Crystallization and preliminary X-ray analysis of α-isopropylmalate synthase from *Mycobacterium tuberculosis*

Nayden Koon, Christopher J. Squire and Edward N. Baker*

School of Biological Sciences, University of Auckland, Auckland, New Zealand

Correspondence e-mail:
ted.baker@auckland.ac.nz

α-Isopropylmalate synthase catalyses the aldol condensation of α-ketoisovalerate and acetyl coenzyme A to produce α-isopropyl-malate. This reaction is the first committed step of leucine biosynthesis, which is interrelated with the pathways for production of the other branched-chain amino acids, valine and isoleucine. The absence of these pathways in mammals suggests that these enzymes could be useful targets for drug design against microbial pathogens. The gene for α-IPMS in *Mycobacterium tuberculosis* (Rv3710) has been cloned, expressed in *Escherichia coli*, both in native and selenomethionine-substituted forms, and crystallized. The SeMet crystals are suitable for high-resolution X-ray structural analysis. These crystals are monoclinic, with unit-cell parameters a = 54.25, b = 154.73, c = 68.82 Å, space group $P2_1$ and two molecules in the asymmetric unit. X-ray diffraction data to 2.0 Å resolution have been collected.

1. Introduction

Leucine, valine and isoleucine are members of the branched-chain amino-acid family and are synthesized by plants and microbes but not by higher organisms. The biosynthetic pathways for these amino acids are interrelated and several of the early enzymes required for processing are shared. For example, the pathways for valine and isoleucine are parallel, with the same enzymes catalyzing equivalent reactions on different substrates. Leucine synthesis occurs *via* the α-isopropylmalate pathway, which branches from valine biosynthesis after α-ketoisovalerate (α-KIV) is produced. α-Isopropylmalate synthase (α-IPMS) is the first dedicated enzyme for leucine synthesis and transfers an acetyl group from acetyl coenzyme A (aCoA) to α-KIV to produce isopropyl-malate *via* an aldol condensation reaction.

It has been shown that disruptions to the branched-chain amino-acid biosynthetic pathways can be inhibitory to microbial growth and virulence (McAdam *et al.*, 1995; Bange *et al.*, 1996; Grandoni *et al.*, 1998; Hondalus *et al.*, 2000). This evidence suggests that enzymes in these pathways could be useful as potential drug targets in the fight against human pathogens; in particular *Mycobacterium tuberculosis*, the causative agent of tuberculosis (TB) and one of the most lethal infectious agents. Open reading frame Rv3710 from the *M. tuberculosis* genome sequence has been annotated as *leuA*, the gene that produces α-IPMS, based on sequence similarities with other organisms (Cole *et al.*, 1998). The gene product of Rv3710 has been shown experimentally to possess α-IPMS activity and has an additional point of

interest in the form of a variable number of tandem repeats, which result in polymorphic forms in different strains (Chanchaem & Palittapongarnpim, 2002). Although extensive studies have been carried out on α-IPMS enzymes from a variety of organisms and much has been elucidated from a functional perspective (for a review of leucine biosynthesis in fungi, see Kohlhaw, 2003), no three-dimensional structure has yet been solved. As part of our participation in the TB Structural Genomics Consortium (http://www.doe-mbi.ucla.edu/TB/), we describe here the successful expression, purification and crystallization of α-IPMS from *M. tuberculosis*.

2. Methods and materials

2.1. Cloning

The open reading frame coding for α-IPMS (Rv3710) was amplified from genomic *M. tuberculosis* (strain H37Rv) DNA using the polymerase chain reaction with the following oligonucleotide primers (Invitrogen): 5′-GC-AACTACCATGGCAACTTCTGAATCGC-3′ and 5′-GCAGTAAGCTTTGTTCTGAAAC-GGCGAGCG-3′. The construct was prepared by ligating the *Nco*I/*Hin*dIII fragment into the expression vector pProEX-HTa (Life Technologies) and was propagated in *Escherichia coli* (DH5α) using ampicillin resistance for selection. Note that in this construct the ATG from the *Nco*I restriction site corresponds to the start codon for *leuA* and thus the first two amino acids of recombinant α-IPMS were mutated to methionine and alanine, respectively. Sequencing was completed at the

Auckland University sequencing facility to confirm that no unintended mutations had been introduced. The final plasmid construct was transformed into *E. coli* BL21(DE3)-CodonPlus-RP strain (Stratagene) for expression.

2.2. Expression and purification

Cells were inoculated from an overnight pre-culture at 1/100 dilution and grown in Luria–Bertani broth containing 100 µg ml^{-1} ampicillin and 34 µg ml^{-1} chloramphenicol. At an OD_{600} of 0.4–0.8, recombinant protein was expressed by induction of the cell culture with a final concentration of 0.1 mM isopropyl β-D-thiogalactopyranoside (IPTG), followed by incubation at 291 K overnight. Cell pellets were lysed by cell disruption (Constant Systems) in a buffer containing 50 mM Tris–HCl pH 9.0 and 250 mM NaCl. The enzyme was purified from crude lysate by immobilized nickel affinity chromatography (INAC, using a 5 ml HiTrap chelating column) in the above buffer and eluted with a stepwise gradient of imidazole (up to 250 mM). Further purification was carried out by gel filtration (Superdex 200 16/60) in a buffer as above but at pH 7.5. Fractions from the major peak were checked for monodispersity using dynamic light scattering before subsequent purification by anion-exchange chromatography (5 ml HiTrap Q Sepharose), with elution with a stepwise gradient of NaCl (up to 350 mM). The His$_6$ tag was cleaved by recombinant tobacco-etch virus (rTEV) protease and removed by INAC. Selenium-incorporated protein was prepared by expression of the construct in DL41(DE3)-CodonPlus-RP cells in minimal media (Hendrickson *et al.*, 1990) containing 25 µg ml^{-1} selenomethionine (SeMet) and appropriate antibiotics. Purification was as above except that all buffers were supplemented with 5 mM β-mercaptoethanol (2 mM during INAC). All columns were supplied by Amersham Pharmacia Biotech.

2.3. Crystallization and data collection

Initial crystallization conditions for native α-IPMS were found from a crystallization screen designed using *CRYSTOOL* (Segelke, 2001). These initial native crystals, grown from 100 mM sodium citrate pH 6.5 with 30% methoxypolyethylene glycol (MPEG) 2000, proved difficult to reproduce, but the original crushed native crystals and their progeny were used to seed all subsequent crystallizations. Diffracting SeMet α-IPMS crystals were obtained by microseeding from these native crystals using the

sitting-drop vapour-diffusion method (2 + 2 µl drops in 96-well plates, Greiner). The protein concentration was ~4.5 mg ml^{-1} and the precipitation buffer contained 100 mM sodium citrate pH 6.25 with 26% MPEG 2000. Crystals appeared overnight and grew as rods or trapezoidal prisms of up to 0.3 mm in length (Fig. 1) over a week or so, but were unstable over longer time periods. SeMet protein crystals grown in the presence of KCl and coenzyme A (5 mM each) were stable for much greater periods (>5 months). All crystals were dipped briefly in a cryoprotectant solution (containing 85% mother liquor, 15% glycerol) and flash-frozen in liquid nitrogen prior to data collection. All X-ray data were collected at 113 K on a Rigaku RU-H3R rotating-anode generator equipped with a MAR345 image-plate detector and were processed using the

HKL program package (Otwinowski & Minor, 1997).

3. Results and discussion

Both native and selenomethionine forms of recombinant protein retained *in vitro* activity as seen in enzymatic studies based on those by Kohlhaw (1988) (data not shown) and confirmed the protein to be α-IPMS. Although good-quality crystals were not readily reproducible from the initial crystallization conditions for native α-IPMS, subsequent microseeding provided a reliable route to both native and SeMet-substituted proteins. The optimal conditions for the SeMet crystals were typically at a lower pH than for the native crystals, as is often observed, but the other conditions were similar for both. A variety of crystal
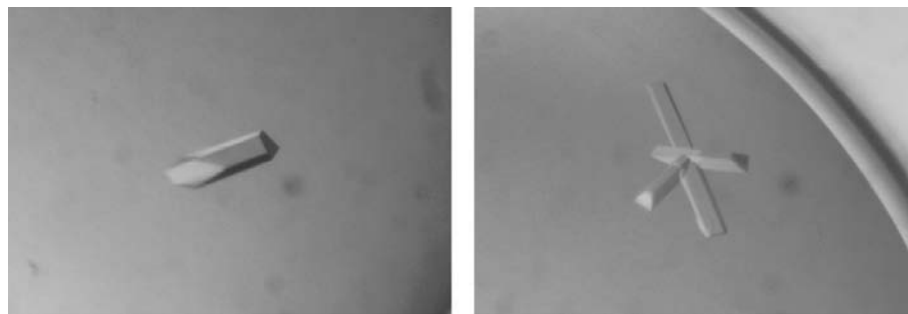
Figure 1
Images of typical α-isopropylmalate synthase crystals of approximately 0.3 mm in length. SeMet crystals were unstable for long periods without additives, but were frequently of higher quality in both appearance and diffraction.
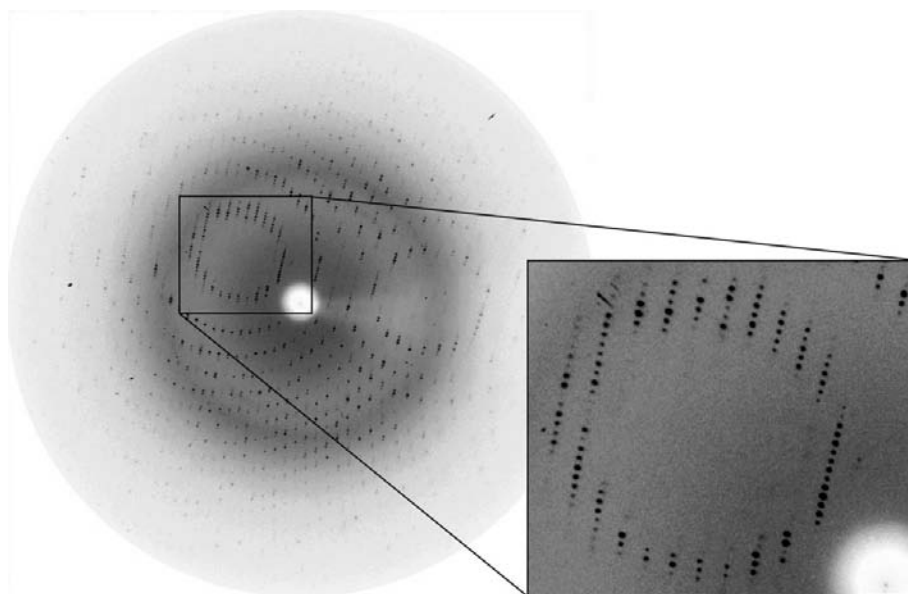
Figure 2
X-ray diffraction pattern obtained from SeMet crystals of α-isopropylmalate synthase on our home source, taken as a 1° oscillation 10 min exposures with a crystal-to-detector distance of 175 mm. The resolution at the edge of the frame is 2.0 Å.

Table 1
Crystal data and data-collection statistics.

Values in parentheses are for the outermost shell of data collected.

Unit-cell parameters (Å, °)	$a = 54.25$, $b = 154.73$, $c = 68.82$, $\beta = 98.05$
Space group	$P2_1$
Matthews coefficient (Å^3 Da^{-1})	2.0
Solvent content (%)	39.3
No. molecules in AU	2
Resolution range (Å)	30–2.0 (2.07–2.0)
Wavelength used (Å)	1.5418
Total No. unique reflections	74305
No. observed reflections [$I/\sigma(I) > 1$]	717235
Completeness (%)	98.4 (96.0)
Multiplicity	9.7
R_{merge}†	0.072 (0.652)
Mean $I/\sigma(I)$	34.6 (3.0)

† $R_{merge} = \sum |I_i - \langle I \rangle| / \sum \langle I \rangle$, where I_i and $\langle I \rangle$ are the measured and mean values of reflection hkl, respectively.

shapes were observed during growth, with the best crystals for diffraction being rods or trapezoidal prisms. Although native crystals grew to a larger size, their data sets gave lower resolution and were often afflicted by severe twinning. SeMet crystals were of better appearance and diffraction quality, but speed was important in purification and growth. If left for longer than about a week, SeMet crystals took on a 'rubbery' consistency and failed to diffract, even though outwardly they appeared unchanged. This problem was solved by the use of KCl and CoA in the crystallization medium. We speculate that the KCl acts as a stabilizing agent, since although it is not an actual substrate of α-IPMS, it is required for full activity (Kohlhaw, 1988). Currently, the highest resolution native data set for α-IPMS is at 2.4 Å resolution and 2.0 Å for the SeMet protein. Full crystallographic details for the latter data set are presented in Table 1 and the diffraction pattern is shown in Fig. 2. The quality of these crystals is suitable for collection of multiwavelength data for MAD phase determination and subsequent structure solution.

References

Bange, F.-C., Brown, A. M. & Jacobs, W. R. Jr (1996). *Infect. Immun.* **64**, 1794–1799.

Chanchaem, W. & Palittapongarnpim, P. (2002). *Tuberculosis*, **82**, 1–6.

Cole, S. T. *et al.* (1998). *Nature (London)*, **393**, 537–544.

Grandoni, J. A., Marta, P. T. & Schloss, J. V. (1998). *J. Antimicrob. Chemother.* **42**, 475–482.

Hendrickson, W. A., Horton, J. R. & LeMaster, D. M. (1990). *EMBO J.* **9**, 1665–1672.

Hondalus, M. K., Bardarov, S., Russel, R., Chan, J., Jacobs, W. R. Jr & Bloom, B. R. (2000). *Infect. Immun.* **68**, 2888–2898.

Kohlhaw, G. B. (1988). *Methods Enzymol.* **166**, 414–423.

Kohlhaw, G. B. (2003). *Microbiol. Mol. Biol. Rev.* **67**, 1–15.

McAdam, R. A., Weisbrod, T. R., Martin, J., Scuderi, J. D., Brown, A. M., Cirillo, J. D., Bloom, B. R. & Jacobs, W. R. Jr (1995). *Infect. Immun.* **63**, 1004–1012.

Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.

Segelke, B. W. (2001). *J. Cryst. Growth*, **232**, 553–562.