Correction

BIOCHEMISTRY

The authors note the following statements should be added to the Acknowledgments: “The authors thank Dr. Ethan Goddard-Borger for synthesizing the substrate 2,4-dinitrophenyl-β-D-cellotriose and potential inhibitors of CtCel124. The authors also thank Steve Withers and Gideon Davies for helpful discussions.” Additionally, the grant number 00562492 should instead appear as DE-FG02-09ER16076.

www.pnas.org/cgi/doi/10.1073/pnas.1104766108
Structural insights into a unique cellulase fold and mechanism of cellulose hydrolysis

Joana L. A. Brás1, Alan Cartmell1,4, Ana Luisa M. Carvalho2,4, Genny Verzé4,6, Edward A. Bayer1, Yael Vazana1, Márcia A. S. Correia1, José A. M. Prates1, Supriya Ratnaparkhe1, Alisdair B. Boraston2, Maria J. Romão3, Carlos M. G. A. Fonseca1, and Harry J. Gilbert2

1Centro de Investigação Interdisciplinar em Sanidade Animal, Faculdade de Medicina Veterinária, Pólo Universitário do Alto da Ajuda, Avenida da Universidade Técnica, 1300-477 Lisboa, Portugal; 2Complex Carbohydrate Research Center, University of Georgia, Athens, Georgia 30602-4712; 3Rede de Química e Tecnologia, Departamento de Química, Faculdade de Ciências e Tecnologia, Universidade Nova de Lisboa, 2829-516 Caparica, Portugal; 4Biocrystallography Laboratory, Department of Biotechnology, University of Verona, 37129 Verona, Italy; 5Department of Biological Chemistry, The Weizmann Institute of Science, Rehovot 76100 Israel; and 6Department of Biochemistry and Microbiology, University of Victoria, Victoria, BC, Canada V8W 3P6

Edited by Arnold L. Demain, Drew University, Madison, NJ, and approved February 1, 2011 (received for review October 8, 2010)

Clostridium thermocellum is a well-characterized cellulose-degrading microorganism. The genome sequence of C. thermocellum encodes a number of proteins that contain type I dockerin domains, which implies that they are components of the cellulose-degrading apparatus, but display no significant sequence similarity to known plant cell wall–degrading enzymes. Here, we report the biochemical properties and crystal structure of one of these proteins, designated CtCel124. The protein was shown to be an endo-acting cellulase that displays a single displacement mechanism and acts in synergy with Cel48S, the major cellulosomal exo-cellulase. The crystal structure of CtCel124 in complex with two cellotriose molecules, determined to 1.5 Å, displays a superhelical fold in which a constellation of α-helices encircle a central helix that houses the catalytic apparatus. The catalytic acid, Glu96, is located at the C-terminus of the central helix, but there is no candidate catalytic base. The substrate-binding cleft can be divided into two discrete topographical domains in which the bound cellotriose molecules display twisted and linear conformations, respectively, suggesting that the enzyme may target the interface between crystalline and disordered regions of cellulose.

The plant cell wall is an important biological and industrial resource. Deconstruction of this composite structure provides nutrients that are utilized by microorganisms from a variety of ecosystems. Indeed, mammalian herbivores derive a significant proportion of their energy from the hydrolysis of plant cell wall polysaccharides by their symbiotic microbiota. The deconstruction of the plant cell wall is also of growing environmental and industrial significance as the demand for renewable sources for bioenergy and substrates for the chemical industry increases (1). The major plant cell wall polysaccharide is cellulose, a β-1,4-glucose polymer (2), which is hydrolyzed by a range of glycoside hydrolases (cellulases). These enzymes display endo (endo-β1, 4-glucanase), exo (cellolibiohydrolases that release cellobiose from cellulose), or endo-processive (cleaves internally and then acts in a processive manner on the generated product) modes of action. The classical paradigm for cellulase hydrolysis is the endoglucanase-cellulohydrolase synergy model, which is based, primarily, on aerobic fungal cellulase systems (see ref. 3 for review). This model, however, does not accurately reflect clostridial systems, where endo-processive GH9 enzymes are central to the degradative process (4) (see below), or Cytophaga hutchinsonii, which appears to lack classical cellolibiohydrolases. Cellulases are currently grouped into 11 of the 12 glycoside hydrolyase sequence-based families (GHs) within the Carbohydrate-Active enZymes (CAZy) database (5). Because there is limited conservation in the catalytic apparatus and the overall fold between the cellulase-containing families, these enzymes are generally thought to have evolved by convergent evolution (3).

Clostridium thermocellum is a well-characterized cellulose-degrading microorganism (6, 7). The bacterium synthesizes a large multienzyme complex, known as the “cellulosome,” which catalyzes the degradation of the plant cell wall (6, 7). Enzymes are recruited into the C. thermocellum cellulosome through the interaction of their type I dockerin modules with the multiple cohesin domains present on the scaffoldin (defined as CipA) (reviewed in refs. 6 and 7). The genome of C. thermocellum encodes 72 proteins containing type I dockerins. These proteins, therefore, are likely to be components of the cellulosome and thus contribute to cellulose or, in a wider context, plant cell wall deconstruction. Synergy experiments have identified two cellulosomal enzymes, an exo-acting GH48 cellolibiohydrolase that acts from the reducing end of cellulose chains and the cellotetraose-producing endo-processive GH9 endoglucanase, Cel9R (8), as central components of the C. thermocellum cellulosome (9). It has been suggested that, by generating cellotetraose as the major product from cellulose, likely through the action of Cel9R, C. thermocellum minimizes the utilization of ATP during import of glucose units (10). The cellulase activity obtained by combining cellulosomal enzymes in vitro, however, is considerably lower than that displayed by the cellulosome presented on the surface of C. thermocellum. It is possible, therefore, that proteins, currently of unknown function, either in the cellulosome or displayed on the surface of C. thermocellum, make a significant contribution to the cellulose-degrading capacity of the bacterium. Although most of the cellulosomal proteins can be assigned to glycoside hydrolase, esterase, or polysaccharide lyase families, 14 of the predicted type I dockerin-containing proteins display little sequence similarity to enzymes in the CAZy database. Thus, some of these non-CAZy C. thermocellum proteins may comprise novel enzymes that target the hydrolysis of components of the plant cell wall such as cellulose. One of these hypothetical proteins, Cte_0435 (hereafter designated as CtCel124), is upregulated when C. thermocellum is cultured on crystalline cellulose (11), suggesting that CtCel124 plays a role in the hydrolysis of the glucose polymer.

Here, we show that CtCel124 is an endoglucanase that acts in synergy with the major exocellulase of the C. thermocellum cel-
lulose to degrade crystalline cellulose. The crystal structure of CtCel124 reveals a substrate-binding cleft in which the bound celloligosaccharides adopt two distinct conformations, indicating that the enzyme targets the interface between crystalline and amorphous regions of cellulose. The active site of the cellulase displays structural conservation to GH23 enzymes, a family that contains inverting lysozymes and lytic transglycosylases.

**Results and Discussion**

**Catalytic Properties of CtCel124.** CtCel124 is highly upregulated when *C. thermocellum* is cultured on crystalline cellulose (11), suggesting the protein may contribute to the metabolism of the polysaccharide. To test this hypothesis, the biochemical properties of the 220 residue C-terminal module of the protein (designated 

\[
\text{CtCel124}_{\text{CD}}
\]

was assessed. The data, summarized in Table 1, show that the enzyme hydrolyzed barley α-glucan, a β-1,3-β-1,4 mixed linked glucan, phosphoric acid swollen cellulose (PASC), and carboxymethylcellulose. The specific activity of the enzyme against β-glucan was only fourfold higher than the value for PASC. The initial reaction products released from PASC ranged from cellotriose to cellohexaose (Fig. 1). Such a profile is typical of endo-acting enzymes, and thus CtCel124 appears to be an endo-β-1,4-glucanase. The difference in activity between the soluble and insoluble polysaccharides is relatively modest compared to, for example, GH5 endoglucanases and GH9 endo-processive endoglucanases, which generally display a much stronger preference for β-glucan (12).

To explore the capacity of CtCel124 to disrupt the plant cell wall, the catalytic module was incubated with sections of *Arabidopsis* stem, which were subsequently stained by Calcofluor White. This stain is predominantly to cellotriose, whereas sections of cellopentaose were incubated with 2 μM of CtCel124, and at various time points the release of cellooligosaccharides was analyzed by HPAEC. (B) The hydrolysis of sections of *Arabidopsis* stem tissue. Sections 1 and ii are untreated, whereas 1 and iv were incubated with 2 μM CtCel124CD for 16 h. Sections 1 and ii were stained with Calcofluor White, whereas sections iii and iv were probed with CBM9 fused to GFP. (C) The kinetics of Avicel hydrolysis by CtCel124CD and Cel48S. Dashed lines indicate a single enzyme activity, and solid lines indicate activity of the two enzymes in combination. In Cel48S CBM3a-Ch, Cel48S (which contains a type I dockerin) was preincubated with CBM3a-Ch (CBM3a fused to a type I cohesin) to generate Cel48S attached to CBM3a through the dockerin-cohesin interaction.
lose (see ref. 3 for review). In addition, these enzymes normally contain cellulose-specific CBMs that potentiate catalysis by recruiting the cellulases to the surface of the insoluble substrate (15). In the C. thermocellum cellulosome, the most abundant exo-acting cellulase is Cel48S, whereas the crystalline cellulose-specific CBM (CBM3a) is supplied by the noncatalytic scaffoldin CipA (6, 7). To explore the possible synergy between C.thermol. and Cel48S, the capacity of the enzymes, individually and in combination, to release reducing sugar from Avicel was assessed. The data (Fig. 1) showed that 1.3-fold more reducing sugar was released when the two enzymes were used in combination, compared to the additive value when the two enzymes were used in isolation. These data indicate that C. thermol. and Cel48S exhibit a degree of synergy when acting on highly crystalline cellulose. When both enzymes were appended to CBM3a, which binds to amorphous regions of cellulose creating new termini from which energy model does not fully explain the capacity of the C. thermocellum cellulosome to completely solubilize crystalline cellulose. It should be emphasized that although C.thermol. contains a type I dockerin, the module displays a preference for the cohesin in the cell-envelope protein OlpC, rather than the cohesins in the cellulosome scaffold protein, CipA (18). These data indicate that the enzyme is predominantly located at the cell surface of the bacterium. Lynd and coworkers showed that the cellulosome, when appended to the surface of C. thermocellum, is more efficient at cellulose degradation than when the complex is released into the culture media (19). Thus, it is possible that this increased activity reflects synergistic interactions between catalytic components of the cellulosome and enzymes, such as Cel124, which are predicted to be directly appended to the surface of the bacterium. For example, C.thermol. may create the chain ends at the amorphous-crystalline interface (see below) that are required by the cellulosomal exo-acting enzymes to hydrolyze cellulose.

Crystal Structure of C.thermol. The X-ray crystal structure of C.thermol. was determined in complex with two cellotriose (which arose through the hydrolysis of cellohexaose during crystallization). The crystal structure revealed a 210 amino acid α-helical protein containing eight α-helices and a small β-sheet comprising three antiparallel β-strands. α-Helix 4 (α-H4) forms the hydrophobic core of the protein, and the other seven helices encircle the core helix (Fig. 2). Thus, C.thermol. appears to display an αβ superhelical fold. Such a structure has not previously been observed in cellulase families, which display the following folds: (β/α)8-barrel (GH5, GH51, and GH44), distorted αβ-barrel (GH6), β-jelly roll fold (GH7 and GH12), (α/α)6-barrel (GH8, GH9, and GH48), β0-barrel (GH45), and sevenfold β-propeller (GH74). It would appear, therefore, that the different cellulase families are in general the result of (functional) convergent evolution, a view reinforced by the superhelical fold displayed by the endoglucanase C.thermol.

Structural Similarity of C.thermol. to Other Glycoside Hydrolases. A BLAST search reveals four proteins in the UNIPROT database that display significant sequence identities (>44%) with e values <e-30 (Fig. S3). We therefore propose that C.thermol. comprises the founding member of a CAZy family designated GH124. These homologous enzymes are derived from highly cellulolytic
organisms that also assemble its plant cell wall degrading apparatus into cellulosomal structures. 3D structural comparisons with the Protein Data Bank (PDB) reveal that black swan lysozyme G (Lyz23), a member of GH23, displays the closest structural similarity to CtCel124<sub>CD</sub>. The proteins have a root-mean-square deviation of 2.1 Å for 115 aligned residues, which display 21% sequence identity (7% identity when the complete catalytic modules were compared). The secondary structural elements of CtCel124, apart from α-1 and α-3, are conserved in Lyz23 (Fig. 3). The conformation and sequence of the critical loop connecting α-4 and α-5, which extends along one face of the substrate-binding cleft of CtCel124 (see below) is not conserved in Lyz23. Furthermore, an additional helix, present in Lyz23 but not in CtCel124, would make steric clashes with the Glc at the +3 subsite. It is evident, therefore, that significant differences in the topology and the residues in the substrate-binding cleft of CtCel124 and Lyz23 explains why these two enzymes display very different substrate specificities.

**Active Site and Catalytic Mechanism of CtCel124.** The active site of CtCel124 displays a high degree of structural similarity with Lyz23, which hydrolyzes glycosidic bonds through a single displacement (inverting) mechanism, and the *Escherichia coli* GH23 lytic transglycosylase Slt70, which cleaves glycosidic bonds through a water-independent substrate-assisted mechanism that ultimately leads to the formation of a 1,6-anhydro product (Fig. 3) (see Fig. S4 for details of the mechanism). Thus, Glu96, Glu117, Tyr187, Asn188, and Tyr203, which compose the central core of the active site of CtCel124, are structurally conserved in the two GH23 enzymes. It is evident, however, that the active sites of both Lyz23 and Slt70 contain an additional, but distinct, acidic residue, Asp97 and Glu583, respectively, which are lacking in the cellulase. The structural similarities and differences between the three enzymes present hypotheses regarding the catalytic mechanism of CtCel124. For example, it is possible that the cellulase cleaves glycosidic bonds through a lytic transglycosylase mechanism in which the 1,6-anhydro bond is hydrolyzed on enzyme to generate the reducing sugar products. This mechanism, however, is not supported by the observation that chemically synthesized 1,6-anhydro-cellotriose is not hydrolyzed by the cellulase, and that neither chitin nor chitosan are substrates for the enzyme (Table 1 and Fig. S5). Although the capacity of Slt70 to display a lytic activity is likely conferred, in part, by the catalytic acid-base residue Glu478 (equivalent to Glu96 in CtCel124), there is no obvious acidic residue in the cellulase or, indeed, in Slt70 capable of stabilizing the negative charge of the C2 N-acetyl group during substrate-assisted catalysis (Fig. S4). Although the structural similarity of CtCel124 and Lyz23 may indicate that the cellulase also acts through an inverting mechanism, the enzyme lacks a residue equivalent to Asp97 in Lyz23, the likely catalytic base (20), which is required to activate the water molecule that attacks C1 (Fig. S4). HPLC analysis of cellotriose, generated from cellobiose by CtCel124, however, shows that the trisaccharide was primarily in its α configuration (Fig. S6), confirming that the cellulase cleaves glycosidic bonds by an acid-base single displacement mechanism leading to anomeric inversion. The lack of a candidate Asp or Glu catalytic base is evident in some inverting glycoside hydrolases. In GH6 enzymes, the identity of the Brønsted base has remained particularly elusive, and a “Grotthus”-style mechanism, in which a remote amino acid activates an active site water via a string of solvent molecules, remains the likely mechanism (21). It is likely, therefore, that CtCel124 also hydrolyzes glycosidic bonds through a Grotthus-style mechanism, although small nucleophilic ions such as thiols, acetate,
or phosphate ions may fulfill the role of the catalytic base by activating the nucleophilic water.

A central feature of the catalytic apparatus of CeCel124 is Glu96, which makes a hydrogen bond with O4 of the Glc at +1 (Figs. 2 and 3). The glutamate is underneath the β face of the +1 Glc and is thus in an ideal orientation to promote leaving group departure by donating a proton to the scissile glycosidic O. This is consistent with the observation that mutation of Glu96 (E96A) completely inactivates the enzyme against polysaccharides and oligosaccharides where the leaving group is poor. Against 2,4-nitrophenyl-cellotrioside (2,4-DNP-cellotrioside), in which the 2,4-dinitrophenolate leaving group does not require protonation (pKa approximately 3.5) (22), the E96A mutation does not decrease activity (Table 1). This is again consistent with the view that Glu96 is the catalytic acid of CeCel124. The descending limb of the pH curve reports a single ionizing group, the catalytic acid, with a pKa of 6.8. The required modulation of the pKa of the Glu96 carboxylic acid (pKa of carboxylic acid groups in solution are approximately 4.0) is likely contributed in part through a polar interaction with the OH of Tyr203, although its apolar environment (Glu96 is in close proximity to Leu119, Tyr187, Tyr203, and the aliphatic chain of Gln117) will ensure that the carboxylic acid group of this residue is mostly protonated at pH 5.5. The equivalent residue in Lyz23 and Slh70, Glu73, and Glu478, respectively, are also likely to function as the catalytic acid during glycosidic bond cleavage.

To summarize, the active site of CeCel124 comprises a basic structural platform that is capable of mediating glycosidic bond cleavage of gluco-configured substrates, likely through a Grothaus-style mechanism. Adornment of this catalytic platform with Asp97 in Lyz70 (Fig. 3) enables the enzyme to exhibit a classical single displacement (inverting) acid-base mechanism. However, it is unclear how the C2 N-acetyl group in Slt70 is activated, which is in agreement with the belief that the N-acetyl group is necessary for activity.

The Substrate-Binding Cleft of CeCel124. The superhelical fold provides a platform for the substrate-binding cleft that extends across the top of the protein (Fig. 2). The cleft houses the two molecules of cellotriose that bind to subsites −4 to −2 and +1 to +3, respectively. No ligand was evident in the −1 subsite. The extended loop, connecting α-H4 and α-H5, forms one wall of the binding cleft, and the other wall consists of several different structural elements that include the C-terminal end of α-H7, the N-terminal region of α-H8, and the N-terminal region of the loop connecting these helices (Fig. 2). Unlike the majority of glycanases, and carbohydrate-binding proteins in general, the substrate-binding cleft of CeCel124 does not contain a significant hydrophobic platform; the cellulase makes numerous direct polar contacts with the two cellotriose molecules (Fig. 2).

A striking feature of the substrate-binding cleft is the different topologies displayed by its positive and negative subsites, respectively (Fig. 2). Subsites −4 to −1 form a deep narrow cleft in which the bound trisaccharide is significantly twisted (Fig. S7). In contrast, subsites +1 to +3 display a more open topology (Fig. 2), and the conformation of the bound trisaccharide adopts an approximate twofold screw axis (Fig. S7). It is not possible to obtain the crystal structure of an enzyme in complex with insoluble polysaccharides such as cellulose. However, the conformation adopted by cellobiose polysaccharides such as cellulotriose, on enzyme, provides insight into the likely topological features of cellulose bound to CeCel124CD. Thus, the helical structure of cellulotriose bound to the negative subsites is distinct from the twofold screw axis displayed by glucon chains in crystalline cellulose. Indeed, the twisted structure of cellulotriose is adopted by cellobiose polysaccharides in solution (23). By contrast, the linear conformation adopted by cellulotriose bound to the distal positive subsites is similar to the structure of the glucon chains in crystalline cellulose (2). Thus, it is likely that the substrate-binding cleft of CeCel124 is tailored to recognize specific substructures of cellulose, which are at the interface between crystalline and paracrystalline (or amorphous) regions of cellulose. Indeed, competition experiments between cellulose-specific CBMs indicate that these proteins recognize specific topological features of the polysaccharide. Thus, CBMs belonging to families 4, 17, and 28 recognize distinct amorphous or paracrystalline regions of cellulose (24, 25), and CBM2a, CBM3a, and CBM1, which bind to crystalline cellulose, also display distinct specificities (26). Many cellulose-degrading bacteria express a large number of endo-β-1,4-glucanases (27, 28), exemplified by C. thermocellum and other organisms is currently unclear. Cellulose, although chemically invariant, displays very different topologies ranging from highly crystalline structures to isolated highly twisted glucan chains (amorphous cellulose). It is possible that at least some of the endoglucanases expressed by a single organism are tailored to recognize specific cellulose substructures found in nature. CeCel124, by targeting the boundary between crystalline and amorphous regions of cellulose, may generate reaction products that comprise substrates for exo-acting cellulases that act on the nonreducing end of crystalline cellulose and the reducing end of isolated cellulose chains, respectively.

Materials and Methods

Cloning, Expression, and Purification of CeCel124. DNA encoding the C-terminal module of CeCel124 (designated CeCel124CD; residues 131–350) was amplified from C. thermocellum strain ATCC 2745 genomic DNA by PCR, using primers listed in Table S2, and the resultant DNA was cloned into PET28a to generate pCel124. CeCel124CD encoded by pCel124 contains a His6-tag. To generate CeCel124CD fused to the noncatalytic carbohydrate (cellulose)-binding module CBM3a (C Cel124CD-CBM3a), overlapping PCR was used deploying primers that amplified the DNA sequences encoding the two modules. Expression of the fusion proteins was achieved by addition of 0.1 M D-thiogalactopyranoside (1 mM final concentration) to midexponential phase cultures of E. coli BL21(DE3) harboring pCel124 with incubation for a further 16 h at 37 °C. The His6-tagged recombinant protein was purified from cell-free extracts by immobilized metal ion affinity chromatography using standard methodology (29). To produce seleno-L-methionine CeCel124, the E. coli methionine auxotroph 8834(DE3) containing pCel124 was cultured as described by Charnock et al. (30). The recombinant protein was purified as described above except all buffers were supplemented with 5 mm β-mercaptoethanol. The production of recombinant Cel485 and the CBM3a-Coh construct (CBM3a fused to the type I cohesion from CipA, the primary scaffoldin of C. thermocellum) were described previously (9, 31, 32).

Enzyme Assays. In enzyme assays using soluble polysaccharides, the reactions were carried out in 50 mM MES buffer, pH 5.5, containing 1 mg/mL BSA and 0.5% of the target polysaccharide. Reactions, which were incubated at 50 °C, were terminated typically, 15 nM of enzyme. At regular time points, aliquots were removed and reducing sugar present was determined (33), using glucose to construct the standard curve. In enzyme assays using Avicel (Sigma Chem. Co.) and PASC (prepared as described previously (15)) as substrates, the reaction mixture consisted of 500 mM enzyme in 100 mM acetate buffer, pH 5.0, 24 mM CaCl2, 4 mM EDTA, and insoluble cellulose at a concentration of 2% (wt/vol). The reactions, which were carried out at 50 °C, were terminated by immersing the sample tubes in ice water. After centrifugation to remove insoluble substrate, the supernatant was assayed for reducing sugar. When assaying Cel485 in the presence of CBM3a-Coh, equimolar quantities of the protein partners were incubated for 2 h at 37 °C (without the substrate) preceding the assay. This ensured that Cel485 was appended to the CBM3a. The hydrolysis of cellobiose polysaccharides (30 μM), conducted in 50 mM MES buffer, was assessed by the rate of substrate depletion using a previously described Dionex®-high-performance anion-exchange chromatography (HPAEC) method to detect cellobiose polysaccharides (15). HPAEC was also used to identify the reaction products released from PASC. The α and β configuration of the anomeric carbon of cellotriose, generated by the hydrolysis of cellobiose by CeCel124CD, was determined as described previously (15, 34).

Binding of the CeCel124CD Mutant E96A to Ligands. Binding was determined by ITC and by depletion binding isotherms. ITC measurements were made.
at 40 °C following standard procedures (30) using a Microcal ITC 
300 calorimeter in 50 mM MES buffer, pH 5.5. During a titration experiment, the protein sample at 88 μM was injected with 20 × 2 μl aliquots of 2 mM cellohexaose. For experiments with RC, prepared as described previously (13), the ligand was in the cell at 29.4 mg/mL and the protein (835 μM) was the titrant. Integrated heat effects, after correction for heats of dilution, were converted to the corresponding binding isotherm using the Microcal Origin, version 2.9. Thermodynamic parameters were calculated using the standard thermodynamic equation: $-RT \ln K_{d} = \Delta G = \Delta H - T \Delta S$. Depletion binding isotherms were performed in 50 mM MES buffer, pH 7.0, at 40 °C. The polysaccharide was centrifuged at 13,000 × g for 1 min, and the $A_{280}$ of the supernatant was measured to quantify the amount of free protein remaining after binding. Bound protein was calculated from the total minus the free protein. The data were analyzed by nonlinear regression using a standard one-site binding model (GraphPad Prism, v5), and the $K_{d}$ and $K_{c}$ values were obtained from the regressed isotherm data.

**Crystallization of CfCel12A$_{w}$ and Data Collection.** Structure determination was by the single-wavelength anomalous dispersion (SAD) method, using the anomalous diffraction of the selenium atoms, incorporated in the protein. CfCel12A$_{w}$ was crystallized with an equal volume (1 μL of protein (60 mg/mL in solution with 10 mM cellohexaose) and reservoir solution 8% (vol/vol) tacsimate pH 5.0 and 20% (vol/wt) PEG 3350). Glycerol (30% vol/wt) was added as the cryoprotectant. The SAD experiment was conducted at the beamline ID14-E4H at the European Synchrotron Radiation Facility at Grenoble, France, using an ADSC Quantum-4 CCD detector. Data were collected at 0.954 Å wavelength. A total of 120 images with 1st oscillation for 5 s were collected. The data were processed with DENZO and the HKL2000 package and scaled with SCALAM (35).

**Phasing, Model Building, and Refinement.** Patterson maps were deconvoluted by calculation of anomalous Patterson maps. The positions of the seleniums were refined, and phases were calculated using SHARPplusOASPH (36). Subsequently, a cycle of model building and refinement was conducted using the program DM (37), where phases were modified and extended to 1.5 Å resolution yielding a figure of merit of 0.91 and an interpretable electron-density map. A model comprising 243 amino acids was built from the initial map with program COOT (38). Water molecules and alternative conformers were added using ARP/WARP (35), and refinement with REFMACS (39) was performed as deemed appropriate from the behavior of the cross-validation $(R_{	ext{cryst}})$ sub-set of reflections (10%). Solvent molecules were added in the final stages of refinement according to hydrogen bond criteria and only if their B factors refined to reasonable values and if they improved the $R_{	ext{cryst}}$. The statistics for structure refinement are shown in Table S3.

**Acknowledgments.** The authors acknowledge financial support from Fundação para a Ciência e a Tecnologia, Portugal, through grants PDTC/BIA-PRO/69732/2006, PDTC/QUI-BIQ/100359/2008, and the individual fellowship SFRH/BD/38667/2007 (to A.C.); the UK Biotechnology and Biological Sciences Research Council (studentship to A.C.); the US Department of Energy (Grant 00562492); and the Israel Science Foundation (Grant 966/09. The authors also acknowledge the European Synchrotron Radiation Facility, Grenoble, France (beamline ID14-E4H) for access and technical support during the data collection.