FACULTY OF SCIENCE
UNIVERSITY OF COPENHAGEN

Structure and Function of microbial and plant Rhamnogalacturonan Lyases

Malene H. Jensen[a], Ulla Christensen[a], Harm Otten[a], Torben Borchert[c], Lars L. H. Christensen[c], Sine Larsen[a,b], and Leila Lo Leggio[a]

[a]*Biophysical Chemistry Group, Department of Chemistry, University of Copenhagen, Denmark*
[b]*European Synchrotron Radiation Facility (ESRF), Grenoble, France*
[c]*Novozymes A/S, Smørmosevej 25, 2880 Bagsværd, Denmark*

Abstract

We present here the first experimental evidence for an active site with bound substrate in a rhamnogalacturonan lyase belonging to family PL4 according to the CAZy classification (http://www.cazy.org, [1]). *Aspergillus aculeatus* Rhamnogalacturonan lyase (AARGL) is involved in the degradation of rhamnogalacturonan-I, an important plant cell wall polysaccharide and the structure for the native enzyme has previously been published (PDB ID: 1NKG, [2]). Based on this structure, enzyme variants H210A and K150A have been produced and characterized both kinetically and structurally showing that His210 and Lys150 are key active site residues. Electron density maps from single crystal x-ray diffraction data collected at MaxLab, Lund, Sweden, after soaking the K150A variant with a rhamnogalacturonan digest showed evidence of bound substrate at the -3/+3 subsites (see Figure 2). Based on the crystallographic and kinetic studies on AARGL, and its structural and sequence comparison to other enzymes in the same and other PL families, we propose a mechanism for the β-elimination on [-,2)-α-L-Rhamno-(1,4)-α-D-Galacturonic acid-(1,-]. The lack of direct involvement of calcium in the mechanism is significantly different to other characterized lyases, so that a different mechanism must be in place.

Keywords: biological macromolecular crystallography, carbohydrate degradation, pectate lyases substrate binding

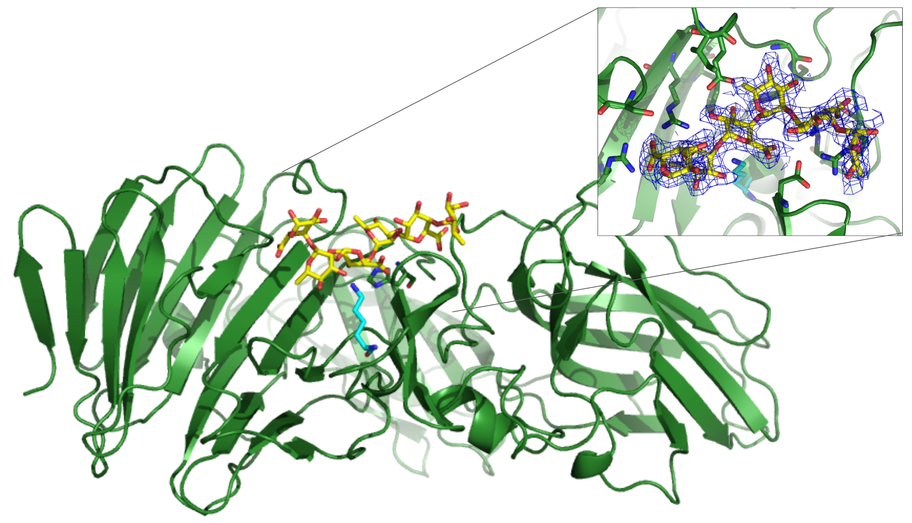
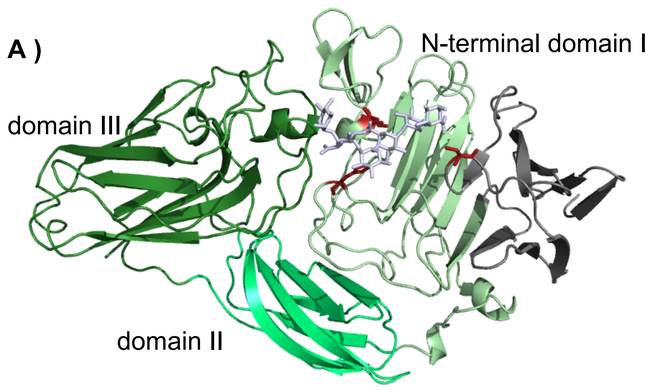


Fig.1. Structure of the *A. aculeatus* RGL with a substrate hexasaccharide [-,2)-α-l-Rhamno-(1,4)-α-d-Galacturonic acid-(1,-] bound. Magnified is the found electron density with the modeled oligosaccharide. The proposed catalytic residue Lys150 is depicted in cyan.

Introduction

Rhamnogalacturonan lyases

Carbohydrate modifying enzymes (CAZYmes) play essential parts in metabolism, antibiotic synthesis, host-pathogen interactions, signal transduction and are important in the synthesis and breakdown of glycosylated proteins, lipids, steroids, and structural polysaccharides [3-7].
CAZYmes are of significant interest in the agricultural, beverage, food, and energy resource industries. Desirable properties for industrial CAZYmes span from high catalytic efficiency to broad or very narrow substrate specificity, all depending on the specific application. An especially sought-after target would be identifying and modifying CAZYmes for the degradation of plant cell walls of pectate rich agricultural or urban waste to generate energy resources. Other possible uses of CAZYmes lie within the production of fine chemicals and medicinal compounds [8].
The **C**arbohydrate **A**ctive en**zy**mes database (http://www.cazy.org, [1]) is a major resource for the study of carbohydrate active enzymes. Unlike the IUMB classification based on activity and specificity, the CAZy classification is based on sequence and structure. This is particularly effective in the post-genomic era, as uncharacterized genes and gene products can easily be incorporated in the database. The database is organized into major functional/activity classes where the enzymes are organized in numbered families. The polysaccharide lyases (PL) classification is particularly relevant to this project and will be used as a point of reference.
The PL enzymes cleave glycosidic bonds, resulting in a new double bond on the non-reducing molecule, the former glycoside acceptor or aglycon. The rhamnogalacturonan lyases are part of this group for which the CAZy database lists over 1150 proteins. They are named after their polysaccharide substrate rhamnogalacturonan I (RG-I), a component of plant cell walls. It is a regular copolymer made of repeating [-,2)-α-L-Rha-(1,4)-α-D-GalUA-(1,-] units.
One of the 39 members of PL4 is the *A. aculeatus* rhamnogalacturonan lyase (AA-RGL). It is an enzyme which is degrading the pectin backbone degrading as mentioned above. It cleaves the rhamnogalacturonan I (RG-I) backbone between Rha and GalUA through a lyase mechanism resulting in a double bond in the non-reducing GalUA residue.



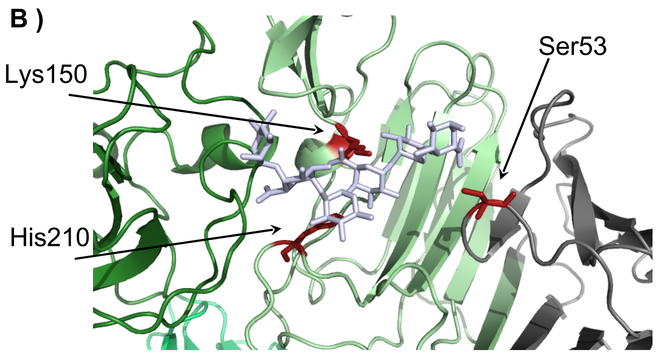


Fig. 2. A) Structure of the *A. aculeatus* RGL (AARGL, PDB ID: 1NGK; PL4). Colored in black is the 90 amino acid long N-terminal fraction of domain I (1-257 aa) missing in the *A. thaliana* enzymes. The remaining domain I is colored in pale green, whereas domain II (268-335) and domain III (337-508) are depicted in lime and forest green respectively. **B)** The putative catalytic residues Lys150 and His210 in proximity to the missing black N-terminal fragment are highlighted in red, as well as Ser53, which is nearest (2.7 Å) to the bluish white tinted substrate.

AA-RGL structure

Crystallography

The structure of AA-RGL, a monomer of 508 amino-acid residues [2], was previously determined at a resolution of 1.5 Å. It is the first three-dimensional characterized enzyme of PL4 and revealed a new modular structure with three domains, consisting predominantly of β-structure. The structure is unique, though individually the three domains showed structural homology to already known protein folds. Based on the structure two residues were proposed to be involved in catalysis, Lys150 and His210 [Fig.3].
The crystals of the native structure were grown in 20% PEG 4000, 9% PEG 400, 0.1 M $(NH_4)_2SO_4$ and 0.1 M sodium acetate pH 4.4. The mutants crystallized in conditions similar to the native enzyme, 20% PEG 4000, 0.1 M $(NH_4)_2SO_4$ and 0.1 M sodium acetate pH 4.3 giving isomorphous crystals. The mutations do not disturb the overall structure, consistent with a direct role of K150 and H210 in catalysis or substrate binding.
Crystals of AA-RGL K150A were soaked in a degradation product from the kinetic assays in which commercially available RG was treated by the native AA-RGL.
The three dimensional structure of the complex between AA-RGL mutant K150A and a stretch of RG backbone was solved to 2.4 Å resolution [Fig.1]. The electron density revealed six sugar rings, rhamnose and galacturonic acid in alternating sequence bound in the proposed active site mainly through interactions with GalUA. The first residue in the non-reducing end of the hexamer is Rha (subsite -3), and the last is GalUA (subsite +3).

Mechanism

Enzyme kinetics and mutagenesis

Putative active site mutants K150A and H210A of AA-RGL were made by site-directed mutagenesis and the proteins purified. The steady state kinetics parameters of AA-RGL catalyzed cleavage of soluble potato RG were determined at 30 °C, pH 6.0 (0.1M Na_2HPO_4 buffer) with 0.425 µg/ml enzyme by measuring the formation of double bonds at λ = 235 nm.
The activity of the AA-RGL mutants K150A and H210A was investigated using insoluble AZCL-galactan from potato measuring the release of azurin linked oligosaccharides at a wavelength of λ = 620 nm [Tab.1].

	Concentration / µg/ml	A_{620} after 72h	Relative activity
Native RGL	0.34	0.41 ± 0.01	1
Mutant K150A	2330	0.030 ± 0.002	1.07×10^{-5}
Mutant H210A	1455	0.713 ± 0.04	4.06×10^{-4}

Tab. 1. Activity of RGL and the mutants K150A and H210A. The assays used 0.1 w/v% AZCL-Galactan at 30 °C. It resulted in k_{cat} = 4.2 ± 0.14 s-1 and K_m = 1.25 ± 0.09 mg/ml (or 0.27 ± 0.02 mM cleavable bonds) .

Known mechanisms

PLs degrade polysaccharides by cleavage of the glycosidic bond with the formation of a double bond on one end in contrast to polysaccharide hydrolases, which do not generate unsaturated products. They are ubiquitous but predominantly of microbial origin [9]. Two types of PL with fundamentally different mechanisms are known. The glucan lyases, which are closely related to the glycoside hydrolase family GH31, generate the double bond on the non-reducing end side of the cleaved oligosaccharide. Pectate, pectin and rhamnogalacturonan lyases (RGL) however generate a double bond on the reducing end side. Currently CAZy lists 18 families of PLs.
The most generally accepted mechanism for PLs involves a Ca^{++} ion and the abstraction of an "activated" proton from a pyranoside-ring C5 atom, which is itself adjacent to a C6 carboxylate moiety. A β-elimination reaction ensues, releasing a 4,5-unsaturated sugar. Such a mechanism demands a C6 carboxylate substitution, such as is found on glucuronic and galacturonic acids [7]. It also is generally associated to high pH (8-10) optimum while AARGL is active at pHs lower than 6.0 [10]. The mechanism for PL4 must be significantly different from the pectate lyase mechanism, which strictly requires calcium ions. This dependency is also shown in another recently characterized PL11 RGL, *Bacillus Subtilis* YesW [11]. In this polysaccharide lyase the overall structure exhibits a 8-bladed β-propeller fold in contrast to the β-sandwich/β-sheet structure of AARGL. Not only deviates the overall secondary structure, but also the binding site on substrate superposition. Only a small resemblance is shown by arginines stabilizing the carboxylate moiety. YesW histidines coordinate the necessary calcium ion instead of the substrate and a tyrosine residue is irreplaceable in the YesW mechanism. In the overlay it is immanent that the two enzymes cover different faces of the substrate.

Plant homologues

In plants, processes for cell-wall synthesis and expansion, starch metabolism, defense against pathogens, symbiosis and signaling involve CAZYmes. This is underlined by a study of the *Arabidopsis thaliana* genome which shows over 730 open reading frames for the two main classes of CAZYmes, glycoside hydrolases and glycosyltransferases. The vast importance of these enzymes in cell-wall formation and degradation is revealed along with the unexpected dominance of pectin degradation, with at least 170 open-reading frames dedicated solely to this task [7].
In *A. thaliana* several open reading frames appear as RGL like enzymes of the PL4 family and are called MYST-multigene family [12]. Alignments show that sequence identities between plant sequences are above 45%, while maximum identities with fungal and bacterial proteins are 24 % and 29 %, respectively [not shown]. The putative catalytic residues from AARGL, Lys150 and His210, are likely to be conserved, though sequence alignments with AARGL are poor in this region. Thus the *A. thaliana* proteins might also be true lyases. The plant sequences lack approximately 90 residues at the N-terminal end though, resulting in a considerably shortened domain I [Fig.2]. This might have an important impact on the activity by affecting a loop region in proximity of the catalytic site. The effect of this shortened domain on enzymatic function is not yet investigated.

References

1. Coutinho, P.M. and B. Henrissat, *Carbohydrate-active enzymes: an integrated database approach.* Recent Advances in Carbohydrate Bioengineering, ed. H.J. Gilbert, et al. 1999, Cambridge: The Royal Society of Chemistry.
2. McDonough, M.A., et al., *Rhamnogalacturonan lyase reveals a unique three-domain modular structure for polysaccharide lyase family 4.* FEBS Lett, 2004. **565**(1-3): p. 188-94.
3. Aspeborg, H., et al., *Carbohydrate-active enzymes involved in the secondary cell wall biogenesis in hybrid aspen.* Plant Physiol, 2005. **137**(3): p. 983-97.
4. Coutinho, P.M. and B. Henrissat, *Life with no sugars?* J Mol Microbiol Biotechnol, 1999. **1**(2): p. 307-8.
5. Coutinho, P.M., et al., *Why are there so many carbohydrate-active enzyme-related genes in plants?* Trends Plant Sci, 2003. **8**(12): p. 563-5.
6. Geisler-Lee, J., et al., *Poplar carbohydrate-active enzymes. Gene identification and expression analyses.* Plant Physiol, 2006. **140**(3): p. 946-62.
7. Henrissat, B., P.M. Coutinho, and G.J. Davies, *A census of carbohydrate-active enzymes in the genome of Arabidopsis thaliana.* Plant Mol Biol, 2001. **47**(1-2): p. 55-72.
8. Buchholz, K. and J. Seibel, *Industrial carbohydrate biotransformations.* Carbohydr Res, 2008.
9. Linhardt, R.J., P.M. Galliher, and C.L. Cooney, *Polysaccharide lyases.* Appl Biochem Biotechnol, 1986. **12**(2): p. 135-76.
10. Kofod, L.V., et al., *Cloning and characterization of two structurally and functionally divergent rhamnogalacturonases from Aspergillus aculeatus.* J Biol Chem, 1994. **269**(46): p. 29182-9.
11. Ochiai, A., et al., *A novel structural fold in polysaccharide lyases: Bacillus subtilis family 11 rhamnogalacturonan lyase YesW with an eight-bladed beta-propeller.* J Biol Chem, 2007. **282**(51): p. 37134-45.
12. Tavares, R., et al., *Organization and structural evolution of four multigene families in Arabidopsis thaliana: AtLCAD, AtLGT, AtMYST and AtHD-GL2.* Plant Mol Biol, 2000. **42**(5): p. 703-17.

The authors acknowledge financial support by SNF (Danish Natural Sciences Research Council), DGF (The Danish National Research Foundation) and Novozymes A/S. We are grateful for synchrotron beamtime and beamline assistance at MAXLAB, Sweden, and travel support through the ARI and Dansync programmes.

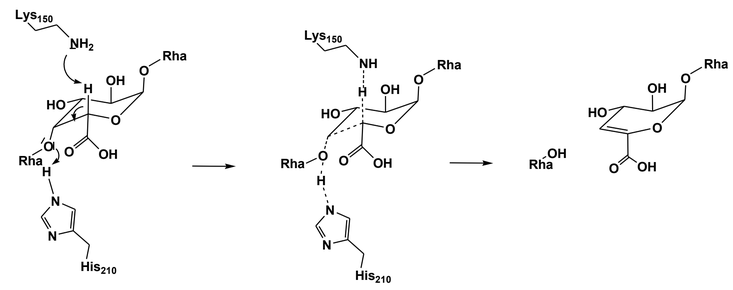


Fig. 3. Proposed reaction mechanism of PL4 lyases [to be published].