The crystal structure of the mycobacterial trehalose monomycolate transport factor A, TtfA, reveals an atypical fold

Kien Lam Ung, Husam Alsarraf, Laurent Kremer, Mickael Blaise

The crystal structure of the mycobacterial trehalose monomycolate transport factor A, TtfA, reveals an atypical fold

Kien Lam Ung[1], Husam M.A.B. Alsarraf[1], Laurent Kremer[1,2] and Mickaël Blaise*[1]
[1]Institut de Recherche en Infectiologie de Montpellier (IRIM), Université de Montpellier, CNRS UMR 9004, 34293 Montpellier, France.
[2]INSERM, IRIM, 34293 Montpellier, France.

*Corresponding author: Tel: (+33)434359447; E-mail: mickael.blaise@irim.cnrs.fr

Running Title: Crystal structure of TtfA

Abstract

Trehalose monomycolate (TMM) represents an essential element of the mycobacterial envelope. While synthesized in the cytoplasm, TMM is transported across the inner membrane by MmpL3 but, little is known regarding the MmpL3 partners involved in this process. Recently, the TMM transport factor A (TtfA) was found to form a complex with MmpL3 and to participate in TMM transport, although its biological role remains to be established. Herein, we report the crystal structure of the *Mycobacterium smegmatis* TtfA core domain. The phylogenetic distribution of TtfA homologues in non-mycolate containing bacteria suggests that TtfA may exert additional functions.

Keywords: mycobacteria, mycolic acids, MmpL3, TtfA, cell wall

1 Introduction

With 10 million new cases and 1.6 million deaths in 2017, tuberculosis (TB) remains a major global threat worldwide. *Mycobacterium tuberculosis*, the etiologic agent of TB , is a resilient microorganism which can persist through long courses of antibiotics. The emergence of drug-resistant bacteria largely contributes to the difficulty in curing TB patients. This involves acquired resistance mechanisms occurring during lengthy antibiotherapy and intrinsic resistance mechanisms that are linked to the atypical mycomembrane, which is impermeable to small hydrophilic molecules, such as antibiotics.

The thick and waxy mycobacterial cell envelope consists of an outermost layer named capsule and composed of (glyco)lipids and polysaccharides[1]. The mycomembrane comprises

essentially mycolic acids, which are very long-chain fatty acids (C80-C90) as well as free and extractable (glyco)lipids, such as dimycolyltrehaloses[2], that are interspersed between mycolic acids and play a key role in the host-pathogen interactions. An arabinogalactan layer is bridging together the mycolic acids and the peptidoglycan[3].The arabinogalactan-mycolyl-peptidoglycan skeleton represents the core structure of the cell wall, beneath of which is found the plasma membrane.

Many enzymes involved in cell wall biosynthesis are unique to the *Mycobacterium* genus and absent in mammals and are thereby, recognized as valuable chemotherapeutic targets[1]. Mycolic acids for instance are essential cell wall components and contribute also to this high cell wall hydrophobicity[4]. First, synthesized in the cytoplasm in the form of trehalose monomycolates (TMM), which are subsequently acetylated, mycolic acids are then transported across the plasma membrane by a specific and atypical lipid transporter, MmpL3[5]. The Ag85ABC complex catalyses the conversion of TMM into trehalose dimycolate (TDM) as well as the transfer of mycolic acids to the arabinogalactan of the mAGP complex, thereby contributing to the elaboration of the mycobacterial membrane.

During the last few years, several independent studies demonstrated that MmpL3 is an essential transporter belonging to the superfamily of the Resistance Nodulation Division (RND) permeases and represents also an attractive drug target[6]. Currently, MmpL3 represents one of the most promising antimycobacterial pharmacological target and a wide panel of chemical scaffolds have been shown to inhibit MmpL3[7] not only in *M. tuberculosis* but also in non-tuberculous mycobacteria, such as *Mycobacterium abscessus*, thus opening a new field in the inhibition of mycolic acid transport.

Two recent studies reported the crystal structure of *Mycobacterium smegmatis* MmpL3 revealing the mode of action of some of these inhibitors and providing a first glance on our understanding of TMM transport[8,9]. A *contrario* to other well studied RND pumps, such as AcrB from *Escherichia coli* that is part of a tripartite complex, the existence of MmpL3 accessory proteins was questioned. This has been partly delineated through the establishment of the MmpL3 interactome that comprises many proteins, thus supporting the occurrence of a very complex cross-talk between the cell envelope biosynthesis and cell division[10]. Additionally, the LpqN lipoprotein was demonstrated to bind to a lipid mimicking TMM and capable to interact with the periplasmic domains of MmpL3 and MmpL11 on the one hand as well as with members of Ag85 complex on the other hand [11]. Importantly, in an elegant study[12], an essential protein MSMEG_0736 interacting with MmpL3 both in *M. smegmatis* and *M. tuberculosis* was discovered. The MSMEG_0736 protein co-eluted in pull-down experiments using MmpL3 as a bait and its essential character for mycobacterial viability was demonstrated. The protein is associated with the plasma membrane where it participates in TMM transport as evidence by genetic studies based on a knock-down

MSMEG_0736 mutant that accumulates TMM while the synthesis of TDM decreases, mirrorring the phenotypes observed in either *mmpL3* conditional mutants[13] or when using specific MmpL3 inhibitors[14]. Additionally, albeit non-essential for bacterial viability the MSMEG_5308 protein was also found in the MSMEG_0736/MmpL3 complex[12]. Collectively, these results indicate that MSMEG_0736/Rv0383c is a key player in TMM transport, hence was named TMM transport factor A, TtfA.

As the exact role and function of TtfA in the TMM transport is not clearly understood and as TtfA does not share any homology to any other known protein, we decided to investigate its structural characterization. Herein, we describe the high-resolution crystal structure of the core domain of TtfA from *M. smegmatis*.

2 Materials and Methods

2.1 Cloning, expression and purification of MSMEG_0736 (TtfA)

The coding sequence for TtfA residues 24-205 was amplified by PCR using *Mycobacterium smegmatis* mc^2155 genomic DNA as a template and the following forward and reverse primers: fw: 5'-atcggtaccgagaacctgtacttccagggtcatatgatcgaccgccgacgagggctgggac-3' ; rv: 5'-gcagctcgagaagcttaattaattactagctagcgttctgcggcaccggcggcagcacc-3'. The PCR product was cloned into pET-30a between the KpnI and XhoI restriction sites, to generate an in-frame fusion with the N-terminal 6xHis-tag and S-tag and designed so that it contains also a Tobacco Etch Virus (TEV) protease cleavage site. The pET30:*MSMEG_0736_(24-205)* was transformed into the *E. coli* B834 (DE3) methionine auxotroph strain (Merck-Millipore), used for the production of both the native and the seleno-methionine proteins.

For the native protein expression, cells were grown under agitation at 180 r.p.m in LB media (6 L) at 37°C supplemented with 50 µg/mL kanamycin (KAN) until OD_{600} reached ~0.8. The cultures were cooled down on ice for 30 min and induced with 1 mM of isopropyl-*β*-D-thiogalactoside (IPTG) (Euromedex) for 16 h at 16 °C. Bacteria were collected by centrifugation at 6,000 *g* for 20 min, resuspended in buffer A (50 mM Tris-HCl pH 8, 0.4 M NaCl, 0.5 mM β-mercaptoethanol, and 1 mM benzamidine) and disrupted by sonication prior to removal of the cell debris by centrifugation at 28,000 *g* at 4°C for 1 h. The clarified supernatant was then supplemented with 10 mM of imidazole pH8, incubated with Ni-Nitrilotriacetic acid sepharose beads for 15 min and loaded onto a gravity column. Extensive washes with sixteen column volumes of buffer B (50 mM Tris-HCl pH 8, 0.4 M NaCl, 20 mM imidazole and 0.5 mM β-mercaptoethanol) were performed prior to protein elution with buffer C (50 mM Tris-HCl pH 8, 0.4 M NaCl, 300 mM

imidazole and 0.5 mM β-mercaptoethanol). The eluate was then incubated with TEV protease (1 mg of TEV protease per 40 mg of TtfA) and dialyzed overnight at 4°C against buffer D (50 mM Tris-HCl pH 8, 0.6 M NaCl, 0.5 mM β-mercaptoethanol and 5% (v/v) glycerol). A Ni-Nitrilotriacetic acid sepharose column was used to remove the uncleaved-tagged protein as well as the His-tagged TEV protease. The tag-free protein was collected in the flow-through fractions and concentrated by ultrafiltration using a 10 kDa cut-off (Centricon; Sartorius). The protein was further purified by size-exclusion chromatography on a Superdex 200 Increase 10/300 GL column (GE Healthcare) and eluted with buffer D. The purity of the protein was estimated by Coomassie Blue-stained SDS/PAGE.

The same *E. coli* strain was also used for labelling the protein with selenomethionine. Cells were first grown in 2 L of LB broth containing 50 µg/mL KAN at 37°C under agitation at 180 r.p.m. for 24 h. This starter culture was then pelleted, washed thrice with 1x PBS buffer and resuspended with pre-warmed medium A containing M9 medium, trace elements, 20% glucose, 1 M $MgSO_4$, 1 M $CaCl_2$, 1mg/L biotin, 1mg/L thiamine, and 50 µg/mL KAN to a final OD_{600} = ~ 0.8. The culture was incubated at 37°C for 1.5 h prior to addition of a mix D/L SeMet at a final concentration of 50 µg/mL (Sigma-Aldrich) and further incubated for 30 min. Finally, protein induction was done by adding 1 mM of IPTG and incubation for 2 days at 16°C. The purification procedure was identical to the one used for native TtfA, as described above, with the exception that the NaCl concentration of buffer B and C was increased to 0.6 M. Protein concentration was determined using a Nanodrop 2000c spectrophotometer (Thermo Fisher Scientific).

2.2 Determination of molecular weight by size-exclusion chromatography

The oligomeric state of TtfA core domain (24-205) in solution was assessed using a Superdex 200 10/300 GL column (GE Healthcare) and elution performed with buffer D. The molecular weight was calculated based on a calibration curve obtained by the elution profile of proteins standard with molecular weight ranging from $12,400$ to $200,000$ Da (Gel Filtration Markers Kit, Sigma-Aldrich). Dextran blue was used to determine the void volume of the column. The apparent mass was determined by plotting the partition coefficient K_{av} against the logarithm of the molecular weight of the standard proteins.

2.3 Crystallization

The MSMEG_0736 native crystals were grown in sitting drops in the Swissci 48-well MRC Maxi Optimization plates (Molecular Dimension) at 18°C by mixing 2 µL of protein solution

concentrated to 8 mg/mL with 2 µL of reservoir solution consisting of 28% PEG 4000 and 0.2 M $(NH_4)_2SO_4$. The MSMEG_0736 protein labelled with selenomethionine was crystallized using the microseeding technique and the crystals of the native protein as seeds. The selenomethionine substituted protein was crystallized at 18°C by mixing 2 µL of protein solution concentrated to 8 mg/mL with 2 µL of a reservoir solution made of 0.1 M Bis-Tris pH 5.5, 27% PEG 3350 and 0.2 M Li_2SO_4. All the crystals were cryo-cooled in liquid nitrogen without any cryoprotection.

2.3 Data processing, structure solution and refinement

Data collection was performed at the Swiss-Light Source, Villigen, Switzerland on the PXIII-X06DA beamline. Data were processed with XDS and scaled and merged with XSCALE[15]. Phasing was done by the single anomalous dispersion (SAD) method using the *AutoSol* module from the *Phenix package*[16]. Five out of the six potential selenium sites in the asymmetric unit were found using a resolution cut-off of 2 Å for the search of the Se atoms. After density modification, the electron density map was of high quality allowing almost automatic complete model building with *autobuild*. The resulting model was then used to performed molecular replacement with the 1.4 Å native dataset using *Phaser*[17] from the *Phenix package*[16]. *Coot*[18] was used for manual rebuilding while structure refinement and validation were performed with the *Phenix package*[16]. The statistics for data collection and structure refinement are displayed in Table 1. Figures were prepared with PyMOL (www.pymol.org).

3 Results and discussion

3.1 Protein expression and purification

Our cloning strategy was based on bioinformatic analysis, secondary structure prediction as well as on previous biological data whereby residues 1-205 were shown to be essential for the function of the protein[12]. The secondary structure prediction also attests that the C-terminal part of the protein that is not well conserved among other species is predicted to be highly disordered **(Figure 1A)**. Furthermore, the TtfA residues 206-278 are dispensable for TMM transport in mycobacteria and for the interaction with MmpL3[12]. We attempted to express several constructs carrying the N-terminal transmembrane helix residues 1-23 but all these proteins were not soluble and found in inclusion bodies. However, discarding residues 1 to 23 allowed us to produce a truncated TtfA containing residues 24 to 205 in a soluble form. Of note, the introduction of a TEV cleavage site and our cloning strategy added three extra residues (GHM) at the N-terminus after TEV cleavage and two extra residues (AS) at the C-terminus . However, due to the propensity of this protein to aggregate, high concentration of NaCl (0.6 M) and 5% glycerol were added in all

buffers to maintain its solubility. Overall, this allowed to purify large amounts of pure and homogeneous TtfA **(Figure 2A)** following a three-step purification procedure as described in the Materials and Methods section.

3.2 Structural analysis

TftA was crystallized and high-resolution datasets were collected for both the native and the seleno-methionine-substituted protein. The structure was solved by SAD using a dataset collected at the Se peak. Further, the structure of the native protein was solved by molecular replacement and refined to 1.4Å **(Table 1)**. The asymmetric unit contains one protein monomer and two sulfate ions. Most of the residues are well ordered and could be modelled, excepted the first 6 residues GHMIDR with R corresponding to residue 26 of the *M. smegmatis* TtfA protein and GHM corresponding to the additional residues added by our cloning strategy. In the C-terminal of the protein, we could not build N205.

Analysis of the crystal packing with the PISA server (https://www.ebi.ac.uk/pdbe/pisa/) fails to predict the existence of stable oligomers. To confirm or infirm this hypothesis, the molecular weight of the protein in solution was assessed by size-exclusion chromatography. While the theoretical mass of TtfA (25-205) is 21.3 kDa, we estimated to 14.4 kDa the apparent molecular weight of the protein **(Figure 2A)**. Overall, these results strongly suggest that the truncated TtfA protein is a monomer in solution.

The overall structure of TtfA is made of a central antiparallel β-sheet flanked by α-helices. The β-sheet is made of nine β-strands formed by the following residues : β1 (residues: 44-47), β2 (71-77), β3 (80-87), β4 (90-97), β5 (106-110), β6 (120-125), β7 (130-134), β8 (161-165) and, β9 (168-173). The central core of the structure is surrounded by four α-helices: α1 (30-41), α2 (137-143), α3 (146-1154) and α4 (179-197). The fold is finally completed by three 3_{10} helices: η1 (52-56), η2 (60-63), η3 (157-160) **(Figure 1A and 2B)**.

We next searched for structures similar to TtfA using the DALI server[19] but failed to retrieve significant hits. Only structures displaying a central antiparallel β-sheet but sharing a very low primary sequence identity matched the TtfA structure. The best Z-score obtained corresponds to the structure of tAvrPphF, the chaperone for the Type III Effector AvrPphF from *Pseudomonas syringae* (PDB entry1s28), which shares only 4% of primary sequence identity and an r.m.s.d. of 3.3 Å. This analysis attests for *a hitherto* unseen and atypical fold that characterizes TtfA.

We next searched for the presence of a potential cavity which could eventually host TMM but we could neither identify any obvious pocket able to accommodate this glycolipid nor a particular hydrophobic area that could bind the long mycolic acid chains of TMM. In the absence of biochemical experiment we cannot totally exclude that TtfA binds lipids ; but the structural analysis

suggesting that TtfA does not bind lipids is supported by the observation of the protein surface that is mainly positively charged on one side, especially towards the N-terminus region and negatively charged on the opposite side **(Figure 2C)**. However this prompted us to investigate the surface charge on the MmpL3 crystal structure which appears to be highly negative and more particularly its periplasmic domains (not shown)[8]. It is tempting to speculate that the TtfA positively-charged area may interact with these MmpL3 extracellular domains. However, it was reported that the C-terminal domain of TtfA (25-205) is cytoplasmic[12], therefore excluding the possibility that TtfA interacts with the MmpL3 periplasmic domains. This suggests that TtfA may not be necessarily a direct binding partner of MmpL3. Another possibility would be that TtfA interacts with the flexible C-terminal part of MmpL3 but we could not explore this hypothesis since the MmpL3 crystal structures available do not comprise the C-terminal domain. However, in other TtfA-containing bacteria such as corynebacteria, MmpL3 homologues do not possess this C-terminal extension. It remains also possible that TtfA interacts *via* its N-terminal transmembrane helix with the transmembrane domains of MmpL3, but this could not be experimentally assayed as we failed to purify a soluble TtfA protein with its complete N-terminal helix.

3.3 Distribution of TtfA among Actinobacteria

The search for homologues of TtfA was performed using the HMMER server. TtfA was only found in bacteria belonging to the Actinobacteria phylum as seen on the simplified phylogenetic tree in **Figure 1B**. As expected, the highest sequence identity was found with homologues from the mycobacterial genus **(Figure 1A)**. Additionally, TtfA homologues were found as reported earlier[12] in other mycolic acid-producing bacteria such as corynebacteria, nocardia or rhodococcus. Unexpectidely, although less conserved, TtfA homologues were also identified in non-producing mycolic acid bacteria, such as pseudonocardia or streptomyces. Primary sequence alignments show for example that in the *Amycolatopsis methanolica* a homologue (UniProtKB – A0A076MZQ3) of TtfA shares 21%/36% of primary sequence identity/similarity with TtfA from *M. smegmatis* **(Figure 1A).** Noteworthingly, *in silico* predictions attest for the presence of a transmembrane helix in the N-terminus of this protein. The genome of *Amycolatopsis methanolica* was next analysed for the presence of genes encoding the two binding partners reported so far for TtfA, i.e. homologues of MmpL3 and MSMEG_5308[12]. This revealed the presence of a homologue of MmpL3 sharing 29%/48% identity/similarity but no homologue for MSMEG_5308 was found. A similar observation was made in *Streptomyces regensis*. Conversely, other genomes such as the one of *Pseudonocardia dioxanivorans* encode for TtfA homologues but are lacking MmpL3-encoding genes. This strongly suggests that TtfA might not be solely dedicated to TMM transport and/or that

TtfA orthologues may exert additional biological functions notably in bacteria that neither produce mycolic acids nor possess MmpL3 homologues.

Acknowledgements

We thank Dr V. Olieric and the staff at SLS beamlines for support during data collection.

Conflicts of interest

The authors declare no potential conflict of interest.

Data accessibility

The coordinates and structure factors have been deposited to the Protein Data Bank under the accession number: 6T84

Funding body

KLU for his Ph.D. fellowship and MB are supported by the National Research Agency [ANR-17-CE11-0008-01 – MyTraM]. HMABA is supported by a fellowship from the Lundbeck Foundation. This work was supported by the Fondation pour la Recherche Médicale (FRM) [grant number DEQ20150331719] to LK.

Authors Contribution

KLU and MB designed research, performed the experiments and data analysis. HMABA and MB collected x-ray data. LK and MB supervised research. MB wrote the first draft and all the authors contributed to the proofreading of the manuscript.

References

1. Kaur D, Guerin ME, Skovierová H, Brennan PJ, Jackson M (2009) Chapter 2: Biogenesis of the cell wall and other glycoconjugates of Mycobacterium tuberculosis. Adv Appl Microbiol 69:23–78.

2. Briken V, Porcelli SA, Besra GS, Kremer L (2004) Mycobacterial lipoarabinomannan and related lipoglycans: from biogenesis to modulation of the immune response. Mol Microbiol 53:391–403.

3. Crick DC, Mahapatra S, Brennan PJ (2001) Biosynthesis of the arabinogalactan-peptidoglycan complex of Mycobacterium tuberculosis. Glycobiology 11:107R–118R.

4. Bhatt A, Molle V, Besra GS, Jacobs WR, Kremer L (2007) The Mycobacterium tuberculosis FAS-II condensing enzymes: their role in mycolic acid biosynthesis, acid-fastness, pathogenesis and in future drug development. Mol. Microbiol. 64:1442–1454.

5. Grzegorzewicz AE, Pham H, Gundi VAKB, Scherman MS, North EJ, Hess T, Jones V, Gruppo V, Born SEM, Korduláková J, et al. (2012) Inhibition of mycolic acid transport across the Mycobacterium tuberculosis plasma membrane. Nat Chem Biol 8:334–341.

6. Quémard A (2016) New Insights into the Mycolate-Containing Compound Biosynthesis and Transport in Mycobacteria. Trends Microbiol. 24:725–738.

7. Xu Z, Meshcheryakov VA, Poce G, Chng S-S (2017) MmpL3 is the flippase for mycolic acids in mycobacteria. Proc. Natl. Acad. Sci. U. S. A. 114:7993–7998.

8. Zhang B, Li J, Yang X, Wu L, Zhang J, Yang Y, Zhao Y, Zhang L, Yang X, Yang X, et al. (2019) Crystal Structures of Membrane Transporter MmpL3, an Anti-TB Drug Target. Cell 176:636-648.e13.

9. Su C-C, Klenotic PA, Bolla JR, Purdy GE, Robinson CV, Yu EW (2019) MmpL3 is a lipid transporter that binds trehalose monomycolate and phosphatidylethanolamine. Proc. Natl. Acad. Sci. U. S. A. 116:11241–11246.

10. Belardinelli JM, Stevens CM, Li W, Tan YZ, Jones V, Mancia F, Zgurskaya HI, Jackson M (2019) The MmpL3 interactome reveals a complex crosstalk between cell envelope biosynthesis and cell elongation and division in mycobacteria. Sci. Rep. 9:10728.

11. Melly GC, Stokas H, Dunaj JL, Hsu F-F, Rajavel M, Su C-C, Yu EW, Purdy GE (2019) Structural and functional evidence that lipoprotein LpqN supports cell envelope biogenesis in M. tuberculosis. J. Biol. Chem.

12. Fay A, Czudnochowski N, Rock JM, Johnson JR, Krogan NJ, Rosenberg O, Glickman MS (2019) Two Accessory Proteins Govern MmpL3 Mycolic Acid Transport in Mycobacteria. mBio 10.

13. Degiacomi G, Benjak A, Madacki J, Boldrin F, Provvedi R, Palù G, Kordulakova J, Cole ST, Manganelli R (2017) Essentiality of mmpL3 and impact of its silencing on Mycobacterium tuberculosis gene expression. Sci. Rep. 7:43495.

14. Dupont C, Viljoen A, Dubar F, Blaise M, Bernut A, Pawlik A, Bouchier C, Brosch R, Guérardel Y, Lelièvre J, et al. (2016) A new piperidinol derivative targeting mycolic acid transport in Mycobacterium abscessus. Mol Microbiol 101:515–529.

15. Kabsch W (2010) Integration, scaling, space-group assignment and post-refinement. Acta Crystallogr. D Biol. Crystallogr. 66:133–144.

16. Adams PD, Afonine PV, Bunkóczi G, Chen VB, Davis IW, Echols N, Headd JJ, Hung L-W, Kapral GJ, Grosse-Kunstleve RW, et al. (2010) PHENIX: a comprehensive Python-based system for macromolecular structure solution. Acta Crystallogr. D Biol. Crystallogr. 66:213–221.

17. McCoy AJ, Grosse-Kunstleve RW, Adams PD, Winn MD, Storoni LC, Read RJ (2007) Phaser crystallographic software. J. Appl. Crystallogr. 40:658–674.

18. Emsley P, Lohkamp B, Scott WG, Cowtan K (2010) Features and development of Coot. Acta Crystallogr. D Biol. Crystallogr. 66:486–501.

19. Holm L, Rosenström P (2010) Dali server: conservation mapping in 3D. Nucl Acids Res 38:W545–W549.

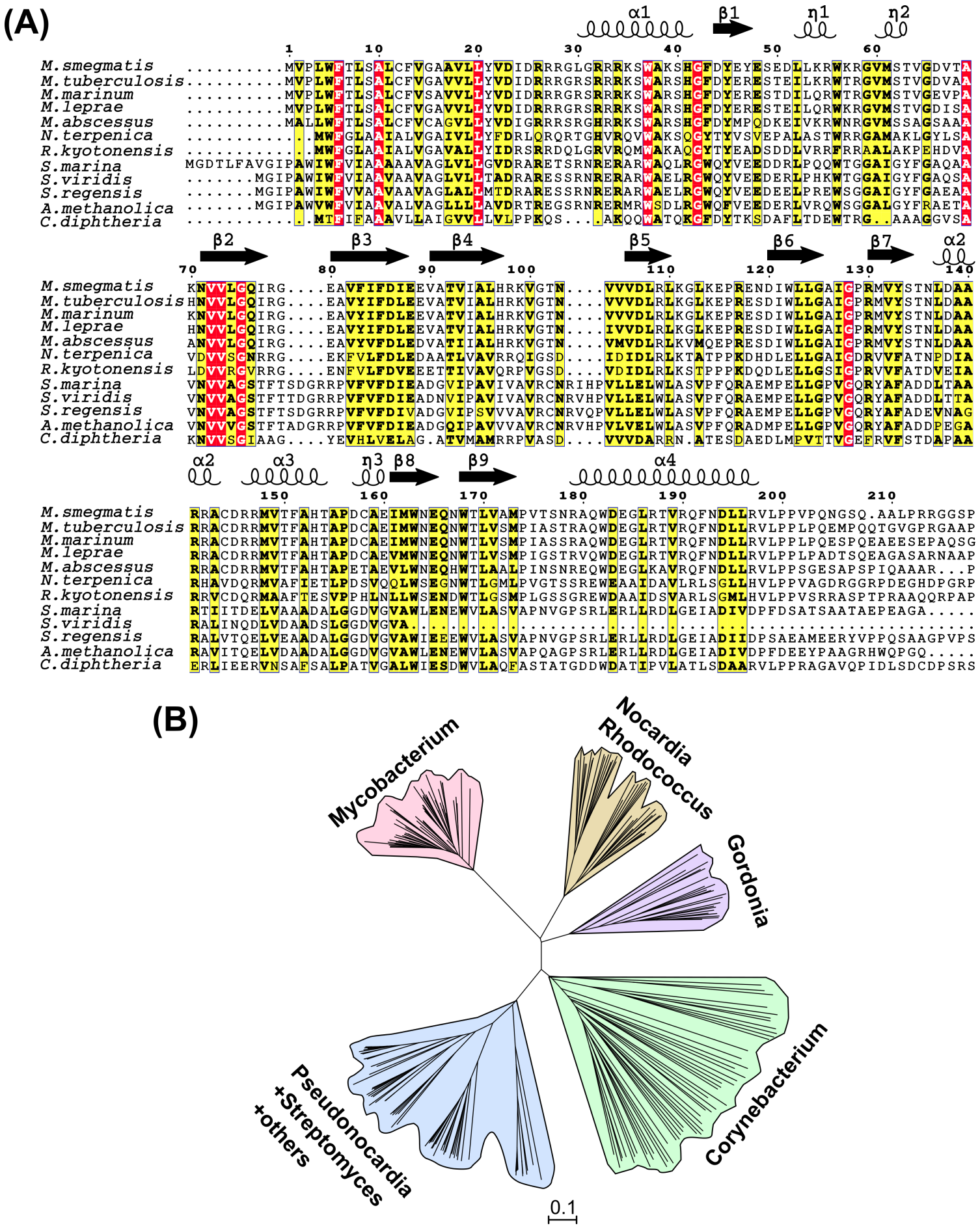
Figures legends

Figure 1: Primary sequence conservation of TtfA. (A) Multiple sequence alignment of TtfA from different species. The figure was generated with the ESPript server (http://espript.ibcp.fr/ESPript/ESPript/). Secondary structure corresponding to the *M. smegmatis* TtfA crystal structure is displayed on top of the alignment. Red colour indicates strict conservation while yellow is for semi-conserved amino acids. The TtfA homologues primary sequences are from the following species: *Mycobacterium smegmatis, Mycobacterium tuberculosis, Mycobacterium marinum, Mycobacterium leprae, Mycobacterium abscessus, Nocardia terpenica, Corynebacterium diphtheriae, Amycolatopsis methanolic, Saccharomonospora marina, Streptomyces regensis, Saccharomonospora viridis and Rhodococcus kyotonensis.* **(B)** Phylogenetic repartition of TtfA homologues in 250 species. Sequences were retrieved from a TtfA homologues search on the HMMER server (https://www.ebi.ac.uk/Tools/hmmer/) using the TtfA sequence from *M. smegmatis* as a search request. ClustalX was used to performed multiple sequence alignment and the iTOL server (https://itol.embl.de/) was used to generate the unrooted tree, which was further manually edited using the Inkscape software.

Figure 2: Biochemical and structural characterization of TtfA. **(A)** Determination of the oligomeric state of TtfA in solution. The elution profile of the proteins used for calibration is displayed as a black line and the elution profile of TtfA is in red. Calibration was established using β-amylase (1) (200 kDa), bovine serum albumin (2) (66 kDa), carbonic anhydrase (3) (29 kDa), and cytochrome C (4) (12.4 kDa) as standards and eluted with estimated volumes of 10.9, 13.1, 15.7 and 17.1 mL, respectively. The void volume was estimated at 8.2 mL. TtfA elutes at 16.9 mL which corresponds to an apparent molecular weight of 14.4 kDa. The purity of TtfA after a three-step purification procedure is shown on the Coomassie Blue-stained SDS polyacrylamide gel electrophoresis, 10 µg of protein were loaded. **(B)** Overall cartoon representation of the TtfA crystal structure. Beta-strands colored in magenta are indicated with the β sign followed by a number, alpha and 3_{10} helices are in slate color and are labeled with the α and η signs respectively. Nt and Ct stand for N-terminal and C-terminal. **(C)** Surface representation of TtfA. The electrostatic potential of the protein was calculated with the PDB2PQR Server, and displayed with the APBS module from Pymol. Red and blue color indicate negative and positive charges respectively.

Table 1: Data collection and Refinement statistics

Data collection statistics	**TtfA_native**	**TtfA_SeMet**
Beamline	X06DA-PXIII	X06DA-PXIII
Wavelength (Å)	0.979	0.979
Resolution range (Å)	46.54-1.4 (1.45-1.4)	46.59-1.45 (1.5-1.45)
Space group	I 2 2 2	I 2 2 2
Unit cell (Å,°)	62.24 70.09 86.01 90 90 90	62.15 70.39 85.69 90 90 90
Total reflections	168405 (15945)	440624 (42702)
Multiplicity	4.5 (4.3)	13.3 (13.3)
Completeness (%)	99.7 (99.4)	98.3 (96.7)
Mean I/sigma(I)	14.7 (1.3)	22.9 (1.9)
Wilson B-factor (Å^2)	17.93	20.1
R-meas	0.057 (1.1)	0.087 (1.4)
CC1/2	0.99 (0.59)	0.99 (0.69)
Data refinement statistics		
Reflections used in refinement	37303 (3670)	
R-work	0.1739 (0.2885)	
R-free	0.1984 (0.3149)	
Number of non-H atoms	1734	
macromolecules	1541	
ligands	10	
solvent	183	
Protein residues	178	
RMS bonds (Å)	0.013	
RMS angles (°)	1.23	
Ramachandran favored (%)	97.73	
Ramachandran allowed (%)	2.27	
Ramachandran outliers(%)	0.00	
Rotamer outliers (%)	0.00	
Average B-factor (Å^2)	28.14	
macromolecules	27.37	
ligands	33.71	
solvent	34.29	
Clashscore	4.15	
PDB accession number	6T84	


(A)
α1
β1
η1
η2
β2
β3
β4
β5
β6
β7
α2
α2
α3
η3
β8
β9
α4
M.smegmatis
M.tuberculosis
M.marinum
M.leprae
M.abscessus
N.terpenica
R.kyotonensis
S.marina
S.viridis
S.regensis
A.methanolica
C.diphtheria
(B)
Mycobacterium
Nocardia
Rhodococcus
Gordonia
Corynebacterium
Pseudonocardia
+Streptomyces
+others
0.1

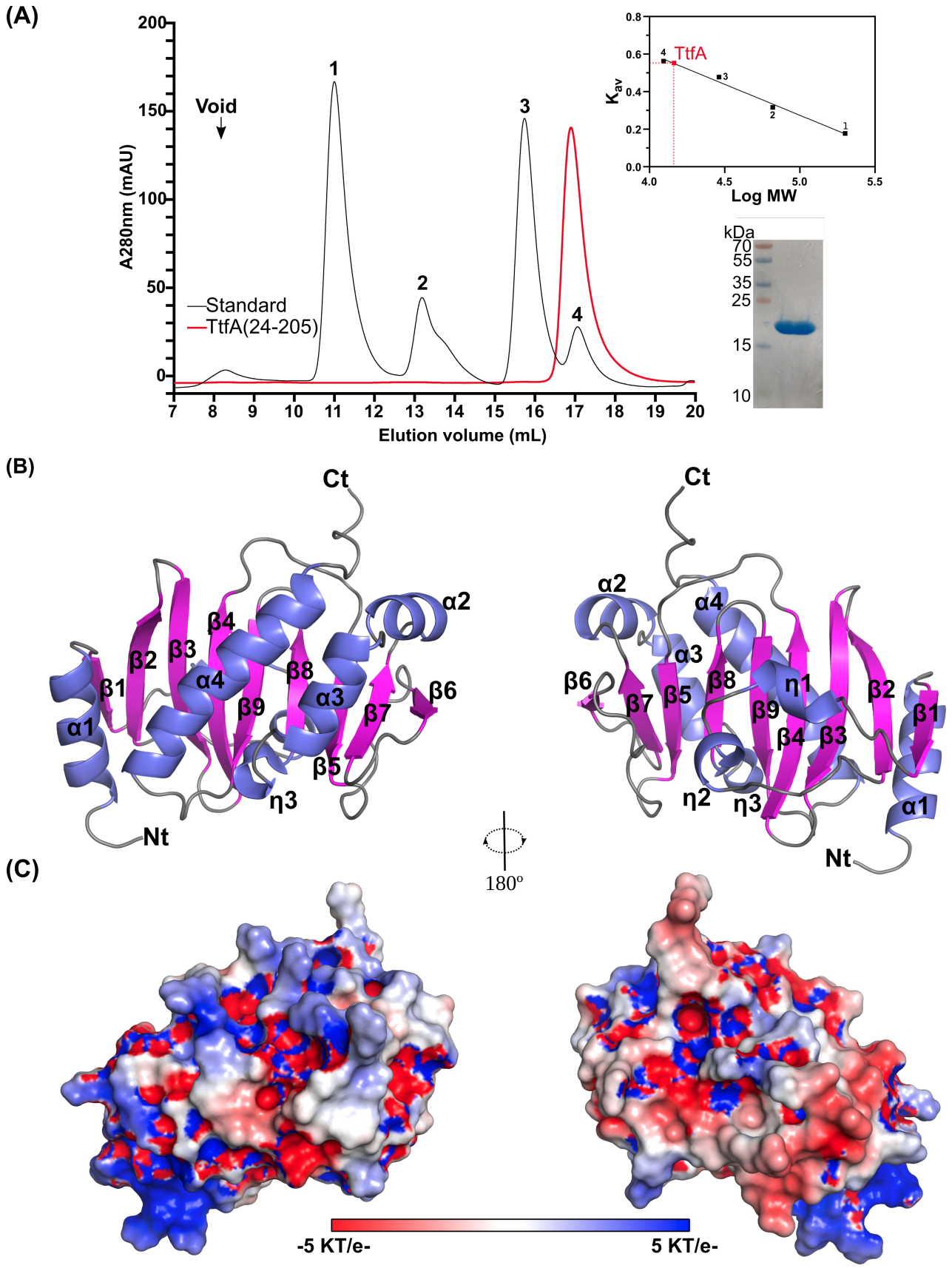

(A)
A280nm (mAU)
Elution volume (mL)
Void
1
2
3
4
Standard
TtfA(24-205)
TtfA
Log MW
kDa
70
55
35
25
15
10
(B)
Ct
Nt
α1
β1
β2
β3
β4
α4
β8
α3
β9
β7
β6
β5
η3
α2
180°
(C)
-5 KT/e-
5 KT/e-

Figure 2