Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

Crystallization and preliminary X-ray analysis of maltose O-acetyltransferase

Leila Lo Leggio,[a] Florence Dal Degan,[b] Peter Poulsen,[b] Susanne Oxenbøll Sørensen,[b] Kenneth Harlow,[b] Pernille Harris[a] and Sine Larsen[a]*

[a]Centre for Crystallographic Studies, Chemical Institute, University of Copenhagen, Universitetsparken 5, DK-2100 Copenhagen Ø, Denmark, and [b]Danisco Cultor Innovation, Langebrogade 1, DK-1001 Copenhagen K, Denmark

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

Received 6 August 2001
Accepted 2 October 2001

Maltose O-acetyltransferase (Mac) is a member of the hexapeptide-repeat family of enzymes, which contains proteins with left-handed parallel β-helix architecture forming homotrimers. Diffraction data for four well diffracting crystal forms were collected. Crystal form I diffracted beyond 1.53 Å resolution but was perfectly merohedrally twinned with an apparent space group *P*622. Crystal forms II and III (space groups *R*3 and *C*2, respectively) could be obtained under very similar conditions by adjusting the buffer pH differently. Crystal forms II and III had several monomers in the asymmetric unit and were difficult to derivatize. However, during soaking with trimethyl lead acetate, the form III crystals dissolved and crystals with a different habit and space group grew in their place (form IV). In three of the crystal forms, a ladder of peaks was visible in the native Patterson maps along the *c* axis. These peaks were interpreted as corresponding to the vectors between the β-strands in the turns of the β-helix. Crystal form IV is suitable for structure determination of Mac exploiting the anomalous scattering of lead.

1. Introduction

Maltose O-acetyltransferase, also referred to as maltose transacetylase (Mac; E.C. 2.3.1.79), from *Escherichia coli* (Brand & Boos, 1991) transfers an acetyl group from acetyl-CoA to a sugar moiety. Originally discovered during studies on the maltose/maltodextrin transport system of *E. coli*, this enzyme has the ability to acetylate a variety of mono- and oligosaccharides. Mac contains several consecutive repeats of a consensus hexapeptide having Ile, Val or Leu in the first position and most often Gly in the second position (Vuorio *et al.*, 1994). The hexapeptide-repeat family contains mostly acyltransferases, but also carbonic anhydrases. Both acceptors and donor substrates are diverse, as are the biological roles identified for the various members of this family.

The first three-dimensional structure in the family to be reported was UDP-*N*-acetylglucosamine acyltransferase from *E. coli* (Raetz & Roderick, 1995). Subsequently, structures of a carbonic anhydrase (Kisker *et al.*, 1996) and other transferases belonging to the family were determined (Beaman *et al.*, 1997, 1998; Brown *et al.*, 1999; Olsen & Roderick, 2001; Sulzenbacher *et al.*, 2001; Kostrewa *et al.*, 2001).

The repeat-containing part of the sequence has been shown to form a left-handed parallel β-helix, with the overall shape of an equilateral triangular prism. All the structures are arranged as trimers of three identical monomers related by threefold rotational symmetry, with the axes of the β-helix of each monomer approximately parallel to the threefold axis relating the three monomers forming the trimer. Recently, however, the *E. coli* serine acetyltransferase has been reported to be a trimer of dimers (Hindson *et al.*, 2000), suggesting that some of the enzymes in the family may have an extra level of quaternary structure organization.

Apart from a *B. subtilis* protein of unknown function, Mac shows the highest sequence similarity (around 46% identity over 183 residues) to NodL, an O-acetyl transferase involved in the biosynthesis of lipo-chitin oligosaccharides which function as signals between plant and rhizobial bacteria to mediate the formation of nitrogen-fixing nodules. This enzyme has been crystallized (Dunn *et al.*, 1996), but no structure has yet been reported. None of the hexapeptide-repeat enzymes of known structure is, to our knowledge, able to acetylate oligosaccharides, a property which makes Mac of biotechnological interest (see PCT patent application WO 97/33974, F. Dal Degan, P. Poulsen, J. Marcussen and S. Oxenbøll). Structure determination has been undertaken as a step to expand the biotechnological potential of Mac and to investigate the evolution of specificity and function within the hexapeptide-repeat family.

Table 1
Crystallization and data-collection conditions and statistics for maltose O-acetyltransferase crystals.

	Form I	Form II	Form III	Form IV†
Reservoir	0.1 *M* MES pH 5.8, 1.53 *M* $(NH_4)_2SO_4$, 10 m*M* $CoCl_2$	0.1 *M* Na citrate pH 6, 12% MPD, 20% PEG 4000	0.1 *M* Na citrate pH 5.6, 3% ethanol, 15% PEG 4000	0.1 *M* Na citrate pH 5.6, 7% ethylene glycol, 18% PEG 6000
Growth temperature (K)	303.15	293	293	293
Cryocooling solution	1:4 mixture of 87% glycerol and reservoir	Reservoir	1:4 mixture of 87% glycerol and reservoir	3:17 mixture of ethylene glycol and reservoir
Beamline	BL711, MAXLAB	BL711, MAXLAB	BL711, MAXLAB	ID14-4, ESRF
Space group	$P622_2$	$R3$	$C2$	$C222_1$
Unit-cell parameters (Å, °)	$a = b = 98.6$, $c = 91.5$, $\alpha = \beta = 90$, $\gamma = 120$	$a = b = 155.5$, $c = 142.2$, $\alpha = \beta = 90$, $\gamma = 120$	$a = 183.5$, $b = 100.7$, $c = 91.4$, $\alpha = \beta = 90$, $\gamma = 101.3$	$a = 64.7$, $b = 106.6$, $c = 176.0$, $\alpha = \beta = \gamma = 90$
Z'	1–2	5–11	6–12	2–5
Resolution range (Å)	29–1.53 (1.56–1.53)	20–3.1 (3.15–3.1)	30–2.7 (2.75–2.7)	15–2.15 (2.23–2.15)
Measurements	235671	63376	132246	120930
Unique reflections	37263	21976	42215	33084
Completeness (%)	92.8 (95.6)	94.8 (90.7)	94 (83.3)	99.7 (99.8)
$R_{\rm merge}$	0.046 (0.215)	0.087 (0.337)	0.082 (0.310)	0.061 (0.396)
Reflections with $I/\sigma(I) < 2$ (%)	1.5 (5.7)	12.7 (36.5)	7.1 (28.8)	15.1 (40.9)

† To fully formed form III crystals grown in a drop consisting of 3 µl reservoir + 1 µl of protein solution were added 6 µl of reservoir and 1 µl 50 m*M* trimethyl lead acetate, after which form III crystals dissolved and form IV crystals began to grow. † Apparent space group used for processing.

The highest sequence identities between Mac and the hexapeptide-repeat enzymes of known structure were around 30–35%. The relatively low sequence identities, as well as the repetitive nature of the structures, make this a difficult problem for molecular replacement, so heavy-atom techniques are additionally required for structure determination.

2. Materials and methods

2.1. Cloning, expression and purification

The gene encoding Mac (*mae*) was identified from the nucleotide sequence of a 4.3 kbp *Eco*RI fragment from λ phage 8C4 (151) (Kohara *et al.*, 1987) as an open reading frame potentially encoding a 183 amino-acid protein (for details, see PCT patent application WO 97/33974, F. Dal Degan, P. Poulsen, J. Marcussen and S. Oxenbøll). In order to overexpress Mac, the *mae* gene was inserted by PCR after the IPTG (isopropyl-β-D-thiogalactopyranoside) inducible phage T7 promoter A1 in expression vector pUHE21-2 to create plasmid pMAC5 in *E. coli* strain NF1830.

Expression and purification were carried out in a similar way to the procedure of Brand & Boos (1991). *E. coli* NF1830-pMAC5 was grown at 310 K in 1 l LB broth containing 100 µg ml^{-1} ampicillin and 25 µg ml^{-1} kanamycin; expression of Mac was then induced with IPTG (2 m*M*). 4 h after induction the cells were harvested by centrifugation, washed in 0.9% NaCl and resuspended in 250 ml 20 m*M* potassium phosphate pH 7.5 containing 0.4 m*M* PMSF (phenylmethylsulfonyl fluoride), 0.4 mg ml^{-1} pepstatin and 1.6 m*M* EDTA (ethylenediaminetetraacetic acid). Cells were disrupted by sonication and the resulting homogenate centrifuged for 1 h at 90 000*g* and filtered through a 0.22 µm filter. This crude extract was applied to a Q-Sepharose 26/120 column (Pharmacia Biotech) equilibrated with 20 m*M* potassium phosphate pH 7.5 (buffer *A*) and the bound protein was eluted by applying a 0–0.3 *M* NaCl linear gradient in the same buffer. The active fractions, as measured by a modified Alpers' assay (Alpers *et al.*, 1965), were pooled and applied to an Affi-Gel Blue column (Biorad) equilibrated in buffer *A* and subsequently washed in the same buffer containing 0.4 *M* NaCl. Protein was eluted with buffer *A* containing 2 *M* NaCl. After removal of the salt by dialysis, the active pool was concentrated and applied to a 6 ml acetyl-CoA-Minileak column in buffer *A*. The affinity resin was made by coupling 200 mg of acetyl-CoA to 5 g of Minileak High (Kem-En-Tek, Denmark) in 10 ml of 1 *M* $NaCO_3$ pH 11 for 20 h at room temperature. The column was washed in buffer *A* and inverted; pure Mac eluted in buffer *A* containing 0.5 *M* NaCl.

2.2. Determination of the oligomerization state

The subunit association of Mac was studied by gel filtration on a Superdex 75 PC 3.2/20 column (Pharmacia Biotech) and by chemical crosslinking with glutaraldehyde (Farmer & Caprioli, 1991) followed by MALDI-TOF mass spectrometry. The Mac concentration in the cross-linking experiment was 8.6 µ*M*.

2.3. Crystallization and data collection

Crystals were grown by the hanging-drop vapour-diffusion method in Linbro culture plates using 1 ml reservoir volume and a drop typically consisting of 4–8 µl of a 15.9 mg ml^{-1} protein solution in 10 m*M* phosphate buffer pH 7.5 and reservoir in a 1:1 ratio. Initial crystallization conditions were identified using the commercially available Hampton Research Screens I and II (Jancarik & Kim, 1991; Cudney *et al.*, 1994) and were later refined to the conditions indicated in Table 1. Data were collected at 100 K at the various synchrotron beamlines as indicated in Table 1 and in the text.

2.4. Data processing and molecular-replacement searches

The diffraction data were processed using *DENZO*/*SCALEPACK* (Otwinowski & Minor, 1997) and reduced to structure-factor amplitudes using the program *TRUNCATE* from the *CCP*4 suite (Collaborative Computational Project, Number 4, 1994). Molecular-replacement searches were performed in *AMoRe* (Navaza, 1994), while native Patterson maps were calculated with the *CCP*4 program *FFT*.

3. Results and discussion

3.1. Oligomerization state of maltose O-acetyltransferase

In gel-filtration studies Mac had an elution volume corresponding to a mass of 47 000 Da, compared with the subunit mass of 19 983 Da measured by MALDI-TOF, suggesting homodimer or homotrimer formation. Further investigation of the mass by MALDI-TOF after cross-linking with glutaraldehyde showed that the main oligomeric species was a homotrimer, although a minor fraction of the protein was in a homohexameric form as in the *E. coli* serine acetyltransferase. As we have several crystal forms available, we expect that they will allow us to see if hexamer packing is conserved and suggest whether the hexamer might have a functional role.

3.2. Crystal form I

Through optimization of condition 25 in the Hampton Research Crystal Screen II needles shaped as hexagonal prisms could be grown. Crystals usually appeared after 3–4 d and grew to a maximum size of 0.2 × 0.2 × 0.5 mm in four to six weeks

(Figs. 1*a* and 1*b*). A full data set for a form I crystal was collected to 1.53 Å resolution (details in Table 1).

The apparent space group is *P*622, with a unit-cell volume most consistent with one monomer per asymmetric unit (two monomers per asymmetric unit would result in 23% solvent content). This implies that a Mac trimer must lie along a crystallographic threefold axis. An analysis of the native Patterson map shows distinct peaks along *w* at distances of around 5 Å from each other (Fig. 2*a*). These are likely to correspond to the vectors between atoms of adjacent β-strands in the parallel β-helix domain and would arise only if the axis of the domain were parallel to the *c* axis; the experimental native Patterson map compared well with one from calculated data using a parallel β-helix model (not shown).

Consistent with this, molecular replacement using the parallel β-helix domain of tetrahydrodipicolinate N-succinyl transferase (Beaman *et al.*, 1997) as the search model showed a clear solution for the rotation function with the monomer positioned almost parallel to the crystallographic sixfold and threefold axes. Despite considerable effort, we were unable to obtain a reasonable solution for the translation function. We collected further data sets from many crystals soaked in heavy-atom derivatives, but either the crystals lost diffraction power or we were unable to find heavy-atom sites.

One source of our problems became apparent when we tried the test for merohedral twinning on Todd Yeates' twinning server (http://www.doe-mbi.ucla.edu/Services/Twinning/, Yeates, 1997). The plot in Fig. 3 strongly suggests that the native data are perfectly merohedrally twinned and the many (over 30) data sets collected for this crystal form all show a similar trend. Channels like the ones visible in Fig. 1(*a*) along the length of the hexagonal needles are known to often be associated with merohedrally twinned crystals. Given the difficulties associated with structure determination of a perfectly merohedrally twinned crystal, our efforts were concentrated in obtaining different crystal forms.

3.3. Crystal forms II and III

Long needles belonging to space group *R*3 (crystal form II; Fig. 1*c*) could be grown using as reservoir condition 40 in the Hampton Research Crystal Screen I [0.1 *M* sodium citrate pH 5.6, 20%(*v*/*v*) 2-propanol, 20%(*w*/*v*) PEG 4000]. The crystals tended to grow from a crystalline gelatinous aggregate which formed a ring around the drop after 1 or 2 d. We attempted to reproduce and improve crystals with this morphology by varying the growth conditions slightly and also by substituting 2-propanol with other small organic molecules. Initially, we could only reliably grow crystals using the solution supplied by Hampton Research, although occasionally crystals would appear in drops set up using solutions prepared in our laboratory. Data were collected from a form II crystal using the conditions indicated in Table 1. The cell edges are consistent with four to ten monomers per asymmetric unit and the data extended to 3.1 Å resolution (details in Table 1).

During our initial attempts to reproduce crystal form II, a third, plate-like crystal form was obtained (crystal form III; Fig. 1*d*) at lower 2-propanol and PEG concentrations. 2-propanol could be substituted by other small organic molecules such as ethanol and ethylene glycol. The crystals tended to grow as clusters but could usually be separated into single plates. Eventually, the reason for the polymorphism was tracked down to a difference in the solution used for adjusting the pH of the

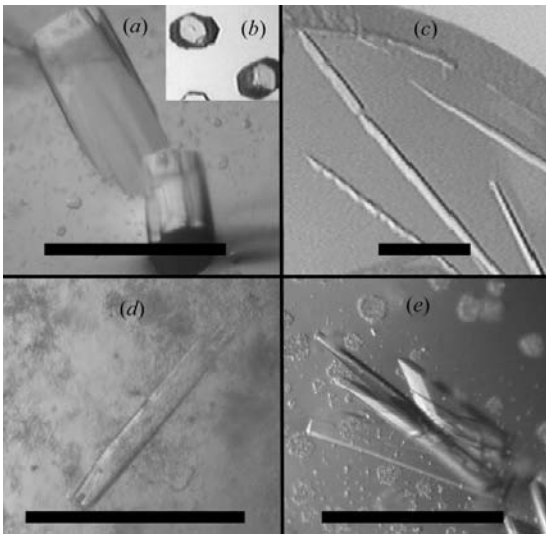

Figure 1
Crystals of maltose O-acetyltransferase. The black line corresponds to 0.5 mm. (*a*) Form I crystals, (*b*) smaller form I crystals, (*c*) form II crystals, (*d*) form IV crystals, (*e*) form III crystals.

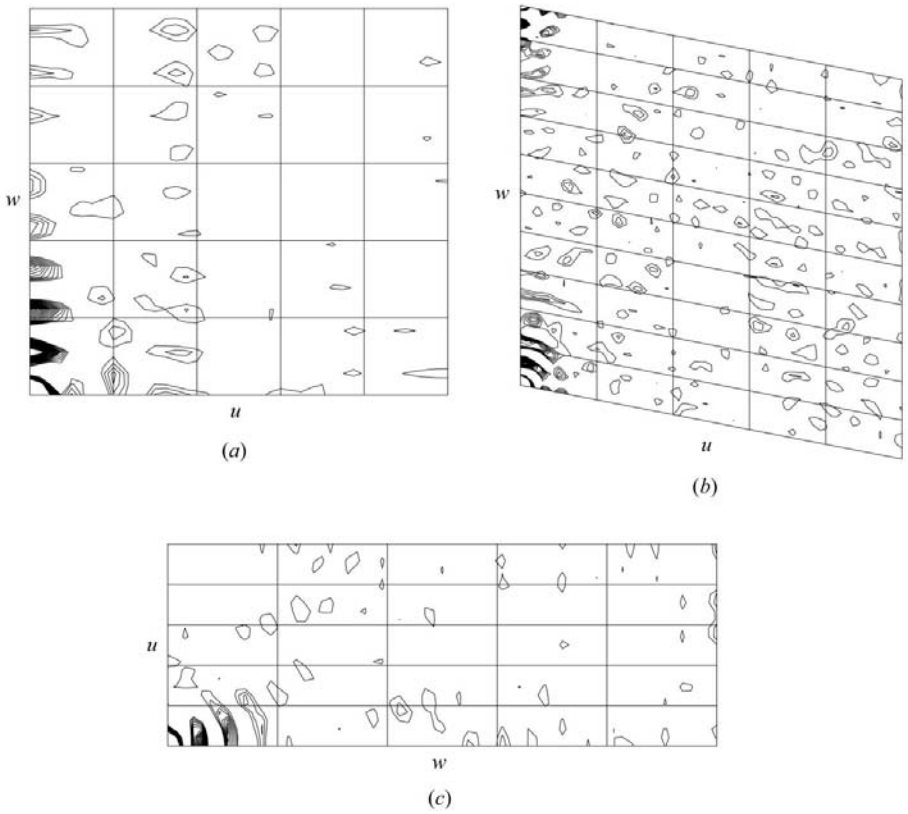

Figure 2
Native Patterson map sections ($v = 0$) for form I (*a*), III (*b*) and IV (*c*) crystals. The maps are contoured at 1σ levels.

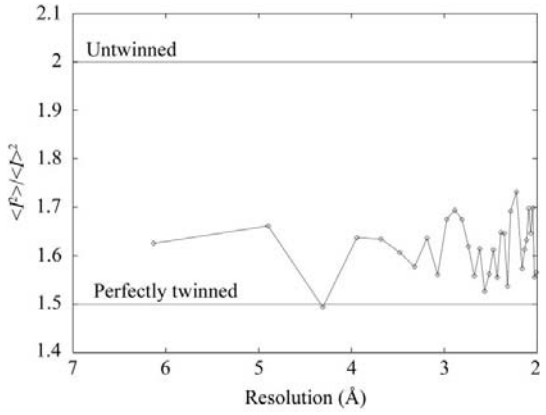

Figure 3
Twinning test for the form I native data set.

buffer, namely HCl for the Hampton solution and citric acid for the 'home-made' replica. It remains to be seen whether the structures of the two crystal forms will explain such a striking effect of a minor reservoir component.

A data set was collected for a form III crystal to 2.7 Å resolution using the conditions indicated in Table 1. Form III crystals belong to space group *C*2, with unit-cell parameters consistent with 6–12 monomers per asymmetric unit (details in Table 1). In this crystal form the parallel β-helix axis is also approximately oriented along the crystallographic *c* axis, as can be seen from the corresponding native Patterson map (Fig. 2*b*).

3.4. Crystal form IV

Crystal forms II and III are expected to have many monomers in the asymmetric unit and therefore present difficulties both for molecular replacement and for heavy-atom derivative interpretation. Many heavy-atom compounds were tested, both in soaks and co-crystallization, but a derivative could not be identified. Using Pt and Hg compounds, the crystals lost all or most of their diffracting power or did not grow in co-crystallization experiments.

However, in one instance we observed that when adding 5 m*M* trimethyl lead acetate to a hanging drop containing *C*2 crystals, the plate-like crystals dissolved but later gave rise to precipitate and morphologically different needle-like crystals (crystal form IV; Fig. 1*e*). The growth of these crystals could be reproduced if exactly the same growth and soaking conditions were used as for the *C*2 form III crystals, as indicated in Table 1. Form IV crystals belonged to space group $C222_1$ and a data set to 2.15 Å resolution was collected. As shown in the native Patterson map section in Fig. 2(*c*), there is also evidence in this case that the parallel β-helix is approximately aligned with the crystallographic *c* axis. Twin tests for crystal forms II, III and IV were negative.

Crystals of similar morphology to the form IV crystals could be grown in the absence of heavy atoms by streak seeding in a drop containing form III crystals with a horse hair that had touched a form IV crystal. However, data collected from one such crystal at beamline X13, EMBL Outstation, Hamburg could not be processed in space group $C222_1$, although the unit-cell parameters were similar to those of form IV crystals. Currently the structure of Mac is being determined exploiting the anomalous signal from lead obtained from the form IV crystals.

We are grateful to MAXLAB (Lund, Sweden), the ESRF (Grenoble, France) and EMBL (Hamburg, Germany) for the provision of the synchrotron time necessary for this project and the EU for financial support towards the trips. We would also like to thank both the beamline staff and the people who have helped with data collection, in particular Johan Gotthardt Olsen, Jens Bukrinsky and Flemming Hansen. We thank Helle Stiller for her technical assistance and acknowledge the financial support of the Danish National Research Foundation and EMBO.

References

Alpers, D. H., Appel, S. H. & Tomkrins, G. M. (1965). *J. Biol. Chem.*, **240**, 10–13.
Beaman, T. W., Binder, D. A., Blanchard, J. S. & Roderick, S. (1997). *Biochemistry*, **36**, 489–494.
Beaman, T. W., Sugantino, M. & Roderick, S. L. (1998). *Biochemistry*, **37**, 6689–6696.
Brand, B. & Boos, W. (1991). *J. Biol. Chem.* **266**, 14113–14118.
Brown, K., Pompeo, F., Dixon, S., Mengin-Lecreulx, D., Cambillau, C. & Bourne, Y. (1999). *EMBO J.* **18**, 4096–4107.
Collaborative Computational Project, Number 4 (1994). *Acta Cryst.* D**50**, 760–763.
Cudney, R., Patel, S., Weisgraber, K., Newhouse, Y. & McPherson, A. (1994). *Acta Cryst.* D**50**, 414–423.
Dunn, S. M., Moody, P. C. E., Downie, J. A. & Shaw, W. V. (1996). *Protein Sci.* **5**, 538–541.
Farmer, T. B. & Caprioli, R. M. (1991). *Biol. Mass Spectrom.* **20**, 796–800.
Hindson, V. J., Moody, P. C. E., Rowe, A. J. & Shaw, W. V. (2000). *J. Biol. Chem.* **275**, 461–466.
Jancarik, J. & Kim, S.-H. (1991). *J. Appl. Cryst.* **24**, 409–411.
Kisker, C., Schindelin, H., Alber, B. E., Ferry, J. G. & Rees, D. C. (1996). *EMBO J.* **15**, 2323–2330.
Kohara, Y., Akiyama, K. & Isono, K. (1987). *Cell*, **50**, 495–508.
Kostrewa, D., D'Arcy, A., Takacs, B. & Kamber, M. (2001). *J. Mol. Biol.* **305**, 279–289.
Navaza, J. (1994). *Acta Cryst.* A**50**, 157–163.
Olsen, L. R. & Roderick, S. L. (2001). *Biochemistry*, **40**, 1913–1921.
Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.
Raetz, C. R. H. & Roderick, S. L. (1995). *Science*, **270**, 997–1000.
Sulzenbacher, G., Gal, L., Peneff, C., Fassy, F. & Bourne, Y. (2001). *J. Biol. Chem.* **276**, 11844–11851.
Vuorio, R., Harkonen, T., Tolvanen, M. & Vaara, M. (1994). *FEBS Lett.* **337**, 289–292.
Yeates, T. O. (1997). *Methods Enzymol.* **276**, 344–358.