



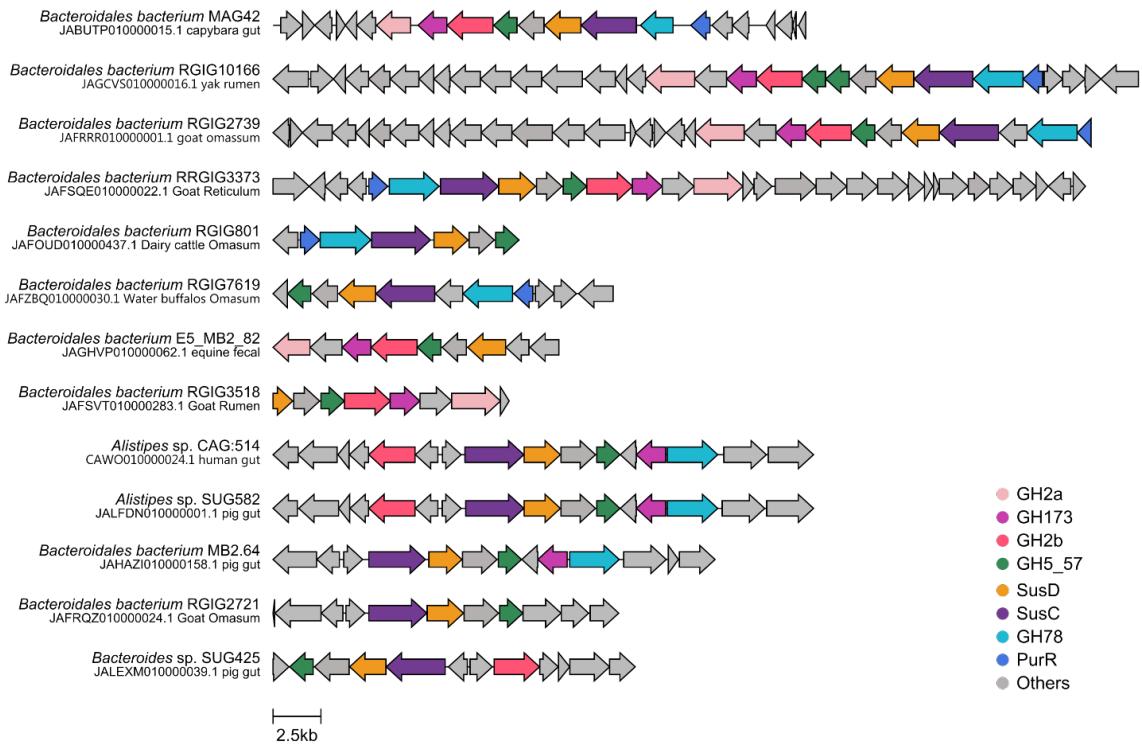
STRUCTURAL  
BIOLOGY

**Volume 78 (2022)**

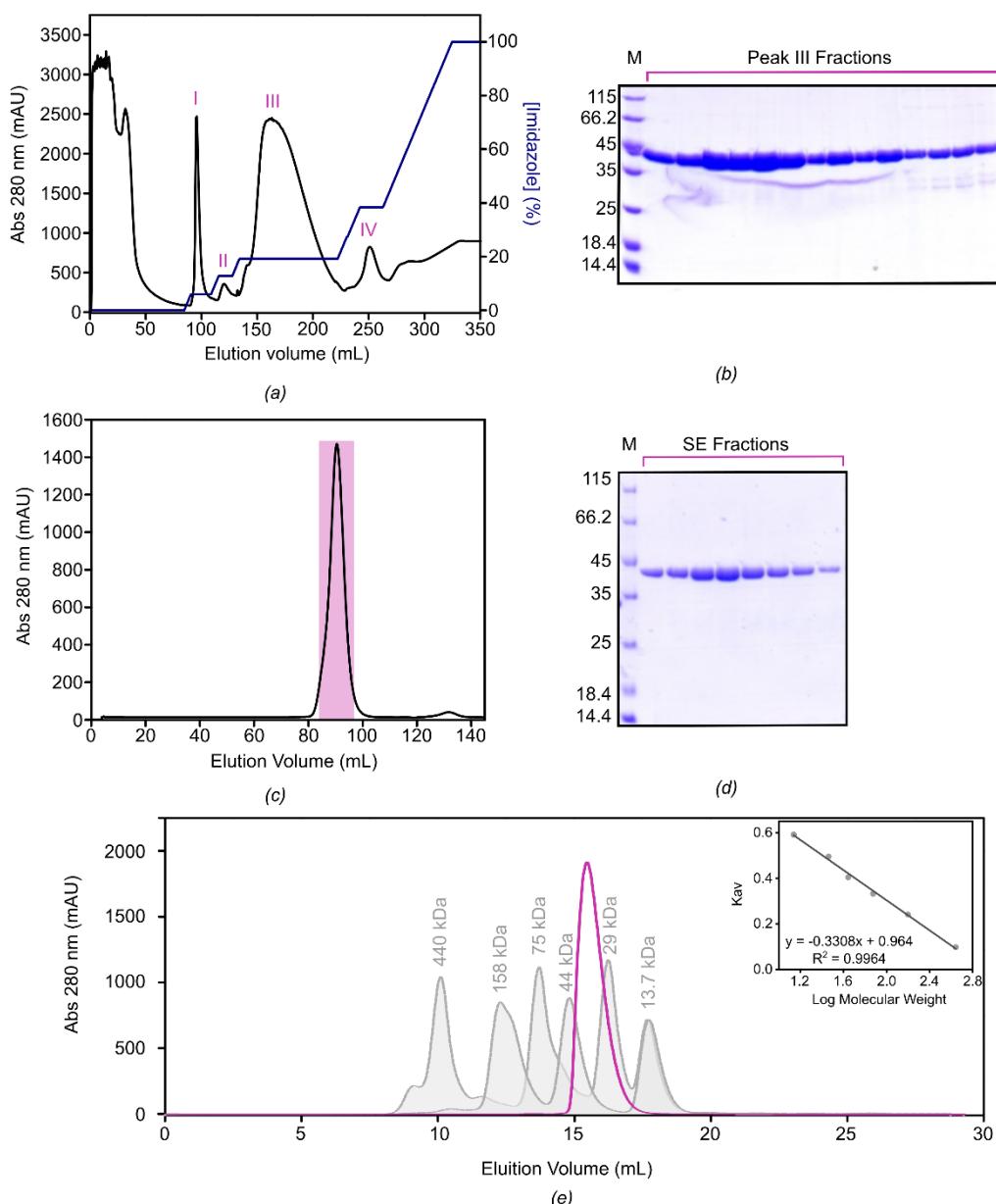
**Supporting information for article:**

**Glycoside hydrolase subfamily GH5\_57 features a highly redesigned catalytic interface to process complex hetero- $\beta$ -mannans**

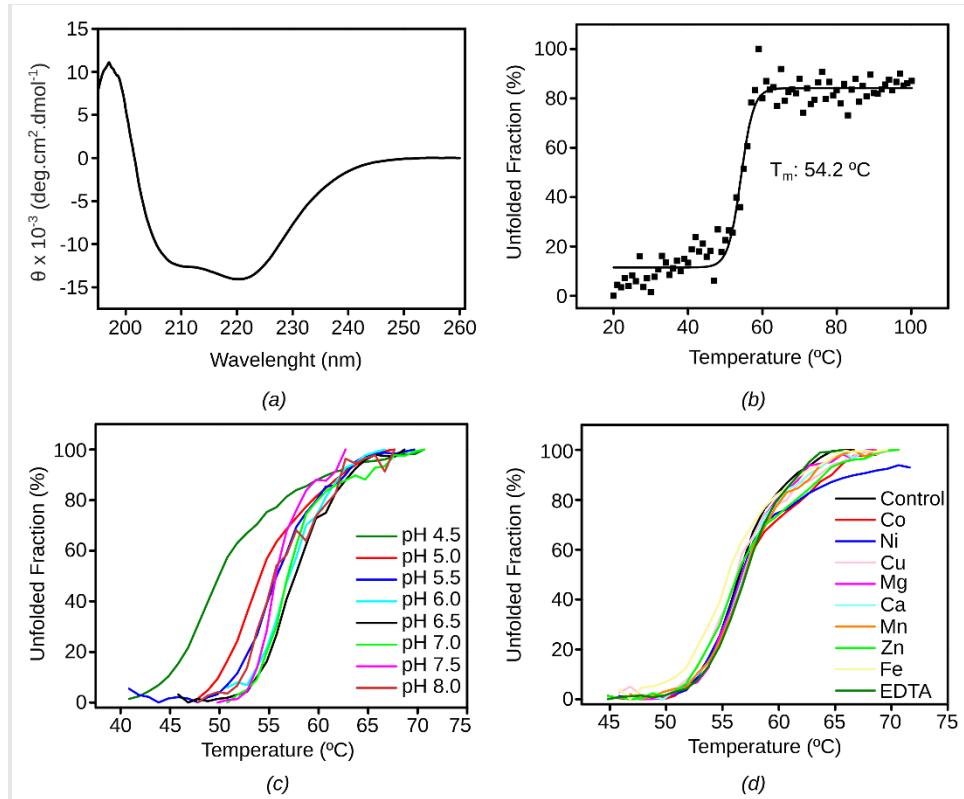
**Marcele P. Martins, Mariana A. B. Morais, Gabriela F. Persinoti, Rafael H. Galinari, Li Yu, Yoshihisa Yoshimi, Fernanda Passos Nunes, Tatiani B. Lima, Shayla F. Barbieri, Joana L. M. Silveira, Vincent Lombard, Nicolas Terrapon, Paul Dupree, Bernard Henrissat and Mário T. Murakami**



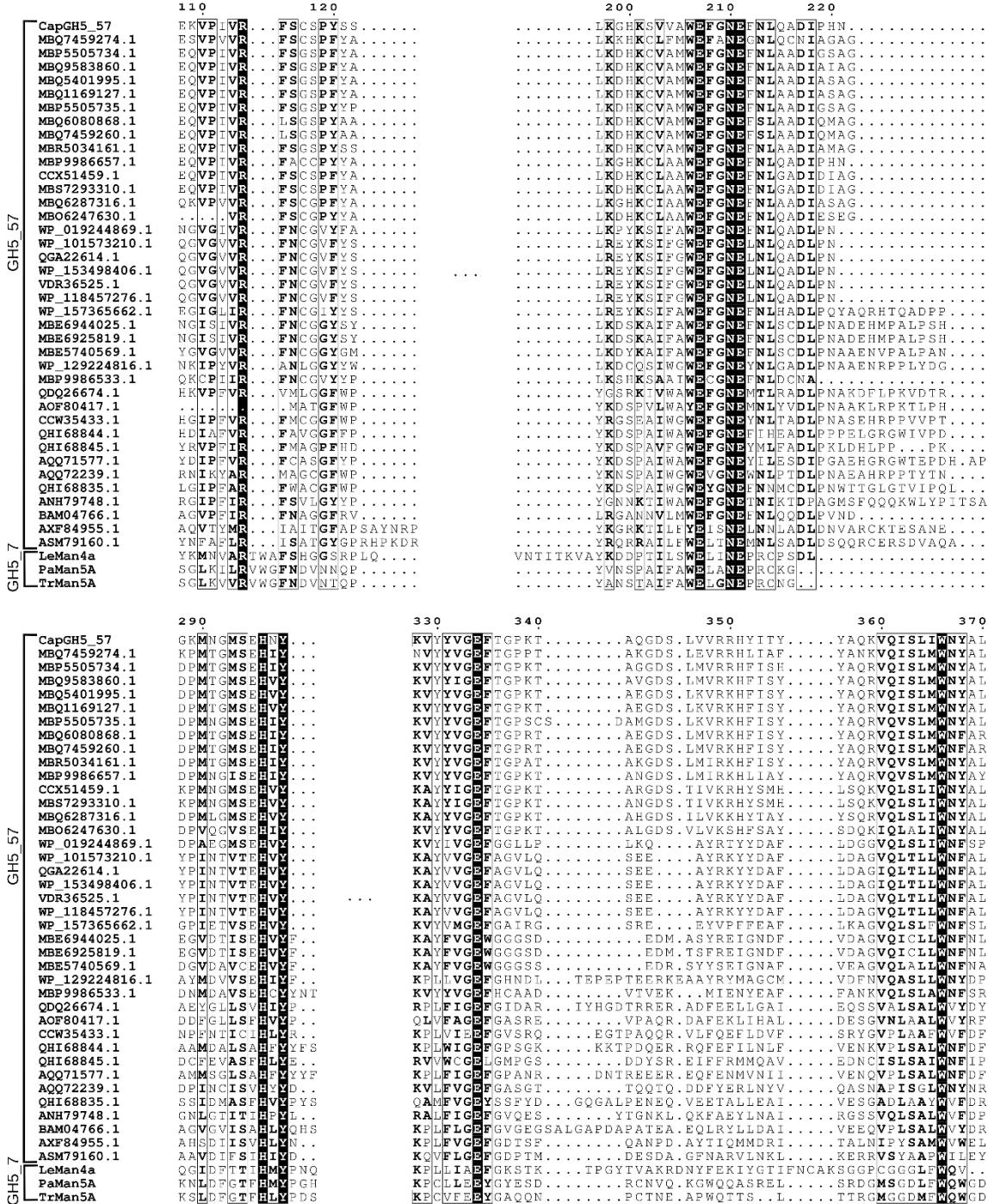
**Fig. S1:** *Bacteroidales bacterium* MAG42 genomic organization of heteromannan PUL and its conservation in other Bacteroidetes genomes.



**Fig. S2: CapGH5\_57 purification.** (a) Nickel-affinity chromatogram displaying four peaks (I, II, III and IV, highlighted in pink). (b) SDS-PAGE analysis of fractions from peak III eluted from affinity chromatography. (c) Size-exclusion (SE) chromatogram, highlighting the selected fractions for further experiments (pink rectangle). (d) SDS-PAGE analysis of fractions highlighted in (c) eluted from size-exclusion chromatography (SE fractions). Lane M, protein molecular weight markers (labelled in kDa). (e) Analytic size-exclusion chromatogram of CapGH5\_57 (purple line). Molecular weight standards peaks and their respective masses are represented in grey. The CapGH5\_57 molecular weight was calculated from the calibration curve using molecular weight standards (*inset*).



**Fig. S3: CapGH5\_57 biophysical characterization.** (a) Circular dichroism spectrum of CapGH5\_57. (b) CapGH5\_57 thermal denaturation profile monitored at 222 nm by circular dichroism spectroscopy. The melting temperature ( $T_m$ ) was calculated from the sigmoidal fit of the denaturation curve. (c) CapGH5\_57 differential scanning fluorimetry (DSF) profile with different pH range and (d) ions and EDTA.



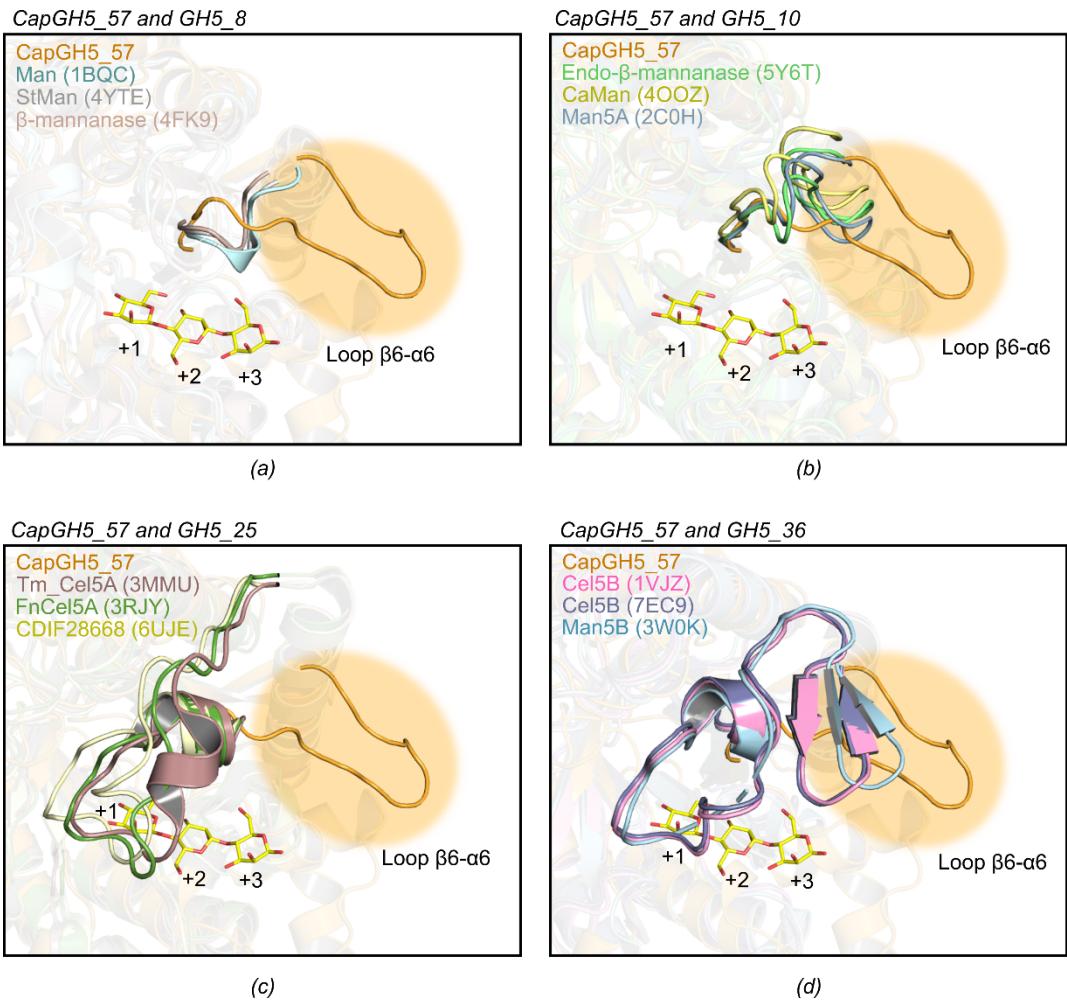
**Fig. S4: Multiple sequence alignment of GH5\_57 subfamily members and GH5\_7 members LeMan4a, PaMan5A and TrMan5A, highlighting the conserved residues.**

Fully conserved residues of GH5 family are shaded in black (Arg114, Asn210, Glu211, His296, Tyr298, Glu334, and Trp366).

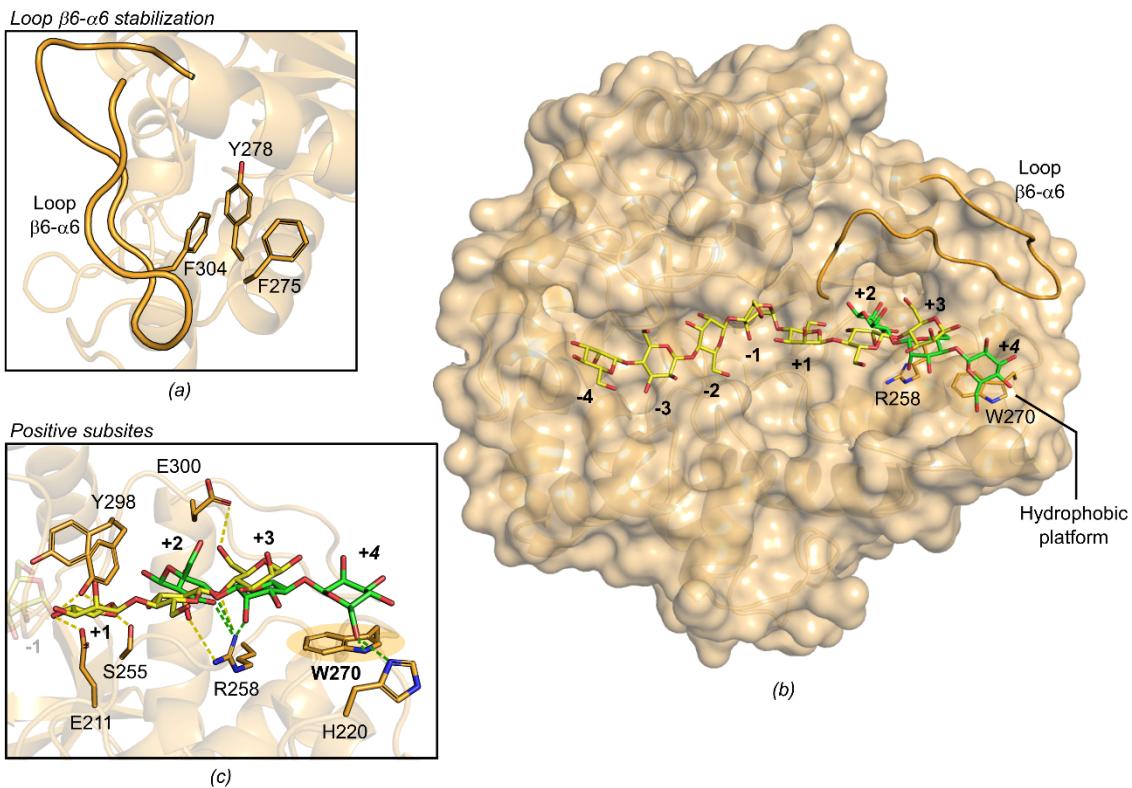
GH5\_57

	270	280	290	300	310	320
CapGH5_57	GSWT.....	A	DTFQQYIDITSIPTP	GKMNCGMSEHVV	FEARTFSD	LGKLGRSEQVIAK
MBQ7459274.1	NSWL.....	D	DSYAQYVEITDLMTP	KPMTCMSEHVV	DDSRVFSD	KGQVNRNDQVAYAKQ
MBP5505734.1	PSWD.....	N	DSYEQYKEVIGVFTP	DPMTCMSEHVV	EDARVFSD	KGTVGRSDQILYAKQ
MBQ9583860.1	PSWD.....	T	DSYAQYKEVTEVFTP	DPMTCMSEHVV	EDARVFSD	KGTVGRSDQILYAKQ
MBQ5401995.1	PSWD.....	N	DSYAQYKEVTEVFNP	DPMTCMSEHVV	EDARVFSD	KGTINRSDQISYAKQ
MBQ1169127.1	PSWD.....	N	DSYAQYKEVTEVFNP	DPMTCMSEHVV	EDARVFSD	KGTINRSDQISYAKQ
MBP5505735.1	PSWD.....	K	DSYEQYVEVSGIFNP	DPMNCMSEHVV	EEAREFSD	KGKVGRSDQISYAKQ
MBQ6080868.1	ASWT.....	T	DSYAEEYVEMCGILNP	DPMTCMSEHVV	EDARVFSD	KGTVNRSEQIAYAKQ
MBQ7459260.1	ASWT.....	T	DSYAEEYVEMCGILNP	DPMTCMSEHVV	EDARVFSD	KGTVNRSEQIAYAKQ
MBR5034161.1	ASWG.....	T	DSPDEYVEACGIFNP	DPMTCMSEHVV	EDARVFSD	KGTVGRSEQISYAKQ
MBP9986657.1	KSWT.....	N	DTPDEYKQITKIEFP	DPMNCMSEHVV	EDKRVFSVDVNNGSELSDRIDQMSYAKQ	
CCX51459.1	KSWA.....	T	DSEFKQYVEITGVMTP	KPMNCMSEHVV	DEPRVFSD	LGEINLNSYQVAKAKE
MBS7293310.1	KSWA.....	T	DSEFRQYVEITGVMTP	KPMNCMSEHVV	DEPRVFSD	LGEINLNSYQVAKAKE
MBQ6287316.1	GSWA.....	D	DSEFRQYVEITGVMTP	DPMLCMSEHVV	EESRKFSD	LGTVNRTYQLIHAKE
MBQ6247630.1	NGWT.....	D	DTPGQYVSITGLLPD	DPVQCMSEHVV	EERRKFSD	IGELDRENQGLEAKE
WP_019244869.1	YTFS.....	V	DTPQGYRAATELFPTP	DPAECMSEHVV	EACREFAD	RGKVTLSEQIAYAVAMG
WP_101573210.1	NSWD.....	T	DTCQYRAITNLNP	YPINIVTEHVV	ETGRKFSD	CGKVSLDRQIREAMY
OGA22614.1	NSWD.....	T	DTCQYRAITNLNP	YPINIVTEHVV	ETGRKFSD	CGKVSLDRQIREAMY
WP_153498406.1	NSWD.....	S	DTCQYRAITNLNP	YPINIVTEHVV	ETGRKFSD	CGKVSLDRQIREAMY
VDR36525.1	NSWD.....	T	DTCQYRAITNLNP	YPINIVTEHVV	ETGRKFSD	CGKVSLDRQIREAMY
WP_118457276.1	NSWD.....	T	DTCQYRAITNLNP	YPINIVTEHVV	ETGRKFSD	CGKVSLDRQIREAMY
WP_157365662.1	DRMT.....	I	DTKQYRAKISRLNP	GPIETVSEHVV	QHGRQFAD	LGKVSLDQEIAIAVE
MBE6944025.1	NSWN.....	Q	DTYEEYKUATAMFTP	DPMVDTISEHVV	QSQRTFG	EISLSAYLSYMMQ
MBE6925819.1	NSWN.....	Q	DTYEEYKUATAMFTP	DPMVDTISEHVV	QSQRTFG	DLSLSQYLSYVMQ
MBE5740569.1	NTWA.....	Q	DSVLEYITGTTP	DGVDAVCEHVV	TSQRTFG	ELTLAEYLQYATA
WP_129224816.1	GAWI.....	Q	DTEEEHKVTERFEP	AYMDVTVSEHVV	RKENYFGR	EMTLEEVLYAYTTE
MBP9986533.1	NHAN.....	H	DSYAQYLEDISNLNP	DNMDAVSEHQYNT	DTPDHINYPNPRAFSD	REIPILVLSEARACCKQ
QDQ26674.1	SMWE.....	S	DTEQFCSSLQRDNP	AEGYGLLSVHTP	GENEGYFG	KKLPDYPNLKATER
AOF80417.1	RSWK.....	S	DTCGEFRTRMLGRHNP	DDFGTFLSFHVP	HHRHQYFG	KASFQGIDLDAIV
CCW35433.1	HTWQ.....	L	DTRAQFBETILRDNP	NPPNTICIHVP	DSKNIYPG	GAKNLGEVVALANO
QHI68844.1	GAWG.....	I	DTRAEEWEIQLQDN	AAMDLSAHVYFSL	DEENHDGG	FMEYGPQQPFMME
QHI68845.1	DCWG.....	H	DTRAQWLEQFASDTP	DCDVAVSFHVE	EAQGSYFK	GEKEPLGEVVAAVQ
AQQ71577.1	NNWE.....	K	DTRCWRSLYVKDN	AMMSGLSAHVYFK	DDRHIHGG	VQEFGPAQAFMME
AQQT2239.1	SSWE.....	T	DTEQKRIIMFKLQNP	DPIINCISVHVD		TARDIEELAK
QHI68835.1	GSWT.....	A	DSYSQAKEPHSWMP	TSSSIDNASFHVPYSM	EDSNGDID	YAGASGVSDILQIYRE
ANH7948.1	GIWG.....	A	DTKSALETNMLWQN	GNLGTITIHYL	NNEGAASY	NTSIGNIUTGAKA
BAM04766.1	GSFR.....	P	DTRDFRENTITDHA	AGVGVI		
AXF84955.1	PEWL.....	D	DSAAQFKKNLLEDVN	AHSIDISVHVV	NINNV	RFGQRNVTDLLTLTKO
ASM79160.1	PEWKBPQAGRSPQ	D	DNEEQFAQYKLKDWH	AVDIFTSHIYD	GHNRSQLN	RLGDKEL.DILNRVQN
LeMan4a			NSYIFGTINFISNNQV	QGIDETTIDHPPNQ	WLPGLTQEAQDKWAS	QWIQVHID
PaMan5A			YQYGEGTDFVKNLQI	KNLDEGTFHVGEGH	GVFTSFGP	GWIKDHAAS
TrMan5A			YQYGEGTDFVKNLQI	KSLDFGTFFHVV	GTNYTWN	GWIQTHAA
				PDS		

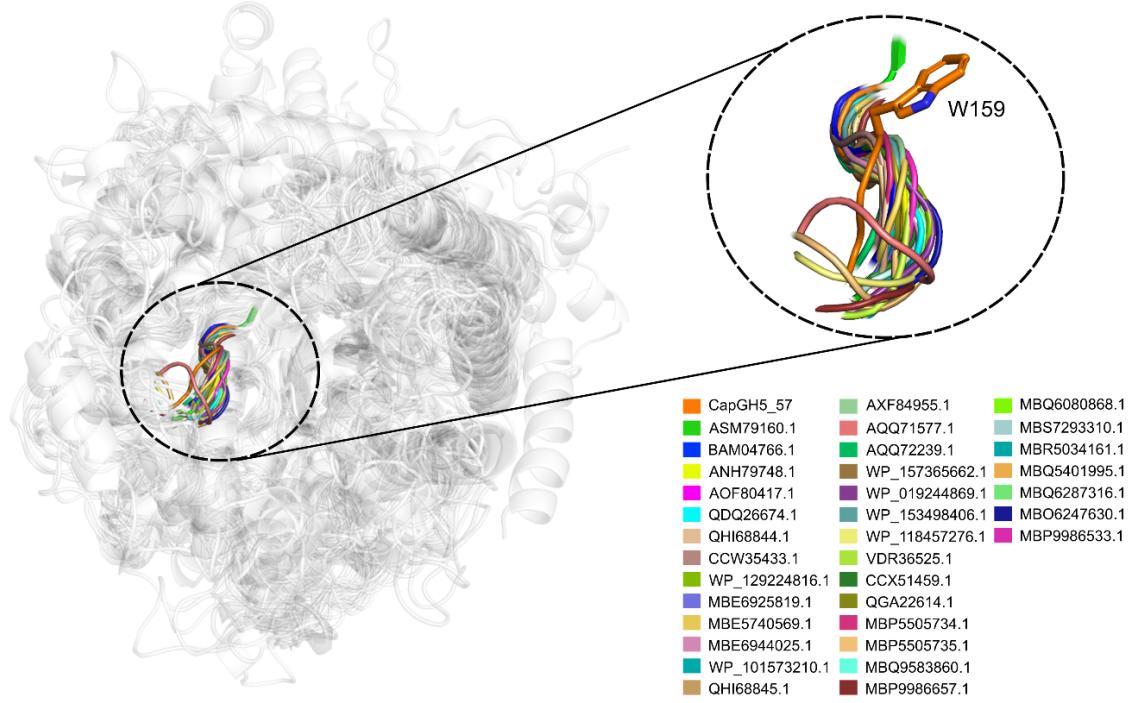
**Fig. S5: Multiple sequence alignment of GH5\_57 subfamily members and GH5\_7 members LeMan4a, PaMan5A and TrMan5A, highlighting the divergent  $\beta$ 6- $\alpha$ 6 region.** Identical residues are shaded in black. The region encompassing the divergent helix ( $\alpha$ 6<sub>a</sub>), within the  $\beta$ 6- $\alpha$ 6 loop (Glu299-Gly311) is highlighted in light blue.



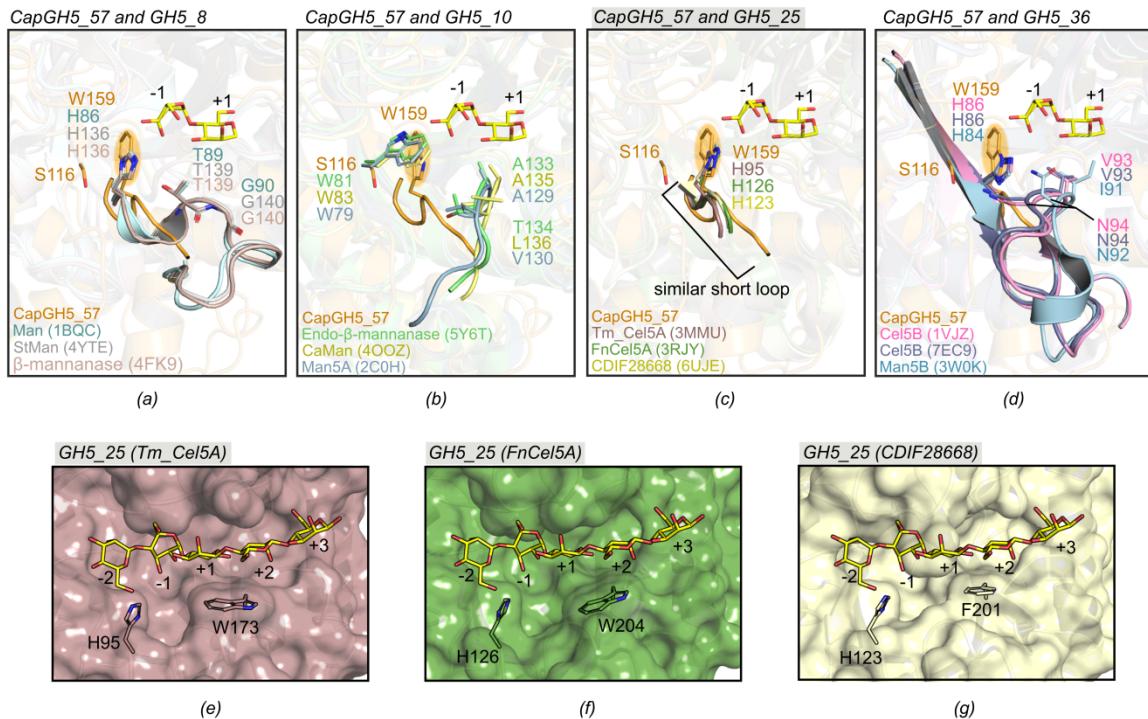
**Fig. S6: Structural comparison of CapGH5\_57  $\beta$ 6- $\alpha$ 6 loop with GH5 subfamilies displaying endo- $\beta$ -mannanase activity.** (a) Structural superimposition of CapGH5\_57 and GH5\_8 (Man PDB code 1BQC (Hilge *et al.*, 1998), StE273Adc PDB code 4YTE (Kumagai *et al.*, 2015) and  $\beta$ -mannanase PDB code 4FK9 (Takasuka *et al.*, 2014)), (b) Structural superimposition of CapGH5\_57 and GH5\_10 (Endo- $\beta$ -mannanase PDB code 5Y6T (Ueda *et al.*, 2018), CaMan PDB code 4OOZ (Kim *et al.*, 2014), and Man5A PDB code 2C0H (Larsson *et al.*, 2006)), (c) Structural superimposition of CapGH5\_57 and GH5\_25 (Tm\_Cel5A PDB code 3MMU (Pereira *et al.*, 2010), FnCel5A PDB code 3RJY (Zheng *et al.*, 2012), and CDIF28668 PDB code 6UJE (Scott *et al.*, 2020)). (d) Structural superimposition CapGH5\_57 and GH5\_36 (Cel5B PDB codes 1VJZ (Joint Center for Structural Genomics, 2004) and 7EC9 (Manoj & Garg, 2022)) and Man5B PDB code 3W0K (Oyama *et al.*, 2013)). The manno-oligosaccharide (yellow sticks) was extracted from the endo- $\beta$ -mannanase StE273Adc (PDB code 4YTE).



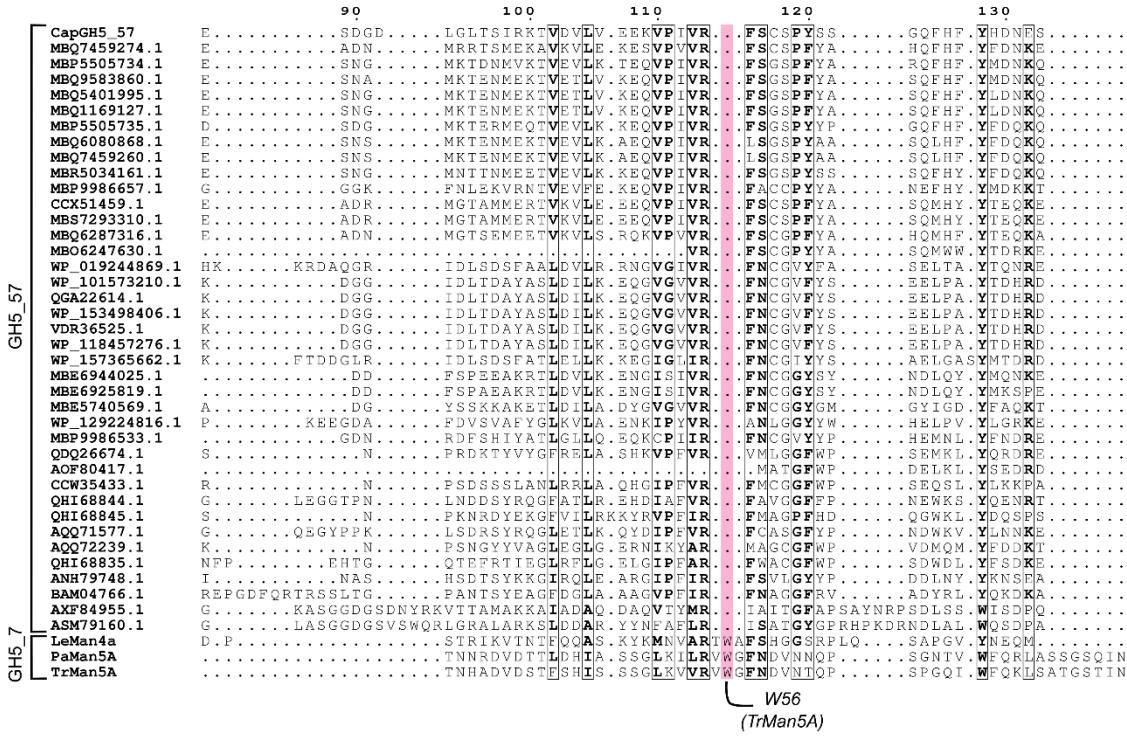
**Fig. S7: CapGH5\_57  $\beta$ 6- $\alpha$ 6 loop stabilization and the hydrophobic platform comprising the distal subsites.** (a) Representation of aromatic cluster that potentially stabilizes the  $\beta$ 6- $\alpha$ 6 loop. (b) Surface representation with the  $\beta$ 6- $\alpha$ 6 loop represented in cartoon and the distal positive subsites highlighted (sticks). The manno-oligosaccharide (yellow sticks) was obtained from structural superimposition with endo- $\beta$ -mannanase StE273Adc (PDB code 4Y7E (Kumagai *et al.*, 2015)). The oligosaccharide represented in green sticks is a fragment of the docked sugar (retrieved from PDB code 4Y7E) in the CapGH5\_57 active site (additional sugar units were omitted for clarity). (c) Positive subsites interactions with the manno-oligosaccharides described in (b). The yellow and green dashed lines represent the interaction of residues with the original (from structural superimposition) and the fragment of the docked oligosaccharide, respectively. The hydrophobic platform at the +4 subsite comprising the Trp270 is highlighted in light orange.



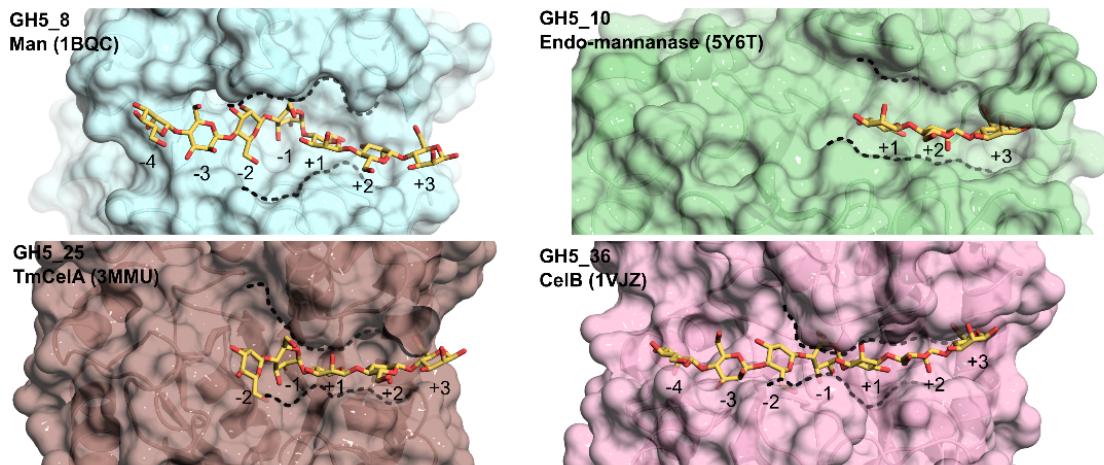
**Fig. S8: Structural comparison between CapGH5\_57 and the catalytic domain of other GH5\_57 members (Table S9), highlighting the loop containing the residue Trp159.** All CapGH5\_57 orthologues were modeled using RoseTTAFold, available in the Robetta server (Baek *et al.*, 2021). The commands “super or cealign” in the PYMOL were used for structural alignment. The loop shortening is highlighted in the right panel and the name of each member is represented by different colors.



**Fig. S9: Structural comparison of the Trp159-containing loop between CapGH5\_57 and other GH5 subfamilies displaying endo- $\beta$ -mannanase activity.** Comparison between CapGH5\_57 and the three most similar members (according to the DALI server) of the subfamilies (a) GH5\_8 (Man, PDB code 1BQC (Hilge *et al.*, 1998); StE273Adc, PDB code 4Y7E (Kumagai *et al.*, 2015); and  $\beta$ -mannanase, PDB code 4FK9 (Takasuka *et al.*, 2014)), (b) GH5\_10 (Endo- $\beta$ -mannanase, PDB code 5Y6T (Ueda *et al.*, 2018); CaMan, PDB code 4OOZ (Kim *et al.*, 2014); and Man5A, PDB code 2C0H (Larsson *et al.*, 2006)), (c) GH5\_25 (Tm\_Cel5A, PDB code 3MMU (Pereira *et al.*, 2010); FnCel5A, PDB code 3RJY (Zheng *et al.*, 2012); and CDIF28668, PDB code 6UJE (Scott *et al.*, 2020)) and (d) GH5\_36 (Cel5B, PDB codes 1VJZ (Joint Center for Structural Genomics, 2004) and 7EC9 (Manoj & Garg, 2022); and Man5B, PDB code 3W0K (Oyama *et al.*, 2013)). The manno-oligosaccharide spanning the -1 and +1 subsites was extracted from the GH5\_8 endo- $\beta$ -mannanase StE273Adc (PDB code 4Y7E (Kumagai *et al.*, 2015)). View of the active site of GH5\_25 members, highlighting the hydrophobic platform at the +1 subsite from the other  $\beta$ -sheet of (e) Tm\_Cel5A (PDB code 3MMU), (f) FnCel5A (PDB code 3RJY), and (g) CDIF28668 (PDB code 6UJE). The manno-oligosaccharide spanning the -2 to +3 subsites was also extracted from endo- $\beta$ -mannanase StE273Adc (PDB code 4Y7E (Kumagai *et al.*, 2015)).



**Fig. S10: Multiple sequence alignment of GH5\_57 members and characterized GH5\_7 members LeMan4a, PaMan5A and TrMan5A, highlighting the Trp conserved in the GH5\_7 subfamily, but absent in the GH5\_57 subfamily. Conserved residues are highlighted in black. The lack of Trp56 (numbered according to TrMan5A, PDB code 1QNO (Sabini *et al.*, 2000) is highlighted in pink.**



**Fig. S11. Overall view of the active site topology of GH5 endo- $\beta$ -mannanases from subfamilies GH5\_8, GH5\_10, GH5\_25 and GH5\_36.** The represented structures correspond to the three most similar members (according to the DALI server (Holm, 2022) from subfamilies GH5\_8 (Man, PDB code 1BQC (Hilge *et al.*, 1998), cyan surface), GH5\_10 (Endo- $\beta$ -mannanase, PDB code 5Y6T (Ueda *et al.*, 2018), green surface), GH5\_25 (TmCel5A, PDB code 3MMU (Pereira *et al.*, 2010), brown surface) and GH5\_36 (Cel5B, PDB code 1VJZ, (Joint Center for Structural Genomics, 2004), pink surface). The manno-oligosaccharide fragments (yellow sticks) were extracted from the endo- $\beta$ -mannanase StE273Adc (PDB code 4Y7E, (Kumagai *et al.*, 2015)).

## Supplementary References

- Baek, M., DiMaio, F., Anishchenko, I., Dauparas, J., Ovchinnikov, S., Lee, G. R., Wang, J., Cong, Q., Kinch, L. N., Dustin Schaeffer, R., Millán, C., Park, H., Adams, C., Glassman, C. R., DeGiovanni, A., Pereira, J. H., Rodrigues, A. V., Van Dijk, A. A., Ebrecht, A. C., Opperman, D. J., Sagmeister, T., Buhlheller, C., Pavkov-Keller, T., Rathinaswamy, M. K., Dalwadi, U., Yip, C. K., Burke, J. E., Garcia, K. C., Grishin, N. V., Adams, P. D., Read, R. J. & Baker, D. (2021). *Science*. **373**, 871–876.
- Hilge, M., Gloor, S. M., Rypniewski, W., Sauer, O., Heightman, T. D., Zimmermann, W., Winterhalter, K. & Piontek, K. (1998). *Structure*. **6**, 1433–1444.
- Holm, L. (2022). *Nucleic Acids Res.* **50**, W210–W215.
- Joint Center for Structural Genomics (2004). Crystal structure of Endoglucanase (TM1752) from Thermotoga maritima at 2.05 Å resolution. doi:10.2210/pdb1VJZ/pdb.
- Kim, M. K., An, Y. J., Song, J. M., Jeong, C. S., Kang, M. H., Kwon, K. K., Lee, Y. H. & Cha, S. S. (2014). *Proteins Struct. Funct. Bioinforma*. **82**, 3217–3223.
- Kumagai, Y., Yamashita, K., Tagami, T., Uraji, M., Wan, K., Okuyama, M., Yao, M., Kimura, A. & Hatanaka, T. (2015). *FEBS J.* **282**, 4001–4014.
- Larsson, A. M., Anderson, L., Xu, B., Muñoz, I. G., Usón, I., Janson, J. C., Stålbrand, H. & Ståhlberg, J. (2006). *J. Mol. Biol.* **357**, 1500–1510.
- Manoj, N. & Garg, P. (2022). Structure of the Thermotoga maritima Family 5 endoglucanase in complex with 1-deoxyxynojirimycin. doi: 10.2210/pdb7EC9/pdb.
- Oyama, T., Schmitz, G. E., Dodd, D., Han, Y., Burnett, A., Nagasawa, N., Mackie, R. I., Nakamura, H., Morikawa, K. & Cann, I. (2013). *PLoS One*. **8**, e80448.
- Pereira, J. H., Chen, Z., McAndrew, R. P., Sapra, R., Chhabra, S. R., Sale, K. L., Simmons, B. A. & Adams, P. D. (2010). *J. Struct. Biol.* **172**, 372–379.
- Sabini, E., Schubert, H., Murshudov, G., Wilson, K. S., Siika-Aho, M. & Penttilä, M. (2000). *Acta Crystallogr. Sect. D Biol. Crystallogr.* **56**, 3–13.

Scott, W., Lowrance, B., Anderson, A. C. & Weadge, J. T. (2020). *PLoS One.* **15**, e0242686.

Takasuka, T. E., Acheson, J. F., Bianchetti, C. M., Prom, B. M., Bergeman, L. F., Book, A. J., Currie, C. R. & Fox, B. G. (2014). *PLoS One.* **9**, e94166.

Ueda, M., Hirano, Y., Fukuhara, H., Naka, Y., Nakazawa, M., Sakamoto, T., Ogata, Y. & Tamada, T. (2018). *Enzyme Microb. Technol.* **117**, 15–22.

Zheng, B., Yang, W., Zhao, X., Wang, Y., Lou, Z., Rao, Z. & Feng, Y. (2012). *J. Biol. Chem.* **287**, 8336–8346.