Acta Cryst. (1998). D**54**, 1026–1029

Crystallization and preliminary X-ray diffraction studies of the heterogeneously glycosylated enzyme rhamnogalacturonan acetylesterase from *Aspergillus aculeatus*

ANNE MØLGAARD,[a] JENS F. W. PETERSEN,[a] SAKARI KAUPPINEN,[b] HENRIK DALBØGE,[b] ANDERS H. JOHNSEN,[c] JENS-CHRISTIAN NAVARRO POULSEN[a] AND SINE LARSEN[a]* at [a]*Centre for Crystallographic Studies, University of Copenhagen, Universitetsparken 5, DK-2100 Copenhagen, Denmark,* [b]*Screening Biotechnology, Novo Nordisk A/S, Novo Allé, DK-2880 Bagsvaerd, Denmark, and* [c]*Department of Clinical Biochemistry, National University Hospital, University of Copenhagen, Blegdamsvej 9, DK-2100 Copenhagen, Denmark. E-mail: sine@xray.ki.ku.dk*

(Received 19 November 1997; accepted 16 March 1998)

Abstract

Well diffracting crystals of rhamnogalacturonan acetylesterase from *Aspergillus aculeatus* have been obtained in two polymorphic modifications despite its heterogeneous glycosylation. The best-diffracting crystals (resolution 1.55 Å) are orthorhombic. The limit of the diffraction pattern of the other (trigonal) form is 2.5 Å. The ability of the enzyme to crystallize appears to depend on the glycosylation of the protein sample. This aspect has been investigated by mass spectrometry, which also showed that the orthorhombic crystals have the same glycosylation as the protein sample used in the crystallization.

1. Introduction

Pectin is a major component of the middle lamella and primary cell walls of higher plants, where it forms the matrix in which the cellulose–xyloglucan framework is embedded. Pectic polysaccharides comprise both smooth unbranched regions of homogalacturonan and ramified (hairy) regions of rhamnogalacturonan consisting of a backbone of alternating rhamnose (Rha) and galacturonic acid (GalUA) residues. Rhamnogalacturonan I is heavily substituted by neutral oligosaccharides such as arabinans, galactans and arabinogalactans at about half of the rhamnose residues. In addition, some of the galacturonic acid residues are methylated or acetylated.

The recent interest in identifying enzymes capable of degrading the rhamnogalacturonan part of pectin was triggered by the observation that the material clogging the ultrafiltration membrane in a final step in the production of fruit juices consisted mainly of ramified pectic substances (Beldman *et al.*, 1996). These could not be degraded by the pectinolytic enzymes that were active towards the unbranched homogalacturonan regions. Two enzymes capable of degrading the rhamnogalacturonan backbone were identified from *Aspergillus aculeatus*, rhamnogalacturonase A (Schols *et al.*, 1990), a hydrolase cleaving the glycosidic α-D-GalUA(1–2)-α-L-Rha bond and rhamnogalacturonase B, a lyase (Kofod *et al.*, 1994; Mutter *et al.*, 1996) which cleaves the α-L-Rha(1–4)-α-D-GalUA bond. A prerequisite for the action of these enzymes is that the acetyl groups have been removed from the substrate. The enzyme rhamnogalacturonan acetylesterase (RGAE) from *A. aculeatus* was recently identified, characterized, cloned and expressed heterologously in *Aspergillus oryzae* (Kauppinen *et al.*, 1995). It specifically deacetylates the backbone of rhamnogalacturonan, thus providing deacetylated substrate for further degradation by the rhamnogalacturonases.

The primary sequence of the mature protein of RGAE was found to comprise 233 amino-acid residues. The sequence of the unglycosylated form corresponds to a molecular mass of 24 607 Da, and contains two potential *N*-glycosylation sites. Mass spectrometry showed both the wild type and the recombinant enzyme to be heterogeneously glycosylated, and that the glycosylation of the recombinant protein (rRGAE) corresponded to a molar mass between 0 and 6 kDa (Kauppinen *et al.*, 1995).

Glycosylated proteins are frequently difficult to crystallize, due to heterogeneity with respect to either chemical composition and/or conformational flexibility (Baker *et al.*, 1994), therefore, problems with crystallization of rRGAE were anticipated. Nevertheless, high-quality crystals were grown from the first purification of the protein. However, it was not possible to grow crystals under these conditions using protein samples prepared from the same fermentation, but where a slightly different purification procedure of the enzyme had been employed. To investigate if there was any relation between these problems and the possible difference in the glycosylation of the protein samples, mass spectrometry was employed.

2. Methods and results

2.1. *Purification and crystallization*

Rhamnogalacturonan acetylesterase (RGAE) was cloned from *Aspergillus aculeatus*, expressed heterologously in *Aspergillus oryzae* and purified from the culture supernatant of an *A. oryzae* transformant as described previously (Kauppinen *et al.*, 1995). The enzyme was eluted from a Q-Sepharose anion-exchange column in a broad peak and in the first purification all fractions containing protein were pooled (purification 1). From this batch high-quality crystals were obtained in two different polymorphic modifications using Li_2SO_4, $(NH_4)_2SO_4$ and PEG as precipitants (Table 1).

The crystallization conditions were screened using the sparse-matrix screen method (Jancarik & Kim, 1991). After optimization, diffraction-quality crystals were obtained at room temperature in 1.4 *M* Li_2SO_4 or 1.4 *M* $(NH_4)_2SO_4$, 0.1 *M* Na acetate pH 5.0 with a protein concentration of 40 OD_{280}, using the hanging-drop vapour-diffusion method with 2 µl protein plus 2 µl reservoir solution in the drops equilibrated against 500 µl reservoir solution. The crystals grown from these conditions were orthorhombic. Trigonal crystals were obtained under the conditions 18% PEG 4000, 20% 2-propanol, 0.1 *M* citrate buffer, pH 4.5, using the same protein concentration. The trigonal crystals grow very slowly (over several months) and are not as reproducible as the orthorhombic crystals. In an attempt to obtain more reproducible

© 1998 International Union of Crystallography
Printed in Great Britain – all rights reserved

Table 1. *Comparison between the two different polymorphs of RGAE*

	Orthorhombic	Trigonal
Crystallization conditions	1.4 *M* Li_2SO_4 or 1.4 *M* $(NH_4)_2SO_4$, 0.1 *M* NaAc, pH 5.0, 4.0 OD_{280} RGAE	18% PEG 4000, 20% 2-propanol, 0.1 *M* citrate buffer, pH 4.5, 40 OD_{280}
	19% PEG 4000, 18% 2-propanol, 0.1 *M* citrate buffer, pH 4.7, 40 OD_{280} RGAE	
Cell dimensions (Å)	$a = 52.14$, $b = 56.87$, $c = 71.89$	$a = 75.36$, $c = 212.30$
Space group	$P2_12_12_1$	$P3_21$ or $P3_221$

conditions an optimization was carried out, and in one case a droplet with both crystal forms present was obtained at the conditions 19% PEG 4000, 18% 2-propanol, 0.1 *M* citrate buffer, pH 4.7, and 40 OD_{280} rRGAE (Fig. 1). A third polymorphic modification was obtained under conditions similar to the trigonal crystals, in the wider pH range 4.0–4.7. These crystals grow quickly, within a few days, and appear to grow as bundles of very thin needles. It has not been possible to obtain a single crystal of sufficient quality to characterize these crystals crystallographically.

A second purification of rRGAE was carried out, but instead of pooling all fractions containing protein from the Q-Sepharose column, only the fraction that was most pure with respect to rRGAE was used (pool 2a). Surprisingly, it was not possible to obtain the orthorhombic crystals using similar conditions as for the protein batch from purification 1. The remaining fractions from purification 2 were then pooled (pool 2b) but it was still not possible to obtain good-quality orthorhombic crystals under either the $(NH_4)_2SO_4$ or the Li_2SO_4 conditions. A third purification was undertaken, where the entire peak was pooled as in the first purification (purification 3). From this purification, crystals were obtained in a sparse-matrix screen test (Jancarik & Kim, 1991) under five different conditions and orthorhombic crystals of good diffraction quality were obtained using 20% PEG 4000, no salt and no buffer. This crystal form has a larger cell than in the orthorhombic crystal form obtained from the first purification, and the resolution limit of 2.2 Å is lower than for the orthorhombic crystals grown in Li_2SO_4 or $(NH_4)_2SO_4$ conditions. Orthorhombic crystals could also be obtained from pools 2a and 2b with PEG as a precipitant.

2.2. *Mass spectrometry*

Mass spectrometry was performed on samples from the different purifications to see whether there might be a difference in the glycosylation of rRGAE that could explain the difficulties in reproducing the initial crystal form. In order to establish whether a purification with respect to glycosylation might occur in the crystallization process, an orthorhombic crystal grown from the first purification of the protein was washed with the reservoir solution, dissolved and its mass spectrum recorded. Owing to the poor reproducibility of the trigonal crystals, we had no crystals of this modification available when the mass spectrometry experiments were performed. The measurements were performed in a matrix-assisted laser desorption/ionization time-of-flight (MALDI-TOF) mass spectrometer (a Biflex instrument from Bruker, Bremen, Germany) in the linear mode with a positive acceleration voltage of 30 kV. The instrument was calibrated using the singly and triply charged ions (MH^+ and MH_3^{3+}) of bovine serum albumin (M_w = 66 266 Da). The samples were prepared by mixing a 0.5 µl aliquot of a solution of rRGAE at a concentration of ~40 pmol µl^{-1} in a 20 m*M* solution of *n*-octylglucoside from Boehringer–Mannheim in a mixture of formic acid, water and acetonitrile (1:3:2 by volume) with 0.5 µl matrix solution composed of 30 m*M* α-cyano-4-hydroxycinnamic acid in a CH_3CN/CH_3OH solution.

Table 2. *Data collection statistics*

	Orthorhombic	Trigonal
Number of reflections used†	222320	124934
Number rejected‡	11361	5507
Number of unique reflections	28693	25036
Complete resolution range (Å)	30.0–1.55	30.0–2.49
Outermost resolution shell (Å)	1.63–1.55	2.62–2.49
Completeness overall (%)	91.0	98.8
Completeness, outermost shell (%)	77.9	93.0
Overall $I/\sigma(I) > 2$ (%)	91.4	87.9
Outermost shell $I/\sigma(I) > 2$ (%)	70.2	71.6
Overall R_{merge} (%)	3.9	8.7
R_{merge}, outermost shell (%)§	29.3	45.2
Multiplicity overall	7.7	5.0
Multiplicity, outermost shell	2.8	5.6

† Number of reflections with a partiality above 0.5. ‡ Number of reflections failing the merging procedure. § $R_{merge} = \sum_{hkl}\sum_i |(I_{hkl} - \langle I_{hkl}\rangle)| / \sum_{hkl}\sum_i \langle I_{hkl}\rangle_i$.

The mass spectra shown in Fig. 2 show that all the protein samples including the dissolved crystal are heterogeneously glycosylated. The measurements are unfortunately not of an accuracy that will enable us to make an exact identification of the differently glycosylated species that are present in the protein samples. There are, however, some distinct differences and similarities between the samples. The signals in the mass spectrum of the dissolved crystal (Fig. 2*a*) are significantly smaller than those recorded for the other samples, but they show the same distribution of peaks as the sample from which it was crystallized (Fig. 2*b*). The mass spectrum from the first purification shows that it contains different glycosylated species with masses that vary between 25.0 and 27.3 kDa with

Fig. 1. Crystals of rRGAE in two of its polymorphic modifications, the trigonal and the initial orthorhombic, grown under the conditions 19% PEG 4000, 18% 2-propanol, 0.1 *M* citrate buffer, pH 4.7 and 40 OD_{280} rRGAE. The length of the orthorhombic crystal is ~1.2 mm.

four major components, the two largest peaks being at 26.2 and 26.7 kDa. Within the experimental accuracy, they are identical to the two major peaks in the equivalent part of the spectrum recorded for the crystal (Fig. 2*a*), but the minor peaks also have a very similar distribution in Figs. 2(*a*) and 2(*b*). Similar agreement is observed for the doubly charged species. From this, we can conclude that the crystallization process unexpectedly did not lead to any increase in the homogeneity of the sample. The mass spectra of the two different pools of protein obtained from the second purification are shown in Figs. 2(*c*) and 2(*d*). The former shows that the protein sample representing the top of the peak contains glycoproteins with molar masses in a narrower range from 24.5 to 26.2 kDa. The purification procedure has apparently produced a more homogeneously glycosylated sample with lower molar masses, the two major peaks being at 25.0 and 25.5 kDa. This shows that with respect to the glycosylation, the second sample is distinctly different to the first. The mass spectrum in Fig. 2(*d*) was measured on the sample obtained by pooling the remaining fractions (pool 2b). This has a range of masses that is shifted upward compared to Fig. 2(*c*), but the prominent peaks at 26.2 and 26.7 kDa in the first purification do not have the same relative intensity in Fig. 2(*d*). It should be mentioned that numerous attempts were made to crystallize the samples with the mass spectra in Figs. 2(*c*) and 2(*d*) under similar conditions to those that were successful for the protein batch from purification 1, but they were all in vain. This indicates that the crystallization of rRGAE is influenced by both the homogeneity and the degree of glycosylation. The mass spectrum in Fig. 2(*e*) was recorded from the protein sample where the same purification procedure was used as for the sample in Fig. 2(*b*). In accordance, the two mass spectra are quite similar, but display small variations in the peak heights of the prominent peaks. In this case, orthorhombic crystals appeared under the same conditions where the other purifications produced the uncharacterized third crystal form.

2.3. *Data collection and reduction*

X-ray intensity data were collected in-house using a Rigaku R-AXIS IIC image-plate system, equipped with a RU-200 rotating-anode generator operating at 50 kV and 180 mA. Monochromatic Cu $K\alpha$ radiation was obtained using a graphite monochromator, and a 0.5 mm collimator was employed.

A native dataset for the orthorhombic crystal form from the first purification was collected at ambient temperature to a maximum resolution of 1.55 Å. The dimensions of the crystal used were 0.3 × 0.3 × 1.0 mm. A total of 170 frames were collected; the first 90 frames with an oscillation range of 2° per frame, $2\theta = 0°$ and a crystal-to-detector distance of 90 mm, and the last 80 frames with an oscillation range of 2° per frame, $2\theta = 24°$ and a crystal-to-detector distance of 98.5 mm. The first 90 frames were exposed for 15 min each and the last 80 frames for 20 min. No significant decomposition of the crystal was detected during the data collection.

The trigonal crystals diffract to a maximum resolution of 2.5 Å and a native dataset was collected at room temperature with a crystal-to-detector distance of 133 mm. The frames were collected with a 1° oscillation per frame and an exposure time of 30 min. A total of 116 frames were collected.

Data reduction was performed with the *HKL* package (Gewirth, 1994) and the *CCP4* package (Collaborative

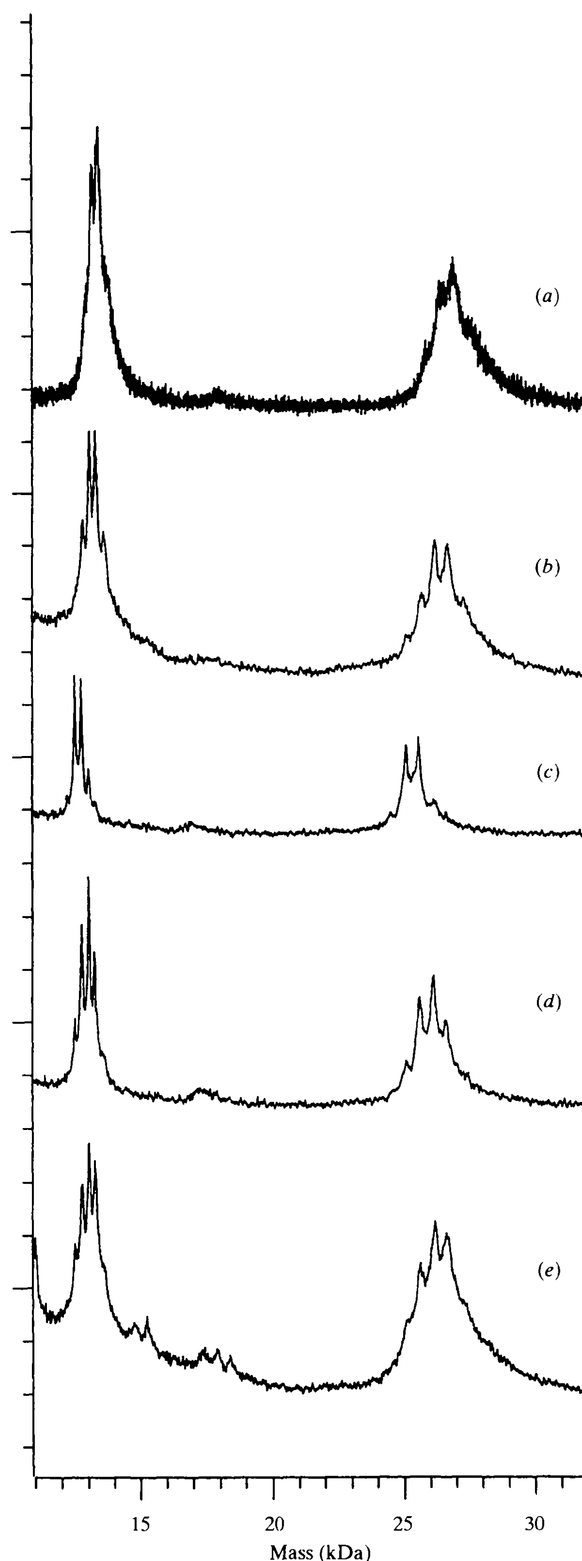

Fig. 2. MALDI-TOF mass spectra recorded for differently purified samples of rRGAE (see text). (*a*) An orthorhombic crystal grown from purification 1, (*b*) purification 1, (*c*) purification 2, pool 2a, (*d*) purification 2, pool 2b, (*e*) purification 3.

Computational Project, Number 4, 1994). The space group of the orthorhombic crystals was determined to be $P2_12_12_1$ from the systematically absent reflections, with $a = 52.14$, $b = 56.87$ and $c = 71.89$ Å. Using Matthews' formula (Matthews, 1968) this corresponds to a solvent content of ~39%, assuming one molecule in the asymmetric unit, and an average glycosylation of ~2 kDa.

From the diffraction symmetry and the systematically absent $000l = 2n + 1$ reflections, the space group of the trigonal form can be assigned as $P3_121$ or $P3_221$. Of the two, the former occurs almost twice as frequently as the latter in the Brookhaven Protein Data Bank (Wukovitz & Yeates, 1995). The cell parameters are $a = 75.36$ and $c = 212.30$ Å. This cell will accommodate two molecules in the asymmetric unit with a water content of ~62% or three molecules with an ~44% water content.

3. Concluding remarks

The difficulties described in crystallizing proteins which are heterogeneously glycosylated were not encountered in the first crystallization attempts of rhamnogalacturonan acetylesterase. It was also shown that the crystallization process did not lead to the expected separation of the differently glycosylated forms of the protein. Subsequent experiments showed that the crystallization conditions are strongly influenced by the homogeneity and the degree of glycosylation of the protein. To understand the complexity in the heterogeneous glycosylation of the protein it may be worthwhile to perform a tryptic degradation of the protein followed by a mass-spectrometric analysis of the products from the different samples. These experiments, combined with detailed structural analysis, should provide insight into the role glycosylation plays in the crystallization of glycoproteins.

It is known that different fermentations can lead to differences in the post-translational modifications, therefore, deglycosylation of the protein seems to be the only feasible route to obtain reproducible crystallization conditions. Only by this procedure can one reliably reduce the microheterogeneity of the protein with respect to chemical composition and conformational flexibility. However, deglycosylation is an alteration of the biologically active enzyme and the covalently bound carbohydrate may play a role in the surface accessibility and substrate specificity of the protein and for this reason we believe that it is worthwhile to attempt crystallization of the unaltered enzyme.

The authors would like to thank Dorthe Boelskifte for help with the crystallization experiments. The work was supported by a grant from the Danish National Research Foundation.

References

Baker, H. M., Day, C. L., Norris, G. E. & Baker, E. N. (1994). *Acta Cryst.* **D50**, 380–384.

Beldman, G., Mutter, M., Searle-van Leeuwen, M. J. F., van den Broek, L. A. M., Schols, H. A. & Voragen, A. G. J. (1996). *Progress in Biotechnology*, Vol. 14, *Pectins and Pectinases*, edited by J. Visser & A. G. J. Voragen, pp. 231–245. Amsterdam: Elsevier Science.

Collaborative Computational Project, Number 4 (1994). *Acta Cryst.* **D50**, 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.

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., Andersen, 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., Geraeds, C. C. J. M., Searle-van Leeuwen, M. F., Kormelink, F. J. M. & Voragen, A. G. J. (1990). *Carbohydr. Res.* **206**, 105–115.

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