Structural and functional analysis of four family 84 glycoside hydrolases from the opportunistic pathogen *Clostridium perfringens*

Benjamin Pluvinage[a], Patricia M. Massel[a], Kristyn Burak[a], Alisdair B. Boraston[a]

Supplementary Data

Table SI: *Cp*GH84 catalytic modules sequence identity

	CpGH84A	CpGH84B	CpGH84C	CpGH84D
CpGH84A		26	24	44
CpGH84B	26		58	34
CpGH84C	24	58		31
CpGH84D	44	34	31	

Table SII: Primer sequences used to clone *Cp*GH84s catalytic modules (*NheI* and *XhoI* restriction sites are underlined).

Primer	Sequence
GH84A Forward	GAT CAA GCT AGC ACA GAT GGT ATT ACA G
GH84A Reverse	CAT ATG CTC GAG TTA ATC AAC AAT AGA GCC
GH84B Forward	GAT CAA GCT AGC GAT GAG GGA TTA AAA AAT
GH84B Reverse	CAT ATG CTC GAG TTA TTC ATG GAC TTT TCT TAT
GH84D Forward	GAT CAA GCT AGC AAT GGC AGC AAA GAA ACA
GH84D Reverse	CAT ATG CTC GAG TTA ATT ATC TCC TAT GAT TTT

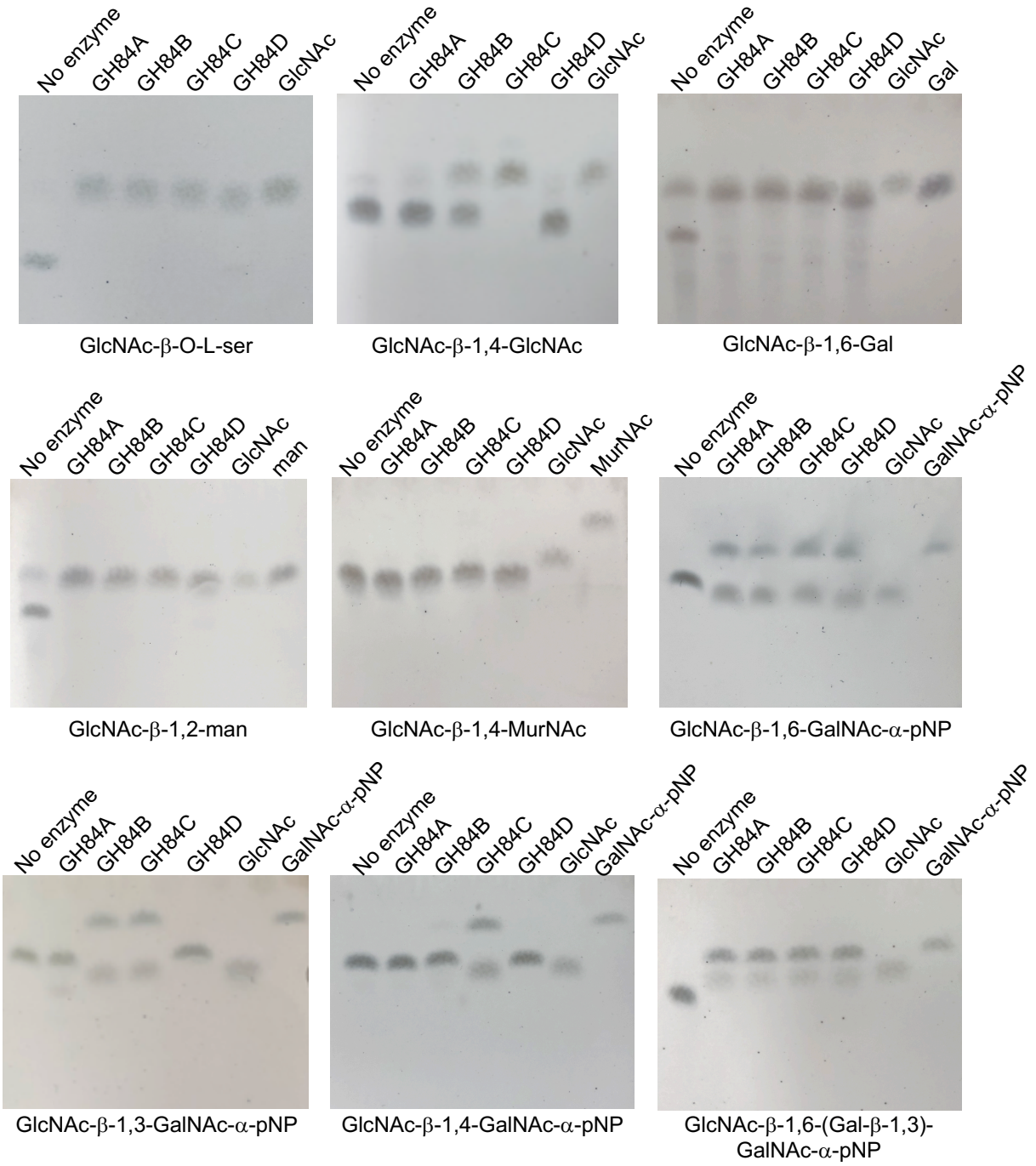

Figure S1: Thin layer chromatography (TLC) analysis of *Cp*GH84s substrate specificities. The substrate tested is indicated below each TLC. GlcNAc: N-acetyl-glucosamine, Gal: galactose, Man: mannose, MurNAc: N-acetylmuramic acid and GalNAc-α-pNP: *para*-nitrophenol-N-acetyl-galactosamine.

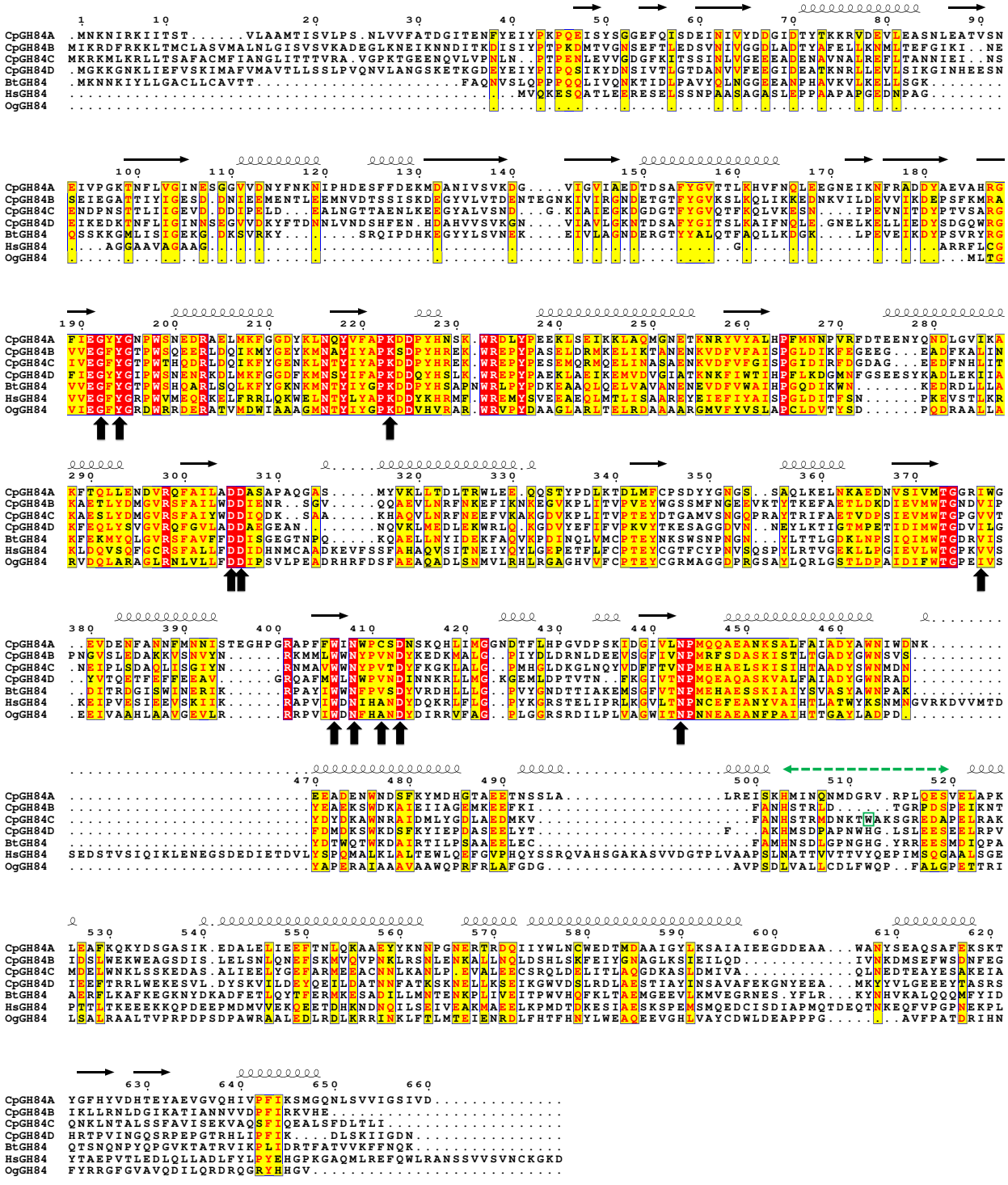

Figure S2: Sequence alignment of the catalytic modules of family 84 glycoside hydrolases. *Cp*GH84A, *Cp*GH84B, *Cp*GH84C and *Cp*GH84D catalytic modules from *C. perfringens* are aligned with the catalytic modules of *Bacteroides thetaiotaomicron* GH84 (*Bt*GH84), human O-glcNAcase (*Hs*GH84) and *Oceanicola granuluosus* GH84 (*Og*GH84). Red indicates identical residues. Yellow indicates strong similarity. Conserved residues forming the catalytic pocket and

catalytic residues are indicated by black arrows and stars, respectively. Numbering indicated corresponds to the amino acid sequence of *Cp*GH84A. The green dashed double arrow indicates the loop from *Cp*GH84C carrying W490 (highlighted by a green box) at its apex. Alignment was generated using ESPript - http://espript.ibcp.fr (Robert, X. and Gouet, P. 2014).

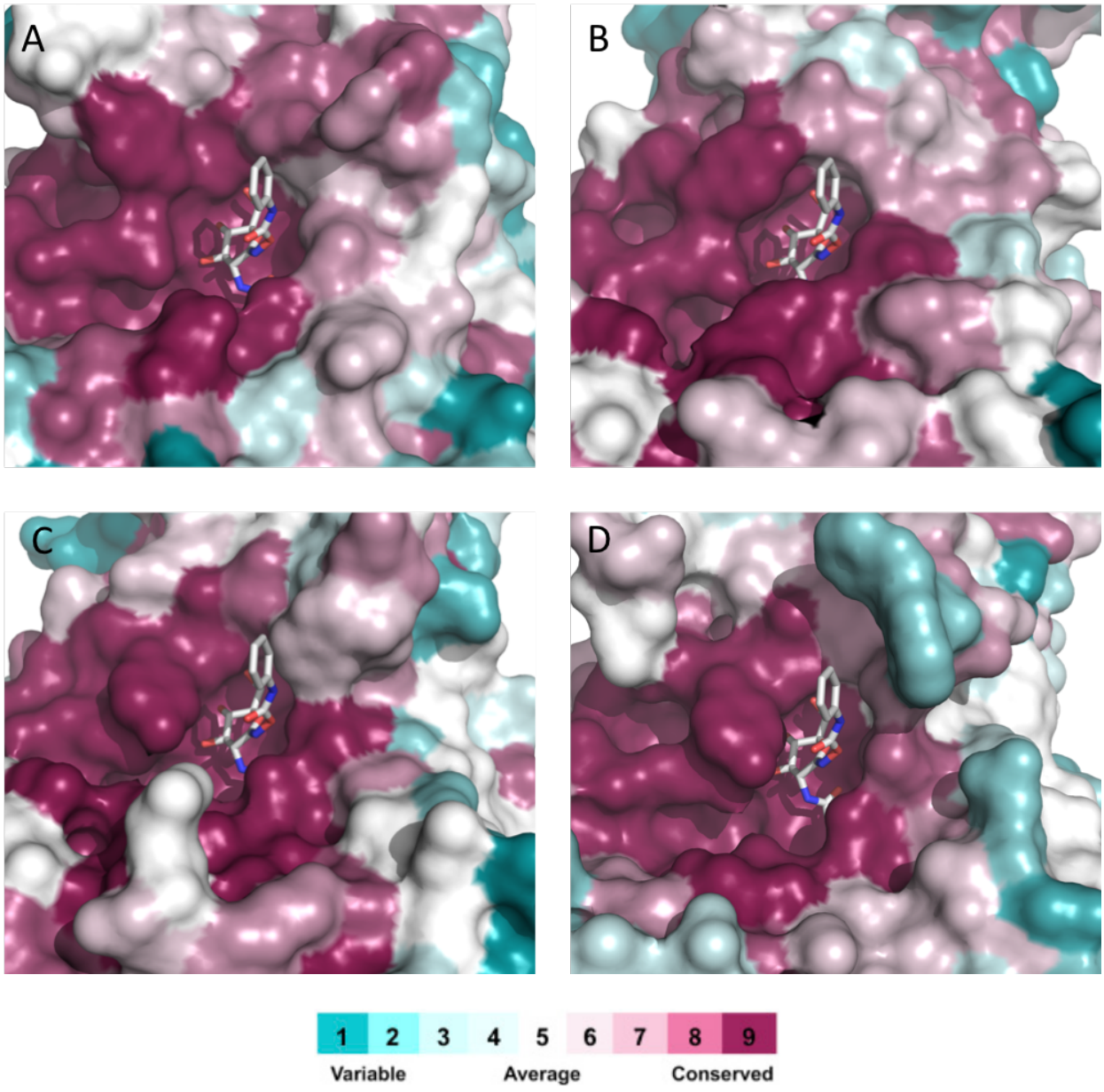

Figure S3: Surface representation of *Cp*GH84 catalytic sites. *Cp*GH84A (A), *Cp*GH84B (B), *Cp*GH84C (C) and *Cp*GH84D (D) catalytic sites are colored based on residue conservation using Consurf (Ashkenazy, H., Abadi, S., *et al.* 2016). All apo structures listed above were aligned with PugNAc (in grey sticks) from the *Cp*GH84C complex (PDB code 2CBJ).

References

Ashkenazy H, Abadi S, Martz E, Chay O, Mayrose I, Pupko T, Ben-Tal N. 2016. ConSurf 2016: an improved methodology to estimate and visualize evolutionary conservation in macromolecules. *Nucleic Acids Res*, 44:W344-350.
Robert X, Gouet P. 2014. Deciphering key features in protein structures with the new ENDscript server. *Nucleic Acids Res*, 42:W320-324.