CaXyn30B from the solventogenic bacterium Clostridium acetobutylicum is a glucuronic acid-dependent endoxylanase

Casey Crooks1*, Liangkun Long1,2 and Franz J. St John1

Abstract
Objective: We previously described the structure and activity of a glycoside hydrolase family 30 subfamily 8 (GH30-8) endoxylanase, CaXyn30A, from Clostridium acetobutylicum which exhibited novel glucuronic acid (GA)-independent activity. Immediately downstream from CaXyn30A is encoded another GH30-8 enzyme, CaXyn30B. While CaXyn30A deviated substantially in the highly conserved β7-α7 and β8-α8 loop regions of the catalytic cleft which are responsible for GA-dependence, CaXyn30B maintains these conserved subfamily 8 amino acid residues thus predicting canonical GA-dependent activity. In this report, we show that CaXyn30B functions as a canonical GA-dependent GH30-8 endoxylanase in contrast to its GA-independent neighbor, CaXyn30A.

Results: A clone expressing the catalytic domain of CaXyn30B (CaXyn30B-CD) exhibited GA-dependent endoxylanase activity. Digestion of glucuronoxylan generated a ladder of aldouronate limit products as anticipated for canonical GA-dependent GH30-8 enzymes. Unlike the previously described CaXyn30A-CD, CaXyn30B-CD showed no activity on arabinoxylan or the generation of appreciable neutral oligosaccharides from glucuronoxylan substrates. These results are consistent with amino acid sequence comparisons of the catalytic cleft and phylogenetic analysis.

Keywords: Glycoside hydrolase, Xylanase, GH30, Glucuronoxylan, Bioconversion, Biorefinery

Introduction
Xylan represents the most abundant form of hemicellulose and consists of β-(1,4)-linked d-xylose units. Effective use of xylan in applications including production of renewable fuels, green chemicals and nutraceuticals would be greatly facilitated by the availability of xylanases with well-defined functions that enable the generation of specific product streams [1, 2]. Depending on the xylan source, the main xylan chain may be decorated with acetyl, α-L-arabinofuranose or 4-O-methylα-d-glucuronic acid substitutions. The hydrolysis of xylan by endoxylanases is significantly influenced by the xylan substitution characteristics. For strict endoxylanases, more xylan chain substitutions typically result in lower observed activity. GH30-8 endoxylanases are dual domain enzymes consisting of a (β/α)8-barrel containing the catalytic determinants with an obligatory stabilizing side β-sandwich structure tightly associated with the (β/α)8-barrel catalytic core through hydrophobic contacts [3-5]. Canonical GH30-8 enzymes exhibit a distinct specificity by requiring the recognition of GA substitutions at the catalytic-2b subsite for hydrolysis [5, 6]. In part, this specificity is provided through a salt bridge interaction between the C6 carboxylate of GA and the guanidinium side-chain of a highly conserved arginine in the catalytic substrate binding cleft [6-8]. This arginine and surrounding motif (RR-motif) in the β8-α8 loop region of these enzymes appears critical in determining

© The Author(s) 2020. This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/. The Creative Commons Public Domain Dedication waiver (http://creativecommons.org/publicdomain/zero/1.0/) applies to the data made available in this article, unless otherwise stated in a credit line to the data.
the substrate specificity of canonical GA-dependent GH30-8 enzymes (Fig. 1). We recently described a novel GH30-8 from *Clostridium acetobutylicum*, CaXyn30A (UniProt ID: Q97T12), that lacks the highly conserved RR-motif and found it exhibited a novel GA-independent activity [9]. Encoded immediately downstream as part of a predicted three gene operon from the non-canonical, GA-independent, CaXyn30A is CaXyn30B (UniProt ID: Q97T13) (Fig. 1). Phylogenetic and structural predictions and the conserved RR-motif of CaXyn30B suggest that this enzyme will exhibit canonical GH30-8 GA-dependent endoxylanase activity exemplified by the previously described GH30-8 endoxylanase from *Bacillus subtilis*, BsXynC, or the gram negative representative, EcXynA, from *Erwinia chrysanthemi* [5, 6]. Here we describe the biophysical functional parameters of this genetically linked GH30-8 enzyme and confirm that it functions primarily as a canonical GH30-8 GA-dependent endoxylanase.

**Main text**

**Methods**

**Chemicals**

All general use chemicals were at least ACS grade. Molecular biology reagents were purchased from New England Biolabs (Ipswich, MA). For xylanase assays beechwood xylan (BX) was obtained from Sigma (St. Louis, MO), wheat arabinoxylan (WAX) and xylohexaose were obtained from Megazyme International (Wicklow, Ireland). Sweetgum wood xylan (SGX) was obtained from the laboratory of Dr. James F. Preston from the University of Florida. The aldotetrauronic acid (GX3) was prepared...
by digesting BX to completion with the GH10 endoxtylanase, PbXyn10A1CD [10, 11], and purifying using preparative TLC. Using the same TLC solvent system and detection described below for analytical studies, preparative TLC was performed using 20 cm x 20 cm, 500 µm thick silica gel H (no binder) plates. The PbXyn10A1CD BX hydrolysate was spotted in bands using the BandIT preparative TLC sample applicator from Miles Scientific (Newark, DE). The plate was developed with two ascension and the sides were removed and developed to guide sample recovery. GX3 migration was identified using GXn standards and was isolated by washing the recovered silica in excess water. This volume was lyophilized, redissolved in water and the volume used in assays was determined empirically by comparison to quantified xylooligosaccharides.

**Cloning of CaXyn30B**  
CaXyn30B (UniProt ID: Q97TI3), encoded on the pSOL1 megaplasmid, was isolated from genomic DNA obtained from *Clostridium acetobutylicum* ATCC 824. PCR amplification of the catalytic domain of CaXyn30B was performed using primers 5'-cttaCCATGctcacttgtaggttagattagcct-c'3' and 5'-ctctacaTCCAAGttgtagtagctgatnacc-c'3' and cloned into pET28b (Novagen) between the Ncol and Xhol sites using the sites as indicated with uppercase in the primers shown. The resulting construct was sequence verified and encodes a 402 amino acid C-terminal hexahistidine tagged expression product, CaXyn30B-CD.

**Protein expression of CaXyn30B-CD**  
CaXyn30B-CD was expressed in *E. coli* using a modified auto-induction method derived from Studier et al. [12]. Briefly, the pET28-CaXyn30B-CD plasmid was transformed into Rosetta 2(DE3) cells and maintained under selection with 50 µg/ml kanamycin and 34 µg/ml chloramphenicol. A single colony was propagated in non-selective induction media (50 mM Na₂HPO₄, 50 mM KH₂PO₄, 50 mM NH₄Cl, 5 mM Na₂SO₄, 2 mM MgSO₄, 0.5 × trace metals mix (Technova), 0.5% glucose, 0.25% sodium aspartate) grown overnight at 37 °C and 250 RPM. A 1:150 inoculum was introduced into auto-induction media (0.5 × terrific broth, 50 mM NH₄Cl, 5 mM Na₂SO₄, 2 mM MgSO₄, 0.5 × trace metals mix, 0.25% sodium aspartate) and grown for 40 h at 18 °C, 250 RPM. Expression cultures were harvested, processed, and protein purified as previously described using Ni²⁺ immobilized metal affinity chromatography [11]. The eluted protein was desalted with two passages through Zeba 7 kDa size exclusion spin column (Thermo Fisher Scientific). Protein size was confirmed and purity was empirically estimated at >95% by SDS-PAGE [13] (not shown).

**Enzyme assays**  
Enzyme activity optimization and specific activity was determined using the Nelson’s test [14, 15] as previously described [8]. In CaXyn30B reactions, BX was used at 10 mg/ml and buffers at 30 mM. Reactions were initiated by addition of 200 ng of CaXyn30B. For pH optimization the reaction temperature was 40 °C. Buffers included sodium acetate for lower pH range, a sodium acetate/MES mixture for a middle pH range and sodium phosphate for a higher pH range. Following optimum reaction pH determination, optimum reaction temperature was determined using sodium acetate/MES buffer at pH 5.75. Specific activity was determined as described above, except that pH 6.0 acetate buffer was used instead of the acetate/MES two component buffer and the reaction was performed at 50 °C for 10 min. One unit is defined as the amount of enzyme which liberates 1 µmol of reducing terminal per minute.

**TLC and MALDI-TOF**  
Hydrolysis reaction products were analyzed by thin layer chromatography (TLC) and Matrix-Assisted Laser Desorption/Ionization Time of Flight Mass Spectrometry (MALDI-TOF MS). TLC and MALDI-TOF MS were performed as previously reported [8, 16]. Briefly, TLC reactions containing 10 mg/ml SGX, 7.5 mg/ml WAX, or 5 mM of the oligomeric substrates (GX3, X4, X5, X6) were digesting in 50 µl volumes (25 µl for GX3) with 20 µg/ml of CaXyn30B, 28 µg/ml BcXyn30A or 5 µg/ml CaXyn30A. CaXyn30B and BcXyn30A were incubated at 40 °C in 30 mM sodium acetate pH 6 buffer while CaXyn30A was incubated using 30 mM sodium acetate pH 4. All TLC reactions were digested for 30 min and then heat inactivated at 95 °C for 10 min and used directly for spotting on Analtech silica G plates spotting 5 µl of the reaction in 1 µl increments under warm air flow to allow drying between applications. Samples for MALDI-TOF were digested overnight using 3 µg/ml in 30 mM pH 6.0 sodium acetate buffer. Samples were heat killed as above and were decationized using Dowex 50WX4 hydrogen form. The sample was added to the MALDI-TOF matrix 2,5-dihydroxybenzoic acid, spotted on the MALDI-TOF plate and dried for analysis. Mass analysis was performed using a Shimadzu MegaTOF at the University of Wisconsin, Biochemistry Instrumentation Facility.

**Results**  
In a recent publication we characterized a confidently classified GH30 subfamily 8 endoxylanase (CaXyn30A,
Q97TI2) from *C. acetobutylicum* which showed strikingly divergent function by exhibiting GA-independent endoxylanase activity as evidenced by the processing of WAX, release of neutral oligosaccharides from xylan polymers, and activity on neutral xylooligosaccharides [9]. This was attributed to amino acid sequence changes in the β8-α8 and β7-α7 loop regions of the enzyme. Other previous studies made a distinction between two primary GH30-8 branches [5]. This occurs between those which derive from Gram-positive bacteria versus those from Gram-negative bacteria. Most interestingly, while CaXyn30A derives from a Gram-positive bacterial host, it is nevertheless of the Gram-negative GH30-8 subtype. The adjacent enzyme, CaXyn30B (Q97TI3) which is the focus of this report, does not have any significant divergent gram-negative type sequence [5], is of the Gram-positive subtype and is expected to be a canonical functioning glucuronoxylanase. Interestingly, primary sequence analysis comparisons of CaXyn30B with CaXyn30A show that overall sequence homology is very high at about 62% identity for the mature protein (Fig. 1). This is unexpected considering their separate grouping into the Gram-type distinguished GH30-8s [5], which typically have identities less than 40%. More detailed analysis reveals that the C-terminal carbohydrate binding module (CBM, CBM13) which both enzymes share is a primary source of the high identities (i.e. 93% identity for this separate domain), such that exclusion of the CBM13 module results in a 51% identity for the remaining protein. This high homology also extends into the catalytic domain stabilizing side beta-structure. Comparisons of the sequence representing only the specificity determining (β/α)₈ TIM-barrel domain of the GH30 enzyme [4, 5, 17] yields a lower level of sequence identity of 37%, as expected for enzymes of different Gram-type grouping.

CaXyn30B-CD demonstrated GA-dependent activity on xylan substrates. MADLI-TOF MS analysis of glucuronoxylan digested by CaXyn30B liberated aldouronates as predicted for GA-dependent GH30-8 enzymes and lacked the accumulation of appreciable amounts of neutral xylooligosaccharides (Fig. 2). TLC analysis of digestion of model xylans showed CaXyn30B to be active on glucuronoxylan (SGX) but not arabinoxylan (WAX), similar to that observed for the defined GA-dependent endoxylanase, *BsXynC*, but unlike that observed for the previously described GA-independent enzyme, CaXyn30A, which also showed activity on arabinoxylan (Fig. 3). Digestion of GX3 yielded GX2 plus xylose, while digestion of xylotetraose showed no detectable activity, confirming the coordination of glucuronic acid in the catalytic-2b subsite (Fig. 3) [5, 8]. More activity than expected was observed in the hydrolysis of xylohexaose by CaXyn30B (Fig. 2). Previous reports of the hydrolysis rate difference between polymeric glucuronoxylan and neutral xylooligosaccharides (typically exemplified by X6) were estimated at three orders lower for the unsubstituted linear xylooligosaccharide [6, 18, 19]. From our studies we estimate that CaXyn30B while still being predominantly GA-dependent, has a greater ability to hydrolyze neutral xylooligosaccharides than previously described GH30-8 GA-dependent endoxylanases. From Nelson’s tests studies using X6 as substrate we estimate that CaXyn30B hydrolyzes X6 only 2-orders lower than a glucuronoxylan. Nevertheless, these results all confirm that CaXyn30B-CD demonstrated GA-dependent activity on xylan substrates. MADLI-TOF MS analysis of glucuronoxylan digested by CaXyn30B liberated aldouronates as predicted for GA-dependent GH30-8 enzymes and lacked the accumulation of appreciable amounts of neutral xylooligosaccharides (Fig. 2). TLC analysis of digestion of model xylans showed CaXyn30B to be active on glucuronoxylan (SGX) but not arabinoxylan (WAX), similar to that observed for the defined GA-dependent endoxylanase, *BsXynC*, but unlike that observed for the previously described GA-independent enzyme, CaXyn30A, which also showed activity on arabinoxylan (Fig. 3). Digestion of GX3 yielded GX2 plus xylose, while digestion of xylotetraose showed no detectable activity, confirming the coordination of glucuronic acid in the catalytic-2b subsite (Fig. 3) [5, 8]. More activity than expected was observed in the hydrolysis of xylohexaose by CaXyn30B (Fig. 2). Previous reports of the hydrolysis rate difference between polymeric glucuronoxylan and neutral xylooligosaccharides (typically exemplified by X6) were estimated at three orders lower for the unsubstituted linear xylooligosaccharide [6, 18, 19]. From our studies we estimate that CaXyn30B while still being predominantly GA-dependent, has a greater ability to hydrolyze neutral xylooligosaccharides than previously described GH30-8 GA-dependent endoxylanases. From Nelson’s tests studies using X6 as substrate we estimate that CaXyn30B hydrolyzes X6 only 2-orders lower than a glucuronoxylan. Nevertheless, these results all confirm

![Fig. 2 MADLI-TOF results identify the anticipated series of aldouronates each which differs by a single anhydroxylose (132 Da) in mass. The primary peak for each cluster is the single sodium adduct, [GXn-Na]+.](image-url)
the anticipated function CaXyn30B as a GA-dependent GH30-8 endoxylanase. The pH optimum of CaXyn30B-CD was determined under three buffer systems spanning pH 3.5–8.0 and found to be pH 5.75 (Additional file 1) in contrast to the linked CaXyn30A which had pH optimum of 4.0 [9]. Temperature optimum was determined at pH 5.75 and found to be 55 °C for a 10 min reaction. The specific activity of CaXyn30B measured on BX was approximately 14 U/mg CaXyn30B (data not shown).

**Discussion**

CaXyn30B is 62% identical to the genetically adjacent previously described endoxylanase, CaXyn30A. Such high degree of homology would suggest conserved functional specificity but the homology is asymmetrically placed. Specifically, the C-terminal Ricin-like CBM13 domain is 93% identical while that of the catalytic domain shows 51% identity. Moreover, identity in the catalytic domain is strongest in the stabilizing side beta-structure that when virtually removed reduces identity of the specificity determining regions of the core TIM-barrel of the catalytic domain to 37%. It is only with consideration of conservation within the β7-α7 and β8-α8 loop regions defining the catalytic cleft including conservation of the RR-motif that one would anticipate the function of CaXyn30B to more closely resemble that of the previously described GA-dependent GH30-8 enzyme from *B. subtilis*, BsXynC, than that of the recently described CaXyn30A.

Such high homology, and it being observed at the nucleic acid level (not shown), is suggestive of a gene duplication event with subsequent diversion in the specificity determining regions. Similarly the high conservation of the CBM13 module suggests similar substrate association albeit with different catalytic specificities. The relative contribution of these two linked enzymes in xylan depolymerization requires further study.

**Limitations**

CaXyn30B was tested on a limited number of substrates and conditions were evaluated and the potential for auxiliary activity in addition to the observed GA-dependent endoxylanase cannot be ruled out.

**Supplementary information**

Supplementary information accompanies this paper at https://doi.org/10.1186/s13104-020-05091-5.

Additional file 1. Optimum reaction conditions determination for CaXyn30B. A) The dependence of activity on pH showing three overlapping alternative buffer compositions. B) Optimum reaction temperature determined for a 10 min reaction period.

**Abbreviations**

Ca: Clostridium acetobutylicum; GH: Glycosyl hydrolase; GA: Glucuronic acid; CBM: Carbohydrate binding module; Bs: Bacillus subtilis; WAX: Wheat arabinoxylan; BX: Beechwood xylan; SGX: Sweetgum xylan; Ct: Clostridium thermocellum; Cp: Clostridium papyrosolvens; Ec: Erwinia chrysanthemi; Pb: Paenibacillus; GX3: Aldotetrauronic acid.
Acknowledgements
MALDI-TOF was performed at the University of Wisconsin Biotechnology Center.

Authors’ contributions
CC and FJSJ designed the study. CC generated CaXyn30B clone, expressed and purified protein. LL performed temperature and pH optimization. FJSJ performed TLC and MALDI-TOF experiments. CC generated original draft of the manuscript. All authors read and approved the final manuscript.

Funding
This study was internally funded by the USDA Forest Service, Forest Products Laboratory, USDA Forest Service, One Gifford Pinchot Drive, Madison, WI 53726, USA. 2 College of Chemical Engineering, Nanjing Forestry University, Laboratory, USDA Forest Service, One Gifford Pinchot Drive, Madison, WI 53726, USA.

Availability of data and materials
Any supporting data or materials are available from the corresponding author upon request.

Ethics approval and consent to participate
Not applicable.

Consent for publication
Not applicable.

Competing interests
The authors declare that they have no competing interests.

Author details
1 Institute for Microbial and Biochemical Technology, Forest Products Laboratory, USDA Forest Service, One Gifford Pinchot Drive, Madison, WI 53726, USA. 2 College of Chemical Engineering, Nanjing Forestry University, Nanjing 210037, China.

Received: 6 February 2020   Accepted: 16 May 2020
Published online: 10 June 2020

References
1. Aachary AA, Prapulla SG. Xylooligosaccharides (XOS) as an emerging prebiotic: microbial synthesis, utilization, structural characterization, bioactive properties, and applications. Compr Rev Food Sci Food Saf. 2011;10:2–16.
2. Dhiman S, Mukhejee G. Recent advances and industrial applications of microbial xylanases: a review. In: Gehlot P, Singh J, editors. Fungi and their role in sustainable development: current perspectives. Singapore: Springer; 2018. p. 329–48.
3. Valenzuela SV, Diaz P, Pastor FJ. Modular glucuronoxylan-specific xylanase with a family CBM35 carbohydrate-binding module. Appl Environ Microbiol. 2012;78:3923–31.

Publisher’s Note
Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Ready to submit your research? Choose BMC and benefit from:
• fast, convenient online submission
• thorough peer review by experienced researchers in your field
• rapid publication on acceptance
• support for research data, including large and complex data types
• gold Open Access which fosters wider collaboration and increased citations
• maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.
Learn more biomedcentral.com/submissions