



HAL
open science

Structure and Activity of Two Metal Ion-dependent Acetylxylylan Esterases Involved in Plant Cell Wall Degradation Reveals a Close Similarity to Peptidoglycan Deacetylases

Edward J. Taylor, Tracey M. Gloster, Johan P. Turkenburg, Florence Vincent, A. Marek Brzozowski, Claude Dupont, François Shareck, Maria S.J. Centeno, José A.M. Prates, Vladimír Puchart, et al.

► To cite this version:

Edward J. Taylor, Tracey M. Gloster, Johan P. Turkenburg, Florence Vincent, A. Marek Brzozowski, et al.. Structure and Activity of Two Metal Ion-dependent Acetylxylylan Esterases Involved in Plant Cell Wall Degradation Reveals a Close Similarity to Peptidoglycan Deacetylases. *Journal of Biological Chemistry*, 2006, 281 (16), pp.10968-10975. 10.1074/jbc.M513066200 . hal-03281241

HAL Id: hal-03281241

<https://hal.science/hal-03281241>

Submitted on 8 Jul 2021

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution 4.0 International License

Structure and Activity of Two Metal Ion-dependent Acetylxylan Esterases Involved in Plant Cell Wall Degradation Reveals a Close Similarity to Peptidoglycan Deacetylases*

Received for publication, December 7, 2005, and in revised form, January 23, 2006. Published, JBC Papers in Press, January 23, 2006, DOI 10.1074/jbc.M513066200

Edward J. Taylor[‡], Tracey M. Gloster[‡], Johan P. Turkenburg[‡], Florence Vincent[‡], A. Marek Brzozowski[‡], Claude Dupont[§], François Shareck[§], Maria S. J. Centeno[¶], José A. M. Prates[¶], Vladimír Puchart^{||}, Luís M. A. Ferreira[¶], Carlos M. G. A. Fontes[¶], Peter Biely^{||}, and Gideon J. Davies^{‡1}

From the [‡]Structural Biology Laboratory, Department of Chemistry, University of York, Heslington, York YO10 5YW, United Kingdom, the [§]Institut National de la Recherche Scientifique, Institut Armand-Frappier, 531 Boulevard des Prairies, Laval, H7V 1B7 Québec, Canada, the [¶]Centro Interdisciplinar de Investigação em Sanidade Animal, Faculdade de Medicina Veterinária, Universidade Técnica de Lisboa, Avenida da Universidade Técnica, 1300-477 Lisbon, Portugal, and the ^{||}Institute of Chemistry, Slovak Academy of Sciences, Dubravská cesta 9, 84538 Bratislava, Slovakia

The enzymatic degradation of plant cell wall xylan requires the concerted action of a diverse enzymatic syndicate. Among these enzymes are xylan esterases, which hydrolyze the *O*-acetyl substituents, primarily at the *O*-2 position of the xylan backbone. All acetylxylan esterase structures described previously display a α/β hydrolase fold with a “Ser-His-Asp” catalytic triad. Here we report the structures of two distinct acetylxylan esterases, those from *Streptomyces lividans* and *Clostridium thermocellum*, in native and complex forms, with x-ray data to between 1.6 and 1.0 Å resolution. We show, using a novel linked assay system with PNP-2-*O*-acetylxyloside and a β -xylosidase, that the enzymes are sugar-specific and metal ion-dependent and possess a single metal center with a chemical preference for Co^{2+} . Asp and His side chains complete the catalytic machinery. Different metal ion preferences for the two enzymes may reflect the surprising diversity with which the metal ion coordinates residues and ligands in the active center environment of the *S. lividans* and *C. thermocellum* enzymes. These “CE4” esterases involved in plant cell wall degradation are shown to be closely related to the de-*N*-acetylases involved in chitin and peptidoglycan degradation (Blair, D. E., Schuettelkopf, A. W., MacRae, J. I., and Aalten, D. M. (2005) *Proc. Natl. Acad. Sci. U. S. A.*, 102, 15429–15434), which form the NodB deacetylase “superfamily.”

Many plant cell wall polysaccharides, including xylan, mannan, and pectin are present in acetylated forms. Acetylation not only modifies the physicochemical properties of polysaccharides, notably increasing the solubility for matrix applications, but also means they are less readily attacked by phytopathogen-derived cell wall-degrading endoglycosidases. To overcome the steric problems provided by acetyl substituents, plant cell wall degrading microorganisms have developed a host of acetyl esterases whose function is to deacetylate the polysaccharides prior to, or concomitant with, its complete hydrolysis by a consortium of *exo*- and *endo*-acting glycoside hydrolases. Such microbial esterases

have, unsurprisingly, found widespread industrial application in both biomass conversion and for the chemoenzymatic synthesis of diverse esters (see for example Refs. 1 and 2 and reviewed in Ref. 3).

Xylan is a chemically and structurally complex plant cell wall polysaccharide, whose complete degradation requires the action of a dedicated enzymatic consortium, Fig. 1A (reviewed in Ref. 4). Esterases are required both to remove the ferulate groups (linked via arabinosides to the xylan backbone (5)) and to deacetylate the *O*-3 and (primarily) *O*-2 positions of the xylan backbone in acetylxylan. Carbohydrate esterases (CE),² which perform these reactions, are classified into 14 sequence-based families defined by Henrissat *et al.* (6) (recently reviewed in Ref. 7). Thus far, structures of ferulate and acetylxylan esterases have revealed a α/β hydrolase fold, and they display a Ser-His-Asp catalytic triad. Examples include the family CE1 (predominantly bacterial) ferulate esterases (8–10), the currently unclassified fungal ferulate esterases (11–13) and the xylan and xylooligosaccharide esterases from families CE5 (14, 15), CE6 (putative esterases, PDB code 2APJ, Centre for Eukaryotic Structural Genomics), and CE7 (16). Enzymes in the largest sequence-based esterase family, CE4, do not, however, display the standard α/β hydrolase fold.

The carbohydrate esterase family CE4 contains over 870 open reading frames.³ This family is notable, not merely for its size, but also as many CE4 members have been reported to be metal ion-dependent. Furthermore, family CE4 contains members with both classical de-*O*-acetylase activity, such as the acetylxylan esterases described here, but also de-*N*-acetylases, such as the “NodB superfamily” involved in the degradation and/or remodeling of diverse *N*-acetylglucosamine-based substrates (Fig. 1B) including chitin (17), chitooligosaccharide rhizobial Nod-factors and peptidoglycan (for example Refs. 18 and 19). Recent work by van Aalten and co-workers (19, 20) has demonstrated examples of the three-dimensional structures of CE4 “peptidoglycan” de-*N*-acetylases. These enzymes possess a single divalent (typically Zn^{2+} or Co^{2+}) metal ion at the active center, similar to that observed for many Zn^{2+} hydrolases (reviewed in Ref. 21). The metal ion gives Lewis-acid assistance to nucleophilic attack by water, with two histidines and an aspartate completing the metal ion coordination.

Here we present the first structures for two distinct de-*O*-acetylases from family CE4, the *Clostridium thermocellum* (CtCE4) and *Streptomyces lividans* (SlCE4) acetylxylan esterases. We show, using a novel

* This work was supported by the Biotechnology and Biological Sciences Research Council, the Engineering and Physical Sciences Research Council, and the Royal Society. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked “advertisement” in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

The atomic coordinates and structure factors (codes 2C71, 2C79, and 2CC0) have been deposited in the Protein Data Bank, Research Collaboratory for Structural Bioinformatics, Rutgers University, New Brunswick, NJ (<http://www.rcsb.org/>).

¹ To whom correspondence should be addressed. E-mail: davies@ysbl.york.ac.uk.

² The abbreviations used are: CE, carbohydrate esterase; MES, 4-morpholineethanesulfonic acid.

³ See <http://afmb.cnrs-mrs.fr/cazy/>.

linked assay system (Fig. 1C) (22), that these enzymes are both metal ion-dependent. Although both show a “chemical” preference for Co^{2+} , the *C. thermocellum* enzyme displays distinctly different ligand coordination to the metal ion, utilizing an aspartate, a histidine, and four water molecules, as opposed to the more classical His-His-Asp family 4 “consensus” displayed by the *S. lividans* enzyme. The *C. thermocellum* enzyme also displays a modified metal ion specificity compared with the *S. lividans* esterase and the recently described peptidoglycan deacetylases (19, 20); CtCE4 is much more active with Ni^{2+} than is observed for enzymes with the classical His-His-Asp “ Zn^{2+} ” coordination.

EXPERIMENTAL PROCEDURES

Protein Expression and Purification—The cloning and expression of the gene fragment encoding the CE4 domain of the *S. lividans* acetylxylan esterase has been described previously (23). Briefly, proteins contained in the supernatant derived from *S. lividans* overexpressing the truncated *SICE4* gene (23) were concentrated by ultrafiltration with a 3-kDa cut-off membrane (Omega). Approximately 200 mg of protein was dialyzed against 20 mM MES-NaOH buffer (pH 6.0) and applied to a CM-Sepharose AP2 column (Waters, 20 × 100 mm) equilibrated with the same buffer. The bound proteins were eluted with a linear gradient of NaCl, and the active fractions were collected and pooled. Following this the active fraction was further purified by gel filtration on a Superdex HR75 beaded column with 100 mM sodium phosphate, pH 6.0, as buffer. Fractions containing pure enzyme were pooled, dialyzed against water, and freeze-dried.

The gene encoding the CE4 domain from the bifunctional cellulosomal xylanase U, Xyn11A, from *C. thermocellum* (designated CtXyn11A-Axe4A) was expressed in *Escherichia coli*. CtXyn11A-Axe4A was amplified from *C. thermocellum* YS genomic DNA using the thermostable DNA polymerase pFU Turbo (Stratagene). The primers used, 5'-CTC-CATATGCCAAATGCGAAACTGGTGGC-3' and 5'-CACCTCGA-GCGGTACAGAGTTATACATTC-3', incorporated NdeI and XhoI restriction sites, which are depicted in bold. The PCR product was cloned into pGEM T-easy (Promega) and sequenced to ensure that no mutations had occurred during the PCR. The recombinant pGEM T-easy derivative was digested with NdeI and XhoI, and the excised CtCE4-encoding gene was cloned into the similarly restricted expression vector pET21a to generate pZC1. pZC1 encodes CtCE4 with a C-terminal His₆-tag. BL21 cells were transformed by the recombinant plasmid, which were subsequently cultured in LB containing 100 $\mu\text{g ml}^{-1}$ ampicillin at 37 °C to midexponential phase (A_{550} 0.6 absorbance units). At this point isopropyl- β -D-thiogalactopyranoside was added to a final concentration of 1 mM, and the culture was incubated for a further 5 h. Cells were harvested by centrifugation, and the cell pellet was resuspended in 50 mM Na-HEPES buffer, pH 7.5, containing 1 M NaCl and 10 mM imidazole. Recombinant CtCE4 was purified by immobilized metal ion affinity chromatography. The enzyme was buffer exchanged, using a PD-10 Sephadex G-25 M gel filtration column (Amersham Biosciences), into 50 mM HEPES buffer, pH 7.5, containing 200 mM NaCl (Buffer A) and concentrated to 20 mg ml^{-1} with an Amicon 10-kDa molecular weight centrifugation membrane. Gel filtration using a HiLoad 16/60 Superdex 75 column (Amersham Biosciences) was performed, and the protein eluted at 1 ml min^{-1} in Buffer A. The purified enzyme was concentrated as described above and washed three times with water using the same centrifugal membranes. The final protein concentration was adjusted to 40 mg ml^{-1} .

Crystallization, Data Collection, and Structure Solution of *C. thermocellum* CE4—Crystals of CtCE4 were grown by vapor-phase diffusion using the hanging drop method with an equal volume (1 μl) of protein

(40 mg ml^{-1} in water) and reservoir solution (0.2 M MgCl_2 and 20% polyethylene glycol 3350 with a final pH \sim 5.8). Crystals were harvested in rayon fiber loops, bathed in Paratone N (Hampton Research) as a cryoprotectant prior to flash freezing in liquid nitrogen. Metal ion complexes were obtained by soaking crystals in 1- μl drops of mother liquor, supplemented with either CdCl_2 or CoCl_2 to a final concentration of 5 mM. CdCl_2 data were collected using an in-house XcaliburTM PX Ultra system, consisting of an Enhance Ultra sealed tube x-ray source (Oxford Diffraction, Oxford, UK) with an Onyx CCD detector, to a resolution of 2.3 Å. All other data were collected at the European Synchrotron Radiation Facility (ESRF), Grenoble, France. Images were processed and integrated using MOSFLM (24) or DENZO (25) and scaled using SCALA or SCALEPACK with other computing using the CCP4 suite (26) unless stated otherwise. Data processing statistics are given in Table 1.

The structure of CtCE4 could not be solved using any of the deposited CE4 structures as a search model in molecular replacement calculations. Phasing information was thus obtained using single wavelength anomalous dispersion with in-house Cd^{2+} derivative data. Heavy atom sites were determined using SHELXD (27), and initial phases were calculated using SHELXE (28). The space group was determined to be $P4_12_12$, with one molecule in the asymmetric unit. Subsequently phases were recalculated using MLPHARE. Phase improvement and extension including the higher resolution CtCE4 data (containing Mg^{2+} , CtCE4- Mg^{2+}) were performed using DM (29). The initial structure was built using ARP/wARP (30) in conjunction with REFMAC (31). Five percent of the observations were set aside for cross-validation and used to monitor the refinement progress and for appropriate weighting of geometrical and temperature factor restraints. COOT (32) was used to make manual corrections to the model between further cycles of refinement using REFMAC. The CtCE4- Mg^{2+} structure was used as the starting model for refinement of the CtCE4 data containing Co^{2+} (CtCE4- Co^{2+}). The Co^{2+} complex data were handled in the same way as the native Mg^{2+} data, with the same reflections maintained for cross-validation. The structures were validated using PROCHECK (33). Final data and refinement statistics for the CtCE4 structures are shown in Table 1.

Crystallization and Data Collection of *S. lividans* CE4—Crystals of *SICE4* were grown by vapor-phase diffusion using the hanging drop method with an equal volume (1 μl) of protein (\sim 20 mg ml^{-1}) and reservoir solution (7–13% w/v polyethylene glycol 20,000, 0.2 M MgCl_2 , 0.1 M Na-acetate, pH 4.5). At this pH the enzyme retains \sim 10% of maximal activity. Data were collected on beamline ID29 at the ESRF at a wavelength of 1.0047 Å. Data were integrated to 1.8 Å resolution. Images were indexed using DENZO (25). Examination of systematic absences confirmed the space group to be $P2_1$ (*SICE4* crystal form I). The data were scaled and merged using SCALEPACK (25). Data processing statistics are given in Table 1.

A second crystal form (*SICE4* crystal form II) was obtained from 7–13% w/v polyethylene glycol 20,000, 0.2 M KBr, 0.1 M Na-acetate, pH 4.5, with protein at \sim 20 mg ml^{-1} . Crystals were mounted and collected to 1.6 Å on beamline ID29 at the ESRF, at a wavelength of 0.979 Å. Data were processed using DENZO and SCALEPACK, as described above. In this case space group identification was ambiguous, as systematic absences along all three crystal axes suggested the space group to be $P2_12_12_1$, but a strong non-origin peak in the native Patterson indicated two molecules of similar orientation separated by a vector of 0.0 0.5 0.37 (fractional coordinates). The presence of this translational vector made assignment of the space group ambiguous in terms of the presence of

Two Metal Ion-dependent Acetylxylan Esterases

FIGURE 1. The structure of xylan and mechanisms for its breakdown. A, schematic diagram of xylan, indicating the diverse enzymes required for its complete hydrolysis. B, *N*-acetylglucosamine, the basic repeat unit of peptidoglycan and chitin, indicating the position of attack of chitin and peptidoglycan deacetylases. C, the basis for the linked assay used in this study, in which esterase-catalyzed acetate release is linked to the generation of a viable and colorimetric substrate for a β -xylosidase.

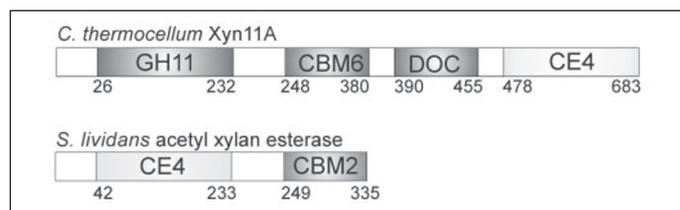
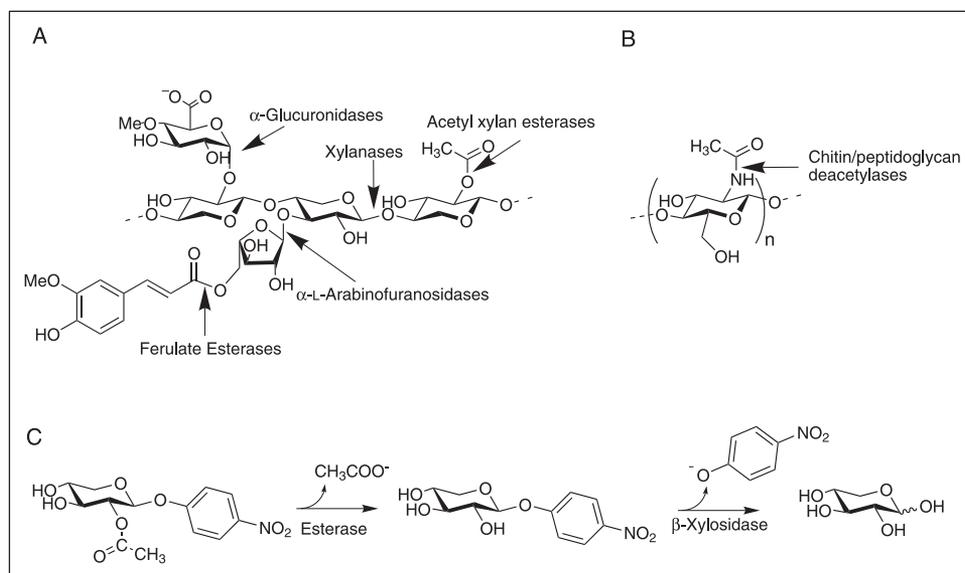


FIGURE 2. Domain organization of the *C. thermocellum* Xyn11A and *S. lividans* acetylxylan esterase whose CE4 domains were cloned and expressed independently for this study.

the screw axis along the *b* axis, which was subsequently only resolved during molecular replacement.

Molecular Replacement and Structure Solution of *SICE4*—A partially refined model of *CtCE4* was taken as a starting model for structure solution of crystal form I of *SICE4* by molecular replacement using the program AMoRe (26). The asymmetric unit was most likely to contain two molecules, giving a Matthews coefficient of $2.5 \text{ \AA}^3 \text{ Da}^{-1}$ and corresponding to a solvent content of $\sim 50\%$. The best two-molecule solution was used as the starting model for refinement and remodeling using ARP/wARP (30) in conjunction with REFMAC (31). The standard ARP/wARP protocol “automated model building starting from existing model” as implemented in the CCP4 interface failed to refine this model, and no peptides were built automatically into the map. As an alternative approach the initial model was used in an extensive RESOLVE/REFMAC protocol as distributed by the program author as RESOLVE_BUILD (34, 35). This built a partial model consisting of 1939 protein atoms, which was input into ARP/wARP and resulted in a model consisting of 2782 protein atoms and 450 “dummy” solvent molecules. Refinement was completed using REFMAC interspersed with manual model building using COOT (32).

SICE4 crystal form II was solved using the model derived from crystal form I using the program MOLREP (36) in the CCP4 suite. Molecular replacement confirmed the space group as $P2_12_12_1$ with two similarly orientated molecules separated by a translation-only vector, as described above. Refinement was performed using REFMAC interspersed with manual model building in COOT. In both crystal forms of *SICE4* a single metal ion was identified in the active center, which was modeled as Zn^{2+} (following metal ion identification by inductively coupled plasma optical emission spectroscopy, data not shown) with 0.5 occupancy. A single acetate molecule coordinated to the metal ion was

readily identified in both crystal forms. Final data and refinement statistics for *SICE4* are shown in Table 1.

Kinetics—To probe metal ion dependence, the catalytic activity of the CE4 enzymes from *S. lividans* and *C. thermocellum* was determined using an enzyme-coupled assay, Fig. 1C, with 2-*O*-acetyl-4-nitrophenyl β -*D*-xylopyranoside as substrate. Reactions were performed at pH 5.5 in 0.1 M Na-phosphate buffer. In this assay, acetate release results in the formation of PNP- β -*D*-xylopyranoside, which is a viable substrate for the cleaving enzyme β -xylosidase (22). Substrate concentrations from 0.1 to 3 times the K_m were used for determination of the catalytic constants of *SICE4*. For *CtCE4*, the K_m was too high to permit saturation, and the kinetics represent k_{cat}/K_m only. Metal ion dependence was performed in two ways, either by incubating the enzyme with 1 mM EDTA and metals subsequently added to a final concentration of 2 mM, or by first dialyzing the enzymes against 4 mM EDTA three times, washing in double-distilled water, and determining kinetics with and without a range of metal ions. Protein concentrations were $\sim 10^5$ -fold below the concentration of the added metal. Preliminary experiments with the β -xylosidase alone showed that enzyme activity was strongly inhibited by Hg^{2+} and Sn^{2+} , and these metals were omitted from subsequent experiments with the acetylxylan esterases. Ni^{2+} partially inhibits the xylosidase, and this was considered when assaying the esterases.

RESULTS AND DISCUSSION

Acetylxylan esterase domains are frequently found in multimodular plant cell wall-degrading enzyme systems. The mature *S. lividans* CE4 comprises residues 42–233 of a twin domain enzyme whose C terminus encodes a family CBM2 carbohydrate binding module (the first 41 residues encodes a cleaved signal peptide) (Fig. 2). *CtCE4* is derived from the xylanase Xyn11A of *C. thermocellum* whose full-length protein (37) comprises an N-terminal family GH11 glycoside hydrolase (endoxylanase) catalytic domain followed by a family CBM6 carbohydrate binding module, a dockerin domain, which directs the whole enzyme to the cellulosome machinery (38), and a C-terminal family 4 carbohydrate esterase (*CtCE4*) (Fig. 2). The esterase domains of both enzymes were cloned and expressed as discrete entities for this study.

FIGURE 3. The structure of *S. lividans* acetylxyln esterase, SICE4. *A*, divergent stereo ribbon representation of the SICE4 dimer. The structure of a monomer is color-ramped from N (blue) to C terminus (red) and shows residues Asp-13, His-62, and His-66 (green) and acetate (gray) in ball-and-stick representation and Zn^{2+} (cyan) and a water molecule (green) as spheres. *B*, divergent stereo ball-and-stick representation of SICE4 active site with residues shown in green and acetate in gray. Zn^{2+} (cyan) and a water molecule are shown as spheres. Observed electron density for the maximum likelihood weighted $2F_o - F_c$ map is contoured at $0.26 \text{ e } \text{\AA}^{-3}$. Figures were made using MOLSCRIPT (44) and BOBSCRIPT (45) and rendered using RASTER3D (46).

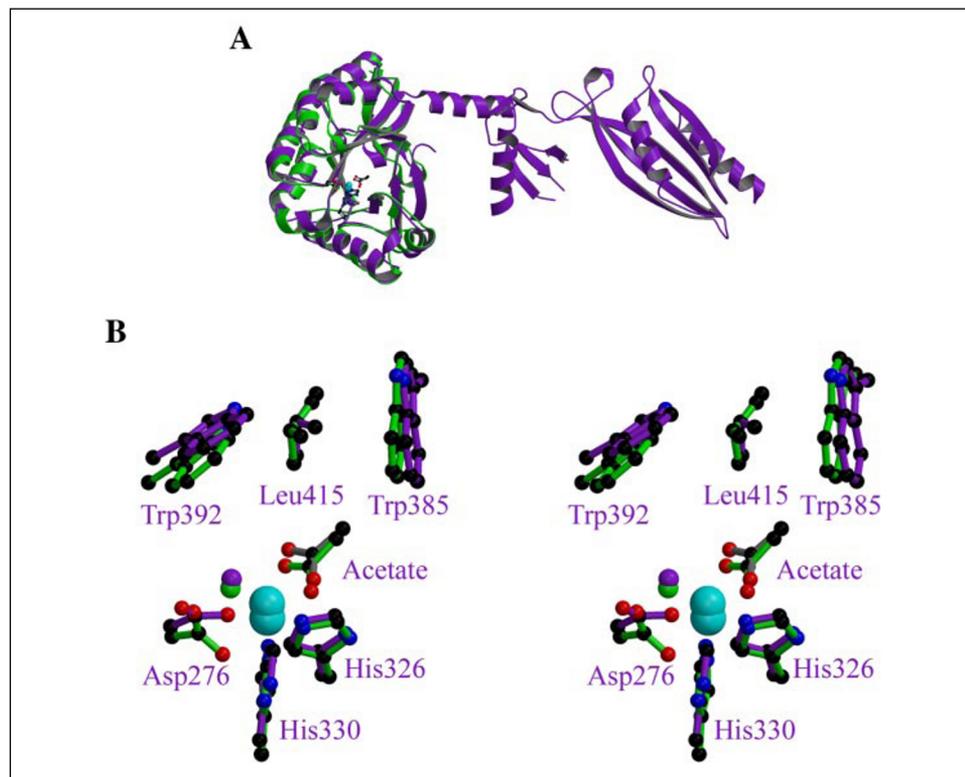
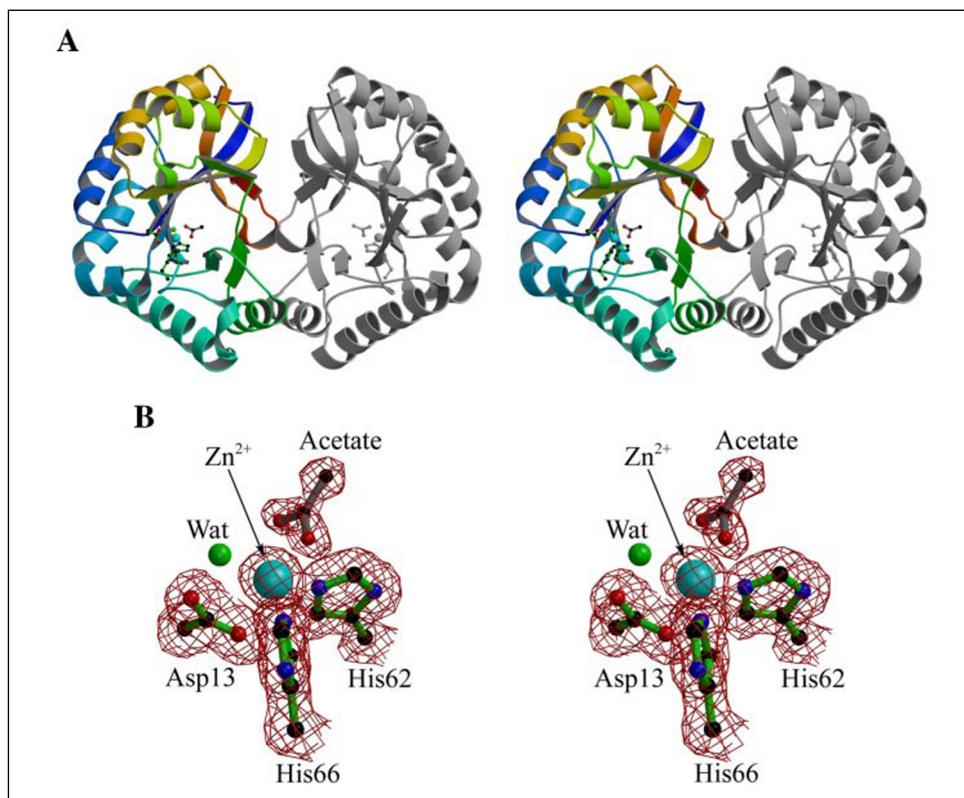


FIGURE 4. Comparison of the *S. lividans* acetylxyln esterase, SICE4, and the peptidoglycan deacetylase from *S. pneumoniae*. *A*, ribbon representation of SICE4 (green) overlapped with the peptidoglycan de-*N*-acetylase from *S. pneumoniae* (purple). Active site residues and acetate molecules (gray) are shown in ball-and-stick representation and Zn^{2+} ions as spheres. *B*, divergent stereo ball-and-stick representation of SICE4 active site residues (green) overlapped with those from the peptidoglycan de-*N*-acetylase from *S. pneumoniae* (purple and labeled); the acetate molecules are shown in green/gray. Zn^{2+} ions (cyan) and water molecules are shown as spheres. Figures were made using MOLSCRIPT (44, 45) and rendered using RASTER3D (46).

Three-dimensional Structure of the Acetylxyln Esterase from S. lividans, SICE4—The structure of the *C. thermocellum* CE4 was solved through the use of an isomorphous Cd^{2+} derivative using data from an in-house sealed tube $CuK\alpha$ source and CCD detector in harness with 1.05 \AA resolution synchrotron data. The subsequent

map interpretation was trivial, with the majority of the chain traced automatically. The CtCE4 structure was used as a molecular replacement model to solve the structure of the acetate complex of the *S. lividans* CE4 enzyme (the enzymes have 32% sequence identity).

Two Metal Ion-dependent Acetylxyylan Esterases

The structure of both esterases are best described as displaying greatly distorted (β/α)₈ barrel folds. The structure of *SICE4* can be traced from residues 2–192 (numbered from the start of the full-length protein lacking its signal peptide) with no breaks. The enzyme forms a dimer, related by a non-crystallographic 2-fold axis of symmetry, Fig. 3A, with a buried surface area of ~2400 Å². A single metal ion, modeled as a partially occupied Zn²⁺, lies on the N-terminal face of the barrel in a shallow surface groove. Crystallographic estimations of its occupancy, based upon density levels and refined temperature factor, suggest ~50% occupancy, with a resultant atomic B value of 12 Å², which is consistent with that of the protein atoms involved in its coordination. The metal ion in both crystal forms of *SICE4* is coordinated by two histidine residues (His-62 and His-66), Asp-13, a single water molecule, and an acetate molecule (Figs. 3B and 7A). The metal ion coordinates in a distorted octahedral arrangement in which the acetate ion makes a bidentate, but asymmetric, coordination to the metal with oxygen to metal distances of

2.0 and 2.5 Å. This arrangement of two histidines and an aspartate is a “classical” coordination for bivalent metal ion-dependent hydrolases (reviewed in Ref. 21) in which the metal ion gives Lewis-acid assistance to nucleophilic attack (described below).

Both the overall three-dimensional structure of *SICE4* and its metal ion coordination is extremely similar to that recently described for the catalytic NodB homology domain of the peptidoglycan de-*N*-acetylase from *Streptococcus pneumoniae* (*SpPgDA*) (19), which has 34% sequence identity (Fig. 4A). The overall structures are strikingly similar with an overall root mean square deviation of only 1.1 Å for 156 equivalent C α positions (calculated using LSQMAN (39)). Indeed, the only significant differences in the structures are at the N terminus, where the *SpPgDA* enzyme contains an additional two domains, and at the C terminus, which is involved in the dimerization of the *S. lividans* acetylxyylan esterase. The similarity extends beyond simple topological considerations into the catalytic center itself, where all metal ion interactions, including the positioning of the acetate are conserved (Fig. 4B). Furthermore, the residues lining the shallow active center substrate-binding groove are highly conserved. The hydrophobic sheath provided by the aromatic platforms of Tyr-103, Trp-124, and Trp-131 are invariant as is Leu-153 (*SICE4* numbering). In addition to the catalytic chemistry, Asp-127, involved in maintaining Trp-131, is also conserved. It is thus not surprising that the *S. lividans* acetylxyylan esterase also functions as a chitin and chitooligosaccharide de-*N*-acetylase with equal efficiency to its activity on xylan (23). *SICE4* and *SpPgDA* also show very similar metal ion preferences, as described below. *SICE4* displays topological similarity to glycoside hydrolase family GH38 (retaining α -mannosidases) with a C α root mean square deviation of 3 Å over 153 equivalent residues, but there is no conservation of metal ion location or coordination, or any conservation of mechanism or function. The classification of both CE4 and GH38 into a PFAM “glycoside hydrolase/deacetylase superfamily

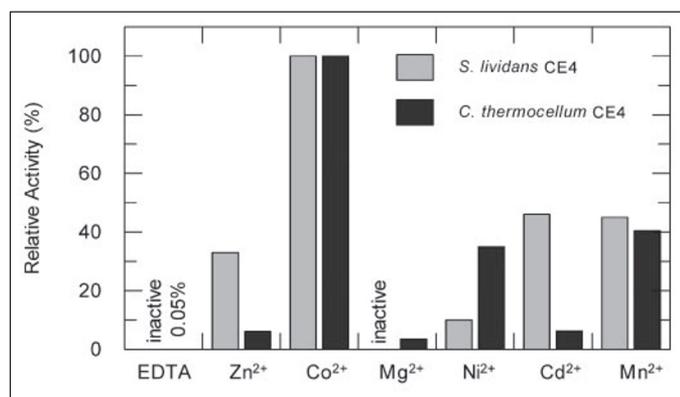


FIGURE 5. Metal ion preferences of the *S. lividans* and *C. thermocellum* family CE4 esterase domains. Activity is expressed as percent of the maximum activity obtained with the optimal metal (Co²⁺).

TABLE 1

Data collection and refinement statistics

Highest resolution shell is shown in parentheses. *CtCE4*-Cd²⁺ and *SICE4*(I) were not refined. r.m.s., root mean square.

	<i>CtCE4</i> -Cd ²⁺	<i>CtCE4</i> -Mg ²⁺	<i>CtCE4</i> -Co ²⁺	<i>SICE4</i> (I)	<i>SICE4</i> (II)
Data collection					
Space group	<i>P</i> 4 ₁ 2 ₁ 2	<i>P</i> 4 ₁ 2 ₁ 2	<i>P</i> 4 ₁ 2 ₁ 2	<i>P</i> 2 ₁	<i>P</i> 2 ₁ 2 ₁ 2 ₁
Cell dimensions					
<i>a</i> , <i>b</i> , <i>c</i> (Å)	106.1, 106.1, 35.5	106.1, 106.1, 35.5	105.8, 105.8, 35.4	57.4, 48.0, 75.0	95.5, 47.9, 89.8
α , β , γ (°)	90, 90, 90	90, 90, 90	90, 90, 90	90, 92, 90	90, 90, 90
X-ray source	XcaliburTM PX CuK α	ESRF ID14.4	ESRF ID14.4	ESRF ID29	ESRF ID29
Wavelength (Å)	1.5418	0.97560	0.9393	1.0047	0.979
Resolution (Å)	19.63–2.30 (2.42–2.30)	30–1.05 (1.07–1.05)	37.40–1.50 (1.58–1.50)	30–1.80 (1.86–1.80)	40–1.60 (1.66–1.60)
<i>R</i> _{merge}	0.037 (0.073)	0.052 (0.40)	0.045 (0.12)	0.07 (0.39)	0.053 (0.26)
$\langle I/\sigma I \rangle$	26.0 (16.1)	20.8 (3.5)	26.4 (13.0)	19 (3.5)	17 (4)
Completeness (%)	98 (94)	95 (91)	100 (100)	100 (95)	98 (91)
Redundancy	4.2 (3.3)	3.8 (3.2)	7.1 (7.2)	5.7 (2.9)	3.4 (2.8)
Refinement					
No. reflections		89153	32768		51014
<i>R</i> _{work} / <i>R</i> _{free}		0.13/0.15	0.13/0.17		0.15/0.19
No. atoms					
Protein		1720	1729		3016
Ligand/ion		1	1		8 (acetate), 2 (ion)
Water		329	335		564
B-factors					
Protein		10.5	10.6		13.6
Ligand/ion		7.2	7.2		11.4 (acetate), 13.5 (ion)
Water		23.7	23.4		30
r.m.s. deviations					
Bond lengths (Å)		0.017	0.018		0.013
Bond angles (°)		1.71	1.67		1.45
% Residues in most favored region of Ramachandran plot		91	92		92
% Residues in disallowed region of Ramachandran plot		0	0.6		0
PDB code		2C71	2C79		2CC0

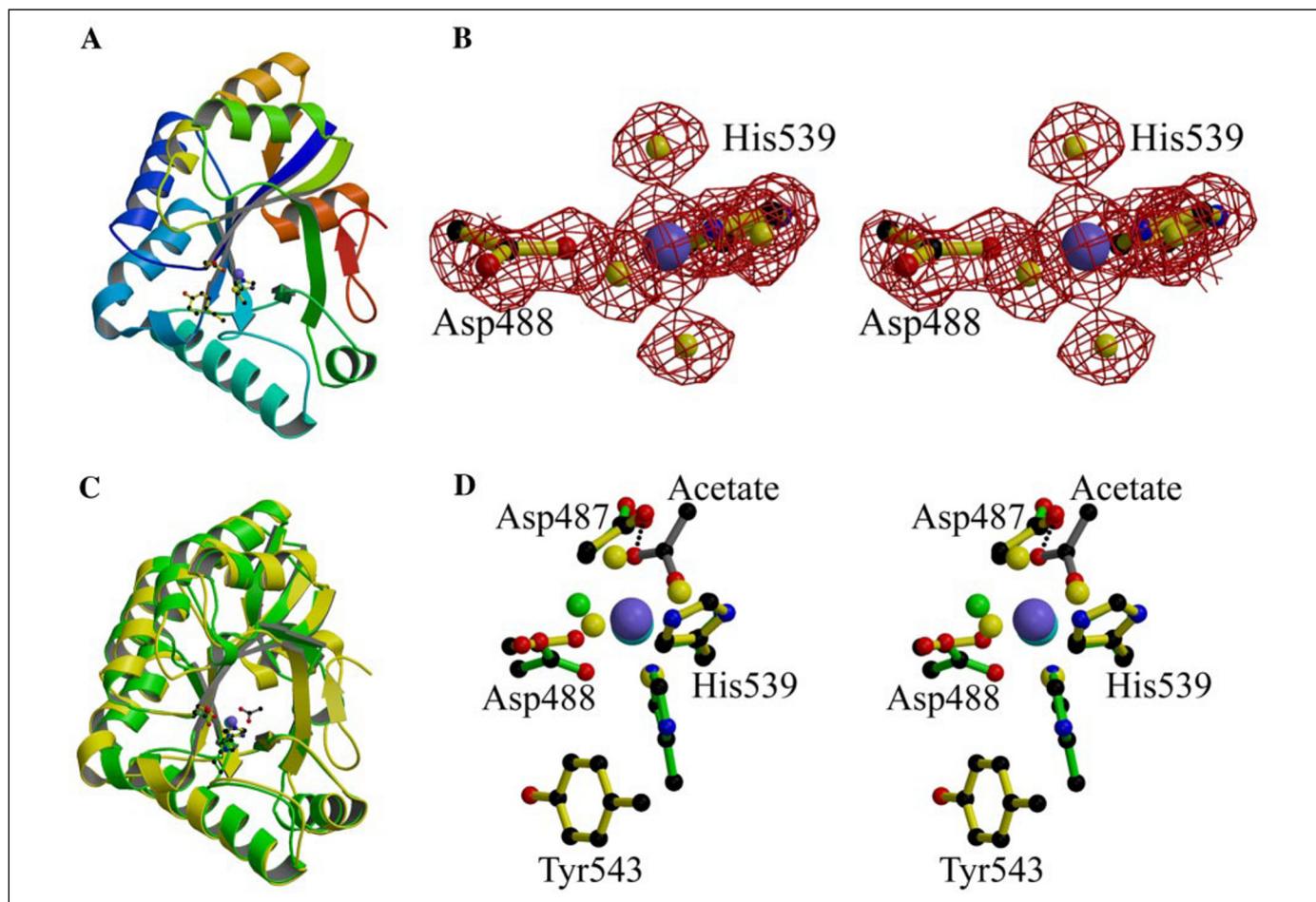


FIGURE 6. The structure of *C. thermocellum* acetylglucan esterase, CtCE4, and a comparison with the *S. lividans* CE4 enzyme. A, ribbon representation of CtCE4. The structure is color-ramped from N (blue) to C terminus (red) and shows residues Asp-488, His-539, and Tyr-543 in ball-and-stick representation and Co^{2+} (dark blue) and a water molecule (yellow) as spheres. B, divergent stereo ball-and-stick representation of SICE4 active site. Zn^{2+} (dark blue) and water molecules (yellow) are shown as spheres. Observed electron density for the maximum likelihood weighted $2F_o - F_c$ map is contoured at 1.5σ ($\sim 0.75 \text{ e \AA}^{-3}$). C, ribbon representation of CtCE4 (yellow) overlapped with SICE4 (green). The active site residues and acetate molecule (gray) are shown in ball-and-stick representation and Co^{2+} (dark blue), Zn^{2+} (cyan), and water molecules are shown as spheres. D, divergent stereo ball-and-stick representation of CtCE4 (yellow) active site residues overlapped with those from SICE4 (green); the acetate molecule is shown in gray. Co^{2+} (dark blue), Zn^{2+} (cyan), and water molecules are shown as spheres. Figures were made using MOLSCRIPT (44) and BOBSCRIPT (45) and rendered using RASTER3D (46).

clan" may well reflect an evolutionary link between the two classes of enzyme, but it provides little predictive function for open reading frame annotation; indeed it obscures this process.

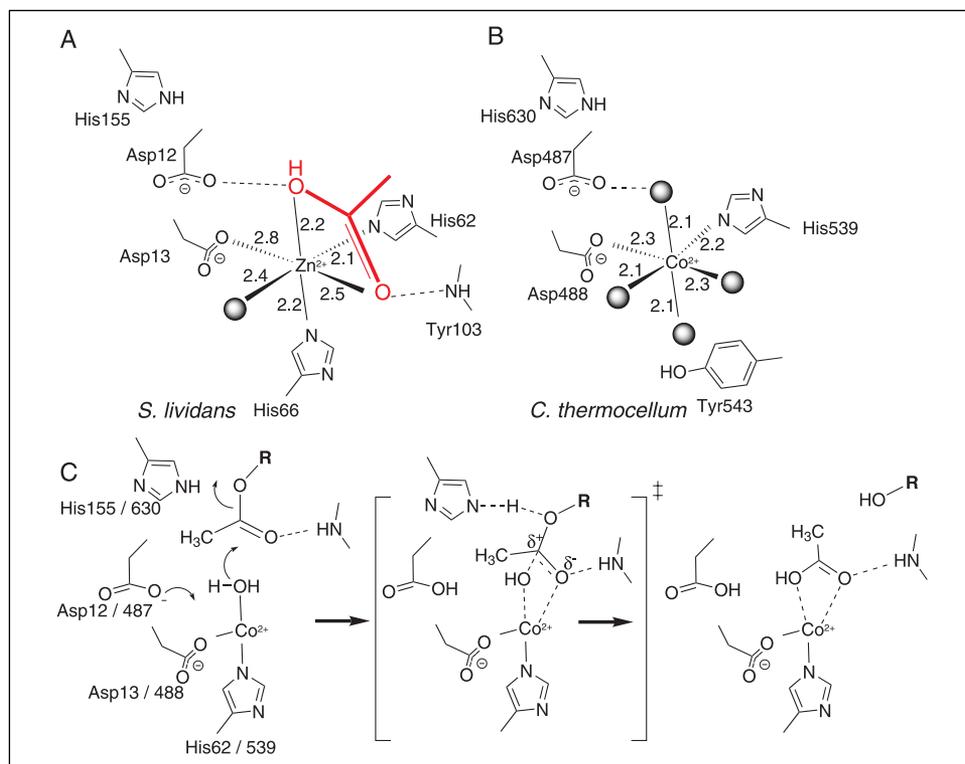
Kinetics and Metal Ion Dependence for CtCE4 and SICE4—Both CtCE4 and SICE4 show no detectable activity on generic esterase substrates including *para*-nitrophenyl acetate. Both enzymes are specific for sugar-based substrates and will precipitate acetylglucan, as a consequence of deacetylation. Kinetics were instead determined using a novel linked enzyme system in which acetate is released from 2-*O*-acetyl-4-nitrophenyl β -D-xylopyranoside, leading to the formation of PNP- β -D-xylopyranoside, which is subsequently a viable substrate for the cleaving β -xylosidase (22) (Fig. 1C). Under these conditions, SICE4 has a k_{cat} of 118 s^{-1} and a K_m of $\sim 1 \text{ mM}$ (thus giving a k_{cat}/K_m of $118 \text{ s}^{-1} \text{ mM}^{-1}$). CtCE4 has a k_{cat}/K_m of $120 \text{ s}^{-1} \text{ mM}^{-1}$, however the K_m was too high ($>12 \text{ mM}$) to determine accurately so precise kinetic constants cannot be evaluated. Although the *Streptomyces* enzyme is able to de-*N*-acetylate chitooligosaccharides (as described previously (23)), CtCE4 is considerably less efficient in this regard, with acetate release from chitooligosaccharides below detectable levels.

Both CtCE4 and SICE4 are metal ion dependent; either addition of 1 mM EDTA or prior dialysis against EDTA completely abolishes the

activity of SICE4 and reduces the activity of the clostridial enzyme to $\sim 0.05\%$ of the maximum (defined here and throughout as the maximal activity with the preferred metal) (Fig. 5). For both enzymes, as with the peptidoglycan deacetylases (19), Co^{2+} is the optimal metal, although bioavailability does not necessarily mean that Co^{2+} is the preferred metal ion *in vivo*. The two enzymes display quite surprising differences in metal ion-dependent activity (which may reflect their distinct metal ion coordination of different ligands/residues observed in the three-dimensional structures). Both enzymes will tolerate Mn^{2+} to give ~ 40 – 45% of maximal activity. As with the peptidoglycan deacetylases (19), the *Streptomyces* enzyme has $\sim 33\%$ of maximal activity with Zn^{2+} , whereas the *C. thermocellum* enzyme displays only 6% of maximal activity with this metal. A similar difference is observed with Cd^{2+} ; 46% of the maximal activity is displayed by SICE4 but only 6% for CtCE4. Conversely, both Ni^{2+} and Mg^{2+} are tolerated better by the clostridial enzyme (35 and 3.5% activity, respectively) compared with the *Streptomyces* enzyme, which is only 10% active with Ni^{2+} and completely inactive with Mg^{2+} . As purified and without EDTA treatment or the addition of supplementary metal ions, SICE4 displays $\sim 54\%$ of its maximal activity, which is consistent with the metal ion occupancy (estimated at 50%) observed in the x-ray structure of the enzyme.

Two Metal Ion-dependent Acetylxyylan Esterases

FIGURE 7. Active site interactions and catalysis in family CE4 acetylxyylan esterases. Schematic diagram of the metal ion coordination of the *S. lividans* family CE4 esterase (acetate shown in red) (A) and the *C. thermocellum* CE4 esterase (B) and a putative reaction mechanism (C) based upon classical Zn^{2+} hydrolase chemistry and the work of van Aalten and colleagues (19) on the streptococcal peptidoglycan deacetylases. The divalent metal plays the role of Lewis acid, with Asp and His residues playing the roles of catalytic base (activating the nucleophilic water) and acid (aiding sugar departure).



Three-dimensional Structure of the Acetylxyylan Esterase from *C. thermocellum*, CtCE4—The difference in metal ion preference and the lack of activity on *N*-acetyl substrates prompted investigation of the three-dimensional structure of the cellulosomal acetylxyylan esterase from *C. thermocellum* Xyn11A (this was actually the first of the two structures determined, as screening for derivatives of the *S. lividans* enzyme failed to yield a solution). CtCE4 was subsequently refined in complex with Mg^{2+} , with data to 1.05 Å resolution, and with the favored metal Co^{2+} , with data to 1.5 Å resolution (Table 1). The chain may be traced from residues 480–684 (numbering based upon the full-length Xyn11A; residue 684 refers to the first residue of the C-terminal His₆-tag attached to CtCE4) with no breaks. CtCE4 is monomeric, as expected for a module that is part of a five-domain protein that is targeted to the cellulosome apparatus of *C. thermocellum*. The overlap with SICE4 gives a root mean square deviation of 1.2 Å for 170 equivalent C α atoms. Yet, although CtCE4 has the same overall topology and fold as SICE4 and the published peptidoglycan deacetylases, Fig. 6, there are marked differences in the topography of the substrate-binding groove, in the chemistry of the active center and in the metal ion coordination itself.

CtCE4 contains a single metal site; both Co^{2+} and Mg^{2+} complexes display octahedral coordination, but in contrast to the vast majority of the sequence family and the previous three-dimensional structures, one of the histidine ligands is absent and coordination is instead formed by Asp-488, His-539, and four solvent water molecules (Figs. 6 and 7). Soaks with acetate did not, in the case of CtCE4, yield a complex with acetate bound. The difference in coordination results from a tyrosine substitution for histidine in CtCE4 compared with SICE4, in which the tyrosine (Tyr-543) perhaps forms an aromatic platform for ligand binding in the substrate-binding cleft. Such a difference is likely, on the basis of sequence alone, to exist only in the closely related *Clostridium cellulovorans* CE4 and the enzyme from *Thermotoga maritima*, which display 60 and 54% identity with the *C. thermocellum* enzyme, respectively. Such differences in metal ion coordination may be responsible, in part at

least, for the unusual differences in metal ion specificity of the *Streptomyces* and *Clostridium* enzymes discussed previously (Fig. 5).

Other differences manifest themselves in the putative substrate-binding cleft. A tryptophan, equivalent to Trp-131 of SICE4, is found in CtCE4 (although with the side chain is rotated through 90°), but other aromatics, equivalent to Tyr-103 and Trp-124 of SICE4, are not present. Such differences in topography may be reflected in the near inactivity of the clostridial enzyme on chitooligosaccharides.

Catalytic Mechanism of CE4 Acetylxyylan Esterases—Metal ion, predominantly Zn^{2+} , hydrolases are well described in the literature (reviewed in Ref. 21). Their mechanism (Fig. 7C) involves the base-catalyzed generation of a hydroxide ion stabilized, in a Lewis-acid sense, by the metal that then undergoes a direct nucleophilic attack at the electrophilic carbon center concomitant with leaving group departure. Departure of the leaving group may, if necessary, receive enzymatic general acid assistance. Both the Mg^{2+} and Co^{2+} complexes of CtCE4 and the acetate/ Zn^{2+} complex of SICE4 are certainly consistent with such a mechanism, and following proposals made for the peptidoglycan de-*N*-acetylases (19, 20), we would suggest that acetylxyylan ester hydrolysis involves general base activation of water, with Asp-488 (CtCE4)/Asp-13 (SICE4) playing the role of the Brønsted base with acid-catalyzed assistance to the departure of the xylose moiety provided by His-539 (CtCE4)/His-66 (SICE4). Site-directed mutagenesis of the SpPgDA enzyme has shown that both His and Asp are indeed essential for catalysis (19). A main-chain amide group would stabilize the developing oxyanion (Tyr-103 in SICE4/Asn-580 in CtCE4).

The metal ion in SICE4 coordinates a single solvent water as part of its distorted octahedral coordination. In the peptidoglycan deacetylase work, van Aalten and colleagues (19, 20) produced a docked model of how a chitotriose substrate might bind in the active center. A feature of this model is the location of the vicinal O-3 hydroxyl group as a Zn^{2+} -coordinating atom, in place of this single solvent water. Such a prediction, that correct binding of substrate might require an O-3-metal²⁺

interaction, is fully consistent with both the specificity of this family of esterases for sugar-derived (as opposed to “generic”) esters, and the observation, in the case of *SlCE4* (40), that catalysis demands a free vicinal OH group adjacent to the ester.

A description of the metal ion geometry or any changes in such coordination during catalysis is difficult. The streptococcal enzyme displays octahedral metal ion coordination in its acetate complex, yet it is tetrahedral in complex with sulfate (19), suggesting that changes in geometry may occur during catalysis. Both the clostridial and *Streptomyces* esterases described here have octahedral geometry despite differences in ligation (water and acetate, respectively) and coordination chemistry. Zn²⁺ hydrolases are frequently described as having “tetrahedral” geometry but, should the O-3 participate in direct coordination, as kinetics (40) and modeling (19) suggest, then tetrahedral coordination is unlikely without displacement of one of the protein-derived ligands.

Summary—The *Streptomyces* and clostridial acetylxyylan esterases provide divergent solutions to the problems of ester hydrolysis on xylan. Both enzymes share the same fold and catalytic mechanism, but marked differences in both metal ion ligand coordination and in the surface topography have evolved. The *S. lividans* enzyme is almost indistinct from known chitin/peptidoglycan deacetylases and is indeed as efficient on those substrates as on xylan itself; the enzyme is, however, most likely to function in xylan degradation *in vivo*. Annotation of this enzyme as an acetylxyylan esterase, as opposed to chitin deacetylase, was originally based upon its chromosomal organization (which is immediately downstream from the xylanase *xyn10b* gene of *S. lividans*) and upon its enhanced expression when the organism is cultured on birchwood xylan (41). Further evidence for a role in xylan degradation comes from the observation that although the enzyme is appended to a CBM2 domain whose function could be the binding of crystalline substrates (cellulose or chitin, reviewed in Ref. 42), the sequence of the CBM is most characteristic of a CBM2b xylan binding domain. In CBM2b xylan-specific domains, a tryptophan is “substituted” for an arginine, a “polymorphism” that allows CBM2b modules to bind to xylan, as opposed to crystalline cellulose or chitin (43).

Family CE4 is the largest of the “true” (as opposed to multifunctional lipase/esterase) carbohydrate esterase/deacetylase families. Its members play key roles ranging from chitin and xylan degradation through to the remodeling of peptidoglycan. We have shown that the xylan esterase members of this growing family are, like other characterized members, metal ion-dependent with a chemical preference for cobalt. Metal ion coordination is, oddly, not conserved, with a histidine present in most members of the family, but which is absent in the clostridial acetylxyylan esterase. In a recent structural genomics deposition of an unknown open reading frame from *Pseudomonas aeruginosa* (PDB code 1Z7A) a tryptophan lies in the equivalent position (although one cannot confirm direct metal ion coordination as the structure was solved in the absence of metal). CE4 members thus show diversity both in the residues that coordinate the metal, as illustrated by the acetylxyylan esterases described here, and in different metal ion specificities. These diversities may, in the future, be exploited in enzymes used in biotechnological processes for biomass conversion, where the metal ion requirements of different enzymes are an important concern.

REFERENCES

- Kosugi, A., Murashima, K., and Doi, R. H. (2002) *Appl. Environ. Microbiol.* **68**, 6399–6402
- Altaner, C., Saake, B., Tenkanen, M., Eyzaguirre, J., Faulds, C. B., Biely, P., Viikari, L., Siika-aho, M., and Puls, J. (2003) *J. Biotechnol.* **105**, 95–104
- Bornscheuer, U. T. (2002) *FEMS Microbiol. Rev.* **26**, 73–81
- Collins, T., Gerday, C., and Feller, G. (2005) *FEMS Microbiol. Rev.* **29**, 3–23
- Mathew, S., and Abraham, T. E. (2004) *Crit. Rev. Biotechnol.* **24**, 59–83
- Henrissat, B., Coutinho, P. M., and Davies, G. J. (2001) *Plant Mol. Biol.* **47**, 55–72
- Davies, G. J., Gloster, T. M., and Henrissat, B. (2005) *Curr. Opin. Struct. Biol.* **15**, 637–645
- Schubot, F. D., Kataeva, I. A., Blum, D. L., Shah, A. K., Ljungdahl, L. G., Rose, J. P., and Wang, B. C. (2001) *Biochemistry* **40**, 12524–12532
- Prates, J. A. M., Tarbouriech, N., Charnock, S. J., Fontes, C., Ferreira, L. M. A., and Davies, G. J. (2001) *Structure* **9**, 1183–1190
- Tarbouriech, N., Prates, J. A. M., Fontes, C., and Davies, G. J. (2005) *Acta. Crystallogr. Sect. D. Biol. Crystallogr.* **61**, 194–197
- McAuley, K. E., Svendsen, A., Patkar, S. A., and Wilson, K. S. (2004) *Acta. Crystallogr. Sect. D. Biol. Crystallogr.* **60**, 878–887
- Hermoso, J. A., Sanz-Aparicio, J., Molina, R., Juge, N., Gonzalez, R., and Faulds, C. B. (2004) *J. Mol. Biol.* **338**, 495–506
- Crepin, V. F., Faulds, C. B., and Connerton, I. F. (2004) *Appl. Microbiol. Biotechnol.* **63**, 647–652
- Hakulinen, N., Tenkanen, M., and Rouvinen, J. (2000) *J. Struct. Biol.* **132**, 180–190
- Ghosh, D., Sawicki, M., Lala, P., Erman, M., Pangborn, W., Eyzaguirre, J., Gutierrez, R., Jornvall, H., and Thiel, D. J. (2001) *J. Biol. Chem.* **276**, 11159–11166
- Vincent, F., Charnock, S. J., Verschuere, K. H. G., Turkenburg, J. P., Scott, D. J., Offen, W. A., Roberts, S., Pell, G., Gilbert, H. J., Davies, G. J., and Brannigan, J. A. (2003) *J. Mol. Biol.* **330**, 593–606
- Hekmat, O., Tokuyasu, K., and Withers, S. G. (2003) *Biochem. J.* **374**, 369–380
- Psylinakis, E., Boneca, I. G., Mavromatis, K., Deli, A., Hayhurst, E., Foster, S. J., Varum, K. M., and Bouriotis, V. (2005) *J. Biol. Chem.* **280**, 30856–30863
- Blair, D. E., Schuettelkopf, A. W., MacRae, J. I., and van Aalten, D. M. (2005) *Proc. Natl. Acad. Sci. U. S. A.* **102**, 15429–15434
- Blair, D. E., and van Aalten, D. M. F. (2004) *FEBS Lett.* **570**, 13–19
- Hernick, M., and Fierke, C. A. (2005) *Arch. Biochem. Biophys.* **433**, 71–84
- Biely, P., Mastihubova, M., la Grange, D. C., van Zyl, W. H., and Prior, B. A. (2004) *Anal. Biochem.* **332**, 109–115
- Caufrier, F., Martinou, A., Dupont, C., and Bouriotis, V. (2003) *Carbohydr. Res.* **338**, 687–692
- Leslie, A. G. W. (1992) in *Joint CCP4 and ESF-EACMB Newsletter on Protein Crystallography*, Vol. 26, Daresbury Laboratory, Warrington, UK.
- Otwinowski, Z., and Minor, W. (1997) *Methods Enzymol.* **276**, 307–326
- Collaborative Computational Project Number 4. (1994) *Acta. Crystallogr. Sect. D. Biol. Crystallogr.* **50**, 760–763
- Schneider, T. R., and Sheldrick, G. M. (2002) *Acta. Crystallogr. Sect. D. Biol. Crystallogr.* **58**, 1772–1779
- Uson, I., and Sheldrick, G. M. (1999) *Curr. Opin. Struct. Biol.* **9**, 643–648
- Cowtan, K. D., and Main, P. (1996) *Acta. Crystallogr. Sect. D. Biol. Crystallogr.* **52**, 43–48
- Perrakis, A., Morris, R., and Lamzin, V. S. (1999) *Nat. Struct. Biol.* **6**, 458–463
- Murshudov, G. N., Vagin, A. A., and Dodson, E. J. (1997) *Acta. Crystallogr. Sect. D. Biol. Crystallogr.* **53**, 240–255
- Emsley, P., and Cowtan, K. (2004) *Acta. Crystallogr. Sect. D. Biol. Crystallogr.* **60**, 2126–2132
- Laskowski, R. A., MacArthur, M. W., Moss, D. S., and Thornton, J. M. (1993) *J. Appl. Crystallogr.* **26**, 283–291
- Terwilliger, T. C., and Berendzen, J. (1999) *Acta. Crystallogr. Sect. D. Biol. Crystallogr.* **55**, 849–861
- Terwilliger, T. C. (2003) *Methods Enzymol.* **374**, 22–37
- Vagin, A., and Teplyakov, A. (1997) *J. Appl. Crystallogr.* **30**, 1022–1025
- Fernandes, A. C., Fontes, C., Gilbert, H. J., Hazlewood, G. P., Fernandes, T. H., and Ferreira, L. M. A. (1999) *Biochem. J.* **342**, 105–110
- Carvalho, A. L., Dias, F. M. V., Prates, J. A. M., Nagy, T., Gilbert, H. J., Davies, G. J., Ferreira, L. M. A., Romão, M. J., and Fontes, C. M. G. A. (2003) *Proc. Natl. Acad. Sci. U. S. A.* **100**, 13809–13814
- Kleywegt, G. J., Zou, J. Y., Divne, C., Davies, G. J., Sinning, I., Stahlberg, J., Reinikainen, T., Srisodsuk, M., Teeri, T. T., and Jones, T. A. (1997) *J. Mol. Biol.* **272**, 383–397
- Biely, P., Mastihubová, M., Cote, G. L., and Greene, R. V. (2003) *Biolchim. Biophys. Acta* **1622**, 82–88
- Shareck, F., Biely, P., Morosoli, R., and Kluepfel, D. (1995) *Gene (Amst.)* **153**, 105–109
- Boraston, A. B., Bolam, D. N., Gilbert, H. J., and Davies, G. J. (2004) *Biochem. J.* **382**, 769–781
- Xie, H. F., Gilbert, H. J., Charnock, S. J., Davies, G. J., Williamson, M. P., Simpson, P. J., Raghothama, S., Fontes, C., Dias, F. M. V., Ferreira, L. M. A., and Bolam, D. N. (2001) *Biochemistry* **40**, 9167–9176
- Kraulis, P. J. (1991) *J. Appl. Crystallogr.* **24**, 946–950
- Esnouf, R. M. (1997) *J. Mol. Graph Model* **15**, 132–134
- Merritt, E. A., and Murphy, M. E. P. (1994) *Acta. Crystallogr. Sect. D. Biol. Crystallogr.* **50**, 869–873