Structure and Function of an Arabinoxylan-specific Xylanase*

Márcia A. S. Correia 1, Koushik Mazumder 1, Joana L. A. Brás 8, Susan J. Firbank 2, Yanping Zhu 5, Richard J. Lewis 8, William S. York 9, Carlos M. G. A. Fontes 12, and Harry J. Gilbert 3

Received for publication, December 29, 2010, and in revised form, March 4, 2011. Published, JBC Papers in Press, March 4, 2011, DOI 10.1074/jbc.M110.217315

THE JOURNAL OF BIOLOGICAL CHEMISTRY VOL. 286, NO. 25, pp. 22510 –22520, June 24, 2011 © 2011 by The American Society for Biochemistry and Molecular Biology, Inc. Printed in the U.S.A.

The enzymatic degradation of plant cell walls plays a central role in the carbon cycle and is of increasing environmental and industrial significance. The enzymes that catalyze this process include xylanases that degrade xylan, a β-1,4-xylose polymer that is decorated with various sugars. Although xylanases efficiently hydrolyze unsubstituted xylans, these enzymes are unable to access highly decorated forms of the polysaccharide, such as arabinoxylans that contain arabinofuranose decorations. Here, we show that a Clostridium thermocellum enzyme, designated CtXy15A, hydrolyzes arabinoxylans but does not attack unsubstituted xylans. Analysis of the reaction products generated by CtXy15A showed that all the oligosaccharides contain an O3 arabinose linked to the reducing end xylose. The crystal structure of the catalytic module (CtGH5) of CtXy15A, appended to a family 6 noncatalytic carbohydrate-binding module (CtCBM6), showed that CtGH5 displays a canonical (α/β)8-barrel fold with the substrate binding cleft running along the surface of the protein. The catalytic apparatus is housed in the center of the cleft. Adjacent to the –1 subsite is a pocket that could accommodate an L-arabinofuranose-linked α-1,3 to the active site xylose, which is likely to function as a key specificity determinant. CtCBM6, which adopts a β-sandwich fold, recognizes the termini of xylo- and gluco-configured oligosaccharides, consistent with the pocket topology displayed by the ligand-binding site. In contrast to typical modular glycoside hydrolases, there is an extensive hydrophobic interface between CtGH5 and CtCBM6, and thus the two modules cannot function as independent entities.

The plant cell wall, which is an important biological and industrial resource, primarily consists of interlocking polysaccharides (for review see Ref. 1). The biological conversion of the polysaccharides within the plant cell wall to their constituent monosaccharides is central to its biological and industrial exploitation (2, 3). An example of this chemical complexity is provided by xylan, which is the major hemicellulosic component of the wall. This polysaccharide includes a backbone of β-1,4-xylose residues in their pyranose configuration (Xylp), which are decorated at O2 with 4-O-methyl-D-glucuronic acid and at O2 and/or O3 with arabinofuranose (Araf) residues, whereas the polysaccharide can also be extensively acetylated. In addition, the Araf side chain decorations can also be esterified to ferulic acid that, in some species, provides a chemical link between hemicellulose and lignin (Fig. 1) (1). The precise structure of xylans varies between plant species, tissues, and during cellular differentiation (4).

Reflecting the chemical complexity of plant structural polysaccharides, microbial plant cell wall-degrading microorganisms express a large number of enzymes, often in excess of 100 biocatalysts, that target specific linkages within these carbohydrate polymers (5–8). The majority of plant cell wall-degrading enzymes are glycoside hydrolases, although polysaccharidases and carbohydrate esterases also contribute to the catalytic process. These enzymes are grouped into families based on sequence and structural and catalytic conservation within the CAZy data base (9). As discussed in the accompanying article (48), many of these enzymes are appended to noncatalytic carbohydrate-binding modules (CBMs) 1 that are also grouped into families on the CAZy data base. The xylan backbone is hydrolyzed by xylanases, the majority of which are located in glycoside hydrolase families 10 and 11, although they are also present in GH8 and GH30 (10, 11). The extensive decoration of the xylan backbone generally restricts the capacity of these enzymes to attack the polysaccharide prior to removal of the side chains (12).

Here, we report the biochemical properties and crystal structure of a GH5 enzyme that is appended to a family 6 CBM (CtCBM6). The enzyme (defined as CtXy15A) is an arabinoxylan-specific xylanase that utilizes Araf/decorations, appended to O3 of the Xylp bound at the active site, as an essential specificity for the reaction process. The amino acid sequence of CtXy15A, the crystal structure, and the biochemical studies suggest that this enzyme is an arabinoxylan-specific xylanase (CtXy15A).

The atomic coordinates and structure factors (code 2y8k) have been deposited in the Protein Data Bank, Research Collaboratory for Structural Bioinformatics, Rutgers University, New Brunswick, NJ (http://www.rcsb.org/).

1 These authors contributed equally to this work.
2 To whom correspondence may be addressed. E-mail: cfontes@fmv.utl.pt.
3 To whom correspondence may be addressed: Institute for Cell and Molecular Biosciences, Newcastle University, Medical School, Framlington Place, Newcastle upon Tyne NE2 4HH, United Kingdom. E-mail: h.j.gilbert@newcastle.ac.uk.

The abbreviations used are: CBM, carbohydrate-binding module; PDB, Protein Data Bank; DP, degrees of polymerization; EI, electron ionization; ESI, electrospray ionization; ITC, isothermal titration calorimetry; HPAEC, high performance anion exchange chromatography.
determinant. The capacity of CtXyl5A to also accommodate arabinose side chains in all the other subsites (in addition to the active site) within the substrate binding cleft enables the enzyme to hydrolyze highly decorated arabinoxylans. The functional significance of the specificity of the arabinoxylanase, in the context of the plant cell wall degrading apparatus of the host bacterium, is discussed.

**EXPERIMENTAL PROCEDURES**

**Cloning, Expression, and Purification of Components of CtXyl5A**—DNA encoding CtGH5, CtGH5-CBM6, and CtCBM6 were amplified using primers, containing Nhel and XhoI restriction sites, which are listed in supplemental Table S1. The amplified DNAs were cloned into Nhel/XhoI-restricted pET21a such that the encoded recombinant proteins contained a C-terminal His6 tag. To express the restricted pET21a such that the encoded recombinant proteins were partially purified by size exclusion chromatography using a Bio-Gel P2 column as described previously (16). The structures of the oligosaccharides were analyzed by NMR, electrospray ionization mass spectrometry (ESI-MS), and HPAEC in combination with selective enzyme treatment. Partially methylated alditol acetate derivatives of the glycosyl residues of the oligosaccharides were prepared and analyzed by gas chromatography-electron impact mass spectrometry GC-EI-MS.

**Preparation of the Partially Methylated Alditol Acetates**—The mixture of oligosaccharides (~500 μg) was per-O-methylated using the method of Ciucanu and Kerek (17). The per-O-methylated oligosaccharides were hydrolyzed with 2 n TFA, reduced, and acetylated to generate partially methylated alditol acetate derivatives (18).

**GC-EI-MS Analysis**—Partially methylated alditol acetate derivatives were analyzed with a Hewlett-Packard 5890 gas chromatograph-mass spectrometer. The partially methylated alditol acetates were separated with an SP 2330 column (30 m × 0.25 mm, 0.25-μm film thickness, Supelco) using the following temperature gradient: 80 °C for 2 min, 80–170 °C at 30 °C/min, 170–240 °C at 4 °C/min, and 240 °C held for 20 min. Samples were ionized by electrons impact at 70 eV.

**MALDI-TOF Mass Spectrometry (MALDI-TOF-MS)**—Positive ion MALDI-TOF mass spectra were recorded using an Applied Biosystems Voyager-DE biospectrometry workstation. Samples (1 μl of a mg/ml solution) were mixed with an equal volume of matrix solution (0.1 M 2,5-dihydroxybenzoic acid and 0.03 M 1-hydroxyisoquinoline in aqueous 50% MeCN) and dried on MALDI target plate. Typically, spectra from 200 laser shots were summed to generate a mass spectrum.

**ESI-MS**—The multiple stage ESI mass spectra were recorded in a Thermo Scientific LTQ XL ion trap mass spectrometer. Per-O-methylated oligoglycosyl alditols in methanol were diluted with 50% acetonitrile/water containing 0.1% TFA. Samples were infused through a fused silica capillary (150 μm inner diameter × 363 μm outer diameter × ~60 cm, Thermo Finnigan) into the source at flow rate of 3 μl/min using the syringe pump provided with the instrument. The electrospray source was operated at a voltage of 5.0 kV, and the capillary heater was set to 275 °C. All the experiments were performed in the positive-ion mode.

**NMR Spectroscopy**—Oligosaccharides (~2 mg) were dissolved in D$_2$O (0.5 ml, 99.9%; Cambridge Isotope Laboratories). $^1$H NMR spectra were recorded with Varian Inova NMR spec-
Structure and Function of CtXyl5A

trometer operating at 500 MHz at 298 K. All two-dimensional spectra were recorded using standard Varian pulse programs.

Isothermal Titration Calorimetry—The binding of CtCBM6 to ligands was quantified by isothermal titration calorimetry (ITC), as described previously (20). Titrations were carried out in 50 mM Na-HEPES buffer, pH 7.5, containing 5 mM CaCl$_2$ at 25 °C. The reaction cell contained protein at 145 μM, and the syringe contained the monosaccharide or oligosaccharide at 5–15 mM, and polysaccharide, when used as the titrant, was at 3–5 mg/ml. The titrations were analyzed using Microcal Origin version 7.0 software to derive, $n$, $K_m$, and $\Delta H$ values, and $\Delta S$ was calculated using the standard thermodynamic equation, $RT\ln K_n = \Delta G = \Delta H - T\Delta S$.

Crystallography—Proteins were crystallized using the hanging drop vapor technique at 20 °C with an equal volume (1 μl) of protein and reservoir solution. Native (10 mg/ml) and selenomethionine (3 mg/ml) CtGH5-CBM6 were crystallized in 16–24% PEG 3000, 150 mM sodium citrate, pH 5.5. A CtGH5-CBM6 construct containing two additional methionines, W391M/W397M, was produced to facilitate structure solution by selenomethionine single wavelength anomalous x-ray scattering. Crystals were cryoprotected by the inclusion of 25% glycerol in the crystallization solution and flash-frozen in liquid nitrogen. Diffraction data were collected at ID14.4 ESRF, Grenoble, France, at the selenium absorption edge to enable K absorption edge to enable structure solution by single wavelength anomalous x-ray scattering. The diffraction data were processed in MOSFLM (21) and SCALA (22), and the heavy atom substructure was solved using SHELX CDE (23) as part of CCP4i, and an initial model was built in Arp/wArp (24), which was completed manually in COOT (25). The complete initial model was used to determine the structure of the wild type protein by molecular replacement and refined at higher resolution from data collected at the Diamond Light Source, UK. The crystal of the reported structure had been soaked in 20 mM "Fraction 1" in an attempt to obtain a structure of the enzyme in complex with carbohydrate, although no sugar molecules, other than glycerol, were observed in the electron density.

All structures were refined to convergence using REFMAC5 (26) with manual corrections being applied in COOT (25). The data collection, phasing, and refinement statistics are displayed in supplemental Table S2, and the PDB code for the protein structure is 2y8k.

RESULTS

Expression and Purification of CtXyl5A—To investigate the function of the CtGH5 and CtCBM6 components of CtXyl5A, the modules were expressed as either individual entities or covalently linked. Although CtCBM6 and CtGH5-CBM6 were expressed in soluble form at high levels in E. coli, CtGH5 was predominantly insoluble, and only a small amount of soluble protein was generated in the enteric bacterium. All three proteins were purified by immobilized metal ion affinity chromatography to electrophoretic homogeneity.

CtXyl5A Is an Arabinoxylanase—Screening the capacity of CtXyl5A to hydrolyze plant structural polysaccharides revealed that the enzyme was able to degrade rye and wheat arabinoxylan, displayed limited activity against oat spelt xylan but was unable to act on glucuronyxylan, birch, or beech xylan (Table 1). The enzyme displayed no activity against a range of mannans, pectins, galactans, arabinans, and β-glucans (data not shown). The individual kinetic constants of CtXyl5A against rye and wheat arabinoxylan could not be determined as the $K_m$ value was greater than the maximum concentration of soluble substrate; however, the catalytic efficiency of the enzyme was similar for both rye and wheat arabinoxylan. The high $K_m$ value may reflect weak affinity for the substrate, or the glycosidic bonds targeted by CtXyl5A occur rarely in the arabinoxylan substrates. The enzyme displayed trace activity against xylolhexaose with a $k_{cat}/K_m$ value estimated to be $<10^4$ min$^{-1}$ M$^{-1}$. These data indicate that CtXyl5A hydrolyzes arabinoxylans but does not act on xylans that contain few arabinofuranose side chains. This is in sharp contrast to typical xylanases, located mainly in GH10 and GH11, which display a preference for the poorly decorated xylans from birch and beech (12). These data show that CtXyl5A displays specificity for arabinoxylans and as such is defined as an arabinoxylanase, an activity not previously reported.

Characterization of the Reaction Products Generated by CtXyl5A from Arabinoxylan—To explore the substrate specificity of CtXyl5A in more detail, the reaction products generated by treating rye arabinoxylan with the enzyme were partially purified by size exclusion chromatography to remove high molecular weight polymers. The fractions containing the majority of the products were pooled (designated henceforth as fraction 1). Fraction 1 was reduced and per-O-methylated, and the products were analyzed by MALDI-TOF-MS. The data revealed that the major reaction products were pentose-containing oligosaccharides with degrees of polymerization (DPs) of 3 (m/z 565), 4 (m/z 725), and 5 (m/z 885), respectively (Fig. 2A). Partially methylated alditol acetate derivatives were then prepared from per-O-methylated fraction 1 and analyzed by GC-EI-MS (Fig. 2B). This semi-quantitative analysis revealed terminal Ara$p$ (methylated at O2, O3, and O5), terminal Xyl$p$ (methylated at O2, O3, and O4), 3-linked Xyl$p$, 4-linked Xyl$p$, and 3,4-linked Xyl$p$. No Xyl$p$ residues decorated at O2 or at

| Proteins | Rye arabinoxylan | Wheat arabinoxylan | Oat spelt xylan |
|----------|-----------------|-------------------|-----------------|
| CtXyl5A  | 1322 ± 357      | 808 ± 76          | 15 ± 1.2        |
| CtXyl5A + 2 mm CaCl$_2$ | 2271 ± 274 | 1638 ± 56 | ND |
| CtXyl5A + 5 mM EDTA | 858 ± 102 | 343 ± 32 | ND |
| CtGH5-CBM6 | 1012 ± 83  | 652 ± 21          | 14 ± 2.2357    |
| W424A (CtGH5-CBM6) | 1656 ± 173 | 728 ± 127 | ND |
| F478A (CtGH5-CBM6) | 983 ± 52 | 713 ± 66 | ND |
| E79A (CtGH5-CBM6) | NA     | NA                | NA              |
| E171A (CtGH5-CBM6) | NA     | NA                | NA              |
| CtGH5    | 1.1 ± 0.45      | 0.6 ± 0.07        | NA              |

$a$ ND means not determined.

*b* NA means no activity detected.

---

**TABLE 1** Catalytic activity of CtXyl5A and its variants

The enzymes were assayed at 60 °C in 50 mM sodium phosphate buffer, pH 7.0, containing substrate at a concentration of 1 mg ml$^{-1}$. The reaction was monitored by the release of reducing sugar (14). The catalytic rate could be used to determine $k_{cat}/K_m$ values as the substrate concentration was $<< K_m$ (the rate of reaction was directly proportional to substrate concentration up to 2 mg ml$^{-1}$). Note that CtXyl5A is the full-length enzyme, and CtGH5-CBM6 and CtGH5 are derivatives of the enzymes containing the catalytic module appended to the CBM6 and the catalytic module, respectively.

---

**JOURNAL OF BIOLOGICAL CHEMISTRY**

VOLUME 286 NUMBER 25 • JUNE 24, 2011

22512
both O2 and O3 were observed. These data indicate that the oligosaccharides consist of a backbone of (1→3)-linked Xylp residues decorated with Araf side chains at O3 of internal or reducing Xylp residues (3,4-linked Xylp), or at O3 of nonreducing terminal Xylp residues in the xylan backbone (27). HPAEC analysis of the CjAbf51A digestion products revealed the presence of arabinose, xylobiose, xylotriose, and xylotetraose (Fig. 2C), indicating that the predominant CtXyl5A products are xylooligosaccharides in which at least one of the Xylp residues bear a mono-Araf side chain. By contrast, GH10 and GH11 xylanases generate predominantly xylose and xylobiose from wheat arabinoxylan, reflecting a preference for undecorated regions of the polysaccharide (12).

The oligosaccharides in fraction 1 were analyzed by several two-dimensional NMR methods, including gCOSY, HSQC, TOCSY, and ROESY. These analyses provided scalar and dipolar correlations that allowed the resonances of the most abundant spin systems to be assigned to specific sugar residues (supplemental Table S3; for a more detailed description of this approach, see, for example, Refs. 19, 28, 29). Upfield shifts typical of reducing residues (19, 28, 29) were observed for two C1 resonances (δ 92.4 and 96.6) in the HSQC spectrum of the CtXyl5A-generated oligosaccharides (Fig. 3A). In combination with other two-dimensional NMR data, this allowed these two resonances to be assigned to H1-Xylp and H3-Xylp residues at the reducing end of the oligosaccharides. However, the exact 1H and 13C shifts of these reducing residues indicate that they are structurally distinct from the unbranched (4-linked) sugars at the reducing end of oligosaccharides, generated by more typical endoxylanases (19, 28, 29). The data reveal the presence of an Araf side chain at O3 (along with a H3-Xylp at O4) of the reducing Xylp residues of the CtXyl5A-generated oligosaccharides. For example, the C3 resonances of the reducing α-Xylp and β-Xylp units exhibit diagnostic downfield glycosylation shifts (δ 77.7 and 77.8), relative to the corresponding unbranched reducing residues produced by more typical endoxylanases (δc 71.2 and 73.8). Furthermore, the ROESY spectrum of fraction 1 (Fig. 3B) revealed strong dipolar interactions between the two most intense α-Araf H1 resonances (δH 5.342 and 5.391) and the reducing α-Xylp and β-Xylp H3 resonances (δc 3.906 and 3.736, respectively), indicating that most of the α-Araf residues are linked to O3 of reducing Xylp moieties. The identification of branched, reducing Xylp residues in fraction 1 is consistent
with the detection of 3,4-linked Xylp residues in the partially methylated derivatives (Fig. 2B). Resonances corresponding to unbranched 4-linked β-Xylp residues at the reducing end of the oligosaccharides (e.g. H1 at δ 4.584, see Fig. 3A) were not detectable in the NMR spectra. Integration of the Xylp and Araf/H1 resonances in the one-dimensional spectrum of the Ctxyl5A-generated oligosaccharides (Fig. 3A) allowed the following quantitative conclusions to be drawn; the oligosaccharides have an average backbone DP of 2.76 and an average overall DP of 4.04; >99% of the oligosaccharides have an α-l-Araf side chain on O3 of the reducing Xylp residue; ~30% of the oligosaccharides have a second α-l-Araf side chain.

To analyze fraction 1 by ESI-MS”, the oligosaccharides in this sample were treated with NaBH$_4$ and the resulting oligoglycosyl alditols were methylated prior to fragmentation. This procedure imparts a distinctive mass label to the newly formed alditol end of the oligosaccharide, facilitating ESI-MS” analysis (19). The data, examples of which are shown in Fig. 4, provided unambiguous evidence supporting the presence of branched reducing residues in the oligosaccharides in fraction 1. This conclusion is exemplified by the analysis of the possible tetrasaccharides in fraction 1. Thus, based on the structure of the polysaccharide substrate, linkage, and NMR analysis of fraction 1, only five different tetrasaccharide structures (Ia, Ib, IIa, IIb, and III) are theoretically possible (Fig. 5). The ESI-MS” analysis provided information regarding the topology of the oligomers but did not define the stereochemistry (identity) of the individual pentose residues. Therefore, the terminal pentose residues at the nonreducing end of the main chain in structures Ia, Ib, IIa, IIb, displayed in Fig. 5, are indicated by the letter P (as the sugar can be either Araf or Xylp residues). However, in Fig. 5, nonterminal backbone residues, and sugars attached to branched backbone units (backbone sugars that are linked to O4 or O3 to other sugars), are known to be Xylp and Araf, respectively. Thus, structure I could be (Araf)-Xylp-Xylp-Xylol (Ia) or Xylp-Xylp-Xylp-Xylol (Ib) in which Araf is an arabinose decoration appended to the following xylose residue, whereas Xylol is the alditol form of the xylose at the reducing end. Structure II could be Araf-Xylp-(Araf)-Xylol (IIa) or Xylp-Xylp-(Araf)-Xylol (IIb), and III is Xylp-(Araf)-Xylp-Xylol. The quasi-molecular (M + Na$^+$) ion at m/z 725, corresponding to these DP4 structures was selected for MS$^2$ (Fig. 4A). The fragmentation pattern is dominated by y ions (19, 30), which contain the alditol end of the oligomer. The y ion (m/z 551) generated by loss of a single terminal pentosyl residue was selected as the precursor for MS$^3$ fragmentation (Fig. 4B). Comparison of this MS$^3$ spectrum (Fig. 4, A and B) to the theoretical fragmentation pattern for all possible m/z 551 ions (Fig. 5) indicates that structures I and III are not present, as these would fragment to form ions at m/z 231, which were not observed. This was confirmed by MS$^4$ analysis (Fig. 4, C and D), in which MS$^3$ fragment ions at m/z 391 and 377 were selected as precursors. Here, the extremely low abundance of ions at m/z 231 confirms the absence of significant amounts of structures I and III (Fig. 5). However, all ions predicted for structure II were observed, notably the high abundance ion at m/z 217, which consists of the alditol residue with two unmethylated hydroxyl groups that were exposed by cleavage of glycosidic bonds during MS$^2$ and MS$^3$ (Fig. 4).

When the DP5 oligoglycosyl alditols in fraction 1 were analyzed by MS”, virtually all of the alditol moieties were branched (supplemental Figs. S1 and S2). ESI-MS” data for the DP5 oligoglycosyl alditols also provide further insight into the extent to which Araf side chains can decorate the xylooligosaccharides produced by Ctxyl5A. Notably, MS$^3$ of the m/z 537 ion (derived from the alditol pentasaccharide) generates an m/z 363 y ion that yields an m/z 217 ion at MS$^5$. As shown in the schematic of supplemental Fig. S2, these species can only be generated if the xylosyl alditol and the adjacent Xylp are both branched. The detection of a m/z 377 ion at MS$^3$, however, demonstrates that the structure Xylp-Xylp-(Araf)-Xylol is also present. Fragmentation of DP5 oligoglycosyl alditols yields an m/z 217 y ion at
MS², although only trace amounts of a m/z 231 ion were evident (supplemental Figs. S3 and S4). This again demonstrates that the xylosyl alditol contains a branch, and thus the structure of the trisaccharide is predicted to be Xylp-(Araf)-Xylol.

**Binding of CtXyl5A to Arabinoxylan**—The terminal reaction products produced by endo-acting glycoside hydrolases reflects an iterative process in which the products from initial hydrolytic reactions serve as substrates in subsequent rounds of catalysis. Analysis of the structure of the terminal reaction products (which are unable to be further hydrolyzed) provides insight into the possible modes of substrate binding to both the negative and positive subsites (see below). The subsite nomenclature of glycoside hydrolases was defined previously by Davies et al. (31). Briefly, the scissile bond is positioned between subsites -1 and +1, and subsites that extend toward the nonreducing and reducing ends of the substrate are assigned increasing negative and positive numbers, respectively. The Xylp residues at the reducing and the nonreducing end of the oligosaccharide products are derived from substrate bound at the -1 and +1 subsites, respectively. As ~99% of the reducing end Xylp residues contain an O3 Araf branch, it is evident that the arabinose decoration of the xylose bound at the -1 subsite is a key specificity determinant of the enzyme. The detection of terminal Xylp (in which O2, O3, and O4 are methylated) and 3-linked Xylp residues, both of which occur at the nonreducing end of the oligosaccharide backbone, indicates that a Xylp with an Araf side chain at O3 can be accommodated in the +1 subsite of CtXyl5A, but a side chain in this position is not a specificity determinant. As both (Araf)-Xylp-(Araf)-Xylol and Xylp-Xylp-(Araf)-Xylol were identified in the tetrasaccharide, an O3-Araf side chain is present on some, but not all, of the Xylp residues bound in the -2 subsite. Thus, although an O3-Araf side chain can be accommodated at the -2 subsite, the arabinose decoration does not define enzyme specificity. The identification of Xylp-(Araf)-Xylp-(Araf)-Xylol in the pentasaccharide reaction products not only confirms that Araf can be present at the -1 and -2 subsites, but it also demonstrates that the +2 and +3 (if it exists) subsites can accommodate Xylp residues bearing arabinose side chains. It should be noted, however, that Xylp-(Araf)-Xylp-(Araf)-Xylp is a potential substrate for the enzyme (binding from subsites -2 to +1), suggesting that this molecule is only hydrolyzed very slowly by the enzyme, possibly because it is unable to access the +2 subsites. This is consistent with the absence of Xylp or (Araf)-Xylp in the reaction products; xylose or decorated xylose can only be generated if the substrate is hydrolyzed when it occupies only +1 of the positive subsites of the enzyme. Thus, to summarize, subsites -2 to +2 of CtXyl5A can accommodate Xylp residues that contain an O3-Araf side.

FIGURE 4. ESI-MS of the tetrasaccharides in fraction 1. The tetrasaccharides in fraction 1, which contains the most abundant products, were analyzed by ESI-MS². A shows the fragmentation of the m/z 725 ion, which includes the tetrasaccharides. B shows the fragmentation of the m/z 551 ion derived from the m/z 725 ion in A. C and D depict the fragmentation pattern of the m/z 391 and m/z 377 ions, respectively, derived from the m/z 551 ion generated in B. The masses of y ions are indicated unless otherwise stated.
chain; however, only at the −1 subsite does the arabinose decoration act as an essential specificity determinant.

*CtCBM6 Specificity*—To investigate whether CtCBM6 is a functional CBM, the capacity of CtGH5-CBM6 to bind to various carbohydrates was assessed by ITC. The data showed that CtGH5-CBM6 bound to cellohexaose and cellobiose with similar affinity (Table 2; example titrations are shown in Fig. 6). By contrast, binding to glucose was too low to quantify. The protein also displayed affinity for the reaction products generated by CtXyl5A and for undecorated xylooligosaccharides. The protein did not appear to bind to various xylans or to 1,3-1,4-glucans. This indicates that CtGH5-CBM6 recognizes the terminal region of these polysaccharides, as the concentration of ligand available to the protein in these polymers, which have DPs 300, would be very low, and thus binding would not be detected. It is possible that the catalytic module, rather than CBM6, mediates binding to the xylo- and cello-oligosaccharides. To test this hypothesis, the ligand binding profile of variants of CtGH5-CBM6, in which either Trp-424 or Phe-478 had been substituted with Ala, was assessed. As discussed below, these two aromatic residues are highly conserved in the CBM6 family and includes the primary binding site in this protein family (32). Both CtGH5-CBM6:W424A and CtGH5-CBM6:F478A, although catalytically active (Table 1), displayed no binding to the xylan- and cellulose-derived oligosaccharides (Table 2). It is therefore evident that the CBM6 component of CtGH5-CBM6 mediates the observed binding to oligosaccharides.

*Crystal Structure of CtGH5-CBM6*—The structure of CtGH5-CBM6 was solved by selenomethionine single wavelength anomalous x-ray scattering, and the resulting structure was used as a starting model for refinement against native data extending to 1.5 Å resolution (supplemental Table S2) (PDB code 2y8k). The polypeptide chain is visible from Ser-37 to Ile-516.

*CtXyl5A*—As expected, the N-terminal CtXyl5A module displays a (β/α)8 barrel architecture, although α-helix 8 points away from the barrel and toward CtCBM6 module (discussed below) (Fig. 7). GH5 enzymes are members of clan GH-A in which the two catalytic residues are invariant glutamates pre-
sentent at the end of β-strands 4 and 7 (33, 34). From the structure of CtGH5-CBM6, the catalytic acid-base is likely to be Glu-171 (end of β-strand 4) and the catalytic nucleophile Glu-279 (end of β-strand 7). The catalytic role of these two residues is confirmed by the observation that the mutants E171A and E279A are inactive (Table 1). A narrow V-shaped cleft, ~25 Å in length, extends along the full length of the protein and sits over the top of the β-barrel. The dimensions of the cleft, in the center of which is the catalytic apparatus, suggest that the protein contains ~5 subsites extending from ~3 to +2.

An analysis of structural homologues of the CtGH5 component of CtGH5-CBM6 by the DaliLite webserver identified a large number of GH5 and clan GH-A enzymes that displayed significant structural similarity to CtGH5. The Pseudoalteromonas haloplanktis cellulase Cel5G (PDB 1tvn), with a root mean square deviation of 2.8 Å over 253 Ca atoms and a Z-score of 24.1, and the Bacillus agaradhaerens cellulase BaCel5A (PDB 1qi2), with a root mean square deviation of 2.9 Å over 254 Ca atoms and a Z-score of 23.6, are representative and close structural homologues. The critical 1 subsite, where the transition state is formed, is similar in the arabinoxylanase and the GH5 cellulases. In addition to the two catalytic glutamates, CtGH5 contains several key residues that have been identified as “strictly conserved” in family GH5 enzymes (35). These residues in the CtGH5 module, which superimpose with amino acids in the active site of BaCel5A (the cellulase residues are shown in parentheses), are as follows: Asn-170 (Asn-138), Glu-171 (Glu-139), Tyr-255 (Tyr-202), Glu-279 (Glu-228), and Phe-310 (Trp-262) (Fig. S4). The catalytic acid base, Glu-171, makes hydrogen bonds with Asn-139 and His-253, and these interac-

### Table 2

**Binding of CtXyl5A derivatives to polysaccharides and oligosaccharides**

The binding of derivatives of CtXyl5A to ligands was measured by ITC. The protein was at the 145 μM in the cell and polysaccharide (3–5 mg/ml) or oligosaccharide (5–15 mM) was in the syringe. ITC was carried out in 50 mM Na/HEPES buffer, pH 7.5, at 25 °C. The concentration of the oligosaccharides generated by the digestion of wheat arabinoxylan (WAX) was fitted to give an n value close to 1.

| CtXyl5A derivative | Ligand | Kᵦ × 10⁴ | ΔG kcal mol⁻¹ | ΔH kcal mol⁻¹ | TΔS kcal mol⁻¹ | n |
|--------------------|--------|----------|---------------|---------------|---------------|---|
| CtGH5-CBM6         | Cellohexose | 2.3 ± 0.1 | -5.2 | -2.8 ± 0.1 | 3.7 ± 0.2 | 1.1 ± 0.1 |
| CtGH5-CBM6         | Xylooligosaccharide | 2.5 ± 0.2 | -5.3 | -2.8 ± 0.2 | 3.7 ± 0.3 | 1.1 ± 0.3 |
| CtGH5-CBM6         | WAX | 2.7 ± 0.3 | -5.1 | -2.9 ± 0.3 | 3.7 ± 0.1 | 0.9 ± 0.1 |
| CtGH5-CBM6         | WAX treated with CtXyl5A | 2.8 ± 0.4 | -5.2 | -2.9 ± 0.4 | 3.7 ± 0.2 | 1.1 ± 0.4 |
| CtGH5-CBM6:E279A   | Cellohexose | 2.4 ± 0.5 | -5.0 | -2.8 ± 0.5 | 3.5 ± 0.2 | 1.0 ± 0.1 |
| CtGH5-CBM6:E279A   | WAX treated with CtXyn10A | 2.5 ± 0.6 | -5.1 | -2.9 ± 0.6 | 3.6 ± 0.1 | 1.1 ± 0.2 |
| CtGH5-CBM6:F478A   | No binding to cellohexose, xylotetraose, WAX treated with CtXyn10A, CtXyl5A, or both CtXyl5A and CtAbF51A |
| CtGH5-CBM6:F478A   | No binding to cellohexose, WAX treated with CtXyn10A, CtXyl5A, or both CtXyl5A and CtAbF51A |
| CtGH5-CBM6        | No binding to cellohexose, xylotetraose, WAX treated with CtXyn10A, CtXyl5A, or both CtXyl5A and CtAbF51A |
| CtGH5-CBM6        | No binding to cellohexose, xylotetraose, WAX treated with CtXyn10A, CtXyl5A, or both CtXyl5A and CtAbF51A |
| CtGH5-CBM6        | No binding to WAX |
| CtGH5-CBM6         | Xyl5A derivative | 2.6 ± 0.7 | -5.3 | -2.9 ± 0.7 | 3.5 ± 0.2 | 1.0 ± 0.1 |
| CtGH5-CBM6         | Xyl5A | 2.7 ± 0.8 | -5.2 | -2.9 ± 0.8 | 3.6 ± 0.1 | 1.1 ± 0.2 |
| CtGH5-CBM6:F478A   | No binding to WAX |
| CtGH5-CBM6:F478A   | No binding to cellohexose, xylotetraose, WAX treated with CtXyn10A, CtXyl5A, or both CtXyl5A and CtAbF51A |
| CtGH5-CBM6:F478A   | No binding to cellohexose, xylotetraose, WAX treated with CtXyn10A, CtXyl5A, or both CtXyl5A and CtAbF51A |
| CtGH5-CBM6:F478A   | No binding to WAX |

**FIGURE 6. Representative ITC data of CtGH5-CBM6 to oligosaccharides.** The ligands (10 mM arabinose) in the syringe was titrated into CtGH5-CBM6 (100 μM) in the cell. The top half of each panel shows the raw ITC heats; the bottom half of each panel shows the integrated peak areas fitted using a single binding model by MicroCal Origin software. ITC was carried out in 50 mM Na-HEPES, pH 7.5, at 37 °C.
Crystallization experiments did not yield diffracting crystals. Consequently, it is difficult to define precisely the structural basis for the unusual substrate specificity displayed by the arabinoxylanase. Superimposing Bacell5A in complex with 2-deoxy-2-fluorocellotriose with Ctxy5A provides some insight into the specificity displayed by the arabinoxylanase. As discussed above, the catalytic apparatus, the residues that interact with O2 and the endocyclic oxygen of the −1 sugar, and the hydrophobic platform are conserved in Ctxy5A (Fig. 8A). It is evident, however, that the arabinoxylanase lacks the residues that in other GH5 enzymes hydrogen bond with O3 of the active site sugar. For instance, His-101 and Tyr-66 in Bacell5A hydrogen bond with O3 of the −1 Glc, whereas the equivalent residues in Ctxy5A are Gly-134 and Cys-95, respectively (Fig. 8A). Indeed, in the −1 subsite of the arabinoxylanase, there is a large pocket around the O3 of the superimposed Glc that could accommodate a sugar decoration such as Araf (Fig. 8B). The pocket contains a tyrosine (Tyr-92) that may make hydrophobic interactions with the arabinose and contains several polar residues, Glu-68, Asn-135, Asn-139, and Asn-170, that could make polar contacts with the sugar. Based on the presence of glycerol and water molecules within this region of the enzyme, an Araf molecule was modeled into the pocket (Fig. 8C).

**CtCBM6**—The structure of the CtCBM6 module displays a β-sandwich fold typical of other family CBM6 members (Fig. 7) (32, 37, 38). The twisted pair of β-sheets, which can be viewed as forming an extended barrel, consist of five and four anti-parallel β-strands, respectively. The structure of CtCBM6 shows strong similarity with numerous CBM6 members. The closest homologue is the CBM6 module (designated CmCBM6) from the Cellvibrio mixtus lichenase CmLic5A (PDB 1uz0; root mean square deviation 1.5 Å over 123 Ca atoms and a Z-score of 18.1). The major binding site in the CBM6 family is in the loops connecting the two β-sheets. This region, referred to as site A (32, 37), may comprise a pocket if terminal sugars are recognized (39) or a cleft for the binding of internal regions of polysaccharides (32). A central feature of site A is a pair of aromatic residues, which bind to the α- and β-face, respectively, of the terminal sugar (or central sugar in the case of xylan-binding modules) and an asparagine, located at the base of the site that makes critical hydrogen bonds with O2, O3, or O4. Specificity is conferred by additional polar and hydrophobic interactions (37). Site A in CtCBM6 displays a pocket-like topology and contains all the key ligand binding residues present in CmCBM6 (Fig. 9) (40). The pair of aromatic residues in CmCBM6, Trp-92 and Tyr-33, which straddle the nonreducing terminal sugar, correspond to Phe-478 and Tyr-202, respectively, in CtCBM6. Furthermore, Glu-20 and Asn-121 in CmCBM6, which make polar contacts with O3 and O4 of the nonreducing terminal sugar in cello- and xylooligosaccharides, superimpose with Glu-411 and Asn-507, respectively, in CtCBM6. Finally, the amide nitrogen of Tyr-33 in CmCBM6 makes a polar contact with O2 and O3 of the terminal sugar, a contact that is likely to be replicated by that of Trp-424 in CtCBM6. The structural conservation between site A in CtCBM6 and CmCBM6 is consistent with the similar ligand specificities displayed by this binding site in the two proteins (Table 2) (39). Thus, both proteins bind to xylo- and gluco-
configured oligosaccharides but do not display affinity for the corresponding polysaccharides. Thus, the structural similarity between CmCBM6 and CtCBM6 is consistent with the view that the Clostridium module targets the terminal regions of oligosaccharides. In CmCBM6 cellobiose-oligosaccharides can bind to site A in both orientations, consistent with the targeting of O1/O4, O2, and O3, but not the endocyclic oxygen or O6, which would adopt different positions in the two orientations. Therefore, it is highly likely that CtCBM6 will also bind to xylo- and cellobiose-oligosaccharides in both orientations. Given that the key interactions with the ligand at site A is with the terminal sugar, it is perhaps surprising that CtCBM6 does not display measurable binding to xylose or glucose. It is possible that the entropic cost of locking the sugar into a pyranose ring conformation may contribute to the weak binding, although it is also possible that the protein makes indirect, water-mediated interactions to the penultimate sugar in the oligosaccharides, as observed in CmCBM6-ligand complex (40).

**Linker Connecting CtGH5 with CtCBM6—CtCBM6 is connected to CtGH5 by a sequence extending from residues Gly-336 to Thr-373. This linker, which adopts a stable conformation based on its B-factor, makes numerous internal polar contacts and forms hydrogen bonds with β-strand 3 and the loop connecting β-strands 3 and 4 of CtCBM6 and α-helices 7 and 8 of CtGH5. Furthermore, the C-terminal region of α-helix 8 and the internal region of α-helix 7 make hydrogen bonds with β-strands 3 and 7 of CtCBM6. The polar contacts between the two modules are augmented by a large number of apolar interactions mediated by the linker sequence. The resultant burial of a significant hydrophobic surface, at the interface between CtGH5 and CtCBM6, likely explains why these two modules (or domains) do not fold independently, as occurs in other glycoside hydrolases that contain catalytic modules and CBMs (41).**

This view is consistent with the observation that CtCBM6, when expressed as a discrete entity (Thr-373 to Ile-516), does not bind to cellohexaose or xylohexaose, and CtGH5 (Asn-32 to Thr-373) exhibits very low catalytic activity and is considerably more thermolabile than CtGH5-CBM6 (supplemental Fig. S5).

**DISCUSSION**

This study reveals a *C. thermocellum* protein that displays arabinoxyylanase activity, an activity not previously reported. The vast majority of xylanases are derived from GH10 and GH11 and target the β-1,4-D-xylose polymeric backbone. These enzymes do not generally distinguish between different xylans, although highly decorated forms of the polysaccharide, such as rye arabinoxylan, are poorly degraded as steric constraints restrict enzyme access (12). Indeed, the only other examples of xylanases that utilize side chains as essential specificity determinants are glucuronoxylan-specific enzymes from GH30. These enzymes make critical interactions with the 4-O-methylglucurononic acid (linked α-1,2 to the xylan backbone) that decorates the xylose at the −2 subsite (42). CtXyl5A is highly unusual in that its essential Ara/f decoration is attached to the xylose positioned in the active site. The only other example of an active site side chain specificity determinant is the α-1,6-Xylp that decorates the −1 Glc in the xyloglucan cellobiohydrolase, OXG-RCBH, from *Geotrichum* sp. (43).

The function of CtXyl5A within the context of *C. thermocellum*, which has the genetic capacity to recruit 72 different enzymes into the cellulosomes (44), including seven GH10 and GH11 xylanases, is intriguing. It is likely that the GH10 and GH11 enzymes target xylans that are sparsely decorated with arabinose side chains. By contrast, CtXyl5A most likely hydrolyzes xylans where tandem Xylp residues contain Ara/f decorations. The recognition of the termini of xylene and glucosyl-configured polymers by CtCBM6 suggests that the arabinoxylanase is targeted to regions of the plant cell wall that is undergoing degradation and is therefore accessible to enzyme attack. Although the primary function of CBMs is to bring their cognate enzymes into close contact with appropriate substrates (45), there is increasing evidence that a subset of these modules, from CBM families 6, 9, and 35, target the termini of polysaccharides and thus may play a similar function to CtCBM6 (37, 46, 47). In conclusion, CtXyl5A displays a specificity that is complementary to endoxylanases from GH10, GH11, and GH30. As such the enzyme will make a contribution to the toolbox of biocatalysts required to degrade plant cell walls to their constituent sugars, which can then be used in the biofuel and bioprocessing industries.
Structure and Function of CtXyIA5A

Acknowledgments—We thank the European Synchrotron Radiation Facility, Diamond Light Source, and the staff at both these facilities for help with data collection. We also thank the Department of Energy Office Center for Plant and Microbial Complex Carbohydrates (supported by Grant DE-FG02-93ER20097) for supporting critical infrastructure and analytical instrumentation required for this research.

REFERENCES

1. Brett, C. T., and Waldren, K. (1996) Physiology and Biochemistry of Plant Cell Walls. Topics in Plant Functional Biology, Chapman and Hall Ltd., London.

2. Himmel, M. E., and Bayer, E. A. (2009) Curr. Opin. Biotechnol. 20, 316–317.

3. Ragauskas, A. J., Williams, C. K., Davison, B. H., Britovsek, G., Cairney, J., Čiček, C. A., Frederick, W. J., Jr., Hallett, J. P., Leak, D. J., Liotta, C. L., Mielzen, J. R., Murphy, R., Templer, R., and Tschapinski, T. (2006) Science 311, 484–489.

4. Ebringerová, A., Hromadkova, Z., and Heinze, T. (2005) in Polysaccharides: I. Structure, Characterization and Use (Heinze, T., ed) pp. 1–67, Springer-Verlag, Berlin.

5. DeBoy, R. T., Mongodin, E. F., Davison, B. H., Grondin, G., Kesser, D., Gilbert, H. J., and Nelson, K. E. (2008) J. Bacteriol. 190, 5455–5465.

6. Weiner, R. M., Taylor, L. E., 2nd, Henrissat, B., Hauser, L., Land, M., Coutinho, P. M., Rancurel, C., Saunders, E. H., Longmire, A. G., Zhang, H., Bayer, E. A., Gilbert, H. J., Larimer, F. Z., Zhulin, I. B., Eborg, N. A., Lamed, R., Richardson, P. M., Borovok, I., and Hutcheson, S. (2008) PLoS Genet. 4, e1000087.

7. Berg Miller, M. E., Antonopolous, D. A., Rincon, M. T., Band, M., Bari, A., Akraiko, T., Hernandez, A., Thimmaphumjar, H., Henrissat, B., Coutinho, P. M., Borovok, I. Jindov, S., Land, R., Rint, H. J., Bai, Y. E., and White, B. A. (2009) PLoS One 4, e6650.

8. Xu, J., Bjursell, M. K., Himrod, J., Deng, S., Carmichael, L. K., Chiang, S., Hooper, L. V., and Gordon, J. I. (2003) Science 299, 2074–2076.

9. Cantarel, B. L., Coutinho, P. M., Rancurel, C., Bernard, T., Lombard, V., and Henrissat, B. (2009) Nucleic Acids Res. 37, D233–D238.

10. Gilbert, H. J. (2010) Plant Physiol. 153, 444–455.

11. Gilbert, H. J., Stålbrand, H., and Brumer, H. (2008) Curr. Opin. Plant Biol. 11, 338–348.

12. Pell, G., Taylor, E. J., Gloster, T. M., Turkenburg, J. P., Fontes, C. M., Ferreira, L. M., Borin, O., Zamboni, B., Darbon, H., Smith, N. L., Black, G. W., Henrissat, B., and Gilbert, H. J. (2001) in press.

13. Henshaw, J. L., Bolam, D. N., Prates, J. A., Cartmell, A., Coutinho, P. M., Henrissat, B., Gilbert, H. J., and Boraston, A. B. (2009) Biochemistry 48, 10395–10404.

14. van Bueren, A. L., Molrand, C., Gilbert, H. J., and Boraston, A. B. (2005) J. Biol. Chem. 280, 530–537.

15. Henshaw, J. L., Bolam, D. N., Pires, V. M., Czjzek, M., Henrissat, B., Ferreira, L. M., and Fontes, C. M., Henrissat, B., Gilbert, H. J., and Boraston, A. B. (2004) Biochemistry 43, 1643–1644.

16. Williams, S. J., Notenboom, V., Wicki, J., Rose, D. R., and Withers, S. G. (2000) J. Am. Chem. Soc. 122, 4229–4230.

17. Abbott, D. W., Ficko-Blean, E., van Bueren, A. L., Rogowski, A., Cartmell, A., Coutinho, P. M., Henrissat, B., Gilbert, H. J., and Boraston, A. B. (2009) Biochemistry 48, 10395–10404.

18. van Bueren, A. L., Molrand, C., Gilbert, H. J., and Boraston, A. B. (2005) J. Biol. Chem. 280, 530–537.

19. Henshaw, J. L., Bolam, D. N., Pires, V. M., Czjzek, M., Henrissat, B., Ferreira, L. M., and Fontes, C. M., Henrissat, B., Gilbert, H. J., and Boraston, A. B. (2004) Biochemistry 43, 1643–1644.

20. Pires, V. M., Henshaw, J. L., Prates, J. A., Bolam, D. N., Ferreira, L. M., Fontes, C. M., Henrissat, B., Planas, A., Gilbert, H. J., and Czjzek, M. (2004) J. Biol. Chem. 279, 21552–21559.

21. Boraston, A. B., Bolam, D. N., Gilbert, H. J., and Davies, G. J. (2004) Biochem. J. 382, 769–781.

22. Vrsanská, M., Kolenová, K., Puchart, V., and Biely, P. (2007) FEBS J. 274, 1666–1677.

23. Yao, K., Kondo, H., Hiyoshi, A., Noro, N., Sugimoto, H., Tsuda, S., Mitsuishi, Y., and Miyazaki, K. (2007) J. Mol. Biol. 370, 53–62.

24. Fontes, C. M., and Gilbert, H. J. (2010) Annu. Rev. Biochem. 79, 655–681.

25. Bolam, D. N., Ciruelas, A., McQueen-Mason, S., Simpson, P., Williamson, M. P., Rixon, J. E., Boraston, A., Hazelwood, G. P., and Gilbert, H. J. (1998) Biochem. J. 331, 775–781.

26. Boraston, A. B., Creagh, A. L., Alam, M. M., Kormos, J. M., Tomme, P., Haynes, C. A., Warren, R. A., and Kilburn, D. G. (2001) Biochemistry 40, 6240–6247.

27. Montanier, C., van Bueren, A. L., Dumon, C., Flint, J. E., Correia, M. A., Prates, J. A., Fibeck, S. J., Lewis, R. J., Grondin, G. M., Ghinet, M. G., Gloster, T. M., Herve, C., Knox, J. P., Talbot, B. G., Turkenburg, J. P., Kerovuo, J., Brzezinski, R., Fontes, C. M., Davies, G. J., and Boraston, A. B., and Gilberth, H. J. (2006) Biochim. Ann. Rev. Biochem. 79, 655–681.

28. Montanier, C., Correia, M. A., S., Flint, J. E., Zhu, Y., McKee, L. S., Prates, J. A. M., Poliari, S. C., Coutinho, P. M., Lewis, R. J., Henrissat, B., and Gilbert, H. J. (2011) J. Biol. Chem. 286, in press.

29. Montanier, C., Correia, M. A., S., Flint, J. E., Zhu, Y., McKee, L. S., Prates, J. A. M., Poliari, S. C., Coutinho, P. M., Lewis, R. J., Henrissat, B., and Gilbert, H. J. (2011) J. Biol. Chem. 286, in press.

30. Montanier, C., Correia, M. A., S., Flint, J. E., Zhu, Y., McKee, L. S., Prates, J. A. M., Poliari, S. C., Coutinho, P. M., Lewis, R. J., Henrissat, B., and Gilbert, H. J. (2011) J. Biol. Chem. 286, in press.

31. Montanier, C., Correia, M. A., S., Flint, J. E., Zhu, Y., McKee, L. S., Prates, J. A. M., Poliari, S. C., Coutinho, P. M., Lewis, R. J., Henrissat, B., and Gilbert, H. J. (2011) J. Biol. Chem. 286, in press.

32. Montanier, C., Correia, M. A., S., Flint, J. E., Zhu, Y., McKee, L. S., Prates, J. A. M., Poliari, S. C., Coutinho, P. M., Lewis, R. J., Henrissat, B., and Gilbert, H. J. (2011) J. Biol. Chem. 286, in press.