Modularity of the hydrophobic core contributes to **functional diversity** in fold A glycosyltransferases (GT-A)

R01 GM130915
R35 GM139656

Modular evolution of the GT-A core with an ancestral phosphate binding cassette at the center

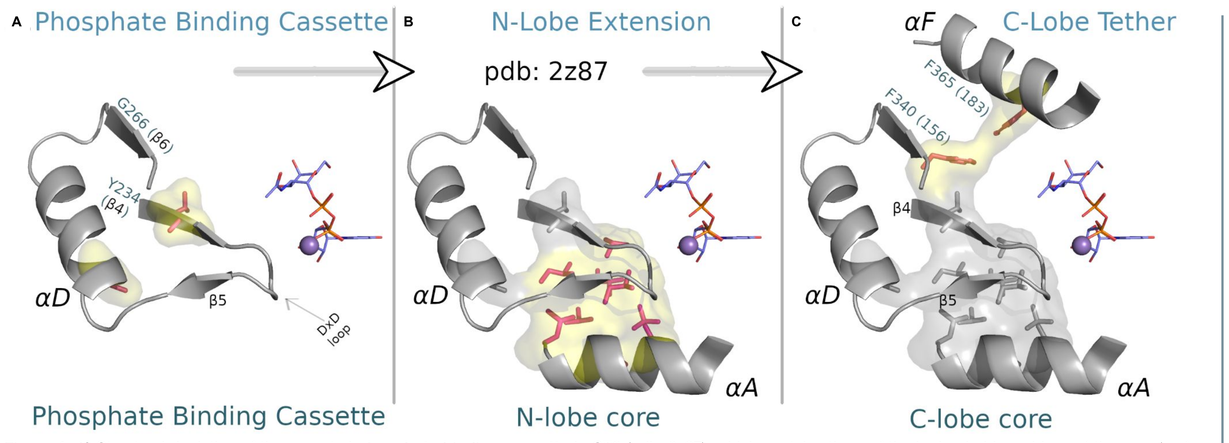


Figure 1: A) Structural depiction of the ancestral **phosphate binding cassette** in GT2 (pdb: 2z87), which contains three of the hydrophobic residues of the core (surface representation). **B-C)** Extension of the hydrophobic core from the phosphate binding cassette, showing the **insertion of an N-lobe core**, common to all Rossmann fold enzymes, and a **GT-A specific C-lobe tether** which connects the αF-helix to the phosphate binding cassette.

Mutational screens reveal B3GNT2-specific variations in the tether contribute to catalytic activity

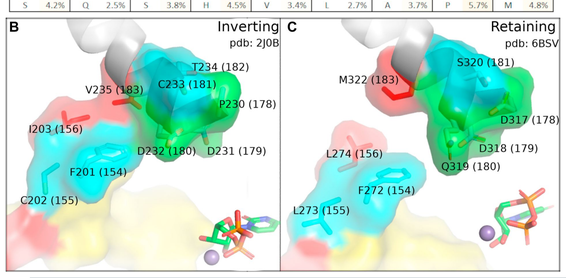
A

Inverting

154	%	155	%	156	%	178	%	179	%	180	%	181	%	182	%	183	%
G	20.5%	A	16.7%	L	15.8%	G	10.8%	E	23.4%	D	39.0%	D	25.7%	D	20.3%	L	20.1%
F	11.5%	L	14.3%	F	15.1%	A	10.4%	D	10.2%	E	18.5%	E	9.6%	E	15.0%	F	11.8%
S	7.3%	V	8.7%	V	12.5%	V	8.0%	F	7.3%	G	4.8%	V	8.2%	V	10.2%	D	7.3%
Y	7.3%	G	8.5%	I	10.2%	L	6.8%	P	6.1%	L	4.3%	M	7.5%	W	6.1%	I	7.1%
T	6.1%	I	7.7%	Y	8.2%	S	6.4%	S	5.7%	S	4.3%	L	5.7%	L	6.0%	E	6.9%
V	5.1%	Y	7.7%	M	6.3%	Y	6.3%	W	5.7%	R	3.6%	S	4.6%	F	5.5%	M	6.7%
A	4.9%	F	7.3%	A	6.0%	P	5.9%	G	5.0%	H	3.5%	A	4.5%	I	4.6%	V	6.2%
L	4.7%	M	6.3%	W	4.9%	F	5.4%	T	4.5%	Y	3.1%	T	4.3%	Y	4.3%	T	5.1%
C	4.5%	C	5.6%	S	4.7%	D	5.2%	A	4.2%	P	2.8%	I	4.1%	R	3.7%	R	4.7%
W	4.4%	S	5.4%	G	4.6%	T	5.0%	L	4.0%	I	2.6%	G	3.6%	T	3.7%	Y	4.4%

Retaining

154	%	155	%	156	%	178	%	179	%	180	%	181	%	182	%	183	%
L	14.0%	F	18.0%	V	15.3%	G	20.3%	D	18.5%	D	18.2%	D	19.5%	D	12.2%	L	15.0%
F	13.3%	L	16.3%	F	13.2%	L	7.9%	G	15.4%	E	17.5%	Q	12.5%	L	11.1%	I	10.5%
G	12.6%	V	13.8%	G	10.8%	D	7.2%	E	12.1%	Q	13.1%	N	8.8%	A	10.5%	E	8.8%
V	7.0%	G	9.9%	L	10.1%	A	6.9%	A	4.7%	A	7.4%	E	8.4%	E	9.1%	D	7.8%
Y	7.0%	I	8.1%	I	8.0%	I	5.5%	F	4.7%	G	5.7%	S	8.4%	Q	8.8%	V	7.1%
A	6.7%	T	6.7%	A	7.0%	V	5.2%	K	4.7%	T	5.1%	L	5.4%	G	7.8%	P	6.8%
I	5.6%	A	4.2%	Y	6.6%	N	4.8%	N	4.7%	N	4.4%	I	5.1%	V	7.4%	F	6.1%
M	5.3%	M	3.2%	M	6.3%	S	4.8%	I	4.4%	V	4.0%	P	5.1%	I	6.4%	G	5.8%
N	4.6%	Y	2.8%	T	5.6%	W	4.8%	L	4.0%	S	3.4%	G	4.7%	H	5.7%	A	5.4%
S	4.2%	Q	2.5%	S	3.8%	H	4.5%	V	3.4%	L	2.7%	A	3.7%	P	5.7%	M	4.8%



D

Comparison	Protein and Buffer Only	Protein with UDP-GlcNAc and MnCl2
Sample	Melt Temp (C)	Melt Temp (C)
WT	49.0	50.5
T336I	47.5	48.3
F309W	45.5	46.5
Y311I T336V	45.3	44.5
Y311F T336I	48.0	47.8
Y311F T336Y	44.5	43.5
Y311F T336V	47.0	47.3

wt			
Acceptor		Donor	
Km (mM)	2.9 ± 3.5	Km (mM)	0.35 ± 0.02
Kcat (min^{-1})	779 ± 48.1	Kcat(min^{-1})	540 ± 18.8

T336I			
Acceptor		Donor	
Km (mM)	5.8 ± 0.91	Km (mM)	0.3 ± 0.05
Kcat (min^{-1})	2072 ± 203	Kcat(min^{-1})	1138 ± 24

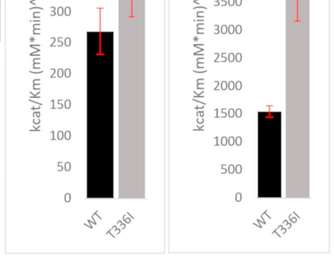


Figure 2: A) Array of top ten residue frequencies from a sequence alignment of inverting and retaining GTs, showing higher conservation and constraints in the C-lobe tether (156, 183) and xED-Asp (180) in inverting GT-As. **B-C)** A comparison of representative inverting and retaining GT-A core packing in the same orientation, showing that the retaining pocket is less packed, as compared with inverting GT-As. **D)** Table of parameters for acceptor and donor saturation for wt and mutant, including Km, kcat, and kinetic efficiency. Kinetic efficiency (Kcat/Km) of B3GNT2 wt vs. T336I upon acceptor saturation.

MD simulations reveal unique conformational motions of catalytic base modulated by the tether

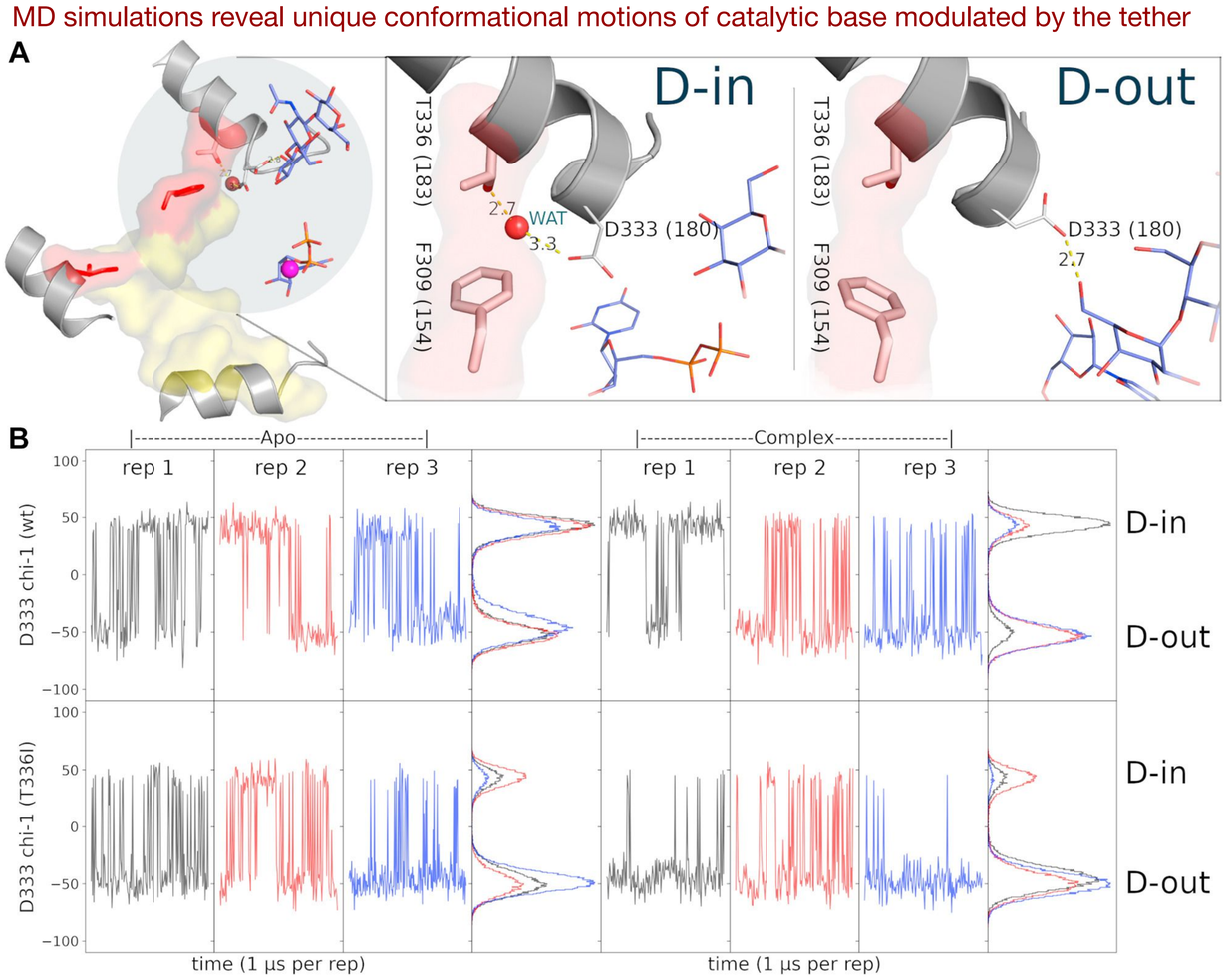


Figure 3: A) Stills from a molecular dynamics simulation of the B3GNT2 wild type complex (pdb: 6WMO), showing **two unique conformations of the xED-aspartate**. The **D-in and D-out** conformations are termed as such depending on their orientation inwards, interacting with the threonine aided through a hydrogen bond interaction with a water molecule, or outwards towards the acceptor-donor complex. **B) 12 MD simulations (three replicates, one microsecond each)** demonstrating the **shift in conformational occupancy of mutant T336I in the D-out conformation.** Replicates show the dynamic switching between the D-in and D-out conformations over the course of the simulation, with the histogram (located after replicate 3 of each set of simulations) showing the **total ratio of D-in:D-out for each replicate**.

DISCUSSION

- The canonical GT-A is **separable into three distinct modules** added over evolutionary time:
  - the PBC, the Rossmann N-lobe, and the GT-A specific C-lobe tether.
- Addition of the **C-lobe tether anchors the F-helix and xED** catalytic base **to the PBC.**
- Mutating to more canonical hydrophobic packing in the tether resulted in **increased catalytic efficiency by favoring the D-out conformation** of the xED-Asp. The D-out increases chance of catalyzing transfer.
- Multiple mutations showed good expression and retained thermostability, thus mutational screens of the hydrophobic core may be a feasible path for the **synthetic design of GTs**.

ABSTRACT

- Hydrophobic cores are fundamental to **protein folding and stability**. However, **how hydrophobic cores shape protein evolution and function is poorly understood**.
- We show that the **GT-A core evolved from an ancestral phosphate binding cassette** (PBC) present in diverse, unrelated, nucleotide phosphate binding proteins. GT-As have also diverged from other nucleotide binding proteins and Rossmann fold enzymes through the **unique tethering** of the PBC to the F-helix, which harbors the catalytic base (xED).
- We find that sequence and structural variation in this **tether is a major contributing factor in the evolution of GT-A fold** catalytic mechanisms and functional specialization.
- Mutational analyses and molecular dynamics simulations in B3GNT2 support a model in which **allosteric control of catalytic base flexibility through variations in the tether** may increase catalytic efficiency by **modulating the conformational occupancy** of the catalytic base between the "D-in" and the **acceptor-accessible "D-out" conformation.**
- Our studies support a model of evolution in which the GT-A core **evolved progressively through elaboration upon an ancient PBC** found in diverse nucleotide-binding proteins, and **malleability of this core** provided the structural framework for **evolving new catalytic and substrate-binding functions in extant GT-A fold enzymes**.

Model of the evolutionary progression of fold A glycosyltransferases

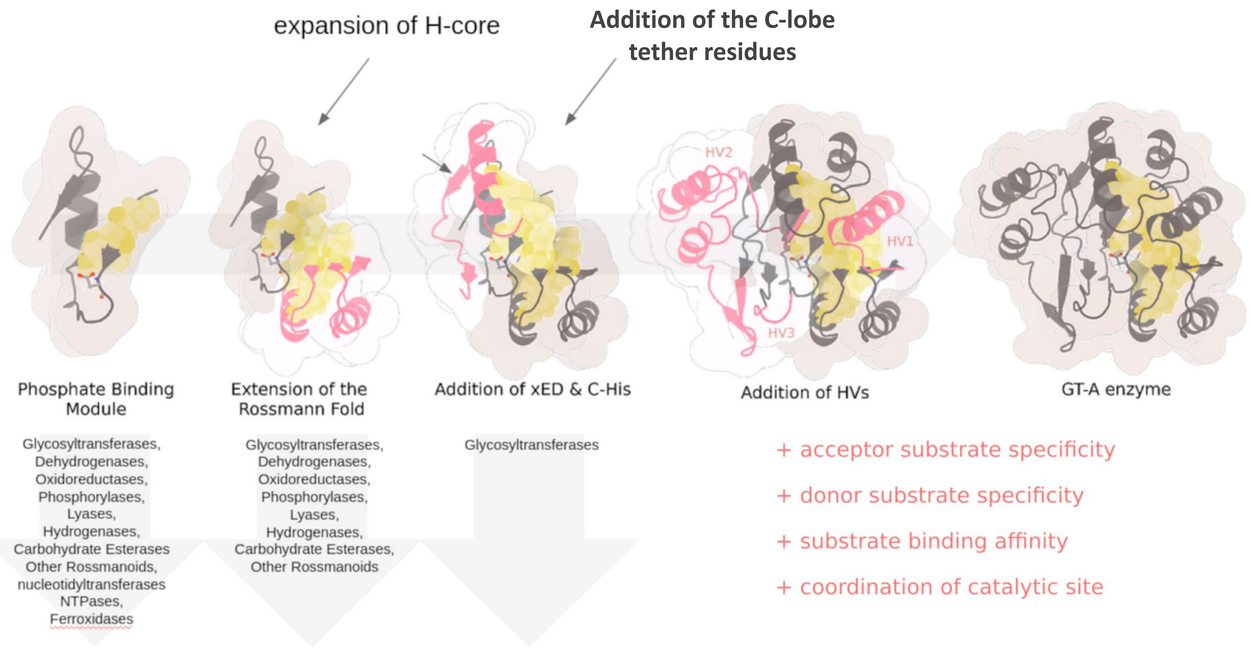
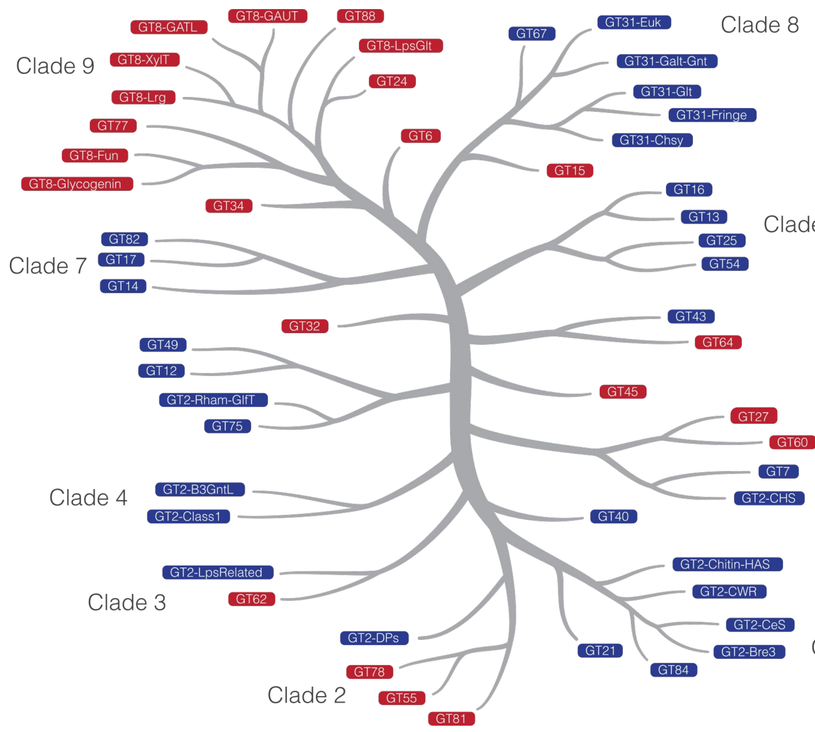


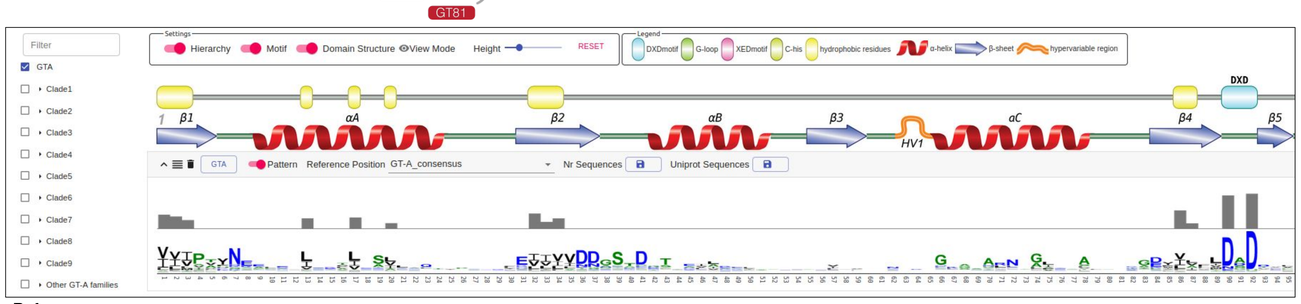
Figure 4: Beginning from the **elementary phosphate-binding cassette**, GT-As gained a Rossmann fold **that extended the hydrophobic core**. Following this, various GT-As make **use of the xED motif as a catalytic base**, the presence of this motif correlates with mechanistic variations. Finally, **family specific hypervariable regions are introduced to further regulate GT-A function**. New additions in pink.

A new tool to explore Glycosyltransferase structure, function, and evolution: **GTXplorer**



Scan for website!

Retaining Enzymes
Inverting Enzymes



References:
1) **Venkat, A.**, Tehrani, D., Taujale, R., Yeung, W., Gravel, N., Moremen, K.W. and Kannan, N., 2022. Modularity of the hydrophobic core and evolution of functional diversity in fold A glycosyltransferases. Journal of Biological Chemistry, 298(8).
2) Moremen, K. W., & Haltiwanger, R. S. (2019). Emerging structural insights into glycosyltransferase-mediated synthesis of glycans. Nature chemical biology, 15(9), 853-864.
3) Taujale, R., **Venkat, A.**, Huang, L. C., Zhou, Z., Yeung, W., Rasheed, K. M., ... & Kannan, N. (2020). Deep evolutionary analysis reveals the design principles of fold A glycosyltransferases. Elife, 9, e54532.
4) Taujale, R., Soleymani, S., Priyadarshi, A., **Venkat, A.**, Yeung, W., Kochut, K. J., & Kannan, N. (2021). GTXplorer: A portal to navigate and visualize the evolutionary information encoded in fold A glycosyltransferases. Glycobiology.

ARCS ADVANCING SCIENCE IN AMERICA

Department of Biochemistry and Molecular Biology
Franklin College of Arts and Sciences
UNIVERSITY OF GEORGIA

Recent paper →

Aarya Venkat[1], Daniel Tehrani[1,2], Rahil Taujale[3], Wayland Yeung[3], Nathan Gravel[4], Kelley Moremen[1,2], and Natarajan Kannan[1,3]

Department of Biochemistry and Molecular Biology[1], Complex Carbohydrate Research Center (CCRC)[2], Institute of Bioinformatics[3], Interdisciplinary Life Sciences[4], University of Georgia, Athens, GA, USA