



**HAL**  
open science

# Crystal structure of MAB\_4123 a putative flavin-dependent monooxygenase from *Mycobacterium abscessus*

Kien Lam Ung, Chloé Poussineau, Julie Couston, Husam M a B Alsarraf,  
Mickaël Blaise

► **To cite this version:**

Kien Lam Ung, Chloé Poussineau, Julie Couston, Husam M a B Alsarraf, Mickaël Blaise. Crystal structure of MAB\_4123 a putative flavin-dependent monooxygenase from *Mycobacterium abscessus*. *Acta crystallographica Section F: Structural biology communications* [2014-..], 2023, 79, pp.128-136. 10.1107/S2053230X2300345X . hal-04092114

**HAL Id: hal-04092114**

**<https://hal.science/hal-04092114>**

Submitted on 9 May 2023

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

## Crystal structure of MAB\_4123 a putative flavin-dependent monooxygenase from *Mycobacterium abscessus*

Kien Lam Ung<sup>1,#</sup>, Chloé Poussineau<sup>1</sup>, Julie Couston<sup>1</sup>, Husam M.A.B. Alsarraf<sup>1,2#</sup> and Mickaël Blaise<sup>1</sup>.

<sup>1</sup>IRIM, Université de Montpellier, CNRS, Montpellier, France.

<sup>2</sup>Department of molecular biology and Genetics, Aarhus University, Aarhus, Denmark

#Present address: Department of molecular biology and Genetics, Aarhus University, Aarhus, Denmark

\* Correspondence: [mickael.blaise@irim.cnrs.fr](mailto:mickael.blaise@irim.cnrs.fr)

### Synopsis

The structural characterization of MAB\_4123, a putative two-component flavin-dependent monooxygenase, from the human pathogen *Mycobacterium abscessus* suggests a role in organosulfur compounds catabolic pathway.

### Abstract

Numerous bacteria from different phylae can perform reactions of desulfurization of organosulfur compounds. In these degradation or detoxification pathways, the two-component flavin-dependent monooxygenases using flavin (FMN or FAD) as a cofactor play important roles as they catalyse the first steps of these metabolic routes. The TdsC or DszC and MsuC proteins belong to this class of enzymes as they process dibenzothiophene (DBT) and methanesulfinate. The elucidation of their X-ray structures in their apo, ligand, and cofactor bound forms brought important molecular insights into their catalytic reaction. Mycobacterial species were also shown to possess a DBT degradation pathway, but no structural information on this two-component flavin-dependent monooxygenases was available. In this study, we present the crystal structure of the uncharacterised MAB\_4123 protein from the human pathogen *Mycobacterium abscessus*. The structure solved at high resolution displays high similarity with homologs from *Rhodococcus*, *Paenibacillus* and *Pseudomonas* species. *In silico* docking approaches suggest that MAB\_4123 binds FMN and may use it as a cofactor. Our structural analysis strongly suggests that MAB\_4123 is a two-component flavin-dependent monooxygenase that could act as a detoxifying enzyme of organosulfur compounds in mycobacteria.

**Keywords:** *Mycobacterium abscessus* ; organosulfur compounds ; two-component flavin-dependent monooxygenase

## 1. Introduction

Monooxygenase proteins use flavins as cofactors to catalyze the breakage of the bond between two oxygen atoms. This reaction results in the addition of one oxygen atom into the substrate and the reduction of the second oxygen into water (Ellis, 2010). In this large enzyme family, the two-component flavin-dependent enzymes are only described in bacteria. The two-component flavin-dependent monooxygenases with the help of either flavin adenine dinucleotide (FAD) or flavin mononucleotide (FMN) oxidize compounds found in the bacterial environment. These enzymes capable of desulfurization of sulfonated compounds (Fig. 1) are of particular interest to biotechnology as they can notably degrade sulfur-compounds from fossil fuels, a source of environmental pollution.

DszC from *Rhodococcus* sp. XP (Liu *et al.*, 2014) or TdsC from *Paenibacillus* sp. A11-2 (Hino *et al.*, 2017) catalyze the two first steps of dibenzothiophene (DBT) degradation (Fig. 1) leading to the formation of DBT sulfone that is further processed by the dibenzothiophene sulfone monooxygenase DszA/TdsA and the 2-(2'-hydroxyphenyl)-benzene sulfinic acid (HPBSi) desulfinate DszB/TdsB (Piddington *et al.*, 1995; Ishii *et al.*, 2000; Kirimura *et al.*, 2004; Ohshiro *et al.*, 2005; Konishi *et al.*, 1997). In this pathway, the DszC and TdsC proteins have been well characterized biochemically and structurally, hence enabling the understanding of their catalytic function at a molecular level (Liu *et al.*, 2014; Hino *et al.*, 2017).

Similarly, the two-component flavin-dependent monooxygenase MsuC from *Pseudomonas fluorescens* was shown to catalyze the second step of the dimethylsulfone (DMSO<sub>2</sub>) degradation pathway (Kertesz *et al.*, 1999; Wicht, 2016). MsuC converts methanesulfinic acid (MSI<sup>-</sup>) to methanesulfonate (MS<sup>-</sup>) (Fig. 1) which is finally processed to sulfite by MsuD that can be assimilated by different bacterial synthesis pathways (Soule *et al.*, 2020). Crystallographic investigation of MsuC confirmed the close structural similarity with the DszC/TdsC enzymes (Soule *et al.*, 2020).

As exemplified above, numerous bacteria possess these organosulfur detoxification pathways enabling organosulfur compounds degradation. Pathogenic bacteria such as *Mycobacterium goodii* a rapidly growing nontuberculous mycobacterium and an opportunistic human pathogen can possess as well this pathway. Of interest, the DszC enzyme from *Mycobacterium goodii* X<sub>7</sub>B (Li *et al.*, 2009) processes DBT using either FMNH<sub>2</sub> or FADH<sub>2</sub> as a cofactor.

So far, the mycobacterial two-component flavin-dependent monooxygenases have not been investigated from a structural point of view. To fill this gap, we investigated the putative two-component flavin-dependent monooxygenase encoded by MAB\_4123 from *Mycobacterium*

*abscessus*, an emerging human pathogen, highly resistant to antibiotherapy (Johansen *et al.*, 2020). We report herein the crystal structure of MAB\_4123 at high resolution.

## 2. Materials and Methods

### 2.1 Cloning, expression and purification of MAB\_4123

The wild-type DNA sequence of *MAB\_4123* (Uniprot: B1MIH5) was amplified from the genomic DNA of *M. abscessus* reference strain 19977 using a pair of primers described in [Table I](#). The resulting PCR fragment was cloned into pET32(a) expression plasmid between the KpnI-XhoI restriction sites. The resulting plasmid encoded for a fusion protein with a thioredoxin tag, hexahistidine, thrombin site, S-tag, and Tobacco Etch Virus (TEV) protease cleavage site at the *N*-terminus of MAB\_4123. This plasmid was transformed into *E. coli* BL21 strain (New England Biolabs, C2527) strain containing the pRARE2 plasmid and used for protein production. Transformed colonies were inoculated into LB medium supplemented with 200  $\mu\text{g}\cdot\text{mL}^{-1}$  ampicillin and 30  $\mu\text{g}\cdot\text{mL}^{-1}$  chloramphenicol and grown overnight at 37 °C under agitation. The above preculture was used to start a 6 L of LB cultures in a shaker flask at 37 °C. When OD<sub>600</sub> reached 0.8, the cultures were chilled in ice for 30 min and 1 mM of Isopropyl- $\beta$ -D-thiogalactoside (Euromedex) was added to induce protein production for 16 h at 18 °C. Cell pellets were harvested and suspended in buffer A (50 mM Tris-HCl pH 8, 0.4 M NaCl, 0.5 mM  $\beta$ -mercaptoethanol, and 1 mM benzamidine). The cell suspension was lysed by sonication, and the lysate was clarified by centrifugation at 28,000 *g* at 4 °C for 1 h. Before loading onto a gravity column 10 mM of imidazole and the Ni-nitrilotriacetic acid Sepharose beads were added to the supernatant followed by 15 min incubation with gentle agitation at 4 °C. Beads were washed with 15 column volumes of buffer B (50 mM Tris-HCl pH 8, 0.4 M NaCl, 15 mM imidazole, and 0.5 mM  $\beta$ -mercaptoethanol) and C (50 mM Tris-HCl pH 8, 1 M NaCl, and 0.5 mM  $\beta$ -mercaptoethanol). The His-tagged MAB\_4123 was eluted with a high concentration of imidazole in buffer D (50 mM Tris-HCl pH 8, 0.2 M NaCl and 0.5 mM  $\beta$ -mercaptoethanol, 250 mM imidazole). TEV protease was added into the eluate following a ratio of 1 mg of TEV protease per 40 mg of proteins (1:22 molar ratio), and the mixture was dialyzed overnight in buffer E (50 mM Tris-HCl pH 8, 0.2 M NaCl, and 0.5 mM  $\beta$ -mercaptoethanol) at 4 °C. After dialysis and TEV cleavage, the solution went through a second Ni-nitrilotriacetic acid column for retaining uncleaved protein, TEV protease, and thioredoxin tag. The flow-through fractions containing tag-free MAB\_4123 were collected and concentrated using a 10 kDa cut-off centricon (Sartorius) and further purified by size-exclusion chromatography on a Superdex 200 Increase 10/300 GL column pre-equilibrated in buffer F (50 mM Tris-HCl pH 8 and 0.2 M NaCl) for crystallization. Protein purity was estimated on a Coomassie-stained SDS-PAGE gel and by comparing the protein of interest versus the presence contaminating bands.

## Table I Macromolecule production information

---

Source organism *Mycobacterium abscessus* ATCC 19977

DNA source: genomic DNA from *Mycobacterium abscessus* ATCC 19977

Forward primer 5'-CTCAC**GGTACC***GAGAACCTGTACTTCCAGGGTCATatgaccaccgccaccacctc*-3' (letters in bold indicate the KpnI restriction site and in italic the sequence encoding the Tobacco Etch Virus protease cleavage site)

Reverse primer 5'-GGATG**CCTCGAG***tcacggcttcttgcctgtg*-3' (letters in bold indicate the XhoI restriction site)

Expression vector pET-32(a)

Expression host *E. coli* BL21 (DE3) transformed with the pRARE2 plasmid

The complete amino acid sequence of the construct after TEV cleavage (in bold extra amino acids from cloning strategy and TEV cleavage)

**GH**MTTATTSPVHAYQGVSDTEFSEWEQVAARVAGELSATALTRDRANQNPIAEIELLRRYGLLSFAT  
AREFGGAGGSLVQALQLGRIIAAADGSIGQLLVYHYSNGVWVWYILGSPTQREYISRGVGGHGWFO  
GSVSNPRDPGITVTRTEEGYRVNGKRTFATGVAVADLITVLLYEAPINAIIPSERDGLRFNDDDWDL  
GQRLTASGSVEFDNVLLRHDEVLTGLDEYSGLDGSRRERDGLRALFSQLIFVHLYLGLIAEGALAAG  
VAYIRDKGRPWEAHSTDVTEDPYHQQLLGRLSAGIAAGVALADSATKEFEQALAFGEAPTEAQW  
GALAIRVDQAKSVATEISLDVTHNIYQATGARSTANSVGLDIYWRNARTHTTHDPLPYRQREIGRH  
LLTDQWSPRSLNGLGRLAETQGGKP

---

## 2.2 Crystallization

Prior to crystallization, 2 mM dibenzothiophene and 4% dimethylsulfoxide (DMSO) were added to the protein solution concentrated at 9.5 mg.mL<sup>-1</sup> in 50 mM Tris pH 8.0 and 0.2 M NaCl. Crystallization drops were set-up in 96 well MRC plates mixing 0.8 µl of protein solution with 0.8 µl of the reservoir solution. Crystals were obtained at 291 K in 1.5 M Ammonium phosphate dibasic and 0.1 M Tris pH 8.5 and cryo-protected in 1.5 M Ammonium phosphate dibasic, 0.1 M Tris pH 8.5 and 20% glycerol (v/v) before being cryo-cooled in liquid nitrogen.

## Table 2 Crystallization

---

Method Vapor drop diffusion method in sitting drops

Plate type Swissci 96-Well 2-Drop MRC Crystallization Plates

Temperature (K) 291

Protein concentration 9.5 mg.mL<sup>-1</sup>

Buffer composition of protein solution 50 mM Tris-HCl pH 8.0, 200 mM NaCl

Composition of reservoir solution 0.1 M Tris-HCl pH 8.5 and 1.5 M Ammonium phosphate dibasic

Volume and ratio of drop 0.8 µL+ 0.8 µL

Volume of reservoir 50 µL

---

### 2.3 Data collection and structure solving

The X-ray dataset was collected on the ID30A1-MASSIF1 beamline (Svensson *et al.*, 2015) at the European Synchrotron Radiation Facility, Grenoble, France. The dataset was recorded on a PILATUS3-2M detector (Dectris) at a wavelength of 0.965 Å (12.842 keV) and a crystal-to-detector distance of 230.6 mm. A total of 3160 frames were collected with an exposure time of 0.11 s, a rotation range of 0.05° and full beam transmission. Data were processed, scaled, and merged with *XDS* (Kabsch, 2010). The structure was solved by molecular replacement using the Molrep (Vagin & Teplyakov, 2010) program with the structure of the dibenzothiophene monooxygenase (DszC) from *Rhodococcus erythropolis* (PDB id 4JEK) as a search model and sharing 37% protein sequence identity with MAB\_4123. The tetramer obtained from the first molecular replacement was used as a search model to perform an additional molecular replacement step using *Phaser* (McCoy *et al.*, 2007). The model was then automatically adjusted with *phenix.autobuild* from the *Phenix software suite* (Liebschner *et al.*, 2019) followed by several cycles of manual adjustment and refinement with *Coot* (Emsley & Cowtan, 2004) and *phenix.refine* (Liebschner *et al.*, 2019) respectively. The statistics for the data collection and structure refinement are reported in [Tables 3 and 4](#).

**Table 3 Data collection and processing**

Values in parentheses are for the outer shell.

Diffraction source	ESRF, ID30A-1
Wavelength (Å)	0.965
Temperature (K)	100
Detector	PILATUS3-2M
Crystal-detector distance (mm)	230.6
Rotation range per image (°)	0.05
Total rotation range (°)	158
Space group	$P2_1$
$a, b, c$ (Å)	131.91, 134.18, 206.80
$\alpha, \beta, \gamma$ (°)	90, 90.68, 90
Mosaicity (°)	0.082
Resolution range (Å)	100-2.1 (2.2-2.1)
Total No. of reflections	1238477 (114732)
No. of unique reflections	413629 (41003)
Completeness (%)	98.75 (98.3)
Multiplicity	3 (2.8)

$\langle I/\sigma(I) \rangle$	10.74 (1.20)
Wilson <i>B</i> -factor (Å <sup>2</sup> )	38.8
<i>R</i> <sub>meas</sub> (%)	9.3 (126.4)
CC1/2	0.998 (0.518)

---

#### Table 4 Structure refinement

Values in parentheses are for the outer shell.

Resolution range (Å)	47.94 - 2.1 (2.175 - 2.1)
Reflections used in refinement	413476 (40988)
Reflections used for <i>R</i> <sub>free</sub>	2016 (199)
<i>R</i> <sub>work</sub>	0.178 (0.314)
<i>R</i> <sub>free</sub>	0.212 (0.336)
Number of non-hydrogen atoms	51940
macromolecules	48833
solvent	2986
ligand	121
Protein residues	6312
R.m.s.d., bond lengths (Å)	0.002
R.m.s.d., angles (°)	0.43
Ramachandran favored (%)	98.60
Ramachandran allowed (%)	1.40
Ramachandran outliers (%)	0
Rotamer outliers (%)	0.69
Clashscore	2.76
Average <i>B</i> -factor (Å <sup>2</sup> )	50.40
Macromolecule	50.54
Solvent	47.76
PDB accession	8CDA

---

## 2.4 Ligand *in silico* docking

FMN was first subjected to molecular dynamics energy minimization by Merck molecular force field (MMFF94) using the *PyRx software* suite (Dallakyan & Olson, 2015). Hydrogen atoms were added both to protein and ligand. The docking was performed with *AutoDock Vina* (Trott & Olson, 2010) run in *PyRx* using the structure of the MAB\_4123 dimer as a receptor and with residue Arg361 defined as flexible. The search was performed with a docking box centered as followed :  $x=-65.8$ ,  $y=-92.4$ , and  $z=-73.8$  with dimension (Å)  $X=25.9$ ,  $Y=30.1$  and  $Z=39.7$ , and exhaustiveness defined to 8.

## 3. Results and Discussion

### 3.1 Purification, crystallization and structure determination of MAB\_4123

The full-length MAB\_4123 gene was cloned into pET-32a and expressed as a fusion protein with the thioredoxin and multi-histidine tags in *E. coli*. The protein was purified through three steps of purification involving two steps on IMAC column and one purification step on size-exclusion chromatography (SEC) (Fig. 2). We could crystallize the protein in presence of a potential ligand, dibenzothiophene (DBT) diluted in dimethyl sulfoxide (DMSO). These crystals belonging to the monoclinic  $P2_1$  space group diffracted to a resolution of 2.1 Å. We solved the phase problem by molecular replacement using the structure of the DszC protein from *Rhodococcus erythropolis* (PDB id 4JEK). The asymmetric unit contains 16 monomers (Fig. 3a). We could model most of the residues apart from the first 11 residues in the N-terminus (Table I) and the last seventeen residues in the C-terminus as well as residues 136-139 and 174-175 in chains H and N respectively. In addition, we placed several phosphate ions found at the interface of two monomers. Packing analysis suggested the existence of four stable tetramers within the asymmetric unit (Figs. 3a and 3b). The homotetramer assembly was further confirmed in solution by SEC (Fig. 2a). Indeed, the MAB\_4123 elution peak on SEC corresponds to an apparent molecular weight of 162 kDa while the theoretical mass of the protein is 46 kDa. Such assembly was also attested in MAB\_4123 close homologs notably TsdC (Hino *et al.*, 2017) and MsuC (Soule *et al.*, 2020).

### 3.2 MAB\_4123 structural analysis

The protein adopts a typical acyl-CoA dehydrogenase fold (Fig. 3). The N-terminal  $\alpha$ -domain (residues 10-129) is composed of five  $\alpha$ -helices followed by the central domain (residues 130-232) made of two antiparallel  $\beta$ -sheets forming a  $\beta$ -sandwich and the C-terminal  $\alpha$ -domain (residues 233-405) comprising five  $\alpha$ -helices.



Using the ENDscript server (Robert & Gouet, 2014), we retrieved close homologous structures to MAB\_4123. As stated in the introduction, DszC and TdsC were shown to be dibenzothiophene monooxygenases while MusC was recently demonstrated to be involved in the conversion of methanesulfinate (MSI<sup>-</sup>) to methanesulfonate (MS<sup>-</sup>) (Fig. 1). These three enzymes present respectively 35%, 34%, and 31% protein sequence identity with MAB\_4123 (Fig. 4a). The MAB\_4123 monomer displays an r.m.s.d. of 1.43 Å over 355 C $\alpha$  with DszC from *Rhodococcus erythropolis* (PDB id 4JEK), an r.m.s.d. of 1.37 Å over 350 C $\alpha$  with TdsC from *Paenibacillus sp.A11-2* (PDB id: 5XB8) and finally an r.m.s.d. of 1.45 Å over 358 C $\alpha$  with MsuC from *Pseudomonas fluorescens* (PDB id: 6UUG) (Fig. 4b). Overall, the structures are very similar with only a few noticeable differences. Firstly, the region between  $\alpha$ 6 and  $\beta$ 7 ranging from residue 217 to 233, folded as a loop and two  $3_{10}$  helices in MAB\_4123, is very similar to the one of MsuC but differs from the ones of DsbC/TdsC which is more folded with the presence of one  $\alpha$ -helix and one  $3_{10}$  helix. The second noticeable difference lies in the length of loop 1 spanning from residues 18CDA31 to 137 which is of similar length in MsuC but slightly longer in DsbC/TdsC as twelve residues composed loop 1 in these proteins. It is also worth mentioning that the electron density map for loop 1 between strands  $\beta$ 1 and  $\beta$ 2 and loop 2 (271-286) between helices  $\alpha$ 6 and  $\alpha$ 7 is well defined in most MAB\_4123 monomers (except for loop 1 in chain H) and in a closed conformation, these characteristics are more related to ligands/cofactor bound structures as proposed for the TdsC protein (Hino *et al.*, 2017). The role of loop 2 is not well defined in these enzymes but loop 1 is near the active site residues. A third difference is the length of the loop separating strands  $\beta$ 4 and  $\beta$ 5 made of five residues in MAB\_4123 while in TdsC/DsbC this loop has 4 additional residues. Structures superimposition attests that MAB\_4123 is a very close homolog of the DsbC/TdsC/MsuC enzymes, thus strongly suggesting that MAB\_4123 belongs also to the flavin-dependent monooxygenase class of enzymes.

### 3.4 Cofactor and putative ligand binding sites.

To further investigate the potential function of MAB\_4123, we assessed if the protein as its close homologs, could bind flavin cofactor and notably FMN. We tried to obtain the structure of MAB\_4123 bound to FMN either by co-crystallization or soaking experiment, but without any success. Alternatively, we performed *in silico* docking of FMN into MAB\_4123 with a search grid that was centered on one monomer and extended partly on the second monomer (Fig. 5a). The search volume encompassed the putative binding site that was defined by structural comparison with the TdsC structure bound to FMN (Hino *et al.*, 2017). This approach led to identifying a potential binding site of FMN in MAB\_4123 at the interface of two monomers as seen in its close homologs. In the best docking pose ( $\Delta G^{\circ} = -10.3$  kcal.mol<sup>-1</sup>) sixteen residues form a binding cavity

for FMN (**Fig. 5b**). Most of these amino acids are strictly conserved and semi-conserved in all MAB\_4123 homologs (**Figs. 3b and 5d**). Residues from one monomer mainly interact with the flavin moiety of FMN while the recognition of the phosphate group is mediated by two Arginines from two different monomers namely Arg137 and Arg361. In agreement with this docking pose, the phosphate group of FMN contacted by Arg137 and Arg361 from each monomer is positioned very similarly to a phosphate ion we could unambiguously identify and model (**Fig. 5c**). Indeed, the extra electron density in the vicinity of these two Arg is well-defined.

Additionally, we noticed two positive electron density peaks in a cavity formed by residues Ser102, Ser245, Trp106, Tyr99, His379, Ser134, and Ser132 (**Fig. 5d**). Interestingly, this unmodelled electron density is situated above the flavin ring of FMN and matches the position of DBT as seen in the cocrystal structure of TdsC bound to DBT and FMN (PDB id 5XDE) (Hino *et al.*, 2017). There is good conservation of the position of the substrates and residues surrounding these extra electron density peaks between the MAB\_4123 and TdsC structures so these two additional electron density peaks could be potential substrates of MAB\_4123. However, we rule out that this could be DBT (which was present in the crystallization drop) as the electron density is discontinued and corresponds more to two small molecules. On top of DBT presence in the crystallization drop, 4% DMSO was also present in the crystallization condition. As it was shown that MsuC, a close homolog of MAB\_4123, converts  $\text{MSI}^-$  to  $\text{MS}^-$ .  $\text{MSI}^-$  is chemically related to DMSO and it is tempting to propose that this density could correspond to DMSO, at least for the extra density near Ser245 and Ser102 (**Fig. 5d**). However, as we could not obtain an exploitable dataset of crystal grown without DMSO, as we cannot guarantee that these positive electron density peaks are DMSO, we, therefore, left them unmodelled. Nonetheless, these observations suggest that MAB\_4123 like MsuC could detoxify methanesulfinate or closely related chemical analogs.

#### 4. Concluding remarks

To conclude, this study demonstrates the structural proximity of MAB\_4123 to the TdsC/DszC and MsuC enzymes that are all members of the two-component flavin-dependent monooxygenase family and involved in the detoxification of organosulfur compounds. The putative role of MAB\_4123 in a such pathway is also very well supported by the genomic organization of the *M. abscessus* genome. Indeed, MAB\_4123 is part of a predicted operon where a potential nitrilotriacetate monooxygenase is encoded by *MAB\_4125*. This protein harbors a bacterial luciferase predicted domain seen in the dibenzothiophene desulfurization enzymes A that is working in pair with DsbC in the dibenzothiophene degradation pathway. Further, several papers have reported dibenzothiophene degradation pathway in mycobacteria (Nomura *et al.*, 2005) and notably the study on *Mycobacterium goodii* X7B describing the DsbC enzyme, which shares 37% sequence

identity with MAB\_4123 (Li *et al.*, 2009). Further, blast search with the MAB\_4123 sequence against the *M. abscessus* genome retrieved only the MAB\_1422c protein as a paralog which displays only 29% identity with MAB\_4123 and moderate sequence identity (<25%) with the DsbC/TdsC and MsuC enzymes. Altogether, it is highly probable, like other nontuberculous mycobacteria that *M. abscessus* possesses an organosulfur detoxification pathway where MAB\_4123 is part of it. However, confirmation of MAB\_4123 function awaits biochemical assay.

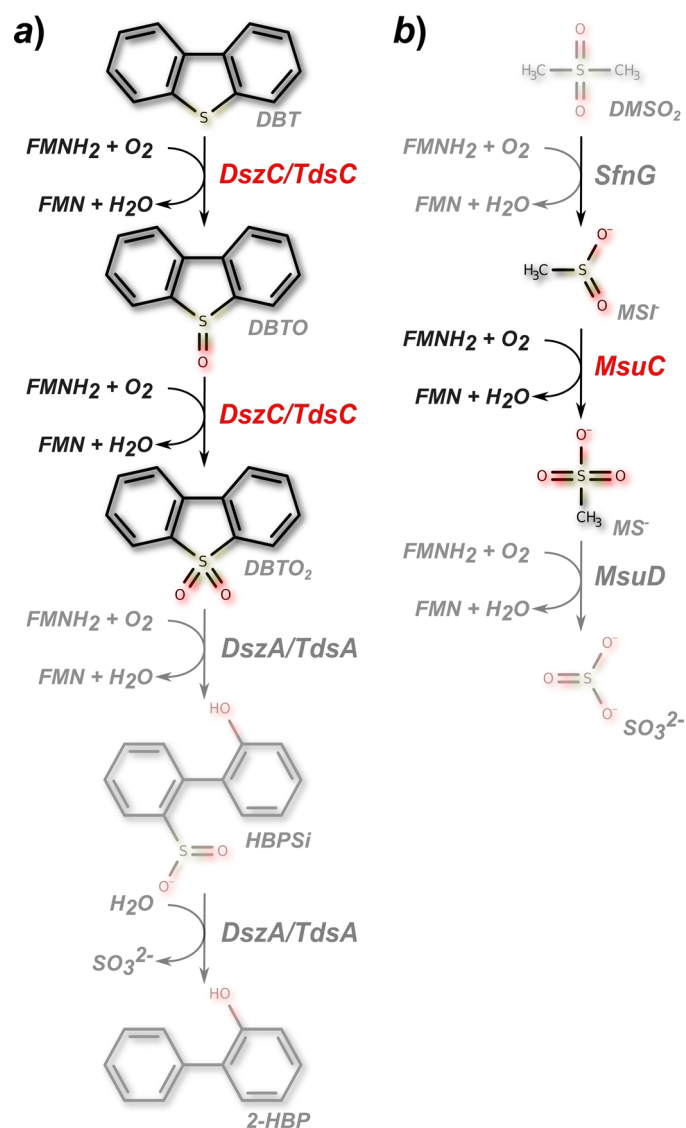
## Acknowledgments

We thank the staff at the MASSIF-1 beamline at ESRF for support during data collection as well as the staff at the PXI and PXIII beamlines at the Swiss Light Source. We thank Dr. L. Kremer and Dr. K. Brodolin for their support at the early stage and throughout the project, respectively. KLU PhD fellowship was supported by the National Research Agency (ANR-17-CE11-0008-01 – MyTraM to MB) and the Lundbeck Foundation, Denmark (R303-2018-2964 to HA). JC's PhD fellowship is financed by CNRS

## BIBLIOGRAPHY

- Dallakyan, S. & Olson, A. J. (2015). *Methods Mol. Biol. Clifton NJ* **1263**, 243–250.
- Ellis, H. R. (2010). *Arch. Biochem. Biophys.* **497**, 1–12.
- Emsley, P. & Cowtan, K. (2004). *Acta Crystallogr. D Biol. Crystallogr.* **60**, 2126–2132.
- Hino, T., Hamamoto, H., Suzuki, H., Yagi, H., Ohshiro, T. & Nagano, S. (2017). *J. Biol. Chem.* **292**, 15804–15813.
- Ishii, Y., Konishi, J., Okada, H., Hirasawa, K., Onaka, T. & Suzuki, M. (2000). *Biochem. Biophys. Res. Commun.* **270**, 81–88.
- Johansen, M. D., Herrmann, J.-L. & Kremer, L. (2020). *Nat. Rev. Microbiol.* **18**, 392–407.
- Kabsch, W. (2010). *Acta Crystallogr. D Biol. Crystallogr.* **66**, 125–132.
- Kertesz, M. A., Schmidt-Larbig, K. & Wüest, T. (1999). *J. Bacteriol.* **181**, 1464–1473.
- Kirimura, K., Harada, K., Iwasawa, H., Tanaka, T., Iwasaki, Y., Furuya, T., Ishii, Y. & Kino, K. (2004). *Appl. Microbiol. Biotechnol.* **65**, 703–713.
- Konishi, J., Ishii, Y., Onaka, T., Okumura, K. & Suzuki, M. (1997). *Appl. Environ. Microbiol.* **63**, 3164–3169.
- Li, J., Feng, J., Li, Q., Ma, C., Yu, B., Gao, C., Wu, G. & Xu, P. (2009). *Bioresour. Technol.* **100**, 2594–2599.

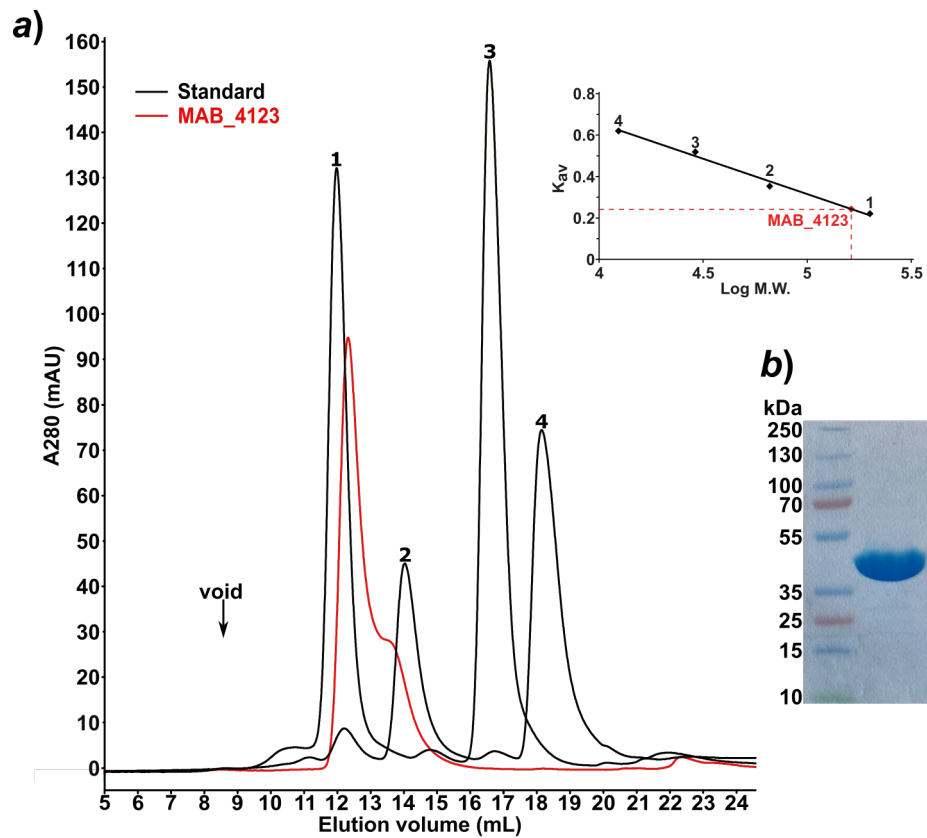
- Liebschner, D., Afonine, P. V., Baker, M. L., Bunkóczi, G., Chen, V. B., Croll, T. I., Hintze, B., Hung, L.-W., Jain, S., McCoy, A. J., Moriarty, N. W., Oeffner, R. D., Poon, B. K., Prisant, M. G., Read, R. J., Richardson, J. S., Richardson, D. C., Sammito, M. D., Sobolev, O. V., Stockwell, D. H., Terwilliger, T. C., Urzhumtsev, A. G., Videau, L. L., Williams, C. J. & Adams, P. D. (2019). *Acta Crystallogr. Sect. Struct. Biol.* **75**, 861–877.
- Liu, S., Zhang, C., Su, T., Wei, T., Zhu, D., Wang, K., Huang, Y., Dong, Y., Yin, K., Xu, S., Xu, P. & Gu, L. (2014). *Proteins Struct. Funct. Bioinforma.* **82**, 1708–1720.
- McCoy, A. J., Grosse-Kunstleve, R. W., Adams, P. D., Winn, M. D., Storoni, L. C. & Read, R. J. (2007). *J. Appl. Crystallogr.* **40**, 658–674.
- Nomura, N., Takada, M., Okada, H., Shinohara, Y., Nakajima-Kambe, T., Nakahara, T. & Uchiyama, H. (2005). *J. Biosci. Bioeng.* **100**, 398–402.
- Ohshiro, T., Ishii, Y., Matsubara, T., Ueda, K., Izumi, Y., Kino, K. & Kirimura, K. (2005). *J. Biosci. Bioeng.* **100**, 266–273.
- Piddington, C. S., Kovacevich, B. R. & Rambosek, J. (1995). *Appl. Environ. Microbiol.* **61**, 468–475.
- Robert, X. & Gouet, P. (2014). *Nucleic Acids Res.* **42**, W320–324.
- Soule, J., Gnann, A. D., Gonzalez, R., Parker, M. J., McKenna, K. C., Nguyen, S. V., Phan, N. T., Wicht, D. K. & Dowling, D. P. (2020). *Biochem. Biophys. Res. Commun.* **522**, 107–112.
- Svensson, O., Malbet-Monaco, S., Popov, A., Nurizzo, D. & Bowler, M. W. (2015). *Acta Crystallogr. D Biol. Crystallogr.* **71**, 1757–1767.
- Trott, O. & Olson, A. J. (2010). *J. Comput. Chem.* **31**, 455–461.
- Vagin, A. & Teplyakov, A. (2010). *Acta Crystallogr. D Biol. Crystallogr.* **66**, 22–25.
- Wicht, D. K. (2016). *Arch. Biochem. Biophys.* **604**, 159–166.



**Figure 1**

**Figure 1 Organosulfur degradation pathway involving DsbC/TdsC and MsuC enzymes.**

a) Dibenzothiophene (DBT) degradation pathway in *Rhodococcus erythropolis*, and *Paenibacillus sp.A11-2*. DBTO stands for dibenzothiophene sulfoxide; DBTO<sub>2</sub>, dibenzothiophene sulfone; HBPSi, 2'-hydroxybiphenyl-2-sulfinate; 2-HBP, 2-hydroxybiphenyl. b) Dimethylsulfone (DMSO<sub>2</sub>) detoxification pathway as shown in *Pseudomonas fluorescens*. MSI<sup>-</sup> stands for methanesulfinate, and MS<sup>-</sup>, methanesulfonate.

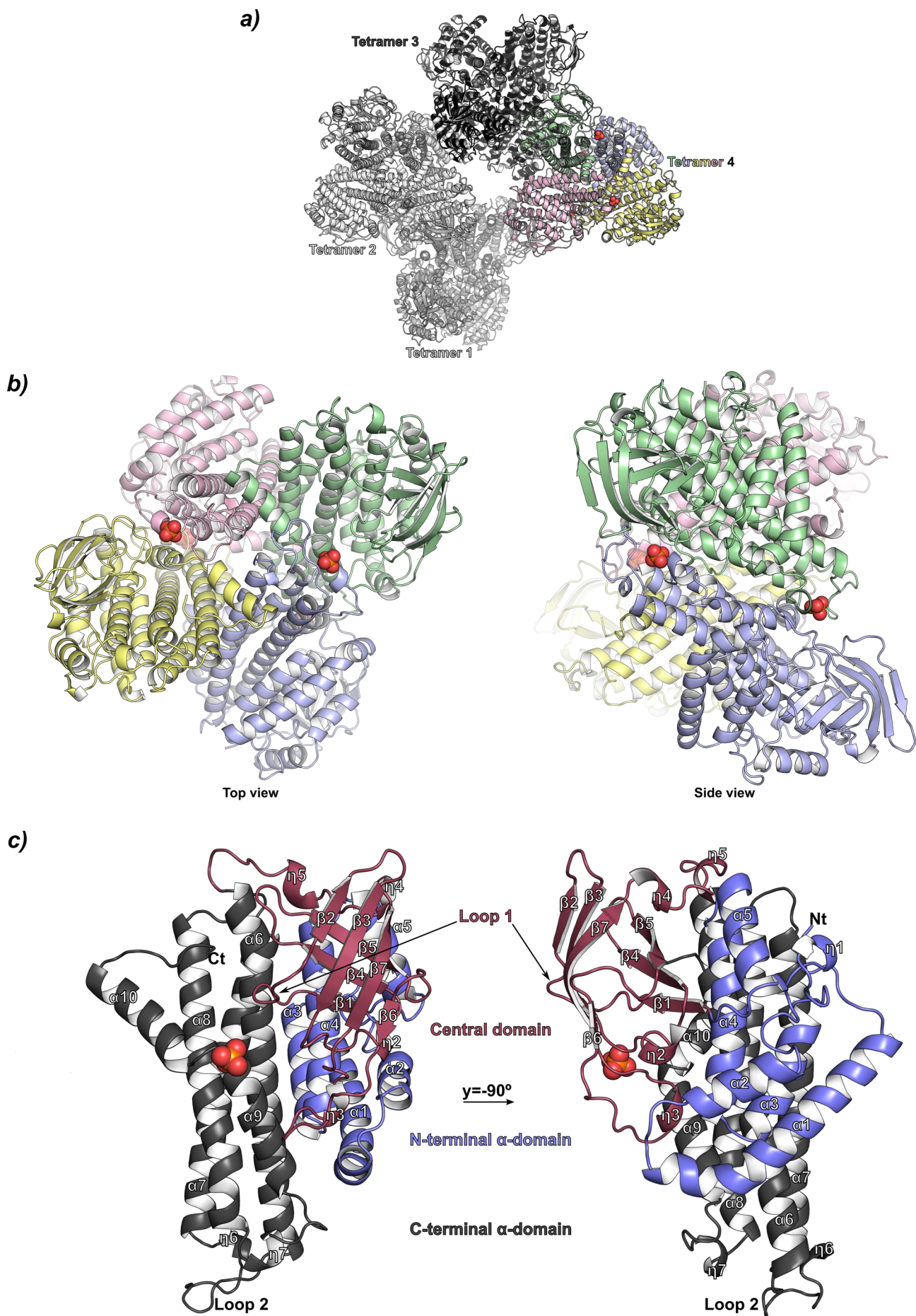


**Figure 2**

**Figure 2 Purification and oligomeric state of MAB\_4123**

a) Size-exclusion chromatogram from the Superdex 200 Increase 10/300 GL column. The void volume (8.58 mL) was determined with dextran blue (black arrow) and the calibration curve (black) was determined using  $\beta$ -amylase ( $V_e=11.98$  mL, 200 kDa), bovine serum albumin ( $V_e=16.58$  mL, 66 kDa), carbonic anhydrase ( $V_e=14.03$  mL, 29 kDa), and cytochrome C ( $V_e=14.03$  mL, 18.14 kDa). Proteins standard (250 to 500  $\mu$ g injected two by two) and MAB\_4123 were eluted in 20 mM Tris pH 8, 0.15 M NaCl and 10% glycerol at a flow-rate of 0.35 mL.min<sup>-1</sup>.  $K_{av}$  was calculated as follows:  $V_e - V_o / V_t - V_o$ , where  $V_e$  is the elution of the protein,  $V_o$  is the void volume and  $V_t$  is the total volume of the column. The calibration curve was obtained by plotting the  $K_{av}$  against the logarithm (Log) of the molecular weight (M.W.) of each protein standard. MAB\_4123 (red curve) eluted at 12.33 mL enabled the estimation of an apparent molecular weight in solution of 162 kDa.

b) Coomassie-stained SDS-PAGE attesting of the high protein purity after the three steps purification procedure. 10  $\mu$ g of protein were loaded on the gel.



**Figure 3**

### **Figure 3 Overall crystal structure of MAB\_4123**

a) Depiction of the asymmetric unit composed of 16 monomers of MAB\_4123 arranged into 4 tetramers. b) Overall structure of the MAB\_4123 homotetramer shown as a cartoon representation and in two different views. The interface of interaction between two monomers is mainly formed by the C-terminal  $\alpha$ -domain. The spheres indicate phosphate ions. c) Cartoon representation of the MAB\_4123 monomer in two different views. The N-terminal  $\alpha$ -domain is shown in blue, the central domain containing the antiparallel  $\beta$ -sheet in raspberry and the C-terminal  $\alpha$ -domain in black. Loops 1 and 2 are indicated and are well-ordered.



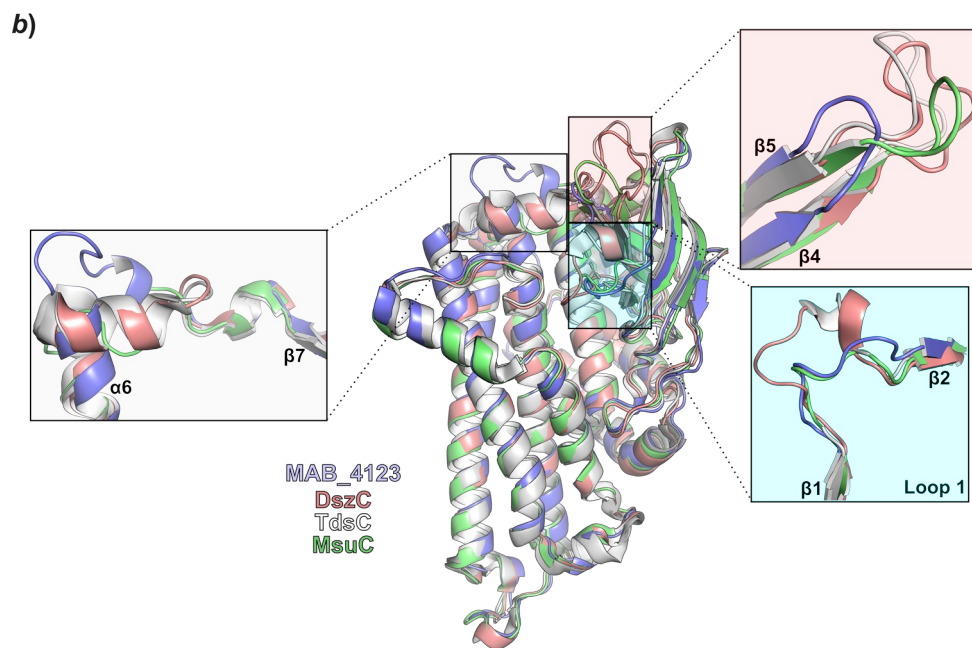
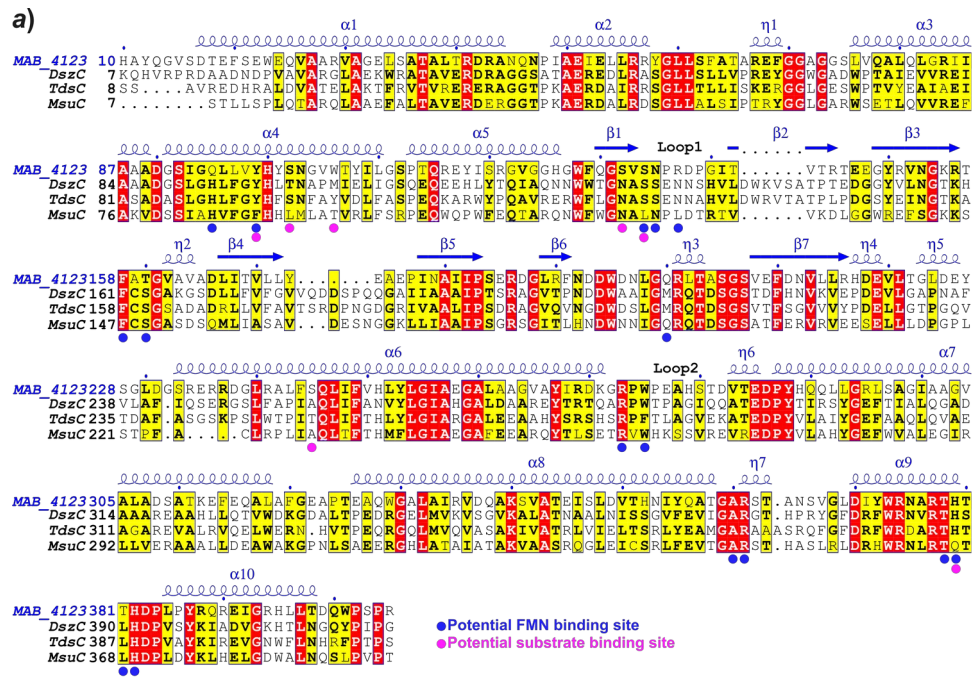


Figure 4

### Figure 4 Structural comparison of MAB\_4123 with close homologs

a) Multiple sequence alignment based on structural homologs of MAB\_4123 and performed with the ENDscript server (<https://endscript.ibcp.fr>) (Robert & Gouet, 2014). Only the three closest structures to MAB\_4123 namely DszC from *Rhodococcus erythropolis* (PDB id 4JEK), TdsC from *Paenibacillus sp.A11-2* (PDB id: 5XB8) and MsuC from *Pseudomonas fluorescens* (PDB id: 6UUG) are displayed. Strictly conserved and semi-conserved residues are coloured in red and yellow respectively. Secondary structures of MAB\_4123 appear on top of the alignment.  $\alpha$ , designs

$\alpha$ -helix;  $\beta$  for  $\beta$ -strand; and  $\eta$  for  $3_{10}$  helix. Residues proposed to be involved in FMN and/or in substrate binding are shown with the blue and magenta circles respectively. b) Structural alignment of MAB\_4123 (blue) with the DszC (salmon), TdsC (white) and MsuC (green) structures. The structures used for this comparison are the same as in a). The main structural differences are highlighted by the colored squared surface.

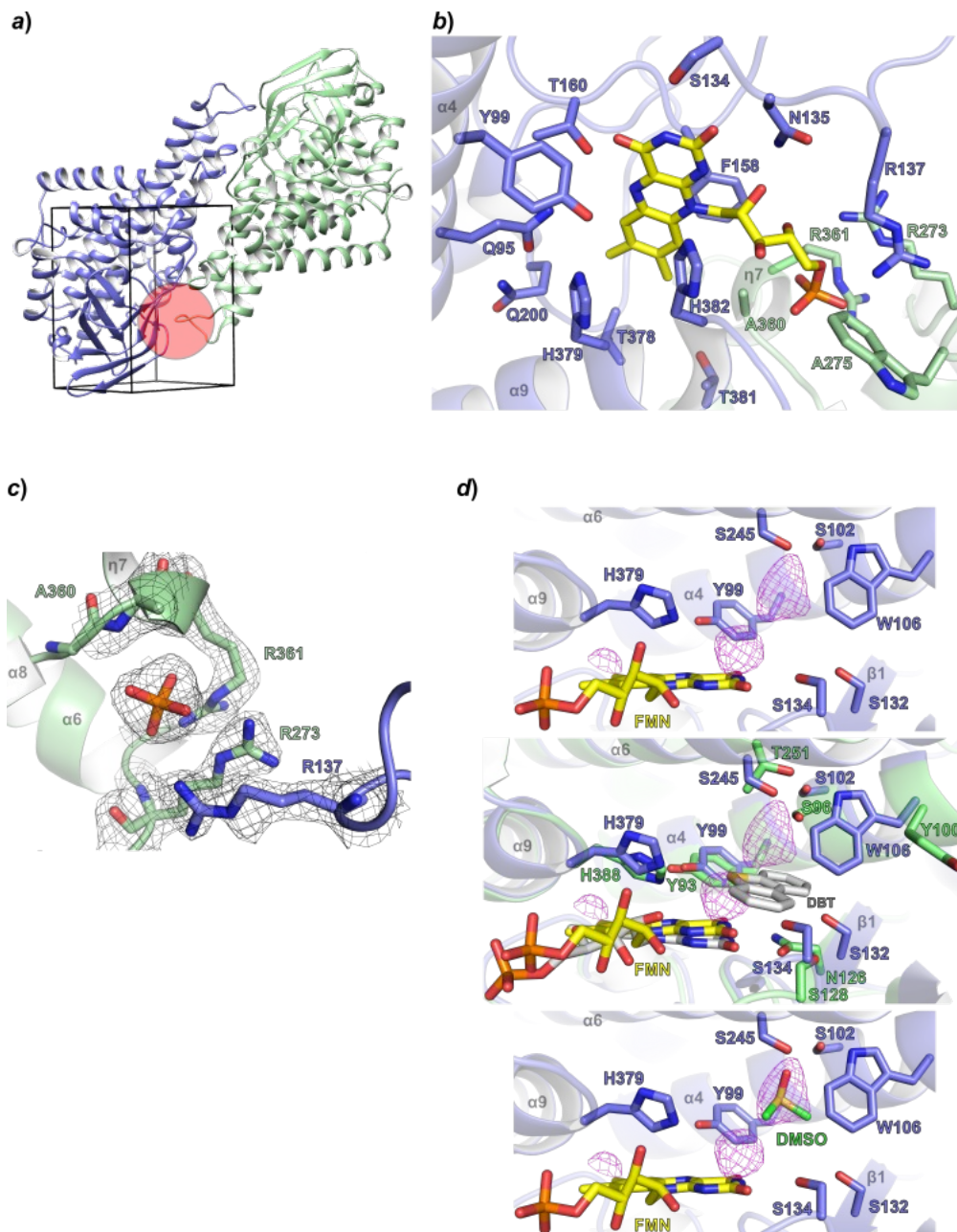


Figure 5

### Figure 5 Identification of MAB\_4123 cofactor and substrates binding sites

a) Representation of the model used to performed *in silico* docking experiment. The search volume involved the MAB\_4123 dimer and used to dock FMN is represented by the grid box in dark. The putative active site location is indicated with the red circle.

b) *In silico* docking of FMN into MAB\_4123. The best docking pose of FMN is displayed in yellow sticks. Residues of MAB\_4123 in close vicinity of FMN (4 Å apart) are displayed as sticks. FMN is potentially recognized at the interface of two MAB\_4123 monomers, colored blue and green. c) Presence of one phosphate ion at each interface of MAB\_4123 monomer. The ion is positioned very similarly to the phosphate group of FMN and recognized by three Arg residues from two different monomers. The 2Fo-Fc map surrounding this region is displayed as a black grid mesh and contoured at a 1  $\sigma$  level. d) Identification of two extra electron density peaks in the putative active site of MAB\_4123. The Fo-Fc map is displayed as a magenta grid mesh and contoured at a 4  $\sigma$  level. The top panel shows MAB\_4123 in which FMN has been docked as in b). The two unmodelled electron density peaks in the difference map are situated above the FMN flavin ring. The middle panel displays the superposition of the TdsC structure (PDB id 5XDE) (green cartoon and sticks) bound to FMN and DBT (both in grey sticks) on MAB\_4123. The low panel shows that DMSO fits the upper positive electron density peak and could be stacked by the side chain of Trp106 and Tyr99. It shows also that the oxygen group of DMSO could interact with the side chains of Ser102 and Ser245.