



HAL
open science

Détection et signalisation du monoxyde de carbone chez des bactéries aérobies - Hémotenseur RcoM-2 et réponses mycobactériennes au CO

Mayla Salman

► **To cite this version:**

Mayla Salman. Détection et signalisation du monoxyde de carbone chez des bactéries aérobies - Hémotenseur RcoM-2 et réponses mycobactériennes au CO. Biologie moléculaire. Université Paris Saclay (COmUE), 2019. Français. NNT : 2019SACLX076 . tel-02484860

HAL Id: tel-02484860

<https://theses.hal.science/tel-02484860>

Submitted on 19 Feb 2020

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.

**Détection et signalisation du monoxyde de carbone
chez des bactéries aérobies - HémO-senseur RcoM-2
et réponses mycobactériennes au CO**

**Thèse de doctorat de l'Université Paris-Saclay
Préparée à l'École Polytechnique**

École doctorale n°573 : **Interfaces**

Interfaces : Approches interdisciplinaires : fondements, applications et innovation

Spécialité : **Biologie**

Par **Mayla SALMAN**

Thèse présentée et soutenue à Palaiseau, le 04 Novembre 2019

Composition du Jury :

M. Gilles TRUAN	INSA, Toulouse	Président
Mme Roberta FORESTI	Université Paris-Est Créteil	Rapportrice
Mme Delphine LECHARDEUR	INRA, Jouy-en-Josas	Rapportrice
Mme URSULA LIEBL	École polytechnique	Directrice de thèse
M. Marten VOS	École polytechnique	Co-directeur de thèse

Remerciements

D'abord j'aimerais remercier François Hache, le directeur du Laboratoire d'Optique et Biosciences (LOB) pour m'avoir accueillie dans son laboratoire et m'avoir donnée l'opportunité d'effectuer ce travail de thèse.

Je tiens à remercier Ursula Liebl, ma directrice de thèse, pour m'avoir aidée à améliorer mes compétences et l'autonomie dans mon travail, et aussi pour toutes les discussions enrichissantes professionnelles et humaines, sa patience, et pour la bonne humeur.

Je remercie Marten Vos, pour m'avoir aidée dans toutes les expériences de spectroscopie et aussi pour son aide dans l'apprentissage de toutes les bases, des principes et des compétences dans le domaine de la biophysique, et pour toutes ses explications.

Je souhaite exprimer mes remerciements aux rapporteurs, Madame Roberta Foresti et Madame Delphine Lechardeur, d'avoir accepté de lire et évaluer cette thèse, ainsi qu'envers Monsieur Gilles Truan pour avoir accepté de participer au jury de soutenance de ma thèse.

Je voudrais remercier Hannu Myllykalio, le directeur de l'équipe « Mécanismes moléculaires de l'adaptation microbienne » pour toutes les discussions, conseils et aides, et je tiens également à remercier Hubert et Stéphane pour leur aide et nos discussions dans plusieurs expériences. Également je remercie tous les membres de l'équipe, Roxane, Yoann, Caroline et Isabelle.

Je remercie Jean-Christophe Lambry pour son aide dans les expériences de modélisation moléculaire et Rivo Ramodiharilafy pour toute son aide dans certaines expériences et dans la résolution des problèmes techniques, ainsi que le partage du bureau et de la paillasse pour plus de quatre ans, et Carolina Villamil Franco pour son aide précieuse aux différentes expériences de spectroscopie. Je remercie tous les membres du LOB pour leur bonne humeur et compagnie, vous étiez vraiment une famille.

Aussi je tiens à remercier Laure Lachapelle et Christelle Français pour tout leur patience et aide pour les démarches administratives.

Et bien sûr je remercie les jeunes du LOB (surtout celles et ceux qui ont travaillé les weekends avec moi) pour tous les bons moments et les soirées avec vous.

Aussi je voudrais remercier mes cousins à Paris, ma tante Karine Fahs, son mari Alain Fahs, Sarah, Dylan et Rami, vous étiez toujours là à côté de moi dans tous les bons et mauvais moments et si je ne me suis jamais sentie seule à l'étranger, c'est grâce à vous. Un merci à vous, pour les soirées et tout le support tout le temps.

Également je remercie tous les amis avec qui je passais des moments inoubliables et des bons weekends.

Finalement, un grand merci à ma famille au Liban, surtout Papa, Maman, mes frères Ali et Mohamad, mes sœurs Rim et Nermine, pour tous vos supports et encouragements depuis mon départ en France afin de continuer mes études supérieures, malgré la distance, vous étiez toujours à côté de moi et vous n'avez jamais cessé de me supporter.

À Papa et Maman, j'espère que vous êtes fiers de moi et de mon travail comme je le suis aujourd'hui !

À toutes les déceptions qui nous ont fait...

À nous comme nous résistons à la laideur de ce monde,

d'où vous ne le savez pas...

***إلى الخيبات التي صنعنا...
إلينا ونحن نقاوم قبح هذا العالم
من حيث لا تعلمون...***

Résumé

Titre : Détection et signalisation du monoxyde de carbone chez les bactéries aérobies : Hémosenseur RcoM-2 et réponses mycobactériennes au CO.

Mots clés : Monoxyde de carbone (CO), hémosenseur à CO RcoM-2, *Mycobacterium tuberculosis* (*Mtb*), gène *cor* (*rv1829*), interaction hème-CO, régulation transcriptionnelle.

Le gaz toxique CO peut agir à faibles quantités comme molécule de signalisation ; détectée par les protéines senseurs à base d'hème.

La première partie de ce travail concerne le régulateur de transcription bactérien CO-dépendant RcoM-2, qui possède une affinité extrêmement élevée pour le CO, tout en étant insensible à l'O₂. RcoM-2 se lie à son ADN cible seulement lorsque du CO est lié à l'hème. Nous avons caractérisé l'interaction hème-CO dans la protéine entière RcoM-2 et l'avons comparée à l'hémo-domaine isolé RcoMH-2. RcoM-2 peut lier du CO avec une affinité effective plus faible que RcoMH-2. Le CO se dissocie avec un taux 20 fois plus élevé par rapport à RcoMH-2, où la liaison du CO est quasi irréversible. Une petite fraction du CO peut échapper de la protéine, ainsi permettant RcoM-2 à agir comme senseur du CO. La présence du domaine de liaison à l'ADN influence les propriétés de fixation du CO à l'hème. Les études sur l'origine moléculaire précise des propriétés dynamiques doivent attendre la structure 3D de RcoM-2.

La détection du CO est cruciale pour *Mycobacterium tuberculosis*, l'agent infectieux de la tuberculose, qui doit répondre aux mécanismes de défense de l'hôte, dont le CO. Le gène *cor* (*rv1829*) pourrait y jouer un rôle. Nous avons démontré que Cor est un dimère très stable, capable de lier un hème de façon stoechiométrique, suggérant ainsi une fonction potentielle de senseur *direct* du CO. Un résidu histidine a été identifié comme ligand potentiel de l'hème. La dynamique interne du CO est très similaire de celle d'autres senseurs à CO bactériens. Cor présente une activité de liaison à l'ADN dépendante de l'hème et du CO, qui est abolie dans le mutant H70A. Nos études ont également montré que le régulateur de transcription mycobactérien Rv0081, induit en réponse aux changements gazeux, peut se lier sur la région régulatrice prédite de *cor*. L'inactivation de *cor* dans le modèle non-pathogène *M. smegmatis* se poursuit et constituera un premier pas vers des futures analyses transcriptomiques.

Abstract

Title: Detection and signaling of carbon monoxide in aerobic bacteria: Heme-based sensor RcoM-2 and mycobacterial CO responses.

Keywords: Carbon monoxide (CO), CO sensor RcoM-2, *Mycobacterium tuberculosis* (*Mtb*), gene *cor* (*rv1829*), heme-CO interaction, transcriptional regulation.

The toxic gas CO can act in low quantities as signaling molecule, detected by heme-based sensor proteins.

The CO-dependent transcription factor RcoM-2 has a very high affinity for CO, while being insensitive to O₂. RcoM-2 binds to DNA only when CO is bound to heme. We characterized the heme-CO interaction in full-length RcoM-2 and compared it with the isolated heme domain RcoMH-2. RcoM-2 can bind CO with lower *effective* affinity than RcoMH-2. CO dissociates with a 20-fold higher rate than in RcoMH-2, where CO binding is almost irreversible. A small fraction of CO can escape from the protein, thus allowing RcoM-2 to act as CO sensor. The presence of the DNA binding domain influences the binding properties of CO to heme. Identification of the precise molecular origin of the dynamic properties must await the 3D structure of RcoM-2.

CO detection is crucial for *Mtb*, the infectious agent of tuberculosis that must overcome the host's defense mechanisms, including CO. The gene *cor* (*rv1829*) has been implicated in these processes. We have shown that Cor is a highly stable dimer that is able to stoichiometrically bind a heme cofactor, suggesting a potential function as *direct* CO sensor. A histidine residue was identified as potential heme ligand. The internal CO dynamics is very similar to other bacterial CO sensors. Cor exhibits DNA binding activity that depends on the presence of heme and CO, which is abolished in the H70A mutant. Our studies also showed that the transcriptional regulator Rv0081, induced in response to gaz changes, can bind to the predicted regulatory region of *cor*. The inactivation of *cor* strain in the non- pathogenic model *M. smegmatis* continues and will be a first step towards transcriptomic analyses.

Abréviations utilisées

A	Alanine
ADN	Acide désoxyribonucléique
Amp	Ampicilline
BET	Bromure d'éthidium
BSA	Sérum albumine bovine
CAP	Catabolite activator protein
CO	Monoxyde de carbone
CO-RM	CO-releasing molecule
DTT	Dithiothréitol
D.O.	Densité optique
EMSA	Electrophoretic mobility shift assay
ERO	Espèces réactives oxygénés
H	Histidine
HO	Hème oxygénase
IPTG	Isopropyl-1-thiol-beta-D-galactopyranoside
Kan	Kanamycine
<i>Mtb</i>	<i>Mycobacterium tuberculosis</i>
ns, ps, fs	nanoseconde, piconseconde, femtoseconde
NO	Monoxyde d'azote
O ₂	Dioxygène
PCB	Polychlorobiphényles
PCR	Polymerase Chain Reaction
PDB	Protein Data Base
SDS-PAGE	Sodium dodécyl Sulfate Polyacrylamide Gel Electrophoresis
UV-vis	Ultraviolet-visible
X-Gal	5-bromo-4-chloro-3-indolyl-beta-D-galactopyranoside

Tables des matières

I-Introduction.....	10
1. Détection des gaz diatomiques par les protéines senseurs à base d'hème.....	11
1.1. L'hème et le fer.....	11
1.2. Les gaz diatomiques.....	12
a) Le dioxygène (O ₂)	12
b) Le monoxyde d'azote (NO)	13
c) Le monoxyde de carbone (CO)	13
1.3. Les senseurs à gaz à base d'hème	15
1.3.1. Les hème-senseurs d'O ₂	17
1.3.2. Les hème-senseurs de NO	19
1.3.3. Les hème-senseurs à CO	21
a) Les senseurs à CO chez les eucaryotes	21
b) Les senseurs à CO chez les bactéries.....	22
b.1) CooA, senseur à CO anaérobie.....	22
b.2) RcoM, senseur à CO aérobie.....	25
1.4. Les mycobactéries.....	29
1.4.1. Généralités et caractéristiques.....	29
1.4.2. <i>Mycobacterium tuberculosis (Mtb)</i>	29
1.4.3. Mécanismes de défense de l'hôte.....	31
1.4.4. Mécanismes de détection des gaz chez <i>Mycobacterium tuberculosis</i>	33
a) Système de régulation DosRS-DosT.....	33
b) Senseurs kinases DosS/DosT.....	34
c) Le système MprA/B.....	36
d) Le régulateur de transcription Rv0081.....	37
e) Détection et réponse au CO dans <i>Mtb</i> ; autres mécanismes.....	38
e.1) Le gène <i>cor</i> (carbone monoxide resistance).....	39
e.2) Cluster de gènes impliquant <i>cor</i> (<i>rv1829</i>).....	41
e.3) Le régulateur de transcription Rv1828.....	42
1.5. Problématique et objectifs de la thèse.....	43
II-Mécanismes de détection et signalisation du CO : l'hème-senseur RcoM-2.....	45
2.1. Introduction	45
2.2. Caractérisation de RcoM-2.....	46
2.2.1. Clonage, expression et purification de RcoM-2.....	46
2.2.2. Analyses de RcoM-2 par spectroscopie d'absorption à l'équilibre.....	47
2.2.3. Cinétique de liaison du ligand CO.....	48
2.2.4. Dissociation du ligand CO.....	50

2.2.5.	Dynamique de l'interaction hème-CO : recombinaison géminée du CO.....	52
2.2.6.	Liaison de RcoM-2 à l'ADN.....	54
III-Mécanismes de détection du CO chez les mycobactéries.....		57
3.1.	Introduction.....	58
3.2.	Caractérisation de la protéine Cor.....	59
3.2.1.	Caractérisation biochimique de Cor.....	59
	a) Surexpression inductible de Cor.....	59
	b) Cor est un dimère très stable.....	61
	c) Modélisation de la protéine Cor.....	62
3.2.2.	Caractérisation spectroscopique de Cor.....	64
	a) Liaison Cor-hème.....	64
	b) Cor fixe du CO et lie de l'hème de façon spécifique.....	66
	c) La présence du CO accélère la liaison de l'hème à Cor.....	69
	d) Dynamique du ligand diatomique CO.....	70
	e) Recombinaison du ligand hémique.....	72
3.2.3.	Études des mutants Cor-H58A et Cor-H70A.....	73
	a) Expression et purification de Cor-H58A et Cor-H70A.....	73
	b) Caractérisation spectroscopique des mutants.....	74
	b.1) Liaison d'hème dans les mutants en absence du CO.....	74
	b.2) Liaison d'hème en présence du CO.....	75
3.3.	Le cluster de gènes « cor »	77
3.3.1.	Alignements et analyses <i>in silico</i>	78
3.3.2.	Expression et purification des régulateurs de transcription.....	80
	a) Le régulateur de transcription Rv0081.....	81
	b) Le régulateur de transcription Rv1828.....	81
3.3.3.	Interaction protéines-ADN.....	82
	3.3.3.1) Spectroscopie d'anisotropie de fluorescence.....	82
	a) Principe de l'anisotropie de fluorescence	82
	b) Fixation de Rv0081 et Rv1828 à la région régulatrice prédite de <i>cor</i>	84
	c) Mesures de fixation de Rv1828 avec la région promotrice <i>rv0081</i>	84

d) Cor lie de l'ADN en présence d'hème et du CO.....	85
d.1) Etude de la liaison de Cor à la région ADN <i>rv1831/rv1832</i>	85
d.2) Etude des mutant de Cor-H58A et Cor-H70A.....	87
3.3.3.2) <u>E</u> lectrophoretic <u>m</u> obility <u>s</u> hift <u>a</u> ssay : EMSA.....	88
a) Principe de l'EMSA (« retard sur gel »)	88
b) EMSA avec Rv0081, Rv1828 et Cor	89
3.4. Cristallogénèse des protéines.....	91
3.5. Rôle physiologique de Cor.....	92
3.5.1. Croissance mycobactérienne en présence du CO.....	93
a) Souche sauvage de <i>M. smegmatis</i>	93
b) Création d'une souche de délétion de <i>cor</i>	95
IV-Conclusions et perspectives.....	99
V-Matériels et méthodes.....	104
5.1. Techniques de biologie moléculaire	104
5.1.1. Milieux de culture et croissance bactérienne.....	104
5.1.2. Cartes des plasmides et gènes utilisés.....	105
5.1.3. Souches des bactéries utilisées.....	107
5.1.4. Extraction de l'ADN génomique de <i>M. smegmatis</i>	107
5.1.5. Amplification et clonage du gène <i>cor (rv1829)</i>	108
5.1.6. Transformation bactérienne.....	110
5.1.7. Extraction, purification de l'ADN plasmidique et digestion enzymatique.....	110
5.1.8. Vérification des géotypages (constructions plasmidiques)	111
5.1.9. Mutagenèse dirigée.....	111
5.1.10. Listes des amorces utilisées.....	112
5.2. Manipulation des protéines et études biochimiques.....	114
5.2.1. Expression des protéines dans <i>Escherichia coli</i>	114
5.2.2. Purification des protéines et détermination de la concentration.....	114
5.2.3. Analyses des protéines purifiées par SDS-PAGE.....	115
5.2.4. Technique d'EMSA.....	115
5.3. Approches spectroscopiques.....	117
5.3.1. Spectroscopie d'absorption.....	117
5.3.2. Anisotropie de fluorescence.....	117
5.3.3. Absorption résolue en temps	118

5.4. Microscopie confocale.....	119
5.5. Ultracentrifugation analytique de la protéine Cor.....	120
5.6. Cristallisation des protéines.....	120
5.7. Modélisation moléculaire de Cor.....	121
VI-Références bibliographiques.....	122
VII-Annexes.....	131

Chapitre I

Introduction

De la bactérie à l'Homme, l'adaptation aux changements environnementaux est essentielle à la survie de tout organisme. Pour s'adapter, des bactéries ont développé des mécanismes spécifiques, capables de détecter et de répondre rapidement aux fluctuations environnementales et facteurs de stress, tels que les changements de lumière, de température, de gradients chimiques ou encore de concentrations de gaz diatomiques. Les organismes anaérobies ont besoin de senseurs qui répondent aux concentrations létales ou aux espèces réactives de l'oxygène (ERO), tandis que les bactéries aérobies (facultatifs) développent fréquemment des systèmes de détection et de réponse permettant de passer d'un état métabolique à un autre en fonction de la concentration et de la disponibilité en oxygène. Ces protéines senseurs répondent généralement en modifiant l'expression de gènes spécifiques conduisant à adapter le métabolisme aux nouvelles conditions, permettant ainsi à l'organisme de survivre.

Des gaz diatomiques comme le dioxygène (O_2), le monoxyde d'azote (NO) et le monoxyde de carbone (CO) peuvent être des molécules vitales, toxiques ou/et de signalisation, selon leurs concentrations. La capacité d'adaptation aux différentes concentrations gazeuses est particulièrement importante pour des bactéries pathogènes intracellulaires afin de se défendre contre les mécanismes de défense de l'hôte.

Afin de détecter et répondre rapidement aux variations de concentrations en O_2 , NO ou CO, de nombreuses bactéries ont développé des systèmes spécifiques de détection, parmi eux les protéines senseurs à base d'hème [1, 2]. Ces senseurs, décrits ci-dessous en plus de détail, constituent une famille de protéines régulatrices conduisant à l'adaptation bactérienne induit par la fixation des ligands gazeux à l'atome de fer de l'hème.

1) Détection des gaz diatomiques par les protéines senseurs à base d'hème

1.1. L'hème et le fer

Le fer est essentiel pour un grand nombre de processus biologiques vitaux. Il est un cofacteur indispensable dans des réactions redox, comme p. ex. dans les clusters Fe-S chez différentes métalloprotéines, et en tant que partie intégrante de beaucoup de systèmes enzymatiques. En tant que fer hémique, il sert de transporteur d'oxygène dans l'hémoglobine des globules rouges, au stockage de l'oxygène dans la myoglobine, au transfert d'électrons par les cytochromes et à l'activation catalytique dans des monooxygénases P450, ainsi que comme site de détection des gaz diatomiques dans des protéines senseurs à base d'hème. L'hème est également une source de fer essentielle pour les bactéries pathogènes dans l'environnement de l'hôte. Dans des conditions physiologiques, presque la totalité du fer des mammifères est liée à des protéines et se trouve dans le compartiment intracellulaire. Les bactéries pathogènes ont développé des mécanismes de compétition pour le fer de l'hôte lors de l'infection [3]. Elles utilisent des protéines capables de saisir l'hème des hémoprotéines. L'hème est ensuite transféré à l'intérieur de la cellule bactérienne, où il est incorporé directement dans les protéines ou dégradé pour libérer le fer [4].

L'hème est un cofacteur ou groupement prosthétique qui contient un atome de fer au centre d'un noyau porphyrine. L'hème ne se trouve pas à l'état libre, mais est toujours associé aux protéines par des acides aminés. L'état de coordination du fer (5- ou 6-coordonné) ou la liaison des ligands diatomiques (CO, NO, et O₂) vont permettre à la protéine de réaliser sa fonction. La forme la plus fréquente de l'hème est l'hème *b* (figure 1.1), présente entre autre dans la myoglobine et l'hémoglobine. Les deux autres formes principales de l'hème sont l'hème *a*, présent surtout dans l'enzyme cytochrome *c* oxydase *aa*₃ et l'hème *c* que l'on trouve dans la protéine de transfert d'électrons cytochrome *c*.

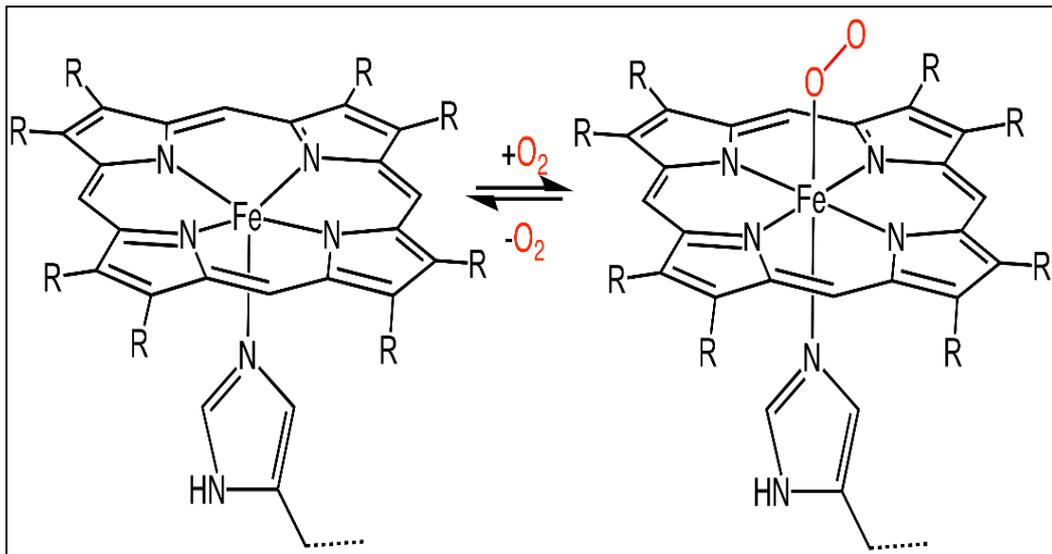


Figure 1.1 : Schéma de la fixation réversible de l'oxygène à l'hème. L'atome de fer (Fe^{2+}) se situe au centre du noyau porphyrinique et est lié par quatre atomes d'azote de la porphyrine. Une valence du fer est liée à un des azotes d'une histidine d'une protéine (proximale). La dernière valence libre du fer (distale) permet de fixer le ligand O_2 (comme dans le cas de l'hémoglobine et myoglobine) [5].

1.2. Les gaz diatomiques

a) Le dioxygène (O_2)

Le dioxygène est un gaz indispensable dans de nombreux processus d'oxydoréduction, surtout dans la respiration cellulaire chez les organismes aérobies. Cette réaction consiste à absorber le dioxygène nécessaire au catabolisme oxydatif et apporte l'énergie pour les cellules afin de fonctionner et se développer.

Le dioxygène peut également générer un stress oxydatif suite à une réduction enzymatique incomplète en produisant des espèces réactives oxygénées (ERO), qui provoquent des dommages à l'ADN, aux protéines et aux lipides, par ex. les ions superoxyde et peroxyde d'hydrogène qui grâce à leurs électrons non appariés possèdent une grande affinité pour l'hème. Ici nous remarquons l'importance de la détection et contrôle de la concentration d'oxygène dans les cellules chez tous les organismes.

b) Le monoxyde d'azote (NO)

Le monoxyde d'azote (ou l'oxyde nitrique) est une molécule diatomique relativement instable qui possède un radical libre qui la rend très réactive. Bien qu'il s'agisse d'un gaz toxique à fortes concentrations, le monoxyde d'azote est une molécule de signalisation importante chez les animaux qui transmet des signaux aux cellules des systèmes cardiovasculaire, nerveux et immunitaire. Sa petite taille lui permet de diffuser à travers les membranes et les parois cellulaires. Le monoxyde d'azote est synthétisé biologiquement suite à l'oxydation de l'acide aminé L-arginine en N-hydroxyarginine, puis en NO grâce à l'enzyme oxyde nitrique synthase (NOS) [6, 7]. Le rôle de l'oxyde nitrique dans la dilatation des vaisseaux sanguins en fait un régulateur important de la pression artérielle. L'oxyde nitrique est également utilisé par le système nerveux comme neurotransmetteur.

Dans le système immunitaire, l'oxyde nitrique est produit par les macrophages. L'oxyde nitrique libéré par les macrophages tue p.ex. des bactéries en perturbant leur métabolisme [6, 8]. L'excès de NO peut également contribuer à une inhibition de certaines enzymes comme la ribonucléotide réductase, requise pour la synthèse de l'ADN [6].

c) Le monoxyde de carbone (CO)

Le monoxyde de carbone est un gaz asphyxiant toxique, généré et libéré dans l'environnement par des activités humaines et des processus naturels (comme la combustion incomplète des composés organiques). Sa toxicité chez l'Homme est dû à sa fixation sur l'hémoglobine (Hb) avec une affinité d'environ 200 fois supérieure à celle de l'oxygène et la formation d'un complexe carboxyhémoglobine (HbCO) dont la libération est 10000 fois plus lente. Ce complexe n'est plus capable de transporter et libérer le dioxygène aux cellules et organes, conduisant à l'hypoxie (voir figure 1.2).

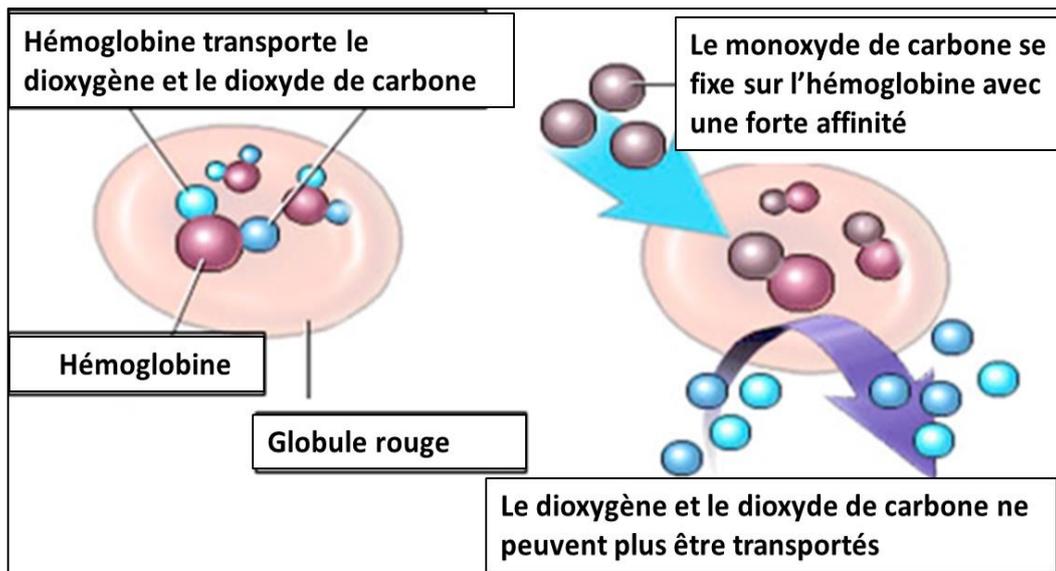


Figure 1.2 : La fixation irréversible du monoxyde de carbone sur l'hémoglobine conduisant à l'hypoxie (adapté de [9]).

Le CO peut être synthétisé également de manière endogène chez l'Homme, par l'hème-oxygénase (HO-1). Cette enzyme catalyse le clivage oxydant de l'hème pour produire des quantités équimolaires de biliverdine, de Fe^{2+} et du CO et fait partie des stratégies antimicrobiennes chez l'hôte [10] (voir sous-chapitre 1.4.3.).

Le CO est une molécule stable et difficile à convertir en d'autres espèces. Il est à noter que le CO n'interagit qu'avec les complexes métalliques, tels que les complexes de fer hémique ou non-hémique ou de fer-molybdène et non avec les résidus d'acides aminés des protéines. Ce comportement chimique du CO est différent de celui de l' O_2 et de NO, qui forment des espèces réactives dérivées telles que l' H_2O_2 ou le peroxy-nitrite, respectivement.

À faibles quantités, le CO peut agir comme molécule de signalisation, qui peut affecter la régulation de gènes [11, 12]. Cela nécessite sa détection rapide et des réponses cellulaires adaptées, ainsi que des mécanismes spécifiques pour concurrencer l'oxygène existant, qui sont peu compris à ce jour. Le monoxyde de carbone est également une molécule intermédiaire dans le métabolisme bactérien grâce à la métalloprotéine carbone monoxyde déshydrogénase (CODH), qui catalyse la réaction suivante :



Selon sa concentration, le CO peut inhiber la respiration aérobie chez *E. coli* en se fixant sur les cytochrome oxydases *bd-I* et *bd-II* et *bo* [13]. Le CO peut inhiber la croissance d'*E. coli* en altérant l'expression de certains gènes impliqués dans la glycolyse et le cycle de Krebs, incluant l'aconitase [12]. Une perturbation des gènes impliqués dans l'acquisition du fer par le CO, surtout le changement de l'activité du facteur de transcription Fur « ferric uptake regulator », a été observée, Fur contrôle l'expression des enzymes qui protègent contre les dommages dûs par des espèces réactives de l'oxygène. Le CO affecte également de façon significative les gènes codant pour des voies de transduction d'énergie par son action sur les régulateurs ArcA et FNR [12] et, contrairement au NO, qui protège les bactéries de certains antibiotiques [14], il n'y a pas de preuve que le CO possède cet effet [12].

1.3. Les senseurs de gaz à base d'hème

Les senseurs à base d'hème constituent une large famille d'hémo-protéines, capables de détecter des molécules de signalisation. Ces protéines jouent un rôle très important dans les processus d'adaptation des organismes (eucaryotes et procaryotes) aux changements gazeux, généralement en modifiant l'expression de certains gènes spécifiques lors de la liaison (ou dissociation) des molécules effectrices à l'hème.

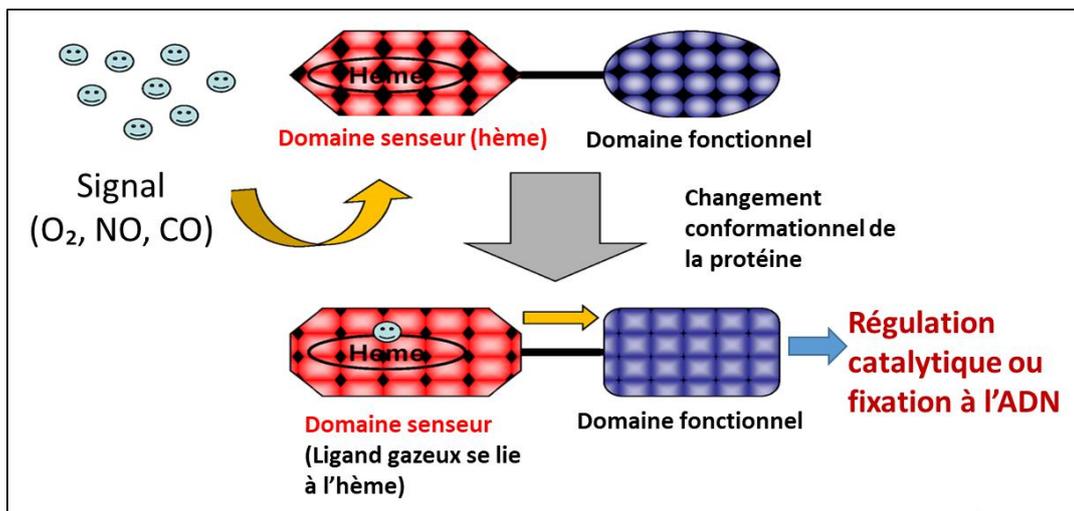


Figure 1.3 : Schéma d'un senseur à gaz à base d'hème. Le domaine senseur (en rouge) détecte le ligand gazeux par l'hème. Des changements structuraux dans l'environnement de l'hème déclenchent la traduction d'un signal intra-protéique vers le domaine fonctionnel (en bleu) qui peut être catalytique ou un domaine de fixation à l'ADN [15].

Les protéines senseurs à base d'hème contiennent typiquement un domaine de détection qui lie le cofacteur hémique. L'hème détecte les changements gazeux et contrôle ainsi un domaine fonctionnel : catalytique ou de fixation à l'ADN (voir figure 1.3).

Parmi les domaines de détection chez les hémoprotéines senseurs, les domaines PAS, structurellement très conservés, sont les plus abondants [15] (voir figure 1.5). Ils permettent la fixation de l'hème et ainsi la détection du ligand gazeux. La fixation du ligand déclenche une réponse adaptative physiologique aux changements de disponibilité de ligand. Les domaines fonctionnels peuvent être catalytiques ou des domaines de fixation à l'ADN ; un grand nombre des hémoprotéines senseurs à gaz étant des facteurs de transcription. Les plus connus des senseurs à gaz à base d'hème sont : le senseur d'O₂ FixL dans la bactérie *Rhizobium meliloti*, une protéine, histidine kinase qui contrôle la fixation de l'azote en réponse aux variations de concentrations d'oxygène [16] ; le senseur de CO, régulateur de transcription CooA dans la bactérie photosynthétique *Rhodospirillum rubrum*, qui régule l'expression des gènes pour l'oxydation du CO [17, 18] ; les senseurs de NO homologues de l'hémoprotéine de la guanylate cyclase soluble (GCs) eucaryote, récemment découverts chez certaines bactéries [19, 20].

<p style="text-align: center;">PAS Family</p> <p style="text-align: center;">Histidine Protein Kinase</p> <p>5 Rhizobial FixLs (bacterial) Nitrogen fixation and alternative oxidases M. thermoautotrophicum FixL (archaeal) C. crescentus FixL (bacterial) Plus 10 bacterial homologs</p> <p style="text-align: center;">Second Messenger</p> <p>A. xylinum PDEA1 (bacterial) Cellulose production E. coli, A. caulinodans (2 bacterial) Plus 12 bacterial homologs</p> <p style="text-align: center;">DNA Binding</p> <p>Mammalian NPAS2 (eukaryotic) Circadian rhythm</p>	<p style="text-align: center;">GCS Family</p> <p style="text-align: center;">Methyl Carrier Protein</p> <p>H. salinarum (archaeal) and B. subtilis HemATs (bacterial) Aerotaxis Plus 13 bacterial homologs</p> <p style="text-align: center;">Second Messenger</p> <p>15 bacterial homologs</p>
<p style="text-align: center;">CooA</p> <p style="text-align: center;">DNA Binding</p> <p>R. rubrum (bacterial) CO metabolism</p>	<p style="text-align: center;">HNOB Family</p> <p style="text-align: center;">Second Messenger, Mononucleotide</p> <p>Mammalian sGC (eukaryotic) Vasodilation, etc.</p> <p style="text-align: center;">Second Messenger, Dinucleotide</p> <p>Rsph2043_Rhsp (bacterial)</p> <p style="text-align: center;">Methyl Carrier Protein</p> <p>Ddes2822Dde and CAC3243 (bacterial)</p> <p style="text-align: center;">Histidine Protein Kinase</p> <p>5 bacterial homologs</p>

Tableau 1.1 : Les quatre principales familles des senseurs à base d'hème (encadrement en noir), leurs sous-groupes (encadrement en rouge) et leurs rôles physiologiques si connus (en vert). GCS, FixL et PDEA1 : senseurs d'oxygène ; CooA et NPAS2 : senseurs de CO ; sGC : senseur de NO (d'après [2]).

Selon leurs domaines de détection, on peut diviser les hème-protéines senseurs en quatre familles principales : Hème-PAS, CoxA, globines (GCs) et Hème-NO binding (HNOB) [2] (voir Tableau 1.1). Leurs sous-groupes sont définis selon leurs domaines enzymatiques et leurs rôles physiologiques divers.

1.3.1. Les hème-senseurs d'O₂

Dans ces protéines senseurs, par exemple FixL, DosS, DosT (voir figure 1.21), les changements conformationnels du domaine senseur induits par la fixation ou dissociation du O₂ seront transmis à un domaine fonctionnel.

Le senseur d'O₂ bactérien le mieux caractérisé est l'hème-protéine rhizobienne FixL. En condition de diminution partielle en oxygène en valeur microaérobie, FixL induit une cascade de réactions cellulaires (figure 1.6) qui induisent l'expression des gènes *fixK* et *nifA* qui sont impliqués dans la fixation de l'azote moléculaire chez *Rhizobium meliloti* [21–23]. La figure 1.4 montre un schéma de FixL.

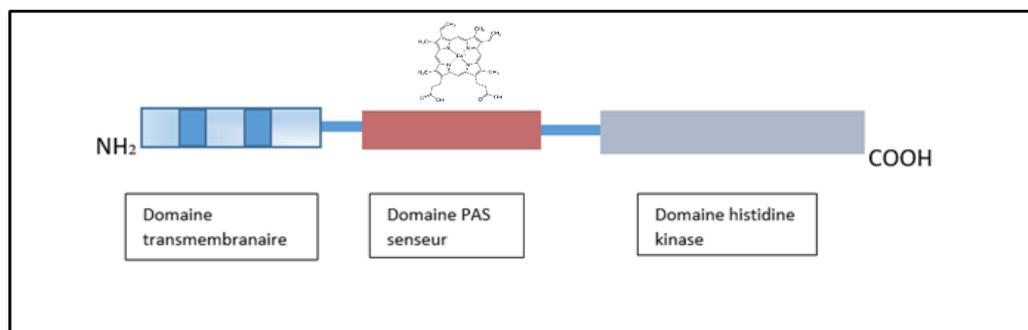


Figure 1.4 : Représentation schématique de la protéine FixL de *Rhizobium meliloti*. L'extrémité N-terminale contient le domaine PAS liant l'hème ; l'extrémité C-terminale le domaine à activité histidine kinase.

FixL détecte le ligand oxygène par l'intermédiaire de l'hème dans un domaine de signalisation structurellement très conservée du type « PAS » (fig. 1.5) et des événements de signalisation intra-protéique régulent son domaine histidine kinase [22].

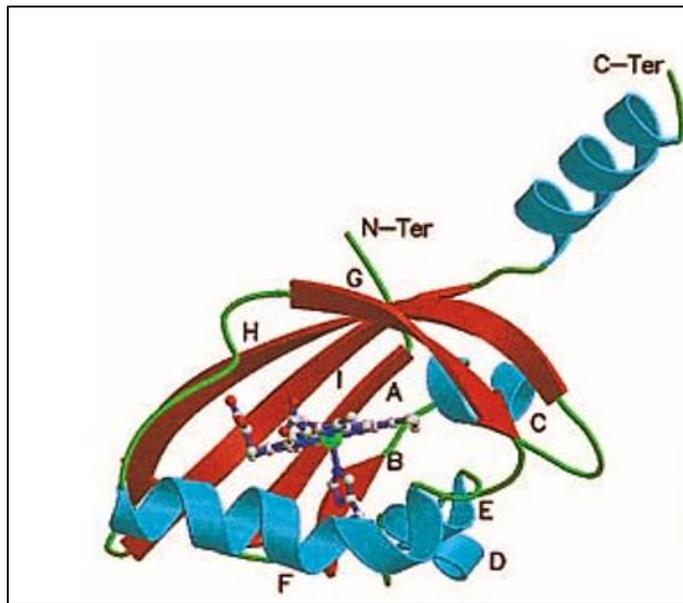


Figure 1.5 : Structure de l'hémo-domaine PAS du senseur à O₂ rhizobien FixL montrant la liaison de l'hème. L'atome de fer est en vert et les atomes oxygène en rouge. La structure tertiaire du domaine senseur est très conservée. L'hélice F est fortement impliquée dans l'accommodation de l'hème et les mécanismes de signalisation. Chaque lettre est associée à une structure (A, B, G, H, I : feuillet β ; C, D, E, F : hélice α), d'après [24].

La protéine FixL et son régulateur de réponse FixJ appartiennent à la famille des régulateurs à deux composants (« two component systems » ; TCS) [22], les systèmes les plus répandus de la signalisation bactérienne [22, 23]. Ce type de voie de signalisation est illustrée dans la figure 1.6 dans le cas de FixL.

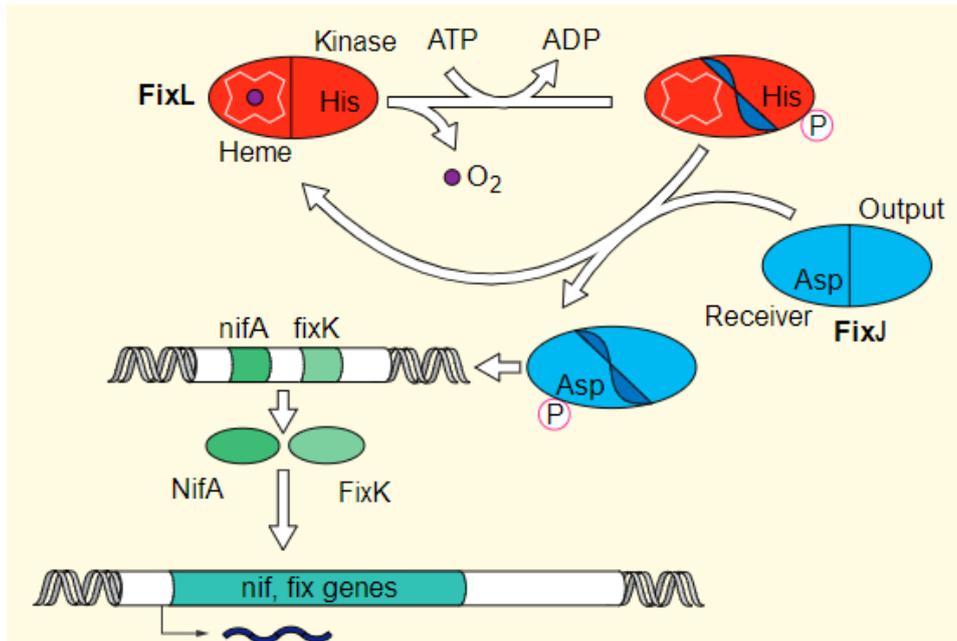


Figure 1.6 : Schéma récapitulatif de la signalisation et adaptation grâce au système à deux composants FixL/FixJ chez *Rhizobium* et *Bradyrhizobium*. La protéine FixL (rouge) possède un domaine senseur d'oxygène en N-terminal et un domaine histidine kinase en C-terminal. À un niveau faible en O_2 , l'hème de FixL relargage d' O_2 . Les changements de conformation ultérieurs dans l'hémo-domaine entraînent l'activation du domaine kinase de FixL par autophosphorylation. Le transfert de phosphate vers l'activateur transcriptionnel FixJ (bleu) déclenche l'activation des gènes *fixK* et *nifA* qui, à leur tour, induisent l'expression des gènes de fixation de l'azote (adapté de [25]).

1.3.2. Les hème-senseurs de NO

L'oxyde nitrique (NO), connu comme molécule de signalisation et toxique à fortes concentrations, est aussi détecté par les senseurs à base d'hème. Parmi les senseurs bactériens à NO connus se trouvent : le régulateur NO-dépendant DNR ("dissimilative nitrate respiration regulator"), qui intervient dans la régulation de la dénitrification chez *Pseudomonas aeruginosa*, et l'homologue de la guanylate cyclase soluble, H-NOX chez *Nostoc punctiforme* (illustré dans la figure 1.7).

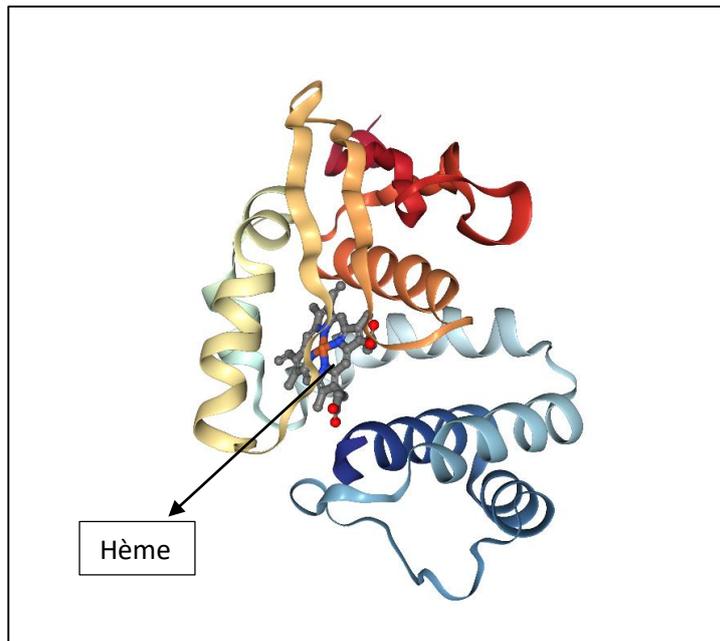


Figure 1.7 : Structure cristalline de H-NOX chez *Nostoc* à 2,1 Å de résolution, avec son substrat hème (PDB, 6MX5) [26].

Le senseur à NO le mieux caractérisé est la guanylate cyclase soluble (GCs) eucaryote (illustré dans la figure 1.8), qui synthétise le second messenger 3'-5' GMPc à partir du GTP en répondant au NO et en forte compétition avec l'O₂ [27]. À noter que le NO se lie au Fe²⁺ de l'hème (dans l'hémo-domaine en N-terminal) de la GCs [28]. Cette enzyme hétérodimère composée de deux sous unités homologues (voir figure 1.8) est impliquée dans différentes voies de signalisation dans les systèmes cardiovasculaires et nerveux, surtout la vasodilatation [29, 30].

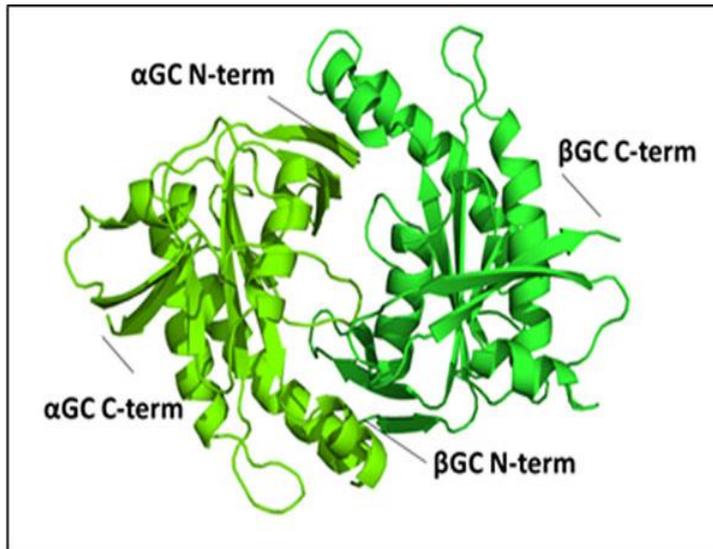


Figure 1.8 : Structure cristalline à 1,9 Å de résolution du domaine catalytique $\alpha\beta$ GC de la guanylate cyclase soluble humaine, un senseur à NO [31].

1.3.3. Les hémosenseurs à CO

À ce jour très peu d'hémo-senseurs à CO sont connus. Chez les eucaryotes le senseur de CO le mieux étudié est la protéine neuronale NPAS2 [32]. Dans les bactéries, le régulateur de transcription CO-dépendant CooA, découvert dans la bactérie photosynthétique *Rhodospirillum rubrum*, est le mieux caractérisé et la structure tridimensionnelle de sa forme inactive a été déterminée [33]. Récemment, dans la bactérie aérobie *Burkholderia xenovorans* un nouveau hémo-senseur à CO, RcoM, a été identifié, mais de nombreux aspects mécanistiques de détection du CO par RcoM restent encore peu compris.

a) Les senseurs à CO chez les eucaryotes :

Le CO est également une molécule de signalisation importante chez les eucaryotes. Autrement que par synthèse exogène (voir 1.2.c), le CO est synthétisé par l'enzyme hème oxygénase à l'aide du NADPH cytochrome P450 réductase [34–36].

Un rôle de neurotransmetteur gazeux a été également attribué au CO [37].

La protéine neuronale NPAS2 est un récepteur du CO [37]. NPAS2 appartient à la famille des facteurs de transcription hélice-boucle-hélice (bHLH)-PAS. Le monomère de NPAS2 contient deux domaines (PAS-A et PAS-B) avec deux hèmes comme sites de fixation de CO [32]. La fixation de CO agit sur la liaison de NPAS2 à l'ADN à condition de la présence de son

partenaire hétérodimère BMAL1 [32]. NPAS2 peut se dimériser avec la protéine BMAL1 et activer la transcription des gènes du rythme circadien *per* et *cry* chez les mammifères [37]. Il est important de noter que l'activité de régulation de la transcription du dimère NPAS2-BMAL1 est inhibée par la liaison du CO à l'hème Fe (II). NPAS2 et le facteur de transcription Clock possèdent des séquences primaires très similaires et partagent le même partenaire dimérique, BMAL1. La figure 1.9 montre un schéma de l'interaction du régulateur Clock avec son partenaire BMAL1 et leurs régulations des gènes *per* et *cry*.

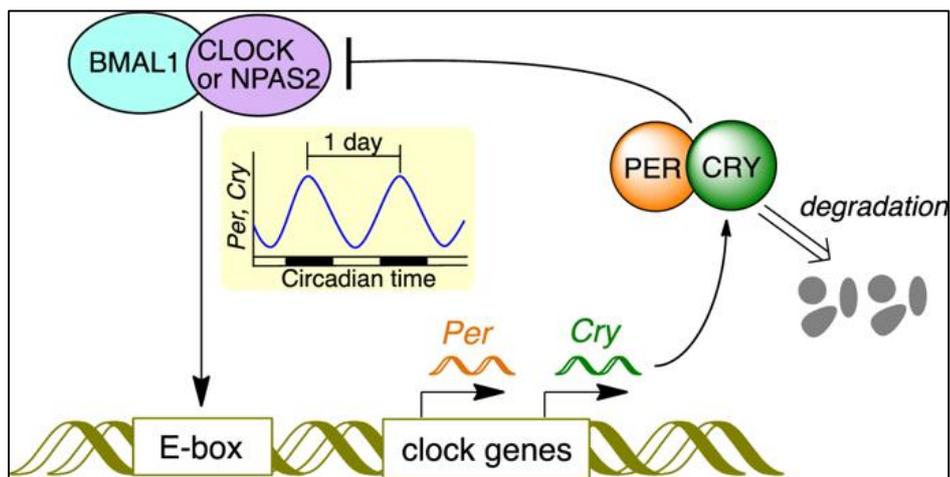


Figure 1.9 : Schéma de l'interaction du facteur CLOCK (séquence primaire très similaire de NPAS2) avec son partenaire dimérique BMAL1, et leurs activations des gènes répresseurs du rythme circadien *per* (Period) et *cry* (Cryptochrome) qui expriment des protéines PER et CRY et répriment l'activité du CLOCK : BMAL1 (adapté de [38]).

b) Les senseurs à CO chez les bactéries

Très peu de senseurs à base d'hème bactériens qui détectent du CO sont connus à ce jour. Les plus étudiés sont le senseur de CO anaérobie CooA et le senseur aérobie RcoM. Ces deux protéines senseurs présentent deux mécanismes non homologues de régulation transcriptionnelle CO-dépendante chez les procaryotes avec une similarité des propriétés fonctionnelles.

b.1) CooA, senseur à CO anaérobie

L'hémo-senseur CooA (« CO oxidation activator ») chez la bactérie photosynthétique *Rhodospirillum rubrum* est le senseur de CO procaryote le plus étudié actuellement. CooA appartient à la famille de régulateurs de transcription CAP (« catabolite activator protein »),

qui activent l'expression de certains gènes en réponse aux signaux externes. CooA a une similarité de 51% avec les protéines de cette famille (28% des séquences identiques), et 45% de similarité avec la famille des protéines fumarate nitrate réductase FNR chez *E. coli* (18% des séquences identiques) [18].

CooA possède une affinité micromolaire pour le CO [39], en accord avec son action sous des conditions d'anaérobie. CooA régule l'expression de l'opéron *coo* [18], qui code pour les CO déshydrogénase et hydrogénase, responsables de l'oxydation du CO, permettant ainsi à *R. rubrum* d'utiliser du CO comme seule source d'énergie et de carbone [40]. CO est l'activateur physiologique de CooA, aboutissant à sa liaison à sa cible d'ADN avec une affinité nanomolaire [41].

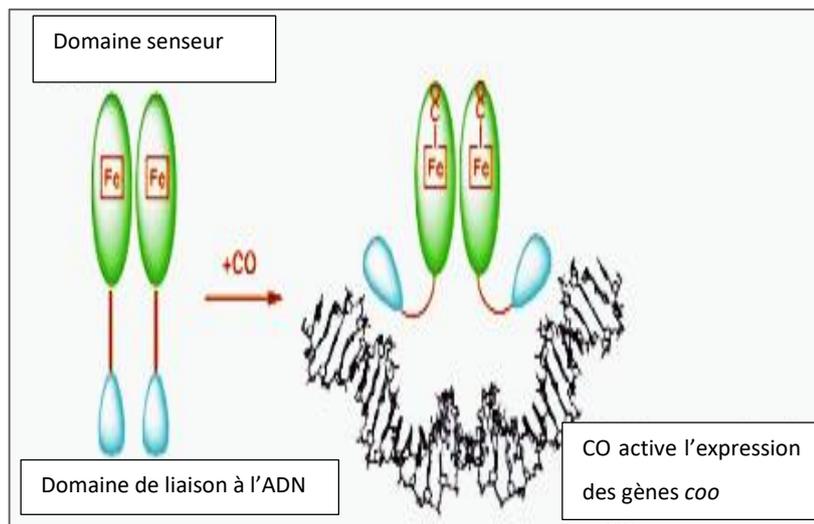


Figure 1.10 : Schéma du mécanisme de régulation par le senseur à un composant CooA homodimère. Domaine senseur avec l'hème de chaque monomère (en vert), domaines de liaison à l'ADN (en bleu) et régions de commutation (en orange). La fixation du CO aboutit à la liaison de CooA à l'ADN afin d'activer la transcription des gènes *coo* (adapté de [42]).

Différent du système à deux composants FixL/FixJ (voir fig. 1.6), CooA est un système à un composant. Ce senseur à CO est un homodimère dont chaque monomère possède en N-terminal un hémio-domaine senseur et en C-terminal un domaine de liaison à l'ADN [33] (voir figure 1.10).

L'homodimère CooA a un mécanisme de régulation allostérique, typique de la famille des protéines CAP [42]. L'hème se trouve en proximité de l'interface du dimère et est 6-coordonné dans la forme inactive. L'hème dans l'état réduit (Fe^{2+}) est lié à la protéine par une histidine proximale, His77. Dans la forme oxydée, ce ligand de l'hème est remplacé par la cystéine Cys75. Le ligand distal est le proline Pro2 de l'autre sous-unité [33], qui sera remplacé par le CO.

La structure 3D de CooA dans sa forme inactive a été déterminée [33]. La structure active n'est pas encore connue, mais plusieurs modèles suggèrent des changements structuraux importants au sein de la protéine, ainsi permettant la fixation à l'ADN, comme illustré dans la figure 1.11) [33, 43].

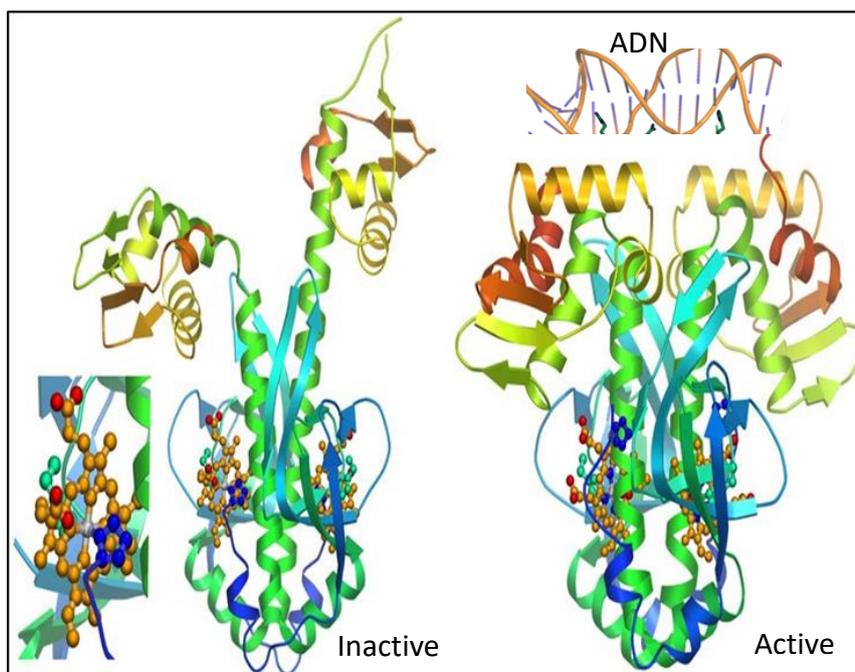


Figure 1.11 : Structures cristallines du senseur à CO CooA de *R. rubrum*. A gauche la forme inactive en absence du CO [33] ; à droite un modèle de la forme active en présence du CO [44]. Le site de fixation à l'ADN est en haut.

b.2) RcoM, senseur à CO aérobie

Le senseur à base d'hème RcoM (« regulator of CO metabolism ») est un facteur de transcription CO-dépendant qui a été identifié chez la bactérie aérobie *Burkholderia xenovorans* [45]. La bactérie *Burkholderia xenovorans* (souche LB 400) est une bactérie aérobie qui se trouve dans le sol, possède la capacité à cataboliser des composés aromatiques (PCB) dans le sol et présente un grand potentiel pour les applications biotechnologiques environnementales, telles que la bio-remédiation des sols contaminés par des PCB. Récemment un nouveau hème-senseur à CO a été identifié chez cette bactérie ; le RcoM.

Le RcoM régule l'expression des gènes *cox*, impliqués dans l'oxydation aérobie du CO [45]. Ce régulateur est capable de lier le CO avec une très forte affinité, mais ne fixe pas d'O₂. Chez la bactérie aérobie *Burkholderia xenovorans* deux homologues sont présents : RcoM-1 et RcoM-2, codés sur différents chromosomes et avec 93% d'identité au niveau protéique [45].

RcoM est un senseur à un composant qui possède un module de repliement PAS portant l'hème en N-terminal et un domaine de liaison à l'ADN du type LytTR en C-terminal [45, 46]. Comme dans le senseur CooA, l'hème est 6-coordonné et le ligand CO externe déplace un ligand interne (acide aminé) de l'hème.

L'alignement de séquences de RcoM-1 et RcoM-2 montre la conservation de His74 dans le domaine PAS et de His218 dans le domaine LytTR [45]. En l'absence d'une structure 3D pour RcoM, la figure 1.12 est une représentation visuelle de l'architecture des domaines PAS et LytTR [47].

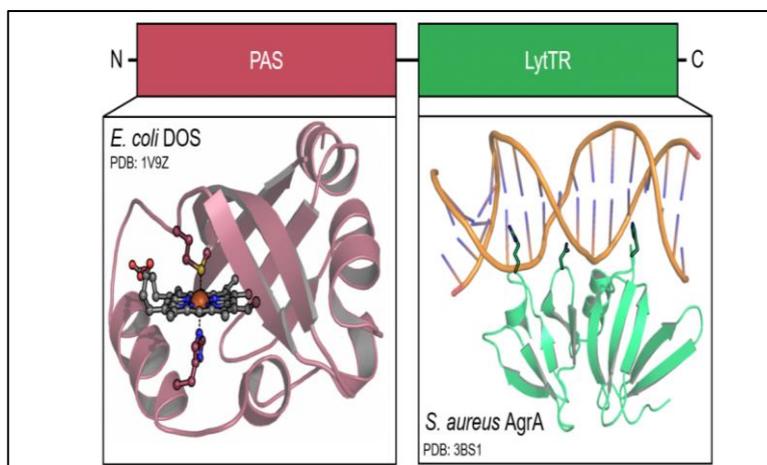


Figure 1.12 : Représentation visuelle de l'architecture 3D d'un domaine PAS en N-terminal (domaine senseur à hème 6-coordonné) et d'un domaine LytTR en C-terminal (domaine de fixation à l'ADN) [47].

Dans **RcoM-1**, les résidus His74 et Met104 ont été identifiés comme des ligands de l'hème (état Fe(II)) [45].

Dans **RcoM-2**, l'His 74 est aussi un ligand de l'hème et un résidu cystéine (Cys94) a été identifié comme l'autre ligand de l'hème dans l'état Fe(III). Lors de la réduction de l'hème (état Fe(II)), Cys94 est remplacé par un acide aminé méthionine (Met104), qui sera remplacé par le CO externe ; situation qui a été également observée pour le senseur Dos d'*E. coli* [45, 48] (voir figure 1.13).

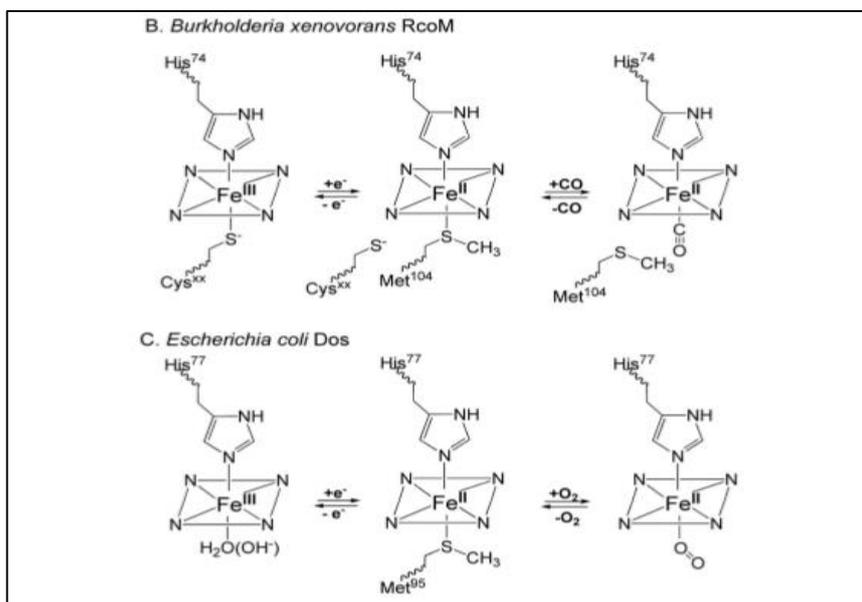


Figure 1.13 : en haut : Dans RcoM-2, un résidu cystéine a été identifié comme ligand de l'hème dans l'état Fe(III). Dans l'état réduit Fe(II)), cette cystéine est remplacée par un acide aminé méthionine (Met104), qui sera remplacé par le CO externe. En bas : Une réaction similaire a été observée pour le senseur Dos de *E. coli* [45, 48].

Les senseurs RcoM possèdent une affinité très élevée, de l'ordre nanomolaire, pour le CO [45, 48], en accord avec leur fonction de détecter des faibles concentrations de CO en présence d'oxygène.

L'hémo-domaine senseur de RcoM-2 (RcoMH-2) possède la propriété unique d'être purifiée dans la forme ferreuse (Fe²⁺) complexée au CO après expression aérobie chez *E. coli*. Ce complexe RcoMH-2-CO est très stable et le ligand CO se libère seulement après illumination forte pour une longue durée sous conditions aérobies [45, 48, 49]. L'observation qu'après

photodissociation RcoMH-2 ne forme pas d'intermédiaire oxydé indique une très faible affinité pour le O₂, essentiel pour un fonctionnement en tant que senseur spécifique de CO sous des conditions aérobies.

Dans toutes les protéines senseurs à hème, la liaison ou dissociation du ligand physiologique gazeux induit une perturbation dans l'environnement de l'hème. L'étude de ces processus primaires permet de déterminer la dynamique du ligand en temps réel et ainsi à mieux comprendre les étapes initiales de détection et de signalisation par ces senseurs et le rôle de l'environnement protéique dans ces processus.

Notre équipe a caractérisé la cinétique de l'interaction hème-CO dans l'hémo-domaine RcoMH-2 et trouvé qu'après photodissociation, le CO se lie sur l'hème très rapidement et de façon quasi complète et irréversible (voir figure 1.14) [49]. Cette observation pose la question comment RcoM-2 peut fonctionner physiologiquement comme senseur de CO et nous a conduit à étudier la protéine entière, incluant le domaine de fixation à l'ADN.

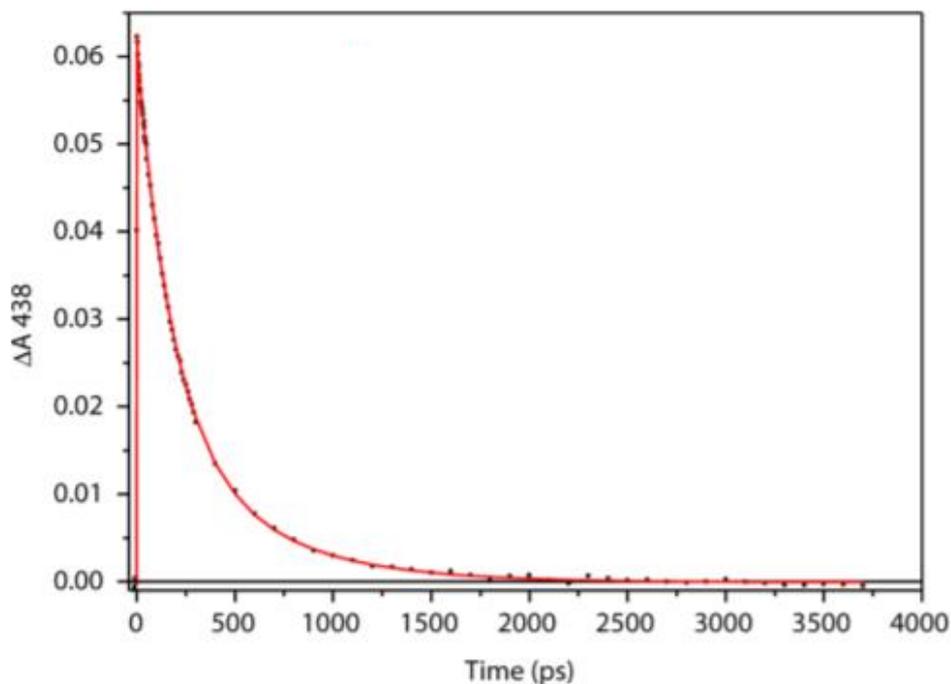


Figure 1.14 : Cinétique de recombinaison géminée du CO dans l'hémo-domaine RcoMH-2 mesurée à 438 nm [49].

Comme indiqué, les protéines RcoM sont des régulateurs de transcription CO-dépendants avec un domaine de liaison à l'ADN du type LytTR (voir figure 1.12). Dans le cas de CooA, il a été trouvé que la liaison du complexe CooA-CO à son ADN cible accélère la recombinaison géminée du CO au sein du domaine senseur [50].

Afin de tester si un effet de la liaison d'ADN sur la cinétique de recombinaison du CO peut être observée chez RcoMH-2, nous avons étudié ces propriétés pour la protéine entière de RcoM-2 dans ce travail de thèse et nous les avons comparées avec celles pour le domaine hémique RcoMH-2.

Comme décrit ci-dessus, à faibles quantités, le CO peut agir comme molécule de signalisation et peut affecter la régulation de gènes. Certaines bactéries, anaérobies ou aérobies, peuvent utiliser le CO comme source de carbone et énergie. Pour un grand nombre de bactéries aérobies, le CO en fortes concentrations exerce un effet toxique, p. ex. par l'inhibition de la chaîne respiratoire.

Il est intéressant à noter que les mycobactéries, bactéries aérobies obligatoires, sont capables de supporter des concentrations très élevées de CO (supérieures à 100 μM) suggérant des mécanismes de détection du CO et des réponses uniques pour combattre la toxicité du CO.

1.4. Les mycobactéries

1.4.1. Généralités et caractéristiques

Les mycobactéries sont des bactéries du genre *Mycobacterium*, des bacilles non sporulés, immobiles, non flagellés mesurant de 1 à 10 µM de long, qui ne se colorent pas facilement, acido-alcoolo-résistants, et aérobies strictes.

Ce genre comprend plusieurs espèces qui se divisent en :

- espèces pathogènes spécifiques comme *M. leprea* agent de la lèpre et *M. tuberculosis*, espèce sur laquelle on se focalise dans ce projet, l'agent de la tuberculose.
- espèces pathogènes opportunistes qui sont des mycobactéries atypiques, comme *M. avium*.

1.4.2. *Mycobacterium tuberculosis* (Mtb)

Mycobacterium tuberculosis (Mtb ou le bacille de Koch), est l'agent infectieux de la tuberculose et cause 1.7 millions de décès par an. Mtb a une large distribution mondiale [50], comme le montre la figure 1.15. Les bactéries se transmettent par voie aérienne et pénètrent principalement dans les voies respiratoires de l'Homme.

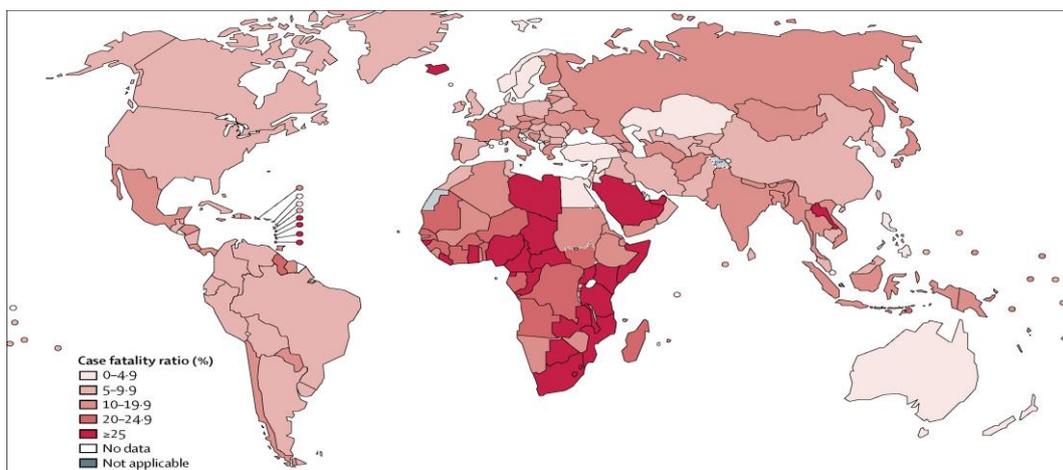


Figure 1.15 : Distribution mondiale de la tuberculose. On constate une concentration en Afrique suite à la pénurie des diagnostics et soins [51].

L'enveloppe cellulaire très hydrophobe des mycobactéries leur confère une résistance aux molécules thérapeutiques [52]. Egalement, elles possèdent un très long temps de réplication, ce qui complique leur éradication.

La souche de référence H37Rv de *Mtb* possède un génome de 4.4 Mpb constitué d'un chromosome unique et circulaire, très riche en GC (65.6%) [53].

Lors de l'infection par *Mtb* suite à une inhalation, les bacilles atteignent les alvéoles pulmonaires et dans la majorité des cas résistent au système immunitaire (inné et adaptatif). Cette résistance conduit à l'état de persistance ou dormance, qui peut durer plusieurs années ; on parle d'une infection latente non contagieuse, durant laquelle il y existe un équilibre entre l'infection bactérienne et le système immunitaire [52, 54].

Des changements de certaines conditions, surtout gazeuses, ou bien une faiblesse du système immunitaire, peuvent réactiver cette infection latente, comme le montre la figure 1.16.

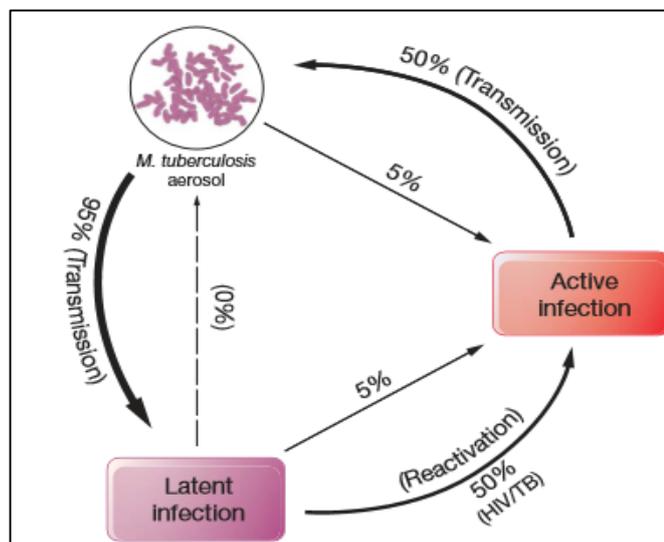


Figure 1.16 : Les différents étapes lors de l'infection tuberculeuse, incluent la transmission du bacille, une phase de latence et une phase d'infection active [54].

1.4.3. Mécanismes de défense de l'hôte

Après infection et début de phagocytose, les bacilles tuberculeux se localisent typiquement dans les macrophages alvéolaires, qui activent des mécanismes de défense afin de détruire ces bactéries. Parmi ces stratégies antimicrobiennes de l'hôte on connaît la production des espèces réactives oxygénées (ERO), l'acidification du milieu, des changements gazeux dont la libération du NO et la production du CO par l'enzyme hème oxygénase (HO) [55, 56].

Chez l'Homme et la souris, il existe trois isoformes de l'hème oxygénase, HO-1, HO-2, HO-3 (codées par les gènes *hmox1*, *hmox2* et *hmox3*, respectivement). Les trois isoformes catalysent l'oxydation de l'hème en biliverdine, qui est ensuite réduite en bilirubine, qui peut être excrétée. L'autre produit de la réaction est le CO (figure 1.17). Enfin, le fer libéré est recyclé et représente la source principale de ce métal dans l'homéostasie de l'hème [10, 57].

HO-2 et HO-3 sont exprimés de manière constitutive, alors que la HO-1 est induite par du stress oxydatif, en cas d'inflammation, et également par infection avec *Mtb* [58]. HO-1 est induit par le facteur de transcription Nrf2 [59–62], et réprimé par le facteur Bach1 [6] (voir figure 1.17). Une induction de l'expression de l'hème oxygénase-1 peut créer des concentrations locales élevées en CO, bénéfiques pour les cellules hôtes car le CO et la biliverdine / bilirubine peuvent agir en tant que molécules de signalisation et constituer une cytoprotection [63].

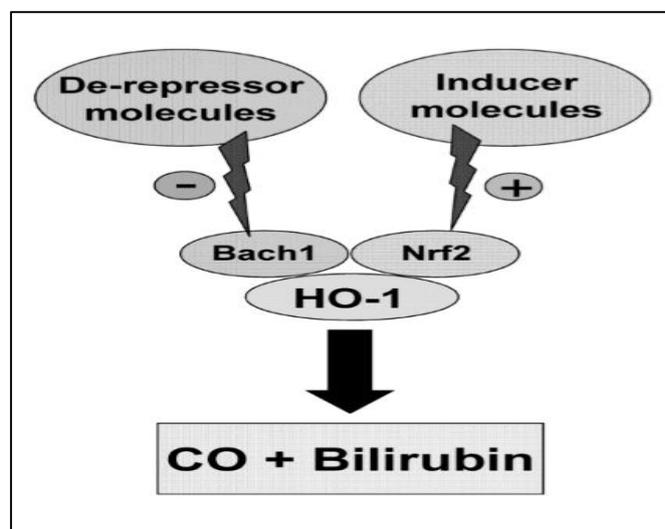


Figure 1.17 : La régulation de l'hème oxygénase HO-1 par les deux facteurs Nrf2/Bach1 et son induction pour catalyser l'hème en CO, Bilirubine et Fer (adapté de [63]).

La structure de la protéine HO-1 en présence de son substrat hème a été déterminée (fig. 1.18).

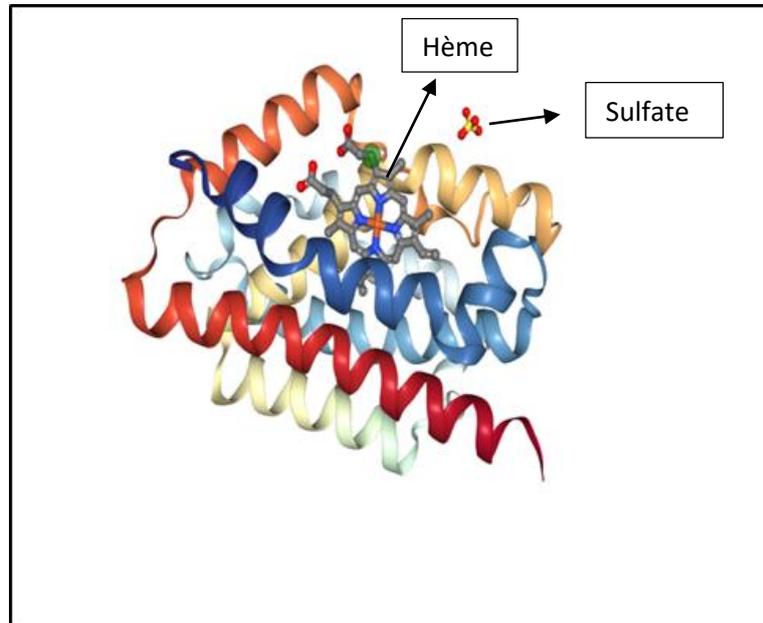


Figure 1.18 : Structure cristalline de l'hème oxygénase HO-1 humaine à 1,5 Å, en complexe avec son substrat hème (PDB 1N45). Un ion sulfate est présent. HO-1 possède une nouvelle structure hélicoïdale, offrant une flexibilité permettant la liaison du substrat et la libération du produit.

La question qui se pose est quels voies et mécanismes *Mtb* a mis en place pour échapper aux défenses diverses de l'hôte, en particulier le CO. Étant donné que *Mtb* réside dans les poumons, où le CO est exhalé de façon continue, il est raisonnable d'assumer que *Mtb* pourrait avoir développé des mécanismes de détection et réponse aux flux de CO. Quels sont les mécanismes de détection de changements gazeux chez *Mtb* et comment la bactérie peut résister aux fortes concentrations du CO ?

Pour mieux comprendre les mécanismes complexes de détection des gaz chez *Mtb*, un aperçu des systèmes connus est présenté.

1.4.4. Mécanismes de détection des gaz chez *Mycobacterium tuberculosis*

Des changements en concentrations gazeuses, en particulier l'hypoxie, jouent un rôle essentiel dans l'infection persistante par *Mtb*. Plusieurs systèmes à deux composants (TCS) sont impliqués, notamment les systèmes DosRS /DosT [64–66], et MprAB [67].

DosS et DosT forment un système de transduction de signal à deux composants avec le régulateur transcriptionnel DosR et sont essentiels à l'adaptation aux changements environnementaux, ainsi que dans l'état persistant de *Mtb* durant la dormance [64-66].

1.4.4.a) Système de régulation DosRS-DosT

Mtb réagit aux changements en concentrations gazeuses grâce à un système de signalisation à deux composants qui détecte des changements dans la tension d'O₂, NO et CO via les domaines de liaison à l'hème GAF de deux histidine kinases senseurs, DosS et DosT, et le régulateur de transcription DosR (voir figure 1.19).

Le système DosRS-DosT s'autorégule et induit un régulon défini de 48 gènes chez *M. tuberculosis* (le régulon de dormance) suite à une exposition à l'hypoxie, à l'oxyde nitrique (NO) et au monoxyde de carbone (CO) [64–66, 68] ; (voir figure 1.19). Des profils de transcription après exposition de *Mtb* au CO ont montré que le CO induisait la transcription de nombreux gènes du régulon de dormance [65].

Chez *Mtb*, le régulon de dormance est bien caractérisé et contient le gène *dosR* lui-même, ainsi que les gènes *dosS* et *dosT*, codant pour des kinases senseurs, le gène *rv0081*, des nitroréductases, la diacylglycérol acyltransférase (DGAT) et de nombreuses protéines de stress universelles [69].

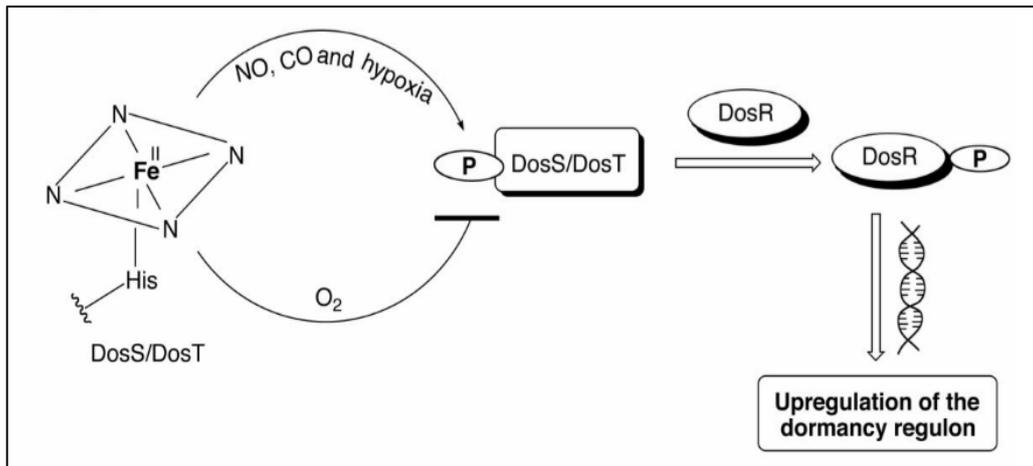


Figure 1.19 : Schéma récapitulatif de la fonction des senseurs kinases DosRS/DosT et du régulon de dormance sous des conditions d'hypoxie ou en présence du NO, CO (adaptée de [70]). Grâce à leurs domaines senseurs à hème (domaines GAF), DosS et DosT détectent les gaz diatomiques conduisant à leur phosphorylation et à l'activation de DosR par phosphorylation. DosR se lie à l'ADN cible et induit l'expression des gènes du régulon de dormance.

La liaison du DosR phosphorylé à l'ADN initie l'expression du régulon de dormance. DosR se lie aux deux régions distinctes dans le promoteur *acr* qui code pour la protéine Acr « alpha-cristalline » qui est nécessaire pour la croissance des *Mtb* sous conditions hypoxiques et exprimée au cours de la latence [65, 71]. Il a été trouvé que l'acétylation de la lysine 182 dans DosR inhibe sa capacité à se lier à l'ADN et à réguler l'expression des gènes au cours de l'hypoxie chez *Mycobacterium tuberculosis* [72].

1.4.4.b) Senseurs kinases DosS/DosT

Les senseurs kinases DosS et DosT détectent les gaz diatomiques par leurs hèmes et font partie des systèmes de régulation que *Mtb* a développé afin de persister et résister contre certains mécanismes de défense de l'hôte [68]. Ces deux kinases montrent une grande similarité, comme le montre la figure 1.20. Elles possèdent des domaines senseurs à hème du type **GAF** en N-terminal, le domaine GAF qui dérive des trois premières familles de protéines identifiées avec ce domaine nommées ; les cGMP-phosphodiesterases dépendantes chez les mammifères, Adenyl cyclases d'*Anabaena* et FhIA d'*E. Coli* [73, 74] et des domaines histidine kinase en C-terminal de structure connu [76, 77].



Figure 1.20 : Représentation schématique des deux senseurs kinases DosS et DosT de *Mtb*. Figure adaptée de [75].

Des études de photodissociation de ligands, suivies par spectroscopie d'absorption résolue en temps, des domaines senseur isolés de ces deux kinases montrent que chez DosT le O₂ présente un rendement faible de recombinaison comparé avec d'autres protéines senseurs à O₂ (1.5%), mais ce rendement est plus élevé chez DosS (18%). Malgré leurs similarités, la tyrosine distale, présente chez les deux senseurs, joue un rôle différent dans la dynamique spécifique des ligands au cours des phases initiales de signalisation intra-protéine grâce à son interaction de liaison hydrogène avec de l'oxygène moléculaire lié à l'hème [78].

La recombinaison géminée du ligand CO à l'hème est plus élevée chez DosT, suggérant que la poche de l'hème est plus fermée chez DosT, tandis que la recombinaison du NO montre une similarité entre DosS et DosT [78]. La protéine native DosS est purifiée en forme (Fe³⁺), indiquant que l'oxygène provoque directement l'oxydation de DosS et suggérant un rôle de DosS comme senseur de potentiel redox. Par contre, DosT est purifié en forme (Fe²⁺- O₂), suggérant un rôle de DosT comme senseur de gaz [68].

Néanmoins, comme montré la figure 1.21, la détection des gaz reste complexe et sujet à discussions.

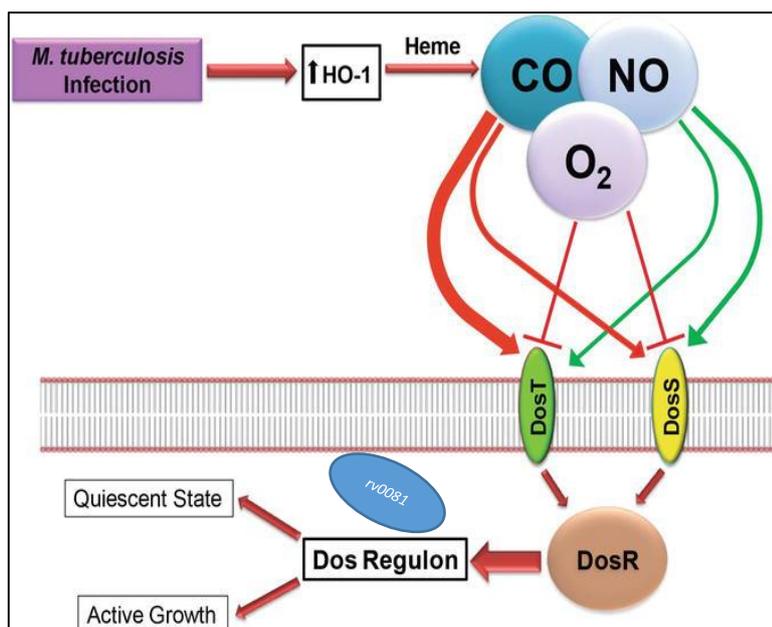


Figure 1.21 : Modèle de régulation de régulon de dormance (adapté de [79]). Suite à l'infection par *Mtb* et la production du CO et les fluctuations gazeuses, le CO et le NO activent les deux senseur kinases DosS et DosT, qui à leurs tours activent DosR. Le O₂ inhibe l'activité de ces deux kinases. La gène *rv0081*, qui code pour le facteur de transcription Rv0081 (en bleu) fait partie du régulon de dormance.

1.4.4.c) Le système MprA/B

Le système MprA/B est un système de transduction de signal à deux composants requis par *M. tuberculosis* pour sa croissance *in vivo* pendant le stade persistant de l'infection. Ce système régule l'expression de nombreux gènes sensibles au stress, comprenant la régulation directe de l'expression des facteurs de transcription importants SigB et SigE (voir figure 1.22). Le régulateur de réponse MprA se lie à des motifs conservés en amont de *sigB* et *sigE* et active leur transcription [67]. L'induction de *mprA* et de *sigE* par MprA en réponse à un stress est médiée par la senseur kinase MprB autophosphorylé, qui transfère le phosphate au régulateur de réponse MprA [67, 80].

Le gène *mprA* lui-même peut aussi directement être induit en réponse aux divers stress, établissant ainsi le rôle de MprA comme régulateur de l'expression des gènes *mprA* et *sigE* (voir figure 1.22).

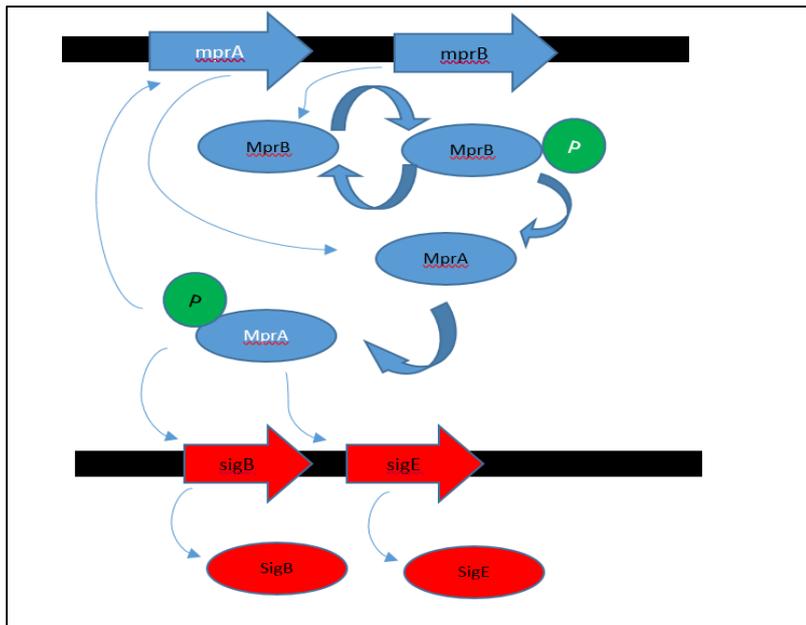


Figure 1.22 : Schéma récapitulatif de régulation par le système de transduction de signal à deux composants MprA/B et des facteurs sigma SigB et SigE chez *M. tuberculosis* ; le régulateur MprB autophosphorylé transfère le phosphate au MprA qui induit à son tour l'expression de *mprA* et *sigE*, le régulateur MprA active la transcription de *sigB* et *sigE*.

Le régulateur MprA peut également induire la transcription d'*acr2* « alpha-cristalline » par une interaction directe avec le promoteur *acr2* chez *Mtb* [81, 82]. *acr2*, transcrit en réponse au stress oxydatif produit par l'hôte chez *Mtb*, est aussi régulé par SigE et permet la persistance de *Mtb* en face au réponse immunitaire et sous conditions hypoxiques [81–83].

1.4.4.d) Le régulateur de transcription Rv0081

Des comparaisons entre les régulateurs MprA/B et DosRS/DosT ont montré qu'un sous-ensemble des facteurs peut être régulé par ces régulateurs comme par ex. Rv0081 et Rv3344. [84, 85].

Le régulateur de transcription Rv0081 qui appartient à la famille ArsR/SmtB des régulateurs transcriptionnels des procaryotes. Ces régulateurs, souvent dépendant des métaux [86], agissent comme répresseurs de la transcription ; ils empêchent l'activité de l'ARN polymérase. Dans les facteurs de transcription ArsR/SmtB qui dépendent des métaux, leurs liaisons diminuent l'affinité du régulateur pour son ADN cible.

rv0081 fait partie du régulon de dormance et intervient dans les changements transcriptionnels sous conditions d'hypoxie (montré dans la figure 1.21 en bleu). Le facteur Rv0081 régule son propre gène *rv0081* négativement par sa liaison à l'élément répété inversé dans sa région promotrice de type 12-2-12 pb [84]. Rv0081 est également un régulateur de nombreux autres gènes comprenant le locus *rv0081-rv0088* et est directement régulé positivement sous conditions hypoxiques par les régulateurs de réponse MprA et DosR, dont les sites de liaison chevauchent totalement ou partiellement ceux reconnus par Rv0081, respectivement. Le gène *rv0081* et des gènes en aval voient leur activité augmentée en réponse à NO [84].

Rv0081, une protéine dimère de 12 kDa [85], montre aussi que plusieurs niveaux de régulation de la liaison à l'ADN pourraient exister chez *Mtb* [85, 87] : autre que le locus *rv0081-rv0088*, sous conditions hypoxiques, Rv0081 régule positivement l'expression du gène *rv3334*, qui réprime la transcription de *rv0081* chez *Mtb* [85], suggérant que Rv0081 et Rv3334 sont actives dans la même voie régulatrice.

Il est intéressant de noter que les modifications de la disponibilité en oxygène entraînent des modifications de l'expression de presque un tiers de gènes dans *Mtb*. Un rôle de « régulateur central » gaz-dépendant dans l'expression des gènes sous conditions d'hypoxie et des changements gazeux a été suggéré pour Rv0081 [88].

1.4.4.e) Détection et réponse au CO dans *Mtb* - autres mécanismes

Comme mentionné plus haut, l'infection de l'hôte par *Mtb* ne cause pas nécessairement une infection active. Dans la majorité des cas, les *Mtb* rentrent dans l'état de dormance où elles persistent dans l'organisme de l'hôte, surtout dans les macrophages des alvéoles pulmonaires. Suite à l'infection par *Mtb*, l'induction de l'hème oxygénase 1 dans les macrophages conduit à la dégradation de l'hème, produisant ainsi du fer, biliverdine et CO. Bien que *Mtb* possède des systèmes de détection des gaz, leur spécificité et les réponses associées ne sont pas claires. En particulier, l'observation que *Mtb* est capable de détecter et résister à des concentrations très élevées de CO (>100 µM) suggère des mécanismes particuliers et uniques que ces espèces ont développés, afin de répondre aux changements importants de la concentration du CO et combattre sa toxicité en présence de l'oxygène.

1- Le gène *cor* (carbone monoxyde resistance)

Récemment un nouveau gène nommé *cor* (« CO resistance » ; *rv1829*) a été identifié dans la souche *Mtb*H37 dont l'inactivation augmente drastiquement la sensibilité de *Mtb* contre le CO [89]. La figure 1.23 montre un modèle de son implication cellulaire, mais aucune donnée mécanistique de *Cor* dans la détection et réponse au CO était connue au début de mon travail de thèse.

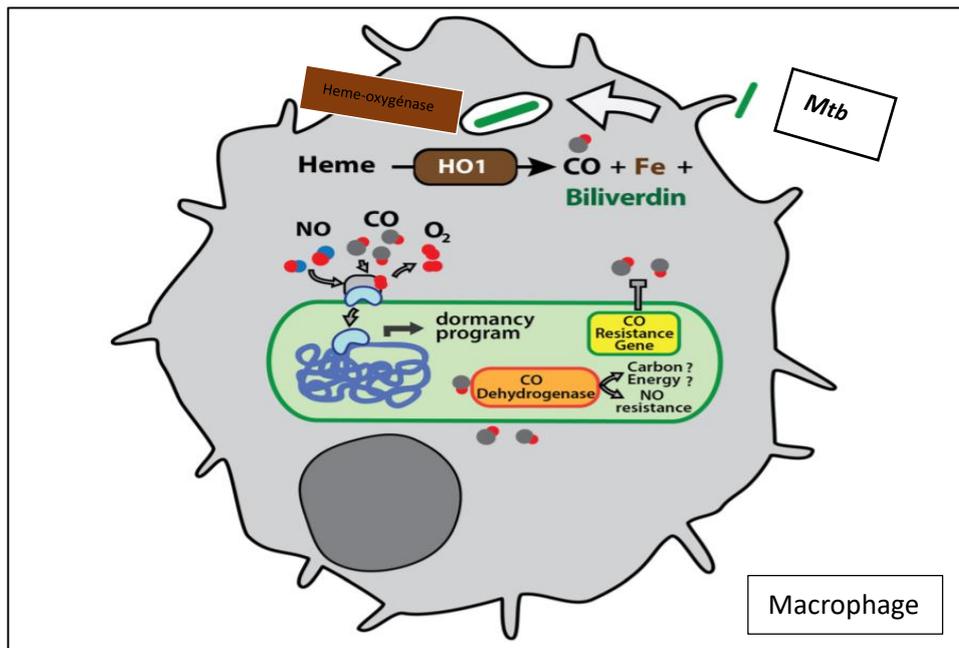


Figure 1.23 : Schéma montrant la production du CO dans le macrophage de l'hôte (comme stratégie anti-microbienne) suite à l'infection par *Mtb*. HO-1 est induite et catabolise l'hème en biliverdine, Fe et CO. CO, NO et de l'hypoxie peuvent activer le programme de dormance chez *Mtb*. Grâce à l'expression du gène *cor*, *Mtb* pourra résister à l'inhibition de la croissance induite par le CO. Certaines mycobactéries peuvent cataboliser le CO par le CO déshydrogénase (CODH) et l'utiliser comme source de carbone pour la croissance. Un rôle de CODH dans la résistance contre le NO produit par l'hôte est également possible (adapté de [58]).

cor (*rv1829*) code pour une protéine de 164 acides aminés, d'une masse moléculaire prédite de 18 kDa. La protéine *Cor* est conservée chez différentes espèces mycobactériennes et également chez *Rhodococcus* sp. et *Streptomyces* sp., comme le montre l'alignement dans la figure 1.24.

100 %	<i>Mycobacterium tuberculosis</i>	MGEVRVVGIRVEQPQNPVLLREANGDRYLPINIGQSEAAATALEQQGVEPPRPLTHDLIRDVIAALGHSLKEVR	76
	<i>Mycobacterium basiliense</i>	MGEVRVVGIRVEQPQNPVLLREANGDRYLPINIGQSEAAATALEQQGVEPPRPLTHDLIRDVIAALGHSLKEVR	
	<i>Rhodococcus fascians</i>	MSEMRVVGIRVEQPQNPVLLRESEDRYLPINIGQTEAAATALEQQGVQPARPLTHDLVKNLISALGHSLKEVR	
	<i>Streptomyces griseus</i>	-----MPSNPQIVLLREVGGDRYLPINIGPGEATAIAFAQQGMAPRPLTHDLFKDVLKALGQELTEVR	
	<i>Mycobacterium xenopi</i>	----MVGIRVEQPQNPVLLREANGDRYLPINIGQSEAAATALEQQGVEPPRPLTHDLIRDVIAALGHSLKEVR	
	<i>Thermotoga maritima</i>	XRXAWKTLALDRVSNTPWILGIEGTRVLPINIGACEGHALALAXEKXEFPRPLTHDLLLSVLESLEARVDKVI	
100 %	<i>Mycobacterium tuberculosis</i>	77 IVDLQEG--TFYADLIF DRNIK VSARPSDSVATIALRVGPIYVEEAVLAQGLLIPDESDEEATTAVREDEVEK ----- 164	
	<i>Mycobacterium basiliense</i>	IVDLQEG--TFYADLIF DRNWK VSARPSDSVATIALRVGPIYVEEAVLAQGLLIPDEGDEDTGSVREDEVEK ----	
	<i>Rhodococcus fascians</i>	IVDLQEG--TFYADLVF DKDIR VSARPSDSVATIALRVGPIYAEFPVLAEGLLMPDE-----REDEVEK ---	
	<i>Streptomyces griseus</i>	ITDLREG--VFYAEVLF AGGVE VSARPSDIAIALRTGTPYVSDGVLDDAGIAIPDEQED-----EVEK --	
	<i>Mycobacterium xenopi</i>	IVDLQEGpsTLT----- ---- -----	
	<i>Thermotoga maritima</i>	IHSLKDN--TFYATLVI[7]EEDEE[5]IDSRPSDAIILAVKTGAPIFVSONLVKHS--IELEVNERDLINSR----- --	

Figure 1.24 : Alignement de Cor (Rv1829) de *M. tuberculosis*, *M. basiliense*, *M. xenopi*, *Rhodococcus fascians*, *Streptomyces griseus*, *Thermotoga maritima*, qui montre une grande conservation des acides aminés entre ces espèces. Les pourcentages d'identité de séquence entre ces espèces sont indiqués à gauche.

Cor contient un domaine de fonction inconnue 151 (DUF151) « domain of unknown function » [89]. La structure 3D d'un homologue potentiel de Cor chez la bactérie *Thermotoga maritima* (33,33% d'identité de séquence avec Cor) s'avérait nouvelle et fournit donc peu d'information sur sa fonction (voir figure 1.25) [89, 90]. Cependant, il a récemment été rapporté que le domaine DUF151 contenant la protéine OmBBD du riz *Oryza minuta* présentait une activité nucléase [90]. Une telle activité n'a pas été détectée dans Cor [89].

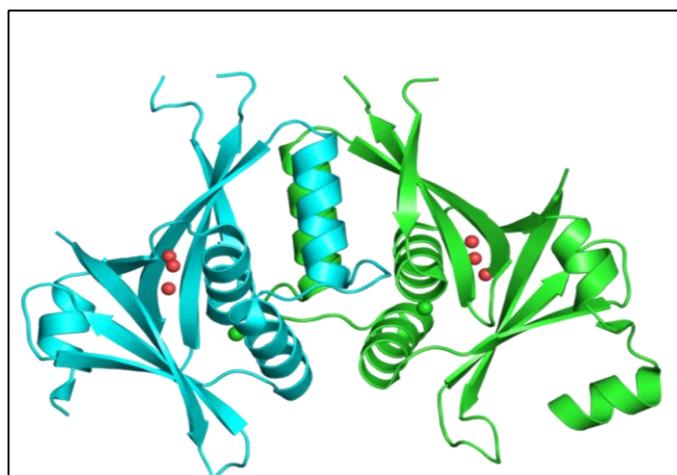


Figure 1.25 : Structure cristalline d'une protéine à DUF 151 (33,33% d'identité de séquence avec Cor) chez *Thermotoga maritima* (pdb 1SJ5 ; tm 0160).

2- Cluster de gènes impliquant *cor* (*rv1829*)

Le gène *cor* (*rv1829*) fait partie d'un cluster de gènes très répandu et conservé chez différentes espèces mycobactériennes (parmi lesquelles *M. tuberculosis*, *M. smegmatis*, et *M. marinum*) (fig. 1.26), parmi ces gènes *rv1828* comme régulateur de transcription gaz/redox-dépendant prédit. Malgré la nature conservée du cluster de gènes, les sites d'initiation de transcription n'ont pas été cartographiés, ainsi l'organisation en opérons n'est pas connue à ce jour.

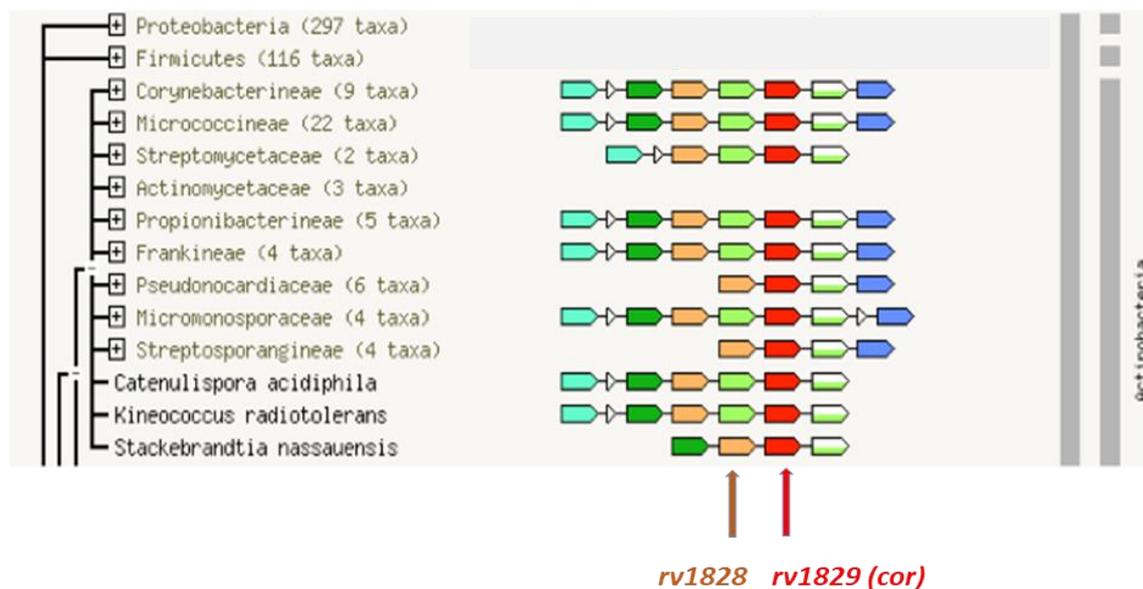


Figure 1.26 : Cluster de gènes impliquant *cor* (*rv1829*) et sa présence dans différentes espèces bactériennes. *rv1828* est un régulateur de transcription prédit. Figure basée sur analyse STRING (<https://string-db.org/>)

En se focalisant sur ce cluster (figure 1.27), nous nous sommes intéressés en particulier au gène *rv1828*, prédit comme régulateur de transcription.

Un autre gène, *rv1832* (*gcvB*) (figure 1.27), qui code pour une protéine du système de clivage de la glycine, catalysant la dégradation de l'acide aminée glycine, a attiré notre intérêt. Ce système est impliqué dans le métabolisme à 1-carbone et régulé par des conditions redox ou hypoxiques.

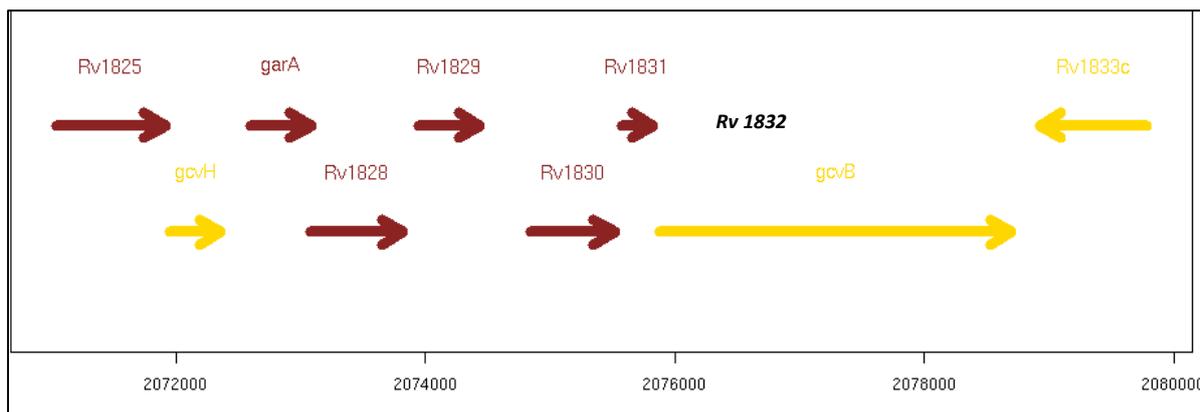


Figure 1.27 : Cluster de gènes impliquant *cor* (*rv1829*), *rv1828*, et *rv1832*. En jaune : gènes prédits être impliqués dans le métabolisme et la respiration. En rouge : gènes conservés hypothétiques (<http://svitsrv8.epfl.ch/tuberculist>).

3- Le régulateur de transcription prédit Rv1828

Le gène *rv1828* est essentiel pour la pathogénicité de *Mycobacterium tuberculosis* [91]. Il code pour une protéine Rv1828 qui a été prédite comme facteur de transcription appartenant à la famille des régulateurs du type MerR (Metal-responsive Transcriptional Regulators). Les protéines de cette famille fonctionnent typiquement comme des répresseurs en l'absence d'ions métalliques et deviennent des activateurs lors de la liaison métallique [92, 93]. Les régulateurs du type MerR sont des homodimères, comprennent typiquement trois domaines : un domaine de liaison à l'ADN, un domaine d'activation permettant la liaison de métal, et un domaine de jonction permettant la dimérisation des deux monomères. Ces régulateurs se lient à une séquence d'ADN spécifique dans une région promotrice entre les éléments -35 et -10 par l'intermédiaire de leur domaine N-terminal.

Rv1828 est une protéine de 247 aa et possède une masse moléculaire de ≈ 26 kDa par monomère. Son domaine N-terminal montre une similarité de séquence de 28% avec les autres membres de la famille MerR, tandis que le domaine C-terminal ne montre aucune similarité avec d'autres régulateurs déjà caractérisés de cette famille, suggérant alors que Rv1828 pourrait appartenir à une nouvelle sous-famille de MerR [91]. La structure 3D de son domaine C-terminal a été déterminée très récemment et montre un dimère (voir figure 1.28) [91].

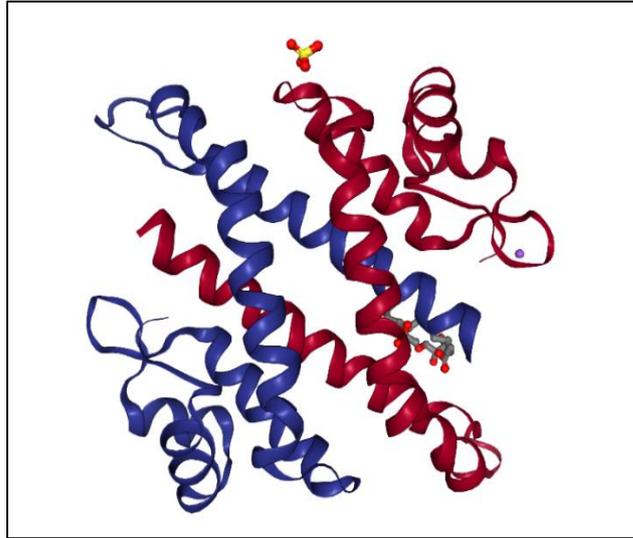


Figure 1.28 : Structure cristalline du domaine C-terminal de Rv1828, montrant les deux monomères (PDB 5YDD).

Des études biochimiques et structurales très récentes ont montré que Rv1828 peut se lier spécifiquement à son propre promoteur (région intergénique *rv1826-rv1827* de 239 pb) grâce à son domaine N-terminal, mais également aux autres régions promotrices (comme p. ex. *rv1827-rv1828*) pour réguler la transcription de leurs gènes [91]. Rv1828 pourrait p. ex. être impliquée dans la régulation du métabolisme du glutamate [91] par la protéine GarA (Rv1827) [94].

Suite aux prédictions *in silico*, nous voulions tester si Rv1828 peut potentiellement être impliquée dans une régulation de *cor* en se liant à sa région régulatrice prédite *rv1828-rv1829*.

1.5. Problématique et objectifs de la thèse

Mon travail de thèse avait comme but de mieux comprendre les mécanismes de détection et signalisation du CO chez certaines bactéries aérobies. Un premier aspect concerne l'hémo-senseur à CO, RcoM-2, chez *Burkholderia xenovorans* et l'interaction hème-CO et protéine-ADN. Un second objectif est de mieux comprendre comment les mycobactéries, en particulier *Mtb*, détectent le CO même en présence d'oxygène.

a) La première partie consiste à comprendre comment l'hémo-protéine entière RcoM-2 peut fonctionner comme senseur à CO et à déterminer les caractéristiques de la cinétique de liaison et dissociation du ligand CO. Ces données ont été comparées avec les propriétés cinétiques du CO dans le domaine senseur à hème seul (RcoMH-2) (réalisées par l'équipe dans des travaux précédents). Ces études se sont appuyées sur les approches de biologie moléculaire, biochimiques et spectroscopiques, dont la photodissociation du ligand suivie d'une détection par spectroscopie d'absorption ultrarapide. Un autre aspect concerne l'étude des effets de la liaison du CO dans l'hémo-domaine sur la fixation de RcoM-2 à l'ADN cible, que j'ai étudié par anisotropie de fluorescence.

Le but de ces expériences est de donner des informations sur le rôle du domaine hémique dans RcoM-2 dans la détection et signalisation du CO chez *Burkholderia xenovorans*.

b) La seconde partie de mon travail de thèse avait pour objectif de mieux comprendre les mécanismes de détection du CO chez *Mtb* par la protéine Cor. À cet effet nous avons d'abord caractérisé cette protéine biochimiquement et spectroscopiquement. Après l'avoir exprimée et purifiée, nous avons démontré que Cor est un dimère fonctionnel très stable. De façon surprenante, et contrairement aux données dans la littérature [89], nous avons trouvé que Cor est capable de fixer un cofacteur hémique et peut donc *à priori* détecter le CO de façon directe.

Nous avons identifié un acide aminé de Cor, capable de lier l'hème sur lequel se fixe le CO. Également, nous nous sommes intéressés à caractériser les aspects dynamiques de l'interaction Cor-hème-CO.

Nous avons également testé l'implication possible de deux régulateurs de transcription, Rv0081 (voir fig. 3.16) et Rv1828 dans la régulation transcriptionnelle du *cor* et testé si la protéine Cor possède de l'activité de fixation à l'ADN dépendant du CO ou non.

RÉSULTATS

Chapitre II

Mécanismes de détection et signalisation du CO : l'hémo-senseur RcoM-2

2.1. Introduction

Le senseur à base d'hème RcoM-2 de la bactérie aérobie *Burkholdéria xenovorans* est un facteur de transcription CO-dépendant [45]. RcoM-2 possède un domaine PAS portant l'hème en N-terminal et un domaine de liaison à l'ADN du type LytTR en C-terminal [45, 46]. L'hème est 6-coordonné et le CO externe déplace un ligand interne de l'hème ; très probablement une méthionine. RcoM-2 possède une affinité extrêmement élevée pour le CO, tout en étant insensible à l'O₂ [45, 48].

Dans l'hémo-domaine RcoMH-2 après photodissociation le CO se relie sur l'hème très rapidement et de façon quasi irréversible (voir figure 1.14) [49]. Ces observations posent la question de la manière dont RcoM-2 peut fonctionner comme senseur à CO, et nous ont conduit à déterminer les caractéristiques de l'interaction protéine-ligand de la protéine entière RcoM-2, comprenant le domaine de fixation à l'ADN. Nous avons exprimé et purifié la protéine, caractérisé l'interaction hème-CO et l'avons comparée avec les données pour le domaine senseur à hème isolé pour caractériser RcoM-2 comme senseur de CO.

Finalement nous avons étudié si la présence de l'ADN influence la cinétique de la liaison hème-CO, et déterminé l'affinité de la protéine pour son ADN cible par anisotropie de fluorescence.

2.2. Caractérisation de RcoM-2

2.2.1. Clonage, expression et purification de RcoM-2

Le gène *rcom-2*, codant pour la protéine RcoM-2 de *Burkholderia xenovorans* (804 pb), a été synthétisé par Genecust, optimisé pour expression dans *Escherichia coli* et cloné dans le plasmide d'expression pQE80L (QIAGEN). Ce plasmide permet l'expression de protéines avec une séquence 6xHis en N-terminal et une purification par chromatographie d'affinité sur des résines de métal chélate. Après induction par l'isopropyl-1-thiol-beta-D-galactopyranoside (IPTG), la protéine entière RcoM-2 (267 acides aminés) a été exprimée dans *E. coli* souche BL21DE3 et purifiée sur les colonnes Protino-NiTED (Macherey-Nagel) (voir Matériel et Méthodes).

La concentration (C) de la protéine purifiée a été déterminée par mesure d'absorption (A) à 280 nm et en appliquant la loi de Beer-Lambert $A = \epsilon.L.C$, avec L le chemin optique et ϵ l'extinction à 280 nm de l'ensemble des acides aminés aromatiques dans la protéine, $7115 \text{ mM}^{-1}\text{cm}^{-1}$. Ensuite la masse moléculaire apparente et la pureté ont été vérifiées par migration sur SDS-PAGE 10%. Le gel d'électrophorèse a été coloré au Bleu de Coomassie, montrant une protéine de masse moléculaire apparente d'environ 30.5 kDa (voir figure 2.1).

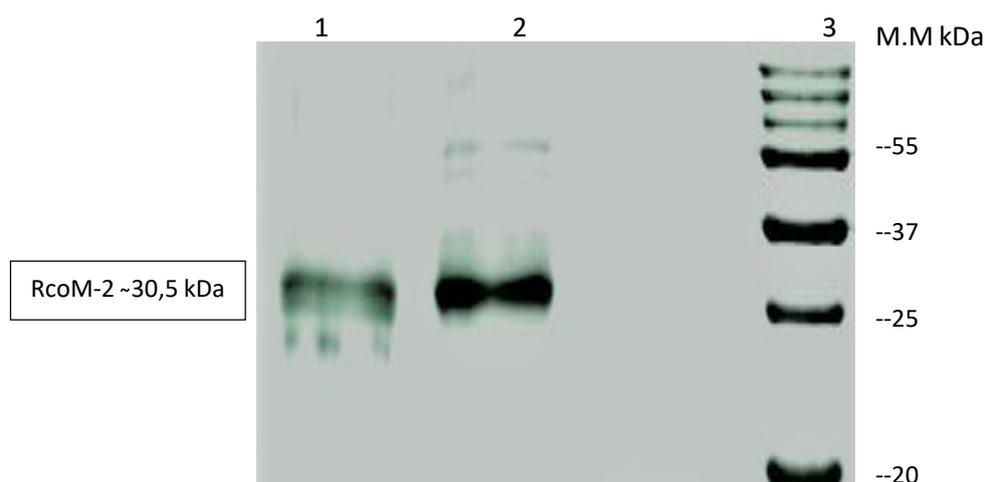


Figure 2.1 : SDS-PAGE 10% coloré au Bleu de Coomassie. La protéine RcoM-2 (dans puits 1 et 2) a une masse moléculaire apparente de ≈ 30.5 kDa. Puits 3 : Marqueur moléculaire.

2.2.2. Analyses de RcoM-2 par spectroscopie d'absorption à l'équilibre

Les spectres d'absorption à l'équilibre de la protéine entière RcoM-2 ont été mesurés sous différentes conditions, comme le montre la figure 2.2.

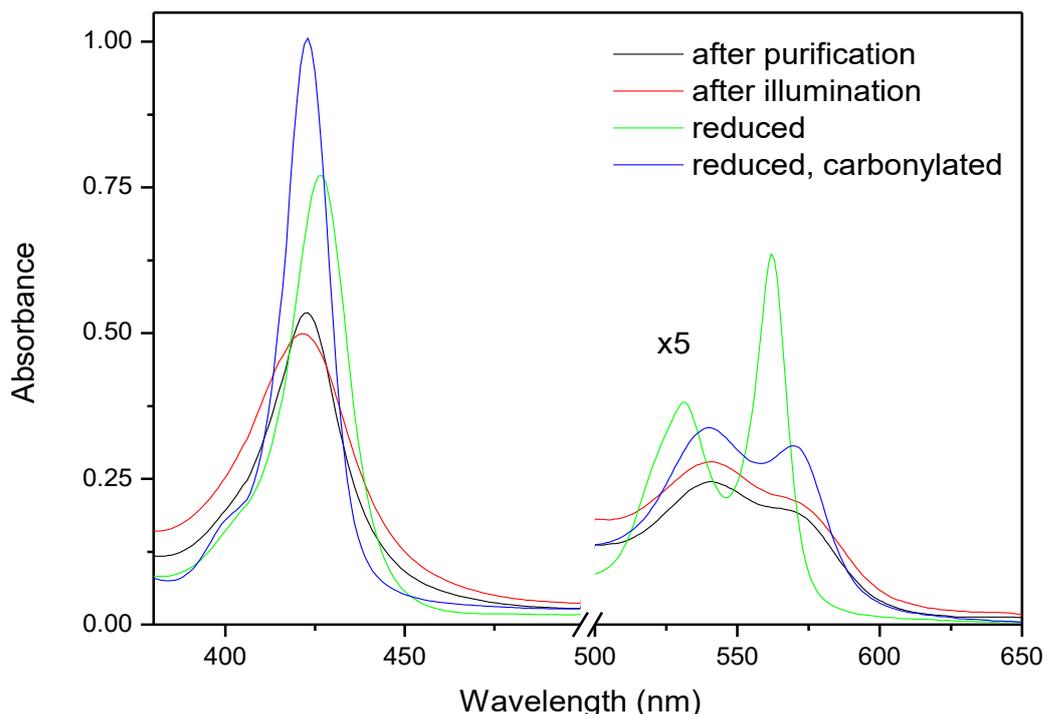


Figure 2.2 : Spectres d'absorption à l'équilibre de RcoM-2, après purification (courbe noire), après illumination prolongée (courbe rouge), après dégazage et réduction avec du dithionite de sodium (courbe verte), et après réduction et exposition à 1 atm. CO (en bleu). Les bandes Soret (400-450 nm) et Q (>500 nm, amplifié par 5 dans la figure) sont représentées.

Les spectres de RcoM-2 purifiée, réduite et carbonylée sont similaires à ceux rapportés précédemment dans la littérature [48] et représentent les complexes hème FeIII His-Cys, FeII His-Met et FeII His-CO, respectivement. Après une illumination prolongée le spectre change quelque peu, avec une bande de Soret légèrement décalée et élargie et un épaulement moins prononcé vers 570 nm (courbe rouge), indiquant qu'une fraction de CO était déjà présente dans la protéine purifiée. Cette fraction est beaucoup plus petite que celle trouvée dans l'hémo-domaine isolé RcoMH-2, qui est presque complètement purifié avec du CO lié [49].

Pour la protéine homologue RcoM-1 purifiée, des variations de liaison du CO ont été trouvées et il a été suggéré qu'elles pourraient être dues à des différences dans l'expression des protéines [46]. Etant donné que nous avons utilisé le même système d'expression et des conditions de croissance d'*E. coli* identiques pour l'expression de RcoM-2 et l'hémo-domaine RcoMH-2, nous attribuons la différence de liaison de CO à la protéine purifiée aux différences intrinsèques d'affinité pour le CO, comme cela sera quantifié ci-dessous. Le spectre de RcoM-2 après réduction par dithionite de sodium, puis liée au CO (courbe bleue) est très similaire au spectre de RcoMH-2 sous les mêmes conditions.

2.2.3. Cinétique de liaison du ligand CO

La liaison du ligand CO à l'hème de la protéine RcoM-2 a été étudiée par spectroscopie « stopped-flow ». La figure 2.3 montre des spectres à différents temps après le mélange de RcoM-2 réduite, avec du CO. Le spectre initial correspond au complexe 6-coordonné His-Met et le dernier au complexe His-CO 6-coordonné. Ces spectres montrent des points isobestiques clairs (le point isobestique est celui où toutes les espèces possèdent la même extinction molaire ϵ) et ainsi il y a aucune indication pour des espèces intermédiaires distinctes.

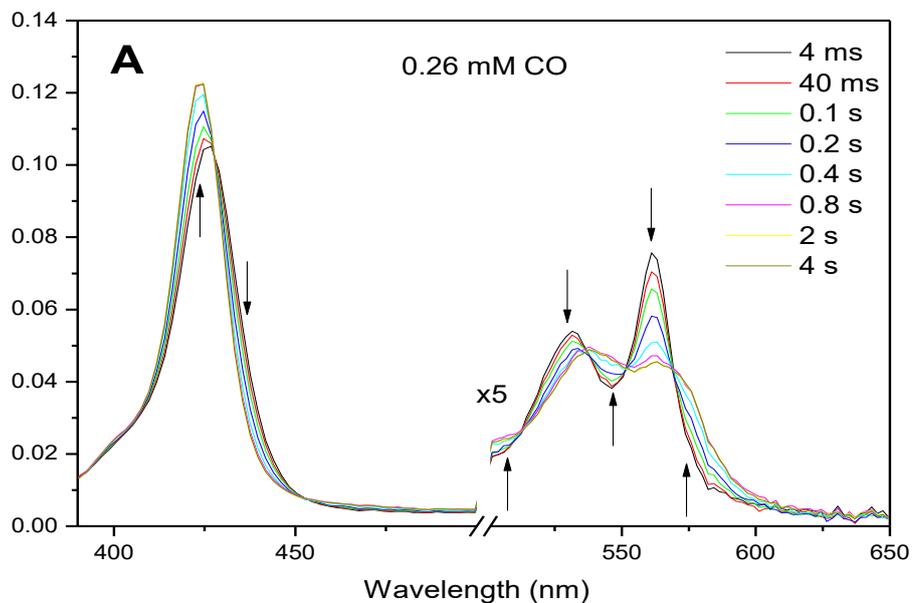


Figure 2.3 : Spectres issues de spectroscopie « stopped flow » de liaison du CO à la protéine RcoM-2, à différents temps de retard après mélange de RcoM-2 réduite, avec du CO jusqu'à une concentration finale en CO de 0,26 mM.

Dans le domaine de concentrations exploré, la cinétique de cette liaison est monophasique avec une vitesse qui varie linéairement avec [CO] (figure 2.4), ce qui donne une constante de vitesse apparente bimoléculaire k_{obs} de $16 \times 10^3 \text{ M}^{-1}\text{s}^{-1}$.

La dépendance de k_{obs} de [CO] attendu a la forme [49] :

$$k_{obs} = \frac{k_{-M}k_{CO}[CO]}{k_M + k_{CO}[CO]}$$

où (k_{-M}) est le taux de dissociation et de déplacement du ligand intrinsèque (méthionine), (k_M) le taux d'association du ligand intrinsèque ($K_M = k_M/k_{-M} \gg 1$, comme la protéine est initialement 6-coordonnée), et k_{CO} est la constante de vitesse bimoléculaire de la liaison du CO à l'hème 5-coordonné.

Dans la plage [CO] testée, aucune saturation de vitesse n'a été observée. Cela montre que la vitesse n'est pas limitée par la dissociation du résidu interne méthionine, et $k_{-M} \gg 5 \text{ s}^{-1}$. k_{obs} correspond alors au k_{CO}/K_M .

Ces résultats sont distincts de ceux rapportés précédemment par l'équipe pour l'hémo-domaine isolé RcoMH-2, où la cinétique était biphasique, et la phase principale montrait une saturation.

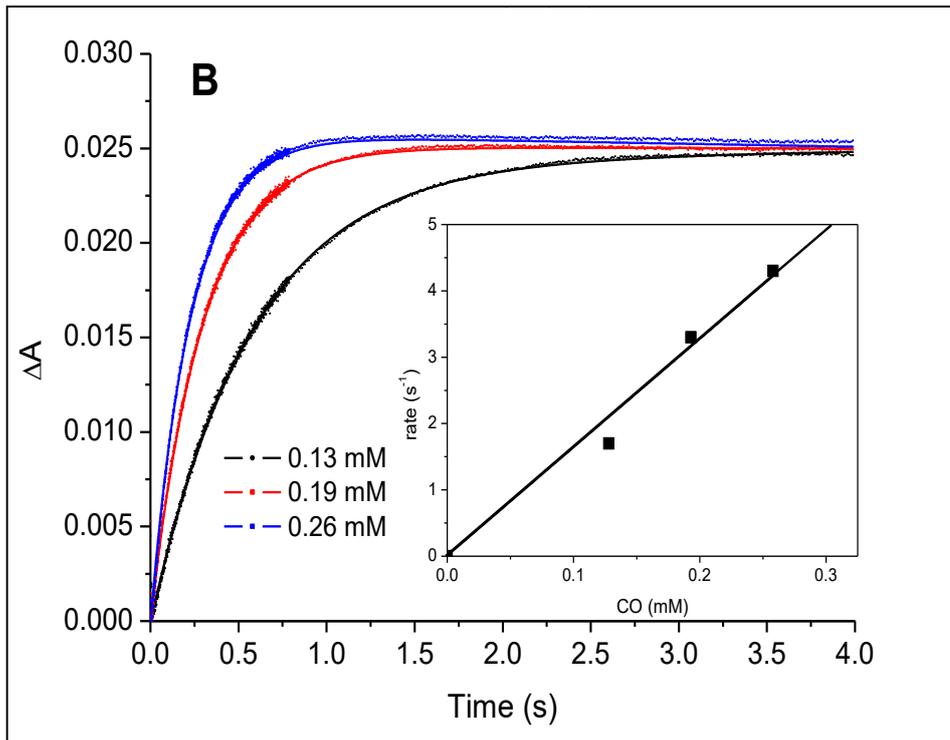


Figure 2.4 : Cinétique à 420 nm à différentes concentrations de CO. Fit linéaire aux taux correspondant au $k_{obs} = 16 \times 10^3 \text{ M}^{-1}\text{s}^{-1}$.

La constante de vitesse apparente bimoléculaire $k_{obs} 16 \times 10^3 \text{ M}^{-1}\text{s}^{-1}$, déterminée pour la protéine entière RcoM-2, est très similaire de $k_{obs} 15 \times 10^3 \text{ M}^{-1}\text{s}^{-1}$ qui a été trouvée pour l'hémo-domaine isolé [49]. Cela indique que la vitesse de liaison du CO est essentiellement la même dans les deux protéines, suggérant que le CO entre dans la protéine directement au niveau du domaine de l'hème et que la voie de migration intra-protéique ne passe pas par le domaine de liaison à l'ADN. Par contre, dans l'hémo-domaine RcoMH-2 un comportement de saturation a été trouvé, correspondant à $k_{-M} = 5 \text{ s}^{-1}$. Dans RcoM-2 $k_{-M} \gg 5 \text{ s}^{-1}$, indiquant que la vitesse de dissociation k_{-M} du ligand hème distal intrinsèque est bien plus grand dans la protéine entière RcoM-2 que dans l'hémo-domaine seul.

2.2.4. Dissociation du ligand CO

La cinétique de dissociation du CO de la protéine RcoM-2 a été déterminée en exposant le complexe RcoM-2-CO à un excès de NO et ensuite en suivant les changements spectraux associés au remplacement du CO par NO en fonction du temps, comme le montre la figure

2.5. Pour éviter la dissociation du CO par la lumière, les échantillons ont été conservés dans l'obscurité entre l'enregistrement des spectres.

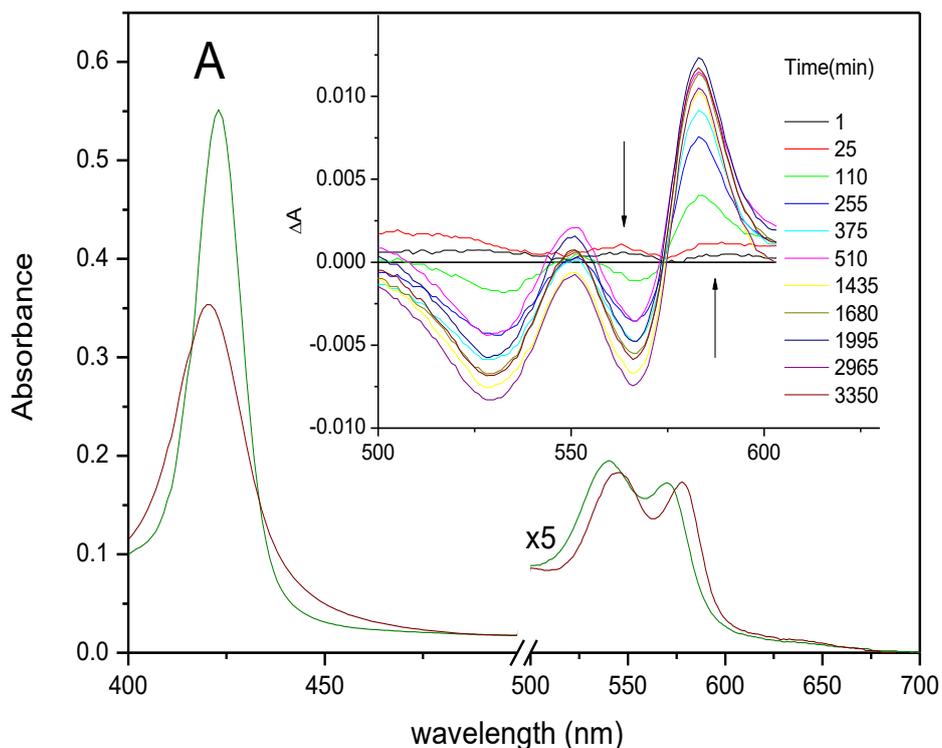


Figure 2.5 : Spectres du complexe RcoM-2-CO (en vert) et après équilibrage avec NO pendant 56 h (en rouge). Inset : Spectres de différence dans la région de la bande Q à différents temps d'exposition au NO.

Les cinétiques du remplacement du CO par NO sont représentées dans la figure 2.6. Elles sont bien décrites par une seule exponentielle avec une constante de temps de 4,4 h, ce qui correspond à un taux de dissociation du CO de $64 \times 10^{-6} \text{ s}^{-1}$.

Cette constante de temps est beaucoup plus courte que l'estimation approximative de l'ordre de jours précédemment avancée par l'équipe pour l'hémo-domaine RcoMH-2 [49]. Nous avons donc utilisé la même méthode pour déterminer avec précision la constante de temps pour RcoMH-2 à 80 h et nous avons constaté qu'elle était en effet près de 20 fois plus longue que celle de la protéine entière.

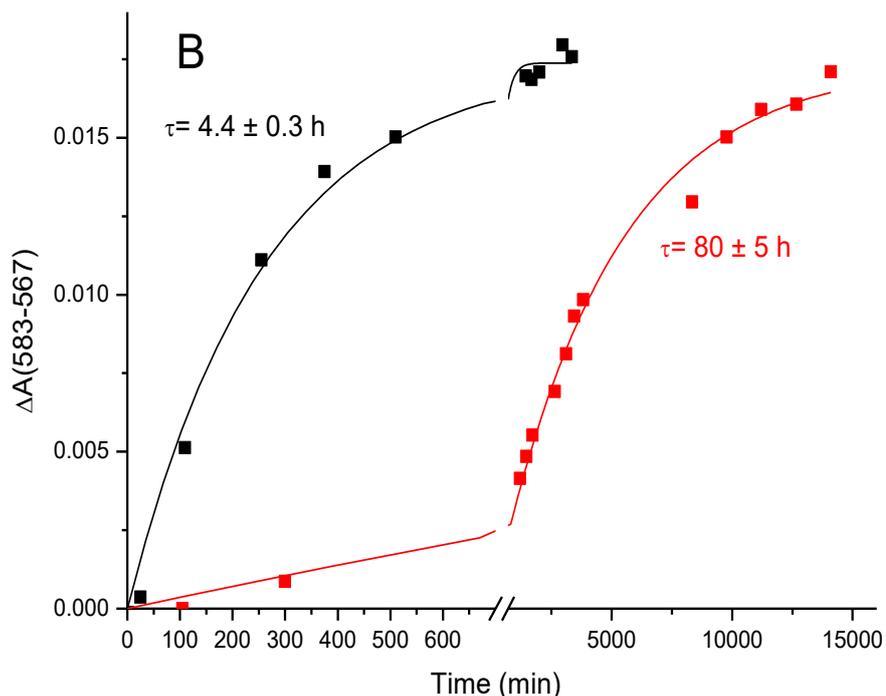


Figure 2.6 : Cinétique du remplacement du CO par NO pour la protéine entière RcoM-2 (en noir) et de l'hémo-domaine RcoMH-2 (en rouge).

Alors que le taux de dissociation du CO de RcoM-2, $64 \times 10^{-6} \text{ s}^{-1}$, est supérieur à celui du domaine de l'hème RcoMH-2, il est inférieur à celui de la plupart des autres senseurs à base d'hème. Cette différence est probablement due à des différences dans la recombinaison géminée hème-CO, qui est beaucoup plus faible dans la plupart des autres senseurs, avec des taux de dissociation plus bas. Il est à noter que le senseur de CO anaérobie CoxA présente aussi un taux de dissociation du CO relativement lent (mais beaucoup moins lent que RcoM-2) et un assez faible rendement d'échappement de CO de quelques pour cent [95].

2.2.5. Dynamique de l'interaction hème-CO : recombinaison géminée du CO

Dans les protéines senseurs à base d'hème, les processus de détection/signalisation sont initialisés par la liaison du ligand gazeux et par la dissociation (thermique) de l'hème. Dans le laboratoire ce dernier processus peut être initialisé par la photodissociation de la liaison hème-CO par une impulsion laser et la recombinaison du ligand peut ensuite être suivie en

temps réel par spectroscopie d'absorption ultra-rapide (coll. M. Vos, LOB). La propagation des perturbations structurales intra-protéique est de l'ordre de $\sim 1,3 \text{ \AA} / 100 \text{ fs}$.

Pour RcoM-2, la cinétique de cette recombinaison est caractérisée par une recombinaison multiphasique avec des constantes de temps de 170 ps et de 660 ps (comme le montre la figure 2.7, courbe rouge). Cette recombinaison est sensiblement plus lente que celle dans l'hémo-domaine isolé (courbe verte).

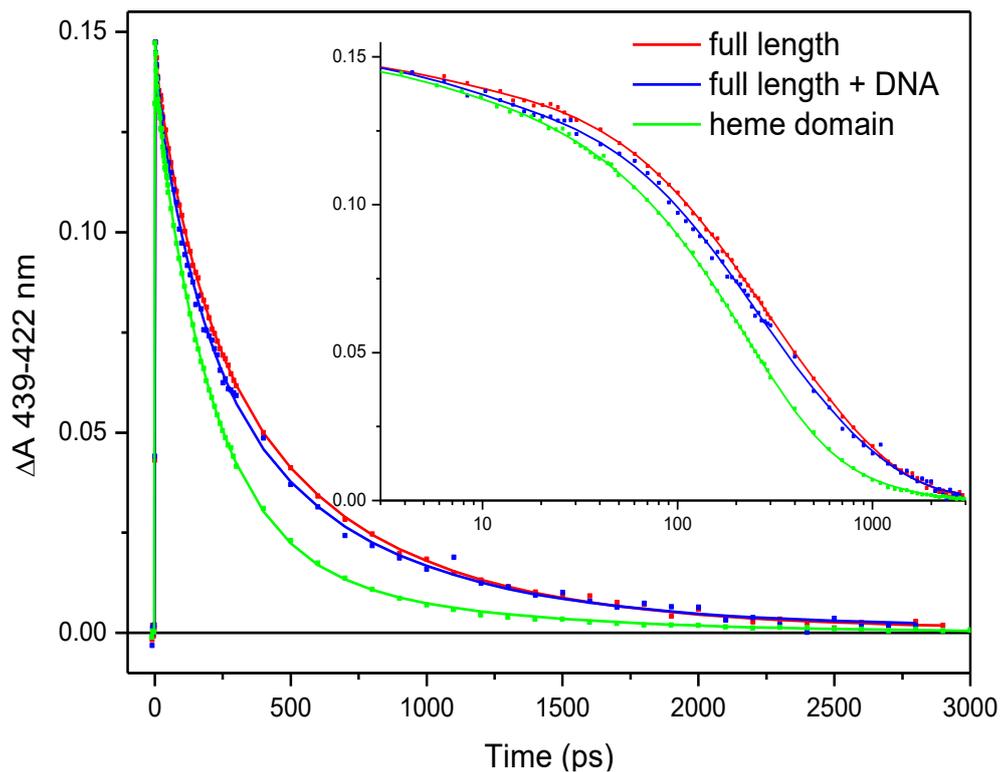


Figure 2.7 : Cinétique de recombinaison du CO de RcoM-2 après photodissociation en l'absence (en rouge) et en présence (en bleu) d'ADN et de l'hémo-domaine RcoMH-2 (en vert). Dans l'encadrement : les données sur une échelle de temps logarithmique.

De plus, pour RcoM-2 une valeur asymptotique faible mais significative de 0,5% a été trouvée, indiquant que du CO peut s'échapper de la poche de l'hème. Cela est en accord avec son rôle de senseur, alors que dans l'hémo-domaine seul, la recombinaison est pratiquement 100% [49]. L'ajout de l'ADN accélère la cinétique de recombinaison du CO légèrement (figure 2.7, courbe bleue).

2.2.6. Liaison de RcoM-2 à l'ADN

Des travaux antérieurs indiquaient que RcoM-1 ne se liait à l'ADN que sous la forme complexée avec du CO, mais l'affinité pour cette liaison ne pouvait pas être déterminée [46]. La technique d'anisotropie de fluorescence (voir Matériel et Méthodes) permet d'étudier de façon quantitative l'affinité d'une protéine pour son ADN cible.

La séquence d'ADN cible utilisée, correspondant à la région « a » du motif de fixation à l'ADN de RcoM-2 [46], a été marquée par le fluorophore Texas Red en 5', et l'anisotropie de fluorescence a été mesurée en fonction de la concentration en protéine RcoM-2 (voir Matériel et Méthodes). Comme le montre la figure 2.8, l'affinité pour la forme RcoM-2-CO (en rouge) est très élevée avec une $K_d < 2$ nM. La saturation en anisotropie observée lors de l'addition de protéines indique une liaison spécifique à la séquence d'ADN. Dans la forme FeIII, aucune liaison à l'ADN n'a eu lieu (en noir).

La constante de liaison à l'ADN observée, $K_d < 2$ nM, est supérieure à celle observée dans les mêmes conditions pour le régulateur de transcription dépendant du NO, DNR (44 nM) [95] et pour l'activateur de transcription anaérobie dépendant du CO, CooA, où des valeurs entre 6 et 16 nM ont été rapportées [50, 95]. Cela est en accord avec une affinité fonctionnelle élevée, similaire à celle observée pour la liaison de RcoM-1 à sa région ADN cible [46] et indique un recrutement efficace de RcoM, spécifique pour le CO.

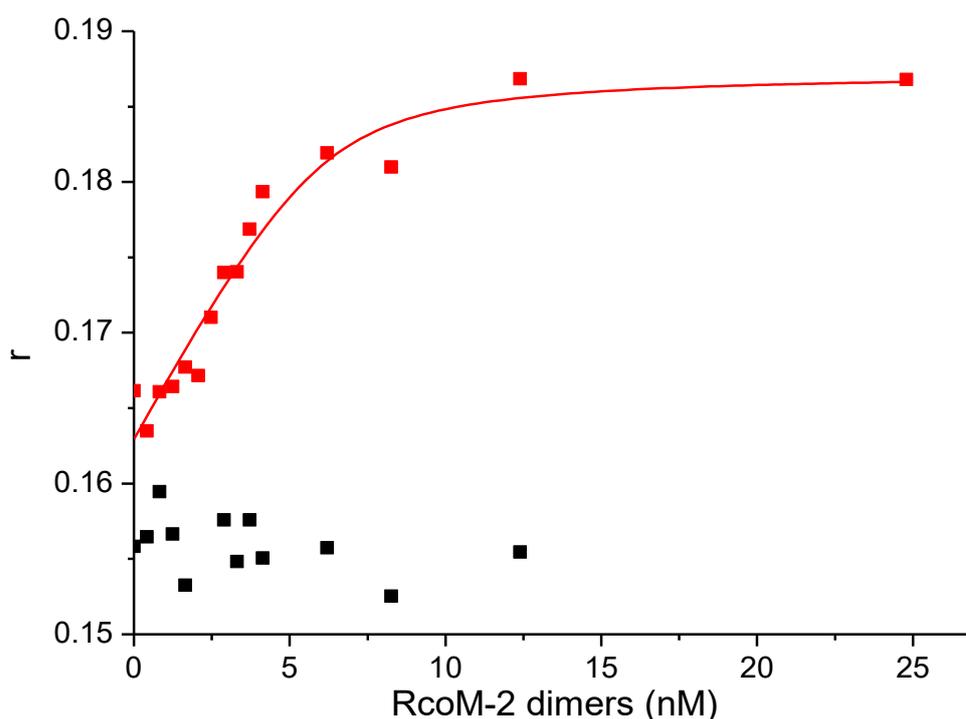


Figure 2.8 : Anisotropie de fluorescence. Fixation du complexe RcoM-2-CO à l'ADN cible (en rouge). La forme FeIII de RcoM-2 ne se lie pas à l'ADN cible (en noir). La concentration d'ADN marqué au Texas Red est de 6.4 nM.

En conclusion, nous avons caractérisé les interactions hème-CO dans le senseur de CO et activateur de transcription RcoM-2. Ces données ont été comparées avec celles de l'hémo-domaine senseur isolé RcoMH-2.

Nous avons trouvé que la protéine entière RcoM-2 fixe du CO avec une affinité effective plus faible que RcoMH-2 (K_d de 4 nM versus 250 pM). Ces valeurs sont parmi les plus hautes affinités connues pour un ligand gazeux et sont déterminées par la poche hémique qui empêche l'échappement du CO. Alors que le rendement de recombinaison géminée du CO est très élevé dans RcoM-2, il est inférieur à celui du domaine hémique isolé, où il est essentiellement de 100%. Le taux de dissociation du CO est 20 fois plus élevé dans la protéine entière. Dans RcoM-2 une fraction du CO (app. 0,5%) peut échapper de la poche d'hème, permettant ainsi à la protéine d'assumer ses fonctions de senseur. Il est à noter que comme senseur aérobie, RcoM-2 doit détecter des concentrations faibles et persistantes en CO en présence d'oxygène. L'échange lent de ligand observé évite également les fluctuations rapides qui pourraient affecter la détection de CO.

Tandis que les propriétés de liaison au ligand du domaine hémique influencent évidemment les propriétés de liaison à l'ADN, il est intéressant à noter que nos résultats démontrent qu'une influence inverse se produit également. On peut supposer que le domaine de liaison à l'ADN exerce une contrainte sur la poche de l'hème senseur qui, en conséquence, devient plus ouverte, donnant lieu à une (faible) probabilité que le CO dissocié de l'hème puisse sortir de la protéine. Les études sur l'origine moléculaire précise des propriétés dynamiques importantes de la poche de l'hème doivent attendre la disponibilité de données structurales sur RcoM.

Nous avons observé un effet faible, mais significatif, de la liaison de l'ADN sur la dynamique de recombinaison du CO dans le site actif (déjà montré dans la figure 2.7). Cette observation laisse penser à une interaction bidirectionnelle entre le domaine senseur et celui de la liaison à l'ADN, et pourrait avoir un rôle de régulation. Un effet qualitativement similaire a été observé pour le senseur à CO CooA [50].

Nous suggérons donc que cette interaction est une caractéristique générale des facteurs de transcription à base d'hème, régulés par un ligand gazeux. Effectivement, les promoteurs régulés par RcoM-2 sont toujours actifs en présence du CO.

Chapitre III

Mécanismes de détection du CO chez les mycobactéries

Comme indiqué auparavant, le CO est un gaz toxique qui peut inhiber la respiration aérobie des bactéries par sa fixation sur des constituants de la chaîne respiratoire (en particulier la cytochrome oxydase). À faibles quantités, le CO peut agir comme molécule de signalisation et affecter la régulation de gènes. Cette signalisation nécessite la détection rapide du CO et des réponses cellulaires adaptées, qui sont encore peu comprises.

À ce jour, peu de senseurs à CO bactériens sont connus. Le régulateur de transcription CO-dépendant CooA anaérobie [33, 39, 41] est le senseur à base d'hème le mieux caractérisé. Le senseur à CO RcoM-2 aérobie, déjà caractérisé dans le chapitre 2, est capable de lier du CO avec une affinité très élevée sans lier l'oxygène, en accord avec une détection du CO en présence d'oxygène.

La signalisation par le CO présente un intérêt particulier dans des bactéries pathogènes intracellulaires, comme p. ex. *M. tuberculosis*, qui doivent surmonter les mécanismes de défense chez l'hôte, dont le CO, mais leurs réponses au CO sont peu connues. Il est intéressant de noter que les mycobactéries sont des bactéries aérobies obligatoires qui peuvent résister à des concentrations élevées de CO, suggérant des mécanismes spécifiques afin de résister au CO et combattre sa toxicité en présence d'oxygène. Dans ces mécanismes, un nouveau gène nommé *cor* (« CO-resistance ») [89] récemment identifié par criblage génétique, mais non caractérisé, pourrait jouer un rôle.

3.1. Introduction

Afin de mieux comprendre les mécanismes de détection et la réponse mycobactérienne au CO, nous avons d'abord caractérisé la protéine Cor (Rv1829) biochimiquement et spectroscopiquement pour obtenir des informations sur son mécanisme fonctionnel. De façon inattendue nous avons découvert que Cor était capable de lier d'une manière spécifique et stoechiométrique un cofacteur hémique, suggérant une fonction de détection directe du CO. En associant la bio-ingénierie pour modifier la protéine par mutagenèse dirigée aux techniques spectroscopiques de fluorescence et d'absorption, un résidu histidine a été identifié, susceptible de lier l'hème.

Le gène *rv1829* (*cor*) fait partie d'un cluster de gènes conservé et peu caractérisé chez les mycobactéries, avec plusieurs régulateurs de transcription prédits. Basé sur les études *in silico* et en utilisant l'anisotropie de fluorescence et l'EMSA nous avons étudié un régulateur transcriptionnel prédit, le Rv1828 (voir 1.4.4.e). Nous avons également pu montrer que Cor présente une activité de liaison à l'ADN qui augmente en présence d'hème et de CO, suggérant une potentielle fonction régulatrice, dépendante de CO. Ces études ont également montré que le régulateur mycobactérien Rv0081, induit en réponse aux changements de concentration gazeuse, peut se lier sur la région régulatrice prédite de *cor*.

Ces résultats suggèrent une régulation transcriptionnelle complexe des gènes du cluster *cor*, qui reste à explorer en plus de détail. Des analyses de cristallogénèse ont également été effectuées pour Cor et Rv0081.

Finalement, un récent projet vise à élucider le rôle physiologique de Cor dans la résistance contre le CO lors de cultures du modèle non-pathogène *Mycobacterium smegmatis*.

3.2. Caractérisation de la protéine Cor

3.2.1. Caractérisation biochimique de Cor

a) Surexpression inducible de Cor

cor (*rv1829*) est un gène de fonction inconnue chez *M. tuberculosis* qui code pour une protéine de 164 acides aminés avec la séquence protéique suivante, qui comprend deux histidines (en bleu) : His58 et His70

N ter-

```
      10          20          30          40          50
MGEVRVVGIR VEQPQNQPVL LLREANGDRY LPIWIGQSEA AAIALEQQGV
      60          70          80          90          100
EPPRPLTHDL IRDLIAALGH SLKEVRIVDL QEGTFYADLI FDRNIKVSAR
      110         120         130         140         150
PSDSVAIALR VGVPIYVEEA VLAQAGLLIP DESDEEATTA VREDEVEKFK
      160
EFLDSVSPDD FKAT          -C ter
```

Après clonage du gène *cor* dans le vecteur d'expression pQE80L, transformation dans la souche d'expression BL21 DE3, et induction par l'isopropyl-1-thiol-beta-D-galactopyranoside (IPTG), la protéine Cor a été purifiée par chromatographie d'affinité sur les colonnes Protino-Ni-TED (voir Matériels et Méthodes).

La concentration de la protéine Cor purifiée a été déterminée par spectroscopie d'absorption à 280 nm ou par dosage colorimétrique Bradford. La masse moléculaire de Cor a été vérifiée par migration sur SDS-PAGE 10%. Le gel d'électrophorèse a été coloré au Bleu de Coomassie, démontrant une bande de masse moléculaire apparente de 18 kDa (voir figure 3.1)

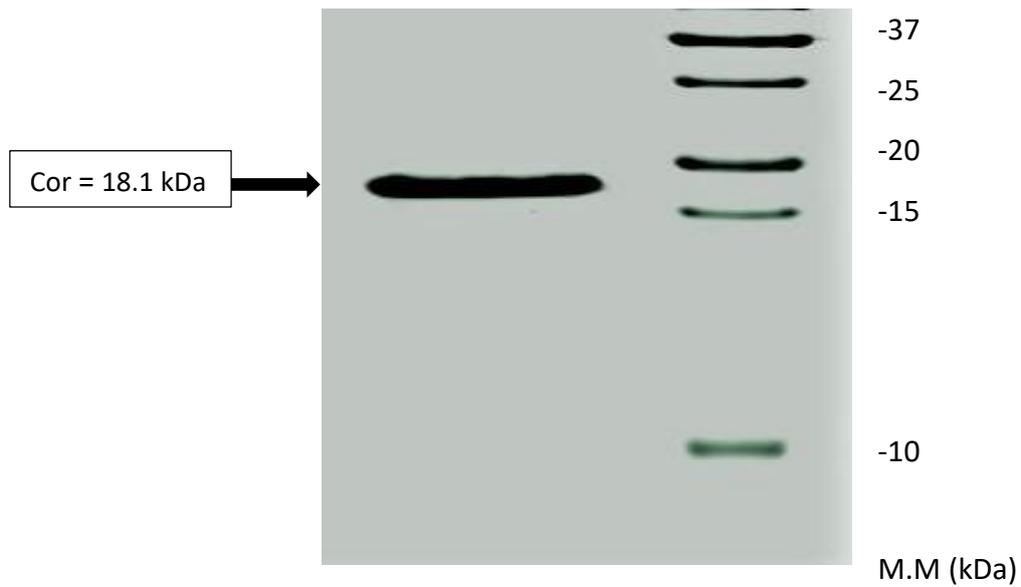


Figure 3.1 : SDS-PAGE 10% coloré au Bleu de Coomassie. La protéine Cor montre une masse moléculaire apparente de 18.1 kDa (puits à gauche). Marqueur moléculaire à droite.

À noter que Cor a été purifiée sous forme d'apoprotéine, sans cofacteur, comme le montre le spectre d'absorption pris après purification (voir figure 3.2).

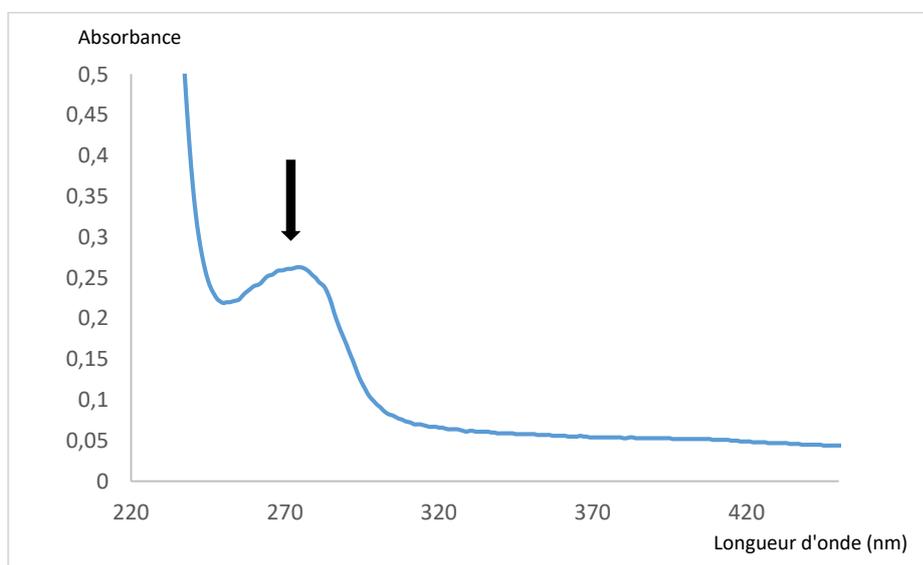


Figure 3.2 : Spectre d'absorption de Cor directement après purification, montrant un pic d'absorption vers 270-280 nm sans absorption dans le visible.

b) Cor est un dimère très stable

Dans la littérature, il avait été suggéré que Cor pourrait être un dimère [89], et nous avons observé que Cor semble former des dimères très stables, même détectables sur SDS-PAGE après dénaturation, avec une bande qui apparaît vers 36 kDa (gel non montré). Afin de caractériser l'état de conformation de la protéine Cor, nous avons réalisé des expériences d'ultracentrifugation analytique (UCA).

L'ultracentrifugation analytique permet de caractériser le comportement des molécules natives en solution, soumises à une force centrifuge. Nous pouvons donc mesurer la vitesse de sédimentation, ce que nous permet d'obtenir le coefficient de sédimentation, la forme, et l'état associatif (monomère ou dimère) d'une macromolécule.

Les résultats de ces expériences réalisées en collaboration avec la « Plateforme de Mesures d'Interaction des Macromolécules » (PIM), I2BC, Université Paris Sud), ont montré que la grande majorité (92 à 94%) des espèces Cor absorbant à 280 nm possède un coefficient de sédimentation 2.5S (voir figure 3.3).

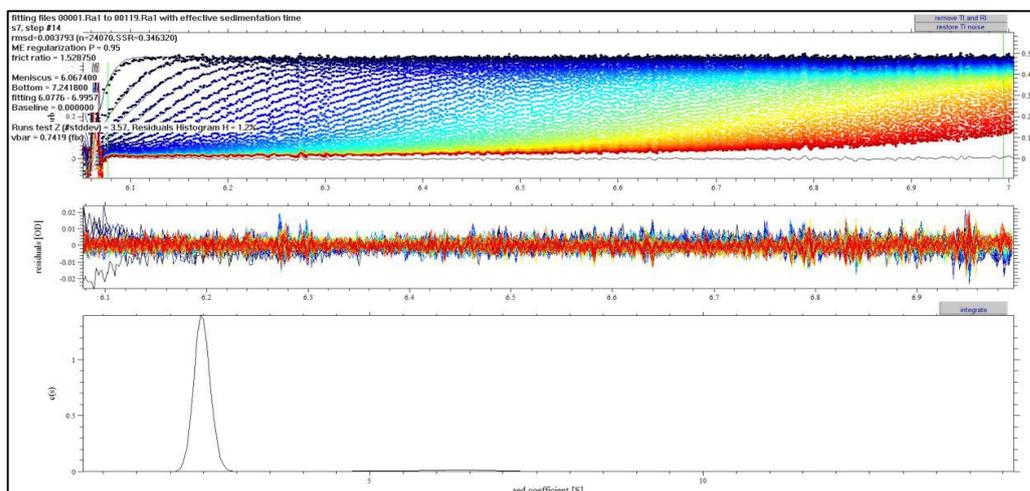


Figure 3.3 : Ultracentrifugation analytique de Cor (Rv1829) montrant un coefficient de sédimentation de 2.5S pour la majorité d'espèces Cor absorbant à 280 nm (collaboration PIM, I2BC, Université Paris Sud).

Le Tableau 3.1 récapitule des expériences à différentes concentrations de Cor et montre des résultats similaires : quelle que soit la concentration de Cor, une espèce

majoritaire (représentant 92 à 94% de l'absorbance totale à 280 nm) est observée, avec un coefficient de sédimentation de 2.5S. Des traces d'espèces avec un plus haut coefficient de sédimentation sont observées également.

	Cor à 40.1 μ M		Cor à 59.3 μ M		Cor à 97 μ M	
	Coef. de sédimentation	% approximatif de représentation	Coef. de sédimentation	% approximatif de représentation	Coef. de sédimentation	% approximatif de représentation
1 ^{ère} population d'espèce	2.5S ($S_{20,W} = 2.6S$)	94	2.5S ($S_{20,W} = 2.6S$)	92	2.5S ($S_{20,W} = 2.6S$)	92
2 ^{ème} population d'espèce	Au-delà de 4S	6	Au-delà de 4S	8	Au-delà de 4S	8

Tableau 3.1: Tableau récapitulatif des espèces protéiques majoritaires obtenues par ultracentrifugation analytique de Cor à 40.1, 59.3 et 97 μ M.

Ces analyses de vitesse de sédimentation permettent de constater la présence d'une espèce protéique majoritaire avec un coefficient de sédimentation de 2.5S, qui correspond à un dimère de Cor avec une forme légèrement allongée ($f/f_0=1.44$), et très stable (comme la bande de dimérisation apparaît même sur SDS-PAGE). Alternativement, cette espèce pourrait correspondre à un trimère avec une forme très fortement allongée ($f/f_0=1.88$) ; néanmoins c'est peu probable.

c) Modélisation de la protéine Cor

Le gène *cor* semble être un gène conservé contenant un domaine de fonction inconnue (DUF151) [89]. L'organisation génomique de la région *cor* se retrouve dans des mycobactéries et des espèces de *Rhodococcus*, *Streptomyces* et *Thermotoga maritima* et d'autres espèces (comme le montre la figure 1.26), suggérant des fonctions partagées. Nous ne disposons pas

de la structure tridimensionnelle de Cor, et dans la PDB se trouve une seule structure d'une protéine à DUF 151, déterminée chez *T. maritima* (PDB 1SJ5). Cette protéine présente une homologie de séquence de 33.3% avec celle de Cor. La structure de cet homologue est nouvelle et fournit donc peu d'informations sur sa fonction.

Nous avons créé un modèle de Cor (en collaboration avec J-C. Lambry, IR INSERM, LOB), en se basant sur la structure de la protéine Duf151 de *T. maritima* (voir 5.7).

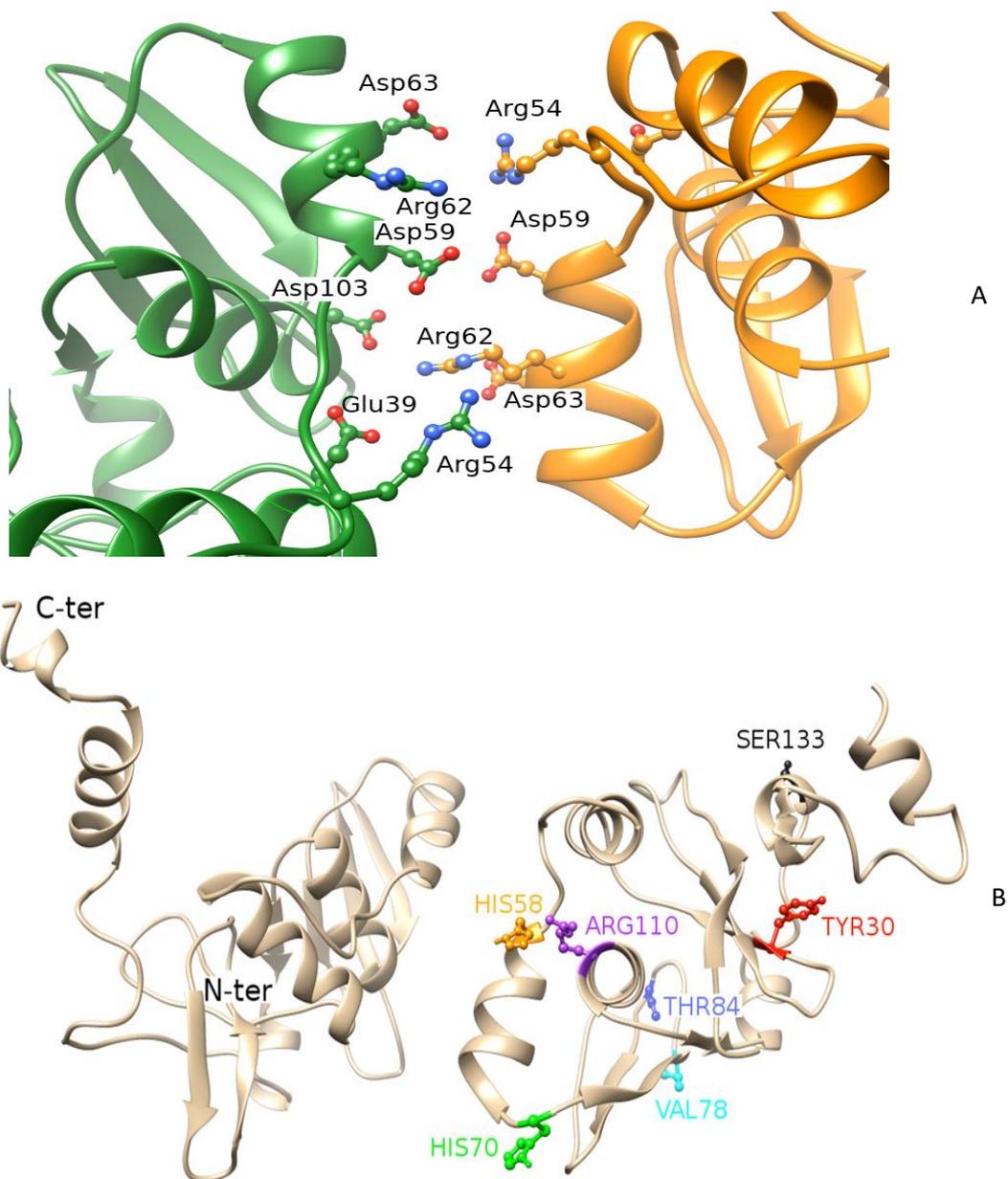


Figure 3.4 : Modèle du dimère de Cor basé sur la structure de la protéine Duf151 chez *T. maritima* (PDB 1SJ5). **(A)** Interface des deux monomères montrant des résidus chargés. **(B)** Représentation des deux histidines, His58 et His70, candidats potentiels de liaison de l'hème (voir partie 3.2.3).

Comme le montre la figure 3.4A, l'interface de dimérisation de Cor présente plusieurs résidus chargés (surtout dans des hélices α), suggérant un effet électrostatique, en accord avec la formation d'un dimère stable. La figure 3.4B montre les deux histidines His58 et His70, candidats potentiels d'une liaison de l'hème, auxquelles nous nous sommes intéressés dans cette thèse (voir 3.2.3).

3.2.2. Caractérisation spectroscopique de Cor

a) Liaison Cor-hème

Comme la fonction précise et le mécanisme fonctionnel de Cor dans la détection du CO n'étaient pas compris, nous avons émis l'hypothèse que Cor pourrait potentiellement lier un cofacteur et donc testé la capacité de Cor à fixer de l'hémine après purification. De façon inattendue et différente des données de la littérature [89], nous avons trouvé que Cor peut fixer de façon spécifique un cofacteur hémique stoechiométriquement (un hème par monomère) et peut ainsi à priori détecter du CO de façon *directe*. La figure 3.5 montre les spectres d'absorption de la titration Cor-hémine.

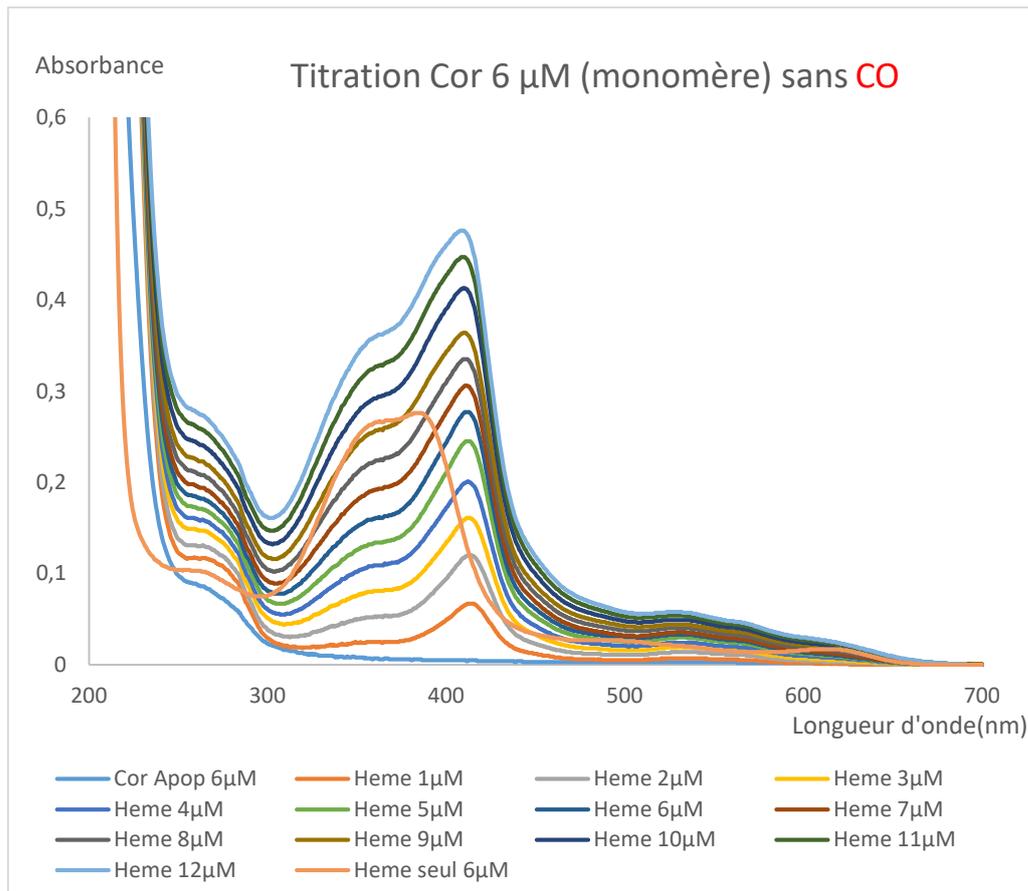


Figure 3.5 : Spectres d'absorption de titration Cor-hémine. Sont montrés : le spectre de la protéine Cor après purification (Cor apoprotéine) (6 μM en monomères) et après l'ajout de l'hémine à différentes concentrations entre 1 μM et 12 μM . L'hémine seule à 6 μM sert comme contrôle. Le temps d'attente après chaque ajout est d'une heure. Le pic du complexe Cor-hémine apparait vers 414 nm, le pic de l'hémine seule et libre vers 395 nm (expérience réalisé sous l'air). Cette forme spectrale suggère que l'hème est sous forme hydroxy-Met, avec une histidine comme ligand protéique [96].

À basses concentrations d'hémine (< 5 μM), après équilibrage d'une heure, un pic est observé à 414 nm. En ajoutant de l'hémine, nous observons que la bande d'absorption commence graduellement à s'élargir et après l'ajout de 5 μM de l'hémine, le pic décale vers le bleu (voir figure 3.6 au-dessous) et l'ajout mène surtout à l'augmentation de l'hème libre (maximum à 395 nm). Cela indique que la protéine Cor monomère commence à être saturée à partir d'ajout quasi-stoechiométrique de l'hème (~ 1 hème/monomère).

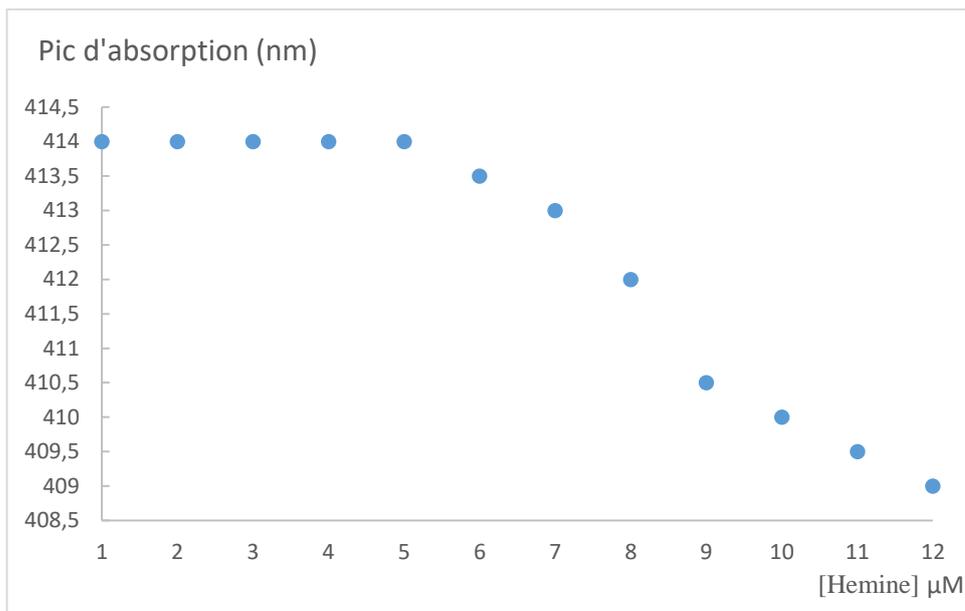


Figure 3.6 : Plot du maximum de la bande Soret en fonction de la concentration de l'hémine, tiré des données de la figure 3.5. La concentration de Cor était de 6 μM en monomères. La courbe est indicative pour la saturation de Cor à partir de l'ajout d'hème quasi-stoechiométrique.

b) Cor fixe du CO et lie de l'hème de façon spécifique

Suite aux titrations de l'hémine, nous avons observé que Cor lie du CO de façon directe, comme le montrent les spectres d'absorption du complexe Cor-hème ligandé au CO dans la figure 3.7a. Ces expériences ont été réalisées avec un complexe de 5 μM Cor (en dimères) et 10 μM hémine, après dégazage et ajout de dithionite de sodium, puis équilibrage avec 1 atm de CO. Les spectres ont été mesurés directement après équilibrage et ne changent plus ensuite. Le spectre de la forme réduite est typique pour un complexe 6-c avec deux acides aminés comme ligands [97]. Le spectre après ajout du CO est typique pour un complexe hème-CO où l'hème est liée à la protéine de façon spécifique, avec un pic particulier à 420 nm. Ainsi, le CO remplace un des ligands internes.

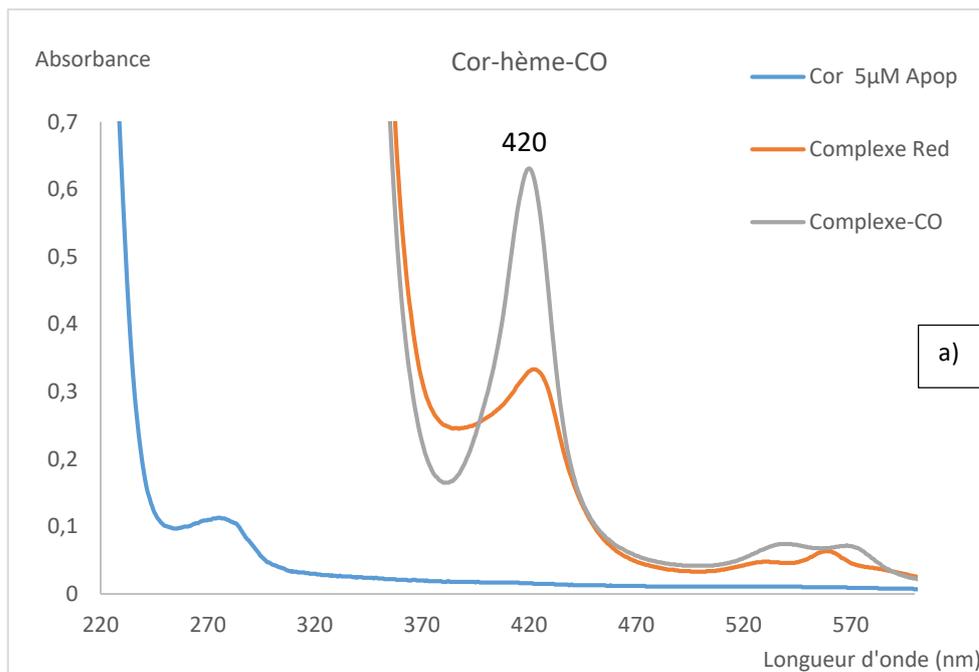


Figure 3.7a : Spectres de Cor (5 μM en dimères) à l'état d'apoprotéine (en bleu), de complexe avec 10 μM hème réduit par dithionite de sodium (en orange), et après exposition au CO (en gris). Le pic à 420 nm indique la fixation du CO.

La figure 3.7b montre des expériences de titration de l'hème en présence de réducteur et de CO. Les spectres ont des formes quasi-identiques, avec un pic à 420 nm, jusqu'à l'ajout d'hème/monomère Cor en quantité stoechiométrique. Ensuite le pic s'élargit vers le bleu correspondant au complexe hème-CO libre ou à une liaison non-spécifique. Le changement net de pente dans l'absorption dans le côté bleu de la bande (405 nm) correspond à l'ajout stoechiométrique d'hème (figure 3.7c) et confirme la liaison directe d'un hème par monomère en présence de CO.

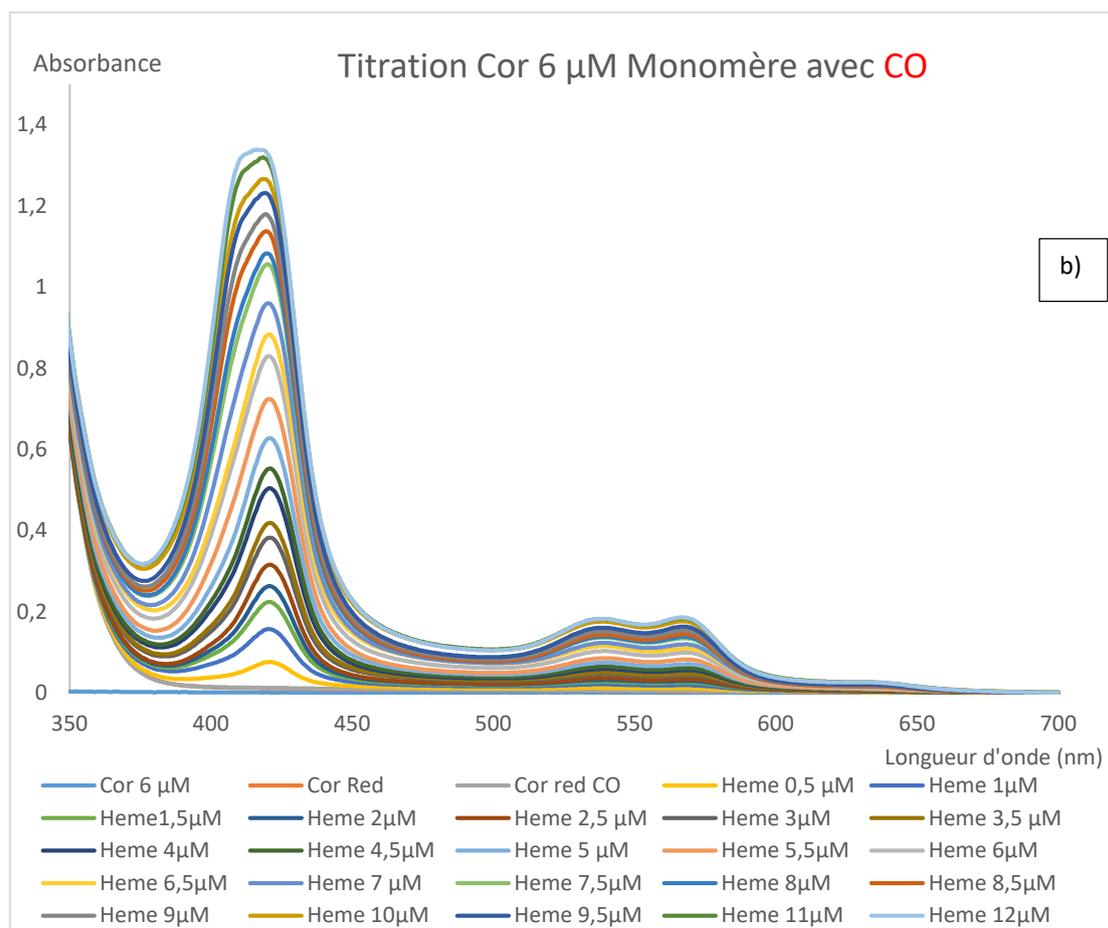


Figure 3.7b : Spectres de titration d'hémine dans des solutions de Cor (6 μM en monomères) en présence de dithionite et de CO.

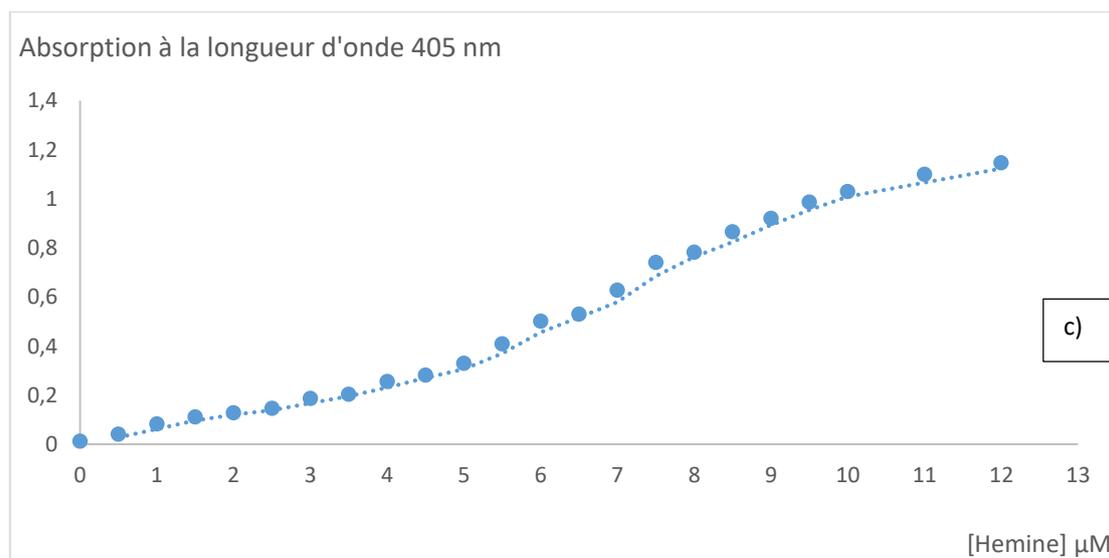


Figure 3.7c : Absorption à 405 nm en fonction de la concentration d'hémine, les valeurs sont tirées des spectres de la figure 3.7b. A noter que le changement de pente est autour de 6 μM d'hémine qui correspond à la concentration de Cor en monomères.

Pour confirmer la spécificité de la fixation de l'hème à Cor, j'ai testé également la liaison de l'hème à l'albumine de sérum bovin (BSA) dans la présence du CO. J'ai observé une fixation du CO à l'hème de façon non-spécifique à 415 nm comme le montrent les spectres d'absorption dans la figure 3.8A. Un deuxième contrôle a été testé avec de l'hème libre qui donne lieu à un pic d'absorption à 405 nm (figure 3.8B).

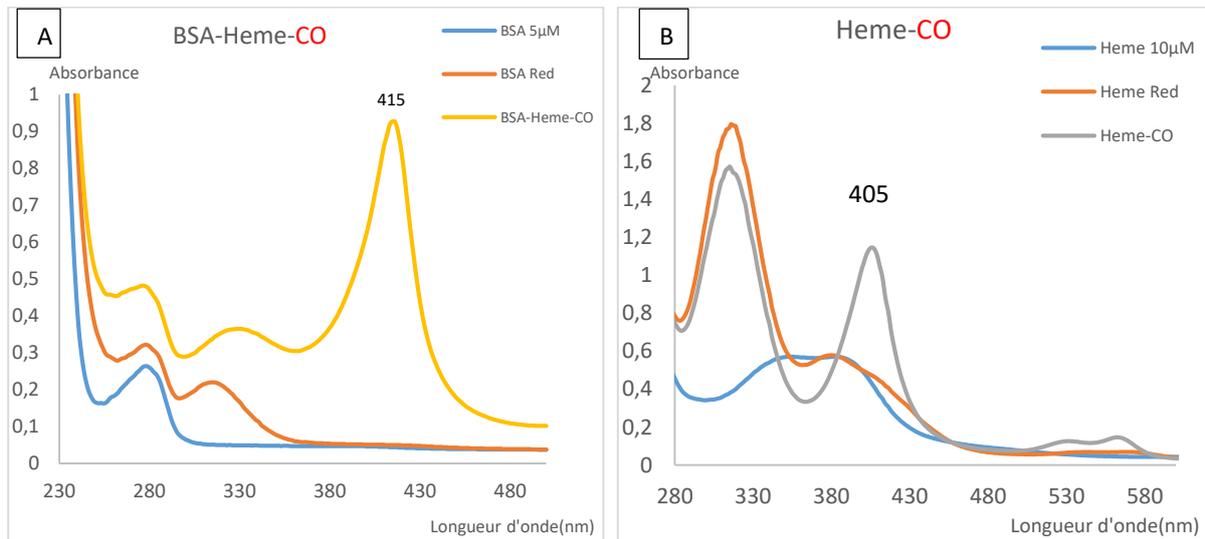


Figure 3.8 : Spectres d'absorption des contrôles. **A)** Les spectres à gauche correspondent à la BSA pure 5 µM (bleu), en présence de dithionite (orange) et en présence de l'hème et du CO (jaune) qui donne lieu à un pic à 415 nm (fixation non spécifique). **B)** Les spectres à droite correspondent à l'hème libre 10 µM (bleu), l'état réduit (orange), et l'hème en présence de CO donne lieu à un pic d'absorption à 405 nm.

Dans l'ensemble, la formation d'un complexe Cor-hème-CO, où l'hème est liée de façon spécifique à la protéine, implique que Cor peut détecter du CO de façon *directe* par un hème.

c) La présence du CO accélère la liaison de l'hème à Cor

Ensuite, pour mieux comprendre quel rôle joue la présence du CO sur la liaison Cor-hème, nous avons testé la fixation de l'hème en fonction du temps en absence du CO et comparé les résultats avec les spectres de liaison du CO, qui se fixe directement. La figure 3.9 montre les spectres d'absorption qui ont été pris toutes les 5 minutes. La liaison Cor-hème est très lente en absence du CO et peut durer jusqu'à une heure. La liaison apparait avec le pic allant de 395 (hème libre) à 414 nm (1h) (voir figure 3.9).

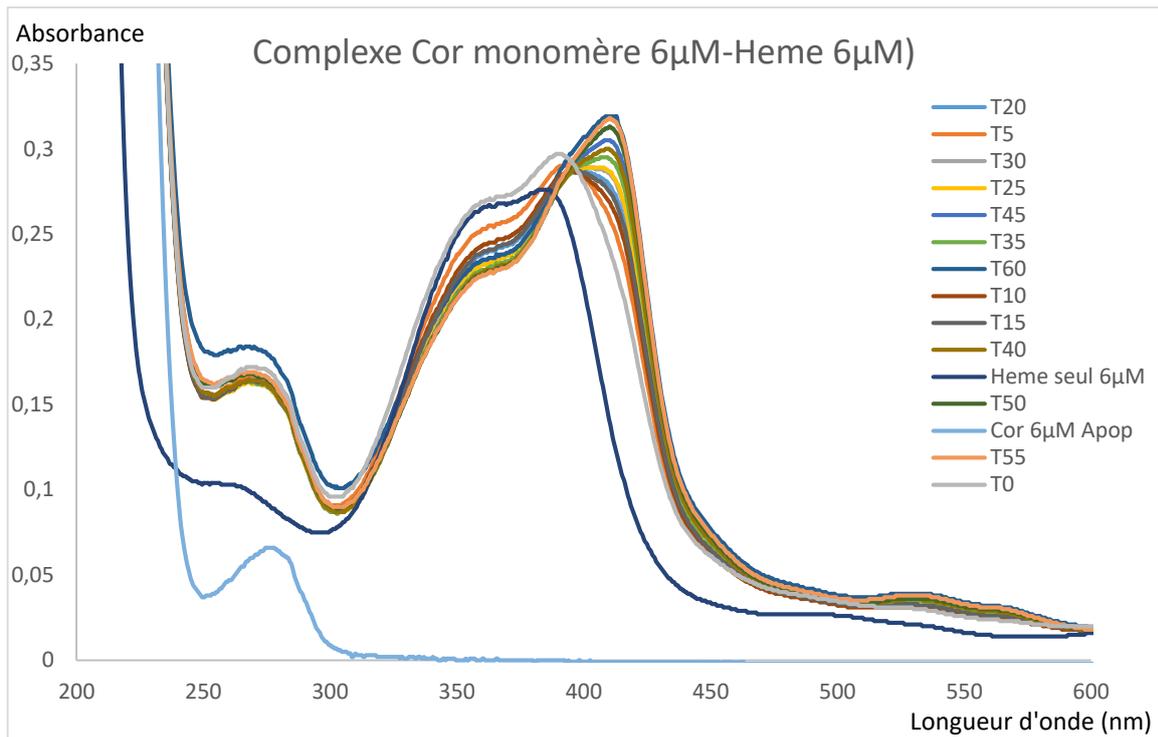


Figure 3.9 : Spectres d'absorption de la liaison Cor-hème sous l'air à température ambiante, en absence du CO. Les spectres sont pris à 5 minutes d'intervalle. La liaison de l'hème est lente et sature après ~une heure.

L'ensemble de ces résultats montre que Cor peut lier un cofacteur hème stoechiométriquement (1 hème/monomère). La fixation de l'hème est spécifique et la présence du CO accélère la liaison Cor-hème (figure 3.7a). Ces observations suggèrent que la protéine Cor peut détecter du CO de façon directe.

Ces résultats constituent une première étape pour nous aider à mieux comprendre comment Cor peut détecter du CO. Pour obtenir plus de détails sur l'interaction hème-CO, nous nous sommes intéressés à l'aspect dynamique caractérisant l'interaction du ligand diatomique CO avec le complexe Cor-hème.

d) Dynamique de la liaison du ligand diatomique CO à l'hème

Dans le but d'étudier l'interaction du CO avec le cofacteur hémique associé à la protéine Cor, nous avons réalisé la photodissociation du CO, suivie d'une détection par spectroscopie d'absorption ultrarapide (expériences faites en collaboration avec M. Vos, LOB).

Une molécule de CO dissociée de l'hème peut se relier au même hème rapidement (dans l'échelle picoseconde/nanoseconde) par une recombinaison dite géminée, ou alors sortir de la protéine. Dans le dernier cas, l'hème peut lier une autre molécule de CO par recombinaison dite bimoléculaire (dans l'échelle microseconde/milliseconde). Pour cette réaction la vitesse est proportionnelle à la concentration de CO en solution. Grâce à la spectroscopie résolue en temps, les vitesses de ces recombinaisons peuvent être étudiées. Ici nous avons étudié la recombinaison géminée par spectroscopie d'absorption ultrarapide.

Dans nos expériences de photodissociation, le ligand CO se relie très rapidement à l'hème associé à la protéine Cor. Cette recombinaison est multiphasique et très efficace et concerne $\approx 80\%$ des molécules CO. La cinétique de cette recombinaison géminée rapide est montrée dans la figure 3.10.

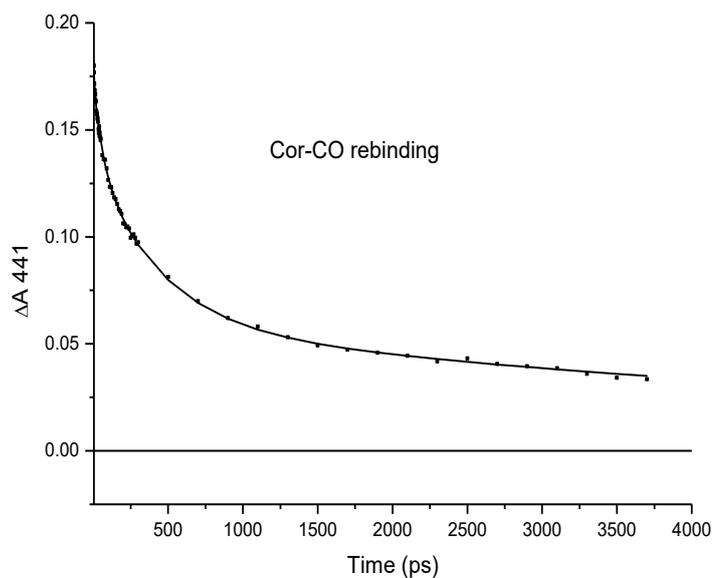


Figure 3.10 : Cinétique de recombinaison de CO à l'hème liée au Cor suivie à 441 nm. L'ajustement correspond à une fonction avec des constantes de temps de 60 ps, 440 ps et 7 ns (ensemble $\sim 80\%$) et un asymptote ($\sim 20\%$).

Cette recombinaison géminée substantielle et multiphasique du CO est qualitativement similaire à celle observée dans d'autres senseurs à CO étudiés dans le laboratoire comme le senseur à CO CooA [95] et l'activateur de transcription CO-dépendant RcoM-2, étudié dans cette thèse (voir chapitre 2). Notamment, dans ces senseurs, la

recombinaison géminée est beaucoup plus étendue que dans des hémoprotéines de stockage/transfert de ligands (myoglobine et hémoglobine [44]), où il est quasi-inexistant.

e) Recombinaison du ligand hémique

Dans un deuxième temps nous avons suivi la recombinaison des ligands internes de Cor, très probablement des acides aminés impliqués dans la liaison de l'hème. Une expérience de photodissociation en absence du CO a été réalisée sur le complexe Cor-hème réduit et la recombinaison du ligand interne Cor-hème a été suivie en temps réel. Le temps de cette recombinaison était de 8 picosecondes, comme le montre la figure 3.11.

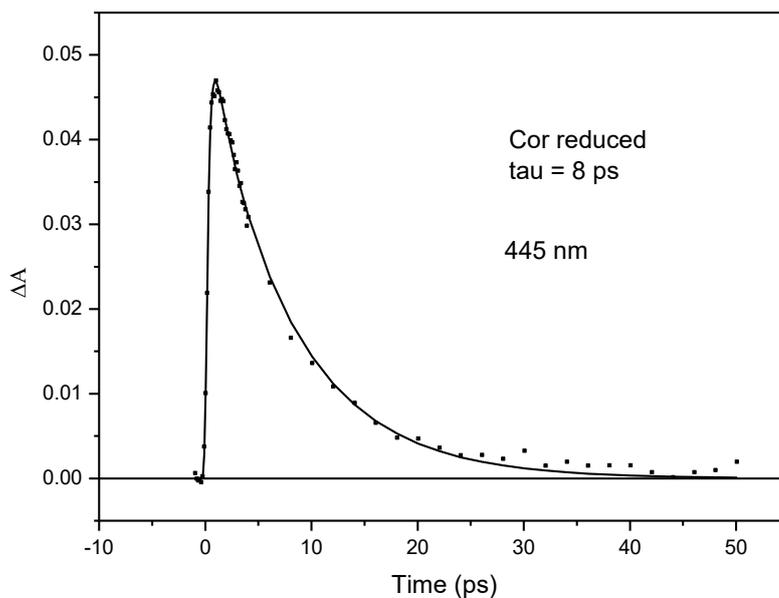


Figure 3.11 : Cinétique de recombinaison du ligand interne de l'hème à la protéine Cor après photodissociation, mesurée à 445 nm avec une constante de temps $\tau=8$ ps.

Ce temps de recombinaison interne de 8 ps est typique pour la recombinaison des résidus internes (acides aminés) dans les protéines senseurs à hème 6-coordonnées (comme par exemple CoxA ou EcDos) [97].

Ces résultats nous ont incités à essayer d'identifier les résidus internes de Cor qui pourraient être impliqués dans la liaison de l'hème.

3.2.3. Études des mutants Cor-H58A et Cor-H70A

a) Expression et purification de Cor-H58A et Cor-H70A

Comme indiqué auparavant dans la partie 3.2.1.a, la séquence protéique de Cor contient deux histidines, His58 et His70 qui constituent des candidats potentiels pour une liaison de l'hème à la protéine (les deux histidines sont représentées dans le modèle de Cor déjà montré dans la figure 3.4B). Pour identifier quel acide aminé pourrait être responsable de la liaison avec l'hème, nous avons muté les deux histidines en alanines par mutagenèse dirigée, dans le but d'affecter la liaison de l'hème à Cor. Ensuite des caractérisations des mutants H58A et H70A ont été réalisées.

Après construire les mutants Cor-H58 et Cor-H70 et purifier les protéines mutantes, leurs concentrations ont été déterminées par spectroscopie d'absorption. Et comme attendu, les deux mutants possèdent la même masse moléculaire apparente sur SDS-PAGE que le type sauvage de Cor (voir figure 3.12).

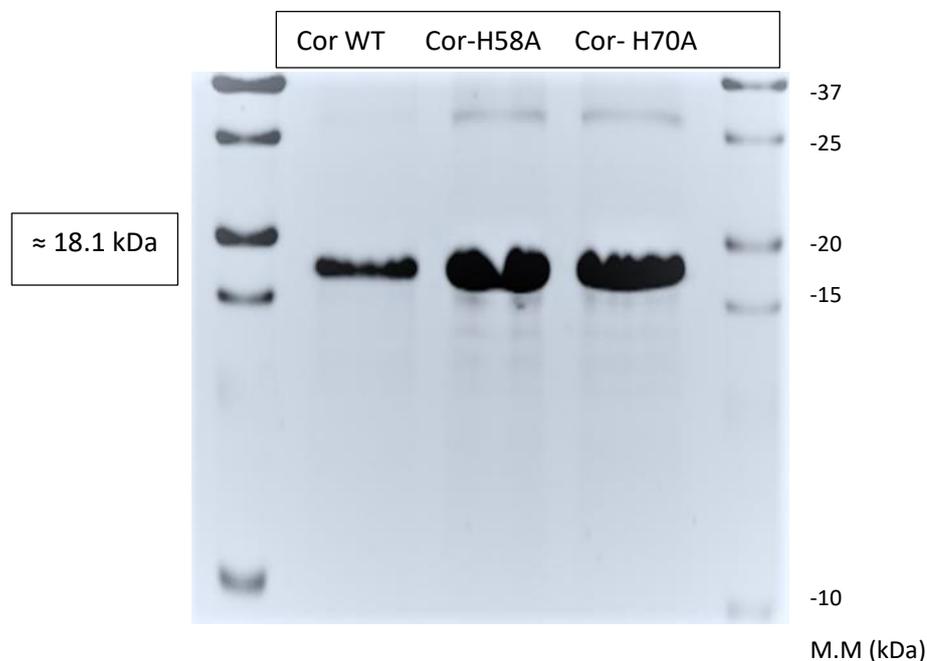


Figure 3.12 : SDS-PAGE 10% coloré au Bleu de Coomassie. Cor type sauvage (WT) (à gauche) comme contrôle et les mutants Cor-H58A (milieu) et Cor-H70A (à droite). Les protéines montrent des masses moléculaires apparentes attendues.

b) Caractérisation spectroscopique des mutants

Dans le but de tester l'implication de l'acide aminé histidine (H) dans la liaison de l'hème et la détection du CO, des expériences de titration protéine-hème en présence et absence du CO ont été réalisées pour les deux mutants Cor-H58A et Cor-H70A

b.1) Liaison d'hème dans les mutants en absence du CO

Pour déterminer s'il y a des différences entre Cor WT et les deux mutants dans la liaison de l'hème, les spectres d'absorption suite à l'ajout d'hémine ont été suivis en absence et présence du CO. Figure 3.13 montre les variations de pic d'absorption reflétant la liaison protéine-hème en fonction du temps. Le mutant Cor-H58A montre une liaison de l'hème proche de celle de WT, tandis que le mutant Cor-H70A lie beaucoup moins de l'hème, indiquant une implication de cet acide aminé dans la liaison de l'hème à la protéine Cor.

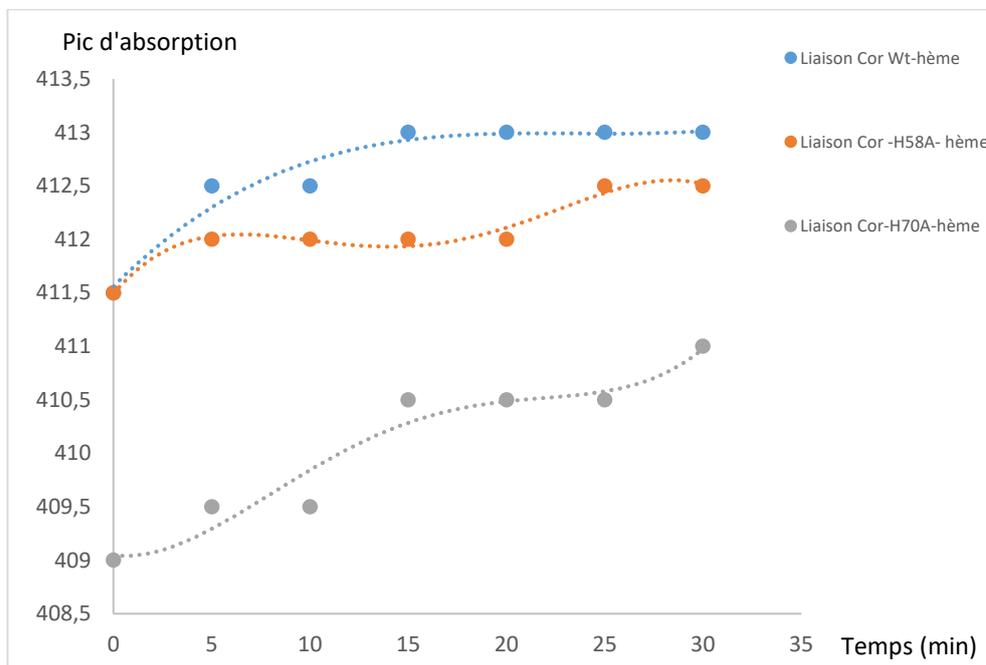


Figure 3.13 : Maximum d'absorption en fonction du temps après ajout d'hémine pour les trois protéines Cor WT (bleu), Cor-H58A (orange) et Cor-H70A (gris).

b.2) Liaison d'hème en présence du CO

Dans cette expérience réalisée en présence du CO, la fixation protéine Cor-hème est directe et plus rapide qu'en l'absence du CO. Cela est en accord avec nos observations antérieures (comme déjà montré dans la figure 3.9 pour le type sauvage, la liaison Cor-hème est très lente en absence du CO et quasi-immédiate en sa présence dans la figure 3.7a). Une différence de la liaison à l'hème est constatée entre la protéine Cor WT et le mutant Cor-H70A, comme le montrent les spectres d'absorption dans la figure 3.14. Le pic étroit à 420 nm est indicatif de la liaison Cor WT-hème-CO) et similaire à celui observé pour Cor H58A-hème-CO (419 nm). Le maximum du pic d'absorption pour le mutant Cor H70A-hème-CO est vers 417 nm et beaucoup moins étroit (figure 3.14). Ces caractéristiques se rapprochent des valeurs d'une fixation non-spécifique vers 415 nm (comme montré dans la figure 3.8a pour la protéine BSA).

Ces observations indiquent que l'histidine His70 joue un rôle dans la fixation de l'hème à Cor. Basé sur nos expériences, le rôle de His58 est moins clair, avec des caractéristiques proches de celles du type sauvage. La construction d'un double mutant His58/His70 est envisagée pour déterminer s'il y aura un effet plus net. En vue de la modélisation préliminaire de Cor présenté dans la figure 3.4B, la détermination de His70 comme un probable site de fixation de l'hème suggère que l'hème se fixe à proximité de la surface de la protéine. Cela pourrait expliquer la labilité de l'hème dans la forme oxydée où l'interaction avec la protéine implique un seul liaison hème-acide aminé.

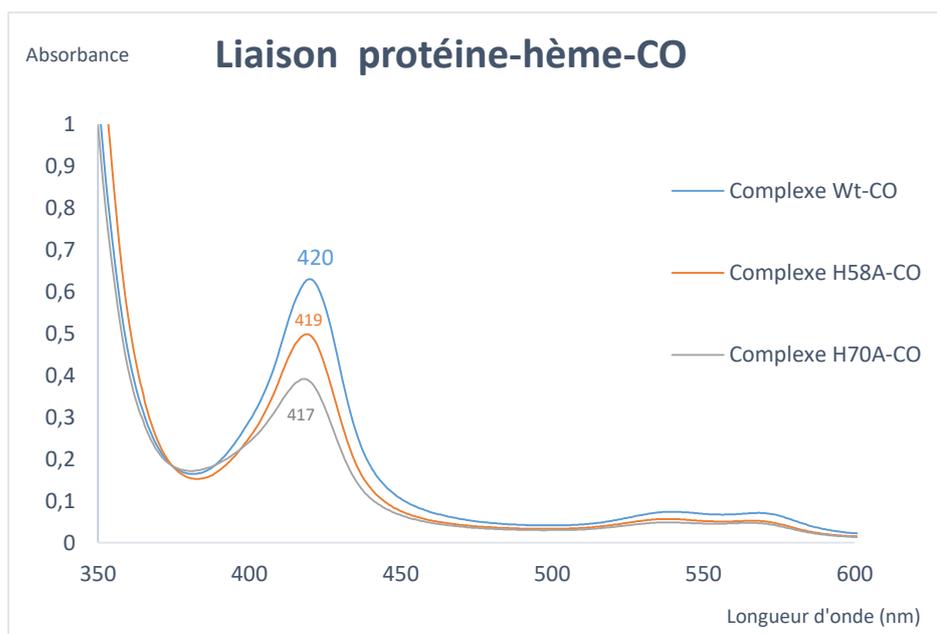


Figure 3.14 : Spectres d'absorption de Cor WT (en bleu) et des deux mutants Cor-H58A (en orange) et Cor-H70A (en gris) en présence de l'hème et du CO. Les pics d'absorption sont indiqués. Les spectres sont stables dans le temps.

Nous avons également regardé l'effet d'autres molécules gazeuses diatomiques. Afin de tester si un oxycomplexe peut être formé dans Cor, nous avons testé si l'hème peut être réduit par l'ascorbate (0,282 mV à pH 7), qui ne réagit pas avec l'O₂ contrairement au dithionite. Ce n'est pas le cas, indiquant un potentiel rédox assez bas. Aussi, le spectre d'absorption de l'hème à l'équilibre de Cor en présence prolongée d'oxygène et d'ascorbate ne montre pas de changements spectraux.

En présence de dithionite et de NO, la formation transitoire d'un complexe instable est observée, ce qui conduit au détachement de l'hème. Ces observations sont en accord avec le fait que Cor ait une spécificité pour le CO, comme suggéré dans [89].

En conclusion, dans cette première partie, nous avons exprimé et purifié Cor (Rv1829) et montré que la protéine est un dimère soluble allongé, très stable. Cor a été purifiée sans cofacteur, mais est capable de lier l'hème sur lequel le CO peut se fixer de façon directe. La présence de CO accélère la liaison de l'hème à la protéine et permet une liaison stœchiométrique (un hème/monomère). L'acide aminé H70 semble impliqué à la liaison de l'hème à Cor.

L'étude des aspects dynamiques de cette protéine nous montre que Cor possède les caractéristiques cinétiques d'un senseur à CO, qualitativement similaires aux senseurs à CO bactériens CooA et RcoM (voir chapitre II).

En conclusion, en se basant sur l'ensemble de ces résultats, la protéine Cor pourrait agir comme senseur direct du CO et ainsi détecter et répondre au CO chez *Mycobacterium tuberculosis*, ce qui nous ouvre les questions sur les mécanismes de régulation et de réponse de cette protéine.

3.3. Le cluster de gènes « *cor* »

Le gène *rv1829* (*cor*) fait partie d'un cluster de gènes très conservé et largement répandu. La base de données NCBI ; BLAST « non-redundant », contient des milliers de séquences contenant ce domaine, en particulier dans des Actinobactéries, y compris d'autres mycobactéries pathogènes (*M. avium*, *M. kansasii*, *M. abscessus* et *M. leprae*) et d'autres espèces bactériennes comme *Rhodococcus sp.* ou des corynébactéries (voir figure 1.26). Malgré sa large distribution, ce cluster de gènes reste très peu caractérisé. Par exemple, aucune donnée fonctionnelle n'était disponible sur la régulation potentielle de *cor* ou encore le rôle d'un autre gène du cluster, *rv1828*, adjacent à *rv1829*.

Dans le but d'obtenir des premières indications sur les rôles de certains gènes du cluster et la régulation potentielle de *cor*, nous nous sommes particulièrement intéressés dans cette thèse à *rv1828*, prédite comme facteur de transcription (figure 3.15) et également à l'implication de *rv0081*, qui code pour un régulateur de transcription et faisant partie du régulon de dormance chez *Mtb* (voir figure 3.16 et 1.4.4.d).

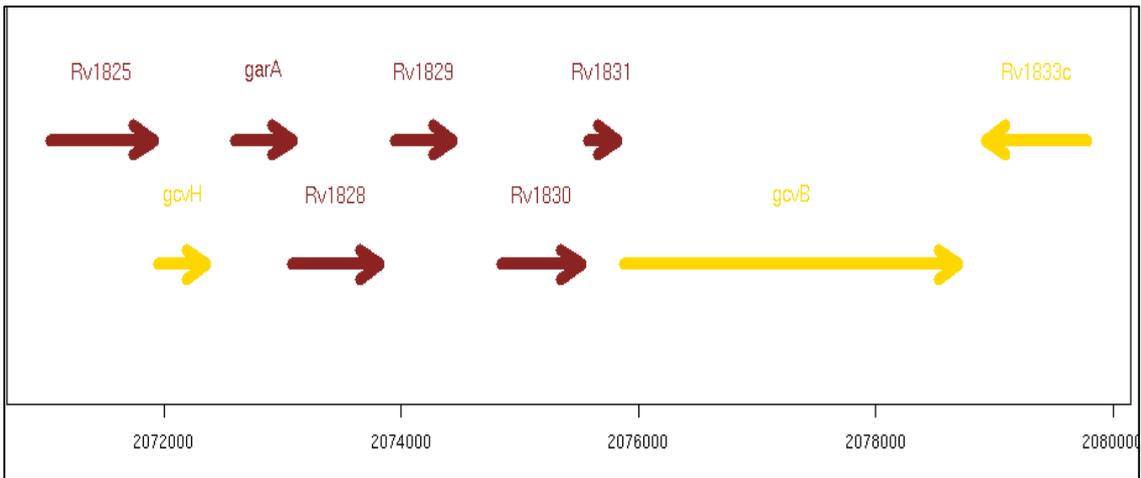


Figure 3.15 : Cluster de gènes dans *MtbH37Rv* impliquant *rv1828*, *rv1829* (*cor*) et *rv1832* (*gcvB*) (<http://svitsrv8.epfl.ch/tuberculist/>). Voir le texte pour détails.

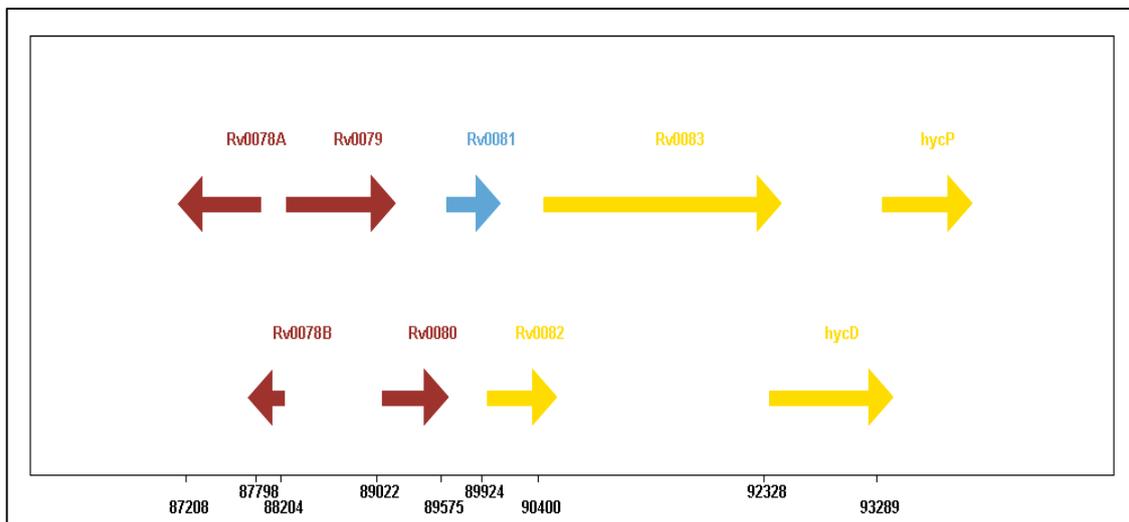


Figure 3.16 : Cluster de gènes dans *MtbH37Rv* impliquant *rv0081* (en bleu). Ce gène fait partie du régulon de dormance chez *Mtb* (<http://svitsrv8.epfl.ch/tuberculist/>).

3.3.1. Alignements et analyses *in silico*

Le gène *rv0081* fait partie du régulon de dormance chez *M. tuberculosis* ([64] ; voir introduction). Il a été constaté que *in vitro* *rv0081* était surexprimé après traitement avec du NO ou CO [84, 88]. Rv0081 est un régulateur transcriptionnel de la famille ArsR/SmtB et régule négativement sa propre expression et celle des gènes en aval (« downstream ») en liant un

élément de répétition inversé dans la région en amont (« upstream »). De plus, Rv0081 se lie à de nombreuses régions promotrices et agit comme centre de régulation dans la phase latente de l'infection tuberculeuse.

Des analyses de la biologie des systèmes et de ChiP-seq sur *Mtb* ont indiqué que *rv1829* soit directement régulé par Rv0081 (voir site Web : <http://networks.systemsbiology.net/mtb/genes/Rv1829>), et cette interaction régulatrice directe en amont du gène *cor* (région intergénique *rv1828-rv1829*) prédite pour Rv0081, comme le montre l'alignement dans la figure 3.17 a été étudiée.

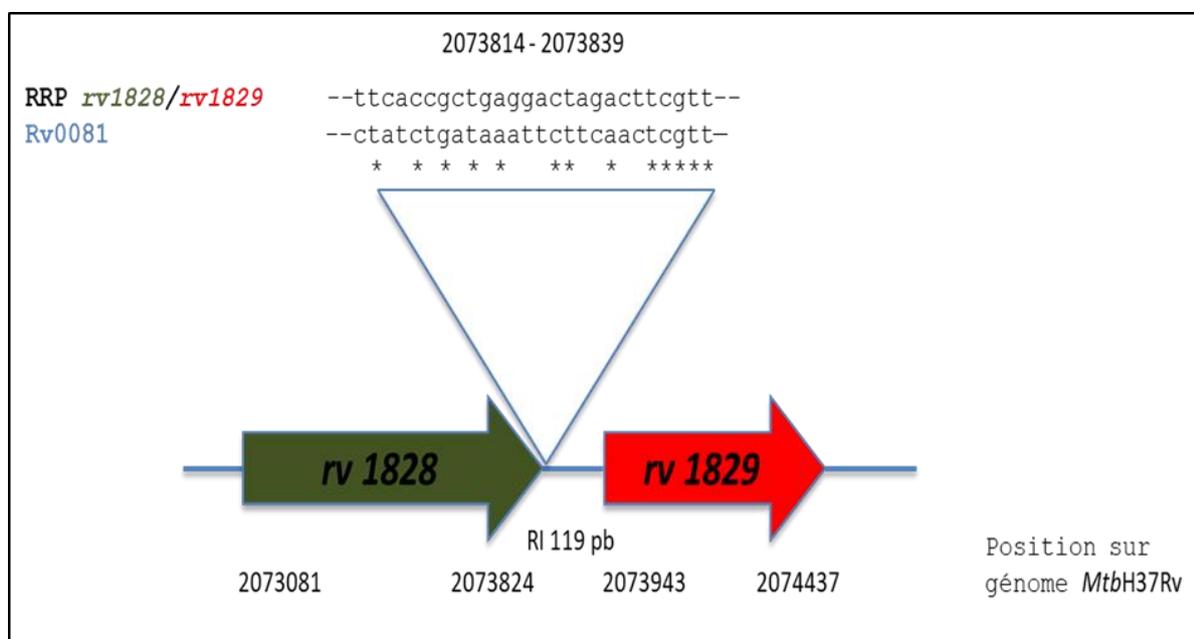


Figure 3.17 : Alignement de séquences de la région régulatrice prédite (RRP) du gène *rv1829* (*cor*) et du motif de fixation d'ADN du régulateur de transcription mycobactérien Rv0081. La région intergénique (RI) *rv1828/rv1829* comprend 119 pb. Alignement de séquences avec CLUSTAL Omega (1.2.4). La numérotation du génome est celle de *MtbH37Rv*.

La protéine Rv1828 est prédite comme régulateur de transcription du type MerR [91], mais sa fonction cellulaire n'est pas connue. Nous avons émis deux hypothèses : vu les similarités de séquence entre les régulateurs de transcription du type ArsR/SmtB et MerR, nous voulons d'abord tester si Rv1828 peut se lier à la région régulatrice connue de *rv0081* (figure 3.18). Étant donné que les gènes d'un même cluster possèdent souvent des fonctions

liées, en deuxième temps nous avons étudié si Rv1828 pourrait jouer un rôle direct dans la régulation de gènes en aval (« downstream ») du cluster.

Nous avons donc testé une interaction potentielle avec la séquence régulatrice directe en amont du gène *cor* (*rv1828-rv1829*).

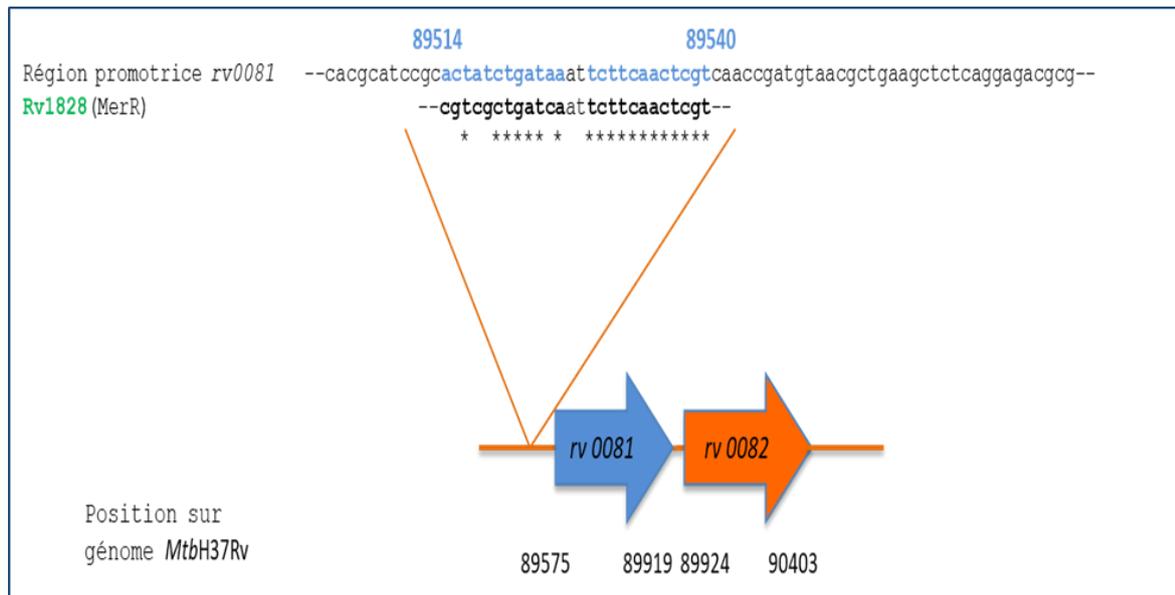


Figure 3.18 : Alignement de séquences de la région promotrice connue de *rv0081* et du motif de fixation d'ADN du prédict régulateur de transcription du type MerR, Rv1828. Alignement de séquences avec CLUSTAL Omega (1.2.4). La numérotation du génome est celle de *MtbH37Rv*.

Finalement, comme la fonction cellulaire de Cor n'est pas connue, nous avons testé si la protéine Cor peut elle-même fixer de l'ADN, de manière dépendante ou non de la présence de l'hème et de CO. Ainsi, nous avons étudié une liaison de Cor à plusieurs motifs de séquences ADN régulatrices (voir 3.3.3.).

3.3.2. Expression et purification des régulateurs de transcription

Pour tester si les protéines Rv0081, Rv1828 et Cor sont capables d'agir potentiellement comme régulateurs directs de transcription, dans un premier temps leurs gènes ont été clonés dans le vecteur d'expression pQE80L, et les protéines correspondantes ont été exprimées de façon hétérologue dans *E. coli*, souche BL21DE3 (voir Matériel et Méthodes).

a) Le régulateur de transcription Rv0081

Suite au clonage et expression, la protéine Rv0081 a été purifiée par chromatographie d'affinité sur colonne Ni-TED, suivi d'une étape d'élimination d'imidazole par les colonnes Econo-Pac 10 DG (Bio-Rad). Après détermination de la concentration de protéines par la méthode de Bradford, Rv0081 a été caractérisée sur SDS-PAGE 15%, comme le montre la figure 3.19. Rv0081 possède une masse moléculaire apparente de 12.3 kDa.

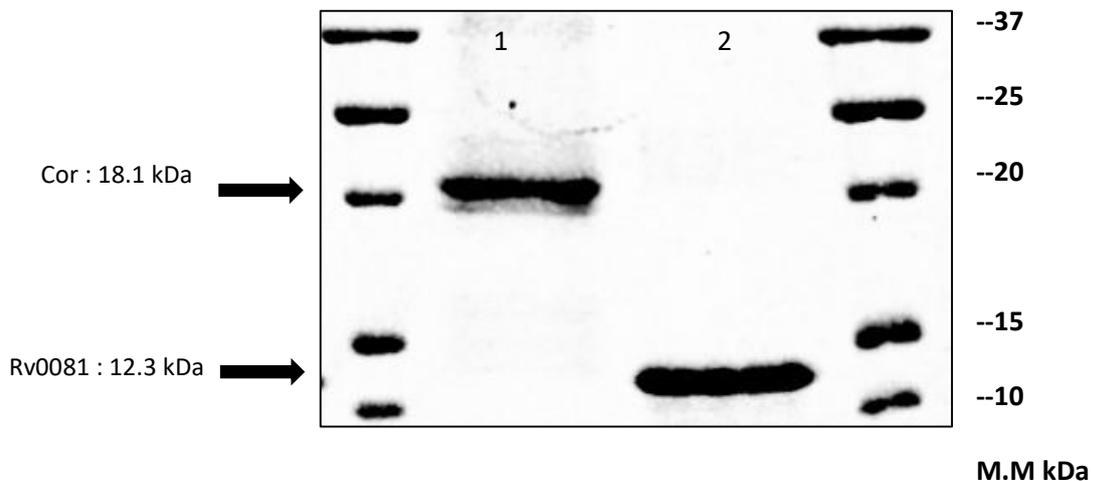


Figure 3.19 : SDS-PAGE 15% coloré au Bleu de Coomassie. Rv0081 (puits 2) purifiée avec une masse moléculaire apparente de ≈ 12.3 kDa. Puits 1 : Cor purifiée.

b) Le régulateur de transcription Rv1828

Cette protéine de 26.4 kDa, prédite comme facteur de transcription, a été purifiée également par chromatographie d'affinité sur colonne Ni-TED et caractérisé sur SDS-PAGE 10%, comme le montre la figure 3.20.

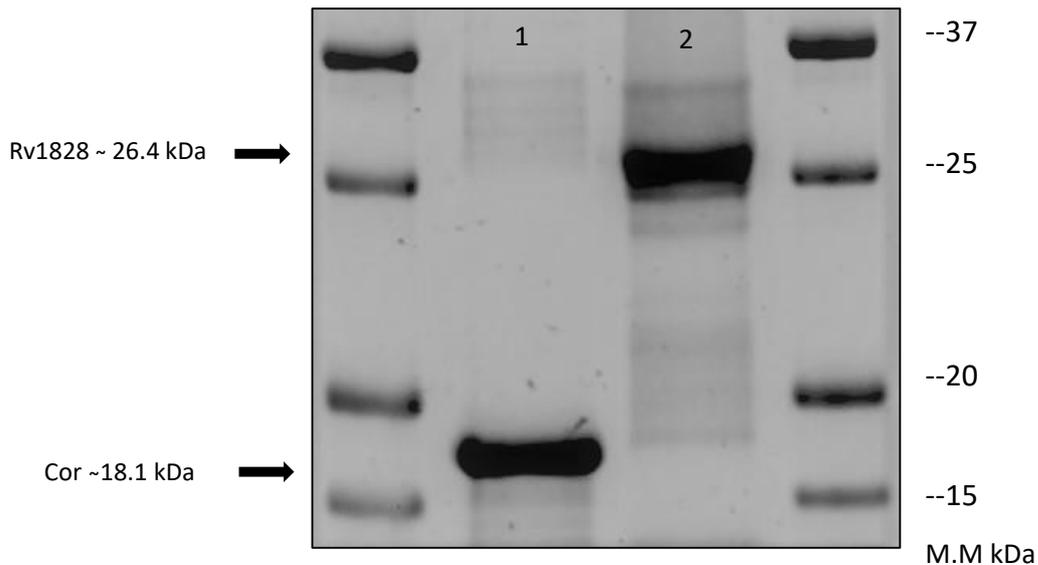


Figure 3.20 : SDS-PAGE 10% coloré au Bleu de Coomassie. Rv1828 (puits 2) montre une masse moléculaire apparente de ≈ 26.4 kDa. Puits 1 : Cor purifiée.

3.3.3. Interaction protéines - ADN

Pour tester la capacité des deux facteurs de transcription Rv0081 et Rv1828 de se lier à la région intergénique *rv1828-rv1829*, et pour tester une liaison possible de Cor à l'ADN, plusieurs régions régulatrices potentielles d'ADN ont été synthétisées. L'appariement des deux oligonucléotides a permis d'obtenir des ADN double brin, dont un brin est étiqueté par le marqueur fluorescent Cy 5.5 en 5'.

La liaison des protéines à leurs ADN cible a été testée sous différentes conditions par deux techniques : 1) de façon quantitative par spectroscopie d'anisotropie de fluorescence ; 2) de façon qualitative par les tests d'EMSA (*electrophoretic mobility shift assay* = retard sur gel).

3.3.3.1) Spectroscopie d'anisotropie de fluorescence

a) Principe de l'anisotropie de fluorescence

L'anisotropie de fluorescence permet de suivre la liaison des protéines (p. ex. des facteurs de transcription) à leurs ADN cibles grâce au changement de vitesse rotationnelle.

L'ADN est étiqueté par un marqueur fluorescent (fluorophore) en 5'. Suite à l'excitation de ce fluorophore par la lumière polarisée, la fluorescence induite est aussi polarisée (anisotropie), dans une mesure qui dépend de la vitesse de réorientation des fluorophores, elle-même dépendant de la taille du complexe liant le fluorophore.

Les amplitudes de la fluorescence émise analysée dans deux directions de polarisation perpendiculaires (I_{vv} et I_{vh}) sont mesurés et la valeur d'anisotropie de fluorescence (r) est calculée, $r = \frac{I_{vv} - I_{vh}}{I_{vv} + 2I_{vh}}$. Lorsqu'une protéine ou macromolécule forme un complexe avec un fluorophore, la valeur r augmente.

La forme et la taille du complexe fluorescent changent lorsque la protéine est ajoutée à la solution de l'ADN avec formation d'un complexe protéine-ADN. Ainsi la formation de complexes protéine-ADN augmente le volume du complexe fluorescent, ce qui ralentit son mouvement de rotation et augmente donc les valeurs d'anisotropie. Comme l'ADN marqué est le plus petit partenaire du complexe, cela implique que lors de la formation d'un complexe, le changement de forme et de taille (et par conséquent le changement d'anisotropie) sera plus important que lors du marquage de la protéine.

Pour ces expériences j'ai utilisé principalement de l'ADN marqué avec le fluorophore Cy 5.5, qui est excité à 680 nm et possède un pic d'émission à 694 nm (voir figure 3.21). Certaines mesures ont été effectuées avec Texas Red, avec λ_{exc} 583 nm, λ_{em} 603 nm.

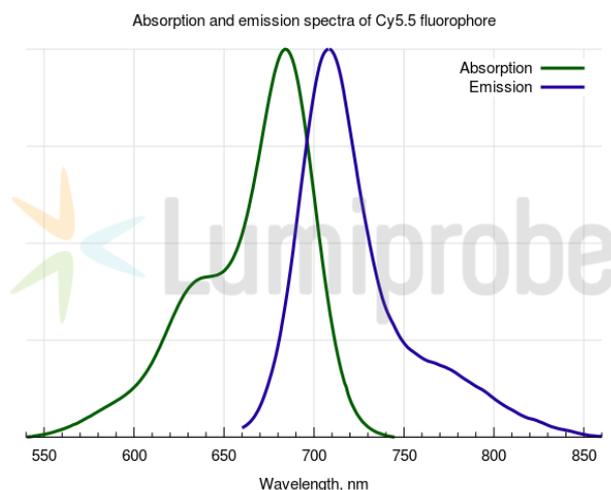


Figure 3.21 : Spectres d'absorption et d'émission du fluorophore Cy 5.5, utilisé dans les expériences d'anisotropie de fluorescence.

b) Fixation de Rv0081 et Rv1828 à la région régulatrice prédite de *cor*

La figure 3.22 montre des mesures d'anisotropie de fluorescence lors de la liaison des deux facteurs de transcription, Rv0081 et Rv1828, à la région régulatrice prédite de *cor* (voir figure 3.17).

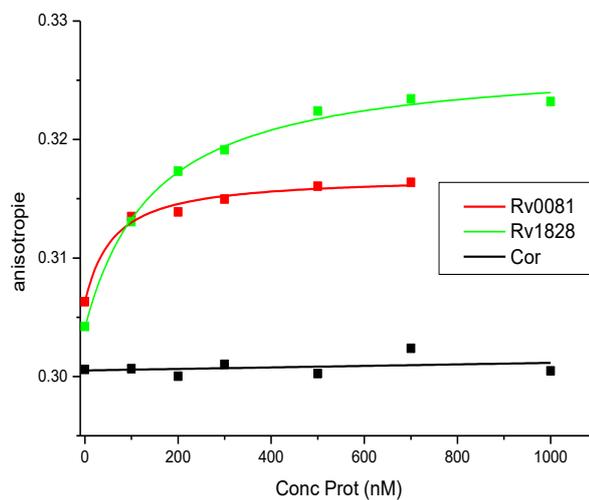


Figure 3.22 : Courbes de fixation des facteurs de transcription Rv0081 (rouge) et Rv1828 (vert) sur la région intergénique régulatrice prédite de *cor* (*rv1828-rv1829*). La protéine Cor elle-même ne montre pas de liaison (noir). La concentration de l'ADN double brin est de 10 nM dans du tampon Tris-HCl pH7.6 (50 mM), NaCl (100 mM) et les expériences ont été réalisées à température ambiante. Les affinités correspondant aux fits sont dans le texte.

Nous avons trouvé que les deux régulateurs Rv0081 et Rv1828 peuvent se lier à la région régulatrice prédite de *cor*, tandis que la protéine Cor ne se lie pas à sa propre région prédite de régulation. Les affinités déterminées sont : K_d (Rv0081) : 55 ± 18 nM ; K_d (Rv1828) : 140 ± 20 nM. Ces courbes de fixation valident alors la prédiction *in silico* pour Rv0081 et Rv1828.

c) Mesures de fixation de Rv1828 à la région promotrice *rv0081*

Les résultats d'anisotropie de fluorescence ont montré que la protéine Rv1828 peut se lier à la région promotrice de *rv0081*, comme le montrent les courbes de fixation dans la figure 3.23.

Le régulateur Rv0081 est utilisé comme contrôle positif étant donné qu'il se lie à sa propre séquence régulatrice [84]. Cor ne montre pas de liaison. Les affinités déterminées étaient : Kd (Rv0081) : 120 ± 20 nM ; Kd (Rv1828) : 110 ± 20 nM.

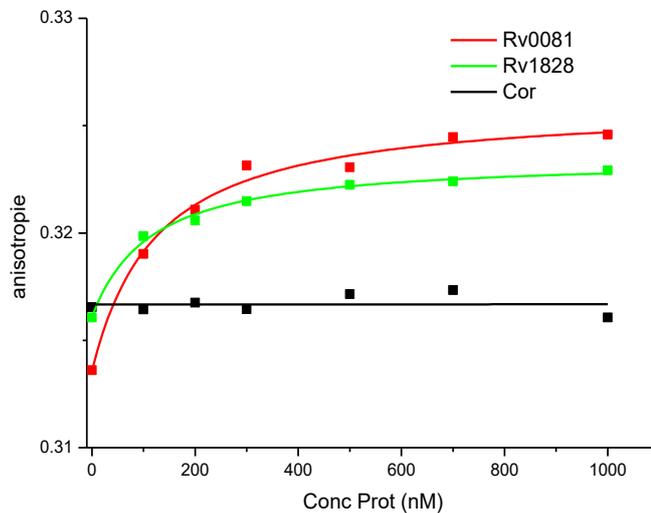


Figure 3.23 : Courbes de fixation des régulateurs Rv1828 (vert), Rv0081 (rouge ; contrôle positif) et Cor (noir) sur la région promotrice *rv0081*. La concentration de l'ADN double brin est de 10 nM dans du tampon Tris-HCl pH7.6 (50 mM), NaCl (100 mM) et les expériences ont été réalisées à température ambiante. Les affinités correspondant aux fits sont dans le texte.

d) Cor lie l'ADN en présence d'hème et du CO

Dans le but de mieux comprendre la fonction de Cor, nous avons testé si Cor elle-même possède une activité de fixation à l'ADN et pourra agir potentiellement comme facteur de transcription dépendant du CO, comme c'est le cas pour RcoM ou CoxA. Nous avons montré qu'en absence du CO, Cor seule ne montre aucune liaison sur les régions d'ADN déjà étudiées auparavant.

d.1) Etude de la liaison de Cor à la région ADN *rv1831/rv1832*

Le gène *rv1832* (*gcvB*) (voir figure 3.15) code pour une protéine du système de clivage de la glycine (« glycine cleavage system »), catalysant la dégradation de l'acide aminé glycine. Ce système est impliqué dans le métabolisme à 1-carbone et sujet d'une régulation gazeuse/rédox. Nous avons donc testé la capacité de Cor à se lier de façon spécifique à la

région ADN régulatrice prédite de *rv1832* (*rv1831-rv1832*). Il est à noter que le technique d'anisotropie de fluorescence possède un avantage très important comparé avec des EMSA : elle nous permet de tester la liaison ADN-protéine en présence du CO.

Les résultats des mesures d'anisotropie de fluorescence sont présentés dans la figure 3.24. Nous avons trouvé une liaison de Cor sur la région d'ADN régulatrice prédite en amont de *rv1832* (*rv1831/1832*), seulement en présence de l'hème et du CO (figure 3.24A). L'hème a été ajouté à la protéine de façon stœchiométrique : 1 hème/monomère. Aucune liaison n'a été observée pour le complexe hème-CO (sans protéine) (fig. 3.24B) ou pour la protéine Cor seule (fig. 3.24C).

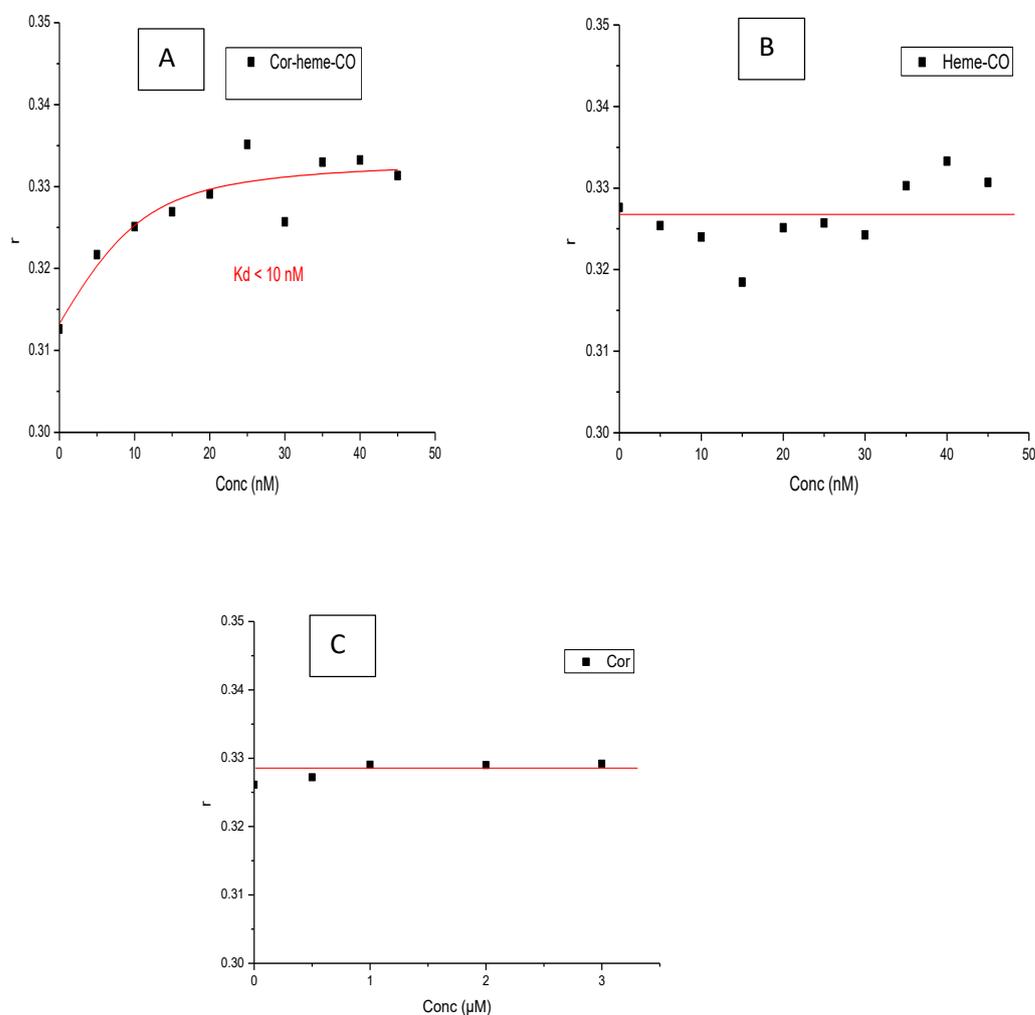


Figure 3.24 : Courbes de fixation de Cor sur la région d'ADN *rv1831/rv1832* (10 nM) dans du tampon Tris-HCl pH7.6 (50 mM), NaCl (100 mM), en présence de l'hème et du CO (A). Le complexe hème-CO sert comme contrôle négatif (B). Cor sans hème et CO ne se lie pas à l'ADN (C). Le fit en A tient compte de la concentration de l'ADN marqué.

Ces résultats nous montrent que *in vitro* Cor possède une activité CO-spécifique de liaison à l'ADN avec une affinité nanomolaire ($K_d < 10$ nM). Bien que préliminaires, ces données sont compatibles avec un rôle potentiel de Cor comme régulateur de transcription CO-dépendant à base d'hème, qui reste à explorer.

d.2) Etude des mutants de Cor H58A et H70A

Nous avons également testé la capacité des mutants Cor-H58A et Cor-H70A à fixer de l'ADN. Comme déjà montré (figure 3.14), Cor-H58 lie de l'hème et du CO de façon similaire à Cor WT, tandis que le mutant Cor-H70A lie moins de l'hème et de CO.

Après avoir préparé les complexes respectifs protéine-hème-CO et testé la possibilité de liaison à l'ADN *rv1831/rv1832*, les résultats de nos mesures d'anisotropie (voir figure 3.25) montrent une liaison de la protéine mutante Cor-H58A à l'ADN similaire à la liaison de Cor-WT. Le mutant Cor-H70A montre une liaison plus faible.

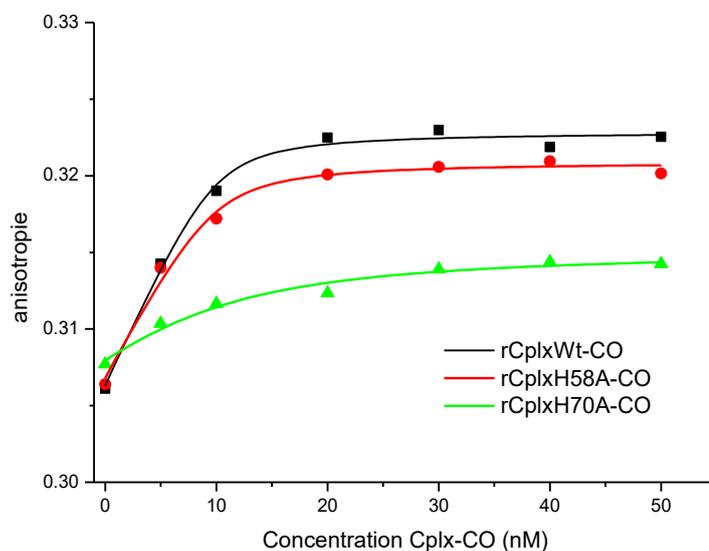


Figure 3.25 : Courbes de fixation des protéines Cor WT (noir) et des mutants H58A (rouge) et H70A (vert), en présence de l'hème et du CO sur la région d'ADN *rv1831/rv1832* (10 nM) dans du tampon Tris-HCl pH7.6 (50 mM), NaCl (100 mM). Les complexes protéine : hème sont stoechiométriques. Les fits tiennent compte de la concentration de l'ADN marqué.

Les affinités déterminées étaient : K_d (Complexe WT-CO) : $< 1 \text{ nM}$; K_d (Complexe H58A-CO) : $< 1 \text{ nM}$; K_d (complexe H70A-CO) : $6 \pm 3 \text{ nM}$.

Ces résultats sont en accord avec une implication de l'histidine H70 dans la liaison de l'hème sur lequel se fixe le CO et montrent que la présence de l'hème et du CO sont indispensables pour une liaison de Cor à l'ADN.

Afin de valider nos prédictions *in silico* et nos résultats d'anisotropie, la technique EMSA (« electrophoretic mobility shift assay ») a été utilisée. Les ADN ont été étiquetés par le fluorophore Cy 5.5 en 5' à un brin d'ADN.

3.3.3.2) Electrophoretic mobility shift assay : EMSA

a) Principe de l'EMSA (« retard sur gel »)

La deuxième technique utilisée pour tester la liaison des protéines à l'ADN cible était l'EMSA (voir Matériels et Méthodes). Cette technique est basée sur l'observation que des complexes protéine-ADN migrent plus lentement qu'un fragment d'ADN linéaire libre sur un gel d'électrophorèse polyacrylamide non-dénaturant. La migration de l'ADN est donc retardée quand elle est liée à une protéine. Un schéma est présenté dans la figure 3.26.

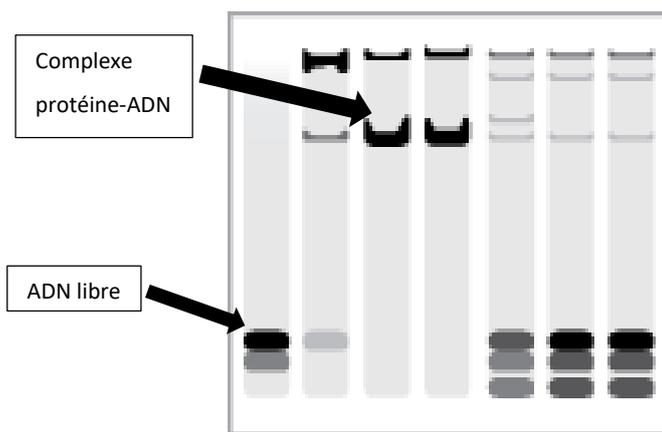


Figure 3.26 : Schéma d'EMSA montrant le retard sur gel quand un complexe protéine-ADN est formé, comparé avec la migration plus rapide de l'ADN libre.

b) EMSA avec Rv0081, Rv1828 et Cor

La technique EMSA a été utilisée pour tester les protéines Rv0081, Rv1828 et Cor avec deux fragments d'ADN cibles des régions régulatrices potentiels : *rv0081* et *rv1828/29*.

Les résultats de cette expérience d'EMSA suite aux réactions de liaison et migration sur gel natif polyacrylamide 10% dans une chambre froide (décrites dans la partie Matériels et Méthodes), sont analysés en utilisant le système Odyssey (LI-COR) à 700 nm (voir figure 3.27).

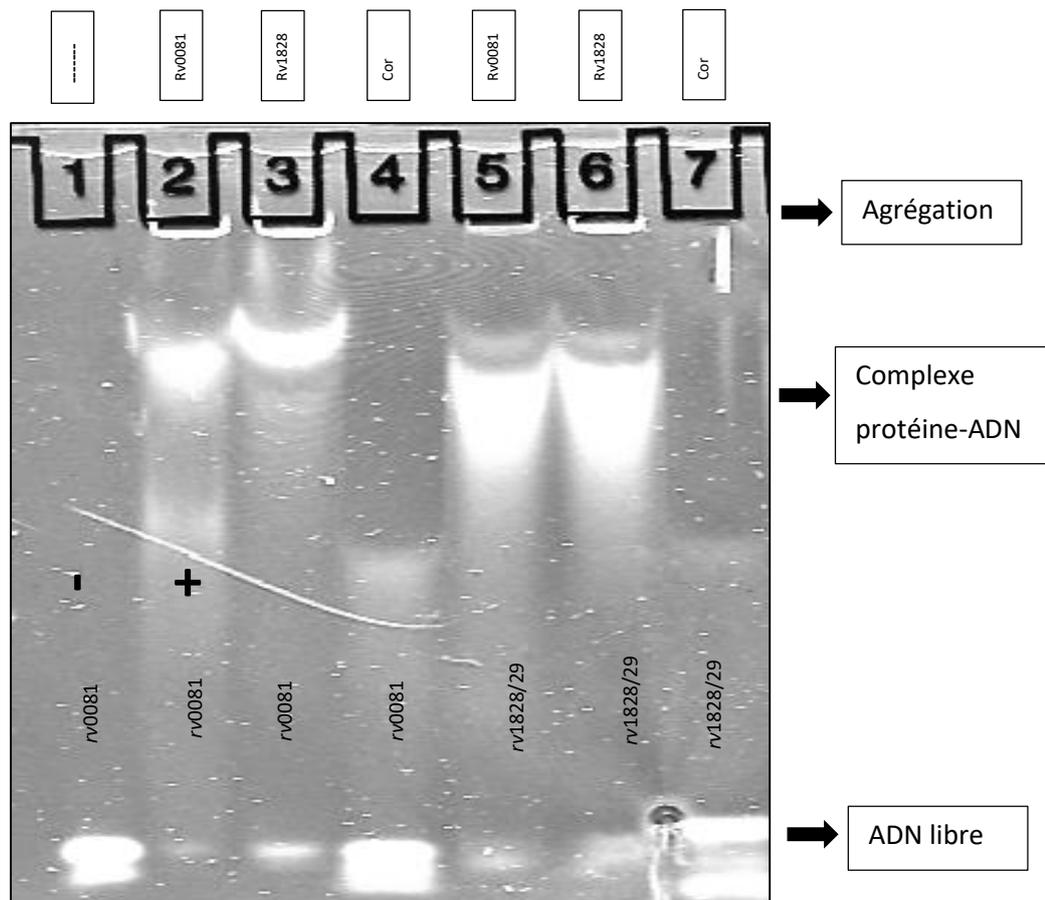


Figure 3.27 : Gel retard (EMSA) en présence d'ADN compétiteur pour les trois protéines testées (concentration 500 ng) : Rv0081, Rv1828 et Cor (avec deux fragments d'ADN testées (concentration 5 nM) : *rv0081*, *rv1828/29*). Puits 1 : contrôle négatif (sans protéine) ; puits 2 : contrôle positif (Rv0081 qui se lie à sa propre région promotrice).

Dans cette expérience un contrôle négatif (puits 1) a été utilisé sans protéine et un contrôle positif avec la région d'ADN *rv0081* (puits 2). Comme montré, l'ADN libre migre

rapidement, comparé au contrôle positif complexé avec Rv0081 qui se fixe sur sa propre séquence promotrice.

Un retard sur gel correspondant à la formation d'un complexe ADN-protéine est observé pour le régulateur Rv1828 avec la région promotrice *rv0081* (puits 3), mais pas pour la protéine Cor en absence de l'hème et CO (puits 4). Un retard sur gel a été également observé pour les régulateurs Rv0081 (puits 5) et Rv1828 (puits 6) avec la région *rv1828/29*, mais pas pour la protéine Cor sans hème et CO (puits 7). Cela montre que ces deux protéines Rv0081 et Rv1828 peuvent se lier à la région régulatrice prédite de *cor*.

Ces résultats d'EMSA sont en accord avec nos résultats d'anisotropie de fluorescence. Par ailleurs, la protéine Cor seule ne montre aucune liaison avec ces trois régions d'ADN, comme les expériences d'EMSA ont été faites en *absence* de l'hème et du CO.

Alors que ces résultats sont encore préliminaires, ils suggèrent la possibilité d'une régulation transcriptionnelle complexe de certains gènes du cluster de *cor*, comme le montre le schéma récapitulatif (voir figure 3.28). Il est intéressant à noter que Cor a besoin de l'hème et du CO pour une activité de liaison à l'ADN. Il faut quand même indiquer que l'activité de liaison observée est relativement faible et ne permet pour l'instant pas des conclusions sur la nature exacte d'une séquence ADN cible de Cor.

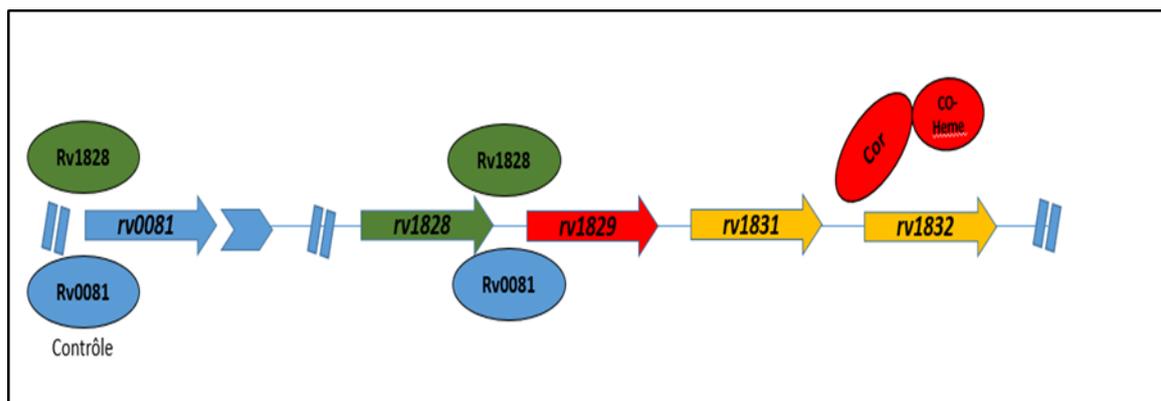


Figure 3.28 : Schéma récapitulatif des résultats des expériences d'anisotropie de fluorescence et d'EMSA. Une liaison de Rv1828 et Rv0081 à la région régulatrice prédite de *cor* (*rv1828/1829*) a été observée. Rv1828 reconnaît la région promotrice de *rv0081* et Rv0081 se fixe à sa propre région promotrice. Cor montre une liaison à l'ADN *rv1831/1832* seulement en présence de l'hème et du CO.

3.4. Cristallogénèse des protéines

Vu qu'il n'y avait pas de données structurales de Cor et Rv0081, j'ai entamé pendant ma thèse la cristallogénèse des deux protéines dans le but de déterminer leurs structures à 3D. Des expériences de cristallisation avec différents kits (décrites dans la partie Matériels et Méthodes) ont été réalisées (en collaboration avec Ines Gallay et Stéphane Planqueel, I2BC, U. Paris Sud).

Bien que nous n'ayons pas réussi à produire de cristaux diffractables pour la protéine Cor, nous avons obtenu des cristaux de Rv0081, qui ont diffractés et sont actuellement en cours de résolution de structure.



Figure 3.29 : Premiers cristaux de Rv0081 avec une résolution de 5.3 Å. Les cristaux ont été obtenus avec le kit JCSG+ D12 g3, comprenant 16% PEG 8000, 20% glycérol et 0,4 M phosphate de potassium à 18°C.

Plus récemment, en utilisant les kits Cryos et Classic lite à 4°C nous avons obtenu des cristaux avec une diffraction de 1.5 Å au synchrotron Soleil (comme le montre la figure 3.30). Ces cristaux sont actuellement en cours de résolution de structure.



Figure 3.30 : Cristal récent de Rv0081 avec une résolution de 1.5 Å. Cristallogénèse avec les kits Classic Lite et Cryos à 4°C dans 20% PEG 10000 et 0,1 M MES à pH 6,5.

Très récemment, une structure 3D de Rv0081 a été publiée avec une résolution de 2.9 Å. Cette structure confirme que Rv0081 appartient bien aux protéines de la famille ArsR/SmtB. Cependant, contrairement à d'autres membres de cette famille, Rv0081 n'a pas de domaine de liaison réversible aux métaux, ce qui suggère un mécanisme alternatif de régulation des gènes et de détection des niveaux d'oxygène [87].

Il sera intéressant de comparer les deux structures et également obtenir des cristaux de Rv0081 liée à son ADN (séquence promotrice).

3.5. Rôle physiologique de Cor

Afin d'élucider le rôle physiologique de la protéine Cor dans la résistance au CO, nous avons testé l'effet du CO sur la croissance mycobactérienne, en utilisant un modèle non-pathogène, *M. smegmatis* souche MC²155. Cette souche présente une grande similarité avec *Mtb* et contient le gène *cor*, ce qui nous permet de réaliser son inactivation et d'étudier plusieurs paramètres, comme la survie bactérienne ou la morphologie cellulaire, par microscopie confocale.

3.5.1. Croissance mycobactérienne en présence du CO

Pour tester l'effet du CO sur les cultures de *M. smegmatis*, nous avons utilisé du CO gazeuse et également des molécules CO-RM-401. Les molécules CO-RM (« CO-releasing molecules ») sont des métaux-carbonyles avec des groupes CO, capables de libérer des molécules CO au moment de la mise dans le milieu par des réactions d'échanges de ligand, une activation enzymatique ou encore une photo-activation. L'avantage de leur utilisation est que la quantité du CO mise en culture peut-être bien contrôlée. Nous avons utilisé dans nos expériences les molécules CO-RM-401 (voir figure 3.31), capables de libérer trois molécules de CO/mole et sensitive à l'oxydation.

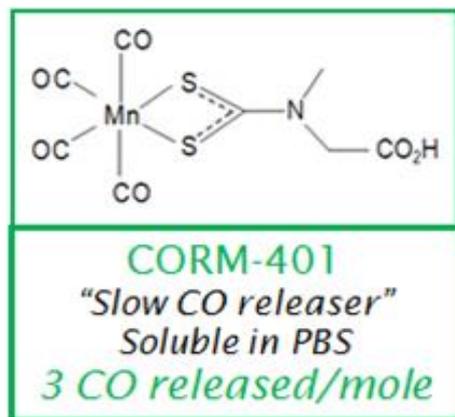


Figure 3.31 : Structure de la molécule CO-RM-401 utilisée comme source du CO dans les milieux de culture de *M. smegmatis*.

L'effet du CO sur les cultures de *M. smegmatis* a été étudié d'abord sur le type sauvage, mais pas encore comparé avec la souche de délétion de *cor*, qui est actuellement en fin de construction (voir ci-dessous).

a) Souche sauvage de *M. smegmatis*

Dans le but de tester l'effet du CO gaz et de CO-RM 401, des cultures de *M. smegmatis* du type sauvage ont été lancées en présence du CO gaz (5 % et 25 %) et de CO-RM-401 (50 μ M) pour 2-3 jours à 37°C. La densité optique (D.O.) à 600 nm des cultures a été mesurée afin

de comparer l'effet sur la croissance cellulaire en premier lieu et ensuite la survie et la morphologie cellulaire en deuxième lieu (voir Tableau 3.2).

D.O (600 nm)	T=16h	T=24h	T=36h	T=48h	T=56h
Contrôle (sans CO)	0.009	0.086	0.426	0.815	1.254
Culture +5% CO	0.005	0.029	0.301	0.551	0.903
Culture +25% CO	0.005	0.01	0.151	0.274	0.602
Culture avec 50µM CORM-401	0.001	0.002	0.006	0.109	0.205

Tableau 3.2 : Tableau de comparaison de la densité optique à 600 nm entre les cultures de *M. smegmatis* en absence et présence du CO gaz et CO-RM 401, en fonction de temps (en heures).

Ces résultats montrent que les mycobactéries peuvent supporter de fortes concentrations (25%) du CO gazeux et continuent également leur croissance en présence de CO-RM 401, mais avec une croissance ralentie.

Nous avons ensuite testé la survie et la morphologie cellulaire de *M. smegmatis* en utilisant de la microscopie confocale. Un colorant fluorescent FM5-95, qui pénètre dans les cellules vivantes et à travers les membranes bactériennes, a été utilisé afin de visualiser les membranes plasmiques sélectivement dans les cultures mycobactériennes et détecter la croissance et présence des mycobactéries.

La figure 3.32 montre les images des cellules de *M. smegmatis* issues de ces cultures en présence/absence du CO et du marqueur fluorescent FM5-95.

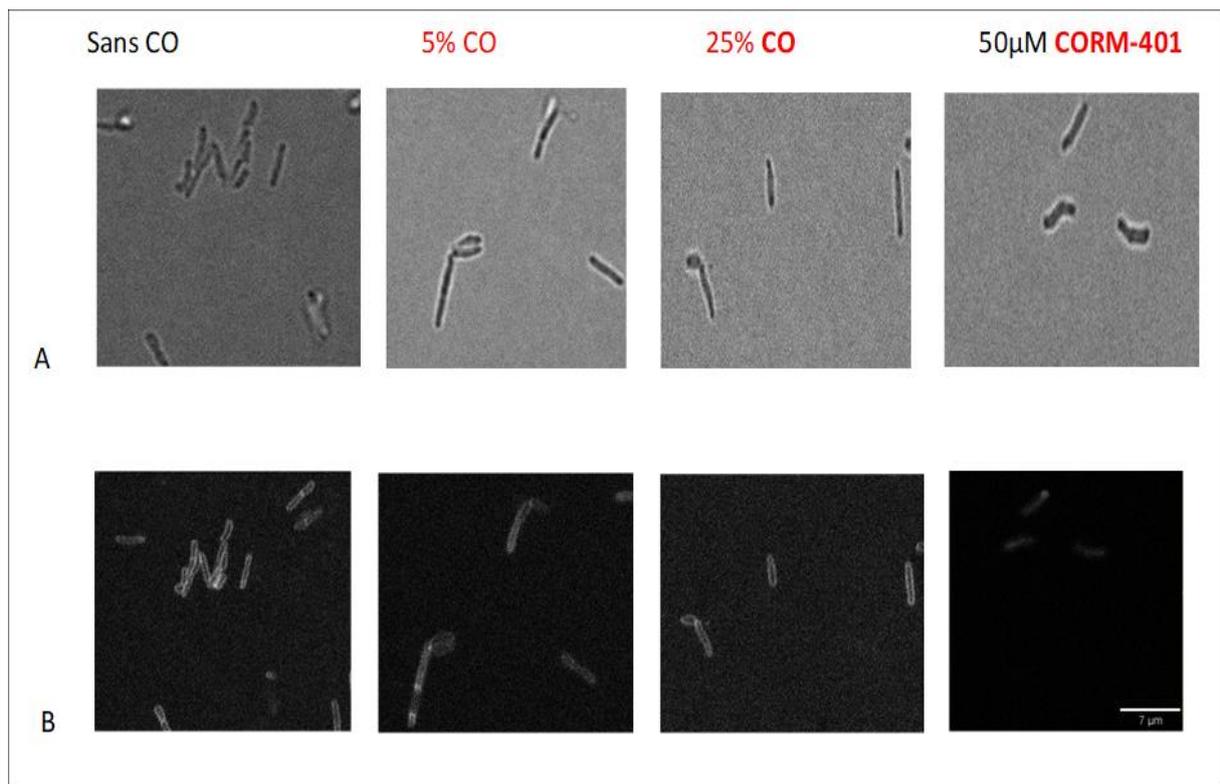


Figure 3.32 : Images des cellules de *M. smegmatis* sur agar. La série A (en haut) présente les images des cellules issues des cultures en présence et absence du CO (sans marqueur FM5-95) ; la série B (en bas) montre les cellules en présence de FM5-95.

Ces images issues de la microscopie confocale pour des cellules de *M. smegmatis* de type sauvage montrent que leurs morphologies cellulaires sont à peu près identiques au contrôle sans CO et ne sont pas affectées par des fortes concentrations en CO.

Suite à ces données, nous nous sommes intéressés à étudier l'effet du CO sur une souche dans laquelle *cor* a été inactivé et le comparer avec le type sauvage.

b) Création d'une souche d'inactivation de *cor*

Dans le but de comparer des paramètres observés pour la souche sauvage de *M. smegmatis* avec une souche d'inactivation de *cor*, nous sommes en train de construire une souche de *M. smegmatis* avec *cor* inactivé. Ces travaux sont très récents et je présente

l'avancement des expériences de cette partie. La construction de la souche de délétion Δcor se réalise en deux étapes :

- 1) Construction du plasmide pGEMT-easy : *cor* : Kan en prévision de l'échange de la cassette Kan dans le plasmide pJV53 :

Amplification du gène *cor* par PCR en utilisant l'ADN génomique de *M. smegmatis* comme matrice avec les oligonucléotides suivants :

5'-AGTACCTGCCGCTCAAAGTC-3' (sens) et

5'-CTGCTGGCTGCTGTGAAAAG-3' (anti-sens)

Le gène *cor* contient un site de restriction *Bgl*II, permettant sa restriction enzymatique spécifique. Le produit amplifié de *cor* (taille 1,2 kb) a été visualisé sur gel d'agarose 1% (voir figure 3.33), puis cloné dans le plasmide pGEMT-easy (3.015 kb). La bonne insertion a été vérifiée par séquençage et le plasmide contenant l'insert *cor* est linéarisé par l'enzyme de restriction *Bgl*II (voir figure 3.34).

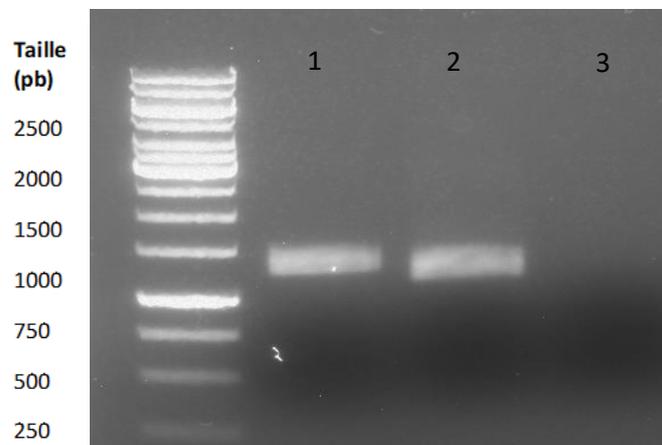


Figure 3.33 : Puits 1 et 2 : produits de PCR de *cor* (1,2 kb), obtenus en utilisant l'ADN génomique de *M. smegmatis*, comme matrice. Puits 3 : contrôle négatif de PCR (sans ADN). Gel d'agarose 1%.

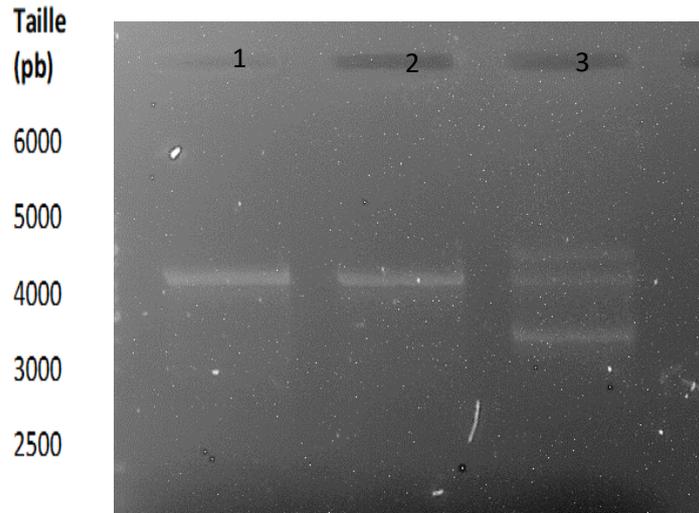


Figure 3.34 : Produits de restriction enzymatique avec *Bgl*II du plasmide pGEMT-easy avec l'insert *cor* (puits 1 et 2) dont la taille attendue est de 4.2 kpb. Contrôle négatif de pGEMT-easy circulaire sans insert *cor* et non-digéré (puits 3) sur un gel d'agarose 1%. Les différentes bandes correspondent aux conformations surenroulées du plasmide.

La cassette de résistance à kanamycine (gène *aphA-3*) a été extraite d'un plasmide pUC18-K2 par une double digestion avec l'enzyme de restriction *Sma*I (figure 3.35).

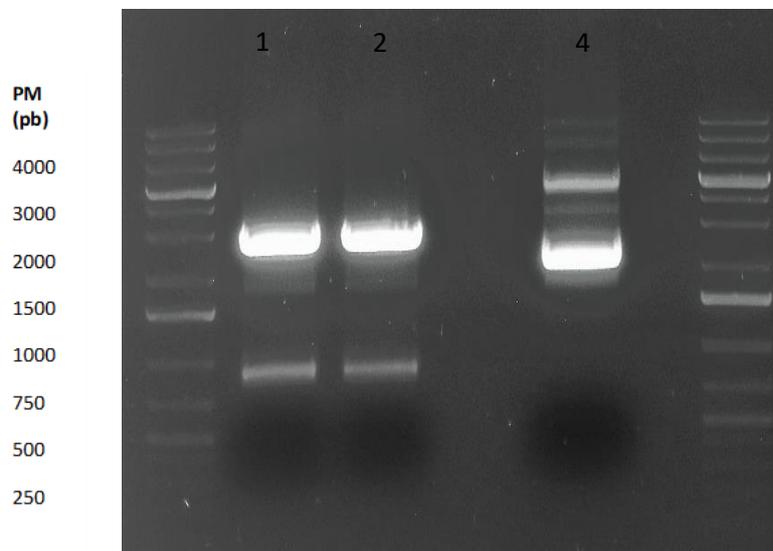


Figure 3.35 : Produits de double digestion du plasmide pUC18-K2 contenant une cassette de résistance à kanamycine (puits 1 et 2). Puits 4 : Plasmide pUC18-K2 non digéré. Gel d'agarose 1%.

La bande sur gel de 850 pb correspond à la cassette de kanamycine extraite. Elle a été purifiée sur gel pour la cloner dans une prochaine étape dans le plasmide pGEMT-easy-*cor*

(linéarisé) par l'utilisation du kit Quick Blunting (New England Biolabs), qui permet de convertir les extrémités incompatibles de 5' ou 3' en 5' phosphorylées, pour ligaturer dans le vecteur du clonage.

2) Système de recombinaison :

Le fragment (*cor* : Kan) sera extrait du plasmide pGEMT-easy pour faire un échange avec le plasmide pJV53 (dans lequel la cassette kanamycine du pJV53 a été déjà échangé en une cassette zeocine) et ensuite une transformation dans *M. smegmatis* sera effectuée. La figure 3.36 montre le schéma récapitulatif de cette stratégie d'inactivation.

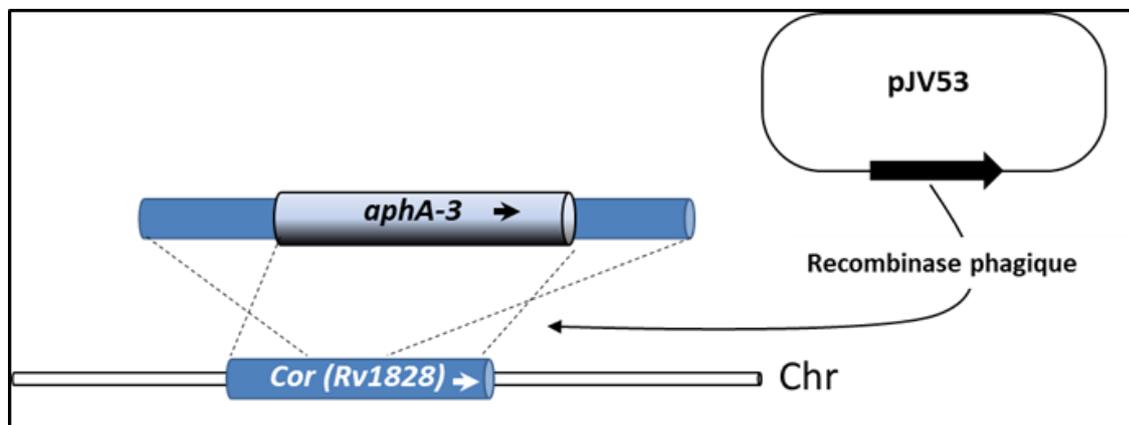


Figure 3.36 : Schéma récapitulatif de la stratégie utilisée pour la création de la souche de délétion Δcor . Inactivation du gène *cor*, extraction de la cassette kanamycine du plasmide pGEMT-easy et échangé dans le plasmide pJV53, permettant une répllication dans *M. smegmatis*.

Ces expériences seront continuées prochainement, dans le but d'étudier l'effet du CO sur la souche de délétion de *cor* et de les comparer avec les résultats déjà obtenus pour le type sauvage de *M. smegmatis*. Cela constituera un premier pas vers (par exemple) des analyses transcriptomiques des deux souches et pourra élucider le rôle physiologique de la protéine Cor dans la résistance contre le CO chez les mycobactéries.

Chapitre IV

Conclusions et Perspectives

La détection et réponse aux changements en concentration gazeuse sont cruciales pour la survie bactérienne. C'est d'importance particulière pour les pathogènes intracellulaires, qui ont besoin de surmonter différents mécanismes de défense chez l'hôte, entre autres du monoxyde de carbone (CO) produit par le métabolisme de l'hôte. Le CO, un gaz très toxique à concentration élevée, a été récemment reconnu à faibles quantités comme molécule de signalisation, pouvant affecter la régulation de gènes. La signalisation nécessite une détection rapide du CO et des réponses cellulaires adaptées, ainsi que des mécanismes spécifiques pour concurrencer le dioxygène dans des bactéries aérobies, qui sont peu compris à ce jour.

Fréquemment les systèmes développés par les bactéries pour détecter des gaz diatomiques sont les protéines senseurs à base de l'hème. Bien que ces senseurs représentent une famille croissante, très peu de senseurs spécifiques du CO sont connus à ce jour. Parmi les hème-senseurs à CO bactériens, le mieux caractérisé est le senseur à CO anaérobie CooA, qui stimule un système d'oxydation du CO. CooA ne lie du CO que lorsque son hème est dans l'état réduit alors que CooA oxydé, présent dans des conditions aérobies, ne répond pas au CO [33, 39, 41]. Par contre, la détection du CO dans les bactéries aérobies est beaucoup moins comprise.

Le senseur à CO RcoM-2

La bactérie du sol aérobie *Burkholderia xenovorans* exprime le senseur à base d'hème RcoM-2 ; un facteur de transcription CO-dépendant. Des études *in vitro* ont montré qu'une réduction de l'hème est nécessaire pour la liaison du CO, RcoM-2 possède une affinité extrêmement élevée pour le CO, tout en étant insensible à l'O₂ [45, 48]. Antérieur au début de ma thèse, l'équipe avait trouvé que dans l'hème-domaine RcoMH-2, après photodissociation, le CO se relie sur l'hème très rapidement et de façon quasi irréversible [49]. Ces observations ont posé la question comment RcoM-2 peut fonctionner comme senseur à CO, nécessitant une sortie du CO de la protéine, et nous ont conduit à déterminer les caractéristiques de la protéine entière RcoM-2, comprenant le domaine de fixation à l'ADN.

Nous avons exprimé et purifié la protéine entière RcoM-2, caractérisé les interactions hème-CO et les comparées avec les données pour l'hémo-domaine isolé RcoMH-2. Plusieurs nouvelles observations ont été faites :

La protéine entière RcoM-2 fixe du CO avec une affinité effective plus faible que RcoMH-2 (K_d de 4 nM vs. 250 pM). Cette affinité est régulée par la dynamique du CO dans la poche de l'hème, qui est influencée par les interactions à longue distance avec le domaine de liaison à l'ADN de ce senseur.

Dans la protéine entière, le CO se dissocie avec un taux 20 fois plus élevé que celui observé dans le domaine de l'hème, où la liaison du CO est quasi irréversible. Une petite fraction du CO peut s'échapper de la poche « senseur », ainsi permettant RcoM-2 d'assumer ses fonctions de senseur (Il est à noter que le taux de dissociation reste quand même très inférieur à celui observé dans le senseur à CO, le CooA [95]). Une explication possible est le processus de recombinaison géminée hème-CO, qui est plus lente et moins étendu dans RcoM-2 que le domaine senseur isolé RcoMH-2. Dans ce contexte, il est important de noter que le senseur aérobie RcoM-2 doit détecter de concentrations faibles et persistantes de CO. La recombinaison lente observée évite également les fluctuations rapides qui pourraient affecter la détection de CO.

En vue de sa fonction de régulateur de transcription CO-dépendant, la liaison du CO au domaine hémique influence évidemment les propriétés de liaison à l'ADN de RcoM-2. Il est intéressant à noter qu'une influence inverse se produit également : la liaison de l'ADN influence la dynamique de recombinaison du CO dans le site actif ; ce que pourrait avoir un rôle de régulation. On peut supposer que le domaine de liaison à l'ADN exerce une contrainte stérique sur la poche de l'hème qui, en conséquence, devient plus ouverte, ainsi permettant le CO dissocié de l'hème de sortir de la protéine. Un effet qualitativement similaire a été observé dans le senseur à CO CooA [50]. Nous suggérons qu'il s'agit d'une caractéristique générale des facteurs de transcription à base d'hème, régulés par un ligand gazeux. Les études sur l'origine moléculaire précise des propriétés dynamiques importantes de la poche de l'hème doivent attendre la disponibilité de données structurales sur RcoM-2.

Un article présentant mes résultats sur RcoM-2 a été publié dans « Biochemistry » (voir Annexe 1).

Détection et réponse au CO dans *Mtb*

La protéine Cor. Dans la bactérie obligatoirement aérobie *Mycobacterium tuberculosis* (*Mtb*), l'agent infectieux de la tuberculose, la signalisation par du CO présente un intérêt particulier. Ces bactéries pathogènes intracellulaires doivent surmonter les mécanismes de défense de l'hôte après infection, dont le monoxyde de carbone (CO), mais leurs réponses au CO sont peu comprises. *Mtb* peut résister aux concentrations élevées du CO, suggérant des mécanismes spécifiques afin de résister au CO et combattre sa toxicité en présence d'oxygène. Dans ces mécanismes, un nouveau gène nommé *cor* (« CO-resistance ») récemment identifié par des approches génétiques, mais non caractérisé [89], pourrait jouer un rôle.

Dans une deuxième partie de mon travail de thèse, j'ai étudié plusieurs aspects de la protéine Cor dans le but de mieux comprendre son rôle dans l'adaptation au CO. J'ai pu montrer que la forme principale de Cor est celle d'un dimère allongé, très stable. Dans l'absence d'une structure à 3D de Cor, nous avons effectué de la modélisation moléculaire sur la structure d'une protéine DUF151 [89] de *T. maritima* (homologie 33.33%). Plusieurs résidus chargés positivement ont été identifiés, qui pourraient potentiellement être impliqués dans une interaction avec de l'ADN.

Comment Cor pourrait-elle agir comme détecteur ou médiateur de la résistance au CO? Il a été postulé que Cor ne contienne pas de site de détection pour du CO, tel qu'un hème, du fer ou un autre métal de transition [89]. Par contre, nous avons identifié deux résidus histidine dans la séquence protéique de Cor et découvert que Cor était capable de lier un cofacteur hémique de façon stoechiométrique (1 hème/monomère). La présence du CO accélère la liaison de l'hème à la protéine. En associant la mutagenèse dirigée aux spectroscopies de fluorescence et de l'absorption, j'ai pu montrer que l'His70 joue un rôle dans la fixation de l'hème à Cor. La construction d'un double mutant His58/His70 est envisagée pour voir s'il y aura un effet plus net. En vue de la modélisation préliminaire de Cor, la détermination de His70 comme un site probable de ligandation de l'hème suggère que l'hème se fixe proche de la surface de la protéine. Cela pourrait expliquer la labilité de l'hème dans la forme oxydée.

Des études de la dynamique du CO montrent que le CO se lie très rapidement et avec une très forte efficacité (80%) à l'hème associé à Cor. Cette recombinaison substantielle est similaire à celle observée dans d'autres senseurs à CO étudiés dans le laboratoire, comme le

senseur à CO CooA [95] et RcoM-2. Elle est et beaucoup plus étendue que dans des hémoprotéines de stockage/transfert de ligands (myoglobine et hémoglobine [44]), où cette recombinaison est quasi-inexistante. Cor possède donc les caractéristiques cinétiques d'un senseur à CO.

Le cluster de gènes *cor*. *cor* fait partie d'un cluster de gènes conservé et peu caractérisé chez les mycobactéries, avec plusieurs régulateurs de transcription prédits. En me basant sur des études *in silico* et en utilisant des techniques biophysiques (anisotropie de fluorescence) et biochimiques (EMSA), j'ai trouvé que le régulateur transcriptionnel Rv1828 est capable de se lier à la région prédite régulatrice de *cor* (*rv1828/1829*). Ces études ont également suggéré que le régulateur « maître » Rv0081, qui fait partie du régulon de dormance DosR chez *Mtb*, induit en réponse à une tension réduite en oxygène et de CO, peut se lier sur la région régulatrice prédite de *cor*. Alors que ces résultats sont encore préliminaires, ils suggèrent la possibilité d'une régulation transcriptionnelle complexe de certains gènes du cluster de *cor*, qui nécessite une exploration plus poussée.

La protéine Cor elle-même présente une activité de liaison à l'ADN en présence d'hème et de CO avec une affinité nanomolaire ($K_d < 10$ nM). Cette activité de liaison est abolie dans le mutant Cor-H70A. Même si encore très préliminaires, ces données seront compatibles avec un rôle potentiel de Cor comme régulateur de transcription CO-dépendant à base d'hème, qui reste à explorer. Il faut néanmoins indiquer que l'activité de liaison observée ne permet pour l'instant pas des conclusions sur la nature exacte d'une séquence ADN cible de Cor ; les résultats d'interactions protéines-ADN issues des tests d'anisotropie de fluorescence et d'EMSA attendent une validation *in vivo*.

Nous avons obtenu des cristaux de Rv0081, qui diffractent à 1.5 Å et sont actuellement en cours de résolution de structure. Très récemment, une structure 3D de Rv0081 a été publiée avec une résolution de 2.9 Å [87]. Il sera intéressant de comparer les deux structures et également obtenir des cristaux de Rv0081 liée à son ADN promotrice.

Mes études sur l'effet du CO sur la croissance des cultures de *M. smegmatis*, un modèle non-pathogène de *Mtb*, ont montré que *M. smegmatis* peut supporter des fortes concentrations du CO (25%) et ont montré des morphologies cellulaires essentiellement identiques aux contrôles sans CO. La création d'une souche d'inactivation de *cor* se poursuit,

dans le but d'étudier l'effet du CO sur la souche d'inactivation de *cor* et le comparer avec les résultats déjà obtenus pour le type sauvage de *M. smegmatis*. Cela constituera un premier pas vers (par exemple) des analyses transcriptomiques et Chip-Seq des deux souches et pourra élucider le rôle physiologique de la protéine Cor dans la résistance contre le CO chez les mycobactéries.

Chapitre V

Matériels et Méthodes

5.1. Techniques de biologie moléculaire

5.1.1. Milieux de culture et croissance bactérienne

Le milieu LB (Luria Burtani) est utilisé pour les cultures bactériennes d'*Escherichia coli* et a aussi été utilisé pour les cultures de *Mycobacterium smegmatis*. LB est composé d'extrait de levure (5g/L), de NaCl (10 g/L) et de Bacto-tryptone (10g/L), et le pH est ajusté à 7.0. Le LB agar (milieu solide) contient du Bacto-agar (15g/L). Les milieux ont été stérilisés par autoclavage de 30 min à 120°C sous une pression de 1 atm.

Les antibiotiques nécessaires sont ajoutés de façon stérile après filtration sur membrane 0,22 µM dans certaines cultures aux concentrations finales de 100 µg/mL d'ampicilline et 40 µg/mL de kanamycine.

Les cultures d'*E.coli* liquides sont cultivées à 37°C à une agitation de 200 rpm.

Les cultures de *M. smegmatis* sont cultivées à 37°C, 200 rpm ; typiquement pendant trois jours. Pour les cultures réalisées en présence des molécules CO-RM-401, les CO-RMs sont ajoutées directement au moment de la mise en culture à une concentration finale de 50 µM. Pour les cultures en présence de CO gazeux (5% et 25%), le volume de CO gazeux correspondant est ajouté en début de culture en fonction du volume de l'Erlenmeyer stérile utilisé, et ensuite fermé avec un bouchon en caoutchouc stérile, permettant ainsi de prélever des aliquotes avec une seringue.

5.1.2. Cartes des plasmides et gènes utilisés

Pendant les expériences de cette thèse, plusieurs plasmides ont été utilisés :

- a) Le vecteur pQE80L (QIAGEN) (Figure 5.1) contenant une étiquette 6xHis et un gène de résistance à l'ampicilline permettant la sélection des transformants dans les cellules compétentes d'*E. coli*. Dans ce plasmide d'expression, les gènes d'intérêt synthétisés (par Genecust) ont été clonés.

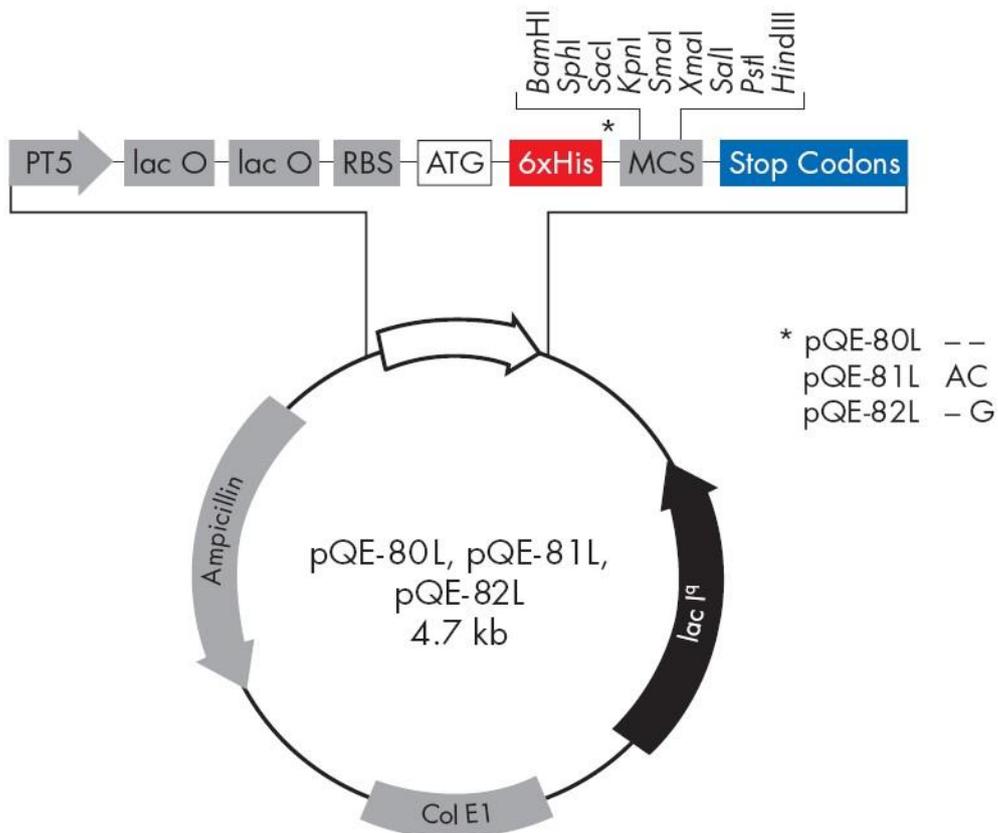


Figure 5.1 : Carte du plasmide d'expression pQE-80L (QIAGEN).

- b) Le vecteur pGEM-T Easy (Promega ; figure 5.2) est utilisé pour le clonage des produits de PCR durant les expériences d'inactivation du gène *cor*. Ce plasmide contient plusieurs sites de restriction dans un site de clonage multiple (MCS), les promoteurs SP6 et T7, nécessaires pour les séquençages, et également un gène de résistance à l'ampicilline.

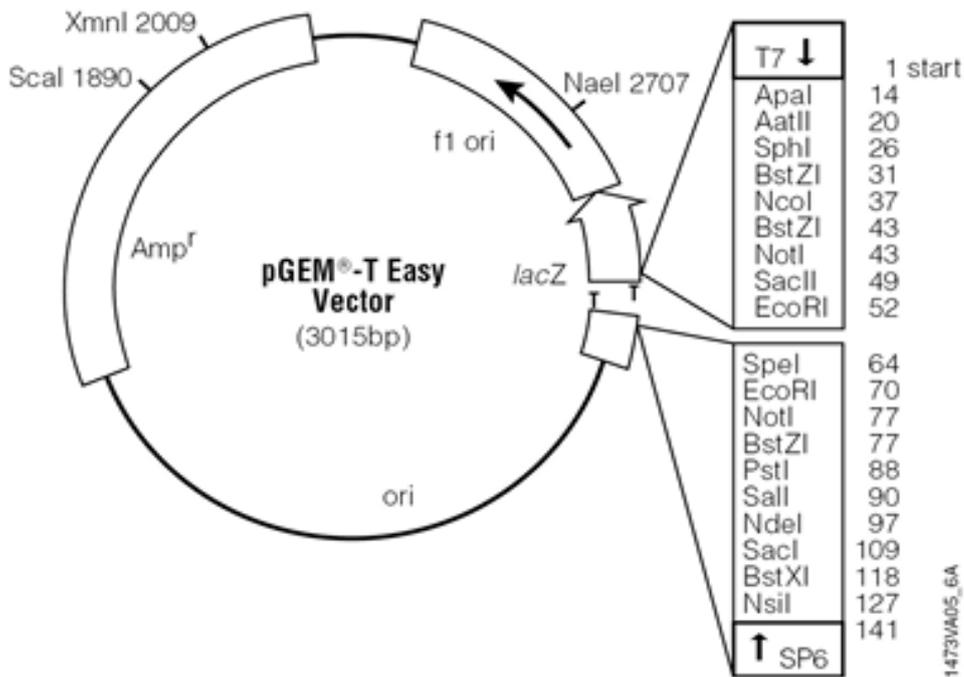


Figure 5.2 : Carte du plasmide pGEM-T Easy (Promega)

c) Le vecteur pUC18K est un dérivé du plasmide pUC18 avec une cassette de résistance à la kanamycine (gène de résistance *aphA-3* de 850 pb) cloné dans les sites de restriction *SmaI* [98].

Cette cassette de résistance à la kanamycine a été extraite et purifiée durant les expériences d'inactivation du gène *cor* pour la création de la souche knock-out du *cor*. La figure 5.3 montre une représentation schématique de la cassette kanamycine dans le plasmide pUC18K.

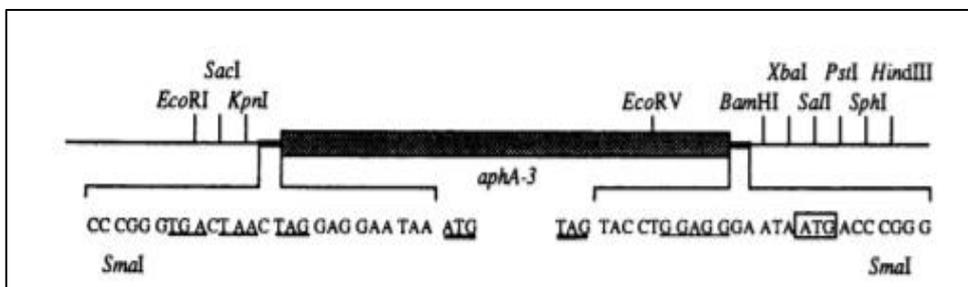


Figure 5.3 : Représentation schématique de la cassette kanamycine (gène *aphA-3*) cloné dans le vecteur pUC18K dans les sites de restriction *SmaI* [98].

Les gènes utilisés durant les expériences, ont été synthétisés par Genecust et clonés dans le vecteur d'expression pQE-80L (voir Tableau 5.1).

Gènes	Taille (pb)	Position génomique	Espèce
<i>rv1829</i>	495	2073943	<i>Mtb</i> H37
<i>rv0081</i>	345	89575	<i>Mtb</i> H37
<i>rv1828</i>	744	2073081	<i>Mtb</i> H37
<i>rcom-2</i>	804	34233-35036	<i>B. xenovorans</i> (st. LB400)

Tableau 5.1 : Gènes utilisées et leurs positions ; synthétisées par Genecust et clonés dans pQE-80L.

5.1.3. Souches des bactéries utilisées

Les souches d'*Escherichia coli* :

- BL21 (DE3) : F⁻*ompT hsdSB (r_B-m_B) gal dcm (λclts857 nin5, ind1, Sam lac UV5-T7gene1)*. Cette souche possède le gène codant pour l'ARN-polymérase du bactériophage T7, sous contrôle du promoteur de l'opéron lactose [99].
- XL1-Blue : F' Tn10 pro+B lacI^s Δ (lacZ) M15/recA1 endA1 gyrAA96 (NaI) thi hsdR17 (r_k⁻ m_k) supE44 relA1a. Cette souche a servi à la transformation et l'amplification des produits de mutagenèse dirigée (Stratagene).
- JM109 : *endA1, recA1, gyrA96, thi, hsdR17 (r_k⁻, m_k⁺), relA1, supE44, Δ(lac-proAB), [F' traD36, proAB, laqI^qZΔM15]* (Promega).

5.1.4. Extraction de l'ADN génomique de *M. smegmatis*

L'ADN génomique de *M. smegmatis* est extrait à partir de 50 ml de culture dans du milieu LB à 37°C / 200 rpm, cultivée jusqu'à une D.O à 600 nm entre 0.8 et 1.

- 1) Les cellules sont centrifugées à 5000g à 4°C pour une durée de 20 min et récupérées afin de faire l'extraction de l'ADN génomique par le Kit GeneJet Genomic DNA Purification (Thermo Scientific).
- 2) Le culot des cellules issu après centrifugation est suspendu dans du tampon de lyse (20 mM Tris/ HCl, pH8, 2 mM EDTA, 1.2 % Triton X-100), puis incubé 30 min à 37°C.
- 3) Ensuite ajouter 200 µL de la solution de lyse et du protéinase K, les mélanger et vortexer pour obtenir une solution homogène.
- 4) Incuber l'échantillon à 56°C pendant la nuit en utilisant un thermomixer jusqu'au lyse complète des cellules.
- 5) Ajouter 20 µL RNase A, mélanger bien et incubé 10 min à température ambiante.
- 6) Ajouter 400 µL de l'éthanol (50%) en agitant bien les échantillons, et ensuite transférer le lysat sur les colonnes de purification insérées dans des tubes de c.
- 7) Les colonnes sont centrifugées 1 min à 6000g, puis 500 µL du tampon de lavage est ajouté avant une nouvelle centrifugation à 8000g. Le 'flow through' est éliminé, et cette étape de lavage est réalisée une deuxième fois.
- 8) Les colonnes sont transférées dans un nouveau tube Eppendorf (1,5 mL) et après ajout du tampon d'éluion et centrifugation à 8000 g pour 1 min, l'ADN génomique est récupéré.
- 9) La concentration est déterminée par mesure sur Nanodrop à 260 nm et l'ADN testé sur gel d'agarose 1%. L'ADN purifié est aliquoté et stocké à -20°C.

5.1.5. Amplification et clonage du gène *cor* (*rv1829*)

Afin de créer le knock-out du gène *cor* dans *M. smegmatis*, nous avons d'abord lancé l'amplification par PCR de *cor* en utilisant les amorces dessinées dans le Tableau 5.4 et en utilisant l'ADN génomique de *M. smegmatis* comme matrice, et l'enzyme Q5 (*Biolabs*) ; une ADN polymérase à très haute fidélité.

Les réactions de PCR ont été réalisées dans un volume final de 50 µL, contenant entre 50-100 ng d'ADN comme matrice, 200 µM de solution de dNTP, et 0.5 µM de chaque oligonucléotide.

Les conditions d'amplification de *cor* sont décrites dans le Tableau 5.2 au-dessous.

Nombre des cycles	Phase	Température	Durée
1 cycle	Dénaturation initiale	98°C	2 minutes
35 cycles	Dénaturation	98°C	10 secondes
	Hybridation	58°C	30 secondes
	Élongation	72°C	30 secondes
1 cycle	Élongation finale	72°C	2 minutes

Tableau 5.2 : Conditions d'amplification du gène *cor* par PCR.

Pour confirmer l'amplification du gène, le produit de PCR est testé sur gel d'agarose 1 % contenant du bromure d'éthidium (BET), par migration pour 40 min à 100 V, ensuite purifié par le kit de GeneJet PCR purification (Thermo Scientific), puis dosé par Nanodrop.

Afin de réaliser le clonage de *cor* dans le vecteur pGEM-T Easy, une étape de rajout d'une queue polyA à chaque extrémité du produit de PCR est nécessaire.

L'addition de polyA est réalisée en préparant 50µL d'une réaction contenant 1 pmol de produit de PCR, 0.2 mM d'ATP, du tampon avec Mg²⁺ et l'enzyme Taq ADN Polymérase (1 U). Ce mélange est incubé 20 min à 72°C.

Cette étape est suivie par celle de l'insertion du *cor* dans le vecteur pGEMT-Easy. A cet effet la réaction a été incubée à température ambiante pour une heure, contenant 10 µL du : produit de PCR avec polyA (1µL), le tampon de ligation (5µL), l'enzyme T4 DNA ligase (1µL), le vecteur pGEM-T Easy (1µL) et eau distillée (2µL). Le rapport entre l'insert *cor* et le vecteur était stoechiométrique (1 : 1).

Les transformations des produits de ligation (insert *cor*-pGEMT-Easy) ont été réalisées en utilisant la technique de transformation par choc thermique (décrite dans la partie 5.1.6.1).

5.1.6. Transformation bactérienne

Les transformations bactériennes ont été réalisées par deux techniques :

1) Transformation par choc thermique (surtout pour l'insertion du gène *cor* dans le vecteur pGEM-T Easy) : sur glace, ajouter 2 µL de la réaction de la ligation (plasmide-gène *cor* ; entre 30-200 ng/µL) à 50 µL de cellules compétentes d'*E. coli* souche JM109 (Promega) dans un tube Eppendorf de 1,5 mL, puis incuber le mélange 20 min sur glace pour ensuite effectuer un choc thermique (45 secondes) dans un bain-marie à 42°C, ensuite l'incuber de nouveau 2 min sur glace.

Ajouter 950 µL de milieu LB liquide stérile et laisser incubé à 37°C pour 1h30 min sous agitation à 250 rpm.

Ensuite étaler 100 µL sur boîte LB agar (+ ampicilline+ 2%X-Gal+ 200 mM IPTG), incubé à 37°C pendant la nuit et sélectionner les colonies blanches (cellules transformées) le lendemain.

2) Transformation par électroporation ; basée sur l'utilisation des bactéries électrocompétentes *E. coli* BL21 DE3 [99] : Sur glace, ajouter 50ng de plasmide à 50µL des bactéries électrocompétentes dans une cuve d'électroporation. L'application d'un choc électrique de 2500V va former temporairement des pores dans la membrane bactérienne, permettant au plasmide de pénétrer dans les cellules. Après l'ajout de 1mL LB et agitation par pipetage, la réaction est transférée dans un tube de culture pour une incubation de 1h30 min à 37°C sous agitation de 250 rpm. Enfin, 100 µL du mélange de transformation sont étalés sur boîte LB agar (+ antibiotique de sélection) et cultivés pendant la nuit dans l'étuve à 37°C.

5.1.7. Extraction, purification de l'ADN plasmidique et digestion

Après avoir réalisé les transformations (par les deux techniques), les cellules transformées sélectionnées sur boîte, ont été mises en culture (LB+antibiotiques) à 37°C et 200 rpm, afin d'extraire et de purifier le plasmide (avec l'insert) par le kit de miniprep (Qiagen) basé sur la lyse alcaline.

5.1.8. Vérifications des géotypages (constructions plasmidiques)

La concentration des plasmides purifiés est ensuite déterminée par Nanodrop et la vérification de l'insertion du gène *cor* a été fait par deux méthodes avant l'envoi au séquençage :

- 1) Réaliser des PCR avec les amorces déjà utilisées pour l'amplification du gène *cor*. Visualisation du produit PCR sur gel d'agarose 1% pour vérifier la taille (insert *cor* amplifié)
- 2) Digestion du plasmide par l'enzyme de restriction *BglII*, qui ne possède aucun site de restriction dans le vecteur pGEM-T Easy et possède un seul site au milieu du gène *cor*.

Le séquençage a finalement confirmé l'insertion du *cor* dans le vecteur pGEM-T Easy.

5.1.9. Mutagenèse dirigée

Dans le but de créer des mutants Cor-H58A et Cor-H70A (pour tester si les deux acides aminés peuvent être impliqués dans la liaison du cofacteur hémique), ces deux histidines ont été mutées en alanine par mutagenèse dirigée.

La mutagenèse dirigée est utilisée afin de muter une séquence d'ADN à un endroit précis. Le *Quikchange Site-Directed Mutagenesis* (Stratagene) a été utilisé pour créer les mutants dirigés.

Le plasmide pQE80L avec le gène *cor* déjà cloné sert de matrice. Les réactions de PCR sont ensuite effectuées en utilisant les amorces avec des séquences mutées (dessinées dans le Tableau 5.5). L'ADN polymérase *Pfu Turbo* incorpore les amorces et permet la formation de brins complémentaires mutés.

Le programme de PCR pour les mutagenèses dirigées est décrit dans le Tableau 5.3 au-dessous.

Nombre des cycles	Phase	Température	Durée
1 cycle	Dénaturation initiale	95°C	30 secondes
25 cycles	Dénaturation	95°C	30 secondes
	Hybridation	55°C	1 minute
	Élongation	68°C	6 minutes

Tableau 5.3 : Programme de PCR pour l'introduction des mutations dirigées.

Suite à l'amplification, une digestion par l'endonucléase de restriction *DpnI* est réalisée, suivie d'une vérification sur gel d'agarose 1%. *DpnI* digère enzymatiquement et sélectivement les brins d'ADN méthylés, parentaux, non mutés. Ensuite les produits d'amplification sont transformés dans les cellules chimio-compétentes de la souche *E. coli* XL1- Blue par choc thermique et les cellules sont étalées sur boîte LB/Amp et incubées.

Les colonies sélectionnées sont ensuite soumises à une extraction d'ADN plasmidique et les plasmides sont vérifiés par séquençage. Les plasmides ainsi vérifiés sont transformés dans les cellules *E. coli* BL21 DE3 par électroporation afin de réaliser l'expression des protéines mutantes.

5.1.10. Listes des amorces utilisées

a) Amorces utilisées pour les expériences d'inactivation du gène *cor*

215 Fw	5'- agtacctgccgctcaaagtc-3'
1493 Rev	5'- ctgctggctgctgtgaaaag-3'
464 Fw	5'-tcaagaacggcgtgatcaca-3'
1652 Rev	5'-ccccagtggtcagatccaac-3'

Tableau 5.4 : Oligonucléotides utilisés dans les réactions PCR d'inactivation de *cor*, (les chiffres indiquent les numéros des nucléotides dans la séquence d'ADN de position (3646-3645-3644) chez *M. smegmatis*).

b) Amorces utilisées pour les mutagenèses dirigées :

Pour mutant Cor-H58A	5'-tccctgatcagatcagcggtcagcggacgtgg-3' 5'-ccacgtccgctgaccgctgatctgatcagga-3'
Pour mutant Cor-H70A	5'-cacctctttgagcgaagccccagcgcagcaatg-3' 5'-cattgctgcgctgggggcttcgctcaaagaggtg-3'

Tableau 5.5 : Oligonucléotides utilisés dans PCR des mutagenèses dirigées.

c) Amorces utilisées pour test d'EMSA et anisotropie de fluorescence

<i>rv 0081</i> Fw	5'-actatctgataaattctcaactcgt-3'
<i>rv 0081</i> Rev	5'-acgagttgaagaatttatcagatagt-3'
<i>rv 1828-rv1829</i> Fw	5'- cttcaccgctgaggactagacttcgttc-3'
<i>rv 1828-rv1829</i> Rev	5'-gaacgaagtctagtcctcagcgggaag-3'
<i>rv1831-rv1832</i> Fw	5'-agagggccgctagtcctctgctttgtcaggagt-3'
<i>rv 1831-rv1832</i> Rev	5'-actcctgacaaagcagaggactagcggcctct-3'
<i>rcom-2</i> Fw	5'-tcgattttcgcgcaaatt-3'
<i>rcom-2</i> Rev	5'-aatttgcgcgaaaatcga-3'

Tableau 5.6 : Oligonucléotides correspondants aux régions régulatrices d'ADN de *rv0081*, région intergénique *rv1828/1829*, région intergénique *rv1831/1832*, et région de fixation de RcoM-2 utilisées dans les expériences d'EMSA et d'anisotropie de fluorescence.

5.2. Manipulation des protéines et études biochimiques

5.2.1. Expression des protéines dans *Escherichia coli*

Les différents gènes (*rv0081*, *rv1828*, *rv1829*, *rcom-2*) codant pour nos protéines d'intérêt (*Rv0081*, *Rv1828*, *Rv1829* (Cor) et *RcoM-2*), sont synthétisés par Genecust et optimisés pour une expression dans *E. coli*. Les gènes mutés ont été clonés dans le vecteur d'expression pQE-80L et transformés dans *E. coli* souche BL21 DE3 par électroporation. Pour l'expression des protéines correspondantes, une colonie isolée de chaque transformation a été diluée dans une pré-culture de 2 ml LB/Amp et laissée croître pendant la journée à 37°C, 200 rpm.

Les pré-cultures sont ensuite diluées dans 50 ml LB/Amp et laissées croître pendant la nuit à 37°C, 200 rpm avant un transfert dans 1 litre LB/Amp et croissance à 37°C, 200 rpm.

Quand la culture a une D.O.₆₀₀ de 0.6-0.7, de l'IPTG est ajouté à une concentration finale de 1 mM et l'induction continue pendant 4 heures à 37°C, 200 rpm.

Récupérer les cellules par centrifugation à 5000 rpm/4°C pour 30 min et reprendre dans 10 ml de tampon LEW (Lysis-Equilibration-Wash Buffer : 50 mM NaH₂PO₄, 300 mM NaCl, pH 8.0) pour purification.

5.2.2. Purification des protéines et détermination de la concentration

Après suspension du culot dans 10 ml de LEW, du lysozyme de concentration finale 1 mg/ml et du DNase de concentration 5 µg/ml sont ajoutés et la suspension est incubée pendant 30 minutes sur glace. Après une sonication de 6 fois 30 sec (avec des intervalles de refroidissement 10 sec), le lysat est centrifugé pour 50 minutes à 4°C à 12000 rpm afin de récupérer le surnageant contenant la protéine d'intérêt.

La purification des protéines recombinantes s'effectue sur des colonnes *Protino* Ni-TED Resin 2000 (Macherey-Nagel) par chromatographie d'affinité. Suite à la fixation, lavage et élution dans un tampon 50 mM NaH₂PO₄, 300 mM NaCl, 250 mM imidazole, pH 8.0, nous faisons une étape d'élimination d'imidazole en utilisant les colonnes d'Econo-Pac 10 DG (Bio-Rad). La protéine ainsi purifiée est suspendue dans un tampon TBS (50 mM Tris/HCl, 150 mM

NaCl, pH 8.0) et concentrée en utilisant les micro-concentrators sur filtre AMICON à 10 kDa d'exclusion de masse moléculaire.

Les concentrations des protéines purifiées sont déterminées par des tests Bradford et par spectroscopie d'absorption grâce aux mesures de l'absorbance à 280 nm. Le calcul de la concentration C est effectué grâce à la loi de Beer Lambert : $A = \epsilon \cdot L \cdot C$ (A : l'absorbance, L : la distance optique de la cuve utilisée en cm, ϵ : le coefficient d'absorption molaire de la protéine en $\text{mol}^{-1} \cdot \text{cm}^{-1}$). ϵ_{280} est trouvé à partir des extinctions connues des acides aminés aromatiques (Tyr, Trp, Phe) présents dans la séquence. Ensuite une analyse sur SDS-PAGE est effectuée.

5.2.3. Analyses des protéines purifiées par SDS-PAGE

Afin de séparer les protéines selon leurs masses moléculaires et d'analyser la pureté des protéines Cor (wt et mutants), RcoM-2, Rv1828 et Rv0081 purifiées, des migrations sur gels d'électrophorèse en conditions dénaturantes (SDS-PAGE) ont été effectuées.

Pour la protéine Rv0081 (masse moléculaire 12.1 kDa) un gel SDS-PAGE 15% (plus résolutif) a été nécessaire avec migration pour 1 heure à 80V.

Pour les protéines Cor (wt et mutants) de masse moléculaire 18.114 kDa, Rv1828 de 26.4 kDa et RcoM-2 de 30.5 kDa, un gel SDS-PAGE 10% avec migration pour 1 heure à 90V a été utilisé.

Ensuite la détection des protéines a été réalisée en colorant les gels au Bleu de Coomassie et décoloration par un tampon de 5% d'acide acétique et 30% d'éthanol.

5.2.4. Technique d'EMSA

La technique du « Electrophoretic Mobility Shift Assay » (EMSA) est basée sur l'observation que des complexes ADN-protéine migrent plus lentement dans une électrophorèse sur gel de polyacrylamide non-dénaturant, comparé avec l'ADN libre linéaire, non lié aux protéines. Étant donné que la migration de l'ADN est « retardée » lorsqu'il est lié

à une protéine, le test est également appelé « retard sur gel ». Il comprend trois étapes principales : (1) réactions de liaison, (2) électrophorèse, (3) analyse d'images.

Appariement des oligonucléotides : Pour les expériences d'EMSA, les oligonucléotides utilisés, synthétisés par Eurogentec (voir Tableau 5.6), doivent être appariés de façon équimolaire. Les deux brins d'oligonucléotides complémentaires (dont un brin a été marqué par Cy 5.5 en 5') sont mélangés, chauffés à 94°C pour 2 minutes et laissés refroidir graduellement à température ambiante pour former de l'ADN double brin.

Le stock des amorces a été préparé dans du tampon Tris/HCl pH7,6 (50 mM), NaCl (100 mM).

Les oligonucléotides marqués ont été utilisés dans les réactions d'EMSA. Différents ratios ADN-protéine ont été testés. Typiquement une réaction de liaison (volume totale de 20 µL) est composée de :

- 10x tampon de liaison (100 mM Tris, 500 mM KCl, 10mM DTT ; pH7,5 : 2 µL
- 25mM DTT/2,5% Tween 20 : 2 µL
- 1 µg/µL Poly (dl.dC) dans 10 mM Tris, 1 mM EDTA; pH 7.5 : 1µL
- 100 µg/mL BSA : 1 µL
- ADN : 5 nM
- Protéine : 500 ng

Compléter à 20 µL par de l'eau ultra-pure.

Le contrôle négatif est préparé sans protéine (ADN seul) ; le contrôle positif pour le facteur de transcription Rv0081 avec sa région promotrice. Comme contrôles, des réactions en présence d'ADN non-marqué ont été également effectuées. Les réactions de liaison sont réalisées à 50°C pour 30 minutes.

Pour l'électrophorèse, 2µL de marqueur (« Orange loading dye » 10x) sont ajoutés et les réactions sont analysées ensuite par migration sur gel d'électrophorèse non dénaturant (5% TBE) avec un faible voltage de 10V/cm dans la chambre froide en absence de lumière. Les gels sont ensuite analysés en utilisant le système LI-COR Odyssey 1249 (Science Tec) avec une détection à 700 nm.

5.3. Approches spectroscopiques

5.3.1. Spectroscopie d'absorption

Les mesures des spectres d'absorption à l'équilibre ont été réalisées en utilisant le spectromètre Shimadzu UV-vis 1700 dans des cuves de 1cm x 1cm.

Le complexe Cor-hème (concentration stoechiométrique) en forme réduite a été préparé par l'ajout d'une solution d'hydrosulfite de sodium (dithionite) par une seringue Hamilton dans la cuve fermée de façon étanche par un bouchon de caoutchouc. Les solutions étaient dégazées sous vide plusieurs fois, et l'air était remplacé par de l'argon. L'ajout de dithionite (et la réduction associée du complexe) est vérifié toujours par la présence d'un pic de l'absorption de dithionite à 310 nm.

Pour les expériences des protéines en forme CO-ligandée, la forme carboxylée a été obtenue à partir de la forme réduite et l'équilibrage avec 1 atm du CO.

5.3.2. Anisotropie de fluorescence

La technique d'anisotropie de fluorescence permet de détecter la liaison des protéines à des régions d'ADN (marquées par un fluorophore en 5') d'une façon quantitative. La constante d'affinité (K_d) de cette liaison est mesurée par un calcul de valeur d'anisotropie r en fonction de la concentration de protéine. Le principe et la mise en œuvre de cette méthodologie sont décrits en détail dans [100].

À noter que cette technique est particulièrement avantageuse dans notre cas car on peut l'utiliser en présence du CO, ce qui n'est pas possible avec la technique EMSA.

Pour les expériences avec l'ADN portant le fluorophore Cy5.5, l'excitation a été réalisée par une lumière centrée à 680 nm avec une fente de 10 nm et l'émission mesurée entre 695 nm et 800 nm, avec une valeur de photo-multiplicateur de 850 V.

Pour les expériences avec l'ADN portant le fluorophore Texas Red, l'excitation a été réalisée par une lumière centrée à 590 nm avec une fente de 10 nm et l'émission mesurée entre 607 nm et 660 nm, avec une valeur de photo-multiplicateur de 800 V.

Ces expériences ont été réalisées dans des cuves de 1cmx1cm, contenant pour le début des expériences une solution d'ADN (db) de 10 nM (pour celles qui sont marqués par Cy 5.5) et 6.4 nM (pour celles qui sont marqués par T.R) dans un tampon de Tris-HCl pH7.6 (50 mM), NaCl (100 mM).

Puis les échantillons de protéines (protéines seules ou complexe : protéine-hème-CO) ont été ajoutés à des concentrations différentes par une seringue (gastight Hamilton), et l'anisotropie de fluorescence a été mesurée par le fluorimètre Cary Eclipse équipé d'un dispositif manuel de polarisation.

Ces expériences ont été réalisées à température ambiante.

Pour chaque concentration de protéines ajoutées, l'intégrale des spectres d'émission de fluorescence a été déterminé pour chaque polarisation, et enfin les valeur d'anisotropie de fluorescence ont été calculées selon les procédures établies [100].

5.3.3. Absorption résolue en temps

Pour les échantillons de RcoM-2 déjà purifiés, la fraction sous forme ferreuse-CO a été transformé en forme ferrique par une illumination pendant 12 h à 4°C, avec une lumière blanche d'une lampe halogène équipée d'une fibre optique avec une intensité d'environ 0.15 W/cm².

Ensuite pour la forme ferreuse non-ligandée, l'échantillon a été dégazée et du dithionite de sodium a été ajouté. Pour la forme carboxylée, la forme réduite a été exposée au CO (1 atm).

La spectroscopie d'absorption résolue en temps a été effectué grâce à deux faisceaux laser à impulsions courtes incidents sur l'échantillon (de durée ~100 fs) ; dont un faisceau est intense et centré à 560 nm dans la bande Q α pour réaliser la dissociation et l'autre faisceau est faible et de lumière blanche, dont les spectres sont enregistrés dans la zone Soret (400-500 nm) [100]. À noter que la variation du retard ou décalage entre ces deux faisceaux nous permet d'étudier la cinétique.

Les expériences d'absorption résolue en temps ont été réalisées à 20°C (coll. M. H. Vos, LOB).

5.4. Microscopie confocale

Dans le but de visualiser et comparer plusieurs paramètres de croissance de *M. smegmatis*, comme p.ex. la morphologie des cellules en présence et absence du CO, des expériences par microscopie confocale ont été réalisées (coll. P. Mahou, LOB). Pour ces expériences, des cultures de *M. smegmatis* ont été testées en absence et présence du CO-RM-401 et du CO gaz.

À partir de ces cultures, 200 μ L sont prélevées et déposées sur l'agar dans des petites boîtes de Pétri (diamètre de 15 mm). Le liquide est ensuite éliminé par aspiration en laissant un film de bactéries attaché sur l'agar. Les bactéries sont ensuite visualisées et des images sont capturées par microscopie confocale (« Zeiss Axiovert 200 inverted wide field microscope fitted with an EM-CCD camera ») à une température de 37°C, comme décrit dans la figure 5.4 ci-dessous. Les images capturées des cellules vivantes de *M. smegmatis* sont récupérées sur le logiciel Image J.

Pour les expériences en présence du marqueur membranaire fluorescent, le dye FM5-95 (2 μ g/ml) avec une maximum d'excitation à 560 nm est ajouté à l'agar au début des expériences, puis les cellules sont déposées comme déjà décrit.

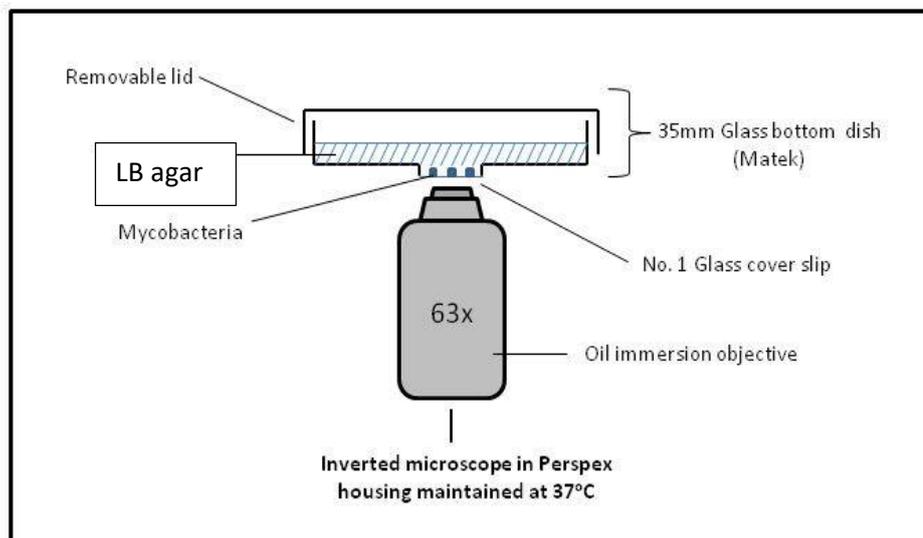


Figure 5.4 : Description de l'utilisation de microscopie confocale pour visualiser les cellules de *M. smegmatis* vivantes (adapté de [101]).

5.5. Ultracentrifugation analytique de la protéine Cor

Dans le but d'étudier le comportement et l'état associatif de la protéine Cor, des expériences d'ultracentrifugation analytique ont été réalisées (en collaboration avec Magali Aumont-Nicaise et Christophe Velours, I2BC, PIM, U. Paris Sud).

Les vitesses de sédimentation des échantillons de Cor à 40 μM , 59.3 μM et 97 μM ont été réalisées par une ultracentrifugeuse analytique XLA70 (Beckman Coulter, Palo Alto, USA) avec un rotor An-50Ti et des cellules équipées de pièces centrales à deux canaux. 400 μL d'échantillon ont été utilisés pour chaque expérience et un scan a été réalisé toutes les 5 minutes. L'ultracentrifugation est réalisée à 42000 rpm (128297 g) à 20°C et les analyses de résultats sont réalisées avec le logiciel Sedfit.

5.6. Cristallisation des protéines

Dans le but d'obtenir des structures à 3D des protéines Rv0081 et Cor (Rv1829), des expériences de cristallogénèse (en collaboration avec Ines Gallay et Stéphane Planqueel, I2BC, U. Paris Sud) ont été réalisées. Plusieurs kits ont été testés avec différentes concentrations de protéines à 4°C et 18°C.

Les protéines ont été purifiées le même matin et manipulées fraîchement (sans congélation antérieure pour assurer une stabilité maximale) avec les kits de cristallisation.

Pour la protéine Cor (18 kDa), les kits utilisés étaient : Classic (4°C), PEG I (4°C), PEG II (4°C), Stura macroSol (4°C), JCSG+ (4°C), MB Class I (4°C), MB Class II (4°C). Les concentrations de Cor testées étaient : 10.2 $\mu\text{g}/\mu\text{L}$, 5.1 $\mu\text{g}/\mu\text{L}$, 2.5 $\mu\text{g}/\mu\text{L}$.

Pour la protéine Rv0081, les kits utilisés étaient : Classic (4°C), JCSG+ (18°C), MB Class I (18°C), MB Class II (18°C), Stura macroSol (18°C), PEG I (4°C), PEG II (4°C). Les concentrations de Rv0081 testées étaient : 5.7 $\mu\text{g}/\mu\text{L}$, 2.8 $\mu\text{g}/\mu\text{L}$, 1.4 $\mu\text{g}/\mu\text{L}$.

Les différents essais de cristallogénèse ont été préparés à température ambiante et ensuite mis en chambre froide (pour les essais à 4°C).

Le tampon utilisé était 30 mM HEPES pH 8, 0.3M NaCl pour Cor et Rv0081. Des plaques de microtitrage à 96 puits ont été utilisées.

La diffraction des cristaux obtenus a été déterminée au synchrotron Soleil.

5.7. Modélisation moléculaire de Cor

Pour modéliser une structure possible de la protéine Cor, le programme Modeller (version 9.14) a été utilisé. Ce programme permet de générer un modèle à partir d'une séquence protéique dont la structure est inconnue en utilisant des séquences de structures connues et présentant des homologies de séquence. La recherche de séquences homologues à la séquence de Cor dans Protein Data Bank a conduit à la structure 1SJ5 (figure 1.25 dans le chapitre 1) qui décrit la structure d'une protéine dimérique de *T. maritima* avec une résolution de 2.80 Å.

Avec un bon taux d'homologie d'environ 33.3% et le choix du modèle qui présente le meilleur score molpdf et qui a un bon diagramme de Ramachandran, nous avons utilisé par la suite le programme Chimera pour construire un dimère par superposition sur les chaînes A et B. Ensuite, la disposition relative des deux monomères a été modélisée avec une approche 'Coarse Grain' (coll. J-C Lambry, LOB).

Chapitre VI

Références bibliographiques

- [1] Farhana, A. ; Saini, V. ; Kumar, A. ; Lancaster, J. R. and Steyn, A. J. C. : Environmental Heme-Based Sensor Proteins : Implications for Understanding Bacterial Pathogenesis, Antioxidants & Redox Signaling, 17 (2012), no. 9, pp. 1232–1245.
- [2] Gilles-Gonzalez, M.-A. and Gonzalez, G. : Heme-based sensors : defining characteristics, recent developments, and regulatory hypotheses, Journal of Inorganic Biochemistry, 99 (2005), no. 1, pp. 1–22.
- [3] Carver, P. L. : The Battle for Iron between Humans and Microbes, Current Medicinal Chemistry, 25 (2018), no. 1, pp. 85–96.
- [4] Sheldon, J.R. and Heinrichs, D. E. : Recent developments in understanding the iron acquisition strategies of gram positive pathogens, FEMS microbiology reviews, 39 (2015), no. 4, pp. 592–630.
- [5] Hémoprotéine Wikipédia, 2017.
- [6] Vadon-Le Goff, S. and Sennequier, N. : Biosynthèse du monoxyde d’azote (NO) : mécanisme, régulation et contrôle., Revue médecine/sciences, 14 (1998), no. 11, p. 1185.
- [7] Marletta, M. A. : Nitric oxide synthase : Aspects concerning structure and catalysis, Cell, 78 (1994), no. 6, pp. 927–930.
- [8] Laskin, J. D. ; Heck, D. E. ; Laskin, D. L. : Multifunctional role of nitric oxide in inflammation, Trends in Endocrinology & Metabolism , Volume 5, Issue 9, November 1994, Pages 377-382.
- [9] Carbon Monoxide Poisoning and Hyperbaric oxygen Therapy, Cancer Care of WESTERN NEW YORK, by Annie, Stuart, (December 2017).
- [10] Soares, M. P. and Bach, F. H. : Heme oxygenase-1 : from biology to therapeutic potential, Science Direct, Volume 15, Issue 2, February 2009, Page 50-58.
- [11] Roberts, G. P. ; Youn, H. and Kerby, R. L. : CO-Sensing Mechanisms, Microbiol Mol Biol Rev. 2004 Sep ; 68(3) : 453–473
- [12] Wareham, L. K. ; Begg, R. ; Jesse, H. E. ; Van Beilen, J. W. A. ; Ali, S. ; Svistunenko, D. ; McLean, S. ; Hellingwerf, K. J. ; Sanguinetti, G. and Poole, R. K. : Carbon Monoxide Gas Is Not Inert, but Global, in Its Consequences for Bacterial Gene Expression, Iron Acquisition, and Antibiotic Resistance, Antioxidants & Redox Signaling, 24 (2016), no. 17, pp. 1013–1028.
- [13] Jesse, H. E ; Nye, T. L. ; McLean, S. ; Green, J. ; Mann, B. E. and Poole, R. K. : Cytochrome bd-I in Escherichia coli is less sensitive than cytochromes bd-II or bo” to inhibition by the carbon monoxide-releasing molecule, CORM-3, Biochim Biophys Acta. 2013 Sep ; 1834(9) : 1693–1703.

- [14] Gusarov, I. ; Shatalin, K. ; Starodubtseva, M. ; and Nudler, E. : Endogenous Nitric Oxide Protects Bacteria Against a Wide Spectrum of Antibiotics, *Science*. 2009 Sep 11 ; 325 (5946) : 1380–1384.
- [15] Shimizu, T. : The Heme-Based Oxygen-Sensor Phosphodiesterase Ec DOS (DosP) : Structure-Function Relationships, *Biosensors*, 3 (2013), no. 2, pp. 211–237.
- [16] David, M. ; Daveran, M. L. ; Batut, J. ; Dedieu, A. ; Domergue, O. ; Ghai, J. ; Hertig, C. ; Boistard, P. ; Kahn, D. : Cascade regulation of *nif* gene expression in *Rhizobium meliloti*, *Cell*. 1988 Aug 26 ; 54 (5) : 671-83.
- [17] Aono, S. ; Matsuo, T. ; Shimono, T. ; Ohkubo, K. ; Takasaki, H. and Nakajima, H. : Single transduction in the transcriptional activator *CooA* containing a heme- Based CO Sensor : isolation of a dominant positive mutant which is active as the transcriptional activator even in the absence of CO, *BIOCHEMICAL AND BIOPHYSICAL RESEARCH COMMUNICATIONS* 240, 783–786 (1997).
- [18] Shelver, D ; Kerby, R. L ; He, Y ; Roberts, G.P. : Carbon monoxide-induced activation of gene expression in *Rhodospirillum rubrum* requires the product of *cooA*, a member of the cyclic AMP receptor protein family of transcriptional regulators, *J Bacteriol*. 1995 Apr ; 177 (8) :2157- 63.
- [19] Iyer, L. M. ; Anantharaman, V. and Aravind, L. : Ancient conserved domains shared by animal soluble guanylyl cyclases and bacterial signaling proteins, *BMC Genomics*. 2003 Feb 3 ; 4 (1) : 5
- [20] Karow, D. S. ; Pan, D. ; Tran, R. ; Pellicena, P. ; Presley, A. ; Mathies, R. A. and Marletta, M. A. : Spectroscopic Characterization of the Soluble Guanylate Cyclase-like Heme Domains from *Vibrio cholerae* and *Thermoanaerobacter tengcongensis*, *Biochemistry* 2004, 43, 31, 10203-10211.
- [21] Stock, J. B. ; Ninfa, A. J. and Stock, A. M. : Protein phosphorylation and regulation of adaptive responses in bacteria, *Microbiol Rev*. 1989 Dec ; 53 (4) : 450-90.
- [22] Gilles-Gonzalez, M. A. ; Ditta, G. S. and Helinski, D. R. : A haemoprotein with kinase activity encoded by the oxygen sensor of *Rhizobium meliloti*, *Nature* 1991 Mars 14 350, 170–172.
- [23] Tuckerman, J. R. ; Gonzalez, G. ; Dioum, E.M. and Gilles-Gonzalez, M. A. : Ligand and oxidation-state specific regulation of the heme-based oxygen sensor FixL from *Sinorhizobium meliloti*, *Biochemistry*, 41 (2002), no. 19, pp. 6170–6177.
- [24] Gong, W. ; Hao, B. ; Mansy, S. S. ; Gonzalez, G. ; Gilles-Gonzalez, M. A. and Chan, M. K. : Structure of a biological oxygen sensor : A new mechanism for heme-driven signal transduction, *Proceedings of the National Academy of Sciences*, 95 (1998), no. 26, pp. 15177–15182.
- [25] Pellequer, J. L. ; Brudler, R. and Getzoff, E. D. : Biological sensors: More than one way to sense oxygen - *Curr Biol*. 1999 Jun 3 ; 9 (11) : R416-8.
- [26] Ma, X. ; Sayed, N. ; Beuve, A. ; Van den Akker, F. : NO and CO differentially activate soluble guanylyl cyclase via a heme pivot-bend mechanism, *The EMBO Journal*, 2007 Jan 24;26(2):578-88. Epub 2007 Jan 11.

- [27] Hardman, J. G and Sutherland, E.W. : Guanyl cyclase, an enzyme catalyzing the formation of guanosine 3',5'-monophosphate from guanosine triphosphate, *J Biol Chem.* 1969 Dec 10 ;244 (23) : 6363-70.
- [28] Karow, D. S. ; Pan, D. ; Davis, J. H. ; Behrends, S. ; Mathies, R. S. and Marletta, M. A. : Characterization of Functional Heme Domains from Soluble Guanylate Cyclase, *Biochemistry.* 2005 Dec 13 ; 44 (49) : 16266–16274.
- [29] Denninger, J. W. and Marletta, M. A. : Guanylate cyclase and the ·NO/cGMP signaling pathway, *Biochimica et Biophysica Acta (BBA) - Bioenergetics*, Volume 1411, Issues 2–3, 5 May 1999, Pages 334-350.
- [30] Moncada, S. and Higgs, A. : Nitric Oxide : Role in Human Disease, *ENCYCLOPEDIA OF LIFE SCIENCES*, 2003, John Wiley & Sons.
- [31] Seeger, F. and Garcin, E. D. : Soluble guanylate cyclase crystal clear : 1st crystal structure of the wild-type human heterodimeric sGC catalytic domains and implications for activity, *BMC Pharmacology Toxicology*, 2013 ; 14 (Suppl 1) : O14.
- [32] Dioum, E. M. ; Rutter, J. ; Tuckerman, J. S. ; Gonzalez, G. and Gilles-Gonzalez, M. A. ; Steven, L. McKnight : NPAS2 : A Gas-Responsive Transcription Factor, *Science*, 20 Dec 2002 : Vol. 298, Issue 5602, pp. 2385-2387.
- [33] Lanzilotta, W. N. ; Schuller, D. J. ; Thorsteinsson, M. V. ; Kerby, R. L. ; Roberts, G. P. and Poulos, T. L. : Structure of the CO sensing transcription activator CoxA, *Nature Structural Biology*, **7** (2000), no. 10, p. 876.
- [34] Tenhunen, R. ; Marver, H. S. and Schmid, R. : Microsomal Heme Oxygenase, *Journal of Biological Chemistry*, December 10, 1969, 6388-6394.
- [35] Wang, J. and Montellano, P. R. : The Binding Sites on Human Heme Oxygenase-1 for Cytochrome P450 Reductase and Biliverdin Reductase, *Journal of Biological Chemistry* May 30, 2003, 20069-20076.
- [36] Nishino, H. and Ishibashi, T. : Evidence for Requirement of NADPH-Cytochrome P450 Oxidoreductase in the Microsomal NADPH-Sterol Δ^7 -Reductase System, *Archives of Biochemistry and Biophysics*, Volume 374, Issue 2, 15 February 2000, Pages 293-298.
- [37] Boehning, D. and Snyder, S. H. : Carbon Monoxide and Clocks, *Science*, 20 Dec 2002 : Vol. 298, Issue 5602, pp. 2339-2340.
- [38] Minegishi, S. ; Sagami, I. ; Negi, S. ; Kano, K. ; Kitagishi, H. : Circadian clock disruption by selective removal of endogenous carbon monoxide, *Scientific Reports* : volume 8, Article number : 11996 (2018).

- [39] Puranik, M. ; Nielsen, S.B. ; Youn, H. ; Hvitved, A.N. ; Bourassa, J.L. ; Case, M.A. ; Tengroth, C. ; Balakrishnan, G. ; Thorsteinsson, M.V. ; Groves, J.T. ; McLendon, G.L. ; Roberts, G.P. ; Olson, J.S. and Spiro, T. G. : Dynamics of Carbon Monoxide Binding to CooA, *Journal of Biological Chemistry*, 279 (2004), no. 20, pp. 21096–21108.
- [40] Kerby, R. L. ; Ludden, P. W. and Roberts, G. P. : Carbon monoxide-dependent growth of *Rhodospirillum rubrum*., *Journal of Bacteriology*, 177 (1995), no. 8, pp. 2241–2244.
- [41] Thorsteinsson, M. V. ; Kerby, R. L. ; Conrad, M. ; Youn, H. ; Staples, C. R. ; Lanzilotta, W. N. ; Poulos, T. J. ; Serate, J. and Roberts, G. P. : Characterization of Variants Altered at the N-terminal Proline, a Novel Heme-Axial Ligand in CooA, the CO-sensing Transcriptional Activator, *Journal of Biological Chemistry*, 275 (2000), no. 50, pp. 39332–39338.
- [42] Chan, M. K. : CooA, CAP and allostery, *Nature Structural Biology*, volume 7, pages 822–824 (2000).
- [43] Borjigin, M. ; Li, H. ; Lanz, N. D. ; Kerby, R. L. ; Roberts, G.P. and Poulos, T. L. : Structure-based hypothesis on the activation of the CO-sensing transcription factor CooA, *Acta Crystallographica Section D : Biological Crystallography*, 63 (2007), no. 3, pp. 282–287.
- [44] Liebl, U. ; Lambry, J.-C. and Vos, M. H. : Primary processes in heme-based sensor proteins, *Biochimica et Biophysica Acta (BBA) - Proteins and Proteomics*, 1834 (2012), no. 9, pp. 1684–1692.
- [45] Kerby, R. L. ; Youn, H. and Roberts, G. P. : RcoM : A New Single-Component Transcriptional Regulator of CO Metabolism in Bacteria, *Journal Bacteriology*, 2008 May ; 190 (9) : 3336–3343.
- [46] Kerby, R.L. and Roberts, G. P. : *Burkholderia xenovorans* RcoMBx-1, a Transcriptional Regulator System for Sensing Low and Persistent Levels of Carbon Monoxide, *Journal of Bacteriology*, 194 (2012), no. 21, pp. 5803–5816.
- [47] A Unique CO-sensing Transcription Factor, RcoM – Burstyn Lab (Research Overview) – The UW–Madison Heme team.
- [48] Marvin, K. A. ; Kerby, R. L. ; Youn, H. ; Roberts, G. P. and Burstyn, J. N. : The transcription regulator RcoM-2 from *Burkholderia xenovorans* is a cysteine-ligated hemoprotein that undergoes a redox-mediated ligand switch, *Biochemistry*, 47 (2008), no. 34, pp. 9016–9028.
- [49] Bouzhir-Sima, L. ; Motterlini, R. ; Gross, J. ; Vos, M. V. and Liebl, U. : Unusual Dynamics of Ligand Binding to the Heme Domain of the Bacterial CO Sensor Protein RcoM-2, *The Journal of Physical Chemistry, B*, 2016, 120-41- pages 10686-10694.
- [50] Benabbas, A. ; Karunakaran, V. ; Youn, H. ; Poulos, T. L. and Champion, P. M. : Effect of DNA Binding on Geminate CO Recombination Kinetics in CO-sensing Transcription Factor CooA, *The Journal of Biological Chemistry*, 287 (2012), no. 26, pp. 21729–21740.

- [51] Floyd, K. ; Glaziou, P. ; Zumla, A. and Raviglione, M. : The global tuberculosis epidemic and progress in care, prevention, and research : an overview in year 3 of the End TB era, *The Lancet Respiratory Medicine*, 6 (2018), no. 4, pp. 299–314.
- [52] Dartois, V. : The path of anti-tuberculosis drugs : from blood to lesions to mycobacterial cells, *Nature reviews. Microbiology*, 12 (2014), no. 3, pp. 159–167.
- [53] Cole, S. T. ; Brosch, R. ; Parkhill, J. ; Garnier, T. ; Churcher, C. ; Harris, D. ; Gordon, S. V. ; Eiglmeier, K. ; Gas, S. ; Barry, C. E. ; Tekaia, F. ; Badcock, K. ; Basham, D. ; Brown, D. ; Chillingworth, T. ; Connor, R. ; Davies, R. ; Devlin, K. ; Feltwell, T. ; Gentles, S. ; Hamlin, N. ; Holroyd, S. ; Hornsby, T. ; Jagels, K. ; Krogh, A. ; McLean, J. ; Moule, S. ; Murphy, L. ; Oliver, K. ; Osborne, J. ; Quail, M.A. ; Rajandream, M.-A. ; Rogers, J. ; Rutter, S. ; Seeger, K. ; Skelton, J. ; Squares, R. ; Squares, S. ; Sulston, J. E. ; Taylor, K. ; Whitehead, S. and Barrell, B. G. : Deciphering the biology of *Mycobacterium tuberculosis* from the complete genome sequence, *Nature*, 393 (1998), no. 6685, p. 537.
- [54] Koul, A. ; Arnoult, E. ; Lounis, N. ; Guillemont, J. and Andries, K. : The challenge of new drug discovery for tuberculosis, *Nature*, 469 (2011), no. 7331, pp. 483–490.
- [55] Gengenbacher, M. and Kaufmann, S. H. E. : *Mycobacterium tuberculosis* : Success through dormancy, *Fems Microbiology Reviews*, 36 (2012), no. 3, pp. 514–532.
- [56] Nathan, C. and Shiloh, M. U. : Reactive oxygen and nitrogen intermediates in the relationship between mammalian hosts and microbial pathogens, *Proceedings of the National Academy of Sciences of the United States of America*, 97 (2000), no. 16, pp. 8841–8848.
- [57] Tenhunen, R. ; Marver, H. S. and Schmid, R. : The enzymatic conversion of heme to bilirubin by microsomal heme oxygenase., *Proceedings of the National Academy of Sciences of the United States of America*, 61 (1968), no. 2, pp. 748–755.
- [58] Zacharia, V.M. and Shiloh, M. U. : Effect of carbon monoxide on *Mycobacterium tuberculosis* pathogenesis, *Medical Gas Research*, 2 (2012), p. 30.
- [59] Brouard, S. ; Otterbein, L. E. ; Anrather, J. ; Tobiasch, E. ; Bach, F. H. ; Choi, A. M. K. and Soares, M. P. : Carbon Monoxide Generated by Heme Oxygenase 1 Suppresses Endothelial Cell Apoptosis, *The Journal of Experimental Medicine*, 192 (2000), no. 7, pp. 1015–1026.
- [60] Sun, J. ; Hoshino, H. ; Takaku, K. ; Nakajima, O. ; Muto, A. ; Suzuki, H. ; Tashiro, S. ; Takahashi, S. ; Shibahara, S. ; Alam, J. ; Taketo, M. M. ; Yamamoto, M. and Igarashi, K. : Hemoprotein Bach1 regulates enhancer availability of heme oxygenase-1 gene, *The EMBO Journal*, 21 (2002), no. 19, pp. 5216–5224.
- [61] Balogun, E. ; Hoque, M. ; Gong, P. ; Killeen, E. ; Green, C. J. ; Foresti, R. ; Alam, J. and Motterlini, R. : Curcumin activates the haem oxygenase-1 gene via regulation of Nrf2 and the antioxidant-responsive element., *Biochemical Journal*, 371 (2003), no. Pt 3, pp. 887–895.

- [62] Alam, J. ; Stewart, D. ; Touchard, C. ; Boinapally, S. ; Choi, A. M. K. and Cook, J. L. : Nrf2, a Cap'n'Collar Transcription Factor, Regulates Induction of the Heme Oxygenase-1 Gene, *Journal of Biological Chemistry*, 274 (1999), no. 37, pp. 26071–26078.
- [63] Motterlini, R. and Foresti, R. : Heme Oxygenase-1 As a Target for Drug Discovery, *Antioxidants & Redox Signaling*, 20 (2013), no. 11, pp. 1810–1826.
- [64] Voskuil, M. I. ; Schnappinger, D. ; Visconti, K. C. ; Harrell, M. I. ; Dolganov, G. M. ; Sherman, D. R. and Schoolnik, G. K. : Inhibition of Respiration by Nitric Oxide Induces a Mycobacterium tuberculosis Dormancy Program, *The Journal of Experimental Medicine*, 198 (2003), no. 5, pp. 705–713.
- [65] Park, H.-D. ; Guinn, K. M. ; Harrell, M. I. ; Liao, R. ; Voskuil, M. I. ; Tompa, M. ; Schoolnik, G. K. and Sherman, D. R. : Rv3133c/dosR is a transcription factor that mediates the hypoxic response of Mycobacterium tuberculosis, *Molecular microbiology*, 48 (2003), no. 3, pp. 833–843.
- [66] Roberts, D. M. ; Liao, R. P. ; Wisedchaisri, G. ; Hol, W. G. J. and Sherman, D. R. : Two Sensor Kinases Contribute to the Hypoxic Response of Mycobacterium tuberculosis, *The Journal of biological chemistry*, 279 (2004), no. 22, pp. 23082–23087.
- [67] He, H. ; Hovey, R. ; Kane, J. ; Singh, V. and Zahrt, T. C. : MprAB is a stress-responsive two-component system that directly regulates expression of sigma factors SigB and SigE in Mycobacterium tuberculosis, *Journal of Bacteriology*, 188 (2006), no. 6, pp. 2134–2143.
- [68] Kumar, A. ; Toledo, J. C. ; Patel, R. P. ; Lancaster, J. R. and Steyn, A. J. C. : Mycobacterium tuberculosis DosS is a redox sensor and DosT is a hypoxia sensor, *Proceedings of the National Academy of Sciences of the United States of America*, 104 (2007), no. 28, pp. 11568–11573.
- [69] Gerasimova, A. ; Kazakov, A. E. ; Arkin, A. P. ; Dubchak, I. ; Gelfand, M. S. : Comparative genomics of the dormancy regulons in mycobacteria, *J Bacteriol.* 2011 Jul ; 193(14) : 3446-52.
Doi :10.1128/JB.00179-11. Epub 2011 May 20.
- [70] Sivaramakrishnan, S. and de Montellano, P. R. O. : The DosS-DosT/DosR Mycobacterial Sensor System, *Biosensors*, 3 (2013), no. 3, pp. 259–282.
- [71] Sherman, D. R. ; Voskuil, M. ; Schnappinger, D. ; Liao, R. ; Harrell, M. I. and Schoolnik, G. K. : Regulation of the Mycobacterium tuberculosis hypoxic response gene encoding α -crystallin, *Proceedings of the National Academy of Sciences of the United States of America*, 98 (2001), no. 13, pp. 7534–7539.
- [72] Yang, H. ; Sha, W. ; Liu, Z. ; Tang, T. ; Liu, H. ; Qin, L. ; Cui, Z. ; Chen, J. ; Liu, F. ; Zheng, R. ; Huang, X. ; Wang, J. ; Feng, Y. and Ge, B. : Lysine acetylation of DosR regulates the hypoxia response of Mycobacterium tuberculosis, *Emerging Microbes & Infections*, 7 (2018), no. 1, p. 34.

- [73] Sousa, E. H. S. ; Tuckerman, J. R. ; Gonzalez, G. and Gilles-Gonzalez, M.-A. : DosT and DevS are oxygen-switched kinases in *Mycobacterium tuberculosis*, *Protein Science*, 16 (2007), no. 8, pp. 1708–1719.
- [74] Sardiwal, S. ; Kendall, S. L. ; Movahedzadeh, F. ; Rison, S. C. G. ; Stoker, N. G. and Djordjevic, S. : A GAF Domain in the Hypoxia/NO-inducible *Mycobacterium tuberculosis* DosS Protein Binds Haem, *Journal of Molecular Biology*, 353 (2005), no. 5, pp. 929–936.
- [75] Kim, M.-J. ; Park, K.-J. ; Ko, I.-J. ; Kim, Y. M. and Oh, J.-I. : Different Roles of DosS and DosT in the Hypoxic Adaptation of *Mycobacteria*, *Journal of Bacteriology*, 192 (2010), no. 19, pp. 4868–4875.
- [76] Cho, H. Y. ; Lee, Y. H. ; Bae, Y.S. ; Kim, E. ; and Kang, B.S. : Activation of ATP binding for the autophosphorylation of DosS, a *Mycobacterium tuberculosis* histidine kinase lacking an ATP lid motif, *The Journal of Biological Chemistry* 2013 May 3 ; 288 (18) : 12437-47.
- [77] Cho, H. Y. ; Cho, H. J. ; Kim, Y. M. ; Oh, J.I. and Kang, B. S. : Structural Insight into the Heme-based Redox Sensing by DosS from *Mycobacterium tuberculosis*, *The Journal of Biological Chemistry*, 284 (2009), no. 19, pp. 13057–13067.
- [78] Vos, M. H. ; Bouzhir-Sima, L. ; Lambry, J.-C. ; Luo, H. ; Eaton-Rye, J. J. ; Ioanoviciu, A. ; Ortiz de Montellano, P. R. and Liebl, U. : Ultrafast ligand dynamics in the heme-based GAF sensor domains of the histidine kinases DosS and DosT from *Mycobacterium tuberculosis*, *Biochemistry*, 51 (2012), no. 1, pp. 159–166.
- [79] Bhat, K. H. and Yaseen, I. : *Mycobacterium tuberculosis* : Macrophage Takeover and Modulation of Innate Effector Responses, *Mycobacterium - Research and Development*, (2018).
- [80] Zahrt, T. C. ; Wozniak, C. ; Jones, D. and Trevett, A. : Functional Analysis of the *Mycobacterium tuberculosis* MprAB Two-Component Signal Transduction System, *Infection and Immunity*, 71 (2003), no. 12, pp. 6962–6970.
- [81] Pang, X. and Howard, S. T. : Regulation of the α -Crystallin Gene *acr2* by the MprAB Two-Component System of *Mycobacterium tuberculosis*, *Journal of Bacteriology*, 189 (2007), no. 17, pp. 6213–6221.
- [82] Yuan, Y. ; Crane, D. D. ; Simpson, R. M. ; Zhu, Y. ; Hickey, M. J. ; Sherman, D. R. and Barry, C. E. : The 16-kDa α -crystallin (Acr) protein of *Mycobacterium tuberculosis* is required for growth in macrophages, *Proceedings of the National Academy of Sciences of the United States of America*, 95 (1998), no. 16, pp. 9578–9583.
- [83] Wilkinson, K. A. ; Stewart, G. R. ; Newton, S. M. ; Vordermeier, H. M. ; Wain, J. R. ; Murphy, H. N. ; Horner, K. ; Young, D.B. and Wilkinson, R. J. : Infection Biology of a Novel α -Crystallin of *Mycobacterium tuberculosis* : *Acr2*, *The Journal of Immunology*, 174 (2005), no. 7, pp. 4237–4243.
- [84] He, H. ; Bretl, D. J. ; Penoske, R. M. ; Anderson, D.M. and Zahrt, T. C. : Components of the Rv0081-Rv0088 Locus, Which Encodes a Predicted Formate Hydrogenlyase Complex, Are Coregulated by

- Rv0081, MprA, and DosR in *Mycobacterium tuberculosis*, *Journal of Bacteriology*, 193 (2011), no. 19, pp. 5105–5118.
- [85] Sun, X. ; Zhang, L. ; Jiang, J. ; Ng, M. ; Cui, Z. ; Mai, J. ; Ahn, S.K. ; Liu, J. ; Zhang, J. ; Liu, J. and Li, Y. : Transcription factors Rv0081 and Rv3334 connect the early and the enduring hypoxic response of *Mycobacterium tuberculosis*, *Virulence*, 9 (2018), no. 1, pp. 1468–1482.
- [86] Campbell, D. R. ; Chapman, K. E. ; Waldron, K. J. ; Tottey, S. ; Kendall, S. ; Cavallaro, G. ; Andreini, C. ; Hinds, J. ; Stoker, N.G. ; Robinson, N. J. and Cavet, J. S. : *Mycobacterial Cells Have Dual Nickel-Cobalt Sensors*, *The Journal of biological chemistry*, 282 (2007), no. 44, pp. 32298–32310.
- [87] Kumar, A. ; Phulera, S. ; Rizvi, A. ; Sonawane, P. ; Panwar, H. S. ; Banerjee, S. ; Sahu, A. and Mande, S. C. : Structural basis of hypoxic gene regulation by the Rv0081 transcription factor of *Mycobacterium tuberculosis*, *bioRxiv*, (2018), p. 465575.
- [88] Galagan, J. E. ; Minch, K. ; Peterson, M. ; Lyubetskaya, A. ; Azizi, E. ; Sweet, L. ; Gomes, A. ; Rustad, T. ; Dolganov, G. ; Glotova, I. ; Abeel, T. ; Mahwinney, C. ; Kennedy, A. D. ; Allard, R. ; Brabant, W. ; Krueger, A. ; Jaini, S. ; Honda, B. ; Yu, W.-H. ; Hickey, M. J. ; Zucker, J. ; Garay, C. ; Weiner, B. ; Sisk, P. ; Stolte, C. ; Winkler, J. K. ; Van de Peer, Y. ; Iazzetti, P. ; Camacho, D. ; Dreyfuss, J. ; Liu, Y. ; Dorhoi, A. ; Mollenkopf, H.-J. ; Drogaris, P. ; Lamontagne, J. ; Zhou, Y. ; Piquenot, J. ; Park, S. T. ; Raman, S. ; Kaufmann, S. H. E. ; Mohny, R. P. ; Chelsky, D. ; Moody, D. B. ; Sherman, D. R. and Schoolnik, G. K. : The *Mycobacterium tuberculosis* regulatory network and hypoxia, *Nature*, 499 (2013), no. 7457, pp. 178–183.
- [89] Zacharia, V. M. ; Manzanillo, P. S. ; Nair, V. R. ; Marciano, D. K. ; Kinch, L. N. ; Grishin, N. V. ; Cox, J. S. and Shiloh, M. U. : *cor*, a Novel Carbon Monoxide Resistance Gene, Is Essential for *Mycobacterium tuberculosis* Pathogenesis, *mBio*, 4 (2013), no. 6, pp. e00721-13.
- [90] You, M. K. ; Shin, H. Y. ; Kim, Y. J. ; Ok, S. H. ; Cho, S. K. ; Jeung, J. U. ; Yoo, S. D. ; Kim, J. K. and Shin, J. S. : Novel Bifunctional Nucleases, OmBBD and AtBBD1, Are Involved in Abscisic Acid-Mediated Callose Deposition in *Arabidopsis*, *Plant Physiology*, 152 (2010), no. 2, pp. 1015–1029.
- [91] Singh, S. ; Sevalkar, R. R. ; Sarkar, D. and Karthikeyan, S. : Characteristics of the essential pathogenicity factor Rv1828, a MerR family transcription regulator from *Mycobacterium tuberculosis*, *The FEBS Journal*, 285 (2018), no. 23, pp. 4424–4444.
- [92] Brown, N. L. ; Stoyanov, J. V. ; Kdd, S. P. ; Hobman, J. L. : The MerR family of transcriptional regulators, *FEMS Microbiology Reviews*, 27 (2003) 145-163
- [93] Pennella, M. A. and Giedroc, D. P. : Structural Determinants of Metal Selectivity in Prokaryotic Metal-responsive transcriptional Regulators, *Biomaterials*. 2005 Aug ; 18 (4) : 413-28.
- [94] Ventura, M. ; Rieck, B. ; Boldrin, F. ; Degiacomi, G. ; Bellinzoni, M. ; Barilone, N. ; Alzaidi, F. ; Alzari, P. M. ; Manganelli, R. and O’Hare, H. M. : GarA is an essential regulator of metabolism in *Mycobacterium tuberculosis*, *Molecular Microbiology*, 90 (2013), no. 2, pp. 356–366.

- [95] Lobato, L. ; Bouzhir-Sima, L. ; Yamashita, T. ; Wilson, M. T. ; Vos, M. H. and Liebl, U. : Dynamics of the heme-binding bacterial gas-sensing dissimilative nitrate respiration regulator (DNR) and activation barriers for ligand binding and escape, *The Journal of Biological Chemistry*, 289 (2014), no. 38, pp. 26514–26524.
- [96] Reichlin, M. : Hemoglobin and Myoglobin in Their Reactions with Ligands. Eraldo Antonini and Maurizio Brunori. North-Holland, Amsterdam, 1971 (U.S. distributor, Elsevier, New York).436 pp., illus. *Frontiers of Biology*, vol. 21, 1972.
- [97] Vos, M. H. ; Battistoni, A. ; Lechauve, C. ; Marden, M. C. ; Kiger, L. ; Desbois, A. ; Pilet, E. ; de Rosny, E. and Liebl, U. : Ultrafast heme-residue bond formation in six-coordinate heme proteins : implications for functional ligand exchange, *Biochemistry*, 47 (2008), no. 21, pp. 5718–5723.
- [98] Ménard, R. ; Sansonetti, P.J. and Parsot, C. : Nonpolar mutagenesis of the ipa genes defines IpaB, IpaC, and IpaD as effectors of *Shigella flexneri* entry into epithelial cells, *Journal of Bacteriology*, 175 (1993), no. 18, pp. 5899–5906.
- [99] Studier, F.W. and Moffatt, B. A. : Use of bacteriophage T7 RNA polymerase to direct selective high-level expression of cloned genes, *Journal of Molecular Biology*, 189 (1986), no. 1, pp. 113–130.
- [100] Lobato, L. : Internal dynamics of heme-based sensor proteins studied using advanced time-resolved optical spectroscopy, 2013, p. 127.
- [101] Joyce, G. ; Robertson, B. D. and Williams, K. J. : A modified agar pad method for mycobacterial live-cell imaging, *BMC Research Notes*, 4 (2011), no. 1, p. 73.

Annexes

Article 1 présente les résultats sur RcoM-2, publié dans « **Biochemistry** ».

Article 2 présente les résultats des expériences indépendantes des travaux de thèse, publié dans « **ChemMedChem** ».

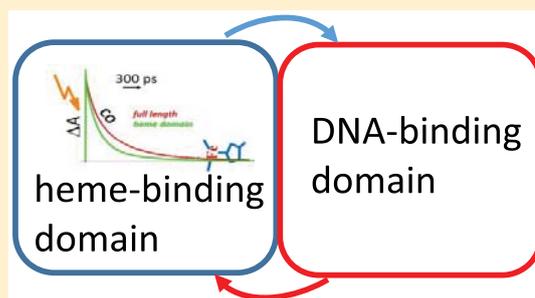
Interaction of the Full-Length Heme-Based CO Sensor Protein RcoM-2 with Ligands

Mayla Salman,[†] Carolina Villamil Franco,^{†,‡} Rivo Ramodiharilafy, Ursula Liebl, and Marten H. Vos*[§]

LOB, CNRS, INSERM, Ecole Polytechnique, Institut Polytechnique de Paris, 91128 Palaiseau, France

Supporting Information

ABSTRACT: The heme-based and CO-responsive RcoM transcriptional regulators from *Burkholderia xenovorans* are known to display an extremely high affinity for CO while being insensitive to O₂. We have quantitatively characterized the heme–CO interaction in full-length RcoM-2 and compared it with the isolated heme domain RcoMH-2 to establish the origin of these characteristics. Whereas the CO binding rates are similar to those of other heme-based sensor proteins, the dissociation rates are two to three orders of magnitude lower. The latter property is tuned by the yield of CO escape from the heme pocket after disruption of the heme–CO bond, as determined by ultrafast spectroscopy. For the full-length protein this yield is ~0.5%, and for the isolated heme domain it is even lower, associated with correspondingly faster CO rebinding kinetics, leading to K_d values of 4 and 0.25 nM, respectively. These differences imply that the presence of the DNA-binding domain influences the ligand-binding properties of the heme domain, thus abolishing the observed quasi-irreversibility of CO binding to the isolated heme domain. RcoM-2 binds target DNA with high affinity ($K_d < 2$ nM) when CO is bound to the heme, and the presence of DNA also influences the heme–CO rebinding kinetics. The functional implications of our findings are discussed.



An important subset of the large family of heme proteins is constituted by heme-based gas sensor proteins.^{1,2} These proteins play crucial roles in the initial adaptation processes of organisms to changing environmental gas conditions, usually by altering the expression of specific genes upon the binding or release of specific small molecules to protein-bound heme. Sensors of the highly reactive signaling and respiratory molecules NO and O₂ have obviously gained considerable attention. More recently, the more inert CO molecule was also established as an important sensor molecule.^{3–6} Atmospheric CO is generated by natural and industrial processes and as a eukaryotic heme breakdown product, and many prokaryotes possess distinct CO-regulated systems that catalyze CO oxidation. In mammalian systems, the transcription factor NPAS2⁷ and the SUR2A receptor-associated K_{ATP} channel⁸ have been reported to be CO-responsive (with CO affinity around 1 μ M), but the physiological relevance of these responses remains debated. In prokaryotic systems, to date, two distinct CO-specific sensor proteins, CoxA and RcoM, have been described; both are single-component transcriptional regulators that are activated upon the binding of CO to an associated heme domain. CoxA, involved in the regulation of anaerobic CO metabolism in *Rhodospirillum rubrum*, is relatively well-studied, and crystallographic models are available.^{9,10} RcoM (regulator of CO metabolism) proteins were discovered a decade ago and are active in the aerobic bacterium *Burkholderia xenovorans*.¹¹ They are less well characterized (in particular, structural data are lacking) but display some surprising properties.^{11–16}

B. xenovorans expresses two RcoM proteins, RcoM-1 and RcoM-2, which are highly homologous.¹¹ Each is composed of an N-terminal PAS heme-binding sensor domain and a C-terminal LytTR DNA-binding effector domain. RcoM-1 has been reported to form homodimers.¹⁷ In the absence of external ligands, in both the ferric and the ferrous states, the heme is six-coordinate with intrinsic amino acids as axial ligands. In its ferrous state, it can bind CO and NO while forming the six-coordinate complex with a proximal histidine.¹⁵ In the ferrous complex, the binding of an external ligand is thought to displace a distal methionine residue,^{13,15} triggering a conformational change that activates the C-terminal DNA-binding domain. In particular, for RcoM-1, it has been demonstrated that target DNA binds in the CO-bound state but not in the ferric state or the ferrous state in the absence of CO.¹⁴ The O₂ complex appears to not be stably formed, which presumably allows the protein to act as a CO sensor under aerobic conditions. Other than for CoxA and the mammalian CO-responsive systems, the CO affinity is extremely high, at least in the low nanomolar regime for full-length RcoM-1¹⁴ and even subnanomolar for the isolated heme sensor domain of RcoM-2.¹² Indeed, the proteins are purified with a high amount of ferrous CO complex present after standard heterologous expression in *Escherichia coli*. The bound CO

Received: July 22, 2019

Revised: August 22, 2019

Published: September 10, 2019



can be permanently removed from the heme only by prolonged illumination under aerobic conditions.^{11,12} Intrigued by these findings, we have previously studied the kinetic properties of CO binding to the RcoM-2 heme domain (termed RcoMH-2). This revealed unique kinetics properties, including a quasi-unity (>99%) yield of heme–CO geminate rebinding on the picosecond time scale upon photocleavage of the heme–CO bond.¹² This indicates that CO can hardly escape from the heme pocket and makes the RcoMH-2 construct a promising oxygen-compatible CO scavenger. Yet, the question arises how RcoM-2 can physiologically function as a true sensor of the CO concentration if CO binding appears quasi-irreversible. Here this question is investigated by performing ligand binding studies on the full-length RcoM-2 protein and also by determining the effects of RcoM binding to DNA. We show substantially different heme–ligand interactions in the heme sensor domain in the presence of the DNA binding domain that lead to more reversible binding of CO.

MATERIALS AND METHODS

The full-length RcoM-2 protein (amino acids 1–267, UniProtKB Q131Y4) was heterologously expressed in *E. coli* BL21 DE3 from a pQE80L expression vector (QIAGEN) using 1 mM isopropyl- β -D-thiogalactopyranoside. RcoM-2 (expected MW 30.5 kDa), containing the N-terminal heme-containing PAS domain (amino acids 15–86) and the C-terminal helix-turn-helix (HTH) LytTR-type DNA-binding response domain (amino acids 161–266), was purified by affinity chromatography on Ni-TED columns (Protino Ni-TED, Macherey Nagel) and eluted with 50 mM NaH₂PO₄, 300 mM NaCl, and 250 mM imidazole, pH 8.0, followed by imidazole removal on Econo-Pac columns (Bio-Rad). The heme-containing PAS sensor domain RcoMH-2 was expressed and purified as described.¹² The purified proteins were suspended in 50 mM Tris-HCl buffer, pH 8.0, 150 mM NaCl, and, unless otherwise indicated, all experiments were performed in this buffer.

The heme-containing PAS sensor domain RcoMH-2 was purified almost entirely in the ferrous CO-bound form, whereas this was only the case for a small fraction of full-length RcoM-2. For experiments on forms other than the CO-bound form, the samples were first fully transformed to the ferric form by illumination for ~12 h with white light originating from a halogen lamp equipped with an optical fiber at an intensity of ~0.15 W/cm² at 4 °C in an air-exposed vial. To obtain the ferrous unliganded form, the ferric protein was degassed using a gas train and reduced with a slight excess of dithionite. The ferrous NO form was fully formed upon the exposure of the ferrous protein to 0.1 atm NO in the cell headspace. To obtain the fully CO-bound form, as-purified protein was reduced with a slight excess of dithionite and exposed to 1 atm CO in the cell headspace.

All steady-state and ultrafast spectroscopy experiments were performed in 1 mm cells that could be sealed gastight with rubber stoppers. Steady-state spectra were obtained using a Shimadzu UV–vis 1700 spectrometer. Ultrafast transient absorption experiments at a 500 Hz repetition rate, with a 570 nm pump pulse and a continuum probe pulse, were performed as in ref 18 using a heme concentration of ~30 μ M. To reduce the baseline fluctuation noise, the data were plotted as difference kinetics between the induced maximum and the bleaching minimum in the Soret absorption regions. To correct for beam overlap drifts on the nanosecond time-scale, kinetics

after photodissociation of CO myoglobin (where no rebinding occurs on the investigated time scale) were used as the reference.

Stopped-flow experiments to monitor CO binding to the ferrous protein were performed in a Biologic SFM 300 instrument equipped with a J&M Tidas diode detector. The optical path length of the measuring cell was 0.8 mm. The instrument was extensively flushed with gaseous nitrogen prior to use. After mixing, the protein concentration was ~25 μ M.

The rate of CO dissociation was determined by the NO replacement method.¹⁹ Here, starting from the fully CO-bound form, CO in the cell head space was replaced by 1 atm NO. The sample was then kept in the dark, and spectra were taken at regular intervals.

Fluorescence anisotropy-based DNA-binding assays with full-length RcoM-2 were performed using 5'-Texas-Red fluorophore-labeled DNA in 1 \times 1 cm optical path length cells. Purified and CO-dissociated protein was suspended in 50 mM Tris-HCl, pH 7.6 and 100 mM NaCl at a concentration of 1 μ M. The ferrous and CO-bound forms were prepared in gastight 1 \times 1 cm optical path length cells, as previously described.

An 18-mer of double-stranded DNA, comprising the equivalent of region a of the RcoM-2 DNA binding region,¹⁴ was obtained by mixing equimolar amounts of the two complementary oligonucleotides 5'-TCGATTTTCGCGCAA-ATT-3' (forward) and 5'-AATTTGCGCGAAAATCGA-3' (reverse) (Eurogentec). The 5'-extremity of the forward oligonucleotide was labeled with Texas Red. The oligonucleotides were heated for 3 min to 80 °C, followed by slow cooling to room temperature to allow complete annealing. The DNA concentration in the assays was 6.4 nM. For the experiments with CO-liganded protein, the DNA stock solution was degassed in the sealed optical cells, and ~0.5 mM dithionite was added. It was, in addition, equilibrated with 1 atm CO atmosphere. RcoM-2 in the different forms was added stepwise using a gastight Hamilton syringe up to concentrations of 25 nM. Fluorescence anisotropy was measured following an as-described procedure²⁰ in a Cary Eclipse fluorometer equipped with a manual polarization device with an excitation beam centered at 590 nm and by scanning the Texas Red emission spectrum. Anisotropy data were fitted to a tight binding formalism^{21,22} to estimate the binding constant (K_d).

All experiments were performed at room temperature.

RESULTS

Steady-State Absorption Spectra. Figure 1 shows the steady-state absorption spectra of RcoM-2 under different conditions. The as-prepared, reduced, and carbonylated spectra are similar to those previously reported by Marvin et al.¹⁵ that were assigned to FeIII His–Cys, FeII His–Met, and FeII His–CO coordinated heme complexes, respectively. Yet we note that upon prolonged illumination, the spectrum changes somewhat, with a modestly blue-shifted and broadened Soret band and a less pronounced shoulder around 570 nm. This finding indicates that a fraction of CO was present in the as-purified protein. This fraction is far smaller than that in the isolated heme domain RcoMH-2 that is almost completely purified in the CO-bound form.¹² For RcoM-1, differences in CO binding in the purified protein have been related to differences in protein expression.¹⁴ We used the same expression system and *E. coli* growth conditions for our full-length and heme domain constructs. Therefore, we assign

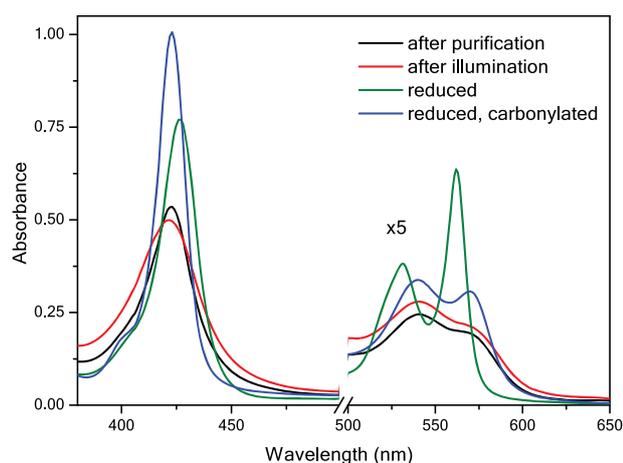


Figure 1. Absorption spectra of RcoM-2 after purification, after prolonged illumination, after degassing and reduction with sodium dithionite, and after exposure to 1 atm CO.

this difference to intrinsic differences in CO affinity, as will be quantified later.

The spectra of the fully reduced and CO-bound form are very similar in the full-length protein and the purified heme domain, although small (~ 1 nm) shifts are observed (Table S1). By contrast, the FeIII (as-prepared illuminated) form is substantially blue-shifted in the RcoMH-2 construct; this can be assigned to the presence of a fraction of the five-coordinate complex.¹²

Bimolecular CO Binding. CO binding to RcoM-2 was monitored by stopped-flow spectroscopy. Figure 2A shows spectra at different delay times after mixing reduced RcoM-2 with CO. The initial spectrum corresponds to the six-coordinate His–Met complex and the final spectrum corresponds to the six-coordinate His–CO complex (cf. Figure 1). The spectra show clear isosbestic points and no indication of the population of distinct intermediate species. The kinetics can be well fit with single-exponential kinetics (Figure 2B) that vary linearly with CO concentration (Figure 2B, inset), yielding an apparent bimolecular rate constant k_{obs} of $16 \times 10^3 \text{ M}^{-1} \text{ s}^{-1}$. In principle, for such a ligand exchange reaction, the rate is described by a hyperbolic function (cf. refs 12 and 23)

$$k_{\text{obs}} = \frac{k_{-M}k_{\text{CO}}[\text{CO}]}{k_M + k_{\text{CO}}[\text{CO}]} \quad (1)$$

where k_{-M} is the rate of dissociation and displacement of the intrinsic (methionine) ligand, k_M is the rate of association of the intrinsic ligand ($k_M = k_{-M}/K_M \gg 1$ as the protein is initially predominantly six-coordinate), and k_{CO} is the bimolecular rate constant of CO binding to the five-coordinate heme. Within the tested [CO] range, no rate saturation was observed, so the rate is not limited by the dissociation of the intrinsic methionine, and $k_{-M} \gg 5 \text{ s}^{-1}$. k_{obs} then corresponds to k_{CO}/K_M .

These findings are distinct from those that we previously reported for the purified heme domain RcoMH-2 but also have some common properties: In RcoMH-2 the kinetics were biphasic, with the main phase displaying a very similar value for k_{CO}/K_M ($15 \times 10^3 \text{ M}^{-1} \text{ s}^{-1}$) but clear saturation behavior corresponding to $k_{-M} = 5 \text{ s}^{-1}$. Thus the motion of the intrinsic

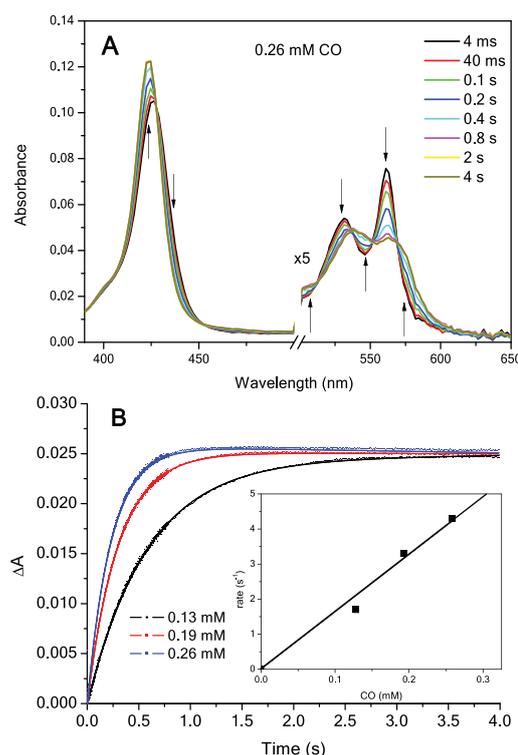


Figure 2. Stopped-flow spectroscopy of CO binding to RcoM-2. (A) Spectra at different delay times after mixing reduced RcoM-2 with CO. The initial spectrum corresponds to the six-coordinate His–Met complex and the final spectrum corresponds to the six-coordinate His–CO complex (cf. Figure 1). The spectra show clear isosbestic points and no indication of the population of distinct intermediate species. The kinetics can be well fit with single-exponential kinetics (Figure 2B) that vary linearly with CO concentration (Figure 2B, inset), yielding an apparent bimolecular rate constant k_{obs} of $16 \times 10^3 \text{ M}^{-1} \text{ s}^{-1}$.

distal heme ligand appears to be less restricted in the full-length protein.

CO Dissociation. The rate of thermal CO dissociation from the protein was determined by exposing the CO-bound protein to excess NO and monitoring the spectral changes associated with replacement of CO by NO as a function of time (Figure 3A). To avoid light-induced CO dissociation, the samples were kept in the dark between monitoring subsequent spectra. The kinetics (Figure 3B) are plotted as a peak-to-bleaching absorption difference to minimize the effect of the baseline variations. They are well-described by single-exponential kinetics with a time constant of 4.4 h, corresponding to a CO off rate of $64 \times 10^{-6} \text{ s}^{-1}$.

This time constant is much shorter than the rough estimation on the order of days that we previously advanced for the heme domain RcoMH-2.¹² We therefore used the same method to precisely determine the time constant for RcoMH-2 at 80 h (CO off rate $3.5 \times 10^{-6} \text{ s}^{-1}$) and indeed found it to be almost 20 times longer than that for the full-length protein.

Heme–CO Geminate Rebinding. Recombination upon photodissociation of the heme–CO bond was investigated using ultrafast pump–probe spectroscopy. The kinetics (Figure 4, red curve) are characterized by a phase of a few picoseconds that reflects mostly photophysical processes²⁴ and multiphasic rebinding that can be fitted (Table 1) with time constants of 170 and 660 ps (with a $\sim 45/55$ amplitude ratio). The rebinding is substantially slower than that in the isolated heme domain (green curve, fit constants 170 and 500 ps, 65/35 amplitude ratio¹²). Moreover, for the full-length protein, a small but significant asymptotic value of 0.5% was found,

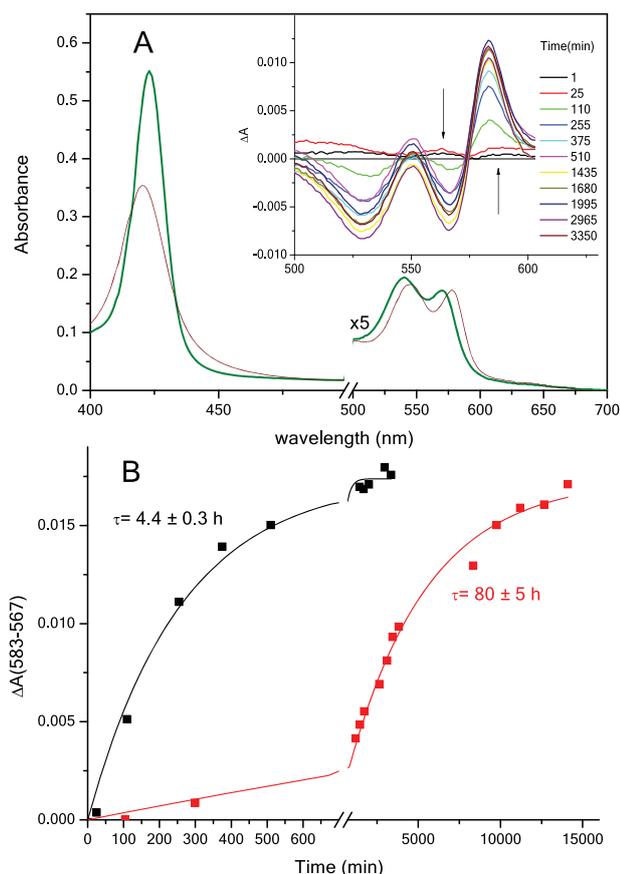


Figure 3. CO dissociation kinetics. (A) RcoM-2 spectra of the CO-bound form (green) and after equilibration of this form with NO for 56 h (red). Inset: difference spectra in the Q-band region at different delay times. (B) Kinetics of CO replacement by NO for full-length RcoM-2 (black) and the heme domain RcoMH-2 (red). Lines are single exponential fits. Note the change in the time scale.

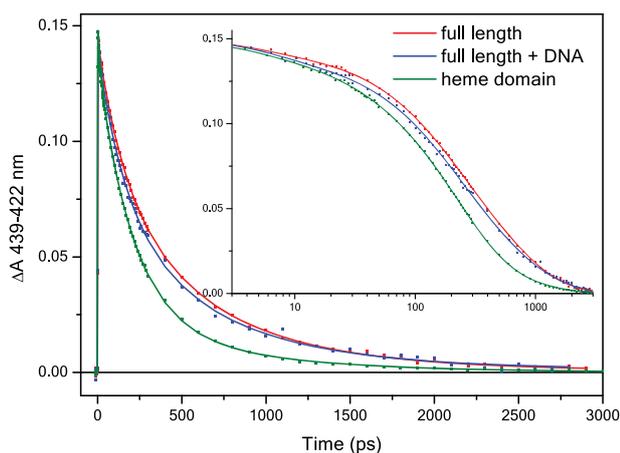


Figure 4. Rebinding kinetics of CO to RcoM-2 after photodissociation in the absence and presence of DNA. The corresponding kinetics for RcoMH-2¹² are also given. Inset: same on a logarithmic time scale.

which reflects CO escape from the heme pocket, whereas in the heme domain rebinding was complete within experimental error.

Table 1. Fit Parameters for the Kinetics of Heme–CO Recombination (Figure 4)^a

	τ_1	A_1	τ_2	A_2	A_3
RcoMH-2	170	0.35	500	0.65	0.00
RcoM-2	180	0.453	660	0.541	0.006
RcoM-2 + DNA	170	0.514	710	0.482	0.004

^aData on the time scale >10 ps were fitted to a function of the form $A_1 e^{-t/\tau_1} + A_2 e^{-t/\tau_2} + A_3$. The amplitudes are normalized to 1, and the time constants are in picoseconds.

Heme–NO Geminate Rebinding. NO rebinding to the full-length protein was also investigated (Figure 5). Faster

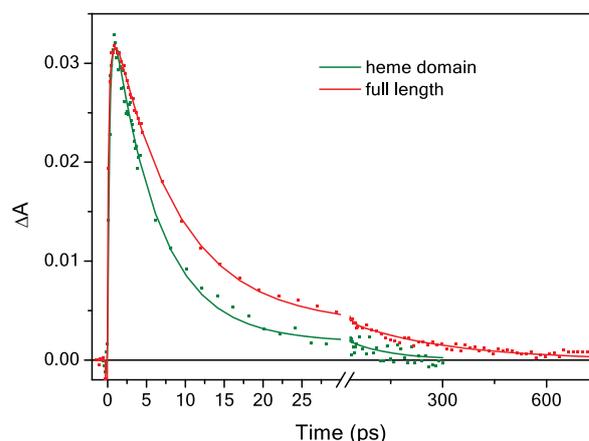


Figure 5. Rebinding kinetics of NO to ferrous RcoM-2 after photodissociation. The corresponding kinetics for RcoMH-2 (ref 12) is also given.

kinetics than with CO, are observed, as is the case in most heme proteins.²⁴ The kinetics can be fit by two phases of 8 and 300 ps (12%) which are substantially slower and more dispersed than those of the corresponding kinetics of the heme domain (6 and 130 ps (5%)¹²). The trend is similar to that of CO recombination and is in full agreement with more conformational freedom for the dissociated ligand in the full-length protein.

DNA Binding. Previous work using fluorescence anisotropy indicated that RcoM-1 binds DNA only in the CO-bound form, but the affinity could not be determined.¹⁴ Here, measuring the fluorescence anisotropy of a 5'-Texas-Red-labeled target DNA sequence as a function of RcoM-2 protein concentration (Figure 6), we show that the affinity for the CO-bound form of this protein is very high ($K_d < 2$ nM) and that the FeIII form of the protein does not bind the target DNA.

Effect of DNA Binding on CO-Bound Protein. Adding unlabeled target DNA to the CO-bound RcoM-2 protein did not have any measurable effect on the shape of the absorption spectrum (Table S1). However, a small but significant effect was observed on the rebinding of CO (Figure 4, blue curve; Table 1) that accelerated overall: The fit constants were similar to those in the absence of DNA, but the amplitude ratio changed to $\sim 50/50$. The fractional escape was found to be unchanged within experimental error. The effect on the decay was in the same direction as that observed for DNA binding to the CooA CO sensor.²⁵ As discussed in detail in the Discussion, this result suggests that DNA binding somewhat favors more rigid conformations of the heme domain.

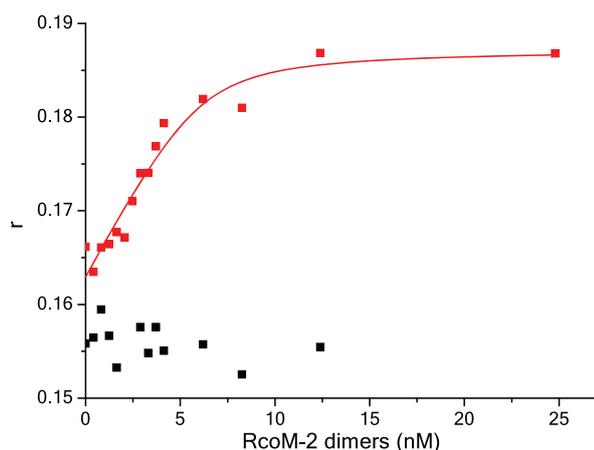


Figure 6. Fluorescence anisotropy of 5'-Texas-Red-labeled target DNA as a function of RcoM-2 concentration in the FeIII state (black) and CO-bound state (red). The fit is to a tight-binding formalism for RcoM-2 dimers and indicates that CO-bound RcoM-2 binds DNA stoichiometrically with a $K_d < 2$ nM.

DISCUSSION

We have investigated the heme–ligand (in particular heme–CO) interactions of full-length RcoM-2 and compared them with those of the isolated heme domain RcoMH-2. In general terms, the binding properties were found to be quite different, with RcoM-2 displaying weaker binding of CO than RcoMH-2. A ~ 20 times higher CO dissociation rate is observed in RcoM-2. This feature can be related to the fact that despite the extremely high CO recombination yield, a sizable fraction ($\sim 0.5\%$) of photodissociated CO escapes the heme pocket in the full-length protein RcoM-2 but not in RcoMH-2. Thus in the presence of the DNA-binding domain, the heme pocket appears to yield more conformational freedom for the ligand. This is in agreement with the observed slower kinetics of CO and NO recombination as well as with the higher rate of displacement of the internal ligand, presumably a methionine residue.¹² Whereas obviously the ligand-binding properties of the heme domain influence the DNA-binding properties and transcriptional activity of the DNA-binding region, interestingly, our results demonstrate that a reverse influence also occurs. One may speculate that the DNA-binding domain

exerts a strain on the heme pocket that, as a result, becomes more open. Unfortunately structural data on RcoM proteins that could shine light on this matter on a molecular level are still missing. In general terms, the influence of DNA binding on the CO recombination kinetics (as discussed later) supports the notion of a bidirectional effect between the two domains that might have a regulatory role.

Whereas the measured CO dissociation rate of RcoM-2, $64 \times 10^{-6} \text{ s}^{-1}$, is higher than that of RcoMH-2, it is still much lower than that of most other heme proteins, including other six-coordinate heme-based sensors (Table 2). Assuming that the intrinsic thermal dissociation rate of the Fe–CO bond is similar in all systems, this difference is likely mostly due to the different yields in heme–CO geminate recombination. From the difference in dissociation rates, we can estimate the CO escape probability in the RcoMH-2 heme domain at $\sim 0.03\%$, which is below our direct detection limit. Indeed, for most other six-coordinate heme-based sensor proteins in Table 2, the geminate rebinding yield is much lower,^{8,18,24,25,28,33,34} and the off rate is two to three orders of magnitude higher than that for RcoM-2. This is on the same order as the CO escape yield measured for RcoM-2 on the nanosecond time scale ($\sim 0.5\%$, Figure 4). Furthermore the relatively slow CO dissociation in CooA (slowest phase $2 \times 10^{-3} \text{ s}^{-1}$ ³¹) correlates with a relatively low CO escape (a few percent^{18,25,33}). Altogether, the CO off rate is modulated by the escape probability from the heme pocket of dissociated CO, and in RcoM proteins, this modulation has led to extremely slow CO release from the protein.

In contrast with the dissociation rate, the effective CO binding rate (at low CO concentrations) is very similar in RcoM-2 and RcoMH-2 (and on the same order as in other heme proteins; Table 2). This finding is indicative of a similar rate of CO reaching the heme pocket, suggesting that in full-length RcoM-2, CO enters the protein directly at the heme domain, and the intraprotein migration pathway does not pass via the DNA binding domain. Taking together the measured binding and dissociation rates, we determined the K_d for CO at 4 nM for RcoM-2 and at 250 pM for RcoMH-2; these high affinities were largely determined by the properties of the heme pocket hindering CO escape. Even for the full-length protein, these are among the highest CO affinities known. As discussed in our previous work¹² they are on the order of those observed

Table 2. CO Affinity (K) and Effective Binding (k_{eff} , Equivalent of k_{CO}/K_M in RcoM) and Dissociation (k_{off}) Rates for Full-Length RcoM-2, the Heme Domain RcoMH-2, and the Other Six-Coordinate Heme-Based Sensor Proteins

	k_{eff} ($\text{M}^{-1} \text{ s}^{-1}$)	k_{off} (s^{-1})	K (M^{-1})
RcoM-2 ^a	16×10^3	64×10^{-6}	0.25×10^{9b}
RcoMH-2 ^c	15×10^3	3.5×10^{-6}	4×10^{9b}
EcDosH	$1.1 \times 10^{3d,e}$	0.011^f	0.1×10^{6d}
	7.8×10^{3g}	0.078^f	
CooA ^h	8×10^3	0.02	0.45×10^6
	1.2×10^3	0.002	
NPAS2 ⁱ			
PASA	370×10^3	$\sim 0.5^f$	$(0.5 \text{ to } 1) \times 10^6$
PASB	40×10^3	0.8^f	0.048×10^6
SUR2A ⁱ	270×10^3	0.05	1.6×10^6
neuroglobin (human) ^k	150×10^3	0.014	11×10^6

^aThis work. ^bComputed from k_{eff} and k_{off} . ^cFrom ref 12. ^dFrom ref 26. ^eUsing stopped flow. Substantially higher k_{eff} values were determined using flash photolysis,^{28–30} see the discussion in ref 28. ^fComputed from k_{eff} and K . ^gFrom ref 27. ^hFrom ref 31. Biphasic kinetics. ⁱFrom ref 7. NPAS2 contains two heme-binding PAS domains. ^jFrom ref 8. ^kFrom ref 32.

for some truncated hemoglobins³⁵ that display slow ligand exchange but where CO does not need to displace an internal ligand.

Our determination of a K_d of 4 nM for CO binding to RcoM-2 is in general agreement with the assessment of a value below ~100 nM by direct titration in the highly homologous RcoM-1 sensor.¹⁴ As has been discussed,¹⁴ these findings and the notion that aerobic RcoM sensors should be sensitive to persistent CO levels in the low nanomolar range, where CO oxidation in *Burkholderia* occurs,³⁶ concord. The observed slow ligand exchange in this range also avoids rapid fluctuations in the sensor activity.

Fluorescence anisotropy-based DNA-binding assays were carried out to quantitatively determine the affinity of RcoM-2 to the 5'-Texas-Red-labeled binding region¹⁴ in its active CO-bound and inactive forms. In the FeIII form, no DNA binding occurs, whereas for CO-bound RcoM-2, the observed anisotropy saturation upon protein addition indicates sequence-specific binding. The observed DNA binding constant K_d of <2 nM is higher than what has been observed under the same assay conditions for the NO-dependent transcriptional regulator DNR (44 nM¹⁸) and the anaerobic CO-dependent transcriptional activator CooA. (Values between 6 and 16 nM have been reported.^{18,37,38}) It is in agreement with a high functional affinity, similar to what has been observed for the binding of RcoM-1 to the a + b + c region¹⁴ and an efficient CO-specific recruitment of RcoM-2.

Finally, our experiments show a small but significant effect of DNA substrate binding on the dynamics of CO recombination in the active site (Figure 4), in general agreement with the bidirectional influence between the sensor and the DNA-binding domains. Upon DNA binding, the overall dynamics accelerate, a finding that is qualitatively similar to that observed in CooA.²⁵ For RcoM-2, where the rebinding kinetics can be described with two exponentials of similar amplitudes, DNA binding results in the enhancement of the faster (170 ps) phase, which is also the dominant phase for the isolated heme domain. This comparison indicates that DNA binding induces a more rigid configuration of the heme domain with less configurational freedom for the dissociated CO. A similar conclusion was reached for CooA based on the narrowing of a wider range of rate distributions upon DNA binding²⁵ (a feature itself not occurring in RcoM-2). Altogether, these results suggest that this is a general feature for heme-based gas-regulated transcription factors.

In conclusion, we have demonstrated that the CO affinity in RcoM-2 is regulated by the ligand dynamics properties in the heme pocket, which are fine-tuned by long-range interactions with the DNA-binding domain. Whereas CO binding to the isolated heme domain is quasi-irreversible, these interactions allow reversible CO binding in the full-length RcoM-2 protein. The resulting CO affinity ($K_d = 4$ nM) is high compared with other CO-regulated systems but brings the sensor in the sensitivity range expected for CO oxidation under aerobic conditions. Studies of the precise molecular origin of the important dynamic properties of the heme pocket must await the availability of structural data on the protein.

■ ASSOCIATED CONTENT

📄 Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acs.biochem.9b00623.

Table S1. Absorption maxima of different forms of RcoM-2 full-length and isolated heme domain (PDF)

Accession Codes

RcoM-2, Q13IY4

■ AUTHOR INFORMATION

Corresponding Author

*E-mail: marten.vos@polytechnique.edu. Tel: +33169335066.

ORCID

Marten H. Vos: 0000-0003-0493-4831

Present Address

‡C.V.F.: LIDYL, CEA, CNRS, Université Paris Saclay, 91191 Gif-sur-Yvette, France.

Author Contributions

†M.S. and C.V.F. contributed equally to this work

Funding

M.S. is supported by a municipal grant from Kafarhatta (Lebanon).

Notes

The authors declare no competing financial interest.

■ REFERENCES

- (1) Shimizu, T., Huang, D., Yan, F., Stranova, M., Bartosova, M., Fojtikova, V., and Martinkova, M. (2015) Gaseous O₂, NO, and CO in Signal Transduction: Structure and Function Relationships of Heme-Based Gas Sensors and Heme-Redox Sensors. *Chem. Rev.* 115, 6491–6533.
- (2) Négrerie, M. (2019) Iron transitions during activation of allosteric heme proteins in cell signaling. *Metallomics* 11, 868–893.
- (3) Boczkowski, J., Poderoso, J. J., and Motterlini, R. (2006) CO-metal interaction: vital signaling from a lethal gas. *Trends Biochem. Sci.* 31, 614–621.
- (4) Gullotta, F., di Masi, A., Coletta, M., and Ascenzi, P. (2012) CO metabolism, sensing, and signaling. *Biofactors* 38, 1–13.
- (5) Heinemann, S. H., Hoshi, T., Westerhausen, M., and Schiller, A. (2014) Carbon monoxide - physiology, detection and controlled release. *Chem. Commun.* 50, 3644–3660.
- (6) Kim, H., Ryter, S., and Choi, A. (2006) CO as a cellular signalling molecule. *Annu. Rev. Pharmacol. Toxicol.* 46, 411–449.
- (7) Dioum, E. M., Rutter, J., Tuckerman, J. R., Gonzalez, G., Gilles-Gonzalez, M.-A., and McKnight, S. L. (2002) NPAS2: A Gas-Responsive Transcription Factor. *Science* 298, 2385–2387.
- (8) Kapetanaki, S. M., Burton, M. J., Basran, J., Urugami, C., Moody, P. C. E., Mitcheson, J. S., Schmid, R., Davies, N. W., Dorlet, P., Vos, M. H., Storey, N. M., and Raven, E. (2018) A mechanism for CO regulation of ion channels. *Nat. Commun.* 9, 907.
- (9) Ascenzi, P., Bocedi, A., Leoni, L., Visca, P., Zennaro, E., Milani, M., and Bolognesi, M. (2004) CO Sniffing through Heme-based Sensor Proteins. *IUBMB Life* 56, 309–315.
- (10) Lanzilotta, W. N., Schuller, D. J., Thorsteinnsson, M. V., Kerby, R. L., Roberts, G. P., and Poulos, T. L. (2000) Structure of the CO sensing transcription activator CooA. *Nat. Struct. Biol.* 7, 876–880.
- (11) Kerby, R. L., Youn, H., and Roberts, G. P. (2008) RcoM: A New Single-Component Transcriptional Regulator of CO Metabolism in Bacteria. *J. Bacteriol.* 190, 3336–3343.
- (12) Bouzhir-Sima, L., Motterlini, R., Gross, J., Vos, M. H., and Liebl, U. (2016) Unusual Dynamics of Ligand Binding to the Heme Domain of the Bacterial CO Sensor Protein RcoM-2. *J. Phys. Chem. B* 120, 10686–10694.
- (13) Bowman, H. E., Dent, M. R., and Burstyn, J. N. (2016) Met104 is the CO-replaceable ligand at Fe(II) heme in the CO-sensing transcription factor BxRcoM-1. *J. Biol. Inorg. Chem.* 21, 559–569.

- (14) Kerby, R. L., and Roberts, G. P. (2012) *Burkholderia xenovorans* RcoMBx-1, a Transcriptional Regulator System for Sensing Low and Persistent Levels of Carbon Monoxide. *J. Bacteriol.* 194, 5803–5816.
- (15) Marvin, K. A., Kerby, R. L., Youn, H., Roberts, G. P., and Burstyn, J. N. (2008) The Transcription Regulator RcoM-2 from *Burkholderia xenovorans* Is a Cysteine-Ligated Hemoprotein That Undergoes a Redox-Mediated Ligand Switch. *Biochemistry* 47, 9016–9028.
- (16) Smith, A. T., Marvin, K. A., Freeman, K. M., Kerby, R. L., Roberts, G. P., and Burstyn, J. N. (2012) Identification of Cys94 as the distal ligand to the Fe(III) heme in the transcriptional regulator RcoM-2 from *Burkholderia xenovorans*. *JBIC, J. Biol. Inorg. Chem.* 17, 1071–1082.
- (17) Pinhancos, R. C., Bowman, H. E., Dent, M. R., Young, B. H., Berndsen, C. E., and Burstyn, J. N. (2017) The Heme-Binding PAS Domain Mediates Dimerization in the CO-Sensing Transcription Factor BxRcoM-1. *FASEB J.* 31, 908.15.
- (18) Lobato, L., Bouzhir-Sima, L., Yamashita, T., Wilson, M. T., Vos, M. H., and Liebl, U. (2014) Dynamics of the heme-binding bacterial gas sensing dissimilative nitrate respiration regulator (DNR) and activation barriers for ligand binding and escape. *J. Biol. Chem.* 289, 26514–26524.
- (19) Antonini, E., and Brunori, M. (1971) *Hemoglobin and Myoglobin in Their Reactions with Ligands*, North-Holland, Amsterdam.
- (20) Creze, C., Ligabue, A., Laurent, S., Lestini, R., Laptanok, S. P., Khun, J., Vos, M. H., Czjzek, M., Myllykallio, H., and Flament, D. (2012) Modulation of the *Pyrococcus abyssi* NucS endonuclease activity by the replication clamp PCNA at functional and structural levels. *J. Biol. Chem.* 287, 15648–15660.
- (21) Copeland, R. (2004) *Enzymes: A Practical Introduction to Structure, Mechanism, and Data Analysis*, John Wiley & Sons.
- (22) Laptanok, S. P., Bouzhir-Sima, L., Lambry, J.-C., Myllykallio, H., Liebl, U., and Vos, M. H. (2013) Ultrafast real time visualization of the active site flexibility of the flavoenzyme thymidylate synthase ThyX. *Proc. Natl. Acad. Sci. U. S. A.* 110, 8924–8929.
- (23) Silkstone, G., Jasaitis, A., Wilson, M. T., and Vos, M. H. (2007) Ligand dynamics in an electron-transfer protein: picosecond geminate recombination of carbon monoxide to heme in mutant forms of cytochrome *c*. *J. Biol. Chem.* 282, 1638–1649.
- (24) Vos, M. H. (2008) Ultrafast dynamics of ligands within heme proteins. *Biochim. Biophys. Acta, Bioenerg.* 1777, 15–31.
- (25) Benabbas, A., Karunakaran, V., Youn, H., Poulos, T. L., and Champion, P. M. (2012) Effect of DNA Binding on Geminate CO Recombination Kinetics in CO-sensing Transcription Factor CooA. *J. Biol. Chem.* 287, 21729–21740.
- (26) Delgado-Nixon, V. M., Gonzalez, G., and Gilles-Gonzalez, M. A. (2000) Dos, a heme-binding PAS protein from *Escherichia coli*, is a direct oxygen sensor. *Biochemistry* 39, 2685–2691.
- (27) Taguchi, S., Matsui, T., Igarashi, J., Sasakura, Y., Araki, Y., Ito, O., Sugiyama, S., Sagami, I., and Shimizu, T. (2004) Binding of Oxygen and Carbon Monoxide to a Heme-regulated Phosphodiesterase from *Escherichia coli*: Kinetics and Infrared Spectra of the Full-Length Wild-Type Enzyme, Isolated PAS Domain, and Met-95 Mutants. *J. Biol. Chem.* 279, 3340–3347.
- (28) Liebl, U., Bouzhir-Sima, L., Kiger, L., Marden, M. C., Lambry, J.-C., Négretrie, M., and Vos, M. H. (2003) Ligand binding dynamics to the heme domain of the oxygen sensor Dos from *Escherichia coli*. *Biochemistry* 42, 6527–6535.
- (29) Ishitsuka, Y., Araki, Y., Tanaka, A., Igarashi, J., Ito, O., and Shimizu, T. (2008) Arg97 at the Heme-Distal Side of the Isolated Heme-Bound PAS Domain of a Heme-Based Oxygen Sensor from *Escherichia coli* (*Ec* DOS) Plays Critical Roles in Autoxidation and Binding to Gases, Particularly O₂. *Biochemistry* 47, 8874–8884.
- (30) Lechavue, C., Bouzhir-Sima, L., Yamashita, T., Marden, M. C., Vos, M. H., Liebl, U., and Kiger, L. (2009) Heme Ligand Binding Properties and Intradimer Interactions in the Full-length Sensor Protein Dos from *Escherichia coli* and Its Isolated Heme Domain. *J. Biol. Chem.* 284, 36146–36159.
- (31) Puranik, M., Nielsen, S. B., Youn, H., Hvitved, A. N., Bourassa, J. L., Case, M. A., Tengroth, C., Balakrishnan, G., Thorsteinnsson, M. V., Groves, J. T., McLendon, G. L., Roberts, G. P., Olson, J. S., and Spiro, T. G. (2004) Dynamics of Carbon Monoxide Binding to CooA. *J. Biol. Chem.* 279, 21096–21108.
- (32) Dewilde, S., Kiger, L., Burmester, T., Hankeln, T., Baudin-Creuz, V., Aerts, T., Marden, M. C., Caubergs, R., and Moens, L. (2001) Biochemical Characterization and Ligand Binding Properties of Neuroglobin, a Novel Member of the Globin Family. *J. Biol. Chem.* 276, 38949–38955.
- (33) Kumazaki, S., Nakajima, H., Sakaguchi, T., Nakagawa, E., Shinohara, H., Yoshihara, K., and Aono, S. (2000) Dissociation and recombination between ligands and heme in a CO-sensing transcriptional activator CooA. *J. Biol. Chem.* 275, 38378–38383.
- (34) Abbruzzetti, S., Faggiano, S., Bruno, S., Spyrikis, F., Mozzarelli, A., Dewilde, S., Moens, L., and Viappiani, C. (2009) Ligand migration through the internal hydrophobic cavities in human neuroglobin. *Proc. Natl. Acad. Sci. U. S. A.* 106, 18984–18989.
- (35) Milani, M., Pesce, A., Nardini, M., Ouellet, H., Ouellet, Y., Dewilde, S., Bocedi, A., Ascenzi, P., Guertin, M., and Moens, L. (2005) Structural bases for heme binding and diatomic ligand recognition in truncated hemoglobins. *J. Inorg. Biochem.* 99, 97–109.
- (36) Weber, C. F., and King, G. M. (2012) The phylogenetic distribution and ecological role of carbon monoxide oxidation in the genus *Burkholderia*. *FEMS Microbiol. Ecol.* 79, 167–175.
- (37) Hines, J. P., Dent, M. R., Stevens, D. J., and Burstyn, J. N. (2018) Site-directed spin label electron paramagnetic resonance spectroscopy as a probe of conformational dynamics in the Fe(III) “locked-off” state of the CO-sensing transcription factor CooA. *Protein Sci.* 27, 1670–1679.
- (38) Thorsteinnsson, M. V., Kerby, R. L., Conrad, M., Youn, H., Staples, C. R., Lanzilotta, W. N., Poulos, T. J., Serate, J., and Roberts, G. P. (2000) Characterization of Variants Altered at the N-terminal Proline, a Novel Heme-Axial Ligand in CooA, the CO-sensing Transcriptional Activator. *J. Biol. Chem.* 275, 39332–39338.

Synthesis and Structure–Activity Relationship Studies of Benzo[*b*][1,4]oxazin-3(4*H*)-one Analogues as Inhibitors of Mycobacterial Thymidylate Synthase X

Jakub Modranka⁺,^[a] Jiahong Li⁺,^[a] Anastasia Parchina⁺,^[a] Michiel Vanmeert,^[a] Shrinivas Dumbre,^[a] Mayla Salman,^[b] Hannu Myllykallio,^[b] Hubert F. Becker,^[b, c] Roeland Vanhoutte,^[e] Lia Margamuljana,^[a] Hoai Nguyen,^[a] Rania Abu El-Asrar,^[a] Jef Rozenski,^[a] Piet Herdewijn,^[a] Steven De Jonghe,^{*,[a, d]} and Eveline Lescrinier^{*,[a]}

Since the discovery of a flavin-dependent thymidylate synthase (ThyX or FDTs) that is absent in humans but crucial for DNA biosynthesis in a diverse group of pathogens, the enzyme has been pursued for the development of new antibacterial agents against *Mycobacterium tuberculosis*, the causative agent of the widespread infectious disease tuberculosis (TB). In response to a growing need for more effective anti-TB drugs, we have built

upon our previous screening efforts and report herein an optimization campaign of a novel series of inhibitors with a unique inhibition profile. The inhibitors display competitive inhibition toward the methylene tetrahydrofolate cofactor of ThyX, enabling us to generate a model of the compounds bound to their target, thus offering insight into their structure–activity relationships.

Introduction

Tuberculosis (TB) is a contagious-infectious disease caused by the bacterium *Mycobacterium tuberculosis*. Currently, it is the ninth leading cause of death worldwide and the leading cause from a single infectious agent.^[1] The prognosis for patients with tuberculosis improved dramatically with the introduction of antitubercular drugs such as rifampicin, isoniazid, pyrazinamide, and ethambutol.^[2] Despite having these treatment options, according to the latest report by the World Health Organization (WHO), *M. tuberculosis* still caused an estimated 1.7

million deaths in 2016. In addition, an estimated 10.4 million people fell ill with tuberculosis.^[1] Recently, two new antitubercular drugs, bedaquiline^[3] and delamanid,^[4] have received marketing approval. However, the emergence of multidrug-resistant (MDR), extensively drug-resistant (XDR) and even totally drug-resistant (TDR) *M. tuberculosis* strains remains a major threat to public health.^[5] Consequently, there is an urgent demand for new anti-TB drugs that act on novel mycobacterial targets.

Thymidylate synthase (ThyA) catalyzes the de novo synthesis of thymidine-5'-monophosphate (TMP) via reductive methylation of 2'-deoxyuridine-5'-monophosphate (dUMP). ThyA uses *N*-5,*N*-10-methylenetetrahydrofolic acid (CH₂H₄folate) both as one-carbon and hydride donor, whereby dihydrofolate (DHF) is formed. The catalytic activity of ThyA is coupled to the action of two other enzymes. Dihydrofolate reductase (DHFR) mediates the reduction of DHF to tetrahydrofolate (H₄folate), which is further subjected to re-methylation catalyzed by serine hydroxymethyltransferase (SHMT) to regenerate the CH₂H₄folate pools. Thymidine-5'-monophosphate is readily phosphorylated to thymidine-5'-triphosphate (TTP), one of the four building blocks required for the biosynthesis of DNA.^[6] Several human pathogens lack the genes encoding for ThyA and DHFR, but are still fully viable in thymidine-deficient media. Genomic analysis led to the discovery of the *thyX* gene which encodes for the ThyX protein, a flavin-dependent thymidylate synthase (FDTs), which is absent in humans.^[7] ThyX uses CH₂H₄folate only as one-carbon donor, with reduced flavin adenine dinucleotide (FADH₂) serving as a hydride donor. ThyX enzymes are NAD(P)H oxidases, as catalysis requires the presence of NAD(P)H to furnish sufficient amounts of FADH₂ via reduction

[a] Dr. J. Modranka,⁺ Dr. J. Li,⁺ Dr. A. Parchina,⁺ M. Vanmeert, Dr. S. Dumbre, L. Margamuljana, H. Nguyen, R. Abu El-Asrar, Prof. J. Rozenski, Prof. P. Herdewijn, Dr. S. De Jonghe, Prof. E. Lescrinier
Medicinal Chemistry, Rega Institute for Medical Science, KU Leuven, Herestraat 49, PO Box 1030, 3000 Leuven (Belgium)
E-mail: Steven.DeJonghe@kuleuven.be
Eveline.Lescrinier@kuleuven.be

[b] Dr. M. Salman, Prof. H. Myllykallio, Prof. H. F. Becker
Laboratory of Optics and Biosciences, INSERM U 696-CNRS UMR 7645, Ecole Polytechnique, Route de Saclay, 91128 Palaiseau Cedex (France)

[c] Prof. H. F. Becker
Faculté des Sciences et Ingénierie, Sorbonne Université, 4 place Jussieu 75005 Paris (France)

[d] Dr. S. De Jonghe
Present affiliation: Laboratory of Virology and Chemotherapy, Rega Institute for Medical Research, KU Leuven, Herestraat 49, PO Box 1043, 3000 Leuven (Belgium)

[e] R. Vanhoutte
Present affiliation: Laboratory of Chemical Biology, KU Leuven, O&N I, Herestraat 49, PO Box 802, 3000 Leuven (Belgium)

[*] These authors contributed equally to this work.

 Supporting information and the ORCID identification number(s) for the author(s) of this article can be found under:
 <https://doi.org/10.1002/cmdc.201800739>

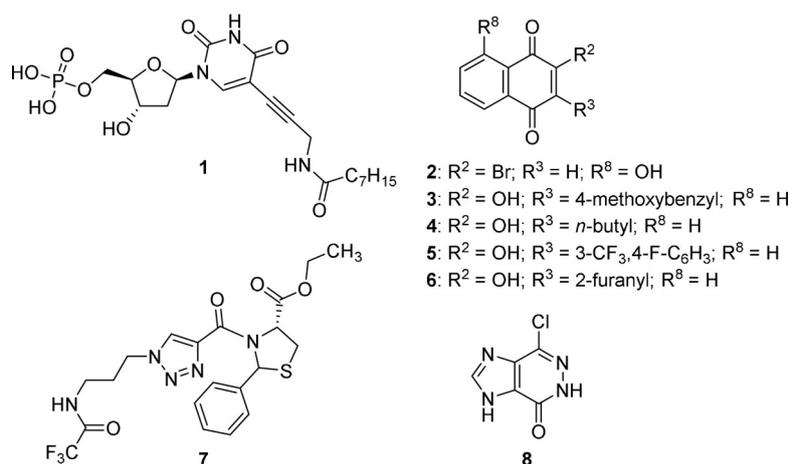


Figure 1. Inhibitors of ThyX.

of FAD.^[8] Furthermore, ThyX shows no structural or sequence similarity to ThyA.^[9] While some pathogens solely rely on ThyX, *M. tuberculosis* carries conserved genes encoding for ThyA as well as for ThyX.^[7] The *thyX* gene is required and essential for mycobacterial growth and survival within macrophages, making ThyX a promising antimycobacterial drug target.^[10] Despite this, only a limited number of ThyX inhibitors are known (Figure 1).^[11] Structural variation of the uracil moiety of the natural substrate dUMP afforded a series of 5-alkynyl dUMP analogues, from which the most potent congener (compound 1), displayed an IC₅₀ value of 0.9 μM against mycobacterial ThyX.^[12]

However, the presence of the polar phosphate moiety precludes its further development into derivatives with in vitro antimycobacterial activity. 2-Bromo-8-hydroxy-1,4-naphthoquinone (compound 2) and 2-hydroxy-3-(4-methoxybenzyl)-1,4-naphthoquinone (compound 3) have been identified as inhibitors of ThyX from *Paramecium bursaria chlorella virus-1* (PBCV-1). In addition, these compounds inhibit ThyX proteins from other microorganisms, including *M. tuberculosis*, *H. pylori*, and *C. trachomatis*. Moreover, compound 2 displays antibacterial activity in a genetically modified *E. coli* strain harboring the *thyX* gene.^[13] Further optimization of this class of compounds as ThyX inhibitors from *H. pylori* led to the discovery of new analogues with improved ThyX inhibition and antibacterial activity.^[14] Moreover, three representatives displayed promising activity in a mouse model of *H. pylori* infection.^[14] The authors did not observe any cytotoxicity, mitochondrial toxicity, or in vivo toxicity with these quinone analogues. However, it is well known that quinones act as Michael acceptors and can cause cellular damage by alkylation of cellular proteins and/or DNA. In addition, quinones are highly redox-active molecules that can lead to the formation of reactive oxygen species (ROS), which can cause severe oxidative stress within cells. Therefore, quinones are not desirable scaffolds in medicinal chemistry.^[15] A library of thiazolidine derivatives has been prepared and evaluated as inhibitors of viral PBCV-1 ThyX. The most potent congener (compound 7) was endowed with an IC₅₀ value of 0.057 μM.^[16] In an effort to obtain non-substrate inhibitors, a virtual screening campaign of the ZINC database

against mycobacterial ThyX was performed. The most potent congener was an imidazo[4,5-*d*]pyridazine analogue (compound 8) that, however, displayed only 29% inhibition at 100 μM.^[17]

Because of the potential of ThyX as antibacterial drug target and the lack of drug-like small-molecule ThyX inhibitors, we recently embarked on a high-throughput screening (HTS) campaign of commercially available compound libraries in the search for novel and drug-like mycobacterial ThyX inhibitors. It led to the discovery of B1-PP146 (compound 9, Figure 2) as a tight-binding mycobacterial ThyX inhibitor with an IC₅₀ value of 0.71 μM.^[18] Screening of structurally related, commercially available analogues led to the discovery of compound 10, endowed with an IC₅₀ value of 0.9 μM. As the *N*-carboxamide-piperazine substructure of compound 10 allows for more and easier structural variation than the *N*-pyrimidinyl-piperazine moiety of 9, compound 10 was selected as starting point for an optimization campaign. In this manuscript, we describe our medicinal chemistry efforts to study the structure–activity relationship of these benzoxazine analogues as inhibitors of mycobacterial ThyX.

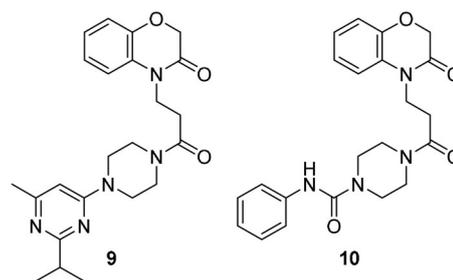
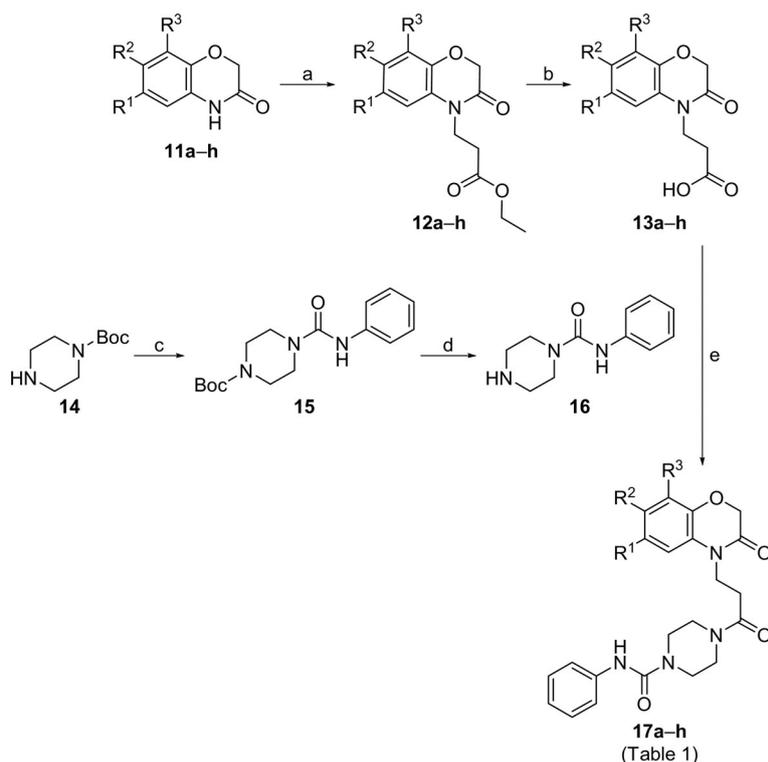


Figure 2. Benzo[*b*][1,4]oxazin-3(4*H*)-ones as mycobacterial ThyX inhibitors.

Results and Discussion

Chemistry

The synthesis of *N*-phenylpiperazine-1-carboxamide analogues 17a–i with structural variation of the phenyl part of the bicy-



Scheme 1. Reagents and conditions: a) ethyl 3-bromopropionate, K_2CO_3 , DMF, 80 °C; b) LiOH, THF, 60 °C; c) PhNCO, THF, RT; d) TFA, 0 °C; e) HCTU, DIPEA, DMSO, RT.

clic benzo[b][1,4]oxazin-3(4H)-one scaffold was accomplished as shown in Scheme 1. Commercially available 2H-benzo[b][1,4]oxazin-3(4H)-one derivatives **11 a-h** were N-alkylated with ethyl 3-bromopropionate under alkaline conditions.^[19] Saponification of the ethyl ester moiety of **12 a-h** with lithium hydroxide furnished carboxylic acids **13 a-h**. N-Phenylpiperazine-1-carboxamide **16** was synthesized by reaction of tert-butyl piperazine-1-carboxylate **14** with phenyl isocyanate, followed by acidic deprotection of the Boc group.^[20] Finally, reaction of acids **13 a-h** and amine **16** with O-(1H-6-chlorobenzotriazole-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate (HCTU) as coupling reagent afforded the desired compounds **17 a-h**.

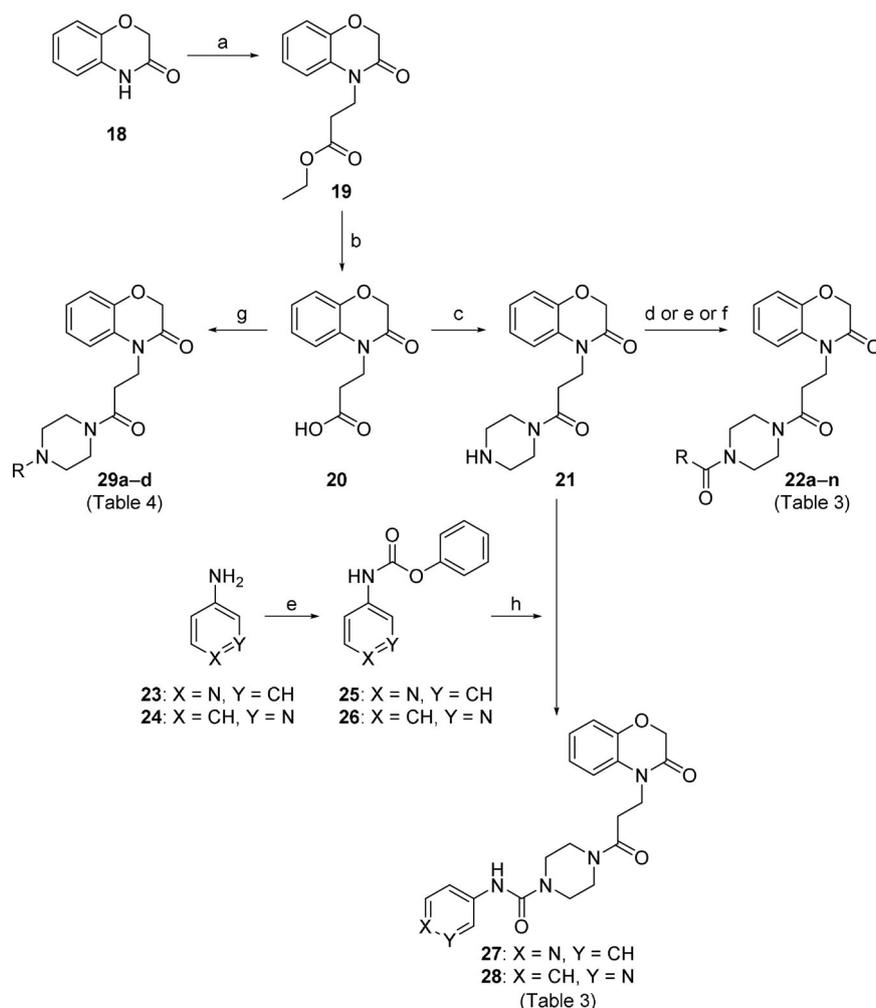
For the synthesis of a library with structural variation of the terminal aryl group, a slightly modified scheme was used (Scheme 2). Alkylation of **18**, followed by alkaline hydrolysis of the ester moiety, furnished acid **20**. N,N'-Dicyclohexylcarbodiimide (DCC)-mediated amide formation afforded compound **21** as key intermediate from which structural variation could easily be introduced. Reaction with a series of isocyanates gave access to a set of urea derivatives **22 a-l**. Alternatively, coupling of **21** with phenyl chloroformate or phenylacetyl chloride furnished carbamate **22 m** and amide **22 n**, respectively. For the synthesis of pyridine-containing urea derivatives **27** and **28**, 4- and 3-aminopyridine (compounds **23** and **24**, respectively) were first coupled with phenylchloroformate, yielding carbamates **25** and **26**, respectively. Condensation with piperazine analogue **21** furnished derivatives **27** and **28**, respectively. Because of the commercial availability of a number of piperazine building blocks, carboxylic acid **20** was easily convert-

ed in one step to the corresponding piperazine analogues **29 a-d**.

The synthesis of analogues with variation in the linker moiety between the benzo[b][1,4]oxazin-3(4H)-one core and the piperazine moiety was effected as shown in Scheme 3. Alkylation of **18** with the appropriate alkyl bromides gave access to compounds **30** and **31**. Hydrolysis of the ethyl ester group, followed by coupling with piperazine and phenylisocyanate afforded final compounds **32** and **33**. For the synthesis of branched-chain derivatives, a similar methodology was applied. Alkylation of **18** with suitable alkyl bromides gave compounds **34** and **35**. Saponification of the ester moiety followed by reaction with N-phenylcarboxamide piperazine **16** yielded the desired derivatives **36** and **37**.

To have the piperazine ring linked to the benzo[b][1,4]oxazin-3(4H)-one scaffold by an alkyl chain (rather than an amide bond), tert-butyl 4-(3-bromopropyl)piperazine-1-carboxylate was reacted with **18** in the presence of potassium carbonate, yielding intermediate **38**. Acidic removal of the Boc protecting group and condensation with phenyl isocyanate afforded target compound **39**.

Modifications in the piperazine motif were introduced as shown in Scheme 4. For the synthesis of the homopiperazine analogue **41**, a stepwise approach starting from compound **20** was followed, consisting of amide coupling and urea formation, yielding the 1,4-diazepine analogue **41**. On the other hand, for the methyl-substituted piperazine analogues **45 a,b**, the aminopiperidine derivatives **45 c,d** and the aminopyrrolidine congeners **45 e,f**, a slightly different methodology was



Scheme 2. Reagents and conditions: a) ethyl 3-bromopropionate, K_2CO_3 , DMF, $90^\circ C$; b) LiOH, THF, $45^\circ C$; c) DCC, HOBt, piperazine, RT; d) RNCO, DMF, RT; e) PhOC(O)Cl, DIPEA, DMF, RT (for **22m**); f) $PhCH_2C(O)Cl$, DIPEA, DMF, RT (for **22n**); g) R-piperazine, HCTU, DIPEA, RT; h) DIPEA, DMF, RT.

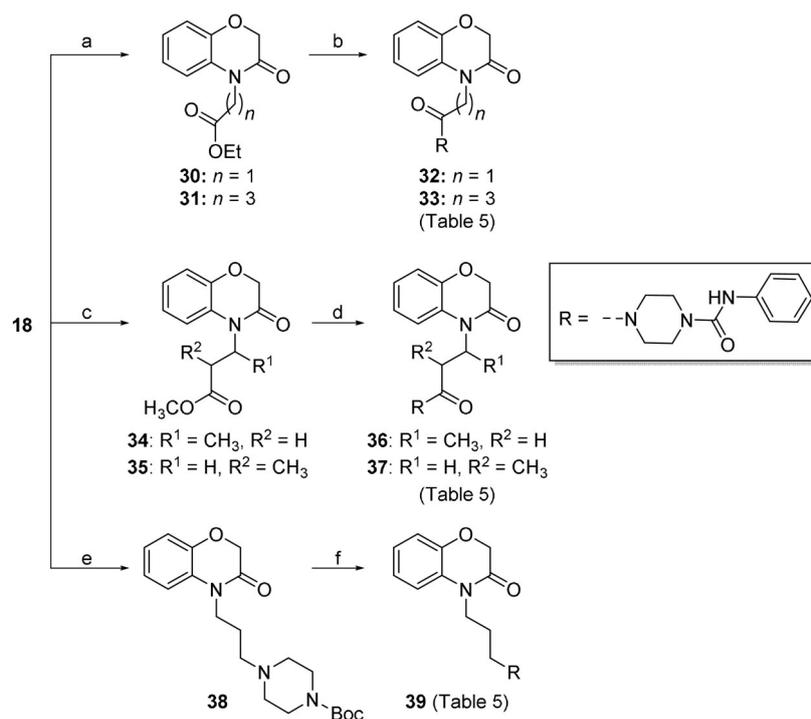
used. The free amino group of the Boc-protected building blocks **42a–f** was coupled with phenylisocyanate, followed by acidic cleavage of the Boc group. Finally, amide coupling of **44a–f** with **20** yielded a series of analogues **45a–f** with structural variation of the piperazine moiety.

Scaffold modified analogues were prepared from appropriate, commercially available building blocks (Scheme 5). The synthesis of pyrido-fused oxazine analogues started from the commercially available *2H*-pyrido[3,2-*b*][1,4]oxazin-3(4*H*)-one **46** and *1H*-pyrido[2,3-*b*][1,4]oxazin-2(3*H*)-one **47**. *N*-Alkylation followed by ester hydrolysis and amide coupling yielded the desired analogues **48** and **49**. A number of derivatives with substituents at position 2 of the oxazine ring were also prepared. 2-Methyl-2*H*-benzo[*b*][1,4]oxazin-3(4*H*)-one **50** is commercially available, whereas building blocks with a mono- (compound **51**) or difluorinated (compound **52**) oxazine moiety were synthesized according to published procedures.^[21] Applying the standard reaction conditions (*N*-alkylation, ester hydrolysis, amide formation) yielded compounds **53**, **54**, and **55**. Final derivatives **58** and **59** were prepared in a similar way starting from benzo[*b*][1,4]thiazin-3(4*H*)-one **56** or 3,4-dihydro-

quinolin-2(1*H*)-one **57**. To have access to a 3,4-dihydroquinoxaline-based analogue, *tert*-butyl 3-oxo-3,4-dihydroquinoxaline-1(2*H*)-carboxylate **60** (synthesized according to a published procedure^[22]) was selected as starting material. Alkylation followed by saponification of the ester residue and formation of an amide bond yielded the desired compound **61**. Unfortunately, upon cleavage of the Boc protecting group under acidic conditions, a concomitant oxidation of the 3,4-dihydroquinoxalin-2(1*H*)-one moiety took place, yielding the quinoxalin-2(1*H*)-one analogue **62**. The benzoxazole congener **64** was easily obtained by reaction of **63** with *N*-phenylpiperazine-1-carboxamide **16**. A benzoxazine analogue lacking the 2-oxo functionality (compound **66**) was accessible via alkylation of **65**, followed by ester hydrolysis and amide coupling.

Mycobacterial ThyX inhibition and structure–activity relationship studies

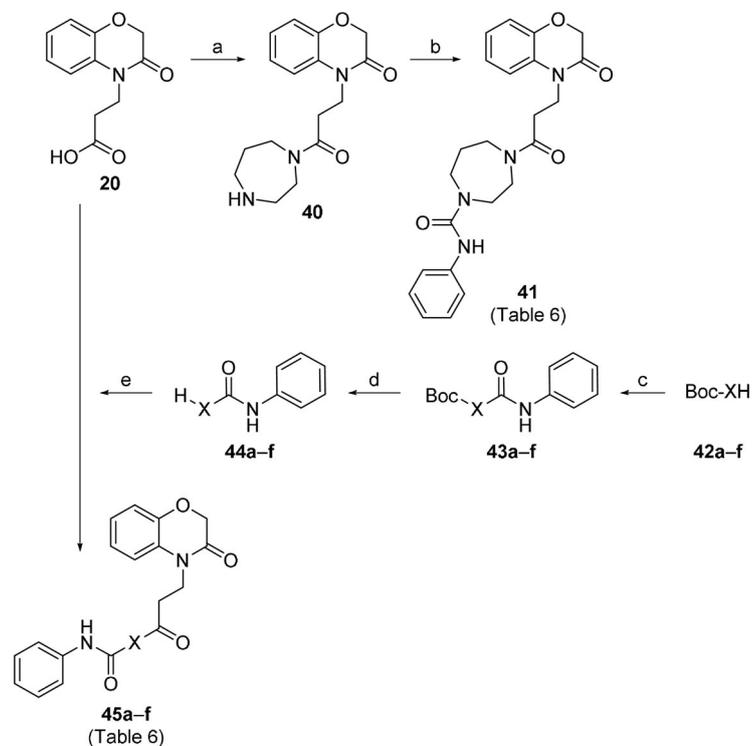
To investigate the SAR in a systematic way, hit compound **10** was divided into different areas, and each of these substructures was subjected to structural modifications (Figure 3). In a



Scheme 3. Reagents and conditions: a) ethyl bromoacetate (for **30**) or ethyl 4-bromobutyrate (for **31**), K_2CO_3 , DMF, $90^\circ C$; b) (i) LiOH, H_2O , THF, $45^\circ C$, (ii) DCC, HOBT, piperazine, DMF, RT, (iii) PhNCO, DMF, RT; c) methyl 3-bromobutanoate (for **34**) or methyl 3-bromo-2-methylpropanoate (for **35**), K_2CO_3 , DMF, $90^\circ C$; d) (i) LiOH, H_2O , THF, $45^\circ C$, (ii) DCC, HOBT, *N*-phenylpiperazine-1-carboxamide, DMF; e) *tert*-butyl 4-(3-bromopropyl)piperazine-1-carboxylate, K_2CO_3 , DMF, $65^\circ C$; f) (i) TFA, RT, (ii) PhNCO, THF, RT.

primary screening, all compounds were tested at a concentration of $5\ \mu M$ as potential inhibitors of mycobacterial ThyX in a

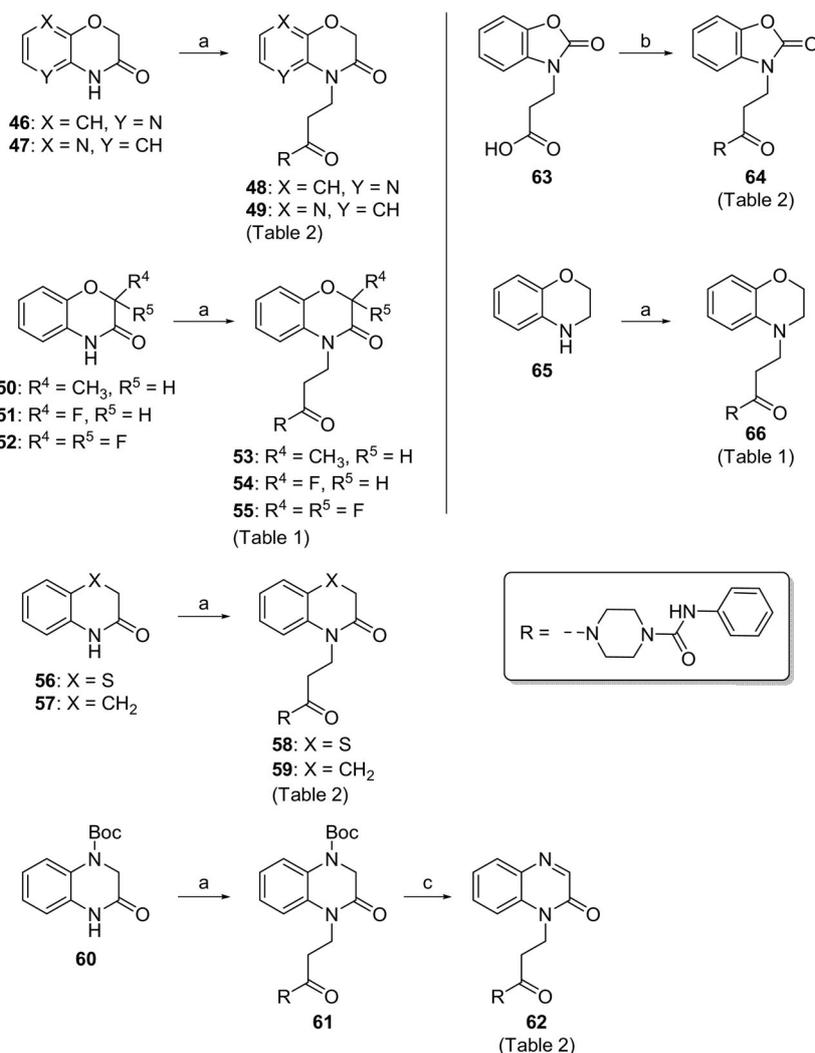
spectrophotometric assay. Folinic acid (also known as leucovorin; $50\ \mu M$) was included as a reference to scale the inhibitory activity of tested compounds. At a concentration of $50\ \mu M$, folinic acid inhibited ThyX enzymatic activity by 80%. Hence, compounds with relative inhibition of more than 1 in this study inhibited the ThyX activity more than 80% at $5\ \mu M$. For the hit compound and its derivatives that had a relative inhibition of at least 1, the concentration to obtain 50% of the maximal inhibition (IC_{50}) was determined using full dose-response curves. All selected compounds were able to fully inhibit the enzymatic activity of ThyX under the selected conditions.



Scheme 4. Reagents and conditions: a) DCC, HOBT, homopiperazine, DMF; b) phenylisocyanate, DMF; c) PhNCO, THF; d) TFA, CH_2Cl_2 ; e) HCTU, DIPEA, DMSO.

SAR of the benzo[*b*][1,4]oxazin-3(4*H*)-one moiety: substitution pattern

The influence of substitution of the phenyl moiety of the benzo[*b*][1,4]oxazin-3(4*H*)-one core on ThyX inhibition was evaluated. The introduction of small electron-withdrawing groups, such as nitro and acetyl groups (compounds **17c** and **17d**, respectively) and electron-donating groups, such as a methyl moiety (compound **17c**) yielded analogues that were completely devoid of mycobacterial ThyX inhibition (Table 1). On the other hand, substituting the phenyl ring with a chlorine at position 6 afforded compound **17a**, which was equipotent with hit compound **10** (relative inhibition of 0.7 at $5\ \mu M$). The



Scheme 5. Reagents and conditions: a) (i) ethyl 3-bromopropionate, K₂CO₃, DMF, (ii) LiOH, THF, (iii) DCC, HOBT, *N*-phenylpiperazine-1-carboxamide, DMF; b) DCC, HOBT, *N*-phenylpiperazine-1-carboxamide, DMF; c) TFA, CH₂Cl₂.

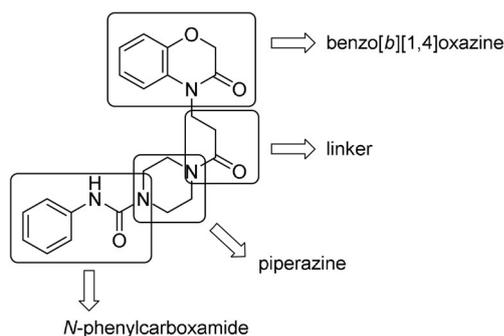


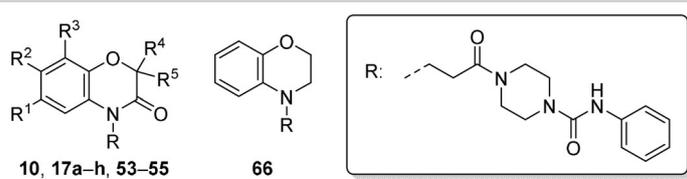
Figure 3. Division of compound 10 into different substructures.

chlorinated regioisomer (compound 17h) was endowed with less inhibitory activity against ThyX (relative inhibition of 0.47 at 5 μM), whereas a fluorine at position 7 afforded compound 17g which was still endowed with reasonable ThyX inhibition (IC₅₀ = 1.75 μM). A dichlorinated analogue (compound 17f) was completely devoid of ThyX inhibition. A sterically demanding

phenyl group (compound 17e) led to complete loss of ThyX inhibition. The substitution pattern of the oxazine moiety was also explored. Introduction of a methyl or fluorine(s) (compounds 53–55) completely abolished ThyX inhibition. Reduction of the lactam functionality of compound 10 into an amine afforded compound 66, which was much less active as a ThyX inhibitor (relative inhibition of 0.35 at 5 μM). Overall, this SAR study reveals that very little variation is tolerated on the benzo[*b*][1,4]oxazin-3(4*H*)-one moiety with respect to ThyX inhibition.

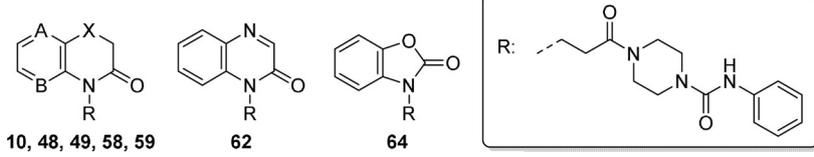
SAR of the benzo[*b*][1,4]oxazin-3(4*H*)-one moiety: scaffold modification

As pyridine is a known bioisostere of benzene, two pyrido-oxazine analogues were prepared. Whereas the pyrido[3,2-*b*][1,4]oxazin-3(4*H*)-one congener 48 was endowed with promising ThyX inhibition (IC₅₀ = 1.72 μM), its isomeric pyrido[2,3-*b*][1,4]oxazin-2(3*H*)-one analogue 49 completely lacks ThyX inhibition, pointing toward a crucial role of the position of the nitro-

Table 1. SAR of the benzo[*b*][1,4]oxazin-3(4*H*)-one moiety: substitution pattern.


Compd	R ¹	R ²	R ³	R ⁴	R ⁵	Relative inhibition @ 5 μm	IC ₅₀ [μM] ^[a]
10 (hit)	H	H	H	H	H	0.70	0.93 ± 0.17
17a	Cl	H	H	H	H	0.72	ND
17b	CH ₃	H	H	H	H	0.28	ND
17c	NO ₂	H	H	H	H	0	ND
17d	COCH ₃	H	H	H	H	0.10	ND
17e	Ph	H	H	H	H	0.04	ND
17f	Cl	H	Cl	H	H	0.19	ND
17g	H	F	H	H	H	0.81	ND
17h	H	Cl	H	H	H	0.47	ND
53	H	H	H	CH ₃	H	0	ND
54	H	H	H	F	H	0.15	ND
55	H	H	H	F	F	0.15	ND
66	-	-	-	-	-	0.35	ND

[a] Values are the mean ± SD of 3–4 independent experiments performed in duplicate; ND: not determined.

Table 2. SAR of scaffold-modified benzoxazine analogues.


Compd	A	B	X	Relative inhibition @ 5 μm	IC ₅₀ [μM] ^[a]
10 (hit)	CH	CH	O	0.70	0.93 ± 0.17
48	CH	N	O	1.24	1.72 ± 0.06
49	N	CH	O	0	ND
58	CH	CH	S	0.16	ND
59	CH	CH	CH ₂	0	ND
62	-	-	-	1.13	1.46 ± 0.05
64	-	-	-	0.06	ND

[a] Values are the mean ± SD of 3–4 independent experiments performed in duplicate; ND: not determined.

gen atom (Table 2). Other scaffold variations focused on the oxazine moiety and included a benzo[*b*][1,4]thiazine (compound **58**) and tetrahydroquinoline scaffold (compound **59**), which were both found to be inactive. The quinoxaline analogue **62** exhibited potent ThyX inhibition, displaying an IC₅₀ value of 1.46 μM. Converting the six-membered oxazine moiety into a five-membered ring yielded the benzo[*d*]oxazol-2(3*H*)-one derivative **64**, which was completely devoid of ThyX inhibition.

SAR of the *N*-phenylcarboxamide moiety

To probe into the optimal substitution pattern of the terminal phenyl ring, a variety of small substituents (e.g., halogens, methyl, methoxy, cyano, and acetyl) were introduced. As can be derived from the data in Table 3, significant structural variety is tolerated at this position, as these different analogues

(**22a–i**) display relative inhibition between 0.44 and 0.95 at 5 μM, relative to 50 μM folic acid under identical conditions. For a selected number of derivatives, IC₅₀ values were obtained. Thienyl and pyridyl are both known isosteres of phenyl, and therefore compounds **22j**, **27**, and **28** were also prepared. Only compound **27** was more active than compound **10** at 5 μM (relative inhibition of 1.19, IC₅₀ = 2.38 μM).

Instead of the terminal phenyl group, a benzyl and phenethyl congener were also prepared. Although elongation with one carbon (compound **22k**) still gives a potent ThyX inhibitor (relative inhibition of 0.83), further elongation by two carbons (compound **22l**) gave rise to diminished ThyX inhibition. To assess the importance of the urea moiety, the corresponding carbamate (compound **22m**) and amide (compound **22n**) were both prepared. The carbamate was still endowed with reasonable ThyX inhibition (relative inhibition of 0.5), whereas the amide is much less active as a ThyX inhibitor.

Table 3. SAR of the *N*-phenylcarboxamide moiety.

Compd	R	Relative inhibition @ 5 μM	IC ₅₀ [μM] ^[a]
10 (hit)		0.70	0.93 ± 0.17
22a		0.09	ND
22b		0.23	ND
22c		0.58	ND
22d		0.62	ND
22e		0.76	ND
22f		0.61	ND
22g		1.06	2.04 ± 0.08
22h		0.44	ND
22i		0.77	ND
22j		0.52	ND
22k		0.83	ND
22l		0.54	ND
22m		0.52	ND
22n		0.31	ND
27		1.19	2.38 ± 0.08
28		1.13	3.56 ± 0.09

[a] Values are the mean ± SD of 3–4 independent experiments performed in duplicate; ND: not determined.

Rather than directly attaching a carbonyl to the piperazine moiety (generating a urea, carbamate or amide), a number of compounds were synthesized in which a methylene linker between the carbonyl group and the piperazine nitrogen was inserted (Table 4). This imparts the amine group with basic properties, which should have an impact on its biological activity, solubility, and permeability. Several congeners within this family were endowed with ThyX inhibitory activity (e.g., compounds **29a** and **29c**, both having a heteroaromatic substituent), and especially the *N*-methyl-*N*-phenylpropanamide derivative **29d** displayed potent ThyX inhibition with an IC₅₀ value of 0.69 μM .

Table 4. SAR of the piperazinyl-acetamide derivatives.

Compd	R	Relative inhibition @ 5 μM	IC ₅₀ [μM] ^[a]
29a		0.89	ND
29b		0.53	ND
29c		0.99	ND
29d		1.37	0.69 ± 0.08

[a] Values are the mean ± SD of 3–4 independent experiments performed in duplicate; ND: not determined.

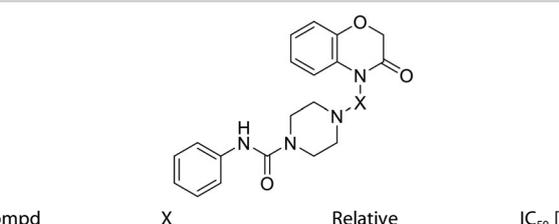
SAR of the propanoyl linker

Subsequently, the importance of the linker between the oxazine and the piperazine moiety was evaluated (Table 5). Shortening (compound **32**), elongation (compound **33**), or branching (compounds **36** and **37**) of the propanoyl linker inevitably led to completely inactive compounds. Reduction of the amide group yielded a derivative with an *n*-propyl linker (compound **39**) that similarly showed a complete loss of ThyX inhibition.

SAR of the piperazine moiety

As shown in Table 6, substituting the piperazine ring of hit compound **10** for a 1,4-diazepine ring (homopiperazine analogue **41**) resulted in a completely inactive derivative. Upon introducing a methyl group on the piperazine ring (compounds **45a** and **45b**), the exact position on the piperazine ring plays a determining role in its ThyX inhibitory activity. The presence of a methyl group closer to the benzoxazine core afforded

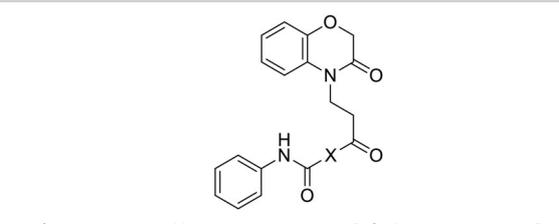
Table 5. SAR of linker modified analogues.



Compd	X	Relative inhibition @ 5 μM	IC ₅₀ [μM] ^[a]
10 (hit)		0.70	0.93 ± 0.17
32		0	ND
33		0	ND
36		0	ND
37		0	ND
39		0	ND

[a] Values are the mean ± SD of 3–4 independent experiments performed in duplicate; ND: not determined.

Table 6. SAR of the piperazine moiety.



Compd	X	Relative inhibition @ 5 μM	IC ₅₀ [μM] ^[a]
10 (hit)		0.70	0.93 ± 0.17
41		0	ND
45 a		0.47	ND
45 b		0.88	ND
45 c		1.23	0.88 ± 0.08
45 d		0.04	ND
45 e		0.58	ND
45 f		0	ND

[a] Values are the mean ± SD of 3–4 independent experiments performed in duplicate; ND: not determined.

compound **45 b** that was endowed with potent ThyX inhibition (relative inhibition of 0.88). On the other hand, when the same methyl group was introduced closer to the urea functionality (compound **45 a**), decreased potency was observed relative to the unsubstituted congener **10**.

To broaden the SAR investigation, the piperazine moiety was replaced by other nitrogen-containing saturated heterocycles. The aminopiperidine congener **45 c** displayed better activity against mycobacterial ThyX (IC₅₀ = 0.88 μM) than the parent piperazine analogue **10**. The closely related aminopiperidine congener **45 d** lacks activity. A similar profile was found in the aminopyrrolidine subseries. When the exocyclic amino group of aminopyrrolidine was derivatized as a urea (compound **45 e**), the compound was endowed with substantial ThyX inhibition (IC₅₀ = 2.95 μM). On the other hand, when the exocyclic amino group was connected to the propanyl linker (compound **45 f**), it was completely devoid of inhibitory activity.

Mechanism of inhibition of mycobacterial ThyX

The inhibition profile of five compounds (**22 a**, **22 c**, **22 e**, **22 i**, **22 k**) was probed as described in the Experimental Section. The results were similar for all tested compounds, and the data obtained for compound **22 e** as a representative example are

shown in Figure 4. The IC₅₀ versus dUMP plot (Figure 4A) suggested that a noncompetitive inhibition pattern can best describe the inhibition by selected molecules against dUMP. The same set of molecules showed a competitive inhibition pattern against CH₂H₄folate with a linear variation of IC₅₀ in function of the CH₂H₄folate concentration. An analogous plot for NADPH revealed a characteristic behavior for uncompetitive inhibition for the same set of molecules regarding NADPH binding. These results suggest that molecules could bind ThyX as a complex of type ThyX–NADPH or ThyX–NADP⁺.

The observation that the IC₅₀ values determined from the dose–response curves were of the same order of magnitude as the total enzyme concentration used suggested a tight-binding inhibition mechanism for the tested molecules, as previously described for 2-hydroxy-1,4-naphthoquinone analogues as tight-binding inhibitors for *PBCV-1* ThyX.^[13]

The qualitative evaluation of the ability of molecules **22 a**, **22 c**, **22 e**, **22 i**, **22 k**, **27**, **28**, **29 d**, **48** and **62** to inhibit ThyX from species besides *M. tuberculosis* (e.g., *H. pylori*, *PBCV-1*, *B. hermsi*, *C. trachomatis*) was also undertaken. None of these analogues inhibited ThyX from *PBCV-1* and *B. hermsi*, whereas weak inhibition at 50 μM was measured for ThyX from *H. pylori* and *C. trachomatis*. Analysis of the different ThyX protein se-

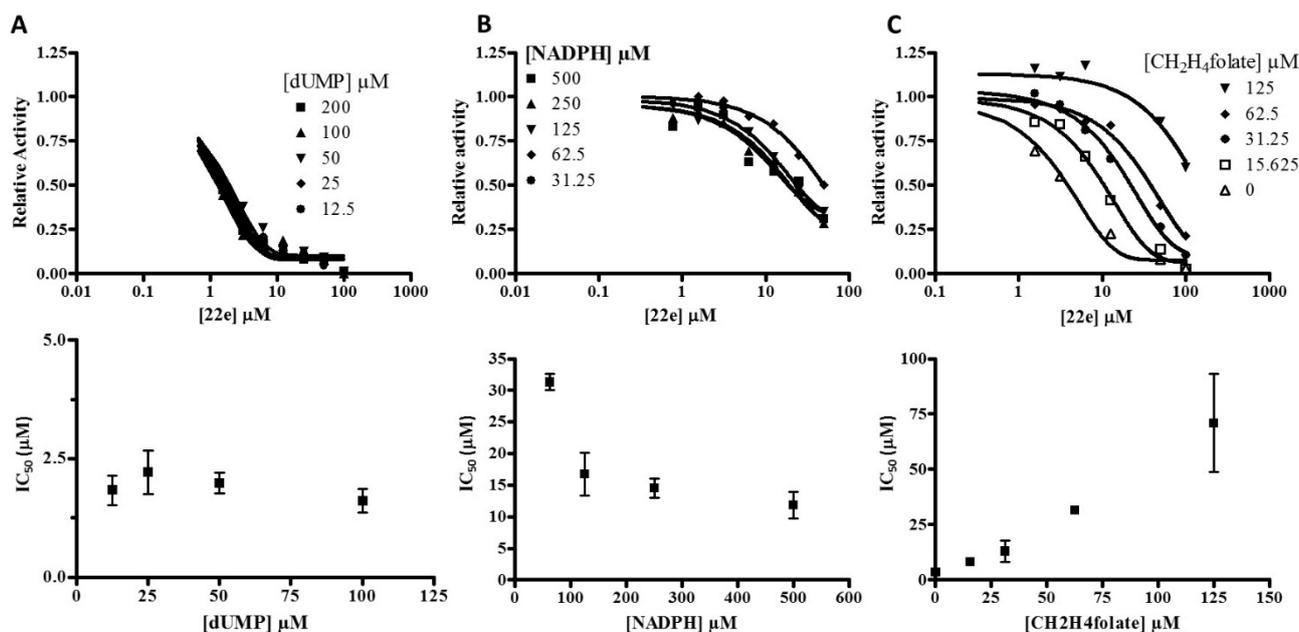


Figure 4. Secondary and final plots of results obtained with compound **22e** for mycobacterial ThyX inhibition using various concentrations of A) dUMP, B) NADPH, and C) methylene tetrahydrofolate.

quences revealed, in the environment of the substrate binding pocket, some interesting differences for presumed amino acids implicated in compound binding.

Molecular docking

Because competition assays indicated that studied compounds share their binding site with tetrahydrofolate, hit compound **10** was docked into the tetrahydrofolate binding site of ThyX (Figure 5). The crystal structure of *T. maritima* ThyX in complex with FAD, dUMP, and folinic acid (PDB ID: 4GTA)^[23] was used as starting structure to perform an initial molecular docking experiment. The aromatic and saturated rings of the benzo[*b*][1,4]oxazine moiety of **10** were posed on the corresponding rings of the pterin moiety of folinic acid, while the piperazine group of compound **10** fitted the aromatic ring in the *p*-aminobenzoic acid moiety of folinic acid. Overlaying the obtained complex with the crystal structure of *M. tuberculosis* ThyX in complex with FAD, 5-Br-dUMP, and glycerol (PDB ID: 2AF6)^[24] revealed that compound **10** is accommodated in a binding site of *M. tuberculosis* ThyX that contains two glycerol molecules in the crystal structure, occupying the binding pockets of the *N*-phenylcarboxamide and benzo[*b*][1,4]oxazine moieties of compound **10**. Starting from this overlay, two complexes of *M. tuberculosis* ThyX were generated after removal of both glycerols mentioned above. More specifically, both complexes contained the original FAD and 5-Br-dUMP together with compound **10** or folinic acid for comparison purposes. After an initial minimization, both complexes were subjected to 20 ns molecular dynamics calculations using the AMBER 16.0 software package^[25] as described in the Experimental Section.

During molecular dynamics, the benzo[*b*][1,4]oxazine part of compound **10** and the pterine scaffold of folinic acid remained

nically stacked to the flavin moiety of the FAD cofactor in the active site. Remarkably, two interconverting binding modes were observed for the glutamate moiety in folinic acid resulting in excessive conformational disorder, thereby explaining the absence of this moiety and some residues in this region of

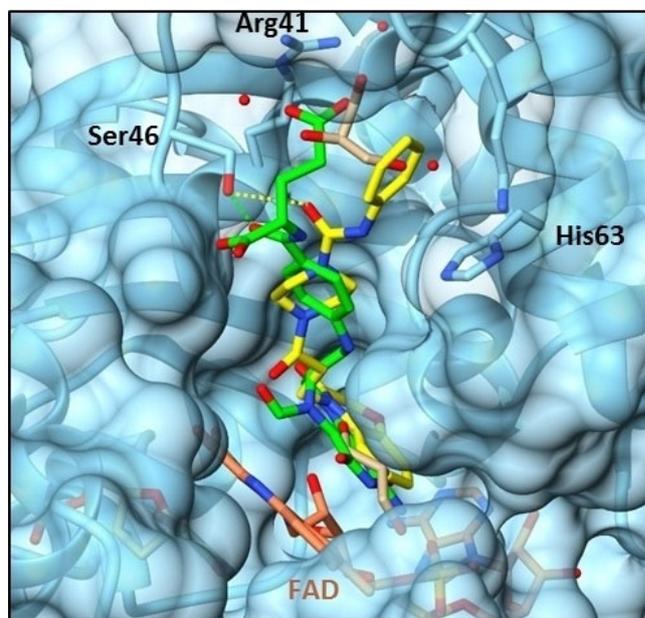


Figure 5. Compound **10** (yellow) in the catalytic site *M. tuberculosis* ThyX crystal structure (PDB ID: 2AF6) including FAD (orange) and 5-Br-dUMP. Modeled folinic acid (green) and glycerol (gold) in the binding site of the crystal structure are overlaid. Side chains of Arg41, Ser46, and His63 in subunit D are depicted. Green and yellow dashed lines indicate putative hydrogen bonds between Ser46 and the carboxamide linker in folinic acid and compound **10**, respectively.

the crystal structure of the *T. maritima* ThyX complex. In the most abundant binding mode as observed in molecular dynamics simulations, the γ -carboxyl group of the glutamate moiety interacts with a conserved Arg residue, thereby replacing a glycerol in the initial crystal structure (*M. tuberculosis*: Arg41 in PDB ID: 2AF6, *T. maritima*: Arg28 in PDB ID: 4GTA, *H. pylori*: Arg42 in PDB ID: 3AH5; Figure 6). In the same binding mode, the *N*-phenylcarboxamide of **10** shows a stacking interaction with the side chain of His63 in *M. tuberculosis* ThyX (Figure 5).

Strikingly, the carboxamide linkers in compound **10** as well as in folinic acid are within hydrogen bonding distance with the terminal hydroxy group of Ser46 in modeled complexes. While Ser46 is conserved in *H. pylori*, it is mutated to Glu46 in ThyX homologues from *B. hermsi* and *PBCV-1* for which no inhibitory activity could be observed. To examine the effect of this mutation on the binding affinity of **10**, Ser46 was mutated in silico to Glu46. The results indicated that this mutation does not allow formation of a hydrogen bond with the carboxamide linker in compound **10** and impaired the stability of the loop that forms the studied binding site. Additionally, a distortion of

binding mode is observed at the piperazine linker. All of these effects contributed to a decreased binding affinity of the compounds, explaining the loss of inhibitory effect on ThyX from *PBCV-1* and *B. hermsi* as described above.

The model of compound **10** bound to *M. tuberculosis* ThyX enabled rationalization of the SAR results (Figure 7). The carbonyl in the benzoxazine moiety forming the δ -lactam ring accepts a hydrogen bond from Ser102, the only residue of subunit D that directly interacts with compound **10** in our model. Intermittently, the same Ser102 donates a hydrogen bond to the oxygen atom in the 1,4-oxazine moiety, yet this bifurcated hydrogen bond was weaker because of the more spread-out orientation of both acceptors. The insertion of a nitrogen atom in the phenyl part of the benzo[*b*][1,4]oxazine moiety yielded pyrido[3,2-*b*][1,4]oxazine **48**, from which the nitrogen atom acts as an isostere for O4 of the pterin ring of the original cofactor and aligns the attached oxazine and aromatic ring stronger with the heterocycle of FAD in a parallel-displaced orientation as observed for compound **10**. Additionally, the more polar character of the pyridine ring of compound **48** allows a hydrogen bond with His69. The absence of the foremen-

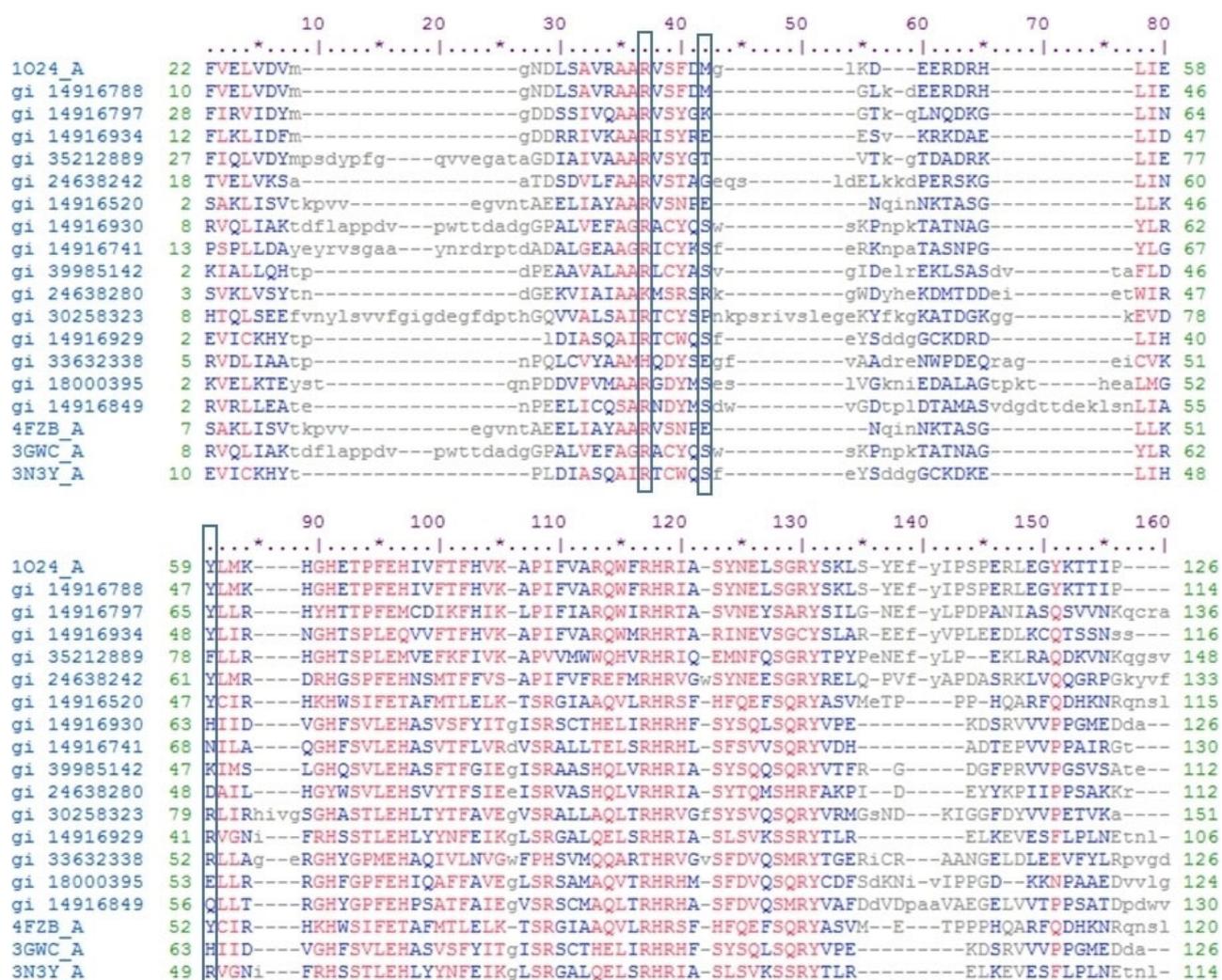


Figure 6. Sequence alignment of ThyX homologues by CDD/SPARCLE.^[26] Residues that are labeled in Figure 5 are boxed.

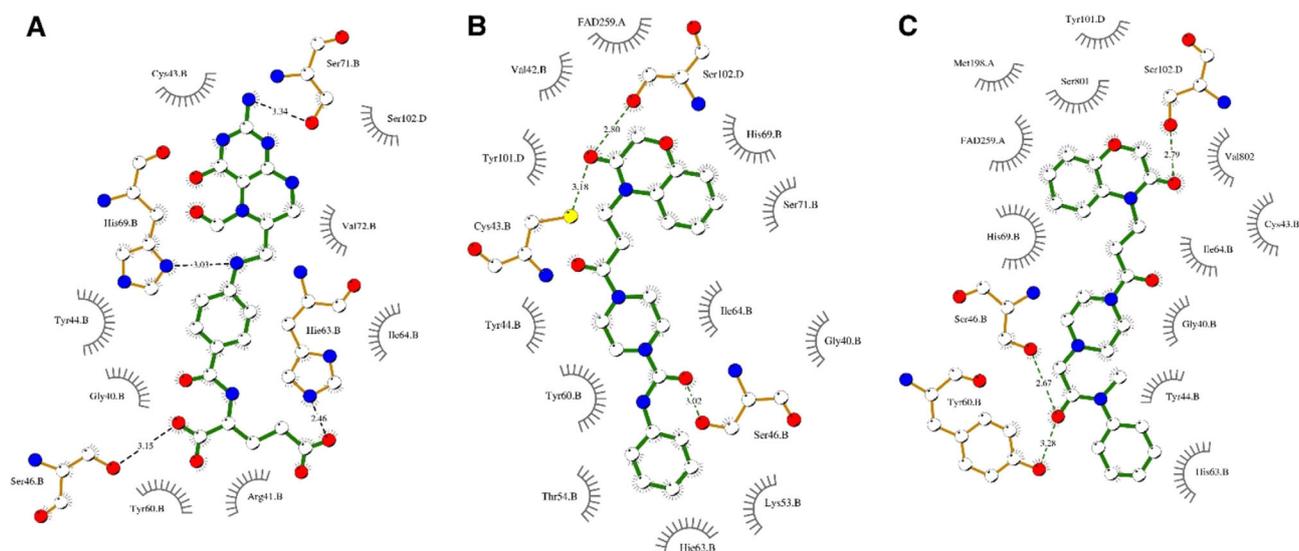


Figure 7. LigPlot+ representation of interacting residues with A) native folic acid, B) hit compound **10**, and C) optimized compound **29d**.

tioned heterocycles in compound **59**, together with its complete lack of ThyX inhibition, support this interaction analysis.

The introduction of a methoxy moiety at the *para* position in **22b** during the SAR of the *N*-phenylcarboxamide moiety was expected to destabilize the flexible closing loop which contributes to the lower activity. For comparison, the *meta*-acetyl variant (compound **22g**) provides a stronger acceptor (no mesomeric effect with carboxamide), which might accept an additional hydrogen bond from Thr54 resulting in a tighter loop closing with relative inhibition of 1.06 compared with 0.70 in compound **10**. The nitrogen atom from the pyridine moieties in compounds **27** and **28** occupies a smaller volume than the acetyl group in **22b** with higher atomic radius and could act as a hydrogen bond acceptor for the conserved Arg41. Such increased loop stabilization correlates with the higher relative inhibition values of 1.19 and 1.13 observed for compounds **27** and **28**, respectively. The benzyl derivative **22k** covers a larger rotational space than the phenyl analogue **10**, providing increased potential stacking interactions with His63 and Tyr54. Moreover, the increased flexibility of compound **22k** resulted in a more stable stacking interaction between the benzoxazine moiety and the heterocycle of FAD, maintaining the His63 stacking interaction and Ser46 hydrogen bond.

With an additional methylene bridge adjacent to the piperazine linker in compound **29d**, an additional source of rotational freedom was introduced, similarly as for compound **22k**. A model of *M. tuberculosis* ThyX in complex with the most active ThyX inhibitor (compound **29d**) was obtained as described in the Experimental Section. The additional *N*-methyl group contributes to hydrophobic van der Waals interactions with Tyr44, thereby positioning the phenyl ring toward optimal perpendicular stacking interaction with His63 and optimizing the geometry of hydrogen bond formation with Ser46. Because of the increased flexibility resulting from the methylene group, a bifurcated hydrogen bond formation of the carbonyl group in the *N*-phenylcarboxamide was viable with Ser46 and Tyr60 which

concurrently positioned the phenyl moiety through stacking. Compared with the binding conformation of compound **10**, this overall network of stabilizing hydrophobic and increased electrostatic interactions in **29d** allows a volumetric occupation of the binding space of the glutamate moiety in folic acid flanked by Arg41 and His63. These results were further quantified by interaction energy calculations, in which a difference of 8.69 kcal mol⁻¹ was observed with a more favorable affinity for **29d** (folic acid: -37.90 kcal mol⁻¹, compound **10**: -38.77 kcal mol⁻¹, compound **29d**: -46.59 kcal mol⁻¹).

Furthermore, an overall slight rotational shift resulted in more planar stacking interaction between the benzoxazine moiety and the FAD cofactor. Finally, an additional hydrogen bond was possible between one hydroxy group from the glycerol linker in FAD and the oxygen atom from the benzoxazine ring. The same reasoning was applied to compounds **29a** and **29c**, explaining the potent inhibitory activity of both compounds (relative inhibition at 5 μM of 0.89 and 0.99, respectively).

In compound **45c**, the exocyclic amino group at the piperidine moiety had a similar effect as the extra methylene in compounds **29a**, **29c** and **29d**. Alterations in the piperazine moiety, albeit with limited conformational space, displayed a positioning role for the binding mode of this central part in the entire compound. Compounds **45d** and **45f** shifted the heterocycle by one atom, resulting in loss of activity and thereby endorsing the positioning function. The methyl side chain in compound **45b** resided in a more hydrophobic pocket displaying hydrophobic interactions with Val72 and Ile64 resulting in a slightly higher relative inhibition (0.88).

In the SAR of the propanoyl linker (compounds **32**, **33**, **36**, **37** and **39**), the loss of inhibitory activity by shortening the linkers provided evidence for the necessary co-occurrence of both His63 stacking and Ser46 and Tyr60 bifurcated hydrogen bond interactions with benzoxazine alignment to the FAD heterocycle.

Conclusions

Starting from the previously identified hit compound **10**, a hit-to-lead optimization campaign was performed to discover novel mycobacterial ThyX inhibitors. The most potent analogue of this series was endowed with an IC₅₀ value of 0.69 μM against mycobacterial ThyX. Molecular modeling was applied to understand the SAR for *M. tuberculosis* ThyX and the variation in activity toward ThyX from other microorganisms. These results open new avenues for structure-based drug design to discover more potent ThyX inhibitors with antibacterial activity. A phenotypic screening for antimicrobial activity of the most potent molecules on specific bacteria will be the first step to address the path forward for these compounds.

Experimental Section

Chemistry

General: All reagents and solvents were purchased from commercial sources and used as obtained. Moisture-sensitive reactions were carried out using oven-dried glassware under a nitrogen or argon atmosphere. ¹H and ¹³C NMR spectra were recorded on a Bruker Avance 300 or 500 MHz spectrometer with tetramethylsilane as internal standard or referenced to the residual solvent signal. The following abbreviations were used to explain multiplicities: s=singlet, d=doublet, t=triplet, q=quartet, m=multiplet, br=broad, dd=doublet of doublets. Coupling constants are expressed in Hz. High-resolution mass spectra (HRMS) were obtained on a quadrupole orthogonal acceleration time-of-flight mass spectrometer (Synapt G2 HDMS, Waters, Milford, MA, USA). Samples were infused at 3 μL min⁻¹, and spectra were obtained in positive (or negative) ionization mode with a resolution of 15000 (FWHM) using leucine enkephalin as the lock mass. Pre-coated aluminum sheets (254 nm) were used for TLC, and spots were visualized with UV light. The products were purified by flash column chromatography on silica gel (60 Å, 0.035–0.070 mm, Acros Organics). The purity of final compounds was determined by analytical RP-HPLC, using one of the following methods. Method A: a Waters 600 HPLC system and a Waters 2996 photodiode array detector, XBridge column (C₁₈, 5 μm, 4.6 mm×150 mm). Elution with a gradient mixture of H₂O containing 0.2% (vol) of TFA (A) and acetonitrile (B). Method B: a Shimadzu HPLC equipped with a LC-20AT pump, DGU-20A5 degasser, and a SPD20A UV/Vis detector; Symmetry C₁₈ column (5 μm, 4.6 mm×150 mm); gradient elution of H₂O/CH₃CN from 95:5 or 70:30 to 5:95 over 25 min; flow rate 1 mL min⁻¹; wavelength 254 nm. Compounds that were submitted for biological testing displayed a purity of 95% or more.

N-Phenylpiperazine-1-carboxamide (16): To a solution of *N*-Boc piperazine **14** (0.50 mmol) in THF (10 mL) was added slowly a solution of phenyl isocyanate (0.57 mmol) in THF (10 mL) at room temperature. After the reaction reached completion, trifluoroacetic acid (10 mL) was added dropwise while stirring at 0 °C. When TLC showed complete deprotection, H₂O (100 mL) was added. The mixture was extracted with CH₂Cl₂ (30 mL). The aqueous phase was adjusted to pH 9 with 1 N NaOH and then extracted with CH₂Cl₂ (100 mL). The organic phase was dried over Na₂SO₄. The solvents were evaporated in vacuo to yield a crude residue, which was used as such for further reaction without any additional purification. ¹H NMR (300 MHz, [D₆]DMSO): δ=2.7 (brs, 2H, CH₂), 3.34 (t, J=4.35 Hz, 4H, 2CH₂), 4.22 (brs, 2H, CH₂), 6.93 (t, J=7.2 Hz, H, CH),

7.33 (t, J=7.55 Hz, 2H, 2CH), 7.45 (d, J=7.8 Hz, 2H, 2CH), 8.53 ppm (s, H, CH); ¹³C NMR (75 MHz, [D₆]DMSO): δ=31.68, 36.66, 67.17, 115.35, 116.77, 122.93, 123.68, 128.31, 145.07, 164.13, 172.35 ppm.

4-(3-(6-Chloro-3-oxo-2H-benzo[b][1,4]oxazin-4(3H)-yl)propanoyl)-N-phenylpiperazine-1-carboxamide (17a): A mixture of 6-chloro-2H-benzo[b][1,4]oxazin-3(4H)-one (91 mg 0.5 mmol), K₂CO₃ (158 mg, 1 mmol) and ethyl 3-bromopropionate (91 mg, 0.51 mmol) in DMF (15 mL), was stirred at 80 °C for 8 h. Then, the mixture was diluted with H₂O (45 mL) and extracted with CH₂Cl₂ (100 mL). The organic layer was separated, washed with H₂O (20 mL) and dried over anhydrous Na₂SO₄. The solvents were removed under reduced pressure. The crude residue was dissolved in THF (10 mL) and a solution of LiOH (210 mg 5 mmol) in H₂O (10 mL) was added. The reaction mixture was stirred at 60 °C for 12 h. After the reaction reached completion, the solution was adjusted to pH 3 by the addition of a 2 N HCl solution. The mixture was extracted with EtOAc (5×20 mL) and dried over anhydrous Na₂SO₄. Evaporation of the solvent yielded the carboxylic acid **13a**. The compound was re-dissolved in DMSO (10 mL). *N*-Phenylpiperazine-1-carboxamide **16** (102 mg 0.50 mmol), HCTU (206 mg, 0.50 mmol) and DIPEA (50 μL) were added. The reaction mixture was stirred overnight at room temperature. The reaction mixture was diluted with CH₂Cl₂ (100 mL) and washed with H₂O (40 mL), dried over Na₂SO₄, concentrated and purified by flash column chromatography to afford the title compound (318 mg, 72%). Purity (Method A): 95.86%. ¹H NMR (300 MHz, CDCl₃): δ=2.75 (t, J=7.6 Hz, 2H, CH₂), 3.50–3.57 (m, 6H, 6CH), 3.75 (t, J=7.5 Hz, 2H, CH₂), 4.25 (t, J=7.8 Hz, 2H, CH₂), 4.60 (s, 2H, CH₂), 6.75 (s, H, NH), 6.94–7.10 (m, 4H, CH), 7.31–7.37 ppm (m, 4H, 4CH); ¹³C NMR (75 MHz, CDCl₃): δ=30.69, 38.26, 41.21, 44.15, 45.27, 68.18, 108.06, 115.08, 118.42, 120.45, 123.77, 124.00, 129.13, 138.90, 140.06, 144.24, 155.60, 164.64, 169.28 ppm; HRMS (ESI): *m/z* [M+H]⁺ calculated for C₂₂H₂₄N₄O₄Cl 443.14859, found 443.1476.

Compounds **17b–h** were synthesized according to the procedure for the preparation of compound **17a**. Exact experimental and spectral data can be found in the Supporting Information.

Ethyl 3-(3-oxo-2,3-dihydro-4H-benzo[b][1,4]oxazin-4-yl)propanoate (19): A solution of **18** (2.98 g, 20 mmol) and K₂CO₃ (40 mmol, 5.52 g) in DMF (45 mL) was stirred at room temperature for 15 min. Then, ethyl 3-bromopropionate (22 mmol, 3.98 g) was added and the mixture was stirred overnight at 90 °C. The precipitate was filtered off and the filtrate was evaporated in vacuo. The crude residue was purified by silica gel flash chromatography (the mobile phase being a mixture of heptane and EtOAc in a ratio of 7:3) affording the title compound as a colorless oil (4.7 g, 95%). ¹H NMR (300 MHz, CDCl₃): δ=7.1–6.95 (m, 4H), 4.58 (s, 2H, O-CH₂-CO), 4.23 (t, 2H, J=7.5 Hz, NCH₂), 4.13 (q, 2H, J=7.1 Hz, CO-CH₂), 2.68 (t, 2H, J=7.7 Hz, N-CH₂-CH₂), 1.23 ppm (t, 3H, J=7.1 Hz, CH₃); ¹³C NMR (75 MHz, CDCl₃): δ=171.1, 164.5, 145.5, 128.3, 124.2, 123.0, 117.4, 114.7, 67.7, 61.0, 37.2, 32.1, 14.2 ppm; HRMS (ESI): *m/z* [M+H]⁺ calculated for C₁₃H₁₆NO₄ 250.10737, found 250.1071.

3-(3-Oxo-2,3-dihydro-4H-benzo[b][1,4]oxazin-4-yl)propanoic acid (20): To a solution of **19** (4.88 g, 19.58 mmol) in THF (70 mL) was added a solution of LiOH (4.93 g, 117.46 mmol) in H₂O (50 mL). The reaction mixture was stirred overnight at 45 °C. The reaction was quenched by the addition of a 5 M HCl solution (24 mL). The organic solvents of the reaction mixture were evaporated and a precipitate was formed. The precipitate was filtered off and dried to yield the title compound as a white-brown powder (3.96 g, 92%). ¹H NMR (300 MHz, [D₆]DMSO): δ=12.39 (s, 1H, OH), 7.24–7.15 (m,

1H), 7.10–6.95 (m, 1H), 4.61 (s, 2H, O-CH₂-CO), 4.11 (t, 2H, *J* = 7.6 Hz, NCH₂), 2.53 ppm (t, 2H, *J* = 7.5 Hz, N-CH₂-CH₂); ¹³C NMR (75 MHz, [D₆]DMSO): δ = 172.2, 164.0, 144.9, 128.2, 123.5, 122.8, 116.6, 115.2, 67.0, 36.5, 31.5 ppm; HRMS (ESI): *m/z* [M-H]⁻ calculated for C₁₁H₁₀NO₄ 220.06152, found 220.0637.

4-(3-Oxo-3-(piperazin-1-yl)propyl)-2H-benzo[b][1,4]oxazin-3(4H)-on (21): A solution of **20** (1.00 g, 4.52 mmol), HOBt (0.692 g, 4.52 mmol) and DCC (1.12 g, 5.42 mmol) in DMF (16 mL) was stirred for 2 h at room temperature. The reaction mixture was cooled down to room temperature and piperazine (0.779 g, 9.04 mmol) was added. The reaction was stirred overnight at room temperature. The precipitate was filtered off and the filtrate was evaporated in vacuo. The crude residue was purified by silica gel flash chromatography (mobile phase being a mixture of CH₂Cl₂/MeOH/NH_{3(aq)} in a ratio of 98:2:0.3), to yield the title compound as a light-yellow oil (0.4 g, 31 %). ¹H NMR (300 MHz, CDCl₃): δ = 7.15–6.95 (m, 4H), 4.57 (s, 2H, O-CH₂-CO), 4.24 (t, 2H, *J* = 7.8 Hz, N-CH₂-CH₂-CO), 3.60–3.37 (m, 4H), 2.90–2.78 (m, 4H), 2.69 (t, 2H, *J* = 7.8, N-CH₂-CH₂-CO), 2.50 (brs, 1H), 1.23 ppm (s, 1H, NH); ¹³C NMR (75 MHz, CDCl₃): δ = 168.8, 164.6, 145.4, 128.3, 124.2, 123.2, 117.3, 115.0, 67.7, 46.7, 46.2, 45.8, 42.6, 40.9, 38.2, 30.7, 29.8 ppm; HRMS (ESI): *m/z* [M+H]⁺ calculated for C₁₅H₂₀N₃O₃ 290.14990, found 290.1505.

N-(3-Chlorophenyl)-4-(3-(3-oxo-2,3-dihydro-4H-benzo[b][1,4]oxazin-4-yl)propanoyl)piperazine-1-carboxamide (22a): To a solution of compound **21** (89.7 mg, 0.31 mmol) in DMF (2 mL) was added 3-chlorophenyl isocyanate (71.4 mg, 0.47 mmol) and the resulting mixture was stirred overnight at room temperature. Reaction completion was monitored by TLC (CH₂Cl₂/MeOH/NH_{3(aq)} 90:10:0.3). The solvents were evaporated in vacuo and the crude residue was purified by silica gel column chromatography (EtOAc 100% as mobile phase) affording the title compound as a white solid (100 mg, 73 %). Purity (Method A): 99.83 %. ¹H NMR (300 MHz, CDCl₃): δ = 7.40 (s, 1H), 7.30–6.93 (m, 8H), 4.57 (s, 2H, O-CH₂-CO), 4.25 (t, 2H, *J* = 7.6 Hz, N-CH₂-CH₂-CO), 3.70–3.60 (m, 2H), 3.58–3.38 (m, 6H), 2.71 ppm (t, 2H, *J* = 7.9 Hz, N-CH₂-CH₂-CO); ¹³C NMR (75 MHz, CDCl₃): δ = 169.1, 164.7, 154.8, 145.4, 140.3, 134.4, 129.9, 128.2, 124.4, 123.4, 120.3, 118.3, 117.4, 114.8, 67.7, 45.2, 44.0, 43.7, 41.2, 37.9, 30.8 ppm; HRMS (ESI): *m/z* [M+H]⁺ calculated for C₂₂H₂₄ClN₄O₄ 443.14804, found 443.1486.

Compounds **22b–n** were synthesized according to the procedure for the preparation of compound **22a**. Exact experimental and spectral data can be found in the Supporting Information.

4-(3-(3-Oxo-2,3-dihydro-4H-benzo[b][1,4]oxazin-4-yl)propanoyl)-N-(pyridin-4-yl)piperazine-1-carboxamide (27): To a solution of 4-aminopyridine **23** (24.67 mg, 0.262 mmol), DIPEA (42.3 mg, 0.328 mmol) in DMF (1 mL) was added phenyl chloroformate (44.5 mg, 0.284 mmol) at 0 °C. The mixture was stirred for 15 min and completion of the reaction was monitored by TLC (CH₂Cl₂/MeOH/NH_{3(aq)} 92:8:0.3). DIPEA (42.3 mg, 0.328 mmol) and a solution of compound **21** (63.2 mg, 0.218 mmol) in DMF (1 mL) were added. The resulting reaction mixture was stirred overnight at room temperature. The solvents were evaporated in vacuo. The residue was dissolved in EtOAc and washed with H₂O. The organic phase was dried over Na₂SO₄ and concentrated in vacuo. The crude residue was purified by silica gel column chromatography (using a mixture of CH₂Cl₂ and MeOH as mobile phase, in a gradient gradually increasing from 5 to 10% MeOH) affording the title compound as an off-white solid (28 mg, 32 %). Purity (Method A): 100 %. ¹H NMR (300 MHz, CDCl₃): δ = 8.38–8.32 (m, 2H), 7.93 (brs, 1H), 7.42–7.35 (m, 2H), 7.10–6.95 (m, 4H), 4.57 (s, 2H, O-CH₂-CO),

4.24 (t, 2H, *J* = 7.6 Hz, N-CH₂-CH₂-CO), 3.70–3.42 (m, 8H), 2.71 ppm (t, 2H, *J* = 7.8 Hz, N-CH₂-CH₂-CO); ¹³C NMR (75 MHz, CDCl₃): δ = 169.1, 164.8, 154.2, 150.1 (2C), 147.1, 145.4, 128.2, 124.4, 123.2, 117.5, 114.8, 113.7 (2C), 67.7, 45.3, 44.2, 43.8, 41.2, 38.0, 30.9 ppm; HRMS (ESI): *m/z* [M+H]⁺ calculated for C₂₁H₂₄N₅O₄ 410.18226, found 410.1817.

Compound **28** was synthesized according to the procedure for the preparation of compound **27**. Exact experimental and spectral data can be found in the Supporting Information.

2-(4-(3-(3-Oxo-2H-benzo[b][1,4]oxazin-4(3H)-yl)propanoyl)piperazin-1-yl)-N-(pyridin-3-yl)acetamide (29a): To a solution of 2-(piperazin-1-yl)-N-(pyridin-3-yl)acetamide (110 mg, 0.5 mmol) in DMSO (10 mL) was added compound **20** (110 mg 0.50 mmol), HCTU (206 mg 0.5 mmol) and DIPEA (50 μL). The reaction mixture was stirred overnight at room temperature. Then, the mixture was diluted with CH₂Cl₂ (100 mL) and washed with H₂O (40 mL). The combined organic layers were dried over Na₂SO₄ and concentrated in vacuo. The crude residue was purified by silica gel flash column chromatography affording the title compound (171 mg, 81 %). Purity (Method A): 100 %. ¹H NMR (300 MHz, CDCl₃): δ = 2.60–2.62 (m, 4H, 2CH₂), 2.73 (t, *J* = 7.8 Hz, 2H, CH₂), 3.19 (s, 2H, CH₂), 3.56–3.58 (m, 2H, CH₂), 3.71–3.72 (m, 2H, CH₂), 4.28 (t, *J* = 7.65 Hz, 2H, CH₂), 4.59 (s, H, CH₂), 7.01–7.12 (m, 4H, 4CH), 7.27–7.31 (m, 1H, 1CH), 8.19–8.23 (m, 1H, 1CH), 8.32–8.37 (m, 1H, 1CH), 8.59 (d, *J* = 2.4 Hz, H, CH), 9.05 ppm (m, H, NH); ¹³C NMR (75 MHz, CDCl₃): δ = 30.74, 38.05, 41.62, 45.53, 53.20, 53.55, 61.92, 67.66, 114.86, 117.35, 123.16, 124.26, 126.92, 128.24, 134.22, 124.11, 141.01, 145.58, 164.63, 168.32, 168.87 ppm; HRMS (ESI): *m/z* [M+H]⁺ calculated for C₂₂H₂₆N₅O₄ 424.19846, found 424.1976.

Compounds **29b–d** were synthesized according to the procedure for the preparation of compound **29a**. Exact experimental and spectral data can be found in the Supporting Information.

Ethyl 2-(3-oxo-2,3-dihydro-4H-benzo[b][1,4]oxazin-4-yl)acetate (30): 1,4-Benzoxazin-3(4H)-one **18** (5.00 g, 33.52 mmol) and K₂CO₃ (13.90 g, 100.57 mmol) were dissolved in DMF (80 mL) and stirred at room temperature for 15 min. Then, ethyl bromoacetate (11.20 g, 67 mmol) was added and the reaction mixture was stirred overnight at 90 °C. The solvents were evaporated in vacuo and the residue was purified by silica gel column chromatography (using a mixture of cyclohexane and EtOAc in a ratio of 7:3 as mobile phase) affording the title compound as a colorless oil (7.70 g, 98 %). ¹H NMR (300 MHz, CDCl₃): δ = 7.05–6.94 (m, 3H), 6.78–6.70 (m, 1H), 4.65 (s, 2H), 4.63 (s, 2H), 4.22 (q, 2H, *J* = 7.14 Hz), 1.26 ppm (t, 3H, *J* = 7.12 Hz); ¹³C NMR (75 MHz, CDCl₃): δ = 167.8, 164.9, 145.1, 128.7, 124.3, 123.0, 117.2, 114.5, 67.5, 61.9, 43.0, 14.2 ppm; HRMS (ESI): *m/z* [M+Na]⁺ calculated for C₁₂H₁₃N₁O₄Na 258.07370, found 258.0749.

2-(3-Oxo-2,3-dihydro-4H-benzo[b][1,4]oxazin-4-yl)acetic acid (30'): To a solution of 30 (7.70 g, 32.73 mmol) in THF (80 mL) was added a solution of LiOH (8.24 g, 196.4 mmol) in H₂O (70 mL). The resulting reaction mixture was stirred overnight at 45 °C. After completion of the reaction, the reaction mixture was acidified by addition of an excess of HCl. The organic solvent (THF) was evaporated in vacuo resulting in the formation of a precipitate in H₂O. The precipitate was filtered off, washed with cold H₂O and dried furnishing the title compound as an off-white powder (6.24 g, 92 %). ¹H NMR (300 MHz, DMSO): δ = 13.09 (brs, 1H), 7.10–6.95 (m, 4H), 4.69 (s, 2H), 4.63 ppm (s, 2H); ¹³C NMR (75 MHz, DMSO): δ = 169.3, 164.5, 144.6, 128.7, 123.7, 122.7, 116.6, 115.3, 66.9, 42.4 ppm; HRMS (ESI): *m/z* [M-H]⁻ calculated for C₁₀H₈N₁O₄ 206.04587, found 206.0464.

4-(2-Oxo-2-(piperazin-1-yl)ethyl)-2H-benzo[b][1,4]oxazin-3(4H)-one (30''): 2-(3-Oxo-2,3-dihydro-4H-benzo[b][1,4]oxazin-4-yl)acetic acid (1.5 g, 7.24 mmol), HOBT·H₂O (1.11 g, 7.24 mmol) and DCC (1.49 g, 7.24 mmol) were dissolved in DMF (25 mL). The mixture was stirred for 2 h at 0 °C. After the addition of piperazine (1.25 g, 14.48 mmol), the reaction was stirred overnight at room temperature. The formed precipitate was filtered off and discarded. The filtrate was concentrated in vacuo and the residue was purified by silica gel column chromatography (CH₂Cl₂/MeOH/NH_{3(aq)}, 92:8:0.3) to yield the title compound as a white powder (742 mg, 37%). ¹H NMR (300 MHz, DMSO): δ = 7.07–6.98 (m, 3H), 6.93–6.88 (m, 1H), 4.78 (s, 2H), 4.66 (s, 2H), 3.53–3.47 (m, 2H), 3.42–3.36 (m, 2H), 2.84–2.77 (m, 2H), 2.73–2.66 ppm (m, 2H); ¹³C NMR (75 MHz, DMSO): δ = 164.4, 164.2, 144.6, 129.1, 123.4, 122.6, 116.4, 115.6, 66.9, 45.6, 45.2, 45.1, 42.4, 42.3 ppm; HRMS (ESI): *m/z* [M + H]⁺ calculated for C₁₄H₁₈N₃O₃ 276.13425, found 276.1350.

4-(2-(3-Oxo-2,3-dihydro-4H-benzo[b][1,4]oxazin-4-yl)acetyl)-N-phenylpiperazine-1-carboxamide (32): Compound **32** was synthesized according to the procedure described for the synthesis of compound **22a**, affording the title compound as a white powder (89 mg, 89%). Purity (Method A): 98.76%. ¹H NMR (300 MHz, CDCl₃): δ = 7.40–7.25 (m, 3H), 7.10–6.95 (m, 4H), 6.82–6.75 (m, 1H), 6.57 (brs, 1H), 4.72 (s, 2H), 4.67 (s, 2H), 3.75–3.60 (m, 6H), 3.55–3.43 (m, 2H), 1.64 ppm (brs, 1H); ¹³C NMR (75 MHz, CDCl₃): δ = 165.4, 165.2, 155.1, 145.3, 138.7, 129.13 (2C), 129.08, 124.4, 123.7, 123.1, 120.3 (2C), 117.3, 115.0, 67.7, 44.9, 44.3, 43.5, 43.2, 41.8 ppm; HRMS (ESI): *m/z* [M + H]⁺ calculated for C₂₁H₂₃N₄O₄ 395.17136, found: 395.1708.

Compounds **33**, **36** and **37** were synthesized according to the procedure for the preparation of compound **32**. Exact experimental and spectral data can be found in the Supporting Information.

4-(3-(3-Oxo-2,3-dihydro-4H-benzo[b][1,4]oxazin-4-yl)propyl)-N-phenylpiperazine-1-carboxamide (39): A mixture of 2H-benzo[b][1,4]oxazin-3(4H)-one **18** (75 mg, 0.5 mmol), K₂CO₃ (158 mg, 1.0 mmol) and *tert*-butyl 4-(3-bromopropyl)piperazine-1-carboxylate (169 mg, 0.55 mmol) in DMF (5 mL) was stirred at 65 °C for 18 h. When the reaction was finished, the mixture was diluted with H₂O (15 mL) and extracted with CH₂Cl₂ (3 × 10 mL). The organic phases were combined, washed with H₂O (15 mL), dried over anhydrous Na₂SO₄ and evaporated. The crude residue was dissolved in a 80% TFA solution (5 mL). The mixture was stirred for 4 h. After evaporation of the solvents, the crude residue was dissolved in THF (5 mL) and phenylisocyanate (66 mg, 0.55 mmol) was added. The mixture was stirred at room temperature overnight. The resulting mixture was diluted with CH₂Cl₂ (20 mL) and washed with H₂O (2 × 20 mL). The combined organic layers were dried over anhydrous Na₂SO₄ and concentrated under reduced pressure. The residue was purified by silica gel flash column chromatography (Et₂O/MeOH 10:1) to afford the product (101 mg, 51%). Purity (Method A): 98.85%. ¹H NMR (300 MHz, CDCl₃): δ = 1.86 (p, *J* = 7.0 Hz, 2H, CH₂), 2.42–2.47 (m, 6H, 3 × CH₂), 3.47–3.51 (m, 4H, 2 × CH₂), 4.02 (d, *J* = 7.0 Hz, 2H, CH₂), 4.59 (s, 2H, CH₂), 6.52 (s, 1H, NH), 6.98–7.10 (m, 5H, 4 × CH_{Ar}), 7.24–7.30 (m, 2H, 2 × CH_{Ar}), 7.33–7.37 ppm (m, 2H, 2 × CH_{Ar}); ¹³C NMR (75 MHz, CDCl₃): δ = 24.41, 39.46, 41.21, 52.94, 55.50, 67.75, 114.99, 117.32, 120.09, 122.85, 123.21, 123.99, 128.64, 128.97, 139.13, 145.52, 155.15, 164.43 ppm.

4-(3-(1,4-Diazepan-1-yl)-3-oxopropyl)-2H-benzo[b][1,4]oxazin-3(4H)-on (40): A solution of compound **20** (0.800 g, 3.62 mmol), HOBT (0.554 g, 3.62 mmol) and DCC (0.746 mg, 3.62 mmol) in DMF (12 mL) was stirred at room temperature for 2 h. Then, the mixture was cooled to 0 °C and homopiperazine (0.725 g, 7.23 mmol) was

added. The reaction mixture was stirred overnight at room temperature. The precipitate was filtered off and washed with CH₂Cl₂. The filtrate was evaporated and the residue was purified by silica gel flash chromatography (the mobile phase being CH₂Cl₂/MeOH/NH_{3(aq)} 92:8:0.3), to yield the title compound as a light-yellow oil (0.506 g, 46%). ¹H NMR (300 MHz, CDCl₃): δ = 7.10–7.02 (m, 1H), 6.98–6.87 (m, 3H), 4.49 (s, 2H, O-CH₂-CO), 4.19 (t, 2H, *J* = 7.7 Hz, N-CH₂-CH₂-CO), 3.60–3.50 (m, 2H), 3.48–3.35 (m, 2H), 2.88–2.82 (m, 2H), 2.79–2.71 (m, 2H), 2.69–2.56 (m, 2H), 1.85 (s, 1H, NH), 1.75–1.65 ppm (m, 2H); ¹³C NMR (75 MHz, CDCl₃): δ = 169.7, 164.3, 145.1, 128.2, 123.9, 123.0, 117.0, 114.8, 67.5, 50.7, 48.6, 44.5, 38.0, 30.7, 30.4, 29.4 ppm; HRMS (ESI): *m/z* [M + H]⁺ calculated for C₁₆H₂₁N₃O₃ 304.16555, found 304.1655.

4-(3-(3-Oxo-2,3-dihydro-4H-benzo[b][1,4]oxazin-4-yl)propanoyl)-N-phenyl-1,4-diazepane-1-carboxamide (41): This compound was synthesized from compound **40**, according to the procedure for the synthesis of compound **22a**. The crude residue was purified by silica gel flash chromatography (using EtOAc as mobile phase), to yield the title compound as a white powder (93% yield). Purity (Method A): 100%. ¹H NMR (300 MHz, CDCl₃): δ = 7.40–7.34 (m, 4H), 7.10–6.95 (m, 5H), 6.72–6.69 (m, 1H), 4.57 (s, 2H, O-CH₂-CO), 4.29–4.15 (m, 2H, N-CH₂-CH₂-CO), 3.77–3.68 (m, 1H), 3.65–3.53 (m, 5H), 3.50–3.40 (m, 2H), 2.75–2.65 (m, 2H, *J* = 7.2 Hz, N-CH₂-CH₂-CO), 1.98 ppm (m, 2H); HRMS (ESI): *m/z* [M + H]⁺ calculated for C₂₃H₂₆N₄O₄ 423.20266, found 423.2027.

(S)-2-Methyl-4-(3-(3-oxo-2H-benzo[b][1,4]oxazin-4(3H)-yl)propanoyl)-N-phenylpiperazine-1-carboxamide (45a): (*S*)-*tert*-Butyl 3-methylpiperazine-1-carboxylate **42a** (100 mg 0.50 mmol) was dissolved in THF (10 mL) and a solution of phenylisocyanate (60 mg 0.57 mmol) in THF (5 mL) was slowly added at room temperature. A light-yellow solid was formed upon reaction completion. Then, trifluoroacetic acid (10 mL) was added dropwise while stirring at 0 °C. When TLC showed completion of the deprotection, the mixture was diluted with H₂O (50 mL). Then, the mixture was washed with CH₂Cl₂ (30 mL) and the aqueous phase was adjusted to pH 9 by the addition of 1 N NaOH. The mixture was extracted with CH₂Cl₂ (50 mL). The combined organic layers were dried over Na₂SO₄. The solvents were evaporated in vacuo, affording crude **44a**. This solid was used in the next reaction without further purification. The solid was re-dissolved in DMSO (10 mL). Compound **20** (110 mg, 0.50 mmol), HCTU (206 mg 0.50 mmol) and DIPEA (50 μL) were added. The mixture was stirred overnight at room temperature. The resulting mixture was diluted with CH₂Cl₂ (100 mL) and washed with H₂O (40 mL). The combined organic layers were dried over Na₂SO₄ and concentrated in vacuo. The crude residue was purified by silica gel flash column chromatography to yield the title compound (210 mg, 50%). Purity (Method A): 98.77%. ¹H NMR (300 MHz, [D₆]DMSO): δ = 1.05–1.19 (m, 3H, CH₃), 2.67–3.14 (m, 5H, CH₂, CH), 3.65–4.38 (m, 6H, 3CH₂), 4.63 (s, 2H, CH₂), 6.94–7.08 (m, 1H, CH), 7.20–7.25 (m, 3H, 3CH), 7.43–7.48 (m, 3H, 3CH), 8.54 ppm (s, H, NH); ¹³C NMR (75 MHz, [D₆]DMSO): δ = 14.71, 15.24, 29.82, 33.08, 37.33, 47.21, 43.80, 44.81, 45.08, 46.51, 48.94, 67.21, 115.32, 115.45, 116.77, 119.95, 121.97, 123.02, 123.69, 128.41, 140.54, 145.04, 154.91, 161.16, 169.20, 169.30 ppm; HRMS (ESI): *m/z* [M + H]⁺ calculated for C₂₃H₂₇N₄O₄ 423.20321, found 423.2025.

Compounds **45b–f** were synthesized according to the procedure for the preparation of compound **45a**. Exact experimental and spectral data can be found in the Supporting Information.

4-(3-(3-Oxo-2,3-dihydro-4H-pyrido[3,2-*b*][1,4]oxazin-4-yl)propanoyl)-N-phenylpiperazine-1-carboxamide (48): A mixture of 2H-pyrido[3,2-*b*][1,4]oxazin-3(4H)-one **46** (75 mg, 0.5 mmol), K₂CO₃

(158 mg, 1.0 mmol), DMF (3 mL), and ethyl 3-bromopropionate (100 mg, 0.55 mmol) was stirred at 80 °C for 8 h. When the reaction was finished, the mixture was diluted with H₂O (15 mL) and extracted with CH₂Cl₂ (3 × 10 mL). The combined organic phases were washed with H₂O (15 mL) dried over anhydrous Na₂SO₄ and evaporated in vacuo. The crude residue was dissolved in THF (5 mL) and a solution of LiOH (42 mg, 1 mmol) in H₂O (5 mL) was added. The mixture was stirred for 18 h at 60 °C. Then, the solution was adjusted to pH 3 with 2 N HCl and the compound was extracted with EtOAc (3 × 20 mL). The combined organic layers were dried over anhydrous Na₂SO₄. The solvents were evaporated in vacuo, to yield the pure carboxylic acid compound. The acid was re-dissolved in DMF (5 mL), and HOBt (76 mg, 0.5 mmol) and DCC (103 mg, 0.5 mmol) were added at 0 °C. After stirring for 2 h, *N*-phenylpiperazine-1-carboxamide **16** (102 mg, 0.5 mmol) was added. The mixture was stirred at room temperature overnight. The resulting mixture was diluted with CH₂Cl₂ (20 mL) and washed with H₂O (2 × 20 mL). The organic phases were dried over anhydrous Na₂SO₄ and concentrated in vacuo. The crude residue was purified by flash column chromatography (using a mixture of Et₂O and MeOH in a ratio of 10:1 as mobile phase) to afford the title compound (135 mg, 66%). Purity (Method A): 100%. ¹H NMR (300 MHz, CDCl₃): δ = 2.76–2.81 (m, 2H, CH₂), 3.46–3.69 (m, 8H, 4 × CH₂), 4.42–4.47 (m, 2H, CH₂), 4.66 (s, 2H, CH₂), 6.62 (s, 1H, NH), 6.94 (dd, *J* = 7.9, 4.9 Hz, 1H, CH_{Ar}), 7.01–7.07 (m, 1H, CH_{Ar}), 7.23 (dd, *J* = 7.9, 1.5 Hz, 1H, CH_{Ar}), 7.27–7.38 (m, 4H, CH_{Ar}), 8.00 ppm (dd, *J* = 4.9, 1.5 Hz, 1H, CH_{Ar}); ¹³C NMR (75 MHz, CDCl₃): δ = 31.49, 36.41, 41.12, 43.62, 44.19, 45.33, 67.50, 119.54, 120.35, 123.51, 123.66, 129.01, 138.89, 140.74, 141.34, 141.45, 155.18, 164.88, 169.49 ppm; HRMS (ESI): *m/z* [M + H]⁺ calculated for C₂₁H₂₄N₅O₄ 410.18228, found 410.1818.

Compounds **49**, **53–55**, **58–59** were synthesized according to the procedure for the preparation of compound **48**. Exact experimental and spectral data can be found in the Supporting Information.

4-(3-(2-Oxoquinoxalin-1(2H)-yl)propanoyl)-*N*-phenylpiperazine-1-carboxamide (62): A mixture of *tert*-butyl 3-oxo-3,4-dihydroquinoxaline-1(2H)-carboxylate **60** (248 mg, 1 mmol), K₂CO₃ (276 mg, 2.0 mmol) and ethyl 3-bromopropionate (199 mg, 1.1 mmol) in DMF (6 mL) was stirred at 80 °C for 8 h. When the reaction was finished, the mixture was diluted with H₂O (30 mL) and extracted with CH₂Cl₂ (3 × 15 mL). The organic phases were combined and washed with H₂O (30 mL). The combined organic phases were dried over Na₂SO₄ and evaporated in vacuo. The crude residue was re-dissolved in THF (10 mL) and a solution of LiOH (84 mg, 2 mmol) in H₂O (10 mL) was added. The mixture was stirred for 18 h at 60 °C. Then, the solution was adjusted to pH 3 with 2 N HCl and the compound was extracted with EtOAc (3 × 20 mL). The combined organic layers were dried over anhydrous Na₂SO₄ and evaporated in vacuo, to yield the pure acid. The acid was dissolved in DMF (10 mL), and HOBt (153 mg, 1 mmol), DCC (206 mg, 1 mmol) were added at 0 °C. After stirring for 2 h, *N*-phenylpiperazine-1-carboxamide **16** (204 mg, 1 mmol) was added and the mixture was stirred overnight at room temperature. The resulting mixture was diluted with CH₂Cl₂ (30 mL) and washed with H₂O (2 × 20 mL). The combined organic layers were dried over anhydrous Na₂SO₄ and concentrated in vacuo. The crude residue was purified by flash column chromatography (Et₂O/MeOH 10:1) affording compound **61**. *N*-Boc deprotection was carried out by using a 3 mL 80% TFA in 15 mL of CH₂Cl₂. After stirring for 3 h at 0 °C, solvents were evaporated and the crude mixture was dissolved in CH₂Cl₂ (20 mL), washed with a saturated aqueous NaHCO₃ solution. The combined organic layers were dried over anhydrous Na₂SO₄, concentrated and purified by flash column chromatography (EtOAc/MeOH 10:1)

to yield the title compound (102 mg, 25%). Purity (Method A): 98.86%. ¹H NMR (300 MHz, CDCl₃): δ = 2.77–2.85 (m, 2H, CH₂), 3.47–3.50 (m, 4H, 2 × CH₂), 3.51 (s, 2H, CH₂), 3.69–3.73 (m, 2H, CH₂), 4.56–4.61 (m, 2H, CH₂), 6.48 (s, 1H, NH), 7.02–7.08 (m, 1H, CH_{Ar}), 7.26–7.41 (m, 5H, 5 × CH_{Ar}), 7.48–7.51 (m, 1H, CH_{Ar}), 7.59–7.65 (m, 1H, CH_{Ar}), 7.90–7.93 (m, 1H, CH_{Ar}), 8.30 ppm (s, 1H, CH); ¹³C NMR (75 MHz, CDCl₃): δ = 30.72, 38.80, 41.33, 43.73, 44.09, 45.30, 113.83, 120.26, 123.69, 124.15, 129.12, 131.05, 131.54, 132.25, 133.82, 138.69, 150.00, 154.96, 155.07, 168.72 ppm. HRMS (ESI): *m/z* [M + H]⁺ calculated for C₂₂H₂₄N₅O₃ 406.18737, found 406.1864.

4-(3-(2-Oxobenzo[d]oxazol-3(2H)-yl)propanoyl)-*N*-phenylpiperazine-1-carboxamide (64): 3-(2-Oxobenzo[d]oxazol-3(2H)-yl)propanoic acid **63** (104 mg, 0.5 mmol) was dissolved in DMF (5 mL) and HOBt (76 mg, 0.5 mmol) and DCC (103 mg, 0.5 mmol) were added at 0 °C. After stirring for 2 h at 0 °C, *N*-phenylpiperazine-1-carboxamide **16** (102 mg, 0.5 mmol) was added. The mixture was stirred overnight at room temperature. The resulting mixture was diluted with CH₂Cl₂ (20 mL) and washed with H₂O (2 × 20 mL). The combined organic layers were dried over anhydrous Na₂SO₄ and concentrated in vacuo. The crude residue was purified by silica gel flash column chromatography (a mixture of Et₂O and MeOH in a ratio of 10:1 as mobile phase) to afford the title compound (170 mg, 86%). Purity (Method B): 99.29%. ¹H NMR (300 MHz, CDCl₃): δ = 2.94 (t, *J* = 6.8 Hz, 2H, CH₂), 3.48–3.59 (m, 6H, 3 × CH₂), 3.71–3.75 (m, 2H, CH₂), 4.26 (t, *J* = 6.8 Hz, 2H, CH₂), 6.59 (s, 1H, NH), 7.10–7.43 ppm (m, 9H, 9 × CH_{Ar}); ¹³C NMR (75 MHz, CDCl₃): δ = 31.10, 38.66, 41.33, 43.63, 44.00, 45.19, 109.20, 110.19, 120.28, 122.67, 123.65, 124.14, 129.09, 131.19, 138.73, 142.78, 154.96, 168.74 ppm.

4-(3-(2,3-Dihydro-4H-benzo[b][1,4]oxazin-4-yl)propanoyl)-*N*-phenylpiperazine-1-carboxamide (66): A mixture of 3,4-dihydro-2H-benzo[b][1,4]oxazine **65** (67 mg, 0.5 mmol), DIPEA (0.24 mL, 1.4 mmol) and ethyl 3-bromopropionate (380 mg, 2.4 mmol) in DMF (3 mL) was stirred at 80 °C for 48 h. When the reaction was finished, the mixture was diluted with H₂O (15 mL) and extracted with CH₂Cl₂ (3 × 10 mL). The organic phases were combined and washed with H₂O (15 mL). The combined organic layers were dried over anhydrous Na₂SO₄ and evaporated in vacuo. The crude residue was dissolved in THF (5 mL) and a solution of LiOH (42 mg, 1 mmol) in H₂O (5 mL) was added. The mixture was stirred for 18 h at 60 °C. Then, the solution was adjusted to pH 3 with 2 N HCl and the compound was extracted with EtOAc (3 × 20 mL). The combined organic layers were dried over anhydrous Na₂SO₄. After solvent evaporation, the pure acid compound was obtained. The acid was re-dissolved in DMF (5 mL), and HOBt (76 mg, 0.5 mmol) and DCC (103 mg, 0.5 mmol) were added at 0 °C. After stirring for 2 h, *N*-phenylpiperazine-1-carboxamide **16** (102 mg, 0.5 mmol) was added. The mixture was stirred overnight at room temperature. The resulting mixture was diluted with CH₂Cl₂ (20 mL) and washed with H₂O (2 × 20 mL). The organic phase was dried over Na₂SO₄ and concentrated in vacuo. The crude residue was purified by silica gel flash column chromatography (using a mixture of Et₂O and MeOH in a ratio of 10:1 as mobile phase) to afford the title compound (96 mg, 49%). Purity (Method A): 95.43%. ¹H NMR (300 MHz, CDCl₃): δ = 2.58–2.63 (m, 2H, CH₂), 3.37–3.50 (m, 8H, 4 × CH₂), 3.65–3.70 (m, 4H, 2 × CH₂), 4.20–4.23 (m, 2H, CH₂), 6.59 (s, 1H, NH), 6.60–6.67 (m, 2H, 2 × CH_{Ar}), 6.77–6.86 (m, 2H, 2 × CH_{Ar}), 7.02–7.07 (m, 1H, CH_{Ar}), 7.24–7.34 ppm (m, 4H, 4 × CH_{Ar}); ¹³C NMR (75 MHz, CDCl₃): δ = 29.39, 41.20, 43.55, 44.10, 45.27, 47.04, 47.67, 64.60, 111.81, 116.67, 117.82, 120.38, 121.83, 123.63, 129.04, 134.49, 138.74, 144.27, 155.03, 170.43 ppm.

Biological assays

Mycobacterial ThyX NADPH oxidase assay: *M. tuberculosis* ThyX was produced and purified as described.^[18] The synthesized compounds were screened using a NADPH oxidation assay for *M. tuberculosis* ThyX activity in 96-well plates.^[27] To determine IC₅₀ values of compounds **22g**, **27**, **28**, **29d**, **45c**, **48** and **52**, the reaction mixture (100 μ L) contained 50 mM HEPES pH7.5, 1 mM MgCl₂, 12 μ M FAD, 0.5% glycerol, 0.05% Triton X-100, 0.5 mg mL⁻¹ BSA, 3 μ M dUMP, 2.5% DMSO, a range of concentrations (ranging from 40 nM to 20 μ M) of each testing compound and 0.5 μ M of purified *MtbThyX*. After incubation for 15 min, the reactions were initiated by adding 600 μ M NADPH and the decrease in absorbance at 340 nm was detected using a CLARIOstar microplate reader (BMG Labtech) after 1 h. The experiment was done in duplicates and repeated 3–4 times. In the control reactions, no enzyme or no inhibitor was used. Reported IC₅₀ values were determined by normalizing the absorbance values to the control points and fitting the data to nonlinear regression (dose–response inhibition) using GraphPad Prism software (version 7).

To understand the detailed kinetic mechanisms involved in the interaction of the molecules with *MtbThyX*, the NADPH oxidase assay of the ThyX enzyme was measured under various combinations of concentrations for compounds **22a**, **22c**, **22e**, **22i** and **22k** (1–100 μ M) and substrate concentrations. All molecules were solubilized in dimethyl sulfoxide (DMSO) and used at a 1% final concentration of DMSO during the test. Standard reaction mixture (100 μ L) contained HEPES 50 mM pH 8, NaCl 30 mM, FAD 50 μ M, β -mercaptoethanol 100 μ M, dUMP 100 μ M, NADPH 250 μ M, and 10 μ M of purified *MtbThyX*. According to the experiments, dUMP, NADPH and CH₂H₄folate were varied across concentration ranges of 5–100 μ M, 50–500 μ M, and 10–125 μ M, respectively. Microtiter plates were prepared and transferred to the microplate reader Chameleon II (Hidex). The reactions were started by automatically injecting NADPH into individual wells, and ThyX activity was determined by following a decrease in absorbance at 340 nm. Samples with added DMSO and enzyme-free reactions were used as positive and negative controls, respectively. The inhibition data were globally fitted to all possible kinetic models including competitive, noncompetitive, and uncompetitive inhibition.

NADPH oxidase assay with ThyX from other species: The same set of molecules was also evaluated as potential inhibitors of ThyX activity from other species (*H. pylori*, *PBcv-1*, *B. hermsi*, and *C. trachomatis*) using the NADPH oxidase assay under standard reaction conditions with test compounds at 50 or 200 μ M in the reaction assay. ThyX activity was also measured by a deprotonation of [5-³H-dUMP] test, with measurement of the tritium release during the reaction.^[14] Molecules were used in 1% DMSO final at the concentrations described above, and reactions were started by the addition of 10 μ M ThyX and stopped after 25 min incubation at 37 °C.

Molecular modeling

Prior to energy minimization, all complexes were neutralized by adding Na⁺ and Cl⁻ ions together with TIP3P water solvation. After minimization, the system was gradually heated from 0 to 300 K over 50 ps, reaching solvent density after another 500 ps. During the following 20 ns simulation, coordinate trajectories were collected every 2 ps after which CPPTRAJ modules of AMBER 16 were implied for trajectories analyses.^[28] The root-mean-square deviation (RMSD) of protein backbone, compound **10** and folic acid were computed. RMSD convergence was reached after 15 ns for both

complexes. The MMGBSA module in AmberTools17 was used to calculate the Gibbs free energy after 20 ns simulation indicating protein affinity for compound **10** compared with folic acid.

Acknowledgements

The authors are grateful to Research Foundation – Flanders (FWO) for support of the project G.0664.12 on “Flavin-dependent thymidylate synthase: a target for new antibiotics”.

Conflict of interest

The authors declare no conflict of interest.

Keywords: antibiotics · benzo[b][1,4]oxazine · drug discovery · tuberculosis

- [1] *Global Tuberculosis Report 2017*, World Health Organization, https://www.who.int/tb/publications/global_report/gtbr2017_main_text.pdf (last accessed February 8, 2019).
- [2] L. G. Dover, G. D. Coxon, *J. Med. Chem.* **2011**, *54*, 6157–6165.
- [3] K. Andries, P. Verhasselt, J. Guillemont, H. W. Göhlmann, J. M. Neefs, H. Winkler, J. Van Gestel, P. Timmerman, M. Zhu, E. Lee, P. Williams, D. de Chaffoy, E. Huitric, S. Hoffner, E. Cambau, C. Truffot-Pernot, N. Lounis, V. Jarlier, *Science* **2005**, *307*, 223–227.
- [4] M. T. Gler, V. Skripconoka, E. Sanchez-Garavito, H. Xiao, J. L. Cabrera-Rivero, D. E. Vargas-Vasquez, M. Gao, M. Awad, S. K. Park, T. S. Shim, G. Y. Suh, M. Danilovits, H. Ogata, A. Kurve, J. Chang, K. Suzuki, T. Tupasi, W. J. Koh, B. Seaworth, L. J. Geiter, C. D. Wells, *N. Engl. J. Med.* **2012**, *366*, 2151–2160.
- [5] D. T. Hoagland, J. Liu, R. B. Lee, R. E. Lee, *Adv. Drug Deliv. Rev.* **2016**, *102*, 55–72.
- [6] C. W. Carreras, D. V. Santi, *Annu. Rev. Biochem.* **1995**, *64*, 721–762.
- [7] H. Myllykallio, G. Lipowski, D. Leduc, J. Filee, P. Forterre, U. Liebl, *Science* **2002**, *297*, 105–107.
- [8] a) T. V. Mishanina, L. Yu, K. Karunaratne, D. Mondal, J. M. Corcoran, M. A. Choi, A. Kohen, *Science* **2016**, *351*–354, 507–510; b) T. V. Mishanina, E. M. Koehn, J. A. Conrad, B. A. Palfey, S. A. Lesley, A. Kohen, *J. Am. Chem. Soc.* **2012**, *134*, 4442–4448; c) E. M. Koehn, A. Kohen, *Arch. Biochem. Biophys.* **2010**, *493*, 96–102.
- [9] I. Mathews, A. M. Deacon, J. M. Canaves, D. McMullan, S. A. Lesley, S. Agarwalla, P. Kuhn, *Structure* **2003**, *11*, 677–690.
- [10] a) A. S. Fivian-Hughes, J. Houghton, E. O. Davis, *Microbiology* **2012**, *158*, 308–318; b) J. Rengarajan, B. R. Bloom, E. J. Rubin, *Proc. Natl. Acad. Sci. USA* **2005**, *102*, 8327–8332; c) C. M. Sassetti, D. H. Boyd, E. J. Rubin, *Mol. Microbiol.* **2003**, *48*, 77–84.
- [11] M. Choi, K. Karunaratne, A. Kohen, *Molecules* **2016**, *21*, E654.
- [12] M. Kögler, B. Vanderhoydonck, S. De Jonghe, J. Rozenski, K. Van Belle, J. Herman, T. Louat, A. Parchina, C. Sibley, E. Lescrinier, P. Herdewijn, *J. Med. Chem.* **2011**, *54*, 4847–4862.
- [13] T. Basta, Y. Boum, J. Briffotiaux, H. F. Becker, I. Lamarre-Jouenne, J. C. Lambry, S. Skouloubris, U. Liebl, M. Graille, H. van Tilbeurgh, H. Myllykallio, *Open Biol.* **2012**, *2*, 120120.
- [14] S. Skouloubris, K. Djaout, I. Lamarre, J. C. Lambry, K. Anger, J. Briffotiaux, U. Liebl, H. de Reuse, H. Myllykallio, *Open Biol.* **2015**, *5*, 150015.
- [15] J. L. Bolton, M. A. Trush, T. M. Penning, G. Dryhurst, T. J. Monks, *Chem. Res. Toxicol.* **2000**, *13*, 135–160.
- [16] F. Esra Önen, Y. Boum, C. Jacquement, M. V. Spanedda, N. Jaber, D. Scherman, H. Myllykallio, J. Herscovici, *Bioorg. Med. Chem. Lett.* **2008**, *18*, 3628–3631.
- [17] R. Luciani, P. Saxena, S. Surade, M. Santucci, A. Venturelli, C. Borsari, G. Marverti, G. Ponterini, S. Ferrari, T. L. Blundell, M. P. Costi, *J. Med. Chem.* **2016**, *59*, 9269–9275.

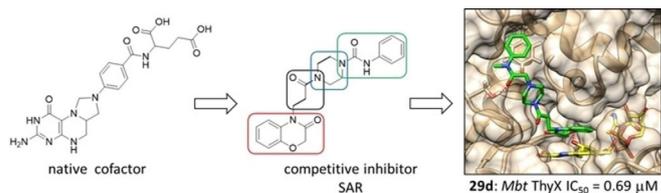
- [18] R. Abu El-Asrar, L. Margamuljana, H. Klaassen, M. Nijs, A. Marchand, P. Chaltin, H. Myllykallio, H. F. Becker, S. De Jonghe, P. Herdewijn, E. Lescrinier, *Biochem. Pharmacol.* **2017**, *135*, 69–78.
- [19] V. Rajachandrashekhara, B. G. Vineela, S. Venkataiaha, P. K. Dubey, *Der Pharma Chemica* **2014**, *6*, 7–10.
- [20] D. S. Johnson, K. Ahn, S. Kesten, S. E. Lazerwith, Y. Song, M. Morris, L. Fay, T. Gregory, C. Stiff, J. B. Dunbar, Jr., M. Liimatta, D. Beidler, S. Smith, T. K. Nomanbhoy, B. J. Cravatt, *Bioorg. Med. Chem. Lett.* **2009**, *19*, 2865–2869.
- [21] F. L. Atkinson, M. D. Barker, S. A. Campos, N. J. Parr, V. K. Patel, (Glaxo Group Ltd., Greenford Middlesex, GB), Int. PCT Pub. No. WO2006129100, **2006**.
- [22] R. E. TenBrink, W. B. Im, V. H. Sethy, A. H. Tang, D. B. Carter, *J. Med. Chem.* **1994**, *37*, 758–768.
- [23] E. M. Koehn, L. L. Perissinotti, S. Moghram, A. Prabhakar, S. A. Lesley, I. I. Mathews, A. Kohen, *Proc. Natl. Acad. Sci. USA* **2012**, *109*, 15722–15727.
- [24] P. Sampathkumar, S. Turley, J. E. Ulmer, H. G. Rhie, C. H. Sibley, W. G. Hol, *J. Mol. Biol.* **2005**, *352*, 1091–1104.
- [25] D. A. Case, D. S. Cerutti, T. E. Cheatham III, T. A. Darden, R. E. Duke, T. J. Giese, H. Gohlke, A. W. Goetz, D. Greene, N. Homeyer, S. Izadi, A. Kovalenko, T. S. Lee, S. LeGrand, P. Li, C. Lin, J. Liu, T. Luchko, R. Luo, D. Mer-
melstein, K. M. Merz, G. Monard, H. Nguyen, I. Omelyan, A. Onufriev, F. Pan, R. Qi, D. R. Roe, A. Roitberg, C. Sagui, C. L. Simmerling, W. M. Botello-Smith, J. Swails, R. C. Walker, J. Wang, R. M. Wolf, X. Wu, L. Xiao, D. M. York, P. A. Kollman, *AMBER* **2017**, University of California, San Francisco.
- [26] A. Marchler-Bauer, Y. Bo, L. Han, J. He, C. J. Lanczycki, S. Lu, F. Chitsaz, M. K. Derbyshire, R. C. Geer, N. R. Gonzales, M. Gwadz, D. I. Hurwitz, F. Lu, G. H. Marchler, J. S. Song, N. Thanki, Z. Wang, R. A. Yamashita, D. Zhang, C. Zheng, L. Y. Geer, S. H. Bryant, *Nucleic Acids Res.* **2017**, *45*, D200–D203.
- [27] K. Djaout, V. Singh, Y. Boum, V. Katawera, H. F. Becker, N. G. Bush, S. J. Hearnshaw, J. E. Pritchard, P. Bourbon, P. B. Madrid, A. Maxwell, V. Mizrahi, H. Myllykallio, S. Ekins, *Sci. Rep.* **2016**, *6*, 27792.
- [28] D. R. Roe, T. E. Cheatham III, *J. Chem. Theory Comput.* **2013**, *9*, 3084–3095.

Manuscript received: November 21, 2018

Revised manuscript received: January 21, 2019

Accepted manuscript online: January 31, 2019

Version of record online: ■ ■ ■ 0000



Optimization campaign: Starting from a previously identified mycobacterial thymidylate synthase X (ThyX) inhibitor based on a benzo[*b*][1,4]oxazin-3(4*H*)-one scaffold, a systematic structure–activity relationship study was performed.

This led to the discovery of a benzo[*b*][1,4]oxazin-3(4*H*)-one analogue with an IC_{50} value of $0.69 \mu\text{M}$. These heterocycles can be used as starting points for the discovery of novel antibacterial agents that act via ThyX inhibition.

J. Modranka, J. Li, A. Parchina, M. Vanmeert, S. Dumbre, M. Salman, H. Myllykallio, H. F. Becker, R. Vanhoutte, L. Margamuljana, H. Nguyen, R. Abu El-Asrar, J. Rozenski, P. Herdewijn, S. De Jonghe, E. Lescrier**



Synthesis and Structure–Activity Relationship Studies of Benzo[*b*][1,4]oxazin-3(4*H*)-one Analogues as Inhibitors of Mycobacterial Thymidylate Synthase X



Titre : Détection et signalisation du monoxyde de carbone chez les bactéries aérobies : Hème-senseur RcoM-2 et réponses mycobactériennes au CO.

Mots clés : Monoxyde de carbone (CO), hème-senseur à CO RcoM-2 *Mycobacterium tuberculosis (Mtb)*, gène *cor (rv1829)*, interaction hème-CO, régulation transcriptionnelle.

Résumé : Le gaz toxique CO peut agir à faibles quantités comme molécule de signalisation ; détectée par les protéines senseurs à base d'hème.

La première partie de ce travail concerne le régulateur de transcription bactérien CO-dépendant RcoM-2, qui possède une affinité extrêmement élevée pour le CO, tout en étant insensible à l'O₂. RcoM-2 se lie à son ADN cible seulement lorsque du CO est lié à l'hème. Nous avons caractérisé l'interaction hème-CO dans la protéine entière RcoM-2 et l'avons comparée à l'hémo-domaine isolé RcoMH-2. RcoM-2 peut lier du CO avec une affinité *effective* plus faible que RcoMH-2. Le CO se dissocie avec un taux 20 fois plus élevé par rapport à RcoMH-2, où la liaison du CO est quasi irréversible. Une petite fraction du CO peut échapper de la protéine, ainsi permettant RcoM-2 à agir comme senseur du CO. La présence du domaine de liaison à l'ADN influence les propriétés de fixation du CO à l'hème. Les études sur l'origine moléculaire précise des propriétés dynamiques doivent attendre la structure 3D de RcoM-2.

La détection du CO est cruciale pour *Mycobacterium tuberculosis*, l'agent infectieux de la tuberculose, qui doit répondre aux mécanismes de défense de l'hôte, dont le CO. Le gène *cor (rv1829)* pourrait y jouer un rôle. Nous avons démontré que Cor est un dimère très stable, capable de lier un hème de façon stoechiométrique, suggérant ainsi une fonction potentielle de senseur *direct* du CO. Un résidu histidine a été identifié comme ligand potentiel de l'hème. La dynamique interne du CO est très similaire de celle d'autres senseurs à CO bactériens. Cor présente une activité de liaison à l'ADN dépendante de l'hème et du CO, qui est abolie dans le mutant H70A. Nos études ont également montré que le régulateur de transcription mycobactérien Rv0081, induit en réponse aux changements gazeux, peut se lier sur la région régulatrice prédite de *cor*. La création d'une souche Δcor dans le modèle non-pathogène *M. smegmatis* se poursuit et constituera un premier pas vers des futures analyses transcriptomiques.

Title: Detection and signaling of carbon monoxide in aerobic bacteria: Heme-based sensor RcoM-2 and mycobacterial CO responses.

Keywords: Carbon monoxide (CO), CO sensor RcoM-2, *Mycobacterium tuberculosis (Mtb)*, gene *cor (rv1829)*, heme-CO interaction, transcriptional regulation.

Abstract: The toxic gas CO can act in low quantities as signaling molecule; detected by heme-based sensor proteins.

The CO-dependent transcription factor RcoM-2 has a very high affinity for CO, while being insensitive to O₂. RcoM-2 binds to DNA only when CO is bound to heme. We characterized the heme-CO interaction in full-length RcoM-2 and compared it with the isolated heme domain RcoMH-2. RcoM-2 can bind CO with lower *effective* affinity than RcoMH-2. CO dissociates with a 20-fold higher rate than in RcoMH-2, where CO binding is almost irreversible. A small fraction of CO can escape from the protein, thus allowing RcoM-2 to act as CO sensor. The presence of the DNA binding domain influences the binding properties of CO to heme. Identification of the precise molecular origin of the dynamic properties must await the 3D structure RcoM-2.

CO detection is crucial for *Mtb*, the infectious agent of tuberculosis that must overcome the host's defense mechanisms, including CO. The gene *cor (rv1829)* has been implicated in these processes. We have shown that Cor is a highly stable dimer that is able to stoichiometrically bind a heme cofactor, suggesting a potential function as *direct* CO sensor. A histidine residue was identified as potential heme ligand. The internal CO dynamics is very similar to other bacterial CO sensors. Cor exhibits DNA binding activity that depends on the presence of heme and CO, which is abolished in the H70A mutant. Our studies also showed that the transcriptional regulator Rv0081, induced in response to gas changes, can bind to the predicted regulatory region of *cor*. The creation of a Δcor strain in the non-pathogenic model *M. smegmatis* continues and will be a first step towards transcriptomic analyzes.