



STRUCTURAL
BIOLOGY

Volume 78 (2022)

Supporting information for article:

Synergic action of an inserted carbohydrate-binding module in a glycoside hydrolase family 5 endoglucanase

Ting-Juan Ye, Kai-Fa Huang, Tzu-Ping Ko and Shih-Hsiung Wu

```

MtGlu5 : -----*-----20-----*-----40-----*-----60-----*-----80-----*-----100-----: 28
503328546 : -----MSAIVLLTCCIVGVLPASAVTPTTNP-----MRFLAVSGLLVLVGPVGCSTQLQTPAP-----: 28
503325617 : -----MKMMNKTQIPSFSPFLMVMVILLLFLFG-CAVRPDDADISVTTDKAGTRLGG-----SLSSTKTVPPFVSV: 62
503546037 : -----MKKSIILYIIVLSLIFMTACDNGNTERS-----MRLKILLLVLLVLIPLACGSSVEFKLE-----: 28
501542033 : -----MKFIYRILVFLFAVIFLPMACNSALDNEE-----MRLKILLLVLLVLIPLACGSSVEFKLE-----: 26
493987263 : -----MKLKVFLLLALLLVFVAGCTSYELKLE-----MRLKILLLVLLVLIPLACGSSVEFKLE-----: 29
647306738 : -----MRLKILLLVLLVLIPLACGSSVEFKLE-----MRLKILLLVLLVLIPLACGSSVEFKLE-----: -
501578355 : -----MRLKILLLVLLVLIPLACGSSVEFKLE-----MRLKILLLVLLVLIPLACGSSVEFKLE-----: 26
504265218 : -----MRLKILLLVLLVLIPLACGSSVEFKLE-----MRLKILLLVLLVLIPLACGSSVEFKLE-----: 36
568335214 : -----MRLKILLLVLLVLIPLACGSSVEFKLE-----MRLKILLLVLLVLIPLACGSSVEFKLE-----: 27
658466537 : -----MRLKILLLVLLVLIPLACGSSVEFKLE-----MRLKILLLVLLVLIPLACGSSVEFKLE-----: 36
428131024 : -----MRLKILLLVLLVLIPLACGSSVEFKLE-----MRLKILLLVLLVLIPLACGSSVEFKLE-----: 33
490201866 : -----MRLKILLLVLLVLIPLACGSSVEFKLE-----MRLKILLLVLLVLIPLACGSSVEFKLE-----: 21
500776899 : -----MRLKILLLVLLVLIPLACGSSVEFKLE-----MRLKILLLVLLVLIPLACGSSVEFKLE-----: 25

```

```

MtGlu5 : -----*-----120-----*-----140-----*-----160-----*-----180-----*-----200-----: 113
503328546 : -----DTGGIVELNRQLGCVNLENALAEAREGEGW-VRLDEEGRBELHREAGFKTHRELVSWTHAGRAAPMIDPAPFSSRQWAWTQCFR: 113
503325617 : -----SDIFKANARLGRGNLNALEAAREGEGW-VVLEADDKNVIRAGFDSVRLPFRWSSNAKTTFPMDPAPARRDWAIGQCFR: 111
503546037 : -----VFGMRRCNINALEAAREGEGW-VTKVDDEKTLHREAGFDSVRLPFRWSSAPADSEFPMDPAPARRDWAIGQCFR: 161
501542033 : -----MLPILRCNINALEAAREGEGW-VVTKVDDEKTLHREAGFDSVRLPFRWSSAPADSEFPMDPAPARRDWAIGQCFR: 105
493987263 : -----AKQEEWAEAGLIN-----QRLGRGNINALEAAREGEGW-VVTKVDDEKTLHREAGFDSVRLPFRWSSAPADSEFPMDPAPARRDWAIGQCFR: 118
647306738 : -----MRLKILLLVLLVLIPLACGSSVEFKLE-----MRLKILLLVLLVLIPLACGSSVEFKLE-----: 74
501578355 : -----MLPILRCNINALEAAREGEGW-VVTKVDDEKTLHREAGFDSVRLPFRWSSAPADSEFPMDPAPARRDWAIGQCFR: 103
504265218 : -----SMDQSVAESDSNSALEYKNIIGRCNINALEAAREGEGW-VVTKVDDEKTLHREAGFDSVRLPFRWSSAPADSEFPMDPAPARRDWAIGQCFR: 130
568335214 : -----RNKILGRGNINALEAAREGEGW-VVTKVDDEKTLHREAGFDSVRLPFRWSSAPADSEFPMDPAPARRDWAIGQCFR: 105
658466537 : -----STEQSVAEIKSKSAFYNKMGICNINALEAAREGEGW-VVTKVDDEKTLHREAGFDSVRLPFRWSSAPADSEFPMDPAPARRDWAIGQCFR: 130
428131024 : -----SKKEGETVHVSKSAFYDKMIGRCNINALEAAREGEGW-VVTKVDDEKTLHREAGFDSVRLPFRWSSAPADSEFPMDPAPARRDWAIGQCFR: 127
490201866 : -----ENKKTAFDYDKMIGICNINALEAAREGEGW-VVTKVDDEKTLHREAGFDSVRLPFRWSSAPADSEFPMDPAPARRDWAIGQCFR: 108
500776899 : -----QSVSNVDKSSAFYNKMGICNINALEAAREGEGW-VVTKVDDEKTLHREAGFDSVRLPFRWSSAPADSEFPMDPAPARRDWAIGQCFR: 116

```

GiN gnALEAP EG WG V e F ik GF vRiPi Wsah Py I F Rv v a

```

MtGlu5 : -----*-----220-----*-----240-----*-----260-----*-----280-----*-----300-----: 212
503328546 : -----RCGNIVVMVHYEINANQAEAEAYLSTFRQCFERYRNGCGS-VYFELINEHGRFNNDNPQWDLAKALRVVRESNESAVVGPVGNLWLRSE: 212
503325617 : -----NNDRVVNNHHYEDMENRSLHHRFLAIEQCLHVDKES-VYFELINEHGRFNNDNPQWDLAKALRVVRESNESAVVGPVGNLWLRSE: 208
503546037 : -----QRCFTLNNHHYEDMENRSLHHRFLAIEQCLHVDKES-VYFELINEHGRFNNDNPQWDLAKALRVVRESNESAVVGPVGNLWLRSE: 257
501542033 : -----NDPVLNHHYEDMENRSLHHRFLAIEQCLHVDKES-VYFELINEHGRFNNDNPQWDLAKALRVVRESNESAVVGPVGNLWLRSE: 200
493987263 : -----NCLAVNHHYEDMENRSLHHRFLAIEQCLHVDKES-VYFELINEHGRFNNDNPQWDLAKALRVVRESNESAVVGPVGNLWLRSE: 215
647306738 : -----CRQVTLNHHYEDMENRSLHHRFLAIEQCLHVDKES-VYFELINEHGRFNNDNPQWDLAKALRVVRESNESAVVGPVGNLWLRSE: 170
501578355 : -----NNDRVVNNHHYEDMENRSLHHRFLAIEQCLHVDKES-VYFELINEHGRFNNDNPQWDLAKALRVVRESNESAVVGPVGNLWLRSE: 200
504265218 : -----NNDRVVNNHHYEDMENRSLHHRFLAIEQCLHVDKES-VYFELINEHGRFNNDNPQWDLAKALRVVRESNESAVVGPVGNLWLRSE: 227
568335214 : -----RCGAVNHHYEDMENRSLHHRFLAIEQCLHVDKES-VYFELINEHGRFNNDNPQWDLAKALRVVRESNESAVVGPVGNLWLRSE: 202
658466537 : -----NNDRVVNNHHYEDMENRSLHHRFLAIEQCLHVDKES-VYFELINEHGRFNNDNPQWDLAKALRVVRESNESAVVGPVGNLWLRSE: 227
428131024 : -----NNDRVVNNHHYEDMENRSLHHRFLAIEQCLHVDKES-VYFELINEHGRFNNDNPQWDLAKALRVVRESNESAVVGPVGNLWLRSE: 224
490201866 : -----NNDRVVNNHHYEDMENRSLHHRFLAIEQCLHVDKES-VYFELINEHGRFNNDNPQWDLAKALRVVRESNESAVVGPVGNLWLRSE: 205
500776899 : -----NNDRVVNNHHYEDMENRSLHHRFLAIEQCLHVDKES-VYFELINEHGRFNNDNPQWDLAKALRVVRESNESAVVGPVGNLWLRSE: 213

```

L iN HH El P W Q a yp l FE NEP l Wn iR np r i lD Vn W a l

```

MtGlu5 : -----*-----320-----*-----340-----*-----360-----*-----380-----*-----400-----: 309
503328546 : -----RUPD-PNIIIVTHVYDLEBETHCGAEEWLN-EVPPVGVVRENQGAFAAGWQNWWSGSRVGFVGEALITYQEGWAGFYLHSDAGVEGYDLRALFRTSP: 309
503325617 : -----TIPETRNIIIVTHVYDLEBETHCGAEEWLN-EVPPVGVVRENQGAFAAGWQNWWSGSRVGFVGEALITYQEGWAGFYLHSDAGVEGYDLRALFRTSP: 251
503546037 : -----ETPEDKNIIIVTHVYDLEBETHCGAEEWLN-EVPPVGVVRENQGAFAAGWQNWWSGSRVGFVGEALITYQEGWAGFYLHSDAGVEGYDLRALFRTSP: 301
501542033 : -----KUPKDEKNIIVTHVYDLEBETHCGAEEWLN-EVPPVGVVRENQGAFAAGWQNWWSGSRVGFVGEALITYQEGWAGFYLHSDAGVEGYDLRALFRTSP: 244
493987263 : -----VUPDDR-IILTHVYDLEBETHCGAEEWLN-EVPPVGVVRENQGAFAAGWQNWWSGSRVGFVGEALITYQEGWAGFYLHSDAGVEGYDLRALFRTSP: 243
647306738 : -----KUPPEERNIIIVTHVYDLEBETHCGAEEWLN-EVPPVGVVRENQGAFAAGWQNWWSGSRVGFVGEALITYQEGWAGFYLHSDAGVEGYDLRALFRTSP: 258
501578355 : -----RUPKDEKNIIVTHVYDLEBETHCGAEEWLN-EVPPVGVVRENQGAFAAGWQNWWSGSRVGFVGEALITYQEGWAGFYLHSDAGVEGYDLRALFRTSP: 213
504265218 : -----KIVN-EKRIIVSHVYDLEBETHCGAEEWLN-EVPPVGVVRENQGAFAAGWQNWWSGSRVGFVGEALITYQEGWAGFYLHSDAGVEGYDLRALFRTSP: 243
568335214 : -----SVPKWEKNIIVTHVYDLEBETHCGAEEWLN-EVPPVGVVRENQGAFAAGWQNWWSGSRVGFVGEALITYQEGWAGFYLHSDAGVEGYDLRALFRTSP: 269
658466537 : -----KIVN-EKRIIVSHVYDLEBETHCGAEEWLN-EVPPVGVVRENQGAFAAGWQNWWSGSRVGFVGEALITYQEGWAGFYLHSDAGVEGYDLRALFRTSP: 245
428131024 : -----KIVD-KNIIVSHVYDLEBETHCGAEEWLN-EVPPVGVVRENQGAFAAGWQNWWSGSRVGFVGEALITYQEGWAGFYLHSDAGVEGYDLRALFRTSP: 269
490201866 : -----KIVD-PYLIVSHVYDLEBETHCGAEEWLN-EVPPVGVVRENQGAFAAGWQNWWSGSRVGFVGEALITYQEGWAGFYLHSDAGVEGYDLRALFRTSP: 266
500776899 : -----KIVD-KNIIVSHVYDLEBETHCGAEEWLN-EVPPVGVVRENQGAFAAGWQNWWSGSRVGFVGEALITYQEGWAGFYLHSDAGVEGYDLRALFRTSP: 247

```

l d Iv fhY P f ETHCGAEEWV G W g

```

MtGlu5 : -----*-----420-----*-----440-----*-----460-----*-----480-----*-----500-----: 408
503328546 : -----VSLQVSCRDDAPAKAVTTSGGVETVNNLSECGNPSRLTDLILQNNSPNARAARFRLELRLRGPSPALLTHQONAAQAEFAQRQAEQNR-REIIVVE: 408
503325617 : -----FERKLTKEFDQALEFSRRTN-REIIVVE-----WKKKATDTEFDLALMAQRYH-REIIVVE: 279
503546037 : -----DQKQFDEILKAVDIAKEHN-RQLIMSE-----WKKKATDTEFDLALMAQRYH-REIIVVE: 329
501542033 : -----DEKAEDELDLMDALSRNKGVLVYME-----DQKQFDEILKAVDIAKEHN-RQLIMSE: 272
493987263 : -----QEQQFDELDLMDALSRNKGVLVYME-----DQKQFDEILKAVDIAKEHN-RQLIMSE: 272
647306738 : -----QEQQFDELDLMDALSRNKGVLVYME-----DQKQFDEILKAVDIAKEHN-RQLIMSE: 286
501578355 : -----QEQQFDELDLMDALSRNKGVLVYME-----DQKQFDEILKAVDIAKEHN-RQLIMSE: 241
504265218 : -----QEQQFDELDLMDALSRNKGVLVYME-----DQKQFDEILKAVDIAKEHN-RQLIMSE: 272
568335214 : -----WEINQASHFYVSDAKRKN-VVHLE-----DQKQFDELDLMDALSRNKGVLVYME: 297
658466537 : -----WEINQASHFYVSDAKRKN-VVHLE-----DQKQFDELDLMDALSRNKGVLVYME: 273
428131024 : -----WEINQASHFYVSDAKRKN-VVHLE-----DQKQFDELDLMDALSRNKGVLVYME: 297
490201866 : -----WEINQASHFYVSDAKRKN-VVHLE-----DQKQFDELDLMDALSRNKGVLVYME: 294
500776899 : -----WEINQASHFYVSDAKRKN-VVHLE-----DQKQFDELDLMDALSRNKGVLVYME: 275

```

i w p f GE

```

MtGlu5 : -----*-----520-----*-----540-----*-----560-----*-----: 470
503328546 : -----FGAYSR-GLDSRVVWVWVRSRSETLKRNFSAWYWEAAG-FGINRTTRQNRTELRKLVPEQP-----: 470
503325617 : -----FGAYSR-GLDSRVVWVWVRSRSETLKRNFSAWYWEAAG-FGINYDRKARREPROLLESTSLDITNPSRP: 350
503546037 : -----FGAYSR-GLDSRVVWVWVRSRSETLKRNFSAWYWEAAG-FGINNAQTHVWVPELTLALIPALP: 392
501542033 : -----FGAYSR-GLDSRVVWVWVRSRSETLKRNFSAWYWEAAG-FGAFVPLNCRREPLQCLVLPSE: 333
493987263 : -----FGAYSR-GLDSRVVWVWVRSRSETLKRNFSAWYWEAAG-FGAFVPLNCRREPLQCLVLPSE: 335
647306738 : -----FGAYSR-GLDSRVVWVWVRSRSETLKRNFSAWYWEAAG-FGAFVPLNCRREPLQCLVLPSE: 346
501578355 : -----FGAYSR-GLDSRVVWVWVRSRSETLKRNFSAWYWEAAG-FGAFVPLNCRREPLQCLVLPSE: 301
504265218 : -----FGAYSR-GLDSRVVWVWVRSRSETLKRNFSAWYWEAAG-FGAFVPLNCRREPLQCLVLPSE: 335
568335214 : -----FGAYSR-GLDSRVVWVWVRSRSETLKRNFSAWYWEAAG-FGAFVPLNCRREPLQCLVLPSE: 360
658466537 : -----FGAYSR-GLDSRVVWVWVRSRSETLKRNFSAWYWEAAG-FGAFVPLNCRREPLQCLVLPSE: 337
428131024 : -----FGAYSR-GLDSRVVWVWVRSRSETLKRNFSAWYWEAAG-FGAFVPLNCRREPLQCLVLPSE: 357
490201866 : -----FGAYSR-GLDSRVVWVWVRSRSETLKRNFSAWYWEAAG-FGAFVPLNCRREPLQCLVLPSE: 354
500776899 : -----FGAYSR-GLDSRVVWVWVRSRSETLKRNFSAWYWEAAG-FGAFVPLNCRREPLQCLVLPSE: 335

```

FGAYsK ad SR WT V ae ayWEfc G FG yd w pL a

MtGlu5: *Miothermus taiwanensis* WR220

503328546: *Isosphaera pallida*

503325617: *Anaerolinea thermophila*

503546037: *Mahella australiensis*

501542033: *Dictyoglomus thermophilum*

493987263: *Caldithrix abyssi*

647306738: *Dictyoglomus turgidum*

501578355: *Caldicoprobacter oshimai*

504265218: *Fervidobacterium pennivorans*

568335214: *Thermotoga maritima*

658466537: *Thermotogae bacterium JGI 0000106-O11*

428131024: *Fervidobacterium gondwanense*

490201866: *Thermosipho africanus*

500776899: *Fervidobacterium nodosum*

Figure S1 Multiple-sequence alignment of *MtGlu5* compared with other GH5 cellulases.

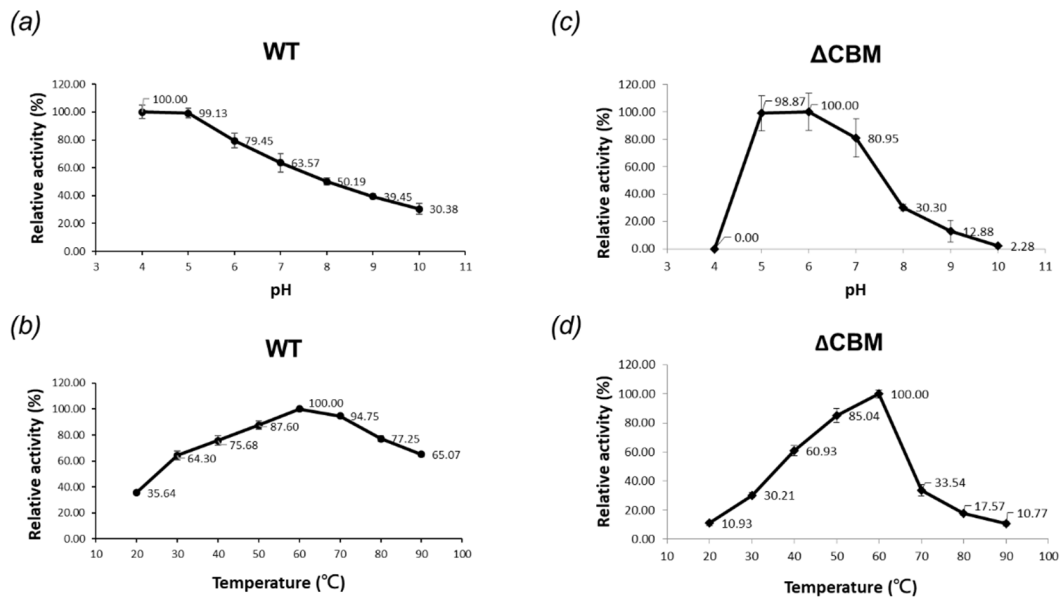


Figure S2 The effect of reaction pH and temperature on enzyme activity on CMC. Data are shown as the means \pm SD of more than three replicates. Both WT (*a, b*) and Δ CBM (*c, d*) present an optimum condition at 60 °C and pH 5.

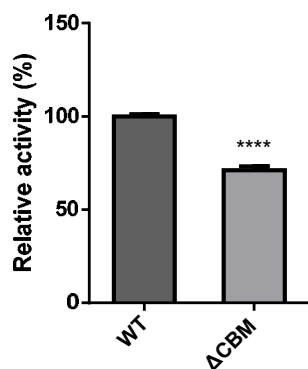


Figure S3 The activity comparison between *MtGlu5* (WT) and Δ CBM toward CMC. The specific activities of *MtGlu5* and Δ CBM toward CMC are 385 ± 5 IU and 275 ± 8 IU. Data are exhibited as the means \pm SD of more than three replicates. **** indicates statistically significant at the level of $P < 0.0001$ compared to WT.

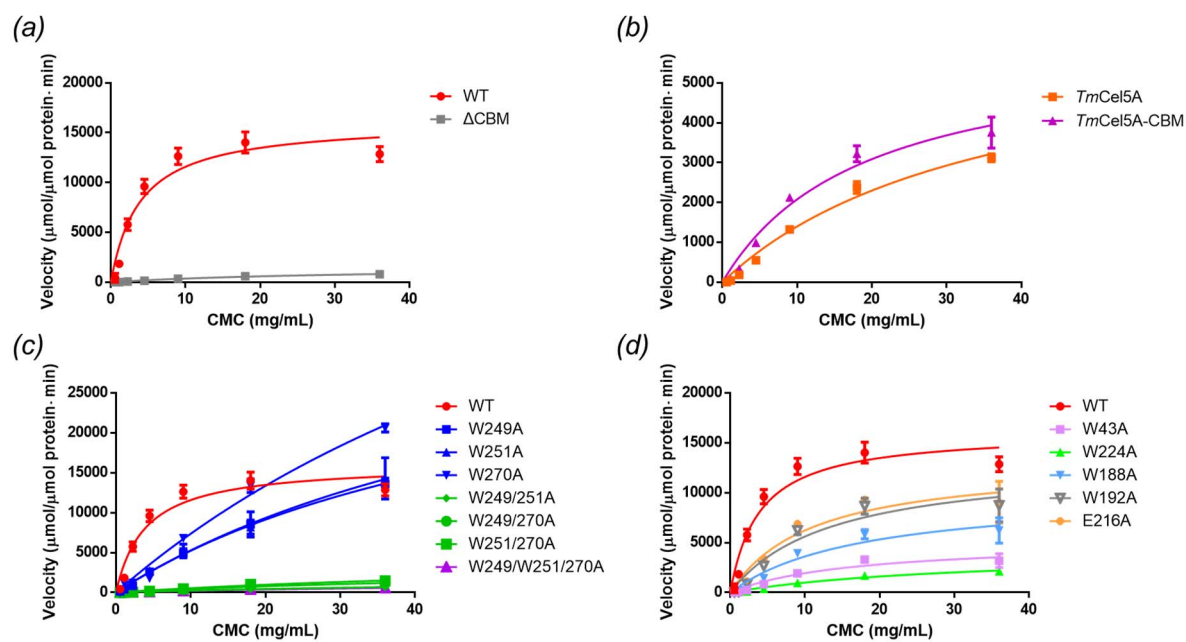


Figure S4 Reaction kinetics curves toward CMC. Comparison of the reaction kinetics curves among *MtGlu5* (WT) with ΔCBM (a), *TmCel5A* with *TmCel5A*-CBM (b), WT with Trp mutagenesis in CBM (c), and WT with Trp mutagenesis in the catalytic domain (d), respectively.

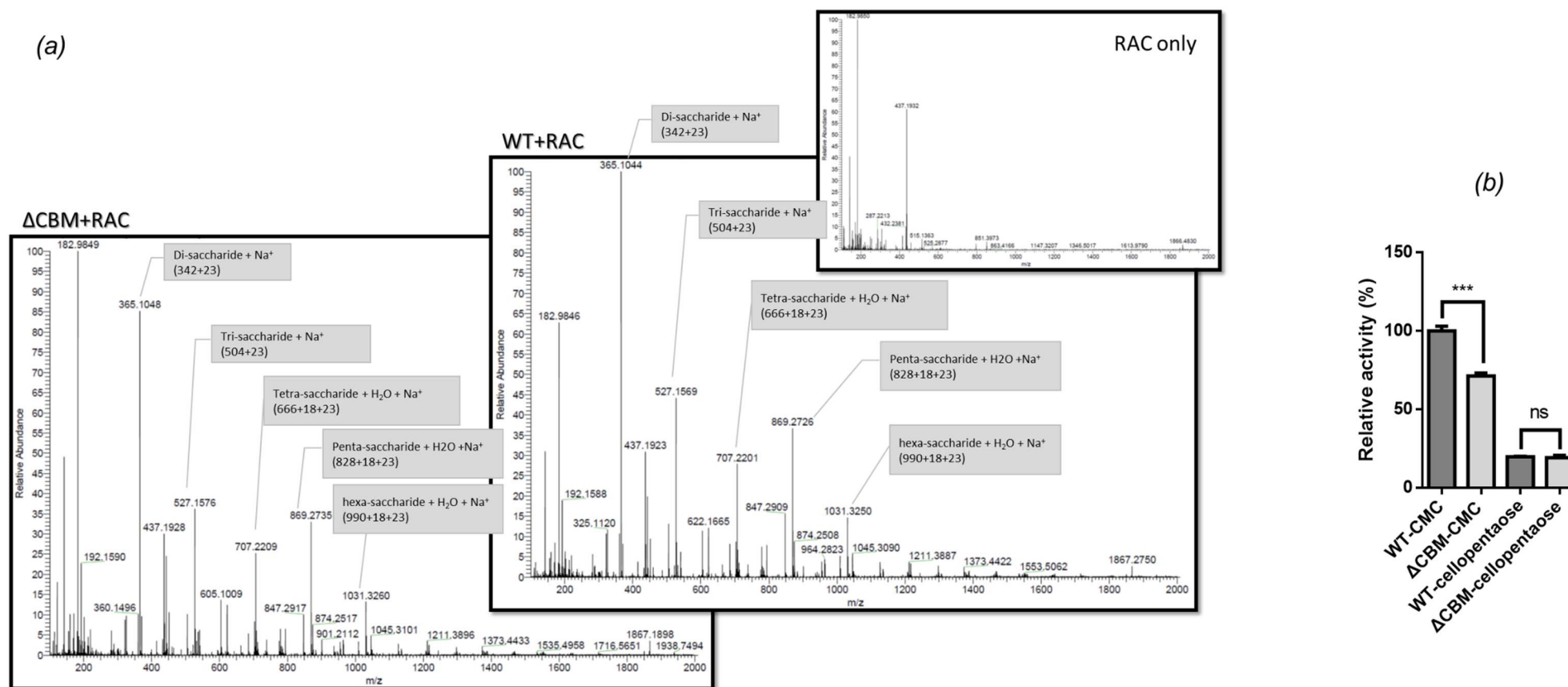


Figure S5 The catalytic centre integrity of WT and Δ CBM. (a) Hydrolyzed products are analyzed by mass spectrometry. Both WT and Δ CBM produce the same products after hydrolysis. m/z 182.98 and m/z 437.19 are considered contaminants in the background. Two enzymes were mixed with RAC and reacted for 24 hrs. The reactions were stopped by boiling. Then the pellet was removed, and hydrolyzed products remained in the supernatant fraction. Desalting was necessary before MS analysis. (b) Enzyme activities toward CMC and cellopentaose. Catalysis that would not use the CBM domain is demonstrated by using

cellopentaose as a substrate. The data show that the enzyme activity of *MtGlu5* (WT) and Δ CBM toward cellopentaose are almost the same, although lower than the activity toward CMC, indicating CDs are not perturbed by the CBM truncation. (Left-half panel) Enzyme activity toward CMC, *** indicates statistically significant at the level of $P < 0.001$ compared to *MtGlu5*. (Right-half panel) Enzyme activity toward cellopentaose, ns represents no statistical significance as compared to *MtGlu*

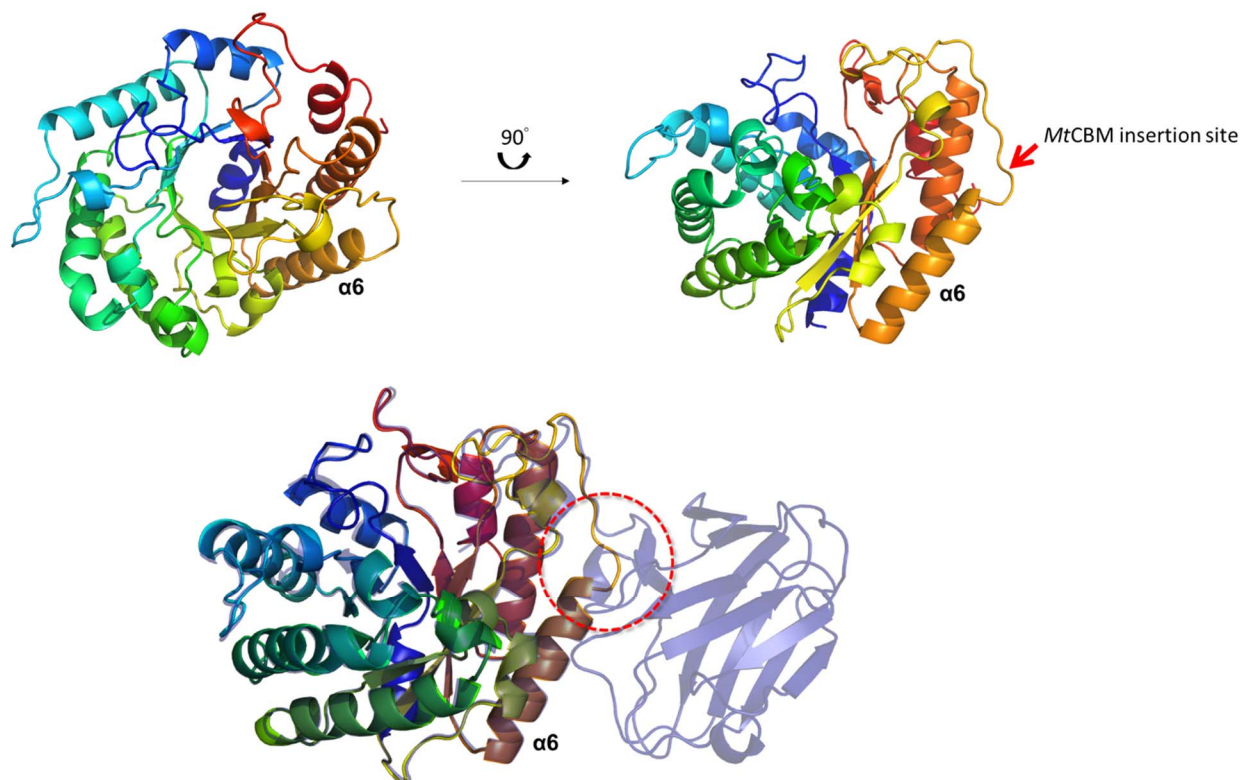


Figure S6 The structure of Δ CBM ramped from the N-terminus (blue) to the C-terminus (red) (PDB: 7VT5). The apo-form Δ CBM structure displayed an intact TIM-barrel fold. The red arrow indicates the location of *Mt*CBM insertion site (a.a 235 of *Mt*Glu5). Superimposition of *Mt*Glu5 and Δ CBM highlights the slight change in helix α 6.

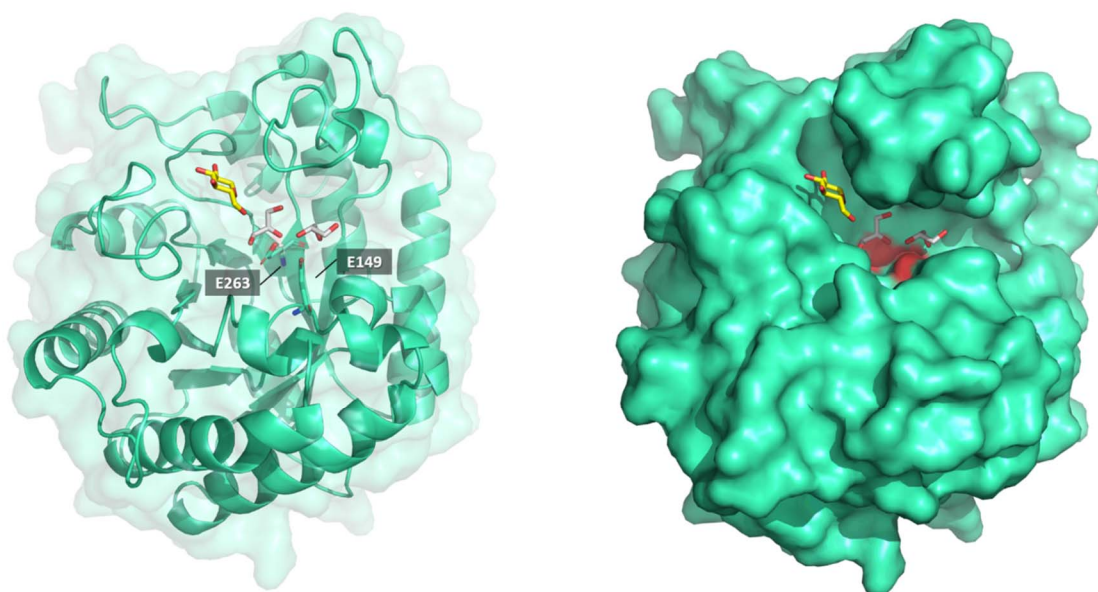


Figure S7 The structure of Δ CBM in complex with glucose (PDB: 7VT6). Glucose is shown in a stick form (yellow) whereas two glycerol molecules that occupied the catalytic site are shown in white sticks.

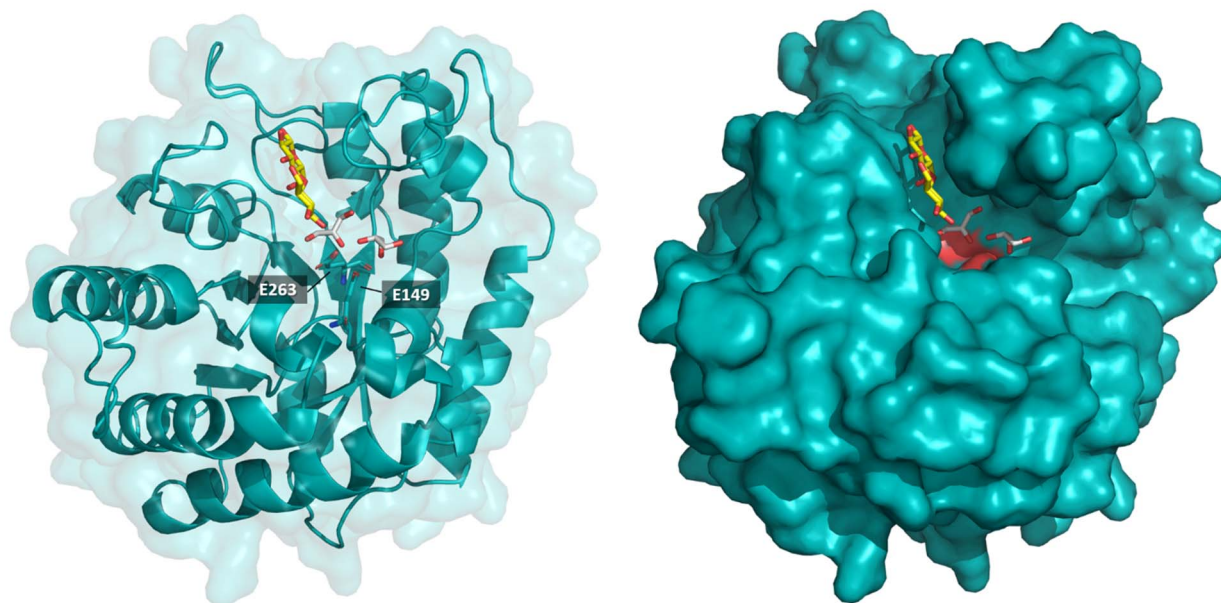


Figure S8 The structure of Δ CBM in complex with cellobiose (PDB: 7VT7). Cellobiose is shown in a stick form (yellow) whereas two glycerol molecules that occupied the catalytic site are shown in white sticks.

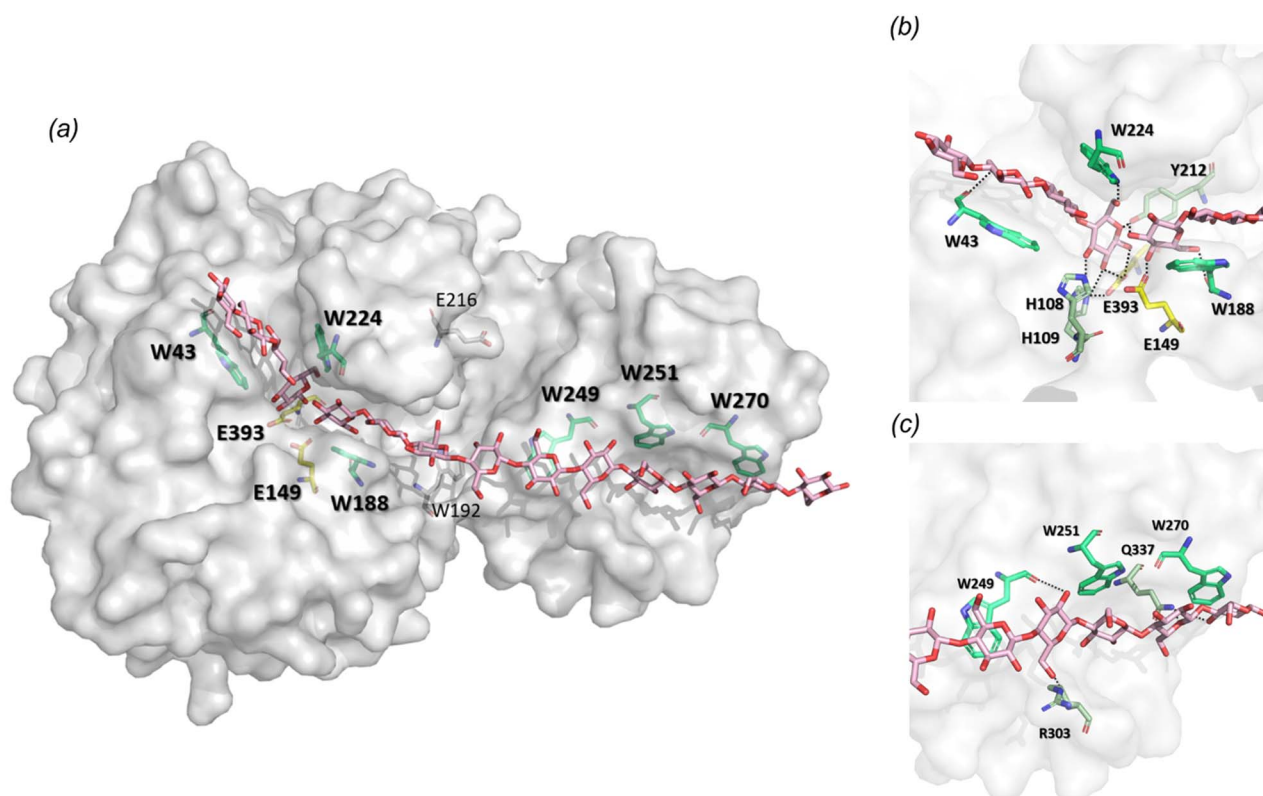


Figure S9 Substrate binding model of *MtGlu5* depicted in surface structure. (a) A 14-sugar-long single polysaccharide chain, shown in stick format (pink), is bound along the putative binding groove full of Trp residues (lime green) on the cleft surface. In this model, the polysaccharide chain is stabilized by the arrangement of the aromatic residues in the cleft of *MtGlu5*. The catalytic residues, E149 and E393, are shown in yellow and stick representation while the less essential residues, W192 and E216, are shown in white. The binding model is superimposed onto *TmCel5A* in complex with cellotetraose (PDB: 3azt) and CBM29-2 in complex with cellohexaose (PDB: 1W8T) then 14-polysaccharides are put in by *coot*. Energy minimization is simulated by the software YASARA Energy Minimization server. Snapshots of hydrogen bonds (black dashed lines) at catalytic pocket (b) and CBM binding surface (c). The surface Trp residues are shown as described and the other active-site residues are shown in olive green. (b) In the catalytic pocket, W43, W224, and W188 show not only stacking interaction but a direct hydrogen bond to the substrate. (c) W249, R303, and Q337 residues make direct interactions with the substrate in the same way as CBM29-2.

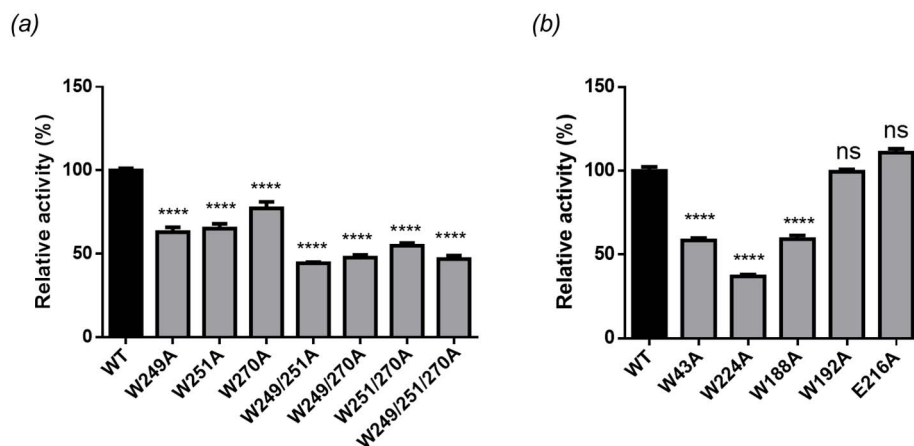


Figure S10 The activity comparison between *MtGlu5* (WT) and Trp mutants on CMC. Data are exhibited as the means \pm SD of more than three replicates. **** indicates statistical significance at the level of $P < 0.0001$ compared to WT, and *ns* is no significant differences compared to WT, respectively. (a) Activity comparison among WT and CBM mutants. (b) Activity comparison among WT and CD mutants.

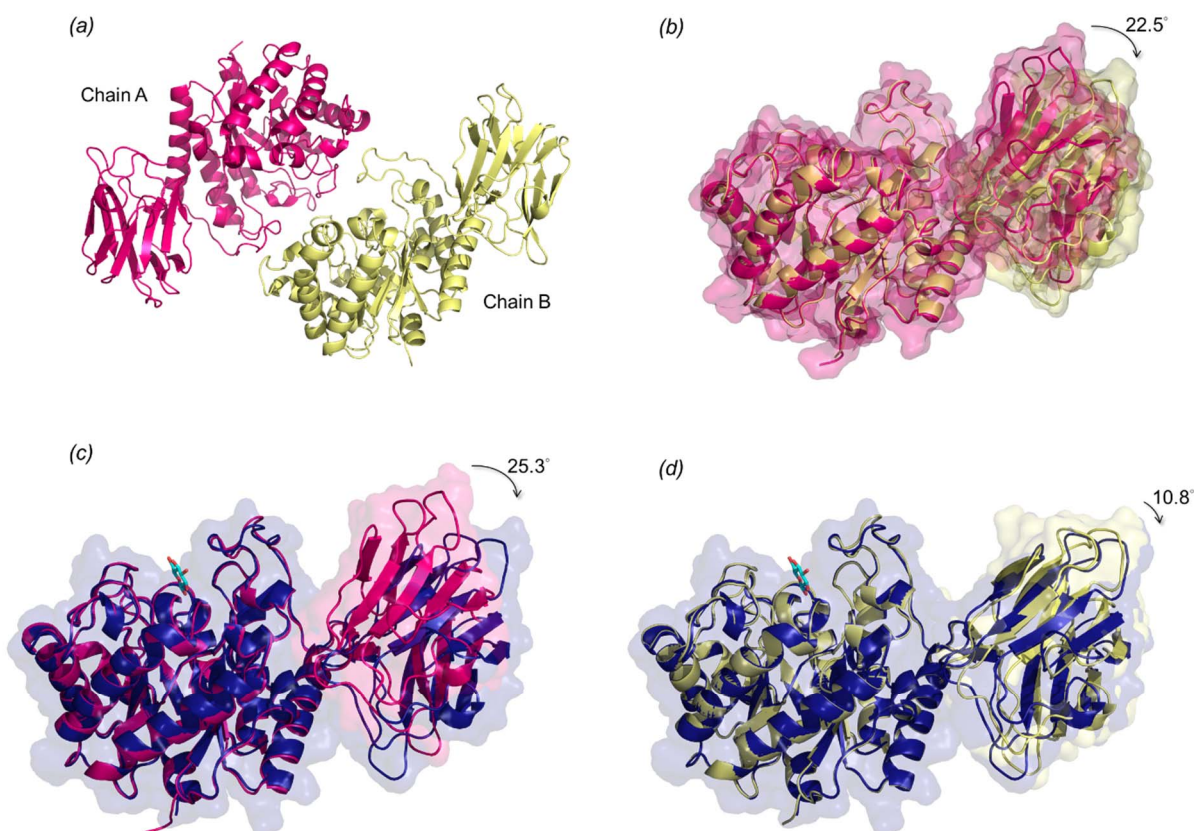


Figure S11 Structure comparison of *iMtGlu5* (PDB: 7VT4) with *MtGlu5* (PDB: 7VT8). (a) crystal structure of *iMtGlu5* shows dimer in the asymmetric unit. (b) By overlaying two chains of *iMtGlu5*, the CBMs of the two *iMtGlu5* molecules differ by a rigid-body rotation of 22.5°. (c) Superimposition of glucose-bound *MtGlu5* and *iMtGlu5* chain A and chain B (d).

Table S1 List of primers.

Construct	Primers (5' to 3')
<i>MtGlu5</i>	F1: TATGATGGGCTGCCAATCCACC F2: TGATGGGCTGCCAATCCACC R1: AGCTTCGGTTGCTCCGGTACAA R2: TCGGTTGCTCCGGTACAA
<i>iMtGlu5</i> Δ CBM	F1: TCGTGGGGCAGTTCGGAGCCTAC F2: CCTACGAGAAGGGCGATCT R1: CTCCGAACTGCCCCACGAAGA R2: AGATGGGGCGCCGGTTCT
W249A <i>iW249A</i>	F1: CAGAACGCATCCTGGGGCAGCCGG F2: AGCCGGGTCGGTTTTGTGGG R1: GCCCCAGGATGCGTTCTGCCAGCCCG R2: CCAGCCCGCCGCGAAGGC
W251A <i>iW251A</i>	F1: TGGTCCGCAGGCAGCCGGGTCG F2: GTCGGTTTTGTGGGTGAAGCT R1: CCGGCTGCCTGCGGACCAGT R2: GTTCTGCCAGCCCGCCGC
W270A <i>iW270A</i>	F1: AAGGCGCAGCCGGTTCTATTTG F2: ATTTGCACTCGGACGCTGGG R1: AGAACCCGGCTGCGCCTTCCTGGTA R2: CCTGGTAGGTGATCTCTAAAGCT
W249/251A <i>iW249/251A</i>	F1: TCCGCAGGCAGCCGGTTCGGTTTTGT F2: GTCGGTTTTGTGGGTGAAGCTT R1: CCGGCTGCCTGCGGATGCGTTCTGCCAGC R2: TCGTTCCTGCCAGCCCGCCGCGAA
W249/270A <i>iW249/270A</i>	Same as W249A using W270 as a template
W251/270A	Same as W251A using W270 as a template

iW251/270A	
W249/251/270A iW249/251/270A	Same as W249/251A using W270 as a template
W43A iW43A	F1: GGGGCCGCAGGGGTCAGGCT F2: CTGGAGGAAGGGTTTTTCGAGCTG R1: CCTGACCCCTGCGGCCCTTCCC R2: TTCCAGGGGGCCTCGAGGGC
W188A iW188A	F1: GTGGGTGCCAACTCGCTGTGGCGGCTG F2: TGGCGGCTGTCCGAGCTGCG R1: CAGCGAGTTGGCACCCACCGCCCC R2: CGGCCCCACAATCACCGCCC
W192A iW192A	F1: TCGCTGGCCCGGCTGTCCGAGCTG F2: GAGCTGCGGCTGCCGGACGAT R1: GGACAGCCGGGCCAGCGAGTTCCAAC R2: GTTCCAACCCACCGGCCCCAC
E216A iE216A	F1: CCCCTGGCGTTCACCCACCAGG F2: CAGGGGGCTGAGTGGCTC R1: GTGGGTGAACGCCAGGGGATCGTA R2: ATCGTAGTAGTGAAGGTAACGA
W224A iW224A	F1: GCTGAGGCTCTCAACCCGTTCCACC F2: GTTCCACCTACCGGTGTGGTCTGG R1: GGGGTTGAGAGCCTCAGCCCCCTG R2: CCCCTGGTGGGTGAACTCCAGGGG
Δ CBM	F1: CACCAGCAAACGCCATAGC F2: GCCCAGGCCATGGAGTTT R1: TATGGCGTTTTGCTGGTGCCAGACCACACCGGTAGG R2: CCAGACCACACCGGTAGG
<i>gMr</i> CBM	F1: TATGTCCCCTATACTAGGTTATTGG F2: TGTCCCCTATACTAGGTTATTGG R1: AGCTTGGTCAGCAAAGCCAGCGG

	<p>R2: TGGTCAGCAAAGCCAGCGG</p> <p>GST-F: TGTCCCCTATACTAGGTTATTGG</p> <p>GST-R: CACCCTGGTTCTCACGGGATCCACGCGGAACCAG</p> <p>CBM-F: CGTGAGAACCAGGGTGCCTTC</p> <p>CBM-R: TGGTCAGCAAAGCCAGCGG</p>
GST	<p>F1: TATGATGTCCCCTATACTAGGTT</p> <p>F2: TGATGTCCCCTATACTAGGTT</p> <p>R1: AGCTTTTTTGGAGGATGGTCGCC</p> <p>R2: TTTTGGAGGATGGTCGCC</p>
<i>Tm</i> Cel5A	<p>F1: TATGATGGGTGTGGATCCGTTTC</p> <p>F2: TGATGGGTGTGGATCCGTTTC</p> <p>R1: AGCTTTTCGATGCTGTCGCCACC</p> <p>R2: TTTCGATGCTGTCGCCACC</p>
<i>Tm</i> Cel5A-CBM	<p>F1: TATGATGGGTGTGGATCCGTTTC</p> <p>F2: TGATGGGTGTGGATCCGTTTC</p> <p>R1: AGCTTTTCGATGCTGTCGCCACC</p> <p>R2: TTTCGATGCTGTCGCCACC</p> <p>Cel5A-F1: TGATGGGTGTGGATCCGTTTC</p> <p>Cel5A-R1: GCCCATTACGACCCAGCCACTTCTC</p> <p>CBM-F: CTGGGTCGTAAATGGGGCCGTGAGAACCAGGGTGCCTTC</p> <p>CBM-R: CTTTTGATCGTCCGGGCTGGTCAGCAAAGCCAGCGG</p> <p>Cel5A-F2: AGCCCGGACGATCAAAAGCACCTGA</p> <p>Cel5A-R2: TTTCGATGCTGTCGCCACC</p>
<i>iTm</i> Cel5A <i>iTm</i> Cel5A-CBM	<p>F1: ATCGGCCAATTCGGCGCGTATC</p> <p>F2: TATCGTAAGGCGGATCTGGAAAGCC</p> <p>R1: CGCGCCGAATTGGCCGATGTAAATC</p> <p>R2: GTAAATCGGACGTTTGTTCCTTGC</p>

Table S2 The processivity of *MtGlu5* towards RAC.

	Enzyme	Reducing sugar (%)		Ratio
		Supernatant	Insoluble	Sol./Insol.
RAC	<i>MtGlu5</i>	72.6	27.4	2.65
	Δ CBM	62.5	37.5	1.67

The processivity is presented as the ratio of the reducing ends in the soluble part to the one in the insoluble part after hydrolysis. We tested the classic methods to verify the processivity by determining the ratio of soluble reducing ends to insoluble reducing sugar ends products using both RAC and filter paper (Data not shown) as substrates, both *MtGlu5* and Δ CBM present the ratio of around 2.65 and 1.67, by the criterion exhibited in the paper (Irwin *et al.*, 1993). Therefore *MtGlu5* is clearly an endo-glucanase.