



**HAL**  
open science

# Designing a drug against tuberculosis using crystallography: fragment-based screening targeting MtDS

Emanuelle Albat

► **To cite this version:**

Emanuelle Albat. Designing a drug against tuberculosis using crystallography: fragment-based screening targeting MtDS. Life Sciences [q-bio]. 2016. dumas-01360742

**HAL Id: dumas-01360742**

**<https://dumas.ccsd.cnrs.fr/dumas-01360742>**

Submitted on 6 Sep 2016

**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.

AGROCAMPUS  
OUEST

CFR Angers

CFR Rennes



UNIVERSITÉ DE  
RENNES 1

UiO : Universitetet i Oslo



Année universitaire : 2015 – 2016

Spécialité : Agronome

Spécialisation (et option éventuelle) :

SCMV

Sciences Cellulaire et Moléculaire du vivant

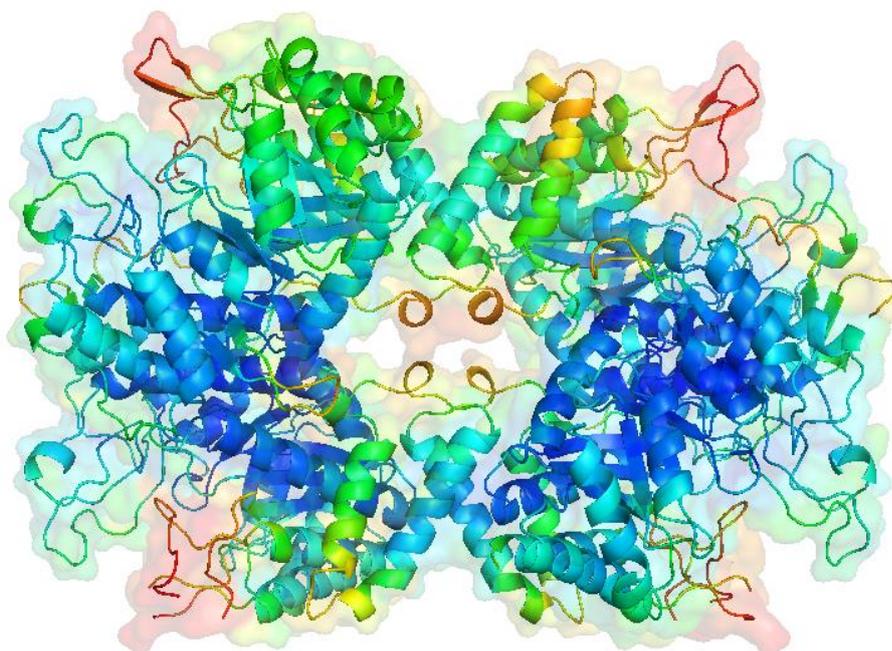
.....

### Mémoire de Fin d'Études

- d'Ingénieur de l'Institut Supérieur des Sciences agronomiques, agroalimentaires, horticoles et du paysage
- de Master de l'Institut Supérieur des Sciences agronomiques, agroalimentaires, horticoles et du paysage
- d'un autre établissement (étudiant arrivé en M2)

# Designing a drug against tuberculosis using crystallography: Fragment-based screening targeting MtDS

Emanuelle Albat



*Soutenu à Oslo*

*le 10 juin 2016*

**Devant le jury composé de :**

- Vincent Lagente
- Luc Paillard
- Pascale Quignon

*Les analyses et les conclusions de ce travail d'étudiant n'engagent que la responsabilité de son auteur et non celle d'AGROCAMPUS OUEST*

### **Confidentialité :**

Non  Oui si oui :  1 an  5 ans  10 ans

Pendant toute la durée de confidentialité, aucune diffusion du mémoire n'est possible<sup>(1)</sup>.

A la fin de la période de confidentialité, sa diffusion est soumise aux règles ci-dessous (droits d'auteur et autorisation de diffusion par l'enseignant).

Date et signature du maître de stage<sup>(2)</sup> :

### **Droits d'auteur :**

L'auteur<sup>(3)</sup> autorise la diffusion de son travail

Oui  Non

Si oui, il autorise

la diffusion papier du mémoire uniquement<sup>(4)</sup>

la diffusion papier du mémoire et la diffusion électronique du résumé

la diffusion papier et électronique du mémoire (joindre dans ce cas la fiche de conformité du mémoire numérique et le contrat de diffusion)

Date et signature de l'auteur :

### **Autorisation de diffusion par le responsable de spécialisation ou son représentant :**

L'enseignant juge le mémoire de qualité suffisante pour être diffusé

Oui  Non

Si non, seul le titre du mémoire apparaîtra dans les bases de données.

Si oui, il autorise

la diffusion papier du mémoire uniquement<sup>(4)</sup>

la diffusion papier du mémoire et la diffusion électronique du résumé

la diffusion papier et électronique du mémoire

Date et signature de l'enseignant :

(1) L'administration, les enseignants et les différents services de documentation d'AGROCAMPUS OUEST s'engagent à respecter cette confidentialité.

(2) Signature et cachet de l'organisme

(3).Auteur = étudiant qui réalise son mémoire de fin d'études

(4) La référence bibliographique (= Nom de l'auteur, titre du mémoire, année de soutenance, diplôme, spécialité et spécialisation/Option)) sera signalée dans les bases de données documentaires sans le résumé

Dépôt numérique de mémoire

ATTESTATION DE CONFORMITE DE LA VERSION NUMERIQUE

Je, soussigné(e),

Nom : .....Albat.....

Prénom : .....Emanuelle.....

Ci-après désigné « l'auteur »

Atteste que la version numérique de mon mémoire de fin d'études dans sa version définitive (incluant les corrections demandées par le jury de soutenance), intitulé

Designing a drug against tuberculosis using crystallography; fragment-based screening targeting MtDS

correspond à la version imprimée du document, déposé à la bibliothèque générale d'AGROCAMPUS OUEST (CFR de référence)

A.....Oslo....., le...20 juin 2016

Signature de l'auteur,

A handwritten signature in black ink, consisting of several overlapping loops and lines, positioned below the signature line.

## RESUME LONG

---

La tuberculose est la plus mortelle des maladies infectieuses avec le VIH de nos jours. Dans la plupart des cas il s'agit d'une maladie pulmonaire, causée par une bactérie, *Mycobacterium tuberculosis*. Les traitements actuellement proposés durent au minimum 6 mois et ont des effets secondaires toxiques notamment au niveau hépatique. De plus, il n'existe aujourd'hui aucun traitement réellement efficace contre les souches résistantes, qui sont un problème croissant, c'est pourquoi il y a un besoin urgent de trouver de nouveaux médicaments.

La voie du Shikimate est une voie de production des acides aminés aromatiques chez les bactéries, les plantes et les champignons, qui est absente chez l'homme. MtDS est une enzyme clef de cette voie, dont elle catalyse la première étape. De plus, elle forme un complexe avec une autre enzyme clef de la production d'acides aminés, nommée MtCM. Cette protéine a une très faible activité seule, mais la formation de ce complexe augmente d'un facteur 100 l'activité de MtCM. Du fait de sa position stratégique dans cette voie métabolique, et le fait que cette voie soit absente chez l'homme, MtDS apparaît comme étant un bon candidat pour être une cible de médicament.

Dans ce projet, le but était de trouver un ligand qui inhiberait MtDS, en utilisant un criblage fondé sur des fragments analysé par cristallographie. D'abord, MtDS a été produite en quantité suffisante, purifiée, puis les conditions optimales pour sa cristallisation ont été établies, enfin les cristaux de MtDS ont été mis en présence de ligands potentiels et le résultat a été observé par cristallographie aux rayons X.

CgDS est l'homologue de MtDS provenant de *Corynebacterium glutamicum*. C'est une protéine très proche de MtDS qui peut aider à la compréhension de cette dernière, et qui est bien plus facile à manipuler de manière générale. De plus, *Corynebacterium glutamicum* ne présente pas de risque biologique contrairement à *Mycobacterium tuberculosis*. MtDS et CgDS ont été exprimées en *E.coli* et purifiées en même temps, par un protocole similaire incluant une chromatographie d'affinité avec une colonne de nickel et une chromatographie d'exclusion stérique. La purification de CgDS requiert une étape supplémentaire, une chromatographie échangeuse d'ions.

Une bibliothèque de fragments aux propriétés pharmacologiques reconnues est utilisée. Les composants de cette bibliothèque sont conservés dans du DMSO. Les cristaux utilisés devaient donc être tolérants au DMSO. Des conditions dans lesquelles MtDS cristallise dans un milieu

contenant du DMSO ont été trouvées grâce à des screening commerciaux. Ces conditions ont été optimisées dans des plaques de 24 puits en utilisant la technique de la goutte suspendue. Le but de cette optimisation est d'obtenir des cristaux suffisamment larges pour pouvoir les manipuler facilement. Pour éviter que le faisceau de rayons X ne détruise les cristaux par échauffement, ceux-ci sont congelés par flux d'azote gazeux à 100K. Afin que la diffraction des rayons X ne soit pas perturbée par des cristaux de glace, leur formation doit être empêchée. Pour ce faire, des agents cryoprotecteurs ont été testés et ajoutés aux conditions lors de leur optimisation.

Ces cristaux ont été trempés pendant 24H dans des mélanges de composés provenant de la bibliothèque de fragments précédemment évoquée. Ces cristaux ont ensuite été récoltés, congelés puis analysés par cristallographie aux rayons X. Deux jeux de données ont pu être collectés à une résolution d'environ 3 Å, et un  $CC_{1/2}$  supérieur à 30%. A présent, ces jeux de données devront être modélisés, grâce à la méthode de remplacement moléculaire, puis ce modèle devra être affiné. Une fois ces étapes réalisées, il est possible de déterminer si un ligand s'est lié à la protéine. Si le résultat est positif, des mesures de l'activité devront aussi être conduites afin de déterminer si le ligand est un inhibiteur.

## ACKNOWLEDGEMENTS

---

I first want to thank Ute Kregel, my supervisor, for accepting me in the lab, giving me the chance to learn about crystallography, and for offering me the opportunity to go to the synchrotron which was an exciting experience.

My gratitude also goes to Helen Thorbjørnsrud, my co-supervisor, for teaching me most of the things I needed to know, answering all my annoying questions and spotting my awful so-french expressions!

The whole protein dungeon team definitely deserves a place in this section, for their warm welcoming into the lab and pieces of advice that were so helpful when I was a bit lost sometimes.

Last but not least, thank you, my significant other and my family, for being a psychological and financial support through harsh times. Thank you for believing in me.

## TABLE OF CONTENTS

---

Introduction .....	1
Material and Methods .....	7
Production and purification of proteins .....	7
Expression in <i>E.coli</i> cells .....	7
Cell lysis .....	7
Purification .....	8
Thermofluor shift assay .....	9
MtDS crystallography experiments .....	9
Screening crystallisation conditions .....	9
Optimisation in 24-well plates .....	9
MtDS seeds and seeding .....	10
Cryoprotection tests .....	10
Fragment-screening experiments .....	11
X-ray data collection .....	11
Data processing .....	11
Results .....	12
MtDS expression and purification .....	12
CgDS expression and purification .....	14
Thermofluor shift assay .....	17
MtDS crystallisation .....	17
Cryoprotectant concentration tests on the cryo stream .....	18
Fragment screening .....	19
X-ray analysis .....	19
Discussion .....	22
References .....	25

## ABBREVIATIONS AND COMMERCIAL NAMES

---

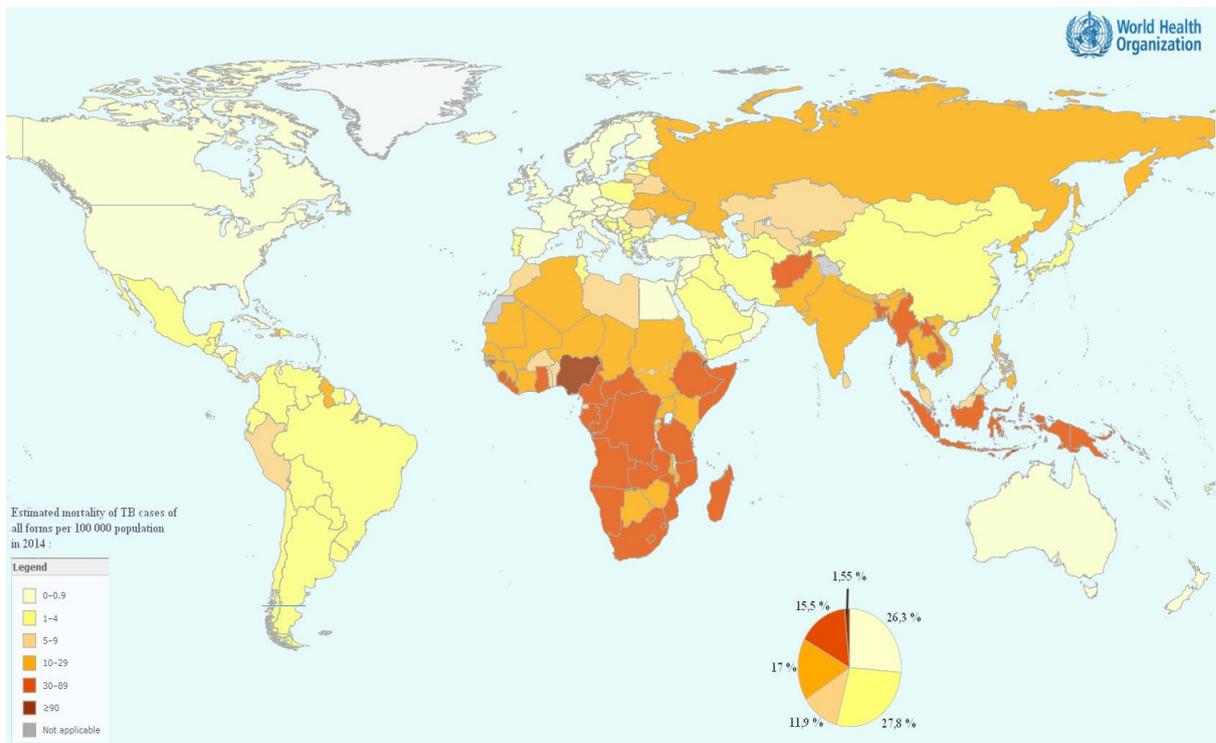
<b><u>AIDS:</u></b>	Acquired immunodeficiency syndrome
<b><u>ÄKTA:</u></b>	Protein purification system by GE Healthcare Life Sciences
<b><u>Amp:</u></b>	Ampicillin
<b><u>BTP:</u></b>	Bis-tris propane
<b><u>CellLytic™:</u></b>	B-cell lysis reagent from Sigma
<b><u>CgCM:</u></b>	<i>Corynebacterium glutamicum</i> chorismate mutase
<b><u>CgDS:</u></b>	<i>Corynebacterium glutamicum</i> DAHP synthase
<b><u>DAHP:</u></b>	3-Deoxy-D-Arabino-heptulosonic acid 7-phosphate
<b><u>E4P:</u></b>	Erythrose-4-phosphate
<b><u>EDTA:</u></b>	Ethylene diaminetetra acetic acid
<b><u>FPLC:</u></b>	Fast protein liquid chromatography
<b><u>HIV:</u></b>	Human immunodeficiency virus
<b><u>IPTG:</u></b>	Isopropyl-D-thiogalactopyranoside
<b><u>LB:</u></b>	Luria-Bertani medium
<b><u>mAU:</u></b>	milli absorbance unit
<b><u>Mtb:</u></b>	<i>Mycobacterium tuberculosis</i>
<b><u>MtCM:</u></b>	<i>Mycobacterium tuberculosis</i> chorismate mutase
<b><u>MtDS:</u></b>	<i>Mycobacterium tuberculosis</i> DAHP synthase
<b><u>PCR:</u></b>	Polymerase chain reaction
<b><u>PEP:</u></b>	Phospho-enol pyruvate
<b><u>Phe:</u></b>	Common three letters writing for phenylalanine
<b><u>PMSE:</u></b>	Phenylmethanesulfonyl fluoride
<b><u>Psi:</u></b>	Pound-force per square inch

<b><u>RT:</u></b>	Room temperature
<b><u>SEC:</u></b>	Size-exclusion chromatography
<b><u>SDS-PAGE:</u></b>	Sodium dodecyl sulfate polyacrylamide gel electrophoresis
<b><u>Tablet of cOmplete™ ULTRA EDTA-free:</u></b>	Cocktail of proteases from Roche
<b><u>TB:</u></b>	Tuberculosis
<b><u>TCEP:</u></b>	Tris(2-carboxyethyl)phosphine
<b><u>Trp:</u></b>	Three letters writing for tryptophan
<b><u>Tyr:</u></b>	Three letters writing for tyrosine
<b><u>UNICORN:</u></b>	Software used with ÄKTA systems
<b><u>WHO:</u></b>	World Health Organization
<b><u>XDS:</u></b>	Software used for the first step of crystallography data processing

## INTRODUCTION

Tuberculosis (TB) has been considered, since last October, as the most deadly infectious disease worldwide alongside AIDS (World Health Organization, 2015). It is a disease affecting the lungs in most cases, and is caused by a bacillus called *Mycobacterium tuberculosis* (Mtb). The World Health Organisation (WHO) estimates that 2 to 3 billion people are infected by TB, but for most of them, it will remain latent for their whole life. When the disease becomes active, if it is left untreated, around 50% patients die. In 2014, there were around 1,4 million of death because of TB (World Health Organization, 2015).

TB is present everywhere in the world. Around 3/5 of the cases are in South Asia and Western Pacific regions, and almost a third are in Africa (Fig. 1).



**Figure 1 - World tuberculosis' mortality map in 2014 (World Health Organization, 2015).** The legend is a percentage of the estimated mortality of TB cases in 2014, light yellow being below 0,5% and dark red being more than 30%.

TB treatments are exceptionally long and with toxic side effects. The actual best treatment known at this point is a 6-month intensive therapy with a 2-month “induction phase” followed by a 4-month of “consolidation phase” (Horsburgh et al., 2015). 5% to 33% of the treated patients have documented hepatotoxic effects. The large difference between these percentages comes from different definitions of “hepatotoxic effects” (Saukkonen et al., 2006). At least 15% patients show adverse drug reactions, making them interrupt or discontinue one or more of the drugs (Lv et al., 2013).

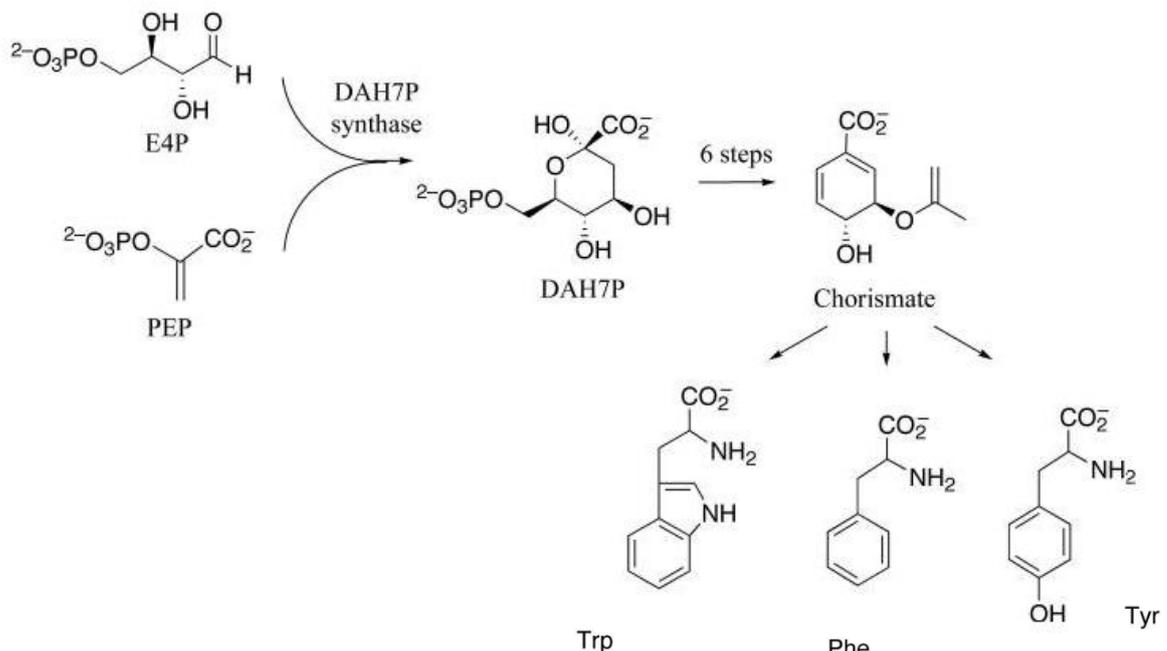
Another problem is the increase of drug-resistant Mtb strains, mainly present in relapsing people. They emerge from both practitioner errors and patients not correctly taking their treatments (Ormerod, 2005). In reoccurring infections, drug-resistant TB represents 20% of the cases whereas in new cases only 3,3% are drug-resistant (World Health Organization, 2015). Only a half of the patients having a drug-resistant infection are successfully treated. There are still debates on the best way to cure patients. The existing treatments proposed are between 6 and 18 month long (Horsburgh et al., 2015).

The worst prognosis for people is co-infection with both TB and HIV. These two pathogens act synergistically, provoking the depression of the immune system (Bruchfeld et al., 2015). There is a depletion of Mtb antigen-specific CD4+ T cells occurring early in HIV infection (Geldmacher et al., 2010). Also, there is an increased number of CD4+ T cells expressing CCR5 (which is an HIV coreceptor) at TB disease sites, contributing to making these cells susceptible to HIV infection (Matthews et al., 2012).

The number of these co-infected people is increasing, leading to a systematic testing of TB in HIV-patients and an aggressive approach in order to improve the situation (Montales et al., 2015). People infected by HIV are 20 to 30 more susceptible to develop active TB, more than a third of the HIV-positive people were infected with TB in 2015 (World Health Organization, 2015).

All these facts show the need of new treatments against TB. A potential drug target from the aromatic amino acid production was found. This pathway is called Shikimate pathway (Fig. 2) for Mtb. The Shikimate pathway is vital to Mtb (Parish and Stoker, 2002) but is absent in animals. Thus it does not exist in humans; if this pathway is blocked in Mtb, that should not have any impact on the patient.

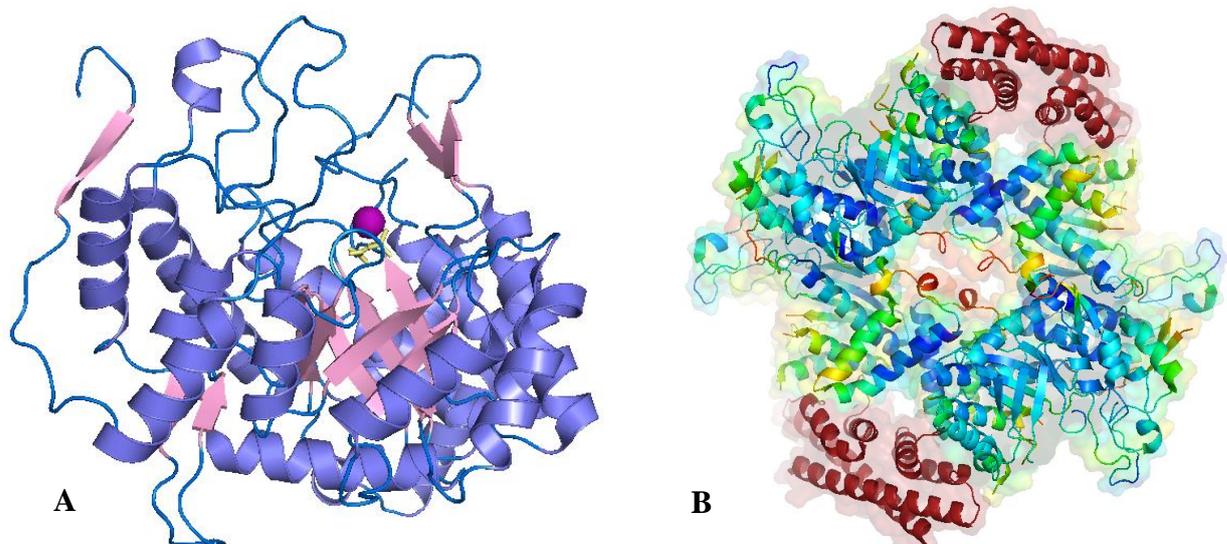
DAHP synthase catalyses the first step of the Shikimate pathway (Fig. 2) which is the transformation of E4P and PEP into DAHP. This project focused mainly on DAHP synthase from Mtb (MtDS). MtDS has a feedback regulation by the aromatic amino-acids. While individual amino-acids poorly affects its activity, the combination of Phe and Trp, Trp and Tyr or Phe and Tyr inhibit MtDS by around 50% or even more depending on whether it is tagged (Webby et al., 2010) or not (Munack et al., 2016). More than that, when all three aromatic amino-acids are bound, the activity of MtDS is completely abolished (Blackmore et al., 2013). The inhibitors do not bind in the active site, but in separate pockets. By allosteric changes, the amino acids decrease the affinity to E4P, preventing the reaction from occurring (Jiao et al., 2012) (Munack et al., 2016).



**Figure 2 - The shikimate pathway, from E4P and PEP to tryptophane, tyrosine and phenylalanine.** DAHP synthase catalyses the first step and at the branch point in the end is chorismate mutase. It catalyses the reaction transforming chorismate into prephenate, needed to produce both Tyr and Phe. Figure modified from Parish and Stoker, 2002.

MtDS is a protein composed of 552 residues. Its global shape is a core formed of a TIM-barrel of 8 beta-strands, each followed by alpha-helices, with additional 5 alpha-helices. The N-terminal beta-strand plays a key role in the dimerization of MtDS (Fig. 3). Manganese ( $Mn^{2+}$ ) presence is required in the active site to be catalytically active.

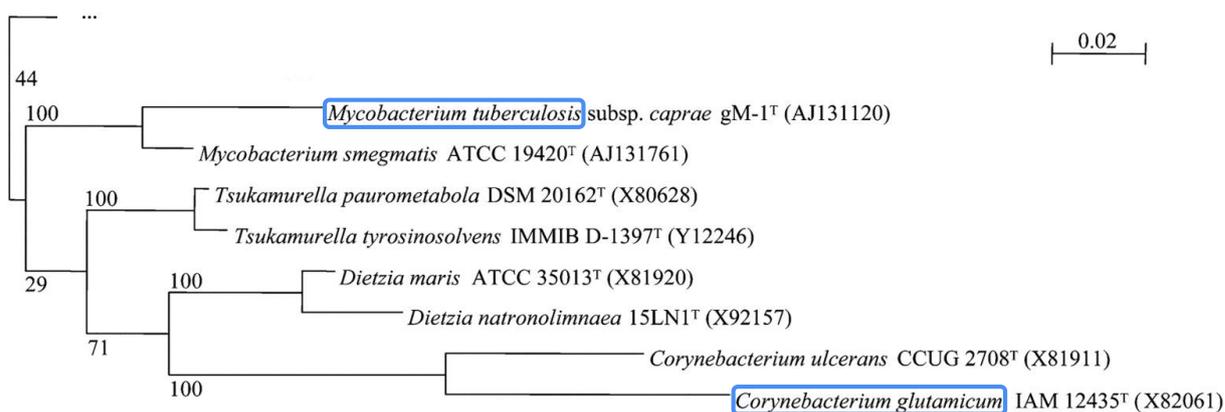
MtDS has a partner called chorismate mutase (MtCM), catalysing the branch point in the reaction pathway (Fig.2 and 3). MtCM alone is not very active (Prakash et al., 2005) (Künzler et al., 2005), the partner molecule MtDS is needed to increase its activity by forming a complex (Sasso et al., 2009).



**Figure 3 - MtDS with and without its partner MtCM.** A/ MtDS protomer. Alpha-helices in blue, beta-strands in light pink.  $Mn^{2+}$  is the pink sphere and in yellow is the PEP. B/ MtDS tetramer (spectrum colours following the B-factors) with two MtCM homodimers (in red).

This complex (Fig. 3) formation increases the activity of MtCM by more than 100-fold (Sasso et al., 2009). Thus, if MtDS is inhibited, MtCM is also inhibited. These enzymes, and particularly MtDS as the activating partner, are essential for the Shikimate pathway, as they are catalysing key steps. All these are excellent reasons to believe that MtDS could make a perfect target for a drug against Mtb.

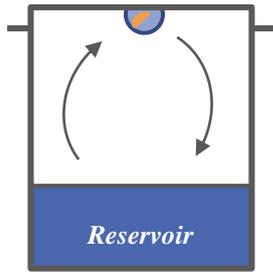
Moreover, *Corynebacterium glutamicum* is a bacteria much easier to handle and does not present any hazardous risk. CgDS (the DAHP Synthase from *Corynebacterium glutamicum*) is much easier to express and to crystallize, thus it was studied in parallel, as a protein that could help us understand MtDS. CgDS is used to produce aromatic amino-acids in industries (Ikeda, 2006). *Corynebacterium* and *Mycobacterium* are two genres belonging to the same suborder *Corynebacterineae*, they are close phylogenetically speaking (Fig. 4) and thus their proteins are also closely related. With a sequence alignment, 66,5% identity is found between MtDS (552 bp) and CgDS (472 bp).



**Figure 4 - Corynebacterineae Suborder.** *Mycobacterium tuberculosis* and *Corynebacterium glutamicum* are highlighted with blue circles. Figure adapted from Adachi et al., 2007.

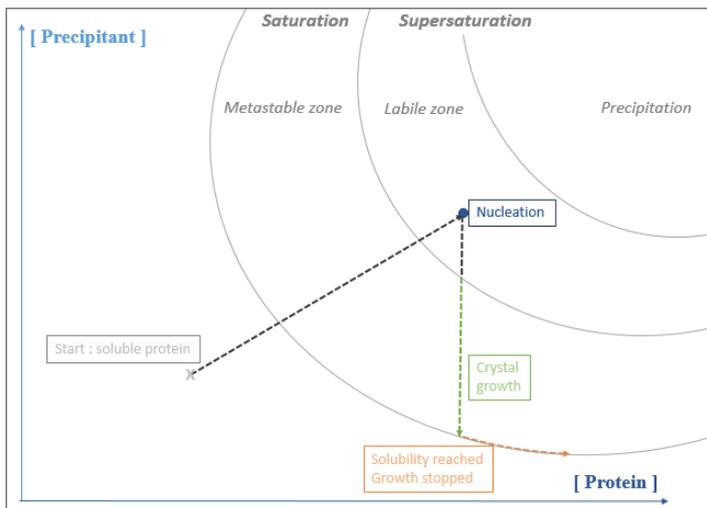
It was also shown that CgCM, which is the MtCM homologue from the same bacteria, can be activated by MtDS (Sasso et al., 2009). That suggests that a complex between CgCM and CgDS is formed, similar to the one known between MtCM and MtDS.

The goal of this study is to find an inhibitor against MtDS. Information like the inhibited/activated status of the protein, where the ligand is bound, to which residues, etc... is needed in order to understand the inhibition mechanism and refine the drug. Crystallography was used to provide these details. A library of drug-like small molecules was brought together with the protein in the hope that one would inhibit it and be a potential future drug against TB. This type of drug-discovery experiment is called a fragment-based screening (Davies and Tickle, 2012).



**Figure 5 - Equilibration between the drop and the reservoir solution**

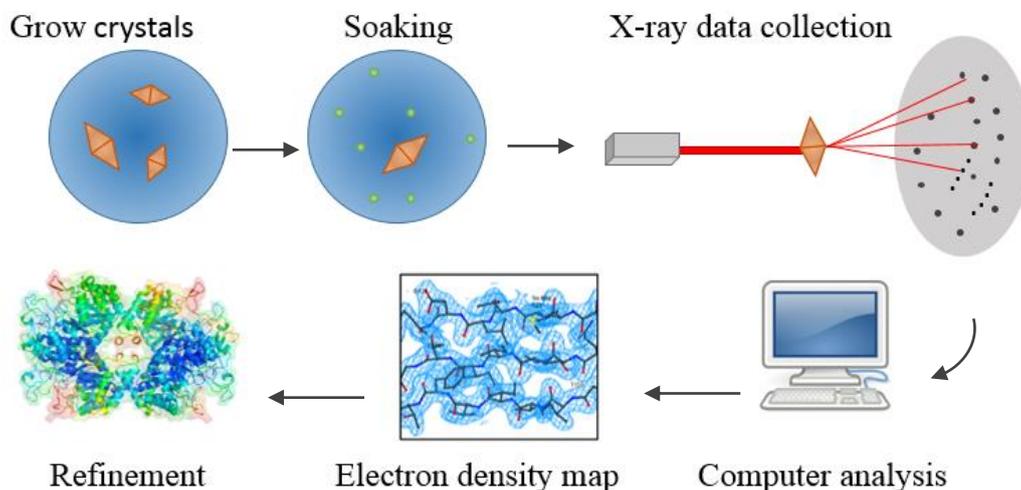
The technique requires protein crystals. The process involves mixing the protein solution with a precipitant. The drop of this mix equilibrates with a reservoir solution (Fig.5) which composition is similar, apart from the presence of the protein. As it equilibrates, the precipitant concentration increases, and the solubility of the protein decrease. There is less and less solute, the protein is saturated. When it reaches the supersaturation, it starts to self-organise (nucleation) then hopefully crystallises (Fig.6).



**Figure 6 - Evolution of protein crystallisation**

The crystal growth can be influenced by the pH, temperature, the type and concentration of precipitant, the type of buffer that is used or the presence of additives. This is not an exhaustive list but it is mainly what is changed in order to find the best crystal growth conditions for a protein. These conditions usually are protein-specific.

A crystal should also be cryo-protected, either by growing in cryoprotective conditions or by being soaked in cryoprotectant just before being flash-frozen. Otherwise, when the sample is frozen, ice crystals may form in addition to the protein crystals and will also make the X-rays diffract. It is important to find an equilibrium between all these parameters in order to have a crystal big enough to work with, otherwise it is difficult to handle.



**Figure 7 - Overview of crystallography methods**

A crystal is composed of repetitive motifs (which here can be MtDS protomer or dimers for example) that would, when hit by X-rays, give a particular diffraction pattern. This diffraction pattern is defined by a number of periodic functions. By observing the intensities and the symmetry between the spots, the amplitude of these functions can be found but the phase is lost. This is called the phase problem. In our case it is solved by using existing models that have similarities with our protein, it is called a Molecular Replacement (MR) method.

With all these information, a Fourier transform is used to have an electron-density map, which will be refined into the final model. This project aims at finding an MtDS inhibitor by a fragment-based screening method. First the protein must be produced, then conditions that suits best for its crystallisation must be found and optimized, then the fragment-based screening can be performed through crystals soaking.

# MATERIAL AND METHODS

---

## PRODUCTION AND PURIFICATION OF PROTEINS

---

The expression in *E.coli* cells and the purification steps were made according to previous protocols (Munack et al., 2016). For the production of the polyhistidine-tagged MtDS, a glycerol stock was used, containing the *E.coli* strain KA12 containing the plasmid pKTDS-HN. For CgDS, which also was histidine-tagged, the glycerol stock was made of KA13 *E.coli* strain containing the pKCGDS-HN plasmid.

### Expression in *E.coli* cells

For MtDS, precultures were grown in LB medium supplemented by 150µg/mL of ampicillin (Amp). This mix was inoculated by KA12-pKTDS-HN and grown overnight at 25°C. For CgDS, the amount of Amp is 100µg/mL and KA13-pKCGDS-HN bacteria were grown at 20°C.

Main culture of MtDS were grown in M9c/P medium (Künzler et al., 2005) supplemented with the 3 aromatic L-amino-acids at 40 µg/mL each. The M9c/P medium was composed of salts supplemented by many additives; different metal ions, glucose and thiamine. Ampicillin at 150µg/mL was also added. 5mL of preculture was used to inoculate 500mL of this medium in 2L-flasks. Cultures were grown for 24h at 25°C.

When the cultures reached OD<sub>600</sub> between 0,4 and 0,6 it was induced by 100µM salicylate and grown for another 24h, still at 25°C. For CgDS, the main culture grew in LB medium supplemented by 100µg/mL Amp at 20°C within hours. Induction with 100 µM IPTG was started when OD<sub>600</sub> = 0,6 and cells were harvested 4h after induction. The cells were harvested by centrifugation at 4000g for 20min at 4°C, rotor JLA 8.1. The pellet was stored at -80°C until being used. Samples of the culture were taken before induction and before harvesting for analysis by SDS-PAGE.

### Cell lysis

From this point, the protein was kept on ice. The same was done for CgDS. Cells expressing MtDS were resuspended in a buffer named BTP++, containing 20mM Bis-Tris Propane, 150mM NaCl, 0,5mM TCEP, 0,2mM PEP and 0,1mM MnCl<sub>2</sub>. BTP++ was supplemented by DNase I, 1 tablet of cComplete™ ULTRA EDTA-free per 50mL (protease inhibitor cocktail) and 150µM PMSF. 10 times the volume of the pellet were added here (in order to prevent the clogging of the homogenizer), meaning that for 16g of pellet, 160mL of the lysis buffer was added. Cell lysis was then realised by three passages through the high-pressure homogenizer with >15 000 psi.

For CgDS, cells were resuspended in 50 mM Tris-HCl pH 8, 300 mM NaCl, 20 mM imidazole, 5% glycerol, 100  $\mu$ M MnCl<sub>2</sub>, 200  $\mu$ M PEP, 1X celLytic™ solution from Sigma, 0,2 mg/mL lysozyme, 1 tablet of cOmplete™ ULTRA EDTA-free per 50 mL and 150 $\mu$ M PMSF. 5 times the pellet volume of solution was added to the cells, and it was vortexed until a completely homogeneous mix was obtained. Then the lysate was incubated at room temperature (RT) for 30 min on a shaker (130 rpm). Then DNase I granules were added and the lysate was incubated for another 20 min on the shaker, at RT.

In both cases, cell debris were pelleted by centrifugation at 48 000 g, 30 min, 4°C. The pellets were stored at -80°C and samples of the lysate, pellet and supernatant were kept for SDS-PAGE analysis, to ensure bacteria were successfully lysed. The lysate was filtered (0,4  $\mu$ m pores).

### Purification

The first step of purification was performed using a 5 mL nickel-column HiTrap Chelating™ HP from GE healthcare Life Sciences with an ÄKTA FPLC system place in a cooling compartment (10°C). The column was stripped and pre-equilibrated with BTP++ containing 20 mM imidazole (His Buffer A). BTP++ containing 500 mM imidazole was His Buffer B. First the lysate was loaded onto the column, and 3 column volume of His Buffer A was used to wash the column. Then a 2-step gradient was used for the elution between His Buffer A and His Buffer B. The elution stage could be followed by UV absorbance at 280 nm with the software UNICORN, and the fractions were analysed by SDS-PAGE stained with coomassie blue. The samples showing a band at the same molecular size than MtDS and of the same purity were pooled together and dialysed overnight in snake skin (10 kDa) in 4 L BTP++ in order to remove the imidazole. It was then concentrated (MWCO 10 kDa) down to 500  $\mu$ L by centrifugation at 4000 g. The sample was then injected onto the pre-equilibrated 25 mL size-exclusion column (SEC) Superdex™ HiLoad 75, with BTP++. As for the nickel-column, the process could be followed thanks to the UV detector, and fractions are analysed by SDS-PAGE. Similar fractions apparently containing MtDS were combined, concentrated and filtered. The enzyme concentration was then measured by nanodrop and was stored at -80°C.

For CgDS, the overall protocol was the same as for MtDS. For the nickel affinity purification, the Buffer A contained 50 mM Tris-HCl pH 8, 300 mM NaCl, 20 mM imidazole, 5% glycerol, 100  $\mu$ M MnCl<sub>2</sub>, 200  $\mu$ M PEP and for Buffer B, only imidazole changed to 500 mM.

After the nickel column, an anion-exchange column was used, a 5 mL HiTrap Q XL column. The buffers for this one were Anion Buffer A (BTP++ without NaCl) and Anion Buffer B

(BTP++ with 1M NaCl instead of 150 mM). As for the nickel-column, the sample was loaded, washing with Anion Buffer A was done, then a gradient (one step only) was realised between Anion Buffer A and Anion Buffer B. Another round was run with the wash and flow-through as they were still showing a strong band of CgDS, probably because the column may have been overloaded (see results). The fractions containing CgDS were pooled together, diluted with BTP++ to avoid a possible precipitation due to high concentration of NaCl then concentrated in cycles, down to 500  $\mu$ L. After filtration, the sample was loaded onto the size-exclusion column. The rest of the protocol was identical as MtDS' one.

## THERMOFLUOR SHIFT ASSAY

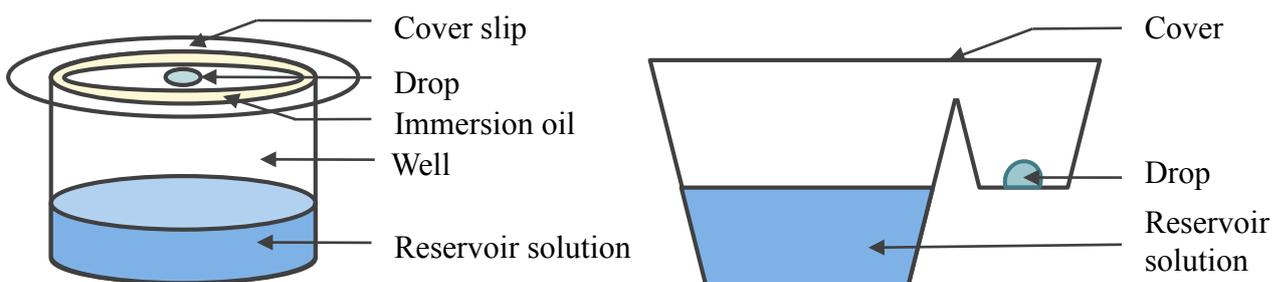
In these experiments, the protein was put in the presence of a mixture of compounds (Zenobia library) stored in 30% DMSO. The library contained 280 compounds that were tested in a 96-well PCR plate. In two wells, 30% DMSO was added instead of the compounds, as a control. Each well contained 4 compounds at 5mM each (or 30% DMSO for the control), the protein at 0,1 mg/mL, 1:1000 Sypro Orange dye, all in BTP++ (1X). The solutions were centrifuged then tested with a LightCycle 480 (Roche). The temperature was programmed to be gradually increased from 20°C until 90°C, increasing by 0,3°C at each step with 3 seconds hold time.

## MTDS CRYSTALLOGRAPHY EXPERIMENTS

### Screening crystallisation conditions

The library compounds were kept in 30% DMSO. Thus, MtDS crystals tolerant to DMSO were required. Their best growing conditions needed to be found. 10% DMSO was added to every plate. For the initial tests, screening plates of 96-well were used, with conditions from commercial screens differing in pH, type of precipitant, additives and concentration of the components. A JCSG-plus screen (Molecular dimension) and a Morpheus screen (Molecular dimension) were used. Sitting drops (Fig.8) were set up by a robot Oryx 4 (Douglas instruments). Conditions that gave crystals are optimised in 24-well plates.

### Optimisation in 24-well plates



**Figure 8 – Hanging and sitting drops.** Left side: hanging drop setting. Right side: sitting drop setting.

A 24-well plate is prepared by placing 500  $\mu\text{L}$  of a reservoir solution in each well. Immersion oil is laid upon the rims of the wells. Then a drop of 1  $\mu\text{L}$  of the protein and 1  $\mu\text{L}$  of the reservoir is pipetted on pre-siliconized cover slips, without mixing them. The cover slip is then flipped over and put on the well, which is then sealed. The drop will equilibrate with the well solution, and that vapour diffusion will allow the crystal growth (see introduction). Crystal growth is followed on a regular-basis with a microscope.

### MtDS seeds and seeding

In order to trigger crystal growth, different techniques can be tested. One of them is microseeding: the idea is to crush an existing crystal into small parts called seeds. These seeds are added into another hanging drop, where no crystal grew yet.

When a small part of a crystal is transferred into a drop, a crystal of the same protein is already present, less of the precipitant is needed to continue the process. It can trigger nucleation and crystal growth. For the creation of the seeds, a metal tool is used to crush the crystals, then they are transferred into 50  $\mu\text{L}$  of the reservoir solution in which the crystals grew by pipetting. A glass bead is added and the mix is vortexed for 2 to 3 min. The bead is removed and dilutions of 1/10, 1/100 and 1/1000 of seeds stock solutions are prepared. Transfer of the whole crystal can also be done; in that case it is called macroseeding.

Streak seeding was also used. In this type of seeding, a hair (like a cat whisker) is used to extricate microseeds from a crystal in a drop, and is then passed through other drops. Whether it is because of the crystals transferred with it or only because disturbance of the previous equilibrium and impurities can trigger nucleation leading to crystal growth.

### Cryoprotection tests

After finding conditions that gave crystals, tests were made to determine the best cryoprotectant. Solutions in which only the cryoprotectant changed were made. They contained 0,1 M HEPES pH 7,5, 10% DMSO, 1,8 M ammonium sulfate, 0,25mM TCEP and then 5, 10 or 15% of cryoprotectant that were MPD, ethylene glycol, glycerol or PEG 400. A clean loop was immersed into the solution and frozen by mounting it in the cryostream. The appearance was visually examined.

## FRAGMENT-SCREENING EXPERIMENTS

The fragments library used (Zenobia Therapeutics) contains 352 compounds divided into four 96-well plates, 100  $\mu\text{L}$  of 200 mM of each compound by well. These compounds are small molecules that are known for having drug-like properties. The compounds are added to the drops in a hanging-drops 24-well plate, similar to a crystallisation plate (500  $\mu\text{L}$  reservoir, 2  $\mu\text{L}$  drop, sealed with immersion oil...). As advised by the company, the soaking solutions that were made as mixtures of 8 compounds each (called “mix 1” to “mix 37”). The largest crystals from the optimisation plates were chosen. The reservoir solutions of the soaking plate were set according to the conditions in which these crystals grew. Only DMSO concentration was changed. It was lowered to 6% instead of 10%, because the compound-mix already contains 30% DMSO, and too much DMSO dissolved the crystals (see results). Then the drops were set with 1  $\mu\text{L}$  of the reservoir, 0,5  $\mu\text{L}$  ultrapure water and 0,5  $\mu\text{L}$  of a compound mix (6,25 mM final concentration for each compound). For each drop, MtDS crystals were fished and transferred into it, the cover slip was put onto the corresponding well, sealed with immersion oil as for a crystallisation plate, and left overnight. After a 24 h soaking, the crystals were fished, frozen on the cryostream, and stored in liquid nitrogen before being sent to the ESRF synchrotron (Grenoble, France) for data collection.

### X-ray data collection

Data sets were recorded at the ESRF synchrotron in Grenoble, France, beamline ID 29. The crystal is centred in the beam and slowly rotated around a defined angle. The image of the scattered X-rays diffraction pattern for each angle was saved.

### Data processing

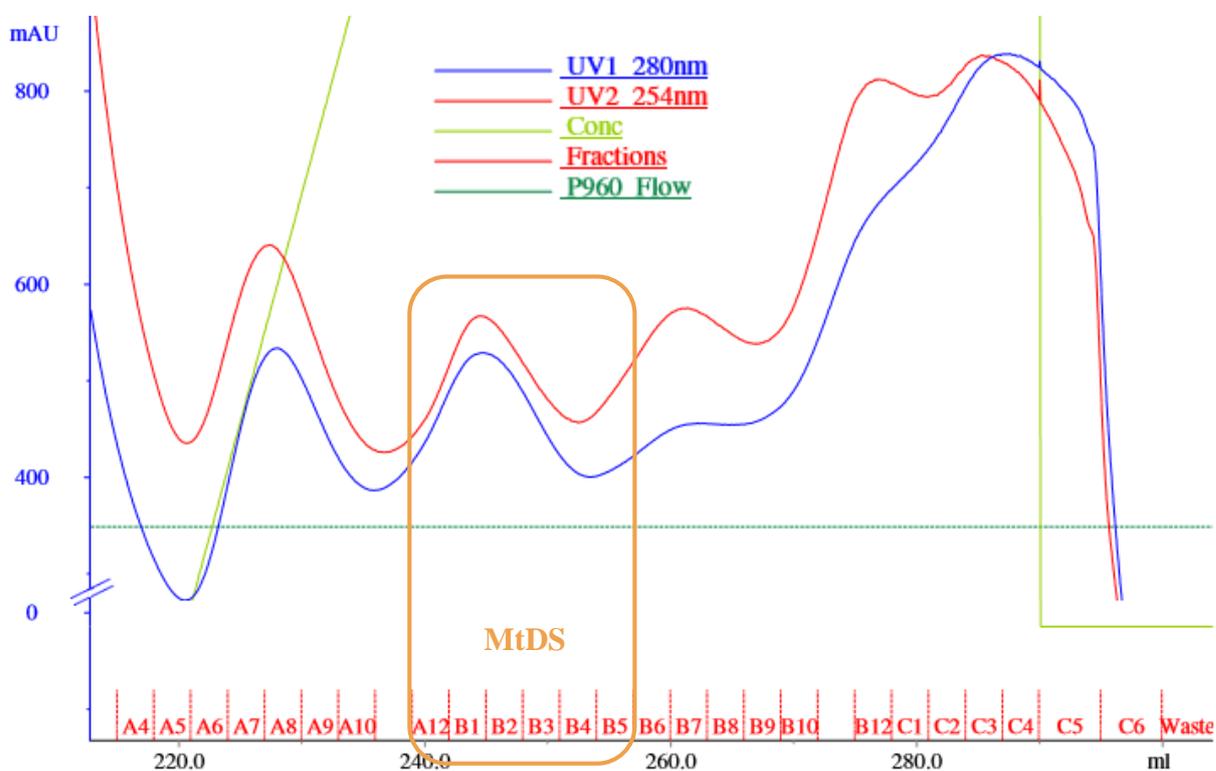
For the first steps, a software called XDS was used in order to do the first steps of the data processing. It starts with indexing of the diffraction pattern, which will determine the unit cell and its orientation, then the program integrates the diffraction intensities. With both the symmetry in the intensities of the diffraction pattern and the systematic absences, the space group is predicted. Then all the images are put on the same scale and merged together and statistics are calculated to evaluate the quality of the dataset.

## RESULTS

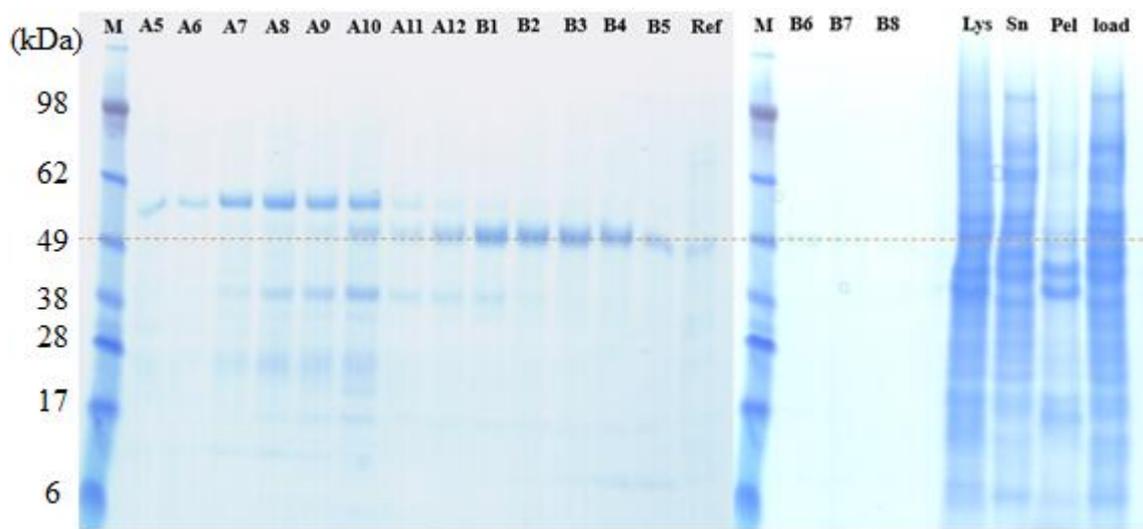
### MtDS expression and purification

The protein generally gives low yields. The best protocol gave with 2 and 8 mg of protein per litre of culture (Munack et al., 2016). MtDS was first expressed with 3 L of the main culture, resulting in 100  $\mu$ L of 3,6 mg/mL of protein meaning an extremely low yield of 0,12 mg/L of culture. Therefore, the expression was retried with 17 L of the main culture.

With 16 g of bacterial pellet, 180 mL of lysate was loaded on the nickel-column, but even though the loading of the lysate looked saturated (4200 mAU for UV 280 nm), not much of the protein could be seen already on the chromatogram (830 mAU for UV 280 nm, Fig. 9).

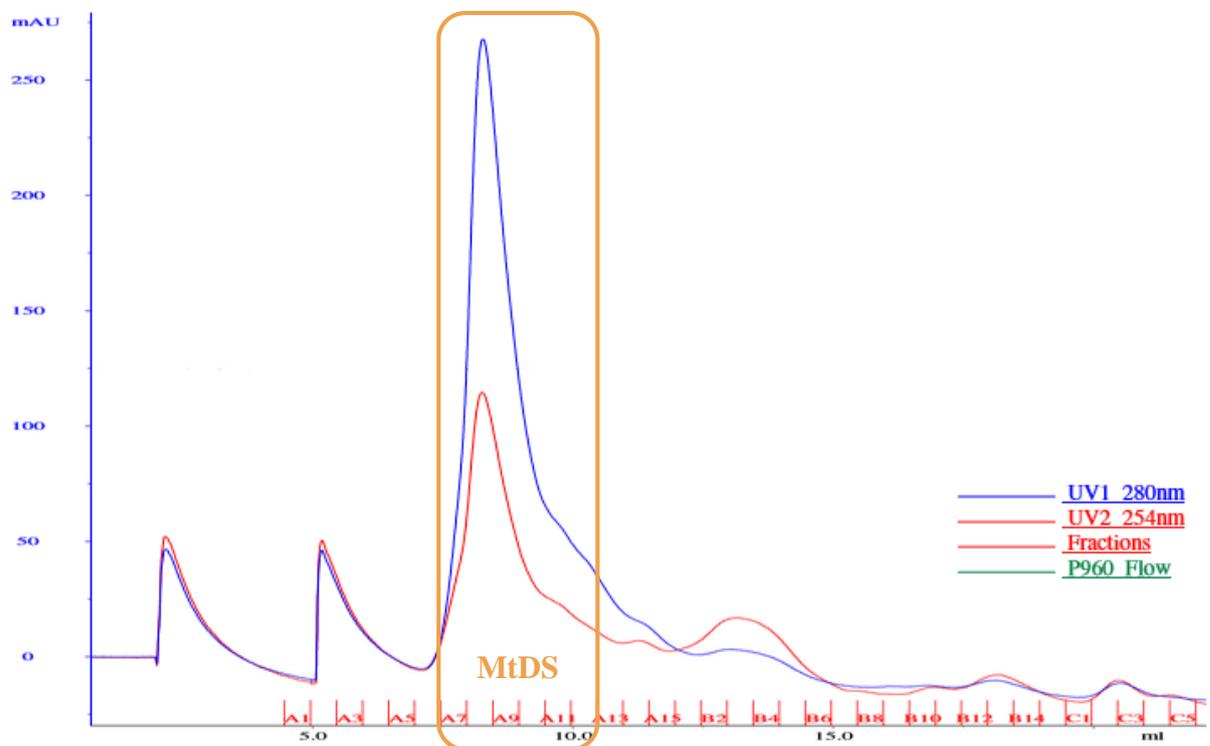


**Figure 9 - Nickel-column purification chromatogram.** The blue and red curves are the UV 280 nm and 254 nm that are related to protein concentration. The light green is the concentration of the His Buffer B. The names of the fractions are indicated in the lower part. In order to facilitate the reading, the washing part is not shown. Different peaks can be observed, corresponding to different proteins present in the lysate. The SDS-PAGE gel (see next figure) revealed that MtDS is present in the fractions A12 to B5 (shown by the orange markings).



**Figure 10 - Nickel-column fractions analysed by SDS-PAGE gel.** The names are taken by reference on the fractions seen on the chromatogram. "Ref" stands for reference; it is pure MtDS from the other purification process. "Lys" is for lysate, Sn for supernatant (after centrifugation of the lysate), pel for pellet (same) and load is a sample of what was loaded onto the column, it should be the same than supernatant. The light orange line shows the MtDS band (52 kDa).

Fractions B1 to B5 were combined. Although A10, A11 and A12 show the MtDS band, they are less pure and show another heavier band (Fig. 10) that was not possible to get rid of, by looking at chromatograms of SEC from previous purifications, thus they were not pooled with the others. The SEC chromatogram was as follows (Fig. 11):

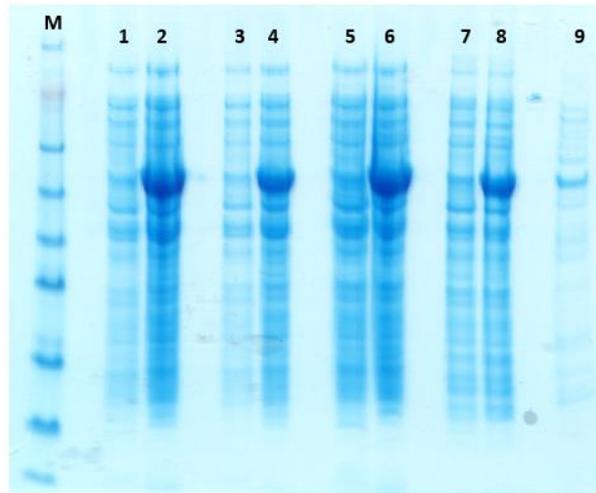


**Figure 11 - Size-exclusion column chromatogram.** Same legend than Fig. 9 without gradient. Also here, in order to facilitate the reading, only the fraction part is shown. MtDS peak appears clearly (orange markings).

Fractions A7 to A12 were combined and concentrated. Only, when the concentration was tested, it appears that the concentrated MtDS contained 200  $\mu$ L of 1,01 mg/mL.

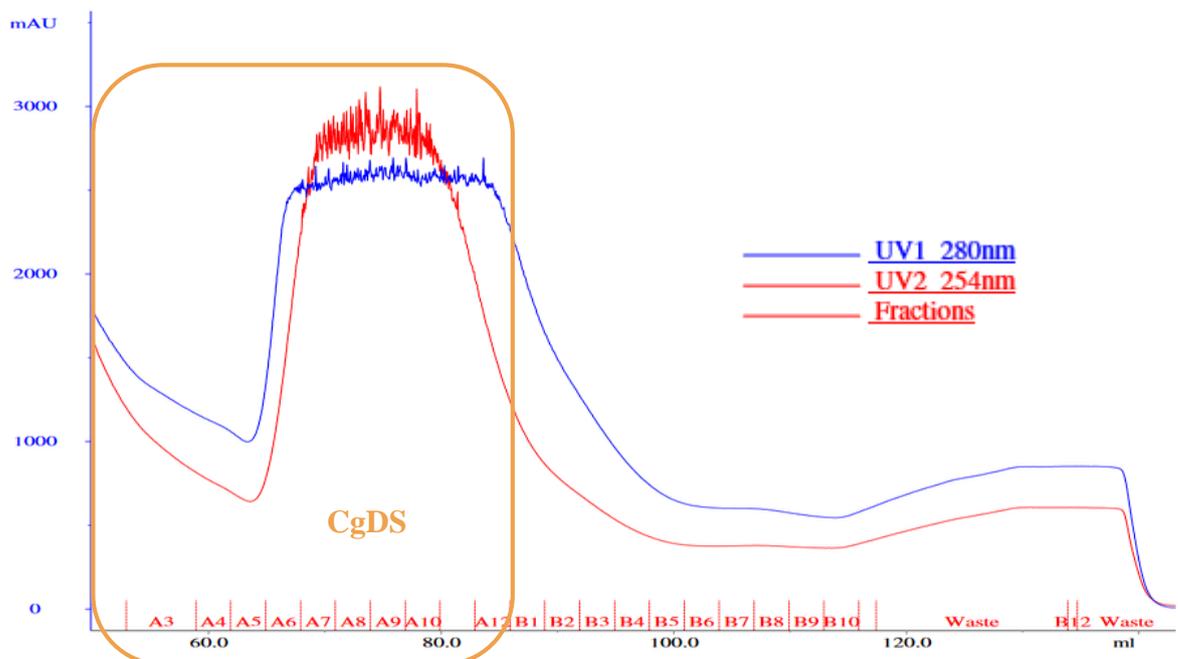
### CgDS expression and purification

The expression was followed by SDS-PAGE gel (Fig. 12).

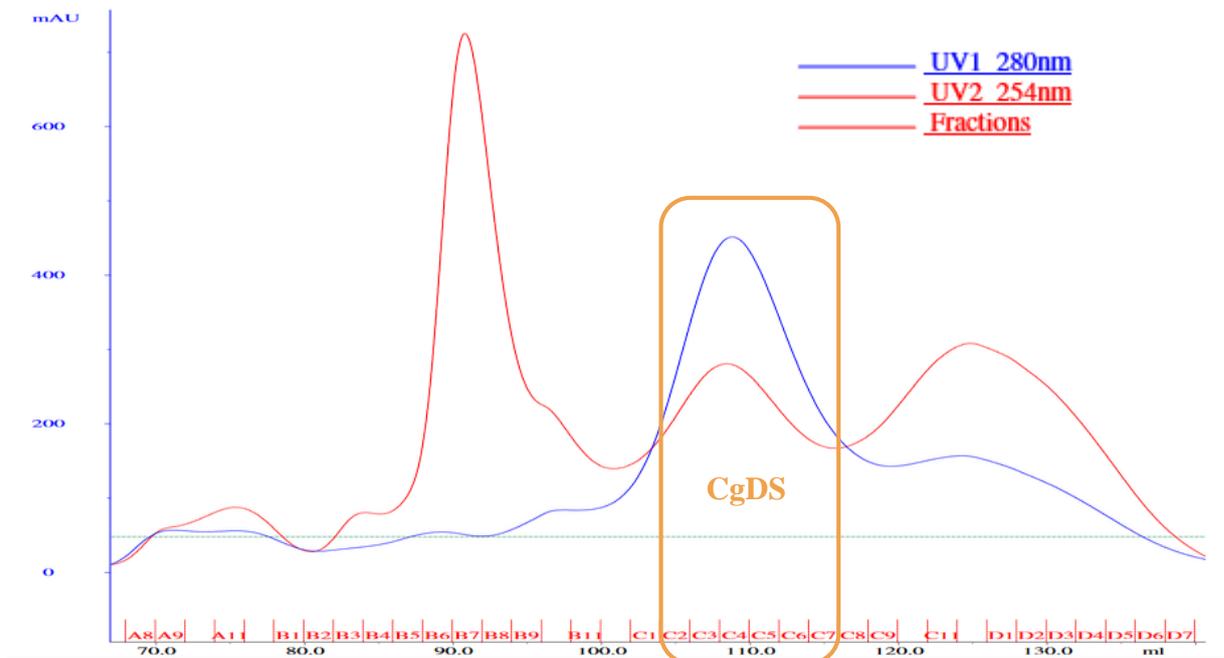


**Figure 12 - CgDS is strongly induced by IPTG.** On this SDS-PAGE gel were loaded whole cells samples from the bacterial cultures. From left to right: the protein ladder, then 1 to 4 are from the same glycerol stock and 5 to 8 from another. In 1, 3, 5 and 7 were loaded cells before induction and in 2, 4, 6 and 8 were loaded cells after induction. Half less sample were loaded in 1 and 2, 5 and 6 compared respectively to 3 and 4, 7 and 8. 9 is the control, 3  $\mu$ L from another CgDS expression.

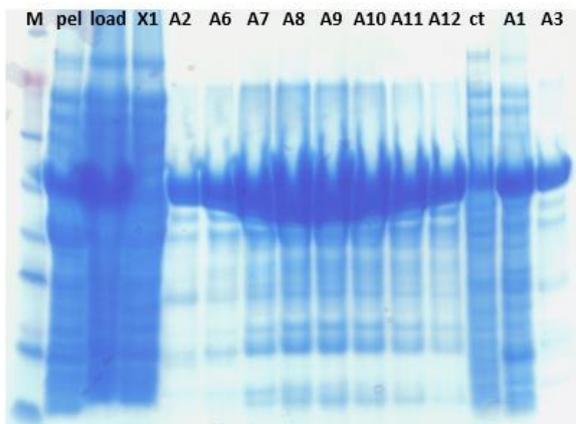
Then the nickel-column followed with a gel revealing apparently great amounts of protein at 52 kDa which is the theoretical molecular weight of CgDS.



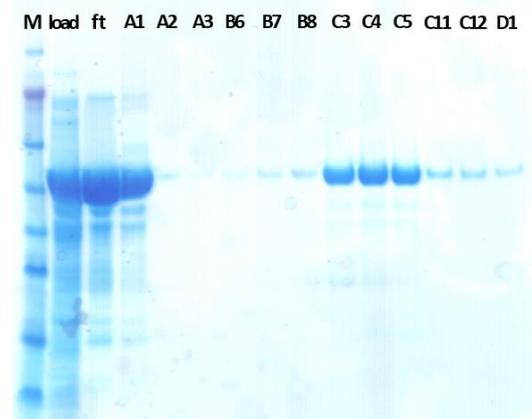
**Figure 13 - CgDS nickel-column chromatogram was overloaded with protein** The CgDS peak marked in orange is very clear and the UV 280 nm at the plateau is 2500 nm which is high.



**Figure 14 - The anion exchange column on CgDS separated different proteins.** This time, different peaks can be seen, corresponding to different proteins. Our protein of interest is more related to UV 280, marked in orange.



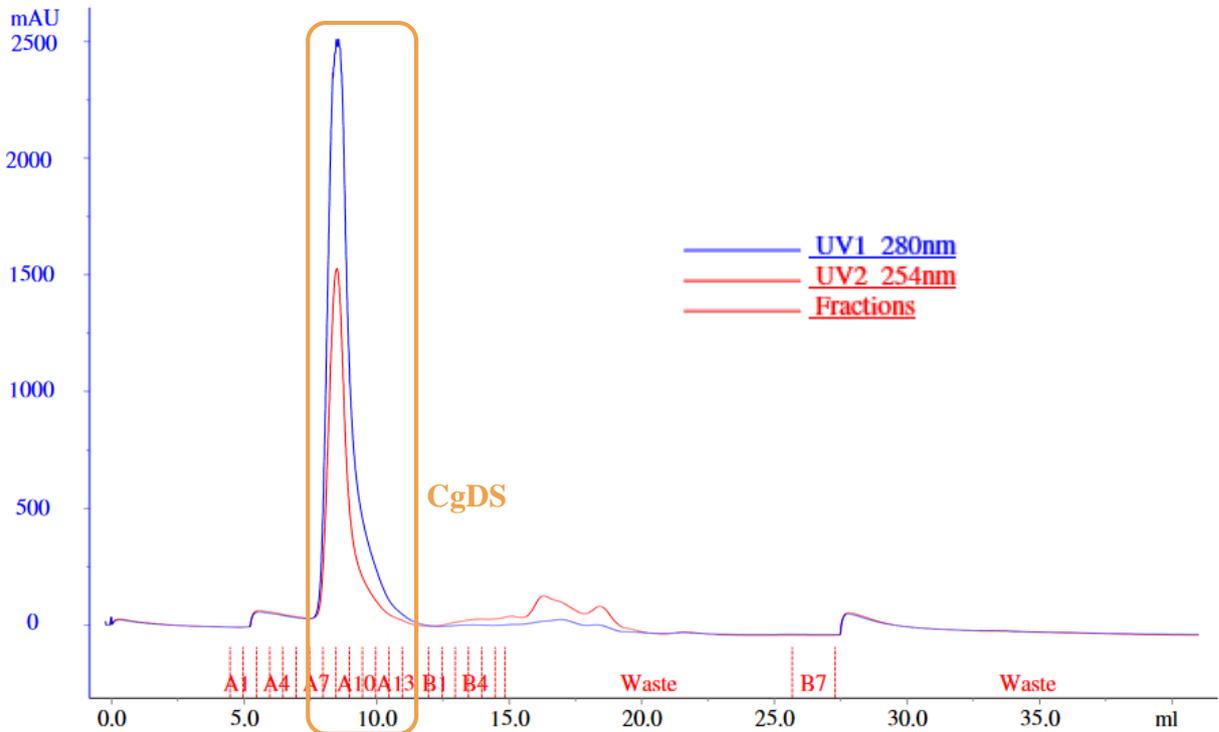
**Figure 15 - SDS-PAGE gel of the nickel-column shows it is overloaded with CgDS.** Samples of the fractions (Fig. 13) were loaded on this SDS-PAGE. M is the ladder, pel stands for the pellet after lysis and centrifugation, the rest are the fractions. The same amount of sample was loaded compared to MtDS, Fig. 10.



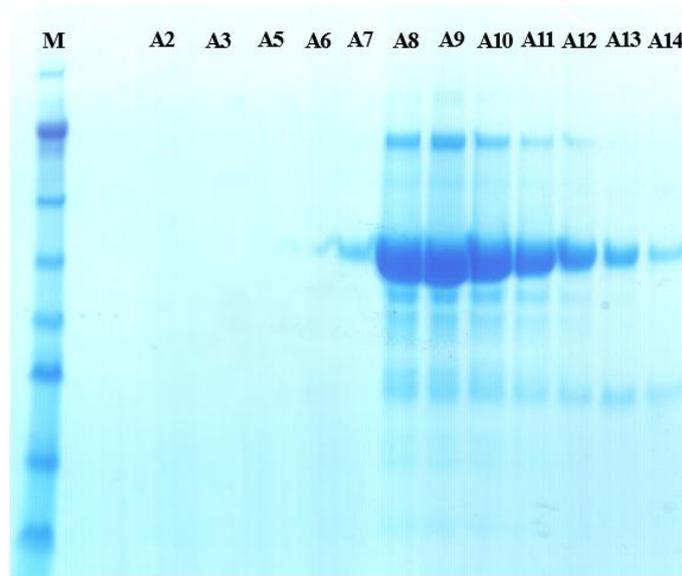
**Figure 16 - The SDS-PAGE gel after the anion-exchange column shows a lot less of CgDS but a lot purer.** « Load » is what was loaded onto the column, « ft » stands for « flow-through » (of the column). The others are the fractions designated (Fig. 14).

For the nickel-column (Fig. 13 and 15), it is obvious on the gel that in A1, A2 and A3, that a lot of the protein went through and did not bind to the column. A1 being very impure, it was kept separated from the other fractions. A2 to A12 are similar and thus were pooled together and loaded on the anion-exchange column (Fig. 14 and 16). There also a lot was lost in the flow-through and washing of the column, probably because it was too much protein for the

column to bind it all. Another anion-exchange was run with the flow-through and wash, and the resulting fractions containing CgDS were pooled with fractions C2 to C7 from the previous run. There was still protein observed in the flow-through and wash that was stored at  $-80^{\circ}\text{C}$ . Then the size-exclusion was done (Fig. 17 and 18).



**Figure 17 - The size-exclusion column shows a clear CgDS peak.** The fractions kept according to the SDS-PAGE gel (Fig. 18) are marked in orange.



**Figure 18 - SDS-PAGE gel after the size-exclusion column.** On this gel were loaded the samples named after the fractions of the size-exclusion column. A7 to A14 were pooled together

From these results, it can be noticed that getting rid of all impurities was not achieved. Nevertheless, as CgDS is highly concentrated compared to the other proteins in solution (see the intensity of the bands Fig. 18), and according to previous work (personal communications) it is pure enough to crystallise it.

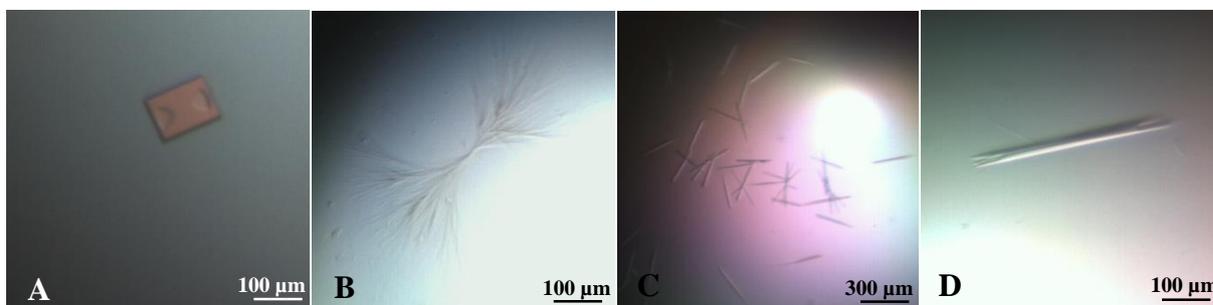
The concentration of the protein was 4,5 mg/mL for 1mL total solution meaning a yield of 1,5 mg of protein per litre of culture.

### Thermofluor shift assay

These experiments were designed to give a preliminary idea of which ligand would likely bind to MtDS and which would not. The denaturation of the protein is followed by fluorescence. When a ligand binds to the protein, a shift should be observed in the denaturation temperature measured during the experiment. Only, an unknown technical error happened and the temperature rose from 20°C to 90°C in 4 minutes instead of 15 minutes. Plus, in the experimental data, it could be observed that some compounds were already fluorescent without protein denaturation.

### MtDS crystallisation

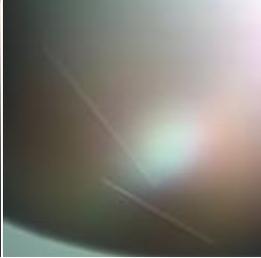
With the initial screening, 192 different conditions were tested. Here are examples of the many types of crystals and crystallisation forms that can be found (photos from my plates):



**Figure 19 - Different types of crystals in MtDS plates.** A/ Crystal that is very likely to be salt according to previous results. B/ This shape is called « needles ». They break very easily, but in a screening they are an encouraging sign. C/ Needle-like crystals. Lots of MtDS crystals looked like this. These are still quite fragile and hard to handle. D/ Thicker crystal, those are less fragile and can be used for data collection.

Conditions giving the best crystals were identified as followed for a protein at 1,95 mg/mL:

Conditions/well	Plate 09.03			Plate 02.03
	A3	A5	D6	D3
Precipitant	1,8M (NH <sub>4</sub> ) <sub>2</sub> SO <sub>4</sub>	1,6M (NH <sub>4</sub> ) <sub>2</sub> SO <sub>4</sub>	1,8M (NH <sub>4</sub> ) <sub>2</sub> SO <sub>4</sub>	1,8M (NH <sub>4</sub> ) <sub>2</sub> SO <sub>4</sub>
Buffer (0,1M)	HEPES pH 7,5	HEPES pH 7,5	HEPES pH 7,5	Tris pH 7,0
Cryoprotectant	5% Glycerol	10% Glycerol	6% Glycerol	5% Glycerol
Microseeding	1/100	1/100	1/1000	0
Additives	10% DMSO 0,25mM TCEP	10% DMSO 0,25mM TCEP	10% DMSO 0,50mM TCEP	10% DMSO
Crystals				

Plate 02.03	Plate 22.03		
D6	A5	B1	C1
1,9M (NH <sub>4</sub> ) <sub>2</sub> SO <sub>4</sub>	1,8M (NH <sub>4</sub> ) <sub>2</sub> SO <sub>4</sub>	1,8M (NH <sub>4</sub> ) <sub>2</sub> SO <sub>4</sub>	1,8M (NH <sub>4</sub> ) <sub>2</sub> SO <sub>4</sub>
Tris pH 7,0	HEPES pH 7,5	HEPES pH 7,5	HEPES pH 7,5
10% Glycerol	15% Glycerol	5% Ethylene glycol	10% Ethylene glycol
0	0	0	0
10% DMSO	10% DMSO 0,25mM TCEP	10% DMSO	10% DMSO
			

**Table 1 - Crystallisation conditions used for fragment-screening experiments.** These crystals conditions are close ones from the others but they could still diffract differently.

These were used for the fragment-screening experiments. MtDS crystals generally appear within 12 h for an average and reach their maximum size of 30x30x180 μm<sup>3</sup> (some can be longer in length than that) between 5-7 days.

### Cryoprotectant concentration tests on the cryo stream

After finding conditions giving crystals, cryo tests were performed. The aim is to find out which additives can prevent ice formation when freezing the crystals. When the inside of the loop looks icy, chances are high that ice crystals formed and may interfere with the diffraction of the protein crystal. The results were noted as « icy » or « glassy » as summed in table 2.

Type of cryoprotectant	% in solution	Test 1	Test 2	Test 3	Exposures before icy
MPD	5	Icy	Icy	Icy	1
	10	Glassy	Glassy	Glassy	4
	15	Glassy	Glassy	Glassy	5
Ethylene glycol	5	Glassy	Glassy	Glassy	2
	10	Glassy	Glassy	Glassy	7
	15	Glassy	Glassy	Glassy	8
Glycerol	5	Icy	Icy	Icy	1
	10	Glassy	Glassy	Glassy	5
	15	Glassy	Glassy	Glassy	5
PEG 400	5	Icy	Icy	Icy	1
	10	Glassy	Glassy	Glassy	3
	15	Glassy	Glassy	Glassy	9

**Table 2 - Cryoprotectant tests.** The « icy » or « glassy » aspect was checked by simple observation of the cryoloop frozen by the nitrogen stream, on the camera.

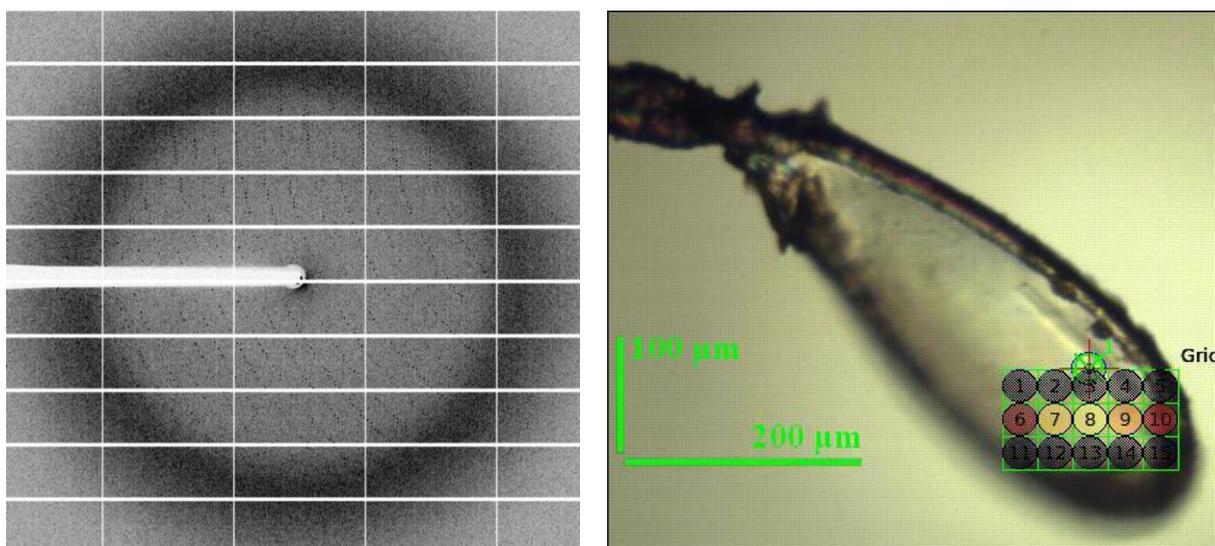
The results from these tests (table 2) indicate that 5% glycerol, that we were quite often using for crystallising until then, was not acting as a cryoprotectant. On the other side, ethylene glycol was the only additive being cryo at a concentration as low as 5%. Plates were set up from these results to try and see if crystals could be grown under these conditions, and some were obtained (see table 1, conditions B1 and C1 from 22.03 plate).

### Fragment screening

To be sure that the crystals would not dissolve in the compound mixtures, preliminary tests were run with 1µL of reservoir solution +0,5 µL of 30% DMSO in water, knowing that all compound mixtures contain 30% DMSO. In this mixture, the crystals dissolved. Then 1,5 µL reservoir + 0,5 30% DMSO was tried, and salt crystals were forming when trying to transfer the crystals. Finally, 1 µL reservoir + 0,5 µL water + 0,5 µL 30% DMSO was tried. Salt crystals were not forming and MtDS crystals were not dissolving, thus this was the best way of setting up the soaking drop.

### X-ray analysis

On the 12 crystals sent to the synchrotron, most of the crystals showed little to no diffraction, and in 2 of the cryoloops, salt crystals had also formed, showing only salt diffraction, but fortunately, 2 of them diffracted to about 3 Å resolution. This data was collected and processed by my co-supervisor Helen Thorbjørnsrud.



**Figure 20 - Diffraction pattern and crystal centring.** The black spots shows the diffraction and the quality of the diffraction.

The dataset statistics were then calculated with XDS (table 3).

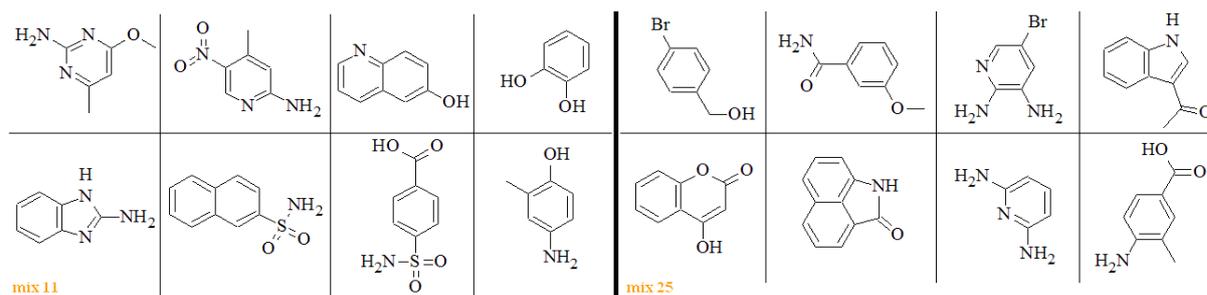
Resolution shells	Completeness of data	R-Factor observed	R-Factor expected	I/ $\sigma$	R-meas	CC <sub>1/2</sub>
7.96	93.4%	7.4%	6.9%	20.91	8.1%	99.4*
5.66	96.3%	13.5%	13.2%	11.90	14.8%	98.3*
4.64	96.6%	15.2%	14.7%	10.90	16.7%	97.7*
4.02	97.7%	18.1%	17.7%	9.58	19.9%	97.0*
3.60	97.8%	29.9%	28.3%	6.11	32.9%	92.7*
3.28	98.2%	55.5%	53.3%	3.35	60.6%	84.2*
3.04	98.8%	96.7%	98.1%	1.85	106.1%	60.5*
2.85	98.3%	160.5%	169.5%	1.11	175.2%	31.6*
2.68	92.4%	223.3%	243.8%	0.62	249.0%	17.9*
total	96.8%	32.0%	32.0%	5.31	35.2%	97.1*

**Table 3 – Statistics on the mix 11 dataset.** The resolution determines the amount of data that is added, others determine if this added data is of good or poor quality. The R-Factor gives the difference between the crystallographic model and the experimental data. The I/ $\sigma$  is the ratio between the signal and the background noise. The R-meas represents the uncertainty. The CC<sub>1/2</sub> is the correlation coefficient in percentage, comparing one half of the dataset to the other half (Diederichs and Karplus, 2013). It is considered as the most relevant statistic nowadays.

Resolution shells	Completeness of data	R-Factor observed	R-Factor expected	I/ $\sigma$	R-meas	CC <sub>1/2</sub>
8.57	97.1%	4.3%	4.5%	19.73	5.2%	99.5*
6.11	97.6%	8.6%	9.1%	10.41	10.5%	98.4*
5.00	96.7%	14.0%	14.6%	7.03	17.2%	94.9*
4.34	98.0%	14.9%	14.4%	7.31	18.3%	94.6*
3.88	96.0%	20.8%	20.3%	5.33	25.6%	91.1*
3.54	98.4%	35.1%	34.0%	3.52	43.1%	81.4*
3.28	96.4%	65.6%	64.6%	1.86	80.5%	58.7*
3.07	98.4%	102.4%	106.7%	1.14	125.3%	39.2*
2.90	96.9%	159.8%	174.6%	0.67	196.1%	21.2*
total	97.3%	26.1%	26.5%	4.53	32.0%	94.6*

**Table 4 – Statistics on the mix 25 dataset.** See table 3 for explanations on the different statistics.

In these two data collections were tested two mix of compounds represented below (Fig. 21).



**Figure 21 - Compounds tested by soaking on which x-ray data were collected.** On the left, all molecules composing the mixture 11, and on the right side, all the molecules composing the mixture 25.

At this point, the electron density map was not generated yet, thus it is not possible to say if something is bound on the protein or not.

## DISCUSSION

---

Different protocols (Munack et al., 2016) were assessed for MtDS production, the one used was indicated as the one with the best reproducibility. Taking into account the relatively short amount of time of this internship, alternative protocols were not tested. Otherwise, the ideal would be to run the three protocols in parallel and then determine which one is the best according to the yield.

After the nickel-affinity purification, the gel (Fig. 10) revealed other bands than the MtDS one, one with a higher molecular weight one and two smaller ones. These bands could be proteins containing histidine residues. They are supposed to be of a much lower affinity to the nickel matrix and the 20mM imidazole in the His Buffer A should be enough to eliminate these. For the lower mass bands, they could also result from MtDS degradation. In order to remove these bands, doing a slower gradient between the two His Buffers may be tried. Adding NaCl to remove the weak electrostatic interactions between non-specific proteins and the column is a common technique. Using a cobalt resin instead of nickel could also improve the results.

The maximum of the protein was kept (see for example Fig 9). All fractions showing the band were combined together, even though impurities could be seen. Yet, the yields are low in general for MtDS and milligrams may be needed to do all the experiments, considering that around 50µg of the protein was used for each 24-well crystallisation plate that were made with the conditions presented.

The two expressions of MtDS resulted in the yields of 120 µg/L (1<sup>st</sup> expression) and 12,5 µg/L (2<sup>nd</sup> expression) of culture. The fact that the yields were so low in a general way could be partly attributed to the glycerol stock. It was made in 2011 and the quality of the stock could have been altered within these 5 years. KA12 should be transformed again with pKTDS-HN and plated in order to verify that statement. 5mL cultures of LB medium supplemented by 150 µg/mL Amp would be inoculated with one colony from the plate.

The yield of 12,5 µg per litre of culture given by the large-scale purification of MtDS is not coherent with the chromatograms that are presented in the results part. Compared to chromatograms of the 1<sup>st</sup> purification using the same protocol (data not shown), they display higher quantities of the protein. Different hypotheses were proposed to explain it. First, it is possible that the protein was in the flow-through after the concentration. A7 and A13 fractions, the concentrated MtDS and the flow-through were tested on SDS-PAGE since then (data not

shown). The flow-through did not contain protein, thus this first hypothesis is wrong. Interestingly, A7 and A13 fractions do not appear to contain MtDS anymore. In this case, it is probable that the protein aggregated during the process and so is lost.

For CgDS, unlike MtDS, an anion-exchange purification step was added. Previous purifications of CgDS had shown that having an anion-exchange purification in between the nickel-affinity and the size-exclusion one was improving the result. The same judgement can be done here; it appears that most impurities were removed by this step (see the difference between Fig. 15 and Fig. 16). For MtDS, considering the loss of protein during the purification process, it was not attempted. The nickel-column and size-exclusion columns are usually sufficient to obtain a rather pure solution for MtDS to crystallise. For both CgDS and MtDS, the size exclusion was performed using a Superdex 75 column instead of the Superdex 200 one. The separation with impurities could have been better on the larger column as the flow rate is slower, but it would have been more time-consuming.

When MtDS is compared to CgDS, from the nickel-column, CgDS peak reached 2600 mAU at UV 280 nm while MtDS peak attained 550 mAU. For the size-exclusion, CgDS peak was still at 2500 mAU and MtDS was then at 250 mAU meaning 10 times lower. From this comparison it is easily shown that CgDS is much easier to purify. The SDS-PAGE also showed much more intense bands for CgDS than for MtDS. As for the yield, it is difficult to compare them as the MtDS yield is not what was expected.

The thermofluor assay would have given a preliminary idea of which molecules can bind to the protein, it would have been an addition to the x-ray data. However, the data collected from this experiment (data not shown) tends to indicate that some of the library compounds were already fluorescent or interacting with the dye in the mixtures before starting measurements. The fluorescence curves cannot be compared. This type of assay is thus not well suited for this library of fragments.

The crystallisation of MtDS did not show one optimal set of conditions (Table 1) but different ones resulting in crystals suitable for X-ray collection. These crystals were similar in length to the ones from Munack S. (personal communication). Some conditions with 5 to 10% ethylene glycol in lower ammonium sulfate (1,6M-1,7M), with microseeding, with or without TCEP were not done and should be tried. Due to the crystal seeds, that could allow the use of less protein. As they offer a nucleation site, it could possibly give bigger crystals, and according to the cryoprotectant tests, the cryoprotection would be better. Then the chances of having ice

crystals into the cryoloop when doing flash-freezing would be reduced. Also with lower salt (ammonium sulfate) concentration, the chance of having salt crystals is also diminished.

The activity of MtDS was not verified after the purifications, but it could also be verified by an enzymatic assay. The enzyme needs to be in an active state for both the thermofluor experiments and the fragment-based screening in order to give interesting results. Knowing that a ligand is bound to the protein is one important step, but after, the crucial part is to look if this ligand inhibits the protein. In that sense, if there is a hit, the protein with its ligand should also be tested by enzymatic assay.

When the X-ray data is analysed by XDS in the first steps of data processing, statistics on the quality of the datasets are generated. By looking at these statistics, for the 1<sup>st</sup> dataset (soaking with mix 11), including all data down to 2.68 Å is not optimal: The  $I/\sigma$  at this resolution is below 1. That means that, when these data are kept, there is more background noise than significant signal. For the 2.85 Å resolution cut-off, the  $CC_{1/2}$  is higher than 30% which is preferable. The completeness of the data is higher than 95% and the R-factor drops from 223% to 160% between 2.68 Å and 2.85 Å. Thus the resolution can be cut-off to 2.85 Å. The reasoning is similar for the second dataset (soaking with mix 25). The 3,07 Å resolution,  $I/\sigma$  is higher than 1, the completeness is around 98%,  $CC_{1/2}$  is at 39,2%, thus the resolution should be cut at 3,07 Å rather than 2,9 Å for which statistics show poor quality.

After this statistic analyses, the phase problem should be solved. In that case, as other MtDS structures have been solved, the MR method would be used. The phases must be found by comparing an existing model to the data collected. Then the images of the chosen resolution should be processed with another software (CCP4) in order to generate the electron-density map thanks to a Fourier transform. A model of the protein should be built into this map. Then the refinement should be done by making the model fit best into the density map. When this step is complete, the final model will be studied in order to know whether a compound is bound or not.

To conclude, during this internship, CgDS and MtDS were expressed and purified, crystallisation conditions for MtDS crystals tolerant to DMSO were found, and fragment-screening through soaking has been performed and analysed by X-ray crystallography. This data still need to be processed.

## REFERENCES

---

- Adachi, K., Katsuta, A., Matsuda, S., Peng, X., Misawa, N., Shizuri, Y., Kroppenstedt, R.M., Yokota, A., and Kasai, H. (2007). *Smaragdicoccus niigatensis* gen. nov., sp. nov., a novel member of the suborder Corynebacterineae. *Int. J. Syst. Evol. Microbiol.* *57*, 297–301.
- Blackmore, N.J., Reichau, S., Jiao, W., Hutton, R.D., Baker, E.N., Jameson, G.B., and Parker, E.J. (2013). Three Sites and You Are Out: Ternary Synergistic Allosteric Controls Aromatic Amino Acid Biosynthesis in *Mycobacterium tuberculosis*. *J. Mol. Biol.* *425*, 1582–1592.
- Bruchfeld, J., Correia-Neves, M., and Källenius, G. (2015). Tuberculosis and HIV Coinfection. *Cold Spring Harb. Perspect. Med.* *5*, a017871.
- Davies, T.G., and Tickle, I.J. (2012). Fragment screening using X-ray crystallography. *Top. Curr. Chem.* *317*, 33–59.
- Diederichs, K., and Karplus, P.A. (2013). Better models by discarding data? *Acta Crystallogr. D Biol. Crystallogr.* *69*, 1215–1222.
- Geldmacher, C., Ngwenyama, N., Schuetz, A., Petrovas, C., Reither, K., Heeregrave, E.J., Casazza, J.P., Ambrozak, D.R., Louder, M., Ampofo, W., et al. (2010). Preferential infection and depletion of *Mycobacterium tuberculosis*-specific CD4 T cells after HIV-1 infection. *J. Exp. Med.* *207*, 2869–2881.
- Horsburgh, C.R., Barry, C.E., and Lange, C. (2015). Treatment of Tuberculosis. *N. Engl. J. Med.* *373*, 2149–2160.
- Ikeda, M. (2006). Towards bacterial strains overproducing L-tryptophan and other aromatics by metabolic engineering. *Appl. Microbiol. Biotechnol.* *69*, 615–626.
- Jiao, W., Hutton, R.D., Cross, P.J., Jameson, G.B., and Parker, E.J. (2012). Dynamic Cross-Talk among Remote Binding Sites: The Molecular Basis for Unusual Synergistic Allostery. *J. Mol. Biol.* *415*, 716–726.
- Künzler, D.E., Sasso, S., Gamper, M., Hilvert, D., and Kast, P. (2005). Mechanistic Insights into the Isochorismate Pyruvate Lyase Activity of the Catalytically Promiscuous PchB from Combinatorial Mutagenesis and Selection. *J. Biol. Chem.* *280*, 32827–32834.
- Lv, X., Tang, S., Xia, Y., Wang, X., Yuan, Y., Hu, D., Liu, F., Wu, S., Zhang, Y., Yang, Z., et al. (2013). Adverse reactions due to directly observed treatment strategy therapy in Chinese tuberculosis patients: a prospective study. *PLoS One* *8*, e65037.
- Matthews, K., Ntsekhe, M., Syed, F., Scriba, T., Russell, J., Tibazarwa, K., Deffur, A., Hanekom, W., Mayosi, B.M., Wilkinson, R.J., et al. (2012). HIV-1 infection alters CD4+ memory T-cell phenotype at the site of disease in extrapulmonary tuberculosis. *Eur. J. Immunol.* *42*, 147–157.
- Montales, M.T., Chaudhury, A., Beebe, A., Patil, S., and Patil, N. (2015). HIV-Associated TB Syndemic: A Growing Clinical Challenge Worldwide. *Front. Public Health* *3*, 281.

Munack, S., Roderer, K., Ökvist, M., Kamarauskaite, J., Sasso, S., van Eerde, A., Kast, P., and Kregel, U. (2016). Remote Control by Inter-Enzyme Allostery: A Novel Paradigm for Regulation of the Shikimate Pathway. *J. Mol. Biol.* 428, 1237–1255.

Ormerod, L.P. (2005). Multidrug-resistant tuberculosis (MDR-TB): epidemiology, prevention and treatment. *Br. Med. Bull.* 73-74, 17–24.

Parish, T., and Stoker, N.G. (2002). The common aromatic amino acid biosynthesis pathway is essential in *Mycobacterium tuberculosis*. *Microbiology* 148, 3069–3077.

Prakash, P., Aruna, B., Sardesai, A.A., and Hasnain, S.E. (2005). Purified Recombinant Hypothetical Protein Coded by Open Reading Frame Rv1885c of *Mycobacterium tuberculosis* Exhibits a Monofunctional AroQ Class of Periplasmic Chorismate Mutase Activity. *J. Biol. Chem.* 280, 19641–19648.

Sasso, S., Ökvist, M., Roderer, K., Gamper, M., Codoni, G., Kregel, U., and Kast, P. (2009). Structure and function of a complex between chorismate mutase and DAHP synthase: efficiency boost for the junior partner. *EMBO J.* 28, 2128–2142.

Saukkonen, J.J., Cohn, D.L., Jasmer, R.M., Schenker, S., Jereb, J.A., Nolan, C.M., Peloquin, C.A., Gordin, F.M., Nunes, D., Strader, D.B., et al. (2006). An Official ATS Statement: Hepatotoxicity of Antituberculosis Therapy. *Am. J. Respir. Crit. Care Med.* 174, 935–952.

Webby, C.J., Jiao, W., Hutton, R.D., Blackmore, N.J., Baker, H.M., Baker, E.N., Jameson, G.B., and Parker, E.J. (2010). Synergistic allostery, a sophisticated regulatory network for the control of aromatic amino acid biosynthesis in *Mycobacterium tuberculosis*. *J. Biol. Chem.* 285, 30567–30576.

World Health Organization (2015). Global Tuberculosis Report 2015.



Diplôme double : Master et Ingénieur  
Spécialité : Agronome  
Spécialisation / option : SCMV  
Enseignant référent : Jean-Marc Fraslin

Auteur : Albat Emanuelle Date de naissance* : 14 octobre 1992	Organisme d'accueil : Oslo i Universitetet Adresse : Sam Sælands vei 26 0371 Oslo, Norvège
Nb pages : <b>26</b> Annexe(s) : Non	Maître de stage : Ute Krengel
Année de soutenance : 2016	

Titre français : Concevoir un médicament contre la tuberculose en utilisant la cristallographie :  
Un criblage basé sur des fragments ciblant MtDS

Titre anglais: Designing a drug against tuberculosis using crystallography:  
Fragment-based screening targeting MtDS

Résumé (1600 caractères maximum) : La tuberculose est la plus mortelle des maladies infectieuses avec le VIH de nos jours. Dans la plupart des cas il s'agit d'une maladie pulmonaire, causée par *Mycobacterium tuberculosis*. Les traitements actuels sont longs et ont des effets secondaires hépatotoxiques. De plus, aucun n'est réellement efficace contre les souches résistantes aux médicaments, c'est pourquoi il y a un besoin urgent de trouver de nouveaux médicaments. La voie du Shikimate est une voie de production des acides aminés aromatiques chez les bactéries qui est absente chez l'homme. MtDS, une enzyme clef de la voie Shikimate, est alors une cible potentielle pour un médicament. Dans ce projet, le but est de trouver un médicament qui cible MtDS en utilisant un criblage basé sur des fragments analysé par cristallographie aux rayons X. MtDS a été exprimée en *E.coli* et purifiée en même temps que son homologue CgDS de *Corynebacterium glutamicum*. CgDS est une protéine très similaire à MtDS qui peut aider à la compréhension de cette dernière, et est plus facile à manipuler de manière générale. Les conditions de cristallisation de MtDS ont été optimisées afin d'obtenir des cristaux qui diffractent. Ceux-là ont été trempés dans des mélanges de composés provenant d'une bibliothèque de fragments ayant des propriétés pharmacologiques, et analysés par cristallographie aux rayons X. Deux jeux de données ont réussi à être collectés à une résolution de 3 Å, et un CC<sub>1/2</sub> supérieur à 30%. Plus d'analyses devront être conduites afin de déterminer si des composants se sont liés à la protéine ou pas.

Abstract (1600 caractères maximum): Tuberculosis is the deadliest infectious disease worldwide alongside HIV nowadays. It is in most cases a lung disease, caused by *Mycobacterium tuberculosis*. Actual treatments are long and have hepatotoxic side effects. Moreover, none is really effective against the drug-resistant strains, this is why there is an urgent need of new medicine. The Shikimate pathway is a way of producing aromatic amino-acids in bacteria that is absent in humans. MtDS, a key enzyme of the Shikimate pathway, is then a potential drug-target. In this project, the aim was to find a drug targeting MtDS by using a fragment-based screening analysed by X-ray crystallography. MtDS was expressed in *E.coli* and purified along with its homologous protein CgDS from *Corynebacterium glutamicum*. CgDS is a protein very similar to MtDS that could help its understanding and is easier to handle in a general way. The conditions for MtDS crystallisation were optimized in order to get diffracting crystals. These were soaked for 24h in compounds mixes of drug-like fragments from a library, and analysed by X-ray crystallography. Two datasets were successfully collected at a 3 Å resolution, with a CC<sub>1/2</sub> superior to 30%. Further analyses need to be conducted in order to determine whether there are compounds bound to the protein or not.

Mots-clés : Cristallographie, *Mycobacterium tuberculosis*, criblage basé sur des fragments, médicament  
Key Words: Crystallography, *Mycobacterium tuberculosis*, fragment-based screening, drug discovery