PROTEINS
STRUCTURE ■ FUNCTION ■ BIOINFORMATICS

Identification of a novel carbohydrate esterase from *Bjerkandera adusta*: Structural and function predictions through bioinformatics analysis and molecular modeling

Laura I. Cuervo-Soto,[1,2] Gilberto Valdés-García,[1] Ramón Batista-García,[1,2] María del Rayo Sánchez-Carbente,[2] Edgar Balcázar-López,[2] Verónica Lira-Ruan,[1] Nina Pastor,[1]* and Jorge Luis Folch-Mallol[2]*

[1] Department of Biochemistry and Molecular Biology, Facultad de Ciencias, Universidad Autónoma del Estado de Morelos. Av. Universidad 1001, Col. Chamilpa, Cuernavaca, Morelos México

[2] Department of Environmental Biotechnology, Centro de Investigación en Biotecnología, Universidad Autónoma del Estado de Morelos. Av. Universidad 1001, Col. Chamilpa, Cuernavaca, Morelos México

ABSTRACT

A new gene from *Bjerkandera adusta* strain UAMH 8258 encoding a carbohydrate esterase (designated as *Baces*I) was isolated and expressed in *Pichia pastoris*. The gene had an open reading frame of 1410 bp encoding a polypeptide of 470 amino acid residues, the first 18 serving as a secretion signal peptide. Homology and phylogenetic analyses showed that BaCesI belongs to carbohydrate esterases family 4. Three-dimensional modeling of the protein and normal mode analysis revealed a breathing mode of the active site that could be relevant for esterase activity. Furthermore, the overall negative electrostatic potential of this enzyme suggests that it degrades neutral substrates and will not act on negative substrates such as peptidoglycan or *p*-nitrophenol derivatives. The enzyme shows a specific activity of 1.118 U mg^{-1} protein on 2-naphthyl acetate. No activity was detected on *p*-nitrophenol derivatives as proposed from the electrostatic potential data. The deacetylation activity of the recombinant BaCesI was confirmed by measuring the release of acetic acid from several substrates, including oat xylan, shrimp shell chitin, *N*-acetylglucosamine, and natural substrates such as sugar cane bagasse and grass. This makes the protein very interesting for the biofuels production industry from lignocellulosic materials and for the production of chitosan from chitin.

Proteins 2015; 83:533–546.

Key words: *Bjerkandera adusta*; family 4 carbohydrate esterase; heterologous expression; lignocellulosic material degradation; deacetylase.

Additional Supporting Information may be found in the online version of this article.

Abbreviations: AOX1, alcohol oxidase 1; AXEs, acetylxylan esterases; *Baces*I, carbohydrate esterase of *Bjerkandera adusta*; BMGY and BMMY, buffered glycerol/methanol-complex medium; CAZy, carbohydrate-active enZYmes; CBM, carbohydrate binding module; CDA1, chitin deacetylase 1; cDNA, complementary DNA; CE, carbohydrate esterases; CE4, family 4 carbohydrate esterases; DIG, digoxigenin; GH, glycoside hydrolase; GHs, glycosyde hydrolases; LB, Luria–Bertani medium; MT, Motif; Mut^+, methanol utilization plus; Mut^s, methanol utilization slow; NCBI, National Center for Biotechnology Information; PCR, polymerase chain reaction; PDA, potato dextrose agar; PDB, Protein Data Bank; RACE, rapid amplification of cDNA ends; rBacesI, recombinant BacesI; rpm, revolutions per minute; YNB, yeast nitrogen base with ammonium sulfate without amino acids; YPD medium, yeast extract, peptone and dextrose

Grant sponsor: CONACyT from the Mexican Government; Grant number: CB-153789-Q.
*Correspondence to: Nina Pastor, Universidad Autónoma del Estado de Morelos. Av. Universidad 1001, Col. Chamilpa, Cuernavaca, 62209, Mor. México.
E-mail: nina@uaem.mx and Jorge Luis Folch-Mallol, Universidad Autónoma del Estado de Morelos. Av. Universidad 1001, Col. Chamilpa, Cuernavaca, 62209, Mor. México. E-mail: jordi@uaem.mx
Received 19 June 2014; Revised 19 December 2014; Accepted 31 December 2014
Published online 13 January 2015 in Wiley Online Library (wileyonlinelibrary.com). DOI: 10.1002/prot.24760

INTRODUCTION

Lignocellulose degradation is a central step for carbon recycling in land ecosystems[1] and involves the combined metabolic activity of numerous organisms. Plant cell walls are composed predominantly of polysaccharides such as cellulose, hemicellulose, pectin (representing up to 70% of its biomass) and lignin, a heterogeneous aromatic polymer [2–4] that forms a tight complex with these polysaccharides in varying proportions in plants.[5] Among plant cell wall polysaccharides, hemicelluloses are interesting because they are part of a barrier against the attack of pathogens and are constituted by different sugars that could have important biotechnological applications. Hemicelluloses have β-(1 → 4)-linked backbones made of xyloglucans, xylans, mannans, and glucomannans, and β-(1 → 3,1 → 4)-glucans in an equatorial configuration.[6] Their backbones have branches composed by monomers such as D-galactose, D-xylose, L-arabinose, and D-glucuronic acid, generating various types of hemicelluloses that are strongly dependent on the tissue and plant species.[4] A special feature in all the types of hemicelluloses is that these are partially esterified with acetic acid. Acetylation not only modifies the physicochemical properties of polysaccharides, for example, increasing their solubility, but it also protects the plant cell wall against microorganisms and prevents hydrolysis of glycosidic linkages by the corresponding hydrolases[7,8] representing an obstacle for enzymatic saccharification of plant hemicelluloses to produce fermentable sugars. To remove acetyl groups, wood-degrading microorganisms, in particular Basidiomycetes, have developed a set of enzymes that mainly include acetylxylan esterases (AXEs), whose function is to hydrolyze the ester linkages of the acetyl groups in positions 2 and/or 3 in natural acetylated xylan and to cooperate with glycoside hydrolases in plant polysaccharide degradation.[9] These enzymes are currently grouped in at least eight families, specifically in carbohydrate esterase (CE) families 1–7 and 16.[8] The enzymes present in CE families 1–7 differ in structure, substrate, and positional specificity, and belong to serine type esterases. An exception are the members of CE4 family, which includes: acetylxylan esterases (EC 3.1.1.72), chitin deacetylases (EC 3.5.1.41), chitooligosacharide deacetylases (EC 3.5.1.), peptidoglycan GlcNAc deacetylase (EC 3.5.1.), and peptidoglycan *N*-acetylmuramic acid deacetylase (EC 3.5.1.), all of which catalyze the *N*- or *O*- deacetylation of substrates such as acetylated xylan, chitin, and peptidoglycan.[10] They all share a universal conserved region in their primary structure named the "NodB homology domain" or the "polysaccharide deacetylase domain,"[11] and are considered aspartate deacetylases, activated with bivalent metal cations coordinated with two histidines,[12] and adopt a distorted $(\alpha/\beta)_7$ barrel fold.[8] Members of this family have been cloned from several microorganisms such as *Volvariella volvacea* (acetylxylan esterase, *Vvaxe*II) (GenBanK FJ536255),[9] chitin deacetylase from *Rhizopus circinans* (AY861444.1),[13] and several others. Three-dimensional structures have been obtained for enzymes from bacteria[10] and for two eukaryotic chitin deacetylases, one from the fungal pathogen *Colletotrichum lindemuthianum* (PDB 2IW0)[14] and the other from *Aspergillus nidulans* (PDB 2Y8U).

In this work, we report a new sequence of a carbohydrate esterase gene from the Basidiomycete fungus *Bjerkandera adusta* (*Baces*I). Blastp results showed high similarity with family 4 carbohydrate esterases and phylogenetic analyses coincided in grouping this gene in the same family. We built a homology model of the CE domain, and explored both its dynamics and ligand binding preferences. The coding sequence was successfully expressed in *Pichia pastoris*, and the esterase activity was confirmed using 2-naphtyl acetate. Its deacetylase activity was further confirmed by the release of acetic acid from commercial substrates such as oat xylan, shrimp shell chitin, and *N*-acetyl glucosamine and natural lignocellulosic materials like sugar cane bagasse and grass, and suggests that this protein has a biotechnological potential. It could be used together with cellulases and other enzymes in biological pretreatment for the saccharification of lignocellulosic materials for biorefinery applications. It could also be useful to treat arthropod or fungal chitin to obtain chitosan, which has received much interest recently for commercial and biomedical applications.

MATERIALS AND METHODS

Strains and growth conditions

For propagation and maintenance, *B. adusta* strain UAMH 8258 was grown on solid PDA medium (DIFCO) (2% potato, 2% dextrose, and 1.5% agar) for 8 days at 28°C and then stored at 4°C until further use. For induction of carbohydrate esterase expression, *B. adusta* was grown for 7 days at 28°C on solid minimal medium (7.8 mg L^{-1} $CuSO_4.5H_2O$, 18 mg L^{-1} $FeSO_4.7H_2O$, 500 mg L^{-1} $MgSO_4.7H_2O$, 10 mg L^{-1} $ZnSO_4$, 50 mg L^{-1} KCl, 1 g L^{-1} K_2HPO_4 2 g L^{-1}, NH_4NO, 2% corn cobs (ground to a homogeneous powder) as carbon source and 1.5% of agar, adjusted to pH 5.0 with phosphoric acid, modified from Ref. [15].

E. coli strain DH5α was grown in Luria-Bertani medium (LB) and used for propagation, stocking and manipulation of plasmid pJET (Thermo-scientific) derivatives. Low-salt LB (tryptone 10 g L^{-1}, sodium chloride 5 g L^{-1}, and yeast extract 5 g L^{-1}) was used for maintenance of *E. coli* DH5α with plasmid pPICzαA (Invitrogen), since high salt concentrations decrease the activity of zeocin, the antibiotic selection for this plasmid.

Pichia pastoris strain X-33 (Invitrogen), was used to express the *Baces*I cDNA. YPD medium (yeast extract 10 g L^{-1}, peptone 20 g L^{-1}, dextrose 20 g L^{-1}, and 15 g L^{-1} of agar) was used for growth and maintenance of the strain, and 100 μg mL^{-1} zeocin was added to YPD to select transformants. Minimal dextrose/methanol medium (MD/MM: 1.34% YNB, 4×10^{-5} biotin, 1.5% agar; 2% dextrose or 0.5% methanol, respectively) were used to identify the phenotypes Mut^{+} and Mut^{s}; BMGY and BMMY media (Buffered Glycerol/methanol-complex Medium: 1% yeast extract, 2% peptone, 100 mM potassium phosphate pH 6.0, 0.34% YNB, 4×10^{-5}% biotin; 1% glycerol or 0.5% methanol, respectively) were used to induce recombinant protein expression.

Cloning of *Baces*I gene

The full-length *Baces*I sequence (see Results section) was obtained using two protocols: GenomeWalker™ Universal Kit (Clontech) and RLM-RACE (rapid amplification of cDNA ends of Ambion® Life Technologies) following the suppliers' instructions. Also, a Northern blot (DIG High DNA Labeling and Detection Starter Kit II, Roche) was performed to predict the size of the transcript, according to the manufacturer's instructions. As a probe, a 500 bp cDNA fragment was amplified through Reverse Transcriptase coupled-PCR (RT-PCR) (RevertAid Reverse Transcriptase Thermo Scientific) from *B. adusta* RNA, with a pair of primers (Supporting Information Table SI) designed within the known partial sequence (see results for the original finding). To generate the 5′-end of the coding region of *Baces*I, total RNA was obtained by the Trizol® method[16] from *B. adusta* grown on solid minimal medium with 2% corn cobs as a carbon source as described above. Its quality was verified in 1.5% agarose gel with 37% formaldehyde, and 10 μg were used to follow the procedure recommended by the RACE kit. To search for the genomic sequence of *Baces*I, a genomic DNA library constructed from *B. adusta* grown on 2% wheat straw as a carbon source was used as template for PCR reactions following the instructions of the GenomeWalker™ Universal Kit (Clontech). Two pairs of reverse primers (Supporting Information Table SI) were designed for the PCR reactions in both protocols. The first pair of oligonucleotides was designed from a known partial sequence. The second pair (GWRev1 and GWRev2) was designed in the sequence generated from the first PCR fragments, which were cloned in pJET and sequenced. Finally, the coding sequence of the *Baces*I without its native signal peptide was amplified using a pair of FwCEpPICzα/RevCEpPICzα primers with cutting sites for *Eco*RI and *Xba*I respectively (Supporting Information Table SI), which allowed the subcloning into the expression vector (pPICzαA). The PCR conditions were as follows: 95°C 5 min (one cycle), 95°C 45 s, 60°C 30 s, 72°C 2 min (30 cycles), and 72°C 7 min (one cycle), on a thermocycler (Axigen-Maxigene) and the amplification was carried out with the *Taq* DNA polymerase enzyme (Thermo Scientific). The amplified product was purified with the Column DNA Gel Extraction Spin Kit (Thermo Scientific), ligated into the vector pJET (pJET-*Baces*I) and transformed into electrocompetent cells of *E. coli* DH5α, for subsequent sequencing.

Phylogenies

Phylogenetic analysis was performed online in the server Phylogeny.fr (http://www.phylogeny.fr/version2_cgi/index.cgi).[16–20] A first phylogeny was constructed considering three protein sequences from each of the sixteen families of carbohydrate esterases (CE) described in the Carbohydrate-Active enZYmes Database (CAZy) (http://www.cazy.org). Sequences from both prokaryotes and eukaryotes were selected and are indicated in Figure 2. Sequences from BLAST searches in the non-redundant protein collection at the National Centre for Biotechnology Information (NCBI) were not considered for this phylogeny to avoid a bias with a particular family of CE.

Another phylogeny was constructed with 105 eukaryotic sequences belonging to CE4 family and the CE sequence of *B. adusta*. All these sequences are deposited in CAZy database and include yeasts, filamentous fungi, plants and protozoa. Also 25 sequences from a BLAST search (≥60% identity) on the non-redundant protein collection at the NCBI were considered for this new reconstruction (in total 130 sequences).

Multiple sequence alignment

Because of the different size of the CE4 sequences, regions containing the CE signatures from six well-characterized and representative proteins of the CE4 family were selected for the multiple alignments (see results and Fig. 3), which were performed using online platforms Clustal Omega (https://www.ebi.ac.uk/Tools/msa/clustalo/) and ESPript 3.0 (http://espript.ibcp.fr/ESPript/ESPript/).

Structure prediction and analysis

The amino acid sequence of the mature protein starting at Q19 was submitted to the I-TASSER server.[21,22] After identifying the structural boundaries of the CE4 domain, the N-terminal (Q19 to G175), CE4 (G176 to

E402) and C-terminal (T403 to F470) domains were submitted independently to I-TASSER.

The model of the CE4 domain was refined by hand, fitting the side-chain rotamers of residues within and around the active site (D197, D198, R285, Y288, Y323, L356, H358 and D359) to those of equivalent residues in CE4 domains deposited in the PDB[23] and identified by I-TASSER as close structural neighbors (2Y8U, 2IW0 and 1W1A), and adding a Zn atom. Hydrogen atoms were added with CHARMM-GUI (http://www.charmm-gui.org) and refined with 100 steps of steepest descent energy minimization with the CHARMM36 force field.[24] The final structure was submitted to energy refinement with 200 steps of steepest descent energy minimization, keeping the gamma carbon of D198 fixed to ensure proper Zn coordination. The refined structure was used for pKa calculations in the PROPKA server,[25–29] using PROPKA 3.1 to account properly for the presence of Zn in the protein. Electrostatic potential calculations were done in the APBS server,[30] using CHARMM charges assigned at pH 7 and default parameters. Normal modes calculations were done in the ElNémo server[31] using default parameters, and retrieving for analysis the first three modes.

Docking assays were done with Autodock-VINA[32] on the refined model, using the default parameters, and including side chain flexibility for residues lining the active site (D197, V223, R226, S250, R285, Y288, W309, D319, N321, D322, Y323, L356, H358, D359, L360, and F361). Ligands were taken directly from the ZINC database[33] (2-naphthyl acetate, *N*-acetyl-D-glucosamine alpha, *N*-acetyl-D-glucosamine beta, ferulate and acetate), from the PDB (chitin, from 1HJW) or built (xylan) and refined using the MMFF94 force field in Avogadro (http://avogadro.cc/wiki/Main_Page) until energy convergence. As positive controls of the docking process, acetate was docked to our model and to 2IW0 (*Colletotrichum lindemuthianum* CE4 with acetate and Zn at the active site).[14] All structural models were rendered in VMD.[34]

Expression of BaCesI in *P. pastoris* X-33

pJET-*Baces*I was digested with *Eco*RI and *Xba*I, and the released *Baces*I cDNA fragment was inserted into pPICzαA (Invitrogen, USA), digested with the same enzymes. This construct was designated as pPICZαA/*Baces*I and was transformed into *E. coli* DH5α. The identity of the insert was verified by restriction digestion and DNA sequencing. This construction was linearized with *Sac*I, and the transformation into *P. pastoris* X-33 cells was performed by electroporation for 5 ms at 2000 volts employing an Eporator apparatus (Eppendorf). Positive transformants were selected for their ability to grow on YPD plates containing zeocin at a final concentration of 100 μg/mL, and *P. pastoris* X-33 transformed with the empty vector pPICZαA was used as a negative control. Integration of the *Baces*I gene into the genome of *P. pastoris* X-33 was confirmed by colony PCR analysis using 5′ and 3′ AOX1 primers, following instructions of EasySelect™ *Pichia* Expression Kit (life technologies Invitrogen).

Positive transformants were grown on MD plates to 28°C for 48 h, and then transferred to MM medium under the same temperature and time conditions to induce the expression of the *Baces*I gene, and to select Mut^+ (Methanol utilization plus) and Mut^s (Methanol utilization slow) phenotypes. A colony identified as Mut^+ and designed as CE1, *P. pastoris* X-33 transformed with the empty vector, and the wild type *P. pastoris* X-33 (controls) were inoculated into 15-mL BMGY medium in a 50-mL flask, and grown at 28°C on a rotary incubator (200–250 rpm) until the OD_{600} reached 2–6 (~16–18 h). Then, cells were harvested by centrifugation at 1500 g for 5 min and resuspended in 50-mL BMMY medium in a 250-mL flask at the same temperature and rpm. *Baces*I expression was induced by adding methanol to a final concentration of 0.5% at 24-h intervals during 4 days. Every day, 5 mL of culture were taken for RNA extraction to monitor the presence of the *Baces*I transcript from CE1 and the control strains by reverse transcription coupled PCR, using primers FwCEpPICzα/RevCEpPICzα. After 96 h the supernatant was then recovered by centrifugation at 1500*g* for 10 min and concentrated in a 30 kDa cut-off amicon (GE healthcare); the protein was stored at 4°C until further use.

rBaCesI activity assays

Initially, recombinant BaCesI (rBaCesI) esterase activity was determined using 2-naphthyl acetate (Sigma) as substrate. A stock solution was prepared as follows: 99 mL of Phosphate buffered saline (PBS 1X: NaCl 8.0 g L^{-1}, KCL 0.2 g L^{-1}, Na_2HPO_4 1.44 g L^{-1} and KH_2PO_4 0.24 g L^{-1}) pH 6.5, 1 mL of triton-X-100, 500 μL of 2-naphthyl acetate (250 μM dissolved in acetone) and 100 μL of Fast Garnet reagent (100 mg mL^{-1} of Fast Garnet powder in PBS-triton-X). The supernatants from 200 mL cultures from *P. pastoris* X-33 transformed with the empty vector, the wild type *P. pastoris* X-33, and the *Baces*I transformed strain were concentrated by ultrafiltration through an Amicon 30 kDa membrane. The final reaction volume was 1 mL of stock solution plus 0.1 mg mL^{-1} protein from concentrated cultures supernatants and incubated at room temperature for 30 min. Absorbance was read at 538 nm every 4 min in a Biomatte 3 spectrophotometer (Thermo Spectronic), and substrate blanks without protein extract were routinely included. In all experiments, total protein was measured by the Lowry method[35] with the Folin reagent (HYCEL), and bovine serum albumin (BSA) was used as a standard. Specific activity was expressed as units per milliliter

per milligram of protein. All measurements were performed in triplicate from independent samples.

Deacetylase activity measurements

rBaCesI activity toward acetylated carbohydrates was also tested. Shrimp shell chitin, oat xylan, and *N*-acetylglucosamine (all from Sigma), and natural substrates, such as ground sugar cane bagasse (*Saccharum officinarum*) collected in the State of Morelos (Mexico) and ground grass (*Pennisetum purpureum*) collected from the State of Veracruz, Mexico, were used to measure the release of acetic acid. Reaction mixtures consisted of 20 mg of each substrate suspended in 300 μL of sodium phosphate buffer pH 7.0 and 0.1 mg mL^{-1} of protein from the concentrated rBaCesI preparation. Mixtures were incubated at 37°C for 16 h with orbital shaking (225 rpm) and centrifuged for 5 min to remove insoluble material. The release of acetic acid was quantified in the supernatant using an enzyme-linked assay kit Enzy-Chrom™ Acetate Assay Kit (BioAssay Systems), and absorbance was read at 570 nm.

Inhibition assays of rBaCesI by ferulic acid

Assays were performed as described above for 2-naphtyl acetate at pH 7, but in the presence of 50, 100, 200, 300, 400, 500, and 1000 μM of ferulic acid.

Statistical analysis

All experiments were performed in triplicate. Similarly, readings from independent experiments were determined in triplicate. Simple classification ANOVA tests were applied to determine significant statistic differences between the different cases. Hartley–Cochran–Bartlett and Kolmogorov–Smirnov and Lilliefors tests were performed to analysis homogeneity of variance and normal distribution in the data.[36]

A *post hoc* analysis that defines the order of the differences found in the ANOVAs was developed. Fisher LSD and Tukey HSD tests were considered for the *post hoc* analyses.[37] All statistical calculations were performed in STATISTICA.

RESULTS

Nucleotide sequence and cloning of BacesI

Upon analysis of 768 sequenced clones from a cDNA library form *B. adusta* (Cuervo *et al.* manuscript in preparation), we identified a sequence with high similarity to fungal carbohydrate esterases. Because it is of interest for our group to study biomass degradation we decided to study this gene, which we termed *Baces*I. Using the protocols mentioned in Materials and Methods, a 1730 bp genomic sequence was obtained, in which a potential promoter of 50 nucleotides was predicted with the Neural Network Promoter Prediction software (http://www.fruitfly.org/cgi-bin/seq_tools/promoter.pl), with a score of 1.0. A TATA box within the promoter region was identified, located 99 bp away from the ATG start codon of the coding sequence. Comparison between the genomic and cDNA sequences revealed the presence of three introns with sizes of 54, 62, and 53 bp showing GT endings in the 5' end and AG at the 3', which is consistent with introns found in fungi[38] (Fig. 1). The cDNA consisted of 1410 bp coding for a protein of 470 amino acids. The deduced protein sequence contains an N-terminal signal secretion sequence of 18 amino acids with a cleavage site between A18 and Q19, which was predicted by the SignalP software (http://www.cbs.dtu.dk/services/SignalP/). The mature protein consisted of 452 amino acids with a calculated molecular mass of 49720 Daltons. The full nucleotide sequence was deposited at GenBank with accession number KJ920937.

The deduced amino acid sequence of BaCesI was subjected to Blastp and showed high identity (> 50%) with carbohydrate esterases of fungal origin (Supporting Information Table SII), whose main functions are annotated to the carbohydrate esterase family 4 according to the CAZy classification (www.cazy.org).

Phylogenetic analysis

Representative sequences of the different CE families were subjected to phylogenetic analysis in order to classify the *B. adusta* CE sequence (Fig. 2). This sequence is grouped into a sub-branch with two other fungal CE4 family sequences annotated as chitin deacetylases: gi 57341404 from *Volvariella* and gi 330689933 from *Rhizopus* (marked in red, Fig. 2). This clustering suggests that *Baces*I sequence belongs to the CE4 family, and agrees with the classification of those proteins reported in the better BLASTp hits (Supporting Information Table SII). However another sequence of CE4 family also from a Basidiomycete (*Flammulina velutipes*, gi 322718529) is directly grouped with a plant representative of family 6 (acetylxylan esterase, EC 3.1.1.72) (CAZy database) (Fig. 2) (see Discussion).

To explore the relationship of the *B. adusta* CE sequence with other proteins deposited in the GenBank, a second phylogeny was performed with 130 CE4 family sequences (Supporting Information Fig. S1). This phylogeny shows that the *B. adusta* sequence is closely related with two proteins from the Basidiomycete *Fomitopsis pinicola* (gi 527299940 and gi 527298270) and with one sequence of the Basidiomycete *Postia placenta* (gi 242214962) (Supporting Information Fig. S1). These proteins are annotated as hypothetical proteins; however, signature domains for chitin deacetylase activity (homologous regions with the catalytic domain of NodB and CDA1) have been identified in these sequences. The presence of these domains strongly suggests that these

Figure 1

Nucleotide and deduced amino acid sequence of *Baces*I from *Bjerkandera adusta* strain UAMH 8258. A TATA box at −99 bp from the ATG start codon is shown double underlined. The cytosine marked as +1 indicates the site of transcription initiation. The ATG start codon is underlined and in bold type. The signal peptide is shown in a grey box. Introns I, II, and III are shown as single underlined sequences. The bold type letters in the introns show the consensus borders for fungal introns. Also in bold type (not underlined) the stop codon is highlighted.

proteins are xylanase/chitin deacetylases, and thus may have the same activity as *B. adusta* CE.

Multiple sequence alignment

Regions containing the CE signatures of CE4 family sequences of *Bacillus subtilis* (PDB 1W17), *Streptococcus neumoniae* (PDB 2C1G), *Aspergillus nidulans* (PDB 2Y8U), *Colletotrichum lindemuthianum* (PDB 2IW0), *Rhizobium meliloti* (gi 89280604) and *Mucor rouxii* (gi 854339) were selected for the multiple alignment considering previous studies (Fig. 3).[11,14,39–45] The best alignment (34.14% of identity) is shown between *B. adusta* CE sequence with the *Mucor*'s CE, while the lowest percent identity (21%) was observed with the *Streptococcus*'s CE. The identity percentages between *B. adusta* CE sequence and the other fungal sequences were 26.67% (*A. nidulans*) and 26.27% (*C. lindemuthianum*). Sequences considered for the alignments show different sizes with respect to BaCesI, so this could be the reason for the low identity percentages between them. However, five distinct highly conserved motifs in all CE4 family that assemble the active site were also found in BaCesI esterase (Fig. 3, see Discussion).[44]

When an alignment was performed with complete sequences we observed that the N-terminus is conserved among fungal CE family 4 sequences (Supporting Information Fig. S2).

Modeling of BaCesI

In our first attempt at structural modeling of BaCesI, we submitted the whole mature protein to the I-TASSER server. The resulting model had a C-score of −1.77, and

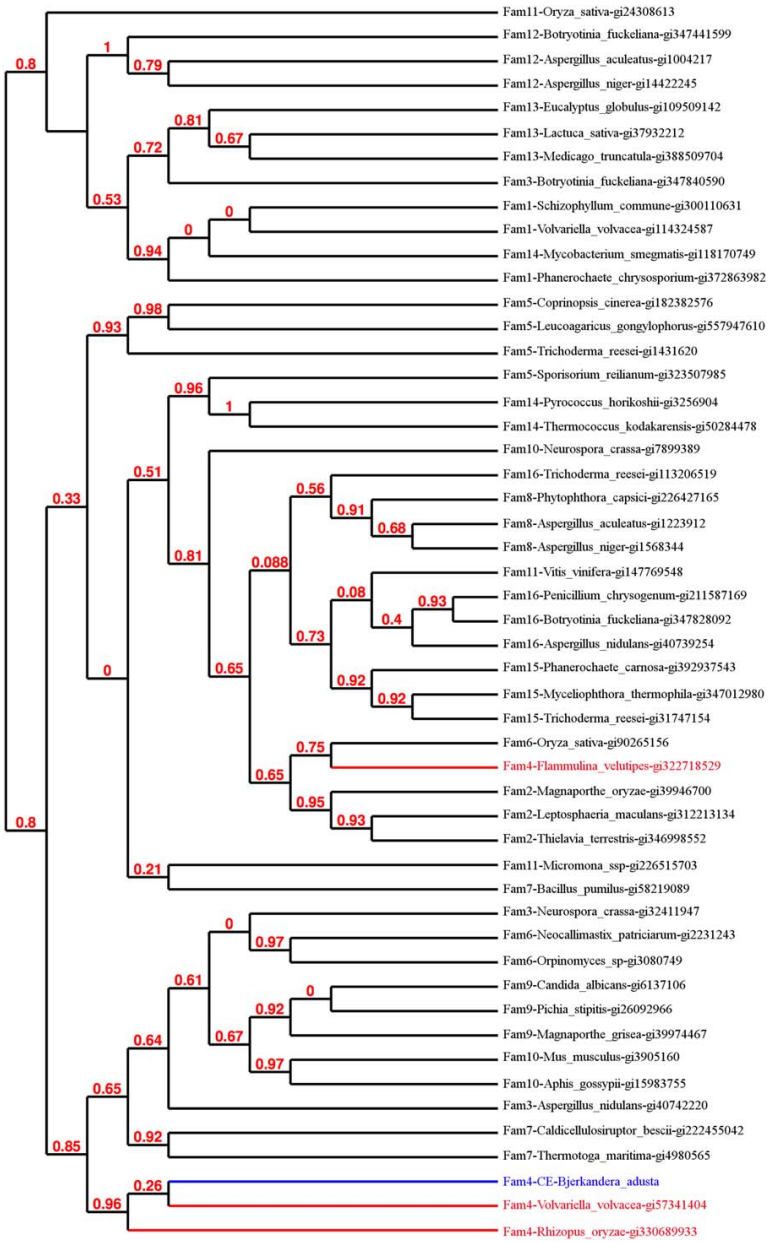

Figure 2

Phylogenetic relationships of the CE of *Bjerkandera adusta* with different CE family members. The sequence of *Bjerkandera adusta* is shown in blue. The sequences corresponding to the CE4 family are in red.

a TM-score of 0.5, mostly due to negligible sequence similarity in the first 175 and the last 63 amino acids of BaCesI with the two selected templates (2Y8U, a chitin deacetylase from *Aspergillus nidulans*, and 2C1G, a peptidoglycan deacetylase from *Streptococcus pneumoniae*[44]). Our attempts to model the *N*-terminal domain proved fruitless, as there are no sequence homologs deposited in the PBD for this region, despite its apparent conservation in fungi (see sequence alignment in Supporting Information Fig. S2). The C-terminus shows more sequence divergence, even amongst fungi, and also lacks homologs with a known structure. Therefore, we centered our

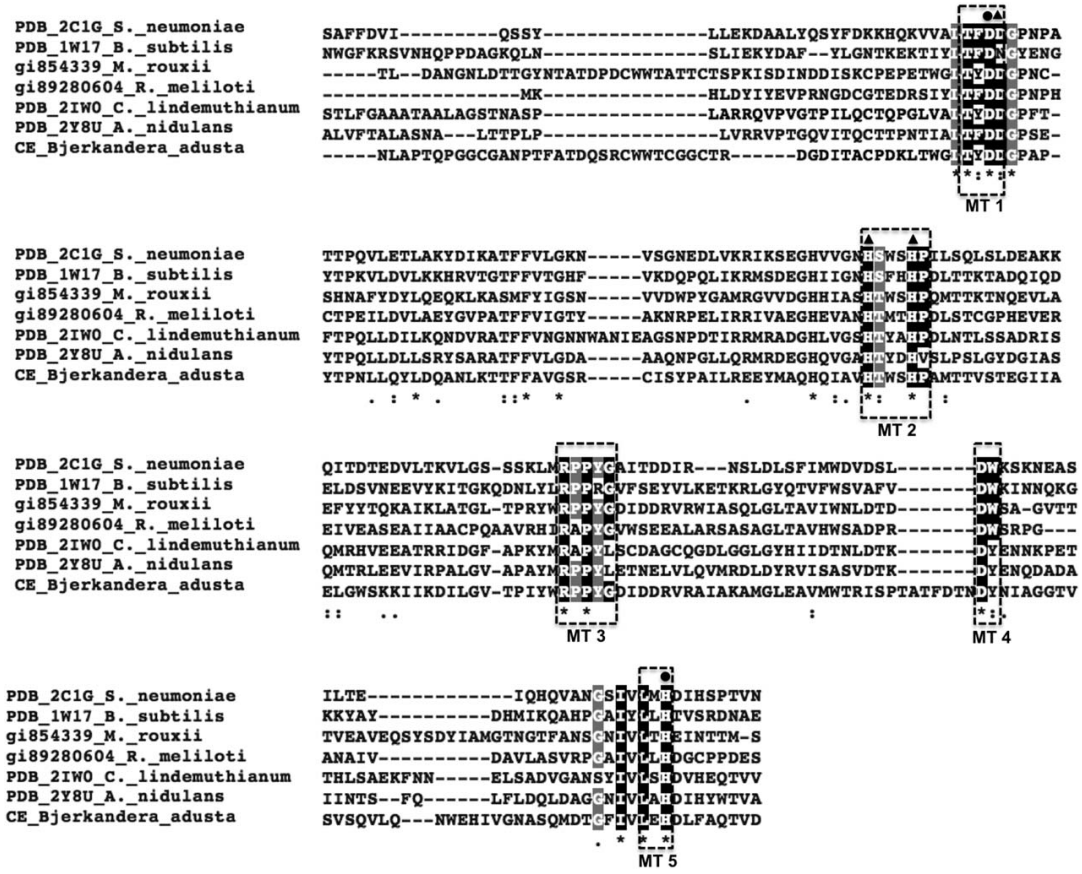

Figure 3

Multiple sequence alignment. CE of *B. adusta* was aligned with six well-characterized sequences from the database. Distinct highly conserved motifs (MT1 to 5) are shown in dotted-line boxes. Amino acid residues involved in catalytic activity are shown with a black circle and triangles.

modeling efforts on the CE4 domain, located between amino acids 176 and 402. This domain was modeled successfully by I-TASSER, yielding a structure with a C-score of 1.31 and a TM-score of 0.9. The closest structures in the PDB are 2Y8U and 2IW0, with sequence identities of 28% and 25% respectively to the CE4 domain of BaCesI. Supporting Information Figure S3A shows the superposition of the model and the two closest structural templates. As expected, the modeled CE4 domain of BaCesI has great structural similarity with its preferred templates (Supporting Information Fig. S3A), only differing in the presence of an extra loop comprising residues 312–319, immediately before the conserved block MT 4 (see below and Fig. 3).

The model of BaCesI was manually refined as described in Materials and Methods. All relevant catalytic residues are conserved, and are in equivalent chemical environments as those of other reported fungal CE4 domains in the PDB,[24] as measured by the local pKas calculated with PROPKA.[25] These are compiled in Supporting Information Table SIII. D198, H247, and H251 are the Zn ligands. D197 was proposed as the water-breaking residue, H358 as the proton donor that ends the hydrolysis reaction, and D322 stabilizes the conformation of H358. Furthermore, other important conserved residues such as R285 (R135 in 2Y8U and R142 in 2IW0) that stabilizes the conformation of D197 (D47 in 2Y8U and D49 in 2IW0), and Y288/Y323 (Y138/Y166 in 2Y8U and Y145/Y173 in 2IW0) that line the active site crevice and could participate in sugar binding and catalysis, are located at the appropriate sites in the energy-refined structure. All these are shown in Supporting Information Figure S3B.

The active site crevice is long and deep, almost cutting the domain in half. Low frequency vibrational modes are associated with functional motions in proteins,[46] and the elastic network model implemented in ElNémo[31] has proven to be a good approximation to these modes.

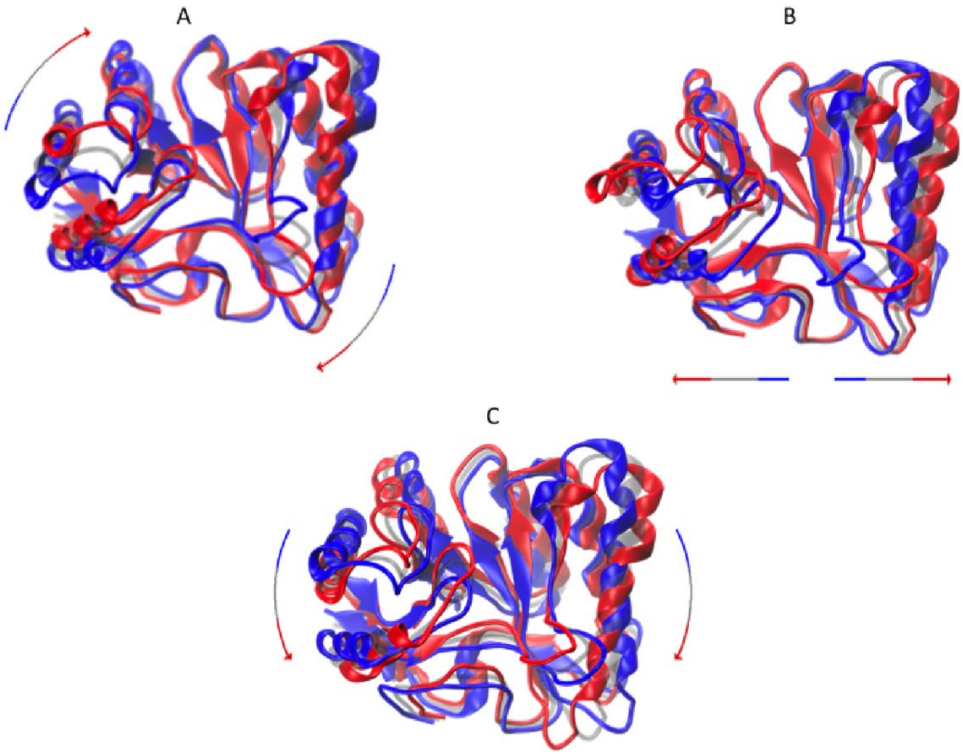

Figure 4

Low frequency motions of BaCesI. The color ribbon follows main chain atoms, and the transparent structure in grey corresponds to the energy-refined model. Full ribbons correspond to the maximal displacements calculated by ElNémo (25) for each mode, one in red and the other in blue. Arrows indicate the direction of movement in each panel, going from the blue to the red structure. (**A**) Twisting motion displacing the left half with respect to the right half in an up and down, antiparallel fashion. (**B**) Clam-like motion opening and closing the active site cleft toward the viewer. (**C**) Mandarin orange-wedge peeling motion, opening and closing the active site cleft toward the top of the figure. This figure was prepared with VMD (35).

Figure 4 shows the first three vibrational modes of BaCesI. The first one consists of a twisting motion of roughly the left half of the molecule with respect to the right half, displacing the alpha helix on the left upwards and downwards [Fig. 4(A)]. The second one is a clam-like motion of one half with respect to the other half, opening and closing the active site cleft [Fig. 4(B)]. The last mode pries open the active site from the top of the figure, as if peeling off the wedges of a mandarin orange [Fig. 4(C)]. We speculate that the second mode is particularly relevant to catalysis, as it results in the largest change in accessibility to the active site, and changes the distance between catalytic residues.

We docked each of the tested substrates on to the energy-refined BaCesI model. From the 20 structures generated by Autodock/VINA for each ligand, we selected those that placed the carboxylate carbon subject to nucleophilic attack close to the Zn atom, H358 and D197. Docking was successful for all ligands, including ferulate. Figure 5 shows successful binding poses for 2-naphthyl acetate and xylan; the pose for ferulate is shown in Supporting Information Fig S4, while the poses for the other ligands are available upon request. Both Y288 and Y323 participate in substrate binding; for the 2-naphtyl-acetate substrate these two residues form stacking interactions with the aromatic moiety.

It has been noted in the literature[8] that *p*-nitrophenol-acetate is a poor substrate for some CE4 carbohydrate esterases. We calculated the electrostatic potential and mapped it at the molecular surface for our BaCesI model and for available crystal structures of fungal and bacterial canonical CE4 enzymes (2IW0 and 2Y8U, both fungal chitin deacetylases; 2C71 and 2CC0, both xylan esterases; and 1W1B, a bacterial peptidoglycan esterase). These are shown in Figure 6, with the active site crevice facing front. We note a strong negative character in chitin deacetylases, a more neutral surface for xylan acetylases, and a positive surface for the peptidoglycan deacetylase.

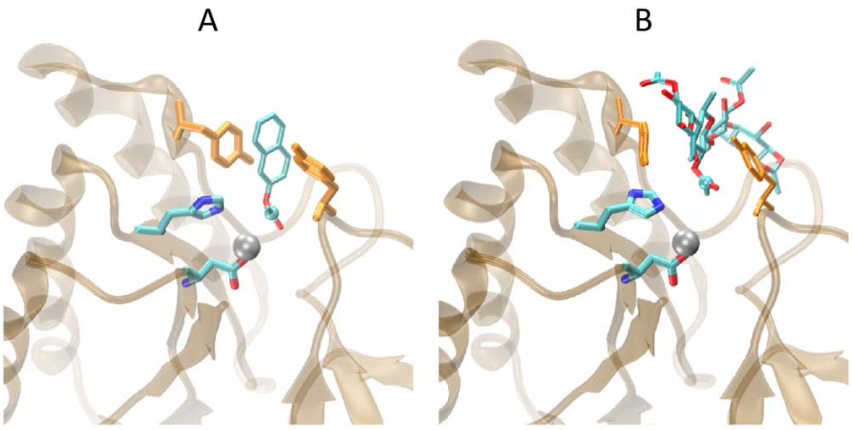

Figure 5

Encounter complexes of BaCesI with selected substrates. BaCesI is shown in transparent ribbon representation, with D197 and H358 in sticks with CPK colors, Y228 and Y323 in orange sticks, and the Zn atom as a silver sphere. The ligands are shown in stick representation, in CPK colors, highlighting the carbon susceptible to nucleophilic attack with a small sphere. (**A**) The 2-naphthyl acetate. (**B**) Xylan (tetraxylose). The orientation of BaCesI is akin to that in Figure 4. This figure was prepared with VMD (35).

Expression of BaCesI in *Pichia pastoris*

The CE1 transformant was selected to test the expression of rBaCes1. The recombinant protein was induced in cultures with methanol at 28°C and the supernatant was collected to determine total protein and esterase activity. The presence of the *Baces*I transcript was assessed by RT-PCR, as described in Materials and Methods. A band of approximately 1400 bp was observed in the transformant CE1, and it was absent in the control strain *(P. pastoris* X-33) (data not shown).

Multiple attempts to purify the enzyme through affinity chromatography using a Ni column were performed. However the protein did not bind to the column and eluted in the flow through. Experiments with the native and denatured protein and Western blot analysis indicated that the His tag was folded inside the protein (data available upon request). We tried a purification step through a carboxy-methyl cellulose column but in this case the protein was inactivated.

Esterase specific activity from a concentrated preparation through ultrafiltration of the culture supernatant was 1.118 U mg^{-1} of protein using 2-naphthyl acetate as a substrate, whereas no activity was detected when compared with concentrated supernatants of the wild type strain X-33 or X-33 transformed with the empty vector under the same conditions. Activity on several *p*-nitrophenyl-esters was undetectable (data not shown). Induction of the expression of rBaCes1 was also carried out in cultures incubated at 26 and 30°C, showing a decrease of activity of >60%, compared with those obtained at 28°C (data not shown).

Deacetylase activity of BaCesI

We performed activity assays on acetylated substrates such as oat xylan, colloidal chitin and N-acetylglucosamine, and natural substrates, (sugar cane bagasse and grass), determining the released concentration of acetic acid as described in Materials and Methods. The results showed an increased release of acetic acid from the substrates treated with the enriched preparation containing rBaCesI compared to control. Acetic acid obtained from chitin was 26 times larger, 87 times larger for oat spelt xylan, whereas from *N*-acetyl glucosamine it was around 89 times larger than the control value. On the natural substrates (bagasse and grass), acetic acid release was around five times greater, compared to the control (Table I). These results confirm the deacetylase activity of BaCesI.

rBaCesI is inhibited by ferulic acid

As predicted by the docking modeling, rBaCesI should be inhibited by ferulic acid, so we tested its activity in the presence of different concentrations of this compound. As shown in Table II, rBaCesI is inhibited by ferulic acid, being practically totally inhibited at 1000 μ*M* of ferulic acid.

DISCUSSION

The extracellular enzymes of the white rot Basidiomycetes represent a potential tool for the pretreatment of lignocellulosic material, this being an important step for

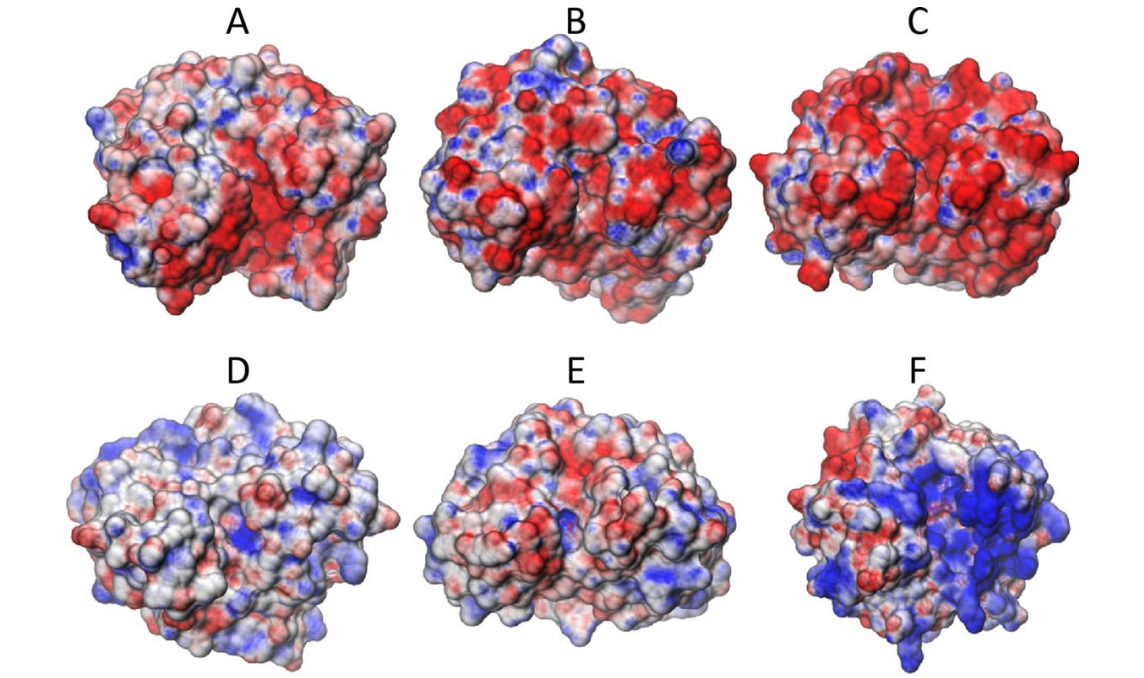

Figure 6

Electrostatic potential at the molecular surface of CE4 enzymes. (**A**) BaCesI, (**B**) 2IW0, (**C**) 2Y8U, (**D**) 2C71, (**E**) 2CC0, and (**F**) 1W1B. Red indicates a negative potential, white a neutral potential and blue a positive potential, ranging from −7e to +7e. This figure was prepared with VMD (35).

obtaining fermentable sugars for biorefinery purposes.[47] Few studies have been published about the cellulolityc enzymes of *B. adusta*.[48,49] In this work we cloned and heterologously expressed in *P. pastoris* a gene encoding for a carbohydrate esterase (*Baces*I) of *B. adusta*, and evaluated its activity and effect on acetic acid liberation from lignocellulosic materials. Bioinformatic analyses of the sequence were performed and gave relevant functional information of the protein.

The carbohydrate esterases (CE) are classified in the Carbohydrate-Active Enzymes (CAZymes) database. It has been proposed that the main function of some of the members, specially those in the CE1 and CE4 families, is to remove acetyl or feruloyl groups present in hemicellulose, enhancing other glycoside hydrolases (GHs) action on cellulose and hemicelluloses.[8,50] Phylogenetic analysis of the BacesI sequence groups it with two family 4 members, but the phylogenetic tree shows certain inconsistencies grouping some CE in different clades, distinct to those proposed by the classification in CAZy (for example, the sequence of a CE4 family member from a Basidiomycete (*Flammulina velutipes*) is directly grouped with a plant representative of family 6 (acetylxylan esterase, EC 3.1.1.72)). Hydrophobic cluster analysis could give more information about this inconsistency. However, it is consistent in the sense that CE4 and CE6 families share the same activity: acetylxylan esterase (CAZy database). On the basis of our Blastp analysis using the BaCesI amino acid sequence, the greatest identity (>50%) was found with members of the CE4 family,

Table I

Release of Acetic Acid by rBaCesI From Natural and Synthetic Substrates

Culture concentrated supernatant	Substrates/acetic acid liberated mmoles[a]				
	Chitin from shrimp shells	Oat spelt xylan	*N*-acetyl-glucosamine	Bagasse	Grass
P. pastoris+pPICzαA (Culture Concentrated supernatant)	1.21 ± 0.0063	0.51 ± 0.0035	0.56 ± 0.0049	4.63 ± 0.0120	5.0 ± 0.0042
rBaCesI (Culture Concentrated supernatant enriched through ultrafiltration)	31.84 ± 0.1034	44.44 ± 0.0111	50.72 ± 0.4332	22.86 ± 0.1107	26.58 ± 0.0338

[a]The measurements were performed using 20 mg of each substrate and 0.1 mg mL^{-1} of protein. The reactions were performed in triplicate for each substrate.

Table II
Effect of Different Concentrations of Ferulic Acid on the Activity of rBaCesI

Ferulic acid (μM)	(U mg^{-1} protein)	%
0	0.815	100
50	0.732	89.81
100	0.715	87.11
200	0.466	57.17
300	0.054	6.62
400	0.054	6.62
500	0.014	1.71
1000	0.001	0.12

especially with fungal sequences. The NodB domain was found in BaCesI and the conserved catalytic site residues (D197, D322, H358) and the metal-binding triad (H247, H251, D198) were identified (Fig. 3). Previous studies[44] showed that there are five highly conserved distinct motifs in all CE4 family members that line and assemble the active site. These motifs play an important role in the metal binding and biocatalysis of the enzyme. The multiple alignment shows these motifs (MT) (identified as MT 1, MT 2, MT 3, MT 4, and MT 5, see Fig. 3) which are also present in the BaCesI sequence. BaCesI in MT 1 shows a change (Y instead of F), which is also conserved in the CE fungal sequences of *M. rouxii* and in *C. lindemuthianum*. In this MT the critical residues are D197 (indicated with black circle, see Fig. 3) and D198 (indicated with black triangle, see Fig. 3) because they contribute to metal and substrate binding, respectively.[44] MT 2 (H(S/T)xxHP) also appears in BaCesI and this is a MT directly involved in metal-binding. H247 and H251 (indicated with black triangles) are the amino acids that bind Zinc, while T248 accepts a hydrogen bond from H251. Blair *et al.*[44] report that this Zinc-binding motif does not exist outside the CE4 family. This evidence strongly suggests that our sequence is a member of CE4 family. MT 3 and MT 5 are highly conserved in BaCesI as shown in Figure 3 and are very important in the active site. H358 (identified with black circle, see Fig. 3) is a critical residue in the catalytic activity because it is proposed to protonate the leaving acetate group.[45] MT 4 forms another site of the active groove in the enzyme, and has a change (Y instead of W) in BaCesI as compared to the bacterial CEs. However, all the considered fungal sequences show the same substitution (Y instead of W), so this conservative change may not be essential for the enzyme activity. The identification of conserved and distinctive motifs (some of them present only in the CE4 family) allows us to classify our sequence as a CE4 family member, and this evidence support the results found in the BLASTp and in the phylogeny analysis. The homology model we obtained for the CE4 domain of BaCesI confirms the proper placement of all catalytic and substrate-binding residues in the three-dimensional structure (Supporting Information Fig. S3B).

A cellulose-binding domain (CBM) was not identified in the sequence of BaCesI, although it is present in some acetylxylan esterases in family CE4. However, CBM are not always present in all carbohydrate active enzymes[51] and are not required for hydrolysis of the substrates.[41] Nevertheless, we note the presence of a long N-terminal domain (roughly 160 amino acids long) that seems to be specific to fungal CE4s (Supporting Information Fig. S2) and with no homology to any annotated domain. The exploration of the function of this domain in BaCesI is in progress.

The three normal modes depicted in Figure 4 suggest that opening and closing of the active site cleft is important for catalysis, in particular, the clam-like mode [Fig. 4(B)], as it displaces the active site residues relative to each other. We posit that BaCesI will not function as a ferulate esterase, as our attempts at docking ferulate suggest that it fits in the active site, but prevents proper closing of the active site cleft (Supporting Information Fig. S4). Indeed, when we tested activity in the presence of ferulate, rBaCesI was inhibited. The suggested mechanical model to explain these findings is that acetate groups resemble nails on a board (the glycan), and each nail can be engulfed by the active site in order to be removed. Ferulate resembles a plank instead of a nail, and its binding to the active site precludes proper closing and positioning of catalytic residues.

The electrostatic potential maps at the molecular surface of BaCesI and other CE4 enzymes with structures deposited in the PDB provide a rationale for the substrate preferences of BaCesI. Chitin deacetylases [Fig. 6(B,C)] are strongly negative, which would not be a problem for neutral substrates such as chitin. AXEs tend to be either negative or neutral [see Fig. 6(D,E)], explaining the observation that these enzymes are poorly active on *p*-nitrophenol acetate and other negatively charged small substrates,[8] which would suffer electrostatic repulsion from the enzyme. This was confirmed as no detectable activity was found when *p*-nitrophenol acetate, *p*-nitrophenol butyrate, *p*-nitrophenol ferulate, *p*-nitrophenol laurate, or *p*-nitrophenol palmitate were used as substrates (data not shown). On the other hand, the only reported enzyme that is active on peptidoglycans displays a positive molecular surface [Fig. 6(F)]. This analysis suggests that BaCesI is a chitin and xylan esterase, as seen in the measured activities reported here, and that it will most probably be inactive on peptidoglycans. This is currently being explored.

According to our results, we can assume that the activity of rBaCesI toward 2-naphtyl-acetate is given by the specific recognition by aromatic residues of the naphtyl moiety [Fig. 5(A)], that takes on the role of the neighboring sugars stabilizing the enzyme-substrate complex, as has been demonstrated for CE4 chitin deacetylase and peptidoglycan *N*-deacetylase.[42,44]

Deacetylase activity of rBaCesI was observed toward commercial substrates such as oat xylan, colloidal chitin

and *N*-acetyl glucosamine, and natural lignocellulosic materials such as sugar cane bagasse and grass. This activity on oat xylan and *N*-acetylglucosamine was evident, since the acetic acid liberation was around 90 times more compared to the control; chitin showed a 30 times increase, while for the natural substrates five times more acetic acid was observed relative to the control (Table I). These results demonstrate that recombinant BaCesI protein exhibits a deacetylase activity typical of CE4 family members, supporting the results of the *in silico* analysis and modeling from the sequence as mentioned above, acting as both an *O* and an *N*-deacetylase.

CONCLUSION

Bioinformatic analysis, molecular modeling and biochemical studies of a carbohydrate esterase from *B. adusta* suggest that the substrate specificity of this type of enzymes depends on the overall electrostatic potential of the enzyme's surface, particularly around the active site. Our analysis shows that for the novel CE4 esterase reported in this work, its substrate preference matches what is expected of a negatively charged active site neighborhood, and that this property can also explain the known substrate preferences of other members of the same family (see Fig. 6). This analysis is not reported elsewhere, to the best of our knowledge. Also, the low frequency movements found for this enzyme seem to be relevant for catalytic activity. The CE esterase described here belongs to Family 4 CEs and has a deacetylase activity on xylan, chitin, and lignocellulosic substrates like sugar cane bagasse and grass.

ACKNOWLEDGMENTS

The authors are grateful to Ayixón Sánchez Reyes for help with part of the statistical analysis and to Dr. Marcela Ayala for her help in the purification attempts of the rBaCesI. They are also grateful to Dra. Laura Álvarez-Berber and Dra. Verónica Narváez-Padilla for fruitful discussions about this work.

REFERENCES

1. Martinez AT, Speranza M, Ruiz FJ, Ferreira P, Camarero S, Guillen F, Martinez MJ, Gutierrez A, del Rio JC. Biodegradation of lignocellulosics: microbial, chemical, and enzymatic aspects of the fungal attack of lignin. Int Microbiol Off J Spanish Soc Microbiol 2005;8: 195–204.
2. Hashimoto K, Kaneko S, Yoshida M. Extracellular carbohydrate esterase from the basidiomycete coprinopsis cinerea released ferulic and acetic acids from xylan. Biosci Biotechnol Biochem 2010;74: 1722–1724.
3. Gilbert HJ. The biochemistry and structural biology of plant cell wall deconstruction. Plant Physiol 2010;153:444–455.
4. van den Brink J, de Vries RP. Fungal enzyme sets for plant polysaccharide degradation. Appl Microbiol Biotechnol 2011;91:1477–1492.
5. Couturier M, Haon M, Coutinho PM, Henrissat B, Lesage-Meessen L, Berrin JG. Podospora anserina hemicellulases potentiate the trichoderma reesei secretome for saccharification of lignocellulosic biomass. Appl Environ Microbiol 2011;77:237–246.
6. Scheller HV, Ulvskov P. Hemicelluloses. Annu Rev Plant Biol 2010; 61:263–289.
7. Taylor EJ, Gloster TM, Turkenburg JP, Vincent F, Brzozowski AM, Dupont C, Shareck F, Centeno MS, Prates JA, Puchart V, Ferreira LM, Fontes CM, Biely P, Davies GJ. Structure and activity of two metal ion-dependent acetylxylan esterases involved in plant cell wall degradation reveals a close similarity to peptidoglycan deacetylases. J Biol Chem 2006;281:10968–10975.
8. Biely P. Microbial carbohydrate esterases deacetylating plant polysaccharides. Biotechnol Adv 2012;30:1575–1588.
9. Ding XLS. Molecular characterization of a new acetyl xylan esterase (AXEII) from edible straw mushroom volvariella volvacea with both de-O-acetylation and de-*N*-acetylation activity. FEMS Microbiol Lett 2009;295:50–56.
10. Oberbarnscheidt L, Taylor EJ, Davies GJ, Gloster TM. Structure of a carbohydrate esterase from bacillus anthracis. Proteins Struct Funct Bioinform 2007;66:250–252.
11. Kafetzopoulos D, Thireos G, Vournakis JN, Bouriotis V. The primary structure of a fungal chitin deacetylase reveals the function for two bacterial gene products. Proc Natl Acad Sci USA 1993;90:8005–8008.
12. Li XL, Skory CD, Cotta MA, Puchart V, Biely P. Novel family of carbohydrate esterases, based on identification of the hypocrea jecorina acetyl esterase gene. Appl Environ Microbiol 2008;74:7482–7489.
13. Gauthier C, Clerisse F, Dommes J, Jaspar-Versali MF. Characterization and cloning of chitin deacetylases from rhizopus circinans. Protein Exp Purif 2008;59:127–137.
14. Blair DE, Hekmat O, Schüttelkopf AW, Shrestha B, Tokuyasu K, Withers SG, van Aalten DMF. Structure and mechanism of chitin deacetylase from the fungal pathogen colletotrichum lindemuthianum. Biochemistry 2006;45:9416–9426.
15. Inglis GD, Popp AP, Selinger LB, Kawchuk LM, Gaudet DA, McAllister TA. Production of cellulases and xylanases by low-temperature basidiomycetes. Can J Microbiol 2000;46:860–865.
16. Dereeper A, Audic S, Claverie JM, Blanc G. BLAST-EXPLORER helps you building datasets for phylogenetic analysis. BMC Evol Biol 2010;10:1–6.
17. Dereeper A, Guignon V, Blanc G, Audic S, Buffet S, Chevenet F, Dufayard JF, Guindon S, Lefort V, Lescot M, Claverie JM, Gascuel O. Phylogeny.fr: robust phylogenetic analysis for the non-specialist. Nucleic Acids Res 2008;36:465–469.
18. Anisimova M, Gascuel O. Approximate likelihood-ratio test for branches: a fast, accurate, and powerful alternative. Syst Biol 2006; 55:539–552.
19. Edgar RC. MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res 2004;32:1792–1797.
20. Guindon S, Gascuel O. A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. Syst Biol 2003; 52:696–704.
21. Zhang Y. ITASSER server for protein 3D structure prediction. BMC Bioinform 2008;9:1–8.
22. Roy A, Kucukural A, Zhang Y. ITASSER: a unified platform for automated protein structure and function prediction. Nat Protocols 2010;5:725–738.
23. Berman HM, Westbrook J, Feng Z, Gillil G, Bhat TN, Weissig H, Shindyalov IN, Bourne PE. The protein data bank. Nucleic Acids Res 2000;28:235–242.
24. Suhre K, Sanejou YH. ElNemo: a normal mode web server for protein movement analysis and the generation of templates for molecular replacement. Nucleic Acids Res 2004;32:610–614.
25. Li H, Robertson AD, Jensen JH. Very fast empirical prediction and interpretation of protein pKa values. Proteins Struct Funct Bioinform 2005;61:704–721.

26. Bas DC, Rogers DM, Jensen JH. Very fast prediction and rationalization of pKa values for protein–ligand complexes. Proteins 2008; 73:765–783.
27. Olsson MHM, Søndergard CR, Rostkowski M, Jensen JH. PROPKA3: consistent treatment of internal and surface residues in empirical pKa predictions. J Chem Theory Comput 2011;7:525–537.
28. Søndergaard CR, Olsson MHM, Rostkowski M, Jensen JH. Improved treatment of ligands and coupling effects in empirical calculation and rationalization of pKa values. J Chem Theory Comput 2011;7:2284–2295.
29. Rostkowski M, Olsson MHM, Søndergaard CR, Jensen JH. Graphical analysis of pH-dependent properties of proteins predicted using PROPKA. BMC Struct Biol 2011;11:1–6.
30. Baker NA, Sept D, Joseph S, Holst MJ, McCammon JA. Electrostatics of nanosystems: application to microtubules and the ribosome. Proc Natl Acad Sci USA 2001;98:10037–10041.
31. Best RB, Zhu X, Shim J, Lopes PEM, Mittal J, Feig M, Mackerell JAD. Optimization of the additive CHARMM all-atom protein force field targeting improved sampling of the backbone φ, ψ and side-chain χ(1) and χ(2) dihedral angles. J Chem Theory Comput 2012; 8:3257–3273.
32. Trott O, Olson AJ. AutoDock vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization and multithreading. J Comput Chem 2010;31:455–461.
33. Irwin JJ, Sterling T, Mysinger MM, Bolstad ES, Coleman RG. ZINC: a free tool to discover chemistry for biology. J Chem Inf Model 2012;52:1757–1768.
34. Humphrey W, Dalke A, Schulten KJ. VMD—visual molecular dynamics. J Mol Graph 1996;14:33–38.
35. Lowry OH, Rosebrough NJ, Farr AL, Randall RJ. Protein measurement with the folin phenol reagent. J Biol Chem 1951;193:265–275.
36. Hoaglin DC, Welsch RE. The hat matrix in regression and ANOVA. Am Stat 1978;32:17–22.
37. Fisher RA, Yates F. Statistical tables for biological, agricultural and medical research. 2nd ed. London: Oliver and Boyd Ltd. 1949;8:112.
38. Kupfer DM, Drabenstot SD, Buchanan KL, Lai H, Zhu H, Dyer DW, Roe BA, Murphy JW. Introns and splicing elements of five diverse fungi. Eukaryotic Cell 2004;3:1088–1100.
39. Vollmer W, Tomasz A. The pgdA gene encodes for a peptidoglycan *N*-acetylglucosamine deacetylase in streptococcus pneumoniae. J Biol Chem 2000;275:20496–20501.
40. Martinou A, Koutsioulis D, Bouriotis V. Expression, purification, and characterization of a cobalt-activated chitin deacetylase (Cda2p) from saccharomyces cerevisiae. Protein Expr Purif 2002;24:111–116.
41. Caufrier F, Martinou A, Dupont C, Bouriotis V. Carbohydrate esterase family 4 enzymes: substrate specificity. Carbohydr Res 2003;338: 687–692.
42. Hekmat O, Tokuyasu K, Withers SG. Subsite structure of the endo-type chitin deacetylase from a deuteromycete, colletotrichum lindemuthianum: an investigation using steady-state kinetic analysis and MS. Biochem J 2003;374:369–380.
43. Blair DE, van Aalten DM. Structures of bacillus subtilis PdaA, a family 4 carbohydrate esterase, and a complex with *N*-acetyl-glucosamine. FEBS Lett 2004;570:13–19.
44. Blair DE, Schuttelkopf AW, MacRae JI, van Aalten DM. Structure and metal-dependent mechanism of peptidoglycan deacetylase, a streptococcal virulence factor. Proc Natl Acad Sci USA 2005;102: 15429–15434.
45. Fukushima T, Kitajima T, Sekiguchi J. A polysaccharide deacetylase homologue, PdaA, in bacillus subtilis acts as an *N*-acetylmuramic acid deacetylase in vitro. J Bacteriol 2005;187:1287–1292.
46. Sanejou YH. Elastic network models: theoretical and empirical foundations. Methods Mol Biol 2013;924:601–616.
47. Saha BC. Hemicellulose bioconversion. J Ind Microbiol Biotechnol 2003;30:279–291.
48. Quiroz RE, Martinez C, Cuervo LI, Segovia L, Folch JL. Loosenin, a novel protein with cellulose-disrupting activity from *Bjerkandera adusta*. Microbial Cell Factories 2011;10:1–9.
49. Quiroz RE, Balcázar E, Dantán E, Folch JL, Martínez C, Martinez A. Characterization of cellulolytic activities of *Bjerkandera adusta* and *Pycnoporus sanguineus* on solid wheat straw medium. Electron J Biotechnol 2009;12:1–8.
50. Cantarel BL, Coutinho PM, Rancurel C, Bernard T, Lombard V, Henrissat B. The carbohydrate-active enzymes database (CAZy): an expert resource for glycogenomics. Nucleic Acids Res 2009;37:233–238.
51. Wong DWS, Chan VJ, McCormack AA, Hirsch J, Biely P. Functional cloning and expression of the schizophyllum commune glucuronoyl esterase gene and characterization of the recombinant enzyme. Biotechnol Res Int 2012;2012:1–7.