International Journal of Biological Macromolecules 362 (2026) 151980

Contents lists available at ScienceDirect

International Journal of Biological Macromolecules

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

CryoEM structural analysis of a thermophilic galactooligosaccharides-producer β-galactosidase unravels an uncommon oligomeric structure

Gennaro Sanità [a,1], Emanuela Maresca [b,1], Stefano Capaldi [c], Angela Casillo [d], Martina Aulitto [b], Federica Donadio [a], Tillmann Pape [e], Maria Michela Corsaro [d], Emanuela Esposito [a], Patrizia Contursi [b,f,*]

[a] Institute of Applied Sciences and Intelligent Systems -EYE LAB - National Research Council, Via Pietro Castellino 111, 80131, Napoli, Italy
[b] Department of Biology, University of Naples Federico II, 80126, Naples, Italy
[c] Department of Biotechnology, University of Verona, Italy
[d] Department of Chemical Sciences, University of Naples Federico II, 80126, Naples, Italy
[e] Structural Molecular Biology Group, Novo Nordisk Foundation Center for Protein Research and Core Facility for Integrated Microscopy (CFIM), Faculty of Health and Medical Sciences, University of Copenhagen, Blegdamsvej 3B, 2200, Copenhagen, Denmark
[f] NBFC – National Biodiversity Future Center, 90133, Palermo, Italy

ARTICLE INFO

Dataset link: PDB ID : 9TA3 / pdb_00009ta3, EMDB ID : EMD-55743 (Original data)

Keywords:
Cryo-EM
Thermophilic β-galactosidase
Galactooligosaccharides

ABSTRACT

Thermostable β-galactosidases represent promising biocatalysts for lactose hydrolysis and production of structurally defined galacto-oligosaccharides (GOS). Here we report the cryo-EM structure of the glycoside hydrolase family 42 (GH42) β-galactosidase from *Heyndrickxia coagulans* MA-13 (*Hc*GalB), determined at 2.97 Å resolution. *Hc*GalB adopts a canonical tripartite architecture and assembles into a barrel-like homo-hexamer composed of two staggered trimers that interact in an unusual top-to-top configuration. This quaternary arrangement contributes not only to structural stability but also to the modulation of substrate channeling and catalytic properties. Molecular docking revealed a surface groove shaped by conserved aromatic residues that might guide the substrate towards the catalytic pocket. Moreover, the structural data provide a mechanistic rationale for the efficient transgalactosylation activity of *Hc*GalB, which predominantly generates β (1 → 3)-linked GOS, along with β(1 → 6) and β(1 → 4) linkages, as confirmed by 2D Nuclear Magnetic Resonance. Overall, these findings expand the structural landscape of GH42 enzymes and identify architecture-specific determinants that can be leveraged to optimize GH42 catalysts for industrial and functional food applications.

1. Introduction

β-Galactosidase (EC 3.2.1.23), a pivotal enzyme in the hydrolysis of β-galactosides, plays a central role in both cellular metabolism and numerous industrial applications [1]. Lactases and other β-galactosidases are classified into four glycosyl hydrolase families (GH): GH1, GH2, GH35, and GH42 [2]. Proteobacteria predominantly express β-galactosidases belonging to the GH1 and GH2 families, whereas GH42 enzymes are typically associated with thermophilic bacteria. Conversely, the GH35 family is expressed across fungi, animals, plants, and bacteria [3,4]. The catalytic versatility of this group of enzymes includes hydrolytic activity, which is commonly used to produce lactose-free products, and transgalactosylation for the synthes of prebiotic galacto-oligosaccharides (GOS) [5].

From a biotechnological perspective, while β-galactosidases from mesophilic sources such as *E. coli* (*lacZ*) have been extensively studied, thermophilic β-galactosidases are gaining attention due to their remarkable stability and activity at elevated temperatures, which are advantageous for industrial biocatalysis [6,7]. Thermostability is largely attributable to unique structural features, including improved hydrophobic packing, shorter loops, increased number of salt bridges, and optimized oligomeric interactions [8–12]. A growing body of evidence

* Corresponding author at: Department of Biology, University of Naples Federico II, 80126, Naples, Italy.
E-mail address: contursi@unina.it (P. Contursi).
[1] These authors contributed equally.

https://doi.org/10.1016/j.ijbiomac.2026.151980
Received 2 December 2025; Received in revised form 7 April 2026; Accepted 10 April 2026
Available online 13 April 2026

highlights the structural plasticity of β-galactosidases, particularly in their oligomeric states, ranging from dimers to hexamers or even more complex assemblies [13,14]. The hexameric structure is a typical characteristic of glycoside hydrolase family 42 (GH42) enzymes, but the deposited structures from extremophilic GH42 enzymes remain limited [15]. Interestingly, some thermophilic β-galactosidases deviate from this paradigm, forming unusual oligomeric architectures such as trimers, tetramers, or asymmetric multimers, often correlated with specific substrate affinities, allosteric regulation, or adaptation to extreme environments [9].

X-ray crystallography generally provides high-resolution structural information to guide protein quaternary structure and its engineering, but important practical issues need to be overcome to achieve crystallisation of the enzyme of interest (i.e., removal of glycosylation, protein purification in sufficient quantities and purity, and difficulty in growing well-diffracting crystals) [16]. In this context, cryo-electron microscopy (cryo-EM) revolutionized structural biology by offering unprecedented insights into protein dynamics in near-native states, especially for large, flexible complexes [17]. Despite the growing interest in GH42 enzymes, high-resolution cryo-EM structures of this family remain scarce.

In general, GH42 β-galactosidases operate via a retaining double-displacement mechanism, involving two sequential nucleophilic substitutions that maintain the anomeric configuration of the cleaved glycosidic bond [13]. In GH42 β-galactosidases, catalysis involves two conserved glutamate residues; one initiates the nucleophilic attack that generates the galactosyl-enzyme intermediate, while the second serves as an acid/base catalyst [18]. The reaction outcome depends on the acceptor, for example, water yields hydrolysis, whereas a sugar molecule drives transgalactosylation, leading to galacto-oligosaccharides (GOS), whose type influences stability and prebiotic efficacy [19]. The balance between transgalactosylation and hydrolysis (T/H ratio) depends on substrate concentration, enzyme flexibility, and, most importantly, active site architecture. Both reactions share the same catalytic pathway in GH42 β-galactosidases, making fine-tuning quite challenging. Enhancing transgalactosylation typically involves reducing water accessibility or increasing affinity for sugar acceptors, often through site-directed mutagenesis [20].

Understanding the diversity and characteristics of GOS produced by a specific β-galactosidase is crucial for assessing its biotechnological potential. In dairy biotechnology, lactose can act both as substrate and acceptor, allowing *in situ* GOS formation in milk or whey, improving nutritional value and supporting beneficial microbiota such as *Bifidobacterium* and *Lactobacillus* [21–24]. This dual activity also reduces lactose content, benefiting lactose-intolerant consumers and making β-galactosidases attractive biocatalysts for functional foods [25]. Accurate characterization of GOS linkages is crucial, since the linkage pattern affect digestibility and persistence in complex matrices and represents an important parameter in the design of mixtures with controlled functionality and shelf-life [26,27]. Nuclear Magnetic Resonance (NMR) spectroscopy, particularly 1D and 2D techniques such as ^{1}H, ^{13}C, COSY, HSQC, and HMBC, represents the method of choice for elucidating the structure and linkage patterns of GOS [28].

Since the relationship between oligomeric architecture and the resulting GOS product profile remains largely unresolved, elucidating the quaternary structure of β-galactosidases in solution by cryo-EM, especially for enzymes capable of producing GOS, offers a valuable framework for identifying structural determinants that can be targeted through rational mutagenesis [29]. This structural information enables the mapping of inter-subunit interfaces, substrate-access channels, and microenvironments that could influence yield of the transgalactosylation, thereby guiding the design of variants with tailored stability or product specificity.

In this work we aimed to extend the characterization of *Hc*GalB from *Heyndrickxia coagulans* MA-13 to unravel its structure and uncover the molecular basis of *Hc*GalB dual catalytic activity (hydrolysis of β-1,4-galactosidic linkages and transgalactosylation), thus assessing its potential in driving GOS production. To this end, we investigated the enzyme structural architecture using cryoEM and characterized through NMR the resulting transgalactosylation products derived from homo- and hetero-condensation reactions.

By combining biochemical assays and structural analysis, we explored how the architecture of the active site influences the ability of *Hc*Bgal to catalyze both hydrolysis and transgalactosylation. Particular attention was given to the nature and diversity of GOS linkages, whose structural resolution provided insights into the active site of the enzyme, offering new avenues for engineering and biotechnological exploitation.

2. Material and methods

2.1. Chemicals

All chemicals and reagents were purchased from commercial suppliers unless otherwise stated. Sodium chloride, tryptone, yeast extract, kanamycin, chloramphenicol, isopropil-β-D-1-tiogalattopiranoside (IPTG), lysozyme, and DNase were obtained from AppliChem (Darmstadt, Germany). Sodium phosphate and sodium citrate were purchased from Sigma-Aldrich (St. Louis, MO, USA), protease inhibitor cocktail tablets from Roche (Basel, Switzerland), and methanol and acetone (MeOH) from ROMIL (Cambridge, UK), respectively. Affinity chromatography columns (HisTrap, solid/support) was obtained from GE Healthcare (Chicago, IL, USA). *E. coli* Rosetta™(DE3) pLysS cells used for protein expression were purchased from Merck/Novagen (Darmstadt, Germany). The following substrates were purchased from Biosynth (Staad, Switzerland): *o*-Nitrophenyl-β-D-galactopyranoside (ONP-β-gal), 4-Nitrophenyl-β-D-xylopyranoside (PNP-β-xyl) and lactose.

2.2. Cloning, expression and purification of HcGalB

The β-galactosidase gene from *H. coagulans* MA-13 (encoding for *Hc*GalB) was previously cloned into pET28b vector and transformed into *E. coli* Rosetta™(DE3) pLysS for C-terminal His-tagged expression as described in Aulitto *et al* 2021. Transformants were selected on Luria Bertani (LB) -kanamycin (50 μg/mL)/chloramphenicol (33 μg/mL) plates, grown from a single colony (50 mL LB, 180 rpm, 37 °C), diluted into 1 L LB (starting from OD_{600}/mL 0.08) and finally induced at OD_{600}/mL 0.5–0.6 with 0.5 mM IPTG overnight. *Hc*GalB was purified by affinity chromatography on a HisTrap column, according to Aulitto *et al.* (2021) with a minor modification in the lysis step performed by a single freeze–thaw cycle. Briefly, cells were resuspended in sodium phosphate 100 mM pH 8, containing lysozyme 1 mg/mL and inhibitor protease, and incubated for 1 h at 37 °C. After the addition of DNase (5 μg/mL), cells were incubated on ice for 40 min, then frozen for 15 min in dry ice, and subsequently thawed in a water bath at 37 °C for 30 min. The cell extract was clarified by a centrifugation step at 18.300 rpm (Fixed Angle Rotor JA 25,50, Beckman) for 30 min at 4 °C.

2.3. Purification of GOS

The homo- and hetero-transgalactosylation reactions using artificial substrates (ONP-β-gal: ONP-β-gal/ONP-β-gal:PNP-β-xyl) were performed according to Aulitto *et al.* 2021 [9]. The reaction mixture was then purified on a reverse phase column (Polar-RP 80 A, Phenomenex, 4 μm, 250 × 10 mm^2, flow 0.8 mL × min − 1) using an Agilent HPLC 1100 series instrument. Samples were eluted using H_2O 60%: MeOH 40% as mobile phase. For all the oligosaccharides isolated, the structural determination was obtained by NMR experiments.

2.4. NMR analysis

NMR analyses of the reaction products were performed on a Bruker 600 MHz equipped with a cryogenic probe. Two-dimensional spectra (1H—1H DQF-COSY, 1H—1H ROESY, 1H—1H TOCSY, 1H—13C HSQC

and 1H—13C HMBC) were recorded in D_2O at 298 K. Chemical shifts are referred to external acetone (δ_H 2.225 ppm, δ_C 31.45 ppm).

2.5. CryoEM sample preparation and imaging conditions

Aliquots of 5 μL of *Hc*GalB (1.5 mg/mL) in citrate sodium 100 mM, pH 5 were applied to glow discharged 200 mesh Quantifoil® grids (R 1.2/1.2). The grids were blotted for 2s using a Vitrobot Mark IV (Thermo Fisher®) in 100% humidity at 4 °C and subsequently plunged into liquid ethane. CryoEM movies were collected at a nominal magnification of 130,000× (corresponding to a pixel size of 0.86 Å/pixel) on a Glacios (Thermo Fisher®) microscope operating at 200 kV, equipped with an Selectris energy filter (slit width 10 eV) and a Falcon 4i Direct Electron Detector (total dose of 45 e−/Å^2). Defoci over a range of – 0.5 μm, − 1 μm, − 1.5 μm and – 2 μm were used. A total of around 3500 micrographs were acquired.

2.6. CryoEM data processing

All data processing, reconstruction and structure refinement were performed with CryoSPARC v.4.6.1 [30]. The data processing workflow is outlined in Fig. S1. After patch-motion correction and CTF determination, a total of 3383 manually curated exposures were selected for further processing. Reference-free autopicking and 2D classification on a subset of micrographs were used to generate 2D class averages for subsequent template picking of the entire dataset, resulting in a total of 799,865 single particles. After several rounds of 2D classification, a total of 629,997 particles were selected for *ab-initio* reconstruction (1 volume, C1 symmetry) and followed by homogeneous refinement. 3D classification without symmetry enforcement was used to reduce composition heterogeneity of the dataset and to select particles that correspond to fully assembled hexameric complexes. After global and local CTF refinement and reference-based motion correction, the selected particles (629,632) were used in non-uniform refinement with D3 symmetry imposed, resulting in a final density map with a resolution of 2.97 Å according to the "gold-standard" Fourier Shell Correlation (FSC = 0.143) criterion [31]. Local resolution of the map was estimated within CryoSPARC (Fig. S1, panel e). For model building, the map was sharpened with a global B factor of −112.5 Å^2.

2.7. Model building and refinement

For model building, a single monomer of the homologous enzyme from *Bacillus circulans* (PDB ID: 3TTY) was used as the starting model. The model was initially docked and fitted by rigid-body optimization into the map using ChimeraX [32]. The amino acid sequence was then manually mutated to match that of the target protein and subjected to several rounds of real-space refinement with phenix.refine and manual rebuilding in Coot [33,34]. Well-ordered water molecules were identified and modelled with phenix.douse and manually reviewed in the density map [35]. The structure of the hexameric assembly was generated applying NCS symmetry to the refined monomer and subjected to a final round of real-space refinement using NCS constraints. The quality of the final model was assessed with Molprobity [36]. A summary of data collection and refinement statistics are reported in Table S1. High resolution figures were prepared with PyMOL (The PyMOL Molecular Graphics System, Version 2.0 Schrödinger, LLC) and ChimeraX.

2.8. Molecular docking and in silico protein structure prediction

The structures of β-lactose, and chromogenic substrate analogs ONP-β-Gal and PNP-β-Xyl were retrieved from PubChem Database (https://pubchem.ncbi.nlm.nih.gov) and imported into Chimera 1.15. The ligands and the protein (a single monomer of *Hc*GalB) were initially prepared by adding polar hydrogen atoms and assigning partial charges using the AutoDock tool [37]. Molecular docking was performed with AutoDock Vina v.1.1.2 using a search grid of 35 × 35 × 35 Å (x, y, and z) centered on the active site of the enzyme [38]. Each docking experiment was repeated two times, generating a total of 20 poses per ligand. The poses were inspected in Chimera and the 10 with the lowest energy (*Z*-score) were selected for analysis.

The structural superimposition reported in paragraph 3.4 was based on protein structures predictions of Aaβ-gal (GenBank: ABI84370.1) and *Bbre*βgal-III (GenBank: BAQ99491.1) using AlphaFold 3 and on BiBga42A PDB deposited structure (8IBT).

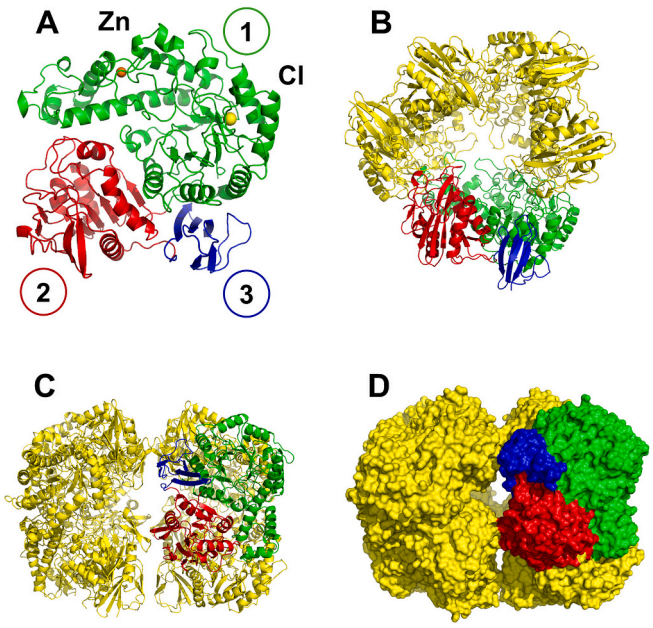

Fig. 1. Overall architecture of *Hc*GalB monomer and oligomeric assemblies. (A) Ribbon representation of the *Hc*Galb monomer, modelled on the structural framework reported by Maksimainen *et al.* Domains 1, 2, and 3 are colored green, red, and blue, respectively [36]. The bound Zn^{2+} and Cl^- ions are shown as orange and yellow spheres, respectively. (B, C) Ribbon representations of the *Hc*GalB homotrimer and homohexamer. The single monomer is shown according to the same colour code as in panel A, while the remaining monomers are colored in yellow. (D) Surface representations of the *Hc*GalB homohexamer, corresponding to the ribbon models shown in panels C. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

2.9. Data availability

The data that support the findings of this study (density maps and atomic coordinates for the *Hc*GalB) are openly available inthe EMDB and RSCB Protein Data Bank with reference numbers (accession codes) EMD-55743 and 9TA3, respectively [39].

3. Results and discussion

3.1. Structural architecture of HcGalB and comparative analysis with homologous β-galactosidases

This work provides a 3D reconstruction of *Hc*GalB through cryoEM, allowing a direct visualization of the domain arrangement and intersubunit interfaces. The structure of the single monomer revealed a canonical tripartite domain organization whilst the quaternary structure analysis highlighted that *Hc*GalB assembles into a barrel-like homohexamer composed of two staggered trimers (Fig. 1).

To gain insights into the structural determinants of enzymatic catalysis of *Hc*GalB, we performed a comparative structural analysis using two β-galactosidases from different phylogenetic and functional

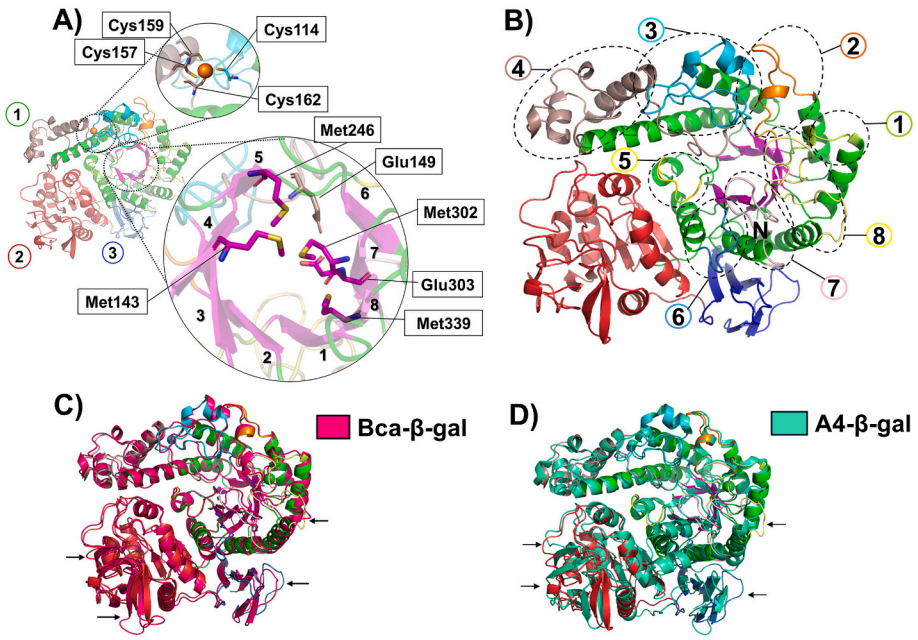

Fig. 2. Structural organization of the β-barrel domain and inter-subunits interactions in *Hc*GalB. (A) Structure of the β-barrel domain of *Hc*GalB, with the eight β-strands numbered from 1 to 8. Sulfur-rich residues (Met143, Met246, Met302, and Met339) are positioned in the core of the barrel, and the four Zn^{2+}-coordinating cysteines (Cys114, Cys157, Cys159, and Cys162) are also shown. (B) Representation of the eight connecting loops is shown in different colors, along with a structural comparison of the loop regions between *Hc*GalB and Bca-β-gal (C) or A4-β-Gal (D), where Bca-β-gal and A4-β-Gal are colored in pink and green cyan, and the arrows highlight differences between the structures. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

backgrounds. The first homolog Bca-β-gal, derived from *Bacillus circulans* (PDB: 3TTY), belongs to the same genus as *H. coagulans* MA-13 (formerly *Bacillus coagulans*), allowing for comparison at the taxonomic and functional level [40]. The second reference structure A4-β-Gal, from the thermophilic bacterium *Thermus thermophilus* A4 (PDB: 1KWK), was selected for its thermostable features, which parallel those reported for *Hc*GalB [41]. Sequence alignments revealed amino acid identity of 45.62% and 32.17% with Bca-β-gal and A4-β-Gal, respectively (not shown), consistent with an overall structural conservation across major domains. However, notable differences in folding patterns and local features, particularly around the active site, were observed (Fig. 2C-D).

Domain 1 (residues 1–384), containing the active site, adopts a typical $(\beta/\alpha)_8$ TIM-barrel fold with a hydrophobic core that includes four methionine residues (Met143, Met246, Met302, Met339) forming a sulfur-rich cluster similar to that of Bca-β-gal (Fig. 2A). Notably, while the *B. circulans* enzyme contains an additional cysteine within this cluster, *Hc*GalB lacks this feature. Nevertheless, it resembles Bca-β-Gal more closely than A4-β-Gal, which possesses only a single methionine in the core of the barrel. This fold is commonly observed in glycoside hydrolases (GHs) belonging to the GH-A clan [42]. In this architecture, eight alternating α-helices and β-strands form a cylindrical structure with a small central cavity (Fig. 2A). The catalytic residues of *Hc*GalB, Glu149 (acid/base) and Glu323 (nucleophile), are located around this central cavity, positioned at the C-terminal ends of the β4 and β7 strands of the TIM barrel, respectively (see also paragraph 3.2).

Loops connecting the β-strands and α-helices in domain 1 displayed both conserved and divergent features (Fig. 2B). Specifically, *Hc*GalB retains the N-terminal segment present in A4-β-Gal and, like both homologs, features a nine-residue C-terminal loop that caps the last α-helix.

The first (Tyr14-Asp20, 6 residues) and the second (Asn42-Ser59, 17 residues) loops contain a single 3_{10}-helix, while the third loop (Thr79-Lys118, 39 residues) displays two 3_{10}-helices and an irregular loop structure that participates in zinc ion binding through Cys114. These structural hallmarks, including the comparable loop length, are likewise observed in Bca-β-gal and A4-β-Gal. Loop 4 (Val146-Pro211), which hosts the remaining three Zn^{2+}-coordinating cysteines (Cys157, Cys159, Cys162) and bears the same secondary elements (two α helices and two 3_{10}-helices) elements, represents a key structural motif that mediates inter-subunitinteractions and contributes to oligomeric stability [43]. The metal Zn^{2+} ion at this site might be involved in the stabilization of the local structural arrangement (Fig. 2A) as previously demonstrated [43]. The fifth loop (Met250-Asp 258) is conserved among the compared structures, while loop 6 (Tyr273-Pro280) is notably longer in A4-β-gal (7 vs 23 residues). Finally, loops 7 and 8 (Gln307-Pro320; Phe341-Ser365) are similar across the three enzymes but lack distinct secondary structure elements.

Similarly to Bca-β-gal and A4-β-Gal, domain 1 of *Hc*GalB is connected to domain 2 (residues 394–600) by a long random coil (residues 385–393). The domain 2 exhibits an α/β fold being composed of ten β strands and five α-helices, similar to that of Bca-β-gal and A4-β-Gal. The β-strands are positioned at the center of the domain, encircled by the α-helical elements. The function of the domain is structural, forming a major interface with the adjacent monomers (within the trimer and the hexamer), thus stabilizing the overall structure assembly as described in detail in paragraph 3.3.

The third domain, comprising 56 amino acids (residues 607–661), is the smallest of the three and adopts an overall jelly-roll fold consisting of an antiparallel β-sandwich formed by two β-sheets, one of which is noticeably shorter than the other. Although the domain topology is broadly conserved between Bca-β-gal and A4-β-gal, *Hc*GalB contains fewer residues than Bca-β-gal. This domain plays a mainly structural role within the hexamer, as highlighted in the hexameric structure (Paragraph 3.3), and it does not bear clefts or cavities on its surface.

3.2. Active site

The identification of the active site residues was guided by structural comparison with the crystallographic templates Bca-β-gal and A4-β-Gal,

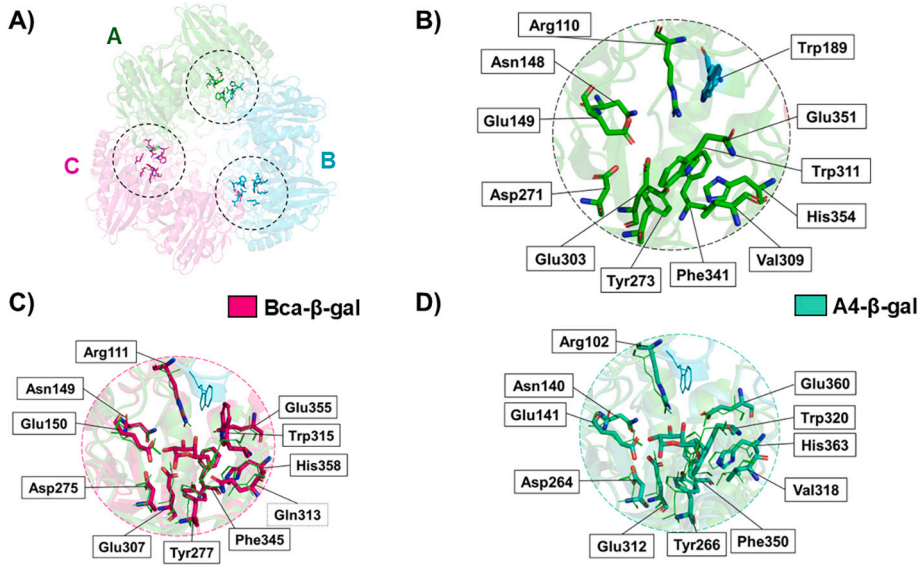

Fig. 3. Active site architecture of *Hc*GalB and comparison with homologous structures. (A) Trimeric structure of *Hc*GalB, with each chain shown in green, cyan, and purple. Active site residues are highlighted in stick representation. (B) Close-up view of the catalytic pocket, showing the key residues involved in the enzymatic function. Trp 189 (blue) belongs to the adjacent monomer (C—D). Structural superposition of *Hc*GalB with the structures Bca-β-gal and A4-β-Gal, highlighting the conserved active site residues across the three structures. The key residues are colored in hot pink for Bca-β-gal and in cyan green for A4-β-Gal, while the key residues reported in panel B of *Hc*GalB are shown in lines. The α-D-galactose ligand is shown centrally positioned within the active site of both structures. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

both resolved in complex with α-D-galactose. The active sites of *Hc*GalB are located at the interfaces between the monomers, so that every two adjacent monomers participate in the formation of each of the three complete active sites in the trimer (Fig. 3A**)**.

Six out of the seven residues involved in galactose binding in Bca-β-gal and A4-β-Gal were also identified in the target structure (Arg110, Asn148, Glu149, Glu303, Glu351, and His354) and are proposed to constitute the core of the active site (Fig. 3B). Specifically, Glu149 (loop 4) and Glu303 (β7) were identified as the acid/base catalyst and nucleophile, respectively. *Hc*GalB also bears the conserved Asp271 (β6), located in close contact with the two glutamates. Interestingly, Maksimainen *et al.* showed that the corresponding Asp275 in Bca-β-gal plays a central role in fine-tuning their catalytic functions [40]. Finally, Trp189 provided *in trans* by the adjacent subunit, is essential for completing the active site pocket and plays a major role in substrate binding and/or recognition (Fig. 3B) [43]. It is well established that carbohydrate moieties often interact through stacking interactions with the planar surfaces of aromatic residues which potentially facilitate substrate positioning and guide the disaccharide towards the catalytic site [40,41]. Therefore, the spatial disposition of three aromatic residues in the Bca-β-gal structure (Tyr 273, Trp 311 and Phe 341), is consistent with a canonical carbohydrate-binding configuration. Among these, Tyr273 (loop 6) in *Hc*GalB corresponds to Tyr277 in Bca-β-gal and to Tyr 266 in A4-β-gal, and it is considered the most likely residue involved in

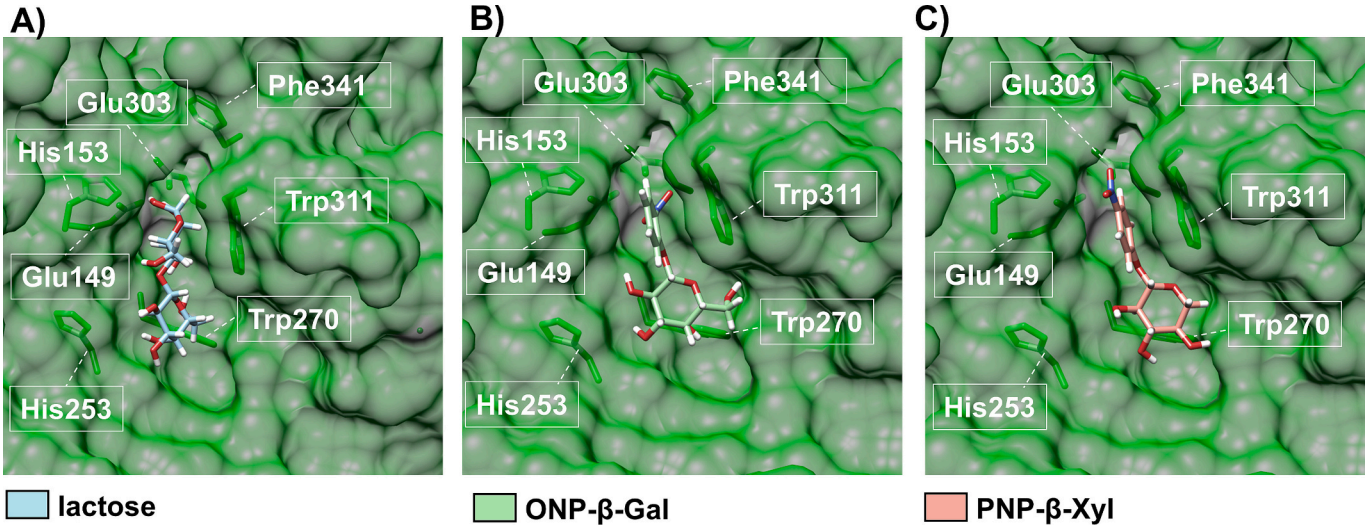

Fig. 4. Docking simulations of *Hc*GalB and its substrates. Docking simulations of (A) lactose (cyan), (B) ONP-β-Gal (green), and (C) PNP-β-Xyl (orange) onto the *Hc*GalB structure reveal a consistent binding pattern near the entrance of the catalytic pocket. All ligands occupy a surface-exposed groove defined by the conserved aromatic residues Trp311 and Phe341, which correspond to Trp315 and Phe345 in the Bca-β-Gal homolog. Glu303 and Glu149 represent the two catalytic residues. Other residues (His153, His253, Trp270) may contribute to shaping the access to the catalytic pocket. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

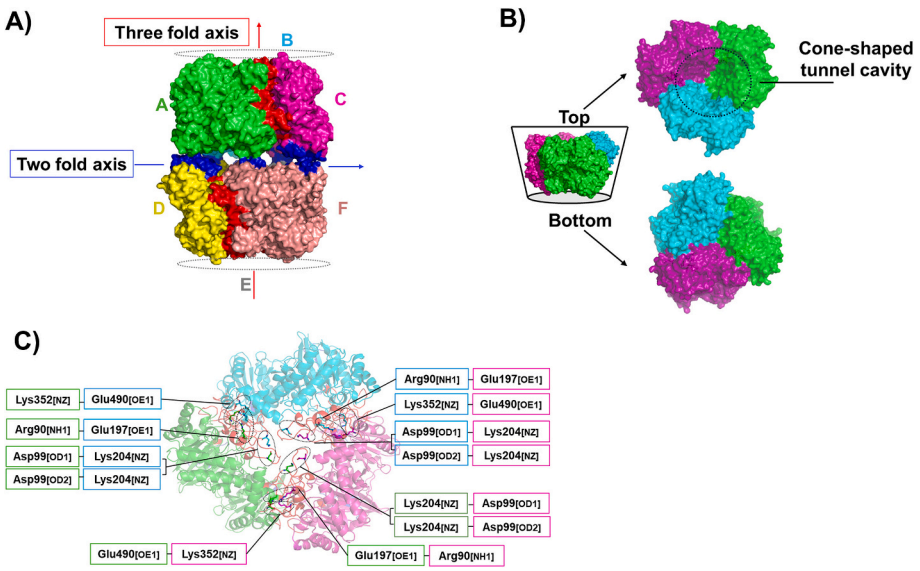

Fig. 5. Quaternary structure of *Hc*GalB. (A) Surface representation of the hexameric assembly of *Hc*GalB, composed of two staggered trimers. The inter-trimer interface regions are highlighted in blue, whereas the intra-trimer interface regions are shown in red. Each trimer consists of three monomers, individually colored to distinguish subunits: trimer 1) monomer A (green), monomer B (cyan), and monomer C (magenta); trimer 2) monomer D (yellow), monomer E (grey), and monomer F (pink); (B) Trimer flowerpot architecture of *Hc*GalB highlighting the top and bottom views that contain the cone-shaped tunnel cavity. (C) Residues (Lys204, Glu197, Glu490, Asp99, Arg90, Lys352) mediating salt-bridge interactions at the interface between monomers within trimer 1 are shown as sticks, keeping the same colour code. Atom labels follow standard PDB notation: NZ (terminal nitrogen of the Lysine side chain), OD1/OD2 (carboxylate oxygens of Asp), OE1/OE2 (carboxylate oxygens of Glu), and NH1 (terminal nitrogen of the Arg side chain). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

stacking with the glucose ring of lactose, thus supporting its functional relevance in substrate recognition. In addition, a single tryptophan residue (Trp311, loop 7) corresponding to Trp315 in Bca-β-gal and to Trp320 in A4-β-gal, is also conserved. This residue is involved in galactose recognition in A4-β-gal. However, the side chain of Trp311 adopts a distinctly different orientation, driven by the altered chemical environment resulting from the substitution of Gln313 in Bca-β-gal with Val309 in *Hc*GalB, a change mirrored by Val318 in A4-β-gal (Fig. 3C–3D).

Notably, Bca-β-gal, contains two tryptophan residues (Trp 270 and Trp 311) implicated in substrate interaction, whilst *Hc*GalB resembles A4-β-Gal in lacking the second tryptophan in this region. Furthermore, a Phe341 (loop8) corresponding to Phe345 in Bca-β-gal and Phe350 in A4-β-gal, is also conserved and located in proximity to the C4 position of galactose, which may contribute to the specificity towards this monosaccharide.

To further validate the hypothesis that a specific aromatic region plays a key role in guiding the substrate towards the catalytic site, molecular docking analyses were performed using the three-dimensional structure of *Hc*GalB. The simulations included both the natural substrate lactose, as well as ONP-βGal and PNP-β-xyl, i.e. synthetic analogues commonly used in enzymatic assays, both of which *Hc*GalB displays high specificity towards [9]. The docking results revealed a consistent binding pattern for the ligands, which were positioned within the same surface region adjacent to the entrance of the catalytic pocket (Fig. 4). All ligands occupy a surface-exposed groove defined by the conserved aromatic residues Trp311 and Phe341, which correspond to Trp315 and Phe345 in the Bca-β-Gal homolog. Glu303 and Glu149 represent the two catalytic residues. Furthermore, other residues, including His153 (loop 4), His253 (loop 5), and Trp270 (β8), may also assist shaping the channel and assist substrate passage.

3.3. Quaternary structure

Previous analyses revealed that *Hc*GalB adopts a hexameric state in solution, as evidenced by size-exclusion chromatography coupled with triple-angle light scattering (QELS) [9]. Six monomers are arranged into a barrel-like structure made of two staggered trimers as revealed by our *Hc*GalB 3D reconstruction. Moreover, the overall architecture of each homotrimer resembles a flowerpot with a cone-shaped tunnel cavity in the center (Movie 1; Figs. 1C–D and 5A–B).

For GH42 β-galactosidases, assembly into trimers is essential for achieving high enzymatic activity [44]. The interface area between the monomers of the trimers, as identified by PISA analysis, measures 2292.43 Å^2 (Fig. 5A–5B**; Table S2**). This value closely matches Bca-β-gal (2278 Å^2) but is slightly lower than A4-β-gal (2385 Å^2), reflecting structural differences among the three proteins. The interface involves residues belonging to the three domains of each monomer (Section 3.1). Domain 1 includes Phe276 and the following regions: Trp87-Gly108, Ser184-Asn210, Ala313-Met323, Ser345-Phe352, Ser360-Arg366; domains 2 and 3 include Glu406-Val415, Asn489-Gly496, Tyr508-Lys517. Notably, six residues (Lys204, Glu197, Glu490, Asp99, Arg90, Lys352) form twelve salt bridges that are crucial for stabilizing the adjacent monomers within the trimer (Fig. 5C), although the total number of salt bridges is lower compared to Bca-β-gal and A4-β-gal.

The two staggered trimers of *Hc*GalB, interact mainly through two regions: Gly526-Gly528 of domain 2 and Glu646-Pro654 of domain 3 of each monomer to form the hexamer. These contacts create interface surfaces of 215.4 and 150.7 Å^2, resulting in a total inter-trimer interface area of 1098.3 Å^2. Interestingly, a central three-fold rotational axis runs vertically through the center of each trimer, while a perpendicular two-fold axis relates opposite subunits across the trimer–trimer interface (Fig. 5A). The hexameric structural organization had previously been observed in Bca-β-gal crystals but never in solution, as indicated by QELS analyses. This suggested that the hexameric arrangement might result from crystal packing rather than representing the predominant solution state [36]. However, a recent study demonstrated that the cold-adapted M-β-gal from *Marinomonas* ef1 is indeed a hexamer both in solution and in the crystalline state [15]. Differently from *Hc*GalB, the structure of M-β-gal consists of two trimeric rings arranged in a stacked configuration. Notably, the cryo-EM reconstruction of *Hc*GalB shows that two trimers assemble in solution in a top-to-top configuration, with

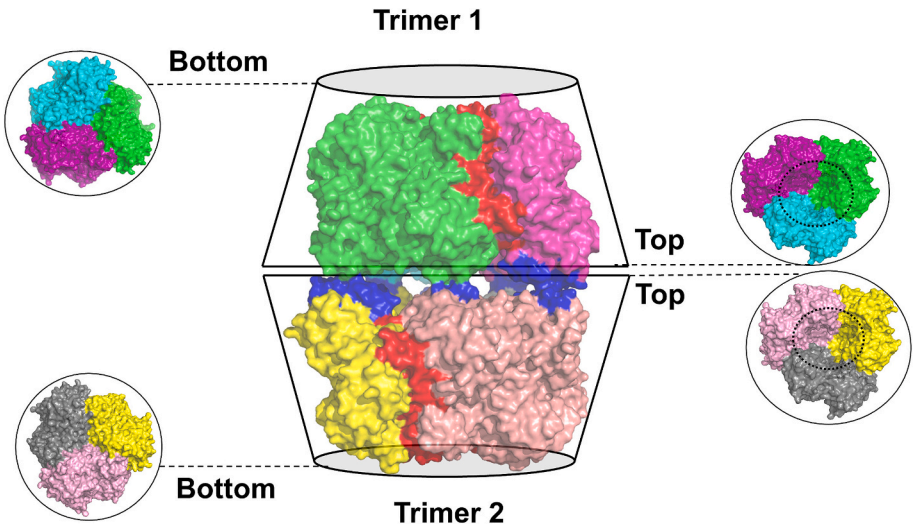

Fig. 6. Hexameric architecture of *Hc*GalB, composed of two flower-pot trimers arranged in a top-to-top configuration, enclosing a central cone-shaped tunnel cavity.

the six active sites facing the inner side of the hexamer (Movie 2, Fig. 6).

To the best of our knowledge, no GH42 β-galactosidase described so far displays such a quaternary organization. Therefore, the architecture observed in *Hc*GalB represents a previously unreported structural solution within the GH42 family, expanding the known repertoire of quaternary assemblies of β-galactosidases. In fact, most GH42 enzymes characterized to date are predominantly trimeric, and only a few assemble into hexamers and none of which adopt a top-to-top arrangement [15,40,41]. Notably, this top-to-top assembly might generate a confined catalytic nanocage with the six active size facing the interior of the hexamer, increasing the local concentration of substrate and/or intermediate oligosaccharides while limiting solvent accessibility. Such micro compartmentalization is expected to favor productive encounters between the galactosyl–enzyme intermediate and alternative sugar acceptors, thereby shifting the reaction equilibrium towards GOS formation rather than hydrolysis (see below). By contrast, the bottom-to-bottom trimeric arrangement described for Bca-β-gal leaves the active sites fully exposed to the solvent, a configuration that facilitates water access and consequently disfavors transgalactosylation. The markedly lower GOS production observed for Bca-β-gal is therefore consistent with its more open quaternary structure. Overall, these findings have also been providing new structure–function insights into the molecular determinants that enable *Hc*GalB to generate a broad repertoire of galactooligosaccharides (Paragraph 3.4).

Table 1
^{1}H and ^{13}C chemical shifts of identified transgalactosylation products of *Hc*GalB.

Product							
	$^{1}H/^{13}C$						
β-Gal(1 → 6)-β-Gal2NP							
	A: t-Gal	4.34	3.44	3.49	3.82	3.53	3.68
		103.2	70.6	72.7	68.6	75.1	60.9
	B: 6-Gal2NP	5.14	3.79	3.70	3.97	4.04	3.97/3.88
		100.9	70.2	72.4	68.5	74.6	68.9
β-Gal(1 → 3)-β-Gal2NP							
	A: t-Gal	4.59	3.56	3.60	3.89	3.82	3.69
		100.7	71.7	72.7	68.6	75.1	61.0
	B: 3-Gal2NP	5.17	3.96	3.88	4.20	3.85	3.70
		100.7	68.5	81.6	68.3	75.4	60.9
β-Gal(1 → 3)-β-Gal(1 → 3)-β-Gal2NP							
	A: t-Gal	4.55	3.54	3.59	3.85	3.62	3.68
		104.2	75.2	72.5	68.5	75.1	60.9
	B: 3-Gal	4.65	3.73	3.77	4.13	3.66	3.68
		103.7	70.3	82.0	68.4	74.8	60.9
	C: 3-Gal2NP	5.17	3.96	3.88	4.20	3.85	3.70
		100.7	68.5	81.6	68.3	75.4	60.9
β -Gal(1 → 3)-β-Gal(1 → 3)-β-Gal(1 → 3)-β-Gal2NP							
	A: t-Gal	4.55	3.54	3.59	3.85	3.62	3.68
		104.2	75.2	72.5	68.5	75.1	60.9
	B: 3-Gal	4.61	3.72	3.77	4.13	3.65	3.68
		104.0	70.2	81.9	68.4	74.8	60.9
	C: 3-Gal	4.65	3.73	3.77	4.13	3.66	3.68
		103.7	70.3	82.0	68.4	74.8	60.9
	D: 3-Gal2N	5.17	3.96	3.88	4.20	3.85	3.70
		100.7	68.5	81.6	68.3	75.4	60.9
β-Gal-(1 → 4)-β-Xyl4NP							
	A: t-Gal	4.41	3.45	3.57	3.84	3.70	3.67/3.72
		101.7	70.6	72.6	75.6	75.6	60.6
	B: 4-Xyl4NP	5.12	3.59	3.66	3.88	4.12/3.55	
		99.7	72.4	73.4	76.1	60.8	

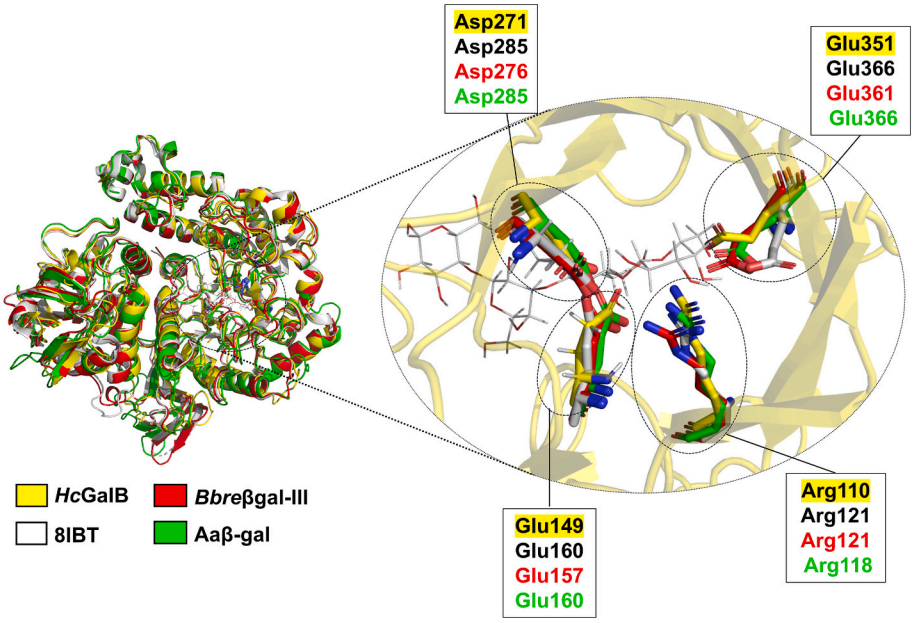

Fig. 7. Hypothesis of residues targeted for site-directed mutagenesis to improve the transglycosylation reaction. Structural alignment highlights the conserved non-catalytic residues located in the active-site pocket, which may contribute to the reduced efficiency of the reaction. Residues from Bbreβgal-III are shown inred, those from Aaβ-gal in green, those from BiBga42A (8IBT) in white, and those from *Hc*GalB inyellow. The lacto-N-tetraose from the BiBga42A structure is shown in whitish. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

3.4. Reaction mechanisms: hydrolysis vs transgalactosylation

Our previous study examined the catalytic behavior (hydrolysis and transgalactosylation) of *Hc*GalB, revealing broad substrate specificity and high hydrolytic efficiency using both lactose and ONP-β-gal as natural and artificial substrates, respectively [9]. Moreover, *Hc*GalB performed both homo-condensation reactions (using ONP-β-gal or lactose as donors and acceptor) as well as hetero condensation reactions (in which PNP-β-glu and PNP-β-xyl are employed as acceptors and ONP-β-gal as a donor) [9]. In this study, we have analyzed the structure of the *Hc*GalB-produced GOS through 2D NMR spectra of the HPLC fractions and results were compared with those already reported [45]. As the lactose was a poor substrate for transgalactosylation reaction (Hydrolysis/Transgalactosylation: 21, data not shown), GOS amount produced from lactose did not allow a detailed structural characterization of such products. Therefore, we focused on the most abundant GOS obtained from synthetic substrates to explore the kind of linkages related to this synthetic activity. Specifically, in homocondensation reactions, ESI–MS analysis identified the formation of oligosaccharides as sodium adducts with *m/z* values of 486.142 (disaccharide), 648.205 (trisaccharide), and 810.277 (tetrasaccharide) [9]. These products consist of one or more galactose units (m/z = 162) added to the galactosyl moiety of ONP-β-gal (m/z = 324). The GOS purified by HPLC and analyzed in this study were confirmed to be a disaccharide, trisaccharide, and tetrasaccharide, as reported in Table 1.

Interestingly, two kinds of disaccharides were found, β-Gal(1 → 6)-β-Gal2Np and β-Gal(1 → 3)-β-Gal2NP, whereas the trisaccharide was associated only with β-Gal(1 → 3)-β-Gal(1 → 3)-β-Gal2NP, and the tetrasaccharides with β-Gal (1 → 3)-β-Gal(1 → 3)-β-Gal(1 → 3)-β-Gal2NP indicating a clear preference in the regioselectivity of *Hc*GalB which favors the formation of β(1 → 3) linkages for chain elongation over β(1 → 6) branch-type disaccharides. This pattern also highlights the distinct catalytic preferences of *Hc*GalB compared to classical commercially available β-galactosidases such as those from *B. circulans* (Biolactase, by Kerry Ingredients and Flavours), which typically produce higher yields of β(1 → 4) and β(1 → 6) GOS [46]. It has been reported that GOS containing β(1 → 3) bonds are metabolized more rapidly by probiotic bacteria than similar GOS containing other linkages, particularly β(1 → 4). Furthermore, there is general consensus that GOS synthesized with β-galactosidases from probiotic strains exert a more pronounced effect on the growth of these beneficial microorganisms [47]. Mechanistically, the regioselectivity of *Hc*GalB over the formation of β(1 → 3) linkages, might be rationalized by the geometry and hydrogen-bonding network of the acceptor-binding subsites in the catalytic pocket: acceptor hydroxyls at O-3 and O-6 may be better positioned for nucleophilic attack on the covalent galactosyl intermediate, whereas the O-2 orientation could be sterically hindered or less favourably stabilized in the transition state. Since the fraction of the products obtained through the homo-condendensation reaction was a complex mixture, a purification by HPLC was performed. Two different fractions were obtained and 2D NMR spectra were collected. The first fraction was revealed to contain two products, namely a disaccharide and a trisaccharide. The disaccharide 1, with anomeric signals at δ 4.34/103.2 ppm and δ 5.14/100.9 ppm, respectively, was recognized to contain a 6-substituted galactose unit due to the downfield shift of the C6 signal at δ 68.9 ppm (Table 1, product 1) [48]. Instead, the second product (disaccharide 2) was constituted by a 3-substitued galactose (Table 1) since the C3 signal of the galactose linked to the 2-nitrophenol was downfield shifted at δ 81.6 ppm. The second HPLC fraction was constituted by four products. Besides the presence of disaccharides 1 and 2, the spectra disclosed that the remaining signals were attributable to those of a trisaccharide (compound 3, Table 1) and a tetrasaccharide (compound 4, Table 1). In both compounds, galactose residues were 3-substituted. The attachment points of the monosaccharides were recognizable by the glycosylation shift of the C3 signals (Table 1) [48].

The biological activity of GOS is strongly influenced by their degree of polymerization and the type of glycosidic linkage: β(1 → 2) linkages, for instance, are abundant in human milk oligosaccharides (HMO) and play a pivotal role in shaping the infant gut microbiota, whereas β(1 → 3), β(1 → 4), and β(1 → 6) linkages contribute to structural diversity and broaden the range of gut bacterial taxa capable of utilizing these substrates as demonstrated in both *in vitro* and *in vivo* studies [49,50] (Sprenger, Mei). β(1 → 2) transfer appears disfavoured or below detection limits for *Hc*GalB. The absence of such linkages is particularly relevant from a functional standpoint, as it indicates a narrower overlap with the structures found in HMOs, yet still supports the production of

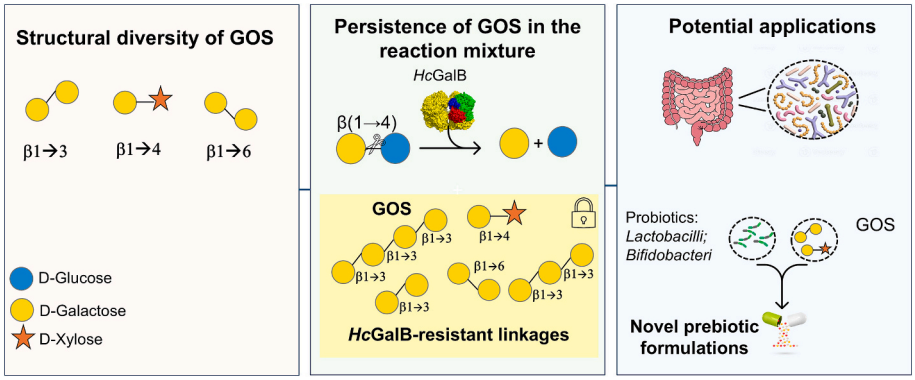

Fig. 8. Future perspectives on the potential applications of *Hc*GalB.

structurally diverse GOS.

In heterocondensation reactions, the only characterized oligosaccharide was found to be β-Gal-(1–4)-β-Xyl-4NP (Table 1, compound **5**), when β-D-Gal-2NP was used as donor and 4-nitrophenyl β-D-xylopyranoside (β-D-Xyl-4NP) as acceptor, respectively. The characterization was obtained by the study of 2D NMR spectra of the compound **5**, which indicated a glycosylation shift for the C4 of the xylose unit [48] . From an applicative standpoint, the presence of GOS containing pentoses, also been reported for YesZ from *Bacillus subtilis*, may be of interest for the development of mixed GOS composed of both hexoses and pentoses [51].

A structural superimposition with three GH42 enzymes (Aaβ-gal from *Alicyclobacillus acidocaldarius*, Bbreβgal-III from *Bifidobacterium breve*, and BiBga42A from *Bifidobacterium longum*) was performed (Fig. 7) [20,51,52]. The selection of these GH42 β-galactosidases was based on previous studies that identified residues suitable for rational mutagenesis aimed at improving transgalactosylation efficiency. Since lactose was found to be a poor substrate for GOS production by *Hc*GalB, future work will focus on enhancing transgalactosylation on this natural and inexpensive substrate. Similar to *Hc*GalB, Aaβ-gal also synthesizes β-(1 → 3)-linked GOS, thus representing an informative comparative model system. In Aaβ-gal, mutation of the invariant non-catalytic Glu361 to Gly increased GOS production up to 177-fold, likely by stabilizing the galactosyl–enzyme intermediate. Moreover, in Bbreβgal-III, the R121C variant reorients the lactose acceptor C3-hydroxyl towards the acid/base residue Glu160, promoting β-(1 → 3)-lactose formation and achieving up to 50% substrate conversion, consistent with the enzyme linkage preference in hydrolysis. Therefore, these residues represent promising targets for future mutagenesis studies aimed at improving *Hc*GalB-mediated GOS synthesis.

4. Conclusions and perspectives

In this work for the first time, we combined cryoEM analysis with biochemical characterization to reveal, the atomic architecture of *Hc*GalB. *Hc*GalB shares structural features with both a thermophilic (A4-β-gal) and a mesophilic (Bca-β-gal) counterpart, consistent with the moderately thermophilic nature of *H. coagulans* MA-13 (T_{opt}: 55 °C), suggesting that the emergence of thermophilic traits may arise from progressive molecular adaptation. *Hc*GalB exhibits a distinctive and stable hexameric, barrel-like assembly formed by two homotrimers arranged in a top-to-top configuration. We suggest that this architecture acts as an intrinsic enhancer transgalactosylation by confining all six active sites towards an internal cavity where catalytic events occur within an occluded volume. Moreover, the uncommon quaternary organization among β-galactosidases may also contribute to the enzyme's thermostability and functional robustness, providing additional opportunities for rational design and industrial application.

A key outcome of this study is the structural diversity and intrinsic stability of the GOS synthesized by *Hc*GalB. The enzyme generates well-defined β(1 → 3)-dominated oligosaccharides, with ancillary β(1 → 6) and β(1 → 4) linkages. Although the complete linkage map of the reaction products could not be fully determined, GOS species persisted in the reaction mixture [9] indicating that the enzyme is unable to hydrolyze the linkages that it forms. This behavior contrasts with that of many β-galactosidases, which readily re-hydrolyze their own transgalactosylation products, thus limiting net GOS accumulation. The formation of hydrolysis-resistant and structurally unconventional products, including hetero-oligosaccharides incorporating C5 sugars, suggests that the enzyme generates substrates that no longer fit the canonical active-site geometry [9]. Overall, *Hc*GalB expands the structural repertoire of GOS beyond classical β(1 → 4)-based mixtures, offering promising perspectives for the development of next-generation, structurally defined prebiotics inspired by HMO-like architectures and for future rational engineering strategies aimed at industrial applications.

Future studies (Fig. 8) should aim to: (i) optimize transgalactosylation conditions with lactose or edible oligosaccharides containing C5 monomers, (ii) map acceptor-binding subsites via site-directed mutagenesis to resolve the structural determinants of regioselectivity, and (iii) engineer *Hc*GalB variants optimized for the high-value, large-scale production of designer prebiotics inspired by the structural logic of HMOs.

CRediT authorship contribution statement

Gennaro Sanità: Writing – original draft, Methodology, Investigation, Conceptualization. **Emanuela Maresca:** Writing – original draft, Methodology, Investigation, Conceptualization. **Stefano Capaldi:** Writing – review & editing, Methodology, Investigation, Data curation, Conceptualization. **Angela Casillo:** Writing – review & editing, Methodology, Investigation, Conceptualization. **Martina Aulitto:** Writing – review & editing, Methodology, Investigation, Conceptualization. **Federica Donadio:** Methodology, Investigation. **Tillmann Pape:** Writing – review & editing, Methodology, Investigation, Data curation, Conceptualization. **Maria Michela Corsaro:** Writing – original draft, Supervision, Investigation, Data curation, Conceptualization. **Emanuela Esposito:** Writing – review & editing, Supervision, Investigation, Data curation, Conceptualization. **Patrizia Contursi:** Writing – review & editing, Writing – original draft, Supervision, Investigation, Funding acquisition, Data curation, Conceptualization.

Fundings

This work was supported by the Ministry of University and Research CUP PRIN 2022 CONTURSI E53D23003010006 and to CN_00000033 MUR decree n. 1034 of 17 June 2022 CUP F13C22000720007, and by investment 1.4 - Call for tender No. 3138 of 16 December 2021, rectified by Decree n.3175 of 18 December 2021 (Project code CN_00000033,

Concession Decree No. 1034 of 17 June 2022, Project title "National Biodiversity Future Center - NBFC") of Italian Ministry of University and Research funded by the European Union – NextGenerationEU.

"PROGETTO CIR01_00023 - "IMPARA - Imaging Dalle Molecole Alla Preclinica Rafforzamento Del Capitale Umano" CUP B28I20000150001 and National Recovery and Resilience Plan (NRRP), Project "Cascata-MNESYS (PE0000006) - "A multiscale integrated approach to the study of nervous system in health and disease" (DN. 1553 11.10.2022) to EDL.

Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: Given my role as Handling Editor, Maria Michela Corsaro had no involvement in the peer review of the article IJBIOMAC-D-25-32,409 (corresponding author Patrizia Contursi) and had no access to information regarding its peer review. Full responsibility for the editorial process for this article was delegated to another journal editor If there are other authors, they declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

Authors are grateful to Dr. Roberto Consonni, National Research Council Istituto di Scienze e Tecnologie Chimiche "Giulio Natta" NMR Laboratory, and Prof. Andrea Strazzulli and Prof. Marco Moracci, University of Naples Federico II, for helpful and constructive scientific discussion.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.ijbiomac.2026.151980.

Data availability

Data will be made available on request.

PDB ID : 9TA3 / pdb_00009ta3, EMDB ID : EMD-55743 (Original data) (PDB)

References

[1] D.H. Juers, B.W. Matthews, R.E. Huber, LacZ β-galactosidase: structure and function of an enzyme of historical and molecular biological importance, Protein science: a publication of the Protein Society. 21 (12) (2012) 1792–1807, https://doi.org/10.1002/pro.2165.

[2] D. Li, Z. Wang, Y. Yu, H. Li, W. Luo, B. Chen, G. Niu, H. Ding, Biochemical insights into a novel family 2 glycoside hydrolase with both β-1,3-galactosidase and β-1,4-galactosidase activity from the Arctic, Mar. Drugs 21 (10) (2023) 521, https://doi.org/10.3390/md21100521.

[3] S. Ruiz-Ramírez, R. Jiménez-Flores, Invited review: Properties of b-galactosidases derived from Lactobacillaceae species and their capacity for galacto-oligosaccharide production, J. Dairy Sci. 106 (12) (2023) 8193–8206, https://doi.org/10.3168/jds.2023-23392.

[4] Y. Zhou, Y. Liu, F. Gao, Z. Xia, Z. Zhang, F.P. Addai, Y. Zhu, J. Chen, F. Lin, D. Chen, β-Galactosidase: Insights into source variability, genetic engineering, immobilisation and diverse applications in food, industry and medicine, Int. J. Dairy Technol. 77 (3) (2024) 651–679, https://doi.org/10.1111/1471-0307.13098.

[5] I. Ignatova, A. Arsov, P. Petrova, K. Petrov, Prebiotic effects of α- and β-galactooligosaccharides: the structure-function relation, Molecules 30 (4) (2025) 803, https://doi.org/10.3390/molecules30040803.

[6] J. Atalah, P. Cáceres-Moreno, G. Espina, J.M. Blamey, Thermophiles and the applications of their enzymes as new biocatalysts, Bioresour. Technol. 280 (2019) 478–488, https://doi.org/10.1016/j.biortech.2019.02.008.

[7] N.M. Mesbah, Industrial biotechnology based on enzymes from extreme environments, Front. Bioeng. Biotechnol. 10 (2022) (2022), https://doi.org/10.3389/fbioe.2022.870083.

[8] S. Hait, S. Mallik, S. Basu, S. Kundu, Finding the generalized molecular principles of protein thermal stability, Proteins 88 (6) (2020) 788–808, https://doi.org/10.1002/prot.25866.

[9] M. Aulitto, A. Strazzulli, F. Sansone, F. Cozzolino, M. Monti, M. Moracci, G. Fiorentino, D. Limauro, S. Bartolucci, P. Contursi, Prebiotic properties of *Bacillus coagulans* MA-13: production of galactoside hydrolyzing enzymes and characterization of the transglycosylation properties of a GH42 β-galactosidase, Microb. Cell Factories 20 (1) (2021) 71, https://doi.org/10.1186/s12934-021-01553-y.

[10] M. Aulitto, F.A. Fusco, G. Fiorentino, S. Bartolucci, P. Contursi, D. Limauro, A thermophilic enzymatic cocktail for galactomannans degradation, Enzym. Microb. Technol. 111 (2018) 7–11, https://doi.org/10.1016/j.enzmictec.2017.12.008.

[11] P. Contursi, K. D'Ambrosio, L. Pirone, E. Pedone, T. Aucelli, Q. She, G. De Simone, S. Bartolucci, C68 from the *Sulfolobus islandicus* plasmid-virus pSSVx is a novel member of the AbrB-like transcription factor family, Biochem. J. 435 (1) (2011) 157–166, https://doi.org/10.1042/BJ20101334.

[12] P. Contursi, B. Farina, L. Pirone, S. Fusco, L. Russo, S. Bartolucci, R. Fattorusso, E. Pedone, Structural and functional studies of Stf76 from the *Sulfolobus islandicus* plasmid-virus pSSVx: a novel peculiar member of the winged helix-turn-helix transcription factor family, Nucleic Acids Res. 42 (9) (2014) 5993–6011, https://doi.org/10.1093/nar/gku215.

[13] D. Talens-Perales, J. Polaina, J. Marín-Navarro, Structural Dissection of the Active Site of *Thermotoga maritima* β-Galactosidase Identifies Key Residues for Transglycosylating Activity, J. Agric. Food Chem. 64 (14) (2016) 2917–2924, https://doi.org/10.1021/acs.jafc.6b00222.

[14] D.H. Juers, T.D. Heightman, A. Vasella, J.D. McCarter, L. Mackenzie, S.G. Withers, B.W. Matthews, A structural view of the action of *Escherichia coli* (lacZ) beta-galactosidase, Biochemistry 40 (49) (2001) 14781–14794, https://doi.org/10.1021/bi011727i.

[15] M. Mangiagalli, M. Lapi, S. Maione, M. Orlando, S. Brocca, A. Pesce, A. Barbiroli, C. Camilloni, S. Pucciarelli, M. Lotti, M. Nardini, The co-existence of cold activity and thermal stability in an Antarctic GH42 β-galactosidase relies on its hexameric quaternary arrangement, FEBS J. 288 (2) (2021) 546–565, https://doi.org/10.1111/febs.15354.

[16] T. Huang, J. Khan, S. Lakhani, A. Li, A. Vyas, J. Hunt, S.A. Espinosa Garcia, B. Liang, The crystallography of enzymes: a retrospective and beyond, Crystals 15 (11) (2025) 966, https://doi.org/10.3390/cryst15110966.

[17] X. Bai, G. McMullan, S.H.W. Scheres, How cryo-EM is revolutionizing structural biology, Trends Biochem. Sci. 40 (1) (2015) 49–57, https://doi.org/10.1016/j.tibs.2014.10.005.

[18] P. Liu, Y. Chen, C. Ma, J. Ouyang, Z. Zheng, β-Galactosidase: a traditional enzyme given multiple roles through protein engineering, Crit. Rev. Food Sci. Nutr. 65 (7) (2025) 1306–1325, https://doi.org/10.1080/10408398.2023.2292282.

[19] C. Guerrero, C. Vera, R. Conejeros, A. Illanes, Transgalactosylation and hydrolytic activities of commercial preparations of β-galactosidase for the synthesis of prebiotic carbohydrates, Enzym. Microb. Technol. 70 (2015) 9–17, https://doi.org/10.1016/j.enzmictec.2014.12.006.

[20] A. Strazzulli, B. Cobucci-Ponzano, S. Carillo, E. Bedini, M.M. Corsaro, G. Pocsfalvi, S.G. Withers, M. Rossi, M. Moracci, Introducing transgalactosylation activity into a family 42 β-galactosidase, Glycobiology 27 (5) (2017) 425–437, https://doi.org/10.1093/glycob/cwx013.

[21] D.P.M. Torres, M.P.F. Gonçalves, J.A. Teixeira, L.R. Rodrigues, Galacto-Oligosaccharides: Production, Properties, Applications, and Significance as Prebiotics, Compr. Rev. Food Sci. Food Saf. 9 (5) (2010) 438–454, https://doi.org/10.1111/j.1541-4337.2010.00119.x.

[22] M.G. Gänzle, Enzymatic synthesis of galacto-oligosaccharides and other lactose derivatives (hetero-oligosaccharides) from lactose, Int. Dairy J. 22 (2) (2012) 116–122, https://doi.org/10.1016/j.idairyj.2011.06.010.

[23] G. Vinderola, M.E. Sanders, S. Salminen, H. Szajewska, Postbiotics: the concept and their use in healthy populations, Front. Nutr. 9 (2022) (2022), https://doi.org/10.3389/fnut.2022.1002213.

[24] G.K.P. Guron, A.T. Hotchkiss, B.H. Bodnar, A. Harron, J.A. Renye, M.J. McAnulty, Oligosaccharide production using b-galactosidase from *Lactobacillus bulgaricus* and Kluyveromyces lactis in sweetened reconstituted nonfat dry milk, J. Dairy Sci. 108 (6) (2025) 5696–5704, https://doi.org/10.3168/jds.2025-26396.

[25] W.A. Abdel Wahab, S.A. Ahmed, A.M.M. Kholif, S.A. El Ghani, H.R. Wehaidy, Low-lactose yoghurt production using β-galactosidase: An integrated study for the enzyme and its application, Int. Dairy J. 151 (2024) 105864, https://doi.org/10.1016/j.idairyj.2023.105864.

[26] C. Martínez-Villaluenga, J. Frias, C. Vidal-Valverde, Alpha-Galactosides: Antinutritional Factors or Functional Ingredients? Crit. Rev. Food Sci. Nutr. 48 (4) (2008) 301–316, https://doi.org/10.1080/10408390701326243.

[27] S.I. Mussatto, I.M. Mancilha, Non-digestible oligosaccharides: a review, Carbohydr. Polym. 68 (3) (2007) 587–597, https://doi.org/10.1080/10408690091189239.

[28] S.S. van Leeuwen, B.J.H. Kuipers, L. Dijkhuizen, J.P. Kamerling, 1H NMR analysis of the lactose/β-galactosidase-derived galacto-oligosaccharide components of Vivinal® GOS up to DP5, Carbohydr. Res. 400 (2014) 59–73, https://doi.org/10.1016/j.carres.2014.08.012.

[29] S. Míguez Amil, E. Jiménez-Ortega, M. Ramírez-Escudero, D. Talens-Perales, J. Marín-Navarro, J. Polaina, J. Sanz-Aparicio, R. Fernandez-Leiro, The cryo-EM Structure of Thermotoga maritima β-Galactosidase: Quaternary Structure Guides Protein Engineering, ACS Chem. Biol. 15 (1) (2020) 179–188, https://doi.org/10.1021/acschembio.9b00752.

[30] A. Punjani, J.L. Rubinstein, D.J. Fleet, M.A. Brubaker, cryoSPARC: algorithms for rapid unsupervised cryo-EM structure determination, Nat. Methods 14 (3) (2017) 290–296, https://doi.org/10.1038/nmeth.4169.
[31] A. Punjani, H. Zhang, D.J. Fleet, Non-uniform refinement: adaptive regularization improves single-particle cryo-EM reconstruction, Nat. Methods 17 (12) (2020) 1214–1221, https://doi.org/10.1038/s41592-020-00990-8.
[32] T.D. Goddard, C.C. Huang, E.C. Meng, E.F. Pettersen, G.S. Couch, J.H. Morris, T. E. Ferrin, UCSF chimeraX: meeting modern challenges in visualization and analysis, Protein Sci. 27 (1) (2018) 14–25, https://doi.org/10.1002/pro.3235.
[33] P.D. Adams, P.V. Afonine, G. Bunkóczi, V.B. Chen, I.W. Davis, N. Echols, J. J. Headd, L.W. Hung, G.J. Kapral, R.W. Grosse-Kunstleve, A.J. McCoy, N. W. Moriarty, R. Oeffner, R.J. Read, D.C. Richardson, J.S. Richardson, T. C. Terwilliger, P.H. Zwart, PHENIX: a comprehensive Python-based system for macromolecular structure solution, Acta Crystallogr. D Biol. Crystallogr. 66 (Pt 2) (2010) 213–221, https://doi.org/10.1107/S0907444909052925.
[34] A.H. Emwas, K. Szczepski, B.G. Poulson, K. Chandra, R.T. McKay, M. Dhahri, F. Alahmari, L. Jaremko, J.I. Lachowicz, M. Jaremko, Nmr as a "gold standard" method in drug design and discovery, Molecules 25 (20) (2020) 4597, https://doi.org/10.3390/molecules25204597.
[35] D. Liebschner, P.V. Afonine, M.L. Baker, G. Bunkóczi, V.B. Chen, T.I. Croll, B. Hintze, L.W. Hung, S. Jain, A.J. McCoy, N.W. Moriarty, R.D. Oeffner, B.K. Poon, M.G. Prisant, R.J. Read, J.S. Richardson, D.C. Richardson, M.D. Sammito, O. V. Sobolev, D.H. Stockwell, T.C. Terwilliger, A.G. Urzhumtsev, L.L. Videau, C. J. Williams, P.D. Adams, Macromolecular structure determination using X-rays, neutrons and electrons: recent developments in Phenix, Acta crystallogr. D Struct. Biol. 75 (Pt 10) (2019) 861–877, https://doi.org/10.1107/S2059798319011471.
[36] V.B. Chen, W.B. Arendall III, J.J. Headd, D.A. Keedy, R.M. Immormino, G.J. Kapral, L.W. Murray, J.S. Richardson, D.C. Richardson, MolProbity: all-atom structure validation for macromolecular crystallography, Acta Crystallogr. D Biol. Crystallogr. 66 (Pt 1) (2010) 12–21, https://doi.org/10.1107/S0907444909042073.
[37] G.M. Morris, R. Huey, W. Lindstrom, M.F. Sanner, R.K. Belew, D.S. Goodsell, A. J. Olson, AutoDock4 and AutoDockTools4: Automated docking with selective receptor flexibility, J. Comput. Chem. 30 (16) (2009) 2785–2791, https://doi.org/10.1002/jcc.21256.
[38] O. Trott, A.J. Olson, AutoDock Vina: Improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading, J. Comput. Chem. 31 (2) (2010) 455–461, https://doi.org/10.1002/jcc.21334.
[39] P.W. Rose, C. Bi, W.F. Bluhm, C.H. Christie, D. Dimitropoulos, S. Dutta, R.K. Green, D.S. Goodsell, A. Prlic, M. Quesada, G.B. Quinn, A.G. Ramos, J.D. Westbrook, J. Young, C. Zardecki, H.M. Berman, P.E. Bourne, The RCSB protein data bank: new resources for research and education, Nucleic Acids Res. 41 (Database issue) (2013) D475–D482, https://doi.org/10.1093/nar/gks1200.
[40] M. Maksimainen, S. Paavilainen, N. Hakulinen, J. Rouvinen, Structural analysis, enzymatic characterization, and catalytic mechanisms of β-galactosidase from *Bacillus circulans sp. alkalophilus*, FEBS J. 279 (10) (2012) 1788–1798, https://doi.org/10.1111/j.1742-4658.2012.08555.x.
[41] M. Hidaka, S. Fushinobu, N. Ohtsu, H. Motoshima, H. Matsuzawa, H. Shoun, T. Wakagi, Trimeric crystal structure of the glycoside hydrolase family 42 beta-galactosidase from *Thermus thermophilus* A4 and the structure of its complex with galactose, J. Mol. Biol. 322 (1) (2002) 79–91, https://doi.org/10.1016/s0022-2836(02)00746-5.
[42] Y.N. Dong, H.Q. Chen, Y.H. Sun, H. Zhang, W. Chen, A differentially conserved residue (Ile42) of GH42 β-galactosidase from *Geobacillus stearothermophilus* BgaB is involved in both catalysis and thermostability, J. Dairy Sci. 98 (4) (2015) 2268–2276, https://doi.org/10.3168/jds.2014-9117.
[43] H.V. Solomon, O. Tabachnikov, S. Lansky, R. Salama, H. Feinberg, Y. Shoham, G. Shoham, Structure-function relationships in Gan42B, an intracellular GH42 β-galactosidase from *Geobacillus stearothermophilus*, Acta Crystallogr. D Biol. Crystallogr. 71 (Pt 12) (2015) 2433–2448, https://doi.org/10.1107/S1399004715018672.
[44] S. Mine, M. Watanabe, Structural insights into the molecular evolution of the archaeal exo-β-d-glucosaminidase, Int. J. Mol. Sci. 20 (10) (2019) 2460, https://doi.org/10.3390/ijms20102460.
[45] J. Zhao, D. Niu, J. Liu, Z. Jin, N.P. McHunu, S. Singh, Z. Wang, Enhancing β-galactosidase performance for galactooligosaccharides preparation via strategic glucose re-tunneling, Int. J. Mol. Sci. 25 (22) (2024) 12316, https://doi.org/10.3390/ijms252212316.
[46] B. Rodriguez-Colinas, A. Poveda, J. Jimenez-Barbero, A.O. Ballesteros, F.J. Plou, Galacto-oligosaccharide synthesis from lactose solution or skim milk using the β-galactosidase from *Bacillus circulans*, J. Agric. Food Chem. 60 (25) (2012) 6391–6398, https://doi.org/10.1021/jf301156v.
[47] V. Füreder, B. Rodriguez-Colinas, F.V. Cervantes, L. Fernandez-Arrojo, A. Poveda, J. Jimenez-Barbero, A.O. Ballesteros, F.J. Plou, Selective Synthesis of Galactooligosaccharides Containing β(1→3) Linkages with β-Galactosidase from *Bifidobacterium bifidum* (Saphera), J. Agric. Food Chem. 68 (17) (2020) 4930–4938, https://doi.org/10.1021/acs.jafc.0c00997.
[48] K. Bock, C. Pedersen, Carbon-13 Nuclear Magnetic Resonance Spectroscopy of Monosaccharides, in: R.S. Tipson, D. Horton (Eds.), Advances in carbohydrate chemistry and biochemistry, Academic Press, 1983, pp. 27–66.
[49] N. Sprenger, H.L.P. Tytgat, A. Binia, S. Austin, A. Singhal, Biology of human milk oligosaccharides: From basic science to clinical evidence, Journal of human nutrition and dietetics: the official journal of the British Dietetic Association. 35 (2) (2022) 280–299, https://doi.org/10.1111/jhn.12990.
[50] Z. Mei Z, J. Yuan, D. Li, Biological activity of galacto-oligosaccharides: A review, Front. Microbiol. 13 (2022) 993052, https://doi.org/10.3389/fmicb.2022.993052.
[51] L.A.B.C. Carneiro, L. Yu, P. Dupree, R.J. Ward, Characterization of a β-galactosidase from *Bacillus subtilis* with transgalactosylation activity, Int. J. Biol. Macromol. 120 (2018) 279–287, https://doi.org/10.1016/j.ijbiomac.2018.07.116.
[52] K. Phirom-On, K.T. Vu-Le, L. Sützl, B. Lehner, D. Whelan, L. Guerent, I. Pasini, M. Schuh, A. de Ruiter, M. Blaukopf, D. Haltrich, C. Oostenbrink, T. Nguyen, Unveiling the Transgalactosylation Switch of a GH42 β-Galactosidase from the Infant Isolate *Bifidobacterium breve* DSM20213, ACS Catal. 16 (3) (2026) 2669–2686, https://doi.org/10.1021/acsomega.5c0681.