

Supplementary material for the following manuscript:

Yeast GH30 xylanase from *Sugiyamaella lignohabitans* is a glucuronoxylanase with auxiliary xylobiohydrolase activity

by Katarína Šuchová, Andrej Chyba, Zuzana Hegyi, Martin Rebroš and Vladimír Puchart.

TcXyn30B

1 10 20 30 40 50

β1X β1

TcXyn30A mrhp...ipilavlgra **sa**..WSYSQT..LSAN **IQ**VNALQR **YQ**EMIGGGCS **SG**AFGW **AC**QQQFPTT
TrXynIV mkss...isvvlallghs **aa**..wsyaTKSQYRAN **IK**INARQT **YQ**TMIGGGCS **SG**AFGI **AC**QQQFGSS
AaXyn30A mrsasviaaaavaslapl **aaa**TDLDPCVDEIVTE **IE**IRPDIK **HQ**QIEGGGFS **SG**AF **RA**QLL...L
TpXynC mhlsl...iidklfllgal **var**vta.TPTSSSGIS **VS**VNPKVQ **FO**EVDFGGS **QAF** **RA**EDIFGKY
TcXyn30Bmvfskvav **aa**sals.lgayaIDAQ **IN**VDLQAR **YQ**SVDFGGS **QAF** **RA**EDIFGKY
BXylD ..ms...iwt...watpsi **ast**vlaNSLSSPTT **IT**VDANEE **KQ**IVDGF **GF**SEAF **RA**ENVFGSA
TtXyn30A mysl...liall...cagt **av**.daqaLQQRQAGTT **LT**VDLSST **YQ**RIDGFGT **SE**AF **RA**VQM...S
SLXyn30A mkls...ltfltlsls.t **al**.....aSAPAPVA **VD**INLAQT **YQ**EIDGFG **GF**SEAF **RA**NDL...Y
TrXynVI ...m...ialv...pfv **ii**..agaatlsvaQTT **IT**VN **VS**GSK **FQ**QIDGFG **GF**SEAF **RA**REF...Q
TcXyn30C mws.....lk...snst **al**avalngi.valgQT **IT**VNPSTT **YQ**TIDGFG **GF**SEAF **RA**GAPI...A
TLXyn30A mssr...ialatl laqv **al**aAPRDQLVARQTT **AT**INLSQT **YQ**TMDGFG **GF**SEAF **RA**NDL...Y

1 10 20 30 40 50

TcXyn30B

60 70 80 90 100 110

α1 β2

TcXyn30A GLTPEN **QEEV** **TKILF** DENI **G**.GLSIVRN **DIGSS** PG.....S **TL**PTCPAT **PAG**PFN **YQ**WD **GS**D
TrXynIV GLSPEN **QQKV** **TQILF** DENI **G**.GLSIVRN **DIGSS** PG.....T **TL**PTCPAT **PQ**DKFD **YV**WD **GS**D
AaXyn30A NVTQDT **QQEL** **LDLIY** NRET **GL**GFSILRN **GIGSS** NS **SY**NDWMN **TL**LPES **PG**SPDGEFN **YV**WD **EY**D
TpXynC GLPAVN **QSLV** **LDLLY** NENV **GAG**FTILRN **GIGSS** NT **SS**SNFMN **SI**EPKN **PG**GPNAIPE **YV**WD **HY**N
TcXyn30B GLSPKN **QSYV** **LDLMY** SEER **GAG**FTILRN **GIGSS** NS **ST**SNLMN **SI**EPFS **PG**SPSTPN **YV**WD **HY**N
BXylD GLSPAN **QQRL** **LDLMY** DENI **GAG**FTILRN **GIGSS** NL **SD**PSDMI **SI**ELTD **PG**LPSSKPT **YTP**.NNN
TtXyn30A RLPEEG **QRRR** **LDVLF** STTN **GAG**LSILRN **GIGSS** PD **MS**SDHMS **SI**APKS **PG**SPNNPLI **YS**WD **GS**D
SLXyn30A HMPEAG **RRKA** **LDLLF** HPTK **GAG**FSTIRN **GIGSS** PN **SS**SNWMN **SI**EPVS **PG**GPNSPPH **YV**WD **RN**D
TrXynVI SANATA **QKAL** **LDLLF** STST **GAG**FSTIRN **RIGSG** GA.....GD **SI**EPNN **PG**SPSAAPN **YV**WD **NN**D
TcXyn30C SASASI **QTQV** **TNYLF** STTT **GAG**LTILRN **RI**AAGSG **SI**EPNA **PG**SPNAQPT **YT**WD **GN**D
TLXyn30A NLPSSQ **RQYA** **LDLLF** STTN **GAG**MTILRN **RIGSG** GT.....GD **SI**EPNS **PG**SPSATPN **YV**WD **GS**D

60 70 80 90 100 110

TcXyn30B

120 130 140 150 160 170

α2 β3

TcXyn30A **SC**QFNLTK **TA**LKYN **PE**LY **VY**AN **AW**SAPGCMKT **VT**GTEN **GG**Q **IC**GV **RG**T.....N **CT**.YDWR **QAY**
TrXynIV **NC**QFNLTK **TA**LKYN **PN**LY **VY**AD **AW**SAPGCMKT **VT**GTEN **LG**Q **IC**GV **RG**T.....D **CK**.HDWR **QAY**
AaXyn30A **SG**QF **FL**AKK **AV**.EY **G**VERI **YI**AN **AW**SAPGFMKNN **AD**DAN **GG**LL **LC**GV **VP**SFDGNSA **CE**.HDWR **QAY**
TpXynC **SG**QF **PL**AQ **QAY**.SR **GL**HTL **YI**AN **AW**SAPGYMKTN **DN**ENN **GG**YL **LC**GVANE.....S **CAS**GNW **QAY**
TcXyn30B **SG**QF **PL**SK **QAY**.AR **GL**PY **YI**AD **AW**SAPGYMKTN **QD**ENW **SG**FL **LC**GI **EG**E.....T **CPS**GDWR **QAY**
BXylD **TG**QL **PL**AK **QAY**.AR **GL**KT **YI**AD **AW**SAPWFMKTN **LN**DND **GG**YL **LC**GV **ED**T.....D **CPT**GSW **QAY**
TtXyn30A **NK**QL **WV**SQ **EA**VHTY **G**VKT **YI**AD **AW**SAPGYMKTN **GN**DAN **GG**TL **LC**GL **SG**A.....Q **CAS**GDWR **QAY**
SLXyn30A **SS**QF **WFA**Q **QAV**.SY **G**VKT **YI**AD **AW**SAPGFMKTN **DN**ENY **GG**YL **LC**GV **RD**T.....N **CAS**GDWR **QAY**
TrXynVI **SG**QL **WFT**Q **QAV**.SY **G**VKT **YI**AD **AW**GAPGFMKTS **GS**DSS **PG**YL **LC**GT **TG**H.....S **CSS**GDWR **QAY**
TcXyn30C **AG**QV **WWS**K **QAR**.AK **G**VKY **YI**AD **AW**SAPAFMKTN **DN**VAN **GG**YL **LC**GT **TG**E.....T **CSS**GDWR **QAY**
TLXyn30A **TI**NQV **WVT**Q **QAV**.SY **G**VKT **YI**AD **AW**SAPGFMKTN **ND**QNN **GG**YL **LC**GV **QNE**.....T **CAS**GNW **QAY**

120 130 140 150 160 170

TcXyn30B

180 190 200 210 220 230

α3 β4 α4

TcXyn30A **AD**YL **VQ**YV **KF**YQAE **GI**D **IS**LGAWN **EP**DFNPVT **YE**SME. **SD**GF **Q**.AK **DF**LE **IL**YP **TV**KKAF **PN**.
TrXynIV **AD**YL **VQ**YV **RF**YKEE **GI**D **IS**LGAWN **EP**DFNPFT **YE**SML. **SD**GY **Q**.AK **DF**LE **VL**YP **TL**KKAF **PK**.
AaXyn30A **AD**YL **VK**YI **QI**YAAE **GI**P **IS**DVAV **IN**EPDLT.TSY **AS**MR. **SNA**E. **VAD**FL **PI**LH **KT**LE **DAG**LD.
TpXynC **AN**YL **VQ**YAR **F**YKES **GV**RV **TH**LGFL **NE**PQFA.AT **YAG**ML. **SNG**T **Q**.AAEF **IR**VL **AK**TV **KK**SGLD.
TcXyn30B **AD**YL **VQ**YV **KF**YAES **GV**PT **TH**LGFL **NE**PQEV.VS **YAS**MG. **SNG**T **Q**.AAEF **VK**IL **GQ**TL **ER**GID.
BXylD **VN**YL **LE**YV **KL**YKEA **GV**TV **TV**NV **GF**LN **NE**PQLN.TS **YAS**MQ. **SNG**T **Q**.AAEF **IR**VL **GR**TI **RE**QGLD.
TtXyn30A **AD**YL **TK**YV **EF**YQES **NV**TV **TH**LGFL **NE**PELT.TS **YAS**MR. **SAS**Q. **AAEF** **IR**IL **YP**TI **QK**SNLT **Y**
SLXyn30A **AN**YL **VQ**YI **KF**YKES **GI**D **IT**HV **GF**LN **NE**PEYN **PG**GY **YAG**ML. **SSG**F **Q**.AADF **IK**VL **YP**TI **QRA**GLST
TrXynVI **AN**FL **VQ**YV **KY**YAAA **GY**NI **TH**LGFL **NE**PDYQ.TT **YS**OM **QI**SSNA **QEA**IS **FI**PL **SS**TV **KA**AGLN.
TcXyn30C **AN**YL **VQ**YI **KD**YANE **GI**TI **ID**FV **GW**LN **NE**PDYS.PN **YD**SML **IT**SGT **Q**.AASF **IP**TL **YN**TI **KS**AGLS.
TLXyn30A **AN**FL **VQ**YI **KD**YKSE **GL**DI **TH**VGFL **NE**PDYV.AS **YS**SM **Q**.**SDG**Y **Q**.AADF **IK**VL **YP**TI **QSA**GLG.

180 190 200 210 220 230

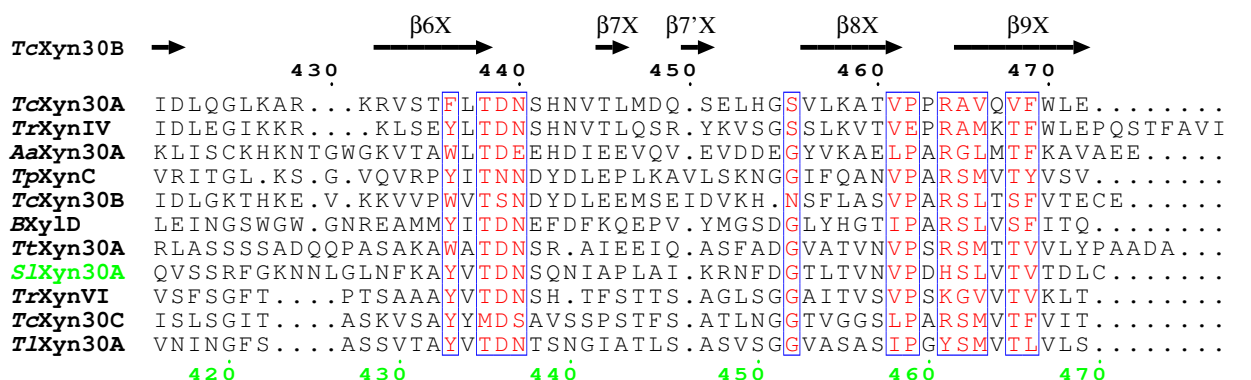
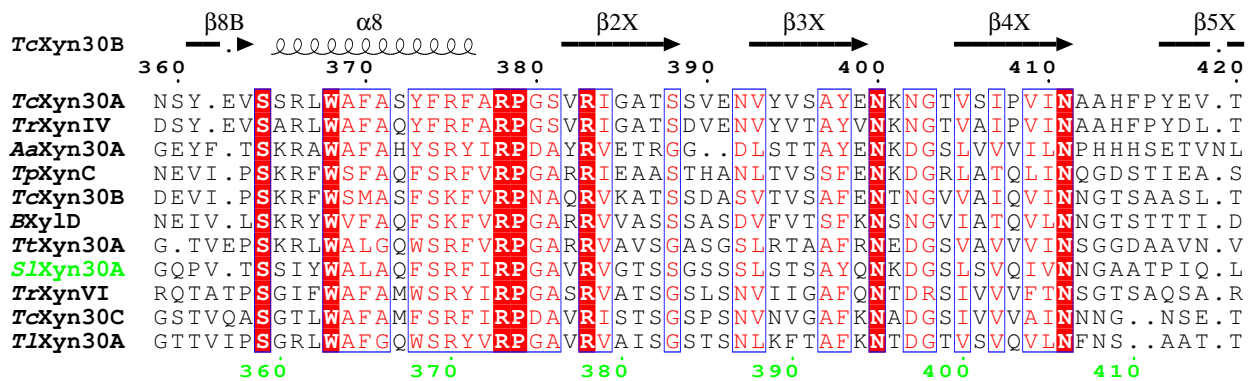
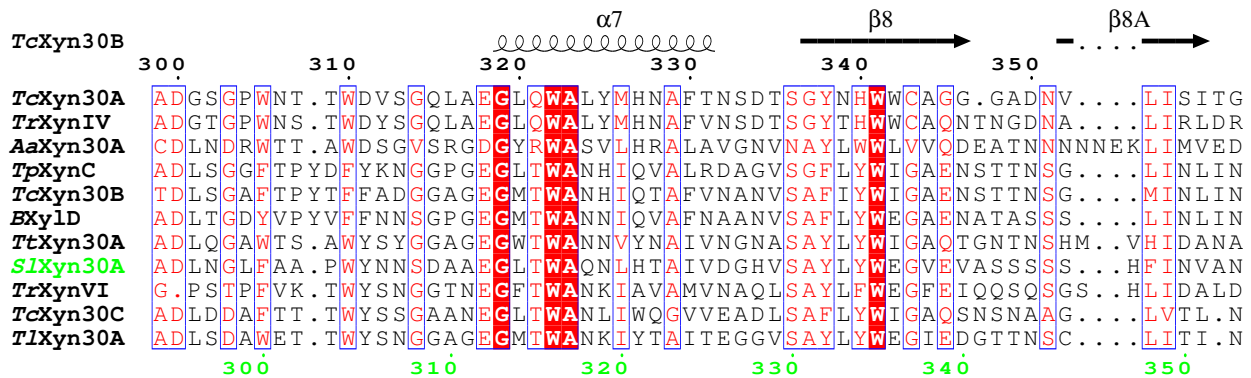
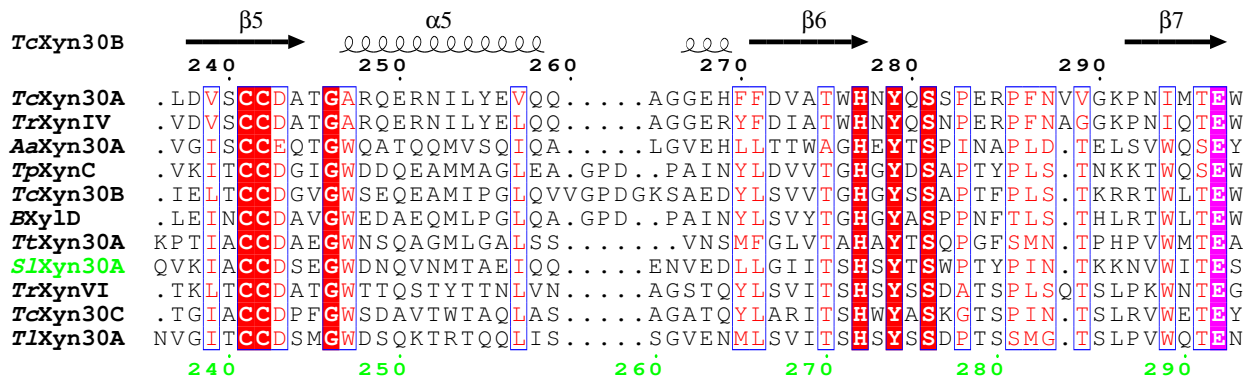


Figure S1. Multiple sequence alignment of *SlXyn30A* and all characterized eukaryotic GH30_7 xylanases, namely *Talaromyces cellulolyticus* *TcXyn30B* (GAM36763.1), *TcXyn30C* (GAM40414.1), *TcXyn30A* (GAM43270.1), *Thermothelomyces thermophila* *TtXyn30A* (AEO55025.1), *Talaromyces purpureogenus* (*Penicillium purpurogenum*) *TpXynC* (AKH40280.1), *Bispora* sp. *XylD* (*BXylD*; ADG62369.1), *Trichoderma reesei* *XynVI* (*TrXynVI*; EGR45006.1), *T. reesei* *XynIV* (*TrXynIV*; AAP64786.1), *Acremonium alcalophilum* *AaXyn30A* [7], and *Talaromyces leycettanus* *TlXyn30A* [10]. The *SlXyn30A* amino acid numbering is shown in green at the bottom. Secondary structure elements and amino acid numbering of *TcXyn30B* is shown on top. Catalytic acid/base and catalytic nucleophile (in *SlXyn30A* Glu199 and Glu292, respectively) are violet highlighted. Conserved arginine (Arg46) which takes a part in MeGlcA recognition is typed in blue and is marked by a blue triangle. Extended region within a β 2- α 2 loop that is presumably responsible for xylobiohydrolase activity is yellow highlighted. Within the extended region an aspartic acid or an asparagine (Asn90) is often found and is typed in brown and marked by a brown triangle. The Asx residue (*TcXyn30B* Asn93 and *TtXyn30A* Asp78) has been demonstrated to be hydrogen bonded to Xylp moiety accommodated in the -2a subsite.

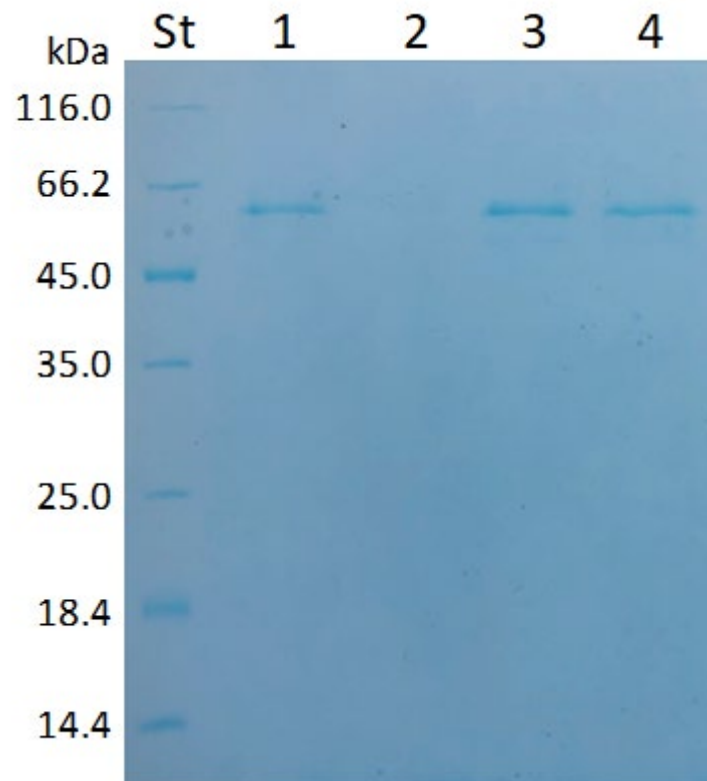


Figure S2: SDS-PAGE electrophoresis of recombinant *P.pastoris* fermentation broths after induction with methanol. St-MW standard, 1-4 fermentation broth of four transformants.