibility and the mechanism of rotation

The docking studies allowed the identification of distinct substrate binding modes in a rigid crystal structure, thus without taking into account protein dynamics. Therefore, the local flexibility of the loops and their influence in the rotation step were investigated subsequently. For this, classical molecular dynamics (MD) simulations on the dimeric form of *Ph*Gal4E_1 were performed. The three systems were simulated, each with three independent replicates (Figure S4): (i) the holoenzyme (enzyme bound to NAD^+), (ii) the Michaelis complex with the substrate GDP-L-Fuc (MC_S) and (iii) the Michaelis complex with the intermediate GDP-4-keto-L-Fuc (MC_I).

RMSF analysis of the three systems indicates high flexibility at the level of the three loops (Figure S11), this supports what was found in the structural analysis with the B-factors. Since loop movement is observed without the substrate (holoenzyme system), this data also reinforces the hypothesis that the intrinsic flexibility of the active site does not rely exclusively on a conformational change induced by substrate binding. Interestingly, when *Ph*Gal4E_1 has the substrate or intermediate bound (especially in the MC_I system), the flexibility in both active sites is slightly reduced. Based on these analyses and previous crystallographic observations, we hypothesized that the enzyme is in constant transition between open and closed conformations in each subunit. Once both subunits are saturated only the closed conformation is predominant and preferred for catalysis, hence the reduced plasticity.

We further investigated the substrate rotation mechanism in both MC_S and MC_I. To systematically study the rotation step we used the dihedral angle (ϕ), defined by the bond between N9 from the guanidine and C1 from the ribose and the distance between C2 and C5 both from the L-fucose and 4-keto-L-fucose moieties in their respective systems (Figure S12). Strikingly, a rigid-body rotation of the L-fucose ring was observed during the MD simulations in the MC_S (at 150 ns). The sugar ring rotates around the C1-O1 glycosidic bond (Fig. 4A, Video S1), which is consistent with previous docking studies. Initially, the ring is in a catalytic conformation (1), then it adopts a conformation with an angle of 120°, positioning itself vertically with respect to the dihedral plane (2) and followed by an additional 60° shift to complete the full 180° rotation (3). In all three replicates, similar rotations of the sugar ring were observed in at least one subunit of the enzyme dimer (Figure S13). Although this sugar rotation is not catalytically consistent with the *Ph*Gal4E_1 mechanism, as the keto-intermediate is expected to rotate, it demonstrates that the active site exhibits considerable flexibility, enabling rigid-body rotation.

Since the mechanism requires a 4-keto-sugar to rotate, we performed MD simulations on the MC_I system. The 4-keto-L-Fuc moiety transitions from a $^1C_4$ chair (typical for L-fucose), to a skew-boat conformation (^2S_O) (Video S2), likely to mitigate steric clashes between the planar C4

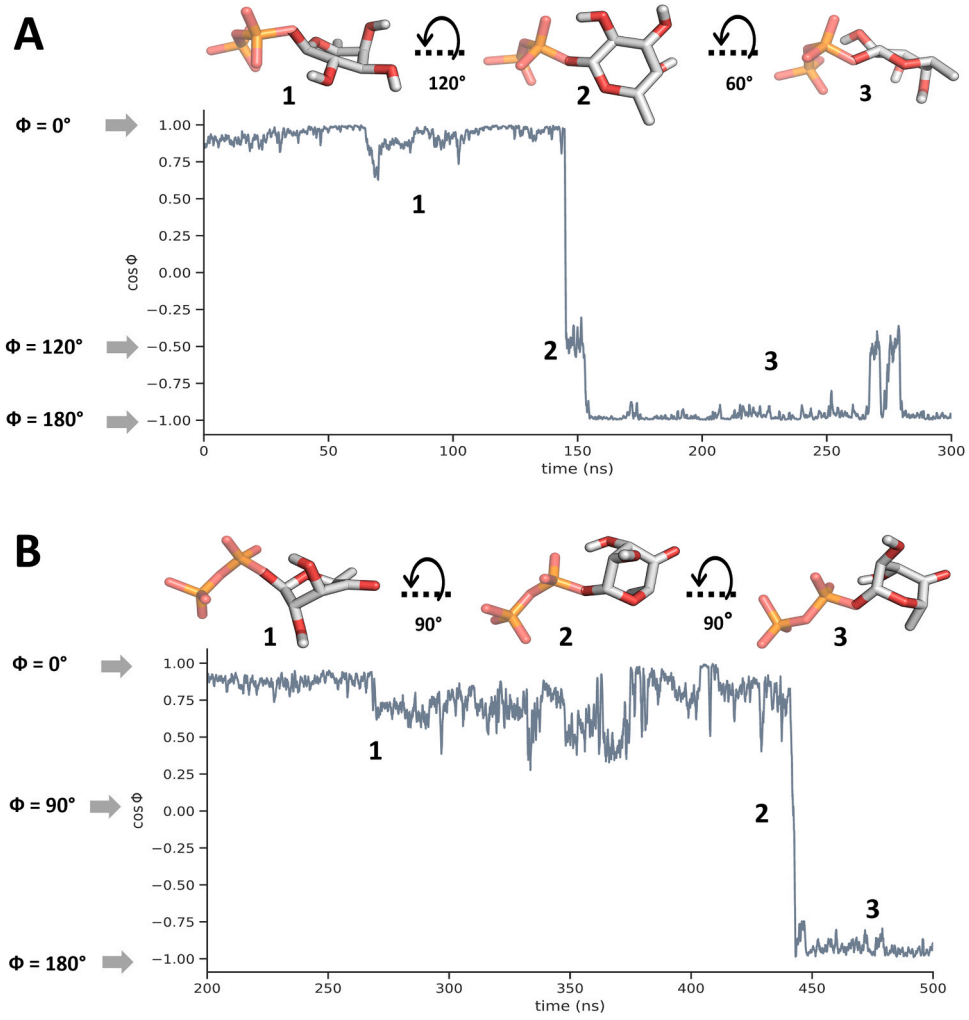

Fig. 4. Molecular dynamics (MD) simulation of *Ph*Gal4E_1 in MC_S and MC_I systems. A) Evolution of the GDP-L-Fuc rotation during the simulation. For simplicity, only the initial 300 ns of simulation are shown (see full extended graphic in Fig. S13). B) Evolution of the GDP-4-keto-L-Fuc rotation during the simulation (see full extended graphic in Fig. S14). Different conformations during the rotation of GDP-L-Fuc and GDP-4-keto-L-Fuc were represented: (1) catalytic, (2) vertical and (3) rotated conformation. The cos of the dihedral angle (ϕ) is depicted in grey and describes the ring rotation from an angle of 0° (cos(ϕ) = 1) to 180° (cos(ϕ) = -1).

carbonyl and the C3-OH. Subsequently, the rotation was observed at 440 ns (Fig. 4B, Video S3). The keto-sugar ring initially adopts a catalytic conformation (1), then rotates to a 90° angle (2), before immediately undergoing a second 90° shift to complete a 180° rotation (3). In all three replicates, similar rotations of the sugar ring were observed in at least one subunit of the enzyme dimer (Figure S14). Notably, a similar ring puckering preceding rotation was previously observed in MD simulations of a UDP-glucuronic acid (UDP-GlcA) 4-epimerase [29]. It was postulated that this conformational rearrangement reduces the accessible volume of the sugar and promote its rotation. Consistent with the prior study, interactions between the β-phosphate and the axial C2-OH (Figure S15) were observed during the rotation, potentially facilitating or stabilizing the rotated conformation.

3.4. *In silico analysis of the enzyme-substrate interactions network together with mutant characterization reveals crucial residues*

The rotation of the GDP-4-keto-L-sugar moiety suggest that there might be some interactions, primarily hydrogen bonds and salt bridges, assisting this motion. We therefore analyzed the different interaction distances between the residues and the intermediate considering a threshold of maximum 4 Å for a weak interaction. Since the GDP moiety is considered as an anchoring system that remains stable during the reaction, only the interactions of the enzyme with the phosphates and fucose were analyzed from the MD trajectory.

Regarding the phosphates, N174 forms weak hydrogen bonds (distance ~ 4 Å) with α-phosphate and stronger interactions (distance < 3.5 Å) with β-phosphate throughout the 500 ns simulation, indicating a role in stabilizing phosphate orientation in *Ph*Gal4E_1 (Figure S16). R83 plays a similar but less pronounced role, primarily forming weak hydrogen bonds and ionic interactions with the α-phosphate, but not with the β-phosphate (Figure S16). This weak coordination in *Ph*Gal4E_1, compared to previously discussed *Ec*Gal4E, allows torsional flexibility in the phosphoanhydride bond between the α- and β-phosphates (Figure S17). These results were consistent with the previous observations from the docking analysis, suggesting that *Ph*Gal4E_1 exhibits high flexibility in the diphosphate backbone surrounding region.

To further elucidate the hydrogen bond network around the 4-keto-L-Fuc ring, simulations identified three critical residues: Y145 (catalytic), P80 (in loop R), and H182 (in loop P). Before the intermediate rotates (390 – 440 ns), the catalytic distance between the Y145-OH group and the C4 of the intermediate is typically under 3.5 Å, ideal for efficient reduction. Likewise, the distance between NADH's reactive hydride and the C4 of the intermediate remains sufficiently close, supporting effective hydride transfer (Fig. 5A). Distance analysis reveals that the carbonyl group of the P80 backbone, serving as a hydrogen bond acceptor for the C3-OH group, helps stabilize the intermediate's position (first 310 ns, Figure S18). The H182 side chain also contributes by forming hydrogen bonds with the C2-OH group (first 250 ns, Figure S18). During rotation, distances between the intermediate and NADH, Y145, P80, and H182 exceed the threshold (Figs. 5A, 5B), leading to the disruption of hydrogen bonds and increasing 4-keto-L-Fuc rotational freedom (Fig. 5C). Notably, residues interacting with the diphosphate backbone and the keto-intermediate are highly conserved within the GDP-Gal4E subgroup (Figure S19). These interactions, mostly weak and situated in flexible loops, readily break, facilitating the intermediate's rotation.

To validate the *in silico* analysis, the relative activity on UDP-Glc, GDP-Glc and GDP-L-Fuc of three alanine mutants (N174A, R83A and H182A) was determined and compared to the wild-type (WT) (Fig. 5D). The mutants were created to disrupt the interactions between the side chains and GDP-4-keto-L-Fuc.

The N174A mutant was almost completely inactive on UDP-Glc, showed decreased activity on GDP-Glc, and retained almost all

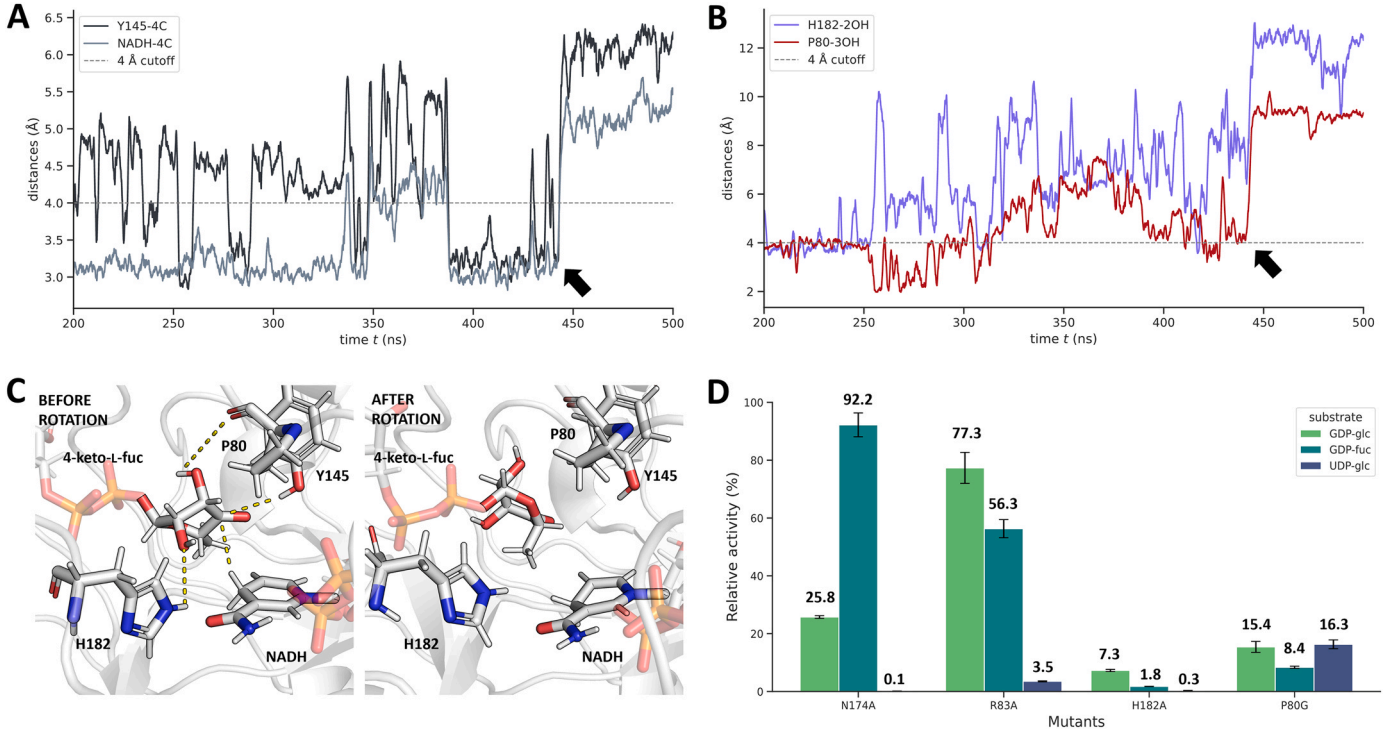

Fig. 5. Interaction distance analysis and site-directed mutagenesis of the selected mutants. A) Catalytic distances during 500 ns of MD simulations between the Y145 (dark grey) and NADH (grey) relative to the keto-intermediate. B) Hydrogen bond evolution of H182 with C2-OH (in purple) and P80 with C3-OH (in red) over 500 ns of simulation. In A and B the black arrow indicates the start of the rotation (440 ns). Only the last 300 ns are represented (see full extended graphic in Fig. S18). C) Interacting hydrogen bond network around the 4-keto-L-Fuc before and after rotation. D) Activity of four site directed mutants on three substrates (2 mM of GDP-Glc, GDP-L-Fuc or UDP-Glc, at pH 6.5 and 60 °C) relative to that of the WT. These are the average values of 3 measurements with values shown at the top of the bars and standard deviations.

activity on GDP-L-Fuc. While the activity on GDP-L-Fuc is hardly affected, the activity on GDP-Glc was reduced to only 25 % (compared to the wild-type). This could suggest that α-phosphate stabilization is more important for GDP-Glc positioning than for GDP-L-Fuc. Additionally, N174 might also interact with the C6-OH group of GDP-Glc (Figure S20), which is absent in GDP-L-Fuc. This could potentially explain the big difference in activity on both GDP-sugars. Conversely, the activity on UDP-Glc is nearly abolished (0.1 %) suggesting a detrimental effect of the N174A mutation on UDP-Glc positioning.

Similarly, the R83A mutation only slightly affected the enzyme's activity on GDP-sugars, retaining relative activities of 77 % on GDP-Glc and 56 % on GDP-L-Fuc. The activity on UDP-Glc on the other hand, was strongly reduced to only 3.5 % of the wild-type activity. This suggests that the stabilization of β-phosphate is also essential for the proper positioning and rotation of UDP-Glc. Significant deleterious effects on the three substrates can be observed for H182A, with 7.3 %, 1.8 % and 0.3 % of the WT activity on GDP-Glc, GDP-Fuc and UDP-Glc, respectively. This indicates the importance of H182 for sugar ring positioning and rotation, regardless of the nature of the NDP and sugar moieties.

As mentioned during the hydrogen bonding analysis, the backbone carbonyl group of P80 is likely crucial for hydrogen bonding with the intermediate, therefore a fourth variant (P80G) was created to further investigate the role of this proline's backbone carbonyl. Prolines are known for their stabilizing effect, as they maintain rigidity in proteins [43], whereas glycines introduce flexibility [44]. Thus, the substitution of the proline in position 80 with a glycine aimed to destabilize the interactions within loop R. This mutation resulted in a reduction of activity for the three substrates (8.4 – 16.3 %, Fig. 5D), which indicates that disruption of this interaction also affects the positioning of the sugar ring in general.

In summary, we observed major effects on UDP-Glc activity compared to those on GDP-Glc and GDP-L-Fuc. This can be attributed to a weaker binding of *Ph*Gal4E_1 on UDP-substrates compared to GDP-substrates (Figure S5). When the weak interactions with the phosphates and the sugar moiety are disrupted, a premature rotation of the sugar ring may take place without prior oxidation. However, this assumption remains to be explored in more detail by performing binding studies. It should be emphasized that H182 and P80 may play a crucial role in stabilizing the sugar moiety, facilitating reaction initiation by sugar oxidation and rotation.

4. Discussion

Summarized, this investigation employed a multidisciplinary strategy to explore the dynamic-function relationships of the *Pyrococcus horikoshii* GDP-sugar 4-epimerase (*Ph*Gal4E_1), yielding significant insights into its catalytic mechanism and promiscuity. Crystal structures of *Ph*Gal4E_1 (holoenzyme and Michaelis complex with GDP-L-fucose) were determined in open and closed states, serving as a foundation for exploring protein flexibility. These structural findings were complemented by *in silico* docking studies with three substrates (GDP-L-Fuc, GDP-Glc, and UDP-Glc), giving insights into *Ph*Gal4E_1 substrate binding and promiscuity. MD simulations reinforced the inherent protein flexibility and uncovered a dynamic hydrogen bond network surrounding the sugar moiety and phosphate groups, pinpointing four key residues as essential for sugar ring positioning and keto-intermediate rotation. While H182 and P80 interact with two neighboring hydroxylgroups of the substrate's sugar moiety (C2-OH and C3-OH, resp.), the other two residues (N174 and R83) interact with the diphosphate backbone. These findings were substantiated through site directed mutagenesis, which disrupted the identified interactions, followed by enzyme activity assays of the resulting variants on the three substrates, confirming the pivotal role of these residues (especially P80 and H182) in the epimerization reaction.

Pioneering endeavors in Gal4E studies, principally conducted by Thoden and Holden, on human and *E. coli* Gal4Es [19,23,45] elucidated its catalytic mechanism, particularly the oxidation and reduction steps, and facilitated the study of the specificity pattern of these enzymes. These investigations opened the door for future studies on Gal4Es from different origins ranging from eukaryotes, bacteria and archaea [14,18,46]. Since then, several Gal4E structures have been reported, exclusively in complex with UDP-sugars, due to their role in the Leloir pathway [47,48], which involves only this type of substrate. This demonstrates how focused Gal4Es research was on UDP-sugars, ruling out the possibility of promiscuity towards other NDP-sugars that also participate in important metabolic pathways, such as other NDP-Glc or even more special NDP-sugars like GDP-L-Fuc. In this study, we presented crystal structures of a Gal4E in complex with a GDP-L-sugar. Through docking and biochemical characterization, we also verified the ability of the archaeal enzyme *Ph*Gal4E_1 to bind GDP-L-Fuc in an optimal catalytic conformation as well as its affinity and activity. Our findings not only challenge the established paradigm of Gal4E functionality restricted to UDP-sugars but also raises additional questions about the potential involvement of these enzymes in alternative pathways.

Furthermore, our study revealed the importance of a hydrogen bond network involving the diphosphate backbone and the sugar moiety of the substrate, which in turn plays a role in facilitating catalysis and potentially contributing to substrate promiscuity. These interactions are essential for maintaining the substrate in an optimal orientation for the sugar oxidation step (first step of the catalytic mechanism). However, breakdown of these interactions is necessary to allow the sugar ring rotation. In the case of *Ph*Gal4E_1, this is achieved by the inherent flexibility of active site loops. Initially, the hydrogen bond interactions stabilize the NDP-sugar in a catalytic position, necessary for oxidation. Then, protein dynamics disrupts the weak interactions to enable the keto-intermediate rotation. Although N174 forms strong interactions with Pβ and weaker ones with Pα, this leads to loose coordination of the diphosphate backbone. Consequently, the phosphates can adopt various conformations, providing the flexibility to accommodate different NDP-sugars in optimal orientations at the active site.

The previous observation contrasts with the well-studied *E. coli* Gal4E, where the two phosphate groups are tightly coordinated by two arginine residues (R231 and R292) [19,49], leading to more stringent substrate specificity. Additionally, the hydrogen bond interaction network around the sugar moiety in *Ph*Gal4E_1 is easily disrupted because P80 and H182 are both located in flexible loops (R and P, respectively). Therefore, we suggest that the weak interactions between these active site residues with the phosphates and the sugar ring enhance the plasticity of *Ph*Gal4E_1 active site, which likely leads to its broader substrate promiscuity. Future studies could explore how this interaction network has evolved, from promiscuous to more specialized Gal4Es, offering insights into the emergence of epimerase specificity.

The heptagonal box model of NS-SDRs [16] offers a useful approach for investigating key regions in Gal4E based on sequence, structure and coevolution analyses. However, this model should be complemented by dynamic studies, as enzymes are characterized by their inherent flexibility and not only the specificity fingerprints. Although significant experimental data has been collected, including structural and biochemical analyses, dynamic-function relationships in Gal4Es have been largely overlooked and only recent investigations have provided insightful information [29,50]. Recent breakthroughs in protein structure prediction [51], coupled with cutting-edge computational techniques like MD simulations [52] and experimental validation, should be applied to a broader range of Gal4E representatives (and by extension other CEP1 and NS-SDR enzymes) to deepen our understanding of the dynamic-function relationships in these intriguing enzymes.

CRediT authorship contribution statement

Tom Desmet: Writing – review & editing, Validation, Supervision, Resources, Funding acquisition, Conceptualization. **Andy-Mark W.H. Thunnissen:** Resources, Investigation, Data curation. **Koen Beerens:**

Writing – review & editing, Validation, Supervision, Conceptualization. **Antoni Planas:** Writing – review & editing, Validation, Supervision, Resources. **Xevi Biarnés:** Writing – review & editing, Validation, Supervision, Software, Resources, Conceptualization. **Alvarez Quispe Carlos Josue:** Writing – review & editing, Writing – original draft, Visualization, Software, Methodology, Investigation, Formal analysis, Data curation, Conceptualization.

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

This research was funded by the Research Foundation – Flanders (FWO) through a doctoral scholarship for C.J.A.Q. (grant 1105922N) and the DeoxyBioCat project (grant G0A7520N, jointly funded by the FWO and Austrian FWF). C.J.A.Q. also thanks FWO for funding for a research stay at IQS-Barcelona (grant V442822N) and to A.P. for grant PID2022–138252OB-I00 from MINECO, Spain. We would also like to acknowledge the European Synchrotron Radiation Facility (ESRF) for provision of synchrotron radiation facilities and thank the scientists at beamline MASSIF-1 for their support. The computational resources (Stevin Supercomputer Infrastructure) and services used in this work were provided by the VSC (Flemish Supercomputer Center), funded by Ghent University, FWO and the Flemish Government – department EWI. Finally, the authors would like to thank UGent's "HTS for SynBio" Core Facility for training, support and access to the instrument park. This Core Facility is supported by UGent-BOF through grant BOF/COR/2022/002.

Appendix A. Supporting information

Supplementary data associated with this article can be found in the online version at doi:10.1016/j.csbj.2025.05.037.

References

[1] Van Overtveldt S, Verhaeghe T, Joosten HJ, van den Bergh T, Beerens K, Desmet T. A structural classification of carbohydrate epimerases: from mechanistic insights to practical applications. Biotechnol Adv 2015;33(8):1814–28. https://doi.org/10.1016/j.biotechadv.2015.10.010.

[2] Beerens K, Desmet T, Soetaert W. Enzymes for the biocatalytic production of rare sugars. J Ind Microbiol Biotechnol 2012;39(6):823–34. https://doi.org/10.1007/s10295-012-1089-x.

[3] Beerens K, Van Overtveldt S, Desmet T. The 'Epimerring' highlights the potential of carbohydrate epimerases for rare sugar production. Biocatal Biotransform 2017;35(3):230–7. https://doi.org/10.1080/10242422.2017.1306738.

[4] Tanner ME. Understanding nature's strategies for enzyme-catalyzed racemization and epimerization. Acc Chem Res 2002;35(4):237–46. https://doi.org/10.1021/ar000056y.

[5] Dolan JP, Cosgrove SC, Miller GJ. Biocatalytic approaches to building blocks for enzymatic and chemical glycan synthesis. JACS Au 2023;3(1):47–61. https://doi.org/10.1021/jacsau.2c00529.

[6] Yin Y, Huang J, Gu X, Bar-Peled M, Xu Y. Evolution of plant nucleotide-sugar interconversion enzymes. PLoS One 2011;6(11). https://doi.org/10.1371/journal.pone.0027995.

[7] Marmont LS, Whitfield GB, Pfoh R, Williams RJ, Randall TE, Ostaszewski A, Razvi E, Groves RA, Robinson H, Nitz M, Parsek MR, Lewis IA, Whitney JC, Harrison JJ, Howell PL. PelX Is a UDP-N-Acetylglucosamine C4-Epimerase Involved in Pel Polysaccharide-Dependent Biofilm Formation. J Biol Chem 2020;295(34):11949–62. https://doi.org/10.1074/jbc.RA120.014555.

[8] Eichler J. Extreme sweetness: protein glycosylation in archaea. Nat Rev Microbiol 2013;11(3):151–6. https://doi.org/10.1038/nrmicro2957.

[9] Andresen S, de Mojana di Cologna N, Archer-Hartmann S, Rogers AM, Samaddar S, Ganguly T, Black IM, Glushka J, Ng KKS, Azadi P, Lemos JA, Abranches J, Szymanski CM. Involvement of the Streptococcus Mutans PgfE and GalE 4-Epimerases in Protein Glycosylation, Carbon Metabolism, and Cell Division. Glycobiology 2023;33(3):245–59. https://doi.org/10.1093/glycob/cwad004.

[10] Pacinelli E, Wang L, Reeves PR. Relationship of Yersinia Pseudotuberculosis O Antigens IA, IIA, and IVB: The IIA Gene Cluster Was Derived from That of IVB. Infect Immun 2002;70(6):3271–6. https://doi.org/10.1128/IAI.70.6.3271-3276.2002.

[11] Wolucka BA, Persiau G, Van Doorsselaere J, Davey MW, Demol H, Vandekerckhove J, Van Montagu M, Zabeau M, Boerjan W. 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 USA 2001;98(26):14843–8. https://doi.org/10.1073/pnas.011578198.

[12] Hoffmeister D, Ichinose K. 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 2000;7:821–31. https://doi.org/10.1016/S1074-5521(00)00029-6.

[13] Kavanagh KL, Jörnvall H, Persson B, Oppermann U. Medium- and short-chain dehydrogenase/reductase gene and protein families: the SDR superfamily: functional and structural diversity within a family of metabolic and regulatory enzymes. Cell Mol Life Sci 2008;65(24):3895–906. https://doi.org/10.1007/s00018-008-8588-y.

[14] Shin SM, Choi JM, Di Luccio E, Lee YJ, Lee SJ, Lee SJ, Lee SH, Lee DW. The structural basis of substrate promiscuity in UDP-Hexose 4-epimerase from the hyperthermophilic eubacterium thermotoga maritima. Arch Biochem Biophys 2015;585:39–51. https://doi.org/10.1016/j.abb.2015.08.025.

[15] Gräff M, Buchholz PCF, Stockinger P, Bommarius B, Bommarius AS, Pleiss J. The short-chain dehydrogenase/reductase engineering database (SDRED): a classification and analysis system for a highly diverse enzyme family. Proteins Struct Funct Bioinforma 2019;87(6):443–51. https://doi.org/10.1002/prot.25666.

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

[17] Allard STM, Giraud MF, Naismith JH. Epimerases: structure, function and mechanism. Cell Mol Life Sci 2001;58(11):1650–65. https://doi.org/10.1007/PL00000803.

[18] Nam YW, Nishimoto M, Arakawa T, Kitaoka M, Fushinobu S. Structural basis for broad substrate specificity of UDP-glucose 4-epimerase in the human milk oligosaccharide catabolic pathway of bifidobacterium longum. Sci Rep 2019;9(1):11081. https://doi.org/10.1038/s41598-019-47591-w.

[19] Thoden JB, Frey PA, Holden HM. Molecular structure of the NADH/UDP-glucose abortive complex of UDP-galactose 4-epimerase from escherichia coli: implications for the catalytic mechanism. Biochemistry 1996;35(16):5137–44. https://doi.org/10.1021/bi9601114.

[20] Frey PA, Hegeman AD. Chemical and stereochemical actions of UDP-galactose 4-epimerase. Acc Chem Res 2013;46(7):1417–26. https://doi.org/10.1021/ar300246k.

[21] Thoden JB, Hegeman AD, Wesenberg G, Chapeau MC, Frey PA, Holden HM. Structural analysis of UDP-sugar binding to UDPp-galactose 4-epimerase from escherichia coli. Biochemistry 1997;36(21):6294–304. https://doi.org/10.1021/bi970025j.

[22] Samuel J, Tanner ME. Mechanistic aspects of enzymatic carbohydrate epimerization. Nat Prod Rep 2002;19(3):261–77. https://doi.org/10.1039/b100492l.

[23] Thoden JB, Holden HM. Dramatic differences in the binding of UDP-galactose and UDP-glucose to UDP-galactose 4-epimerase from escherichia coli. Biochemistry 1998;37(33):11469–77. https://doi.org/10.1021/bi9808969.

[24] Thoden JB, Wohlers TM, Fridovich-Keil JL, Holden HM. Crystallographic evidence for Tyr 157 functioning as the active site base in human UDP-galactose 4-epimerase. Biochemistry 2000;39(19):5691–701. https://doi.org/10.1021/bi000215l.

[25] Kumar SU, Sankar S, Kumar DT, Younes S, Younes N, Siva R, Doss CGP, Zayed H. Molecular dynamics, residue network analysis, and cross-correlation matrix to characterize the deleterious missense mutations in GALE causing galactosemia III. Cell Biochem Biophys 2021;79(2):201–19. https://doi.org/10.1007/s12013-020-00960-z.

[26] Timson DJ, Lindert S. Comparison of dynamics of wildtype and V94M human UDP-galactose 4-epimerase-a computational perspective on severe epimerase-deficiency galactosemia. Gene 2013;526(2):318–24. https://doi.org/10.1016/j.gene.2013.05.027.

[27] Friedman AJ, Durrant JD, Pierce LCT, Mccorvie TJ, Timson DJ, Mccammon JA. The molecular dynamics of trypanosoma brucei UDP-galactose 4′-epimerase: a drug target for African sleeping sickness. Chem Biol Drug Des 2012;80(2):173–81. https://doi.org/10.1111/j.1747-0285.2012.01392.x.

[28] Iacovino LG, Savino S, Borg AJE, Binda C, Nidetzky B, Mattevi A. Crystallographic snapshots of UDP-glucuronic acid 4-epimerase ligand binding, rotation, and reduction. J Biol Chem 2020;295(35):12461–73. https://doi.org/10.1074/jbc.ra120.014692.

[29] Borg AJE, Esquivias O, Coines J, Rovira C, Nidetzky B. Enzymatic C4-epimerization of UDP-glucuronic acid: precisely steered rotation of a transient 4-keto intermediate for an inverted reaction without decarboxylation. Angew Chem - Int Ed 2023;62(4):1–7. https://doi.org/10.1002/anie.202211937.

[30] Ishiyama N, Creuzenet C, Lam JS, Berghuis AM. Crystal Structure of WbpP, a genuine UDP-N-acetylglucosamine 4-epimerase from pseudomonas aeruginosa: substrate specificity in UDP-hexose 4-epimerases. J Biol Chem 2004;279(21):22635–42. https://doi.org/10.1074/jbc.M401642200.

[31] Demendi M, Ishiyama N, Lam JS, Berghuis AM, Creuzenet C. Towards a better understanding of the substrate specificity of the UDP-N-Acetylglucosamine C4 Epimerase WbpP. Biochem J 2005;389(1):173–80. https://doi.org/10.1042/BJ20050263.

[32] Burget EG, Verma R, Mølhøj M, Reiter WD. The biosynthesis of L-arabinose in plants: molecular cloning and characterization of a golgi-localized UDP-D-Xylose 4-Epimerase Encoded by the MUR4 Gene of Arabidopsis. Plant Cell 2003;15(2): 523–31. https://doi.org/10.1105/tpc.008425.
[33] Borg AJE, Dennig A, Weber H, Nidetzky B. Mechanistic characterization of UDP-glucuronic Acid 4-epimerase. FEBS J 2021;288(4):1163–78. https://doi.org/10.1111/febs.15478.
[34] Yoo HG, Kwon SY, Karki S, Kwon HJ. A new route to DTDP-6-Deoxy-l-Talose and DTDP-l-Rhamnose: DTDP-l-Rhamnose 4-Epimerase in Burkholderia Thailandensis. Bioorg Med Chem Lett 2011;21(13):3914–7. https://doi.org/10.1016/j.bmcl.2011.05.030.
[35] Alvarez Quispe C, Da Costa M, Beerens K, Desmet T. Exploration of archaeal nucleotide sugar epimerases unveils a new and highly promiscuous GDP-Gal4E Subgroup. Curr Res Biotechnol 2022;4:350–8. https://doi.org/10.1016/j.crbiot.2022.08.003.
[36] Tokuriki N, Tawfik DS. Protein dynamism and evolvability. Science 2009;324: 203–7. https://doi.org/10.1126/science.1169375.
[37] Petrovic D, Risso VA, Kamerlin SCL, Sanchez-Ruiz JM. Conformational dynamics and enzyme evolution. J R Soc Interface 2018;15(144):20180330. https://doi.org/10.1098/rsif.2018.0330.
[38] Beerens K, Gevaert O, Desmet T. GDP-Mannose 3,5-Epimerase: a view on structure, mechanism, and industrial potential. Front Mol Biosci 2022;8(784142):1–13. https://doi.org/10.3389/fmolb.2021.784142.
[39] Molcanov K, Kojić-Prodić B. Towards understanding π-stacking interactions between non-aromatic rings. IUCrJ 2019;6(Pe 2015):156–66. https://doi.org/10.1107/S2052252519000186.
[40] Lau STB, Tanner ME. Mechanism and active site residues of GDP-Fucose Synthase. J Am Chem Soc 2008;130(51):17593–602. https://doi.org/10.1021/ja807799k.
[41] Corbella M, Pinto GP, Kamerlin SCL. Loop dynamics and the evolution of enzyme activity. Nat Rev Chem 2023;7(8):536–47. https://doi.org/10.1038/s41570-023-00495-w.
[42] Albesa-Jové D, Romero-García J, Sancho-Vaello E, Contreras FX, Rodrigo-Unzueta A, Comino N, Carreras-González A, Arrasate P, Urresti S, Biarnés X, Planas A, Guerin ME. Structural snapshots and loop dynamics along the catalytic cycle of glycosyltransferase GpgS. Structure 2017;25(7):1034–1044.e3. https://doi.org/10.1016/j.str.2017.05.009.
[43] Yu H, Zhao Y, Guo C, Gan Y, Huang H. The role of proline substitutions within flexible regions on thermostability of luciferase. Biochim Biophys Acta - Proteins Proteom 2015;1854(1):65–72. https://doi.org/10.1016/j.bbapap.2014.10.017.
[44] Yan BX, Sun Qing. Y. Glycine residues provide flexibility for enzyme active sites. J Biol Chem 1997;272(6):3190–4. https://doi.org/10.1074/jbc.272.6.3190.
[45] Schulz JM, Watson AL, Sanders R, Ross KL, Thoden JB, Holden HM, Fridovich-Keil JL. Determinants of function and substrate specificity in human UDP-Galactose 4′-Epimerase. J Biol Chem 2004;279(31):32796–803. https://doi.org/10.1074/jbc.M405005200.
[46] Sakuraba H, Kawai T, Yoneda K, Ohshima T. Crystal structure of UDP-galactose 4-epimerase from the hyperthermophilic archaeon pyrobaculum calidifontis. Arch Biochem Biophys 2011;512(2):126–34. https://doi.org/10.1016/j.abb.2011.05.013.
[47] Hou J, Tian S, Yang L, Zhang Z, Liu Y. A systematic review of the uridine diphosphate-galactose/glucose-4-epimerase (UGE) in plants. Plant Growth Regul 2021;93(3):267–78. https://doi.org/10.1007/s10725-020-00686-1.
[48] Holden HM, Rayment I, Thoden JB. Structure and function of enzymes of the leloir pathway for galactose metabolism. J Biol Chem 2003;278(45):43885–8. https://doi.org/10.1074/jbc.R300025200.
[49] Bhattacharyya U, Dhar G, Bhaduri A. An arginine residue is essential for stretching and binding of the substrate on UDP-glucose-4-epimerase from escherichia coli. Use of a stacked and quenched uridine nucleotide fluorophore as probe. J Biol Chem 1999;274(21):14573–8. https://doi.org/10.1074/jbc.274.21.14573.
[50] Rapp C, Borg A, Nidetzky B. Interplay of structural preorganization and conformational sampling in UDP-glucuronic acid 4-epimerase catalysis. Nat Commun 2024;15(1). https://doi.org/10.1038/s41467-024-48281-6.
[51] Akdel M, Pires DEV, Pardo EP, Jänes J, Zalevsky AO, Mészáros B, Bryant P, Good LL, Laskowski RA, Pozzati G, Shenoy A, Zhu W, Kundrotas P, Serra VR, Rodrigues CHM, Dunham AS, Burke D, Borkakoti N, Velankar S, Frost A, Basquin J, Lindorff-Larsen K, Bateman A, Kajava AV, Valencia A, Ovchinnikov S, Durairaj J, Ascher DB, Thornton JM, Davey NE, Stein A, Elofsson A, Croll TI, Beltrao P. A structural biology community assessment of AlphaFold2 applications. Nat Struct Mol Biol 2022;29(11):1056–67. https://doi.org/10.1038/s41594-022-00849-w.
[52] Yu H, Dalby PA. A Beginner's Guide to Molecular Dynamics Simulations and the Identification of Cross-Correlation Networks for Enzyme Engineering, 1st ed., 643. Elsevier Inc; 2020. https://doi.org/10.1016/bs.mie.2020.04.020.
[53] Sparta KM, Krug M, Heinemann U, Mueller U, Weiss MS. Xdsapp2.0. J Appl Crystallogr 2016;49:1085–92. https://doi.org/10.1107/S1600576716004416.
[54] Evans PR, Murshudov GN. How good are my data and what is the resolution? Acta Crystallogr Sect D Biol Crystallogr 2013;69(7):1204–14. https://doi.org/10.1107/S0907444913000061.
[55] Winn MD, Ballard CC, Cowtan KD, Dodson EJ, Emsley P, Evans PR, Keegan RM, Krissinel EB, Leslie AGW, McCoy A, McNicholas SJ, Murshudov GN, Pannu NS, Potterton EA, Powell HR, Read RJ, Vagin A, Wilson KS. Overview of the CCP4 suite and current developments. Acta Crystallogr Sect D Biol Crystallogr 2011;67(4): 235–42. https://doi.org/10.1107/S0907444910045749.
[56] McCoy AJ, Grosse-Kunstleve RW, Adams PD, Winn MD, Storoni LC, Read RJ. Phaser crystallographic software. J Appl Crystallogr 2007;40(4):658–74. https://doi.org/10.1107/S0021889807021206.
[57] Jumper J, Evans R, Pritzel A, Green T, Figurnov M, Ronneberger O, Tunyasuvunakool K, Bates R, Žídek A, Potapenko A, Bridgland A, Meyer C, Kohl SAA, Ballard AJ, Cowie A, Romera-Paredes B, Nikolov S, Jain R, Adler J, Back T, Petersen S, Reiman D, Clancy E, Zielinski M, Steinegger M, Pacholska M, Berghammer T, Bodenstein S, Silver D, Vinyals O, Senior AW, Kavukcuoglu K, Kohli P, Hassabis D. Highly accurate protein structure prediction with AlphaFold. Nature 2021;596(7873):583–9. https://doi.org/10.1038/s41586-021-03819-2.
[58] Emsley P, Cowtan K. Coot: model-building tools for molecular graphics. Acta Crystallogr Sect D Biol Crystallogr 2004;60(12 I):2126–32. https://doi.org/10.1107/S0907444904019158.
[59] Afonine PV, Grosse-Kunstleve RW, Echols N, Headd JJ, Moriarty NW, Mustyakimov M, Terwilliger TC, Urzhumtsev A, Zwart PH, Adams PD. Towards automated crystallographic structure refinement with Phenix.Refine. Acta Crystallogr Sect D Biol Crystallogr 2012;68(4):352–67. https://doi.org/10.1107/S0907444912001308.
[60] Adams PD, Afonine PV, Bunkóczi G, Chen VB, Davis IW, Echols N, Headd JJ, Hung LW, Kapral GJ, Grosse-Kunstleve RW, McCoy AJ, Moriarty NW, Oeffner R, Read RJ, Richardson DC, Richardson JS, Terwilliger TC, Zwart PH. PHENIX: a comprehensive python-based system for macromolecular structure solution. Acta Crystallogr Sect D Biol Crystallogr 2010;66(2):213–21. https://doi.org/10.1107/S0907444909052925.
[61] Chen VB, Arendall WB, Headd JJ, Keedy DA, Immormino RM, Kapral GJ, Murray LW, Richardson JS, Richardson DC. MolProbity: all-atom structure validation for macromolecular crystallography. Acta Crystallogr Sect D Biol Crystallogr 2010;66(1):12–21. https://doi.org/10.1107/S0907444909042073.
[62] Eberhardt J, Santos-Martins D, Tillack AF, Forli S. AutoDock Vina 1.2.0: new docking methods, expanded force field, and python bindings. J Chem Inf Model 2021;61(8):3891–8. https://doi.org/10.1021/acs.jcim.1c00203.
[63] Maier JA, Martinez C, Kasavajhala K, Wickstrom L, Hauser KE, Simmerling C. Ff14SB: improving the accuracy of protein side chain and backbone parameters from Ff99SB. J Chem Theory Comput 2015;11(8):3696–713. https://doi.org/10.1021/acs.jctc.5b00255.
[64] Kirschner K, Yongye A, Tschampel S. Glycam06: a generalizable biomolecular force field. Carbohydr J Comput Chem2 2008;29(4):622–55. https://doi.org/10.1002/jcc.20820.GLYCAM06.
[65] Petrová P, Koča J, Imberty A. Potential energy hypersurfaces of nucleotide sugars: Ab initio calculations, force-field parametrization, and exploration of the flexibility. J Am Chem Soc 1999;121(23):5535–47. https://doi.org/10.1021/ja983854g.
[66] Ryde U. On the Role of Glu-68 in alcohol dehydrogenase. Protein Sci 1995;4(6): 1124–32. https://doi.org/10.1002/pro.5560040611.
[67] Frisch MJ, Trucks GW, Schlegel HB, Scuseria GE, Robb MA, Cheeseman JR, Scalmani G, Barone V, Mennucci B, Petersson GA. Gaussian. Revision B.01, 09. Wallingford CT: Gaussian, Inc; 2009.
[68] Mayne CG, Saam J, Schulten K, Tajkhorshid E, Gumbart JC. Rapid parameterization of small molecules using the force field toolkit. J Comput Chem 2013;34(32):2757–70. https://doi.org/10.1002/jcc.23422.
[69] Romero-García J, Francisco C, Biarnés X, Planas A. Structure-function features of a mycoplasma glycolipid synthase derived from structural data integration, molecular simulations, and mutational analysis. PLoS One 2013;8(12):1–14. https://doi.org/10.1371/journal.pone.0081990.