Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

Crystallization and preliminary X-ray diffraction studies of an α-methylacyl-CoA racemase from *Mycobacterium tuberculosis*

Prasenjit Bhaumik,[a] Petri Kursula,[a] Ville Ratas,[a] Ernst Conzelmann,[b] J. Kalervo Hiltunen,[a] Werner Schmitz[b] and Rik K. Wierenga[a]*

[a]Biocenter Oulu and Department of Biochemistry, University of Oulu, Linnanmaa, Box 3000, FIN-90014 Oulu, Finland, and [b]Biozentrum der Universität, Am Hubland, D-97074 Würzburg, Germany

Correspondence e-mail: rik.wierenga@oulu.fi

Received 6 August 2002
Accepted 13 November 2002

α-Methylacyl-CoA racemase is a key enzyme in the metabolism of 2-methyl-branched fatty acids and, in mammals, in the conversion of cholesterol to bile acids. The enzyme from *Mycobacterium tuberculosis* has been purified to homogeneity and crystallized by the hanging-drop vapour-diffusion method. The crystals of the unliganded racemase belong to space group $P6_122$ or $P6_522$, with unit-cell parameters $a = b = 122.0$, $c = 256.4$ Å. Data sets were collected at 100 K. The crystals diffract to 2.8 Å using synchrotron radiation.

1. Introduction

The enzyme α-methylacyl-CoA racemase (AMACR) catalyzes the racemization of a wide spectrum of α-methyl-branched carboxylic acids as their coenzyme A thioesters (Schmitz *et al.*, 1994, 1995). Its major physiological role is thought to be in the biosynthesis of bile acids, since mitochondrial hydroxylation at C-26 of the cholesterol side chain specifically yields the (25*R*)-diastereomer of di- and trihydroxycoprostanoic acids (THCA) (Shefer *et al.*, 1978), while the peroxisomal oxidases which initiate β-oxidative cleavage of the side chain are highly stereospecific for the (25*S*)-isomer (Pedersen *et al.*, 1996). Thus, inversion of the configuration at C-25 is required to connect the two pathways. Other physiological substrates are methyl-branched fatty acids of dietary origin such as pristanic acid (2,6,10,14-tetramethylpentadecanoic acid), which is the α-oxidation product of phytanic acid, a metabolite of the chlorophyll component phytol (for a review, see Steinberg, 1995) and possibly also endogenously synthesized isoprenoids (Fig. 1). Inherited deficiency of AMACR causes an adult-onset sensory motor neuropathy similar to mild Refsum disease variants (Ferdinandusse *et al.*, 2000). AMACR is also pharmacologically interesting because it is involved in the biotransformation of 2-arylpropionic acids (2-methyl-arylacetic acids), which are used as non-steroidal anti-inflammatory and analgesic drugs (ibuprofen), from the inactive (2*S*)-enantiomer to the pharmacologically active (2*R*)-enantiomer (Shieh & Chen, 1993). In humans and in mouse, the enzyme is distributed in both peroxisomes and mitochondria (Kotti *et al.*, 2000). In both intracellular localizations the leader peptide is not cleaved off.

In contrast to most other known racemases, AMACR does not require any cofactor (Schmitz *et al.*, 1994). Little else is known about the catalytic mechanism and no structural information about this enzyme is available. No other proteins with similar amino-acid sequence could be found in the Protein Data Bank (PDB) and, at the cDNA level, in expressed sequence tag (EST) libraries of human and mouse.

Purification and characterization of AMACR from human and rat have been reported (Schmitz *et al.*, 1994, 1995). Here, we report the purification and crystallization of this racemase from *Mycobacterium tuberculosis*. In the *M. tuberculosis* genome, three genes, termed *mcr*, *far* and *Rv3727*, code for proteins with amino-acid sequences similar to that of mammalian AMACRs. Two of them, *mcr* and *far*, yield enzymatically active proteins

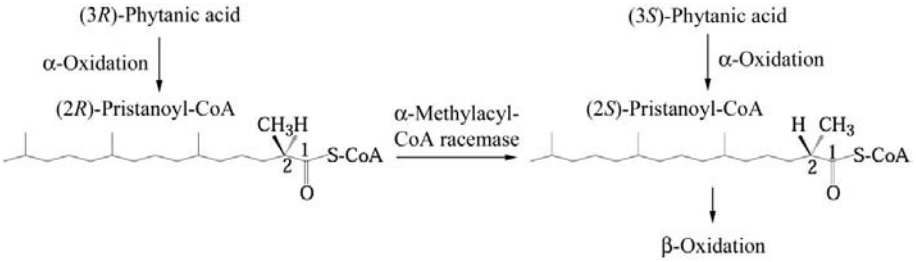

Figure 1
Involvement of α-methylacyl-CoA racemase in the degradation of phytanic acid. AMACR has a broad substrate specificity as it also catalyses the racemization of the 2-methyl group of ibuprofenoyl-CoA and of 3,7,12-trihydroxycoprostanoyl-CoA (see text).

when expressed in *Escherichia coli* (W. Schmitz, unpublished work), with MCR having the highest specific activity. Alignment of the MCR protein sequence (Cole *et al.*, 1998) with rat (SWISS-PROT Accession No. P70473, with updated N-terminal sequence) and human (SWISS-PROT Accession No. Q9UHK6) AMACR using *CLUSTALX* (Thompson *et al.*, 1997), with only four minor gaps, reveals a sequence identity of 153 out of 353 common amino acids (43%) (Fig. 2). It therefore seems reasonable to assume that the bacterial AMACR may serve as a model structure for the AMACRs of higher organisms. Here, we describe the crystallization and preliminary X-ray diffraction analysis of the α-methylacyl-CoA racemase encoded by the *mcr* gene of *M. tuberculosis*.

2. Materials and methods

2.1. Expression and purification

Purification of the rat liver enzyme (Schmitz *et al.*, 1994) and the human liver enzyme (Schmitz *et al.*, 1995) have been described previously. Purification of α-methylacyl-CoA racemase from *M. tuberculosis* (MCR) followed the same strategy.

The plasmid pET3a containing the *mcr* gene was transformed into the *E. coli* BL21(DE3)pLysS strain for overexpression. A single colony was chosen and inoculated in 2 ml Luria–Bertani (LB) broth containing 50 µg ml^{-1} ampicillin and 34 µg ml^{-1} chloramphenicol. The inoculated medium was incubated overnight with shaking at 310 K. 50 ml LB broth containing 50 µg ml^{-1} ampicillin and 34 µg ml^{-1} chloramphenicol was inoculated with 100 µl overnight-grown culture and incubated with shaking at 310 K overnight. The culture was then centrifuged (15 min, 4550*g*) and the cell pellet resuspended in 4 ml pre-warmed (at 310 K) LB broth containing 50 µg ml^{-1} ampicillin and 34 µg ml^{-1} chloramphenicol. The suspended culture was then diluted 1000-fold into 4 l of LB broth containing 50 µg ml^{-1} ampicillin and 34 µg ml^{-1} chloramphenicol. The culture was incubated with shaking at 310 K for 2 h and overexpression of α-methylacyl-CoA racemase was initiated by adding isopropyl-β-D-thiogalactoside to 0.4 m*M* at 303 K. Overexpression was continued for 5 h. Cells were harvested by centrifugation and stored at 253 K overnight.

All the purification steps were performed at 277 K unless mentioned otherwise. Protein purification was initiated by suspending cells in 100 ml cold lysis buffer (50 m*M* potassium phosphate buffer containing 0.3 *M* NaCl and 0.01% NaN_3 pH 8.0). Cells were disrupted by sonication. The lysate was centrifuged at 47 800*g* for 20 min to remove cell debris. Solid ammonium sulfate was added slowly under continuous stirring to the supernatant to 40% saturation. The precipitated protein was spun off (47 800*g*, 20 min) and discarded. The supernatant was brought to 60% saturation with ammonium sulfate. Precipitated proteins were collected by centrifugation (47 800*g*, 30 min) and resuspended in 60 ml 10 m*M* potassium phosphate buffer containing 0.01% NaN_3 pH 6.0 (hereafter referred to as PP buffer). The protein solution was dialyzed exhaustively against PP buffer. The retentate was applied to a DEAE-Sepharose column (80 ml) equilibrated with PP buffer. After washing out the unbound protein with three column volumes of 0.2 *M* NaCl in PP buffer, the enzyme was eluted with 200 ml of a linear gradient of 0.2–0.4 *M* NaCl in PP buffer (fraction volume 4 ml). MCR-containing fractions were combined and concentrated to 20 ml by ultrafiltration (30 kDa cutoff). A solution of 3 *M* $(NH_4)_2SO_4$ in PP was added to the concentrated sample to a final concentration of 1.5 *M*. The sample was applied to an octyl-Sepharose column (50 ml) equilibrated with PP buffer. After washing out the unbound proteins with two column volumes of 1.5 *M* $(NH_4)_2SO_4$ in PP buffer, the enzyme was eluted with 100 ml of a linear gradient of 1.5–0.0 *M* $(NH_4)_2SO_4$ in PP buffer. MCR-containing fractions were combined. The protein solution was desalted and concentrated by ultrafiltration (30 kDa cutoff).

2.2. Solution studies

The purified protein (1 mg ml^{-1}) was stored at 277 K in PP buffer. An enzyme assay was performed to measure the activity of the purified enzyme. Formation of (^{3}H)-H_2O from (2-^{3}H)-pristanoyl-CoA was measured as described previously (Schmitz *et al.*, 1994). Briefly, the enzyme source was incubated with the labelled substrate in 50 µl 50 m*M* Tris–HCl pH 8.0 at 310 K for 30 min. The reaction was stopped with 450 µl 1% trichloroacetic acid and (^{3}H)-H_2O was separated from unreacted substrate on reverse-phase silica gel (RP-18) and quantified by liquid-scintillation counting. Dynamic light-scattering experiments were performed to check the monodispersity of the purified protein sample. These

```
MCR   MAGPLSGLRVVELAGIGPGPHAAMILGDLGADVVRIDRPSSVDGISRDAMLRNRRIVTAD 60
HR    MA--LQGISVVELSGLAPGPFCAMVLADFGARVVRVDRPGSRYDVSR--LGRGKRSLVLD 56
RR    MA--LRGVRVLELAGLAPGPFCGMILADFGAEVVLVDRLGSVNHPSH--LARGKRSLALD 56
      **  * *: *:**:*:.***...*:*.*:** ** :** .*    *:  : *.:* :. *

MCR   LKSDQGLELALKLIAKADVLIEGYRPGVTERLGLGPEECAKVNDRLIYARMTGWGQTGPR 120
HR    LKQPRGAAVLRRLCKRSDVLLEPFRRGVMEKLQLGPEILQRENPRLIYARLSGFGQSGSF 116
RR    LKRSPGAAVLRRMCARADVLLEPFRCGVMEKLQLGPETLRQDNPKLIYARLSGFGQSGIF 116
      **   *  :  ::  ::***:* :* ** *:* ****   : * :*****::*:**:*

MCR   SQQAGHDINYISLNGILHAIGRGDERPVPPLNLVGDFGGGSMFLLVGILAALWERQSSGK 180
HR    CRLAGHDINYLALSGVLSKIGRSGENPYAPLNLLADFAGGGLMCALGIIMALFDRTRTGK 176
RR    SKVAGHDINYVALSGVLSKIGRSGENPYPPLNLLADFGGGGLMCTLGILLALFERTRSGL 176
      .: *******::*.*:*  ***..*.* .****:.**.**.::  :**: **::*  :*

MCR   GQVVDAAMVDGSSVLIQMMWAMRATGMWTDTRGANMLDGGAPYYDTYECADGRYVAVGAI 240
HR    GQVIDANMVEGTAYLSSFLWKTQKSSLWEAPRGQNMLDGGAPFYTTYRTADGEFMAVGAI 236
RR    GQVIDANMVEGTAYLSTFLWKTQAMGLWAQPRGQNLLDGGAPFYTTYKTADGEFMAVGAI 236
      ***:** **:*:: *  ::*  :  .:*  .** *:******:* **. ***.::*****

MCR   EPQFYAAMLAGLGLDAAELPPQNDRARWPELRALLTEAFASHDRDHWGAVFANSDACVTP 300
HR    EPQFYELLIKGLGLKSDELPNQMSMDDWPEMKKKFADVFAKKTKAEWCQIFDGTDACVTP 296
RR    EPQFYTLLLKGLGLESEELPSQMSIEDWPEMKKKFADVFARKTKAEWCQIFDGTDACVTP 296
      *****  :: ****.: *** * .   ***::  :::.** : : .*  :* .:******

MCR   VLAFGEVHNEPHIIERNTFYEANG-GWQPMPAPRFSRT--ASSQPRPPAATIDIEAVLTD 357
HR    VLTFEEVVHHDHNKERGSFITSEEQDVSPRPAPLLLNTPAIPSFKRDPFIGEHTEEILEE 356
RR    VLTLEEALHHQHNRERGSFITDEEQHACPRPAPQLSRTPAVPSAKRDPSVGEHTVEVLKD 356
      **:: *. :. *  **.:*   :     * *** : .*   .*  * *    .   :* :

MCR   WDG---------------------- 360
HR    FGFSREEIYQLNSDKIIESNKVKASL 382
RR    YGFSQEEIHQLHSDRIIESNKLKANL 382
      :.
```

Figure 2
Comparison of the AMACR amino-acid sequences of *M. tuberculosis* (MCR), human (HR) and rat (RR). Identical amino acids are marked '*'; high-similarity and low-similarity amino acids are indicated by ':' and '.', respectively.

measurements were performed with a DynaPro instrument (Protein Solutions). The concentration of the protein sample used for the measurement was 3 mg ml^{-1}. The temperature was maintained at 277 K during the experiment.

2.3. Crystallization

Crystallization was performed at 295 K using the hanging-drop vapour-diffusion method (Zeelen *et al.*, 1994). In all experiments the crystallization drops, containing 1 µl protein solution and 1 µl reservoir solution, were equilibrated against 1 ml reservoir solution. The first crystallization screening was performed with a protein concentration of 7 mg ml^{-1}. Several conditions produced needles. Further optimization of these conditions revealed that the best precipitants for MCR crystallization are ammonium phosphate and ammonium sulfate. Crystallization trials with these precipitants under different conditions produced better small bipyramidal crystals with ammonium phosphate than with ammonium sulfate. Further optimization of several conditions at different pH values produced the best crystals at pH 7.0. Crystallizations were performed using different protein concentrations. The best crystals grow in 1.26 *M* ammonium phosphate pH 7.0 using a protein concentration of 24 mg ml^{-1}. These crystals (Fig. 3) grow to their maximum size within three weeks.

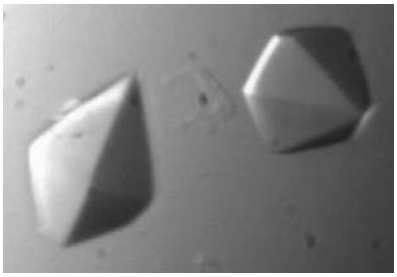

Figure 3
Crystals of α-methylacyl-CoA racemase from *M. tuberculosis*. Approximate dimensions are 0.17 × 0.1 × 0.1 mm.

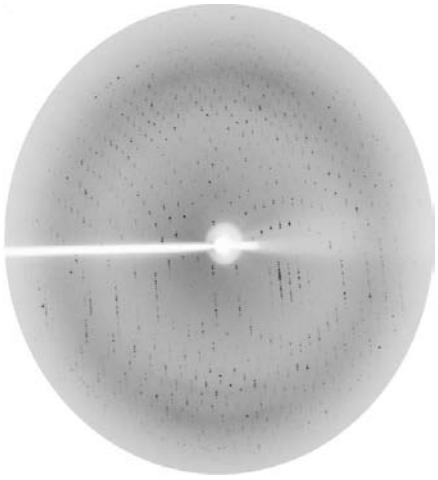

Figure 4
Diffraction pattern of MCR crystal recorded at the beamline I711, Max-lab, Lund with a MAR CCD detector. The wavelength was 1.097 Å. The high-resolution limit is 2.8 Å.

2.4. X-ray diffraction study

Diffraction data were collected under liquid-nitrogen cryoconditions at 100 K using 1.26 *M* ammonium phosphate solution pH 7.0, also containing 30% glycerol as cryoprotectant. Crystals were flash-cooled by rapidly moving them into a cold nitrogen stream. A data set was collected by the rotation method with 0.5° rotation per frame at a wavelength of 1.097 Å at beamline I711, Max-lab, Lund, Sweden. Fig. 4 shows a diffraction pattern obtained from these crystals. The image frames of this data set were indexed and integrated using the programs *DENZO* and *SCALEPACK* (Otwinowski & Minor, 1997). Table 1 summarizes the properties and data-collection statistics of this crystal form.

3. Results and discussion

MCR has been purified to homogeneity. The molecular weight of the polypeptide as judged from SDS–PAGE is 39 kDa. Dynamic light-scattering measurements showed the protein sample to be monodisperse. The calculated molecular weight from these light-scattering measurements is 75 kDa, indicating the racemase to be a dimer in solution. The specific activity was found to be 0.2 units mg^{-1}, very similar to that observed for human and rat AMACRs (Schmitz *et al.*, 1994, 1995). Processing of the diffraction data showed that the crystals belonged to the hexagonal lattice, with unit-cell parameters $a = b = 122.0$, $c = 256.4$ Å. Integration of the data with *SCALEPACK* revealed possible space groups to be *P*622, $P6_122$, $P6_522$, $P6_222$, $P6_422$ or $P6_322$. From the systematic absences of the 00*l* reflections, it can be concluded that the space group is $P6_122$ or $P6_522$. The V_M value (Matthews, 1968) for this crystal form is 4.0 Å^3 Da^{-1} for a dimer in the asymmetric unit. The self-rotation function, as calculated using *GLRF* (Tong & Rossmann, 1990), finds one significant peak for a non-crystallographic twofold axis, in good agreement with the notion that MCR has been crystallized with a dimer in the asymmetric unit.

Currently, no structure has been reported of any protein that has significant sequence similarity with MCR. It is therefore not possible to use molecular replacement to solve the structure of MCR. Attempts are under way to solve the structure using various approaches for obtaining initial phasing. Crystallographic analysis of this protein will provide a basis for elucidating the structure of other mammalian AMACRs and for understanding the catalytic reaction mechanism of these enzymes.

Table 1
Space group and data-collection statistics.

Values in parentheses are for the highest resolution shell.

Temperature (K)	100
Wavelength (Å)	1.097
Space group	$P6_122$ or $P6_522$
Unit-cell parameters (Å)	$a = b = 122.0$, $c = 256.4$
Resolution (Å)	35.0–3.0 (3.11–3.0)
Mosaicity (°)	0.374
Observed reflections	249334
Unique reflections	23160 (2281)
Redundancy	10.8 (10.5)
$I/\sigma(I)$	18.3 (6.7)
Completeness (%)	98.9 (99.7)
R_{merge} (%)	12.0 (39.7)

We would like to thank the staff at Max-lab, Lund for their help during the data collection at beamline I711. This work was supported by grants from the Academy of Finland and from the Sigrid Juselius Foundation.

References

Cole, S. T. *et al.* (1998). *Nature (London)*, **393**, 537–544.
Ferdinandusse, S., Denis, S., Clayton, P. T., Graham, A., Rees, J. E., Allen, J. T., McLean, B. N., Brown, A. Y., Vreken, P., Waterham, H. R. & Wanders, R. J. A. (2000). *Nature Genet.* **24**, 188–191.
Kotti, T., Savolainen, K., Helander, H. M., Yagi, A., Novikov, D. K., Kalkkinen, N., Conzelmann, E., Hiltunen, K. J. & Schmitz, W. (2000). *J. Biol. Chem.* **275**, 20887–20895.
Matthews, B. W. (1968). *J. Mol. Biol.* **33**, 491–497.
Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.
Pedersen, J. I., Veggan, T. & Björkhem, I. (1996). *Biochem. Biophys. Res. Commun.* **244**, 37–42.
Schmitz, W., Albers, C., Fingerhut, R. & Conzelmann, E. (1995). *Eur. J. Biochem.* **231**, 815–822.
Schmitz, W., Fingerhut, R. & Conzelmann, E. (1994). *Eur. J. Biochem.* **222**, 313–323.
Shefer, S., Cheng, F. W., Batta, A. K., Dayal, B., Tint, G. S., Salen, G. & Mosbach, E. H. (1978). *J. Biol. Chem.* **253**, 6386–6392.
Shieh, W. R. & Chen, C. S. (1993). *J. Biol. Chem.* **268**, 3487–3493.
Steinberg, D. (1995). *The Metabolic Basis of Inherited Disease*, ch. 7, edited by C. R. Scriver, A. L. Beaudet, W. S. Sly & D. Vallé, pp. 2351–2369. New York: McGraw–Hill.
Thompson, J. D., Gibson, T. J., Plewniak, F., Jeanmougin, F. & Higgins, D. (1997). *Nucleic Acids Res.* **24**, 4876–4882.
Tong, L. A. & Rossmann, M. G. (1990). *Acta Cryst.* A**46**, 783–792.
Zeelen, J. P., Hiltunen, J. K., Ceska, T. A. & Wierenga, R. K. (1994). *Acta Cryst.* D**50**, 443–447.