Expression, purification and crystallization of the (3R)-hydroxyacyl-ACP dehydratase HadAB complex from Mycobacterium tuberculosis

Yu Dong a,b,1, Jun Li a,1, Xiaodi Qiu a,b, Chuanqiang Yan a,b, Xuemei Li a,b

a National Laboratory of Biomacromolecules, Institute of Biophysics, Chinese Academy of Sciences, 15 Datun Road, Beijing 100101, People’s Republic of China
b University of Chinese Academy of Sciences, Beijing 100049, People’s Republic of China

1 These authors contributed equally to this work.

A R T I C L E   I N F O

Keywords:
Mycobacterium tuberculosis
HadAB
Mycolic acid
Dehydratase
Co-expression
Cry stallization
Drug target

A B S T R A C T

The (3R)-hydroxyacyl-ACP dehydratase HadAB, involved in the biosynthetic pathway for mycolic acid (MA) of Mycobacterium tuberculosis, catalyzes the third step in the fatty acid (FA) elongation cycle, which is an ideal and actual target for anti-tubercular agent. Though HadAB is predicted to be a member of the hotdog superfamily, it shares no sequence identity with typical hotdog fold isoenzyme FabZ. To characterize the significance of HadAB from the perspective of structural biology, large amount of pure HadAB complex is required for biochemical characterization and crystallization. Here, we used a unique expression and purification method. HadA and HadB were cloned separately and co-expressed in Escherichia coli. After GST affinity chromatography, two steps of anion exchange chromatography and gel filtration, the purity of the protein as estimated by SDS–PAGE was >95%. Using hanging-drop vapor-diffusion method, crystals were obtained and diffracted X-rays to 1.75 Å resolution. The crystal belongs to space group P412121 with unit-cell parameters a = b = 82.0 Å, c = 139.8 Å, α = β = γ = 90°.

© 2015 The Authors. Published by Elsevier Inc. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

1. Introduction

Tuberculosis (TB), caused by Mycobacterium tuberculosis (Mtbb), is the second most common cause of death due to a single infectious pathogen with an enormous global medical burden [1,2]. The complicated cell envelope of Mtb contains many kinds of lipids which play vital roles in the physiology and pathogenicity of this bacterium [3,4]. Mycolic acids (MAs), 2-alkyl, 3-hydroxy long-chain (C60–C90) fatty acids, constitute the major and specific lipid component of the envelope. They either link covalently to the cell wall peptidoglycan via the underlying arabinogalactan to form Mycolic acid–Arabinogalactan-Peptidoglycan complex (MAPc) or exist in the mycomembrane as trehalose mono- and dimyculates [5]. This provides a thick layer of lipid in the cell wall and protects the Mtb from dry environment, poisonous chemicals and host’s immune system [6,7].

The biosynthetic pathway of these MAs involves two types of fatty acid-synthesizing systems (FASs): the multifunctional FAS-I and the dissociated FAS-II [8] (Fig. 1). The latter is composed of a series of discrete soluble enzymes which act successively and repetitively to elongate the FA chains produced by FAS-I [7]. There are four kinds of enzymes catalyze each cycle of elongation in Mtb: the β-ketoacyl-ACP synthetases (KasA and KasB, Rv2245 and Rv2246, respectively. Enzyme Classification No.: 2.3.1.41), the β-ketoacyl-ACP reductase (MabA, Rv1483. EC: 1.1.1.100), the β-hydroxyacyl-ACP dehydratases (HadAB and HadBC, Rv0635–Rv0636 and Rv0636–Rv0637. EC Number is not registered), and the trans-2-enoyl-ACP reductase (InhA, Rv1484. EC: 1.3.1.9). Deficiency or inactivation of anyone above will cease the biosynthesis of MA and absence of these enzymes in human make them very ideal and actual targets for drug discovery [9,10].

The (3R)-hydroxyacyl-ACP dehydratase HadAB, composed of two subunits HadA (Rv0635) and HadB (Rv0636), catalyzes the third step in the FA elongation cycle by dehydrating β-hydroxyacyl-ACP to trans-2-enoyl-ACP, which is the last piece to be identified in the mycobacterial FAS-II [11]. HadAB would take part, like KasA, in the early FA elongation cycles, leading to the formation of the intermediate-size (C32–C42) meromycolic chains, while HadBC, like KasB, would elongate further the intermediate-size meromycolic chains to full-size molecules (C52–C64) during the late elongation cycles [11,12]. It has been verified that NAS-21, NAS-91, some kinds of flavonoids and an anti-TB prodrug Thiaceatazone (TAC) can inhibit the activity of HadB (Rv0636) as well as the growth of bacteria at different levels by mutation and overexpression methods [13–16]. Therefore HadAB complex is a significant drug target.

Though HadAB is predicted to be a member of hotdog superfamily, it resides only in Corynebacterineae and presents a different...
or distantly related catalytic sequence motif compared to the typical isoenzyme FabZ (EC: 4.2.1.59) which widely represented in other bacteria [11,17,18]. Currently, all of the resolved FabZ structures have similar hexametric hotdog fold structures, displaying a classic "trimer of homodimer" organization [19–22]. Because of the particular long substrate and failing to identify a specific mycobacterial homologue by BLAST searches of E. coli FabA (EC: 4.2.1.59 and 5.3.3.14) [17], another typical member of hotdog family [23], it seems possible that HadAB possesses new structural features to execute dehydration process. Determination of the crystal structure of HadAB would allow thorough understanding of the differences and similarities between HadAB and FabZ and lay the foundation for new anti-TB drug development.

In this work, because common or tandem expression could not get enough protein with high purity and stability for crystallization, a method for co-expression of the recombinant HadAB was developed. The gene of HadA and HadB was cloned respectively into different vectors, co-transformed into E. coli cells and the enzyme HadAB complex was expressed. Then we described the purification, crystallization and preliminary X-ray diffraction analysis of HadAB from Mycobacterium tuberculosis. Our work provided a particular method of protein complex expression and purification and laid the foundation for structure determination of HadAB and more similar enzymic complexes.

2. Materials and methods

In this work, we used pGEX-6p-1, Glutathione Sepharose resin, HiTrap™ Q HP anion exchange column, RESOURCE™ Q anion exchange column and Superdex 75 10/300 GL gel-filtration column from GE Healthcare (USA); pRSFDuet-1 from Novagen (USA); SDS and isopropyl-β-d-thiogalactoside (IPTG) from Sigma (USA); tryptone, yeast extract and protein standards for electrophoresis from Thermo Fisher Scientific (USA); tris, glycine, DTT and Coomassie G-250 and R-250 from Amresco (USA); ultrafiltration equipment from Millipore (USA); crystallization screening kits from Hampton Research (USA); DNA extraction kit from QIAGEN...
2.1. Molecular cloning

The Rv0635 gene encoding HadA was amplified from M. tuberculosis H37Rv genomic DNA by using ExTaq polymerase (TaKaRa) as well as the forward primer 5’- TTCTATATGTCGGCGCTGTA GCGCAGACAT-3’ and the reverse primer 5’- CGCAGCTGAGTC GCGCAGCCCATCAGAAA-3’. Ndel and Xhol restriction sites are underlined. Amplification were carried out by a thermal cycler (Bio-Rad) according to the following protocol: step 1, 95°C for 30 s; step 2, 94°C for 40 s; step 3, 60°C for 40 s; step 4, 72°C for 32 s; repeat from step 2 for 30 cycles. The PCR products were subjected to a final amplification step of 5 min at 72°C. They were separated by horizontal electrophoresis in 1% agarose gel. The product corresponding to HadA (about 500 bp) was excised from the gel using the DNA extraction kit (QIAGEN). Then the purified PCR product was hydrolyzed with endonucleases NdeI and XhoI and ligated according to the manufacturer’s protocol with expression vector pRSFDuet-1 (Novagen), which was preliminarily treated with the same endonucleases. The Rv0636 gene from the same source which encodes HadB was PCR-amplified using primers pBm_F (5’- CGCGGATCC ATGGCGCTGCGTGAGTTCAG-3’) and pXh_R (5’- CGCAGCTGAGTC GCGCAGCCCATCAGAAA-3’). BamHI and XhoI restriction sites are underlined. Amplification protocol is the same as HadA. After restriction digestion, the gene was ligated into the expression vector pRSFDuet-1 (Novagen), which was preliminarily treated with the same endonucleases.

Table 2

<table>
<thead>
<tr>
<th>Data collection</th>
<th>P41212</th>
</tr>
</thead>
<tbody>
<tr>
<td>Space group</td>
<td>P41212</td>
</tr>
<tr>
<td>Unit-cell parameters</td>
<td></td>
</tr>
<tr>
<td>a, b, c (Å)</td>
<td>82.0, 139.8</td>
</tr>
<tr>
<td>a = b = 90.0, c = 90.0</td>
<td>1.00000</td>
</tr>
<tr>
<td>Resolution (Å)</td>
<td>50.00–1.75 (1.81–1.75)</td>
</tr>
<tr>
<td>Total observation</td>
<td>462,997</td>
</tr>
<tr>
<td>Unique reflections</td>
<td>48,837</td>
</tr>
<tr>
<td>Data completeness (%)</td>
<td>99.9 (100.0)</td>
</tr>
<tr>
<td>Rmerge b</td>
<td>0.058 (0.424)</td>
</tr>
<tr>
<td>l(</td>
<td>d</td>
</tr>
<tr>
<td>Redundancy</td>
<td>9.5 (9.6)</td>
</tr>
</tbody>
</table>

a Values in parentheses are for the highest resolution shell.

b Rmerge = ΣhΣi|ih − 〈ih〉|/ΣhΣi〈ih〉, where 〈ih〉 is the mean intensity of the observations of ih of reflection h.

(Germany): restrictionase BamHI, Ndel, Xhol, T4 DNA ligase and corresponding buffer solution from TaKaRa(China).

All other reagents were of high purity or reagent grade (China).

2.1. Molecular cloning

The Rv0635 gene encoding HadA was amplified from M. tuberculosis H37Rv genomic DNA by using ExTaq polymerase (TaKaRa) as well as the forward primer 5’- TTCTATATGTCGGCGCTGTA GCGCAGACAT-3’ and the reverse primer 5’- CGCAGCTGAGTC GCGCAGCCCATCAGAAA-3’. Ndel and Xhol restriction sites are underlined. Amplification were carried out by a thermal cycler (Bio-Rad) according to the following protocol: step 1, 95°C for 30 s; step 2, 94°C for 40 s; step 3, 60°C for 40 s; step 4, 72°C for 32 s; repeat from step 2 for 30 cycles. The PCR products were subjected to a final amplification step of 5 min at 72°C. They were separated by horizontal electrophoresis in 1% agarose gel. The product corresponding to HadA (about 500 bp) was excised from the gel using the DNA extraction kit (QIAGEN). Then the purified PCR product was hydrolyzed with endonucleases Ndel and Xhol and ligated according to the manufacturer’s protocol with expression vector pRSFDuet-1 (Novagen), which was preliminarily treated with the same endonucleases. The Rv0636 gene from the same source which encodes HadB was PCR-amplified using primers pBm_F (5’- CGCGGATCC ATGGCGCTGCGTGAGTTCAG-3’) and pXh_R (5’- CGCAGCTGAGTC GCGCAGCCCATCAGAAA-3’). BamHI and Xhol restriction sites are underlined. Amplification protocol is the same as HadA. After restriction digestion, the gene was ligated into the expression vector pRSFDuet-1 (Novagen), which was preliminarily treated with the same endonucleases.
The column was washed with 30–50 column volumes of Buffer A. Sepharose resin (GE Healthcare) pre-equilibrated with Buffer A. The supernatant was loaded onto a column with 2 mL Glutathione HCl, pH 8.0, 300 mM NaCl, and 10% (v/v) glycerol. It was further purified using a 5 mL HiTrap™ Q HP anion exchange column (GE Healthcare) with an AKTA purifier (GE Healthcare) machine at 16 °C. The column was pre-equilibrated with Buffer B and the protein sample was loaded onto the column with 1.5 mL min⁻¹ flow rate. After thorough washing with Buffer B for 2–3 column volumes, protein bound to the column was eluted with a linear gradient (0.05–0.5 M) of NaCl concentration in 8 column volumes with 1.5 mL min⁻¹ flow rate. The linear gradient of NaCl concentration was created by mixing Buffer B and Buffer C (50 mM Tris–HCl, pH 8.0, 1 M NaCl and 10% (v/v) glycerol) by AKTA purifier machine. Fractions containing protein were analyzed by SDS–PAGE. Concentration of the separating gel and concentrating gel was 16% and 4%, respectively. Gels were stained with Coomassie R-250. Those fractions that contained pure target protein were pooled, desalted with Buffer B and then further purified by using a 1 mL RESOURCE™ Q anion exchange column (GE Healthcare). After processing in the same way, the pure target protein was pooled, concentrated to a volume less than 500 µL and injected into a 24 mL Superdex™ 75 10/300 GL gel-filtration column (GE Healthcare) pre-equilibrated with Buffer D (50 mM Tris–HCl, pH 8.0, 150 mM NaCl, and 10% (v/v) glycerol). The protein sample was eluted with 0.5 mL min⁻¹ flow rate and collected by fractions. After SDS–PAGE analysis, fractions containing pure and homogenous HadAB was pooled and concentrated by ultrafiltration using Amicon centrifugal device (Millipore). The protein concentration was calculated by measuring the sample absorbance at 280 nm and adjusting with its extinction coefficient of 1.15. Purified HadAB was finally concentrated to 10 mg mL⁻¹ for crystallization trials. The purification conditions of HadAB in each step was summarized in Table 1.

### 2.3. Crystallization

Initial crystallization screening was carried out at 16 °C by the hanging-drop vapor-diffusion method in 96-well plates (XtalQuest) using a Mosquito™ liquid-handling system (TPP Labtech). Each hanging drop was set up by mixing 200 nL protein reservoir solution. Final optimized crystals for data collection were obtained from a condition consisting of 6 mg mL⁻¹ protein, 26% (w/v) PEG, 4000, 100 mM HEPES, pH 7.5.

### 2.4. Data collection and processing

Data collection was performed at 100 K on beamline BL17U at Shanghai Synchrotron Radiation Facility (SSRF) after crystals were soaked in the reservoir solution supplemented with 20% (v/v) glycerol for 10 s and flash frozen in liquid nitrogen. Crystals diffracted X-rays to 1.75 Å resolution and a data set was collected. HKL2000 [25] software was used for data processing.

### 3. Results and discussion

Our previous work indicated that, when HadB was expressed alone, most of the His-tagged product was inclusion body, and
the purified protein was relatively unstable in solution. Meanwhile, although HadAB complex could be obtained by a tandem expression method, we were unable to get high quality crystals from this sample by reason of the His-tag fused onto HadA. To improve our sample, co-expression was performed for GST tagged HadB and none tagged HadA. Purer HadAB complex was successfully overexpressed in *E. coli* and obtained. It is more conducive to crystallization and qualified diffraction data collection. After expression, the protein was purified to homogeneity by a series of chromatographic techniques which included GST affinity purification at first using a Glutathione Sepharose column (GE Healthcare), ion exchange chromatography next using a HiTrap Q HP and a Resource Q anion-exchange column (GE Healthcare) and size-exclusion chromatography finally using a Superdex 75 10/300 GL gel filtration column (GE Healthcare). GST affinity chromatography could remove most of impure proteins. However, this process will lose a lot of target protein because of insufficient Precision Protease digestion (Fig. 2). After two kinds of anion exchange chromatography purification, redundant HadA and most of the molecular chaperone could be separated (Fig. 3a and b). If either HiTrapSM Q HP or RESOURCEQ was used, even twice, purification effect will be reduced. Remarkably, the HadAB complex elution performed as a perfect single peak during gel filtration chromatography (Fig. 4). The final purity of sample after gel filtration as estimated by SDS–PAGE was >95%. An equal molar ratio of HadA and HadB for complex formation could be deduced from SDS–PAGE as well (Fig. 4). An elution peak volume of 9.3 mL corresponded to an approximate molecular mass of 65 kDa based on calibration of the column using standard molecular weight markers. Given the theoretical molecular weight of HadAB complex is 32.5 kDa, our result indicated that HadAB exists as a tetramer at 16 °C. After initial crystallization screening, crystals were obtained in one of the conditions consisting of 22% (w/v) PEG 3000, 100 mM HEPES, pH7.5 after 2 days of incubation at 16 °C. After multiple optimization of crystallization conditions, crystals grew up to 50 μm (Fig. 5a). This crystal diffracted X-rays to 1.75 Å resolution on beamline BL17U at SSRF (Shanghai, China) (Fig. 5b). A data set was collected and the statistics of data collection are listed in Table 2. The crystal belonged to space group P4_2_2, with unit-cell parameters a = b = 82.0 Å, c = 139.8 Å, α = β = γ = 90.0°. A Matthews coefficient of 3.56 Å³ Da⁻¹ [26,27], corresponding to a solvent content of 65.49%, indicated the presence of both one molecule of HadA and HadB per asymmetric unit. In subsequent studies, an anomalous data set collected using crystals of selenomethionine labeled HadAB was used for phasing by single-wavelength anomalous diffraction (SAD) method. A total of four selenium sites were located by SHELX program suit [28] and the phase was solved. A density map was created by using intensities in the collected data and the solved phase. The 3D structural model of HadAB was built according to the density and refinement is currently in progress. Expectingly, the crystal structure of HadAB will help us understand the mechanism of substrate dehydration as well as biosynthesis of mycolic acids, and also lay the foundation for anti-TB drug discovery.

Conflict of interest

The authors declare that they have no conflict of interest.

Acknowledgements

This work was supported by Grants from the Strategic Priority Research Program of the Chinese Academy of Sciences (Grant No. XDB08020200), State Key Development Program for Basic Research of the Ministry of Science and Technology of China (973 Project Grant Nos. 2014CB542800, 2014CB940203, 2011CB915001 and 2011CB910304), and National Science Foundation (China) Grant 813100237.

References