THE JOURNAL OF BIOLOGICAL CHEMISTRY VOL. 289, NO. 11, pp. 7362–7373, March 14, 2014
Published in the U.S.A.

Mechanistic Strategies for Catalysis Adopted by Evolutionary Distinct Family 43 Arabinanases*

Received for publication, November 27, 2013, and in revised form, January 14, 2014 Published, JBC Papers in Press, January 27, 2014, DOI 10.1074/jbc.M113.537167

Camila R. Santos[‡1], **Carla C. Polo**[‡1], **Maria C. M. F. Costa**[‡], **Andrey F. Z. Nascimento**[‡], **Andreia N. Meza**[‡], **Junio Cota**[§], **Zaira B. Hoffmam**[§], **Rodrigo V. Honorato**[‡], **Paulo S. L. Oliveira**[‡], **Gustavo H. Goldman**[§], **Harry J. Gilbert**[¶], **Rolf A. Prade**[§], **Roberto Ruller**[§], **Fabio M. Squina**[§], **Dominic W. S. Wong**[‖], **and Mário T. Murakami**[‡2]

From the [‡]*Brazilian Biosciences National Laboratory and* [§]*Brazilian Bioethanol Science and Technology Laboratory, Brazilian Center of Research in Energy and Materials, Campinas/SP, 13083-970, Brazil, the* [¶]*Institute for Cell and Molecular Biosciences, The Medical School, Newcastle University, Newcastle Upon Tyne, Tyne and Wear NE1 7RU, United Kingdom, and the* [‖]*Western Regional Research Center, Agricultural Research Service, United States Department of Agriculture, Albany, Calfornia 94710*

Background: Arabinanases are key enzymes involved in hemicellulose degradation.
Results: Crystallographic, mutational, and biochemical assays of three arabinanases reveal the molecular mechanisms governing their catalysis and activation.
Conclusion: Accessory domain and metal ion are essential for catalysis. Structural adaptations in the catalytic interface confer unique action modes to ruminal arabinanases.
Significance: This work provides new molecular strategies for arabinan hydrolysis.

Arabinanases (ABNs, EC 3.2.1.99) are promising catalysts for environmentally friendly biomass conversion into energy and chemicals. These enzymes catalyze the hydrolysis of the α-1,5-linked L-arabinofuranoside backbone of plant cell wall arabinans releasing arabino-oligosaccharides and arabinose, the second most abundant pentose in nature. In this work, new findings about the molecular mechanisms governing activation, functional differentiation, and catalysis of GH43 ABNs are presented. Biophysical, mutational, and biochemical studies with the hyperthermostable two-domain endo-acting ABN from *Thermotoga petrophila* (TpABN) revealed how some GH43 ABNs are activated by calcium ions via hyperpolarization of the catalytically relevant histidine and the importance of the ancillary domain for catalysis and conformational stability. On the other hand, the two GH43 ABNs from rumen metagenome, ARN2 and ARN3, presented a calcium-independent mechanism in which sodium is the most likely substituent for calcium ions. The crystal structure of the two-domain endo-acting ARN2 showed that its ability to efficiently degrade branched substrates is due to a larger catalytic interface with higher accessibility than that observed in other ABNs with preference for linear arabinan. Moreover, crystallographic characterization of the single-domain exo-acting ARN3 indicated that its cleavage pattern producing arabinose is associated with the chemical recognition of the reducing end of the substrate imposed by steric impediments at the aglycone-binding site. By structure-guided rational design, ARN3 was converted into a classical endo enzyme, confirming the role of the extended Arg^{203}–Ala^{230} loop in determining its action mode. These results reveal novel molecular aspects concerning the functioning of GH43 ABNs and provide new strategies for arabinan degradation.

Endo-1,5-α-L-arabinanases (ABNs)[3] (E.C. 3.2.1.99) are key enzymes in the breakdown of plant cell wall arabinans and their products such as arabinose and arabino-oligosaccharides have potential applications as low-calorie sweetener and prebiotics, respectively (1, 2). These enzymes are also considered promising tools in the environmentally friendly biomass conversion into energy and chemicals (3) because pentoses can also be used as carbon sources and hemicellulosic oligosaccharides can competitively inhibit the cellulolytic system, lessening the hydrolysis efficiency. Indeed, recent studies showed that the supplementation of fungal enzyme mixtures with ABNs and arabinofuranosidases (ABFs, E.C. 3.2.1.55) improved the hydrolysis rate of plant biomass (4–6). Moreover, genomics and proteomics approaches have revealed an impressive expansion of the GH43 family, which includes ABNs and ABFs, in a number of organisms specialized in plant cell wall degradation. This observation highlights the importance of this CAZy family for depolymerization of lignocellulosic material and, probably, reflects a broader range of specificities than described so far.

GH43 ABNs consist of a five-bladed β-propeller catalytic core that can be found as a discrete protein (7–10) or fused to an ancestral ancillary domain with a β-barrel-like architecture (11). Although the accessory β-sandwich domain, found in GH43 β-xylosidases and ABFs, contributes to the active site formation of its parental protein (12–14), in ABNs the β-barrel-like domain does not participate in the groove formation. To date, there are no structural and functional studies with deletion mutants of two-domain ABNs to determine the correlation

* This work was supported by Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP) Grants 10/51890-8 and 13/13309-0, Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) Grants 476043/2011-5 and 308092/2012-0, and the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES).

[1] Both authors contributed equally to this work.

[2] To whom correspondence should be addressed: Giuseppe Maximo Scolfaro 10000, ZIP 13083-970, Campinas/SP, Brazil. Tel.: 55-19-3512-1106; Fax: 55-19-3512-1100; E-mail: mario.murakami@lnbio.cnpem.br.

[3] The abbreviations used are: ABN, arabinanase; GH, glycoside hydrolase; ABF, arabinofuranosidase; TpABNΔC, TpABN devoid of the accessory domain; TRX, thioredoxin.

of this ancestral domain with catalysis and stability of these enzymes. Moreover, all structurally characterized ABNs so far (7–11) contain a divalent ion bound to a solvated channel located at the inner region of the propeller-fold, which recently has been demonstrated to be essential for catalysis (11). Nevertheless, the structural basis for the ion role remains unresolved and there has been no evidence for the existence of divalent ion-independent GH43 ABNs.

The synergism between endo- and exo-acting hydrolases is a *sine qua non* for complete saccharification of complex and recalcitrant polysaccharides (15, 16) such as arabinan that consists of a backbone of α-1,5-linked L-arabinofuranosyl residues decorated with α-1,3- and α-1,2-linked arabinofuranosides and can be found freely or attached as a side chain of xylogalacturonan and rhamnogalacturonan (17). However, only two GH43 exo-ABNs have been characterized so far. BsArb43A releases a trisaccharide from the non-reducing end of arabinan and the structural basis for its action mode is the partial blocking at the −3 subsite (8). ARN3 is an exo-ABN that releases arabinose from both arabinan and arabinooligosaccharides for which there is no structural information (18). Most GH43 ABNs (8, 10, 19–21) show preference for linear substrates, despite their natural substrate, arabinan, is commonly a branched chain. ARN2, a functionally characterized two-domain ABN presented similar affinity for both linear and branched arabinan (22), which is interesting and also relevant to biomass degradation, but the structural basis for this feature is still unknown.

Thus, to shed light on the structural determinants for the distinct molecular mechanisms governing the catalysis of this important family of glycosidases, we have investigated three GH43 ABNs: two endo-acting ABNs, TpABN and ARN2, which are two-domain enzymes, from *Thermotoga petrophila* RKU-1 (23–24) and from a bovine rumen metagenomic library (22), respectively; and, a single-domain exo-ABN (ARN3) recovered from the same metagenomic DNA library (18). Biochemical and biophysical studies with TpABN contributed to elucidate the structural basis for calcium dependence of some GH43 ABNs and the role of the ancestral ancillary module for function and stability of two-domain ABNs. The structure, mutagenesis, and kinetics of ARN3 revealed that its exo-acting mode (18) is related to the chemical recognition of the reducing ends of arabinan imposed by a steric barrier at the aglycone-binding region. Moreover, structural data of ARN2 showed that its unique ability to efficiently degrade branched substrates as observed by Wong and colleagues (22) is due to a larger catalytic interface with higher solvent accessibility than that observed in other ABNs with preference for linear arabinan. These molecular findings provide a further understanding in the functional aspects, stability, and regulation of GH43 ABNs and represent an advance toward new strategies for hemicellulose degradation.

EXPERIMENTAL PROCEDURES

Molecular Cloning and Mutagenesis—The pET28a/TpABN, pET32b/ARN2, and pET32b/ARN3 constructs are described in previous works (18, 22, 24). The TpABN devoid of the accessory domain (TpABNΔC, residues Ala^{28}-Leu^{371}) was amplified by PCR using the oligonucleotides 5′-GCTAGCGAACAACCCACCTTTCGATG-3′ and 5′-GGATCCTTATTCTTCCACTCTTATTCCCC-3′ and the pET28a/TpABN vector as template. The amplified sequence was cloned into pGEM-T Easy vector (Promega) and further subcloned into pET28a. Amino acid substitutions of TpABN and ARN3 were generated by the QuikChange site-directed mutagenesis kit (Stratagene) using pET28a/TpABN and pET32b/ARN3 vectors as templates and the primers 5′-GAAAACATCCCAATGCACTCGATCCTTGCGTTTTTTATGATAA-3′ and 5′-TTATCATAAAAAACGCAAGGATCGAGTGCATTGGGATGTTTTC-3′ for TpABN/G72C; 5′-CACAGTTCCATGGAGGCGCCATACATCCTTTAC-3′ and 5′-GTAAAGGATGTATGGCGCCTCCATGGAACTGTG-3′ for TpABN/G224A; 5′-GTTCCATGGAGGGGCCAGCCATCCTTTACAGTCCTG-3′ and 5′-CAGGACTGTAAAGGATGGCTGGCCCCTCCATGGAAC-3′ for TpABN/Y226A; 5′-TGCCATCGAGGCTCCGGCTATCATCAGGGAAGGG-3′ and 5′-CCCTTCCCTGATGATAGCCGGAGCCTCGATGGCA-3′ for ARN3/F237A; 5′-GTTCGATGCGACCAATCCTTGGGTCCACGATCCTGTGATGG-3′ and 5′-CCATCACAGGATCGTGGACCCAAGGATTGGTCGCATCGAAC-3′ for ARN3/D37W; and 5′-GGGCTAACAACGGTGCACCGAAGCTGGTCATC-3′ and 5′-GATGACCAGCTTCGGTGCACCGTTGTTAGCCC-3′ for ARN3/S329P. The chimera SRGEEP-ARN3 was constructed according to Erster and Liscovitch (25). In this protocol, the entire plasmid is amplified by inverse PCR (26) and sequence replacement is dictated by the primers. The forward primer (5′-GAAGAACCAAATGCCATCGAGGCTCCGTTTATCATCAGGG-3′) contains, at the 5′ portion, the second half of the inserted sequence and, at the 3′ portion, the region downstream to the insertion site. The reverse primer (5′-ACCACGGCTTGCGATGGTCTTCGGCTCCCCCTTTG-3′) contains the first half of the inserted sequence and the region upstream to the insertion site. The pET32b/ARN3 vector was used as template and the Platinum *Pfx* polymerase (Invitrogen) for chain amplification. The resulting product is a linear plasmid that, after analysis by agarose gel electrophoresis, was phosphorylated using T4 Polynucleotide Kinase (ThermoScientific), self-ligated using T4 ligase (Invitrogen), and used to transform *Escherichia coli* DH5α cells. The loop replacement was confirmed by DNA sequence analysis. The wild-type (WT) and mutant constructs of thermostable enzyme, TpABN, were produced with a His$_6$ tag at its N terminus (24). Both ABNs from the bovine rumen microflora metagenomic library, ARN2 and ARN3, were produced as inclusion bodies when expressed with a His$_6$ tag; then a fusion with thioredoxin (TRX) was employed aiming at enhancing solubility when using *E. coli* as the heterologous expression system (22, 18). Mutant versions of ARN2 and ARN3 were also expressed with the TRX fusion tag.

Expression and Purification—All of the proteins were expressed in BL21(DE3) cells supplemented with pRARE2 plasmid for 4 h at 30 °C with 0.5 mM isopropyl 1-thio-β-D-galactopyranoside at $A_{600\ nm}$ of 0.6–0.8. The cells were collected, resuspended in lysis buffer (20 mM sodium phosphate, pH 7.5, 500 mM NaCl, 5 mM imidazole, 1 mM PMSF, and 5 mM benzamidine), and disrupted by lysozyme treatment (80 μg/ml, 30 min, on ice), followed by sonication (~6 pulses of 15 s on ice using a Vibracell VCX 500, Sonics and Materials). The extracts

were centrifuged at 10,000 × *g* for 30 min and the supernatants were loaded onto nickel-charged 5-ml HiTrap chelating columns (GE Healthcare) with a flow rate of 1 ml/min. The proteins were eluted using a nonlinear gradient of imidazole ranging from 5 to 500 mM. The fractions containing the target protein were concentrated using Amicon Ultra centrifugal filters (Millipore) and submitted to size exclusion chromatography using a Superdex 75 16/60 column (GE Healthcare) and a flow rate of 0.5 ml/min. The TRX tag was cut out of the ARN3 construct before the size exclusion chromatography experiment by limited proteolysis using 0.01 mg/ml of trypsin for 30 min at room temperature. All chromatographic steps were performed in an ÄKTA FPLC system (GE Healthcare). Sample homogeneity was confirmed by SDS-PAGE and protein concentration was estimated by absorbance at 280 nm using molar extinction coefficients derived from the sequence of the constructs.

Crystallization, Data Collection, Structure Determination, and Refinement—Crystals of TpABN were promptly obtained without His_6 tag removal. Crystals of ARN3, a single-domain ABN, were obtained only after sample treatment with trypsin that efficiently cleaved out the macromolecular tag. Trials to remove the TRX tag from ARN2 yielded a heterogeneous sample with low stability, suggesting an interaction between the enzyme and macromolecular tag. Based on this observation, we preceded the crystallization with the TRX-ARN2 chimeric protein. The purified TpABN, TRX-ARN2, and ARN3 proteins were concentrated to 29, 5.8, and 12.9 mg/ml, respectively, using Amicon Ultra centrifugal filter units (Millipore) and dialyzed against 20 mM Tris-HCl, pH 7.5 (TpABN), or 20 mM HEPES, pH 7.5 (TRX-ARN2 and ARN3). Crystallization experiments were performed by the vapor diffusion method using a HoneyBee 963 robot (Genomic Solutions). Sitting drops were prepared by mixing 0.5 μl of the protein solution with an equal volume of mother liquor and equilibrated against 80 μl of the reservoir at 18 °C. Formulations based on the following commercial crystallization kits were tested: SaltRX, Crystal Screen, Crystal Screen 2 (Hampton Research), Precipitant Synergy, Wizard I and II (Emerald BioSystems), PACT and JCSG (Qiagen/Nextal), totalizing 544 solutions. Crystals of TpABN with a ligand at the active site were obtained within 1 week in 100 mM sodium citrate, pH 5.5, 40% (*v/v*) PEG400, and 5% (*w/v*) PEG3350. The reduction of the crystallization condition pH to 5.0 yielded crystals with better diffraction quality.

Crystals were directly flash-cooled in a 100 K nitrogen gas stream because PEG400 is a cryoprotectant compound. TpABN crystals in the native state were obtained by changing the dialysis buffer, prior to protein concentration, from 20 mM Tris-HCl, pH 7.5, to 20 mM HEPES, pH 7.5. Small crystals of TRX-ARN2 were obtained within 1 day in 2.1 M sodium malonate and bigger crystals were formed in hanging drops by reducing the precipitant concentration to 1.8 M. Crystals of ARN3 were obtained within 45 days in 100 mM Tris-HCl, pH 8.5, 2.4 M ammonium phosphate and better crystals were obtained reducing the precipitant concentration to 2.1 M. Crystals of ARN2 and ARN3 were cryoprotected using 30% (v/v) glycerol before flash cooling. X-ray diffraction data were collected at the W01B-MX2 beamline (Brazilian Synchrotron Light Laboratory, Campinas, Brazil). Data processing was carried out with HKL2000 or XDS (27, 28). All structures were solved by the molecular replacement method, using the program MolRep (29). The native and ligand-complex structures of TpABN were initially solved by molecular replacement in $P2_12_12$. However, the high values of R_{factor}/R_{free} (26.1/30.1 for TpABN and 26.2/29.7 for TpABN-ligand complex) after model refinement were not compatible with high data quality and resolution indicating an incorrect space group and/or a packing disorder (30, 31). Besides that, the high solvent content (63%) was very suspicious for a well diffracting crystal, but all molecular replacement attempts to find another molecule in the asymmetric unit were unsuccessful. Therefore, data sets were processed in a lower symmetry space group ($P2_1$) and three molecules were found in asymmetric units with interpretable maps and lower R_{factor}/R_{free} values (16.0/19.7 for TpABN and 15.3/19.7 TpABN-ligand complex). The third molecule (chain C) presented a poor map in comparison with chains A and B, probably due to several orientations of this molecule along the crystal lattice. The absence of significant off-origin peaks in the native Patterson map and a few diffused reflections in some frames suggest the presence of rotational order-disorder structure (30). Besides the order-disorder problem, chain C lies on a 2-fold symmetry axis and, because the molecule does not have an intrinsic symmetry, it is not possible to expand symmetry to $P2_12_12$. The YxiA protein from *Bacillus licheniformis* (PDB code 3LV4), sharing 54% identity with TpABN was used as a search model to solve the TpABN structure. The ABN from *Geobacillus stearothermophilus* (AbnB, PDB code 3CU9) was used as search model for ARN3 (43% of identity). The TRX-ARN2 structure was solved using TpABN (identity of 34%) and *E. coli* TRX structures as search models. The structures were refined using the program REFMAC5 (32). After each cycle of refinement, the model was inspected and manually adjusted into the $(2F_o - F_c)$ and $(F_o - F_c)$ electron density maps using the program COOT (33). Solvent molecules were manually added at positive peaks above 3.0 σ in the difference Fourier maps. The local and global quality of the structures was evaluated using the Molprobity server (34). Processing and refinement statistics are reported in Table 1. Interestingly, analysis of the ARN2 crystal lattice showed that the macromolecular tag assisted the crystallization process by providing new molecular surfaces entropically favorable for crystal packing. In fact, the main crystalline interface corresponds to a buried area of 1330 Å^2 and involves inter-chain interactions between TRX and ARN2 from the 2-fold symmetrically related molecules found in the asymmetric unit (Fig. 1*A*). Moreover, when the two molecules of ARN2 from the asymmetric unit were superimposed, the TRX did not superimpose, indicating that the two chimeras adopt different orientations, possibly to maximize the intermolecular contacts (Fig. 1*B*). These observations corroborate that TRX crystalline contacts played a key role in the crystal formation. This strategy was also successfully described for other proteins (35–43) and the use of a short linker connecting the protein target with the macromolecular is essential to produce a rigid and crystallizable chimera (41, 42).

Enzyme Assays—Standard arabinanase activity assays were performed in 50 mM citrate/phosphate/glycine buffer, contain-

TABLE 1

Data collection and structure refinement statistics for TpABN, TpABN-ligand, TRX-ARN2, and ARN3 structures

Values in parentheses correspond to the highest resolution shell.

	TpABN	TpABN-ligand	TRX-ARN2	ARN3
Data collection				
Space group	$P2_1$	$P2_1$	$P2_12_12_1$	$P6_322$
Unit cell parameters (Å) (°)	41.67, 86.91, 194.85, 90.15	41.67, 86.77, 194.37, 90.07	99.91, 114.32, 167.20	140.12, 140.12, 159.01
Resolution (Å)	55.00–1.75 (1.86–1.75)	55.00–1.76 (1.87–1.76)	50.00–1.90 (1.97–1.90)	40.00–2.90 (3.00–2.90)
R_{merge} (%)	5.6 (16.2)	5.4 (18.3)	9.3 (50.5)	15.6 (77.3)
$I/\sigma(I)$	22.8 (8.8)	22.5 (7.9)	16.2 (2.6)	11.8 (2.8)
Completeness (%)	99.6 (98.3)	99.5 (97.3)	98.3 (94.0)	97.5 (99.2)
Multiplicity	7.1 (6.4)	6.6 (5.9)	6.0 (4.7)	7.0 (6.9)
Mosaicity (°)	0.1	0.2	0.5	0.6
Refinement				
Resolution (Å)	40.77–1.75	43.38–1.76	44.37–1.90	35.03–2.90
No. of unique reflections	132,487	128,860	140,105	19,562
R_{work}/R_{free} (%)	16.0/19.7	15.3/19.7	16.1/21.0	21.7/26.7
Overall B factor (Å^2)	26.6	29.4	33.4	54.9
No. of protein atoms	10,731	10,831	9,901	4,864
No. of water molecules	731	699	843	19
Ligands	Calcium, PEG	Calcium, TRIS	Sodium, GOL	PO4
Root mean square deviations				
Bond lengths (Å)	0.019	0.022	0.013	0.007
Bond angles (°)	2.042	2.225	1.507	1.125
Ramachandran plot				
Favored region (%)	96.3	94.1	97.6	91.8
Allowed regions (%)	3.5	5.4	2.2	6.9
Disallowed region (%)	0.2	0.5	0.2	1.3
PDB entry code	4KC7	4KC8	4KCA	4KCB

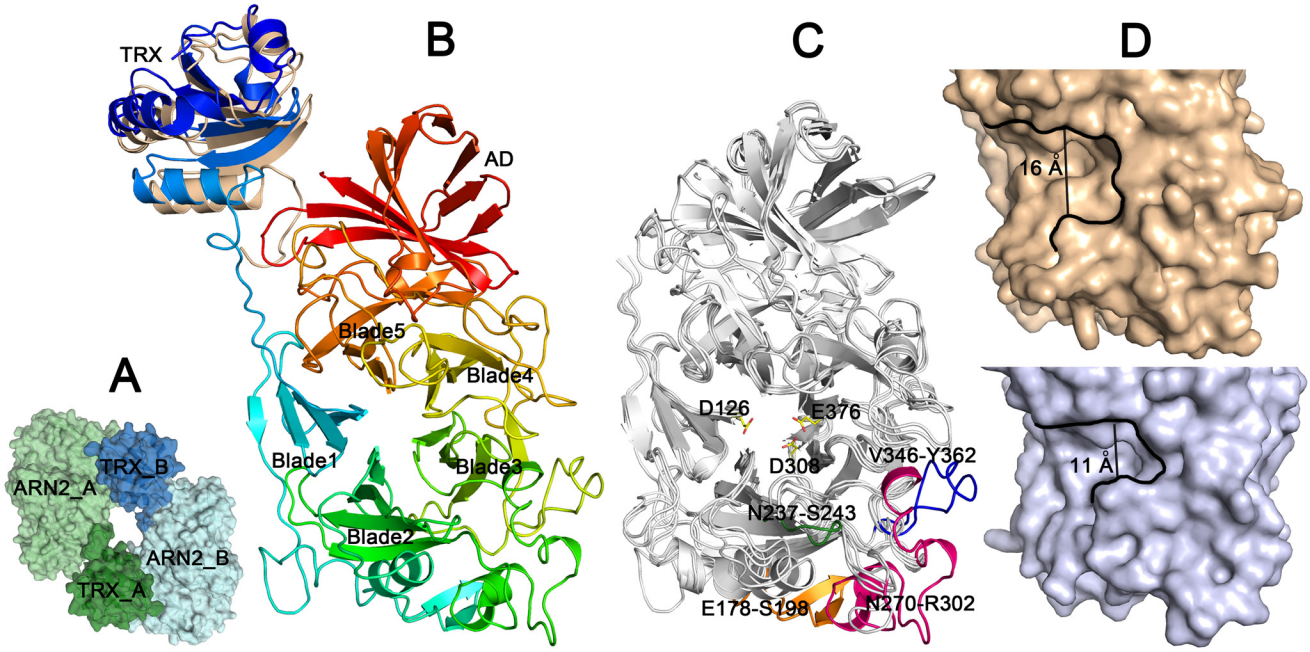


FIGURE 1. **Structural dissimilarities between the ruminal endo-acting enzyme (*ARN2*) and classical two-domain ABNs.** *A*, crystal packing of TRX-ARN2 showing the main interaction interface between the two monomers. One monomer is colored in *green* and the other in *blue*. *B*, arrangement of the chimeric structure colored from the N to C terminus (*blue* to *red*). Superposition of the two chimeras shown in *A* highlights the distinct arrangement of TRX (colored in *wheat* for the second molecule) in relationship to ARN2. *C*, structural superposition of TpABN, BsArb43B, and YxiA on the ARN2 structure, with loop insertions and catalytic triad of ARN2 depicted. *D*, surface topography of the substrate binding channel of ARN2 (*wheat*) in comparison with TpABN (*light blue*).

ing 0.25% (w/v) debranched arabinan (Megazyme International, Ireland) and 8 μg/ml of enzyme. The amount of reduction sugar was determined by the DNS method (44). The enzymes and solutions were previously incubated with Chelex® 100 (Bio-Rad) to remove any trace of metal ion. Chelators (EDTA and EGTA) and metal ions (chloride salts of Ca^{2+}, Mg^{2+}, Mn^{2+}, Ni^{2+}, and Co^{2+}) were used in concentrations of 5 and 10 mM, respectively. The reactions were carried out at pH 6.0 and 70 °C for TpABN/mutants and pH 6.0 and 45 °C for ARN2/ARN3. Temperature profiles were obtained from reactions incubated at pH 6.0 for 10 min at temperatures ranging from 20 to 85 °C. pH curves were obtained from reactions incubated at 70 °C for TpABN/mutants and 45 °C for ARN2/ARN3/mutants for 10 min at pH values from 2 to 10. The apparent kinetic parameters K_m, V_{max}, k_{cat}, and k_{cat}/K_m were calculated by nonlinear regression analysis of the Michaelis-Menten plot. Debranched arabinan was used as substrate in a concentration range from 0.8 to 25.0 mg/ml. Assays were conducted in 50 mM

citrate/phosphate/glycine buffer, pH 7.0, at 45 °C for 10 min. Mathematical calculations were made using GraphPad Prism 5.0 (GraphPad Software). These assays were performed in sextuplicate.

Capillary Zone Electrophoresis—WT and mutant versions of ARN3 (SRGEEP-ARN3, ARN3/F237A, ARN3/D37W, ARN3/S329P, ARN3/D37W/S329P) were incubated with 0.25% (w/v) linear arabinan (Megazyme) for 15 min at 20 °C. Alternatively, 7 mM arabinoheptaose (Megazyme) was used as substrate and the reaction was stopped after 5 and 10 min of incubation at 20 °C. The reaction products were derivatized with 8-aminopyrene-1,3,6-trisulfonic acid by reductive amination as described previously (45). Capillary electrophoresis experiments were performed in a P/ACE MDQ system (Beckman Coulter) with laser-induced fluorescence detection using a fused-silica capillary (TSP050375, Polymicro Technologies) of 50-mm internal diameter and 31-cm length. Samples were injected by application of 0.5 p.s.i. for 0.5 s. Electrophoresis conditions were 15 kV/70–100 mA with the cathode at the inlet, 0.1 M sodium phosphate, pH 2.5, as running buffer and a controlled temperature of 20 °C. The capillary was rinsed with 1 M NaOH followed by running buffer with a dipcycle to prevent carry over after injection. Oligosaccharides labeled with 8-aminopyrene-1,3,6-trisulfonic acid were excited at 488 nm and emission was collected through a 520-nm band pass filter. Retention times varied slightly in different runs due to small volumes of the capillary and variations in the buffer strength and temperature. The resultant peaks were assigned by comparison with electrophoretic behavior of standards: arabinose and 1,5-α-L-arabinooligosaccharides (Megazyme).

Spectroscopy Characterization—Far UV circular dichroism spectra were recorded on a Jasco J-810 spectropolarimeter (Jasco International Co., Ltd., Tokyo, Japan) coupled to a Peltier temperature controller and a 1-mm quartz cuvette. The proteins were previously dialyzed against 20 mM sodium phosphate, pH 7.5, and used at a concentration of 10 μM. The data collection parameters were set to scan rate of 50 nm/min, response time of 4 s, sensitivity of 100 mdeg, accumulation of 10, and delay time for spectrum collection of 60 s. Unfolding experiments were monitored at 230 nm. For thermal denaturation experiments the cuvette was heated at a rate of 1 °C/min.

Computational Analysis—The ARN3 structure was prepared for molecular dynamics simulations using the program YASARA (46). All hydrogen atoms and other missing atoms were created using force field parameters from YAMBER3. A simulation box was defined at 15 Å around all atoms of each structure. Protonation was performed based on pH 7. Cell neutralization was reached filling the box with water molecules and Na/Cl counterions. A short molecular dynamics simulation was performed for solvent relaxation, deleting water molecules until the water density of 0.997 g/ml was reached. A short steepest descent energy minimization was carried until the maximum atom speed dropped below 2200 m/s. Then 500 steps of simulated annealing were performed at 0 K. Finally, a 10-ns simulation at 298 K and with non-bonded cutoff of 7.86 Å was performed (47). A snapshot was saved every 25 ps. Each of the structures was simulated with and without a calcium ion at the coordination site. To analyze the backbone flexibility and stability, the root mean square deviation of each Cα atom was calculated for each structure. Using the algorithm KVFinder (developed by R. V. Honorato, P. S. L. Oliveira, and others) we developed a custom routine that was able to calculate the volume of the coordination site for each simulation snapshot. The matrix-filling procedure for cavity prospection of KVFinder is able to detect and account for slight fluctuations in the coordination site.

RESULTS

Molecular Basis of the Accessory Domain Influence on Stability and Catalytic Activity—The crystal structure of the hyperthermostable GH43 ABN from *T. petrophila* (TpABN), was determined at 1.7-Å resolution, revealed a canonical 5-bladed β-propeller catalytic core with a C-terminal β-barrel-like domain, which lacks sequence similarity with characterized carbohydrate-binding modules. The only structural study addressing a two-domain ABN attempted to produce truncated versions of the protein lacking this C-terminal domain, but no expression was observed in the soluble fraction (11). Herein, to characterize the influence and importance of this accessory domain for stability and catalysis, we designed a truncated construct of TpABN devoid of the C-terminal accessory domain (TpABNΔC, Ala^{28}–Leu^{371}) that was successfully produced in *E. coli*. The truncated protein TpABNΔC was correctly folded, as suggested by circular dichroism analysis (Fig. 2*A*), but presented decreased thermal stability, with a melting temperature (T_m) of 66.4 °C (Fig. 2*B*), whereas the full-length protein TpABN showed a T_m of 88.7 °C (Fig. 2*B*). The unusual circular dichroism spectra of TpABN constructs, with a positive band at 230 nm, are the result of multiple tryptophan side chain absorption in the β-propeller core (48). Structural analysis using the Protein Interactions Calculator (49) showed a large inter-domain area (2,297 $Å^2$) stabilized by extensive polar and hydrophobic contacts (Table 2) that are affected by ionic strength, because the WT protein presented an increase of 7 °C in the T_m under moderate ionic strength (Fig. 2*B*). In single-domain ABNs, the corresponding region is populated by charged and polar residues. These results suggest a direct structural role of the accessory domain in two-domain ABNs and explain the poor solubility and stability of the deletion mutant. Interestingly, despite the native-like properties of the catalytic core, the enzymatic activity of the enzyme was abolished after deletion of the accessory domain, indicating an essential role in catalysis as well. Structural data revealed that the accessory domain participates in the conformational stabilization of the Gly^{243}–Gly^{250} loop (GGLDYRGG, Fig. 2*C*). This loop is adjacent to the active site and forms the substrate-binding groove. The accessory domain maintains the Gly^{243}–Gly^{250} loop in a conformation favorable for substrate recognition and binding, and in the absence of the accessory domain, this loop may adopt a different or non-ordered conformation, disrupting the substrate channel and displacing the substrate. The leucine residue, located at the tip of this loop (Leu^{245} in TpABN, Fig. 2*C*), is conserved in all two-domain ABNs (Fig. 3) and possibly plays a role in the substrate interaction because it establishes aliphatic contacts with carbon atoms from a sugar mimetic (results not shown). In single-domain ABNs, a conserved disulfide bond

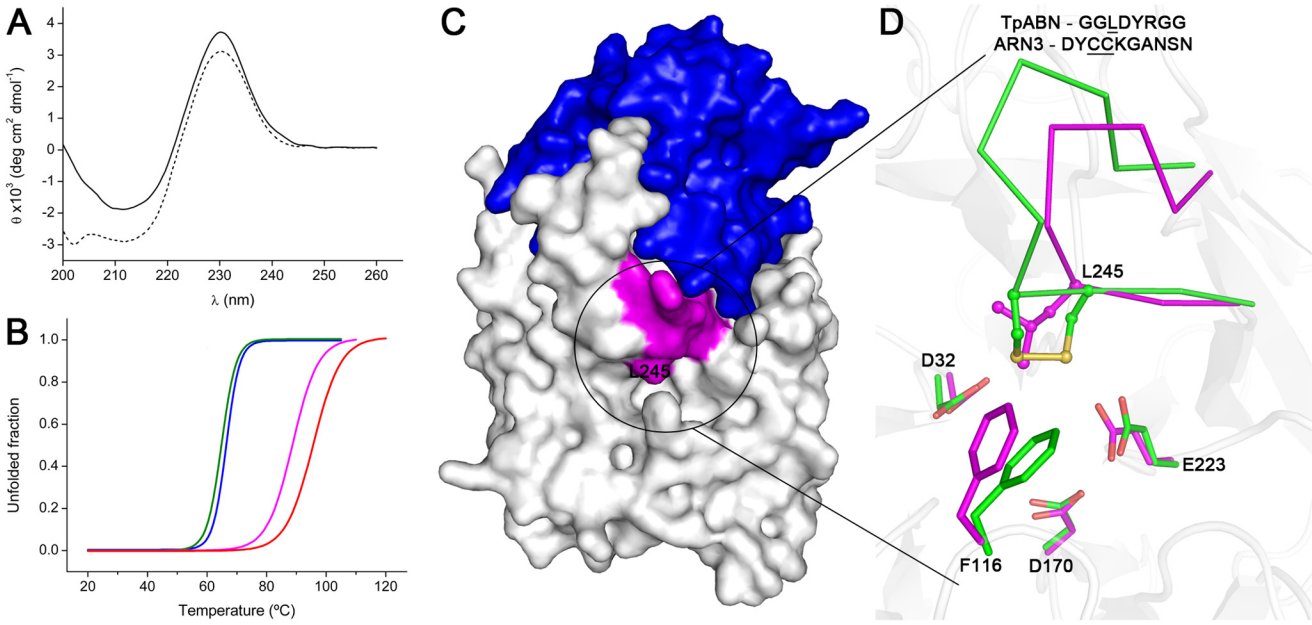


FIGURE 2. **Structural role of the accessory domain in two-domain GH43 ABNs.** *A*, circular dichroism spectra of the hyperthermophilic enzyme (TpABN, *continuous line*) and its deletion mutant (TpABNΔC, *dashed line*). *B*, unfolding curves of TpABN (*magenta* and *red*) and TpABNΔC (*blue* and *green*) without addition of NaCl (*magenta* and *blue*) and in the presence of 250 mM NaCl (*red* and *green*). *C*, molecular surface of TpABN with the catalytic β-propeller in *white* (GGLDYRGG loop in *magenta*) and the β-barrel like accessory domain in *blue*, indicating a highly complementary surface. *D*, structural overlay highlighting the fully conserved disulfide bridge in single-domain ABNs (represented by ARN3, *green*) and the presence of a leucine residue at the corresponding position in two-domain ABNs (represented by TpABN, *magenta*).

TABLE 2

Inter-domain interaction analysis of TpABN using the Protein Interaction Calculator web server (49)

HYDROPHOBIC INTERACTIONS						HYDROGEN BONDS					
DOM	POS	RES	DOM	POS	RES	DOM	POS	RES	DOM	POS	RES
CD	23	P	AD	443	W	CD	27	W	AD	445	Y
CD	23	P	AD	450	W	CD	254	R	AD	387	K
CD	25	F	AD	443	W	CD	268	P	AD	459	N
CD	25	F	AD	445	Y	CD	269	E	AD	459	N
CD	27	W	AD	445	Y	CD	269	E	AD	460	Q
CD	242	F	AD	389	I	CD	285	N	AD	387	K
CD	247	Y	AD	393	I	CD	289	K	AD	385	H
CD	268	P	AD	457	L	CD	289	K	AD	388	E
CD	283	I	AD	389	I	CD	289	K	AD	461	G
CD	288	A	AD	457	L	CD	289	K	AD	463	S
CD	291	V	AD	465	W	CD	292	G	AD	385	H
CD	294	F	AD	383	I	CD	293	N	AD	391	D
CD	294	F	AD	465	W	CD	293	N	AD	394	K
CD	295	I	AD	393	I	CD	295	I	AD	394	K
CD	296	L	AD	383	I	CD	296	L	AD	381	Q
CD	296	L	AD	467	I	CD	297	S	AD	395	Q
CD	307	F	AD	445	Y	CD	298	E	AD	381	Q
CD	307	F	AD	465	W	CD	298	E	AD	398	R
CD	310	V	AD	393	I	CD	300	N	AD	395	Q
CD	335	P	AD	393	I	CD	351	N	AD	436	K
CD	349	F	AD	438	V	CD	352	E	AD	418	K
CD	349	F	AD	440	L	CD	353	D	AD	436	K
CD	355	W	AD	457	L	CD	358	M	AD	455	T
CD	357	V	AD	438	V	CD	358	M	AD	465	W
CD	357	V	AD	457	L	CD	361	F	AD	442	Q
CD	358	M	AD	465	W	CD	366	E	AD	442	Q
CD	360	P	AD	453	V	CD	371	L	AD	417	R
CD	360	P	AD	465	W						
CD	360	P	AD	467	I	IONIC INTERACTIONS					
CD	361	F	AD	445	Y	CD	248	R	AD	391	D
CD	361	F	AD	451	V	CD	267	D	AD	387	K
CD	361	F	AD	453	V	CD	269	E	AD	387	K
CD	361	F	AD	467	I	CD	289	K	AD	388	E
CD	362	P	AD	445	Y	CD	298	E	AD	398	R
CD	368	V	AD	440	L	CD	352	E	AD	418	K
CD	371	L	AD	372	P	CD	352	E	AD	436	K
CD	371	L	AD	376	I	CD	353	D	AD	436	K
CD	371	L	AD	439	L						

(Cys^{254}–Cys^{255} in the ruminal enzyme, ARN3) is observed at the equivalent position of the leucine residue (Fig. 2*D*). This disulfide bridge also participates in the formation of the substrate-binding channel and is important to stabilize the Gly^{243}–Gly^{250} loop (numbering based on the sequence of TpABN) in the proper geometry for substrate binding, enabling single-domain ABNs to be functional without the presence of an accessory domain. The conservation of this disulfide bridge in *Cellvibrio japonicus* exo-arabinanase CjArb43A (Fig. 3) indicates that exo-acting single-domain ABNs share a common mechanism of active site formation with endo-acting single-domain ABNs.

Calcium Ion Caged into β-Propeller Core Is Essential for Catalysis—In all single- and two-domain ABNs with crystallographically determined structures (7–11), a calcium ion was modeled in the highly hydrated cavity located at the inner core of the structure, which connects the catalytic interface with the opposite face of the propeller (Fig. 4*A*). Functional analysis showed that TpABN, as well as BsArb43B (11), was inhibited by chelating agents and only recovered its maximum enzymatic activity in the presence of calcium (Fig. 4*B*). In these calcium-dependent ABNs, the ion interacts with the catalytically relevant His^{314} (TpABN sequence numbering) and six water molecules forming a distorted pentagonal bipyramidal molecular geometry (Fig. 4*C*). These water molecules interact with main chain carbonyl oxygen atoms of Pro^{33}, Ala^{96}, Asp^{170}, Pro^{171}, Gly^{224}, Pro^{225}, and Asn^{315}, which permits the calcium ion to form water-mediated interactions with all five blades of the propeller. In addition, the side chains of Ser^{34} and Gln^{98} interact with W1, W2, and W6 water molecules (Fig. 4*C*), but these interactions seem not to be essential for calcium binding because in the BsArb43B structure, the corresponding residues, Ser^{40} and Asp^{103}, adopt different orientations and are distant from the calcium-coordinating water molecules. A possible structural role of the calcium ion in GH43 enzymes is not evident because thermal stability assays in the presence or absence of the ion did not result in significant changes in the T_m ($\Delta t < 2$ °C).

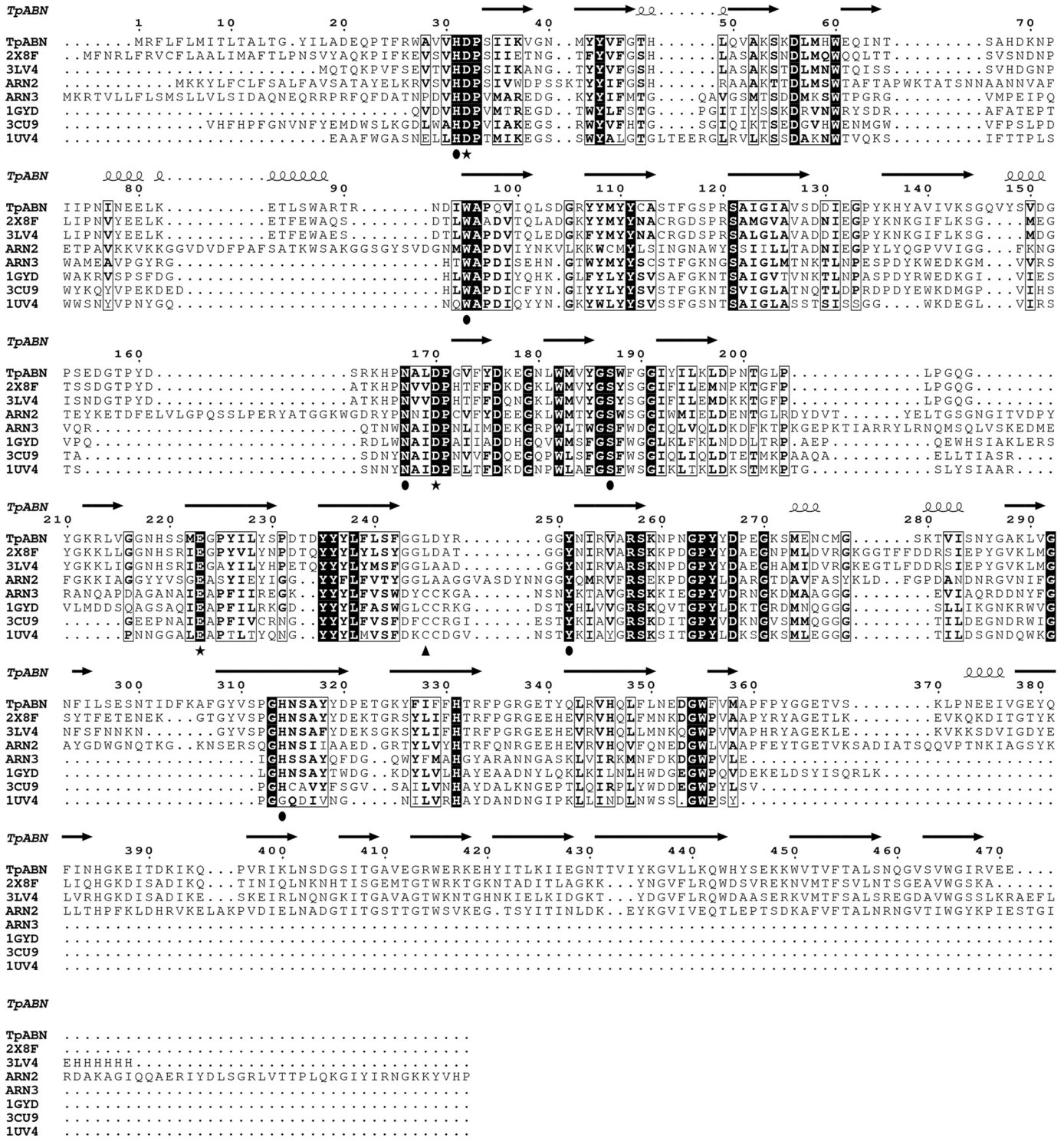


FIGURE 3. **Sequence alignment of structurally characterized GH43 ABNs.** TpABN (this paper), 2X8F (BsArb43B, 11), 3LV4 (YxiA, Northeast Structural Genomics Consortium), and ARN2 (this paper) are two-domain ABNs. ARN3 (this paper), 1GYD (CjArb43A, 7), 3CU9 (AbnB, 10), and 1UV4 (BsArb43A, 8) are single-domain ABNs. The three catalytic residues are marked with *stars* at the *bottom* of the alignment and other residues important for substrate binding and catalysis are marked with *balls*. The cysteine residues of single-domain ABNs and the corresponding leucine of two-domain ABNs are marked with *arrows*.

Ruminal Enzymes Display a Calcium-independent Catalytic Mechanism—Interestingly, the catalytic activity of the ruminal enzymes, ARN2 and ARN3, did not undergo any change by the presence of calcium ions or chelating agents (Fig. 4*B*), indicating a calcium-independent mechanism. In fact, the ARN2 crystalline structure did not present the calcium ion bound to the hydrated cavity. Based on electron density, coordination sphere, and atom distances, we deduced the presence of a sodium ion with distorted pentagonal bipyramidal coordination at the corresponding position (Fig. 4*D*). The geometry is slightly different from calcium coordination, but all direct and water-mediated interactions are fully conserved. The presence

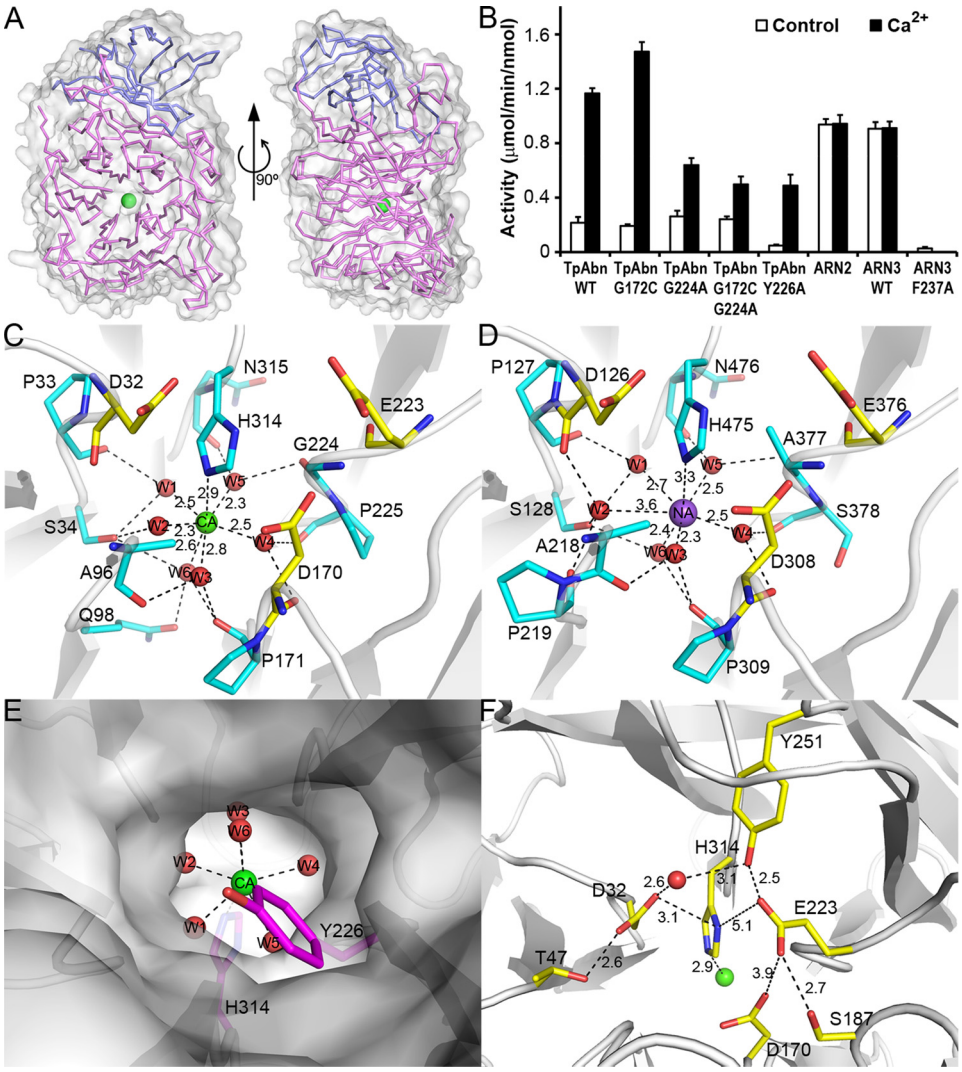


FIGURE 4. **Calcium dependence differs among GH43 ABNs.** *A,* position of the calcium ion (*green*) in the middle of the catalytic propeller (*pink*) at different views. *B,* enzymatic activity of the hyperthermophilic two-domain ABN (*TpABN*) and the two ruminal enzymes (*ARN2* and *ARN3*) in the absence of any metal ion or in the presence of calcium ion. *C,* the pentagonal bipiramidal coordination of the calcium ion in the TpABN structure. *D,* the sodium ion that replaces calcium in the ARN2 structure. *E,* aromatic residue working as a gateway of the hydrated pocket. *F,* interactions involved in the activation of ABNs by calcium/sodium ions.

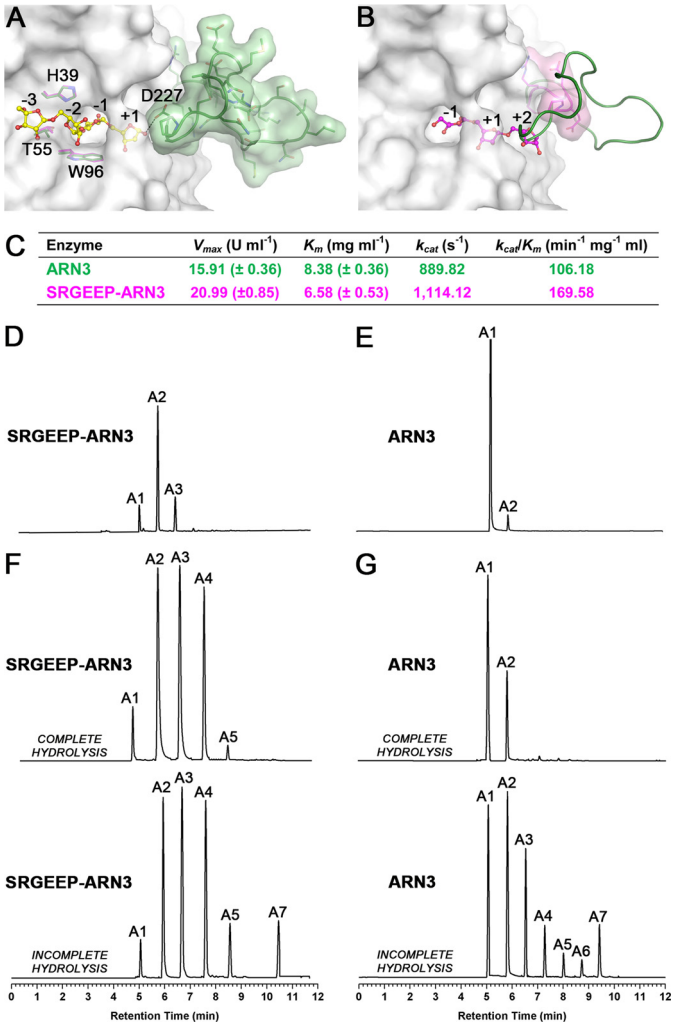
of a potassium ion was discarded due to larger ionic radius and typical bond lengths with oxygen atoms, which is not compatible to the hydrated pocket. Although the water molecule has a similar electron density to sodium ion, the number of interactions and geometry bring it less plausible. Thus, the sodium is the most likely substituent for calcium ions in ARN2 and ARN3 enzymes. To elucidate the structural determinants for calcium dependence, mutants G172C, G224A, and G172C/G224A of TpABN were produced according to comparative structural analysis of the ion-binding pocket. Nonetheless, all mutants showed similar residual activity to the WT enzyme under calcium deprivation (Fig. 4*B*), indicating that these elements are not responsible for the calcium-independent behavior observed in the ruminal enzymes, ARN2 and ARN3. Other molecular aspects were also analyzed including electrostatic surface charge distribution, the presence or absence of the accessory domain, and volume of the cavity. However, these analyses did not indicate a clear correlation of these parameters with calcium dependence. Thus, collectively these structural and functional evidences suggest a global effect, resulting from multiparameter attributes, which select some GH43 ABNs to be activated by calcium or sodium ions.

Aromatic Gateway Controls Solvent Dynamics at the Hydrated Pocket—The molecular complexity regarding the catalytic machinery of GH43 ABNs is not limited to the direct role of His^{314} and the presence of metallic cofactors. The solvent dynamics at the hydrated pocket is also an important issue for the proper functioning of these enzymes. The ion coordination is almost exclusively formed by water molecules, thus it might be labile to solvent perturbations in the hydrated pocket. To investigate this feature of the ion-binding pocket, the fully conserved aromatic residue located at the entrance of the hydrated pocket (Tyr^{226} in TpABN (Fig. 4*E*) and Phe^{237} in ARN3) was mutated to alanine causing a significant reduction of the catalytic activity in both enzymes (Fig. 4*B*). In ARN3, the mutation completely abolished the catalytic activity, whereas in TpABN the activity was reduced up to ~60% in the presence of calcium ions (Fig. 4*B*). This residual activity observed in TpABN can be

interpreted as a result of the stronger coordination of calcium ions in comparison to sodium ions, being less labile to solvent perturbations caused by deletion of the aromatic residue. Therefore, these results suggest that the conserved aromatic residue acts as a gateway controlling the solvent exchange rate at this pocket and, independent of calcium or sodium ions coordinating the catalytically relevant histidine, the solvent dynamics at this pocket were shown to be relevant for catalysis.

An Unusual Catalytic Interface Accommodates Branched Substrates—The crystallographic structure of ARN2, a ruminal endo-acting two-domain enzyme, showed a classical overall fold of the GH43 family (Fig. 1*B*) with all the catalytically relevant residues (10) conserved including the three acidic residues involved in the catalytic reaction (base, Asp^{126}, acid, Glu^{376}, and pK_a modulator, Asp^{308}) and the ion-polarizable histidine (His^{125}). ARN2 contains four main amino acid segment insertions in its sequence compared with other two-domain ABNs such as TpABN, YxiA from *B. licheniformis,* and endo-arabinanase 2 from *Bacillus subtilis* (BsArb43B). Interestingly, three are located at the opposite side of the accessory domain (Glu^{178}–Ser^{198}, Asn^{270}–Arg^{302}, and Val^{346}–Tyr^{362}) forming an unexpected motif (Fig. 1*C*), which along with the Asn^{237}–Ser^{243} loop, generates a unique catalytic interface (Fig. 1*D*). These structural differences result in a larger catalytic interface with higher solvent accessibility in comparison to other structurally characterized GH43 ABNs, supporting the fact that ARN2 efficiently degrades branched arabinans (22), the natural substrate of these enzymes. Indeed, the majority of ABNs characterized to date display tight specificity for linear arabinan chains, exhibiting very low activity against branched forms of the polysaccharide.

Structural Basis for the Exo-operating Mode of ARN3—The exo-acting GH43 ABNs, CjArb43A and ARN3, attack the same substrate, arabinan; but they release different products, which are arabinotriose for CjArb43A (8) and arabinose for ARN3 (18). The exo-activity of CjArb43A is associated with strong affinity for the free O5 atom of the non-reducing terminal arabinose at the −3 subsite (8). The removal of two residues that make hydrogen bonds with the hydroxyl group at the C5 position in the ligand-complexed structure (Asp^{35} and Gln^{316}) induced the enzyme to degrade in an endo mode (8). The residues forming the −3 subsite of ARN3 are Asp^{37} and Ser^{329}, similar to Asp^{35} and Gln^{316} of CjArb43A. Therefore, to test if those residues also contribute to the exo action mode of ARN3, three mutants were produced, D37W, S329P, and D37W/S329P. However, all of them presented a similar cleavage pattern to that observed for the WT protein, releasing arabinose from arabinooligosaccharides (results not shown). If ARN3 releases arabinose in a similar manner as CjArb43A, there would be unique residues recognizing the chain end after the −1 subsite, but it is not the case because all residues near the −2 subsite (His^{39}, Thr^{55}, and Trp^{96} in ARN3) are fully conserved in endo-acting ABNs (Fig. 5*A*). SaAraf43A, an ABF from *Streptomyces avermitilis* (50), also releases arabinose as ARN3; however, ARN3 cannot be considered an ABF because it does not cleave pNP-Araf (18). Moreover, the exo-acting manner of SaAraf43A was attributed to the strict pocket architecture of



Enzyme	V_{max} (U ml^{-1})	K_m (mg ml^{-1})	k_{cat} (s^{-1})	k_{cat}/K_m (min^{-1} mg^{-1} ml)
ARN3	15.91 (± 0.36)	8.38 (± 0.36)	889.82	106.18
SRGEEP-ARN3	20.99 (±0.85)	6.58 (± 0.53)	1,114.12	169.58

FIGURE 5. **The exo-acting mechanism of the ruminal single-domain ABN (ARN3).** *A*, catalytic interface of ARN3 (*white*) showing the residues that form the −2 subsite of the exo-enzyme ARN3 (*green*) and the endo-enzyme AbnB (*magenta*), the modeled oligosaccharide in the ARN3 structure (*yellow ball* and *sticks*), and the Arg^{203}–Ala^{230} loop (*green* surface representation). *B*, topographical adaptations of the SRGEEP-ARN3 active site. The shortened loop is represented as a *transparent magenta* surface and the substrate from the AbnB complex structure is in *magenta ball* and *sticks*. For comparison, the original loop of ARN3 is shown as a *green line*. *C*, kinetic analysis of ARN3 and SRGEEP-ARN3 using debranched arabinan as substrate. Capillary zone electrophoresis analysis of cleavage products of linear arabinan and arabinoheptaose by SRGEEP-ARN3 (*D* and *F*, respectively) and ARN3 (*E* and *G*, respectively).

the −1 subsite (50), whereas ARN3 has a substrate-binding channel with additional −2 and −3 subsites (Fig. 4*A*).

The crystal structure of ARN3 revealed an intriguing structural feature at the aglycone-binding region that is a long insertion in the loop connecting the second and third blades of the β-propeller (Fig. 5*A*). This loop typically comprises 4 or 6 residues in other GH43 ABNs (Fig. 5*B*); however, in ARN3, it is 28 residues long (Arg^{203}–Ala^{230}). It is partially disordered (Arg^{207}–Asp^{227}) in the crystal structure of ARN3, but the proximity with positive subsites suggest a role as a physical barrier in the substrate-binding channel, implying that the exo-activity of ARN3 is related to the reducing end of the arabinan chain. All known exo-acting GH43 enzymes attack the non-reducing end of the polysaccharide (8, 12, 13, 50–52) and few exo-enzymes

that attack the reducing end have been reported to date. For instance, the exo-oligoxylanase (EC 3.2.1.156) from *Bacillus halodurans* C-125 (Rex) is a GH8 enzyme that cleaves xylo-oligosaccharides to release xylose from the reducing end (53). The exo-activity of this enzyme is determined by a physical barrier and chemical recognition. The +2 subsite is blocked by a kink in the loop before the α10 helix and His^{319}, at the tip of this loop, forms a direct hydrogen bond with the β-hydroxyl group of the xylose at the +1 subsite, the reducing end of the polysaccharide (54). Molecular dynamics simulations of ARN3 in the native state and complexed with substrate showed no structural changes in this loop, indicating a naturally occluded aglycone region of the substrate binding channel. Moreover, in the substrate-complex model, the Asp^{227} side chain is hydrogen bonded to the O1 atom from the hydroxyl group of arabinose (Fig. 5*A*), which limits the substrate channel to the +1 subsite and explains the cleavage pattern with arabinose as final product. Based on these observations, we propose that elongation of this loop results in topographical adaptations of the catalytic interface, which limits the channel to the +1 subsite and determines the exo-mode of action of this novel enzyme. This ability to attack reducing ends of the substrate within a GH43 member is of extreme relevance for outright chemical recognition and synergism in arabinan degradation. To date, all characterized GH43 exo-acting enzymes cleave from the non-reducing end of the polysaccharide and the use of both exo-enzymes that preferentially attack either the reducing or the non-reducing end is an important synergism observed for exocellulases (55).

Redesigning the Action Mode of ARN3—According to our structural analysis, the Arg^{203}–Ala^{230} loop is related to the action mode of the ruminal single-domain enzyme (ARN3). To test this hypothesis, we redesigned the catalytic interface of the protein, replacing this long loop insertion by the amino acid segment corresponding to classical endo-acting ABNs (sequence "SRGEEP"). The chimera SRGEEP-ARN3, heterologously produced, was fully functional and presented a native-like structure. The chimera preserved the temperature and pH optimum (results not shown), but increased the catalytic efficiency in more than 50% (Fig. 5*C*). The loop replacement exposed additional subsites of the substrate binding channel, increasing the affinity for the substrate, as observed by the lower K_m of the chimera (Fig. 5*C*). Moreover, extraction of the physical barrier represented by the loop favored the turnover of the substrate, indicated by the higher k_{cat}. The main result of the loop exchange is the modification in the cleavage pattern. SRGEEP-ARN3 generates principally arabinobiose and arabinotriose as the product from linear arabinan (Fig. 5*D*), whereas cleavage by the WT enzyme produces exclusively arabinose (Fig. 5*E*). Arabinoheptaose was also used as substrate for the WT enzyme and the SRGEEP-ARN3 chimera. As expected, ARN3 successively removed arabinose from the oligosaccharide, sequentially producing hexaose, pentaose, tetraose, triose, and biose (Fig. 5*F*). The chimera also produced a range of oligosaccharides (Fig. 5*G*), but the ratio of the products formed at the beginning of the reaction indicates that the enzyme randomly cleaved the oligosaccharide. Moreover, differently from the WT enzyme, di-, tri-, and tetrasaccharides instead of mono- and disaccharides, accumulated in late stages of the reaction with SRGEEP-ARN3 (Fig. 5, *F* and *G*, respectively). Therefore, the replacement of the Arg^{203}–Ala^{230} region by a short segment in the chimera SRGEEP-ARN3 clearly led to a change in the enzyme cleavage pattern, indicating that this long loop is responsible for the exo-activity of this ruminal single-domain enzyme.

DISCUSSION

A Mechanistic Model for GH43 ABN Activation—GH43 ABNs operate by a reaction mechanism with inversion of anomeric configuration using the canonical three acidic residues, Asp^{32} (the base), Glu^{233} (the acid), and Asp^{170} (the pK_a modulator) (Fig. 4*F*). Asp^{32} is deprotonated to activate the catalytic water that promotes a single-displacement attack on the anomeric carbon (56) and Glu^{223} protonates the glycosidic oxygen making the aglycone a better leaving group. For that, the Glu^{223} pK_a is maintained higher than usual by Asp^{170} (12). However, in contrast to classical glycoside hydrolases, ABNs contain a fourth key residue for catalysis, an interfacing histidine (His^{314} in TpABN) that mediates the interaction with the metallic cofactor (Fig. 4*F*). This residue is strategically located between the two catalytic residues (3.1 and 5.1 Å from Asp^{32} and Glu^{223}, respectively) and interacts via its Nϵ2 atom with calcium ion (or sodium ion) located at the hydrated pocket. The rotameric configuration of the histidine ring is maintained perpendicular to the catalytic acid residues by interaction with the ion, preventing a specific interaction with these residues that could disrupt the geometry of the active site. This network of interactions and the fact that histidine mutation abolishes the catalytic activity of GH43 enzymes (11) support the crucial role of this residue in the catalysis. Moreover, the replacement of His by Gln, which should preserve the ability to make hydrogen bonds with the substrate, abolished the enzymatic activity (11), indicating a particular correlation between the metallic cofactor and histidine. These observations suggest that the positively charged ion might have a hyper-polarizing effect on the imidazole side chain of His^{314} via the Nϵ2 atom resulting in an electrophilic character of the His^{314}-Nδ1 atom. Consequently, it might influence the charge distribution in the active site microenvironment and the protonation states of catalytic acidic residues, contributing to the functional state of catalytic residues. This hyperpolarizing effect also explains the fact that the His $\rightarrow$ Gln mutation terminated the catalytic activity of BsArb43B (11), because the amide side chain of Gln is not polarizable by metallic ions. Therefore, these findings indicate that the ion-induced hyperpolarization of histidine is essential to maintain an intact active site configuration for catalysis.

New Routes for Arabinan Degradation—This study revealed the molecular bases of the unique functional properties for outright chemical recognition and synergism in arabinan degradation of the ruminal ABNs, ARN2, and ARN3. ARN2 efficiently degrades naturally abundant branched substrates due to profound topographical adaptations in the active site interface conferring higher accessibility, whereas all known orthologs have preference for linear substrates. ARN3 is the first GH43 member with the ability to attack the reducing ends of the substrate. This rare exo-acting operation mode is achieved by imposing a physical barrier to the aglycone subsites, as con-

firmed by structure-guided redesign of the active site, which converted the ruminal exo-enzyme into a classical endo-acting glycosidase. The combined use of exo-enzymes that preferentially attack either the reducing or the non-reducing ends of the substrate is an important synergism already observed for exo-cellulases and now also viable for breakdown of hemicellulosic polysaccharides. Both enzymes with their unique action modes provide new routes for arabinan degradation that can impact on technologies for conversion of plant cell-wall polysaccharides into energy and chemicals.

Acknowledgments—We thank Laboratório Nacional de Luz Síncrotron and Laboratório Nacional de Biociências for the provision of time on the MX2 beamline, and both crystallization (Robolab) and spectroscopy (LEC) facilities.

REFERENCES

1. Osaki, S., Kimura, T., Sugimoto, T., Hizukuri, S., and Iritani, N. (2001) L-Arabinose feeding prevents increases due to dietary sucrose in lipogenic enzymes and triacylglycerol levels in rats. *J. Nutr.* **131,** 796–799
2. Al-Tamimi, M. A., Palframan, R. J., Cooper, J. M., Gibson, G. R., and Rastall, R. A. (2006) *In vitro* fermentation of sugar beet arabinan and arabino-oligosaccharides by the human gut microflora. *J. Appl. Microbiol.* **100,** 407–414
3. Jorgensen, H., Kristensen, J. B., and Felby, C. (2007) Enzymatic conversion of lignocellulose into fermentable sugars. Challenges and opportunities. *Biofuels Bioprod. Bioref.* **1,** 119–134
4. Berlin, A., Maximenko, V., Gilkes, N., and Saddler, J. (2007) Optimization of enzyme complexes for lignocellulose hydrolysis. *Biotechnol. Bioeng.* **97,** 287–296
5. Gao, D., Uppugundla, N., Chundawat, S. P., Yu, X., Hermanson, S., Gowda, K., Brumm, P., Mead, D., Balan, V., and Dale, B. E. (2011) Hemicellulases and auxiliary enzymes for improved conversion of lignocellulosic biomass to monosaccharides. *Biotechnol. Biofuels* **4,** 5
6. Delabona, P. da S., Cota, J., Hoffmam, Z. B., Paixão, D. A., Farinas, C. S., Cairo, J. P., Lima, D. J., Squina, F. M., Ruller, R., and Pradella, J. G. (2013) Understanding the cellulolytic system of *Trichoderma harzianum* P49P11 and enhancing saccharification of pretreated sugarcane bagasse by supplementation with pectinase and α-L-arabinofuranosidase. *Bioresour. Technol.* **131,** 500–507
7. Nurizzo, D., Turkenburg, J. P., Charnock, S. J., Roberts, S. M., Dodson, E. J., McKie, V. A., Taylor, E. J., Gilbert, H. J., and Davies, G. J. (2002) *Cellvibrio japonicus* α-L-arabinanase 43A has a novel five-blade β-propeller fold. *Nat. Struct. Biol.* **9,** 665–668
8. Proctor, M. R., Taylor, E. J., Nurizzo, D., Turkenburg, J. P., Lloyd, R. M., Vardakou, M., Davies, G. J., and Gilbert, H. J. (2005) Tailored catalysts for plant cell-wall degradation. Redesigning the exo/endo preference of *Cellvibrio japonicus* arabinanase 43A. *Proc. Natl. Acad. Sci. U.S.A.* **102,** 2697–2702
9. Yamaguchi, A., Tada, T., Wada, K., Nakaniwa, T., Kitatani, T., Sogabe, Y., Takao, M., Sakai, T., and Nishimura, K. (2005) Structural basis for thermostability of endo-1,5-α-L-arabinanase from *Bacillus thermodenitrificans* TS-3. *J. Biochem.* **137,** 587–592
10. Alhassid, A., Ben-David, A., Tabachnikov, O., Libster, D., Naveh, E., Zolotnitsky, G., Shoham, Y., and Shoham, G. (2009) Crystal structure of an inverting GH43 1,5-α-L-arabinanase from *Geobacillus stearothermophilus* complexed with its substrate. *Biochem. J.* **422,** 73–82
11. de Sanctis, D., Inácio, J. M., Lindley, P. F., de Sá-Nogueira, I., and Bento, I. (2010) New evidence for the role of calcium in the glycosidase reaction of GH43 arabinanases. *FEBS J.* **277,** 4562–4574
12. Brüx, C., Ben-David, A., Shallom-Shezifi, D., Leon, M., Niefind, K., Shoham, G., Shoham, Y., and Schomburg, D. (2006) The structure of an inverting GH43 β-xylosidase from *Geobacillus stearothermophilus* with its substrate reveals the role of the three catalytic residues. *J. Mol. Biol.* **359,** 97–109
13. Brunzelle, J. S., Jordan, D. B., McCaslin, D. R., Olczak, A., and Wawrzak, Z. (2008) Structure of the two-subsite β-d-xylosidase from *Selenomonas ruminantium* in complex with 1,3-bis[tris(hydroxymethyl)methylamino]propane. *Arch. Biochem. Biophys.* **474,** 157–166
14. McKee, L. S., Peña, M. J., Rogowski, A., Jackson, A., Lewis, R.J., York, W. S., Krogh, K. B., Viksø-Nielsen, A., Skjøt, M., Gilbert, H. J., and Marles-Wright, J. (2012) Introducing endo-xylanase activity into an exo-acting arabinofuranosidase that targets side chains. *Proc. Natl. Acad. Sci. U.S.A.* **109,** 6537–6542
15. Boisset, C., Pétrequin, C., Chanzy, H., Henrissat, B., and Schülein, M. (2001) Optimized mixtures of recombinant *Humicola insolens* cellulases for the biodegradation of crystalline cellulose. *Biotechnol. Bioeng.* **72,** 339–345
16. Srisodsuk, M., Kleman-Leyer, K., Keränen, S., Kirk, T. K., and Teeri, T. T. (1998) Modes of action on cotton and bacterial cellulose of a homologous endoglucanase-exoglucanase pair from *Trichoderma reesei*. *Eur. J. Biochem.* **251,** 885–892
17. Wong, D. (2008) Enzymatic deconstruction of backbone structures of the ramified regions in pectins. *Protein J.* **27,** 30–42
18. Wong, D. W., Chan, V. J., and Batt, S. B. (2008) Cloning and characterization of a novel exo-α-1,5-L-arabinanase gene and the enzyme. *Appl. Microbiol. Biotechnol.* **79,** 941–949
19. McKie, V. A., Black, G. W., Millward-Sadler, S. J., Hazlewood, G. P., Laurie, J. I., and Gilbert, H. J. (1997) Arabinanase A from *Pseudomonas fluorescens* subsp. *cellulosa* exhibits both an endo- and an exo-mode of action. *Biochem. J.* **323,** 547–555
20. Takao, M., Akiyama, K., and Sakai, T. (2002) Purification and characterization of thermostable endo-1,5-α-L-arabinase from a strain of *Bacillus thermodenitrificans*. *Appl. Environ. Microbiol.* **68,** 1639–1646
21. Inácio, J. M., and de Sá-Nogueira, I. (2008) Characterization of abn2 (yxiA), encoding a *Bacillus subtilis* GH43 arabinanase, Abn2, and its role in arabino-polysaccharide degradation. *J. Bacteriol.* **190,** 4272–4280
22. Wong, D. W., Chan, V. J., and McCormack, A. A. (2009) Functional cloning and expression of a novel endo-α-1,5-L-arabinanase from a metagenomic library. *Protein Pept. Lett.* **16,** 1435–1441
23. Squina, F. M., Prade, R. A., Wang, H., and Murakami, M. T. (2009) Expression, purification, crystallization and preliminary crystallographic analysis of an endo-1,5-α-L-arabinanase from hyperthermophilic *Thermotoga petrophila*. *Acta Crystallogr. Sect. F Struct. Biol. Cryst. Commun.* **65,** 902–905
24. Squina, F. M., Santos, C. R., Ribeiro, D. A., Cota, J., de Oliveira, R. R., Ruller, R., Mort, A., Murakami, M. T., and Prade, R. A. (2010) Substrate cleavage pattern, biophysical characterization and low-resolution structure of a novel hyperthermostable arabinanase from *Thermotoga petrophila*. *Biochem. Biophys. Res. Commun.* **399,** 505–511
25. Erster, O., and Liscovitch, M. (2010) A modified inverse PCR procedure for insertion, deletion, or replacement of a DNA fragment in a target sequence and its application in the ligand interaction scan method for generation of ligand-regulated proteins. *Methods Mol. Biol.* **634,** 157–174
26. Hemsley, A., Arnheim, N., Toney, M. D., Cortopassi, G., and Galas, D. J. (1989) A simple method for site-directed mutagenesis using the polymerase chain reaction. *Nucleic Acids Res.* **17,** 6545–6551
27. Otwinowski, Z., and Minor, W. (1997) Processing of x-ray diffraction data collected in oscillation mode. *Methods Enzymol.* **276,** 307–326
28. Kabsch, W. (2010) *XDS*. *Acta Crystallogr. D Biol. Crystallogr.* **66,** 125–132
29. Vagin, A., and Teplyakov, A. (2010) Molecular replacement with MOLREP. *Acta Crystallogr. D Biol. Crystallogr.* **66,** 22–25
30. Pletnev, S., Morozova, K. S., Verkhusha, V. V., and Dauter, Z. (2009) Rotational order-disorder structure of fluorescent protein FP480. *Acta Crystallogr. D Biol. Crystallogr.* **65,** 906–912
31. Wang, J., Kamtekar, S., Berman, A. J., and Steitz, T. A. (2005) Correction of x-ray intensities from single crystals containing lattice-translocation defects. *Acta Crystallogr. D Biol. Crystallogr.* **61,** 67–74
32. Murshudov, G. N., Vagin, A. A., and Dodson, E. J. (1997) Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr. D Biol. Crystallogr.* **53,** 240–255
33. Emsley, P., Lohkamp, B., Scott, W. G., and Cowtan, K. (2010) Features and

development of Coot. *Acta Crystallogr. D Biol. Crystallogr.* **66,** 486–501
34. Chen, V. B., Arendall, W. B., 3rd, Headd, J. J., Keedy, D. A., Immormino, R. M., Kapral, G. J., Murray, L. W., Richardson, J. S., and Richardson, D. C. (2010) MolProbity. All-atom structure validation for macromolecular crystallography. *Acta Crystallogr. D Biol. Crystallogr.* **66,** 12–21
35. Ke, A., and Wolberger, C. (2003) Insights into binding cooperativity of MATα1/MATα2 from the crystal structure of a MATα1 homeodomain-maltose binding protein chimera. *Protein Sci.* **12**, 306–312
36. Kobe, B., Center, R. J., Kemp, B. E., and Poumbourios, P. (1999) Crystal structure of human T cell leukemia virus type 1 gp21 ectodomain crystallized as a maltose-binding protein chimera reveals structural evolution of retroviral transmembrane proteins. *Proc. Natl. Acad. Sci. U.S.A.* **96,** 4319–4324
37. Center, R. J., Kobe, B., Wilson, K. A., Teh, T., Howlett, G. J., Kemp, B. E., and Poumbourios, P. (1998) Crystallization of a trimeric human T cell leukemia virus type 1 gp21 ectodomain fragment as a chimera with maltose-binding protein. *Protein Sci.* **7,** 1612–1619
38. Ware, S., Donahue, J. P., Hawiger, J., and Anderson, W. F. (1999) Structure of the fibrinogen γ-chain integrin binding and factor XIIIa cross-linking sites obtained through carrier protein driven crystallization. *Protein Sci.* **8,** 2663–2671
39. Kuge, M., Fujii, Y., Shimizu, T., Hirose, F., Matsukage, A., and Hakoshima, T. (1997) Use of a fusion protein to obtain crystals suitable for X-ray analysis. Crystallization of a GST-fused protein containing the DNA-binding domain of DNA replication-related element-binding factor, DREF. *Protein Sci.* **6,** 1783–1786
40. Rosenbaum, D. M., Cherezov, V., Hanson, M. A., Rasmussen, S. G., Thian, F. S., Kobilka, T. S., Choi, H. J., Yao, X. J., Weis, W. I., Stevens, R. C., and Kobilka, B. K. (2007) GPCR engineering yields high-resolution structural insights into β2-adrenergic receptor function. *Science* **318,** 1266–1273
41. Zou, Y., Weis, W. I., and Kobilka, B. K. (2012) N-terminal t4 lysozyme fusion facilitates crystallization of a G protein coupled receptor. *PLoS One* **7,** e46039
42. Corsini, L., Hothorn, M., Scheffzek, K., Sattler, M., and Stier, G. (2008) Thioredoxin as a fusion tag for carrier-driven crystallization. *Protein Sci.* **17,** 2070–2079
43. Stoll, V. S., Manohar, A. V., Gillon, W., MacFarlane, E.L., Hynes, R. C., and Pai, E. F. (1998) A thioredoxin fusion protein of VanH, a D-lactate dehydrogenase from *Enterococcus faecium*. Cloning, expression, purification, kinetic analysis, and crystallization. *Protein Sci.* **7,** 1147–1155
44. Miller, G. L. (1959) Use of dinitrosalicylic acid reagent for determination of reducing sugar. *Anal. Biochem.* **31,** 426–428
45. Chen, F. T., and Evangelista, R. A. (1995) Analysis of mono- and oligosaccharide isomers derivatized with 9-aminopyrene-1,4,6-trisulfonate by capillary electrophoresis with laser-induced fluorescence. *Anal. Biochem.* **230,** 273–280
46. Krieger, E., Koraimann, G., and Vriend, G. (2002) Increasing the precision of comparative models with YASARA NOVA. A self-parameterizing force field. *Proteins* **47,** 393–402
47. Krieger, E., Nielsen, J. E., Spronk, C. A., and Vriend, G. (2006) Fast empirical pK_a prediction by Ewald summation. *J. Mol. Graph. Model.* **25,** 481–486
48. Woody, R. W. (1994) Contributions of tryptophan side chains to the far-ultraviolet circular dichroism of proteins. *Eur. Biophys. J.* **23,** 253–262
49. Tina, K. G., Bhadra, R., and Srinivasan, N. (2007) PIC. Protein interactions calculator. *Nucleic Acids Res.* **35,** W473–W476
50. Fujimoto, Z., Ichinose, H., Maehara, T., Honda, M., Kitaoka, M., and Kaneko, S. (2010) Crystal structure of an Exo-1,5-α-L-arabinofuranosidase from *Streptomyces avermitilis* provides insights into the mechanism of substrate discrimination between exo- and endo-type enzymes in glycoside hydrolase family 43. *J. Biol. Chem.* **285,** 34134–34143
51. Jiang, D., Fan, J., Wang, X., Zhao, Y., Huang, B., Liu, J., and Zhang, X. C. (2012) Crystal structure of 1,3Gal43A, an exo-β-1,3-galactanase from *Clostridium thermocellum*. *J. Struct. Biol.* **180**, 447–457
52. Ichinose, H., Kotake, T., Tsumuraya, Y., and Kaneko, S. (2006) Characterization of an exo-β-1,3-D-galactanase from *Streptomyces avermitilis* NBRC14893 acting on arabinogalactan-proteins. *Biosci. Biotechnol. Biochem.* **70,** 2745–2750
53. Honda, Y., and Kitaoka, M. (2004) A family 8 glycoside hydrolase from *Bacillus halodurans* C-125 (BH2105) is a reducing end xylose-releasing exo-oligoxylanase. *J. Biol. Chem.* **279**, 55097–55103
54. Fushinobu, S., Hidaka, M., Honda, Y., Wakagi, T., Shoun, H., and Kitaoka, M. (2005) Structural basis for the specificity of the reducing end xylose-releasing exo-oligoxylanase from *Bacillus halodurans* C-125. *J. Biol. Chem.* **280,** 17180–17186
55. Barr, B. K., Hsieh, Y. L., Ganem, B., and Wilson, D. B. (1996) Identification of two functionally different classes of exocellulases. *Biochemistry* **35,** 586–592
56. Sinnott, M. L. (1990) Catalytic mechanisms of enzymatic glycosyl transfer. *Chem. Rev.* **90**, 1171–1202