Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

Crystallization and preliminary X-ray analysis of arabinanase A from *Pseudomonas fluorescens* subspecies *cellulosa*

M. Scott,[a] R. W. Pickersgill,[a] G. P. Hazlewood,[b] H. J. Gilbert[c] and G. W. Harris[a]*

[a]Institute of Food Research, Earley Gate, Whiteknights Road, Reading RG6 6BZ, England, [b]Laboratory of Molecular Enzymology, The Babraham Institute, Babraham, Cambridge CB2 4AT, England, and [c]Department of Biological Sciences, University of Newcastle-upon-Tyne, Newcastle-upon-Tyne NE1 7RU, England

Correspondence e-mail: gillian.harris@bbsrc.ac.uk

Received 16 March 1998
Accepted 4 September 1998

Crystals of 1,5-α-arabinanase A from *Pseudomonas fluorescens* subspecies *cellulosa* have been obtained by vapour diffusion. The crystals belong to the space group $P6_122$ with unit-cell parameters $a = b = 91.6$, $c = 179.4$ Å with one molecule in the asymmetric unit. The native crystals and, to a much greater extent, heavy-atom soaked crystals are sensitive to radiation which necessitates cryocooling. Suitable cryocooling conditions have been established, though a shrinkage of the unit cell is observed, with $a = b = 88.8$ and $c = 176.9$ Å.

1. Introduction

L-Arabinan is a common structural component of plant cell walls. It consists of a backbone of α-1,5-linked L-arabinofuranosyl units, some of which are substituted with α-L-arabinose side chains (Tanaka *et al.*, 1981). There are two classes of enzymes which are capable of hydrolysing L-arabinan: α-L-arabinofuranosidases (E.C. 3.2.1.55), which cleave the arabinose side chains, and endo-1,5-α-arabinanases (E.C. 3.2.1.99), which then attack the arabinan backbone (Tagawa & Terui, 1968; Dunkel & Amadò, 1994). Recently an arabinofuranosidase from *Aspergillus niger* strain 3M43 has been crystallized (Scott *et al.*, 1997), but to date no crystallization of an arabinanase has been reported. Here, we report the crystallization and crystallographic characterization of 1,5-α-L-arabinanase A from *P. fluorescens* subspecies *cellulosa* (arabinanase A).

Arabinanase A consists of a single catalytic domain. This non-modular structure is in common with mannanase A from *P. fluorescens* subspecies *cellulosa* (Braithwaite *et al.*, 1995) which has recently been crystallized (Scott *et al.*, 1998), but is in contrast to other glycosyl hydrolases from *P. fluorescens* subspecies *cellulosa* which are modular in structure and consist of a catalytic domain linked by serine-rich sequences to one or more non-catalytic domains (Hazlewood & Gilbert, 1992).

Genes encoding three 1,5-α-L-arabinanases have been cloned and sequenced to date. Whereas both the 1,5-α-L-arabinases from *A. niger* (Flipphi *et al.*, 1993) and from *Bacillus subtilis* (Sakamoto *et al.*, 1997) are endo-enzymes, hydrolysing the 1,5-α-glycosidic bonds in a random fashion, 1,5-α-L-arabinanase A from *P. fluorescens* subspecies *cellulosa* (McKie *et al.*, 1997) exhibits both an endo- and an exo- mode of action, as this enzyme is also capable of successively releasing arabinotriose units from the non-reducing end of an arabinan chain.

Glycosyl hydrolases have been classified into families on the basis of amino-acid sequence similarity (Henrissat, 1991). Comparison of the primary sequences of arabinanase A and other proteins in the SWISS-PROT database shows that arabinanase A displays the greatest sequence identity with 1,5-α-L-arabinase from *A. niger* (McKie *et al.*, 1997), placing it in glycosyl hydrolase family 43 (Henrissat & Bairoch, 1993, 1996). As there are to date no known structures for glycosyl hydrolases belonging to family 43, the crystal structure of arabinanase A will thus reveal the protein fold for this family of glycosyl hydrolases.

Family 43 enzymes are known to cleave the substrate with an inversion of anomeric configuration. The active-site residues have not been identified, but would be expected to be a nucleophile/base and a proton donor separated by a distance of about 10 Å, as opposed to the 5 Å separation expected in the case of enzymes hydrolysing by means of a retaining mechanism. In conjunction with this biochemical information, the crystal structure of arabinanase A will enable the identification of the catalytic residues, facilitating further biochemical and mutagenesis studies.

2. Experimental

2.1. Crystallization

The expression, purification and characterization of arabinanase A has been published elsewhere (McKie *et al.*, 1997); arabinanase A

purified from a recombinant strain of *Escherichia coli* had an M_r of 34000.

The purified enzyme was dialysed into water containing 2 m*M* sodium azide (NaN_3) and was concentrated in a 3 ml stirred-cell concentrator (Amicon) fitted with a 10 kDa membrane (Flowgen). Initial crystallization trials (protein concentration of 12 mg ml^{-1}) were set up using the Hampton Research Crystal Screens I and II and the hanging-drop vapour-diffusion technique in Linbro plates over 1 ml reservoirs, allowing many different crystallization conditions to be surveyed (Jancarik & Kim, 1991). After about a week at 291 K, crystals grew in droplets of the screen where polyethylene glycol (PEG) was the precipitant. After optimization of these conditions, the best crystals were obtained from drops containing 4 µl of protein solution and 2 µl of reservoir consisting of 9% PEG 8000 or 6000, 0.1 *M* HEPES pH 7.50 and 5% MPD. The crystals were rod-like with hexagonal faces (Fig. 1).

2.2. Data collection (room temperature)

A native data set was collected from a crystal of arabinanase A at room temperature (293 K) at the LURE-DCI synchrotron (Orsay, France) station D41 operating at a wavelength of 1.386 Å and using an imaging-plate system with a diameter of 180 mm (MAR Research, Hamburg). A total of 60 frames of 1° oscillations were measured using a crystal-to-detector distance of 170 mm. The crystal diffracted to a resolution of 2.9 Å at the start; however, the crystals proved to be susceptible to radiation damage and the resolution limit had fallen to at most 3.3 Å by the end of the collection.

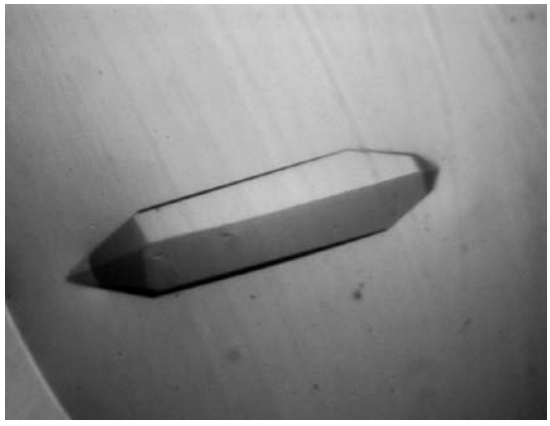

Figure 1
Hexagonal crystal of arabinanase A from *P. fluorescens* subspecies *cellulosa* grown in a hanging drop using polyethylene glycol as the precipitant.

Table 1
Statistics for merged and scaled native arabinanase data sets in space group $P6_122$ at 3.5 Å resolution, compared with data sets merged in space group $P61$ at 3.5 Å resolution.

R linear = $\sum|(I-\langle I\rangle)|/\sum I$; R square = $\sum(I-\langle I\rangle)^2/\sum I^2$. In all sums single measurements are excluded.

	Room-temperature data		Cryocooled data	
Space group	$P6_122$	$P6_1$	$P6_122$	$P6_1$
R linear (overall)	0.070	0.067	0.143	0.138
R square (overall)	0.064	0.062	0.122	0.117
Completeness (%)	99.8	99.9	99.8	99.9
Number of reflections	60877		105319	
Number of independent reflections	6107	10770	5718	10054
Number of independent reflections measured more than once	6060	10674	5690	10025
Reflections with 1–6 observations (%)	33.1		19.2	
Reflections with 7–8 observations (%)	54.2		10.1	
Reflections with 9–19 observations (%)	12.5		70.5	
$\langle I/\sigma(I)\rangle$	15.7		6.7	
Mosaicity	0.13		0.46	
R linear (3.72–3.50 Å)	0.133		0.280	
R square (3.72–3.50 Å)	0.115		0.240	

2.3. Cryocooling and data collection

A crystal of native arabinanase A was transferred to a solution of 20% PEG 8000, 0.1 *M* HEPES pH 7.50 and 5% MPD with 20% glycerol and 10% 2-propanol. The additional PEG, glycerol and 2-propanol acted as cryoprotectant allowing the crystals to be successfully cryocooled. The data were collected at 100 K using an Oxford Cryosystems Cryostream cryocooling system. Individual crystals were scooped up in mohair cryo-loops and flash-frozen in the 100 K gaseous nitrogen stream. In this way, a cryocooled native data set was collected from a crystal of arabinanase A at the EMBL X31 beamline at the DORIS storage ring, DESY, Hamburg, operating at a wavelength of 1.00 Å and using an imaging-plate system with a diameter of 180 mm (MAR Research, Hamburg). The crystal diffracted to a resolution limit of 3.2 Å. A total of 100 frames of 1° oscillations were measured using a crystal-to-detector distance of 275 mm. In contrast to the room-temperature data collection, no deterioration of the crystal was observed during data collection.

2.4. Data processing

Both data sets were processed with *DENZO* from the *HKL* suite of programs (Otwinowski, 1993). *DENZO*, in conjunction with the display program *XDISPLAYF* (Minor, 1993), was also used for autoindexing. Crystals of arabinanase A have a primitive hexagonal lattice. The data were scaled and merged using the program *SCALEPACK* from the *HKL* suite (Otwinowski, 1993). Details of the statistics for the two data sets are given in Table 1.

3. Results and discussion

Inspection of the systematic absences for prospective primitive hexagonal space groups identified the space group as either $P6_1$ (or $P6_5$) or $P6_122$ (or $P6_522$) on the basis of the reflection condition (0, 0, $l = 6n$). The data were merged and scaled in both $P6_1$ and $P6_122$ (Table 1). Since the improvement in the R factor observed when merging together less equivalent reflections in space group $P6_1$ was not significant, we conclude that the space group is $P6_122$ (or $P6_522$).

The unit-cell parameters for the room-temperature crystal were $a = b = 91.6$ and $c = 179.4$ Å. A shrinkage of the unit cell is observed for the cryocooled crystal, with unit-cell parameters $a = b = 88.8$ and $c = 176.9$ Å. Both data sets were scaled and merged in space group $P6_122$ (Table 1). Assuming one molecule in the asymmetric unit, the values of the crystal packing parameter V_m for the room-temperature and cryocooled crystals are 3.20 and 2.96 Å^3 Da^{-1}, respectively, which are within the range commonly observed for protein crystals (Matthews, 1968).

When data were collected from an arabinanase crystal at room temperature, deterioration of the crystal was observed, with a resultant fall-off of resolution limit and data quality. This was reflected in an increase in the R factor (for a resolution limit of 3.5 Å) for the individual images from

around 0.06 at the start to around 0.10 towards the end of the 60 frames collected. Owing to the non-uniform fall-off in resolution over the data set and the resultant deterioration of overall data quality, the data were cut to a resolution limit of 3.5 Å. Merging statistics are given in Table 1.

The deterioration of arabinanase crystals when exposed to synchrotron radiation was found to be even more severe for crystals which had been soaked in various heavy-metal reagents, and it was not possible to collect 60° of data from such crystals at room temperature, resulting in very poor quality and incomplete data being obtained. Hence, cryocooling of the arabinanase crystals is considered essential for the success of the project.

Cryocooling conditions have been established for arabinanase crystals and a data set has been collected from a cryocooled native arbinanase crystal. A shrinkage of the unit cell of about 2.5 Å in all directions is observed for the cryocooled crystal. Merging statistics for the data cut to a resolution limit of 3.5 Å are given in Table 1. With cryocooling, the arabinanase crystal does not show any appreciable deterioration during data collection and the *R* factor for individual images remains essentially constant throughout the 100 frames of data collected. The mosaicity of the arabinanase crystal does increase with cryocooling, with a post-refinement value of 0.46. This suggests that the cryocooling conditions may not be optimal.

The merging *R* factor for the data from the cryocooled crystal is found to be higher than that for the data from the room-temperature crystal. The diffraction from the cryocooled arabinanase crystal is relatively weak, but the two arabinanase data sets are not directly comparable as the crystals were obtained from different protein/crystallization batches (slightly different crystallization conditions) and the data were collected using different data-collection strategies. Although both the arabinanase A native data sets are 99.8% complete, the redundancy is much higher for the cryocooled data set than for the room-temperature data set (see Table 1). As a result, the merging statistics might be expected to show some bias in favour of the room-temperature data set with lower redundancy. However, the higher merging *R* factor and increased mosaicity for the cryocooled data set point towards the cryocooling conditions not yet being optimal. Future work will seek to fine-tune the cryocooling conditions in order to reduce the mosaicity and improve the data quality if possible.

Given the rapid decay of arabinanase crystals soaked in a variety of heavy-metal reagents in the synchrotron beam at room temperature and the shrinkage of the unit cell of arabinanase crystals on cryocooling, a native data set from cryocooled arabinanase crystals was essential for the project to proceed, even though the cryocooled data set is not as good quality as the room-temperature data set. Heavy-metal derivatives of crystals of arabinanase will be collected under cryocooling conditions to enable the structure to be solved. If it does not prove possible to improve the quality of the cryocooled native data set, the final structure will be refined against the higher quality room-temperature native data set.

There has been a resurgence of interest in the structure and mechanism of glycosyl hydrolases in recent years. An analysis of the information available from the structures representative of 22 of the 60 or so glycosyl hydrolase families (Davies & Henrissat, 1995) has shown that the substrate specificity and mode of action of these enzymes is governed by their detailed three-dimensional structures. The structure of arabinanase A will provide insight into structure–function relationships of arabinanase and other family 43 glycosyl hydrolases, as well as the differences and similarities to other families of glycosyl hydrolases.

We thank LURE (Orsay, France) and EMBL/DESY, Hamburg for the provision of beamtime, Professor R. Fourme (LURE) and Dr Paul Tucker (EMBL/DESY, Hamburg) for advice and assistance with station operation and Dr J. Jenkins for assistance with data collection. We thank the European Union for support of the work at EMBL Hamburg through the HCMP Access to Large Installations Project, Contract Number CHGE-CT93-0040. MS, RWP and GWH are supported by the BBSRC.

References

Braithwaite, K. L., Black, G. W., Hazlewood, G. P., Ali, B. R. S. & Gilbert, H. J. (1995). *Biochem. J.* **305**, 1005–1010.
Davies, G. & Henrissat, B. (1995). *Structure*, **3**, 853–859.
Dunkel, M. P. H. & Amadò, R. (1994). *Carbohydr. Polym.* **24**, 247–263.
Flipphi, M. J. A., Panneman, H., van der Veen, P., Visser, J. & de Graaff, L. H. (1993). *Appl. Microbiol. Biotechnol.* **40**, 318–326.
Hazlewood, G. P. & Gilbert, H. J. (1992). *Xylans and Xylanases*, edited by J. Visser, pp. 259–273. Amsterdam: Elsevier Science.
Henrissat, B. (1991). *Biochem. J.* **280**, 309–316.
Henrissat, B. & Bairoch, A. (1993). *Biochem. J.* **293**, 781–788.
Henrissat, B. & Bairoch, A. (1996). *Biochem. J.* **316**, 695–696.
Jancarik, J. & Kim, S.-H. (1991). *J. Appl. Cryst.* **24**, 409–411.
McKie, V. A., Black, G. W., Millward-Sadler, S. J., Hazlewood, G. P., Laurie, J. I. & Gilbert, H. J. (1997). *Biochem. J.* **323**, 547–555.
Matthews, B. W. (1968). *J. Mol. Biol.* **33**, 491–497.
Minor, W. (1993). *XDISPLAYF* Program, Purdue University, West Lafayette, Indiana, USA.
Otwinowski, Z. (1993). *Data Collection and Processing. Proceedings of the CCP4 Study Weekend*, edited by L. Sawyer, N. Isaacs & S. Bailey, pp. 56–62. Warrington: Daresbury Laboratory.
Sakamoto, T., Yamada, M., Kawasaki, H. & Sakai, T. (1997). *Eur. J. Biochem.* **245**, 708–714.
Scott, M., Connerton, I. F., Harris, G. W., Gravesen, T. N., Madrid, S. M. & Mikkelsen, J. D. (1997). *Acta Cryst.* D**53**, 222–223.
Scott, M., Pickersgill, R. W., Hazlewood, G. P., Bolam, D., Gilbert, H. J. & Harris, G. W. (1998). *Acta Cryst.* D**54**, 129–131.
Tagawa, K. & Terui, G. (1968). *J. Ferment. Technol.* **46**, 693–700.
Tanaka, M., Abe, A. & Uchida, T. (1981). *Biochim. Biophys. Acta*, **658**, 377–386.