Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

A stepwise optimization of crystals of rhamnogalacturonan lyase from *Aspergillus aculeatus*

Renuka Kadirvelraj,[a] Pernille Harris,[a] Jens-Christian N. Poulsen,[a] Sakari Kauppinen[b] and Sine Larsen[a]*

[a]Centre for Crystallographic Studies, University of Copenhagen, Universitetsparken 5, DK 2100 Copenhagen, Denmark, and [b]Novozymes A/S, Novo Allé, DK2880 Bagsværd, Denmark

Correspondence e-mail: sine@ccs.ki.ku.dk

Received 10 April 2002
Accepted 20 May 2002

Recombinant rhamnogalacturonan lyase from *Aspergillus aculeatus* has been crystallized by a stepwise procedure and X-ray diffraction data have been collected. The crystals were grown using hanging-drop vapour-diffusion and microseeding techniques. Crystals were obtained showing a flat plate morphology. The crystallization conditions were 20% PEG 4000, 9% PEG 400, 0.1 *M* $(NH_4)_2SO_4$ and 0.1 *M* sodium acetate pH 4.4. These crystals diffracted to a resolution of 1.5 Å. The unit-cell parameters are $a = b = 77.0$, $c = 170.8$ Å with the possible space group $P4_32_12$ or $P4_12_12$. There is most likely to be one molecule in the asymmetric unit, leading to a calculated solvent content of approximately 47% for the crystals.

1. Introduction

The primary cell walls of land plants form a single continuous extracellular matrix throughout the body of the plant, consisting of different types of polysaccharides, proteins and lignins in varying amounts. The polysaccharide composition of the matrix varies according to the species, tissue and age of the plant, but the main components are cellulose, xyloglucan and pectin. Pectin constitutes an independent cementing matrix prevalent in fruits. It is the most complex of the cell-wall polysaccharides, consisting of covalently linked regions of the acidic polymers homogalacturonan and rhamnogalacturonan (Schols *et al.*, 1990). Homogalacturonan is an α(1–4)-linked unbranched (smooth) polysaccharide of D-galacturonic acid residues (GalUA) and rhamnogalacturonan is a branched (hairy) polysaccharide composed of subgroups rhamnogalacturonan I and II (RG-I and RG-II). The backbone of the main component in rhamnogalacturonan, RG-I, is made up of alternating rhamnose (Rha) and galacturonic acid residues (GalUA) with the dimer repeat unit (1,2)-α-L-Rha-(1,4)-α-D-GalUA; about half of the rhamnose residues are substituted at O4 with side chains of neutral sugars such as galactan, arabinan and arabinogalactan (O'Neill *et al.*, 1990; Carpita & Gibeaut, 1993), which explains why RG-1 is referred to as the ramified or hairy part of pectin. The less abundant RG-II has a polygalacturonic acid backbone that is substituted with rhamnose and rare sugars such as 2-*O*-methyl fucosyl and apiosyl (Whitcombe *et al.*, 1995).

Considering the complex composition of pectin, it is understandable that a variety of enzymes are required for its degradation. The homogalacturonan region is degraded by a synergistic action of pectin and pectate lyases, pectin methyl esterases and pectin acetyl esterases (Shevchik & Hugouvieux-Cotte-Pattat, 1997). In a similar way, rhamnogalacturonan is synergistically degraded by rhamnogalacturonan acetyl esterase (RGAE), rhamnogalacturonase A (RG-hydrolase; Kauppinen *et al.*, 1995) and rhamnogalacturonan lyase (RG-lyase; Mutter *et al.*, 1996). Naturally occurring rhamnogalacturonan can be acetylated at the C2 and/or C3 position of GalUA (Schols *et al.*, 1990; Carpita & Gibeaut, 1993). Acetylation prevents the degradation catalysed by RG-hydrolase and RG-lyase by sterically hindering the action of the enzymes. The deacetylation is catalyzed by RGAE, thus paving the way for the lyases and hydrolases to act on and degrade the backbone of RG-I (Kauppinen *et al.*, 1995). Of the enzymes that degrade RG-I, only RG-hydrolase and RGAE from *Aspergillus aculeatus* have been structurally characterized to date (Petersen *et al.*, 1997; Mølgaard *et al.*, 2000).

RG-lyase from *A. aculeatus* has been purified, characterized and overexpressed in *A. oryzae* (Schols *et al.*, 1990; Kofod *et al.*, 1994). RG-hydrolase and RG-lyase act on different glycosidic linkages in the RG-I backbone. RG-hydrolase is an endo-acting hydrolase, hydrolysing the glycosidic linkage α-D-GalUA-(1,2)-α-L-Rha in the RG-I backbone and belongs to the CAZY family 28 of glycosyl hydrolases (Coutinho & Henrissat, 1999). RG-lyase cleaves the linkage α-L-Rha-(1,4)-α-D-GalUA by a β-elimination process, introducing a double bond at the non-reducing end of the GalUA unit in RG-I (Fig. 1). The

smallest RG-I substrate hydrolyzed by RG-hydrolase is a nonamer and the smallest entity cleaved by RG-lyase is a decamer of the substrate (Mutter *et al.*, 1996, 1998).

The crystallization and preliminary X-ray characterization of RG-lyase from *A. aculeatus* is presented in this paper. The crystallization required a systematic step-by-step approach in order to proceed from the initial dendritic growth of thin needles through spherulites and layered plates to well formed single crystals of high diffraction quality (Fig. 2). The temperature optimum for the enzyme is 323 K and its pH optimum is 6.0, close to its pI of 5.2 (Kofod *et al.*, 1994). The mature enzyme is comprised of 508 amino-acid residues, corresponding to a molecular mass of 54 202.9 Da. It shares 21% sequence identity with RG-hydrolase but shows no immunological cross-reactivity (Kofod *et al.*, 1994). A mass-spectrometric analysis of the purified recombinant RG-lyase showed a very narrow mass distribution around a molecular weight of 54.242 kDa; when compared with the calculated weight of the mature enzyme this indicates that the RG-lyase, in contrast to the RG-hydrolase (Petersen *et al.*, 1997), is unglycosylated. The crystal structure of RG-lyase will be the first crystal structure of a family 4 polysaccharide lyase (Coutinho & Henrissat, 1999). It is expected to provide new insights into the substrate binding and catalytic activity of rhamnogalacturonan-degrading enzymes. So far no crystal structure is available in this family, which to date includes a putative rhamnogalacturonan lyase from *Neurospora crassa*; seven open reading frames from the *Arabidopsis thaliana* genome project have also been assigned to this family.

2. Materials and methods

2.1. Purification and crystallization

RG-lyase from *A. aculeatus* was purified from *A. oryzae* transformant producing RG-lyase. In our hands, RG-lyase was significantly contaminated with TAKA-amylase when purified as described by Kofod *et al.* (1994) and therefore a TAKA-amylase-deficient strain of *A. oryzae* was used (JaL228; J. Lehmeck, Novozymes A/S) for heterologous expression of RG-lyase. The subsequent purification of RG-lyase starting from culture filtrate was performed essentially as described previously (Kofod *et al.*, 1994). The purified recombinant RG-lyase was more than 95% pure as estimated from SDS–PAGE and native IEF gels. Crystallization conditions were initially screened using the sparse-matrix method (Jancarik & Kim, 1991) and crystalline precipitates were obtained from conditions containing polyethylene glycol as the precipitating agent. Incomplete factorials (Carter & Carter, 1979) were designed and set up around these conditions. The hanging-drop vapour-diffusion method was used with 2 µl of protein mixed with 2 µl of reservoir solution in drops equilibrated against 500 µl of reservoir solution at room temperature. This resulted in thin needle clusters and dendrites (Fig. 2*a*) from the condition 26% PEG 4000, 0.3 *M* $(NH_4)_2SO_4$, 0.1 *M* sodium acetate pH 4.5 with a protein concentration of 17 mg ml^{-1}. A small change in the conditions to 24% PEG 4000, 0.1 *M* $(NH_4)_2SO_4$, 0.1 *M* sodium acetate pH 4.8 resulted in the formation of spherulites and crystalline precipitate (Fig. 2*b*). At this stage, trial screens were set up at room temperature, 283 and 277 K. Layered plates along with

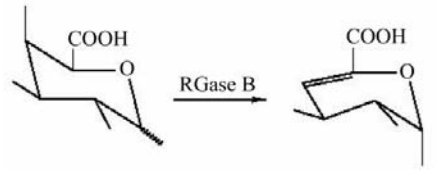

Figure 1
RG-lyase mode of action leading to the formation of a double bond in the GalUA unit in the non-reducing end of RG-I.

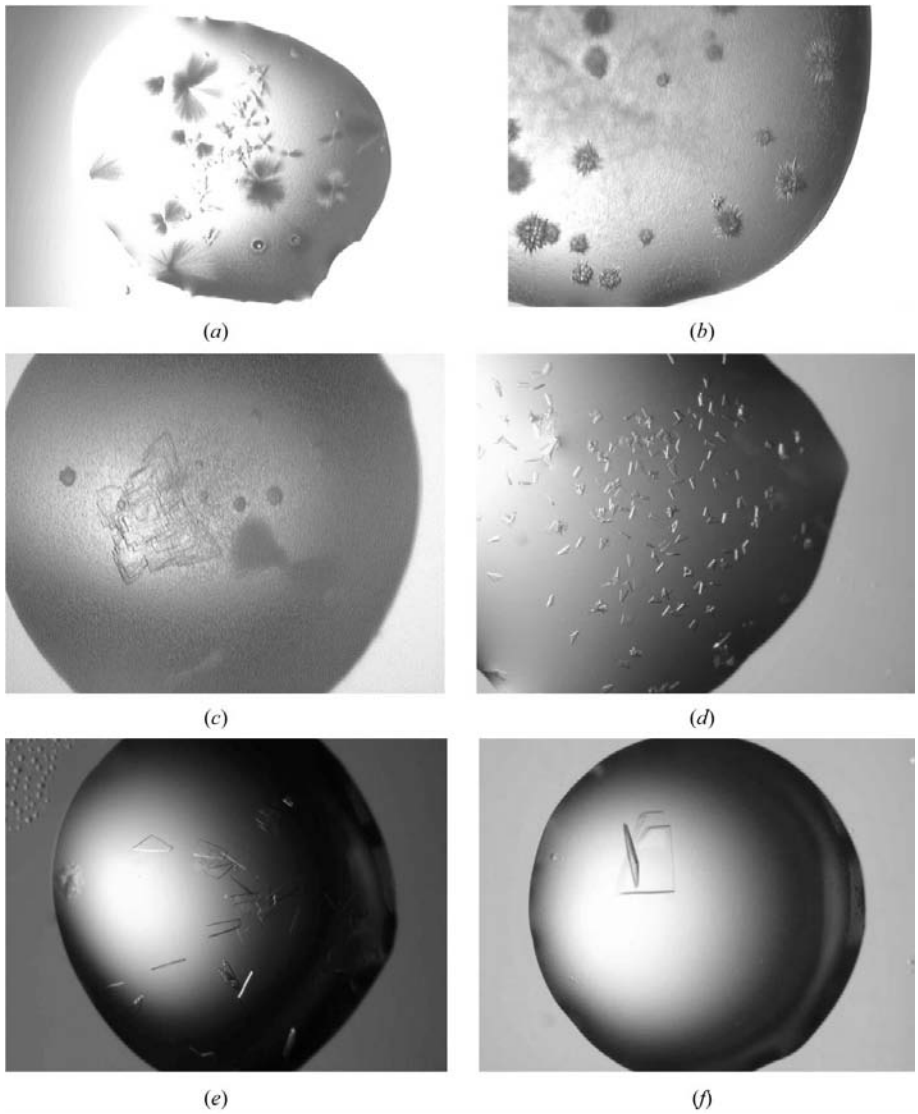

Figure 2
Photographs of RG-lyase crystals. (*a*) Needle clusters, dendrites of RG-lyase from 26% PEG 4000, 0.3 *M* $(NH_4)_2SO_4$, 0.1 *M* sodium acetate pH 4.5. (*b*) Spherulites from 24% PEG 4000, 0.1 *M* $(NH_4)_2SO_4$, 0.1 *M* sodium acetate pH 4.8. (*c*) Stacked plates from 24% PEG 4000, 0.1 *M* $(NH_4)_2SO_4$ and 0.1 *M* sodium acetate pH 4.4 at 277 K. (*d*) Single crystals from microseeding using a tenfold diluted seed stock with the condition 22% PEG 4000, 0.1 *M* $(NH_4)_2SO_4$ and 0.1 *M* sodium acetate pH 4.4 at 277 K. (*e*) Larger crystals obtained by further dilution of seed stock, protein and precipitant concentration. (*f*) Single crystals from the condition 20% PEG 4000, 9% PEG 400, 0.1 *M* $(NH_4)_2SO_4$ and 0.1 *M* sodium acetate pH 4.4 at 277 K.

spherulites were obtained from 24% PEG 4000, 0.1 *M* $(NH_4)_2SO_4$ and 0.1 *M* sodium acetate pH 4.4 at 277 K (Fig. 2*c*). The plates were transferred to a seed bead kit (Hampton Research HR2-320) along with 50 µl of reservoir solution and crushed by rapid vortexing to generate a seed stock. This concentrated seed stock was serially diluted, stored at 277 K and used to micro-seed pre-equilibrated drops containing 1 µl of protein mixed with 2 µl of reservoir solution. The protein concentration used was 15 mg ml^{-1} and the crystallization conditions were 22% PEG 4000, 0.1 *M* $(NH_4)_2SO_4$ and 0.1 *M* sodium acetate pH 4.4 at 277 K. This resulted in the formation of plate-like single crystals of average dimensions 0.1 × 0.15 × ∼0.03 mm (Fig. 2*d*). Larger crystals of the protein were obtained by increasing the droplet volume to 6 µl, decreasing the PEG 4000 concentration to 20% and varying the number of seeds introduced into the drop by diluting the seed stock to optimum dilution (Fig. 2*e*). Dilution and using fresh seed stock further helped to avoid the problem of five or more crystals growing together in a rosette or in a stack of layered plates (Fig. 2*f*).

3. Preliminary X-ray characterization

3.1. Native RG-lyase

Though the crystals obtained were of good diffraction quality at room temperature, they were fragile and unstable in the X-ray beam at room temperature. A series of cryoprotectants were prepared using different concentrations of glycerol, ethylene glycol, paratone-N, 2-methyl-2,4-pentanediol and sucrose. In all cases, the crystal mosaicity was high (>0.9°) and reflections beyond ∼2.5 Å were diffuse. The best cryoprotection was finally obtained by using a reservoir solution with 15% ethylene glycol. Preliminary X-ray data were collected to a resolution of 2.5 Å with an in-house MAR 345 image-plate system equipped with a Cu $K\alpha$ rotating-anode generator operating at 50 kV and 100 mA. The diffraction patterns were consistent with 4/*mmm* Laue symmetry and the systematic absences indicated that the space group was either $P4_12_12$ or $P4_32_12$.

The diffraction spots at the higher resolution limits were smeared owing to the high mosaicity of the crystals (>0.9°). This problem was tackled by converting the reservoir solution into a cryoprotectant by adding different amounts of low-molecular-weight PEG solutions to PEG 4000. Well diffracting crystals were obtained at 277 K by microseeding 24 h pre-equilibrated drops formed by 2 µl protein solution mixed with 6 µl reservoir solution containing 20% PEG 4000, 9% PEG 400, 0.1 *M* $(NH_4)_2SO_4$ and 0.1 *M* sodium acetate at pH 4.4 with a protein concentration of 15 mg ml^{-1} (Fig. 2*f*). The mosaicity of these crystals with average dimensions 0.2 × 0.25 × 0.05 mm was in the range 0.3–0.5°. The crystals diffracted to a resolution of 2 Å at the in-house facility.

Table 1
Data-collection and processing statistics for native RG-lyase crystals.

Values in square brackets are for the highest resolution shells.

Derivative	Native	Orange platinum	Mercury acetate
Beamline	ID14-4, ESRF Grenoble	In-house	In-house
Detector	ADSC Quantum CCD	MAR 345 image plate	MAR 345 image plate
Space group	$P4_12_12$ or $P4_32_12$	$P4_12_12$ or $P4_32_12$	$P4_12_12$ or $P4_32_12$
Unit-cell parameters (Å)			
$a = b$	77.00 (0.2)	77.11 (22)	77.35 (24)
c	170.81 (0.23)	170.51 (66)	171.5 (12)
Wavelength (Å)	0.9393	Cu $K\alpha$	Cu $K\alpha$
Resolution range (Å)	19.92–1.5 [1.55–1.50]	25.00–3.00 [3.11–3.00]	25.00–2.65 [2.74–2.65]
No. of reflections	1437589	121575	262943
No. of unique reflections	83086	10997	15846
Multiplicity	5–6	11	16
$\langle I/\sigma(I)\rangle$	16.71 [4.5]	16.4 [6.6]	50.5 [14.2]
Mosaicity (°)	0.4	0.8	0.5
Oscillation per image (°)	0.5	0.5	1.0
Completeness (%)	100.0 [100.0]	100.0 [100.0]	99.9 [99.5]
$R_{\rm merge}$† (%)	0.108 [0.292]	0.176 [0.361]	0.097 [0.241]

† $R_{\rm merge} = \sum_{hkl}\sum_i |I(hkl)_i - \langle I(hkl)\rangle|/\sum_{hkl}\sum_i I(hkl)_i$.

Native RG-lyase synchrotron data were collected to a maximum resolution of 1.5 Å at the ID14-4 beamline at ESRF, Grenoble (Fig. 3). Owing to the flat plate morphology of the crystals, the long *c* axis gave problems with overlapping reflections. This was avoided by bending the loop 45° prior to mounting the crystal. Data were indexed, integrated and scaled using the *HKL* suite (Gewirth, 1994; Otwinowski & Minor, 1997) and further processing was performed with the *CCP*4 package (Collaborative Computational Project, Number 4, 1994). A summary of the data-collection and processing statistics is given in Table 1. The solvent content was estimated using the Matthews formula (Matthews, 1968) assuming the molar mass of the unglycosylated enzyme and gave a volume-to-mass ratio $V_{\rm M}$ of 2.34 Å^3 Da^{-1} and a solvent content of 47%.

3.2. Derivative RG-lyase

RG-lyase has no significant sequence homology with any protein of known three-dimensional structure and therefore a search for heavy-atom derivatives was carried out. This was hampered by the fragile nature of the crystals, which tended to crack on addition of heavy-atom derivative solutions or to diffract poorly. Two heavy-atom reagents proved successful in binding to the protein as well as giving crystals of reasonable diffraction quality (2.5–3.5 Å). These were chloro(2,2′:6′,2″-terpyridine)platinum(II) chloride, commonly known as orange platinum, and mercury acetate.

Orange-platinum derivative crystals were prepared by soaking the native crystals for ∼2.5 h in a solution with a final concentration of 1.5 m*M* orange platinum. Crystals were briefly back-soaked in reservoir solution before mounting. It was found that unlike the native, the platinum-derivative data correspond to a monoclinic cell, but exhibited pseudo-tetragonal symmetry and absences corresponding to $P4_32_12$ or $P4_12_12$.

The mercury acetate derivative crystals were prepared by soaking native crystals for around 1 h in a solution containing a final concentration of 1 m*M* mercury acetate. Crystals were back-soaked briefly in the reservoir solution and mounted. MAD data were collected at the EMBL Outstation at DESY, Hamburg[1]. Four mercury sites were found using the program *SOLVE* (Terwilliger & Berendzen, 1999). Heavy-atom positions were refined using *MLPHARE* (Collaborative Computational Project, Number 4, 1994) and *DM* (Cowtan & Main, 1996) was used for density modification. Roughly two-thirds of the backbone structure could be built using *TURBO* (Roussel & Cambillau, 1992), but the rest was too weakly defined in the map and prohibited further progress.

Eventually, we obtained two derivatives that were useful for MIR phasing. The

[1] Supplementary data have been deposited in the IUCr electronic archive (Reference ZA0061). Services for accessing these data are described at the back of the journal.

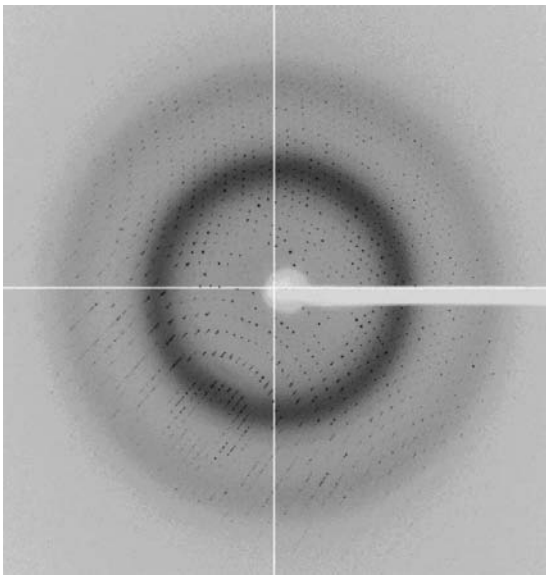

Figure 3
A representative 0.5° oscillation image of data collected on a native RG-lyase crystal at ID14-4, ESRF, Grenoble. Diffraction was observed to a resolution of 1.5 Å.

previously successful mercury acetate soak was prepared with a 2.5 m*M* concentration of mercury acetate to obtain a stronger signal from the Hg atoms.

The orange-platinum derivative was prepared using a tenfold lower concentration of orange platinum (0.15 m*M*) in the soak solution to avoid the changes in the space group upon soaking.

The derivative data sets were collected at the in-house rotating-anode generator. To obtain the best possible data, a full 360° data set was collected. Data were indexed, integrated and scaled using the *HKL* suite (Gewirth, 1994; Otwinowski & Minor, 1997) and further processing was performed with the *CCP*4 package (Collaborative Computational Project, Number 4, 1994). A summary of the data-collection and processing statistics is given in Table 1. It should be noted that the orange-platinum derivative crystals with tetragonal symmetry were obtained by lowering the concentration of the heavy-atom soak.

The positions of four mercury sites and one platinum site were obtained using *SOLVE* (Terwilliger & Berendzen, 1999). The quality of the solution from *SOLVE* gave a *Z* score of 76 and a figure of merit of 0.55, indicating a good solution. *RESOLVE* (Terwilliger, 2000) was run and a preliminary inspection of the map showed overall agreement with the previously built model. The new map extends further out, which gives us hope that we will be able to trace the whole enzyme from this map.

This work has been supported financially by a grant from the Danish National Research Foundation. We thank Dr Gordon Leonard at ESRF, Grenoble, Mr Flemming Hansen and Dr Jens Bukrinsky for help with the data collections and the synchrotron facilites ESRF and EMBL in Hamburg for providing beam time for the project. Helpful discussions with Dr Anders Kadziola and critical reading of and corrections to the manuscript by Drs Michael McDonough and Anne Mølgaard are gratefully acknowledged.

References

Carpita, N. C. & Gibeaut, D. M. (1993). *Plant J.* **3**, 1–30.
Carter, C. W. Jr & Carter, C. W. (1979). *J. Biol. Chem.* **254**, 12219–12223.
Collaborative Computational Project, Number 4 (1994). *Acta Cryst.* D**50**, 760–763.
Coutinho, P. M. & Henrissat, B. (1999). *CAZy – Carbohydrate-Active enZYmes.* http://afmb.cnrs-mrs.fr/~cazy/CAZY/index.html.
Cowtan, K. D & Main, P. (1996). *Acta Cryst.* D**52**, 43–48.
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., 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.
Mølgaard, A., Kauppinen, S. & Larsen, S. (2000). *Structure*, **8**, 373–383.
Mutter, M., Colquhoun, I. J., Beldman, G., Schols, H. A., Bakx, E. J. & Voragen, A. G. J. (1998). *Plant Physiol.* **117**, 141–152.
Mutter, M., Colquhoun, I. J., Schols, H. A., Beldman, G. & Voragen, A. G. J. (1996). *Plant Physiol.* **110**, 73–79.
O'Neill, M., Albersheim, P. & Darvill, A. (1990). *The Pectic Polysaccharides of Primary Cell Walls*, pp. 415–441. London: Academic Press.
Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.
Petersen, T. N., Kauppinen, S. & Larsen, S. (1997). *Structure*, **5**, 533–544.
Roussel, A. & Cambillau, C. (1992). *TURBO-FRODO Biographics and AFMP*. Architecture et Fonction des Macromolécules Biologique, Marseilles, France.
Schols, H. A., Posthumus, M. A. & Voragen, A. G. J. (1990). *Carbohydr. Res.* **206**, 117–129.
Shevchik, V. E. & Hugouvieux-Cotte-Pattat, N. (1997). *Mol. Microbiol.* **24**, 1285–1301.
Terwilliger, T. C. (2000). *Acta Cryst.* D**56**, 965–972.
Terwilliger, T. C. & Berendzen, J. (1999). *Acta Cryst.* D**55**, 849–861.
Whitcombe, A. J., O'Neill, M. A., Steffan, W., Albersheim, P. & Darvill, A. (1995). *Carbohydr. Res.* **271**, 15–29.