Tailored catalysts for plant cell-wall degradation: Redesigning the exo/endo preference of Cellvibrio japonicus arabinanase 43A


*Institute for Cell and Molecular Biosciences, University of Newcastle upon Tyne, Newcastle upon Tyne NE2 4HH, United Kingdom; and §Structural Biology Laboratory, Department of Chemistry, University of York, Heslington, York YO10 5YW, United Kingdom

Communicated by Arnold L. Demain, Drew University, Madison, NJ, January 4, 2005 (received for review August 17, 2004)

Enzymes acting on polymeric substrates are frequently classified as exo or endo, reflecting their preference for, or ignorance of, polymer chain ends. Most biotechnological applications, especially in the field of polysaccharide degradation, require either endo- or exo-acting hydrolases, or they harness the essential synergy between these two modes of action. Here, we have used genomic data in tandem with structure to modify, radically, the chain-end specificity of the Cellvibrio japonicus exo-arabinanase CjArb43A. The structure of Bacillus subtilis endo-arabinanase 43A (BsArb43A) in harness with chain-end recognition kinetics of CjArb43A directed a rational design approach that led to the conversion of the Cellvibrio enzyme from an exo to an endo mode of action. One of the exo-acting mutants, D35L/Q316A, displays similar activity to WT CjArb43A and the removal of the steri block mediated by the side chains of Gin-316 and Asp-53 at the –3 subsite. This study provides a template for the production of tailored industrial catalysts. The introduction of subtle changes informed by comparative 3D structural and genomic data can lead to fundamental changes in the mode of action of these enzymes.

exo/endo glycoside hydrolases | rational design

The most abundant sources of renewable carbon in the biosphere are plant structural polysaccharides (PSPs). Annually, ~1011 tons of these polymers are synthesized (1–3), with an energy content equivalent to 640 billion tons of oil (4). These saccharides represent a valuable industrial substrate in the energy, pulp and paper, detergent, textile, human food, and animal feed sectors (1, 2, 5–7). There is considerable interest in the enzymatic degradation and modification of these substrates to more efficient and economical processes. Thus, the enzymatic degradation and modification of these substrates is of enormous biotechnological importance. A key feature of all enzyme-based processes is the balance between exo (chain-end specific) and endo (internal cleaving) enzymes. Although many biotechnological applications, particularly in the paper and pulp, textile, detergent, human food, and animal feed sectors use endo-acting enzymes (5–7), others, such as the complete saccharification of complex and highly recalcitrant polysaccharides, harness the synergy of endo- and exo-acting hydrolases (9, 10). The structural basis for endo- or exo-acting glycoside hydrolases is critical in the rational design of enzymes that are tailored for specific industrial applications.

Based primarily on studies of cellulases and amylases, the structural bases for exo and endo activity have seemed to be clear. Endo-acting enzymes generally contain an open extended substrate binding cleft that can interact at random positions along the saccharide chains (11–13). The topographies of the active site of exo-acting glycoside hydrolases are also consistent with their mode of action. Enzymes that release monosaccharides have pocket topographies, and thus, only the terminal sugar residues of polymeric chains can enter the active site (14). Cellulases offer a third topography, the “tunnel,” through which processing glucan chains are threaded (15). Recently, a more complex and subtle basis for exo and endo preference has emerged. Thus, selected enzymes that contain open clefts can display typical exo activity, indicating that binding-energies and steric factors within individual subsites contribute to the mode of action of these biocatalysts in a similar manner to the more gross and obvious structural features (16). This latter phenomenon should simplify specific engineering to tailor the substrate binding preferences of these industrially important biocatalysts, as exemplified by the work described here.

Linear arabinan, a pectic polysaccharide comprising α-1,5-linked L-arabinofuranosides (17, 18), is hydrolyzed by arabinanases (19), which find numerous applications, especially in the human food sectors (20–22). Before this study, the sole 3D structure for an arabinanase was that of the unusual and highly active exo-acting gluco-oligosaccharide hydrolyase family 43 arabinanase from Cellvibrio japonicus, CjArb43A (23, 24).

Here, we describe a multifaceted approach, based on exploitation of the genomic resource, subsite and chain-end recognition kinetics, 3D structure solution of relevant enzymes, and directed mutagenesis to retarget CjArb43A toward internal glycosidic bonds in arabinan. The redesigned endo-enzyme remains a more powerful catalyst than unmodified endo-acting arabinanases. We demonstrate the feasibility of engineering industrially relevant modes of action into highly active carbohydrate modifying enzymes, thus increasing their role in biotechnological processes.

Experimental Procedures

Sources of Sugars Used. Arabinooligosaccharides with a degree of polymerization (d.p.) of <7 and polysaccharides were obtained from Megazyme International (Bray, Ireland). To generate arabinooligosaccharides with a d.p. of >7, linear arabinan (5 g) was partially hydrolyzed with BsArb43A, such that most products had a d.p. of 5–15. The reaction products were subjected to size-exclusion chromatography by using two Biogel P-2 columns (120 × 2.5 cm) in series and then four Biogel P-4 columns (120 × 2.5 cm) in series. The final products (1–3 mg) were >95% pure. To generate reduced arabinooligosaccharides, the saccharides were reacted with sodium borohydride (25), followed by size-exclusion chromatography.

Abbreviation: d.p., degree of polymerization.

Data deposition: The atomic coordinates for BsArb43A have been deposited in the Protein Data Bank, www.pdb.org (PDB ID code 1UV4).

1M.R.P. and E.J.T. contributed equally to this work.

To whom correspondence should be addressed. E-mail: h.j.gilbert@newcastle.ac.uk.

© 2005 by The National Academy of Sciences of the USA

www.pnas.org/cgi/doi/10.1073/pnas.0500051110

PNAS | February 22, 2005 | vol. 102 | no. 8 | 2697–2702
Bacterial Strains, Culture Conditions, and Plasmids. The bacterial strains that we used were Escherichia coli TUNER (Novagen) and XL1-Blue. These organisms were cultured in Luria broth at 37°C with aeration, unless indicated otherwise. The plasmids pVM1 and pEJTI, which are recombinants of pET22b, encode mature CjArb43A and BsArb43A, respectively. pVM1 was described (23), whereas pEJTI was generated as follows. The gene (bsarb43A) encoding mature BsArb43A was PCR-amplified from the genome of Bacillus subtilis by using primers ET1 and ET2 (Table 2, which is published as supporting information on the PNAS web site), which bind to nucleotides 180–198 of bsarb43A and nucleotides 10–28 upstream of the published translational stop codon of the arabinanase gene. The amplified DNA was cloned into the NdeI and Xhol sites of pET22b. The sequence of bsarb43A contained an adenine nucleotide insertion at position 922 (present in all sequenced DNA molecules derived from three independent PCRs), which is not evident in the published sequence (which, thus, contains an error) and altered the C-terminal sequence of BsArb43A to ANDNGIPKLLINDL-NWSSGWPST (residues in italics differ from the published sequence). The corrected version of bsarb43A has been deposited in the GenBank database (accession no. Z75208).

Site-Directed Mutagenesis of BsArb43A and CjArb43A. Amino acid substitutions of BsArb43A and CjArb43A were generated by the QuikChange site-directed mutagenesis kit (Stratagene) by using pEJTI and pVM1 as template DNA, respectively, and the appropriate primers (Table 2). To generate derivatives of the arabinanases in which specific loops were exchanged, inverse PCR was used with the phosphorylated primers given in Table 2. The PCR product was purified, ligated, and transformed into E. coli XL1-Blue.

Expression and Purification of BsArb43A and CjArb43A. E. coli strain TUNER, harboring the appropriate recombinant plasmid, was cultured in LB containing 50 μg/ml ampicillin at 37°C to an A600 of 0.6, and expression of the arabinanase was induced by addition of 0.2 mM isopropyl β-thiogalactopyranoside and incubation for an additional 20 h at 16°C. CjArb43A and BsArb43A were purified from cell-free extract, and material released by osmotic shock, respectively, by anion exchange and gel filtration (24).

Enzyme Assays. Activity of BsArb43A and CjArb43A against linear arabinan was assayed at 27°C in 50 mM sodium phosphate buffer as described by McKie et al. (23), except that reducing sugar was determined by using the Somogyi-Nelson reagent (26). The activity of the two arabinanases against arabinooligosaccharides was determined by analyzing the reaction products released by osmotic shock, respectively, by anion exchange and gel filtration (24).

CjArb43A Releases Arabinotriose from the Nonreducing End of Arabinan. CjArb43A isexo-acting releasing arabinotriose from linear α-1,5-arabinans (23) (Fig. 1). The enzyme has high activity compared with other GH4 arabinanases, with a kcat of 207 s⁻¹ and Km of 1.5 mg·ml⁻¹ (23). The 3D structure of CjArb43A in complex with arabinohexaose revealed a five-bladed β-propeller structure (24) and three subsites on either side of the catalytic center (i.e., subsites extend from −3 to +3), consistent with the liberation of arabinotriose as the major product of exo activity. The symmetry of arabinan meant that unambiguous assignment of the oligosaccharide chain direction was not possible, nor was it known at which end of the polysaccharide substrate degradation occurred. CjArb43A does not possess the “tunnel,” “blind-canyon,” or “pocket” active-center topographies observed inexo-β-glucanases and amylases (14, 33); instead, the catalytic residues lie in a long open groove across the enzyme surface, more akin to the topography displayed byendo-enzymes.

The most obvious structural difference with enzymes annotated as “endo-arabinanases” is the presence of a large-loop insertion (residues 205–217) that seals the putative aglycone end of the substrate-binding groove of CjArb43A, which is replaced by shorter loops in reportedendo-arabinanase sequences. However, the exchange of this loop for the equivalent loop in BsArb43 (described in more detail below) results in an enzyme that is 350 times less active than the parent and displays exo activity, as evidenced by the release of arabinotriose only (data not shown), indicating that the presence or absence of this loop is unrelated to the mode of attack of the enzyme.

To address the ambiguity of the direction of the bound substrate chain in CjArb43A (and, hence, the features responsible for chain-end recognition), reduced arabinooligosacchar-
ides were used as substrate. The initial cleavage of arabinocytotol by CjArb43A generates arabinopentatol and arabinotriose (Fig. 2), demonstrating that the enzyme attacks linear arabinan from the nonreducing end, indicating that binding in the glycone (−1 to −3) region of the substrate-binding groove confers the exo mode of action displayed by the Cellvibrio enzyme. Given that this is the opposite end of the cleft to the large insertion described above, it was important to compare at a detailed level the differences between classical endo-arabinanases and the CjArb43A exo-arabinanase in this region.

Characterization of BsArb43A. Previous studies designated BsArb43A as a GH43 endo-α-1,5-arabinanase based on sequence similarity with enzymes that display typical endo activity. To establish whether BsArb43A is a typical endo-α-1,5-arabinanase, its gene (bsarb43A) was amplified by PCR from the B. subtilis genome and cloned, and the encoded enzyme was overexpressed in E. coli and purified to electrophoretic homogeneity.

BsArb43A displays catalytic properties typical of an endo-α-1,5-arabinanase. It initially generates a range of oligosaccharides from linear arabinan, and these oligosaccharides become smaller as the reaction proceeds (Fig. 1), although it displays no activity against heavily substituted arabinans or a range of other polysaccharides. The activity of BsArb43A against linear α-1,5-linked arabinooligosaccharides increases with the d.p. of the substrate, with k_{cat}/K_{m} values of 5.8 × 10^2 min\(^{-1}\)mM\(^{-1}\) for arabinotriose to 2.9 × 10^6 for arabinopentaose and 3.7 × 10^6 min\(^{-1}\)mM\(^{-1}\) for arabinohexaoase and 6.0 × 10^6 min\(^{-1}\)mM\(^{-1}\) for linear arabinan. The difference in the hydrolysis rate between arabinohexosae and arabinopentaose is modest, implying that the enzyme contains five to six significant arabinose binding sites. The product profiles generated from arabinooligosaccharides reveal that arabinopentaose is hydrolyzed exclusively into arabinobiose and arabinotriose (data not shown), demonstrating that the enzyme contains at least two subsites either side of the site of cleavage. The catalytic activity of BsArb43A against both polysaccharides and oligosaccharides is ∼10-fold lower than CjArb43A.

The 3D Structure of BsArb43A. The crystal structure of BsArb43A was determined by molecular replacement, at a 1.5-Å resolution, using CjArb43A as the search model (Fig. 3a). BsArb43A is composed of a single five-bladed β-propeller. It has a cylindrical shape with diameter and height of ∼40 Å. The unusual fold displayed by GH43 enzymes is also evident in a GH32 invertase (34) and GH68 fucosidase (35). The BsArb43A propeller is based upon a 5-fold repeat of blades that adopt the classical “W” topology of four antiparallel β-strands that comprise a β-sheet. The axial cavity contains two ions modeled as Ca\(^{2+}\). The β-propeller fold is “closed” by the fifth blade, comprising strands from the N and C terminus. Such closure of β-propeller proteins is colloquially termed “molecular velcro” (24).

BsArb43A contains a deep V-shaped cleft that extends along the full length of the protein. Located in the central part of this cleft is the constellation of carboxylate amino acids Asp-14, Asp-133, and Glu-185, which overlays with the CjArb43A residues Asp-38, Asp-158, and Glu-221 and constitutes the catalytic apparatus (Fig. 3b). The cleft, which is 28.3 Å long, is consistent with its capacity to accommodate five residues and its maximal activity against arabinopentaose.
Function of Conserved Residues in GH43 Arabinanases. The function of amino acids conserved in the active sites of both \( \text{BsArb43A} \) and \( \text{CjArb43A} \) was assessed in the \( \text{Bacillus} \) enzyme by using alanine-scanning mutagenesis (full data set is given in Table 1). D14A, D133A, and E185A are inactive, consistent with the proposed catalytic function of Asp-14, Asp-133, and Glu-185. \( \text{BsArb43A} \) Trp-74 and Phe-151 are in equivalent positions to the respective \( \text{CjArb43A} \) amino acids Trp-94 and Phe-176. Substituting each of these residues in \( \text{BsArb43A} \) causes a significant decrease in activity against linear arabinan and the oligosaccharides arabinotetraose and arabinohexaose. In particular, note that W74A is \( \times 10^4 \) less active than the WT arabinanase against both oligosaccharides and the polysaccharide linear arabinan, which reflects its location at the critical \(-1\) subsite of the enzyme. Likewise, mutation of Phe-151, implicated in the +1/+2 subsites results in an \( \approx 50\) - and 1,000-fold reduction in activity against linear arabinan and arabinooligosaccharides, respectively.

Manipulation of the Glycone Region of the Substrate Binding Groove. A key unresolved question regarding \( \text{CjArb43A} \) is the mechanism by which the enzyme displays exo activity through recognition of the nonreducing-chain end of arabinan. The glycone region of the substrate binding groove reveals two major features that are different between the \( \text{Bacillus} \) and \( \text{Cellvibrio} \) arabinanase. (i) \( \text{CjArb43A} \) Pro-55 is replaced in \( \text{BsArb43A} \) by a five-residue loop, LTER, that alters the structure of the \(-3\) subsite; and (ii) \( \text{CjArb43A} \) contains two polar residues, Gln-316 and Asp-35, that interact with O5 of the sugar at the \(-3\) subsite, which in the \( \text{Bacillus} \) enzyme are Ala-277 and Leu-11, respectively.
To interrogate the functional significance of these two differences, the LTEER loop was inserted into \textit{Cj}\textsubscript{Arb43A} (LTEER/CjArb43A), and the mutations D35L and Q316L were introduced into the \textit{Cellvibrio} enzyme. LTEER/CjArb43A was 130-fold less active than WT CjArb43A. In contrast, the D35L and Q316A mutations, either individually or in combination (D35L/Q316A), caused only a very modest decrease in activity, with the mutants displaying catalytic efficiencies that were 77% (D35L), 75% (Q316A), and 50% (D35L/Q316A) of the WT enzyme (Table 1). These data indicate that the loop insertion influences the integrity of the −3 subsite, which, by inference, plays an important role in enzyme activity. In contrast, the hydrogen bonds that Gln-316 and Asp-35 make with the O5 of the −3 sugar are not critical to the function of the distal glycone subsite. Analysis of the reaction products generated by LTEER-CjArb43A D35L, Q316A, and D35L/Q316A from linear arabinan reveal a significant change in the mode of action of these enzymes, compared with the WT arabinanase.

WT CjArb43A exclusively yields arabinotriose as the product from linear arabinan. During the initial phase of the reaction, LTEER-CjArb43A generates a range of oligosaccharides with d.p. values of 2–11 (Fig. 4). Similarly, D35L/Q316A also releases oligosaccharides with d.p. values of 2–10, although the trisaccharide is produced in higher quantities than the other reaction products (Fig. 4). Although arabinotriose is the dominant reaction product produced by D35L and Q316A, even these single mutants produce significantly more oligosaccharides with a d.p. of >3 than WT CjArb43A (Fig. 4). To provide more quantitative information on the range of reaction products, WT CjArb43A and the four mutants were incubated with linear arabinan, and the ratio of arabinotriose and reducing sugar generated was compared during the initial stages of the reaction (<2% of substrate hydrolyzed). There is a dramatic decrease in the ratio of arabinotriose/reducing sugar for LTEER-CjArb43 and D35L/Q316A, compared with the WT \textit{Cellvibrio} enzyme, unambiguously demonstrating a switch from exo to an endo mode of action (Fig. 5).

To explore the endo/exo mode of action of the mutant proteins more rigorously, we compared their capacity to hydrolyze oligosaccharides that occupy the full length of the substrate-binding groove. Example data for reduced arabinooctaose are shown in Fig. 2. Although the WT enzyme initially generated arabinotriose and an oligosaccharide with a d.p. of \( n - 3 \) (where \( n \) is the d.p. of the substrate) from arabinooctaose and arabinononaose, the mutants produced a range of reaction products from these substrates. These

![Fig. 4. Reaction products generated by WT and mutants of GH43 arabinanases.](image-url)
data demonstrate the exquisite specificity of WT CjArb43A for the glycosidic bond between sugar residues 3 and 4 from the nonreducing end, and, thus, the −3 subsite displays very tight specificity for the terminal sugar. By contrast, the more complex reaction products generated by the mutant enzymes from the reduced oligosaccharides, which include arabinotetraose, demonstrate that substrate extends distal of the −3 subsite. Thus, the CjArb43A derivatives display endo activity against oligosaccharides in addition to linear arabinan.

Together, the data show that the −3 subsite of CjArb43A is the major factor in exo activity. This introduction of the LTEE(R)Arb43A (data not shown), in which the inserted loop is highly disordered.

The mechanism by which the −3 subsite of CjArb43A mediates arabinotriose release from the nonreducing end of the substrate is an interesting question. It is possible that the −3 subsite displays strong specificity for the free O5 of the nonreducing terminal arabinose. The observation that Asp-35 and Glu-316 hydrogen bond with the hydroxyl at C5 supports this proposal (Fig. 3c). However, the removal of these interactions, although it changes the mode of action of the enzyme, causes only a very modest reduction in catalytic activity, suggesting that these two hydrogen bonds do not make a substantial contribution to the binding energy of the −3 subsite. Thus, it seems more likely that Asp-35 and Glu-316 do not harness profound binding energy but instead make a steric barrier to any sugar appended to O5 of the arabinofuranose located at the −3 subsite. The importance of Asp-35 and Glu-316 in conferring exo activity is illustrated by the observation that these residues are not present in any of the characterized GH43 endo-acting arabinanases.

Conclusions

Historically, exo–endo activity conversions have achieved limited success. Warren and colleagues (36) demonstrated that removal of the loop comprising the ceiling of the tunnel in a GH7 cellobiohydrolase from Cellulomonas fimi Cel7B increased the endo character of the enzyme, as reflected in viscosity measurements. CjArb43A, in contrast to such cellulases, acts as an exo enzyme through subtle changes to its open active-center cleft. Here, we show that steric restriction at the −3 subsite confers exo activity and, remarkably, endo activity can be introduced into the enzyme through modification of this region of the active site without compromising catalytic efficiency. This study demonstrates how comparative 3D structural information can inform rational design strategies that lead to industrially relevant, fundamental changes in the mode of action of enzymes, and thus, it provides a template for the production of engineered industrial catalysts.

We thank the staff of the Daresbury Synchrotron Radiation Source for assistance. This work was supported by the Biotechnology and Biological Sciences Research Council.