Structural characterization of β-xylosidase XynB2 from *Geobacillus stearothermophilus* CECT43, a member of the glycoside hydrolase family GH52

Jose Antonio Gavira[a], Lellys M. Contreras[b], Hassan Mohamad Alshamaa[a,c], Josefa María Clemente-Jiménez[b], Felipe Rodríguez-Vico[b], Francisco Javier Las Heras-Vázquez[b] and Sergio Martínez-Rodríguez[a,c,*].

[a]Laboratorio de Estudios Cristalográficos. IACT. CSIC-Universidad de Granada, Granada, Spain. [b]Dpto. Química y Física, Universidad de Almería, Almería, Spain. [c]Dpto. Bioquímica y Biología Molecular III e Inmunología, Universidad de Granada, Granada, Spain.

* **Corresponding author:**
Sergio Martínez Rodríguez. Dpto. Bioquímica y Biología Molecular III e Inmunología, Universidad de Granada. Facultad de Medicina. Universidad de Granada. Avda. de la Investigación nº 11. 18071, Granada, Spain. e-mail: sergio@ugr.es

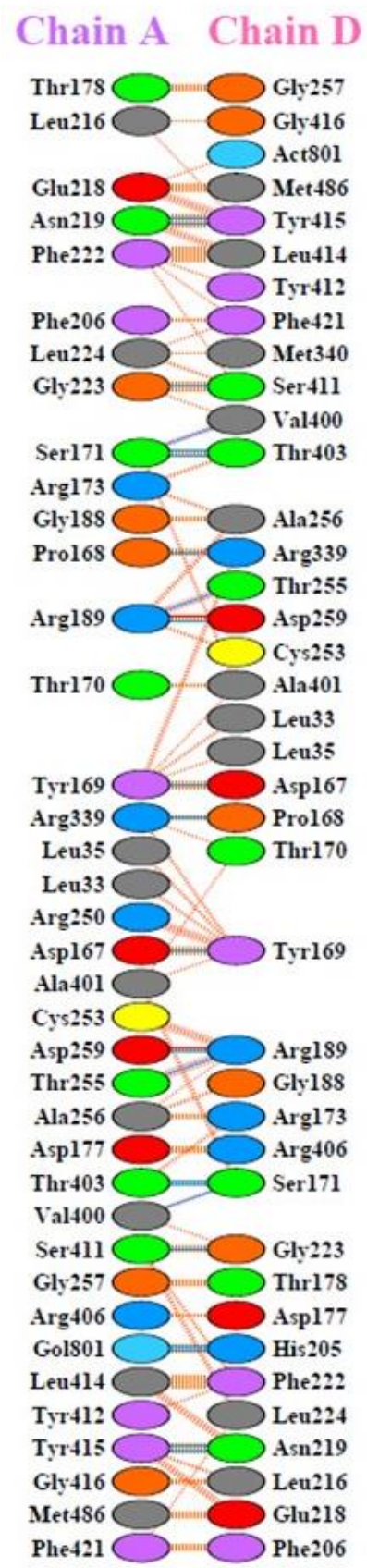
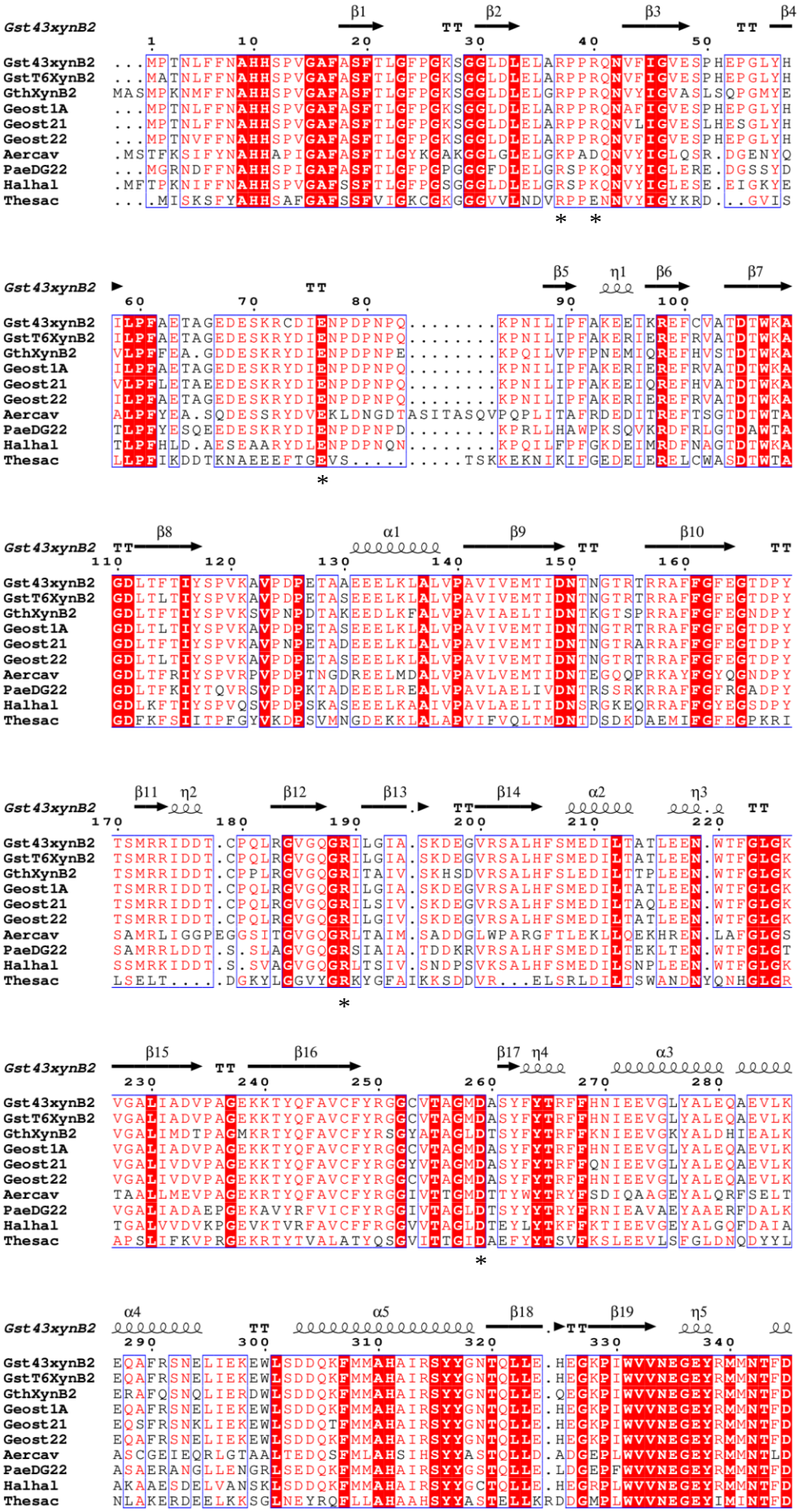


Figure S1. Contacts found in the Gst43XynB2 interface, calculated by PDBsum server. Salt bridges, hydrogen bonds and non-bonded contacts appear in red, blue and orange, respectively. Gol: glycerol. Act: acetate.


Gst43xynB2
Gst43xynB2
GstT6XynB2
GthXynB2
Geost1A
Geost21
Geost22
Aercav
PaeDG22
Halhal
Thesac

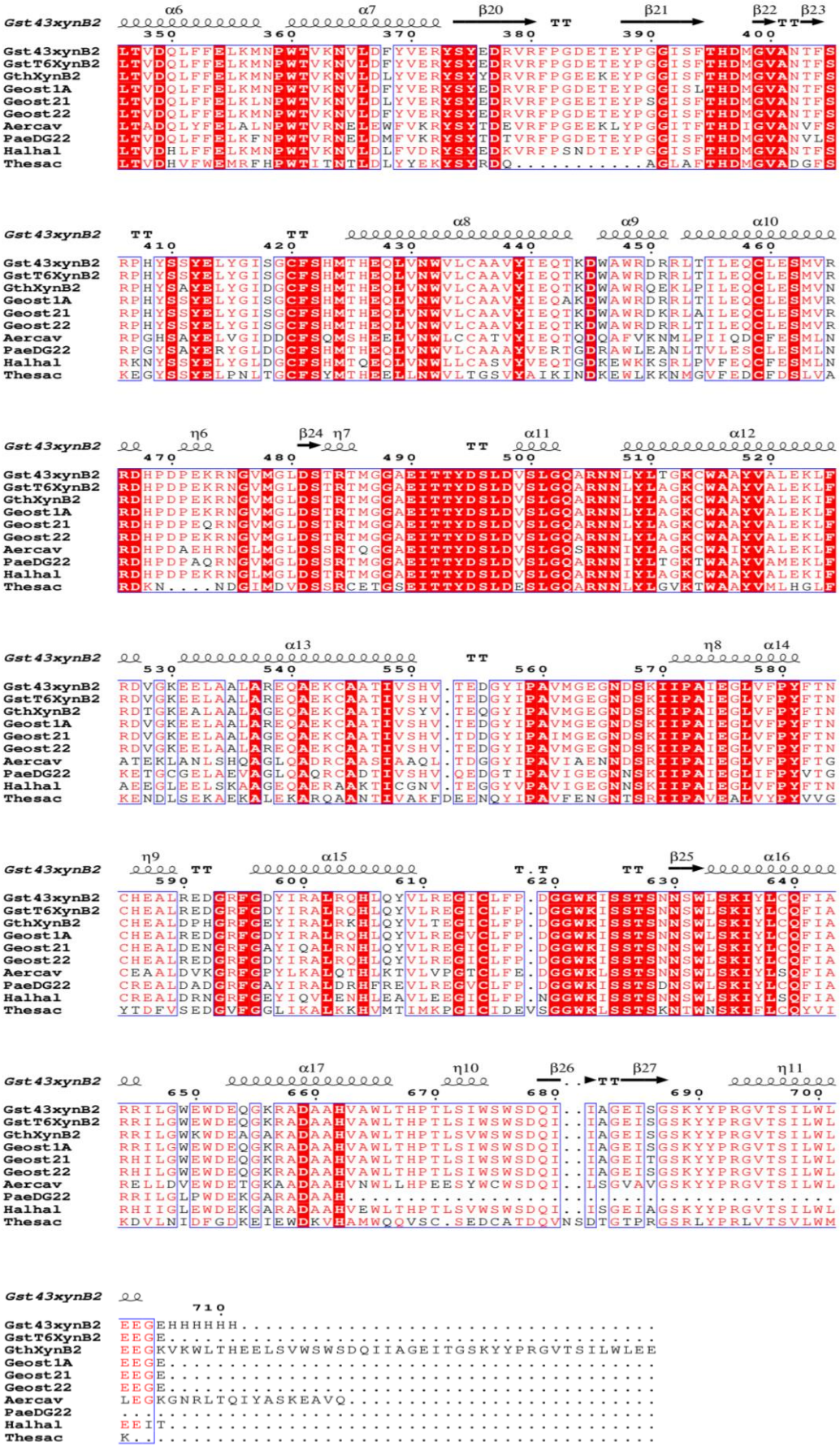

Gst43xynB2
GstT6XynB2
GthXynB2
Geost1A
Geost21
Geost22
Aercav
PaeDG22
Halhal
Thesac

Figure S2. Sequence alignment of β-xylosidases of contrasted activity available at the CAZY database (http://www.cazy.org/GH52_characterized.html, accessed November, the 8th, 2023). Gst43XynB2, *Geobacillus stearothermophilus* CECT43, Genbank acc. No. WOK24302; GstT6XynB2, *Geobacillus stearothermophilus* T-6 NCIMB 40222, Uniprot Q09LZ0; GthXynB2, *Parageobacillus thermoglucosidasius* NBRC 107763 / TM242, Uniprot A0A067XG64; Geost1A, *Geobacillus stearothermophilus* 1A05583. Uniprot M1GNL7; Geost21, *Geobacillus stearothermophilus* 21, Uniprot P45702; Geost22, *Geobacillus stearothermophilus* 21, Genbank acc. No. KFL17027.1; Aercav, *Aeromonas caviae* ME-1, Uniprot Q9Z487; PaeDG22, *Paenibacillus* sp. DG-22, Uniprot B6C867; Halhal, *Halalkalibacterium halodurans* C-125, Uniprot Q9KB21; Thesac, *Thermoanaerobacterium saccharolyticum* strain DSM 8691, Uniprot I3VVB3. Residues conserved in the interface of Gst43XynB2 are marked with asterisks (*).

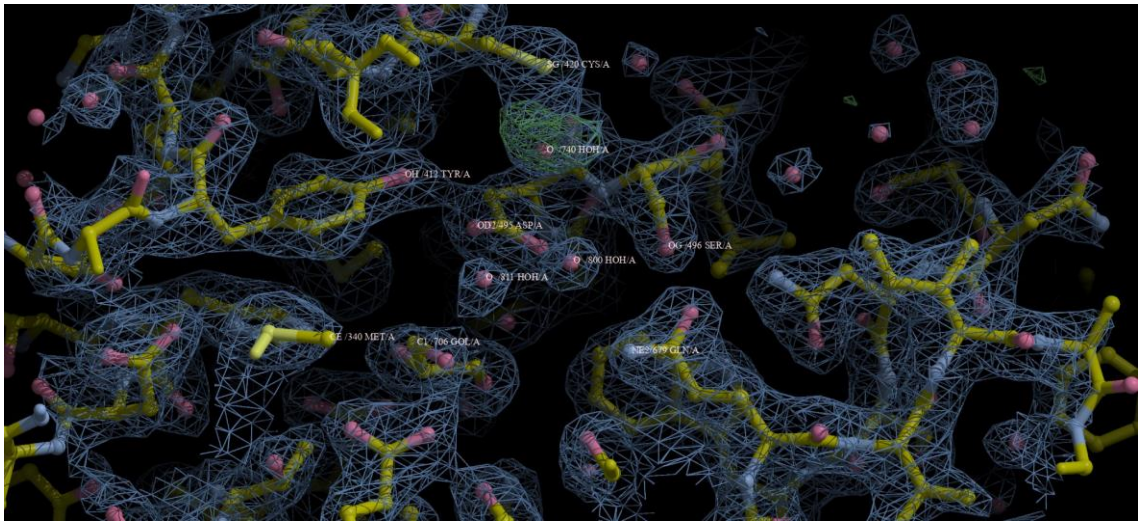


Figure S3. Electron densities in the Gst43XynB2 +1 site contoured at 1.09 σ. Modification of Cys420 by radiation damage is compatible with extra densities observed for this residue.

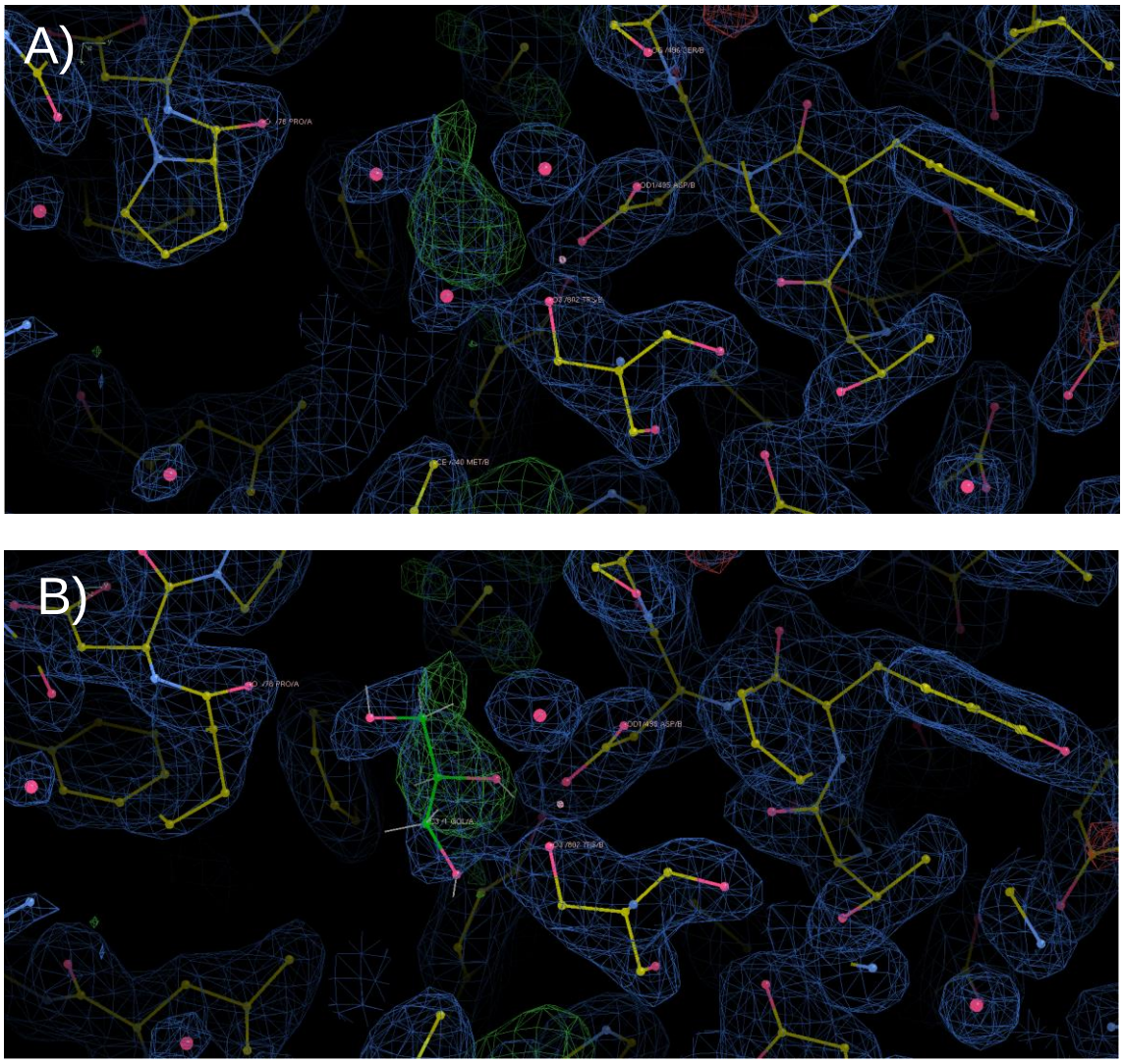


Figure S4. Extra electron densities found in the GstT6XynB2 +1 binding site A) Original fitting in GstT6XynB2 +1 site (PDB 4RHH, chain B). B) Placing of a glycerol molecule in the same region. No refinement was carried out.

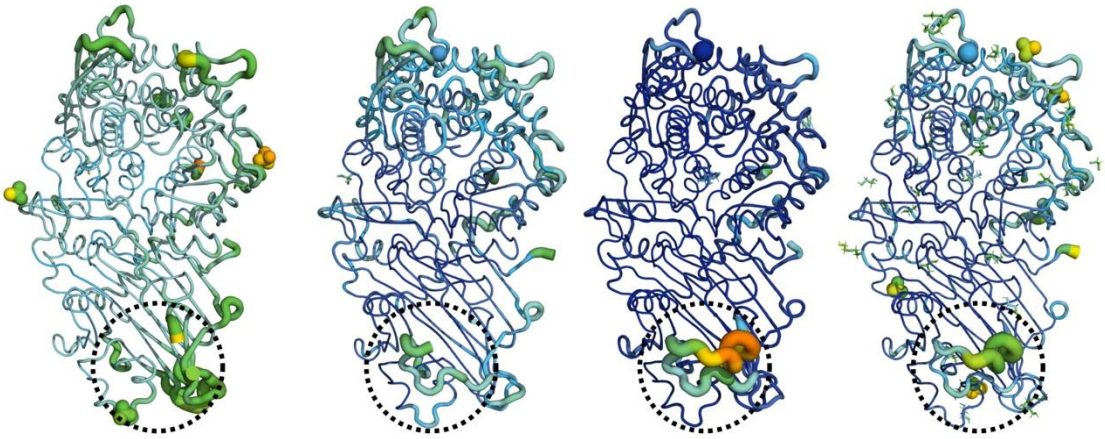


Figure S5. Representation in putty mode of the different *Geobacillus* XynB2 structures (from left to right, PDBs 4C1O, 4C1P, 4RHH and 8QME). The structures are coloured according to their B-factors. The position of the loop closing the catalytic centre (containing $Pro100^{GthXynB2}$ / $Pro78^{Gst43XynB2}$) is highlighted.

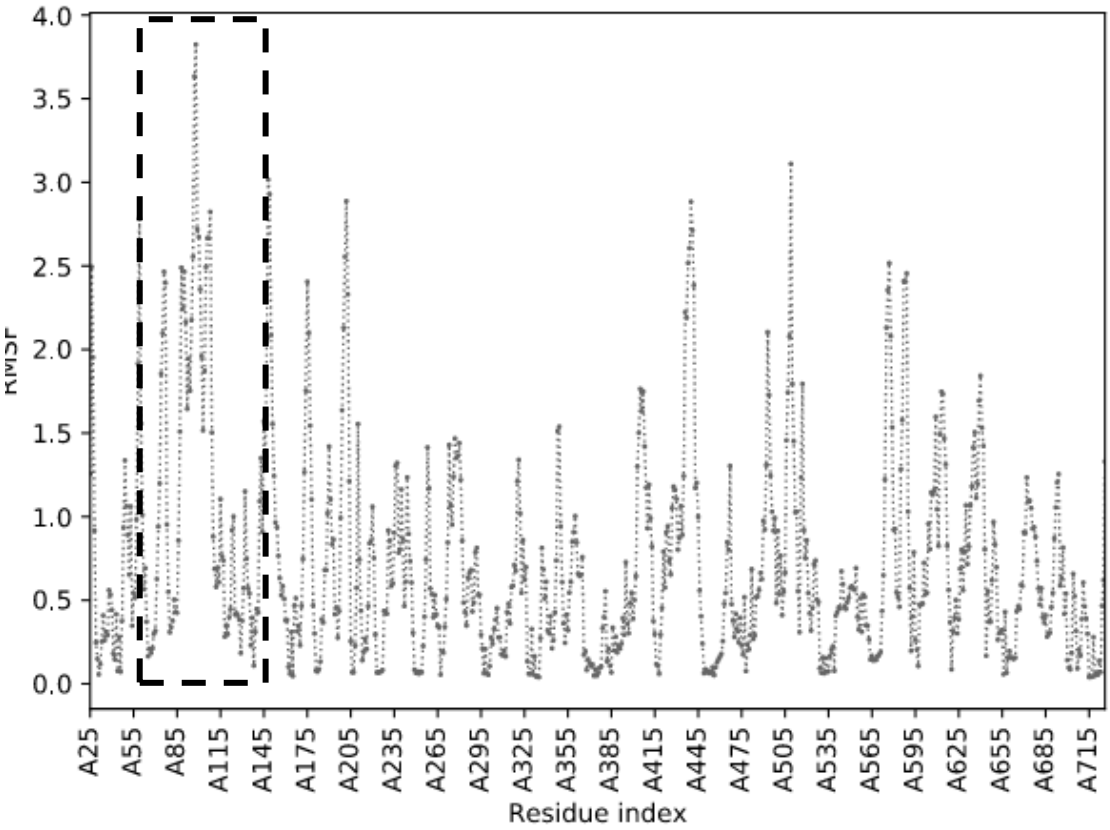


Figure S6. Fluctuation map obtained with the CABS-flex 2.0 server using the GthXynB2 structure as input (PDB 4C1O). The loop containing $Pro100^{GthXynB2}$ ($Pro78^{Gst43XynB2}$) is highlighted.

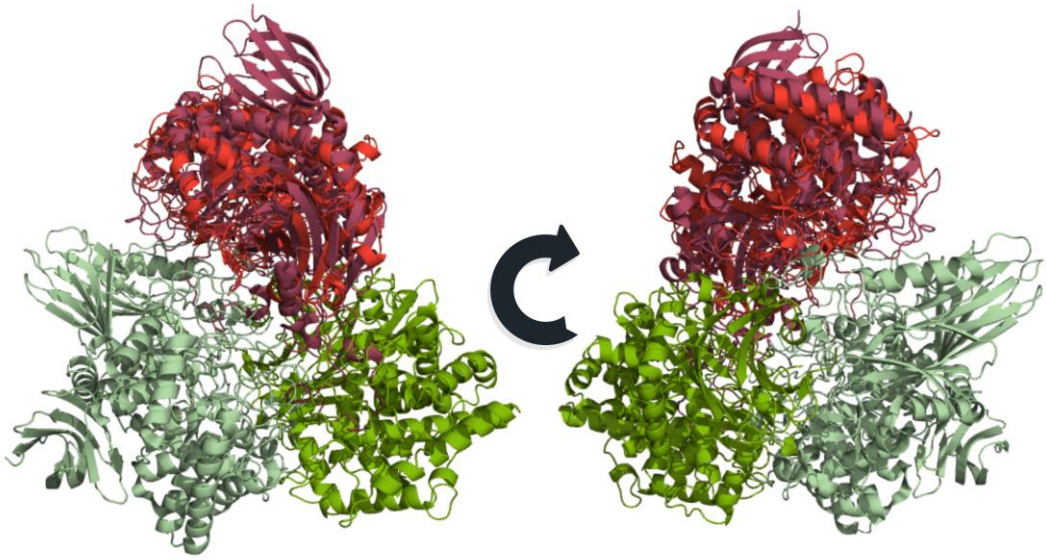


Figure S7. Different dimerization interfaces found in the GH52 and the GH116 families, using Gst43XynB2 (PDB 8QME) and ExoMA2 (PDB 8IC6) dimers as representatives. Both dimers were superposed using chain A (in red). As it is shown, the second protomer of each dimer (in different green colors) does not occupy the same position, being roughly rotated 100º.

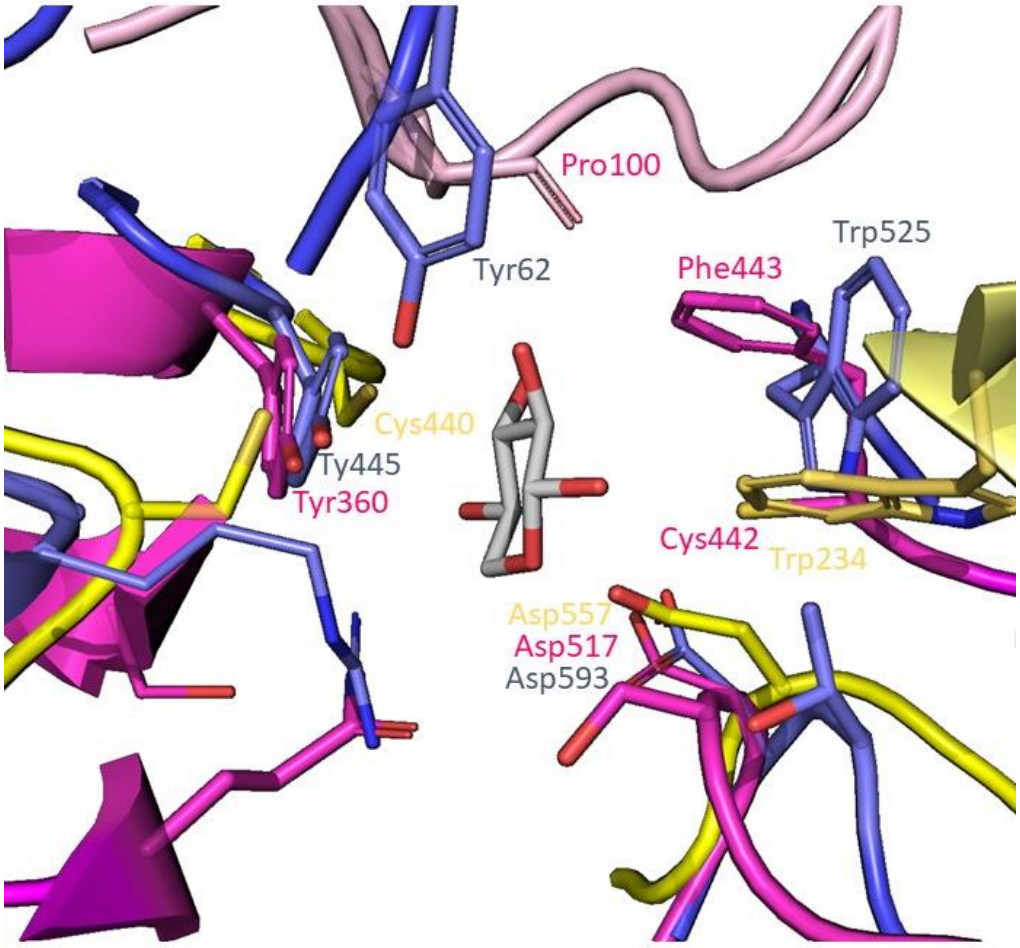


Figure S8. Superposition of the +1 binding site of GthXynB2 (PDB 4C1P, pink tones), TxGH116 (PDB 5BX5, blue tones) and ExoMA2 (PDB 8IC7, yellow tones). Xylose from PDB 4C1P, occupying the +1 site, is shown in white tones.

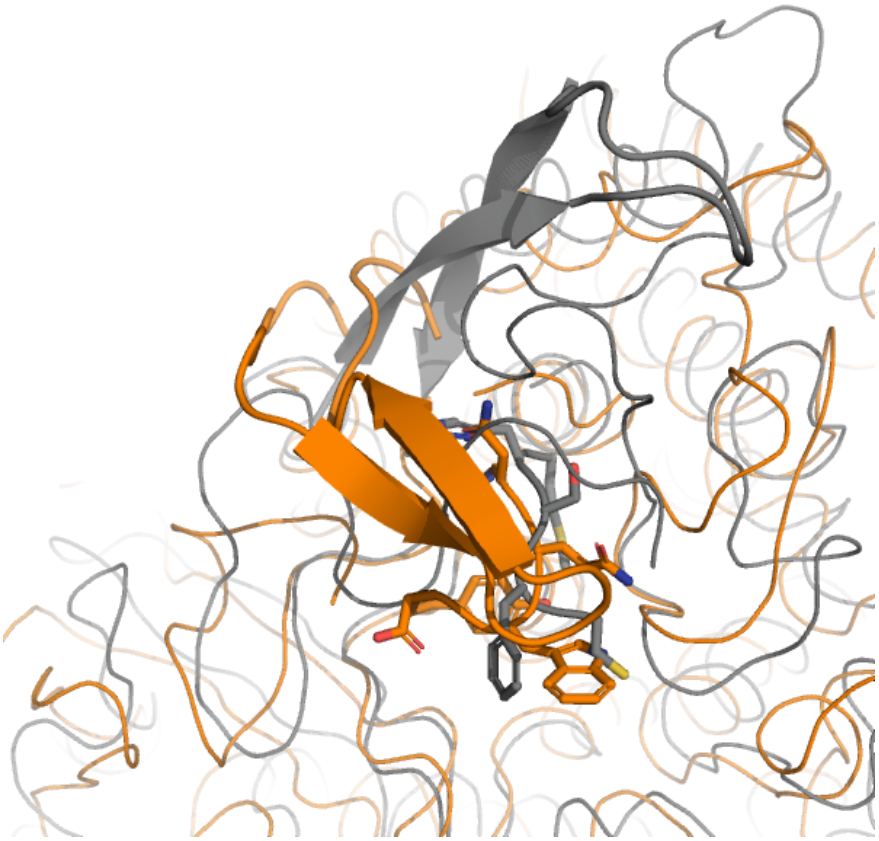


Figure S9. Common β-hairpin appearing in TxGH116 (PDB 5BX5, orange) and Gst43XynB2 (PDB 8QME, gray) which are totally displaced between the two structures (residues 480-504^{TxGH116}/373-395^{Gst43XynB2}).

	GthXyn	**GthXynB**	**GstT6XynB2**	**Gst43XynB2**
PDB ID	**4C1P**	**4C1O**	**4RHH**	**8QME**
Resolution (Å)	2.63	1.70	2.15	2.25
Ligand -1 subsite	xylobiose	Glycerol	Tris	Glycerol
	Glu357	Glu357	Gly335[c,d]	**Glu335[a]**
	Tyr360	Tyr360	Tyr338[c]	Tyr338[c]
	Met362	Met362	Met340[c]	Met340[c]
	Thr365	Thr365	Thr343	Thr343
	Asp367	Asp367	**Asp345[a]**	**Asp345[a]**
	Leu368	Leu368	Leu346[c]	Leu346[c]
	His418[a]	**His418[a]**	**His396[a]**	**His396[a]**
	Tyr434	Tyr434	Tyr412[c]	Tyr412[c]
	Met446	Met446	Met424	Met424
	Thr515[a]	**Thr515[a]**	**Thr493[a]**	**Thr493[a]**
	Tyr516[c]	Tyr516	Tyr494[c]	Tyr494[c]
	Asp 517	Asp 517	**Asp495[a]**	Asp495
	Trp654	Trp654	Trp632	Trp632
	Gln701	Gln701[c]	Gln679	Gln679[c]
	Ser710	Ser710	**Ser688[a]**	Ser688[c]
	Tyr713	Tyr713	**Try691[a]**	Try691[c]
	Arg715[a]	**Arg715[a]**	**Arg693[a]**	**Arg693[a]**
Ligand +1 subsite	xylobiose	Glycerol	None	None
Residues	**Pro100[a,b]**	**Pro100[a,b]**	Pro78*	Pro78[c]
	Pro102[b]	Pro102[b]	Pro80[c]	Pro80[c]
	Tyr360[d]	Tyr360	Tyr338[c]	Tyr338[c]
	Met362	Met362[c]	Met340[c]	Met340[c]
	Tyr434	Tyr434	Tyr412[c]	Tyr412[c]
	Cys442	Cys442	Cys420[c]	Cys420[c]
	Phe443	Phe443	Phe421[c]	Phe421[c]
	Asp517	**Asp517[a]**	Asp495[c]	Asp495[c]
	Ser518	Ser518	Ser496[c]	Ser496[c]
	Trp654[c]	Trp654	Trp632[c]	Trp632[c]
	Gln701[a]	**Gln701[a]**	Gln679[c]	Gln679[c]
	Ser710	Ser710[c]	Ser688[c]	Ser688[c]

Table S1. Correspondence of residues in the different *Geobacillus* XynB2 structures with any atom at less than 4Å from ligands found in binding subsites +1 and -1. [a]Direct polar contacts calculated by Pymol, also highlighted in bold. [b]Residues belonging to a second protomer. [c]Counterpart residues which do not appear or appear at more than 4Å from a ligand in the corresponding structure. [d]This structure is a E335G active-site mutant.