Journal of Structural Biology 193 (2016) 172–180

Contents lists available at ScienceDirect

Journal of Structural Biology

journal homepage: www.elsevier.com/locate/yjsbi

Crystal structure of Rv2258c from *Mycobacterium tuberculosis* H37Rv, an *S*-adenosyl-L-methionine-dependent methyltransferase

Ha Na Im [a], Hyoun Sook Kim [b,c], Doo Ri An [a], Jun Young Jang [b], Jieun Kim [a], Hye-Jin Yoon [b], Jin Kuk Yang [d], Se Won Suh [a,b,*]

[a] Department of Biophysics and Chemical Biology, College of Natural Sciences, Seoul National University, Seoul 08826, Republic of Korea
[b] Department of Chemistry, College of Natural Sciences, Seoul National University, Seoul 08826, Republic of Korea
[c] Research Institute of Pharmaceutical Sciences, College of Pharmacy, Seoul National University, Seoul 08826, Republic of Korea
[d] Department of Chemistry, College of Natural Sciences, Soongsil University, Seoul 06978, Republic of Korea

ARTICLE INFO

Article history:
Received 21 December 2015
Accepted 5 January 2016
Available online 6 January 2016

Keywords:
Rv2258c
Mycobacterium tuberculosis
Small-molecule methyltransferase
Sinefungin
S-adenosyl-L-homocysteine

ABSTRACT

The *Mycobacterium tuberculosis* Rv2258c protein is an *S*-adenosyl-L-methionine (SAM)-dependent methyltransferase (MTase). Here, we have determined its crystal structure in three forms: a ligand-unbound form, a binary complex with sinefungin (SFG), and a binary complex with *S*-adenosyl-L-homocysteine (SAH). The monomer structure of Rv2258c consists of two domains which are linked by a long α-helix. The N-terminal domain is essential for dimerization and the C-terminal domain has the Class I MTase fold. Rv2258c forms a homodimer in the crystal, with the N-terminal domains facing each other. It also exists as a homodimer in solution. A DALI structural similarity search with Rv2258c reveals that the overall structure of Rv2258c is very similar to small-molecule SAM-dependent MTases. Rv2258c interacts with the bound SFG (or SAH) in an extended conformation maintained by a network of hydrogen bonds and stacking interactions. Rv2258c has a relatively large hydrophobic cavity for binding of the methyl-accepting substrate, suggesting that bulky nonpolar molecules with aromatic rings might be targeted for methylation by Rv2258c in *M. tuberculosis*. However, the ligand-binding specificity and the biological role of Rv2258c remain to be elucidated due to high variability of the amino acid residues defining the substrate-binding site.

1. Introduction

Mycobacterium tuberculosis is a highly successful intracellular pathogen, infecting nearly one-third of the world's population. It causes tuberculosis (TB), claiming the lives of millions of people in the world every year. The advent of multidrug-resistant TB cases, the HIV epidemic, imperfect diagnostic assays, limited vaccine efficacy, and non-availability of new anti-TB drugs pose a global health problem (Lin and Flynn, 2010). Therefore, worldwide efforts are being made to develop new anti-TB drugs and more effective vaccines to combat TB. As an important step, the genome sequence of *M. tuberculosis* H37Rv strain was reported in 1998 (Cole et al., 1998). However, we still have little or limited functional information on a significant portion of approximately 4090 genes in *M. tuberculosis* (Lew et al., 2011). Identifying the molecular and biological functions of the proteins that are encoded by the *M. tuberculosis* genome would provide the groundwork for the development of new anti-TB drug targets.

Methyltransferases (MTases) mediate a wide variety of cellular processes, such as cell signaling, metabolite synthesis, and gene regulation in nearly all living organisms. They comprise a large family of over 300 members and transfer a methyl group most frequently from *S*-adenosyl-L-methionine (SAM) to various acceptor substrates, which include small molecules, lipids, proteins, and nucleic acids, yielding a methylated product with *S*-adenosyl-L-homocysteine (SAH) as a by-product. The Rv2258c protein from *M. tuberculosis* H37Rv and its close orthologs in mycobacteria are annotated as SAM-dependent methyltransferases and possible transcriptional regulatory proteins. To gain structural insights into

Abbreviations: MTase, methyltransferase; SAM, *S*-adenosyl-L-methionine; SAH, *S*-adenosyl-L-homocysteine; SFG, sinefungin; r.m.s., root mean square; TB, tuberculosis.

* Corresponding author at: Department of Chemistry, College of Natural Sciences, Seoul National University, Seoul 08826, Republic of Korea.
E-mail address: sewonsuh@snu.ac.kr (S.W. Suh).

http://dx.doi.org/10.1016/j.jsb.2016.01.002
1047-8477/

the function of Rv2258c, specifically, to provide a structural basis to decipher substrate binding and specificity, we have determined its crystal structure in three forms: ligand-unbound Rv2258c, a binary complex with sinefungin (SFG), and a binary complex with SAH. A monomer of Rv2258c consists of two domains, which are linked by a long connecting α-helix. The N-terminal domain is essential for dimerization and the C-terminal domain has the Class I MTase fold. The structure of Rv2258c is distinct from eleven other mycobacterial SAM-dependent MTases that have been structurally characterized but the overall fold of Rv2258c resembles those of small-molecule *O*-MTases. Rv2258c forms a homodimer in the crystal, with the N-terminal domains facing each other. Size-exclusion chromatography confirms that Rv2258c also exists as a homodimer in solution. Our structure reveals that the Rv2258c dimer has a large cavity for binding a methyl-accepting substrate adjacent to the SFG (or SAH) binding site. It also shows that dimerization is essential to form such a cavity, as it is contributed not only by two domains and the connecting α-helix of one monomer but also by an α-helix in the N-terminal domain of the other monomer. Due to high variability of the amino acid residues defining the substrate-binding site, further experiments are required to establish the methyl-accepting substrate and the biological role of Rv2258c.

2. Materials and methods

2.1. Protein expression and purification

To obtain well-diffracting crystals, six constructs encompassing residues 1–353 (full-length), 1–320, 4–353, 6–353, 9–353, and 28–353 of the Rv2258c protein from *M. tuberculosis* H37Rv were generated. The genes encoding these constructs were amplified by PCR using the genomic DNA of *M. tuberculosis* H37Rv as the template and were cloned into the expression vector pET-28b(+) (Novagen) using NdeI and XhoI restriction enzyme sites. The resulting recombinant proteins are fused with hexa-histidine containing tags at both N- and C-termini (MGSSHHHHHHSSGLVPRGSH and LEHHHHHH, respectively). They were overexpressed in *Escherichia coli* Rosetta 2(DE3) cells. The cells were grown at 37 °C in Luria Broth culture medium containing 30 μg/ml kanamycin. Protein expression was induced by 0.5 mM isopropyl β-D-thiogalactopyranoside and the cells were incubated for additional 20 h at 30 °C. The cells were harvested by centrifugation at 5600*g* for 15 min at 4 °C and subsequently lysed by sonication in ice-cold buffer A (20 mM Tris–HCl at pH 7.9, 500 mM sodium chloride, and 50 mM imidazole), which was supplemented with 10% (*v*/*v*) glycerol and 1 mM phenylmethylsulfonyl fluoride. The crude lysate was centrifuged at 36,000*g* for 1 h at 4 °C to discard the cell debris. The supernatant was applied to an affinity chromatography column of HiTrap Chelating HP (GE Healthcare), which was previously equilibrated with buffer A. The column was eluted with buffer B (20 mM Tris–HCl at pH 7.9, 500 mM sodium chloride, and 500 mM imidazole), with the Rv2258c protein being eluted at 120–180 mM imidazole concentration. The eluted protein was further purified by gel filtration on a HiLoad 16/60 Superdex 200 prep-grade column (GE Healthcare), which was previously equilibrated with buffer C (20 mM Tris–HCl at pH 7.0 and 200 mM sodium chloride). The protein purity was analyzed by SDS–PAGE. Fractions containing the recombinant Rv2258c were pooled and concentrated to 17 mg/ml (0.42 mM monomer concentration) for crystallization using an Amicon Ultra-15 Centrifugal Filter Unit (Millipore). Four of the above constructs, 1–353 (full-length), 4–353, 6–353, and 9–353, were expressed in a soluble form in *E. coli* and were purified for crystallization trials. The 1–320 and 28–353 constructs were expressed in an insoluble form.

2.2. Crystallization

Crystals were grown at 23 °C by the sitting drop vapor diffusion method using a Mosquito robotic system (TTP Labtech). The 1–353 (full-length) and 4–353 constructs produced crystals diffracting poorly to 4 Å only, while the 9–353 construct did not yield crystals. Only the 6–353 construct gave well-diffracting crystals. To obtain ligand-free crystals of the 6–353 construct ('Rv2258c-unbound'), a sitting drop was prepared by mixing 0.2 μl of the protein solution in buffer C and 0.2 μl of the reservoir solution [200 mM sodium malonate at pH 6.0 and 20% (*v*/*v*) PEG 3350]. The sitting drop was equilibrated against 100 μl of the reservoir solution. Crystals grew up to approximate dimensions of 0.2 mm × 0.3 mm × 0.2 mm within a few days.

In an effort to achieve phasing by single-wavelength anomalous diffraction (SAD), the selenomethionine-substituted Rv2258c protein (6–353 construct) was expressed in *E. coli* but it did not produce crystals. Instead, we prepared a platinum derivative of Rv2258c-unbound crystals by soaking them for 90 min in 5 μl of a heavy atom-containing cryoprotectant solution, which was prepared by supplementing the reservoir solution with 30% (*v*/*v*) glycerol and 5 mM K_2PtCl_4.

Attempts to co-crystallize the ligand-bound Rv2258c protein were not successful due to the tendency of the protein to aggregate in the presence of the ligands. Therefore, crystals of Rv2258c complexed with either sinefungin (SFG), an analog of the co-substrate *S*-adenosyl-L-methionine (SAM), or the by-product *S*-adenosyl-L-homocysteine (SAH) ('Rv2258c-SFG' and 'Rv2258c-SAH', respectively) were obtained by soaking crystals of Rv2258c-unbound for 1 min in 5 μl of a cryoprotectant solution, which was prepared by supplementing the reservoir solution with 30% (*v*/*v*) glycerol and 19.9 mM SFG (or 12.5 mM SAH).

2.3. X-ray data collection

A set of X-ray diffraction data from a ligand-free crystal of Rv2258c ('Rv2258c-unbound') was collected to 1.83 Å on an Area Detector Systems Corporation Q270 CCD detector at the beamline BL-7A of Pohang Light Source, Korea. All the raw data were processed and scaled using the program suite HKL2000 (Otwinowski and Minor, 1997). Crystals were flash-frozen in a nitrogen gas stream at 100 K. The crystal of Rv2258c-unbound belongs to the space group *C*2, with unit cell parameters of $a = 109.1$ Å, $b = 140.6$ Å, $c = 97.1$ Å, and $\beta = 98.5°$. Assuming the presence of three monomers of the recombinant Rv2258c in the asymmetric unit, the Matthew's parameter and the solvent content are 3.26 $Å^3$/Da and 62.3%, respectively. Data collection statistics are given in Table 1.

Several sets of SAD data were collected at 100 K from different platinum-derivatized crystals at a wavelength of 1.0717 Å using the Area Detector Systems Corporation Q315r CCD detector at the beamline BL-5A of Photon Factory, Japan. Data collection statistics are given in Table 1.

X-ray diffraction data from crystals of Rv2258c-SFG and Rv2258c-SAH were collected at 100 K using an Area Detector Systems Corporation Q270 CCD detector at the beamline BL-7A experimental station of Pohang Light Source, Republic of Korea and on an Area Detector Systems Corporation Q315r CCD detector at the beamline BL-5A of Photon Factory, Japan, respectively. The crystal of the Rv2258c-SFG diffracted to 1.90 Å and belongs to the space group *C*2, with unit cell parameters of $a = 108.9$ Å, $b = 140.9$ Å, $c = 96.6$ Å, and $\beta = 98.3°$. The crystal of Rv2258c-SAH diffracted to 2.90 Å and belongs to the space group *C*2, with unit cell parameters of $a = 109.6$ Å, $b = 140.5$ Å, $c = 96.1$ Å, and $\beta = 97.9°$. Both crystals contain three monomers per asymmetric unit, giving a Matthew's parameter and solvent fraction of