

A STRUCTURAL SURVEY OF CATALYTIC RESIDUES AND RING CONFORMATIONS IN GLYCOSIDE HYDROLASES FROM THE CAZY DATABASE

Wijitra Jitonnorn,^{1,2,3} Panthip Tue-ngeun,⁴ Sopida Wongwas,⁴ James R. Ketudat Cairns⁵, Pitak Nasomjai,^{2,3} and Jitrayut Jitonnorn^{2,3*}

Received: October 19, 2025; Revised: December 22, 2025; Accepted: December 29, 2025

Abstract

In this study, we performed a comprehensive statistical analysis and a detailed survey of carbohydrate ligands bound within the active sites of carbohydrate-active enzymes (CAZymes) from the CAZy database (www.cazy.org). Specifically, we analyzed 3D structures of CAZyme-carbohydrate ligand complexes across 184 glycoside hydrolase families. By utilizing only coordinate files, we were able to retrieve, analyze, and interpret key structural and functional features. These include previously underexplored aspects, such as Cremer-Pople and Altona-Sundaralingam analyses of sugar ring conformations at the subsite -1 sugar ring, identification of catalytic residues (acid/base and nucleophile), and population trends. This study provides structural insights into catalytic residue patterns and ring conformational preferences in CAZymes, thereby supporting future efforts in functional analysis, mechanistic interpretation, and elucidation of catalytic itineraries.

Keywords: Carbohydrate-Ligand Interactions; 3D Structure Analysis; Glycoside Hydrolase Families; Sugar Conformational Analysis

¹ Demonstration School, University of Phayao, Phayao 56000, Thailand. E-mail: wijitra.me@up.ac.th, ORCID: 0009-0004-3060-315X

² Unit of Excellence in Computational Molecular Science and Catalysis, University of Phayao, Phayao 56000, Thailand. E-mail: wijitra.me@up.ac.th, ORCID: 0009-0004-3060-315X; pitak.na@up.ac.th, ORCID: 0009-0000-3047-4301; jitrayut.ji@up.ac.th, ORCID: 0000-0002-5041-3343

³ Division of Chemistry, School of Science, University of Phayao, Phayao 56000, Thailand. E-mail: jitrayut.ji@up.ac.th, ORCID: 0000-0002-5041-3343

⁴ Program of Chemistry, Faculty of Science and Technology, Uttaradit Rajabhat University, Uttaradit, 53000, Thailand. E-mail: ptuengeun@uru.ac.th, ORCID: 0009-0007-0662-4243; sopida.won@uru.ac.th, ORCID: 0000-0002-5770-9745

⁵ Center for Biomolecular Structure, Function and Application and School of Chemistry, Institute of Science, Suranaree University of Technology, Nakhon Ratchasima 30000, Thailand. E-mail: cairns@sut.ac.th, ORCID: 0000-0003-3042-1626

* Corresponding author

DOI: <https://doi.org/10.55766/sujst11867>

Introduction

Carbohydrate-active enzymes (CAZymes) play a pivotal role in the metabolism, synthesis, and modification of complex carbohydrates, underpinning diverse biological processes and industrial applications (Bains *et al.*, 2024; Wardman and Withers, 2024). These enzymes include glycoside hydrolases (GHs), glycosyltransferases, polysaccharide lyases, carbohydrate esterases, and auxiliary activities, which are collectively cataloged in the CAZy database (Lombard *et al.*, 2014). The CAZy database (www.cazy.org) has become an indispensable resource for researchers, offering a wealth of information on CAZyme families, with links to sequence data, and an expanding repository of three-dimensional (3D) structures in public databases. Other commonly used databases that contain the 3D CAZyme structures in CAZy and analyze them for functional and structural information include dbCAN (Yin *et al.*, 2012), UniProt (The UniProt Consortium, 2023), and Catalytic Site Atlas (Porter *et al.*, 2004). However, none of these provide detailed structural information about carbohydrate conformations and their variations at the catalytic site of CAZymes.

Understanding CAZyme-carbohydrate interactions is crucial for elucidating the mechanisms of substrate recognition, binding, and catalysis, enabling new approaches to screening novel CAZymes (Bains *et al.*, 2024; Wardman and Withers, 2024; Davies *et al.*, 2005). The structural features of CAZymes, including their active site architecture and the conformation of bound carbohydrate ligands, are key to their specificity and catalytic efficiency (Davies and Henrissat, 1995). Recent advances in crystallography and computational tools have enabled detailed studies of CAZyme-ligand complexes, yet significant gaps remain in our ability to interpret structural data comprehensively (Ardévol and Rovira, 2015; Jitonom *et al.*, 2023). Notably, current tools for assessing sugar conformations at catalytic centers—particularly the -1 subsite—are limited in their ability to analyze large and diverse datasets. Improved methods for evaluating these conformations would significantly enhance the analysis of CAZyme-ligand interactions and provide deeper insights into enzyme mechanisms across families.

In this study, we address missing features by conducting a comprehensive analysis of three-dimensional (3D) CAZyme structures, focusing on 184 glycoside hydrolase (GH) families. Using structural coordinate files, we extracted key insights into enzyme functionality at the catalytic center, including Cremer-Pople and Altona-Sundaralingam

analyses of the -1 sugar conformations and the characterization of catalytic residues. By identifying trends and previously unknown structural features across enzyme families, this work provides a deeper understanding of structure-function relationships in these important enzymes, which can support future CAZyme-related research.

Materials and Methods

The general procedure employed in this study is summarized in Figure 1, which presents a step-by-step schematic representation of the workflow. The process can be divided into four main steps: input, 3D data survey, data analysis, and output. Detailed descriptions of relevant methods are provided below.

Data Collection and Selection

3D structural data of carbohydrate-active enzymes (CAZymes) were retrieved from the CAZy database (www.cazy.org) (Cantarel *et al.*, 2009; Lombard *et al.*, 2014) (Figure 2). To focus our analysis on glycoside hydrolase (GH) families due to their relevance in glycosidic bond cleavage (Davies and Henrissat, 1995), we downloaded structural coordinate files of CAZyme-carbohydrate ligand complexes from the Protein Data Bank (PDB) (Berman *et al.*, 2000), based on the PDB entries listed in CAZy. Structures lacking a carbohydrate ligand in the active site were excluded using visual inspection and VMD tools to ensure the dataset specifically represented protein-carbohydrate interactions. Any glycosylated enzymes, as well as those where the glycosylation moiety docks in the active site, were excluded from this study. A total of 184 GH families with carbohydrate ligand complexes were identified, encompassing over 1,400 PDB entries. This dataset highlights the diversity of GH-mediated carbohydrate recognition and catalysis, providing a foundation for further analysis.

Structural Analysis and Catalytic Residue Identification

The retrieved coordinate files were processed using PyMOL (Schrödinger, 2015) and VMD (Humphrey *et al.*, 1996) for initial inspection and annotation of structural features. Active sites were identified based on the proximity of ligands to conserved catalytic residues, including acid/base catalysts and nucleophiles. These residues were used as reference and supported by comparison with

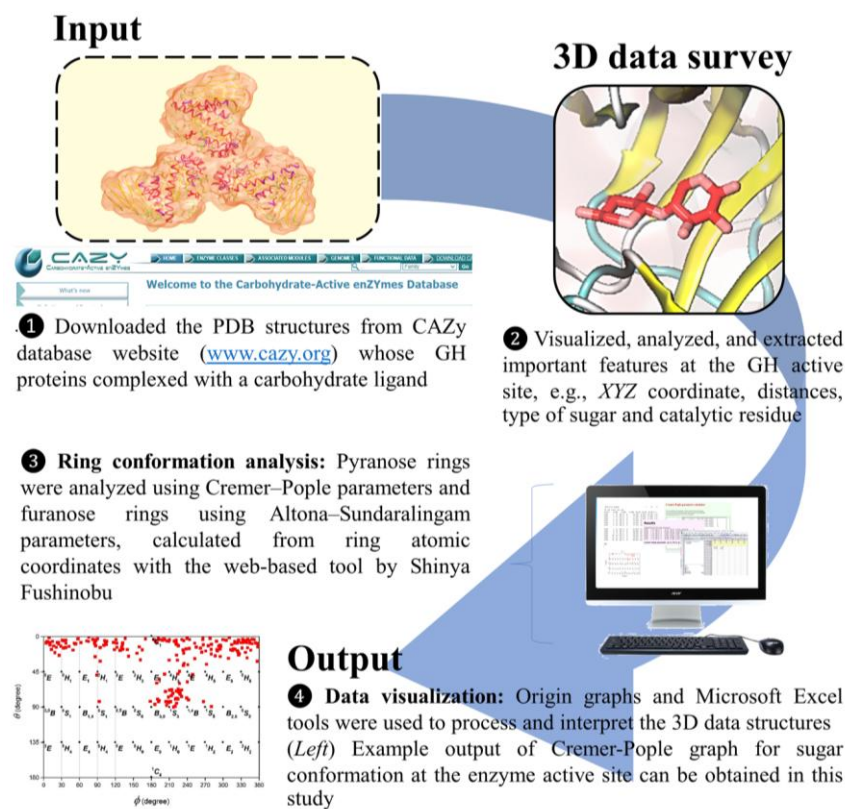


Figure 1. Schematic representation of the steps used in this work

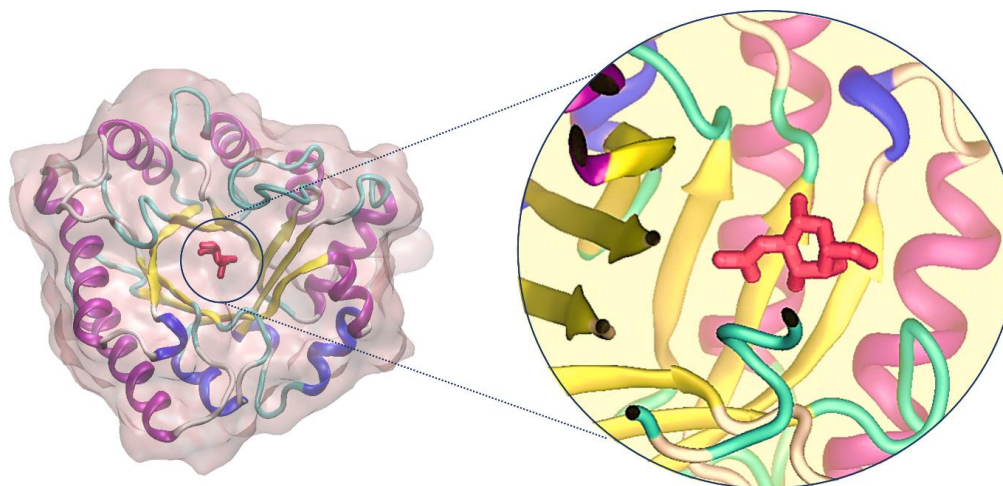


Figure 2. Representative 3D structure of a CAZyme bound to a carbohydrate ligand, highlighting interactions at the enzyme active site, especially subsite -1. The data were obtained from the CAZY database (www.cazy.org)

previously characterized GHs from the literature. Selection of candidate acid/base and nucleophile residues was based on their spatial positioning relative to the substrate and known catalytic residues in the enzyme's active site, as observed

in X-ray crystal structures. Subsite -1 was operationally defined as the sugar-binding position in which the anomeric carbon of the substrate is placed for catalysis. Residues were considered to interact productively with the subsite -1 if they were

appropriately oriented toward the sugar ring and fulfilled a geometric distance criterion.

Specifically, carboxylate-containing residues (Asp or Glu) whose oxygen atoms were located within 3.5 Å of the anomeric carbon were retained as candidate catalytic nucleophiles. This cutoff reflects a physically reasonable distance for nucleophilic participation in glycosidic bond cleavage or formation. To avoid ambiguity arising from non-catalytic sugar binding, only ligands occupying the active site with a well-defined glycosidic orientation were analyzed. In cases where the enzyme was crystallized with an excess of sugar, such as glucose, resulting in multiple copies of the saccharide bound at various sites within the structure, these additional binding sites were not considered. These residues were mapped onto the active sites to analyze trends across GH families.

Analysis of Sugar Ring Conformation

Sugar conformations at the -1 position of GHs (Davies and Henrissat, 1995) were analyzed by calculating their Cremer-Pople and Altona-Sundaralingam parameters to determine their puckering states. This analysis characterized carbohydrate conformations and their mechanistic relevance in enzyme catalysis (Ardèvol and Rovira, 2015). Cremer-Pople parameters describe the puckering of six-membered pyranose rings based on their spherical polar coordinates (meridian angle ϕ , azimuthal angle θ , and radius Q) and a Stoddard (or “pseudorotational”) diagram, which projects puckering conformations onto a sphere. According to IUPAC nomenclature, 38 basic conformations describe pyranose ring puckering. For complexes containing five-membered furanose rings, the Altona-Sundaralingam parameters were used, which represent the magnitude of puckering (ϕ_m) and the phase angle (P) (Altona and Sundaralingam, 1972; Taha *et al.*, 2013). These parameter values were obtained using a web-based tool developed by Shinya Fushinobu (<http://enzyme13.bt.a.u-tokyo.ac.jp/CP/> and <http://enzyme13.bt.a.u-tokyo.ac.jp/CP/altonas.html>).

Results and Discussion

Distribution and Specificity of CAZyme-Carbohydrate structures

The structural diversity of CAZymes within different GH families reflects the adaptation of many protein folds to accommodate the various carbohydrate ligands. Figure 3 highlights the number of 3D structures of CAZymes from 184 GH families with furanose or pyranose ligands

at the -1 subsite. The included ligands are classified as monosaccharides, disaccharides, or oligosaccharides (DP>2), comprising 38.27%, 27.16%, and 34.57% for furanose and 36.02%, 24.36%, and 39.62% for pyranose, respectively. A total of 1,450 PDB structures with ligands were analyzed, representing 54% of all CAZy GH structures available at the time of access (10 November 2024).

In the dataset with furanose ligands ($N=81$), GH32 stands out with the highest number of structures (21), followed by GH68 (13) (Figure 3(A)). This trend underscores the prevalence of these families in accommodating furanose substrates, which is consistent with their biological roles in processing furanose-containing carbohydrates such as fructans (Lammens *et al.*, 2009; Alméida-Díaz *et al.*, 2011). A statistical analysis of the data in Figure 3 further reveals that furanose ligands are predominantly monosaccharides (38.3%), with disaccharides (27.2%) and oligosaccharides with DP>2 (34.6%) also contributing significantly.

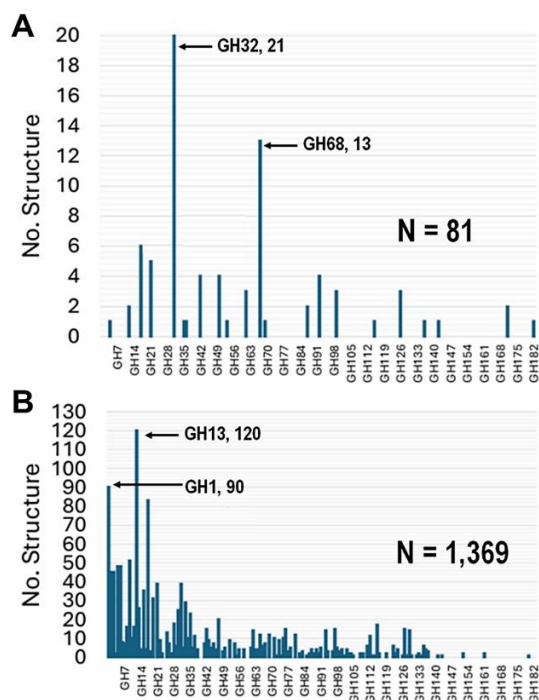


Figure 3. Bar charts showing the distribution of furanose (A) and pyranose (B) ligands at the -1 subsite across CAZyme structures from 184 GH families. Top-ranking GH families are indicated with arrows. Only canonical carbohydrate ligands were included; mimetics were excluded.

The larger dataset with pyranose ligands ($N=1,369$) shows a broader structural representation (Figure 3). GH13 and GH1 dominate with 120 and 90 structures, respectively. This reflects the widespread occurrence of pyranose-containing carbohydrates, such as glucose and galactose, in biological systems (Sinnott, 2007). This data includes a slightly higher proportion of oligosaccharides (39.6%) compared to monosaccharides (36.0%) and disaccharides (24.4%). This distribution highlights the central role of pyranose-utilizing CAZymes in breaking down carbohydrate polymers, consistent with their importance in metabolic and industrial processes.

Comparing the two datasets reveals notable differences in the GH families and substrate distributions associated with furanose and pyranose ligands. The prominence of GH32 and GH68 in the furanose dataset contrasts with the dominance of GH13 and GH1 in the pyranose dataset, underscoring distinct substrate specificity among GH families. Furthermore, the higher representation of oligosaccharides in the pyranose dataset aligns with the structural characteristics of pyranose carbohydrates, which are commonly found in polysaccharides with straight backbones composed of pyranose rings with alternating ring orientations that give favorable internal hydrogen bonding.

These findings underscore the structural and functional diversity of CAZymes, which is essential for their role in nature and industrial applications. Understanding the substrate preferences and structural diversity of GH families provides a foundation for tailoring CAZymes to specific biotechnological needs, such as the synthesis of functional oligosaccharides or the depolymerization of complex polysaccharides (Bains *et al.*, 2024; Wardman and Withers, 2024).

Amino Acid Distribution in CAZymes: Nucleophiles and Acid/Base Residues in Furanose and Pyranose Binding

The bar charts in Figure 4 highlight the distribution of key amino acids involved in the catalytic mechanism of GHs from various families, specifically focusing on their roles as nucleophiles and acid/base residues. The data were obtained from CAZyme PDB structures binding furanose and pyranose carbohydrate ligands at subsite -1.

Figure 4(A) illustrates the distribution of amino acid types at the positions of catalytic residues across top-ranking glycoside hydrolase (GH) families that utilize furanose ligands at subsite -1. GH32 exhibits a similar carboxylate-based mechanism, although with reversed roles of Glu and Asp. Interestingly, GH68 shows a balanced use of

Glu and Asp for both catalytic positions, indicating functional interchangeability between these two residues in this family. In contrast, Figure 4(B) presents amino acid distributions in GH families that act on pyranose ligands. GH1 shows a strong preference for Glu, serving as both the nucleophile and acid/base. In GH13, Asp is the principal nucleophile, while Glu appears in a substantial fraction as the acid/base catalyst. Note that Glu remains the most frequent nucleophile and maintains some consistency with its furanose-active counterparts, while Asn, Gln and Ser are likely from mutations used to prevent hydrolysis of the substrate in the crystal.

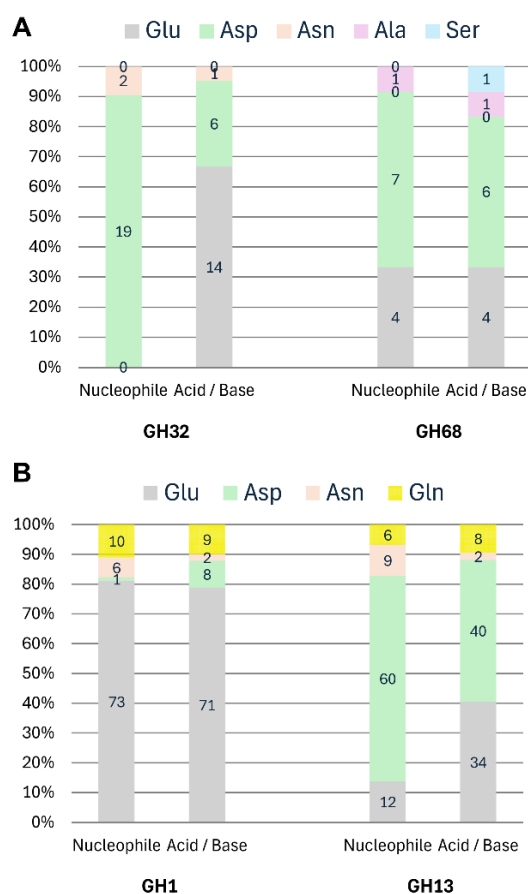


Figure 4. Top-ranking GH families and the most frequently found amino acids as nucleophile and acid/base, using CAZymes PDB structures whose carbohydrate ligands contain (A) furanose and (B) pyranose rings at subsite -1. Note that Asn, Ala, Gln and Ser residues are generally found due to mutation of the catalytic residues to avoid substrate hydrolysis in the crystal structure

These results suggest a clear difference between families utilizing furanose versus pyranose ligands, with variations in the specific amino acids involved in catalysis. The prevalence of glutamic acid as both nucleophile and acid/base residue across the majority of families underscores its central role in carbohydrate hydrolysis.

Conformational Analysis of the Furanose and Pyranose Ring in the Binding Subsite -1

Figure 5 presents the analysis of sugar conformations at the catalytic site from GH families with the highest population. The plots of Altona-Sundaralingam and Cremer-Pople maps show the distribution of furanose and pyranose conformations of carbohydrate ligands at the catalytic site observed in the crystal structures.

The Altona-Sundaralingam and Cremer-Pople parameter maps (Figure 5) reveal the primary and secondary conformation populations of the -1 subsite sugar ring. The solid arrows indicate the most frequently observed conformations, while the dashed arrows represent secondary conformations that are less prevalent. The furanose ligands in this dataset predominantly adopt

the ${}^1T_2/E_2/{}^3T_2$ conformations for GH32 and the $E_2/{}^3T_2/{}^3E$ conformations for GH68 (Figure 5(A)), while the pyranose ligands show a preference for the stable chair conformation (4C_1), with distortions into other boat or skew conformations (Figure 5(B)). Analysis of less populated GH families shows that many share dominant sugar conformations seen in major families, while a few display unique or less common conformations. This reflects both conformational preferences of the monosaccharide and the catalytic itinerary by which it is converted to products, along with small, enzyme family-specific variations that fine-tune sugar conformations at the -1 subsite. Since the glycosidic bond is generally broken faster than the new bond is formed, the transition state typically has oxocarbenium ion-like character in which the endocyclic oxygen donates electron density to the anomeric carbon, leading to a partial double bond. This means that the endocyclic oxygen, anomeric carbon and all atoms attached to them must be in the same plane due to the sp^2 hybridization. For a pyranose, the conformations that allow this are 4H_3 and 3H_4 half chairs, the related 4E and E_4 envelopes and ${}^{2,5}B$ and $B_{2,5}$ boats,

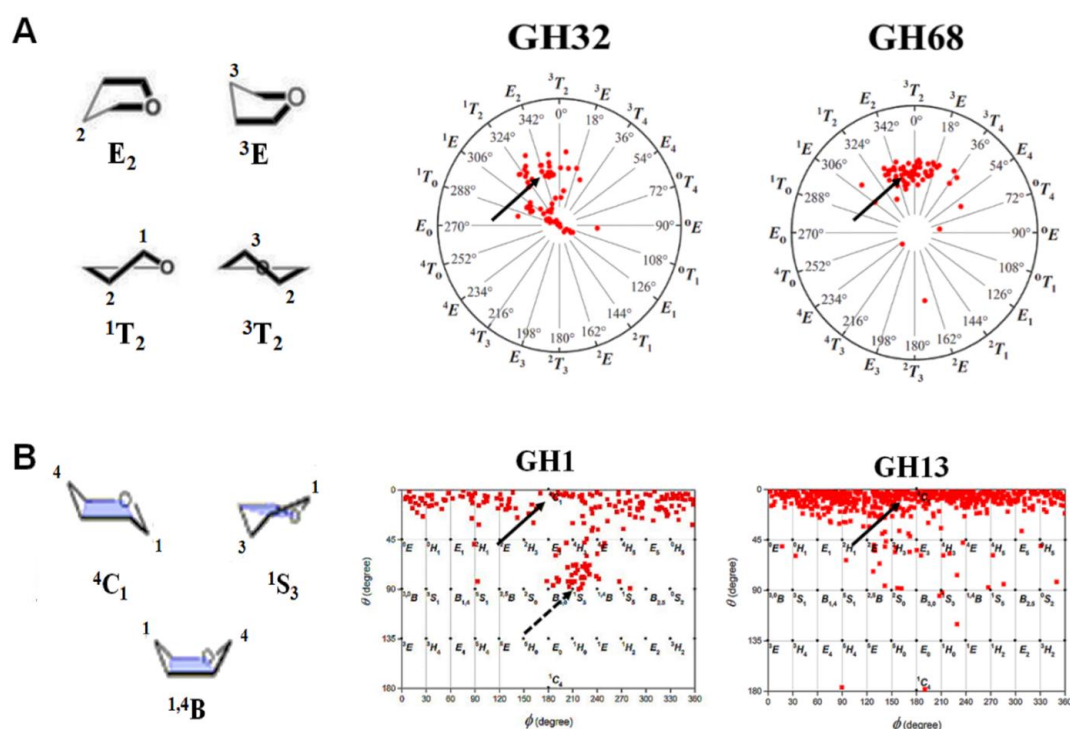


Figure 5. Plots of (A) Altona-Sundaralingam and (B) Cremer-Pople maps for the furanose and pyranose conformations at the -1 sugar ring of GHs with the highest population. Sugar conformations corresponding to the primary (solid arrows) and secondary (dashed arrows) populations are also included

with different monosaccharides favoring different shapes. Many glycoside hydrolases distort the subsite -1 sugar residue toward this transition state upon binding, leading to, for instance, 1S_3 conformations on the way to 4H_3 . However, monosaccharide products and covalent intermediates are generally in relaxed chair conformations, and some families that show late distortion of the sugar toward the transition state may also bind their substrates in relaxed chairs (Pengthaisong *et al.*, 2023). These results highlight the structural flexibility of both sugar types, which adopt dominant yet variable conformations at the -1 subsite. This also means that individual enzymes can subtly adjust sugar shape or position to suit their specific catalytic needs (Morais *et al.*, 2023). The predominance of these conformations indicates their favorable role in binding and catalysis.

Conclusions

We report a structural survey of carbohydrate ligands bound at GH active sites based on CAZyme-carbohydrate structures from 184 GH families in the PDB. Using Altona-Sundaralingam and Cremer-Pople parameter calculators, we systematically characterized furanose and pyranose conformations at the -1 subsite, revealing features not captured in the CAZy database. Furanose ligands predominantly adopt ${}^1T_2/E_2/{}^3T_2$ conformations in GH32 and $E_2/{}^3T_2/{}^3E$ in GH68, while pyranose ligands favor the 4C_1 chair conformation with distortions toward skew-boat or boat forms in enzyme-substrate complexes. These results define conformational preferences relevant to sugar binding and catalysis and offer a valuable reference for predicting catalytic itineraries and transition-state distortions in GH enzymes. We also identified catalytic acid/base and nucleophile residues across GH families, noting that in substrate-complex structures these are often replaced with nonreactive amino acids. Analysis of ligand size and residue substitution patterns reveals family-specific catalytic strategies, with glutamic acid dominating as the nucleophile for both furanose and pyranose ligands.

Overall, our analysis will support future CAZyme studies, particularly in conserved sequence analysis, mechanistic interpretation, and catalytic itinerary assignment in GHs. The dataset also offers direct utility for enzyme engineering and structure-based prediction, as statistically derived relationships between ring conformation, ligand type, and catalytic residue identity can be

incorporated as mechanistic constraints or training features in computational models for predicting enzyme function, substrate specificity, and mutational effects.

Funding

This work was supported by the Thailand Science Research and Innovation Fund and the University of Phayao under the Fundamental Fund 2026 (Grant No. 2262).

Author Contributions

W.J. and J.J. conceived and designed the study and analysed the data. W.J., P.T., S.W. and P.N. acquired the data. J.J. and J.R.K.C. supervised the study. W.J. and J.J. wrote the manuscript with support from J.R.K.C., and all authors approved the final version.

Ethics Statement

Ethical approval was not required for this study as it utilized previously published, publicly available datasets and did not involve human participants or animal subjects.

Declaration of Generative AI Use

The authors declare that no Generative AI tools were used in the preparation of this manuscript.

Conflict of Interest

The authors declare that they have no conflicts of interest.

Data Availability Statement

The datasets generated during and/or analyzed during the current study are available upon reasonable request.

Acknowledgements

The authors thank Pawanpat Vijitprapasin, Pranida Chucham, Benyapa Kansutham, and Pranpreeya Uparakul for their assistance with data collection.

References

- Alméciga-Díaz, C. J., Gutierrez, Á. M., Bahamon, I., Rodríguez, A., Rodríguez, M. A., & Sánchez, O. F. (2011). Computational analysis of the fructosyltransferase enzymes in plants, fungi and bacteria. *Gene*, 484(1–2), 26–34. <https://doi.org/10.1016/j.gene.2011.05.024>
- Altona, C., & Sundaralingam, M. (1972). Conformational analysis of the sugar ring in nucleosides and nucleotides: A new description using the concept of pseudorotation. *Journal of the American Chemical Society*, 94(23), 8205–8212. <https://doi.org/10.1021/ja00778a043>
- Ardévol, A., & Rovira, C. (2015). Reaction mechanisms in carbohydrate-active enzymes: Glycoside hydrolases and glycosyltransferases. Insights from ab initio quantum mechanics/molecular mechanics dynamic simulations. *Journal of the American Chemical Society*, 137(24), 7528–7547. <https://doi.org/10.1021/jacs.5b01156>
- Bains, R. K., Nasser, S. A., Wardman, J. F., & Withers, S. G. (2024). Advances in the understanding and exploitation of carbohydrate-active enzymes. *Current Opinion in Chemical Biology*, 80, Article 102457. <https://doi.org/10.1016/j.cbpa.2024.102457>
- Berman, H. M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T. N., Weissig, H., Shindyalov, I. N., & Bourne, P. E. (2000). The Protein Data Bank. *Nucleic Acids Research*, 28(1), 235–242. <https://doi.org/10.1093/nar/28.1.235>
- Cantarel, B. L., Coutinho, P. M., Rancurel, C., Bernard, T., Lombard, V., & Henrissat, B. (2009). The Carbohydrate-Active enZymes database (CAZy): An expert resource for glycogenomics. *Nucleic Acids Research*, 37(Suppl. 1), D233–D238. <https://doi.org/10.1093/nar/gkn663>
- Davies, G., & Henrissat, B. (1995). Structures and mechanisms of glycosyl hydrolases. *Structure*, 3(9), 853–859. [https://doi.org/10.1016/S0969-2126\(01\)00220-9](https://doi.org/10.1016/S0969-2126(01)00220-9)
- Davies, G. J., Gloster, T. M., & Henrissat, B. (2005). Recent structural insights into the expanding world of carbohydrate-active enzymes. *Current Opinion in Structural Biology*, 15(6), 637–645. <https://doi.org/10.1016/j.sbi.2005.10.008>
- Humphrey, W., Dalke, A., & Schulten, K. (1996). VMD: Visual molecular dynamics. *Journal of Molecular Graphics*, 14(1), 33–38. [https://doi.org/10.1016/0263-7855\(96\)00018-5](https://doi.org/10.1016/0263-7855(96)00018-5)
- Jitonnorn, W., Wanjai, T., Thinkumrob, N., Friedman, R., & Jitonnorn, J. (2023). Multiscale QM/MM simulations identify the roles of Asp239 and 1-OH···nucleophile in transition state stabilization in *Arabidopsis thaliana* cell-wall invertase 1. *Journal of Chemical Information and Modeling*, 63(15), 4827–4838. <https://doi.org/10.1021/acs.jcim.3c00796>
- Lammens, W., Le Roy, K., Schroeve, L., Van Laere, A., Rabijns, A., & Van den Ende, W. (2009). Structural insights into glycoside hydrolase family 32 and 68 enzymes: Functional implications. *Journal of Experimental Botany*, 60(3), 727–740. <https://doi.org/10.1093/jxb/ern333>
- Lombard, V., Ramulu, H. G., Drula, E., Coutinho, P. M., & Henrissat, B. (2014). The carbohydrate-active enzymes database (CAZy) in 2013. *Nucleic Acids Research*, 42(D1), D490–D495. <https://doi.org/10.1093/nar/gkt1178>
- Morais, M. A. B., Nin-Hill, A., & Rovira, C. (2023). Glycosidase mechanisms: Sugar conformations and reactivity in endo- and exo-acting enzymes. *Current Opinion in Chemical Biology*, 74, Article 102282. <https://doi.org/10.1016/j.cbpa.2023.102282>
- Pengthaisong, S., Piniello, B., Davies, G. J., Rovira, C., & Ketudat Cairns, J. R. (2023). Reaction mechanism of glycoside hydrolase family 116 utilizes perpendicular protonation. *ACS Catalysis*, 13(9), 5850–5863. <https://doi.org/10.1021/acscatal.3c00620>
- Porter, C. T., Bartlett, G. J., & Thornton, J. M. (2004). The Catalytic Site Atlas: A resource of catalytic sites and residues identified in enzymes using structural data. *Nucleic Acids Research*, 32(Suppl. 1), D129–D133. <https://doi.org/10.1093/nar/gkh028>
- Schrödinger, LLC. (2015). *The PyMOL molecular graphics system* (Version 1.8) [Computer software].
- Sinnott, M. L. (2007). *Carbohydrate chemistry and biochemistry: Structure and mechanism*. Royal Society of Chemistry. <https://doi.org/10.1039/9781782626329>
- Taha, H. A., Richards, M. R., & Lowary, T. L. (2013). Conformational analysis of furanoside-containing mono- and oligosaccharides. *Chemical Reviews*, 113(3), 1851–1876. <https://doi.org/10.1021/cr300249c>
- The UniProt Consortium. (2023). UniProt: The Universal Protein Knowledgebase in 2023. *Nucleic Acids Research*, 51(D1), D523–D531. <https://doi.org/10.1093/nar/gkac1052>
- Wardman, J. F., & Withers, S. G. (2024). Carbohydrate-active enzyme (CAZyme) discovery and engineering via (ultra)high-throughput screening. *RSC Chemical Biology*, 5, 595–616. <https://doi.org/10.1039/D4CB00024B>
- Yin, Y., Mao, X., Yang, J., Chen, X., Mao, F., & Xu, Y. (2012). dbCAN: A web resource for automated carbohydrate-active enzyme annotation. *Nucleic Acids Research*, 40(W1), W445–W451. <https://doi.org/10.1093/nar/gks479>