The structure of a *Streptomyces avermitilis* α-L-rhamnosidase reveals a novel carbohydrate-binding module CBM67 within the six-domain arrangement

Zui Fujimoto‡1,3, Adam Jackson§1, Mari Michikawa¶, Tomoko Maehara¶, Mitsuru Momma‡, Bernard Henrissat‖‖, Harry J. Gilbert§4, and Satoshi Kaneko¶2

From ‡Biomolecular Research Unit, National Institute of Agrobiological Sciences, 2-1-2 Kannondai, Tsukuba 305-8602, Japan, §Institute for Cell and Molecular Biosciences, Newcastle University, Medical School, Newcastle upon Tyne NE2 4HH, United Kingdom ¶Food Biotechnology Division, National Food Research Institute, 2-1-12 Kannondai, Tsukuba, Ibaraki 305-8642, Japan, and ‖‖Lab Architecture et Fonction des Macromolécules Biologiques, Aix-Marseille Université, CNRS, UMR 7257, Case 932, 163 Avenue de Luminy, 13288 Marseille cedex 9, France

Z.F, and A.J. contributed equally to this work.

Running title: Crystal structure of *Streptomyces avermitilis* α-L-rhamnosidase

Address correspondence to : Satoshi Kaneko, National Food Research Institute, 2-1-12 Kannondai, Tsukuba, Ibaraki 305-8642, Japan; Phone. +81-29-838-8063; Fax+81-29-838-7996; E-mail: sakaneko@affrc.go.jp, or Zui Fujimoto, National Institute of Agrobiological Sciences, 2-1-2 Kannondai, Tsukuba, Ibaraki 305-8602, Japan; Phone and Fax. +81-29-838-7877; E-mail: zui@affrc.go.jp or Harry J. Gilbert, Newcastle University, Medical School, Newcastle upon Tyne NE2 4HH, United Kingdom. Phone. +44-(0)191-2228800; E-mail: harry.gilbert@newcastle.ac.uk.

**Background:** α-L-Rhamnosidases hydrolyze α-linked L-rhamnose from rhamnoglycosides or polysaccharides.

**Result:** The crystal structure of *Streptomyces avermitilis* α-L-rhamnosidase belonging to glycoside hydrolase family 78 was determined.

**Conclusion:** The L-rhamnose complexed structure revealed the catalytic mechanism of the enzyme and a calcium-dependent carbohydrate-binding module.

**Significance:** Efficient catalysis of an exo-rhamnosidase requires a novel carbohydrate-binding module that binds terminal rhamnose sugars.

α-L-Rhamnosidases hydrolyze α-linked L-rhamnoses from oligosaccharides or polysaccharides. We determined the crystal structure of the glycoside hydrolase family 78 *Streptomyces avermitilis* α-L-rhamnosidase (SaRha78A) in its free and L-rhamnose complexed forms, which revealed the presence of six domains N, D, E, F A and C. In the ligand complex, L-rhamnose was bound in the proposed active site of the catalytic module, revealing the likely catalytic mechanism of SaRha78A. Glu-636 is predicted to donate protons to the glycosidic oxygen, and Glu-895 is the likely catalytic general base, activating the nucleophilic water, indicating that the enzyme operates through an inverting mechanism. Replacement of Glu-636 and Glu-895 resulted in significant loss of α-rhamnosidase activity. Domain D also bound L-rhamnose in a calcium-dependent manner, with a $K_D$ of 135 μM. Domain D is thus a non-catalytic carbohydrate binding module (designated SaCBM67).

Mutagenesis and structural data identified the amino acids in SaCBM67 that target the features of L-rhamnose that distinguishes it from the other major sugars present in plant cell walls. Inactivation of SaCBM67 caused a substantial reduction in the activity of SaRha78A against the polysaccharide composite gum arabic, but not against aryl rhamnosides, indicating that SaCBM67 contributes to enzyme function against insoluble substrates.

L-Rhamnose is widely distributed in plants and bacteria as a component of polysaccharides, such as rhamnogalacturonan and arabinogalactan-protein (1, 2), while the sugar is also found in some glycol conjugates such as naringin and rutin (3, 4). Naringin, a major flavonoid in grapefruit, contains α-1,2 linked L-rhamnose. Rutin is a citrus flavonoid found in buckwheat, and it contains α-1,6 linked L-rhamnose. α-L-Rhamnosidases (EC
3.2.1.40) catalyze the hydrolysis of α-L-rhamnosyl-linkages in L-rhamnose containing compounds. According to the CAZy database (5, 6); http://www.cazy.org/), currently, α-L-rhamnosidases are classified into three glycoside hydrolase families (GHs); GH28, GH78, and GH106. To date, only two crystal structures of α-L-rhamnosidases have been determined; α-L-rhamnosidase B (BsRhaB) from Bacillus sp. GL1 (7) and the putative α-L-rhamnosidase BT1001 from Bacteroides thetaiotaomicron VPI-5482 (8). Both enzymes belong to GH78. BsRhaB is composed of five distinct domains and BT1001 contains four domains. The catalytic module of GH78 enzymes is an (α/α)6-barrel, but the detailed catalytic mechanism of GH78 enzymes is unclear, and the role of the surrounding domains are unknown.

In a previous study, we characterized a GH78 α-L-rhamnosidase from Streptomyces avermitilis (SaRha78A), showing that the enzyme hydrolyzed aryl rhamnosides and rhamnose-containing polysaccharides (9). In this article, we provide the crystal structure of SaRha78A in apo form and in complex with L-rhamnose. The data revealed the catalytic mechanism of the enzyme and identified a novel non-catalytic carbohydrate-binding module (CBM), SaCBM67, the founding member of CBM67. Analysis of the ligand specificity of SaCBM67 showed that the protein module bound L-rhamnose through a calcium-dependent mechanism in a short binding cleft. Inactivation of SaCBM67 through mutagenesis or chelation of calcium, showed that the CBM made a substantial contribution to the activity of SaRha78A against polysaccharides, but not against aryl rhamnosides.

**EXPERIMENTAL PROCEDURES**

**Crystallization of SaRha78A** – Recombinant SaRha78A was expressed in Escherichia coli and purified by histidine tag affinity chromatography as described by Ichinose et al. (9). The purified protein solution in 2 mM Tris-HCl buffer, pH 7.0, containing 20 mM NaCl, was concentrated to 10 mg ml⁻¹ (A₂₈₀, mm=2.2 U) by ultrafiltration using a YM-30 membrane (Millipore, Billerica, MA, USA), and filtered through a 0.1 µm membrane. The protein was crystallized by the sitting-drop vapour diffusion method using a precipitant solution composed of 10 % (w/v) PEG6000, 5 % (v/v) MPD and 0.1 M HEPES pH 6.8. Plate-shaped crystals with the dimensions 500 × 50 × 10 mm were consistently obtained using 50 µl of the reservoir solution with a drop consisting of 1 µl protein solution and 1 µl reservoir solution. Selenomethionine (SeMet)-labeled SaRha78A was expressed in LeMaster medium using the methionine-auxotrophic strain B834(DE3) (10) and crystallized in the same condition used for the native enzyme.

**Data collection and structure determination** – Diffraction experiments for native crystals were conducted at the beamline BL41XU of the synchrotron facility SPring-8, Hyogo, Japan. Crystals were soaked into the reservoir solution containing 13% (v/v) glycerol, scooped in a 0.5-mm nylon loop (Hampton Research, Aliso Viejo, CA, USA) and flash-frozen in a nitrogen stream at 95 K. Diffraction data were collected at a wavelength of 0.97915 Å with a Quantum 210 CCD detector (Area Detector Systems Corp., Poway, CA, USA). Diffraction experiments for the SeMet-substituted crystals were conducted at the beamline BL-5 of the Photon Factory, High Energy Accelerator Research Organization, Tsukuba, Japan. Diffraction data were collected at a wavelength of 0.97909 Å with a Quantum 315 CCD detector (ADSC). For structural analyses of the enzyme in complex with L-rhamnose (Wako Pure Chemical Industry), SaRha78A crystals were soaked in to a drop containing 1% (w/v) L-rhamnose in the precipitant solution for 10 min before the diffraction experiment. Diffraction data of the L-rhamnose complex to 1.9 Å resolution were collected at the beamline BL-NW12 of the Photon Factory Advanced Ring. Diffraction data were collected at a wavelength of 1.0000 Å with a Quantum 210 CCD detector (ADSC). All data were integrated and scaled...
using the program DENZO and SCALEPACK in the HKL2000 program suite (11).

The crystal structure was determined by the single-wavelength anomalous dispersion method using SeMet-labeled crystals. In total 12 selenium atom positions were determined, and initial phases were calculated using the program SOLVE/RESOLVE (12, 13). The initial model building was conducted by the auto-modeling program ARP/wARP (14) within the CCP4 program suite (15). Manual model building and molecular refinement were performed using Coot (16) and Refmac5 (17, 18).

For the analysis of L-rhamnose–protein complex, structural determination was conducted using the ligand-free structure as the starting model and the bound L-rhamnose was observed in the difference electron density map. Data collection and refinement statistics are given in Table 1. Stereochemistry of the models was analyzed with the program Rampage (19). Structural drawings were prepared by the program PyMol (DeLano Scientific LLC, Palo Alto, CA, USA).

Sugar binding analysis by isothermal titration calorimetry – To produce SaCBM67, the appropriate region of the SaRha78A gene (encoding residues Pro133–Ala296) was amplified using the following primers: forward: 5’-CAT ATG CCT TCC CTG GAG GGT AGT TCG TGG-3’ and reverse, 5’-AAG CTT TCA CGC GAC CCG GCC CCA CGG TCC TGC -3’. The amplified DNA was cloned into pET28 vector (Novagen) at NdeI and HindIII restriction enzyme sites (underlined). The protein was expressed by the same method as SaRha78A, and purified by immobilized metal ion affinity chromatography using the N-terminal histidine tag. The D179A and N180A mutants of SaCBM67 were constructed using the PCR-based QuikChange™ site directed mutagenesis kit (Strategene) of the SaRhaCBM plasmid using the following primers: mutant D179A, 5’-CATGCCATACGCAGCGACGCGTCTAGCGTCTCC-3’ and 5’-AGGGCCGCTAGACCCGCGTGCAGCGTGC-3’; mutant N180A, 5’-GGAGACGCCTAGACCCGCTCGGCCGC TGATGCC-3’ (mutated bases underlined).

Binding of SaCBM67 to its target ligands was determined by isothermal titration calorimetry using a MicroCal VP-ITC. Titrations were carried out at 25 °C in 20 mM Na-HEPES buffer, pH 7.5, containing 2mM CaCl₂, unless otherwise stated. The reaction cell contained protein at 700-800 μM (depending on the ligand), while the syringe contained 20 mM of potential ligands (L-fucose, L-mannose and L-rhamnose purchased from Sigma Aldrich). The data were analyzed using MicroCal Origin version 7.0, which yielded the change in enthalpy (∆H), association constants (Kₐ) and stoichiometry of binding (n), enabling changes in entropy (∆S) and Gibbs free energy (∆G) to be calculated using the equation –RTlnKₐ = ∆G = ∆H – T∆S.

Mutant Generation and Enzyme assays – Site-directed mutagenesis was used to generate the SaRha78A mutants of catalytically important residues. The pET30-SaRha78A plasmid, encoding full length SaRha78A (9), was used as a template and PCR with KOD -plus- neo polymerase (Toyobo, Osaka, Japan) was performed using the following sets of primers: E636D-F, 5’-CGGCCCGTGATGATCGGCTCGGCTGGA-3’, and E636D-R, 5’-CAGCCGAGCCGATCATCACGGGCCGGAGTCCGCTGGA-3’; E636Q-F, 5’-CGGCCCGTGATCAACGGCTCGGCTGGA-3’, and E636Q-R, 5’-CAGCCGAGCCGATCATCAGGGCTGGCTGGA-3’; E895Q-F, 5’-ACCACCATGTGGCAGCGCTGGGACTCCAT-3’, and E895Q-R, 5’-AGTCCCAGCGCTGCCACATGGTGGTGGA-3’; E895D-F, 5’-ACCACCATGTGGGATCGCTGGGACTCCAT-3’, and E895D-R, 5’-AGTCCCAGCGCTGCCACATGGTGGTGGA-3’. The underlined sequences represent mutation sites. Recombinant SaRha78A mutants were expressed in Escherichia coli using the program DENZO and SCALEPACK in the HKL2000 program suite (11).
Tuner (DE3) (Novagen, Merck KGaA, Darmstadt, Germany) and purified, and the kinetic parameters of the wild type and mutants of SaRha78A for p-nitrophenyl-\(\alpha\)-L-rhamnopyranoside (PNP-\(\alpha\)-L-Rha) were determined as described previously (9). The activity of the wild type and mutants of SaRha78A for gum arabic were determined as follows. Gum arabic was treated with 0.05 N H\(_2\)SO\(_4\) at 100°C for 1 h for acid hydrolysis, and subjected to enzyme digestion by \(\alpha\)-L-arabinofuranosidase (20) and \(\beta\)-L-arabinopyranosidase (21), to increase enzyme accessibility to the polysaccharide. After removal of released L-arabinose by dialysis, the partially digested gum arabic was freeze dried and then used as the substrate. The reactions were performed in McIlvaine buffer (pH 5.0) containing 1–5% (w/v) substrates, 0.1% (w/v) bovine serum albumin, 2 mM CaCl\(_2\), and 0.9 nM–10.0 mM enzyme at 40°C for up to 360 min. The amount of L-rhamnose released was quantified by high-performance anion-exchange chromatography with pulsed amperometric detection. The assay was performed in duplicate. The kinetic parameters \(k_{\text{cat}}\) and \(K_m\) were determined by Lineweaver–Burk plot from three independent experiments and at five substrate concentrations.

RESULTS

Overall structure of SaRha78A – The crystal structure of SaRha78A was determined by the single-wavelength anomalous dispersion method using SeMet crystal data. The structure of the apo form of SaRha78A and the protein in complex with L-rhamnose were determined. Structure refinement statistics are summarized in Table 1. The quality and accuracy of the final structures are indicated by the observation that no residues fell within the disallowed region of the Ramachandran stereochemistry plot. The recombinant SaRha78A molecule is composed of a single polypeptide chain of 1043 amino acids, where the C-terminal His-tag, \text{KLAAALEHHHHHHHH}^{1043}, were derived from the expression vector and purification tag, respectively. The N-terminal two residues Met-1 and Ser-2 and the C-terminal 11 residues, Ala-1033 to His-1043, were not identified due to lack of electron density. The final model consisted of one SaRha78A molecule in the asymmetric unit bound to a single calcium ion. The ligand free structure contained three Tris molecules, and the L-rhamnose complex structure contained three sodium ions.

The protein forms a multi-domain structure comprised of six distinct domains, one \(\alpha\)-domain and five \(\beta\)-domains (Fig. 1). They were designated as domain N, domain D (designated SaCBM67 as it fulfills a carbohydrate binding function, see below), domain E, domain F, domain A (designated SaRha78CM as it comprises the catalytic module, see below) and domain C from the N-terminus to the C-terminus. Domain N (Ala-3—Gly-114) consists of 10 \(\beta\)-strands folding into 4- and 3-stranded \(\beta\)-sheets and adopt a fibronectin-type 3 fold. Domain D (SaCBM67 Pro-133—Pro-297) and domain E (Ala-115—Ala-132, Val-298—Gly-426 and Glu-444—Gly478) display a \(\beta\)-jelly roll fold consisting of 11 and 13 \(\beta\)-strands, respectively (SaCBM67 is described in detail below). Domain E provided two \(\beta\)-strands (Pro-427—Lys-443) that were inserted into one of the \(\beta\)-sheets of Domain F (Pro-479—Ala-596 and Pro-427—Lys-443), which also displayed a \(\beta\)-jelly roll fold. Domain F consisted of 16 \(\beta\)-strands organized into two parallel \(\beta\)-sheets.

SaRha78\textsubscript{CM} (Val-606—Ala-931) comprised an (\(\alpha/\alpha\))\textsubscript{6}-barrel, which displays structural similarity to the catalytic module of BsRhaB, the prototypic GH78 rhamnosidase from Bacillus sp. (7), and is thus likely to fulfill the same function in SaRha78A. SaRha78\textsubscript{CM} contains 13 \(\alpha\)-helices and lacks any \(\beta\)-structures. Domain C (Gly-932—Val-1030 and Pro-597—Asn-605) adopt a \(\beta\)-sandwich structure in which the two \(\beta\)-sheets consist of 5- and 6- anti-parallel \(\beta\)-strands, respectively.

Domains N, E, F and C are located around SaRha78\textsubscript{CM}. Domain F and domain C interact with SaRha78\textsubscript{CM} through large interfaces, while domain N and domain E also abut onto the catalytic module, but
through a small contact area. SaCBM67 protrudes from domain E through two linker peptides, and makes no contact with the other domains. The domain arrangement of SaRha78A is partially different to the other two GH78 proteins, BsRhaB and BT1001, for which crystal structures are available. When the three proteins were superimposed it is evident that a domain equivalent to domain N is not present in BT1001 or BsRhaB, while SaRha78 CM domain F and domain C are common to all three proteins and superimposed with an r.m.s.d. of 2.3 Å (SaRha78 CM), 2.1 Å (domain F) and 1.5 Å (domain C) between SaRha78A and BT1001 and 2.0 (SaRha78 CM), 2.0 Å (domain F) and 1.3 Å (domain C) between SaRha78A and BsRhaB (Fig. 2). Domain E also superimposed with an equivalent domain in BsRhaB with an r.m.s.d. of 1.8 Å, but the equivalent domain is displaced in BT1001. SaCBM67 were essentially in a different position in SaRha78A and BsRhaB, but was absent in BT1001 (Fig. 2).

Crystal structure of SaRha78A in complex with L-rhamnose – To explore the mechanism of enzyme action the structure of SaRha78A in complex with L-rhamnose, its reaction product, was determined. The bound L-rhamnose molecules were observed in two biologically significant regions of the enzyme (Fig. 1A); in the active site of the catalytic module SaRha78 CM (Rha1501) and bound to SaCBM67 (Rha1511). The average B-factors of the cyclic six atoms of the L-rhamnose rings were 20.9 and 17.9 Å² for Rha1501 and Rha1511, respectively. The overall structure of the SaRha78A–L-rhamnose complex was almost identical to that of ligand-free SaRha78A with an r.m.s.d. of 0.21 Å, implying little effect of ligand binding upon the overall structure.

Catalytic module: Rha1501, located in the active site of SaRha78 CM, adopted an intermediate structure between skewboat 5S, and boat 25B conformations (Figs. 1, 3A, 3B), and the bound L-rhamnose was observed in mainly α-anomeric configuration. The bound L-rhamnose was sandwiched between two aromatic residues, Trp-747 and Trp-640, which thus make extensive hydrophobic interactions with the pyranose ring of the sugar. The pocket topology is completed by additional hydrophobic residues, Trp-695, Tyr-744, and Phe-851. The C-6 methyl group, which comprises the signature feature of L-rhamnose, distinguishing the sugar from L-mannose, was buried in a hydrophobic hollow comprising Trp-695, Tyr-744 and Trp-747. The likely polar contacts between Rha1501 and the protein are as follows: O1 is within hydrogen bonding distance with Glu-636 Oε2, and makes water mediated polar contacts with Glu-636 Oε1 and Arg543 Nη2. O2 hydrogen bonded with Asp-630 Oδ2, His-916 Nε2 and Arg-634 Nη2, where the closest contact was with Asp-630 (2.4 Å), O3 makes polar contacts with Asp-643 Oδ2 and His-916 Nε2 and, through water mediated interactions, with Trp-747 Nε1 and Glu-895 Oε2. This water molecule was situated in close proximity to the C1 atom of Rha1501 with a distance of 2.9 Å. O4 hydrogen bonded with Asp-643 Oδ1 and Trp-695 Nε1. Thus, Rha1501 made eight direct and four solvent mediated hydrogen bonds with SaRha78A, and all these amino acids are conserved in BsRhaB. The O3 and O4 atoms of Rha1501 were located at the bottom of the pocket, and the O1 atom was at the entrance and was solvent exposed, explaining why SaRha78A is an exo-acting enzyme that removes L-rhamnose residues from the non-reducing end of oligosaccharide or polysaccharide substrates.

SaCBM67: SaCBM67 displays a β-jelly roll fold in which the convex (β-sheet 1) and concave (β-sheet 2) β-sheets comprise six and five anti-parallel strands, respectively. The order of the β-strands is as follows: β-sheet 1; β-6, β-3, β-8, β-9/β-1 (forms an interrupted β-strand), β-11; β-sheet 2; β-10, β-2, β-7, β-4, β-5 (Fig. 1B). Although this fold is typical of the vast majority of CBMs (see (22) for review), unusually, the two sheets are not fully solvent exposed; β-sheet 2 is partially occluded by a loop extending from Thr-260 to Lys-282, while the two loops that link SaCBM67 with the rest of the
enzyme lie over β-sheet 1. Rha1511 was positioned in a blind canyon interacting primarily with loops connecting β-strands β-3 and β-4, β-5 and β-6, β-7 and β-8. A central feature of the L-rhamnose binding site is a calcium ion that makes co-ordinate bonds with O3 and O4 of the sugar, while the metal interacts with the protein through Asp-179 Oδ1, Asn-180 Oδ1, Asn-228 Oδ1, Pro-233 main-chain O, and a water mediated contact with Ser-230 main-chain O (Figs. 3C, 3D). The bound L-rhamnose also makes direct hydrogen bonds with Trp-203 Nε1, Asn-180 Nδ2, and Asp-179 Oδ2 through O2, O3 and O4 atoms, respectively. It is worthy to note that the peptide linkage between two calcium coordinating residues Asp-179 Oδ1 Asn-180 Oδ1 was in cis-configuration. Two water molecules also mediated interactions between L-rhamnose and the protein. The C-6 methyl group pointed towards a small hydrophobic pocket comprising Trp-203, Pro-233, Pro-291 and Trp-292. The bound L-rhamnose adopted a relaxed 1C4 chair conformation with O1 atom in α-anomeric configuration. While O2, O3 and O4 pointed at the protein surface, O1 was solvent exposed, indicating that SaCBM67 binds to L-rhamnose residues at the non-reducing termini of complex carbohydrates.

Catalytic residues of SaRha78A and the specificity and function of SaCBM67 – To verify the proposed identity of the catalytic residues of SaRha78A, aspartic acid and glutamine mutants of Glu-636 and Glu-895 were constructed. As expected, the activities of all mutants were at least 40-times lower than those of the wild-type enzyme (Table 2). The values of kcat/Km for SaCBM67 bound L-rhamnose with a Kd of 7.2 × 10^3 M^-1 and free energy of binding ΔG of −5.3 kcal/mol. SaCBM67 did not bind to L-mannose, or L-fucose, demonstrating that the stereochemistry of the sugar at C4 and/or C2 are important specificity determinants. SaCBM67 bound to L-mannose, albeit with a two-fold reduction in affinity. The C-6 methyl group of the bound L-rhamnose was partially buried in a hydrophobic pocket, however, there was sufficient solvation in the region of C6 to accommodate a hydroxyl group. Since L-mannose seldom exists in natural polysaccharides, SaCBM67 binds primarily to L-rhamnose in biological systems.

SaCBM67 did not bind to L-rhamnose in the presence of EDTA, which chelates calcium, demonstrating that the metal ion plays a key role in ligand recognition. Similarly the D179A and N180A mutations, which remove calcium-mediated and direct hydrogen bonds with L-rhamnose, abrogate ligand binding, confirming the importance of calcium in the binding of SaCBM67 to its ligand. To summarize, these data are entirely consistent with the structural data in showing that key residues, and EDTA, demonstrated that the metal ion plays a key role in ligand recognition.

To investigate the role of SaCBM67 in the function of SaRha78A, the activities of the wild type enzyme, in the presence of EDTA, and the mutants D179A and N180A were explored. The data (Table 2) showed that inactivation of SaCBM67, through either calcium chelation with EDTA or mutation of key residues, did not influence activity against aryl-rhamnosides, but caused a substantial reduction (~50-fold) in activity against the rhamnose-containing polysaccharide composite, gum arabic. These results demonstrate that SaCBM67 plays a central role in the action of SaRha78A against polysaccharides; the mechanism for this catalytic potentiation is discussed below.

DISCUSSION
In this study, we determined the crystal
structure of SaRha78A, which was the third example of a GH78 protein. The structure in conjunction with biochemical studies should that the three proteins contain four highly conserved domains. Domain N of SaRha78A comprising a fibronectin-type 3 fold is not evident in the other GH78 proteins, while the other five domains of the *Streptomyces* enzyme are present in BsRhaB or in both BsRhaB and BT1001. Despite the structural similarities of the proteins, the total amino acid identity was less than 15 % compared to the other two proteins. These differences in the domain arrangement in the three α-L-rhamnosidases points to unusual evolutionary pathways within GH78.

**Catalytic mechanism of GH78 enzymes** – Previous studies have established that GH78 enzymes hydrolyse glycosidic bonds through an acid-base assisted single displacement or inverting mechanism (23). Analysis of the crystal structure of the GH78 enzyme BsRhaB soaked with rhamnose revealed electron density in a deep pocket which, by analogy with related inverting (α/α)6-barrel glycoside hydrolases, is likely to comprise the active site. Although the electron density was sufficiently large to comprise a sugar, rhamnose, in its relaxed 1C4 conformation, could not be modeled into the observed density (7). In the structure of the SaRha78A–L-rhamnose complex, an L-rhamnose molecule (Rha1501) was observed in the proposed active site pocket. The modeled L-rhamnose adopted an intermediate conformation between a skewboat 5S1 and boat 2,5B. Previously, in some inverting glycoside hydrolases, including a *Clostridium thermocellum* GH8 endoglucanase (CtCel8A) that also displays a (α/α)6 fold, the sugar bound at the active site (−1 subsite) adopted distorted 2S0,2,5B ring conformations (24–26). Further molecular dynamics studies indicated that the sugar ring in its 2S0 conformation represented the Michaelis complex, which would then adopt a 2,5B conformation at the transition state, and, finally, the reaction product would display a 5S1 conformation (27, 28). By analogy, Rha1501 bound in the active site of SaRha78A appears to be migrating between the transition state and product conformations adopted during the reaction trajectory. Thus, it would appear that a 2,5B-Boat is the conformation adopted by the transition state of glycans hydrolyzed by GH78 rhamnosidases. It is interesting to note that although Cui *et al.* (7) were unable to model an L-rhamnose in its relaxed confirmation into the active site electron density, the authors suggested that this may be because the bound sugar adopted a distorted conformation, as observed here.

Generally, two acidic amino acids, either aspartate or glutamate, are employed by most inverting GHs to catalyze hydrolysis (23, 29). One carboxylate protonates the scissile glycosidic oxygen atom and the other coordinates the nucleophilic water molecule. In the SaRha78A–L-rhamnose complex, O1 of Rha1501 was within hydrogen bonding distance with Glu-636 Oε2 atom. The C-1 hydroxyl group was observed in an α-anomeric configuration, and the O1 atom position could be considered as the scissile bond position of the substrate. Therefore, Glu-636 appeared to comprise the catalytic proton donor (acid) of the enzyme. A water molecule makes a strong hydrogen bond with O1 of Rha1501, occupying the β-anomeric position, and thus is likely to comprise the solvent nucleophile utilized by the “inverting” enzyme. This water molecule makes strong hydrogen bonds with Trp-747 Nε1 and Glu-895 Oε2, and thus Glu-895 appears to be the catalytic general base. These catalytic residues and nucleophilic water are conserved in CtCel8A (24, Fig. 4). Thus, based on the criteria of conservation of fold, catalytic apparatus and catalytic mechanism, we propose that GH78 be included in clan GH-M that currently contains families GH8 and GH48.

Consistent with the proposed catalytic role of Glu-895, the glutamate is conserved in the other two GH78 enzymes for which structures are available. While, Glu-636, the catalytic acid, corresponds to Glu-572 in BsRhaB, in BT1001 the equivalent residue is Asp-337. When Glu-572 and Glu-841 in BsRhaB (corresponding to Glu-636 and
Glu-895 in SaRha78A) were mutated to glutamine, catalytic efficiency decreased by 4-5 orders of magnitude, supporting their proposed catalytic function. Furthermore, the observation that BT1001 is catalytically inactive (data not shown) is entirely consistent with the replacement of Glu-636 in SaRha78A (catalytic acid) with an aspartate, which would be too distant from the nucleophilic water to activate the solvent molecule. Reflecting a pH optimum of 6.0 the catalytic acid of SaRha78A requires a pKa modulator which, typically, is a carboxylate residue. The only candidate pKa modulator in SaRha78A is Asp-630, which is within hydrogen bonding distance with Glu-636. Consistent with the role of this aspartate is the observation that mutation of the equivalent residue in BsRhaB, Asp-567, causes a 10⁵-fold reduction in catalytic activity (7).

L-Rhamnose-binding SaCBM67 – A DALI search indicated that BsRhaB (PDB code 2OKX, 7), but not BT1001, contains a domain that is structurally equivalent to SaCBM67 with a Z-score of 20.6. SaCBM67 also displays weak structural homology with a CBM32 (Z-score of 5.0, PDB code 2VCA, 30), CBM60 (Z-score of 4.4, 2XFE, 31), CBM36 (Z-score of 4.1, 2DCJ, 32) and CBM35 (Z-score of 3.9, 2W87, 33), which recognize their different ligands through either an exo- (CBM32 and CBM35) or endo-mode (CBM36 and CBM60) of binding. These CBMs all comprise a β-jelly roll structure and, in addition to a structural calcium (which is absent in CBM67), contain a second calcium atom that is integral to ligand recognition. In SaCBM67 calcium also plays a key role in L-rhamnose recognition, although the metal binding site is displaced compared to the other CBMs (Fig. 5).

CBM67 members are distributed not only in many bacterial GH78 α-L-rhamnosidases, but also in some Basidiomycete lectins, family 1 pectate lyases, peptidases and proteins of unknown functions. Protein alignment of candidate members of CBM67 (Fig. 6) identified five sub-families within the constructed phylogenetic tree (Fig. S1). In addition to SaCBM67, the only member of CBM67 for which a function is known is the lectin from Pleurotus cornucopiae. The lectin, which shows highest affinity against N-acetyl-d-galactosamine (34, 35), contains two CBM67-like sequences in tandem with sequence identities against SaRhaCBM67 of 25% and 35% for the N- and C-terminal modules, respectively. Hemagglutinating activity of this lectin is inhibited by EDTA and restored by calcium chloride, indicating that calcium binding is necessary for sugar binding. The calcium coordinating residues Asp-179 and Asn-228 of SaCBM67 are conserved in the two lectin modules, while Asn-180, which is conserved in the C-terminal module, is replaced by an aspartate in the N-terminal module. Thus, the calcium binding site is conserved in the CBM67 ligand recognition site in the lectin, a putative polysaccharide lyase family 1 pectate lyase from Sorangium cellulosum (GenBank accession number CAN95071) and α-L-rhamnosidase from Paenibacillus sp. Y412MC10 (ACX62649), and both modules in a putative peptidase S8 and S53 from Pseudoalteromonas atlantica T6c (ABG39167). While the L-rhamnose binding residues in SaCBM67 are conserved in CBM67 from the Paenibacillus α-L-rhamnosidase, suggesting that both proteins bind to L-rhamnose, these amino acids are not retaining in other CBM67 members. In particular, the loop extending from Glu-197—Gly-202 in SaCBM67, which contains Asp-199 that makes a water-mediated interaction with the bound L-rhamnose, is missing in the Basidiomycete lectin, pectate lyase and peptidase, which might cause different sugar specificity. Indeed, while SaCBM67 binds to L-rhamnose, the protein module does not recognize GlcNAc (data not shown), the ligand recognized by the P. cornucopiae lectin. Prior to the discovery of the CBM67 family, the only other, non-enzymatic, L-rhamnose binding proteins were animal lectins for which NMR and crystal structures are available (36, 37). These proteins display
an α/β fold and ligand binding is not metal dependent. Thus, the mechanism of ligand recognition in SaCBM67 is distinct from other proteins that bind to the hexose sugar.

SaRha78A is an example of an enzyme that displays an exo-mode of action in which both the active site of the catalytic module, and the appended CBM, bind to the same terminal sugar. Indeed, the ~50-fold, increase in catalytic activity afforded by the CBM is substantially greater than the 2-4 fold enhancement mediated by typical endo-binding CBMs (38, 39; see 22 for review). Gum arabic is a highly complex mixture of polysaccharides and glycoproteins, in which the predominant structure is an arabionogalactan in which the β-1,3-galactan backbone contains branches that are capped with L-rhamnose or L-arabinofuranose residues (40). This presents a structure in which both the GH78 catalytic module and CBM67 of SaRha78 can bind to different terminal L-rhamnose residues of the same polysaccharide molecule. The ensuing avidity effect will result in much tighter binding of SaRha78 to gum arabic, compared to either the CBM or the catalytic module as discrete entities, leading to the observed, CBM-mediated, enhanced catalytic efficiency. This emerging model for how exo-acting CBMs potentiate the activity of exo-acting glycosidases is supported by a recent study showing that a CBM, which binds to terminal fructose residues, mediates a substantial increase in the activity of an exo-acting β-fructosidase against levan, a highly branched fructose-containing polysaccharide (41).

REFERENCES


Acknowledgement — We would like to thank the beamline researchers and staffs at Photon Factory and SPring-8. We thank Dr. Hitomi Ichinose for her technical support for SaCBM67 construction and sample preparation for the crystallization.

FOOTNOTES

* This work was supported, in part, by JSPS KAKENHI Grant Number 22580110.
The atomic coordinates and structure factors (codes 3W5M and 3W5N) have been deposited in the Protein Data Bank through PDBj, Research Collaboratory for Structural Bioinformatics, Rutgers University, New Brunswick, NJ (http://www.rcsb.org/).

1 Z.F, and A.J. contributed equally to this work.

2 Address correspondence to : Satoshi Kaneko, National Food Research Institute, 2-1-12 Kannondai, Tsukuba, Ibaraki 305-8642, Japan; Phone. +81-29-838-8063; Fax +81-29-838-7996; E-mail: sakaneko@affrc.go.jp, or Zui Fujimoto, National Institute of Agrobiological Sciences, 2-1-2 Kannondai, Tsukuba, Ibaraki 305-8602, Japan; Phone and Fax. +81-29-838-7877; E-mail: zui@affrc.go.jp or Harry J. Gilbert, Newcastle University, Medical School, Newcastle upon Tyne NE2 4HH, United Kingdom. Phone. +44-(0)191-2228800; E-mail: harry.gilbert@newcastle.ac.uk.

3 The abbreviations used are: BsRhaB: *Bacillus* sp. GL1 α-L-rhamnosidase B; BT1001: *Bacteroides thetaiotaomicron* VPI-5482 α-L-rhamnosidase; CBM: carbohydrate-binding module family; GH: glycoside hydrolase family; PNP-α-L-Rha: p-nitrophenyl-α-L-rhamnopyranoside; r.m.s.d.: root mean square differences or deviations; SaCBM67: carbohydrate-binding module family xx of *Streptomyces avermitilis* α-L-rhamnosidase; SaRha78A: *Streptomyces avermitilis* α-L-rhamnosidase catalytic module

SeMet: selenomethionine

FIGURE LEGENDS

**Figure 1.** Structure of SaRha78A. A, Stereoview of the ribbon model of the SaRha78A–L-rhamnose complex structure. Each domain is drawn in different color; Domains N, D, E, F, A C are in blue, cyan, green, yellow, orange and red colors, respectively. Calcium ion as pink sphere, and the bound L-rhamnose molecules in gray stick models. B, Topology diagram of SaRha78A. α-Helices, 310-helices and β-strands are shown in filled cylinders, shaded cylinders and filled arrows, respectively. Domains are colored in accordance with Fig. 1A. β-strands are numbered in domain D.

**Figure 2.** Stereoview of the superimposed ribbon-models of GH78 proteins. SaRha78A, BsRhaB (PDB code, 2OKX; (7), BT1001 (PDB code, 3CIH) are drawn in orange, sky blue, and light green, respectively.

**Figure 3.** The bound L-rhamnose structures in the SaRha78A–L-rhamnose complex. Bound L-rhamnose molecules are shown in gray stick models. Estimated hydrogen bonds are shown in cyan, water-mediating hydrogen bonds in green, ion coordination in magenta break lines. A, Stereo view of the catalytic pocket. Two catalytic residues are in pale-red color. B, 2Fcalc–Fobs electron density map of the bound L-rhamnose in the catalytic pocket. Contour level is 1σ. C,
Stereo view of the sugar-binding pocket in domain D (SaCBM67). Calcium ion as pink sphere. 

**D**

, Electron density map of the bound L-rhamnose and calcium ion shown with contour level of 

1 σ.

**Figure 4.** Superimposed model of the catalytic site of SaRha78A (orange) on that of 

*Clostridium thermocellum* endoglucanase CelA E95Q mutant complexed with cellopentaose 

(gray, 24). Two catalytic residues, the nucleophilic waters in the inverting reaction mechanism, 

bound sugars at subsite –1 are shown in stick models.

**Figure 5.** Stereoview of the superimposed ribbon models of SaCBM67 and the different CBM 

structures. Calcium ions are drawn as spheres, and the bound sugars in stick models. SaCBM67, 

green; CBM32 in sialidase toxin from *Clostridium perfringens* (PDB code, 2V72; 42), pink; 

CBM35 in xylanase from *Cellvibrio japonicus* (PDB code, 2W87; 33), lightorange; CBM36 in 

xylanase from *Paenibacillus polymyxa* ATCC 842 (PDB code, 1UX7; 43), lightblue; CBM60 

from uncultured bacterium (PDB code, 2XFD; 31), lightgreen.

**Figure 6** Amino acid sequence alignment of SaCBM67 with selected CBM67 modules from 

related proteins. PcLec_N, *Pleurotus cornucopiae* lectin N-terminal domain (Genbank accession 

number BAD16583); PcLec_C, *P. cornucopiae* lectin C-terminal domain; ScPL1, *Sorangium 

cellulosum* pectate lyase (CAN95071); PsRha78, *Paenibacillus* sp. Y412MC10 

α-L-rhamnosidase (ACX62649); PaPep_N, *Pseudoalteromonas atlantica* T6c N-terminal 

domain (ABG39167); PaPep_C, *P. atlantica* T6c C-terminal domain. Residues conserved in at 

least five sequences are shaded. Residues coordinating calcium ion are indicated by ©, other 

residues in contact with the bound L-rhamnose are labelled with ®. The secondary structure 

elements of SaCBM67 are indicated by arrows for β-strands and coils for 3_10-helices.
Table 1. Data collection and refinement statistics of α-L-rhamnosidase

Values in parentheses refer to the highest resolution shell. r.m.s.d., root mean square deviations

<table>
<thead>
<tr>
<th>Data collection</th>
<th>Native Apo</th>
<th>1-Rhamnose complex</th>
<th>SeMet- Peak</th>
</tr>
</thead>
<tbody>
<tr>
<td>PDB code</td>
<td>3W5M</td>
<td>3W5N</td>
<td></td>
</tr>
<tr>
<td>Space group</td>
<td>(P2_1)</td>
<td>(P2_1)</td>
<td>(P2_1)</td>
</tr>
<tr>
<td>Cell parameters (Å)</td>
<td>(a = 52.8, b = 128.4, c = 75.2)</td>
<td>(a = 53.1, b = 128.6, c = 75.3)</td>
<td>(a = 52.9, b = 128.8, c = 75.0)</td>
</tr>
<tr>
<td>(β)</td>
<td>(99.3)</td>
<td>(99.8)</td>
<td>(99.4)</td>
</tr>
<tr>
<td>X-ray source</td>
<td>Spring-8 BL41XU</td>
<td>PF-AR BL-NW12</td>
<td>PF BL-5</td>
</tr>
<tr>
<td>Wavelength (Å)</td>
<td>0.97915</td>
<td>1.00000</td>
<td>0.97909</td>
</tr>
<tr>
<td>Resolution (Å)</td>
<td>50–1.8 (1.86–1.80)</td>
<td>50–1.8 (1.86–1.80)</td>
<td>50–2.0 (2.07–2.00)</td>
</tr>
<tr>
<td>No. reflections</td>
<td>640 389</td>
<td>694 122</td>
<td>499 517</td>
</tr>
<tr>
<td>No. unique reflections</td>
<td>92 680 (9 140)</td>
<td>91 813 (9 161)</td>
<td>66 810 (6 533)</td>
</tr>
<tr>
<td>Completeness (%)</td>
<td>99.8 (98.5)</td>
<td>100.0 (99.9)</td>
<td>99.8 (97.9)</td>
</tr>
<tr>
<td>Multiplicity</td>
<td>6.9 (5.7)</td>
<td>7.6 (7.3)</td>
<td>7.5 (6.6)</td>
</tr>
<tr>
<td>(R)-merge</td>
<td>0.099 (0.632)</td>
<td>0.084 (0.558)</td>
<td>0.085 (0.387)</td>
</tr>
<tr>
<td>Average I/σ</td>
<td>36.9 (4.3)</td>
<td>27.5 (3.9)</td>
<td>25.8 (4.5)</td>
</tr>
</tbody>
</table>

Refinement

| Resolution (Å) | 35.7–1.8 (1.85–1.80) | 40.6–1.8 (1.85–1.80) |
| \(R\)-factor  | 0.180 (0.275) | 0.163 (0.230) |
| \(R\)-free   | 0.214 (0.378) | 0.194 (0.287) |
| No. waters    | 758 | 1099 |
| Average \(B\)-value (Å²) | 30.3 | 19.6 |
| r.m.s.d. from ideals \(L\)-values (Å) | 0.010 | 0.009 |
| angles (°)    | 1.40 | 1.26 |
| Ramachandran plot (%) | 98.1 / 1.9 / 0.0 | 97.6 / 2.4 / 0.0 |
| (favored/allowed/disallowed) | |

14
<table>
<thead>
<tr>
<th>Enzyme</th>
<th>Substrate</th>
<th>CaCl₂ *1</th>
<th>$K_m$ (mM)</th>
<th>$k_{cat}$ (s⁻¹)</th>
<th>$k_{cat}/K_m$ (mM⁻¹ s⁻¹)</th>
</tr>
</thead>
<tbody>
<tr>
<td>WT</td>
<td>PNP-α-L-Rha</td>
<td>–</td>
<td>0.026 ± 0.003</td>
<td>26.4 ± 1.7</td>
<td>1014</td>
</tr>
<tr>
<td>WT</td>
<td>PNP-α-L-Rha</td>
<td>+</td>
<td>0.024 ± 0.002</td>
<td>27.2 ± 1.2</td>
<td>1115</td>
</tr>
<tr>
<td>WT</td>
<td>Gum arabic</td>
<td>+</td>
<td>238 ± 61 *2</td>
<td>0.003 ± 0.001</td>
<td>1.4×10⁻⁵ ± 7×10⁻⁶ *3</td>
</tr>
<tr>
<td>E636D</td>
<td>PNP-α-L-Rha</td>
<td>–</td>
<td>0.015 ± 0.002</td>
<td>0.07 ± 0.004</td>
<td>4.66</td>
</tr>
<tr>
<td>E636Q</td>
<td>PNP-α-L-Rha</td>
<td>–</td>
<td>0.22 ± 0.03</td>
<td>2.49 ± 0.18</td>
<td>11.3</td>
</tr>
<tr>
<td>E895D</td>
<td>PNP-α-L-Rha</td>
<td>–</td>
<td>0.022 ± 0.002</td>
<td>0.54 ± 0.20</td>
<td>24.5</td>
</tr>
<tr>
<td>E895Q</td>
<td>PNP-α-L-Rha</td>
<td>–</td>
<td>0.012 ± 0.003</td>
<td>0.29 ± 0.13</td>
<td>24.9</td>
</tr>
<tr>
<td>D179A</td>
<td>PNP-α-L-Rha</td>
<td>+</td>
<td>0.028 ± 0.002</td>
<td>27.5 ± 1.3</td>
<td>975</td>
</tr>
<tr>
<td>D179A</td>
<td>Gum arabic</td>
<td>+</td>
<td>N.D. *4</td>
<td>N.D. *4</td>
<td>2.7×10⁻⁷ ± 4×10⁻⁹ *3</td>
</tr>
<tr>
<td>N180A</td>
<td>PNP-α-L-Rha</td>
<td>+</td>
<td>0.026 ± 0.004</td>
<td>24.4 ± 0.78</td>
<td>939</td>
</tr>
<tr>
<td>N180A</td>
<td>Gum arabic</td>
<td>+</td>
<td>N.D. *4</td>
<td>N.D. *4</td>
<td>2.7×10⁻⁷ ± 2×10⁻⁸ *1</td>
</tr>
</tbody>
</table>

*1 Reaction mixture contained final 2 mM CaCl₂ indicated as +.
*2 The figure in the Table is given in mg/ml.
*3 The figures in the Table are given in mg/mI⁻¹s⁻¹.
*4 Not determined as $K_m$ was too high to measure.
Table 3. Binding parameters for the recognition of sugars by SaCBM67 and mutants

<table>
<thead>
<tr>
<th>Ligand</th>
<th>RaCBM67</th>
<th>EDTA</th>
<th>N</th>
<th>$K_a$ (M$^{-1}$)</th>
<th>$\Delta H$ (kcal/mol)</th>
<th>$T\Delta S$ (kcal/mol)</th>
<th>$\Delta G$ (kcal/mol)</th>
</tr>
</thead>
<tbody>
<tr>
<td>L-Rhamnose</td>
<td>Wild type</td>
<td>0</td>
<td>1.00 ± 0.03</td>
<td>7.2 × 10$^3$ ± 742.0</td>
<td>-4.9 ± 0.18</td>
<td>0.4 ± 0.02</td>
<td>-5.3 ± 0.06</td>
</tr>
<tr>
<td></td>
<td>Wild type</td>
<td>5 mM</td>
<td>No Binding</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>L-Rhamnose</td>
<td>D179A</td>
<td>0</td>
<td>No Binding</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>L-Rhamnose</td>
<td>N180A</td>
<td>0</td>
<td>No Binding</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>L-Mannose</td>
<td>Wild type</td>
<td>0</td>
<td>1.05 ± 0.04</td>
<td>3.6 × 10$^3$ ± 430</td>
<td>-4.6 ± 0.03</td>
<td>0.2 ± 0.01</td>
<td>-4.8 ± 0.04</td>
</tr>
<tr>
<td></td>
<td>D179A</td>
<td>0</td>
<td>No Binding</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>L-Mannose</td>
<td>N180A</td>
<td>0</td>
<td>No Binding</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>L-Fucose</td>
<td>Wild type</td>
<td>0</td>
<td>No Binding</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The concentration of SaCBM67 in the cell was 700 µM for L-rhamnose and 800 µM for L-mannose. The concentration of both ligands in the syringe was 20 mM.
Figure 1

Crystal structure of Streptomyces avermitilis GH-78 α-L-rhamnosidase
Crystal structure of *Streptomyces avermitilis* GH-78 α-L-rhamnosidase

**Figure 2**

[Diagram showing the crystal structure with labeled domains: Domain A, Domain B, Domain C, Domain D, Domain E, Domain F, Domain N]
Figure 3

Crystal structure of Streptomyces avermitilis GH-78 α-L-rhamnosidase
Figure 4

Crystal structure of Streptomyces avermitilis GH-78 α-L-rhamnosidase
Figure 5

Crystal structure of Streptomyces avermitilis GH-78 α-L-rhamnosidase
Crystal structure of Streptomyces avermitilis GH-78 α-L-rhamnosidase

Figure 6

SaCBM67 134:---SLE---GSSWIWFPEGE-PANSAAPAATRWFRERTV---DLPDD---ITGAT---------LAIS

PcLec_N 49:PAFLN---ISKYWTGENTVPEANNPVGTARFKNIT---TSACGGCATCAT---------IVVA

PcLec_C 214:---PLA---QSKWIATWSA---ATTAPAASNAFKTI---DD-CTKVACAT---------VLIS

ScPlI 59:SPTLP---ELA-----CDGEPPAPELPPG-AGFGE---ELASSPLVSAASYLGAEGGLAIT

PsRha78 1:---MKDHWQASWIGDFPNTAN---VXLEARRSF---DLQEAP-MA-GR---------LYVS

PaPep_N 455:-----------------SALLNTAKAVDSGSLTFGRSHAIGAAAEMLMPKEIDQQLS---------IRIS

PaPep_C 695:AGTPDA-PGEFVTVEVTDTGG---------EQAQLEL---------SIAGAPITSEMA---------ITIS

© β-4            β-5            ® ©            ® ©            β-6            β-7

SaCBM67 178:ADNYVAVSDGAEVARTDLEADNECGWRPAPIVDLVDVVHGNTLVAASNAVS---------

PcLec_N 98:ADDSSTTVVVNGVAIGSAG---------WTTPGQYFAP---LNPSSNLFAIAVGNVA---------

PcLec_C 258:ADNHYKLVVNGQAVSGDS---------FGRAEAYSIP-KLHPTLTLAIDAKNDEG---------

ScPlI 110:GRDTPRFFVYNGDLIAAS---------ESRAPFVPLSSLPGNVIAVSVHAAAGT---------

PsRha78 44:ANQEYLLYVNTVVGRGSPADQQ---WYYDEIDVTPLEPKGRNVLAVAYNFGSTDIVTG---------

PaPep_N 499:VDNSEELYFGVWLLGSSTN---------WTGKLYNV---------NVTQVNVLAVKAMDV---------

PaPep_C 742:VDNFEDTYFNGVQFLGQSN---------WYAKSYTV---------ELVSQNVLAVQKADV---------

© β-8            β-9            β-10

SaCBM67 233:-----------------PAGWICVVLTLTASGERK-IFSDASWK-----------------S-TDHEP

PcLec_N 144:-------------------RAALMATINHYSDTGETFITDESWK-------------------T-VRGAV

PcLec_C 305:-------------------PAGVIAHITYRTDQTNQTIVTDGSWK-------------------ASQT

ScPlI 157:-------------------PAAVHLDELERSHGSG---------SDWK-------------------LSTAP

PsRha78 103:QMHCQGLIALEIERNGVGCVAATDADWKRPRWMAQRVQHKWNGFREVYLAQQ-----------

PaPep_N 545:-------------------LAALIAQLDLDQDP-FYSDEIWK-------------------VSTEF

PaPep_C 788:-------------------IAALIAKIETEQQ---------VIVSDSDWK-------------------ISTQT

© β-11            ©

SaCBM67 265:ADGREWDPFDSCWSAPAAAIAAWGAG---------PWGRV-------------------

PcLec_N 177:PGQFQLPATDSSTWTFAMLQFPIS---FWGNP-------------------

PcLec_C 337:PGFQETFPFDSDSWTATVQVNGYIA---PWGSAVAIIPPA-------------------

ScPlI 183:EGDWATAYDGCGTARTELGSAGDI---PCCDPEA-------------------

PsRha78 162:EDGWEESAYDSSSWSGAVIAKAMDPDSPWPLL-------------------

PaPep_N 572:VEGQOQTDFFDSQWQAASYGYYGVW---PWYKKVKWPPSSQAQLW-------------------

PaPep_C 815:FDWWNTQAFDSSWGNARAYGYSYGN---PWRSRSVLNAGAKWIWS-------------------

22
The structure of a Streptomyces avermitilis α-L-rhamnosidase reveals a novel carbohydrate-binding module CBM67 within the six-domain arrangement
Zui Fujimoto, Adam Jackson, Mari Michikawa, Tomoko Maehara, Mitsuru Momma, Bernard Henrissat, Harry J. Gilbert and Satoshi Kaneko

*J. Biol. Chem.* published online March 13, 2013

Access the most updated version of this article at doi: 10.1074/jbc.M113.460097

Alerts:
- When this article is cited
- When a correction for this article is posted

Click here to choose from all of JBC’s e-mail alerts

Supplemental material:
http://www.jbc.org/content/suppl/2013/03/13/M113.460097.DC1

This article cites 0 references, 0 of which can be accessed free at
http://www.jbc.org/content/early/2013/03/13/jbc.M113.460097.full.html#ref-list-1