Acta Cryst. (1997). D53, 105–107

Crystallization and preliminary X-ray studies of rhamnogalacturonase A from *Aspergillus aculeatus*

THOMAS N. PETERSEN,[a] STEPHEN CHRISTGAU,[b] LENE V. KOFOD,[b] SAKARI KAUPPINEN,[b] HENRIK DALBØGE,[b] ANDERS H. JOHNSON[c] AND SINE LARSEN[a]* *at Centre for Crystallographic Studies, H. C. Ørsted Institute, University of Copenhagen, Universitetsparken 5, DK-2100 Copenhagen, Denmark, [b]Screen Biotechnology, Novo Nordisk A/S, Novo Allé, 2880 Bagsvaerd, Denmark, and [c]Department of Clinical Biochemistry, Rigshospitalet, University of Copenhagen, Blegdamsvej 9, DK-2100 Copenhagen, Denmark. E-mail: sine@xray.ki.ku.dk*

(Received 2 May 1996; accepted 3 September 1996)

Abstract

Recombinant rhamnogalacturonase A from *Aspergillus aculeatus* has been crystallized and X-ray diffraction data has been collected. Crystals were grown by the hanging-drop vapour-diffusion technique, under the conditions 10% PEG 8000, 0.05 *M* KH_2PO_4 and 0.1 *M* sodium acetate buffered at pH 3.5. The crystals diffract beyond 2.0 Å resolution and belong to one of the orthorhombic space groups $I2_12_12_1$ or $I222$, with the unit-cell parameters $a = 62.9$, $b = 125.4$ and $c = 137.0$ Å. There is one molecule in the asymmetric unit and a solvent content of approximately 54%. The enzyme is highly glycosylated corresponding to 5.9 kDa.

1. Introduction

A major component of the plant cell wall is pectin, which is a complex network of polysaccharides. Pectin consists of smooth regions of homogalacturonan and hairy regions of rhamnogalacturonan. The backbone in rhamnogalacturonan is a polysaccharide, composed of alternating rhamnose and galacturonic acid residues, linked by alternating α1–2 and α1–4 glycosidic linkages, (1–2)-α-L-Rha(1–4)-α-D-GalUA. A branch point at the C4 position of rhamnose can serve as an attachment site for carbohydrate side chains like arabian, galactan and arabinogalactan. This makes the region highly ramified and therefore the term 'hairy region' is used to denote this part of pectin.

In contrast of the large number of known enzymes that are able to degrade the smooth regions of pectin, only limited attention has been given to enzymes acting on the hairy regions. (Kofod *et al.*, 1994). Rhamnogalacturonase A (RGase A) from *Aspergillus aculeatus* belongs to a new family of glycosyl hydrolases (Kofod *et al.*, 1994). It is a plant cell-wall degrading enzyme, that hydrolyses within the hairy region of pectin. The cDNA encoding RGase A has been cloned and overexpressed in *Aspergillus oryzae* to produce recombinant enzyme in large quantities (Kofod *et al.*, 1994). RGase A is an endo-acting enzyme, hydrolysing the glycosidic linkage α-D-GalUA(1–2)-α-L-Rha in the backbone of rhamnogalacturonan. The primary sequence of mature RGase A comprises 422 amino-acid residues, giving a molecular weight of 44 165 Da (Kofod *et al.*, 1994). Another enzyme acting on the same substrate is rhamnogalacturonase B (RGase B). It has been characterized (Schols, Posthumas & Voragen, 1990; Kofod *et al.*, 1994), but crystals have not yet been obtained. Rhamnogalacturonase B is a lyase, functionally different to RGase A. It degrades the same substrate by cleavage of the glycoside bond α-L-Rha(1–4)-α-D-GalUA, through β-elimination (Mutter, Coloquhoun, Schols, Beldman & Voragen, 1996). The natural substrate, rhamnogalacturonan, can be acetylated at the C2 or C3 position of GalUA (Schols *et al.*, 1990; Carpita & Gibeaut, 1993). Acetylation hampers the degradation catalyzed by RGase A and RGase B. Therefore the acetyl groups need to be removed, a reaction which is catalyzed by the enzyme rhamnogalacturonan acetyl esterase (Kauppinen *et al.*, 1995). The enzymatic degradation of pectin is of great industrial importance and enzyme mixtures including rhamnogalacturonases are used for different applications in the processing of fruits and vegetables. It has been shown that the degradation of modified hairy regions from apple, occur through synergetic action of a rhamnogalacturonase and rhamnogalacturonan acetyl esterase (Kauppinen *et al.*, 1995; Voragen, Schols & Beldman, 1992).

2. Methods and results

2.1. *Crystallization*

Heterologous expression of rhamnogalacturonase A from *Aspergillus aculeatus* in *Aspergillus oryzae*, has been described earlier (Kofod *et al.*, 1994). Crystallization conditions were screened by the sparse-matrix method (Jancarik & Kim, 1991). Hanging drops were equilibrated at room temperature by vapour diffusion. A droplet volume of 3 µl protein and 3 µl reservoir solution was mixed and equilibrated against 0.5 ml of a reservoir solution. Initially, a protein concentration of 22 OD_{280} was used. However, very few drops gave a precipitate indicating that the protein was very soluble, probably because of a high degree of glycosylation. A twofold increase in the protein concentration gave crystals in one of the drops after two weeks under the conditions 10% PEG 8000, 0.05 *M* KH_2PO_4 and 0.1 *M* sodium acetate buffered at pH 3.5. The crystals were suitable for X-ray diffraction experiments, but even larger crystals were obtained by increasing the droplet volume to 10 µl. Separate single crystals were very difficult to obtain as they grow either into a rosette of five to ten elongated crystals or into a stack of minor crystals attached to one major crystal as shown in Fig. 1. This growth habit made it difficult to select and mount a single crystal, however the problems were overcome by gentle removal of unwanted crystal growth. The average dimensions of the crystals were 0.2 × 0.2 × 0.5 mm.

2.2. *Molecular mass determination*

The molecular mass of the enzyme and the degree of glycosylation has been determined using a matrix-assisted laser desorption/ionization time-of-flight (MALDI–TOF) mass spectrometer (a 'Biflex' instrument from Bruker, Bremen, Germany). A 0.5 µl aliquot of a solution of RGase A at a

concentration of 3 OD_{280} in 0.1% F_3CCOOH and 30% CH_3CN was mixed with 0.5 µl matrix solution composed of 30 m*M* α-cyano-4-hydroxycinnamic acid in a CH_3CN/CH_3OH solution. Bovine serum albumin was used for calibration of the spectrometer. The molecular mass of the glycosylated enzyme was determined to 50.1 kDa (Fig. 2), which gives a carbohydrate amount of 5.9 kDa as the difference between the experimental mass and the mass calculated from the primary sequence of the enzyme.

2.3. *Data collection and processing*

Data collection was performed with an in-house Rigaku R-AXIS IIC image-plate system equipped with a RU-200 rotating-anode generator operating at 50 kV and 180 mA. A 0.5 mm collimator was employed and a graphite monochromator was used to select monochromatic Cu $K\alpha$ radiation. A native data set was collected at 288 K with a 2° oscillation per frame and crystal-to-detector distance of 100 mm. Each frame was exposed for 40 min and a total of 90 frames were recorded. Integrated intensities were obtained using the *HKL* package (Gewirth, 1994) and further data processing was performed using the *CCP*4 package (Collaborative Computational Project, Number 4, 1994). The diffraction symmetry and the systematically absent reflections, $h+k+l=2n+1$, showed that the space group can be $I2_12_12_1$ or $I222$. The unit-cell parameters are $a=62.9$, $b=125.4$ and $c=137.0$ Å. An experimental diffraction pattern is shown in Fig. 3. Though the diffraction pattern extended to 1.8 Å resolution the data analysis showed that measurements between 1.8 and 2.0 Å were of much lower quality than the rest of the data, therefore a cut-off was chosen at 2.0 Å. Some data-collection statistics are shown in Table 1. The solvent content has been determined, using Matthews formula (Matthews, 1968) with molecular mass of 50.1 KDa. This gives a volume to mass ratio V_m of

Table 1. *Data-collection statistics*

The rows are to be interpreted as follows: total number of reflections with a partiality above 0.5. Number of reflections failing the merging procedure. Unique number of reflections in the data set. Overall and outermost completeness is the percentage of obtained reflections compared to the theoretically obtainable measurements. Lowest and outermost resolutions are lowest resolution limit and highest resolution shell, respectively. Overall and outermost $I/\sigma(I)>2$ are the percentage of intensities higher than two standard deviations. R_{merge} is the internal R factor, $R_{merge}=\sum_{hkl}\sum|I(hk_iI)-\langle I(hkl)\rangle|/\sum_{hkl}\sum_i I(hkl)_i$. Multiplicity is the number of redundant measurements.

No. of reflections used	192450
No. rejected	427
No. of unique reflections	34076
Complete resolution range (Å)	30.0–2.0
Outermost resolution shell (Å)	2.1–2.0
Completeness overall (%)	92.7
Completeness outermost resolution shell (%)	76.7
Overall $I/\sigma(I)>2$ (%)	88.4
Outermost shell $I/(\sigma)>2$ (%)	75.2
Overall R_{merge} (%)	6.1
Outermost R_{merge} (%)	27.9
Multiplicity overall	5.6
Multiplicity outermost resolution shell	4.6

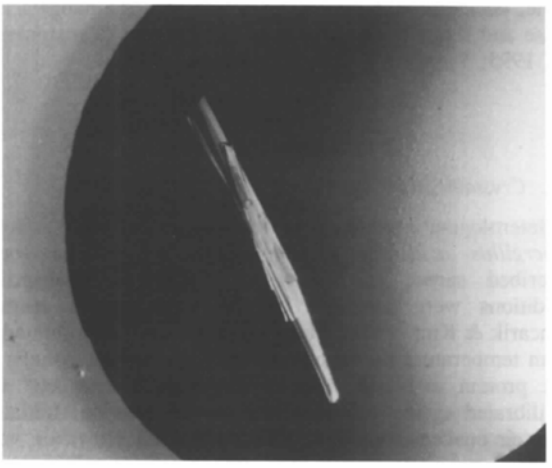

Fig. 1. Crystals of recombinant rhamnogalacturonase A from *Aspergillus aculeatus*. The maximum dimension of the crystal is *ca* 0.5 mm.

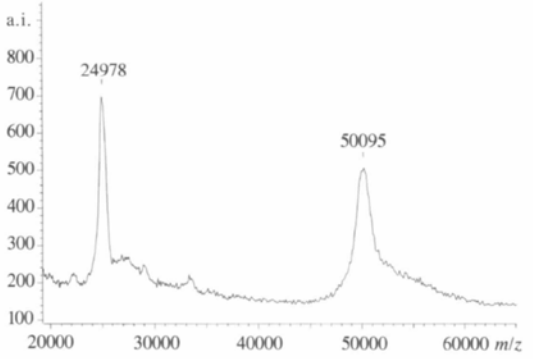

Fig. 2. A MALDI–TOF analysis of recombinant RGase A. The two peaks corresponds to molecular ions carrying two and one positive charges, respectively. The molecular mass of the enzyme is 50.1 ± 0.1 kDa.

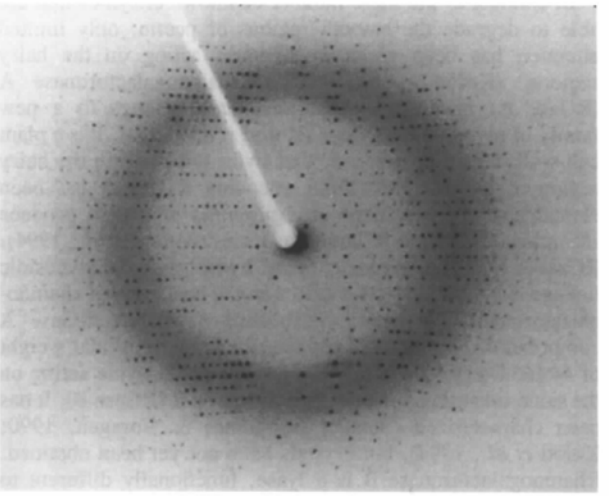

Fig. 3. An oscillation diffraction pattern from crystals of rhamnogalacturonase A. It shows the first frame recorder after 40 min exposure with a 2° oscillation of the protein crystal.

2.7 Å^3 Da^{-1} and a solvent content of 54%, assuming one molecule in the asymmetric unit. We have noted that all known protein structures with one molecule in the asymmetric unit and with the same space group ambiguity crystallize in space group $I222$ and not in $I2_12_12_1$ (Wukovitz & Yeates, 1995). Therefore, we assume $I222$ to be the correct space group in the subsequent structure determination.

We are grateful for the funding from the Danish National Research Foundation, which enables this research.

References

Carpita, N. C. & Gibeaut, D. M. (1993). *Plant J.* **3**, 1–30.

Collaborative Computational Project, Number 4 (1994). *Acta Cryst.* D**50**, 760–763.

Gewirth, D. (1994). *The HKL Manual: an Oscillation Data Processing Suite for Macromolecular Crystallography*. Yale University, New Haven, Connecticut, USA.

Jancarik, J. & Kim, S.-H. (1991). *J. Appl. Cryst.* **24**, 409–411.

Kauppinen, S., Christgau, S., Kofod, L. V., Halkier, T., Dörreich, K. & Dalbøge, H. (1995). *J. Biol. Chem.* **270**, 27172–27178.

Kofod, L. V., Kauppinen, S., Christgau, S., Anderson, L. N., Heldt-Hansen, H. P., Dörreich, K. & Dalbøge, H. (1994). *J. Biol. Chem.* **269**, 29182–29189.

Matthews, B. W. (1968). *J. Mol. Biol.* **33**, 491–497.

Mutter, M., Colquhoun, I. J., Schols, H. A., Beldman, G. & Voragen, A. G. J. (1996). *Plant Physiol.* **110**, 73–77.

Schols, H. A., Posthumus, M. A. & Voragen, A. G. J. (1990). *Carbohydr. Res.* **206**, 117–129.

Voragen, A. G. J., Schols, H. A. & Beldman, G. (1992). *Fruit Process.* **2**, 98–102.

Wukovitz, S. W. & Yeates, T. O. (1995). *Nature Struct. Biol.* **2**(12), 1062–1067.