

ÉCOLE DOCTORALE DES SCIENCES DE LA VIE ET DE LA SANTÉ
UMR 7156 Génétique Moléculaire, Génomique, Microbiologie

THÈSE présentée par :

Louis-François MEY

soutenue le : **18 décembre 2023**

pour obtenir le grade de : **Docteur de l'Université de Strasbourg**
Discipline/ Spécialité : Sciences de la Vie et de la Santé/ Aspects
moléculaires et cellulaires de la biologie

**Adaptations à la déshalogénation
et à l'utilisation du chlorométhane par la voie
cmu chez *Methylobacterium extorquens***

THÈSE dirigée par :

Mme BRINGEL Françoise, directrice Directrice de Recherche CNRS, Université de Strasbourg
M. VUILLEUMIER Stéphane, co-directeur Professeur des Universités, Université de Strasbourg

RAPPORTEURS :

Mme PRIGENT-COMBARET Claire Directrice de Recherche CNRS, Université de Lyon 1
M. AMATO Pierre Chargé de Recherche CNRS, HDR, Institut de Chimie de Clermont-Ferrand

AUTRES MEMBRES DU JURY :

Mme GEOFFROY Valérie Professeure des Universités, Université de Strasbourg
M. DÖRING Volker Directeur de Recherche, CEA Evry

Remerciements	
Liste des Figures	
Liste des Tableaux	
Liste des abréviations	
Chapitre 1. Introduction	15
1.1. Le chlorométhane.....	15
1.1.1. Un gaz destructeur d’ozone.....	15
1.1.2. Sources et puits du CH ₃ Cl	15
1.2. La méthylotrophie	18
1.2.1. <i>Methylobacterium extorquens</i> , une espèce modèle pour l’étude de la méthylotrophie	18
1.3. Voies d’utilisation du méthanol	21
1.4. Voies d’utilisation du CH ₃ Cl	22
1.4.1. La voie <i>cmu</i> d’utilisation du CH ₃ Cl comme source de carbone et d’énergie....	22
1.4.2. Voies <i>cmu</i> -indépendantes de dégradation du CH ₃ Cl	24
1.5. Le transfert horizontal de gènes, un moteur de l’évolution bactérienne.....	27
1.5.1. Principaux mécanismes de transfert horizontal	27
1.5.2. Les limitations de l’acquisition de gènes par HGT	30
1.5.3. Indices pour un transfert horizontal des gènes <i>cmu</i>	31
1.6. Limites à la croissance avec le CH ₃ Cl après acquisition des gènes de la voie <i>cmu</i> par transfert horizontal	34
1.6.1. Les stress de la croissance avec des composés chlorés	34
1.7. L’évolution expérimentale pour identifier les verrous limitant une croissance bactérienne.....	38
1.7.1. Trois types d’évolution expérimentale	39
1.7.2. L’évolution expérimentale chez <i>M. extorquens</i>	41
1.8. Objectifs de la thèse	44
Chapitre 2. Adaptations de souches méthylotrophes à la croissance avec le CH ₃ Cl après acquisition des gènes de la voie <i>cmu</i>	47
2.1. Stratégie pour la mise en place de l’évolution expérimentale	47
2.1.1. Choix du plasmide pJM105 porteur de 9 gènes de la voie <i>cmu</i>	47
2.1.2. Les souches parentales choisies pour l’évolution expérimentale	48
2.1.3. Une évolution expérimentale en deux temps	49
2.1.4. Les contrôles mis en place pour l’évolution expérimentale.....	51
2.1.5. Stratégie pour l’identification des mutations bénéfiques à la croissance avec le CH ₃ Cl	51

2.2.	Des souches inadaptées à la croissance avec le CH ₃ Cl ont acquis la capacité de croître avec ce composé	54
2.2.1.	Les gènes de la voie <i>cmu</i> d' <i>Hyphomicrobium</i> présents sur le plasmide pJM105 sont exprimés en contexte <i>M. extorquens</i>	54
2.2.2.	Profil au fil des transferts des cultures indépendantes vers la capacité de croître avec le CH ₃ Cl.....	54
2.2.3.	Diversité des souches isolées après évolution expérimentale.....	59
2.2.4.	L'intégrité du plasmide pJM105 est affectée chez les souches évoluées en contexte A0 seulement.....	63
2.2.5.	Séquençage du génome des souches évoluées et analyse de la couverture relative des réplicons	65
2.2.6.	Un Southern Blot qui montre l'intégration de pJM105 dans le génome des souches	68
2.2.7.	Identification des sites de coïntégration du plasmide pJM105 dans le plasmide p2META chez les souches évoluées issues de 105A.....	71
2.2.8.	Un site d'interruption de <i>trfA</i> différent entre les souches évoluées dérivant d'AM1	74
2.3.	Répertoire de l'ensemble des mutations apparues chez les souches évoluées ...	75
2.3.1.	Des mutations nombreuses et diverses	75
2.3.2.	Un nombre limité de loci mutés communs à des cultures évoluées indépendantes	83
2.3.3.	<i>ftfL</i> , le seul locus muté à la fois chez DM4 et AM1.....	84
Chapitre 3.	La perte de fonction de FtfL, une adaptation majeure	87
3.1.	La formyl-tétrahydrofolate ligase, une enzyme centrale pour le métabolisme en C1	87
3.2.	Elucidation du rôle des mutations du gène <i>ftfL</i> dans l'adaptation à la croissance avec le CH ₃ Cl	88
3.2.1.	Etude en contexte évolué de l'effet des mutations dans <i>ftfL</i> sur l'enzyme et les capacités de croissance avec des composés en C1	88
3.2.2.	Recherche de corrélation entre les mutations dans <i>ftfL</i> et la croissance avec le CH ₃ Cl	89
3.3.	Analyse détaillée de l'effet de la perte d'activité de FtfL chez les souches évoluées.....	90
3.3.1.	Les mutations dans <i>ftfL</i> changent les capacités de croissance méthylotrophe des souches évoluées	90
3.3.2.	L'activité formyl-tétrahydrofolate ligase n'est pas détectée chez les souches évoluées portant un gène <i>ftfL</i> muté	92

3.3.3. Localisation des variations de séquence dans la structure de FtfL et leur impact potentiel.....	95
3.4. Les mutations dans <i>ftfL</i> pourraient être liées à l'évolution de la croissance avec le CH ₃ Cl.....	97
3.4.1. Les mutations dans <i>ftfL</i> se sont produites en début d'expérience d'évolution expérimentale de la culture D7	97
3.4.2. Un gène <i>ftfL</i> muté corrèle avec une croissance plus rapide avec le CH ₃ Cl	99
3.5. Construction ciblée de mutants du gène <i>ftfL</i> pour démontrer un lien de cause à effet par rapport à la capacité de croissance avec le CH ₃ Cl	100
3.6. Hypothèse pour expliquer l'effet bénéfique d'une inactivation de FtfL sur la croissance avec le CH ₃ Cl : la dualité entre FtfL et PurU dans le métabolisme de la voie <i>cmu</i>	104
3.7. PurU, une enzyme essentielle pour la croissance avec le CH ₃ Cl chez <i>M. extorquens</i>	105
3.8. Faible occurrence du gène <i>purU</i> dans le génome de souches isolées de l'environnement et porteuses des gènes <i>cmu</i>	108
3.9. Conclusions.....	112
Chapitre 4. Des mutations touchant la voie <i>cmu</i> chez les souches évoluées dérivées de la souche AM1.....	113
4.1. Des variations ponctuelles d'acides aminés dans le domaine N-terminal de CmuA	113
4.2. Le bénéfice de mutations dans CmuA est mineur par rapport à celui conféré par l'inactivation de <i>ftfL</i> pour la croissance avec le CH ₃ Cl.....	115
4.3. Les changements d'acides aminés dans CmuA chez les souches évoluées AM1 ont pris place à des positions conservées de la séquence protéique	115
4.4. Détermination de l'effet des modifications génétiques plasmidiques sur l'activité CH ₃ Cl-déshalogénase.....	117
4.4.1. Choix des conditions de culture et de dosage des activités de déshalogénation	118
4.4.2. Une activité de déshalogénation similaire des souches évoluées mutantes pour CmuA et des souches curées correspondantes exprimant un gène <i>cmuA</i> de type sauvage	120
4.5. Conclusions.....	123
Chapitre 5. Conclusions et Perspectives	125
5.1. L'inactivation de FtfL suite à l'acquisition des gènes de la voie <i>cmu</i> est-elle une réponse adaptative utile dans l'environnement ?	126
5.2. La modification des conditions de l'évolution expérimentale permettrait-elle la découverte d'autres mécanismes d'adaptation à la croissance avec le CH ₃ Cl ?	127

5.3.	Une augmentation de l'expression des gènes <i>nuo</i> est-elle bénéfique à la croissance avec le CH ₃ Cl ?	128
5.4.	Analyse des mutations FtlL-indépendantes partagées en contexte génétique D0	130
	Par ailleurs, une délétion chromosomique de gènes de régulation est apparue au cours de l'évolution expérimentale en contexte <i>M. extorquens</i> DM4	131
5.5.	Comment les mutations du plasmide pJM105 impactent-elles la voie <i>cmu</i> d'utilisation du CH ₃ Cl chez les souches évoluées en contexte <i>M. extorquens</i> AM1 ?	132
5.5.1.	L'incompatibilité des plasmides pJM105 et p2META explique-t-elle la formation systématique de coïntégrat sans autre bénéfice que la conservation des gènes <i>cmu</i> ?	133
5.5.2.	Quel est le bénéfice de mutations plasmidiques affectant les gènes <i>cmu</i> ?	134
5.6.	Un nouvel outil pour l'étude génétique de la croissance avec le CH ₃ Cl : une opportunité à saisir pour la recherche du régulateur de la voie <i>cmu</i>	135
Chapitre 6. Matériel et Méthodes.....		139
6.1.	Matériel biologique	139
6.1.1.	<i>Souches bactériennes</i>	139
6.1.2.	<i>Plasmides</i>	141
6.1.3.	<i>Amorces</i>	142
6.2.	Cultures bactériennes	145
6.2.1.	<i>Milieus de cultures</i>	145
6.2.2.	<i>Sources de carbone</i>	145
6.2.3.	<i>Antibiotiques</i>	146
6.2.4.	<i>Suivi de cultures bactériennes</i>	147
6.2.5.	<i>Estimation de la biomasse par mesure de la densité optique à 600nm</i>	147
6.2.6.	<i>Calcul du temps de génération</i>	148
6.2.7.	<i>Test de croissance avec le formiate</i>	149
6.2.8.	<i>Vérification de l'essentialité de gènes compris dans le cluster <i>cmu</i></i>	149
6.3.	Méthodes de biologie moléculaire.....	149
6.3.1.	<i>Extraction d'ADN génomique</i>	149
6.3.2.	<i>Extraction d'ADN plasmidique</i>	150
6.3.3.	<i>Dosage d'ADN</i>	150
6.3.4.	<i>Electrophorèse d'ADN</i>	150
6.3.5.	<i>Amplification d'ADN par Polymerase chain reaction (PCR)</i>	151
6.3.6.	<i>Hybridation ADN/ADN par la méthode dite du Southern Blot</i>	152
6.3.7.	<i>Mutagenèse dirigée par échange allélique</i>	156

6.3.8.	<i>Séquençage Sanger de produits PCR</i>	159
6.3.9.	<i>Séquençage du génome à partir de l'ADN extrait par Illumina</i>	159
6.3.10.	<i>Transformation d'ADN plasmide chez E. coli par choc thermique</i>	160
6.3.11.	<i>Transfert des plasmides d'E. coli vers M. extorquens par conjugaison tri-parentale</i>	160
6.3.12.	<i>Electrotransformation d'ADN plasmidique chez M. extorquens</i>	161
6.4.	<i>Dosages enzymatiques</i>	162
6.4.1.	<i>Test d'activité enzymatique de FtfL</i>	162
6.4.2.	<i>Test d'activité de la CH₃Cl deshalogénase</i>	164
6.5.	<i>Évolution expérimentale de bactéries vers la capacité à croître avec le CH₃Cl.</i>	165
6.5.1.	<i>Transferts en série de cultures sur un grand nombre de générations</i>	165
6.5.2.	<i>Isolement de souches évoluées</i>	167
6.5.3.	<i>Évaluation des mutations associées au plasmide pJM105 des souches évoluées</i>	168
6.6.	<i>Méthodes analytiques in silico</i>	170
6.6.1.	<i>Analyse des génomes et recherche des mutations des souches évoluées avec la plateforme MaGe de Microscope</i>	170
6.6.2.	<i>Mise en évidence des grandes délétions au sein du génome des souches évoluées</i>	171
6.6.3.	<i>Recherche des liens paired-end entre le plasmide pJM105 et les réplicons des souches évoluées</i>	171
6.6.4.	<i>Détermination de la couverture relative des lectures de chaque réplicon d'un génome par rapport à celle du chromosome</i>	172
6.6.5.	<i>Prédiction de Structure de FtfL</i>	172
6.6.6.	<i>Tests statistiques sur données de suivis de croissance ou tests de déshalogénation</i>	173
	<i>Bibliographie</i>	
	<i>Communications scientifiques</i>	

Remerciements

Tout d'abord merci beaucoup à mes deux encadrants. Merci à Françoise Bringel, ma directrice de thèse, pour tout son encadrement, son implication et sa disponibilité au cours de ces 4 ans. Merci à Stéphane Vuilleumier, mon co-directeur de thèse, pour sa précieuse contribution à la réflexion et l'élaboration des travaux présentés au cours de ce manuscrit.

Un grand merci à nos deux collaborateurs du Génoscope, Stéphanie Fouteau et David Roche. Leur immense travail et leur expertise en bioinformatique (ainsi que leur disponibilité) ont grandement contribué à l'avancement de cette thèse.

Un grand merci aux membres du jury d'avoir accepté d'évaluer mon travail. Merci à mes deux rapporteurs, Claire Prigent-Combaret et Pierre Amato. Un grand merci également à Valérie Geoffroy d'avoir accepté d'être l'examinatrice interne. Enfin, merci à Volker Döring d'avoir également accepté d'être membre de mon jury et d'avoir fait partie de mon comité de suivi de thèse aux côtés de Purificación López-García.

Je remercie toute l'équipe Adaptations et Interactions des Micro-organismes dans l'Environnement pour leur aide au cours de ces 4 ans. Emilie, Thierry, Catherine et les docteurs en devenir Carmen, Adrien, Lucile et Enrico. Merci pour la super ambiance dans le bureau, quand je le fréquentais encore !

Merci à toute l'unité de recherche UMR7156 GMGM de m'avoir accueilli.

Merci beaucoup à Domitille, ancienne post-doc, qui m'a aussi beaucoup aidé pour la partie bioinformatique de mon travail (sans qui le processus aurait été trèèèèèè long et ardu). Surtout merci beaucoup de m'avoir apporté un soutien moral qui a vraiment su faire la différence. Quelques mots encourageants ça peut vous ragaillardir un homme !

En parlant de soutien moral, L'autre tête de cet aigle bicéphale c'est évidemment mon grand Elie ! Merci de ton oreille attentive sur mes problèmes, mes réflexions sur la vie, l'environnement, la géopolitique ! De façon plus générale, merci d'être toujours resté aux nouvelles sur mon état mental et physique.

Merci aussi à ceux qui sont partis mais qui soit m'ont épaulé dans mon travail, soit ont contribué à la bonne ambiance du bureau. Merci à Eugénie, Yousra (et son grain de folie), Abhijit (pour la découverte de la culture indienne), Christelle, Maria, Victoria, Henintsoa, Wen et les nombreux stagiaires qui ont aussi été de passage, en particulier Klarida et Zoé qui m'ont permis de m'essayer à l'encadrement moi-même.

Merci enfin à ma famille pour son soutien. Mon père, ma mère, mon grand-père, ma grand-mère ainsi que ma sœur, Emmanuelle, bien occupée avec le petit Jacob né l'année dernière.

En route pour de nouvelles aventures !

Liste des Figures

Figure 1.1 Sources et puits du CH ₃ Cl	16
Figure 1.2 Les réplicons des génomes de souches de <i>M. extorquens</i> AM1, PA1, DM4 et CM4	20
Figure 1.3 Gènes et enzymes de la voie <i>cmu</i> du métabolisme méthylotrophe de <i>M. extorquens</i> CM4	23
Figure 1.4 Le transfert horizontal chez les procaryotes.....	28
Figure 1.5 Organisation des gènes <i>cmu</i> chez plusieurs souches bactériennes	32
Figure 1.6 Le plasmide pCMU01 d'utilisation du CH ₃ Cl chez <i>M. extorquens</i> CM4	33
Figure 1.7 Les différents types d'évolution expérimentale	40
Figure 2.1 La stratégie d'évolution expérimentale	50
Figure 2.2 Suivi de la densité optique au cours des transferts de l'expérience d'évolution expérimentale	57
Figure 2.3. Corrélation entre la concentration en ions chlorure et la DO à 600 nm au cours de l'expérience d'évolution expérimentale	58
Figure 2.4 Schéma récapitulatif de l'origine des souches évoluées séquencées	60
Figure 2.5 Mesure de la DO ₆₀₀ et de la production d'ions Cl ⁻ après croissance des souches évoluées isolées avec du CH ₃ Cl	61
Figure 2.6 Southern Blot mettant en évidence l'intégration du plasmide pJM105 dans le génome des souches évoluées dérivant de 105A.....	69
Figure 2.7 Modèle du cointégrat entre pJM105 et p2META et données de PCR le démontrant	73
Figure 2.8 La différence d'insertion d'ISMex25 au niveau de <i>trfA</i> chez les souches évoluées issues d'AM1.....	74
Figure 2.9 Région génomique du gène <i>ftfL</i> et localisation des mutations dans les souches évoluées	85
Figure 3.1 La réaction enzymatique catalysée par la formyl-tétrahydrofolate ligase (FtfL)....	88
Figure 3.2 Optimisation du dosage d'activité de FtfL par utilisation de trois concentrations de tétrahydrofolate	92
Figure 3.3 Prédiction de la structure du tétramère de FtfL de DM4 par Alphafold avec les variations ponctuelles retrouvées dans les souches évoluées.	96
Figure 3.4 Comparaison des capacités de croissance des souches évoluées et parentales avec le CH ₃ Cl, le succinate et le méthanol selon la présence de l'allèle sauvage ou muté du gène <i>ftfL</i>	99
Figure 3.5 La culture avec le CH ₃ Cl des mutants <i>ftfL</i> construits montre l'effet de l'inactivation de FtfL sur la croissance avec ce composé	103
Figure 3.6 L'hypothèse du rôle de FtfL dans l'adaptation à la croissance avec le CH ₃ Cl	105

Figure 3.7 PurN et PurT peuvent toutes deux catalyser la formylation du glycineamine-ribonucléotide (GAR) en formylglycineamine-ribonucléotide (fGAR) durant la synthèse des purines.....	109
Figure 4.1 Les souches évoluées AM1 présentent toutes un changement d'acide aminé dans le domaine N-terminal de la protéine CmuA	114
Figure 4.2 Alignement des séquences protéiques de CmuA et comparaison des positions présentant un changement d'acide aminé dans les souches évoluées AM1	116
Figure 4.3 Comparaison de l'activité de déshalogénation de MC1 cultivée en présence de CH ₃ Cl avec ou sans succinate	119

Liste des Tableaux

Tableau 1.1 La présence simultanée du gène <i>cmuA</i> et des gènes de la voie de dégradation du CH ₃ Cl de <i>L. methylhalidovorans</i> MB2 chez plusieurs souches	26
Tableau 2.1 Comparaison des capacités de croissance des souches évoluées	62
Tableau 2.2 Souches évoluées curées : phénotype et conditions d'obtention	64
Tableau 2.3 Données du séquençage Illumina paired-end 150 pb des génomes des souches étudiées	66
Tableau 2.4 Couverture relative des réplicons des souches évoluées par rapport au chromosome	67
Tableau 2.5 Inventaire de l'ensemble des mutations retrouvées dans les souches évoluées issues de 105A	76
Tableau 2.6 Inventaire de l'ensemble des mutations retrouvées dans les souches évoluées issues de 105D	77
Tableau 2.7 Loci mutés communs à plusieurs cultures de souches évoluées	83
Tableau 3.1 Comparaison des capacités de croissance des souches évoluées selon la nature des altérations des protéines FtfL	91
Tableau 3.2 Activité de l'enzyme FtfL mesurée chez des souches porteuses du gène <i>ftfL</i> sauvage ou muté	93
Tableau 3.3 Conservation de la capacité de croître avec le méthanol au cours de l'évolution expérimentale du réplicat D7	98
Tableau 3.4 Croissance de transconjugants de souches évoluées en CH ₃ Cl avec ou sans le gène <i>fold</i> ou <i>purU</i>	107
Tableau 3.5 Présence et conservation de la séquence protéique des gènes <i>purU</i> , <i>purN</i> , <i>ftfL</i> chez plusieurs souches possédant les gènes <i>cmu</i> dont le génome est séquencé	110
Tableau 4.1 Activités de déshalogénation des souches évoluées dérivant d'AM1 ainsi que de leurs équivalents curés	121
Tableau 6.1 Inventaire des souches	139
Tableau 6.2 Inventaire des plasmides	141
Tableau 6.3 Inventaire des amorces	142

Liste des abréviations

ϵ : coefficient d'extinction

Ca^{2+} : ion calcium

CH_2Cl_2 : dichlorométhane

CH_3Cl : chlorométhane

Cl^- : chlorure

cmu : « chloromethane utilization »

cycle TCA : cycle des Acides TriCarboxyliques

dcm : « dichloromethane utilization »

DO_{600} : DO à 600 nm

g/mg/ μg /ng : gramme/ milligramme/ microgramme/ nanogramme

GST : glutathion-S transférases

H_4F : tétrahydrofolate

H_4MPT : tétrahydrométhanoptérine

HGT : « Horizontal Gene Transfer »

IS : « Insertion Sequence »

Kan : Kanamycine

kb/ pb : kilobases/ paire de bases

kDa : kilodalton (Dalton = kg/mol)

La^{3+} : ion lanthanide

LB : « Lysogeny Broth » ; milieu riche pour la croissance d'*Escherichia coli*

L/mL : Litre/ millilitre

M : molaire (mole.L^{-1})

MaGe : Magnifying Genomes : Plateforme Microscope d'analyse de génomes

M3Cl0 : milieu « *Methylobacterium* Mineral Medium » dépourvu d'ions Cl^-

min : minute

mol : moles

MW : « Molecular Weight »

ORF: « Open Reading Frame »

PCR : « Polymerase Chain Reaction »

PQQ : PyrroQuinoline Quinone

ppt : « parts per trillion »

PE : « Paired-End »

qPCR : « quantitative PCR »

RT-qPCR : « Real Time quantitative PCR »

SNP : Single Nucleotide Polymorphism

SU : Sous-Unité

sacB : gène codant pour une levansucrase métabolisant le saccharose en produit toxique

Sac^R/Sac^S : résistant au saccharose/ sensible au saccharose

rpm : tours par min

LTEE : « Long Term Evolution Experiment »

Tc : tétracycline

Tc^R/ Tc^S : résistant à la tétracycline/ sensible à la tétracycline

TSS : « Transcription Starting Site »

UV : rayons Ultra-violet

Chapitre 1. Introduction

1.1. Le chlorométhane

1.1.1. Un gaz destructeur d'ozone

Le chlorométhane de formule chimique CH_3Cl , aussi appelé chlorure de méthyle, Fréon 40 ou R40, est un composé organochloré. Les halogènes sont le 17^{ème} groupe du tableau périodique comprenant le fluor, le chlore, le brome, l'iode, l'astate et le ténness. C'est une famille de composés très réactifs dont la couche de valence contient 7 électrons, et qui forment ainsi facilement des anions halogénures.

Le CH_3Cl est le composé organohalogéné le plus abondant dans l'atmosphère terrestre (Keppler et al. 2005). C'est un gaz d'une durée de vie dans l'atmosphère d'environ un an. Sa concentration dans l'air est d'environ 540 ppt (parts per trillion). Il est responsable de 16% de la destruction de la couche d'ozone par des composés chlorés (Carpenter et al. 2014). En effet, sa photolyse libère des radicaux libres qui vont catalyser la destruction de l'ozone en particulier dans la stratosphère, et ainsi diminuer sa capacité à absorber les rayons UVC et UVB nocifs pour le vivant (Bringel et al. 2019). Le CH_3Cl est répertorié par le Centre International de Recherche sur le Cancer (CIRC) dans la classe 3 : composés non classifiables quant à leur carcinogénicité. À haute concentration, le CH_3Cl est cependant toxique en provoquant des lésions des voies respiratoires si inhalé (sous forme de gaz), et des lésions de la peau et par contact oculaire (sous forme liquide) (Tsai 2017). Les organes cibles sont le système nerveux central, le foie, les reins ainsi que le système reproducteur.

1.1.2. Sources et puits du CH_3Cl

Les sources du CH_3Cl sont nombreuses, principalement biotiques mais des sources abiotiques existent aussi (Figure 1.1).

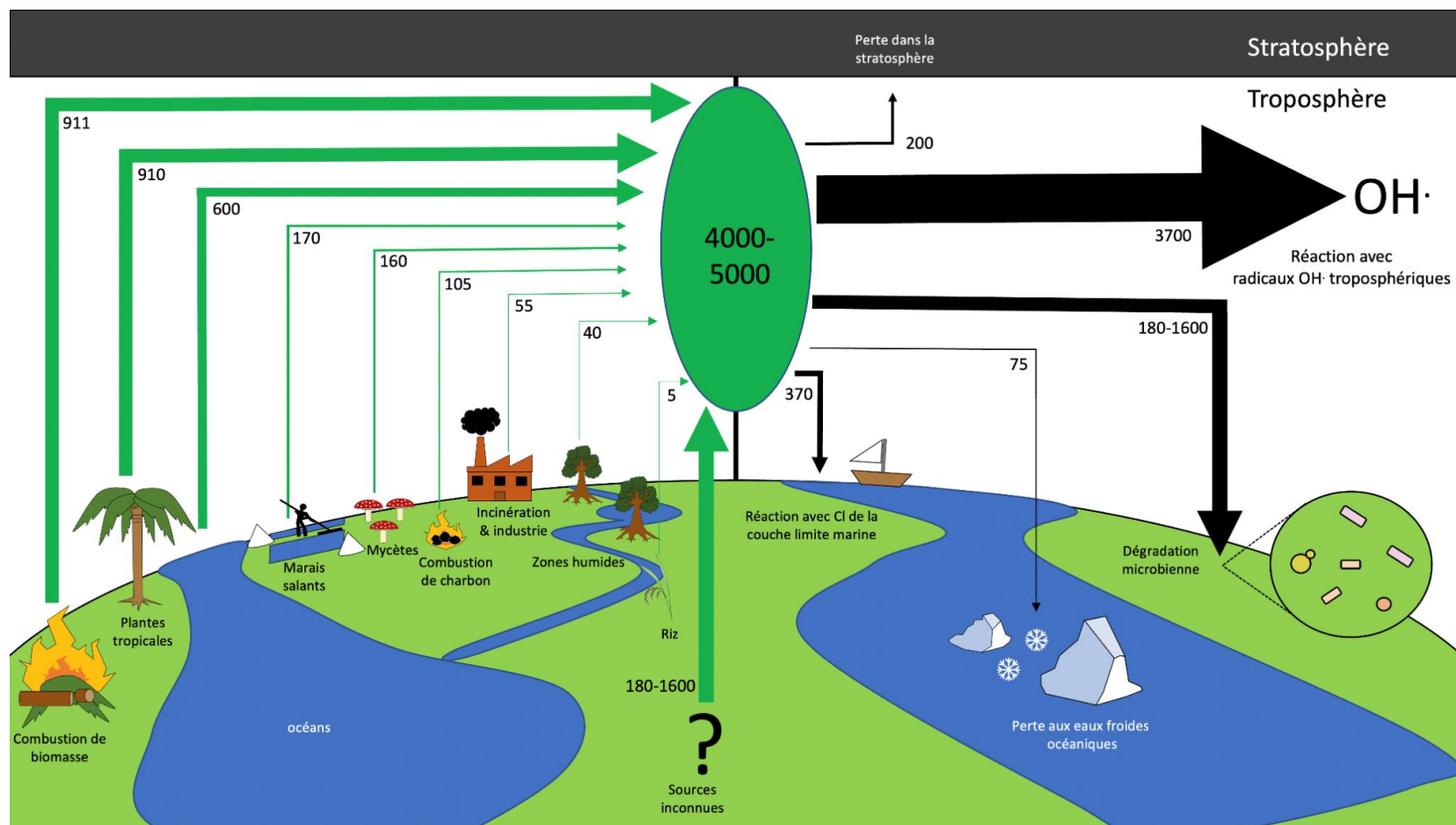


Figure 1.1 Sources et puits du CH_3Cl

Les quantités sont représentées en Gg (10^9 g)/an. On estime la production à environ 4 à 5 Tg (4000 à 5000 Gg) par an (Keppler et al 2005). Les flèches en vert représentent les émissions de CH_3Cl , leur épaisseur variant en fonction de la quantité émise. La bulle au centre représente le bilan total des émissions. Les flèches en noir représentent la quantité dégradée ou absorbée pour chacun des puits.

Ses sources sont principalement d'origine naturelle, comme les plantes tropicales et subtropicales (Jaeger et al. 2018 ; Yokouchi et al. 2002 ; Yokouchi et al. 2007), les mangroves (Manley et al. 2007 ; Kolusu et al. 2018) ou les mycètes dégradant le bois (Harper et al. 1985 ; Harper et al. 1990). Le CH_3Cl est produit par les plantes, notamment par méthylation enzymatique d'un ion chlorure (Attieh et al. 1995 ; Rhew et al. 2003). Le gène *HOL1* (pour « Harmless to Ozone Layer ») codant pour une méthyltransférase est le responsable principal (>80%) de la production de CH_3Cl chez *Arabidopsis thaliana* et est aussi retrouvé dans de nombreuses plantes vasculaires (Rhew et al. 2003). Son rôle dans le fonctionnement de la plante n'a pas été caractérisé pour la production de CH_3Cl . En revanche, il a été montré que lors d'une blessure, l'expression du gène *HOL1* est induite dans la plante, avec formation de méthyl-thiocyanate qui inhibe la croissance de phytopathogènes tels que *Pseudomonas syringae* (Nagatoshi & Nakamura 2009).

Le CH_3Cl d'origine abiotique est principalement issu de la conversion de groupes méthoxyl (donneur de méthyle) de la pectine et de leur réaction avec des ions chlorure (Hamilton et al. 2003 ; Keppler et al. 2004 ; Sailaukhanuly et al. 2014), un procédé favorisé par les hautes températures (Hamilton et al. 2003). L'alkylation d'un ion chlorure durant l'oxydation de matière organique par un accepteur d'électron comme le Fe(III) représente un autre procédé de formation abiotique majeur de CH_3Cl mis en évidence dans les sols et les sédiments (Keppler et al. 2000).

Le CH_3Cl est également produit industriellement de par ses nombreuses applications : en premier lieu dans le secteur de la manufacture de silicone puis, et aussi dans la fabrication de la méthylcellulose, des amines quaternaires, du polyisobutylène et de produits agricoles (Tsai 2017).

Le principal puits du CH_3Cl est dans l'atmosphère, par réaction abiotique du CH_3Cl avec des radicaux $\cdot\text{OH}$ (Carpenter et al. 2014), mais de nombreux autres puits existent (détaillés dans la Figure 1.1, Keppler et al. 2005).

L'ampleur de la part de la dégradation par des micro-organismes dans le budget global du CH_3Cl n'est pas connue. Des bactéries dégradantes ont été isolées d'environnements contrastés tels que les sols, les plantes, et les milieux aquatiques d'eau douce et marins (Bringel et al. 2019). Ce sont des anaérobies strictes (*Acetobacterium dehalogenans* MC qui

utilise le CH₃Cl comme source d'énergie (Traunecker et al. 1991)), des bactéries aérobies facultatives (*Pseudomonas aeruginosa* NB1 (Freedman et al. 2004)), et des bactéries aérobies dites méthylotrophes capables d'utiliser le CH₃Cl comme unique source de carbone et d'énergie (cf. Bringel et al. 2019 pour une revue). Dans le cas particulier de *Methylobacterium album* BG8, une bactérie méthanotrophe, la croissance avec d'autres composés en C1 comme le méthanol est améliorée en présence d'une faible concentration en CH₃Cl (Han & Semrau. 2000).

1.2. La méthylotrophie

La méthylotrophie est définie comme la capacité d'utiliser les composés carbonés réduits et sans liaisons carbone-carbone comme source de carbone et d'énergie (Anthony et al. 1982). Les substrats des micro-organismes méthylotrophes sont notamment des composés à un seul atome de carbone comme le méthane, le méthanol, les amines méthylées, les méthanes halogénés et les sulfures méthylés. Les micro-organismes méthylotrophes jouent un rôle central dans le cycle du carbone, de l'azote et du soufre (Chistoserdova et al. 2011), et peut-être aussi du chlore. Les méthylotrophes facultatifs sont capables d'utiliser autant des substrats dépourvus de liaisons carbone-carbone et des composés multicarbonés, contrairement aux méthylotrophes stricts. Des organismes méthylotrophes ont été détectés dans des habitats divers tels que les sols, les océans (Schaefer et al. 2002), les sédiments (Antony et al. 2012), les nuages (Amato et al. 2007) et la phyllosphère (parties aériennes des plantes, Bringel et Couée 2015).

1.2.1. *Methylobacterium extorquens*, une espèce modèle pour l'étude de la méthylotrophie

Les Alphaprotéobactéries de l'espèce *Methylobacterium extorquens* sont des méthylotrophes facultatifs étudiés de manière intensive pour leurs capacités à dégrader différents composés en C1. Il existe à ce jour des outils génétiques pour plusieurs d'entre elles, comme des vecteurs d'expression (Marx & Lidstrom 2001 ; Marx & Lidstrom 2004 ; Chubiz et al. 2013 ; von Borzyskowski et al. 2015) transmis par conjugaison (Fulton et al. 1984) ou par électroporation (Toyama et al. 1998), des plasmides permettant de faire de la mutagénèse aléatoire par transposon (Marx et al. 2003a ; 2017 ; Muller et al. 2011a/b ; Metzger et al. 2013 ; Ochsner et

al. 2017), de la mutagenèse dirigée par échange allélique (Marx 2008 ; Gruffaz et al. 2014) et plus récemment par le système CRISPR Cas9 (Mo et al. 2020). Il y a également des développements pour la biologie synthétique appliqués à la bioéconomie de composés en C1 (von Borzyskowski et al 2015, von Borzyskowski 2018 ; Carillo et al. 2019).

Historiquement, la souche phare a longtemps été *M. extorquens* AM1. D'abord découverte en 1961 comme contaminant d'un milieu contenant de la méthylamine (Peel & Quayle 1961), elle a été le modèle d'étude de la voie de dégradation du méthanol (Ochsner et al. 2015) qui est conservée chez les autres souches de l'espèce. La souche *M. extorquens* PA1 représente maintenant une alternative à AM1 pour l'étude des métabolismes de composés en C1 comme le méthanol (Ochsner et al. 2015), et ses gènes sont conservés par rapport à AM1 (Marx et al. 2012 ; Nayak & Marx 2014a). La structure génomique chez la souche PA1 est plus simple car elle ne possède aucun plasmide (Marx et al. 2012) contrairement à la souche AM1 (Vuilleumier et al. 2009). Par ailleurs, la souche PA1 a été isolée de la phyllosphère d'*Arabidopsis thaliana* (Knief et al. 2010) et n'a pas la longue histoire de domestication de *M. extorquens* AM1 qui a entraîné sa divergence de la souche originelle (Carroll et al. 2014a). Les souches AM1 et PA1 sont incapables de dégrader des méthanes chlorés.

La souche *M. extorquens* DM4 isolée à partir de boues d'épuration d'eaux usées industrielles (Gälli & Leisinger 1985) est une souche de référence pour la dégradation du dichlorométhane (CH_2Cl_2), un autre composé organochloré (Vuilleumier et al. 2009 ; Chaignaud et al. 2017 ; Biki-Triki et al. 2018 ; Maucourt et al. 2022). L'enzyme DcmA appartenant à la famille des glutathion-S transférases (GST) catalyse la déchloration d'une molécule de CH_2Cl_2 libérant deux protons et deux ions chlorure, avec production de formaldéhyde qui va entrer dans le métabolisme méthylotrophe. Un intermédiaire de réaction très toxique, le S-chlorométhyl-glutathion, est responsable de la génotoxicité associée à la dégradation du CH_2Cl_2 catalysée par DcmA par formation de liaisons covalentes avec l'ADN (adduits à l'ADN) (Kayser et al. 2000 ; Kayser & Vuilleumier 2001).

M. extorquens CM4 est la seule souche du genre *Methylobacterium* connue à ce jour pour être capable d'utiliser le CH_3Cl en tant que substrat de croissance. La souche a été isolée du sol d'une usine pétrochimique en Russie (Doronina et al 1996 ; McDonald et al. 2001). La voie de dégradation méthylotrophe du CH_3Cl la mieux caractérisée y a été étudiée (Vannelli et al. 1998, 1999 ; Studer et al. 2002 ; Roselli et al. 2013). Cette voie métabolique, appelée voie *cmu*, est décrite en détail ci-après (voir paragraphe 1.4.1).

Les chromosomes des souches AM1, DM4 et CM4 sont hautement synténiques (c'est-à-dire que les gènes sont retrouvés dans le même ordre) (Vuilleumier et al. 2009 ; Chaignaud et al. 2017). Les génomes des souches AM1, CM4, DM4 et PA1 partagent une majorité des gènes (80% d'identité d'acides aminés) (Chaignaud et al. 2017). Des 10252 CDS ("coding sequence") constituant le pangénome de ces 4 souches (c'est-à-dire l'ensemble des différents gènes présents dans ces souches), 3927 constituent le core genome (les gènes partagés par l'ensemble des souches) (correspondant à 62% des CDS d'AM1, 69 % pour DM4, 63% pour CM4 et 76% pour PA1). Le génome de la souche AM1 a la particularité de compter un mégaplasmide d'une taille avoisinant 130 kb qui est globalement en synténie avec une région du chromosome de la souche DM4. A l'inverse, les plasmides (Figure 1.2) sont spécifiques des souches. Ils contiennent principalement des gènes dont la fonction est inconnue ou liés à leur machinerie de réplication, avec de rares gènes de fonctions prédites (exemple du plasmide p2META dont les gènes p2META1_0029/p2META1_0030 codent pour un système de résistance au cuivre) (Vuilleumier et al. 2009).

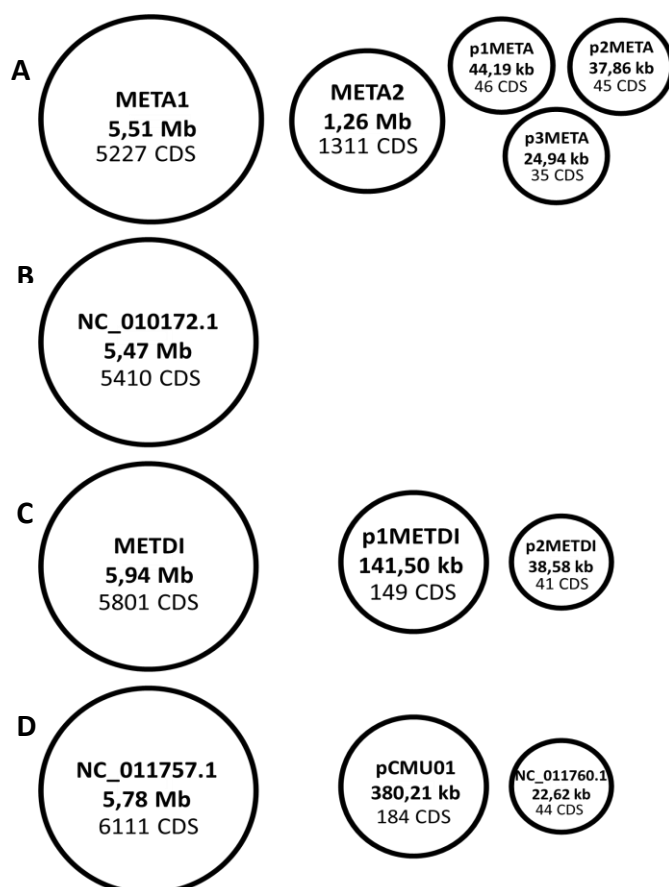


Figure 1.2 Les réplicons des génomes de souches de *M. extorquens* AM1, PA1, DM4 et CM4

A) les cinq réplicons du génome d'AM1 (référence MaGe <https://mage.genoscope.cns.fr>) (contenu GC moyen 68,5%)

B) Le seul chromosome de PA1 (contenu GC moyen 68,2%)

C) Les trois réplicons chez DM4 (contenu GC moyen 68,0%)

D) Les trois réplicons chez CM4 (contenu GC moyen 68,1%).

Les plasmides ne sont pas représentés à l'échelle.

Dans le génome de *M. extorquens* DM4, on trouve un fragment de 5,5 kb flanqué de deux IS et contenant les 4 gènes *dcm* (pour dichlorométhane) *dcmA*, *dcmB*, *dcmC*, *dcmR* et une séquence d'insertion appelée IS1357. Cet îlot *dcm* est niché dans un îlot génomique de 126 kb (Muller et al 2011a, 2011b) dont le pourcentage de GC de 60% diffère de celui du reste du génome de la souche (68%). L'îlot *dcm* portant les gènes de dégradation du CH₂Cl₂ est conservé (identité de séquences des protéines DcmRAB > 80%) chez d'autres souches méthylophiles dégradant le CH₂Cl₂ (Schmid-Appert et al. 1997 ; Vuilleumier et al. 2001 ; Bringel et al. 2017 ; Hayoun et al. 2020).

1.3. Voies d'utilisation du méthanol

Le méthanol est le substrat de référence des études de méthylophilie. Très répandu dans l'environnement (Yurimoto et al. 2021), c'est aussi un intermédiaire dans l'oxydation biologique du méthane (Koo & Rosenzweig 2021). L'oxydation du méthanol est assurée par différentes enzymes chez les bactéries méthylophiles. Chez *M. extorquens*, la méthanol déshydrogénase (MDH) est une enzyme pyrroquinoline quinone (PQQ) dépendante de structure hétéro-tétramérique $\alpha_2\beta_2$ et dont la sous-unité α (contenant un atome de Ca²⁺ et un groupe PQQ) est encodée par *mxoF* et la sous-unité β par *mxoI* (Vuilleumier et al. 2009). Les électrons de cette oxydation par MDH sont transmis à un accepteur unique, le cytochrome CL (*mxoG*) qui le transmet au cytochrome de type I CH, puis à l'oxygène par une cytochrome c oxydase. Plus récemment, une autre méthanol déshydrogénase, XoxF a été découverte qui fonctionne avec un ion de terres rares à la place de l'ion Ca²⁺ (Skovran & Martinez-Gomez 2015). Il a été démontré qu'en présence de La³⁺, il y a un passage de l'expression du cluster *mxo* au cluster *xox* (contenant *xoxF*, *xoxG* codant pour un cytochrome c ainsi que *xoxI* de fonction inconnue mais qui est un homologue à *mxoI* appartenant au cluster *mxo*) (Vu et al. 2016 ; Ochsner et al. 2019).

Le formaldéhyde formé par l'oxydation du méthanol peut soit réagir spontanément avec le tétrahydrofolate (H₄F) pour former du méthylène-H₄F, soit être oxydé en formiate par des enzymes utilisant d'autres cofacteurs : la tétrahydrométhanoptérine (H₄MPT) puis le méthanolurane (MFR) (Ochsner et al. 2015) (Figure 1.3).

1.4. Voies d'utilisation du CH₃Cl

1.4.1. La voie *cmu* d'utilisation du CH₃Cl comme source de carbone et d'énergie

La voie *cmu* a été découverte et caractérisée de façon extensive chez *M. extorquens* CM4 (Vannelli et al. 1998 ; Vannelli et al. 1999 ; Studer et al. 2002). Une mole de CH₃Cl est oxydée en une mole de CO₂ avec production d'une mole d'acide chlorhydrique tout en consommant 1,5 moles d'oxygène. La souche CM4, lorsqu'elle est cultivée avec du CH₃Cl, peut également déshalogéner d'autres méthanes halogénés comme le bromométhane (CH₃Br) et l'iodométhane (CH₃I). Contrairement au CH₃Cl, le CH₃Br ne peut pas être utilisé comme source de carbone pour la croissance par la souche CM4 (Vannelli et al. 1998), probablement à cause de sa toxicité. Quelques bactéries utilisent la voie *cmu* pour la croissance avec d'autres méthanes halogénés en plus du CH₃Cl (McDonald et al. 2005).

Deux méthyltransférases sont essentielles à la première étape de la voie *cmu* et sont encodées respectivement par les gènes *cmuA* et *cmuB* (Vannelli et al. 1999 ; Figure 1.3). Elles ont été initialement identifiées comme deux polypeptides de 67 et 35 kDa induits dans des cultures en CH₃Cl. Des mutants des gènes *cmuAB* obtenus par mutagenèse aléatoire d'un minitransposon de type Tn5 ont perdu la capacité de croître avec le CH₃Cl mais sont toujours capables d'utiliser d'autres composés en C1 tels que le méthanol, la méthylamine et le formiate (Vannelli et al. 1998). Les deux méthyltransférases CmuA et CmuB forment un complexe enzymatique essentiel à la formation *in vitro* de méthyl-H₄F à partir de CH₃Cl (Studer et al. 2001). CmuA transfère d'abord le méthyle du CH₃Cl à un cofacteur corrinoïde, qui pris en charge par CmuB, le transfère du cofacteur corrinoïde à un cofacteur H₄F (Studer et al. 1999 ; Studer et al. 2001).

Chez *M. extorquens* CM4, les gènes *cmuA* et *cmuB* sont localisés dans différents groupes (clusters) de gènes. Chez d'autres souches dégradant le CH₃Cl appartenant à d'autres genres comme *Hyphomicrobium* (McAnulla et al. 2001a ; Borodina et al. 2004), les gènes *cmu* co-localisent dans un seul cluster comme chez *Hyphomicrobium* sp. MC1 (Figure 1.3A). Le gène *cmuA* est le gène le plus conservé de la voie *cmu*. Son utilisation comme biomarqueur de la voie *cmu* (revue de Bringel et al. 2019) a permis sa détection chez des bactéries du sol (Borodina et al. 2005 ; McDonald et al. 2005), de milieux marins (Schäfer et al. 2005 ; Cox et al. 2012) et de la phyllosphère (Nadalig et al. 2011).

Le méthyl-H₄F formé par CmuAB est pris en charge par MetF2, une 5,10-méthylène-H₄F réductase qui l'oxyde en méthylène-H₄F. Cette molécule entre dans le cycle de la sérine pour la formation de biomasse ou est oxydée en CO₂ pour dégager des équivalents réduits. Dans ce deuxième cas, la molécule de méthylène-H₄F est convertie en deux étapes en formyl-H₄F par une enzyme bifonctionnelle nommée FdD, une méthylène-H₄F déshydrogénase/méthényl-H₄F cyclohydrolase. Le cofacteur H₄F est libéré du formyl-H₄F par PurU, une formyl-H₄F hydrolase, ce qui aboutit à la formation de formiate. Enfin, le formiate est oxydé en CO₂ par une formiate déshydrogénase composée de 4 sous-unités (plusieurs copies des gènes *fdhABCD* sont présentes dans le génome de CM4). La voie *cmu* permet la récupération d'énergie à partir de nicotinamide adénine dinucléotide phosphate (NAD(P)⁺) réduit. D'autres gènes (*cmuC*, *paaE*, *hutI*, *fmdB*) sont présents dans les clusters *cmu* (Figure 1.3A) mais n'ont pas été caractérisés à ce jour. Le gène *cmuC* code pour une potentielle méthyltransférase et est essentiel pour l'assimilation du CH₃Cl et dispensable pour la déshalogénation du CH₃Cl (Vannelli et al. 1999). Le gène *paaE* code pour une possible oxydo-réductase qui pourrait agir en réactivant le cofacteur corrinoïde de la forme Co(II) en Co(I) (Roselli et al. 2013). Le gène *fmdB* code pour un potentiel régulateur transcriptionnel et le gène *hutI* une possible imidazolone propionase peut-être impliquée dans la formation de l'anneau imidazole de la cobalamine (Woodall et al. 2001). Ces gènes sont souvent localisés à proximité des trois gènes *cmuABC* suggérant une cotranscription (Doronina et al. 2004). Les gènes *fmdB*, *paaE* et *hutI* ont été montrés comme bénéfiques pour la croissance avec le CH₃Cl en contexte *M. extorquens* (Michener et al. 2016).

1.4.2. Voies *cmu*-indépendantes de dégradation du CH₃Cl

Le génome de *Leisingera methylohalidovorans* MB2, une souche marine capable de croître avec le CH₃Cl, ne possède pas les gènes *cmu* (Schaefer et al. 2002 ; Buddrhuhs et al. 2013). Il existe donc au moins une autre voie d'utilisation du CH₃Cl en plus de la voie *cmu* (Nadalig et al. 2014). Jing Luo, au cours de sa thèse dans notre laboratoire, a mis en évidence une potentielle méthyltransférase (LEIME_2500 appelée HMT pour hydroxyl methyltransférase) chez *L. methylohalidivorans* MB2 (Luo 2021). Une recherche de séquence des protéines CmuA et HMT dans les génomes de souches dégradant le CH₃Cl a permis de montrer que

généralement, lorsque les souches possèdent un gène homologue à celui de *L. methylohalidivorans* MB2, les gènes de la voie *cmu* ne sont pas détectés, et inversement (Tableau 1.1). Les deux systèmes peuvent cependant cohabiter, comme chez *Aminobacter ciceronei* IMB-1.

Tableau 1.1 La présence simultanée du gène *cmuA* et des gènes de la voie de dégradation du CH₃Cl de *L. methylhalidivorans* MB2 chez plusieurs souches

Souche	Identité (%) au niveau protéique avec ^a				
	<i>M. extorquens</i> CM4 CmuA	<i>Leisingera methylhalidivorans</i> MB2			
		LEIMEv1_2498	LEIMEv1_2499	LEIMEv1_2500 ^d	LEIMEv1_2501
<i>Leisingera methylhalidivorans</i> MB2	NA^b	100%	100%	100%	100%
<i>Aminobacter ciceronei</i> IMB-1	80.2%^c JACJHZ01_180069	46.0% JACJHZ01_150140	49.3% JACJHZ01_150141	91.7% JACJHZ01_150142	44.3% JACJHZ01_150143
<i>Roseovarius</i> sp. UBA9371	NA	NA	NA	87.2% DNBV01_950008	43.3% DNBV01_950007
<i>Roseovarius</i> sp. 217	NA	NA	NA	88.9% AAMV_v1_30294	41.8% AAMV_v1_30293
<i>Roseovarius nitratireducens</i> TFZ	NA	NA	NA	87.5% PDNI01_770124	43.3% PDNI01_770123
<i>Roseovarius</i> sp. MCTG156(2b)	NA	NA	NA	88.3% JQLS01_83747	44.7% JQLS01_83748
<i>Roseovarius</i> sp. A-2	NA	NA	NA	88.6% BDIY01_30158	40.1% BDIY01_30157
<i>Ruegeria</i> sp. ANG-S4	NA	NA	NA	87.8% JWLK01_10158	45.2% JWLK01_10157
<i>Maritimibacter</i> sp. UBA9533	NA	NA	NA	88.7% DOUP01_1970008	45.4% DOUP01_1970007
<i>Methylobacter</i> sp. BBA5.1	NA	NA	NA	65.7% JQKS01_v1_150004	NA
<i>Methylobacter whittenburyi</i> ACM 3310	NA	NA	NA	65.7% MWIT_v1_10136	NA
<i>Methylobacter marinus</i> A45	NA	NA	NA	65.4% MMARv2_30320	NA
<i>Aminobacter lissarensis</i> DSM 17454	80.2% JACJHY01_180087			NA	
<i>Hyphomicrobium</i> sp. MC1	84.0% HYPMCv2_2273		35% MetF2 HYPMCv2_2268	NA	

Souches confirmées expérimentalement comme dégrader le CH₃Cl (gras)

^a Nomenclature des gènes de MaGe <https://mage.genoscope.cns.fr>

^b NA=Non Applicable

^c cluster de gènes similaire à *Hyphomicrobium* MC1

^d *panB* 23.5% identité avec LEIMEv1_2500 chez *Leisingera methylhalidivorans* MB2

LEIMEv1_2499, 5,10-methylene-H₄F reductase

LEIMEv1_2500, putative 2-oxoacid hydroxymethyltransferase

LEIMEv1_2501, putative SAM-binding methyltransferase

1.5. Le transfert horizontal de gènes, un moteur de l'évolution bactérienne

Le transfert horizontal de gène (HGT pour “Horizontal Gene Transfer”) est défini comme un transfert de matériel génétique entre deux organismes ne partageant pas de lien de parenté (un transfert d'un parent à son descendant est qualifié de vertical). Comme les bactéries sont des organismes asexués se reproduisant par division cellulaire (les cellules filles ont le même patrimoine génétique que la cellule mère, mis à part les mutations acquises lors de la réplication), le transfert horizontal est un facteur majeur de diversification et de brassage génétique chez ces organismes unicellulaires (Soucy et al. 2015). Nombre de gènes de virulence chez les bactéries pathogènes ou de résistance aux antibiotiques ont été acquis par des événements de HGT (Gyles et Boerlin 2014 ; Lerminiaux & Cameron 2018). Par suite de sa divergence avec *Salmonella*, *E. coli* MG1655 a acquis des gènes par HGT qui représentent aujourd'hui 18% du génome (755 gènes sur un total de 4288) (Lawrence et Ochman 1998) sur une période de 14 millions d'années (Ma) avec un taux de 16 kb/Ma. Un HGT de gènes codant pour des enzymes et des réactions périphériques semble plus fréquent que pour des gènes de ménage (Popa & Dagan 2011),

Les transferts horizontaux sont plus fréquents entre bactéries de contenu GC similaire, notamment entre espèces apparentées. En effet, pour les HGT impliquant des mécanismes de recombinaison homologue, une similarité de séquence favoriserait l'intégration de gènes exogènes (Popa & Dagan 2011 ; Varner & Gunsch 2021).

1.5.1. Principaux mécanismes de transfert horizontal

Plusieurs mécanismes de transfert horizontal existent pour assurer le transfert de matériel génétique entre bactéries (Soucy et al. 2015) (Figure 1.4), avec trois mécanismes principaux :

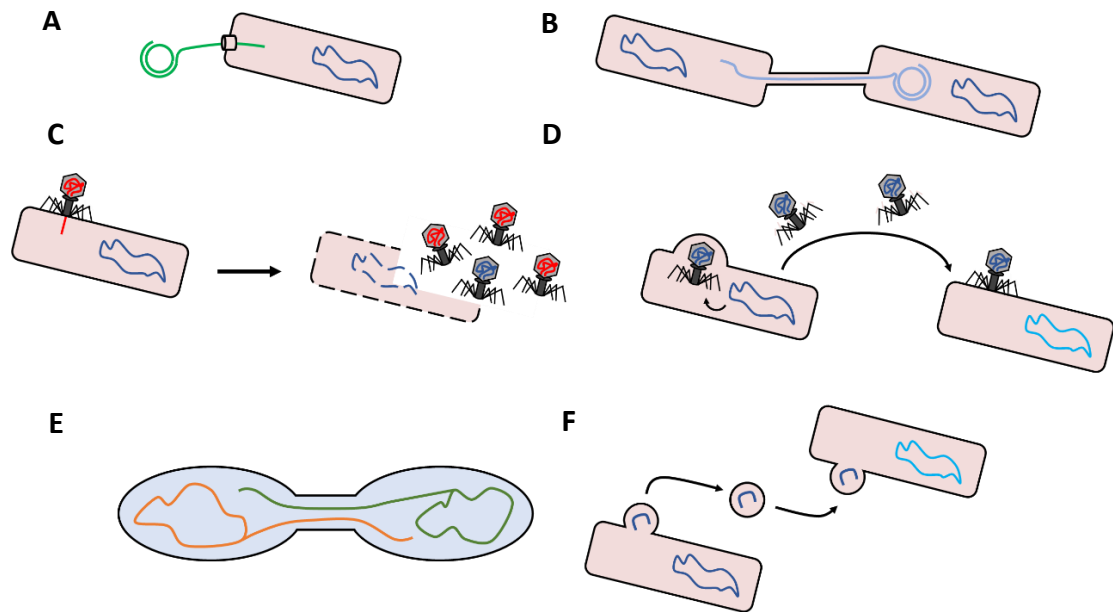


Figure 1.4 Le transfert horizontal chez les procaryotes

La transformation (A) permet l'acquisition d'ADN nu issu de l'environnement. La conjugaison (B) résulte du transfert entre deux partenaires ayant un contact direct. La transduction (C) résulte de l'infection d'un phage qui va embarquer de l'ADN issu de son hôte pour le transmettre à une autre. Les Agents de Transferts de Gènes sont des petites particules de type bacteriophage-like produites par des bactéries donneuses (D) qui ont encapsidé de l'ADN pour le transférer à une souche réceptrice. D'autres mécanismes impliquent le transfert de gènes par fusion de cellules (E) chez les archées ou bien la vésiduction où du matériel génétique peut être transmis par des vésicules (F). Adapté de Soucy et al. 2015.

- La transformation. Il s'agit de l'entrée d'un fragment d'ADN exogène nu à travers l'enveloppe de bactéries dites compétentes, qui sera ensuite intégré dans le génome hôte. Ce mécanisme a été signalé à la fois chez les bactéries et chez les archées (Johnston et al. 2014 ; Chimileski et al. 2014).
- La conjugaison. Elle nécessite un contact physique entre une souche donneuse et une souche réceptrice par l'intermédiaire d'un pilus de conjugaison (voir Virolle et al 2020 pour le mécanisme). Le transfert de gènes par conjugaison n'est pas restreint aux échanges entre bactéries. En particulier, *Agrobacterium tumefaciens* peut effectuer un transfert de matériel génétique de bactérie à plante en provoquant des tumeurs (Guo et al. 2019). Les éléments conjuguatifs peuvent être des plasmides conjuguatifs ou des éléments conjuguatifs intégratifs (abrégiés ICEs en anglais) (Johnson & Grossman 2015). Le transfert de gènes par conjugaison a été mis en évidence pour des environnements

très variés comme des écosystèmes animaux (intestins humains, de souris, et de larves d'insecte), dans la rhizosphère de plantes, au niveau des feuilles de plante, des sols et eaux non pollués, des environnements pollués de sols, des boues activées, et des eaux usées (Davison 1999).

- La transduction. Il s'agit du transfert de matériel génétique (ADN bactérien) entre deux bactéries par l'intermédiaire d'un vecteur viral. Concrètement, un bactériophage prédatant une bactérie peut encapsider en plus de son propre génome, une partie du génome de la bactérie hôte. Lorsque le bactériophage réalise un nouveau cycle d'infection, l'ADN est transféré à une cellule réceptrice. Il existe deux types de transduction : une transduction généralisée durant laquelle un phage lytique va entraîner la dégradation du génome de la cellule hôte. Un fragment du génome de l'hôte est alors encapsidé en plus du génome viral et les virions sont libérés par lyse de la cellule. Lors d'un cycle lysogénique, un prophage qui s'était intégré dans le génome de son hôte peut s'exciser en emportant de l'ADN de l'hôte flanquant son propre génome. La transduction est un phénomène qui a été observé chez les bactéries et les archées. La transduction est le mécanisme où la distance entre les deux partenaires peut être la plus grande (Popa & Dagan 2011) par rapport à la transformation limitée par la stabilité de l'ADN nu et la conjugaison qui nécessite un contact physique entre les deux partenaires.
- Les agents de transferts de gène (abrégé en anglais GTA pour "Gene Transfert Agents"). Les GTA dérivent de prophages ayant perdu la capacité de cibler leur propre gène pour l'encapsidation (Lang et al. 2012). Les GTA peuvent encapsider des fragments aléatoires de petite taille du génome de l'hôte et les transférer à des cellules réceptrices proches de bactéries ou d'archées.
- D'autres moyens de transfert horizontal de gènes impliquent des vésicules membranaires extracellulaires (ou "membrane-bound extracellular vesicles") et des nanotubes. Des échanges de matériel génétique entre bactéries ou archées pourraient impliquer des structures tubulaires membranaires appelées nanotubes qui forment des ponts entre les cellules (Dubey & Ben Yehuda, 2011). Chez *Bacillus subtilis*, la structure CORE d'ancrage des flagelles servirait aussi à l'assemblage de nanotubes. En absence de CORE, les échanges de matériel entre cellules de *Bacillus subtilis* seraient abolis (Bhattacharya et al 2019). Chez les archées (*Euryarchaeota* et *Crenarchaeota*),

des mécanismes de fusions cellulaires (et de petits ponts) permettent des transferts bidirectionnels où transite le matériel génétique (Naor & Gophna 2013 ; Schleper et al. 1995). La vésiduction réfère au transfert de matériel génétique par des membranes vésiculaires (Abe et al. 2020 ; Soler & Forterre 2020).

1.5.2. Les limitations de l'acquisition de gènes par HGT

L'utilisation d'une nouvelle voie métabolique dont les gènes ont été acquis par transfert horizontal peut se heurter à divers freins dans la cellule réceptrice. Le destin des gènes transmis par HGT dépend du bénéfice qu'ils apportent ou du coût pour le fitness de la souche réceptrice (Bershtein et al. 2016). L'effet sur le fitness d'une bactérie réceptrice modèle a été évalué dans une étude où des protéines issues de 79 génomes procaryotes avaient été exprimés dans un vecteur d'expression chez *E. coli* sous la dépendance d'un promoteur inducible. L'expression de beaucoup des gènes étrangers avait eu un effet négatif pour la souche réceptrice *E. coli*, notamment du fait d'un effet de toxicité lié au dosage (Sorek et al. 2007). Dans le cas de transfert de gènes homologues dont une version est déjà présente dans la cellule, un effet négatif sur le fitness de la souche réceptrice est observé lorsque les gènes sont exprimés avec comme conséquence un déséquilibre de dosage de gènes (l'expression des gènes transférés perturbe la concentration du produit de ceux déjà présents dans la cellule), ou un désordre de structures protéiques (une protéine avec une structure tertiaire instable ou flexible forme des interactions non-spécifiques avec d'autres molécules qui peut conduire à des repliements dysfonctionnels voire la formation d'agrégats) (Acar Kirit et al. 2020). Ces effets à court-terme conditionnent la probabilité de maintien des gènes transférés à long-terme (Acar Kirit et al. 2020). Néanmoins, le HGT peut avoir un effet neutre sur le fitness de la cellule hôte (Wellner et al. 2008 ; Knöppel et al. 2014).

Des facteurs intrinsèques à la composition de l'ADN transféré peuvent influencer le succès d'un HGT. Ainsi, le pourcentage en GC, le biais d'usage des codons et la stabilité de l'ARN sont autant de facteurs pouvant directement impacter l'efficacité de l'expression des gènes transférés au niveau de la transcription et de la traduction. Au niveau cellulaire, d'autres facteurs pourraient affecter le succès du HGT comme les réseaux de régulation, les flux métaboliques, les interactions protéine-protéine ou encore les processus maintenant

l'homéostasie protéique dans la cellule. Dans ce contexte, on parle de connectivité pour définir le nombre d'interactions entre un gène transféré, le produit pour lequel il code et les gènes et produits des gènes de la cellule hôte (Jones et al. 2022). L'hypothèse la plus simple est qu'un gène acquis par transfert horizontal avec une faible connectivité a plus de chances d'être fixé dans la cellule hôte qu'un gène ayant une haute connectivité. La fonction associée au gène transféré et l'environnement fluctuant de l'hôte complexifie cette connectivité en un réseau d'interactions dynamiques en constante évolution, qui peut ponctuellement modifier l'effet neutre, bénéfique ou délétère de l'acquisition de gènes par HGT, et par conséquent aussi le destin de la cellule hôte.

1.5.3. Indices pour un transfert horizontal des gènes *cmu*

Plusieurs éléments suggèrent que les gènes de la voie *cmu* ont été acquis par transfert horizontal de gènes (Figure 1.5). Tout d'abord, les gènes de la voie *cmu* ont été retrouvés chez plusieurs souches bactériennes phylogénétiquement éloignées comme des Alphaprotéobactéries (*Methylobacterium*, *Hyphomicrobium*, *Aminobacter*, *Celeribacter*, *Roseovarius*, *Desulfobacula*), des Deltaprotéobactéries (*Desulfomonile tiedjei*) ou même des bactéries à Gram positif (*Thermincola potens*, *Desulfurispora thermophila*, *Desulfotomaculum alcoholivorax*). Dans la plupart de ces souches, les gènes de la voie *cmu* ont conservé une organisation similaire en un seul cluster de gènes et avec une forte conservation en termes d'identité de séquence (Nadalig et al. 2014 ; Bringel et al. 2019).

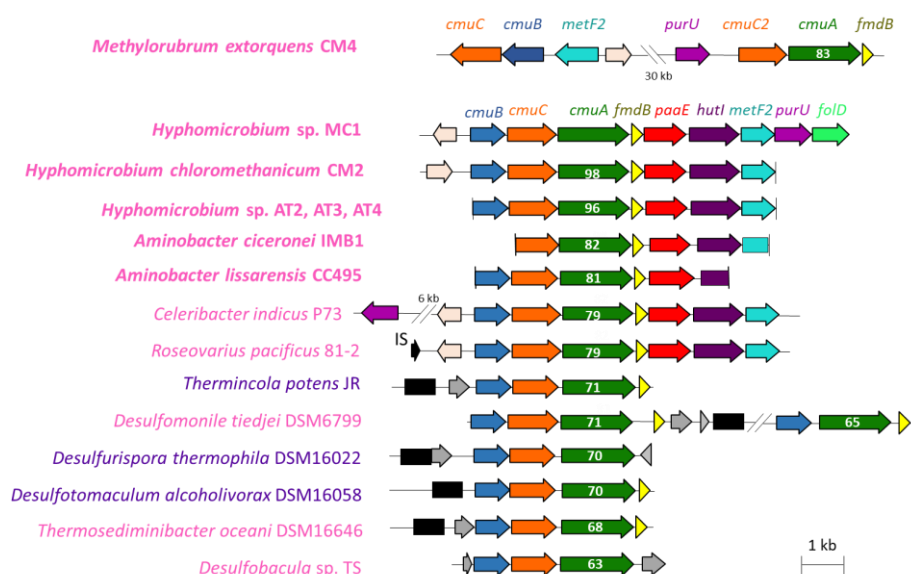


Figure 1.5 Organisation des gènes *cmu* chez plusieurs souches bactériennes

Les flèches représentent des gènes codant des protéines, de même couleur pour des homologues. En noir sont représentés des éléments mobiles, en rose pâle des régulateurs transcriptionnels et en gris des gènes non impliqués dans le métabolisme du CH₃Cl. Pour *cmuA*, la similarité de la séquence en acides aminés par rapport à la référence *Hyphomicrobium* sp. MC1 est aussi indiquée. Le nom des souches est en rose pour les Gram-positives et en violet pour les Gram-négatives. Un nom en gras indique que la souche a été expérimentalement validée comme pouvant croître avec le CH₃Cl. D'après Bringel et al. 2019 et Nadalig et al. 2014.

M. extorquens CM4 fait figure d'exception : I) Cette souche possède les gènes *cmu* sous la forme de deux clusters éloignés de 30 kb et non d'un seul (Vannelli et al. 1998) ; II) l'ordre des gènes est unique ; et III) c'est la seule souche possédant les gènes *cmu* sur un plasmide (Figure 1.6). Le plasmide pCMU01 porte en plus des gènes de la voie *cmu* d'autres gènes pouvant faciliter sa croissance avec le CH₃Cl, avec notamment 13 gènes *cob* codant probablement pour la synthèse du facteur corrinoïde essentiel à l'activité CmuA-dépendante de déshalogénation du CH₃Cl (en plus des gènes *cob* du core génome portés par le chromosome de CM4), ainsi que d'autres gènes impliqués dans le métabolisme du folate (Roselli et al. 2013). Le plasmide pCMU01 arbore les gènes du système de réplication RepABC exclusif aux Alphaprotéobactéries (Roselli et al. 2013) ainsi que des composants d'un système de sécrétion/conjugaison de type IVB suggérant que le plasmide est conjugatif et qu'il peut être transmis entre Alphaprotéobactéries (Roselli et al. 2013). Enfin, les gènes *cmu* sont souvent retrouvés associés à des éléments mobiles (Bringel et al. 2019) (Figure 1.5).

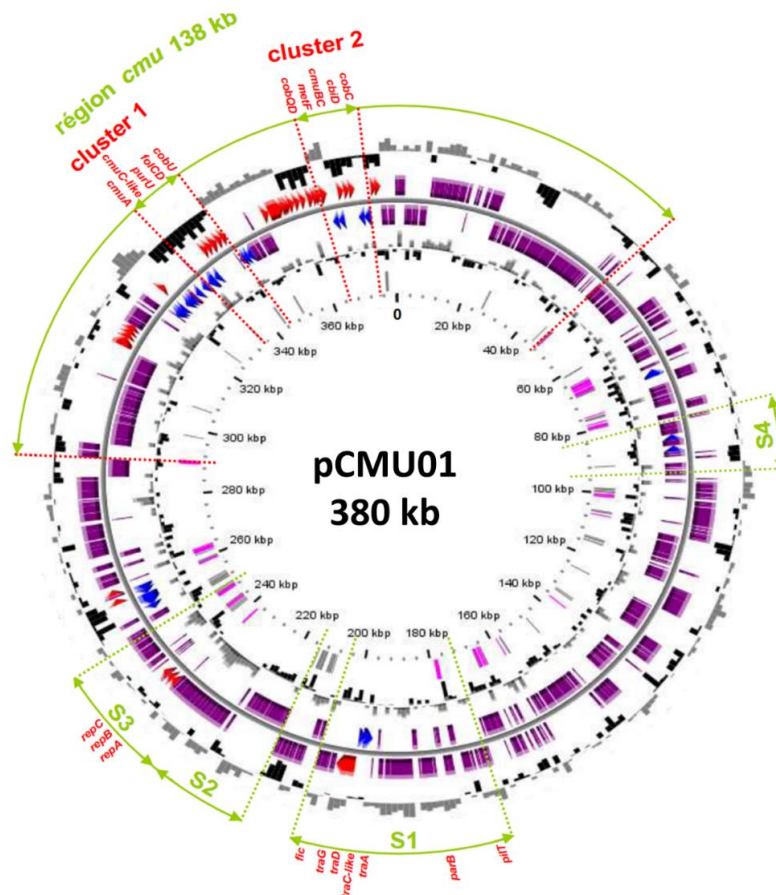


Figure 1.6 Le plasmide pCMU01 d'utilisation du CH₃Cl chez *M. extorquens* CM4

Les clusters 1 et 2 contiennent les gènes essentiels pour l'utilisation du CH₃Cl (Vannelli et al. 1999). Les régions S1 à S3 sont homologues à une région de 65kb du plasmide p1METDI de *M. extorquens* DM4. La région S4 montre une homologie avec une région dupliquée du chromosome de *M. extorquens* DM4. Les cercles de l'extérieur vers l'intérieur représentent : 1, le pourcentage de déviation GC (GC fenêtre – GC moyen) dans une fenêtre de 1 kb ; 2, les CDS prédits dans le sens des aiguilles d'une montre ; 3, les CDS prédits transcrits dans le sens inverse des aiguilles d'une montre ; 4, le GC skew (G+C/ G-C dans une fenêtre de 1 kb ; 5, les éléments transposables (rose) et les pseudogènes (gris). Repris de Roselli, 2009.

La plateforme MaGe (<https://mage.genoscope.cns.fr>) permet de détecter les îlots génomiques portant des gènes acquis par transfert horizontal en effectuant des comparaisons entre génomes de souches proches (grâce notamment à l'outil RGPFinder (Vallenet et al. 2009). L'utilisation de cet outil a permis de montrer que pour les deux souches appartenant au genre *Aminobacter* (Figure 1.5) comparées aux génomes de souches apparentées (80% similaire soit 3 génomes pour chacune des deux souches) dont le génome est disponible sur la plateforme, les gènes *cmu* sont localisés dans des régions identifiées comme îlots génomiques avec les coordonnées MaGe (*A. lissarensis* CC495 positions 5349442- 5488356 ; *A. ciceronei* IMB-1 positions 3715345- 3755695). Ces îlots génétiques portent des gènes de

transfert de type *tra* et *trb* impliqués dans la construction d'un système de sécrétion/conjugaison de type IV ainsi que des gènes RepA, RepB et RepC caractéristiques de la réplication de plasmides, qui sont des indices pour suggérer une acquisition des gènes *cmu* associée à des réplicons conjugatifs.

1.6. Limites à la croissance avec le CH₃Cl après acquisition des gènes de la voie *cmu* par transfert horizontal

1.6.1. Les stress de la croissance avec des composés chlorés

Une souche méthylotrophe peut-être inadaptée à la croissance avec des méthanes chlorés après acquisition de gènes de la voie d'utilisation de ces composés (Michener et al. 2014b ; Michener et al. 2016). En effet, le métabolisme catalysé par les voies d'utilisation de méthanes chlorés génère de nouveaux stress à laquelle la souche doit faire face et qui peuvent inhiber sa croissance.

1.6.1.1 Limites à la croissance avec les méthanes chlorés

Dans le cas de la dégradation aérobie du CH₂Cl₂, la voie commence par une déshalogénation catalysée par la glutathion-S-transférase DcmA avec libération d'ions chlorure, de protons, et la production de composés génotoxiques comme le formaldéhyde et surtout d'un intermédiaire métabolique très réactif, le S-chlorométhylglutathion. Il en résulte un stress intracellulaire acide avec aussi une augmentation de la concentration en ions chlorure. Ces derniers peuvent perturber l'équilibre du pool d'anions central à la régulation de la turgescence et du pool de potassium intercellulaire (Evans et al. 2000). Par ailleurs, le S-chlorométhylglutathion et le formaldéhyde peuvent former des adduits à l'ADN par N-alkylation des nucléotides (Kayser & Vuilleumier 2001) et par conséquent endommager l'ADN. Ces stress représentent autant d'obstacles à une croissance efficace par utilisation du CH₂Cl₂. De fait, l'introduction du gène *dcmA* essentiel à la croissance de *M. extorquens* DM4 avec le CH₂Cl₂ ne suffit pas à conférer la capacité de croissance à différentes souches appartenant au genre *Methylobacterium* (qui ne dégradent pas naturellement le CH₂Cl₂) (Michener et al. 2014a). Par ailleurs, la phylogénie ne permet pas de prédire la capacité d'une souche à pouvoir utiliser la voie *dcm*, puisque les souches les plus proches de *M. extorquens* DM4 ont montré

une moins bonne performance que les souches les plus distantes avec le CH₂Cl₂ (Michener et al. 2014a).

De manière similaire, il a été montré que l'acquisition par HGT des gènes de la voie *cmu* en laboratoire n'est pas suffisante pour acquérir la fonction de croissance avec le CH₃Cl (Michener et al. 2016). Dans ces travaux, les gènes *cmu* de *Hyphomicrobium* sp. MC1 avec leur propre promoteur ont été insérés dans le plasmide pCM62, un vecteur d'expression répliquable chez *Methylobacterium* (Marx & Lidstrom 2001) et dérivé du plasmide pRK2 du groupe d'incompatibilité IncP (Figurski & Helinski 1979). Le plasmide résultant nommé pJM105 (Michener et al. 2016) a été transféré par conjugaison à plusieurs souches du genre *Methylobacterium* incapables d'utiliser le CH₃Cl pour leur croissance, afin de mimer en laboratoire le transfert horizontal des gènes pouvant se produire naturellement. Les transconjugants ont montré une très faible croissance avec le CH₃Cl, du fait entre autres du stress acide et d'accumulation d'ions chlorure intracellulaire similaire à celui du CH₂Cl₂, même si a priori moins important (un seul proton et chlorure par carbone pour le CH₃Cl, deux pour le CH₂Cl₂).

1.6.1.2 Réponse commune à l'utilisation de méthanes chlorés

Les transcriptomes des souches *M. extorquens* DM4 et CM4 cultivées respectivement avec du CH₂Cl₂ et du CH₃Cl, et avec le méthanol comme composé de référence de la méthylotrophie, ont été comparés dans une étude précédente (Chaignaud et al 2017). Douze gènes du core génome de *M. extorquens* étaient plus exprimés à la fois chez DM4 et CM4 en condition d'utilisation des méthanes chlorés uniquement. Parmi ces gènes, seuls deux ont des fonctions prédites.

- Le gène *degP* code pour une sérine protéase périplasmique et protéine chaperonne moléculaire du contrôle qualité du repliement des protéines du périplasme en réponse au stress. Or un stress acide favorise l'accumulation d'ions chlorure dans le périplasme, qui en absence de protéine chaperonne, provoque l'agrégation des protéines périplasmiques (Fu et al. 2018 ; Kim et al. 2021).
- Le gène *hpp* code pour une pyrophosphate synthase couplée à la translocation de protons. Cette protéine va utiliser l'énergie générée lors de l'hydrolyse du pyrophosphate nécessaire pour l'extrusion de protons à travers la membrane tout en créant un potentiel

de protons (Chaignaud et al. 2017 ; Belogurov and Lahti, 2002). Il a été démontré que HppA est impliquée dans la réponse au stress salin chez *Rhodospirillum rubrum* (Marqués et al. 2004).

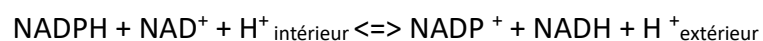
Quant au gène *clcA* codant pour un antiporteur H^+/Cl^- (METDI3481), il s'avère essentiel à la croissance avec le CH_2Cl_2 pour l'expulsion des ions chlorure au détriment de la force proton motrice. Ce gène a une transcription constitutivement élevée chez la souche DM4 (Michener et al. 2014b). Son homologue chez la souche CM4 (Mchl_3945, 99% d'identité au niveau protéique) n'est pas exprimé de manière différentielle, contrairement à une copie plasmidique du gène *clc* (Mchl_4138, 69% d'identité protéique avec Mchl_3945) qui est plus fortement transcrit pendant la croissance avec le CH_3Cl (Chaignaud et al. 2017). Ainsi, différents homologues du gène *clc* pourraient être impliqués dans la réponse adaptative au CH_2Cl_2 et au CH_3Cl , comme proposé précédemment (Michener et al. 2016).

1.6.1.3 Des stress spécifiques au métabolisme de la voie *cmu* d'utilisation du CH_3Cl

Parmi les conditions limitant la croissance avec le CH_3Cl , non partagés avec le CH_2Cl_2 ou le méthanol, plusieurs paramètres sont à envisager :

- La concentration relative en cofacteurs oxydés $NADP^+$ et NAD^+ pourrait affecter la croissance avec le CH_3Cl . Pour rappel, l'oxydation du CH_3Cl jusqu'au formiate génère deux molécules de $NAD(P)H$ alors que celle du méthanol en génère qu'une seule (comme avec le CH_2Cl_2). Dans la voie *cmu*, ce sont les enzymes MetF2 et FOLD qui catalysent des étapes H_4F -dépendantes de l'oxydation du CH_3Cl alors que ce sont les enzymes MtdA et Fch qui interviennent lors de l'oxydation du méthanol et du CH_2Cl_2 . Si MetF2 et FOLD ont une préférence différente (NAD^+ ou $NADP^+$) par rapport à MtdA et Fch, le pool de cofacteurs oxydés pourrait limiter la croissance avec le CH_3Cl . C'est l'hypothèse la plus crédible au vu des résultats du transcriptome (Chaignaud et al. 2017) et du protéome (Roselli et al. 2013) de *M. extorquens* CM4 où une expression significativement plus élevée en condition CH_3Cl qu'en méthanol du cluster *pnt* a été observée.

Le cluster *pnt* code pour une transhydrogénase transmembranaire qui catalyse la réaction suivante :



Cette transhydrogénase pourrait ainsi permettre de rééquilibrer la répartition de cofacteurs oxydés de manière à favoriser l'utilisation de la voie *cmu*.

- Le ratio des formes oxydées et réduites des cofacteurs pourrait aussi tendre à diriger le flux du méthylène-H₄F produit lors du catabolisme du CH₃Cl (Figure 1.3) vers le cycle de la sérine (biomasse) qui nécessite l'oxydation des cofacteurs. Pour l'utilisation du méthylène-H₄F vers la formation de formiate (énergie), la cellule semble activer le système de clivage de la glycine (abondance accrue au niveau ARNm et protéines détectée chez *M. extorquens* CM4 cultivée avec du CH₃Cl par rapport au méthanol (Roselli et al. 2013 ; Chaignaud et al. 2017). Ce système catalyse la réaction :



Lors de la croissance avec le méthanol ou le CH₂Cl₂, le flux métabolique lié au transport des groupements méthyle associés au H₄F fonctionne plutôt vers le cycle de la sérine et non dans le sens de l'oxydation du carbone vers la formation de formiate comme c'est le cas pour la croissance avec le CH₃Cl (Michener et al. 2016 ; Nayak & Marx 2014b). Ainsi l'expression de la voie *cmu* entraîne sans doute une perturbation majeure du métabolisme méthylotrophe chez les souches réceptrices.

- Le pool des cofacteurs (H₄F et cofacteur corrinoïde) de la réaction de déshalogénation du CH₃Cl est le point d'entrée de la voie *cmu* dans le métabolisme méthylotrophe de la cellule hôte et ainsi un facteur clé de la croissance avec le CH₃Cl. Chez *M. extorquens* CM4, le plasmide pCMU01 héberge en plus des gènes de la voie *cmu*, des gènes *cob* qui chez *Pseudomonas denitrificans* codent pour la voie de synthèse aérobie de la cobalamine (Rodionov et al. 2003) ainsi que des gènes codant pour des transporteurs du cobalt et de corrinoïdes (Roselli et al. 2013). Le gène *bluB2* (codant pour une protéine avec 38% d'identité avec celle produite à partir du gène *bluB* chromosomique) catalyse la production de 5,6-diméthylbenzimidazole, un ligand de la cobalamine ainsi que la production d'érythrose-4P, un précurseur du H₄F (Roselli et al. 2013), reliant ainsi la biosynthèse des deux cofacteurs essentiels à l'activité de déshalogénation du CH₃Cl. Les protéines codées par la plupart de ces gènes sont significativement plus abondantes durant la croissance de *M. extorquens* CM4 avec le CH₃Cl par rapport à la croissance avec le méthanol (Roselli et al. 2013).

En résumé, la croissance par utilisation de méthanes chlorés génère des perturbations multiples. Il est difficile de prédire lesquelles empêchent la croissance et quelles sont les cibles génomiques d'adaptations correspondantes après acquisition des gènes du catabolisme des méthanes chlorés.

1.7. L'évolution expérimentale pour identifier les verrous limitant une croissance bactérienne

L'évolution expérimentale (ou évolution dirigée) est une approche très utilisée en laboratoire. Les micro-organismes comme les levures, les bactéries ou les virus sont particulièrement adaptés à ce type d'expériences. Ils se divisent rapidement ce qui permet l'obtention d'un grand nombre de générations avec une population comprenant de nombreux individus et ceci dans des conditions de culture contrôlées et reproductibles au laboratoire (Van den Bergh, 2018). Les conditions de l'expérience exercent une pression de sélection pour l'adaptation d'un organisme (Van den Bergh 2018). L'adaptation se fait par l'acquisition de mutations bénéfiques. L'apparition de mutations peut être spontanée ou induite (Watford & Warrington, 2023). Chez les bactéries, les mutations spontanées se déroulent avec un taux d'une substitution de 1 base sur 10^5 à 10^8 par génération (Watford & Warrington 2023). Des erreurs de réplication de l'ADN Pol III peuvent conduire à une perte (délétion) ou un ajout d'un nucléotide (insertion) (Brown 2002). Un mauvais appariement peut être à l'origine d'une mutation sur les brins d'ADN (Brown 2002). C'est le cas lors de dysfonctionnements de la machinerie de réplication si l'ADN Pol III s'apparie à un nucléotide complémentaire incorrect à celui sur le brin parent, ou si le système de relecture (« proofreading ») de la polymérase ne détecte pas l'erreur. D'autre part, la tautomérisation de bases azotées durant la réplication peut causer un mésappariement qui va aboutir à une erreur de réplication. Les bases azotées existent sous plusieurs formes, appelées tautomères (Karalkar et al. 2017). Les modifications les plus courantes des bases azotées sont sous forme keto (C=O) et amino (C-NH₂) et rarement sous forme imino (C=NH) ou enol (=C-OH). Le changement d'un tautomère d'une nucléobase peut conduire à des appariements non-canoniques de paires de bases, et entraîner des erreurs de réplication. Des séquences avec des nucléotides répétés provoquent parfois le glissement de la polymérase lors de la réplication avec addition ou perte de bases (Fan & Chu 2007). Une

bibliographie abondante documente que des mutations sont induites suite à l'exposition d'un organisme à des agents mutagènes tels que des agents chimiques, physiques (des rayons UV) ou génétiques (des éléments génétiques mobiles comme par exemple des transposons s'intégrant dans une séquence) (Hutchinson et al 2019 ; Watford & Warrington, 2023).

Les conditions d'évolution expérimentale utilisent un environnement constant avec une pression de sélection qui peut imposer par exemple l'utilisation d'un nutriment comme une source de carbone (Van den Bergh 2018) alors que les autres paramètres (température, aération) sont gardés constants. La plus longue expérience d'évolution expérimentale a débuté en 1988 dans le laboratoire de Richard Lenski (« Long Term Evolution Experiment » ou LTEE) (Lenski 1991), avec des cultures d'*E. coli* maintenues dans un milieu contenant du glucose et du citrate. Des cultures indépendantes ou lignées sont propagées en parallèle dans les mêmes conditions de façon à faire émerger des adaptations chez des populations indépendantes. Cette expérience est toujours en cours et a dépassé 75 000 générations (Lenski 2023).

La présence d'une étiquette génétique qui confère un phénotype singulier au micro-organisme (gène codant une protéine fluorescente (Michener et al. 2014b) ou une résistance à un antibiotique, un phage ou conférant une couleur aux colonies) est souvent utilisée pour détecter et limiter des contaminations éventuelles (Van den Bergh 2018).

1.7.1. Trois types d'évolution expérimentale

Dans une expérience standard, le milieu de culture est régulièrement renouvelé afin de permettre de nouvelles divisions cellulaires. Trois modes d'évolution expérimentale sont utilisés (Figure 1.7) (Barrick & Lenski 2013 ; Van den Bergh 2018).

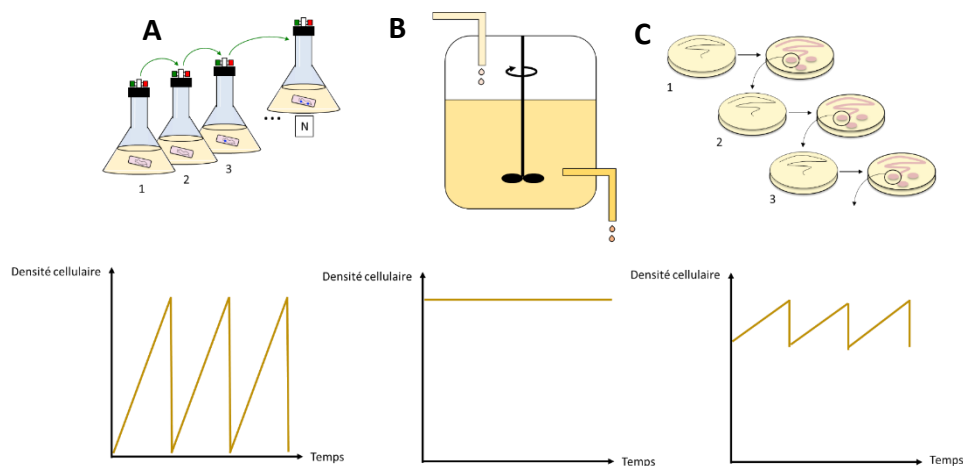


Figure 1.7 Les différents types d'évolution expérimentale

Pour chacun des 3 modes d'évolution expérimentale, le haut de la figure illustre le mode opératoire et le bas schématise sous forme de graphe, l'évolution théorique de la densité cellulaire au cours de l'expérience.

(A) Transferts en série d'une fraction de la population de façon périodique dans du milieu frais. (B) Culture continue en d'une population alimentée avec un apport à flux constant de milieu neuf et une sortie de suspensions cellulaires. (C) Repiquage régulier de colonies sur un modèle similaire à un transfert en série.

- **Culture discontinue (batch culture)** par transfert en série (Figure 1.7A) où une culture de micro-organismes est régulièrement réinoculée dans du milieu frais pour démarrer la culture suivante. C'est notamment par cette méthode de transferts en série qu'ont été maintenues les populations durant la LTEE du laboratoire de Richard Lenski. Elle est facile à mettre en œuvre pour suivre plusieurs lignées de cultures en parallèle. Au cours de chaque transfert, toutes les phases de croissance (phase de latence, phase exponentielle, phase stationnaire) sont successivement atteintes. Les cellules sont constamment dans une alternance de festins et de famines, passant d'un environnement où les nutriments sont facilement disponibles et faciles à utiliser à un environnement où ces derniers s'épuisent et des déchets s'accumulent. Ceci altère la pression de sélection et est source de dérive génétique et de stochasticité étant donné qu'un seul échantillon aléatoire (statistiquement non-représentatif de l'ensemble des variants) commence la culture suivante (Van den Bergh et al. 2018).
- **La culture continue** (Figure 1.7B) en chémostat (ou un de ses dérivés : turbidostat ou morbidostat p. ex.) où du milieu frais est constamment ajouté à une vitesse de dilution

fixe tout en retirant le milieu usé à la même vitesse. Cette méthode permet de garder les conditions de culture constantes à la différence des transferts en série. Les micro-organismes vont se développer en ayant une vitesse de croissance spécifique équivalente à la vitesse de dilution (Van den Bergh et al. 2018). Néanmoins, c'est une installation relativement complexe qui réduit le nombre de réplicats réalisables en parallèle, plus facilement sujette à la contamination et la formation de biofilm, et qui nécessite une surveillance plus importante. Une culture en continu pourrait être moins représentative des environnements dynamiques auxquels pourraient être soumis les micro-organismes dans l'environnement (Gresham & Dunham 2014).

- **Accumulation de mutations** (Figure 1.7C) par repiquage périodique de colonies sur un nouveau milieu solide. Ce mode pousse le concept de transfert en série à sa limite. Les populations passent par un goulot d'étranglement (« bottleneck ») d'un ou plusieurs individus, ce qui va faire que la probabilité qu'une quelconque mutation soit fixée est complètement aléatoire (Barrick & Lenski 2013). A l'exception des mutations létales ou particulièrement délétères, les mutations vont s'accumuler à une vitesse proche du taux de mutation de la souche, contrairement aux deux premières approches où la dynamique génétique est menée par l'apparition de mutations bénéfiques.

1.7.2. L'évolution expérimentale chez *M. extorquens*

L'évolution expérimentale a été utilisée chez *M. extorquens* depuis des décennies, d'abord chez la souche AM1 puis chez PA1, tout particulièrement par le groupe de Christopher Marx aux Etats-Unis (Marx 2019). *M. extorquens* a été évolué pour acquérir une capacité de croissance à de fortes concentrations de méthanol (Belkhefja et al. 2019 ; Cui et al 2018), avec le CH₂Cl₂ (Michener et al. 2014b), et la tolérance au 1-butanol (Hu et al 2016). Parmi toutes les expériences menées chez *M. extorquens*, trois exemples en lien plus étroit avec mon projet de thèse sont présentés ci-après.

- Cas 1 : inactivation de FtfL lors de la croissance avec le succinate (Lee et al. 2009). Huit lignées de *M. extorquens* AM1 ont été cultivées avec du succinate et huit autres avec du méthanol comme unique source de carbone pendant 1500 générations pour améliorer leur croissance avec ces substrats. Trois isolats par lignée ont été obtenus. L'évolution avec le succinate est corrélée à la perte la capacité de croître avec le méthanol et d'autres

composés tels que la méthylamine ou le succinate (6 populations évoluées sur 8). Tous ces isolats incapables de croître en méthanol ont montré une mutation non-identique dans *ftfL* (Carroll et al. 2014b) qui code une formyl-tétrahydrofolate ligase (Marx et al. 2003b). FtfL sert comme porte d'entrée au formiate dans le métabolisme associé au H₄F et à la formation de biomasse (Figure 1.3B). Environ 60 % du carbone total lors de la croissance avec le méthanol est assimilé grâce à FtfL (Carroll et al. 2014b ; Van Dien et al. 2003a ; Peyraud et al. 2011). FtfL est une enzyme pouvant transformer du formiate et du H₄F en formyl-H₄F mais peut aussi catalyser la réaction inverse (Figure 1.3B). Les auteurs ont fait l'hypothèse que durant la croissance avec du succinate, FtfL de par son activité de conversion du formyl-H₄F en formiate, entraînerait une diminution du pool de formyl-H₄F produit dans la cellule. Cette molécule est nécessaire pour des processus comme la synthèse des purines. Ainsi son inactivation permettrait de conserver le pool de formyl-H₄F. Une autre hypothèse formulée par les auteurs serait que l'augmentation du pool de formyl-H₄F du fait de l'inactivation de FtfL pourrait permettre une plus grande expression des gènes du cycle de la sérine, car cette molécule va permettre au régulateur transcriptionnel QscR de se fixer au promoteur des gènes de ce cycle (Kalyuzhnaya et Lidstrom 2005). Ceci serait avantageux pour augmenter la capacité de conversion de composés multicarbonés comme le succinate en d'autres composés.

- Cas 2 : évolution de la capacité de *M. extorquens* AM1 à s'adapter à l'utilisation d'une nouvelle voie d'oxydation du formaldéhyde. Une souche de *M. extorquens* AM1 ne pouvant plus utiliser la voie endogène H₄MPT-dépendante (Chou et al. 2011) a reçu le plasmide pCM410 arborant la voie d'oxydation du formaldéhyde de *Paracoccus denitrificans* par clonage des gènes *flhA* et *fghA* sous la dépendance d'un promoteur constitutif de *M. extorquens* (*P_{mxaF}*). Chez *Paracoccus denitrificans*, le formaldéhyde (CH₂O) réagit spontanément avec le glutathion (GSH) pour former de l'hydroxyméthyl-GSH (GS-CH₂OH) qui va être oxydé en formyl-GSH (GS-CHO) par une GS-CH₂OH déshydrogénase NAD-dépendante (encodée par *flhA*), formant du NADH. Le groupe formyl est ensuite hydrolysé en formiate par une GS-CHO hydrolase (encodée par *fghA*). Le formiate formé est ensuite oxydé en CO₂ ou utilisé pour la formation de biomasse par la voie H₄F-dépendante (Chou et al. 2011). Cette souche (appelée EM pour « *engineered Methylobacterium* ») avait évolué par culture avec le méthanol pendant 600 générations. À la suite de cette expérience, plusieurs mutations avaient été détectées dans le génome

de l'isolat ayant le meilleur fitness. Une mutation chromosomique dans un promoteur a eu pour effet d'augmenter l'expression de l'enzyme GshA (*gshA* codant une γ -glutamylcystéine synthétase) de la première étape de la synthèse du cofacteur GSH. Une autre mutation chromosomique du promoteur de *pntA* (codant une transhydrogénase) a augmenté l'expression de cette enzyme impliquée dans l'interconversion du NADH et NADP⁺ en NAD⁺ et NADPH qui régulerait le pool cellulaire de NADPH perturbé par la nouvelle voie. Une délétion de 11 pb entre *flhA* et *fghA* qui réduit l'expression des deux gènes (de 55 à 73%) a également été observée. D'autres mutations entraînent une diminution de l'expression des gènes *flhA* et *fghA* du plasmide pCM410 chez l'ensemble des isolats obtenus, à la suite d'évènements mutationnels différents (délétion d'un site de fixation au ribosome, insertion d'une IS au niveau d'un promoteur ou formation d'un coïntégrat entre pCM410 et un plasmide natif de la souche). On notera aussi que des mutations bénéfiques ont eu lieu à la fois au niveau des gènes exogènes de la nouvelle voie métabolique (plasmide pCM410) qu'au niveau du génome de l'hôte. Cette expérience est un bon exemple de convergence évolutive : lorsque plusieurs populations indépendantes ont été soumises à une même pression de sélection, elles ont développé plusieurs stratégies différentes qui tendent vers le même but (diminuer l'expression des gènes du plasmide pCM410).

- Cas 3 : adaptations à la croissance avec le CH₂Cl₂ chez *Methylobacterium* après acquisition d'un plasmide porteur du gène *dcmA* avec son promoteur natif. Après 150 générations, toutes les souches isolées présentent des mutations dans leur chromosome visant à pallier le stress causé par les ions chlorure produits durant la croissance avec le CH₂Cl₂. En effet, des mutations touchant différents mécanismes d'export des ions chlorure ont été détectées, comme des mutations dans le promoteur de *clcA* (antiporteur H⁺/Cl⁻), des mutations dans *secY* (translocase de protéines, la transformant en un canal laissant passer les ions chlorure), des mutations dans *bsa* (canal à chlorure) ou *edgA* aussi potentiellement impliqué dans l'export de chlorure (Michener et al. 2014b). Le gène *clcA* avec son promoteur muté et conférant une surexpression a été cloné sur le plasmide portant *dcmA*. L'acquisition de ce seul plasmide confère la capacité de croissance à l'ensemble des souches testées initialement incapable d'utiliser le CH₂Cl₂. Ainsi le principal frein à la croissance avec le CH₂Cl₂ est l'accumulation des ions chlorure dans l'espace intracellulaire. Cependant, même si le CH₂Cl₂ et le CH₃Cl sont deux composés chlorés dont la dégradation

libère des ions chlorure, augmenter l'expression de *clcA* tout en fournissant pJM105, n'a pas permis d'augmenter le fitness des souches pour la croissance en CH₃Cl (Michener et al. 2016). Les auteurs en ont conclu que les freins à la croissance avec le CH₃Cl étaient différents de ceux observés avec du CH₂Cl₂.

1.8. Objectifs de la thèse

Plusieurs indices suggèrent que les gènes *cmu* sont acquis par transfert horizontal, notamment la conservation de l'organisation des gènes *cmu* chez divers taxons ou leur association avec des éléments génétiques mobiles. Néanmoins, l'utilisation de cette voie métabolique nécessite vraisemblablement des adaptations post-transfert car l'acquisition des gènes de la voie *cmu* ne confère pas la capacité de croître avec le CH₃Cl (Michener et al. 2016). La base moléculaire de ces adaptations est jusqu'à ce jour inconnue. Je me suis ainsi posé la question des freins limitant la croissance avec le CH₃Cl et de la nature des adaptations requises pour une croissance méthylotrophe par la voie *cmu* d'utilisation du CH₃Cl.

Pour répondre à cette question, j'ai utilisé une approche d'évolution expérimentale par transfert en série avec le CH₃Cl comme principale pression de sélection, en partant de deux souches de *Methylobacterium extorquens* proches de CM4 mais non-adaptées à la croissance avec le CH₃Cl auxquelles j'ai transféré les gènes *cmu* portés par le plasmide pJM105. Ces souches ont été soumises à l'évolution expérimentale afin d'obtenir des souches évoluées ayant acquis des mutations bénéfiques leur permettant de croître avec le CH₃Cl, d'identifier ces mutations par séquençage de leurs génomes, puis les fonctions cellulaires correspondantes. En utilisant cette approche, j'ai ainsi souhaité répondre à plusieurs questions spécifiques :

- Est-il possible de faire évoluer des souches méthylotrophes incapables d'utiliser le CH₃Cl comme substrat de croissance après acquisition des gènes de la voie *cmu* vers la capacité de croissance avec le CH₃Cl comme unique source de carbone et d'énergie ?
- Les adaptations à la croissance avec le CH₃Cl dépendent-elles du contexte génomique ?
- Les mutations bénéfiques pour la croissance avec le CH₃Cl correspondent-elles à une pluralité de réponses adaptatives ? Ou à l'inverse, comme cela a été observé pour le CH₂Cl₂ (Michener

et al 2014), une adaptation en réponse au stress chlorure suffit-elle pour permettre aux souches de croître avec le CH₃Cl ?

Je présente mes travaux de thèse destinés à répondre à ces questions dans trois Chapitres de Résultats. Le Chapitre 2 traite de la mise en place, du suivi, de l'isolement et de l'analyse des données de séquençage de souches évoluées issues de l'évolution expérimentale. Le Chapitre 3 présente l'analyse des adaptations trouvées chez les souches évoluées pour l'assimilation du formiate. Le chapitre 4 est consacré à l'étude des mutations rencontrées dans les gènes de la voie *cmu* et du plasmide portant ces gènes. Le Chapitre 5 présente les conclusions et perspectives de mon travail, et les matériels et méthodes utilisés sont regroupés dans le Chapitre 6. L'ensemble des références bibliographiques est rassemblé en fin de mémoire.

Chapitre 2. Adaptations de souches méthylotrophes à la croissance avec le CH₃Cl après acquisition des gènes de la voie *cmu*

Ce Chapitre décrit la mise en place de l'évolution expérimentale avec comme principale pression de sélection le CH₃Cl dans le contexte de deux souches de *M. extorquens* (AM1 et DM4) après acquisition des gènes *cmu* de *Hyphomicrobium* sp. MC1. Il détaille ensuite les paramètres du suivi des transferts successifs sur plus de 200 générations jusqu'à l'obtention de souches évoluées isolées. Au total, 13 souches évoluées ont été obtenues et caractérisées en termes de phénotype. Les génomes des souches évoluées ont été séquencés et comparés entre eux et avec les souches parentales dans le but d'identifier des mutations potentiellement bénéfiques pour la croissance avec le CH₃Cl apparues au cours de l'évolution expérimentale.

2.1. Stratégie pour la mise en place de l'évolution expérimentale

Les choix des souches parentales et les conditions de cultures après chaque transfert jouent sans doute un rôle dans le profil des mutations conduisant à la croissance avec le CH₃Cl. C'est pourquoi je commence ce Chapitre par expliciter ces choix.

2.1.1. Choix du plasmide pJM105 porteur de 9 gènes de la voie *cmu*

Pour mener à bien l'évolution des souches AM1 et DM4, je leur ai transféré le plasmide pJM105 utilisé dans des travaux d'évolution expérimentale précédents auxquels l'équipe d'accueil avait été associée (Michener et al. 2016). Ce plasmide dispose des gènes de la voie *cmu* de *Hyphomicrobium* sp. MC1 (Vuilleumier et al. 2011).

Les gènes *cmu* de *Hyphomicrobium* sp. MC1 sont regroupés en un seul cluster contrairement à ceux de *M. extorquens* CM4, une autre Alphaprotéobactérie dont le métabolisme méthylotrophe est proche (Vuilleumier et al. 2011). Michener et collaborateurs avaient également construit et utilisé le plasmide pJM50 avec les gènes *cmu* de la souche CM4 regroupés artificiellement par clonage. Une différence majeure entre ces deux plasmides est l'absence dans le plasmide pJM50 de 3 gènes (*fmdB*, *hutI* et *paaE* de fonction inconnue)

généralement associés aux gènes *cmu* chez les souches possédant les gènes *cmu* (Bringel et al. 2019). Lorsque la croissance par utilisation du CH₃Cl a été comparée chez des transconjugants portant l'un ou l'autre de ces plasmides, seule la présence du plasmide pJM105 a permis une croissance très faible. De fait, la délétion un par un des gènes *fmdB*, *hutI* et *paaE* du plasmide pJM105 entraîne une perte de fitness des transconjugants avec le CH₃Cl allant de 18 à 47% par rapport au plasmide pJM105 entier (Michener et al. 2016). C'est pourquoi, en définitive, seul le plasmide pJM105 portant l'ensemble des gènes *cmu* du cluster de la souche MC1 (*cmuA*, *cmuB*, *cmuC*, *metF2*, *purU*) ainsi que des gènes bénéfiques bien que dispensables (*fmdB*, *hutI*, *paaE*, *folD*) (Michener et al. 2016 ; Vannelli et al. 1999) a été choisi comme point de départ de mes travaux.

2.1.2. Les souches parentales choisies pour l'évolution expérimentale

J'ai décidé d'apporter le plasmide pJM105 à trois souches appartenant à l'espèce *Methylobacterium extorquens* : AM1, DM4 et PA1. Leur génome avait été annoté par un réseau d'experts à l'international en méthylotrophie (Vuilleumier et al. 2009 ; Marx et al. 2012). Les versions de ces souches considérées pour ce travail, notées ici A0, D0 et P0, ont été obtenues dans l'équipe de Chris Marx et portent toutes la mutation Δcel qui permet une croissance en milieu liquide homogène (Delaney et al. 2013). Cette délétion d'une région de 7,2 kb comprend 3 gènes : *celB* codant une cellulose-synthase ainsi que les gènes *celA* et *celC* également impliqués dans la synthèse de cellulose (Delaney et al.-2013). La délétion de ces trois gènes permet de réduire le processus d'agrégation des cellules en milieu liquide, rendant ainsi plus facile l'estimation de la biomasse par mesure de la DO₆₀₀. Elles possèdent également une fusion de gène non-essentiel avec un marqueur fluorescent : pour A0, il s'agit d'une fusion *katA::mCherry*, D0 *hpt::Venus* en plus d'une délétion de l'îlot *dcm* contenant les gènes d'utilisation du CH₂Cl₂ et pour P0 *hpt::mCherry*. Ces fusions de gène représentent un autre outil permettant de suivre la croissance de ces souches (par fluorescence) notamment par des tests de compétition (Michener et al. 2016), même si je ne m'en suis pas servi au cours de cette thèse.

Ces souches sont des méthylotrophes facultatifs au même titre que *M. extorquens* CM4, et phylogénétiquement très proches de cette dernière (Kato et al. 2005 ; Chaignaud et al. 2017). À l'exception du CH₃Cl, elles utilisent d'autres-composés en C1 tels que le méthanol ou la

méthylamine} en plus de composés multicarbonés comme le succinate (Large et al, 1961) et l'acétate (Schneider et al. 2012). *M. extorquens* CM4 possédait ainsi vraisemblablement déjà le métabolisme méthylotrophe permettant d'utiliser des composés comme le méthanol avant l'acquisition des gènes *cmu*. Si cette hypothèse est correcte, il est alors possible d'envisager qu'une souche appartenant à l'espèce *M. extorquens* puisse s'adapter à l'utilisation du CH_3Cl par la voie *cmu* par suite de l'acquisition des gènes correspondants. J'ai souhaité tester expérimentalement cette hypothèse en reproduisant le transfert des gènes *cmu* avec pour but d'identifier des mécanismes d'adaptation subséquente.

Plusieurs souches ont été utilisées afin de s'affranchir du contexte génétique d'origine et distinguer d'éventuelles adaptations souche-spécifiques. En effet, bien que les souches AM1, DM4 et PA1 soient très proches l'une de l'autre, elles comportent des différences : *M. extorquens* DM4 est une souche qui peut également utiliser le CH_2Cl_2 , un autre composé organochloré que le CH_3Cl , comme substrat de croissance, à la différence des souches AM1 et PA1. Ainsi, *M. extorquens* DM4 pourrait déjà être adaptée à certains stress associés au métabolisme de déshalogénation comme la production d'ions chlorure, contrairement aux deux autres souches. Les trois souches peuvent être discriminées par génotypage par PCR en ciblant des gènes du génome spécifique de chacune des souches, par exemple le cluster de gènes *mau* (pour l'utilisation de la méthylamine) pour AM1 et l'îlot *dcm* pour DM4. Il s'est cependant avéré au cours du travail que la souche que je croyais appartenir au contexte génétique de la souche PA1 était en réalité la même souche que D0. Ainsi, l'évolution expérimentale a été effectuée avec le double de réplicats DM4 par rapport aux réplicats AM1.

2.1.3. Une évolution expérimentale en deux temps

J'ai choisi de réaliser l'évolution expérimentale par des transferts en série en mode batch, ce qui a pu être rapidement mis en place au laboratoire au contraire d'un mode de culture en chémostat. Chaque semaine, un aliquot de la culture précédente a servi à ensemercer une nouvelle culture dans du milieu frais. Je me suis inspiré de l'étude de Michener et al. (2014b) qui décrit l'évolution de la souche AM1 vers l'utilisation du CH_2Cl_2 , en procédant en deux phases, avec en première phase du méthanol fourni en plus du CH_2Cl_2 lors des premiers transferts car les souches étaient encore incapables de se développer avec le composé. J'ai de mon côté commencé par une première phase dans laquelle les cultures successives, en plus

du CH_3Cl comme principale source de carbone dans le milieu (15mM), pouvaient utiliser le succinate fourni à faible concentration (0,2 mM) pour leur croissance. En effet, comme illustré dans la Figure 2.1A, les souches encore inadaptées au CH_3Cl n'effectuent pas de divisions cellulaires sources de mutations. Lors des transferts hebdomadaires de cette première phase, la dilution réalisée a été de 1/16 à chaque transfert. De cette manière, les souches ont pu effectuer au moins 4 divisions avant d'atteindre la même densité cellulaire qu'à la culture précédente. Lors de la deuxième phase, les transferts ont été réalisés avec une dilution de 1/64 (permettant aux souches de réaliser 6 divisions).

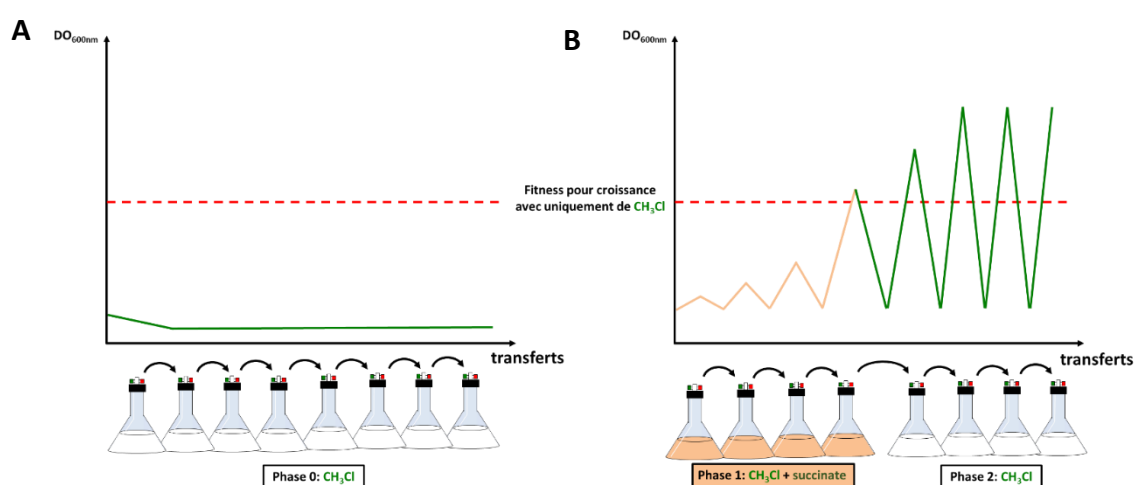


Figure 2.1 La stratégie d'évolution expérimentale

Les graphiques présentent l'évolution théorique de la densité cellulaire (représentée par la DO_{600}) au cours des transferts successifs. Un seuil théorique est fixé : le dépasser signifie que la population a accumulé suffisamment de mutations bénéfiques pour parvenir à se développer avec le CH_3Cl comme seule source de carbone et d'énergie. En A est présentée une stratégie avec une seule phase avec uniquement du CH_3Cl (phase 0) et en B une stratégie en deux phases. La phase 1 avec du succinate en quantité limitante (0,2 mM) et du CH_3Cl (15 mM), et la phase 2 avec uniquement du CH_3Cl .

J'ai privilégié le succinate au méthanol car les bactéries se développent plus vite avec ce composé. En gardant cette source de carbone à une concentration plus faible que celle du CH_3Cl , j'ai mis en place les conditions permettant aux cellules portant les mutations permettant la croissance avec le CH_3Cl d'être sélectionnées. En effet, une concentration de succinate de 0,2 mM ne permet qu'une croissance limitée (environ $\text{DO}_{600}=0,02$ après une semaine de culture contre $\text{DO}_{600}=0,5$ pour 15 mM de CH_3Cl). Le but était que les souches puissent atteindre un seuil à partir duquel elles seraient capables de croître avec le CH_3Cl

comme unique source de carbone et d'énergie (Figure 2.1B). Bien qu'il ait été impossible de déterminer ce seuil avec précision, j'ai décidé de fixer à 0,1 de DO₆₀₀ le seuil de passage vers une seconde phase de transferts où le succinate ne serait plus ajouté.

En plus du CH₃Cl, il existe une autre pression de sélection pour le maintien de pJM105 portant les gènes d'utilisation du CH₃Cl. Au même titre que le plasmide pCM62 dont il est issu, pJM105 peut être maintenu dans les transconjugants en exploitant la présence de deux gènes apportant une résistance à la tétracycline, un antibiotique bactériostatique. La tétracycline se fixe sur l'ARN 16S de la petite sous-unité 30S du ribosome, empêchant la fixation de l'ARNt amino-acylé au niveau du site A durant l'élongation (Chopra & Roberts 2001). Le gène *tetA* code pour une pompe à efflux expulsant l'antibiotique à l'extérieur de la cellule et *tetR* code pour un répresseur de la transcription de *tetA* (permettant une régulation fine de son expression) (Møller et al. 2016).

2.1.4. Les contrôles mis en place pour l'évolution expérimentale

Pour s'assurer que le promoteur natif des gènes *cmu* chez *Hyphomicrobium* sp. MC1 était bien reconnu en contexte *Methylobacterium* et qu'il n'y avait pas de problème d'expression, pJM105 a été introduit dans la souche mutante 22B3 de *M. extorquens* CM4 (Vannelli et al. 1998). Une insertion d'un transposon de type Tn5 dans le gène *cmuA* rend-la souche incapable de croître avec le CH₃Cl. J'ai souhaité établir si la présence de pJM105 permettrait de rétablir la croissance avec le CH₃Cl de cette souche, signe que les gènes *cmu* issu de MC1 peuvent être transcrits (et traduits) chez *M. extorquens*.

De plus, et en parallèle des cultures des deux souches possédant pJM105, une culture ayant le vecteur d'origine sans insert pCM62 a été suivie pour chacun des deux contextes génétiques. Ces contrôles négatifs ont permis de s'assurer que la croissance observée durant l'évolution expérimentale était bien due à la présence des gènes *cmu* apportés par pJM105 et donc de l'utilisation du CH₃Cl.

2.1.5. Stratégie pour l'identification des mutations bénéfiques à la croissance avec le CH₃Cl

Isolement de souches évoluées

À la suite de l'évolution expérimentale, j'ai recherché les mutations bénéfiques en comparant les séquences de génomes de souches évoluées isolées de réplicats indépendants plutôt qu'en comparant l'ensemble des mutations apparues dans une même culture. En effet, il est possible que plusieurs sous-populations ayant des stratégies évolutives différentes puissent émerger au sein d'une même culture. Isoler des souches évoluées permet ainsi de relier plus facilement plusieurs mutations à une même stratégie évolutive, ce qui aurait été difficile en séquençant toutes les mutations apparues au sein d'une même culture. D'autre part, j'ai fait l'hypothèse qu'il était également possible que des coopérations entre plusieurs sous-populations se mettent en place au sein d'une même culture. De fait, une cellule d'une culture se développant avec le CH₃Cl ne posséderait pas forcément toutes les adaptations nécessaires à la croissance avec ce composé. En choisissant cette approche, j'ai ainsi cherché à me rapprocher du contexte des souches comme *M. extorquens* CM4 et *Hyphomicrobium* sp. MC1 pour isoler des souches qui se suffisent à elles-mêmes pour la croissance avec le CH₃Cl.

Pour réaliser ces isollements, la technique dite d'étalement direct sur boîte permet d'obtenir rapidement des colonies isolées capables de croître avec du CH₃Cl. Néanmoins, j'ai estimé que cette méthode ne permettrait peut-être pas d'obtenir les souches pouvant croître seules avec le CH₃Cl en milieu liquide car elle n'écarte pas complètement la possibilité de relations complexes de dépendance et de coopération entre plusieurs sous-populations. Aussi, pour réduire les possibilités d'interactions entre cellules et me mettre dans des conditions plus proches des transferts en milieu liquide faites lors de l'évolution expérimentale, j'ai également utilisé une autre approche impliquant une étape supplémentaire d'isolement par « Extinction Dilution », consistant d'abord à mettre en culture des dilutions en cascade de cultures d'évolution expérimentale. A partir de la culture inoculée avec la dilution la plus forte ayant abouti à une croissance, j'ai ensuite procédé à l'isolement sur milieu solide comme pour l'étalement direct. Des souches évoluées ont été isolées par la méthode directe avec ou sans l'étape d'« Extinction Dilution » et ceci pour chacune des cultures d'évolution expérimentale.

Caractérisation de la capacité de croissance des souches évoluées

Afin de caractériser les capacités des souches évoluées isolées à utiliser le CH₃Cl, j'ai déterminé leurs temps de génération pour les principaux substrats de croissance d'intérêt (CH₃Cl, méthanol, succinate), car j'avais pour objectif de relier de façon plus fine les mutations mises

en évidence à un phénotype de croissance. Tester plusieurs substrats permet en plus d'évaluer si les différentes mutations apparues chez les souches évoluées ont un effet spécifique sur la croissance avec le CH₃Cl ou si elles perturbent le métabolisme plus globalement, touchant aussi la capacité de croissance avec d'autres composés.

*Evaluation de la présence de mutations bénéfiques portées par des gènes *cmu**

J'ai poursuivi la caractérisation des souches par une expérience visant à déterminer si des mutations cruciales pour l'évolution à la croissance avec le CH₃Cl s'étaient produites au niveau du plasmide pJM105 (par exemple des mutations influençant la transcription des gènes *cmu*), ici encore dans le but de relier mutation et phénotype. Pour cela, j'ai appliqué une stratégie qui a consisté à remplacer le plasmide pJM105 présent dans les souches par la version de départ (avant que les souches ne passent par l'évolution expérimentale) pour observer si la suppression des éventuelles mutations sur le plasmide affecte la croissance avec le CH₃Cl. Il a fallu tout d'abord obtenir des versions des différentes souches évoluées ayant perdu le plasmide pJM105 (souches évoluées curées). Pour cela, les souches évoluées ont été cultivées sans pression de sélection pour le maintien du plasmide pJM105, en absence de CH₃Cl ou de tétracycline. Puisque ce dernier représente un coût pour la cellule, le rendre inutile pour cette dernière devrait favoriser sa perte. Les souches évoluées ont été cultivées d'abord sur milieu liquide avec uniquement du succinate comme substrat de croissance et sans présence de tétracycline (on parle ici de curage). Une sélection sur milieu solide a ensuite permis de sélectionner les souches se développant avec le succinate mais ayant perdu la capacité de le faire en présence de tétracycline et de CH₃Cl comme source de carbone. On apporte ensuite le plasmide pJM105 afin de tester si les souches sont capables de croître avec le CH₃Cl.

Recherche des loci mutés communs à plusieurs souches

Le séquençage des souches évoluées a permis une comparaison de leur génome à celui des souches de départ A0 et D0 et de mettre en évidence les mutations qui sont apparues au cours de l'évolution expérimentale. Avec l'aide de nos collaborateurs du Génoscope ainsi que de Domitille Jarrige, post-doc en bioinformatique de l'équipe, j'ai dressé le profil de mutations de chacune des souches évoluées. J'ai recherché les loci mutés chez plusieurs souches évoluées

provenant de différents réplicats de l'évolution expérimentale (ou bien qui pourraient affecter la même fonction cellulaire). En effet, si plusieurs cultures indépendantes d'une même souche sont soumises à la même pression de sélection (en l'occurrence ici le CH₃Cl comme principal substrat de croissance), elles devraient indépendamment développer des adaptations similaires en réponse aux mêmes contraintes. Les paragraphes qui suivent font l'analyse des résultats obtenus dans les expériences qui ont été décrites.

2.2. Des souches inadaptées à la croissance avec le CH₃Cl ont acquis la capacité de croître avec ce composé

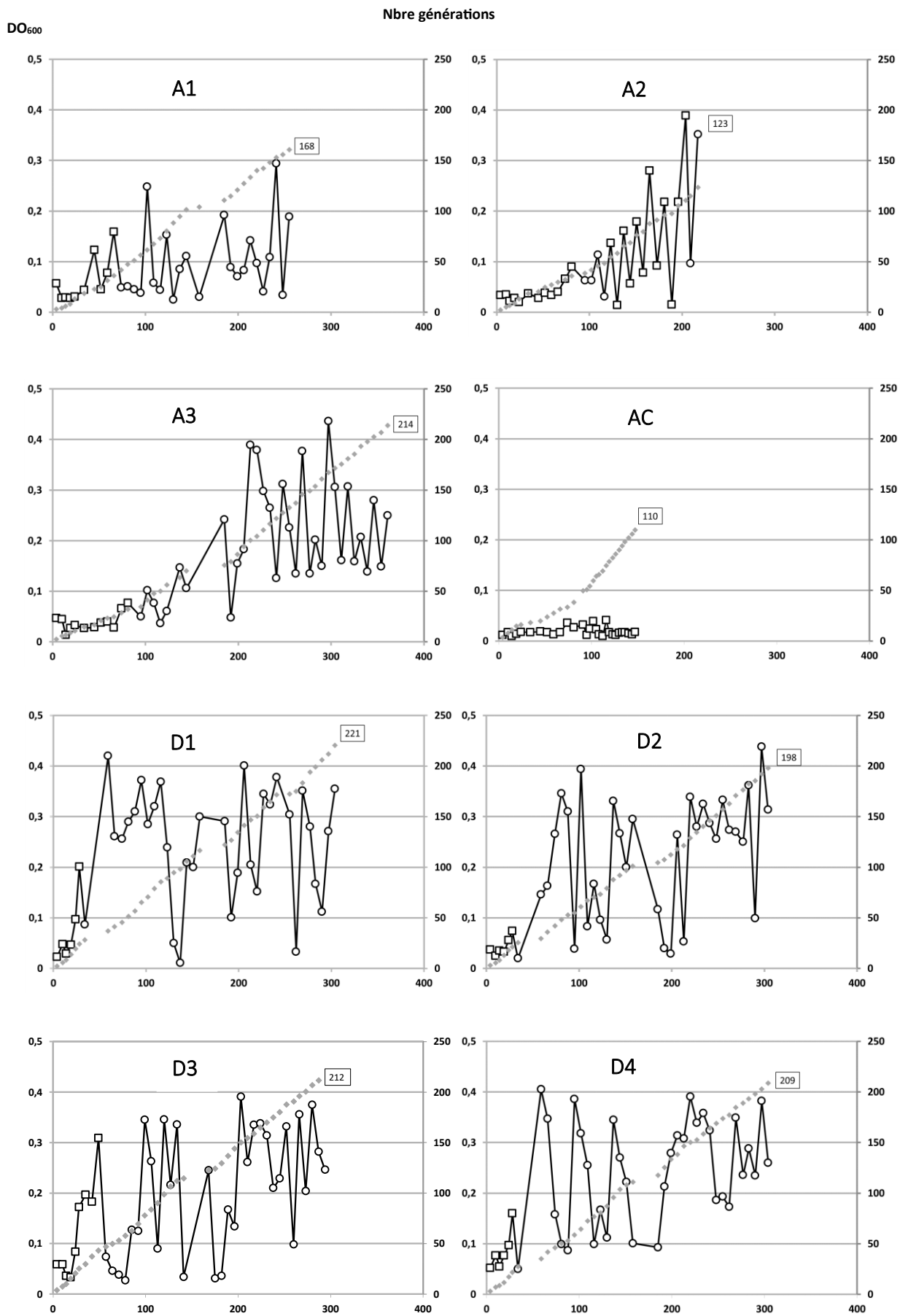
2.2.1. Les gènes de la voie *cmu* d'*Hyphomicrobium* présents sur le plasmide pJM105 sont exprimés en contexte *M. extorquens*

Le transfert de pJM105 à la souche mutante 22B3 de *M. extorquens* CM4 (Vannelli et al. 1999) a restauré la croissance de la souche avec le CH₃Cl, à la différence du plasmide pCM62 qui n'a pas restauré la croissance de la souche. Ainsi, une souche appartenant au contexte *M. extorquens* peut reconnaître un promoteur hétérologue natif de *Hyphomicrobium* sp. MC1. J'en ai conclu que fournir pJM105 à A0 et D0 ne devrait pas poser de problème pour l'expression des gènes *cmu*. L'incapacité des souches 105A et 105D obtenues à croître avec le CH₃Cl comme unique source de carbone ayant ainsi été confirmée, l'expérience d'évolution expérimentale a pu être initiée.

2.2.2. Profil au fil des transferts des cultures indépendantes vers la capacité de croître avec le CH₃Cl

Pour tous les réplicats, les cultures portant le plasmide pJM105 ont évolué vers une meilleure croissance avec le CH₃Cl dès la phase 1, contrairement aux contrôles portant le plasmide pCM62 (Figure 2.2). Les réplicats de cultures en contexte AM1 et DM4 ne présentent pas les mêmes profils d'évolution en phase 1. En effet, la plupart des réplicats de cultures en contexte DM4 dépassent une DO₆₀₀ de 0,1 dès le sixième transfert. En revanche, pour les réplicats de culture en contexte AM1, il a fallu attendre pour cela le 7^{ème}, 12^{ème} et 13^{ème} transferts pour les cultures A1, A2 et A3, respectivement. Dans les deux contextes, les cultures des témoins AC ou DC possédant le plasmide vide pCM62 n'ont montré qu'une très faible croissance ne

dépassant jamais une DO_{600nm} supérieure à 0,02. Cette faible croissance correspond comme attendu à l'utilisation de la faible quantité de succinate (0,2 mM) dans le milieu.



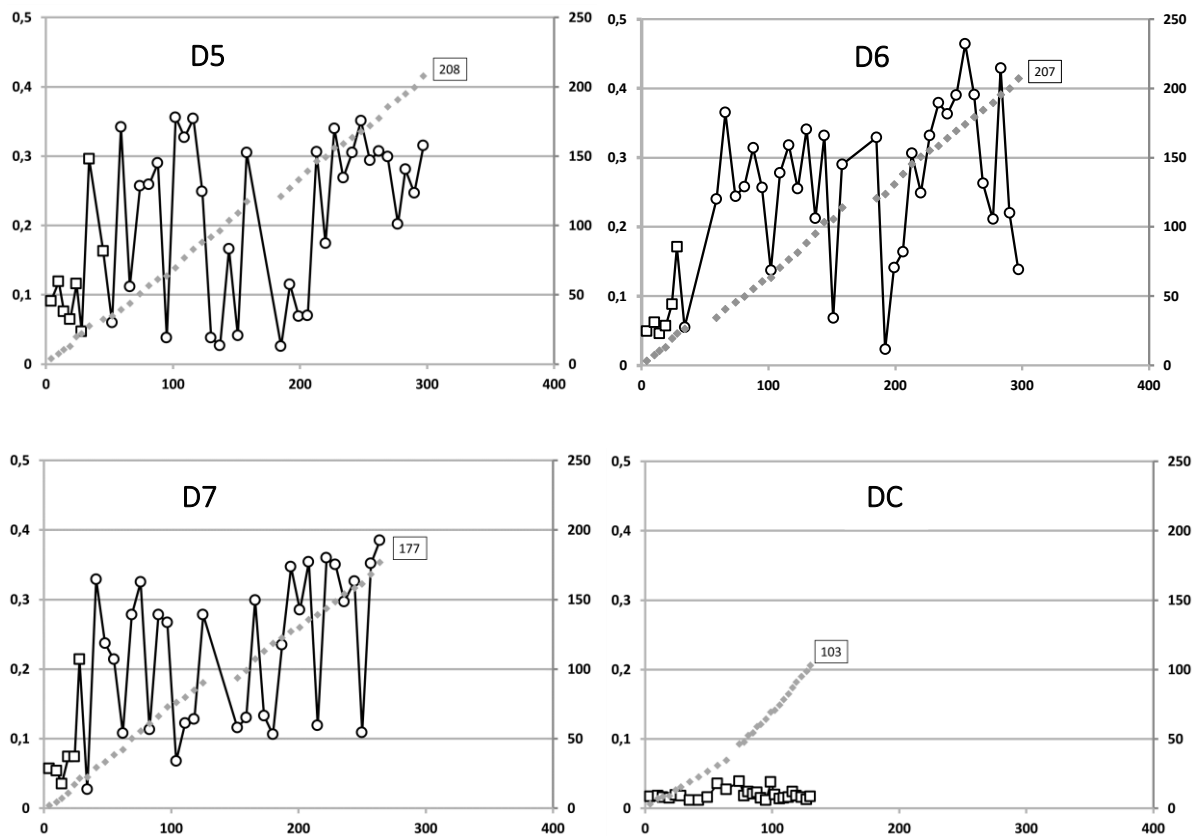


Figure 2.2 Suivi de la densité optique au cours des transferts de l'expérience d'évolution expérimentale

La densité optique (DO) à 600nm a été mesurée juste avant un transfert (axe des ordonnées à gauche) et est représentée en fonction du temps de culture en jours (en abscisse). A1 à A3 représentent les cultures indépendantes issues de 105A, la souche parentale AM1 portant le plasmide pJM105. AC est un témoin négatif de la même souche AM1 portant le plasmide pCM62. D1 à D7 représentent les cultures indépendantes issues de 105D, la souche parentale DM4 portant le plasmide pJM105, et DC la souche DM4 portant le plasmide vide pCM62. Les mesures sont indiquées par des carrés pour la phase 1 de l'évolution expérimentale, et par des cercles pour la phase 2. Le nombre total de générations (losanges gris) est indiqué en ordonnées du côté droit. Le nombre estimé de générations pour chaque réplicat en fin d'expérience est encadré.

L'évolution des mesures de chlorure au cours de l'expérience est montrée en Figure 2.3. Que ce soit pour les réplicats AM1 ou les réplicats DM4, on observe une corrélation entre les deux paramètres, suggérant que l'évolution de la croissance constatée au cours de l'expérience est bien due à la déshalogénéation et à l'utilisation du CH_3Cl .

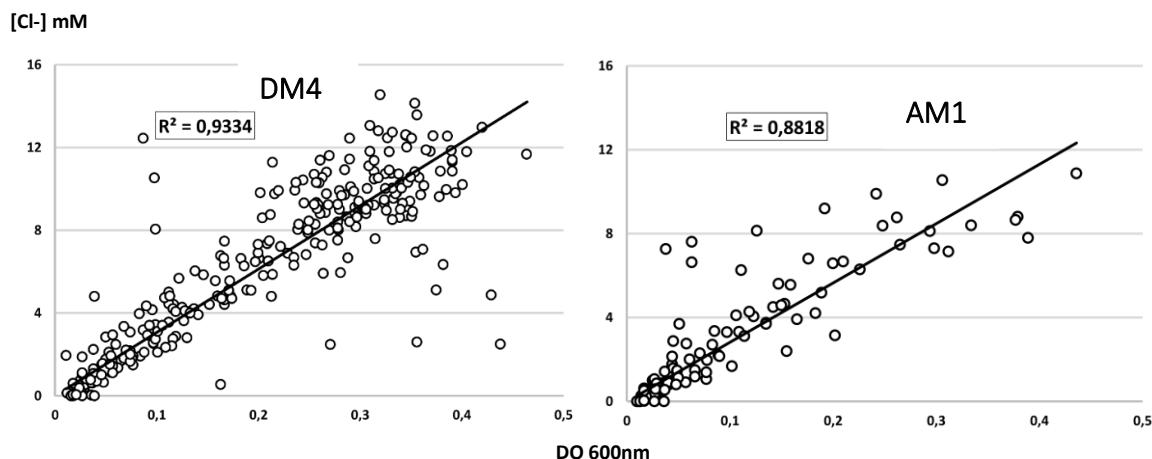


Figure 2.3. Corrélation entre la concentration en ions chlorure et la DO à 600 nm au cours de l'expérience d'évolution expérimentale

La concentration en ions chlorure (Cl^-) (en ordonnées) a été mesurée avant transfert et mise en relation avec la densité optique à 600 nm (DO_{600}) correspondante (en abscisse), pour des cultures indépendantes issues de DM4 (à gauche) et d'AM1 (à droite). Dans les deux cas, le R^2 a été déterminé et il s'approche de 1.

Déterminer quand arrêter les transferts pour isoler des souches évoluées est un point délicat. Le choix a été fixé à quelque 200 générations, sur la base de l'expérience de Michener (Michener et al. 2014b). Idéalement, l'objectif serait d'atteindre la même DO_{600} à la fin de chaque transfert, indiquant que le potentiel de croissance de la culture a été atteint. En réalité, la densité cellulaire finale mesurée a varié au cours des cultures successives (Figure 2.2) bien que le pas de temps entre les mesures reste constant. Ceci avait déjà été relevé lors d'une expérience d'évolution expérimentale précédente (Michener et al. 2014b). La variation de densité cellulaire à chaque transfert pourrait être due à une population non-homogène dans la culture de par la non-fixation de mutations bénéfiques qui supplantent toutes les autres. Cela pourrait être dû en partie à un phénomène souvent rapporté dans les études d'évolution expérimentale : l'interférence clonale (Barrick & Lenski 2013), impliquant la compétition entre sous-populations ayant acquis des mutations bénéfiques de façon indépendante au sein d'une même culture. Il pourrait y avoir également l'apparition et la disparition continue au cours des transferts successifs de populations du fait de la mise en place de coopérations. Par exemple, des bactéries issues de deux sous-populations différentes peuvent s'entraider dans la dégradation du CH_3Cl , lorsqu'un des deux partenaires ne peut faire fonctionner la voie *cmu* que jusqu'à un certain point en libérant un intermédiaire métabolique dans le milieu capté par un deuxième partenaire qui effectue d'autres transformations pour sa croissance. Si un

des deux partenaires disparaît, l'autre ne pourra plus se développer. De manière similaire, des relations antagonistes entre deux sous populations, l'une qui inhibe ou limite la croissance d'une autre, peuvent également être envisagées. J'ai donc considéré qu'a priori, les cultures successives de l'expérience d'évolution expérimentale avaient pu générer une grande diversité de populations pouvant suivre des chemins évolutifs différents. A noter que même si l'expérience a été conduite pour 200 générations environ, on a observé qu'après quelques transferts déjà (5 à 7 pour les réplicats DM4), la variation de la DO_{600} mesurée s'est toujours située entre 0,1 et 0,4, et ce donc dès le début de la phase 2. Ainsi, les mutations bénéfiques déterminantes se seraient peut-être produites assez tôt déjà au cours de l'expérience, sans beaucoup de marge de progression par la suite. L'isolement de souches auquel j'ai ensuite procédé avait entre autres pour objectif d'aborder ces questions par le phénotypage des souches isolées obtenues et leur séquençage.

2.2.3. Diversité des souches isolées après évolution expérimentale

Un total de 162 souches a été isolé de l'ensemble des réplicats des cultures de l'évolution expérimentale à l'issue de la phase 2. Pour chaque réplicat biologique indépendant, les génomes d'au moins une souche isolée par « Extinction Dilution » (notée E, Figure 2.4), ainsi que trois souches évoluées isolées directement par étalement sur milieu gélosé solide (notées S, Figure 2.4), ont été séquencés. J'ai ainsi pu comparer le profil génétique de souches isolées indépendamment d'une même culture. Finalement, un total de 13 souches évoluées (4 souches évoluées AM1 et 9 souches évoluées DM4) a été séquencé par technologie Illumina paired-end 150 pb.

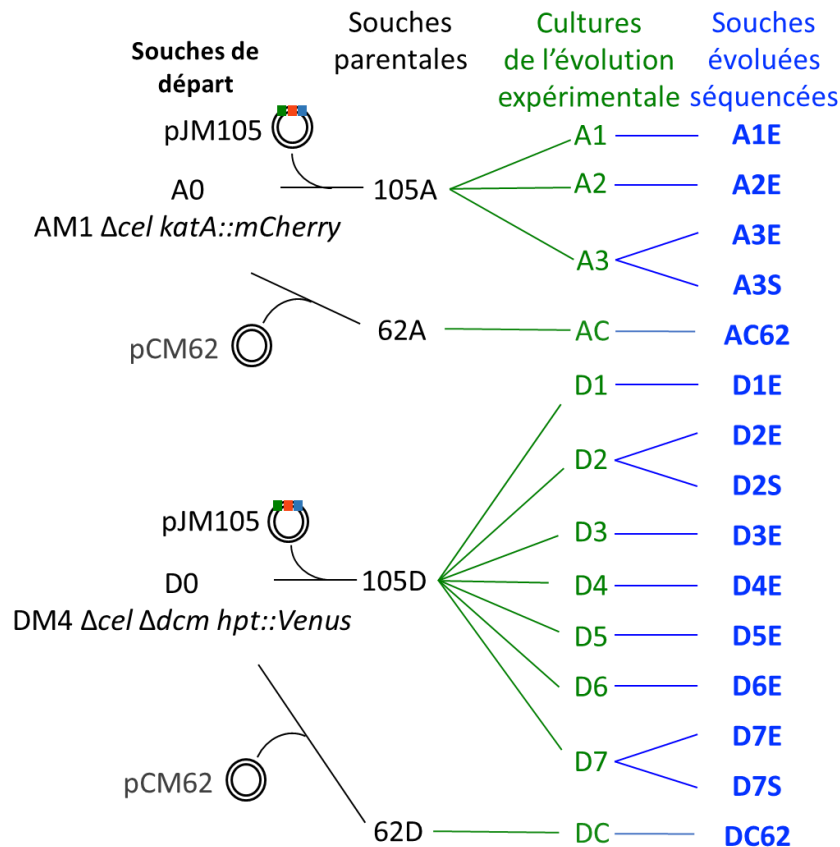


Figure 2.4 Schéma récapitulatif de l'origine des souches évoluées séquencées

L'expérience d'évolution expérimentale a été menée pour 10 cultures indépendantes (en vert) issues de deux souches parentales (en noir) portant le plasmide pJM105. Au total, 13 souches évoluées isolées (en bleu) ont été choisies pour un séquençage de leur génome en plus des souches parentales 105A et 105D. La méthode d'isolement est notée soit E (pour Extinction Dilution) soit S (pour étalement solide direct) (voir Matériel et Méthodes). Pour chacun des contextes génétiques, une culture témoin de la souche portant le plasmide vide pCM62 a permis d'obtenir une souche de contrôle de contexte évolué dont le génome a également été séquencé (en gris).

En plus des 13 souches évoluées possédant pJM105, le génome de deux souches possédant pCM62 et issues des cultures témoin, appelées AC62 et DC62, a également été séquencé, afin de mettre en évidence d'éventuelles mutations qui seraient apparues dans le contexte de l'expérience et indépendamment de l'utilisation du CH_3Cl .

La culture avec le CH_3Cl des 13 souches isolées a confirmé leur capacité à utiliser ce composé comme unique source de carbone (Figure 2.5). Cette croissance est bien due à la dégradation du CH_3Cl comme l'indique la présence d'ions chlorure dans le milieu du fait de la réaction de déshalogénéation de la voie *cmu*. L'incapacité des souches de départ cultivées de la même

manière à utiliser le CH_3Cl comme substrat de croissance a été à nouveau confirmée alors qu'elles sont capables de croître avec le succinate (résultats non montrés).

Ainsi, l'expérience d'évolution expérimentale a permis d'obtenir des souches évoluées ayant acquis la capacité de croître avec le CH_3Cl comme unique source de carbone et d'énergie par utilisation de la voie *cmu*. Le temps de génération des souches évoluées a été déterminé en milieu liquide et comparé à celui des souches de départ lorsque ces dernières étaient cultivées avec du CH_3Cl , du succinate et du méthanol (Tableau 2.1).

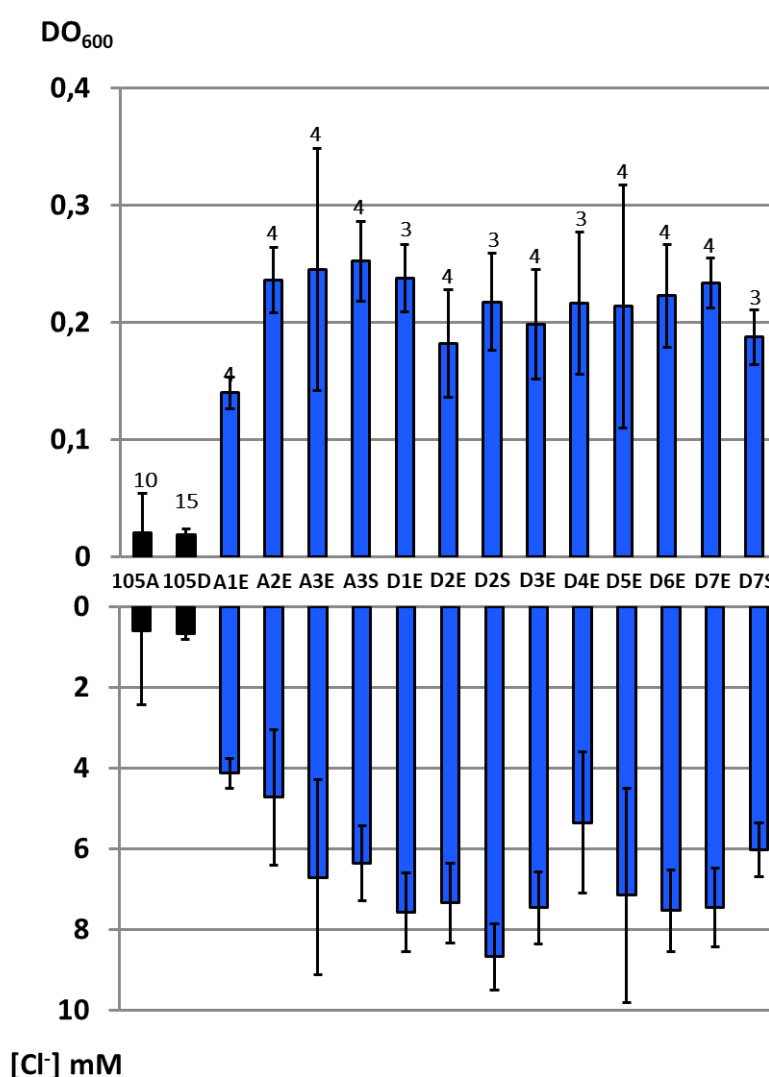


Figure 2.5 Mesure de la DO_{600} et de la production d'ions Cl^- après croissance des souches évoluées isolées avec du CH_3Cl

En haut, la DO_{600} mesurée après culture des deux souches de départ (en noir) et des différentes souches évoluées (en bleu) après 12 jours de culture. En bas, la concentration en ions Cl^- du surnageant des mêmes cultures. Le nombre de réplicats biologiques pour chaque souche est indiqué en haut.

Tableau 2.1 Comparaison des capacités de croissance des souches évoluées

Souche	Temps de génération ^a (en h) selon la source de carbone		
	CH ₃ Cl	Succinate	Méthanol
	15 mM	10 mM	40 mM
Souches parentales			
105A	NA ^b	4,7 ± 0,5	6,0 ± 1,4
105D	NA	4,7 ± 0,2	5,5 ± 0,7
Souches évoluées			
A1E	30,8 ± 1,4	4,7 ± 0,7	6,0 ± 0,6
A2E	26,5 ± 2,5	4,1 ± 0,0	5,4 ± 0,6
A3E	14,1 ± 0,4	5,2 ± 0,8	NA
A3S	15,3 ± 1,3	4,5 ± 0,1	NA
D1E	11,6 ± 0,2	4,8 ± 0,7	NA
D2E	17,2 ± 0,6	4,0 ± 0,2	5,3 ± 0,7
D2S	16,4 ± 0,3	5,0 ± 0,4	NA
D3E	13,2 ± 1,6	6,2 ± 1,8	NA
D4E	23,4 ± 1,9	3,7 ± 0,1	5,8 ± 0,6
D5E	16,4 ± 0,1	5,5 ± 0,5	NA
D6E	16,1 ± 1,4	4,7 ± 0,5	NA
D7E	11,8 ± 0,7	5,3 ± 0,2	NA
D7S	16,2 ± 2,6	4,1 ± 0,1	6,8 ± 0,2

^a Cultures réalisées en milieu liquide M3C10 avec ajout de tétracycline 12,5 µg/mL ; ± écart type standard ; 3 réplicats biologiques sauf pour D4E (4), D5E (2) et A1E et A2E (2) avec du CH₃Cl

^b non-applicable: pas de croissance observée

Le temps de génération avec le CH₃Cl pour les souches évoluées varie de 11,6 ± 0,2 h à 30,8 ± 1,4 h, laissant entrevoir des chemins évolutifs différents indépendants du contexte génétique des souches AM1 ou DM4 dont elles sont issues. Ainsi, des souches issues des deux contextes parentaux peuvent croître plus rapidement avec le CH₃Cl comme unique source de carbone et d'énergie (D1E, D7E, D3E, A3E) alors que d'autres ne peuvent le faire que très lentement (A1E, A2E, D4E).

Ces résultats contrastent avec les temps de génération assez proches observés avec le succinate, tant pour les souches évoluées que pour les souches de départ. Ceci suggère que

les mutations acquises au cours de l'expérience d'évolution expérimentale affectent uniquement la croissance avec le CH₃Cl et non le métabolisme central. Pour la croissance avec le méthanol, en revanche, la situation est différente. La majorité des souches évoluées ont entièrement perdu la capacité de croissance avec ce composé, ce qui sera discuté dans la suite du Chapitre. Les souches pouvant toujours se développer avec le méthanol présentent des temps de génération similaires à ceux des souches de départ, confirmant que l'hétérogénéité des phénotypes de croissance observés est spécifique au CH₃Cl.

2.2.4. L'intégrité du plasmide pJM105 est affectée chez les souches évoluées en contexte A0 seulement

Le curage et le remplacement du plasmide pJM105 dans les souches évoluées isolées par le plasmide pJM105 d'origine a permis de rapidement répondre à la question de savoir si des mutations majeures étaient apparues sur pJM105. Le plasmide pJM105 se perd facilement en contexte *M. extorquens* lorsque l'on cultive les bactéries en absence de l'antibiotique tétracycline à laquelle le plasmide pJM105 confère la résistance, sauf si la séquence du plasmide est conservée indépendamment de ses mécanismes de réplication par suite d'un évènement d'intégration dans un réplicon plus stable (chromosome ou plasmide endogène). Pour l'ensemble des souches évoluées dérivant de D0, j'ai observé une forte instabilité du plasmide dès la première culture liquide sans pression de sélection (perte plus ou moins importante, Tableau 2.2). Dans tous les cas, la réintroduction de pJM105 dans les souches évoluées curées dérivant de D0 a rétabli leur croissance avec le CH₃Cl. J'ai ainsi fait l'hypothèse qu'il n'y avait pas eu de mutations significatives au niveau du plasmide pJM105 dans les souches évoluées dérivant de DM4, et que les mutations permettant la croissance avec le CH₃Cl sont portées par le génome.

Tableau 2.2 Souches évoluées curées : phénotype et conditions d'obtention

Souche curée ^a	Croissance avec le CH ₃ Cl après réintroduction de pJM105 ^b	Souche évoluée d'origine		Ratio de colonies curées / colonies repiquées (%)	Nombre de colonies repiquées ^e
A1EL	-	A1E	Après 4 transferts ^d	1	145
A2EL	-	A2E		2	145
A3EL	ND ^c	A3E		20	145
A3SL	+	A3S		28	145
D1EL	+	D1E	Après 1 seul transfert	46	48
D2EL	+	D2E		52	48
D2SL	+	D2S		42	48
D3EL	+	D3E		42	48
D4EL	+	D4E		12	48
D5EL	+	D5E		78	36
D6EL	+	D6E		61	36
D7EL	+	D7E		70	36
D7SL	+	D7S		52	31

^a Tc^S et incapable de croître avec le CH₃Cl; se développe avec le succinate

^b Tc^R et capable de croître avec le succinate

^c ND = Non Déterminé

^d 4 cultures successives sans pression de sélection ont été nécessaires pour observer des clones Tc^S de la souche ayant perdu la capacité de croître avec le CH₃Cl

^e après culture liquide M3C10 avec succinate

Au contraire, pour obtenir des versions curées des souches évoluées dérivant d'AM1, plusieurs cultures successives (4^{ème} transfert) en milieu succinate sans tétracycline ont été nécessaires pour isoler des souches Tc^S ayant potentiellement perdu le plasmide pJM105 (Tableau 2.2). La proportion des souches Tc^S par rapport aux souches Tc^R porteuses du plasmide est de l'ordre de 2% pour A2E et A1E même après 4 transferts, proportion bien en deçà des souches évoluées isolées issues de la culture A3 (souches A3S et A3E). La stabilité du phénotype Tc^R des souches évoluées dérivées de 105A suggèrent que des événements d'intégration de pJM105 au sein du génome ont eu lieu dans des réplicats biologiques indépendants de cultures de l'expérience d'évolution expérimentale. Pour A3SL, la réintégration de pJM105 a permis de restaurer la croissance avec le CH₃Cl comme unique source de carbone. Ceci n'a pas

été le cas pour A2EL et A1EL, révélant ici une différence nette entre les différentes souches (néanmoins d'autres tests de croissance seraient nécessaires pour le confirmer). Peut-être les mutations présentes sur le plasmide de A1E et A2E sont-elles importantes pour la croissance de ces souches avec le CH₃Cl. On ne peut pas non plus exclure que d'autres événements (comme des réarrangements chromosomiques) aient pu avoir lieu lors du curage dans ces deux souches ayant impacté leur croissance avec le CH₃Cl. En conclusion, dans les souches évoluées dérivant d'AM1 le plasmide pJM105 a vraisemblablement été intégré au niveau de leur génome. Comme décrit ci-après, les résultats du séquençage ont permis de mettre en lumière les mutations apparues sur pJM105 ainsi que son nombre de copies dans la cellule.

2.2.5. Séquençage du génome des souches évoluées et analyse de la couverture relative des réplicons

La couverture de séquençage est très élevée pour l'ensemble des souches séquencées notamment des souches évoluées avec une couverture de plus de 600X, ce qui, avec le score qualité (Tableau 2.3) a d'autant plus renforcé l'analyse des mutations qui sont apparues par suite de l'évolution expérimentale.

Tableau 2.3 Données du séquençage Illumina paired-end 150 pb des génomes des souches étudiées

ADN génomique	Séquences brutes (en Mb)	Nombre de lectures (paired-end) x10⁶	Score qualité >=Q30 (%) ^a	Couverture de séquençage (X)^b
Souches parentales				
105A	783,6	5,2	90,2	131
105D	825,2	5,5	90,7	138
Souches évoluées témoin				
AC62	748,7	5,0	90,4	125
DC62	613,5	4,1	89,8	102
Souches évoluées				
A1E	3628,1	24,0	89,0	605
A2E	3740,3	24,8	89,0	623
A3E	3561,5	23,6	89,2	594
A3S	3797,3	25,1	89,3	633
D1E	3556,5	23,5	90,0	593
D2E	3440,3	22,8	90,2	573
D2S	3968,7	26,3	90,2	661
D3E	3893,6	25,8	90,5	649
D4E	4093,6	27,1	90,1	682
D5E	4084,1	27,0	90,1	681
D6E	4366,8	28,9	90,5	728
D7E	4101,0	27,2	91,7	684
D7S	4202,9	27,2	92,1	700

^a Q30 = associé à un taux d'erreur pour une base de 1 %

^b sur la base d'une taille de génome de 6 Mb

J'ai évalué la couverture relative de chacun des réplicons par rapport au chromosome, qui a servi d'approximation à l'estimation du nombre de copies des réplicons de pJM105 pour les souches évoluées. Il s'est agi d'une part de déterminer si l'évolution expérimentale a impacté le nombre de copies des réplicons. Par ailleurs, pour confirmer l'hypothèse d'un évènement d'intégration de pJM105 chez les souches évoluées AM1, la couverture de pJM105 devrait être la même que celle du réplicon dans lequel s'est déroulée l'intégration. Les résultats obtenus sont renseignés dans le Tableau 2.4.

Tableau 2.4 Couverture relative des réplicons des souches évoluées par rapport au chromosome

Couverture relative des réplicons en contexte AM1 ^a						
ancêtre		Souches évoluées				
réplicon	105A	A1E	A2E	A3E	A3S	AC62
META1	1,00	1,00	1,00	1,00	1,00	1,00
META2	0,99	0,76	0,86	0,57	0,58	0,75
p1META	1,81	0,39	0,20	0,33	0,33	0,40
p3META	2,55	0,43	0,26	0,44	0,41	0,46
p2META	2,39	0,40 ^c	0,15 ^c	1,68 ^c	1,24 ^c	0,40
pJM105	3,25	0,36 ^c	0,10 ^c	1,63 ^c	1,11 ^c	0,94 ^d

Couverture relative des réplicons en contexte DM4 ^b											
ancêtre		Souches évoluées									
réplicon	105D	D1E	D2E	D2S	D3E	D4E	D5E	D6E	D7E	D7S	DC62
METDI	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	0,39	1,00
p1METDI	0,55	0,25	0,26	0,30	0,51	0,38	0,34	0,49	0,41	0,39	0,47
p2METDI	1,60	0,52	0,53	0,53	0,50	0,80	0,41	0,77	0,61	0,60	1,20
pJM105	3,08	1,50	1,51	1,36	1,25	2,90	1,16	2,95	2,39	2,34	2,36 ^d

^a Couverture du réplicon / couverture du chromosome META1 pour chaque souche

^b Couverture du réplicon / couverture du chromosome METDI1 pour chaque souche

^c Liens paired-end démontrés entre p2META et pJM105

^d pCM62, le plasmide dans lequel les gènes *cmu* ont été clonés pour donner le plasmide pJM105 (Michener et al 2014)

Pour les souches évoluées possédant pJM105, une réduction globale de l'ensemble des réplicons par rapport aux souches de départ est observée pour les souches dérivant d'AM1 comme celles dérivant de DM4. Une réduction similaire est également observée chez les souches AC62 et DC62, peut-être liée au contexte de l'expérience indépendamment de l'utilisation du CH₃Cl. Chez les souches dérivant d'AM1, une même couverture relative pour pJM105 et pour le plasmide p2META natif d'AM1 est un fait marquant qui pourrait indiquer que le plasmide pJM105 et le plasmide endogène p2META ont formé un coïntégrat chez les

souches évoluées dérivant d'AM1 qui pourraient ainsi désormais dépendre de la même machinerie de réplication.

2.2.6. Un Southern Blot qui montre l'intégration de pJM105 dans le génome des souches

Afin de démontrer l'intégration de pJM105 dans le génome des souches évoluées AM1, j'ai réalisé une expérience de Southern Blot afin de mettre en évidence que pJM105 ne se retrouve plus sous forme libre chez les souches évoluées AM1. Des sondes spécifiques de pJM105 vont révéler après hybridation une bande de taille différente selon les cas (Figure 2.6). L'ADN total des différentes souches évoluées dérivant d'AM1, de leurs équivalents curés, ainsi que de la souche parentale 105A ont été digérés soit avec l'enzyme BamHI (Figure 2.6A) soit avec l'enzyme Apal (Figure 2.6B).

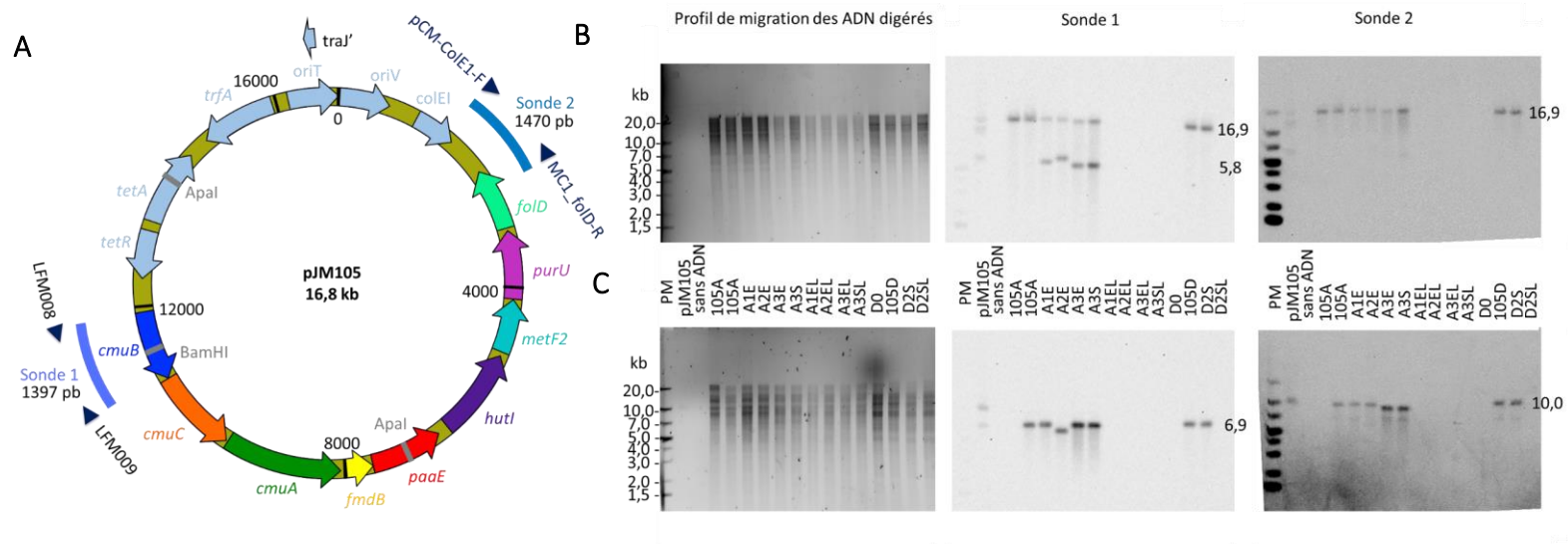


Figure 2.6 Southern Blot mettant en évidence l'intégration du plasmide pJM105 dans le génome des souches évoluées

En A, une représentation du plasmide pJM105, avec les deux sondes ainsi que les sites de restriction utilisés pour ce Southern Blot, avec les sites de reconnaissance des enzymes de restriction utilisées sur pJM105.

Les ADN ont été digérés avec *Bam*HI (B) ou *Apal* (C), séparés par électrophorèse sur gel d'agarose 0,7% (gauche), puis hybridés avec la sonde 1. Après déshybridation, une réhybridation avec la sonde 2 a été effectuée pour la détection des fragments correspondants. Les sondes ont été synthétisées par PCR avec de l'U-11-digoxigénine et les bandes révélées par un anticorps anti-DIG conjugué à une phosphatase alcaline.

BamHI linéarise le plasmide pJM105 par coupure d'un site unique de reconnaissance (position 11465 dans le gène *cmuB*), produisant un fragment linéaire de 16,8 kb. Comme la sonde 1 chevauche cet unique site BamHI, si le plasmide se trouve sous sa forme libre dans la cellule (comme c'est le cas chez la souche 105A, la souche 105D et la souche D2S), une seule bande d'environ 17 kb est attendue par hybridation avec la sonde 1. Comme anticipé, aucune bande n'est détectée dans les souches curées (A1EL, A2EL, A3EL et A3SL). Une bande proche de 17 kb est visible pour l'ensemble des souches possédant pJM105, avec en plus, une bande d'environ 6 kb chez les souches évoluées en contexte AM1, indiquant qu'un autre site BamHI associé à la séquence du plasmide pJM105 est apparu dans ces souches. Ce résultat vient en appui de l'hypothèse d'une intégration du plasmide pJM105 dans le génome des souches évoluées à partir de 105A. À environ 5 kb du site de coupure BamHI du plasmide pJM105 on trouve le gène *trfA*, qui pourrait se situer à proximité d'un site BamHI apporté par un autre réplicon que pJM105. Le profil d'hybridation des ADN digérés par Apal avec la sonde 1 (non coupée par BamHI) ne donne qu'une seule bande proche de 6,9 kb pour les souches évoluées issues de 105A et la souche D2S issue de 105D. Cela indique que la séquence du plasmide pJM105 n'est pas modifiée dans cette région (hormis pour la souche A2E dont la taille de la bande est légèrement plus petite).

Avec l'enzyme de restriction Apal, la digestion du plasmide pJM105 génère deux fragments de 10kb et 7kb correspondant aux deux sites de coupure dans pJM105 au niveau des gènes *tetA* et *paaE* (le premier révélé avec la sonde 1 et le deuxième avec la sonde 2). En admettant que le site de coïntégration du plasmide pJM105 se situe au niveau du gène *trfA*, l'hybridation avec la sonde 1 ne devrait révéler qu'une seule bande de 7 kb chez les souches évoluées dérivant d'AM1 ainsi que celles ayant le plasmide libre, car le gène *trfA* est en dehors de ce fragment de 7 kb. La taille des bandes révélée est conforme à cette hypothèse. Pour la sonde 2, on s'attendrait à une bande de 10 kb pour le plasmide libre et peut-être à un fragment plus grand pour les souches évoluées AM1 contenant la région du gène *trfA* où se serait effectuée l'intégration. Une bande avoisinant 10 kb a bien été détectée pour le pJM105 libre, mais aussi pour les souches évoluées dérivant d'AM1, suggérant la présence d'un site Apal en dehors du plasmide pJM105 à proximité du lieu d'insertion de pJM105. Pour les souches curées, on confirme aussi la perte du pJM105 avec la sonde 2.

En définitive, les données résultant des hybridations avec les deux sondes sont en accord avec la possible intégration de pJM105 dans le génome des souches évoluées issues du contexte 105A.

2.2.7. Identification des sites de coïntégration du plasmide pJM105 dans le plasmide p2META chez les souches évoluées issues de 105A

Pour confirmer l'hypothèse que pJM105 s'est bien intégré dans p2META, des liens paired-end entre le contig représentant pJM105 et un autre contig ont été recherchés dans les données brutes de séquençage Illumina. De tels liens ont uniquement été détectés pour des séquences correspondant à pJM105 avec pour bordure le gène *trfA* et la séquence d'insertion, ISMex25 retrouvée en deux exemplaires sur le plasmide natif p2META dans la souche d'origine. Ces éléments sont des éléments en faveur de l'intégration de pJM105 au niveau de *trfA* et avec un site différent selon les souches évoluées (voir Figure 2.8) au sein du génome des souches évoluées AM1, vraisemblablement au niveau du réplicon p2META.

Pour confirmer le site de coïntégration entre pJM105 (au niveau du gène *trfA*, un gène essentiel pour la réplication de pJM105) et le réplicon p2META (au niveau de l'ISMex25_1), des PCR ont été réalisées en ciblant ces séquences. Différents jeux d'amorces ont été tout d'abord validés en contexte non-évolué. Pour le couple d'amorces ciblant le gène *trfA* (amorces LFM010 et LFM011 ; Figure 2.7B), l'amplification d'un produit de taille supérieure à celui de la souche parentale 105A a été observée pour l'ensemble des souches évoluées (2 kb contre 1,1kb pour 105A ; Figure 2.7B). Le produit d'amplification chez A3E a été séquencé et des homologues de séquences (BLAST) ont été recherchés dans le génome complet d'AM1 : 100% d'identité de séquence ont été trouvés avec la séquence d'ISMex25. Pour le couple d'amorces ciblant une amplification de l'allèle sauvage d'ISMex25_1, une bande de 1,1 kb a été obtenue pour la souche parentale 105A, mais pas pour les souches évoluées, comme attendu pour une intégration à proximité de cette IS empêchant une amplification avec ces amorces du fait de la trop grande taille du produit attendu qui en résulte (Figure 2.7B). Enfin, les régions flanquantes entre les plasmides pJM105 et p2META ont été amplifiées en utilisant une combinaison d'amorces LFM010-LFM012 (obtention d'un produit de 1,7 kb pour A1E et A2E et de 2,5 kb pour A3E/A3S), et LFM011-LFM013 (obtention d'un produit de 2,2 kb pour A1E et A2E et de 1,4 kb pour A3E/A3S). Une amplification a été obtenue uniquement pour les

souches évoluées AM1. L'analyse des séquences de ces bandes a permis de construire un modèle du coïntégrat (Figure 2.7A) localisé au niveau du gène *trfA* de pJM105 (interrompu de ce fait) et de l'ISMex25 du plasmide p2META. En plus de la formation du coïntégrat entre pJM105 et p2META, une séquence d'insertion de type ISMex4 s'est interposée entre le gène *tetR* et en amont de gène *cmuB* chez la souche évoluée A2E. Je l'ai localisée grâce à une PCR avec les amorces *cmuB105-tetR105* (dont le produit est normalement de 0,7 kb mais qui était de plus de 2 kb pour A2E). ISMex4 est une séquence d'insertion présente en 7 exemplaires sur le chromosome de la souche AM1. Une de ces 7 copies serait d'ailleurs apparue au cours de la lente domestication de la souche (Carroll et al. 2014a). Elle a été également placée sur le coïntégrat de la souche A2E.

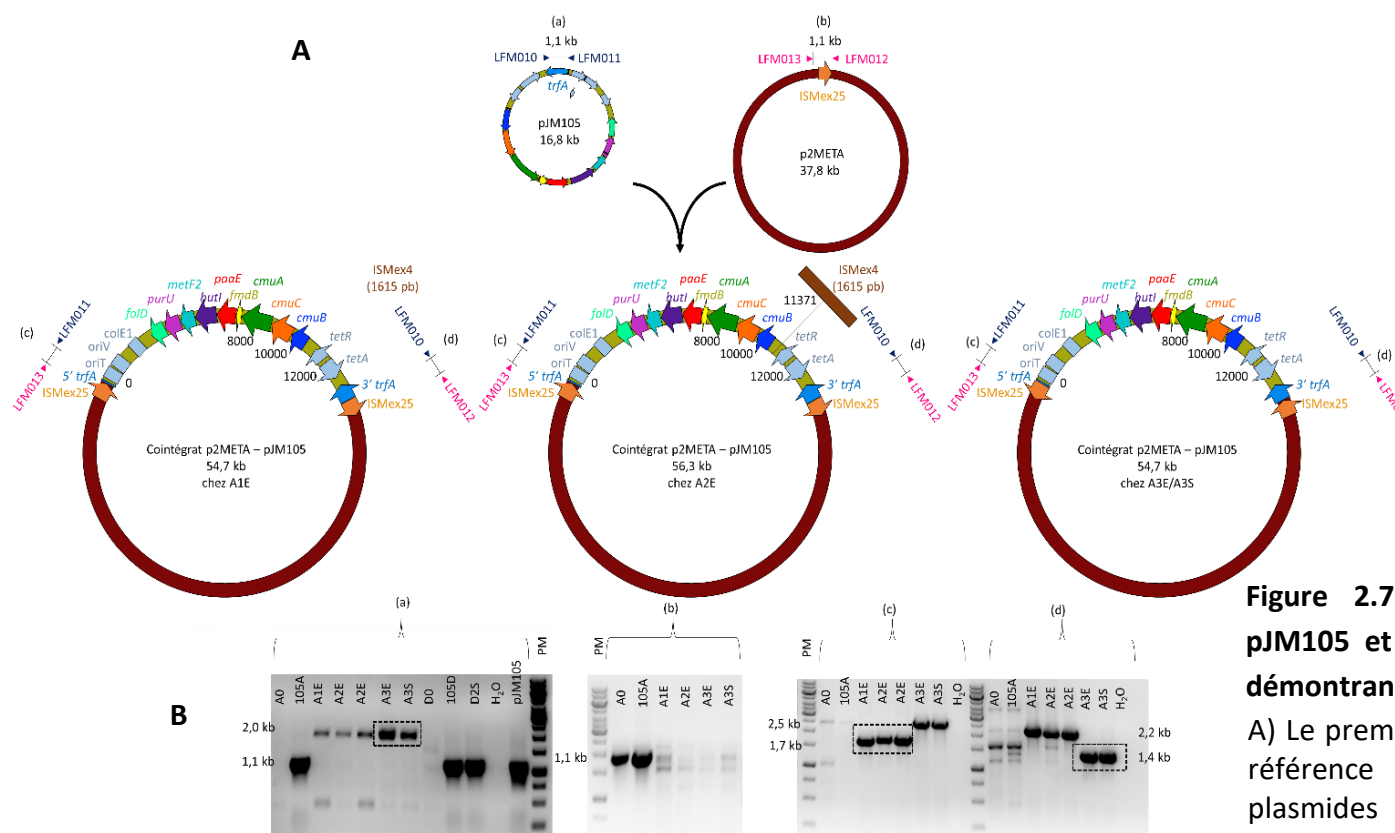


Figure 2.7 Modèle du coïntégrat entre pJM105 et p2META et données de PCR le démontrant

A) Le premier nucléotide est placé selon la référence Genbank de pJM105. Les plasmides ne sont pas représentés à l'échelle.

B) Les PCR qui ont permis de déterminer les allèles sauvages des gènes *trfA* sur pJM105 et de l'ISMex25 sur p2META (PM=Marqueur de taille).

(a) avec les amorces ciblant *trfA* sur pJM105 (LFM010-LFM011)

(b) avec les amorces ciblant ISMex25 sur p2META (LFM012-LFM013)

(c) avec les amorces LFM010-LFM012

(d) avec les amorces LFM011-LFM013

	Produits d'amplification PCR (taille en kb)			
ADN matrice	(a) LFM010-LFM011	(b) LFM012-LFM013	(c) LFM010-LFM012	(d) LFM011-LFM013
A0	1,1	1,1	bandes aspécifiques	bandes aspécifiques
105A	1,1	1,1	bandes aspécifiques	bandes aspécifiques
A1E	2,0	bandes aspécifiques	1,7	2,2
A2E	2,0	bandes aspécifiques	1,7	2,2
A3E/A3S	2,0	bandes aspécifiques	2,5	1,4

2.2.8. Un site d'interruption de *trfA* différent entre les souches évoluées dérivant d'AM1

Le gène *trfA* n'a pas été interrompu à la même position pour chacune des souches évoluées : à l'extrémité 5' du gène *trfA* chez A3E et A3S, et à l'extrémité 3' chez A1E et A2E (voir Figure 2.8). Le gène *trfA* est un gène essentiel pour la réplication du plasmide (Lin et Helinski, 1992). Ainsi, il est probable que le coïntégrat dépende de la machinerie de réplication de p2META, ce qui explique qu'on retrouve la même couverture relative entre pJM105 et p2META chez les souches évoluées alors qu'en contexte 105A, pJM105 est en plus grand nombre de copies que p2META (environ 3 pour pJM105 et 2 pour p2META, voir Tableau 2.4). Outre la réduction de leur coût de maintien dans la cellule, un plus faible nombre de copies pourrait peut-être impacter l'expression des gènes *cmu*, ce qui sera discuté plus en détail dans le Chapitre 4.

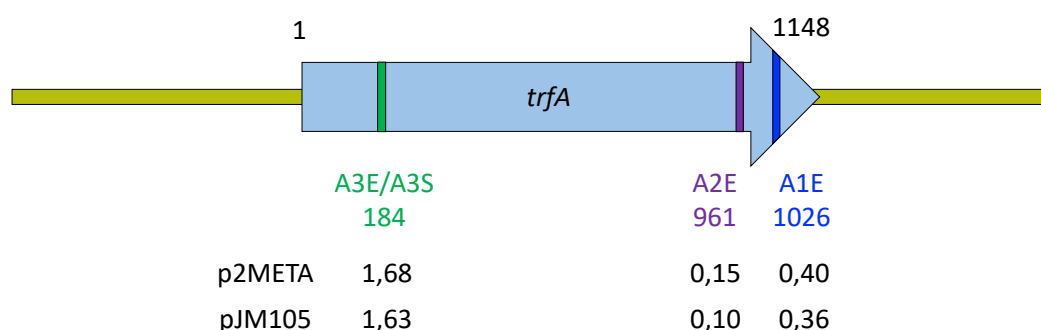


Figure 2.8 La différence d'insertion d'ISMex25 au niveau de *trfA* chez les souches évoluées issues d'AM1

Les sites d'interruption de *trfA* (gène de 1848 kb) sont indiqués pour chacune des souches évoluées dérivant d'AM1. Les couvertures relatives de p2META et de pJM105 (par rapport au chromosome) sont également indiquées (issues du Tableau 2.4).

On note une différence marquée entre A1E et A2E et A3E/A3S. Les deux premières souches présentent toutes deux une couverture relative très faible de p2META et de pJM105. Chez A3E/A3S, *trfA* a été interrompu dans la partie N-terminal de la protéine TrfA. Or, il a été mis en évidence par le passé que le domaine C-terminal de la protéine TrfA porte le domaine catalytique de la protéine (Lin & Helinski 1992). La couverture relative de p2META-pJM105 est supérieure chez A3E/A3S par rapport à A1E et A2E. Une IS (ISMex25) accolée à cette partie 3' du gène pourrait apporter un promoteur permettant d'exprimer la partie 3' de *trfA* de

A3E/A3S et permettrait l'expression de *trfA* à partir d'un codon ATG interne à *trfA* avec conservation du cadre de lecture. Si c'est le cas, une protéine TrfA tronquée d'une partie 3' sera produite chez A3E/A3S avec une activité résiduelle qui pourrait permettre une réplication du coïntégrat.

Le séquençage et l'analyse des mutations qui en a découlé a permis de mettre en évidence la nature de certaines des différences entre les souches évoluées, comme décrit dans la section suivante.

2.3. Répertoire de l'ensemble des mutations apparues chez les souches évoluées

2.3.1. Des mutations nombreuses et diverses

En comparant les données Illumina des séquences des génomes non-assemblés de chacune des souches (évoluées et parentales), les mutations de type SNP et petites insertions/délétions (indel) acquises par chacune d'entre elles ont été mises en évidence. Les délétions plus larges (>6 nt) ont été mises en évidence en identifiant les bornes de réarrangement via les gènes dont la couverture de séquençage était très faible ou nulle. Les Tableaux 2.5 et 2.6 présentent une synthèse de ces deux analyses et liste pour chacun des gènes référencés chez AM1 et DM4 (nommés d'après leur référence sur MaGe <https://mage.genoscope.cns.fr>) les changements de séquence correspondants.

Tableau 2.5 Inventaire de l'ensemble des mutations retrouvées dans les souches évoluées issues de 105A

Étiquette	Gène	Produit	Position	Taille nt (aa)	Souches évoluées			
					A1E	A2E	A3E	A3S
META1_0055	<i>rpsA</i>	30S ribosomal subunit protein S1	58701-60413	1713 (571)			C>T 60167 D83N	
META1_0329	<i>ftfL</i>	formate-tetrahydrofolate ligase	340743-342416	1674 (558)			G>T 342037 T127N	G>T 342037 T127N
META1_tRNA49	-	Met tRNA	601764-601840	77 (NA)	T>C 601846 (intergenic)			
META1_2420	<i>clpX</i>	ATP-dependent Clp protease ATP-binding subunit clpX	2543598-2544869	1272 (424)	C>A 2544492 P299T			
META1_2510	<i>petB</i>	ubiquinol-cytochrome c reductase complex, cytochrome b subunit	2645366-2646658	1293 (431)			A>G 2645457 H31R	A>G 2645457 H31R
META1_2897	-	protein of unknown function	3020145-3020981	837 (279)				C>A 3020616 R158S
META1_3092	-	putative MgtC family protein	3232518-3233198	681 (227)				T>G 3232633 L39R
META1_4441	-	putative hybrid histidine kinase with multiple PAS and a response regulator receiver domain	4543186-4545513	2328 (776)		A>T 4544663 V384D		
META1_4919	-	transcriptional regulator (Ros/MucR family)	5046789-5047319	531 (177)	G>C 5047105 S72W			
META1_5170	-	conserved protein of unknown function	5323086-5323625	540 (180)			G>C 5323194 C144W	G>C 5323194 C144W
META2_0008	-	putative beta-ketoacyl synthase	8472-15308	6837 (2278)			ΔG 9133 fs (A223L 288*)	ΔG 9133 (fs A223L- 288*)
ALS39161.1	pJM105-p2META cointegrate	replication initiation protein	14788-15936	1149 (383)	14910 <i>trfA</i> :: ISMex25 14132	14975 <i>trfA</i> ::ISMex25 14132	15752 <i>trfA</i> ::ISMex25 14132	<i>trfA</i> 15752::ISMex25 14132
-	Intergenic <i>cmuB-tetR</i>	-	-	-	wt	Ins ISMex4 (12371-12373)	wt	Wt
ALS39166.1	<i>cmuA</i>	Methyltransferase/corrinoid binding protein CmuA	8007-9851	1845 (614)	A>C 8975 S293A	A>G 9754 M33T	G>A 9751 T34I	G>A 9751 T34I

Tableau 2.6 Inventaire de l'ensemble des mutations retrouvées dans les souches évoluées issues de 105D

Étiquette	Gène	Produit	Position	Taille nt (aa)	Souches évoluées								
					D1E	D2E	D2S	D3E	D4E	D5E	D6E	D7E	D7S
METDI0025	<i>mdcC</i>	malonate decarboxylase acyl carrier protein	28486 – 28794	309 (103)	wt	wt	wt	wt	wt	wt	wt	C>T 28594 R37C	wt
METDI0225	-	putative DNA-binding protein; putative AP2/ERF domain	232539-233249	711 (237)	wt	wt	wt	wt	Δ232492–392332 METDI0225–0426 loss	wt	wt	wt	wt
METDI0294	-	putative phosphate butyryltransferase	281586-282983	1398 (466)	wt	wt	wt	wt	Δ232492–392332 METDI0225–0426 loss	wt	wt	G>A 281880 V99I	wt
METDI0317	-	putative two component response regulator (FixJ domain-like)	299121-299513	393 (131)	wt	wt	wt	wt	Δ232492–392332 METDI0225–0426 loss	wt	wt	wt	Δ298810–307071 METDI0317–0325 loss
METDI0407	-	transposase of ISMdi20 IS630 family (ORF 2)	381426-382001	576 (192)	wt	wt	wt	wt	Δ232492–392332 METDI0225–0426 loss	wt	A>C 381891 G37G	wt	wt
METDI0459	-	putative membrane protein of unknown function	420575-421168	594 (198)	wt	wt	wt	wt	wt	G>T 420663 T169K	wt	wt	wt
METDI0483	<i>ftfL</i>	formate-tetrahydrofolate ligase	443332-445005	1674 (558)	Δ443771–443777 fs (E412A-418*)	wt	G>T 443548 C486*	C>T 444804 G68S	wt	Δ445023–446402 (prom) METDI0483 loss	ΔC 444674 fs (G111A-143*)	C>T 444623 G128D	wt

					Souches évoluées								
Étiquette	Gène	Produit	Position	Taille nt (aa)	D1E	D2E	D2S	D3E	D4E	D5E	D6E	D7E	D7S
METDI0484	-	putative metal-binding protein	445244-445465	222 (74)	wt	wt	wt	wt	wt	Δ445023–446402 METDI0484 loss	wt	wt	wt
METDI0485	-	conserved protein of unknown function; EAL domain	445752-447233	1482 (494)	wt	wt	wt	wt	wt	Δ445023–446402 METDI0485 loss	wt	wt	wt
METDI0589	<i>rpoD</i>	sigma D (sigma 70) factor of RNA polymerase	555680-557680	2001 (667)	wt	wt	wt	wt	wt	wt	C>T 556264 E473K	wt	wt
METDI0871	-	tartrate dehydrogenase/ decarboxylase TtuC (fragment)	804620-805303	684 (228)	wt	wt	Δ804541–826410 METDI071–0905 loss	wt	wt	wt	wt	wt	wt
METDI0892	-	putative transcriptional regulator (Crp/Fnr family) with fused CheY-like response receiver and cyclic nucleotide binding domains	817453-818583	1131 (377)	Δ818423–819257 METDI0892 loss	wt	Δ804541–826410 METDI071–0905 loss	wt	wt	wt	wt	wt	wt
METDI0893	-	putative transcriptional regulator Crp/Fnr family with N-terminal cNMP-binding domain	818853-819566	714 (238)	Δ818423–819257 METDI0893 loss	wt	Δ804541–826410 METDI071–0905 loss	wt	wt	wt	wt	wt	wt

					Souches évoluées								
Étiquette	Gène	Produit	Position	Taille nt (aa)	D1E	D2E	D2S	D3E	D4E	D5E	D6E	D7E	D7S
METDI1566	<i>nuoA</i>	NADH-quinone oxidoreductase chain A	1431801- 1432166	366 (122)	InsC 1432295 (prom) ND	wt	wt	ΔC 1432295 (prom) ND	wt	ΔG 1432353 (prom) ND	wt	ΔC 1432295 (prom) ND	wt
METDI1692	-	protein of unknown function	1556327- 1556569	243 (81)	wt	wt	ΔA 1556329 *241NF*	wt	wt	wt	wt	wt	wt
METDI2155	<i>hemE</i>	uroporphyrinoge n decarboxylase	2041068- 2042180	1113 (371)	wt	G>A 2041204 R46Q	G>A 2041204 R46Q	wt	wt	wt	wt	wt	wt
METDI2542	-	glycosyl transferase	2450607- 2451866	1260 (420)	wt	wt	wt	wt	wt	C>G 2450697 M910I	wt	wt	wt
METDI2949	<i>rpoA</i>	DNA-directed RNA polymerase alpha subunit	2876608- 2877627	1020 (340)	wt	wt	wt	wt	G>T 2876724 L302I	wt	wt	wt	T>A 2876717 E304V
METDI2958	<i>rplR</i>	50S ribosomal protein L18	2883506- 2883868	363 (121)	wt	wt	wt	wt	wt	wt	wt	C>G 2883548 K107N	wt
METDI2963	<i>rplE</i>	50S ribosomal protein L5	2885185- 2885751	567 (189)	C>A 2885691 V21F	wt	wt	wt	wt	wt	wt	wt	wt
METDI3287	-	putative transglycosylase	3218528- 3219673	1146 (382)	G>A 3218948 G141S	wt	wt	wt	wt	wt	wt	wt	wt
METDI3473	<i>treZ</i>	malto- oligosyltrehalose trehalohydrolase	3394691- 3396496	1806 (602)	ΔG 3394909 fs (D74T-90*)	wt	wt	wt	wt	wt	wt	wt	wt
METDI4077	-	conserved exported protein of unknown function	4010584- 4010931	348 (116)	wt	wt	wt	wt	wt	wt	wt	Ins A 4010909 fs (L9P-41*)	wt

					Souches évoluées								
Étiquette	Gène	Produit	Position	Taille nt (aa)	D1E	D2E	D2S	D3E	D4E	D5E	D6E	D7E	D7S
METDI4412	-	ABC transporter periplasmic protein	4328512-4330068	1557 (519)	wt	Δ4329491–4343041 METDI4412–4430 loss	wt	wt	wt	wt	wt	wt	wt
METDI4835	-	protein of unknown function	4724315-4724533	219 (73)	wt	wt	wt	wt	wt	Δ4724335–4741034 METDI4835–4851 loss	wt	wt	wt
METDI4923	-	short-chain dehydrogenase/reductase SDR	4819470-4820213	744 (248)	wt	wt	wt	wt	wt	G>T 4819508 P236T	wt	wt	wt
METDI5012	<i>rpoC</i>	RNA polymerase beta prime subunit	4914553-4918770	4218 (1405)	wt	A>G 4918486 T1312A	wt	wt	wt	wt	wt	wt	wt
METDI5245	<i>ate</i>	putative arginyl-tRNA--protein transferase	5171228-5171974	747 (249)	wt	wt	wt	wt	wt	T>A 5171423 C66S	ΔC 5171864 fs (P213R-23*)	wt	wt
METDI5725	-	putative molybdopterin biosynthesis-related protein	5710202-5711428	1227 (409)	wt	wt	wt	wt	wt	wt	G>C 5711324 V375L	wt	wt
METDI5787	<i>rplS</i>	50S ribosomal protein L19	5773926-5774330	405 (135)	wt	wt	wt	T>G 5774008 V28G	wt	wt	wt	wt	wt
p1METDI0017–p1METDI0042	NA	25 CDS region (including p1METDI40065D)	18125-50907	32782 (NA)	Δ28032–50836 p1METDI0025 – 0042 loss	Δ21653–50835 p1METDI0021 – 0042 loss	Δ18224–50837 p1METDI0017 – 0042 loss	wt	Δ27173–50837 p1METDI0025 – 0042 loss	wt	Δ32744–50836 p1METDI0029 – 0042 loss	Δ23007–50836 p1METDI0022 – 0042 loss	Δ33257–50837 p1METDI0030 – 0042 loss
p1METDI40065D	-	putative glycosyltransferase (family GT2)	40065-41210	1146 (382)	Δ28032–50836 p1METDI0025 – 0042 loss	Δ21653–50835 p1METDI0021 – 0042 loss	Δ18224–50837 p1METDI0017 – 0042 loss	ΔT 40763 fs (S236L-28*)	Δ27173–50837 p1METDI0025 – 0042 loss	wt	Δ32744–50836 p1METDI0029 – 0042 loss	Δ23007–50836 p1METDI0022 – 0042 loss	Δ33257–50837 p1METDI0030 – 0042 loss

					Souches évoluées								
Étiquette	Gène	Produit	Position	Taille nt (aa)	D1E	D2E	D2S	D3E	D4E	D5E	D6E	D7E	D7S
p1METDI54264D	-	putative lipase; SGNH-hydrolase domain	54264- 55397	1134 (378)	wt	wt	wt	wt	wt	G>C 54815 R184S	wt	wt	wt
-	pJM105 intergen <i>cmuB- tetR</i>	-	-	-	wt	wt	A>T 12248	wt	wt	wt	wt	wt	wt

Ins : Insertion

Δ : délétion

prom : promoter

Gras : effet sur la protéine

* : codon stop

fs (A-B*): Changement d'acide

aminé de A jusqu'à * en B

Bleu : Îlot génomique

Jaune : Dans le plasmide pJM105

On constate qu'au total, 64 mutations sont apparues dans 45 gènes dans les souches évoluées avec pJM105. Aucun des loci présentés dans le Tableau n'est affecté dans les souches DC62 et AC62, à part p1METDI54264D, muté chez D5E et également muté chez DC62 (G>A 54390). Outre cette mutation, DC62 ne présente qu'une autre mutation, dans un autre gène porté par p1METDI, p1METDI0118 (G>A 118809 codant pour un gène nommé *traG*). La souche AC62 possède trois mutations dans des loci qui ne sont pas touchés chez les souches évoluées (T>G 1063694, C>A 2894980), cette dernière juste en amont du cluster *mau* d'utilisation de la méthylamine) ainsi que C>A 3094073 touchant META1_2961, une histidine kinase de la famille HWE). Alors que certaines souches telles que D1, présentent plusieurs SNP et petites indels (2 SNP et 3 petites indels), d'autres souches comme D4E ou D7S ne présentent qu'un seul SNP au niveau du gène *rpoA* codant pour la sous-unité alpha de l'ARN polymérase. Des mutations dans les gènes codant les SU de l'ARN polymérase ADN-dépendante sont fréquemment retrouvées dans les expériences d'évolution dirigée (Malhotra et al. 1996 ; Utrilla et al. 2016 ; Gleizer et al. 2019 ; Phaneuf et al. 2019). Si les mutations ne sont pas liées directement avec le CH₃Cl, elles pourraient néanmoins avoir eu un rôle à jouer, comme détaillé dans le Chapitre 3. De manière similaire, des mutations dans des gènes codant pour des protéines appartenant au ribosome ont également été mises en évidence (*rplR*, *rplE*, *rplS* codant tous des protéines de la grande sous-unité du ribosome, ou *rpsA* chez A3E codant pour une protéine de la petite sous-unité du ribosome) et sont également souvent retrouvés dans les expériences d'évolution dirigée.

Par ailleurs, les souches D7S et surtout D4E présentent également des délétions de taille importante mais la plupart des gènes perdus par ces délétions sont des gènes de fonction inconnue ou des IS.

En revanche les souches évoluées dérivant d'AM1 ne présentent aucune délétion de ce type. Chez ces souches, des mutations au niveau de régulateurs transcriptionnels ont été identifiées, notamment chez A2E (histidine kinase pouvant appartenir à un système à 2 composantes) et A1E (dans un régulateur transcriptionnel appartenant à la famille Ros/MucR déjà été mis en évidence par protéomique de la souche DM4 cultivée avec le CH₂Cl₂ (Muller et al. 2011b). Il serait difficile d'identifier le phénotype spécifiquement lié à l'utilisation du CH₃Cl de ces mutations, les régulateurs transcriptionnels ayant souvent des effets pléiotropiques agissant sur un grand nombre de cibles. Ces mutations seront néanmoins abordées à nouveau dans le Chapitre 4.

Ceci étant, 5 loci de mutation communs à plusieurs cultures de souches évoluées ont été identifiés (Tableau 2.7) : Le gène *ftfL*, le promoteur du gène *nuoA*, le plasmide p1METDI, le gène *trfA* et enfin le gène *cmuA*. Ces loci ont de ce fait été analysés en priorité, et en détail.

2.3.2. Un nombre limité de loci mutés communs à des cultures évoluées indépendantes

Tableau 2.7 Loci mutés communs à plusieurs cultures de souches évoluées

Souche	Chromosome ^a		Plasmide		
	<i>ftfL</i>	<i>nuoA</i> (promoteur)	p1METDI ^b	Coïntégré pJM105-p2META ^c	
A1E	wt	wt	NA	<i>trfA</i> 14910::ISMex25 14132	<i>cmuA</i> A>C 8975
A2E	wt	wt	NA	<i>trfA</i> 14975::ISMex25 14132	<i>cmuA</i> A>G 9754
A3E/A3S	G>T 342037	wt	NA	<i>trfA</i> 15752::ISMex25 14132	<i>cmuA</i> G>A 9751
D1E	Δ443771–443777	insC 1432295	Δ28032–50836	NA	wt
D2E	wt	wt	Δ21653–50835	NA	wt
D2S	G>T 443548	wt	Δ18224–50837	NA	wt
D3E	C>T 444804	ΔC 1432295	wt	NA	wt
D4E	wt	wt	Δ27173–50837	NA	wt
D5E	Δ445023–446402 (promoteur)	ΔG 1432353	wt	NA	wt
D6E	ΔC 444674	wt	Δ32744–50836	NA	wt
D7E	C>T 444623	ΔC 1432295	Δ23007–50836	NA	wt
D7S	wt	wt	Δ33257–50837	NA	wt

^a coordonnées d'après la référence METDI de *M. extorquens* DM4 et META1 dans MaGe <https://mage.genoscope.cns.fr>

^b coordonnées d'après la référence p1METDI de *M. extorquens* DM4 sur MaGe <https://mage.genoscope.cns.fr>

^c coordonnées d'après la référence pJM105 (KT950741.1) sur GenBank <https://www.ncbi.nlm.nih.gov/genbank/> et en souligné d'après la référence p2META de *M. extorquens* AM1 sur MaGe <https://mage.genoscope.cns.fr> ; pJM105 présent dans l'ensemble des souches diffère de la référence Genbank au niveau de trois nucléotides : 1252 C>T, 12236 G>T, 13254 A>C

Tout d'abord, on peut relever que les mutations de quatre des cinq loci sont retrouvées dans des souches dérivées uniquement d'un des deux contextes génétiques, suggérant qu'elles sont potentiellement spécifiques au contexte génétique de cette souche.

- Pour l'ensemble des souches évoluées appartenant au contexte AM1 et uniquement pour elles, des événements de coïntégration indépendants ont eu lieu entre le plasmide p2META et le plasmide pJM105 (par interruption du gène *trfA*). En plus de ceci, toutes ces souches évoluées portent une mutation dans le gène *cmuA* codant

pour la méthyltransférase effectuant la première étape de la voie *cmu*. Il se pourrait que les deux répondent à une limite uniquement retrouvée chez AM1. Cela sera discuté plus en détail dans le Chapitre 4.

- Pour la plupart des souches DM4 et uniquement pour elles, on observe une grande délétion (1379 pb) au niveau du plasmide p1METDI, un des plasmides natifs de DM4. Cette région comporte plusieurs transposases, sources d'instabilité génétique.
- Toujours chez les souches du contexte DM4 uniquement, quatre souches possèdent des mutations au niveau du promoteur de *nuoA*, qui est également le promoteur de l'ensemble des gènes *nuo* (Maucourt et al. 2022) Le complexe Nuo est une NADH déshydrogénase (Falk-Krzesinski & Wolfe 1998), l'équivalent du premier complexe de la chaîne respiratoire chez les mitochondries qui procède à l'oxydation du cofacteur NADH réduit par une quinone, avec extrusion de protons. Or l'analyse protéomique a montré que l'abondance de protéines appartenant au complexe Nuo chez *M. extorquens* CM4 est accrue en culture avec le CH₃Cl par rapport au méthanol (Roselli et al. 2013). Ainsi ces mutations pourraient peut-être avoir un lien dans la croissance avec le CH₃Cl. Cela est discuté plus en détail dans le Chapitre 5 (Perspectives).
- Pour une majorité de souches (8 des 13 souches évoluées des deux contextes DM4 et AM1), différentes mutations dans le gène *ftfL* codant pour une formiate-tétrahydrofolate ligase ont été identifiées.

2.3.3. *ftfL*, le seul locus muté à la fois chez DM4 et AM1

Sur l'ensemble des loci avec des mutations, un seul est retrouvé muté à la fois dans les souches évoluées dérivant de DM4 et celles dérivant de AM1, celui du gène *ftfL* codant pour une formiate-tétrahydrofolate ligase. Ce gène présente une mutation dans six des neuf souches

évoluées dérivant de DM4 et dans deux des quatre souches issues de AM1. Une zone privilégiée où se sont accumulées les mutations n'a pas été détectée (Figure 2.9A).

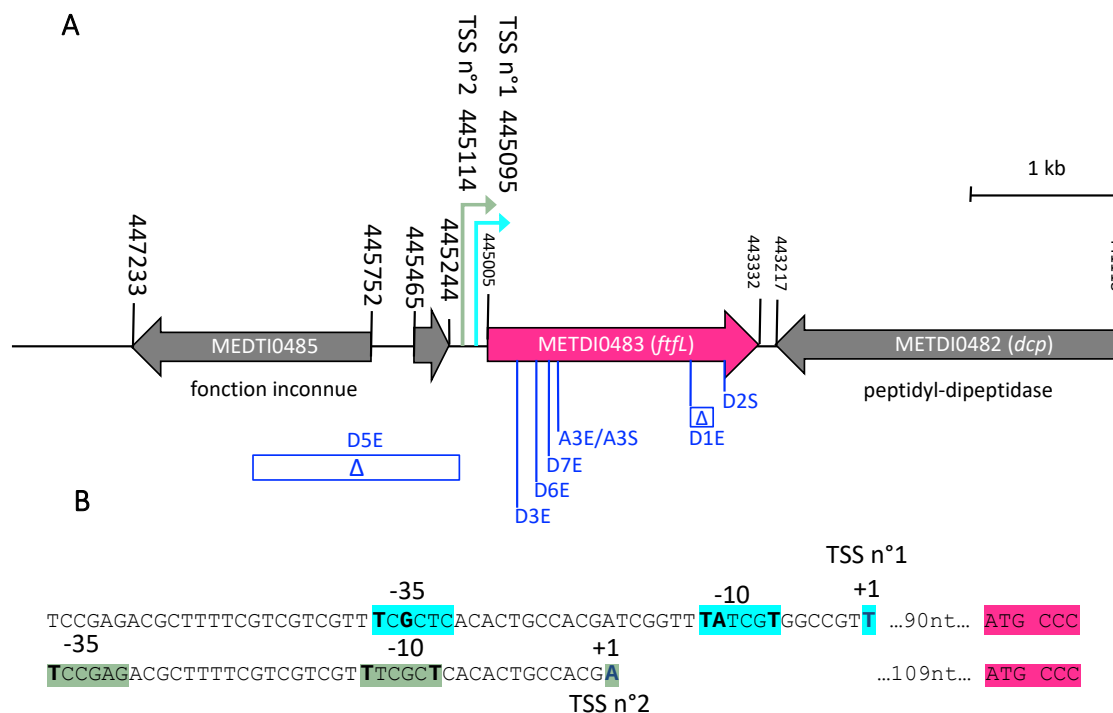


Figure 2.9 Région génomique du gène *ftfl* et localisation des mutations dans les souches évoluées

- Les sept mutations retrouvées dans le gène *ftfl* dans les souches évoluées sont renseignées, de même que la fonction des gènes voisins (coordonnées d'après le chromosome METDI de *M. extorquens* DM4 référencé sur MaGe (<https://mage.genoscope.cns.fr>). Le signe Δ indique une délétion.
- Régions promotrices prédites en amont du gène *ftfl* à partir des sites d'initiation de la transcription (Transcription Start Site, TSS) identifiés expérimentalement (Maucourt et al. 2022).

Dans le cas de D5E, il n'y a pas de mutation dans le gène *ftfl* lui-même mais une délétion en amont d'une région qui comprend la région promotrice avec les deux TSS prédits (sites de démarrage de la transcription ; Maucourt et al. 2022) (Figure 2.9 B). La délétion du promoteur suggère une absence de transcription du gène et de production de l'enzyme FtfL dans cette souche. Les mutations dans le gène *ftfl*, et une discussion de leur rôle dans la croissance avec le CH₃Cl, sont traitées dans le Chapitre suivant (Chapitre 3).

En résumé, il a été possible de faire évoluer DM4 et AM1 vers une meilleure utilisation du CH₃Cl et les nombreuses mutations retrouvées chez les souches évoluées laissent entrevoir qu'il existe plusieurs chemins évolutifs et notamment certains exclusifs à AM1 ou DM4.

Chapitre 3. La perte de fonction de FtfL, une adaptation majeure

Une majorité de souches évoluées (8 sur un total de 13 souches séquencées) a acquis une mutation dans le gène *ftfL* au cours de l'expérience d'évolution expérimentale. Il s'agit de l'unique locus muté à cette fréquence chez les souches évoluées et à être ciblé dans des souches évoluées provenant des deux contextes génétiques (6 souches sur 9 dérivant de DM4 et 2 souches sur 4 dérivant d'AM1). En outre, chez les souches de contrôle AC62 et DC62 ne possédant pas les gènes *cmu* et donc incapables de transformer le CH₃Cl, le gène *ftfL* est resté intact. Ce constat m'a poussé à investiguer plus en détail comment des mutations du gène *ftfL* pouvaient être associées à l'adaptation des souches à la croissance avec le CH₃Cl médiée par la voie *cmu*. Mon hypothèse a été qu'une mutation du gène *ftfL* pourrait modifier la fonction de l'enzyme FtfL et ainsi altérer le flux métabolique de carbone en réponse à une nouvelle contrainte à laquelle toutes les souches appartenant à l'espèce *M. extorquens* seraient exposées par suite de l'acquisition des gènes *cmu* par transfert horizontal.

3.1. La formyl-tétrahydrofolate ligase, une enzyme centrale pour le métabolisme en C1

La formyl-tétrahydrofolate ligase (FtfL, aussi appelée Fhs ou encore FTHFS pour formyl-tétrahydrofolate synthétase) est référencée EC 6.3.4.3 selon la nomenclature des enzymes du comité international de l'union de biochimie et de biologie moléculaire (<https://enzyme.expasy.org>). Elle catalyse la formation de formyl-tétrahydrofolate à partir de tétrahydrofolate (H₄F) et de formiate (Figure 3.1) avec consommation d'un équivalent ATP. L'enzyme peut néanmoins catalyser la réaction dans le sens opposé également, ce qui sera donc associé à la production d'ATP (Rabinowitz & Pricer 1962).

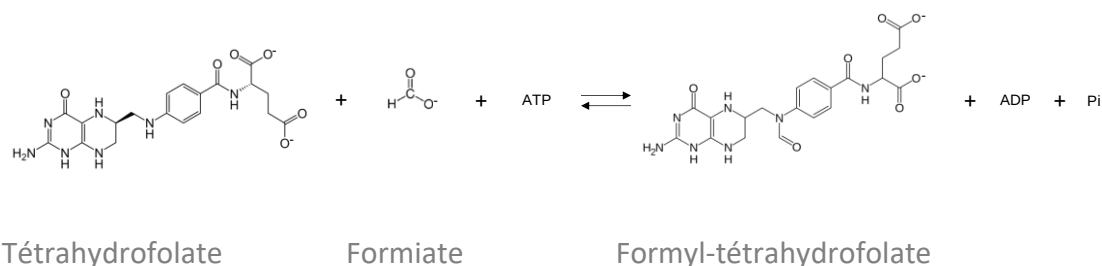


Figure 3.1 La réaction enzymatique catalysée par la *formyl-tétrahydrofolate* ligase (FtfL)

Cette enzyme est impliquée dans différents processus clés du métabolisme central, tels que la synthèse des purines et des pyrimidines (Celeste et al. 2012). Par ailleurs, le gène *ftfL* participe à la voie de fixation du CO₂ dite de Wood-Ljungdahl (Celeste et al. 2012 ; Ragsdale & Pierce 2008) et à l'acétogénèse. Il représente ainsi un marqueur clé des bactéries acétogènes dans les échantillons environnementaux (Lovell & Leaphart 2005 ; Singh et al. 2019). Chez *M. extorquens*, qui n'utilise pas la voie de Wood-Ljungdahl, FtfL est essentielle à la croissance avec le méthanol (Marx et al. 2003b). Pour rappel (Chapitre 1), chez *M. extorquens*, le méthanol est oxydé en formaldéhyde puis en formiate dans une voie enzymatique dépendante du cofacteur H₄MPT. Le formiate peut ainsi être converti par FtfL en formyl-H₄F, qui va ensuite être réduit successivement en méthényl-H₄F puis en méthylène-H₄F, pour l'assimilation du carbone via le cycle de la sérine (Marx et al. 2005 ; Crowther et al. 2008). En d'autres termes, l'assimilation du carbone issu du méthanol requiert l'enzyme FtfL. Or il s'avère que 8 des 13 souches évoluées ayant acquis une mutation dans *ftfL* (Tableaux 2.5 et 2.6), ont également perdu la capacité de croître avec le méthanol, suggérant une inactivation de l'enzyme FtfL chez ces souches évoluées.

3.2. Elucidation du rôle des mutations du gène *ftfL* dans l'adaptation à la croissance avec le CH₃Cl

3.2.1. Etude en contexte évolué de l'effet des mutations dans *ftfL* sur l'enzyme et les capacités de croissance avec des composés en C1

Pour vérifier si les mutations dans *ftfL* ont bien entraîné une inactivation de l'enzyme, j'ai en premier lieu déterminé si les souches possédant une mutation dans ce gène avaient également perdu la capacité de croître avec le formiate, un autre composé nécessitant FtfL pour l'assimilation du carbone. Le formiate permet le passage de la voie d'oxydation du

méthanol H₄MPT-dépendante à la voie d'assimilation du carbone H₄F-dépendante (Marx et al. 2003b). Par la suite, pour déterminer si les mutations dans *ftfL* ont abouti à l'inactivation totale ou partielle de l'enzyme, j'ai dosé l'activité FtfL de plusieurs souches ayant une version sauvage ou mutée du gène, et comparé à la souche de départ dont les souches évoluées sont issues. J'ai ensuite localisé les mutations sur la structure 3D protéique de FtfL pour essayer de comprendre comment les mutations observées sur le gène *ftfL* ont eu un effet sur l'activité enzymatique de FtfL.

3.2.2. Recherche de corrélation entre les mutations dans *ftfL* et la croissance avec le CH₃Cl

L'observation récurrente de mutations dans *ftfL* chez les souches évoluées ne garantit pas pour autant qu'elles soient liées à l'évolution des souches pour l'utilisation du CH₃Cl. Pour étudier ce point, trois approches ont été utilisées :

- Un inventaire de la chronologie d'acquisition des mutations dans *ftfL* et leur corrélation éventuelle avec une amélioration de la croissance avec le CH₃Cl durant l'expérience d'évolution expérimentale. Ceci est possible en déterminant la proportion de bactéries ayant perdu la capacité de croissance avec le méthanol, comme indicateur de la présence d'une mutation dans *ftfL* au cours de l'expérience, à partir d'aliquots de culture conservés sous forme congelée avant chaque nouveau transfert.
- Une comparaison des temps de génération des souches avec le CH₃Cl entre souches évoluées selon la présence d'une mutation dans *ftfL*, de façon à évaluer une éventuelle corrélation de l'inactivation de FtfL avec une croissance plus rapide avec le CH₃Cl.
- Une approche de mutagénèse dirigée du gène *ftfL* dans différents contextes génétiques pour évaluer si une mutation dans *ftfL* est à elle seule suffisante pour conférer la croissance avec le CH₃Cl (après transfert des gènes *cmu*).

Cette dernière approche s'est basée sur le plasmide suicide pCM433 (Marx et al. 2008) (cf. Chapitre 6, Matériels et Méthodes). J'ai tout d'abord apporté la mutation *ftfL*_{D7E} retrouvée dans la souche évoluée D7E (C>T 444623) à la souche de départ D0 dont sont issues toutes les souches évoluées DM4. J'ai introduit la même mutation dans une souche évoluée ayant conservé la version sauvage de *ftfL* dans le but d'évaluer si l'apport de la mutation améliorerait la croissance déjà existante avec le CH₃Cl. J'ai également procédé à une

réversion du gène muté *ftfL* d'une souche évoluée vers une version sauvage *ftfL_{wt}* afin d'évaluer si cela entraînerait un effet négatif pour la croissance avec le CH₃Cl. Pour construire ces différents mutants, j'ai tout d'abord cloné les allèles *ftfL_{D7E}* et *ftfL_{wt}* dans le plasmide suicide pCM433, que j'ai ensuite introduit dans les souches de *M. extorquens* d'intérêt par intégration au locus *ftfL* par recombinaison homologue comme décrit dans le Chapitre 6 de Matériels et Méthodes.

L'ensemble des résultats de ces expériences sont présentés dans les paragraphes suivants et m'ont permis d'avancer plusieurs hypothèses quant au rôle des mutations *ftfL* dans l'adaptation des souches évoluées à une meilleure croissance avec le CH₃Cl en utilisant la voie *cmu*.

3.3. Analyse détaillée de l'effet de la perte d'activité de FtfL chez les souches évoluées

3.3.1. Les mutations dans *ftfL* changent les capacités de croissance méthylotrophe des souches évoluées

La croissance des souches évoluées a été comparée avec celle des souches parentales en condition de croissance méthylotrophe par utilisation du méthanol ou du formiate qui requièrent la présence d'un gène *ftfL* actif (Marx et al. 2005 ; Crowther et al. 2008). Comme contrôles, les souches ont été cultivées en condition succinate, où l'intégrité de *ftfL* n'est pas nécessaire (Ochsner et al. 2017). Le succinate a permis la croissance de toutes les souches testées comme attendu. En revanche, on constate que toutes les souches portant une mutation dans *ftfL* ont perdu la capacité de croître avec le méthanol et avec le formiate, contrairement aux souches ayant conservé la version sauvage du gène (Tableau 3.1) qui peuvent toujours croître avec le méthanol. On peut ainsi conclure que les mutations dans *ftfL* chez les souches évoluées ont entraîné leur perte de croissance avec des composés nécessitant FtfL pour l'assimilation de carbone (voir Figure 1.3).

Tableau 3.1 Comparaison des capacités de croissance des souches évoluées selon la nature des altérations des protéines FtfL

Souche	Mutation <i>ftfL</i>		Croissance		
	Type et position sur le génome ^a	Effet sur la protéine (aa)	Méthanol 16 mM	Formiate 16 mM	Succinate 4 mM
Souches de départ					
A0	-	WT	+	+	+
D0	-	WT	+	+	+
Souches évoluées					
A1E	-	WT	+	ND ^c	+
A2E	-	WT	+	ND	+
A3E	G>T 342037	T127N	-	-	+
A3S	G>T 342037	T127N	-	-	+
D1E	Δ443771–443777	fs (E412A-418*) ^b	-	-	+
D2E	-	WT	+	ND	+
D2S	G>T 443548	C486*	-	-	+
D3E	C>T 444804	G68S	-	-	+
D4E	-	WT	+	ND	+
D5E	Δ445023–446402 (promoteur)	expression altérée	-	-	+
D6E	ΔC 444674	fs (G111A -143*)	-	-	+
D7E	C>T 444623	G128D	-	-	+
D7S	-	WT	+	ND	+

^a Par rapport à la référence *M. extorquens* DM4 et AM1 de MaGe <https://mage.genoscope.cns.fr>

^b fs (A-B*): Changement d'acide aminé de A jusqu'à * en B

^c ND = Non déterminé

Certaines des mutations ponctuelles d'un seul nucléotide observées entraînent un simple changement d'acide aminé pour quatre des souches évoluées (A3E, A3S, D3E, D7E), sachant que A3E et A3S sont issues du même réplicat biologique de l'évolution expérimentale contrairement à D3E et D7E. Chez d'autres souches évoluées (D1E, D2S et D6E), une mutation ponctuelle provenant d'événements distincts a entraîné un décalage du cadre de lecture ('frameshift' en anglais) suivi d'un codon stop résultant en une protéine tronquée et de séquence différente selon les souches (Tableau 3.1 : D1E une délétion de 6 nt ; D2S une simple

substitution ; et D6E une délétion d'une seule cytidine). Dans la souche D5E, la perte de la région promotrice de *ftfL* par délétion de 1379 nt en amont du codon d'initiation de *ftfL* (Figure 2.9) aboutit probablement à une perte d'expression du gène. En conclusion, l'altération des capacités de croissance des souches évoluées résulte de protéines FtfL de séquence modifiée ou mal produites dont l'activité de formyl-H₄F ligase est défavorablement altérée.

3.3.2. L'activité formyl-tétrahydrofolate ligase n'est pas détectée chez les souches évoluées portant un gène *ftfL* muté

Pour confirmer que les mutations dans le gène *ftfL* sont bien responsables de la perte de croissance avec les composés en C1, j'ai quantifié l'activité FtfL des souches grâce à un test enzymatique réalisé sur extrait cellulaire (repris de Rabinowitz & Pricer 1962 ; Marx et al 2003b ; Carroll et al 2014). Les souches testées ont été cultivées en présence d'un mélange de succinate et de méthanol (ce dernier pour garantir la production de l'enzyme (Large & Quayle 1963) en phase exponentielle pour la préparation des extraits cellulaires. Ce test enzymatique repose sur la quantification au cours du temps de la production de formyl-H₄F à partir de formiate, d'ATP et de H₄F. La Figure 3.2 montre l'optimisation du test pour déterminer la concentration en H₄F qui ne soit plus limitante pour la réaction.

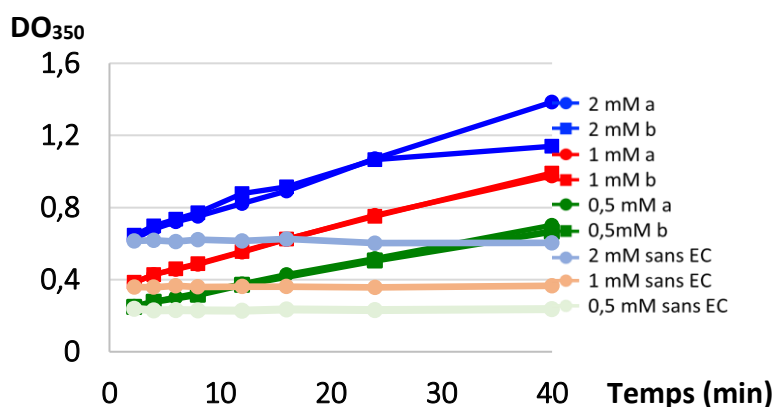


Figure 3.2 Optimisation du dosage d'activité de FtfL par utilisation de trois concentrations de tétrahydrofolate

Un extrait cellulaire a été obtenu à partir de la souche 62D (souche portant un allèle *ftfL* sauvage), cultivée dans un milieu M3C10 avec un mélange de succinate 4 mM, de méthanol 16 mM et de tétracycline 12,5 µg/mL. Deux réplicats techniques ont été réalisés pour chacune des trois concentrations de H₄F testées : 2 mM (en bleu), 1 mM (en rouge) et 0,5 mM (en vert). Pour chacune des concentrations, le témoin négatif (eau à la place de l'extrait cellulaire = sans EC) est indiqué en pâle.

Une concentration optimale de 2 mM de H₄F a été ensuite utilisée pour les tests d'activité FtfL. L'activité formyl-tétrahydrofolate ligase n'a pas été testée en présence du plasmide pJM105, mais entre la souche parentale D0 et les souches évoluées curées du plasmide pJM105 et porteuses du plasmide pCM62 pour maintenir la sélection à la tétracycline. En effet, le plasmide pJM105 porte le gène *purU* codant pour une enzyme catalysant la réaction opposée à FtfL, avec formation de formiate à partir de formyl-H₄F (Figure 1.3). Bien que l'enzyme PurU n'ait pas été décrite comme pouvant fonctionner dans les deux sens (Chistoserdova et al. 2011 ; Nagy et al. 1995), j'ai préféré éviter une possible interférence entre l'activité des deux enzymes PurU et FtfL.

Tableau 3.2 Activité de l'enzyme FtfL mesurée chez des souches porteuses du gène *ftfL* sauvage ou muté

Souche portant pCM62	Mutation <i>ftfL</i>		Activité FtfL (nmoles.min ⁻¹ . mg protéines ⁻¹) ^b
	Type et position sur le génome ^a	Variant protéique (aa)	
Souche de départ			
D0	wt	WT	263,8 ± 56,2 (N=4)
Souches évoluées curées du plasmide pJM105			
D2EL	wt	WT	372,6 ± 43,6 (N=4)
D2SL	G>T 443548	C486*	6,2 ± 10,7 (N=3)
D4EL	wt	WT	64,9 ± 19,8 (N=7) ^c
D7EL	C>T 444623	G128D	11,1 ± 17,3 (N=3)
D7SL	wt	WT	66,0 ± 71,6 (N=3)

^a Par rapport à la référence *M. extorquens* DM4 de MaGe <https://mage.genoscope.cns.fr>

^b () = nombre de réplicats biologiques ; activité dosée après culture sur milieu M3C10 avec succinate 4 mM, méthanol 16 mM et tétracycline des souches curées L après réintroduction de pCM62

^c Activité aussi testée après une culture M3C10 avec méthanol 16 mM et tétracycline : 315,7 ± 43,3 (N=3)

La mutation dans *ftfL* conduit à une perte d'activité de l'enzyme dans toutes les souches testées. Cependant, toutes les souches possédant un allèle *ftfL* sauvage n'ont pas une activité équivalente à la souche parentale D0 (pCM62). C'est le cas de la souche D4EL (pCM62) qui bien que portant un allèle sauvage *ftfL*, a une activité FtfL inférieure à la souche D0 (pCM62). Pourtant, la souche évoluée D4E se développe avec un temps de génération similaire avec le méthanol que la souche de départ ou la souche évoluée D2E (5,5 h ± 0,7 pour 105D et 5,8h ±

0,6 pour D4E, Tableau 2.1). Cette activité plus faible de l'enzyme FtfL chez D4EL (pCM62) pourrait résulter d'une production plus faible de l'enzyme lorsque cette souche est cultivée en présence de succinate. Pour le vérifier, l'activité FtfL a été testée sur des extraits cellulaires de la souche D4EL (pCM62) cultivée avec un mélange de succinate et de méthanol ou avec le méthanol comme unique source de carbone. L'activité FtfL observée a en effet été de $315,7 \pm 43,3$ nmoles.min⁻¹.mg de protéines⁻¹ pour la souche cultivée avec le méthanol seul, contre $64,9 \pm 19,8$ nmoles.min⁻¹.mg de protéines⁻¹ en présence de succinate.

Comme la souche D4EL (pCM62), la souche D7SL (pCM62) présente elle aussi une activité FtfL faible qui pourrait également résulter d'une inhibition de l'activité FtfL lors de la croissance en présence de succinate. Ceci n'a pas été confirmé au cours de ce travail. On note que les souches D4EL (pCM62) et D7SL (pCM62) portent des délétions (Tableau 2.6) dont une partie est commune. Cette région est flanquée d'éléments génétiques mobiles (IS) et est dupliquée chez DM4 avec plus de > 87% d'identité nucléique sur 7,7kb. La région perdue code pour les protéines METDI0317 à METDI0325 comprenant un système de régulation à deux composantes potentiellement sensible à de faibles concentrations en oxygène (METDI0317 à METDI0319 présentant 44 à 57% d'identité avec les protéines du système FixLJ de bactéries diazotrophiques fixatrices d'azote (Anthamatten & Hennecke 1991 ; revue récente par Price et al. 2021). Un gène codant pour un cytochrome c de la classe 1 (METDI0322) et deux gènes de fonction inconnue portant un domaine codant pour des protéines de réponses au stress (METDI0320 et METDI023) font partie des autres gènes perdus dans ces délétions.

En résumé, outre ces cas particuliers, on observe pour la plupart des souches soit des mutations 'frameshift' entraînant une protéine FtfL tronquée, soit un changement d'un seul acide aminé par rapport à la séquence protéique sauvage de FtfL. J'ai cherché à évaluer leur impact potentiel sur l'activité de FtfL par analyse de leur localisation dans la structure 3D de FtfL.

3.3.3. Localisation des variations de séquence dans la structure de FtfL et leur impact potentiel

La protéine FtfL de type sauvage est sous forme homotétramérique pour l'enzyme de *M. extorquens* (Figure 3.3B) (Marx et al. 2003b ; Kim et al. 2020). Chez *Clostridium cylindrosporum*, il a été montré que la dissociation du tétramère entraîne une perte d'activité (Scott & Rabinowitz 1967) et que des cations monovalents empêchent cette dissociation. Les domaines B (Asp129-Leu252) et C (Lys444-Phe557) ont des rôles dans l'interaction entre monomères. Une longue séquence C-terminale dans le domaine C permet la formation d'un dimère stable (Kim et al 2020) et trois hélices alpha et deux brins beta dans le domaine B avec 19 résidus permettent l'interaction entre deux dimères.

Toutes les mutations ponctuelles d'acide aminés rencontrées dans les souches évoluées sont localisées dans le domaine A de la protéine FtfL qui porte le site actif (Figure 3.3.A). Pour D3E, le changement d'acide aminé (G68S) se situe juste au niveau du site actif (Figure 3.3C), proche de la variation P64R décrite précédemment (Caroll et al. 2014b). Pour D7E, un simple SNP entraîne la variation G128D à l'interface entre deux sous-unités. Dans la version sauvage du tétramère, deux résidus glycine sans chaîne latérale et neutres appartenant chacun à une sous-unité différente se font face alors que dans la protéine mutée, les deux résidus d'acide aspartique plus volumineux et chargés négativement perturberont sans aucun doute l'ensemble de la structure en aboutissant à une perte d'activité.

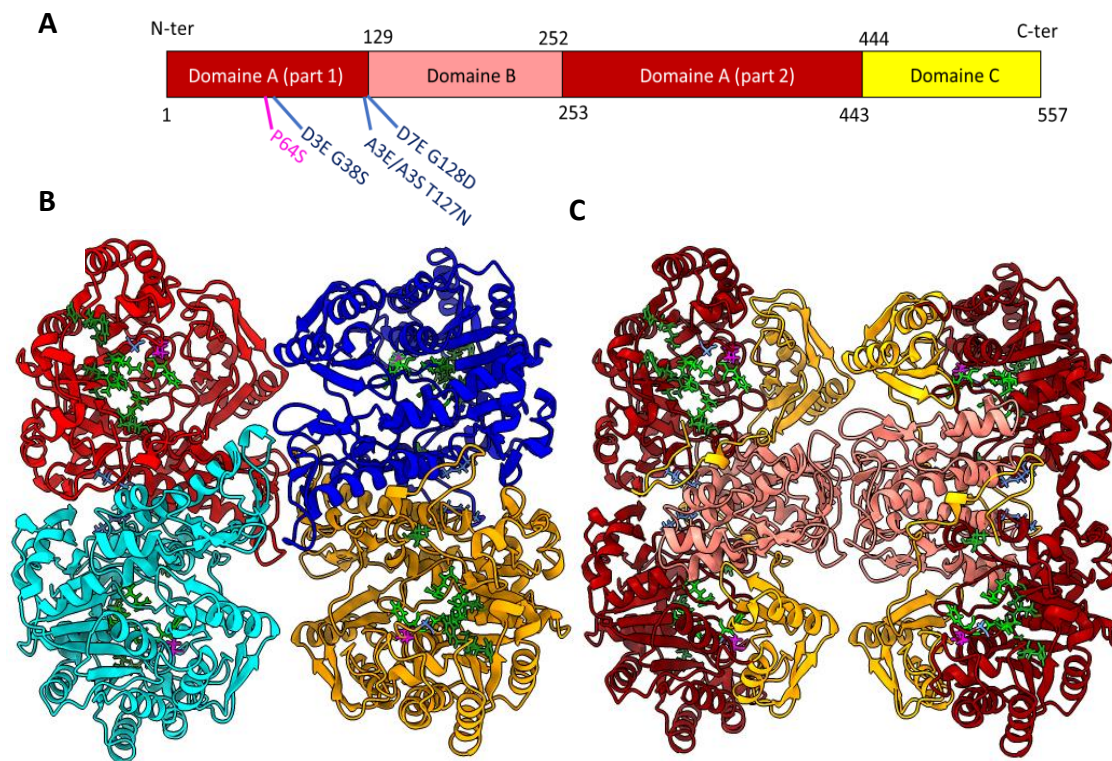


Figure 3.3 Prédiction de la structure du tétramère de FtfL de DM4 par Alphafold avec les variations ponctuelles retrouvées dans les souches évoluées.

- A)** La protéine FtfL chez *M. extorquens* a une taille de 557 aa et comporte trois domaines (Kim et al. 2020). Chaque domaine est indiqué par une couleur différente (conservée entre A et C). Le site actif est contenu dans le plus grand domaine A, qui comprend deux parties de la séquence (M1-G128 et Q253-T443 interrompues par le domaine B (D129-L252)). Le domaine B, comme le domaine C (K444-F557) permettent la liaison avec les autres monomères de FtfL.
- B)** Structure du tétramère de FtfL de la souche DM4 prédite par Alphafold et visualisée avec ChimeraX. Les 4 monomères composant le tétramère de FtfL sont de couleurs différentes. Le site actif de chacun des monomères est indiqué en vert (d'après Kim et al. 2020).
- C)** La structure donnée en partie B avec les trois domaines de chacun des monomères de même couleur qu'en A. Les résidus ayant fait l'objet de mutations dans les souches évoluées D3E (G68), D7E (G128) et A3E/A3S (T127) sont indiqués en bleu clair. Le résidu 64P (modifié dans Carroll et al. 2014) a été mis en évidence en magenta.

En conclusion, les différentes mutations retrouvées dans le gène *ftfL* chez les souches évoluées ont entraîné l'inactivation de l'enzyme correspondante rendant les souches incapables de croître avec le méthanol ou le formiate sans effet sur la croissance avec le succinate. Puisque les souches ont été en contact avec le succinate durant la première partie de l'expérience d'évolution expérimentale, il se pourrait que des mutations dans les gènes de la voie d'assimilation du méthanol soient apparues par dérive génétique du fait qu'aucune pression de sélection n'ait été exercée sur les gènes de la voie de dégradation du méthanol. Néanmoins, le fait d'avoir retrouvé des mutations de façon répétée dans un seul gène de cette voie laisse plutôt penser que l'acquisition des mutations *ftfL* pourrait avoir eu un effet bénéfique pour la croissance en présence du CH₃Cl. Des expériences complémentaires ont ainsi été mises en place pour le vérifier.

3.4. Les mutations dans *ftfL* pourraient être liées à l'évolution de la croissance avec le CH₃Cl

3.4.1. Les mutations dans *ftfL* se sont produites en début d'expérience d'évolution expérimentale de la culture D7

Lors de l'expérience d'évolution expérimentale, le passage de la phase 1 à la phase 2 a été possible lorsque les cultures ont accumulé des premières mutations bénéfiques leur permettant de croître avec le CH₃Cl seul. J'ai vérifié pour une des cultures de l'expérience si des mutations dans *ftfL* étaient survenues dès la phase 1. J'ai évalué ceci pour le réplicat D7 à partir duquel j'ai isolé deux souches évoluées (D7E avec une mutation *ftfL* et D7S ayant conservé l'allèle sauvage), grâce auxquelles j'ai pu reconstituer la dynamique d'apparition de ces deux sous-populations. J'ai retracé la présence de ces mutations en partant d'aliquots de cultures conservées à -80°C et j'ai déterminé la proportion de souches ayant perdu la capacité de croître avec le méthanol, comme indicateur de l'état fonctionnel de FtfL. Trois temps ont été choisis, en fin de phase 1 et deux temps au cours de la phase 2 (voir Tableau 3.3).

Tableau 3.3 Conservation de la capacité de croître avec le méthanol au cours de l'évolution expérimentale du réplicat D7

Nombre de générations	Numéro du transfert	Stade de l'évolution expérimentale	Croissance des colonies ^a		Nombre de colonies assimilant le méthanol/succinate (%)
			Succinate 10 mM	Méthanol 40 mM	
22	6	Fin de phase 1	180	12	6,7
103	22	Milieu de phase 2	363	127	35,0
190	39	Fin de phase 2	366	8	2,2

^a Croissance sur milieu solide (contenant de la tétracycline 12.5 µg/mL) de colonies issues d'un aliquot à 80°C additionné de glycérol

Dès la fin de phase 1 (6^{ème} transfert), la majorité des souches de la culture D7 a perdu la capacité de croître avec le méthanol, ce qui correspond à la première culture où l'on observe une augmentation importante de la DO₆₀₀ (0,214 comparé aux transferts précédents ne dépassant pas 0,075 ; Figure 2.2). Au cours de la phase 2, la proportion des souches ayant perdu la capacité de croître avec le méthanol est variable mais reste inférieure à 35%. Ainsi, la perte de la fonctionnalité du gène *ftfL* n'est pas totale car une sous-population est toujours capable d'assimiler le méthanol en conservant l'enzyme FtfL fonctionnelle. Deux souches évoluées provenant de D7 avaient été isolées, D7E et D7S. Les deux proviennent vraisemblablement de deux sous-populations différentes avec D7S provenant d'une sous-population ayant conservé *ftfL* sauvage. La proportion de cette sous-population augmente à nouveau en milieu de phase 2 puis rediminue en fin de phase 2. Ceci pourrait traduire un équilibre dynamique entre les différentes sous-populations, peut-être en raison des mutations bénéfiques qui se sont produites au cours de l'évolution expérimentale. D'une manière similaire, D2S et D2E sont issues du même réplicat D2 de l'évolution expérimentale et une des deux souches a conservé FtfL fonctionnelle alors que l'autre non.

En résumé, les mutations dans *ftfL* ont été trouvées dans une majorité des souches et sont possiblement apparues dès la phase 1 dans l'expérience d'évolution expérimentale, comme démontré dans le cas de la culture D7. L'acquisition des mutations dans *ftfL* pourrait ainsi

coïncider avec l'amélioration de leur croissance avec le CH₃Cl, ce qui conforterait l'hypothèse que les mutations sur *ftfL* sont liées à l'évolution de la croissance avec le CH₃Cl.

3.4.2. Un gène *ftfL* muté corrèle avec une croissance plus rapide avec le CH₃Cl

L'effet de la mutation du gène *ftfL* a été évalué par rapport au temps de génération des différentes souches évoluées pour le CH₃Cl, avec comme témoin la croissance avec le succinate et le méthanol. Toutes les souches évoluées se développant rapidement avec le CH₃Cl (temps de génération faible) portent une mutation dans *ftfL* sans pour autant modifier les capacités de croissance avec le succinate de manière significative (Figure 3.4 ; comparer D7E* et D7S ou A3E* et A3S*).

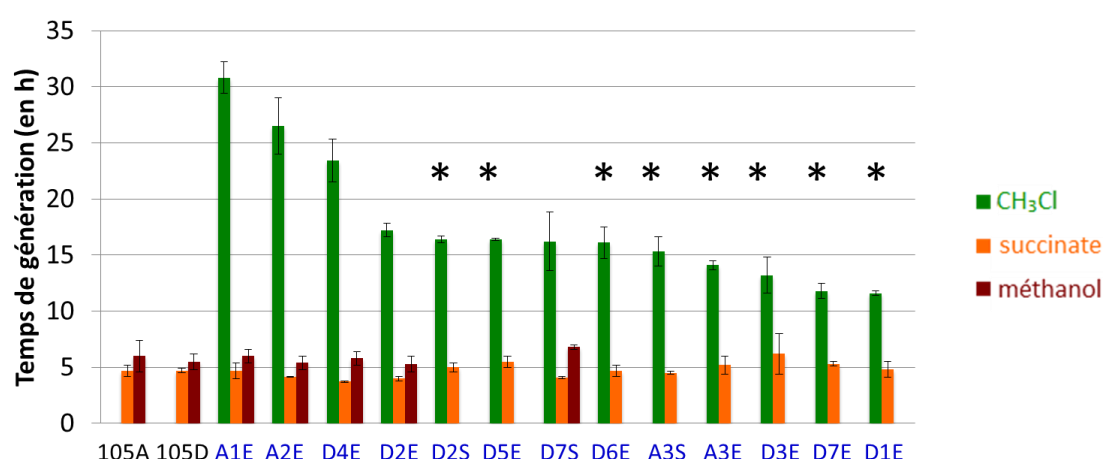


Figure 3.4 Comparaison des capacités de croissance des souches évoluées et parentales avec le CH₃Cl, le succinate et le méthanol selon la présence de l'allèle sauvage ou muté du gène *ftfL*

Les temps de génération (en heures) des 13 souches évoluées (nom en **bleu**) ainsi que des deux souches de départ 105A et 105D sont indiqués pour le CH₃Cl (**vert**), le succinate (**orange**) et le méthanol (**bordeaux**). Ils ont été déterminés à partir de trois réplicats biologiques (les écarts-types sont indiqués). Les souches ont été placées sur l'axe des abscisses en ordre décroissant du temps de génération avec le CH₃Cl. Les souches possédant une mutation dans *ftfL* aboutissant à la perte de fonction de l'enzyme sont notées par un astérisque, et correspondent à l'absence de croissance avec le méthanol.

On constate que les souches évoluées se développant le plus rapidement en CH₃Cl portent un gène *ftfL* muté, et vice-versa. Cela suggère un rôle clé de l'inactivation de FtfL dans l'adaptation à la croissance avec le CH₃Cl dans les conditions testées, que ce soit en entraînant une protéine tronquée ou une déstabilisation de la forme tétramérique de FtfL.

3.5. Construction ciblée de mutants du gène *ftfL* pour démontrer un lien de cause à effet par rapport à la capacité de croissance avec le CH₃Cl

Les trois mutants construits pour le locus *ftfL* ont été validés en combinant le séquençage Sanger de leurs allèles *ftfL* après amplification PCR, la caractérisation du phénotype de croissance vis-à-vis du méthanol, et le dosage de l'activité de formyl-tétrahydrofolate ligase (Figure 3.5A). L'étape suivant la validation des mutants a été de leur fournir les gènes de la voie *cmu* du plasmide pJM105 afin d'évaluer leurs capacités de croissance avec le CH₃Cl (Figure 3.5B).

La construction des mutants répondait à trois objectifs.

- déterminer si la seule inactivation de l'enzyme FtfL est suffisante pour permettre à la souche parentale D0 d'acquérir la croissance avec le CH₃Cl en présence des gènes *cmu*. J'ai construit la souche D0_*ftfL*_{D7E} par remplacement de l'allèle sauvage *ftfL*_{wt} de la souche de départ D0 par l'allèle muté *ftfL*_{D7E} de la souche D7E (C>T 444623). On note ici que la souche D7E est l'une des souches évoluées les plus performantes pour sa croissance avec le CH₃Cl (Figure 3.4). Le mutant D0_*ftfL*_{D7E}(pJM105) se développe avec le CH₃Cl contrairement à la souche D0. Ainsi, **l'inactivation de FtfL est suffisante pour conférer la capacité de croissance avec le CH₃Cl après acquisition du plasmide pJM105 en contexte D0.** Toutefois le temps de génération de D0_*ftfL*_{D7E}(pJM105) n'est pas encore optimal car on voit qu'il n'est pas égal à D7E, signifiant qu'il existe encore une marge de progression (19,6 ± 4,8 h pour D0_*ftfL*_{D7E}(pJM105) contre 12,0 ± 0,8h pour D7EL (pJM105)). **En plus de l'inactivation de FtfL, d'autres gènes mutés ont participé à l'amélioration de la croissance de plusieurs des souches évoluées.** Par exemple, quatre d'entre elles (D1E, D3E, D7E et D5E) possèdent une mutation dans le promoteur des gènes *nuo*.
- évaluer si l'apport d'une version sauvage de *ftfL* ralentit la croissance de la souche évoluée D7E porteuse d'un allèle muté de *ftfL*. J'ai ainsi construit la souche D7EL_*ftfL*_{wt}. Comme attendu, on observe que la souche D7EL_*ftfL*_{wt} (pJM105) a une capacité de croissance bien moindre (temps de génération au-delà de 23h) par rapport à la souche D7EL (pJM105) dont elle est issue (temps de génération d'environ 12h) (Figure 3.5B) montrant que **l'inactivation de FtfL a été bénéfique pour la croissance de la souche évoluée D7E avec le CH₃Cl.**

- tester l'effet inverse de l'acquisition d'un allèle muté *ftfL*_{D7E} dans le contexte d'une souche évoluée *ftfL*_{wt} avec une histoire évolutive conférant une croissance peu performante. En effet, D4E est la souche évoluée issue du contexte D0 qui se développe le plus lentement avec le CH₃Cl (Figure 3.4). Si la présence d'un allèle muté a un effet bénéfique dominant, on s'attendrait à ce que le remplacement d'un allèle muté (*ftfL*_{D7E}) par l'allèle sauvage (*ftfL*_{wt}) accélère sa croissance par utilisation du CH₃Cl. J'ai construit la souche D4EL_*ftfL*_{D7E} et comparé sa croissance avec la souche D4EL. Bien qu'une grande variation entre les différents réplicats de mesure du temps de génération en CH₃Cl soit observée chez la souche D4EL (pJM105), la mesure moyenne avoisine celle du temps de génération avec la souche construite D4EL_*ftfL*_{D7E} (pJM105) (Tableau 2.1 : 23,4 ± 1,9). On en conclut que l'inactivation de *ftfL* n'a pas eu d'effet positif notable sur la capacité de croître en CH₃Cl dans le contexte de la souche D4E. Dans l'étude de Carroll et al. (2014b), les auteurs avaient montré de manière similaire que l'introduction d'une mutation dans *ftfL* dans des souches évoluées qui n'en possédaient pas n'était pas forcément bénéfique (Carroll et al. 2014b) pour améliorer la croissance de *M. extorquens* avec le succinate. Une telle mutation est restée sans effet pour certaines des souches, alors qu'elle était désavantageuse pour d'autres. Ceci renvoie au concept d'épistasie dans lequel l'effet d'une mutation est dépendant d'autres mutations préexistantes. Je fais ici l'hypothèse que l'inactivation de FtfL n'a pas eu d'impact positif sur la croissance de D4EL avec le CH₃Cl car la souche évoluée D4E aurait suivi un chemin évolutif qui aurait rendu l'acquisition d'une mutation inactivant FtfL sans effet bénéfique. Cela pourrait être le cas également pour D7S. Peut-être que la délétion de gènes dans les deux souches y a participé, les facteurs de transcription ayant des effets pléiotropiques. De plus, les trois souches évoluées issues de D0 qui portent un allèle *ftfL*_{wt} ont en commun le fait de présenter une mutation dans la partie C-terminale de la SU alpha de l'ARN polymérase ADN-dépendante codée par *rpoA* (L302I chez D4E et E304V chez D7S) et dans la SU bêta de la même enzyme codée par *rpoC* (T1312A) chez D2E (Tableau 2.6). La partie C-terminale de la protéine RpoA de la SU alpha de l'ARN polymérase ADN-dépendante est non-essentielle à la transcription basale mais néanmoins impliquée dans la liaison à de nombreux activateurs de transcription chez *E. coli* (Zhang et Darst 1998) ; RpoA d'*E. coli* et de *M. extorquens* partagent 49% d'identité protéique. Elles pourraient faire partie d'une stratégie FtfL-indépendante (bien que la souche D6E qui possède une mutation dans *ftfL* possède également une mutation dans

rpoD qui code un facteur sigma). Ce type de mutation est fréquemment rencontré dans les expériences d'évolution dirigée et pourrait, indépendamment du CH₃Cl, contribuer à une meilleure croissance des souches évoluées (Dalldorf et al. 2023).

A

Souche	Mutation <i>ftfL</i>	Activité FtFL (nmoles.min ⁻¹ .mg protéines ⁻¹) ^b	Croissance avec méthanol 16 mM
D0	wt	263,8 ± 56,2 (N=4)	+
D0 <i>ftfL</i> _{D7E}	C>T 444623 ^a	0 (N=2)	-
D4EL	wt	NA ^c	+
D4EL <i>ftfL</i> _{D7E}	C>T 444623	NA ^c	-
D7EL	C>T 444623	11,1 ± 17,3 (N=3)	-
D7EL <i>ftfL</i> _{D0}	wt	442,3 ± 52,8 (N=3)	+

^a coordonnées selon la référence *M. extorquens* DM4 de MaGe <https://mage.genoscope.cns.fr>

^b () = nombre de réplicats biologiques effectués

^c NA = Non Applicable, inhibition de l'activité FtFL en présence de succinate

B

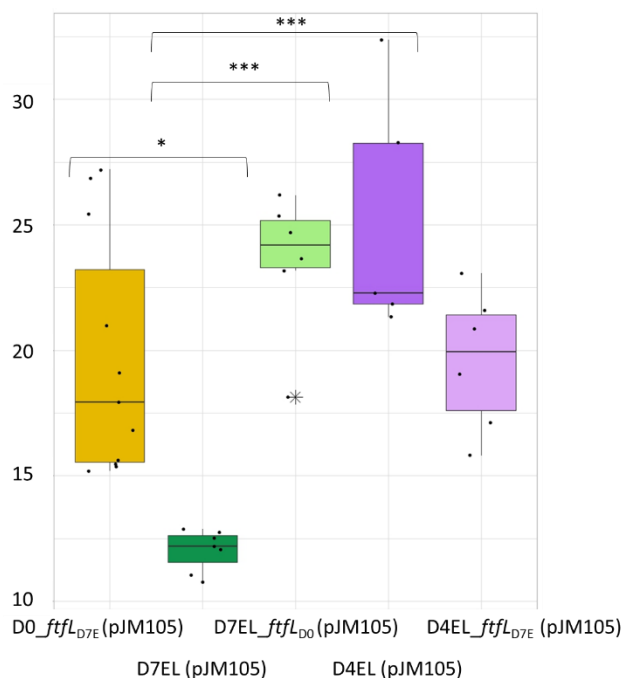


Figure 3.5 La culture avec le CH₃Cl des mutants *ftfL* construits montre l'effet de l'inactivation de FtFL sur la croissance avec ce composé

A) Récapitulation de l'état du gène *ftfL* et des données phénotypiques associées (activité FtFL, croissance avec méthanol) pour les mutants construits en parallèle des souches parentales dont elles sont issues.

B) Détermination des temps de génération des mutants avec le CH₃Cl (et des souches desquelles elles sont issues). Un point représente un temps de génération déterminé pour une seule culture. Le temps de génération n'a pas été déterminé pour la souche D0 car elle ne peut pas se développer avec le CH₃Cl. Un test statistique Kruskal-Wallis suivi d'un test de Dunn a été effectué pour déterminer si les différences entre les souches étaient significatives (* : p<0,05 ; *** : p<0,001)

Ayant démontré que l'inactivation de FtfL est suffisante pour permettre la croissance avec la voie *cmu* de la souche parentale D0, j'ai recherché une hypothèse permettant d'expliquer l'avantage d'inactiver cette enzyme essentielle à la croissance avec le méthanol.

3.6. Hypothèse pour expliquer l'effet bénéfique d'une inactivation de FtfL sur la croissance avec le CH₃Cl : la dualité entre FtfL et PurU dans le métabolisme de la voie *cmu*

L'enzyme PurU apportée par les gènes *cmu* ne catalyserait que la dégradation du formyl-H₄F en formiate et H₄F. Or, FtfL peut catalyser en théorie la formation de formyl-H₄F à partir de formiate et de H₄F mais également la réaction inverse. Cependant, il a été démontré que lors de la croissance avec le méthanol, FtfL a principalement une activité de formation de formyl-H₄F et cette activité est indispensable pour la formation de biomasse (Marx et al. 2003b) car le formyl-H₄F va être ultimement transformé en méthylène-H₄F qui entrera dans le cycle de la sérine. Néanmoins, lors de la croissance avec le CH₃Cl, FtfL n'est pas essentielle à la formation de biomasse à partir de CH₃Cl car le méthylène-H₄F entrant dans le cycle de la sérine est généré directement par une première oxydation du méthyl-H₄F, lui-même produit par l'activité des méthyltransférases CmuAB à partir du CH₃Cl, par la méthylène-H₄F réductase MetF2 de la voie *cmu* (Figure 3.6). Je postule qu'au contraire, FtfL a une activité contre-productive lors de la croissance avec le CH₃Cl : cette enzyme empêcherait l'oxydation totale du carbone issu du composé à partir du formiate pour générer du CO₂ (générant un équivalent réduit), en reconvertissant le formiate généré par PurU en formyl-H₄F, le tout en consommant un ATP. Ainsi, l'activité de FtfL serait désavantageuse pour la formation d'énergie à partir du CH₃Cl. De plus, cette reconversion du formiate en formyl-H₄F implique le cofacteur H₄F également nécessaire pour effectuer la première étape de la voie *cmu*, qui pourrait de ce fait devenir limitant pour le métabolisme de la souche.

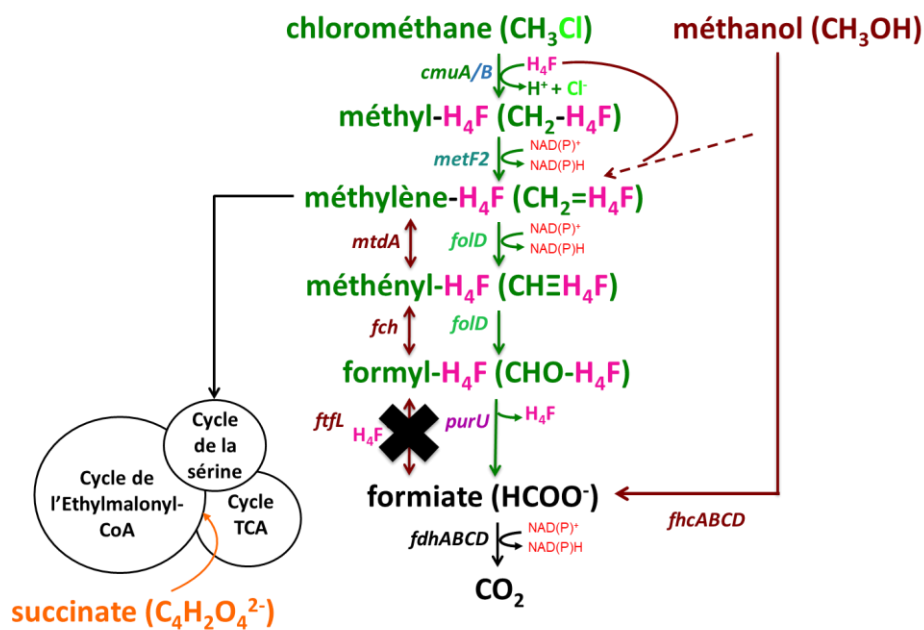


Figure 3.6 L'hypothèse du rôle de FtfL dans l'adaptation à la croissance avec le CH_3Cl

La voie *cmu* est indiquée en vert. La réaction catalysée par FtfL a été inactivée dans plusieurs des souches évoluées. Ceci pourrait rediriger le flux uniquement vers la formation de formiate à partir de formyl- H_4F par *purU* pour une dernière oxydation par la formiate-deshydrogénase en CO_2 .

Ainsi, en éliminant FtfL, les souches évoluées peuvent à la fois synthétiser de la biomasse et de l'énergie à partir du CH_3Cl (Figure 3.6). Néanmoins, cela sous-entend que FtfL ne fonctionne que dans le sens de la formation de formyl- H_4F , et que dans le cas présent, seule PurU effectue la réaction inverse, alors que FtfL est normalement capable d'effectuer cette réaction dans les deux sens. La question de savoir si l'activité FtfL peut se substituer à l'activité PurU comme précédemment proposé (Michener et al. 2016) s'est alors posée, ce qui serait en opposition avec l'hypothèse que j'ai formulée.

3.7. PurU, une enzyme essentielle pour la croissance avec le CH_3Cl chez *M. extorquens*

Pour vérifier si FtfL peut remplacer PurU, plusieurs approches étaient possibles. Il existe une approche enzymatique de dosage de l'activité de PurU (Nagy et al. 1995). Il m'était difficile de mettre en œuvre rapidement un test d'activité PurU notamment car nous ne disposons pas au laboratoire de formyl- H_4F substrat de la réaction. Une approche génétique a été choisie car j'avais à ma disposition différentes versions de pJM105 avec des délétions de chacun des gènes du cluster *cmu* (Michener et al. 2016). Notamment, le plasmide pJM112 contient une

version de l'opéron *cmu* avec une délétion de *purU* (vérifié par PCR, avec les amorces LFM024-LFM025 ciblant *purU* de pJM105 donnant 1,2 kb pour pJM105 et 0,3 kb pour pJM112, non-montrées). Je disposais également du plasmide pJM111 avec un opéron *cmu* portant une délétion du gène *fold* codant pour l'enzyme effectuant les deux réactions précédant celles de PurU et fournissant donc à cette dernière le pool de formyl-H₄F (vérifié avec les amorces LFM026-LFM027 ciblant *fold* de pJM105 et donnant 1,15 kb pour pJM105 et 0,2 kb pour pJM111).

Je suis parti du principe que si FtfL peut assurer le rôle de PurU, les souches évoluées avec *ftfL* sauvage devraient se développer avec le CH₃Cl si on leur fournit pJM112 à la place de pJM105. Les souches avec FtfL non fonctionnelle en seraient incapables. Le plasmide pJM112 a ainsi été introduit par conjugaison à plusieurs des souches évoluées (les versions curées des souches D2E, D2S, D7E, D7S, A3E) ayant soit conservé la version sauvage de FtfL soit ayant une version mutée. A l'inverse, FtfL ne devrait pas être en mesure de remplacer Fold. Ainsi, l'état de FtfL n'impactera pas la croissance des souches curées si on leur fournit pJM111 (j'ai ainsi utilisé ces souches comme contrôle de l'expérience en plus des souches possédant pJM105). Les résultats de cette expérience sont présentés dans le Tableau 3.4

Tableau 3.4 Croissance de transconjugants de souches évoluées en CH₃Cl avec ou sans le gène *fold* ou *purU*

Souche ^a	Caractéristique du plasmide				Croissance CH ₃ Cl 15 mM ^b
	Nom	<i>cmuABC, fmdB</i> <i>paaE, hutI metF2</i>	<i>fold</i>	<i>purU</i>	
Souches de départ					
A0, D0	pJM111	+	-	+	-
	pJM112	+	+	-	-
	pJM105	+	+	+	-
	pCM62	-	-	-	-
Souches évoluées					
A3EL*, D2EL, D2SL* D7EL*, D7SL	pJM111	+	-	+	+
	pJM112	+	+	-	-
	pJM105	+	+	+	+
	pCM62	-	-	-	-

^a * indique une mutation dans le gène *ftfL*

^b déterminé par présence (+) ou absence (-) de trouble après 12 jours de culture (culture en présence de tétracycline 12,5 µg/mL).
L'ensemble des souches testées se sont développées avec le succinate comme substrat de croissance.

Toutes les souches, peu importe l'état de FtfL, sont incapables de croître avec le CH₃Cl avec pJM112 (alors qu'elles en étaient capables lorsqu'elles possédaient pJM105) soulignant le **côté essentiel de PurU dans la croissance avec le CH₃Cl**. Ceci est en adéquation avec Vannelli et al. 1999, qui avaient obtenu des résultats similaires avec *M. extorquens* CM4 mais qui rentre en contradiction avec les résultats de Michener (Michener et al. 2016). Néanmoins, leurs résultats étaient obtenus dans le contexte d'un flux de carbone limité.

Au vu des résultats des expériences avec pJM111, **Fold est une enzyme dispensable** puisque toutes les souches possédant pJM111 se sont développées avec le CH₃Cl, ici encore indépendamment de l'état de FtfL. L'activité de l'enzyme Fold, une enzyme bifonctionnelle, pourrait être assurée par deux enzymes : la NADP-dépendante méthylène H₄F déshydrogénase MtdA et la méthényl-H₄F cyclohydrolase Fch nécessaires à la formation de biomasse lors de la croissance avec le méthanol (Vorholt et al. 1998 ; Pomper et al. 1999 ;

Vorholt 2002) (Figure 3.6). Une telle hypothèse avait déjà été émise (Michener et al. 2016) mais encore une fois pour expliquer les résultats d'une expérience où les flux de carbone étaient très faibles. MtdA/Fch et FOLD ne pourraient que peu catalyser les réactions inverses. L'inactivation de MtdA ou de Fch ne peut pas être compensée par FOLD lors de la croissance avec le méthanol de *M. extorquens* AM1 (Marx & Lidstrom 2004). Ainsi, si les deux enzymes MtdA et Fch ne sont que partiellement capables de remplacer les deux activités de FOLD, on pourrait s'attendre à une vitesse de croissance avec le CH₃Cl grandement ralentie lorsque *fold* est absent. Pour confirmer cette prédiction, en plus de caractériser quantitativement la capacité de croissance des souches (Tableau 3.8), une mesure du temps de génération devrait être effectuée.

En conclusion, les résultats obtenus sont en adéquation avec l'hypothèse d'une interférence de flux métabolique entre des réactions réverses catalysées par les enzymes FtfL et PurU, sachant que seule PurU est essentielle pour l'utilisation du CH₃Cl par la voie *cmu*.

3.8. Faible occurrence du gène *purU* dans le génome de souches isolées de l'environnement et porteuses des gènes *cmu*

Si PurU est indispensable pour la dégradation du CH₃Cl médiée par la voie *cmu* de souches affiliées à *M. extorquens*, peut-être n'est-ce pas le cas pour d'autres contextes génétiques. Par ailleurs, en plus de son implication dans la voie *cmu* d'utilisation du CH₃Cl, PurU est une enzyme qui sert dans le métabolisme des purines (Figure 3.7). En effet, dans la voie de biosynthèse des purines, un des intermédiaires de réaction, le fGAR (formylglycineamine-ribonucléotide) est formé par formylation de GAR (glycineamine-ribonucléotide) (Figure 3.7). Cette formylation peut être effectuée par PurN à partir de formyl-H₄F, ou par PurT à partir de formiate dont le pool est formé par PurU (Nagy et al. 1995). *E. coli* possède les deux gènes et peut effectuer les deux réactions (Nygaard & Smith 1993). Les souches de *M. extorquens* utilisées dans cette étude ne possèdent que PurN (vérifié grâce à la plateforme MaGe <https://mage.genoscope.cns.fr>). FtfL est très certainement l'une des sources du formyl-H₄F utilisé par PurN chez *M. extorquens*, ce qui n'est pas le cas de *E. coli* où *ftfL* est absent (Aluri et al. 2015). Comme FtfL, PurU fonctionne sous forme de tétramère (Sampei et al. 2013). PurT et FtfL utilisent le même substrat (le formiate). Il a été démontré qu'un antagonisme existait entre la présence des gènes *purT* et *ftfL*. Les bactéries possédant *purT* ne possèdent que très

rarement *ftfL* et inversement (Aluri et al. 2015). Par ailleurs, *E. coli* exprimant l'un ou l'autre de ces gènes montre un meilleur fitness que la souche exprimant les deux (Aluri et al. 2015). Je me suis donc demandé s'il existait un antagonisme similaire entre *ftfL* et *purU* pour la croissance avec le CH₃Cl (Tableau 3.5).

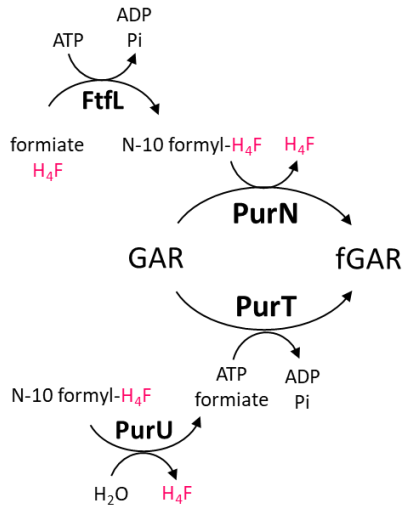


Figure 3.7 PurN et PurT peuvent toutes deux catalyser la formylation du glycineamine-ribonucléotide (GAR) en formylglycineamine-ribonucléotide (fGAR) durant la synthèse des purines

PurN catalyse la formylation du GAR en fGAR à partir de formyl-H₄F (qui peut être fourni par FtfL). Une autre voie alternative existe dans laquelle le GAR est formylé en fGAR à partir d'ATP et de formiate (ce dernier peut être fourni par PurU). D'après Zhang et al. 2008.

Tableau 3.5 Présence et conservation de la séquence protéique des gènes *purU*, *purN*, *ftfL* chez plusieurs souches possédant les gènes *cmu* dont le génome est séquencé

Souches ^a	Conservation (%) de la séquence protéique (aa)									
	<i>purU</i> ^b	Référence gène	<i>purN</i> ^c	Référence gène	<i>ftfL</i> ^c	Référence gène	<i>cmuA</i> ^b	Référence gène	<i>cmuB</i> ^b	Référence gène
<i>Hyphomicrobium</i> sp. MC1	100	HYPMCv2_2267	61 (75)	HYPMCv2_4125	67 (79)	HYPMCv2_4379	100	HYPMCv2_2273	100	HYPMCv2_2275
<i>Methylobacterium extorquens</i> CM4	62 (71) ^d	Mchl_5699	98 (99)	Mchl_2337	99 (99)	Mchl_0447	84 (92)	Mchl_5697	64 (79)	Mchl_5727
<i>Aminobacter ciceronei</i> IMB-1	-	-	55 (70)	JACJHZ01_250013	65 (78) 66 (78)	JACJHZ01_180048 JACJHZ01_200074	83 (91)	JACJHZ01_180069	73 (84)	JACJHZ01_180067
<i>Aminobacter lissarensis</i> CC495	-	-	55 (69)	JACHGI01_30035	61 (75) 66 (79)	JACHGI01_130001 JACHGI01_160046	83 (91)	JACHGI01_160025	73 (85)	JACHGI01_160027
<i>Celeribacter indicus</i> P73	55 (68) 49 (66) 51 (66) 46 (64)	P73_3955 P73_0126 P73_3498 P73_4396	53 (67)	P73_2485	-	-	81 (88)	P73_3965	71 (83)	P73_3963
<i>Roseovarius pacificus</i> 81-2	-	-	60 (71)	FRBR_v1_150317	66 (80)	FRBR_v1_290081	81 (89)	FRBR_v1_300036	69 (83)	FRBR_v1_300034
<i>Thermincola potens</i> JR	-	-	42 (67)	TherJR_0829	41 (57)	TherJR_0208	72 (84)	TherJR_0143	73 (85)	TherJR_0145
<i>Desulfomonile tiedjei</i> DSM6799	-	-	28 (53)	Desti_5520	40 (58)	Desti_2763	66 (89) 72 (82)	Desti_5437 Desti_5447	70(84) 70(85)	Desti_5438 Desti_5449
<i>Desulfurispora thermophila</i> DSM16022	-	-	46 (67)	AQWNv1_20249	40 (56)	AQWNv1_40042	71 (83)	AQWNv1_40006	71 (86)	AQWNv1_40004
<i>Desulfotomaculum alcoholivorax</i> DSM16058	-	-	42 (63)	DESALv110045	39 (56)	DESALv1110033	71 (83)	DESALv160093	73 (87)	DESALv160091
<i>Thermosediminibacter oceani</i> DSM16646	-	-	36 (58)	Toce_1499	60 (76)	JIBV01_v1_7850001	69 (82)	Toce_1533	71 (86)	Toce_1535
<i>Desulfobacula</i> sp. TS	-	-	28 (43)	JIBV01_v1_1180004	43 (62)	ACZV_v1_40720	60 (75)	JIBV01_v1_3490007	69 (84)	JIBV01_v1_3490005

^a d'après Bringel *et al.* 2019

^b par rapport la référence de *Hyphomicrobium* sp. MC1 de MaGe <https://mage.genoscope.cns.fr>

^c par rapport à la référence de *M. extorquens* DM4 de MaGe <https://mage.genoscope.cns.fr>

^d % d'identité ; (% de positifs)

^d % d'identité ; (% de positifs)

gras : souches confirmées comme pouvant croître avec le CH₃Cl

Le gène *purU* est présent dans le cluster *cmu* chez les souches dégradantes isolées de l'environnement telles que *M. extorquens* CM4 et *Hyphomicrobium* sp. MC1, mais elles feraient figure d'exception parmi les souches présentées dans le Tableau 3.5. Ainsi, les souches bactériennes répertoriées dans Bringel et al. 2019 qui possèdent les gènes *cmu* ne possèdent pas *purU* ni *purT* (résultats non montrés pour *purT*) dans la quasi-totalité des cas (Tableau 3.5). L'ensemble des souches possèdent néanmoins *purN* pour la synthèse des purines. La croissance avec le CH₃Cl de toutes les souches dans le Tableau n'a pas été vérifiée expérimentalement. Pour les souches pour lesquelles c'est le cas (en gras dans le Tableau), les deux souches appartenant au genre *Aminobacter* ne possèdent pas *purU*. Peut-être alors que le produit d'un autre gène que *purU* effectue la réaction de déformylation. Peut-être même qu'il s'agit de FtfL, dont le gène est présent chez la plupart de ces souches qui effectue cette réaction. D'après Kim et al. 2020, la vitesse de réaction de la ligation entre le formiate et le H₄F réalisée par FtfL apparaît comme très dépendante de la concentration en formiate (K_m de 2 mM par rapport au formiate contre 0,2 mM pour le H₄F et 150 μ M pour l'ATP). Les auteurs font l'hypothèse que lorsque la concentration en formiate est forte, FtfL va plutôt catalyser la condensation du formiate et du H₄F en formyl-H₄F. Lorsqu'elle est faible, l'enzyme va plutôt catalyser la réaction inverse (la séparation du formyl-H₄F en formiate et en H₄F). Dans les souches ne possédant pas PurU, si la concentration en formiate reste faible, peut-être FtfL va-t-elle plutôt catalyser la déformylation du formyl-H₄F en formiate, qui permet ensuite une dernière oxydation du composé ainsi que la libération du H₄F. Pour les deux souches d'*Aminobacter* présentées dans le Tableau 3.5, il s'est également avéré que pour chacune d'entre elles, *ftfL* est un gène présent à proximité du cluster de gènes *cmu* (retrouvé au sein de l'îlot génomique abritant également les gènes *cmu*) et qu'ainsi ce gène participe au fonctionnement de la voie *cmu*. Il se pourrait également qu'il s'agisse d'une autre enzyme que FtfL. Après tout, j'ai démontré que FtfL ne peut pas remplir la fonction de PurU chez *M. extorquens*. Les gènes *purU* et *ftfL* sont présents tous les deux chez *M. extorquens* CM4 et *Hyphomicrobium* sp. MC1. Peut-être que l'environnement dans lequel vivent les souches (où l'on trouve à la fois du CH₃Cl et des composés comme le méthanol) les a contraintes à conserver les deux gènes afin de maximiser leurs capacités de croissance. De fait, l'expérience d'évolution expérimentale que j'ai menée n'avait que le CH₃Cl comme principale pression de sélection.

3.9. Conclusions

J'ai mis en évidence que l'inactivation de *FtfL* représente une adaptation majeure de la croissance avec le CH_3Cl , possiblement pour forcer vers l'oxydation totale du carbone issu du CH_3Cl dans les souches. Le phénomène observé est ce que l'on appelle en anglais « trade off » (littéralement "compromis", c'est-à-dire que l'amélioration d'un phénotype va entraîner une détérioration d'un autre phénotype) et qui est fréquemment rencontré dans les études d'évolution expérimentale (Marx 2019). Ici en l'occurrence, l'acquisition de la capacité de croissance avec le CH_3Cl s'est accompagnée d'une perte de croissance avec d'autres composés en C1. On peut aussi dessiner d'autres voies d'adaptation *ftfL*-indépendantes. Trois souches évoluées issues de D0 contiennent un gène *ftfL_{wt}* (D2E, D4E et D7S). Elles partagent différentes mutations ponctuelles touchant une des SU de l'ARN polymérase ADN-dépendante, qui ont été largement documentées comme permettant une reprogrammation rapide du transcriptome. Cela permettrait d'aider la croissance des souches soumises à évolution expérimentale au laboratoire (Conrad et al. 2010 ; Dalldorf et al. 2023). De plus, D4E et D7S partagent une délétion de plusieurs gènes et qui pourrait aussi avoir contribué à l'adaptation de ces souches.

Enfin et comme discuté dans le Chapitre suivant, d'autres chemins d'adaptation à la croissance en CH_3Cl ont vraisemblablement été suivis pour chacune des souches évoluées dérivant d'AM1, avec une mutation dans le gène *cmuA* impactant potentiellement la voie *cmu* de manière directe.

Chapitre 4. Des mutations touchant la voie *cmu* chez les souches évoluées dérivées de la souche AM1

Deux types récurrents de modification génétique observés uniquement dans le génome des souches évoluées en contexte AM1 ont été identifiés.

- D'une part, les souches évoluées dans le contexte AM1 présentent toutes des mutations ponctuelles dans le gène *cmuA* dont le produit, une CH₃Cl-corrinoïde méthyltransférase, effectue avec CmuB, une méthyltransférase tétrahydrofolate-dépendante, la première étape de la voie *cmu*. Ces deux gènes sont ainsi essentiels pour la croissance avec le CH₃Cl. Ainsi, une inhibition totale de l'activité CmuAB n'est pas attendue mais plutôt une modulation de son activité car c'est une enzyme essentielle à la croissance avec le CH₃Cl. En effet, les souches évoluées peuvent toujours effectuer la déshalogénation et croître avec le CH₃Cl (Figure 2.5).
- D'autre part, la formation de coïntégrats entre les plasmides pJM105 et p2META, un plasmide natif de 37,9 kb de la souche AM1 a également été mise en évidence (cf. Chapitre 2 ; et plus particulièrement les sous-sections 2.2.5 à 2.2.7). On peut s'attendre à ce que ces modifications génétiques impactent la voie *cmu* en modifiant l'activité de la CH₃Cl déshalogénase, soit directement par les mutations ponctuelles dans le gène *cmuA*, soit de son expression par suite des changements structuraux du plasmide portant les gènes de la voie *cmu*. Dans ce Chapitre, je présente les expériences effectuées pour analyser l'effet de ces modifications génétiques sur l'adaptation des souches évoluées AM1 pour la croissance avec le CH₃Cl.

4.1. Des variations ponctuelles d'acides aminés dans le domaine N-terminal de CmuA

Toutes les mutations dans *cmuA* observées chez les souches évoluées du contexte AM1 conduisent à des changements d'acides aminés à différentes positions du domaine N-terminal de CmuA (Figure 4.1A). Un modèle prédictif de la structure tridimensionnelle de CmuA a été obtenu par Alphafold. La représentation en 3D de la structure de CmuA prédite permet de constater que les acides aminés concernés par les mutations observées sont prédits comme étant proches spatialement dans une région centrale du domaine N-terminal de CmuA

(Figure 4.1B). A ce jour, toutefois, les sites d'activité méthyltransférase de la protéine CmuA n'ont pas encore été caractérisés expérimentalement. Néanmoins, le domaine N-terminal est similaire chez *M. extorquens* CM4 à d'autres méthyltransférases alors que le domaine C-terminal à des protéines fixant un cofacteur corrinoïde (Vannelli et al. 1999) et les sites sont proches de résidus dont les équivalents chez *M. extorquens* CM4 pourraient lier un atome de zinc et être impliqués dans l'activité de CmuAB (Studer 2001).

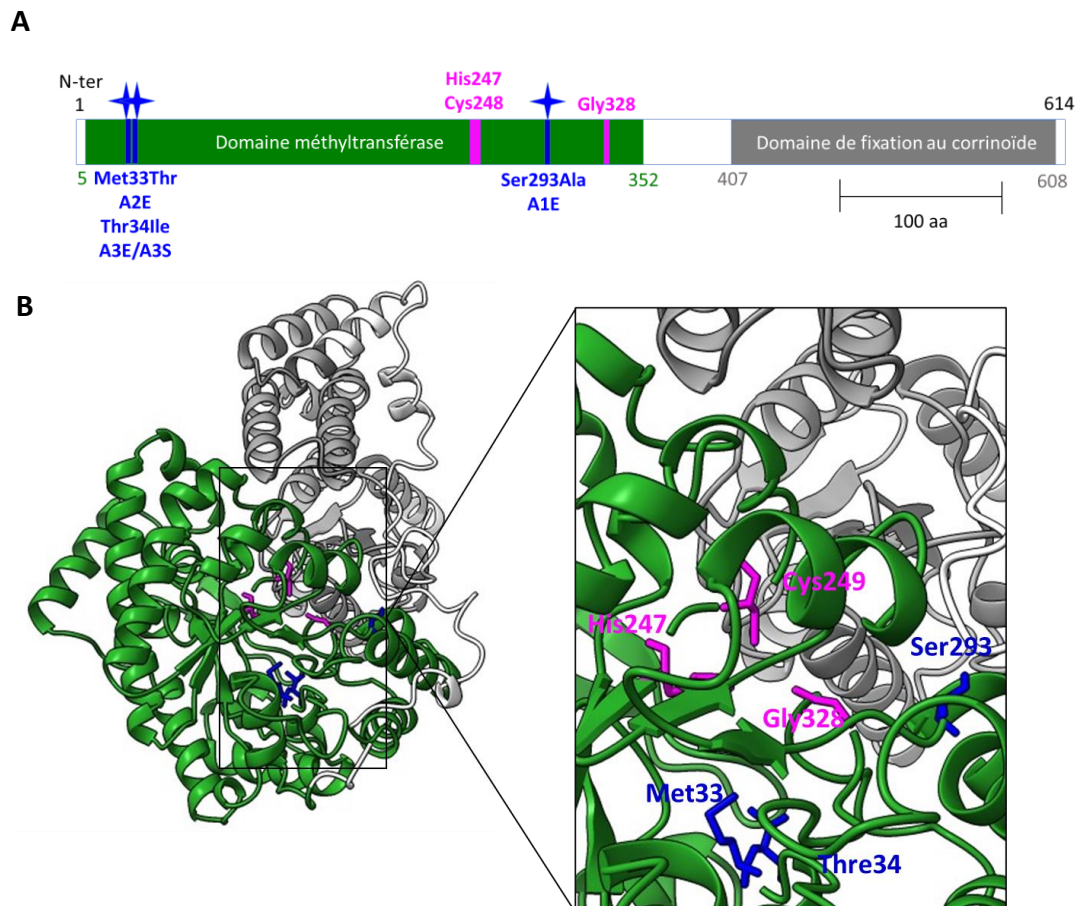


Figure 4.1 Les souches évoluées AM1 présentent toutes un changement d'acide aminé dans le domaine N-terminal de la protéine CmuA

En **A**, l'organisation en deux domaines de la protéine CmuA est représentée. Longue de 608 acides aminés, elle comprend un domaine N-terminal (montré en couleur verte) et un domaine C-terminal où se lie le cofacteur corrinoïde (en couleur grise). Toutes les souches AM1 évoluées portent une mutation entraînant un changement d'acide aminé dans la première partie du gène *cmuA* codant pour le domaine N-terminal. Le nom des souches évoluées portant ces mutations est indiqué. En magenta sont indiqués des résidus pouvant lier un atome de zinc pouvant être impliqué dans l'activité catalytique de l'enzyme. En **B**, la structure de CmuA de *Hyphomicrobium* sp. MC1 (Uniprot <https://www.uniprot.org>, numéro d'accèsion F8J7J8) a été prédite avec le logiciel AlphaFold. Les acides aminés modifiés chez les souches évoluées dérivant d'AM1 par rapport à la protéine CmuA sauvage de la souche MC1 sont indiqués.

Ces éléments sont une première indication forte que les mutations dans *cmuA* ont vraisemblablement un effet sur l'activité enzymatique de déshalogénation du CH₃Cl, ce qui pourrait avoir un effet phénotypique sur la croissance des souches avec le CH₃Cl.

4.2. Le bénéfice de mutations dans CmuA est mineur par rapport à celui conféré par l'inactivation de *ftfL* pour la croissance avec le CH₃Cl

Dans un premier temps, j'ai cherché à mettre en évidence une corrélation entre la présence des mutations dans le gène *cmuA* et le temps de génération des souches avec le CH₃Cl. La croissance avec le CH₃Cl des 4 souches évoluées du contexte AM1 a été comparée entre elles. Les souches A3E et A3S présentent des temps de génération plus courts ($14,1 \pm 0,4$ et $15,3 \pm 1,3$ heures, respectivement) que les souches A1E et A2E ($30,8 \pm 1,4$ et $26,5 \pm 2,5$ heures, respectivement). Sur l'ensemble des souches évoluées, A1E et A2E sont les deux souches avec la croissance avec le CH₃Cl la plus lente (Tableau 2.1). Pour rappel, les souches A3E et A3S portent toutes deux une mutation dans *ftfL* qui pourrait expliquer leur temps de génération plus faible avec le CH₃Cl (Chapitre 3). Ainsi, en dehors de la mutation bénéfique apportée par l'inactivation du gène *ftfL*, l'effet observé des mutations *cmuA* dans des régions différentes du domaine méthyltransférase de CmuA trouvées uniquement chez les souches évoluées A1E et A2E suggère qu'elles ne contribuent pas de manière significative à une croissance rapide avec le CH₃Cl.

Ainsi, la présence d'une mutation inactivant *ftfL* semble dominer l'effet de mutations ponctuelles dans *cmuA* dans le contexte de coïntégrats des souches évoluées issues du contexte AM1. Il est probable que les mutations dans *cmuA* ont un effet sur l'activité de l'enzyme notamment si elles conduisent à des substitutions de résidus du site actif. Si ces positions dans la séquence protéique sont impliquées dans l'activité de CmuA, on peut s'attendre à ce que les résidus correspondants soient conservés chez les souches utilisant le CH₃Cl. C'est ce que j'ai cherché à évaluer.

4.3. Les changements d'acides aminés dans CmuA chez les souches évoluées AM1 ont pris place à des positions conservées de la séquence protéique

La séquence protéique de CmuA des souches évoluées AM1 a été comparée à celles de 26 séquences entières et alignées de CmuA. Seules 9 de ces séquences correspondent à celles de

souches dont l'activité de déshalogénéation du CH₃Cl a été démontrée expérimentalement (indiquées en caractères gras dans la Figure 4.2).

Les résidus modifiés dans la protéine CmuA des souches évoluées du contexte AM1 sont localisés dans des positions conservées (position 33, méthionine changée en thréonine ; position 293, serine changée en alanine) voire semi-conservées (position 34, thréonine changée en isoleucine). Ils ne sont pas retrouvés dans les séquences des protéines caractérisées expérimentalement.

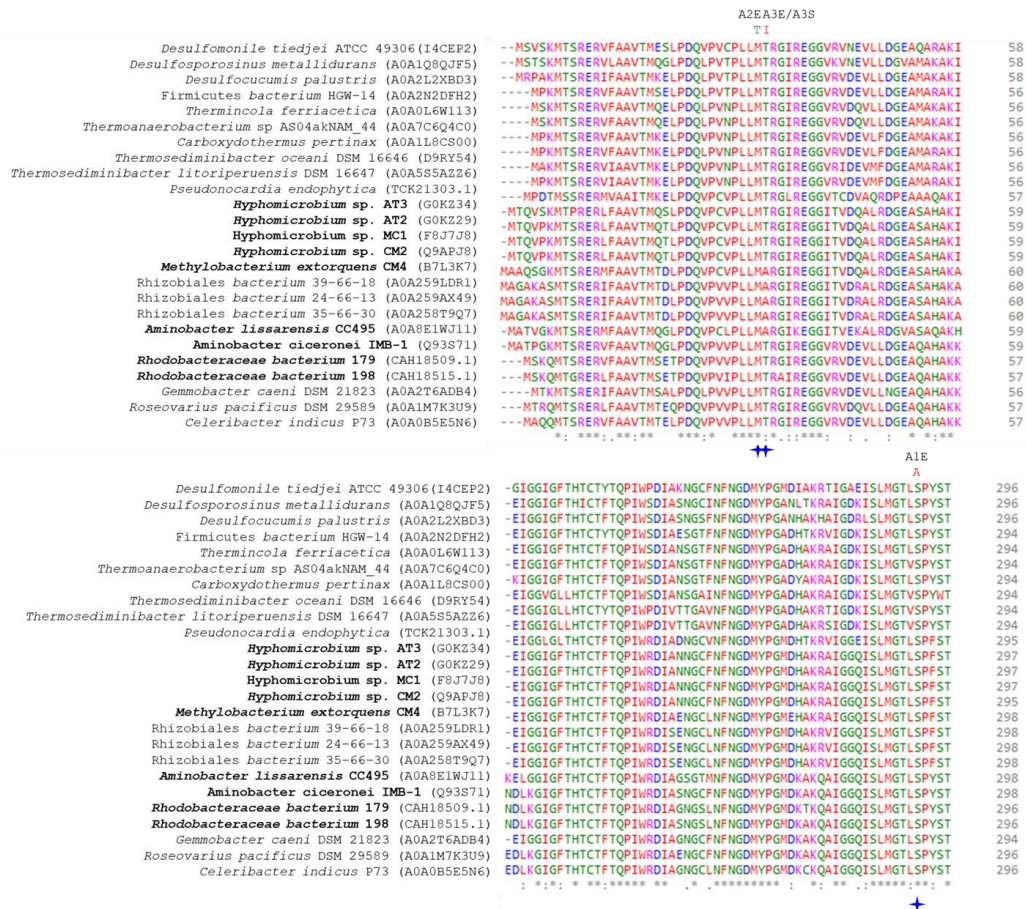


Figure 4.2 Alignement des séquences protéiques de CmuA et comparaison des positions présentant un changement d'acide aminé dans les souches évoluées AM1

Les 26 séquences protéiques de CmuA récupérées dans la base de données NCBI pour les souches possédant les gènes *cmu* (Bringel et al. 2019) ont été alignées avec le logiciel Clustal Omega. Pour les souches évoluées du contexte AM1 où les changements d'acides aminés sont survenus, une étoile bleue est indiquée sous les symboles des résidus conservés (*: acide aminé identique sur l'ensemble des séquences ; : acide aminé identique pour une majorité de séquences). L'emplacement et l'identité des résidus mutés sont indiqués (position 293 pour A1E, 33 pour A2E et 34 pour A3E/A3S).

Pour A1E et A3E/A3S, un résidu avec une chaîne latérale polaire remplace un résidu hydrophobe ; c'est l'inverse pour A2E. Les résidus polaires peuvent former des liaisons

hydrogène contrairement aux résidus hydrophobes, ce qui pourrait affecter la structure du domaine méthyltransférase de CmuA. Si l'on considère les séquences des 9 souches confirmées expérimentalement comme capables de se développer avec le CH₃Cl (Figure 4.2, en gras), on retrouve à chaque fois les mêmes acides aminés à ces positions, sauf le résidu thréonine en position 34 qui est une alanine chez *Aminobacter lissarensis* CC495 et *M. extorquens* CM4 (en position 34). Par ailleurs, la sérine en position 293 a été retrouvée sur l'ensemble de 47 séquences de *cmuA* additionnelles étudiées précédemment (Chaignaud et al. 2018). Si les résidus à ces positions sont aussi conservés, il est probable qu'ils jouent un rôle dans la stabilité et/ou l'activité de la protéine CmuA et que leur modification pourrait avoir un effet sur le fonctionnement optimal de l'enzyme.

4.4. Détermination de l'effet des modifications génétiques plasmidiques sur l'activité CH₃Cl-déshalogénase

Les mutations ponctuelles dans des régions conservées de la protéine CmuA chez les souches évoluées AM1 n'inactivent pas l'activité de l'enzyme puisque ces souches sont capables de déchlorer et de croître avec le CH₃Cl (Figure 2.5). Mon hypothèse est que ces mutations conduisent à une diminution d'activité de déshalogénation du CH₃Cl chez les souches évoluées, contrairement aux souches curées portant le plasmide pJM105 libre et originel, peut-être pour ralentir le flux de carbone en provenance du CH₃Cl qui entraînerait l'apparition de stress trop brutaux pour la cellule (production d'ions Cl⁻, stress acide entre autres). Cette perte d'activité résulterait de la combinaison de variations ponctuelles du domaine méthyltransférase de la protéine CmuA et d'une expression altérée des gènes *cmu* à la suite de la formation du coïntégrat pJM105-p2META. En appui de cette hypothèse, il avait déjà été montré que la formation d'un coïntégrat entre p2META et un plasmide similaire à pJM105 (pCM410) porteur des gènes de la voie d'oxydation du formaldéhyde de *Paracoccus denitrificans* chez *M. extorquens* AM1 à la suite d'une évolution expérimentale visant à améliorer la croissance de la souche avec le méthanol par l'utilisation de cette voie avait conduit à une baisse de l'expression des gènes portés par pCM410 (Chou & Marx 2012 ; Lee et al. 2013).

Pour tester l'impact des mutations obtenues, j'ai dosé l'activité déshalogénase dans les souches évoluées du contexte AM1 et leurs dérivés curés après réintroduction du plasmide pJM105 natif sous forme libre et sans mutation dans *cmuA*.

4.4.1. Choix des conditions de culture et de dosage des activités de déshalogénéation

L'activité de la première étape de la voie *cmu* catalysée par la déshalogénase formée par CmuAB a été évaluée selon un protocole précédemment décrit (Nadalig et al. 2011) impliquant la quantification de la production d'ions chlorure en suspension de cellules, sachant que pour chaque molécule de CH₃Cl déshalogénée, un ion chlorure est libéré.

Ceci étant, l'expression des gènes *cmu* requiert la présence de CH₃Cl (Studer et al. 2001), de sorte que le dosage de l'activité de déshalogénéation nécessite des bactéries cultivées en présence de CH₃Cl. Or, les souches de départ 105A et 105D et également les souches curées A1EL (pJM105) et A2EL (pJM105) incapables de se développer avec ce composé faisaient partie des souches à évaluer. Afin d'assurer des conditions de culture homogènes pour l'ensemble des échantillons, les souches ont donc été cultivées dans un mélange de CH₃Cl (afin d'induire l'expression des gènes *cmu*) et de succinate, un substrat qui permet de soutenir la croissance de l'ensemble des souches testées (pour rappel, les souches A3S, D7E portant des mutations dans *ftfL* sont incapables de se développer avec le méthanol). Il avait été montré précédemment que lorsque *M. extorquens* CM4 était cultivée avec un mélange de méthanol et de CH₃Cl, l'induction des gènes *cmu* était intermédiaire (158 ± 20 nmoles de méthylène-H₄folate formé.min⁻¹.mg de protéines⁻¹) entre une condition avec le méthanol seul (pas d'activité observée) et le CH₃Cl seul (245 ± 44 nmoles.min⁻¹.mg de protéines⁻¹) (Studer et al. 2001). Il est possible que la situation soit similaire avec le succinate et le CH₃Cl, où la présence de succinate est associée à une expression plus faible des gènes *cmu* qu'avec le CH₃Cl seul.

J'ai d'abord testé l'activité de déshalogénéation du CH₃Cl chez la souche *Hyphomicrobium* sp. MC1, le plasmide pJM105 portant les gènes *cmu* sous le contrôle du promoteur natif de cette souche (Michener et al. 2016). J'ai testé l'activité de déshalogénéation du CH₃Cl chez MC1 dans deux conditions, avec le CH₃Cl comme seul substrat carboné pour la croissance, et avec un mélange de CH₃Cl et de succinate, bien que le succinate ne soit pas un substrat privilégié pour *Hyphomicrobium* sp. MC1 (Hartmans et al. 1986 n'avaient pas obtenu de croissance). Les résultats sont présentés dans la Figure suivante.

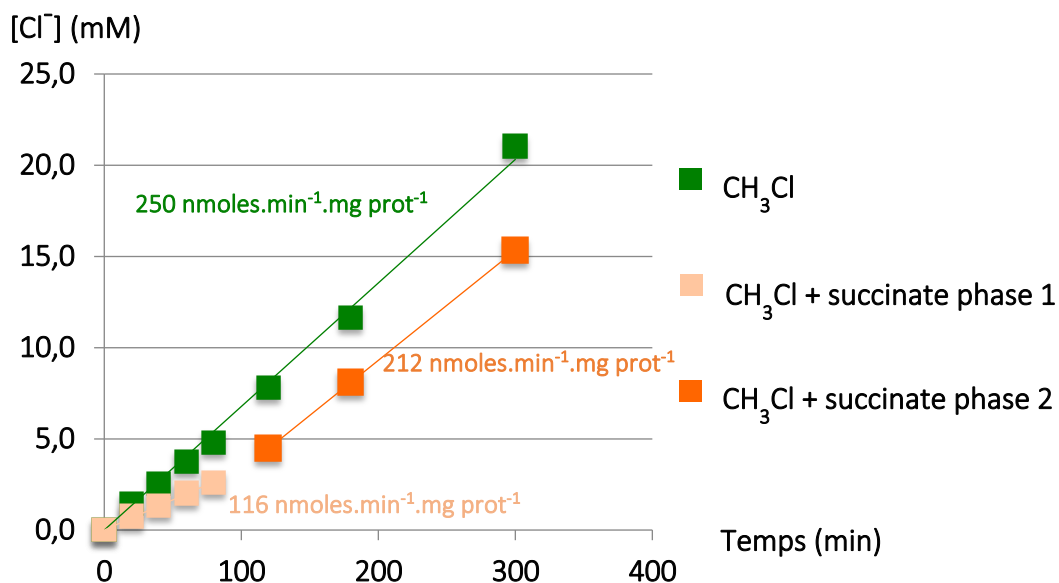


Figure 4.3 Comparaison de l'activité de déshalogénation de MC1 cultivée en présence de CH₃Cl avec ou sans succinate

La cinétique de variation de la concentration d'ions chlorure a été mesurée à intervalles réguliers après ajout du CH₃Cl à une suspension de la souche MC1 cultivée soit avec du CH₃Cl soit avec un mélange de succinate et de CH₃Cl puis récoltée en phase exponentielle de croissance puis lavée pour éliminer les traces de succinate au moment du dosage. Un seul réplicat de bactéries a été cultivé avec le CH₃Cl (15 mM) seul (en vert), ou avec un mélange de CH₃Cl et de succinate (4 mM) (orange). Deux phases successives d'activité sont observées si les bactéries sont préparées à partir d'une culture développée avec un mélange de succinate et de CH₃Cl.

La cinétique de déshalogénation d'*Hyphomicrobium* sp. MC1 après culture en mélange CH₃Cl et succinate présente deux phases (Figure 4.3) : une première phase où l'activité est plus faible (116 nmoles.min⁻¹.mg de protéines⁻¹), puis une deuxième phase où l'activité se rapproche de celle d'une culture développée avec le CH₃Cl seul. On observe que la présence de succinate dans le milieu de culture est associée à un temps de latence avant l'augmentation de l'activité qui pourrait s'expliquer par une induction retardée ou plus faible de l'expression des gènes *cmu* en présence de succinate.

Je n'ai pas trouvé de données publiées sur une interférence potentielle du succinate sur la régulation (induction par le CH₃Cl) de l'expression des gènes *cmu* chez une souche dégradant le CH₃Cl. Néanmoins, les observations de l'expérience décrite en Figure 4.3 m'ont conduit à mesurer l'activité de déshalogénation au moins deux heures après le début de l'incubation dans les expériences qui ont suivi.

4.4.2. Une activité de déshalogénation similaire des souches évoluées mutantes pour CmuA et des souches curées correspondantes exprimant un gène *cmuA* de type sauvage

J'ai dosé l'activité CmuAB de l'ensemble des souches après cultures dans un mélange succinate et CH₃Cl.

L'activité de déshalogénation de trois des souches évoluées dérivant d'AM1 (A1E, A2E et A3S) et celle de leurs équivalents curés après réintroduction du pJM105 d'origine sous forme libre et présentant un allèle de type sauvage de *cmuA* a été mesurée. Si une différence était observée entre l'activité de déshalogénation de la souche curée et de la souche mutée d'origine, elle serait alors sûrement due à la présence des deux types de mutation présents sur le plasmide pJM105 des souches évoluées AM1 (mutation dans *cmuA*, formation de coïntégrat).

J'ai également dosé l'activité de 105D, la souche de départ en contexte DM4, afin de comparer son activité avec la souche D7E pour évaluer la possibilité d'une modification de l'activité de déshalogénation en contexte DM4.

Tableau 4.1 Activités de déshalogénation des souches évoluées dérivant d'AM1 ainsi que de leurs équivalents curés

Souche	Variant CmuA	Support génétique des gènes <i>cmu</i>	Nombre de copies de pJM105/chromosome	Temps de génération (heures) CH ₃ Cl 15 mM ^a	Activité de déshalogénation (nmoles Cl ⁻¹ .min ⁻¹ .mg de protéines ⁻¹)
105A	WT	pJM105	3,25	pas de croissance	27 ± 8 (3)
105D	WT	pJM105	3,08	pas de croissance	24 ± 2 (2)
A1E	Ser293Ala	p2META::pJM105	0,36	30,8 ± 1,4 (4) ^b	70 ± 1 (2) ^b
A1EL (pJM105)	WT	pJM105	ND	AC ^c	86 ± 14 (2)
A2E	Met33Thr	p2META::pJM105	0,10	26,5 ± 2,5 (4)	140 ± 10 (2)
A2EL (pJM105)	WT	pJM105	ND	AC	81 ± 11 (2)
A3S	Thr34Ile	p2META::pJM105	1,11	15,3 ± 1,3 (3)	134 ± 29 (4) ^e
A3SL (pJM105)	WT	pJM105	ND	ND ^d	119 ± 21 (4)
D7E	WT	pJM105	2,39	11,8 ± 0,7 (3)	191 ± 10 (4)
MC1	WT	chromosome	NA ^g	6,7 ± 0,5 (5)	240 ± 15 (2) ^f

^a Cultures avec addition de tétracycline 12,5 µg/mL

^b () Nombre de réplicats biologiques

^c AC = Aucune croissance

^d ND = La souche croît mais le temps de génération n'a pas été déterminé

^e Test de Tukey suite à une ANOVA n'a pas mis en évidence de différence d'activité de déshalogénation entre A3S et A3SL(pJM105) mais en a mis une en évidence entre A3S/A3SL(pJM105) et 105A ou A3S/A3SL(pJM105) et D7E

^f mg de protéines déterminé avec kit BCA et non pas avec Qubit protéines comme les autres

^g NA = Non Applicable

La comparaison des activités de déshalogénation (Tableau 4.1) a permis de constater plusieurs points remarquables du point de vue de l'expérience d'évolution expérimentale effectuée :

a) **Une augmentation de l'activité de déshalogénation du CH₃Cl est observée chez les souches évoluées par rapport aux souches de départ dans les conditions testées.** Les souches de départ 105A et 105D présentent une activité de déshalogénation similaire (autour de 25 nmoles.min⁻¹.mg protéines⁻¹) et nettement inférieure à l'ensemble des souches testées, y compris des souches évoluées et de leurs équivalents curés possédant pJM105.

b) **L'augmentation de l'activité de déshalogénation des souches évoluées issues de AM1 apparaît moindre que celle observée pour la souche évoluée D7E issue du**

contexte DM4. A ce stade, il n'est pas possible de déterminer si cela reflète une différence de contexte génomique entre AM1 et DM4 car aucune autre souche évoluée en contexte DM4 n'a été testée (Tableau 4.1).

c) **Les mutations plasmidiques du coïntégrat pJM105-p2META n'ont pas eu d'effet bénéfique détectable pour les 2 souches évoluées de AM1.** Dans deux cas, aucune différence d'activité n'est observée entre une souche évoluée et son équivalent curé : A1E comparée à A1EL (pJM105) et A3S comparée à A3SL (pJM105). Il est possible que la souche A3SL (pJM105) ait perdu une partie de p2META durant le curage : l'amplification de l'ISMex25 n'a pas été possible pour cette souche-là à la différence des autres souches curées, alors qu'un autre gène plus éloigné de p2META (p2META010) a pu être amplifié. Cela sera discuté plus en détail dans le Chapitre 5 de Perspectives.

d) **L'activité de déshalogénation accrue chez la souche évoluée A2E pourrait résulter de l'insertion d'une IS de type ISMex4 en amont de *cmuB*.** La souche évoluée A2E a une activité de déshalogénation mesurée bien supérieure à celle de son équivalent curé A2EL (pJM105) (Tableau 4.1). Or le coïntégrat pJM105-p2META porte une IS de type ISMex4 en amont de *cmuB*, qui est le premier gène du cluster *cmu*. Il a déjà été montré que l'insertion d'une IS en amont de gènes d'intérêt pouvait apporter un promoteur et ainsi augmenter leur expression (Chou et al. 2009). Un dosage par RT-qPCR des transcrits des gènes *cmuA* et *cmuB* comparé chez A2E, la souche parentale A0 et la souche A2EL (pJM105) serait nécessaire pour montrer que l'augmentation d'activité de déshalogénation du CH₃Cl observée est associée à une expression transcriptionnelle accrue des gènes *cmu*.

Outre l'acquisition de mutations de la voie *cmu* portées par les coïntégrats pJM105-p2META chez les souches évoluées en contexte AM1, **d'autres mutations chromosomiques pourraient également moduler l'expression des gènes *cmu*** et ainsi avoir un effet sur l'activité de déshalogénation et potentiellement améliorer leur fitness en croissance avec le CH₃Cl. Par exemple, la souche A1E a acquis une mutation dans un régulateur transcriptionnel (famille Ros/MucR ; META1_4919). Cette famille comprend principalement des répresseurs

transcriptionnels (Janczaek 2022). Si la mutation inactive ce régulateur (changement d'acide aminé S72W) ou au moins l'altère et qu'il s'agit en effet d'un répresseur transcriptionnel, l'expression de nombreux gènes pourrait augmenter, impacter la résistance aux stress voire augmenter l'expression des gènes *cmu*. La souche A2E possède également une mutation dans un régulateur transcriptionnel (META1_4441 de type Histidine-Kinase) qui pourrait avoir des effets pléiotropiques. Quant à la souche A3S, bien qu'il n'y ait pas de mutations dans des régulateurs transcriptionnels, elle possède une mutation dans deux gènes (META1_2897 et META1_5170) de fonction inconnue et qui pourraient peut-être aussi agir sur l'expression des gènes *cmu*.

4.5. Conclusions

Les souches évoluées en contexte AM1 portent des mutations du gène *cmuA* dans la première partie du gène (correspondant au domaine N-terminal de la protéine), ainsi que du support génétique des gènes de la voie *cmu* (plasmide pJM105 intégré dans le plasmide endogène p2META combiné à des événements d'insertion d'IS). Concernant les premières, le fait d'avoir trouvé des mutations directement dans un gène *cmu* aussi central, et avec une telle récurrence chez le contexte AM1, m'a fait émettre l'hypothèse qu'elles ont contribué à diminuer l'efficacité du catabolisme du CH₃Cl de la voie *cmu*. Ces deux types de modifications génétiques pourraient respectivement moduler l'expression de l'ensemble des gènes *cmu* et agir directement sur l'activité de CmuA qui va conditionner le flux de carbone de l'ensemble de la voie métabolique. Néanmoins, la mesure de l'activité de déshalogénation des souches évoluées AM1 et de leurs équivalents curés après réintroduction de pJM105 ne suggère pas d'effet de la présence de ces mutations sur l'activité. Dans le Chapitre 5 de perspectives, je discuterai de différentes hypothèses qui pourraient néanmoins contribuer à expliquer ces résultats et comment elles pourraient être vérifiées par l'expérience.

Chapitre 5. Conclusions et Perspectives

Au cours de ma thèse, j'ai cherché à élucider la nature des adaptations de *M. extorquens* à la croissance avec le CH₃Cl après acquisition par transfert horizontal des gènes de la voie *cmu*. Par évolution expérimentale, j'ai démontré qu'il était possible de faire évoluer deux souches de *M. extorquens* (AM1 et DM4), initialement inadaptées à la croissance avec le CH₃Cl, vers une meilleure utilisation de ce composé.

Le séquençage du génome de 13 souches évoluées comparé à celui du génome des souches de départ (105A et 105D) a révélé 64 mutations dans 45 gènes : des mutations ponctuelles comme des SNP, ou des délétions ou insertions d'un ou plusieurs nucléotides, ainsi que d'autres événements génétiques (réarrangement entre plasmides, grandes délétions). Les quelques loci mutés partagés entre plusieurs génomes de souches évoluées pointent des pistes évolutives convergentes.

Les résultats phares sont :

- L'inactivation de FtfL, une enzyme essentielle à la croissance avec le méthanol, améliore la croissance avec le CH₃Cl après acquisition de pJM105
- Les gènes *nuo* pourraient favoriser l'utilisation du CH₃Cl comme source d'énergie
- Des chemins évolutifs FtfL-indépendants ont été détectés et diffèrent selon le contexte génétique de la souche receveuse des gènes de la voie *cmu*
- Les souches évoluées en contexte *M. extorquens* AM1 ont toutes des mutations du plasmide pJM105 (*cmuA* muté et cointégration avec un plasmide endogène)
- Quelques souches évoluées en contexte *M. extorquens* DM4 ont des mutations chromosomiques de fonctions de régulation
- Des souches génétiquement modifiables dégradant le CH₃Cl par la voie *cmu* ont été obtenues.

Dans ce chapitre, les perspectives découlant de ces résultats sont discutées.

5.1. L'inactivation de FtfL suite à l'acquisition des gènes de la voie *cmu* est-elle une réponse adaptative utile dans l'environnement ?

Les souches évoluées ont dans leur majorité des mutations dans le gène *ftfL* qui inactivent la formyl-H₄F ligase FtfL encodée par le gène. L'activité FtfL est dommageable pour la croissance avec le CH₃Cl dans les conditions de l'expérience d'évolution expérimentale effectuée dans ce travail. Au contraire, elle est essentielle à la croissance avec du méthanol et d'autres composés en C1 impliquant l'oxydation de formaldéhyde en formiate (Marx et al. 2003b ; Vannelli et al. 1999). Dans le contexte de l'apport de gènes *cmu*, l'inactivation de FtfL suffit à lever l'incapacité de croissance en CH₃Cl de *M. extorquens* D0 initialement incapable de croître avec le CH₃Cl. J'ai fait l'hypothèse que l'inactivation de l'enzyme FtfL est bénéfique pour l'assimilation du CH₃Cl car son activité s'oppose à celle de la formyl-H₄F hydrolase PurU (dont le gène est apporté par le plasmide pJM105). De plus, FtfL entraînerait une utilisation supplémentaire du cofacteur H₄F nécessaire au fonctionnement de la déshalogénéation du CH₃Cl, avec pour conséquence de rendre limitant ce cofacteur indispensable à la première étape de la voie *cmu* (voir paragraphe 3.6 du Chapitre 3).

L'inactivation du gène *ftfL* est certes une adaptation bénéfique pour la croissance avec le CH₃Cl dans les conditions testées au laboratoire, mais l'est-elle dans les conditions rencontrées dans l'environnement ? *M. extorquens* CM4 et *Hyphomicrobium* sp. MC1 isolées de l'environnement sont des souches modèles de la dégradation *cmu*-dépendante du CH₃Cl. Ces souches sont capables de croître avec le méthanol en plus du CH₃Cl, indiquant que FtfL est restée fonctionnelle dans ces souches. Ainsi l'inactivation du gène *ftfL* ne semble pas un chemin évolutif retenu chez des souches dégradant le CH₃Cl isolées de l'environnement, contrairement aux souches évoluées. L'acquisition de mutations dans le gène *ftfL* pourrait même restreindre les chances de survie en conditions naturelles comme discuté ci-après. En effet, l'inactivation de FtfL chez une souche méthylotrophe s'accompagne d'une réduction de sa flexibilité métabolique d'utilisation de divers substrats carbonés (méthanol, formiate) plus abondants dans la nature que le CH₃Cl, la rendant moins compétitive par rapport aux souches ayant un métabolisme méthylotrophe intègre. Alors que le méthanol est un composé abondant émis principalement par les plantes (100 Tg par an ; Galbally & Kirstine 2002 ; Yurimoto et al. 2021), les concentrations environnementales de CH₃Cl, de l'ordre du

nanomolaire au picomolaire, peuvent être utilisées par des méthylotrophes du sol (Borodina et al. 2005 ; Yoshida et al. 2006 ; Jaeger et al. 2018). Dans des sols forestiers, l'utilisation du méthanol permet d'accroître celle du CH_3Cl en permettant le développement de communautés méthylotrophes dégradant le CH_3Cl (Chaignaud et al. 2018). Lorsque le méthanol est abondant dans l'eau de mer lors d'efflorescences phytoplanctoniques (Heikes et al. 2002), le méthylotrophe marin HTCC 2181 utilise le CH_3Cl comme source d'énergie (Halsey et al. 2012). Cette situation d'utilisation du CH_3Cl comme source d'énergie et un autre composé en C1 comme source de carbone n'aurait pas été sélectionnée au cours de l'évolution expérimentale où seulement une concentration élevée de CH_3Cl (15 mM) éloignée des conditions environnementales a été testée, et sans pression de sélection pour l'utilisation d'autres substrats. Ainsi, une adaptation entraînant la perte de croissance avec un des principaux substrats de croissance méthylotrophe de la souche n'apparaît pas comme la piste qui pourrait être privilégiée dans l'environnement, où les souches sont exposées à une multitude de substrats carbonés. Les conditions de l'évolution expérimentale avec un apport d'une seule source de carbone en concentration bien supérieure à la concentration retrouvée dans l'environnement (CH_3Cl est de l'ordre du 450 pptv) ne reflètent probablement pas les conditions auxquelles les communautés méthylotrophes utilisatrices de CH_3Cl sont exposées dans l'environnement. Ainsi, le design de l'évolution expérimentale a probablement favorisé l'acquisition des mutations inactivant *ftfL*.

5.2. La modification des conditions de l'évolution expérimentale permettrait-elle la découverte d'autres mécanismes d'adaptation à la croissance avec le CH_3Cl ?

Il aurait été difficile d'utiliser une concentration aussi faible que celle dans l'environnement dans une expérience d'évolution expérimentale en milieu liquide. L'utilisation de plusieurs concentrations de CH_3Cl dans des cultures en évolution en parallèle permettrait de tester l'effet de la concentration en CH_3Cl sur le type de mutations acquises par les souches évoluées. Par ailleurs, l'inactivation de *FtfL* est apparue dès la première phase de l'expérience d'évolution expérimentale, alors que le succinate était présent en plus du CH_3Cl (démonstré pour la lignée D7, Chapitre 3). Si le méthanol avait été utilisé à la place du succinate, la pression de sélection pour le maintien d'une enzyme *FtfL* fonctionnelle (essentielle à la croissance avec du méthanol mais pas avec du succinate) aurait vraisemblablement défavorisé l'acquisition de

mutations dans le gène *ftfL* et privilégié des voies alternatives comme chez les souches D2E, D7S ou D4E (paragraphe 5.4, voir discussion adaptations *ftfL*-indépendantes). Nous ne sommes donc pas à l'abri d'avoir observé ici ce que l'on pourrait qualifier de "quick fix", c'est-à-dire le chemin évolutif le plus rapide qui permet aux souches de surmonter la pression de sélection et d'utiliser au plus vite le CH₃Cl.

L'acquisition de pJM105 par une souche hôte représente l'apport d'un ensemble de 9 gènes (*cmuBCA fmdB paaE hutI metF2 purU fold*) qui a probablement impacté le profil de mutations des souches évoluées, en particulier du fait de la présence de *purU*. L'état fonctionnel ou non de FtfL n'a pas eu d'effet détectable quant à l'essentialité de PurU dans la voie *cmu* d'utilisation du CH₃Cl chez *M. extorquens* (Chapitre 3 paragraphe 3.7). On peut néanmoins se demander s'il aurait été possible d'obtenir des souches évoluées après acquisition du plasmide pJM112, un plasmide dérivé de pJM105 sans le gène *purU* (Michener et al. 2016). On se rappelle en effet que *purU* n'est pas systématiquement détecté dans le génome de souches possédant les gènes *cmu*. Deux souches déchlorant le CH₃Cl, *Aminobacter ciceronei* IMB-1 et *Aminobacter lissarensis* CC495, sont dépourvues de *purU* mais portent deux copies de gènes *ftfL* dont l'une à proximité des gènes *cmu*. À l'opposé, *Celeribacter indicus* P73 possède 4 copies du gène *purU* en plus des gènes de la voie *cmu*, mais est dépourvue de gène *ftfL* (Tableau 3.5). Étant donné que les enzymes FtfL et PurU catalysent une réaction métabolique méthylotrophe de sens opposé (Figure 3.6), et que FtfL peut théoriquement aussi catalyser la même réaction que PurU, on peut émettre l'hypothèse que FtfL pourrait remplacer PurU selon les contextes génétiques. Pour découvrir des fonctions qui lèveraient l'essentialité du gène *purU*, une expérience d'évolution expérimentale en contexte *M. extorquens* de souches possédant pJM112 serait pertinente.

5.3. Une augmentation de l'expression des gènes *nuo* est-elle bénéfique à la croissance avec le CH₃Cl ?

Une mutation inactivant *ftfL* est bénéfique pour la croissance par utilisation du CH₃Cl, sans toutefois conférer une croissance optimale (Chapitre 3, le mutant D0_*ftfL*_{D7E} et la souche évoluée D7EL (pJM105) ont un temps de génération de 19,6 ± 4,8 h contre 12,0 ± 0,8h, respectivement), indiquant qu'une progression est encore possible. D'autres mutations ont

sans doute contribué à l'évolution des souches évoluées issues de D0 et se développant le plus rapidement avec le CH₃Cl, mais elles restent cependant à identifier et à caractériser.

Les souches D1E, D3E, D5E et D7E (Tableau 2.6) possèdent toutes des mutations dans P_{nuo} , un promoteur expérimentalement confirmé se situant en amont du gène *nuoA* (Maucourt et al. 2022). Or, certaines protéines appartenant au complexe Nuo ont une abondance accrue chez *M. extorquens* CM4 cultivée avec du CH₃Cl par rapport aux cellules cultivées avec du méthanol (Roselli et al. 2013). Les sous-unités de la NADH déshydrogénase (ou NADH-quinone oxydoréductase) Nuo forment un complexe qui est l'équivalent chez les mitochondries du premier complexe de la chaîne respiratoire effectuant l'oxydation des cofacteurs NADH réduits avec extrusion de protons dans l'espace intermembranaire (Falk-Krzesinski & Wolfe 1998). Je fais l'hypothèse qu'une plus grande expression des gènes *nuo* est bénéfique pour la croissance avec le CH₃Cl en permettant un meilleur recyclage des équivalents réduits vers leur forme oxydée. Pour le démontrer, la quantification des transcrits des gènes *nuo* par RT-qPCR pourrait être comparée à partir d'ARN extraits des souches évoluées porteuses d'une mutation dans P_{nuo} , de la souche parentale et de souches de référence de la dégradation du CH₃Cl après croissance avec le CH₃Cl (et succinate comme référence). A partir des mêmes cultures, la détermination de la concentration des formes oxydées des nucléotides (NAD⁺ et NADP⁺) et des formes réduites (NADH et NADPH) (kits commerciaux disponibles) permettrait de voir si le pool en cofacteurs est impacté par l'expression accrue de P_{nuo} en condition de croissance en CH₃Cl.

A ce jour, le caractère essentiel du complexe Nuo pour la croissance avec le CH₃Cl n'est pas connu alors qu'il est dispensable pour la croissance avec le méthanol chez *M. extorquens* PA1 (Ochsner et al 2017). De même, chez *M. extorquens* DM4, la délétion de sous-unités du complexe Nuo maintient une croissance de type sauvage en méthanol et en succinate, contrairement à la croissance avec le CH₂Cl₂ (résultats non publiés de l'équipe AIME). Si les gènes *nuo* sont essentiels à la croissance avec le CH₃Cl, on s'attend à une absence de croissance d'un mutant $D0\Delta nuo_ftfL_{D7E}$ (pJM105) contrairement à un mutant $D0_ftfL_{D7E}$ (pJM105). Si les mutations dans P_{nuo} des souches évoluées ont eu un effet bénéfique pour leur croissance avec le CH₃Cl, on pourrait aussi s'attendre à une croissance accélérée après remplacement de P_{nuo_wt} par P_{nuo_D7E} (mutation du promoteur P_{nuo} chez D7E) dans la souche

D0_ *ftfL*_{D7E} (pJM105). Néanmoins, comme ces mutations ne se sont produites que chez des souches évoluées ayant subi une inactivation de FtfL, il est possible qu'elles ne soient pas suffisantes pour permettre la croissance avec le CH₃Cl. Introduire la mutation P_{*nuo*_D7E} chez D0 (possédant toujours *ftfL* sauvage) permettrait de tester si la seule mutation du promoteur des gènes *nuo* suffit à conférer la croissance avec le CH₃Cl, et ceci indépendamment de l'état de *ftfL*.

5.4. Analyse des mutations FtfL-indépendantes partagées en contexte génétique D0

Les trois souches issues de D0 (D2E, D4E et D7S) ayant conservé le gène *ftfL* sauvage possèdent toutes une mutation dans la partie C-terminale d'une sous-unité de l'ARN polymérase ADN-dépendante. Chez D4E et D7S, il y a changement d'un acide aminé de RpoA, la sous-unité alpha, respectivement L302I et E304V. Chez D2E, la sous-unité bêta (RpoC) présente un changement d'acide aminé T1312A (Tableau 2.6). Les gènes *rpo* sont souvent mutés chez des souches issues d'évolution dirigée (Phaneuf et al. 2019). Ces mutations participeraient à une reprogrammation rapide du transcriptome où une seule mutation a un effet pléiotropique de diminution de l'expression de gènes de la réponse aux stress tout en augmentant l'expression de gènes qui conduisent à une croissance plus rapide (Dalldorf et al. 2023). Ce type de mutations touchant l'ARN polymérase ADN-dépendante de souches évoluées au laboratoire modifie l'équilibre entre une réponse ciblée face aux stress et une réponse rapide touchant des fonctions de croissance. On note que la souche D6E (gène *ftfL* muté) présente une mutation dans le gène de l'ARN polymérase RpoD codant pour un facteur sigma. De telles mutations pourraient ainsi contribuer à une meilleure croissance des souches indépendamment de l'état fonctionnel de FtfL. Elles joueraient un rôle dans l'accélération de la croissance des souches n'ayant pas l'avantage d'une enzyme FtfL inactivée et qui serait encore plus lente sans la présence de ces mutations touchant l'ARN polymérase ADN dépendante. Puisque ces modifications de sous-unités de cette enzyme sont souvent observées lors d'expériences d'évolution dirigée, elles semblent peu inféodées à une pression de sélection particulière, sauf quelques cas particuliers (Dalldorf et al. 2023). C'est pourquoi une étude plus approfondie de ces mécanismes n'est pas envisagée comme une perspective immédiate de mon travail de thèse.

Par ailleurs, une délétion chromosomique de gènes de régulation est apparue au cours de l'évolution expérimentale en contexte *M. extorquens* DM4. La délétion de 8 gènes a été détectée chez la souche évoluée D7S (chromosome position 298810-307071) et fait partie d'une délétion plus large de 221 gènes chez D4E (chromosome position 232492-392332). La délétion commune d'une taille de 8,3 kb est flanquée d'éléments génétiques mobiles (IS) et comprend :

- i. Trois gènes d'un système de régulation à deux composantes : METDI0317 à MEDTI0319 présentant 44 à 57% d'identité au niveau protéique avec le système FixLJ de bactéries diazotrophiques fixatrices d'azote (Anthamatten et Hennecke 1991 ; revue Price et al. 2021). FixLJ est un système hème-dépendant de régulation à deux composantes qui est sensible aux faibles teneurs en O₂ (microaérobie) et induit la biosynthèse de cytochromes chez *Bradyrhizobium japonicum* (Nellen-Anthamatten et al. 1998).
- ii. Un gène codant pour un cytochrome c présumé de classe I (METDI0322).
- iii. deux gènes de fonction inconnue portant un domaine codant pour des protéines de réponses au stress (METDI0320 et METDI0323).
- iv. un gène codant pour une ATPase pour le transport de cations de type P présumée (METDI0324).

La perte d'un ou de plusieurs gènes de cette région de 8,3 kb pourrait avoir un effet bénéfique pour la croissance avec le CH₃Cl, en modifiant par exemple l'expression de fonctions clé de la réponse adaptative au CH₃Cl, peut-être en relation avec la biosynthèse de cytochrome et l'utilisation de CH₃Cl comme source d'énergie. Une première expérience pour valider cette hypothèse serait de construire par mutagenèse dirigée un mutant délété de cette région de 8,3kb (appelé Δ_{D7S}) chez la souche parentale D0. Une comparaison des caractéristiques de croissance par utilisation de différentes sources de carbone dont le CH₃Cl entre le mutant D0 Δ_{D7S} , la souche D0 et la souche évoluée D7S, permettrait de tester si la délétion Δ_{D7S} suffit à elle seule à conférer la capacité de croissance avec le CH₃Cl.

Cette région est dupliquée dans le chromosome de *M. extorquens* DM4 (identité de séquence >85% entre METDI0317 à MEDTI0324 et METDI4569 à METDI4566). Cette région est également dupliquée chez *M. extorquens* AM1, et pourtant la mutation n'a été observée que dans le contexte DM4. Peut-être observerait-on une stratégie similaire chez AM1 si plus de réplicats d'évolution expérimentale avaient été réalisés.

5.5. Comment les mutations du plasmide pJM105 impactent-elles la voie *cmu* d'utilisation du CH₃Cl chez les souches évoluées en contexte *M. extorquens* AM1 ?

Les souches évoluées AM1 présentent toutes des modifications génétiques avec formation d'un coïntégrat entre pJM105 et un plasmide endogène, ainsi que des mutations dans *cmuA* codant l'enzyme effectuant la première étape de la voie *cmu*. La comparaison des activités de déshalogénation entre les souches évoluées AM1 et leurs équivalents curés possédant le plasmide pJM105 originel n'a pas montré de différence significative. De plus, l'activité de déshalogénation chez les souches évoluées AM1 (ou leurs équivalents curés possédant le plasmide pJM105 originel) est bien supérieure à la souche de départ 105A. L'ensemble de ces observations suggère que les mutations plasmidiques sont mineures par rapport aux mutations chromosomiques des souches évoluées en contexte AM1. Pour le confirmer, il aurait fallu disposer d'une souche parentale porteuse d'un coïntégrat et tester si cette seule acquisition lui permet de croître avec le CH₃Cl. Pour ce faire, le transfert du coïntégrat d'une taille attendue de 55 kb peut être tenté par transformation de cellules électrocompétentes de A0 à partir de l'ADN plasmidique extrait de souches évoluées portant le coïntégrat. L'obtention de transformants Tc^R pourrait néanmoins être compromise par des efficacités d'extraction et d'électroporation faibles étant donné la taille du coïntégrat.

Une expérience directe pour tester l'effet des mutations ponctuelles de *cmuA* sur l'activité de la CH₃Cl-déshalogénase serait de remplacer l'allèle *cmuA* sauvage par un allèle muté sur pJM105 et de l'introduire chez la souche de départ A0. L'activité de déshalogénation du CH₃Cl de la souche pourrait ensuite être testée après induction par le CH₃Cl (culture en succinate suivie d'une exposition au CH₃Cl). En particulier, l'enzyme (avec ou sans variation dans sa séquence protéique) peut être purifiée et son activité dosée selon un protocole décrit précédemment (Studer et al. 2001).

Je fais l'hypothèse que les modifications plasmidiques apparues en début d'évolution chez les souches évoluées du contexte *M. extorquens* AM1 réduisent la vitesse de fonctionnement de la voie *cmu* et ceci par réduction de l'activité de la CH₃Cl déshalogénase (mutation ponctuelle dans la partie N-terminale de CmuA) et peut-être aussi par réduction de l'expression des gènes *cmu* (suite à la formation du coïntégrat). Cette hypothèse serait discutée notamment dans le cadre de propositions d'expériences (voir paragraphes 5.5.2).

5.5.1. L'incompatibilité des plasmides pJM105 et p2META explique-t-elle la formation systématique de coïntégrat sans autre bénéfice que la conservation des gènes *cmu* ?

La formation de coïntégrat entre un plasmide 'exogène' et le plasmide endogène p2META observée ici dans la souche AM1 avait déjà été signalée dans des expériences d'évolution expérimentale (Chou & Marx 2012). La coïntégration entre pJM105 et p2META a conduit à l'interruption du gène *trfA* qui est essentiel à la réplication de pJM105 (Lin & Helinski 1992). En conséquence, la conservation au cours de l'expérience des gènes de la voie *cmu* portés par le coïntégrat plasmidique dépend du système réplcatif de p2META. Ainsi, pour éviter la formation de coïntégrat à l'issue de l'expérience d'évolution expérimentale, l'utilisation d'un vecteur de clonage des gènes *cmu* compatible avec les réplicons natifs de la cellule parentale et n'appartenant pas au groupe de réplication IncP (dont fait partie pJM105) est envisageable sur la base de nouveaux vecteurs développés récemment chez *M. extorquens* AM1 (Carrillo et al. 2020). Une évolution expérimentale avec par exemple les gènes *cmu* portés par le plasmide pABC-DM4K ou pABC-CM4K (origines de réplication de respectivement *M. extorquens* DM4 et CM4) permettrait de voir s'il n'y a en effet pas de formation d'un coïntégrat lorsque l'on supprime l'incompatibilité entre les plasmides endogènes de la souche receveuse par HGT d'un plasmide porteur des gènes *cmu*.

La formation systématique de coïntégrat entre pJM105 et p2META pose la question de l'essentialité des gènes portés par p2META. La difficulté à obtenir des souches curées Tc^S ayant perdu les gènes *cmu* (Tableau 2.2) suggère qu'au moins une portion du plasmide p2META est essentielle à la croissance des souches en succinate, la source de carbone utilisée pour les essais de curage de pJM105. La perte partielle de séquence de p2META du coïntégrat chez les souches curées pourrait être vérifiée. Des essais d'amplification PCR en utilisant un jeu d'amorces spécifique à p2META ont été effectués pour les souches évoluées AM1 ainsi que leurs équivalents curés. Dans le cas de la souche curée A3SL, le jeu d'amorces ciblant ISMex25_1 n'a pas donné d'amplification (contrairement à une PCR ciblant le gène p2META010), alors qu'une amplification s'est produite pour les deux sites chez les autres souches évoluées curées AM1. Cela pourrait signifier qu'une partie de p2META a été excisée en même temps que pJM105 lors de l'obtention de la souche curée A3SL. Pour confirmer la présence de séquences du p2META dans le génome de souches évoluées curées et non curées

(notamment de A3SL par rapport à A3S et A0), une approche de Southern Blot avec un jeu de plusieurs sondes permettant de couvrir l'entièreté de p2META pourrait être appliquée afin de délimiter la présence ou l'absence de segments de p2META chez les souches d'intérêt.

L'obtention des souches curées à partir des souches évoluées AM1 a été difficile, peut-être du fait de gènes essentiels à la croissance portés par p2META. Cette difficulté n'a pas été observée pour les souches évoluées DM4 qui ne portent pas un réplicon incompatible avec le plasmide pJM105. On ne peut pas exclure que d'autres événements génétiques mutationnels aient pris place lors du curage du coïntégrat des souches évoluées AM1. Or le plasmide pJM105 originel réintroduit chez les souches curées A1EL et A2EL, contrairement à A3SL, ne permet pas de rétablir la croissance avec le CH₃Cl. Après confirmation des phénotypes de croissance des souches curées, un séquençage du génome de souches évoluées curées d'AM1 permettrait d'identifier des événements (réarrangements intra- ou inter-réplicon, mutations chromosomiques ou plasmidiques) qui se seraient produits au cours de l'étape de curage des souches évoluées d'AM1 afin de comparer leurs génomes. De fait, aucune différence significative entre l'activité de déshalogénéation n'a été observée entre A1EL (pJM105) et A1E, indiquant que l'expression des gènes *cmu* n'est pas différente entre la souche curée et la souche non curée et suggère une perte de gènes essentiels à la croissance avec le CH₃Cl.

5.5.2. Quel est le bénéfice de mutations plasmidiques affectant les gènes *cmu* ?

Le bénéfice de mutations ralentissant la vitesse de fonctionnement de la voie *cmu* serait de permettre une exposition moindre et échelonnée des stress associés à l'utilisation du CH₃Cl qui permettrait à la cellule en condition de croissance ralentie d'y faire face. Cette hypothèse implique que les mutations plasmidiques se seraient produites en début d'expérience d'évolution expérimentale, lorsque la souche 105A a été confrontée pour la première fois au CH₃Cl, en jouant le rôle de première ligne d'adaptation des souches. Des mutations chromosomiques acquises ultérieurement favoriseraient ensuite une croissance plus rapide en CH₃Cl par augmentation de l'expression des gènes de la voie *cmu*. En faveur de cette hypothèse, j'ai dosé une activité de déshalogénéation chez les souches évoluées AM1 isolées en fin d'expérience d'évolution expérimentale bien supérieure à la souche de départ 105A. Ceci aurait pour effet de minorer l'effet des mutations plasmidiques acquises lors des premiers

transferts et expliquerait qu'aucune différence significative d'activité de déshalogénation n'ait été détectée entre souches évoluées et souches curées pour A1E et A3S.

Puisque l'hypothèse formulée implique que les mutations plasmidiques ont été acquises en début d'expérience d'évolution expérimentale, on pourrait le vérifier en déterminant la chronologie d'acquisition des mutations de *cmuA* dans les cultures A1, A2 et A3 au cours de l'expérience. Il serait ici possible d'extraire de l'ADN de cultures conservées à -80°C correspondants à plusieurs points de l'évolution expérimentale. Par amplification qPCR à partir des ADN extraits, l'utilisation d'amorces discriminant l'allèle muté de l'allèle sauvage permettrait d'établir un ratio de l'abondance relative de l'allèle *cmuA* sauvage et muté normalisé par rapport à celle d'un gène de référence (*tetR*), et ce tout au long de l'expérience. L'hypothèse prédit qu'en fin d'expérience, l'expression des gènes *cmu* serait accrue dans des cultures de souches évoluées par rapport aux souches parentales. L'abondance des transcrits du gène *cmuA* pourrait aussi être comparée entre des souches parentales et évoluées après induction de l'expression des gènes *cmu* par le CH₃Cl (suspension cellulaire exposée au CH₃Cl après culture en succinate) par quantification RT-qPCR. Pour déterminer l'impact de l'augmentation de la transcription des gènes *cmu* sur le fitness, on pourrait également cloner les gènes *cmu* de *Hyphomicrobium* sp. MC1 dans un plasmide sous contrôle d'un promoteur fort et constitutif de *M. extorquens*. Un tel plasmide serait par la suite introduit dans la souche D0_ *ftfL*_{D7E} pour en mesurer le temps de génération lors de la croissance avec le CH₃Cl et le comparer à D0_ *ftfL*_{D7E}(pJM105). Si l'hypothèse est correcte, on s'attend à une amélioration du fitness de la souche exprimant les gènes *cmu* à partir du promoteur fort par rapport au promoteur natif P_{*cmu*}.

5.6. Un nouvel outil pour l'étude génétique de la croissance avec le CH₃Cl : une opportunité à saisir pour la recherche du régulateur de la voie *cmu*

Les souches évoluées que j'ai pu obtenir au cours de cette étude constituent un outil supplémentaire pour l'étude du fonctionnement de la voie *cmu*. Jusqu'à présent, les méthodes de mutagenèse ciblée par remplacement allélique ne sont pas applicables aux souches modèle d'étude de la dégradation du CH₃Cl que sont *M. extorquens* CM4 et *Hyphomicrobium* sp. MC1, limitant les possibilités d'études approfondies des étapes

métaboliques de la voie *cmu*. Au contraire, il est possible de manipuler génétiquement les souches AM1 et DM4 (Marx et al. 2008 ; Gruffaz et al. 2014). Ainsi, comme je l'ai fait pour *ftfL*, un échange allélique peut désormais être effectué chez des souches de *M. extorquens* capables de croître avec le CH₃Cl. Une première cible parmi les gènes chromosomiques candidats pour une étude approfondie de leur rôle dans la croissance avec le CH₃Cl, serait une mutation des gènes *nuo* (voir paragraphe 5.3).

Plusieurs gènes au sein du cluster *cmu* codent pour des protéines de fonction non caractérisée alors même que leur rôle est essentiel (ou du moins bénéfique) à la croissance ou la déshalogénation du CH₃Cl comme montré précédemment (Vannelli et al. 1999 ; Michener et al. 2016). Le gène *fmdB* du cluster *cmu* peut être apparenté à *fmdB* retrouvé chez *Methylophilus methylotrophus* (29% d'identité par rapport au gène de MC1 sur 84% de la longueur qui encode un possible régulateur transcriptionnel pour la formamidase elle-même encodée par *fmdA* (Wyborn et al. 1996). Pour tester son rôle de régulateur transcriptionnel du cluster *cmu*, une version de pJM105 arborant une délétion $\Delta fmdB$ est disponible (Michener et al. 2016, plasmide pJM109). Introduire ce plasmide dans plusieurs des souches évoluées curées permettrait de déterminer si l'abondance des transcrits *cmu* (déterminée par RT-qPCR) ou les capacités de croissance avec le CH₃Cl (déterminée par la mesure du temps de génération) sont modifiées en fonction de la présence ou non du gène. Une meilleure connaissance de la régulation du promoteur P_{cmuA} permettrait potentiellement l'amélioration du seuil de détection de CH₃Cl par le biorapporteur basé sur son utilisation (Ul Haque et al. 2013).

Enfin et de manière plus large, et même si l'ensemble de mes travaux de thèse ont été réalisés en étudiant la voie *cmu*, il est clair qu'il existe également d'autres voies de dégradation du CH₃Cl (Nadalig et al. 2014), dont les modes d'expression et de régulation sont encore à caractériser. Le moment venu, les adaptations post-transfert après acquisition d'autres voies d'utilisation pourront être comparées à celles identifiées ici pour la voie *cmu*. On peut ainsi imaginer que l'utilisation exhaustive des approches d'évolution expérimentale telles qu'explorées dans le présent travail permettront de mettre en évidence, à l'exemple de la croissance avec le CH₃Cl, la diversité des solutions évolutives ouvertes aux bactéries pour

optimiser leurs capacités de croissance, dans un contexte donné et en fonction de leur contexte génomique, lui-même reflet de l'historique de leur évolution passée.

Chapitre 6. Matériel et Méthodes

6.1. Matériel biologique

6.1.1. Souches bactériennes

Les souches sont conservées à -80°C dans un cryotube avec 10% de glycérol. Une colonie est inoculée dans 10 mL de milieu M3C10 avec 4 mM de succinate (source de carbone et présence d'antibiotique selon la souche à conserver). La culture en phase stationnaire est concentrée 10X après une centrifugation 10 min à 8000 rpm et mélangée à 250 µL de glycérol 50% dans un cryotube.

Tableau 6.1 Inventaire des souches

Affiliation & nom	Nom stock	Génotype/origine	Référence
<i>Escherichia coli</i>			
DH5α	LFM800	<i>fhuA2 lacΔU169 phoA glnV44 Φ80' lacZΔM15 gyrA96 recA1relA1 endA1 thi-1 hsdR17</i>	Taylor et al. 1993
Helper	LFM801	HB101 (F ⁻ , <i>thi-1</i> , <i>hsdS20</i> (r _B ⁻ , m _B ⁻), <i>supE44</i> , <i>recA13</i> , <i>ara-14</i> , <i>leuB6</i> , <i>proA2</i> , <i>lacY1</i> , <i>galK2</i> , <i>rpsL20</i> (str ^r), <i>xyl-5</i> , <i>mtl-1</i>) portant le plasmide pRK2073	Leong et al. 1982
JM109	LFM802	<i>endA1</i> , <i>recA1</i> , <i>gyrA96</i> , <i>thi</i> , <i>hsdR17</i> (r _k ⁻ , m _k ⁺), <i>relA1</i> , <i>supE44</i> , Δ(<i>lac-proAB</i>), [F' <i>traD36</i> , <i>proAB</i> , <i>lacI</i> ^q ZΔM15].	Yanisch-Perron et al. 1985
<i>Hyphomicrobium</i> sp.			
MC1	LFM803	WT (sauvage)	Vuilleumier et al. 2011
<i>Methylobacterium extorquens</i>			
105A	LFM807	souche ancestrale A0 portant le plasmide pJM105	Cette thèse
105D	LFM808	souche ancestrale D0 portant le plasmide pJM105	Cette thèse
62A	LFM809	souche ancestrale A0 portant le plasmide pCM62	Cette thèse
62D	LFM810	souche ancestrale D0 portant le plasmide pCM62	Cette thèse
A0	LFM804	AM1 Δ <i>cel katA::mCherry</i>	Michener et al. 2014a
CM4	LFM805	WT (sauvage)	Vannelli et al. 1999
D0	LFM806	DM4 Δ <i>cel hpt::Venus</i> Δ <i>dcm</i>	Michener et al. 2014a
D0_Δ <i>ftfL</i> _{D7E}	LFM837	D0 muté portant l'allèle <i>ftfL</i> _{D7E} (C>T 444623) ^a	Cette thèse
A1E	LFM811	isolat de l'évolution expérimentale de 105A, culture A1	Cette thèse
A1EL	LFM824	isolat Tc ^S dérivé de A1E incapable de croître en CH ₃ Cl	Cette thèse
A2E	LFM812	isolat de l'évolution expérimentale de 105A, culture A2	Cette thèse

A2EL	LFM825	isolat Tc ^S dérivé de A2E incapable de croître en CH ₃ Cl	Cette thèse
A3E	LFM813	isolat de l'évolution expérimentale de 105A, culture A3	Cette thèse
A3EL	LFM826	isolat Tc ^S dérivé de A3E incapable de croître en CH ₃ Cl	Cette thèse
A3S	LFM814	isolat de l'évolution expérimentale de 105A, culture A3	Cette thèse
A3SL	LFM827	isolat Tc ^S dérivé de A3S incapable de croître en CH ₃ Cl	Cette thèse
D1E	LFM815	isolat de l'évolution expérimentale de 105D, culture D1	Cette thèse
D1EL	LFM828	isolat Tc ^S dérivé de D1E incapable de croître en CH ₃ Cl	Cette thèse
D2E	LFM816	isolat de l'évolution expérimentale de 105D, culture D2	Cette thèse
D2EL	LFM829	isolat Tc ^S dérivé de D2E incapable de croître en CH ₃ Cl	Cette thèse
D2S	LFM817	isolat de l'évolution expérimentale de 105D, culture D2	Cette thèse
D2SL	LFM830	isolat Tc ^S dérivé de D2S incapable de croître en CH ₃ Cl	Cette thèse
D3E	LFM818	isolat de l'évolution expérimentale de 105D, culture D3	Cette thèse
D3EL	LFM831	isolat Tc ^S dérivé de D3E incapable de croître en CH ₃ Cl	Cette thèse
D4E	LFM819	isolat de l'évolution expérimentale de 105D, culture D4	Cette thèse
D4EL	LFM832	isolat Tc ^S dérivé de D4E incapable de croître en CH ₃ Cl	Cette thèse
D4EL_ <i>ftfL</i> _{D7E}	LFM839	D4EL muté pour l'allèle <i>ftfL</i> _{D7E} (C>T 444623) ^a	Cette thèse
D5E	LFM820	isolat de l'évolution expérimentale de 105D, culture D5	Cette thèse
D5EL	LFM833	isolat Tc ^S dérivé de D5E incapable de croître en CH ₃ Cl	Cette thèse
D6E	LFM821	isolat de l'évolution expérimentale de 105D, culture D6	Cette thèse
D6EL	LFM834	isolat Tc ^S dérivé de D6E incapable de croître en CH ₃ Cl	Cette thèse
D7E	LFM822	isolat de l'évolution expérimentale de 105D, culture D7	Cette thèse
D7EL	LFM835	isolat Tc ^S dérivé de D7E incapable de croître en CH ₃ Cl	Cette thèse
D7EL_ <i>ftfL</i> _{D0}	LFM838	D7EL muté portant l'allèle <i>ftfL</i> _{D0} (T>C 444623) ^a	Cette thèse
D7S	LFM823	isolat de l'évolution expérimentale de 105D, culture D7	Cette thèse
D7SL	LFM836	isolat Tc ^S dérivé de D7S incapable de croître en CH ₃ Cl	Cette thèse

^a coordonnées d'après la référence METDI de *M. extorquens* DM4 de MaGe ((<https://mage.genoscope.cns.fr>))

6.1.2. Plasmides

Les plasmides sont conservés dans la souche *E. coli* DH5α ou JM109 à -80°C et sous forme d'ADN purifié à -20°C. L'ensemble des plasmides est décrit dans le tableau 6.2.

Tableau 6.2 Inventaire des plasmides

Nom	Génotype ou description	Référence
pCM62	<i>colE1 oriV oriT(IncP) traJ' trfA tetAR</i>	Marx & Lidstrom 2001
pCM433	<i>colE1 ori, oriT, tet, cat, sacB</i>	Marx 2008
pJM105	pCM62 portant un insert de 7 kb du groupe de gène <i>cmu</i> de <i>Hyphomicrobium</i> sp. MC1 contenant les gènes <i>cmuBCA fmdB paaE hutI metF purU fold</i>	Michener et al. 2016
pJM111	pJM105 $\Delta fold$	Michener et al. 2016
pJM112	pJM105 $\Delta purU$	Michener et al. 2016
pME8348	pCM433 linéarisé par les enzymes de restriction Apal et Agel et portant l'allèle <i>ftfL</i> sauvage (insert de 819 kb correspondant aux positions 443270-444089 amplifié par PCR chez D0 et coupé avec les mêmes enzymes)	Cette thèse
pME8349	pCM433 linéarisé par les enzymes de restriction Apal et Agel et portant l'allèle muté <i>ftfL_{D7E}</i> (insert de 819 kb correspondant aux positions 443270-444089 amplifié par PCR chez D7E et coupé avec les mêmes enzymes)	Cette thèse
pRK2073	<i>colE1 ori, kan::Tn7, tra, trb</i> , dérivé de pRK2013. Plasmide de la souche Helper pour des conjugaisons tri-parentales	Leong et al. 1982

6.1.3. Amorces

Les solutions stock d'amorces sont à une concentration de 100 µM dans des cryotubes et conservées à -20°C. Les solutions à 20 µM (dilution des solutions stock dans de l'eau ultrapure) sont utilisées pour la préparation des milieux réactionnels de PCR. Les séquences des amorces et leurs utilisations sont décrites en détail ci-dessous (Tableau 6.3).

Tableau 6.3 Inventaire des amorces

Cible	Description	Nom de l'amorce	Séquence (5'-3')	T _m (°C)	Taille de l'amplicon (pb)
ColE1- <i>folD</i> de pJM105	hybride séquence ColE1- <i>folD</i> de pJM105 (1359-1378) ^a pour synthèse de la sonde 2	pCM-ColEI-F	AGTCGTGTCTTACCGGGTTG	59,4	1470
	hybride séquence ColE1- <i>folD</i> de pJM105 (2828-2809) ^a pour synthèse de la sonde 2	MC1_foID-R	GCCTTACGGTGTGATGAAGC	59,4	1470
<i>cmuB</i>	hybride <i>cmuB</i> en amont du site de coupure BamHI de pJM105 (10769-10788) ^a pour synthèse de la sonde 1	LFM008	GCGTGAATATCTGGCGATTT	55,2	1397
	hybride <i>cmuB</i> en aval du site de coupure BamHI de pJM105 (12165-12146) ^a pour synthèse de la sonde 1	LFM009	ACCGTACGCATTTGTCTTCC	57,3	1397
<i>cmuB-tetR</i> de pJM105	hybride <i>cmuB</i> d' <i>Hyphomicrobium</i> MC1 (11949-11968) ^a pour identifier la présence de pJM105	cmuB105	TGCTACCGACCAGTAAGCCG	61,4	733
	hybride <i>tetR</i> (12662-12681) ^a pour identifier la présence de pJM105	tetR105	TCGCGGTGATTGTCGATGGA	59,4	733
Délétion D4E	hybride en amont de la grande délétion de D4E Δ232492–392332 (231742-231761) ^b	LFM045	TGATCCACAACGACTGGCCC	61,4	1108 ou 1182 ^f
	hybride en aval de la grande délétion de D4E Δ232492–392332 (392690-392671) ^b . Avec LFM045, amplification uniquement chez D4E	LFM046	CTGGCGTCGCGATCTTGAAG	61,4	1108

	hybride au niveau de la délétion de D4E Δ 232492–392332 (232924-232905) ^b chez D0 mais pas chez D4E. Avec LFM045, amplification uniquement chez D0	LFM047	CGAACAACACCTCGGGCATG	61,4	1182
<i>folD</i>	hybride en aval de <i>folD</i> de pJM105 (2257-2277) ^a pour vérifier la délétion Δ <i>folD</i> chez pJM111	LFM026	ACCCATCTTCCATCGCTACCG	61,8	1155
	hybride en amont de <i>folD</i> de pJM105 (3393-3412) ^a pour vérifier la délétion Δ <i>folD</i> chez pJM111	LFM027	CAAAGGCCGGGACATCGAAC	61,4	1155
<i>ftfL</i>	hybride en aval de <i>ftfL</i> (443270-443288) ^b	LFM014	GCAACGATCCGCCCAAGTC	61,0	1930
	hybride en amont de <i>ftfL</i> (445204-445185) ^b	LFM015	CCCAAACCACCTGTCCCATG	61,4	1930
<i>ftfL</i> (intragénique)	hybride <i>ftfL</i> (443270-443288) ^b avec site Apal en 5' pour construction pME8348 & pME8349	ApalLFM015	TAACAGGGCCCCCAAACCACCTGTCCCATG	73,5	841
	hybride <i>ftfL</i> (444070-444089) ^b avec site AgeI en 5' pour construction pME8348 & pME8349	AgeILFM030	ATCATACCGGTGTCGATGAACTTCTCCGCGC	70,8	841
<i>îlot dcm</i>	hybride ISMdi21 (2557517-2557538) ^b présent chez DM4, absent chez AM1	ISMdi21f	ATGCAATCCTGCATGATCTCTC	58,4	1209
	hybride IS1354 (2558726-2558705) ^b présent chez DM4, absent chez AM1	IS1354r	AGATGGCCAGCAATGAAGCAAG	60,3	1209
ISMex25_1	hybride en aval de ISMex25_1 (12559-12578) ^c pour caractériser le coïntégrat pJM105-p2META	LFM012	CGCAAGCAATGGTCTCGGAG	61,4	1908
	hybride en amont de ISMex25_1 (14448-14467) ^c pour caractériser le coïntégrat pJM105-p2META	LFM013	TACCCGAAGTGCTGACCCTG	61,4	1908
<i>mauA- mauD</i>	hybride <i>mauA</i> (2899267-2899248) ^d présent chez AM1, absent chez DM4	LFM002	TCGTGCAGTGGTAGGTCATC	59,4	810
	hybride <i>mauD</i> (2898457-2898476) ^d présent chez AM1, absent chez DM4	LFM003	GACATCCGCTACGTCGTCTC	61,4	810
METDI0025	hybride METDI0025 (28417-28436) ^a en amont de la mutation C>T 28594 uniquement présente chez D7E	LFM039	AGGATCTCGTGCCTGGTCC	61,4	738
	hybride METDI0025 (29155-29136) ^a en aval de la mutation C>T 28594 uniquement présente chez D7E	LFM040	CCGCCCCGTATCGAACAGGAT	61,4	738
p2META010	hybride p2META010 (8696-8715) ^c pour vérifier la présence de p2META	LFM020	CTGCGACGGACATCCTGAAG	61,4	804
	hybride p2META010 (9500-9481) ^c pour vérifier la présence de p2META	LFM021	CCCAGGCGATGTACCCAAAG	61,4	804

pCM433 (cassette de clonage)	hybride en amont de la cassette de clonage pCM433 (968-986) ^e pour vérification de la présence d'un insert (pME8348 & pME8349)	V433F	ACATTTCCCCGAAAAGTGC	54,5	228 ou 1032 ^g
	hybride en aval de la cassette de clonage pCM433 (1196-1177) ^e pour vérification de la présence d'un insert (pME8348 & pME8349)	V433R	GAGCGTGACAATCACGAAAC	57,3	228 ou 1032 ^g
<i>purU</i>	hybride en aval de <i>purU</i> de pJM105 (3192-3211) ^a pour vérifier la délétion $\Delta purU$ chez pJM112	LFM024	CTCTTGCAATCGAGGCGGAC	61,4	1224
	hybride en amont de <i>purU</i> de pJM105 (4397-4416) ^a pour vérifier la délétion $\Delta purU$ chez pJM112	LFM025	GCCATGCGTGACCAAGATCC	61,4	1224
<i>trfA</i>	hybride <i>trfA</i> (14814-14833) ^a pour caractériser le cointégrat pJM105-p2META	LFM010	CATTGACCCAGGCGTGTTC	61,4	1084
	hybride <i>trfA</i> (15897-15878) ^a pour caractériser le cointégrat pJM105-p2META	LFM011	CTGATCGACGCGGGGTTTC	61,4	1084

^a coordonnées d'après la référence sur Genbank (KT950741.1) de pJM105

^b coordonnées d'après la référence METDI de *M. extorquens* DM4 de MaGe (<https://mage.genoscope.cns.fr>)

^c coordonnées d'après la référence p2META de *M. extorquens* AM1 de MaGe (<https://mage.genoscope.cns.fr>)

^d coordonnées d'après la référence META de *M. extorquens* AM1 de MaGe (<https://mage.genoscope.cns.fr>)

^e coordonnées d'après la référence sur Genbank (EU118176.1) de pCM433

^f produit de 1108 nt avec LFM046 et 1182 nt avec LFM047

^g produit de 228 nt pour pCM433 et 1032 nt pour pME8348 et pME8349

6.2. Cultures bactériennes

6.2.1. Milieux de cultures

Les souches de *M. extorquens* et de *Hyphomicrobium* sp. MC1 sont cultivées dans du milieu M3 (pour *Methylobacterium* Minimum Medium) (Roselli *et al.*, 2013). Ce milieu est dépourvu d'ions Cl^- (noté M3Cl0). Sa composition pour 1 L est : KH_2PO_4 1,9 g Euromedex® ; Na_2HPO_4 5,1 g Sigma-Aldrich® ; $\text{MgSO}_4 \cdot 7 \text{H}_2\text{O}$ 0,1 g Sigma-Aldrich® et $(\text{NH}_4)_2\text{SO}_4$ 0,2 g Sigma-Aldrich®. Le pH du milieu est ajusté à 7,2 avec du NaOH et du HCl. Pour le milieu solide, 15 g d'agar-agar Euromedex® sont rajoutés. Après autoclavage 15 min à 120°C, le milieu gélosé est maintenu en surfusion à 55°C. Le milieu autoclavé refroidi à environ 45°C est supplémenté d'1 mL de solution de $\text{Ca}(\text{NO}_3)_2$ à 25 g.L⁻¹ et d'1 mL d'une solution d'éléments traces dépourvue d'ions Cl^- dont la composition est pour 100 mL : 0,5 mL H_2SO_4 à 95% ; 0,1 g $\text{Fe}(\text{II})\text{SO}_4 \cdot 7\text{H}_2\text{O}$; 19,5 mL d'eau ultra pure et de 10 mL MnSO_4 à 10 g.L⁻¹ ; 10 mL $\text{Na}_2\text{MoO}_4 \cdot \text{H}_2\text{O}$ à 25 g.L⁻¹ ; 10 mL H_3BO_3 à 1 g.L⁻¹ ; 10 mL $\text{CuSO}_4 \cdot 5 \text{H}_2\text{O}$ à 4,1 g.L⁻¹ ; 10 mL $\text{ZnSO}_4 \cdot 7\text{H}_2\text{O}$ à 5,27 g.L⁻¹ ; 10 mL NH_4VO_3 à 1 g.L⁻¹ ; 10 mL $\text{Co}(\text{NO}_3)_2 \cdot 6\text{H}_2\text{O}$ à 2,5 g.L⁻¹ et 10 mL $\text{NiSO}_4 \cdot 7\text{H}_2\text{O}$ à 1,07 g.L⁻¹. Juste avant utilisation, une source de carbone est ajoutée dans le milieu.

6.2.2. Sources de carbone

Le CH_3Cl (MW= 50,49 g.mole⁻¹) est prélevé sous forme de gaz à l'aide d'une seringue stérile dans une bouteille en verre reliée à la bonbonne de CH_3Cl sous pression (Linde AG, Pullach, Allemagne). Le composé à l'état gazeux est injecté par l'intermédiaire d'une aiguille stérile portée par une seringue au travers de bouchons Mininert à valve Thermo Scientific™ étanches au gaz dans les erlenmeyers ou des tubes hungate à vis scellés. Pour une concentration finale à 15 mM, le volume de CH_3Cl injecté est calculé ainsi :

$$V = nRT/p = \frac{(\text{Volume de liquide en L} \times 15 \text{ mmol}) \times 8.3 \text{ J.mol.K} \times 293 \text{ K}}{101325 \text{ Pa}}$$

V= volume en mL

n= quantité de matière en moles

R= constante des gaz parfaits en J.mol⁻¹.K⁻¹

T= température en Kelvin

Pour les cultures sur milieu solide, les boîtes de Pétri sont placées dans des jarres de verre scellées par un cerceau en métal ajustable au couvercle en verre pour une fermeture hermétique au gaz. Le couvercle dispose d'un embout où est vissé un bouchon Mininert par lequel le CH_3Cl est injecté de sorte à représenter 10% du volume de la jarre.

Un indicateur coloré de pH (bleu de bromothymol (BBT, 80 mg/L)) peut être ajouté dans les milieux solides afin de détecter l'acidification produite par déchloration du CH_3Cl (virage du bleu au jaune si acidification).

Les autres sources de carbone (à l'état liquide à température ambiante) sont ajoutées sous forme de solutions stérilisées par filtration (unité de filtration 0,2 μm Nalgene™ Rapid-Flow™) dans le milieu de culture contenu dans des erlenmeyers fermés par un bouchon en coton cardé pour les cultures supérieures à 10 mL et dans des tubes Falcon™ pour celles de moins de 10 mL. **Le succinate** est rajouté pour avoir une concentration finale de 4 ou 0,2 mM pour le milieu liquide et 10 mM pour le milieu solide à partir d'une solution stock stérile à 1 M préparée à partir de sodium succinate hexahydraté ($\text{MW}=270,1 \text{ g.mole}^{-1}$ Sigma-Aldrich®). **Le méthanol** est rajouté pour avoir une concentration finale à 16 mM pour le milieu liquide et 40 mM pour le milieu solide à partir d'une solution stock à 800 mM dans de l'eau ($78,11 \text{ g.mole}^{-1}$ Fisher Chemical™). **Le formiate** est rajouté pour avoir une concentration de 16 mM en culture liquide à partir d'une solution stérile stock à 1 M préparée à partir de formiate de sodium ($\text{MW}=68,01 \text{ g.mole}^{-1}$ Sigma-Aldrich®).

Les souches de *E. coli* utilisées sont cultivées en milieu LB (Bertani 1951) dont la composition est pour 1L : 10 g tryptone ; 5 g d'extrait de levure ; 10 g NaCl.

6.2.3. Antibiotiques

Des solutions d'antibiotiques sont ajoutées stérilement dans les milieux. **La tétracycline** est rajoutée dans les milieux solides et liquides de manière à avoir une concentration finale de 12,5 $\mu\text{g/mL}$ à partir d'une solution stock à 50 mg/mL dans de l'eau préparée à partir de poudre de tétracycline-HCl (Sigma-Aldrich®). De **la kanamycine** est rajoutée de manière à avoir une concentration finale de 50 $\mu\text{g/mL}$ à partir d'une solution stock à 50 mg/mL préparée dans de l'eau milliQ à partir de poudre de kanamycine-sulfate (Sigma-Aldrich®). Après préparation, les solutions stock sont filtrées (filtre de 0,2 μm) et conservées à -20°C en aliquots de 300 μL jusqu'à utilisation.

6.2.4. *Suivi de cultures bactériennes*

Toutes les cultures liquides de *M. extorquens* et de *Hyphomicrobium* sp. MC1 sont incubées à 30°C en aérobie, sous agitation à 120 rpm (incubateur Infors HT Multitron Standard).

Toutes les cultures liquides de *E. coli* sont incubées à 37°C en aérobie, sous agitation à 200 rpm (incubateur Infors HT Multitron Standard).

6.2.5. *Estimation de la biomasse par mesure de la densité optique à 600nm*

La densité optique à 600 nm (DO_{600}) est déterminée dans un volume de culture de 1 mL (diluée lorsque nécessaire pour obtenir une $DO_{600} < 0,9$) dans une cuve de spectrophotométrie en plastique. La mesure est réalisée avec un spectrophotomètre Jenway 7315.

Dosage des ions chlorure produits lors de l'utilisation du CH_3Cl : La dégradation du CH_3Cl peut être quantifiée en réalisant le dosage des ions Cl^- libérés dans le surnageant de cultures en milieu M3ClO au cours de la croissance. Un dosage colorimétrique est effectué selon la méthode de Jörg et Bertau (2004) en microplaque 96 puits Nunclon™ Thermo Scientific™. Les ions Cl^- réagissent dans un milieu hautement acide (ici de l'acide perchlorique $HClO_4$) avec du nitrate de fer ($Fe(NO_3)_3$), aboutissant à la formation de chlorure de fer $[FeCl]^{2+}$ dont la concentration peut être déterminée par mesure de la DO_{340} . Pour l'ensemble d'une microplaque, un mélange réactionnel est préparé en mélangeant 6,6 mL d'acide perchlorique 60%, 1 mL de $Fe(NO_3)_3$ à 0,2 M et 0,4 mL d'eau ultra pure. Chaque puits est rempli avec 80 μ L de milieu réactionnel auquel est rajouté 10 μ L de l'échantillon à tester (un surnageant de culture). Après une incubation de 15 min à température ambiante, la DO à 340nm est lue dans un lecteur de microplaques (Synergy HT-1 Biotek d'Agilent). Pour chaque plaque, une gamme étalon de solutions de NaCl de concentration s'échelonnant de 0 à 20 mM est faite avec un pas de 1 mM. La valeur de DO_{340} obtenue avec la solution à 0 mM est soustraite à toutes les autres valeurs obtenues. Le dosage mesuré pour trois réplicats par échantillon permet de calculer une moyenne.

6.2.6. Calcul du temps de génération

Le temps de génération traduit la vitesse à laquelle une souche peut se diviser en phase exponentielle de croissance. Le temps de génération est déterminé en réalisant un suivi de croissance par mesure régulière de la DO_{600} au cours de la phase exponentielle de croissance. L'évolution de la DO_{600} (échelle log10 axe des y) est représentée en graphique en fonction du temps en heures (axe des x). L'équation de la courbe de tendance est déterminée : $y=ae^{bx}$

Le temps de génération (g) en heures correspond à $\ln(2)/b$

Le temps de génération moyen est déterminé pour au moins trois réplicats biologiques.

Le taux de croissance des souches est en fonction du génotype des souches et des conditions de croissance. Ci-après est décrit le protocole permettant le suivi de croissance en fonction de la source de carbone de souches portant le plasmide pJM105.

Avec le CH_3Cl : Une colonie isolée issue d'un milieu M3Cl0 solide avec tétracycline et soit succinate soit CH_3Cl en présence de BBT est suspendue dans 200 μ L de milieu M3Cl0. De cette suspension, 170 μ L servent à inoculer une préculture de 50 mL de milieu M3Cl0 avec de la tétracycline et 15 mM de CH_3Cl . Après une semaine d'incubation à 30°C sous agitation (120 rpm), la DO_{600} est mesurée et la préculture est diluée à 0,02 le soir. Le lendemain matin, le suivi de la croissance commence et dure pendant une semaine à raison de 2 à 3 mesures de DO_{600} par jour. En parallèle, 10 μ L de la suspension servent à inoculer 3 mL de milieu M3Cl0 avec 4 mM de succinate et de la tétracycline, afin de vérifier la viabilité de la colonie en milieu liquide.

Avec le succinate ou le méthanol : Une colonie isolée issue d'un milieu M3Cl0 solide avec tétracycline et succinate est suspendue le matin dans une préculture de 25 mL de milieu M3Cl0 avec 10 mM de succinate et de la tétracycline. Les précultures sont incubées pendant 24 heures puis diluées pour avoir une DO_{600} de 0,04 dans le même milieu. L'évolution de la DO_{600} est suivie sur la journée à raison de 5 mesures (chacune espacée de 2 h) ainsi que d'une dernière mesure le lendemain matin. Pour la détermination du temps de génération avec le méthanol, le même protocole est appliqué que pour le succinate, à la différence que le méthanol à 40 mM remplace le succinate à 10 mM et que le temps d'incubation de la préculture est de 48 heures au lieu de 24 heures.

6.2.7. Test de croissance avec le formiate

Pour vérifier si des souches ayant perdu la capacité de croître avec le méthanol ont également perdu la capacité de croître avec le formiate, des colonies des souches à tester issues d'un milieu solide M3Cl0 avec succinate et tétracycline sont suspendues chacune dans 200 µL de M3Cl0. De ces suspensions, des inoculations en parallèle sont effectuées avec 50 µL dans 3 mL de M3Cl0 contenant de la tétracycline et soit 4 mM de succinate, soit 16 mM de méthanol, ou soit 16 mM de formiate. La DO₆₀₀ des cultures est mesurée au bout de 7 jours.

6.2.8. Vérification de l'essentialité de gènes compris dans le cluster *cmu*

Pour vérifier l'essentialité d'un gène de la voie *cmu* à la croissance avec le CH₃Cl dans une souche donnée, un variant du plasmide pJM105 comprenant une délétion de ce gène (Michener et al. 2016) est transféré par conjugaison tri-parentale (voir paragraphe 6.3.11) aux souches à tester. Les transconjugants obtenus sont cultivés en duplicats en parallèle en tube hungate dans 3 mL de milieu M3Cl0 avec de la tétracycline et du CH₃Cl et dans 3 mL de milieu M3Cl0 avec de la tétracycline et 4 mM de succinate. Après 12 jours d'incubation, la présence de croissance est vérifiée.

6.3. Méthodes de biologie moléculaire

6.3.1. Extraction d'ADN génomique

Pour le séquençage du génome des souches évoluées, une colonie de chacune d'entre elles (issue d'une boîte de milieu solide M3Cl0 avec succinate et tétracycline) est mise en culture dans 25 mL de M3Cl0 avec de la tétracycline et du CH₃Cl. Après une période d'incubation de 5 jours, les cultures sont centrifugées 10 min à 8000 rpm. Pour les souches de départ, les souches sont mises en culture dans 25 mL de milieu M3Cl0 avec 4 mM de succinate et de la tétracycline et les cultures sont également centrifugées. L'extraction est réalisée à partir du culot avec le kit Masterpure™ de Lucigen®, selon les instructions du fabricant. Après une dissolution de l'ADN une nuit à 4°C dans 35 µL de TE, la concentration en ADN est mesurée par Qubit™ et 5 µL de ces solutions sont migrés dans un gel d'agarose pour vérifier leur qualité. Pour d'autres usages, l'ADN génomique des souches est extrait à partir d'une culture

avec du succinate (et avec ou sans tétracycline) avec le kit Nucleospin™ Microbial DNA Macherey-Nagel™ d'après les instructions du fabricant.

6.3.2. Extraction d'ADN plasmidique

Une colonie de *E.coli* portant le plasmide à extraire est inoculée dans 20 mL de milieu LB liquide (avec l'antibiotique adéquat). Après 1 jour d'incubation, lorsque la culture est en phase stationnaire de croissance, la culture est centrifugée (8000 rpm, 10 min) et le culot peut être conservé à -20°C. L'ADN plasmidique est extrait avec le kit Nucleospin™ Plasmid Macherey-Nagel™ (selon les instructions du fabricant et le nombre de plasmides dans la cellule).

6.3.3. Dosage d'ADN

La quantité d'ADN est dosée avec le kit Qubit™ dsDNA Quantification Assay HS (High Sensitivity) ou BR (Broad Range) Invitrogen™. Ce dosage repose sur la présence de colorants fluorescents dans le mélange réactionnel se fixant de façon spécifique à l'ADN et émettant un signal fluorescent uniquement après cette fixation. Un volume de 199 µL de mélange réactionnel (préparé selon les instructions du fabricant) est mélangé à 1 µL d'échantillon d'ADN. Après 2 min d'incubation à température ambiante, la concentration d'ADN est lue avec le lecteur Qubit™ Fluorometer 2.0 Invitrogen™ après calibration avec les deux solutions standard fournies dans le kit (d'après les instructions du fabricant).

6.3.4. Electrophorèse d'ADN

L'ADN (au moins 10 ng) dans un volume de 10 µL est mélangé à 2 µL de tampon de charge TriTrack DNA Loading Dye 6X Thermo Scientific™, puis chargé dans un gel d'agarose 1% contenant du Bromure d'éthidium. Après migration dans du tampon Tris Acétate EDTA (TAE) entre 40 à 120V pendant au moins 30 minutes à 4 heures, les fragments d'ADN sont visualisés sous rayons UV avec un transilluminateur UVT-28M Herolab et leur taille est déterminée par rapport à un marqueur de taille Thermo Scientific™ GeneRuler.

6.3.5. Amplification d'ADN par Polymerase chain reaction (PCR)

Pour réaliser des PCR de routine (par exemple pour l'amplification de gènes visant à l'identification d'un contexte génétique), une ADN polymérase de type DreamTaq (DreamTaq DNA Polymérase (5 U/μL) Thermo Scientific™) est utilisée. Pour des expériences comme des clonages dans le but de construire des mutants nécessitant une haute-fidélité, une polymérase de type Phusion (Phusion™ High-Fidelity DNA Polymerase (2U/μL) Thermo Scientific™) est utilisée.

La PCR est faite pour un mélange réactionnel final de 20 μL en ajoutant :

Réactifs ^a	Volume (μL) pour une réaction ^a
H ₂ O ultrapure	18,4 (6,2)
DMSO 100%	0,75 (0)
Tampon DreamTaq x10 (Tampon Phusion HF ou GC 5X)	2,5 (2)
dNTPs 10 mM (dNTP 20 mM)	0,25 (0,2)
Amorce forward 20 μM	1 (0,25)
Amorce reverse 20 μM	1 (0,25)
DreamTaq 5U/μL (Phusion 2U/μL)	0,1 (0,1)
ADN (extrait lyse cellulaire)	1 (1)
Total	20 (10)

^a Réactifs et volumes pour la polymérase DreamTaq sont indiqués en premier et les réactifs et volumes pour la polymérase Phusion sont indiqués entre parenthèses

Les réactions ainsi préparées suivent un programme pendant 30 cycles sur un thermocycleur Biometra Tone Analytik Jena AG.

Programme PCR pour une amplification de 30 cycles (un cycle commence à partir de l'étape de dénaturation jusqu'à l'étape d'élongation) :

Pré-dénaturation	5 min 95°C (3 min 30 98°C) ^a
dénaturation	30 sec 95°C (30 sec 98°C)
hybridation	30 sec 57°C (30 sec 57°C)
élongation	1 min 30 72°C (2 min 72°C)
Post-élongation	10 min 72°C (10 min 72°C)

^a Les temps et températures de chaque étape sont indiqués pour la DreamTaq en premier et entre parenthèses pour la polymérase Phusion

Pour vérifier si la PCR a mené à une amplification, 5 à 10 µL du volume de la réaction sont déposés en gel d'agarose et une électrophorèse permet de séparer les fragments selon leur taille.

6.3.6. Hybridation ADN/ADN par la méthode dite du Southern Blot

Le Southern Blot (Southern 1975) permet la détection d'une séquence d'ADN homologue à une sonde d'ADN avec la détermination d'une taille précise obtenue par digestion des échantillons d'ADN par des enzymes de restriction et séparés par migration sur gel par électrophorèse. Après électrophorèse, l'ADN est transféré sur une membrane, puis hybridé à une sonde complémentaire à la séquence recherchée et la sonde peut être révélée par la suite (Un anticorps anti-DIG va se fixer à la dioxigénine qui est conjuguée à certaines bases de la sonde. Cet anticorps est conjugué à une phosphatase alcaline et sa présence sur la membrane peut être révélée grâce à un substrat chemoluminescent de l'enzyme qui va émettre de la lumière suite à sa transformation.

Synthèse de la sonde : La sonde 1 est synthétisée et marquée au cours d'une amplification PCR en présence de digoxigénine-11-dUTP (Roche).

La PCR est faite pour un mélange réactionnel final de 50 µL en ajoutant :

Réactifs	Volumes (en μL pour une réaction)
H ₂ O ultrapure qsp 50 μL	38,8
Tampon PCR 10X DreamTaq	5
dNTPs 10 mM	0,5
DIG-11-dUTP	1
Amorce 'forward' 20 μM	2
Amorce 'reverse' 20 μM	2
DreamTaq DNA polymérase 5U/ μL	0,2
ADN 20 ng/ μL	0,5
Total	50

Avec pour programme pour 30 cycles (un cycle commence à partir de l'étape de dénaturation jusqu'à l'étape d'élongation) :

Pré-dénaturation	5 min 95°C
Dénaturation	30 sec 95°C
Hybridation	30 sec 56°C
Elongation	32 sec 72°C
Post-élongation	10 min 72°C

Le produit PCR est purifié pour le débarrasser d'éléments comme les nucléotides non polymérisés avec une colonne Illustra™ Microspin S-400HR GE Healthcare, d'après les instructions fournies par le fabricant. Le volume obtenu (à une concentration de 50 ng/ μL) est complété avec du tampon d'hybridation de manière à avoir un volume final de 50 μL . Les sondes 1 et 2 ont été préparées en les amplifiant avec les couples d'amorces LFM008/ LFM009 et pCM-ColEI-F/ MC1_fold-R, respectivement.

Extraction d'ADN et digestion par des enzymes de restriction : L'ADN génomique des souches est extrait à partir de 55 mL d'une culture préparée à partir de milieu M3C10 additionné de 8 mM de succinate et avec ou sans tétracycline selon les cas. Après 4 jours d'incubation, les cultures sont divisées en 2 et centrifugées 10 min à 8000 rpm pendant 10 min. Un des deux culots sert à faire l'extraction. Les ADN génomiques sont extraits avec le kit NucleoSpin™

Microbial DNA Macherey-Nagel™ d'après les instructions du fabricant (l'extraction nécessite un culot faisant jusqu'à 40 mg de poids mouillé). Les ADN sont digérés par des enzymes de restriction en utilisant les kits FastDigest de Thermo Scientific™.

La digestion est faite pour un mélange réactionnel final de 25 µL en ajoutant :

Réactifs	Volumes (en µL) pour une réaction
H ₂ O	13
Tampon Fast Digest 10X Green Buffer	2,5
Enzyme Thermo Scientific™ Fast Digest	1
ADN (60 ng/µL dans de l'eau ultra pure)	8,5 (correspondant à au moins 500 ng)
Total	25

Après incubation au bain-marie 15 min à 37°C, les ADN digérés sont migrés sur gel d'agarose 0,6% à 60 volts pendant environ 2h30, en parallèle du marqueur de taille Thermo Scientific™ GeneRuler 1kb plus.

Traitement des ADN et transfert vers une membrane : Le gel d'agarose est lavé dans un bain d'eau déminéralisée pendant 5 min, puis 15 min dans 500 mL de solution de dépurination (125 mL de HCl 1 M (0,25 M final) ; 375 mL de H₂O ultrapure), puis deux bains successifs dans 500 mL de solution de dénaturation (300 mL de NaCl 5M (1,5M final) ; 100 mL NaOH 5 M (0,5M final) ; H₂O ultrapure qsp 1L) et enfin encore deux bains de 30 min dans 500 mL de solution de neutralisation (300 mL NaCl 5 M (1,5 M final) ; 500 mL Tris 1M (0,5 M final) pH 7,6 ; 2 mL EDTA 0,5 M (1 mM final) ; H₂O qsp 1L). Un montage est réalisé pour permettre le transfert par capillarité de l'ADN vers une membrane de nitrocellulose avec de haut en bas : plusieurs feuilles de papier Whatman imbibées dans du SSC 20X (350,4 g de NaCl (3 M final) ; 176,5 g tri-sodium citrate (0,3 M final) ajustement à pH 7 avec HCl 1M ; qsp 2 L), le gel, la membrane, d'autres feuilles de Whatman imbibées dans du SSC 20X, 15 cm de papier absorbant et enfin un poids. Après 24 h, la membrane est retirée, protégée par du film alimentaire et exposée face ADN à des rayons UV pour fixer les ADN à la membrane.

Hybridation ADN/ADN : La membrane est roulée dans un tube d'hybridation en présence de 20 mL de SSC 2X. Le rouleau est ensuite placé quelques minutes dans le système rotatif d'un four à hybridation Hybridiser HB-1D Techne®. Le SSC 2X est remplacé par 20 mL de solution de préhybridation (12,5 mL de phosphate de sodium 1M pH 7,2 (0,25 M final) ; 100 µL EDTA 0,5 M (1 mM final) ; 10 g SDS (20% final) ; 0,25 g de réactif de blocage (Roche référence 11096176001) ; H₂O ultrapure qsp 50 mL) et placé en rotation dans le four pendant 2h à 65°C. L'ADN de la sonde est dénaturé en l'incubant 5 min à 100°C, immédiatement placé dans la glace avant d'être mélangé à 10 mL de solution de préhybridation préchauffée à 65°C. Cette solution d'hybridation remplace le mélange de préhybridation pour une incubation dans le four à hybridation à 65°C. Après une nuit, la membrane est retirée et lavée avec du tampon de lavage (20 mL phosphate de sodium 1M pH 7,2 (20 mM final) ; 2 mL EDTA (1mM final), 10 g SDS ; H₂O ultrapure qsp 1L) préchauffé à 65°C dans un bac en verre avec faible agitation. La membrane est encore lavée par trois bains successifs de tampon de lavage de 20 min chacun à température ambiante, puis un bain de 5 min à température ambiante dans du tampon de lavage 2 (5,8 g d'acide maléique (0,1 M final) puis ajustement à pH 8 ; 87,66 g de NaCl puis ajustement à pH 8 ; 1,5 mL de Tween 20 (0,3% final) ; H₂O qsp 500 mL). L'ADN de la sonde marquée qui a hybridé l'ADN fixé à la membrane peut être maintenant révélé.

Révélation de l'ADN marqué : La membrane est incubée 60 min à température ambiante dans 20 mL de tampon de blocage (tampon de blocage 0,25 g de réactif de blocage Roche référence 1109617600 ; tampon de lavage 2 qsp 50 mL) dans un sachet en plastique fermé, remplacé par 10 mL de tampon dans lequel 0,7 µL d'anti-DIG-AP-conjugate (Roche référence 11093274910) ont été ajoutés. Après une incubation à température ambiante de 30 min, la membrane est retirée du sachet, lavée dans un bac de verre par 4 lavages successifs, chacun 10 min à température ambiante à agitation douce, avec 75 mL de tampon de lavage 2. La membrane remise dans un nouveau sachet en plastique, est équilibrée pendant 2 à 5 min avec 20 mL de tampon de détection (0,61 g de Tris-amino (0,1M final) ; 0,29 g de NaCl (0,1 M final) puis ajustement à pH 9,5 ; H₂O qsp 50 mL) avec une agitation douce. Le volume est retiré et 5 mL de tampon de détection ainsi que 1 mL de « CDP-Star working solution » (10 µL CDP-Star (Roche référence 11685627001) ; 1 mL de tampon de détection) sont ajoutés pour 5 min à température ambiante et agitation douce. La membrane est récupérée, enveloppée dans une pochette plastique et la révélation est effectuée avec une machine ChemiDoc MP Imaging

System Bio-Rad. A ce stade, la membrane ne doit pas être séchée si une nouvelle hybridation avec une deuxième sonde est prévue. Si c'est le cas, la première sonde doit être déshybridée.

Réutilisation de la membrane pour une deuxième hybridation avec une sonde : La membrane conservée dans du SSC 1X est lavée pendant 1 min dans de l'eau ultrapure, puis lavée au cours de deux bains successifs de 15 min chacun à 37°C avec une solution alcaline de déshybridation (0,2M NaOH contenant 0,1% de SDS). La membrane est à nouveau lavée avec du SSC 2X pendant 5 min avec faible agitation et prête à être hybridée avec une nouvelle sonde en suivant le même protocole que pour la sonde 1.

6.3.7. Mutagenèse dirigée par échange allélique

Un échange allélique par recombinaison ciblée d'un gène sauvage par un gène muté par exemple, est possible seulement chez certaines souches de *M. extorquens* (ex. AM1, PA1, DM4 mais pas chez CM4). La première étape est un clonage de l'allèle muté flanqué d'environ 0,5 kb de part et d'autre de la région génomique de la souche à muter dans la cassette de clonage de pCM433, un plasmide suicide chez *M. extorquens*. En effet, pCM433 est dépourvu d'une origine de réplication fonctionnelle chez *M. extorquens* de sorte à forcer son intégration par recombinaison homologue au niveau de l'une des deux régions flanquantes du gène à muter. Un deuxième évènement de recombinaison pour la perte du plasmide permettra le criblage d'excisants portant l'allèle du gène échangé, criblage grandement facilité par une sélection positive associée au gène *sacB* de pCM433 (Marx et al. 2008). Le protocole est décrit ci-dessous pour le remplacement allélique du gène *ftfL*.

Construction des plasmides pME8348 et pME8349 portant $ftfL_{wt}$ et $ftfL_{D7E}$, respectivement. Les deux plasmides sont construits de la même manière en amplifiant par PCR avec une enzyme de haute fidélité (Thermo Scientific™ Phusion™) un fragment de *ftfL* (position du génome chez la souche DM4, 444070-445204) avec les amorces LFM015 (présence d'un site de coupure Apal en 5') et LFM030 (présence d'un site de coupure AgeI en 5'). Pour pME8349, l'ADN génomique de la souche D7E sert de matrice, alors que pour pME8348 c'est celui de D0. Les amplicons sont purifiés avec le kit NucleoSpin™ Gel et PCR Clean-up Macherey Nagel™ d'après les instructions du fabricant. Une digestion enzymatique simultanée avec les deux

enzymes de restriction (Agel et Apal) est faite avec le kit FastDigest Thermo Scientific™ avec 600 ng de pCM433 et 260 ng d'insert (correspondant à 3 fois la quantité de plasmide en moles, ratio qui favorise le clonage de l'insert).

La digestion est faite pour un volume réactionnel final de 35 µL en ajoutant :

Réactifs	Mélange Vecteur et Insert	Mélange Vecteur uniquement
H ₂ O	22,2	30
Tampon FastDigest Green Buffer	4	4
pCM433 à concentration 300 ng/µL	2	2
Insert	7,8	0
Total	36	36

Comme contrôle de l'efficacité du clonage, un milieu réactionnel sans insert est réalisé. Après constitution de ces deux mélanges, 5 µL sont retirés de chaque (qui serviront de contrôles non digérés) et 2 µL de solution de Apal et 2 µL de solution de Agel sont ajoutés. Les mélanges ainsi constitués sont incubés 18 min dans un bain-marie à 37°C. Un volume de 5 µL des réactions est déposé sur gel d'agarose 0,8% et migré par électrophorèse afin de vérifier que la digestion s'est bien effectué puis les produits de cette digestion sont purifiés par une colonne Illustra™ Microspin S-400HR GE Healthcare, selon les instructions du fabricant. Un tiers du volume obtenu est utilisé pour la ligation. Cette dernière est effectuée avec le kit T4 DNA Ligase (5U/µL) Thermo Scientific™.

La ligation est faite pour un mélange réactionnel de 40 µL en ajoutant :

Réactifs	Volumes (en µL) pour une réaction
H ₂ O	qsp 39,4 µL
10 X Tampon T4 ADN ligase	4
ADN	1/3 volume (environ 200 ng d'ADN plasmidique et 87 ng d'insert)
Total	39,4

De ce mélange, 10 μL sont extraits (pour servir de témoin non ligué). Aux 29,4 μL restants sont ajoutés 0,6 μL de T4 ADN ligase 5 U/ μL (correspondant à 3 U). La réaction de ligation s'effectue sur la nuit à 4°C. Pour inactiver l'enzyme, le mélange réactionnel est incubé 10 min à 65°C, et 5 μL des produits sont déposés sur gel d'agarose 0,8% et migrés par électrophorèse pour vérifier que la ligation a bien fonctionné. Les solutions de plasmides ligués sont utilisées pour réaliser une transformation par choc thermique chez la souche JM109 d'*E. coli* de Promega (selon les instructions du fabricant). De cela, 75 colonies transformées pour chacun des deux plasmides sont repiquées sur milieu LB solide avec 12,5 $\mu\text{g/mL}$ de tétracycline. Une cinquantaine de souches sont testées par PCR pour vérifier la présence de l'insert.

Une PCR est réalisée pour un mélange réactionnel de 25 μL en ajoutant :

Réactifs	Volumes (en μL) pour une réaction
H ₂ O ultrapure	19,4
DMSO 100%	0,75
Tampon PCR x10	2,5
dNTPs 10 mM	0,25
V433 F 20 μM	1
V433R 20 μM	1
DreamTaq DNA Polymérase (5 U/ μL)	0,1
Pour chaque réaction, une colonie à tester a été directement suspendue dans le mélange.	

Le temps d'élongation lors de la PCR est de 2 min. Quatre transformants TcR présentant le bon profil (produit PCR de 1,3 kb alors que le plasmide pCM433 vide donne un fragment amplifié de 0,23 kb) sont sélectionnés et leur plasmide extrait. Le plasmide est envoyé pour séquençage Sanger (voir paragraphe 6.3.8) avec les amorces V433F et V433R pour vérifier la séquence des inserts.

Les plasmides pME8348 et pME8349 sont introduits par électroporation chez *M. extorquens* et les colonies TcR sont phénotypées pour vérifier qu'il s'agit bien de transformants TcR simple intégrants et sensibles à la présence de saccharose, comme décrit précédemment (Marx et al.

2008). Les colonies TcR sont suspendues dans 100 µL de M3ClO d'une plaque 96 puits puis inoculées à la surface de milieux M3ClO solide (à l'aide d'une plaque de repiquage stérilisable adaptée) sur trois milieux : M3ClO avec succinate, M3ClO avec succinate et tétracycline et M3ClO avec succinate et 5% de saccharose. Les colonies TcR qui se développent sur les deux premiers milieux solides mais pas celui contenant du saccharose correspondent au phénotype attendu de simple intégrant. Ces dernières sont étalées sur milieu solide M3ClO contenant du succinate 10 mM et différentes concentrations de saccharose (2,5%, 5 ou 10%) pour favoriser la croissance de colonies ayant excisé le plasmide. Les colonies résistantes à la présence du saccharose (peu importe la concentration) sont testées pour leur perte de résistance à la tétracycline par étalement sur milieu solide M3ClO avec succinate, M3ClO solide avec succinate et tétracycline, M3ClO solide avec succinate et 5% de saccharose et M3ClO solide avec 40 mM de méthanol. Les colonies portant l'allèle *ftfL*_{D7E} de phénotype Sac^R Tc^S et incapables d'assimiler le méthanol sont retenues et sont mises à conservation à -80°C. Les colonies portant l'allèle *ftfL*_{wt} de phénotype Sac^R Tc^S et capables d'assimiler le méthanol sont mises à conservation à -80°C. L'ADN génomique de ces souches est extrait avec le kit NucleospinTM Microbial DNA Macherey-NagelTM d'après les instructions du fabricant. Afin de vérifier la présence de la mutation D7E (ou le retour à la version sauvage du gène dans le cas du mutant D7EL_*ftfL*_{D0}), le gène *ftfL* est amplifié avec les amorces LFM014 et LFM015 et le produit envoyé pour séquençage Sanger avec les mêmes amorces.

6.3.8. Séquençage Sanger de produits PCR

Le séquençage Sanger d'ADN est réalisé par la société Microsynth AG en suivant leurs instructions. Un séquençage sur une longueur de 1Kb est obtenu. L'utilisation de deux amorces s'hybridant à l'extrémité du fragment permet un séquençage des deux brins. L'analyse des séquences est faite avec le logiciel SnapGene Viewer (SnapGene® software de Dotmatics ; disponible sur snapgene.com) ou BioEdit (Hall 1999).

6.3.9. Séquençage du génome à partir de l'ADN extrait par Illumina

L'ADN génomique (600 à 2300 ng par échantillon) a été envoyé à la société GenoScreen qui a construit les librairies avec le kit "Illumina Nextera® XT DNA Library Prep" d'après les

instructions du fabricant. La distribution de taille des fragments d'ADN des librairies a été contrôlée par électrophorèse capillaire sur Bioanalyzer kit DNA High Sensitivity (Agilent) et l'ADN a été dosé (Dosage Thermo Scientific™ SYBR Green). Un run de séquençage de type Illumina® paired-end en 2x150pb permet d'obtenir en général plus de 80% de lecture avec un score >Q30 (associé à un taux d'erreur de 1/1000 soit 0.1%).

Les séquences aux extrémités des lectures simple-brin correspondant aux doubles indexages des librairies faits pour relier les séquences aux échantillons et permettant l'hybridation des adaptateurs utilisés par la technologie de séquençage Illumina, ont été enlevées avec l'outil Trimmomatic (Bolger et al. 2018) et assemblées avec l'outil Unicycler v0.4.6 (Wick et al. 2017).

6.3.10. Transformation d'ADN plasmide chez *E. coli* par choc thermique

À 50 µL de cellules chimiocompétentes (traitement au chlorure de calcium ; Dagert & Ehrlich 1979) décongelées sont mélangés 1 µL de solution d'ADN plasmidique ou bien 1 µL d'eau ultrapure servant de témoin négatif. Le mélange est placé dans la glace pendant 2 min puis dans un bain-marie à 42°C pendant 2 min. Puis, 1 mL de milieu LB liquide est ajouté à chacun des mélanges qui sont ensuite incubés à 37°C à 100 rpm pendant une heure. Les mélanges sont étalés sur milieu solide LB avec tétracycline. Une incubation à 37°C pendant une nuit permet d'obtenir une colonie isolée.

6.3.11. Transfert des plasmides d'*E. coli* vers *M. extorquens* par conjugaison tri-parentale

Deux jours précédant l'expérience, une préculture de la souche réceptrice est préparée par inoculation d'une colonie (issue d'un milieu solide M3C10 avec méthanol) dans 25 mL de milieu M3C10 avec 40 mM de méthanol et la préculture est mise en incubation.

La veille, une préculture de la souche *E. coli* DH5α contenant le plasmide à transmettre (on parle de souche donneuse) est préparée en suspendant une colonie (issue d'une boîte de milieu solide LB avec tétracycline) dans 25 mL de milieu LB avec tétracycline. Pareillement, une préculture de la souche *E. coli* helper, contenant le plasmide pRK2073, est préparée en suspendant une colonie de la souche (issue d'une boîte de milieu solide LB avec kanamycine) dans 25 mL de milieu LB liquide avec kanamycine. Le plasmide pRK2073 apporte *en trans* le produit des gènes *tra* nécessaires au transfert des plasmides mobilisables du groupe IncP

(Figurski & Helinski 1979). Les précultures d'*E. coli* sont incubées la nuit à 37°C, diluées le lendemain matin dans du milieu frais LB avec l'antibiotique adéquat de manière à avoir 50 mL à DO₆₀₀ avoisinant 0,015. Parallèlement, la préculture de *M. extorquens* est diluée au ½ dans du milieu M3CI0 frais avec du méthanol à la même concentration que précédemment.

Les cultures sont incubées sur la journée jusqu'à ce que les cultures atteignent la phase exponentielle de croissance (*E. coli* DO₆₀₀ avoisinant 0,5 et *M. extorquens* DO₆₀₀ avoisinant 0,4). Les cultures de *E. coli* et *M. extorquens* sont alors centrifugées 10 min à 8000 rpm à 4°C et les culots sont lavés avec 5 mL de milieu M3CI0 sans source de carbone. Les suspensions sont recentrifugées et les culots obtenus suspendus dans un volume de M3CI0 sans source de carbone de manière à avoir une DO₆₀₀ de 0,5/mL pour les suspensions d'*E. coli* et une DO₆₀₀ de 1/mL pour celles de *M. extorquens*. Un croisement tri-parental consiste à mélanger 1 mL de suspension de souche réceptrice, 1 mL de souche donneuse et 1 mL de souche helper. Des témoins négatifs sont réalisés en mélangeant soit la souche réceptrice avec uniquement la souche donneuse, soit en mélangeant la souche réceptrice avec uniquement la souche helper. Chaque mélange est centrifugé 10 min à 8000 rpm et les culots suspendus dans le surnageant résiduel sont déposés en goutte sur milieu solide M3CI0 sans source de carbone. Après séchage des gouttes, les boîtes sont incubées pendant au moins une nuit. Les transconjugants sont récoltés dans 250 µL de milieu M3CI0 sans source de carbone pour être étalés sur milieu solide M3CI0 avec tétracycline et méthanol. Trois jours plus tard, deux colonies de transconjugants TcR de chaque croisement sont réisolées sur une nouvelle boîte du même milieu solide. Pour la culture des souches réceptrices, le méthanol peut être remplacé par du succinate (4 mM en milieu liquide et 10 mM en milieu solide).

6.3.12. Electrotransformation d'ADN plasmidique chez *M. extorquens*

Pour préparer des cellules électrocompétentes, une colonie issue d'un milieu solide M3CI0 avec succinate est inoculée dans 25 mL de milieu M3CI0 liquide avec 25 mM de succinate. Le lendemain, ces précultures sont diluées de manière à démarrer des cultures de 200 mL à une DO₆₀₀ initiale de 0,01 dans du milieu M3CI0 avec 30 mM de succinate. Les cultures sont suivies jusqu'à avoir une DO₆₀₀ à environ 1. Après cela, les cellules sont placées dans la glace pendant 30 min pour arrêter leurs divisions. Les cellules sont ensuite centrifugées à 4000 rpm pendant 15 min à 4°C dans des tubes prérefroidis. Après élimination du surnageant, deux lavages avec

un même volume d'eau ultrapure stérile glacée sont réalisés. Un troisième lavage est réalisé avec la moitié du volume de l'eau glacée stérile additionnée de 10% de glycérol. Les culots sont ensuite resuspendus dans un dixième du volume de glycérol 10% et 100-200 µL conservés à -80°C dans des cryotubes prérefroidis.

Dans un tube Eppendorf refroidi, 200 ng de solution d'ADN plasmidique sont mélangés à 50 µL de cellules compétentes, déposés dans une cuvette d'électroporation stérile prérefroidie (Gene Pulser®/MicroPulser™ 0,2 cm Bio-Rad prérefroidie). La cuvette est insérée dans une machine d'électroporation Eppendorf Electroporator 2510. L'électroporation se fait par un pulse unique d'un voltage de 2,5 kV pendant 5,4 - 5,6 ms. Tout de suite après le pulse, 1 mL de Nutrient Broth sans NaCl préchauffé à 30°C est rajouté dans la cuvette et mélangé délicatement avec la suspension cellulaire électroporée. Le contenu de la cuvette est ensuite transféré dans un tube Falcon™ de 15 mL et incubé pendant 2h. Les cellules sont étalées sur milieu solide M3Cl0 avec du succinate et de la tétracycline et les boîtes sont incubées jusqu'à apparition de colonies de transformants TcR.

6.4. Dosages enzymatiques

6.4.1. Test d'activité enzymatique de FtfL

L'activité de l'enzyme FtfL d'un extrait cellulaire est dosée par la formation au cours du temps de formyl-H₄F à partir de formiate et de H₄F. Le protocole du test d'activité a été décrit précédemment (Caroll et al. 2014), lui-même repris de Marx et al. 2003b et Rabinowitz et Pricer 1962.

Préparation de culots cellulaires : Des précultures des souches à tester sont préparées en suspendant une colonie issue d'un milieu solide M3Cl0 avec tétracycline et succinate dans 10 mL de milieu M3Cl0 liquide avec 4 mM de succinate et de la tétracycline dans un tube Falcon™ de 15 mL. Après un jour d'incubation, les précultures sont diluées à une DO₆₀₀ de 0,04 dans 30 mL en milieu M3Cl0 avec tétracycline et un mélange de 4 mM de succinate et de 16 mM de méthanol. Après incubation (sur la nuit ou plus), lorsque les cultures atteignent une DO₆₀₀ entre 0,4 et 0,45, chacune des cultures (20 mL) est lavée par centrifugation pendant 10 min à 8000 rpm et le culot obtenu est resuspendu dans le même volume de Tris 100 mM pH8.

La suspension est alors divisée en deux suspensions de 10 mL qui sont également centrifugées 10 min à 8000 rpm et les culots obtenus sont transférés dans des tubes de 1,5 mL pour être conservés à -80°C avant utilisation.

Préparation d'extraits cellulaire par lyse des culots et dosage des protéines : Après décongélation sur la glace, un de chacun des deux culots est resuspendu dans 1,2 mL de Tris 1M pH8. Les suspensions ainsi préparées sont divisées en 2 pour mélanger 600 µL de suspension à 1,2 g de billes de verre d'un diamètre 0,3 mm. Ces mélanges sont lysés par agitation pendant 6 min et 30 secondes dans des blocs prérefroidis au bead-beater Retsch® 2MM réglé à la puissance maximale. Après avoir replacé les échantillons dans la glace pendant 2 min, les lysats sont centrifugés 5 min à 8000 rpm à 4°C. Les deux surnageants issus des lysats correspondant au même culot sont rassemblés.

Quantification des protéines par Qubit : La quantité de protéines est déterminée à partir d'un lysat cellulaire par mélange de 180 µL de mélange réactionnel à 20 µL d'échantillon avec le kit Qubit™ Protein Assay Invitrogen™, d'après les instructions du fabricant. Le lecteur Qubit™ Fluorometer 2.0 Invitrogen™ est calibré en utilisant des solutions étalon du kit. Pour chaque échantillon, la quantification est réalisée 3 fois afin de calculer une moyenne.

Dosage de l'activité FtfL : Le test consiste à quantifier la formation de formyl-H₄F à partir de formiate et de H₄F au cours du temps. Dans une rangée verticale d'une plaque 96 puits, 6 puits sont chacun remplis d'un mélange réactionnel constitué de 7 µL de Tris 1M pH8, 7 µL de MgCl₂ 100 mM (MW= 95,211 g/mol), 7 µL d'ATP 50 mM (ATP disodique (MW=551,14 g/mol) neutralisé avec une solution KOH à 1M), 7 µL de formiate de sodium 1M pH 8,0 et de 22 µL d'eau, en plus de 6 µL d'extrait cellulaire. La réaction est démarrée par l'ajout de 14 µL de H₄F 10 mM (MW=445,43 g/mol pureté à 65% ; neutralisation avec KOH 1M ; conservation sous atmosphère anoxique à -80°C après préparation d'aliquots dans une boîte à gants sous atmosphère azotée) pour un volume final de la réaction de 70 µL dans chacun des puits. La réaction est arrêtée par l'ajout de 140 µL de HCl 0,36 M, successivement à différents temps : 2 min pour le premier puits, 4 min pour le deuxième avec le même pas de temps de 2 min jusqu'au 6^{ème} puits à 12 min. Le traitement acide va transformer le formyl-H₄F en méthényle-H₄F qui peut être dosé à 350 nm. Pour chacun des extraits cellulaires, ce même test est réalisé

deux fois (deux réplicats techniques dont une moyenne est déterminée). La souche D0 (pCM62) sert de témoin positif d'une activité FtlL de type sauvage. Pour le témoin négatif, l'eau remplace l'extrait cellulaire. Après 10 minutes d'incubation à température ambiante, la DO à 350 nm est lue avec le lecteur de microplaques Synergy HT-1 (Biotek) d'Agilent. Le coefficient directeur de la pente du témoin négatif est retiré à celui de la pente de l'échantillon. Le coefficient d'extinction du méthényle-H₄F ($\epsilon = 24.9 \text{ mM}^{-1} \text{ cm}^{-1}$) et la longueur du trajet optique de 0,78 cm permet la détermination de l'activité de l'échantillon dont l'activité spécifique est calculée en nmoles.min⁻¹.mg de protéines⁻¹.

6.4.2. Test d'activité de la CH₃Cl deshalogénase

Le dosage de l'activité de déshalogénéation du CH₃Cl est basé sur la détermination des ions Cl⁻ produits au cours du temps d'après le protocole décrit précédemment (Nadalig et al. 2011).

Préparation des cellules en suspension : Des précultures des souches à tester sont réalisées en suspendant une colonie (issue d'un milieu solide succinate et tétracycline) dans 25 mL de milieu M3Cl0 avec tétracycline et 4 mM de succinate et 15 mM de CH₃Cl. Après 3 jours d'incubation, les précultures sont diluées de manière à avoir une DO₆₀₀ initiale de 0,025 dans 55 mL de milieu frais. Le lendemain matin, lorsque les cultures atteignent un DO₆₀₀ d'au moins 0,1 elles sont centrifugées 10 min à 8000 rpm à 30°C et reprises dans 20 mL de tampon phosphate 50 mM pH=7,2. Le lavage est répété et le culot cellulaire est suspendu dans le même tampon de manière à avoir une DO₆₀₀ de 1.

Test d'activité de la déshalogénéation du CH₃Cl et quantification des protéines : Dans un tube hungate avec un bouchon Thermo Scientific™ Mininert, 3 mL des suspensions sont mis en présence de 2 mL de CH₃Cl et incubés à 30°C avec une agitation de 120 rpm. Des prélèvements de 300 µL de suspensions cellulaires sont réalisés à intervalles réguliers (respectivement à 2h, 3h, 4h, 5h et 6h). Comme témoin négatif, 500 µL de chaque suspension cellulaire sont incubés pendant 6h dans les mêmes conditions dans des tubes hungate mais en absence de CH₃Cl injecté. Un échantillon de 300 µL de ces contrôles négatifs est également prélevé à 6h. Un autre témoin est obtenu avec 1 mL de chacune des suspensions non incubées (temps zéro considéré comme un 'blanc' sans activité). Après centrifugation 5 min à 15000 rpm, le

surnageant des échantillons est conservé à -20°C pour doser les ions Cl^- . À partir d'un graphe, la concentration en Cl^- des surnageants est calculée en fonction du temps (nmoles de $\text{Cl}^- \cdot \text{min}^{-1}$) après déduction de la valeur du « blanc » de chaque échantillon. Les témoins sans CH_3Cl permettent de s'assurer qu'il n'y a pas d'augmentation de la concentration en Cl^- en l'absence de CH_3Cl .

Pour déterminer la quantité de protéines, le culot obtenu après centrifugation de 1 mL de chacune des suspensions est mélangé à 0,5 mL de tampon de lyse (0,1 g/mL saccharose, 0,1 M NaCl, 0,1% SDS) et incubé pendant 1h à 100°C. La concentration en protéines est dosée à partir du lysat (kit Qubit™ Protein Assay Invitrogen™ d'après les instructions du fabricant et lecteur Qubit™ Fluorometer 2.0 Invitrogen™). Pour chaque échantillon, le test est réalisé 3 fois afin de calculer une moyenne et permet de ramener l'activité en $\text{nmoles} \cdot \text{min}^{-1} \cdot \text{mg}$ de protéines⁻¹. Alternativement, la concentration en protéines est déterminée avec le kit BCA Protein Assay Sigma-Aldrich® d'après les instructions du fabricant, bien que la première méthode soit préférée.

6.5. Évolution expérimentale de bactéries vers la capacité à croître avec le CH_3Cl

6.5.1. Transferts en série de cultures sur un grand nombre de générations

Mise en culture liquide des souches parentales : Pour la construction des souches parentales, les plasmides pJM105 et pCM62 sont transmis par conjugaison tri-parentale aux deux souches A0 et D0 cultivées avec du méthanol (pour chaque plasmide et chaque souche, deux croisements). Pour chacun des deux contextes génétiques, quatre colonies de transconjugants portant pJM105 sur milieu solide M3ClO supplémenté de méthanol et de tétracycline (issues de deux croisements indépendants) et correspondant à quatre réplicats biologiques servent à inoculer 5 mL de milieu M3ClO avec 10 mM de succinate et de la tétracycline. Une colonie de transconjugant portant pCM62 pour chacun des deux contextes sert également à inoculer une culture qui servira de contrôle négatif. Ces précultures sont incubées pendant 4 jours jusqu'à atteindre une DO_{600} d'environ 0,5.

Première culture (n0) de l'évolution expérimentale : Chaque préculture sert à inoculer la première culture de l'évolution expérimentale (n0) par dilution au 1/100^{ième} dans des

erlenmeyers étanche au gaz (bouchon Mininert) de 250 mL contenant 20 mL de milieu M3ClO supplémenté de 0,2 mM de succinate et de tétracycline. Avant l'ajout de 7,2 mL de CH₃Cl correspondant à 15 mM, 1 mL de culture est prélevé pour mesurer la DO₆₀₀ avant croissance. La mesure de la DO₆₀₀ avant et après croissance permet le calcul du nombre de générations de chaque culture de l'évolution expérimentale. Après quatre jours d'incubation, la DO₆₀₀ est mesurée et la culture n0 sert à inoculer une nouvelle culture d'évolution expérimentale n+1. Pour cela, un aliquot de 1,6 mL de la culture n0 est transféré dans du milieu neuf de même composition que pour la culture n0 pour un volume final de 20 mL, ce qui correspond à une dilution 1/16^{ième}.

Transferts en série et passage de la phase 1 en phase 2 : Les transferts sont réalisés tous les 7 jours. Il y a deux phases de transferts qui se distinguent par la présence ou l'absence d'un supplément de succinate. Lors de la première phase, la dilution utilisée pour un transfert est de 1/16^{ième}. Le passage en phase 2 est décidé lorsque la culture dépasse une DO₆₀₀ d'au moins 0,1. En phase 2, la dilution finale est de 1/64^{ième} et le CH₃Cl est l'unique source de carbone. Au démarrage de la phase 2, un transfert est également réalisé en parallèle avec 0,2 mM de succinate en plus du CH₃Cl comme culture de réserve pour relancer le transfert si nécessaire. Les transferts sont menés jusqu'à atteindre environ 200 générations.

Échantillonnage au cours des transferts : Avant le transfert dans du milieu neuf, à chaque culture, des échantillons sont conservés pour des études ultérieures.

- ➔ **Conservation à -80°C d'un aliquot de culture.** Un aliquot de chaque culture est conservé à -80°C en mélangeant 1 mL de la culture à 200 µL d'une solution stérile de glycérol à 80%. Alternativement, avant ajout de glycérol, 1 mL de culture est centrifugé 15 min à 13 000 rpm pour concentrer 10X et la suspension est mélangée à 100 µL de glycérol 20% stérile dans un cryotube.
- ➔ **Dosage de chlorure.** 15 mL de culture sont centrifugés 10 min à 8000 rpm. Pour le dosage du chlorure, 2 mL de surnageant sont conservés à -20°C.
- ➔ **Conservation de culot de culture à -80°C.** Le culot des 15 mL de culture est transféré dans un tube Eppendorf de 1,5 mL et conservé à -80°C pour une éventuelle extraction d'ADN pour un suivi des mutations au cours des transferts.

➔ **Mesure de la DO₆₀₀ pour le calcul du nombre de générations à chaque transfert.** Afin de déterminer le nombre de générations atteint à la fin de l'évolution expérimentale, un calcul du nombre de générations accompli à la fin de chaque culture est réalisé ainsi :

$$g = \frac{(\log(\text{DO}_{600} \text{ finale}) - \log(\text{DO}_{600} \text{ initiale}))}{0.301}$$

avec DO₆₀₀ initiale et DO₆₀₀ finale correspondant à la mesure de DO₆₀₀ avant et après croissance, respectivement.

6.5.2. Isolement de souches évoluées

Afin d'obtenir des colonies évoluées génétiquement homogènes sur milieu gélosé, deux techniques d'isolement sont utilisées : Une méthode dite d'étalement direct et une autre méthode dite d'« extinction-dilution », avec une étape intermédiaire de culture en milieu liquide. Les deux approches commencent par des dilutions en cascade au dixième des dernières cultures d'évolution expérimentale en mélangeant 20 µL de la dilution précédente dans 180 µL de milieu M3ClO dans une microplaque stérile 96 puits à fond plat Nunclon™ Thermo Scientific™.

Isolement de clones évolués par étalement direct : Pour chaque réplicat de l'évolution expérimentale, un échantillon est dilué en cascade. Puis, 50 µL de chacune des dilutions 10⁻³ à 10⁻⁵ sont étalés avec des billes en verre sur milieu solide M3ClO avec tétracycline et BBT. Après une incubation de plus de 8 jours avec du CH₃Cl, 8 colonies acidifiantes sont sélectionnées par réplicat pour archivage à -80°C.

Isolement de clones évolués par Extinction Dilution : Après deux dilutions en cascade (a et b), un volume de 50 µL de chacune des dilutions 10⁻³ à 10⁻⁷ est utilisé pour l'inoculation de deux cultures en milieu liquide de 3 mL de milieu M3ClO avec tétracycline : l'une injectée avec 10 mM de CH₃Cl et l'autre contenant 4 mM de succinate. Cette dernière sert de contrôle de la présence de bactéries viables dans les dilutions utilisées. En guise de témoin positif, une culture de *M. extorquens* CM4 permet de tester la croissance en présence de CH₃Cl (sans tétracycline dans le milieu). Un autre contrôle est d'inoculer des dilutions 10⁰ et 10⁻¹ avec chacune des deux souches de départ 105D et 105A afin de vérifier que seules les souches évoluées peuvent croître avec le CH₃Cl. Enfin, un témoin négatif (non-inoculé) sert à vérifier l'absence de contamination du milieu utilisé. Après 7 jours d'incubation, la DO₆₀₀ des cultures

est mesurée. Pour chaque série, la culture en CH₃Cl issue de la dilution la plus forte et présentant toujours une croissance bactérienne est utilisée pour l'isolement de clones par étalement sur milieu solide. Comme pour l'étalement direct, la culture est diluée en cascade et les dilutions 10⁻³ à 10⁻⁶ sont étalées en double sur milieu M3Cl0 solide avec tétracycline et BBT et sont incubées avec du CH₃Cl. Après 8 jours d'incubation, 4 colonies jaunes isolées sont mises en cultures pour une conservation à -80°C.

Test de croissance des souches évoluées isolées et mise en conservation à -80°C : Chacune des colonies sélectionnées est suspendue dans 200 µL de milieu M3Cl0. Un aliquot (50 µL de suspension par culture) sert à inoculer deux cultures de 3 mL de milieu liquide M3Cl0 avec tétracycline contenant soit 10 mM de CH₃Cl, soit 4 mM de succinate. Après 7 jours d'incubation, les DO₆₀₀ des cultures en CH₃Cl sont mesurées et 2 mL de culture sont centrifugés 15 min à 13 000 rpm. Un volume de 1,8 mL de surnageant est conservé à -20°C et sert pour le dosage du chlorure. Les culots sont remis en suspension dans 200 µL de surnageant et mélangés à 200 µL d'une solution stérile de glycérol à 20% avant conservation à -80°C.

6.5.3. *Évaluation des mutations associées au plasmide pJM105 des souches évoluées*

Pour vérifier si le plasmide pJM105 abrite des mutations essentielles pour la croissance avec le CH₃Cl des souches évoluées, un remplacement du plasmide présent dans les souches évoluées pour la version originelle permet de tester si la croissance des souches avec le composé va en être compromise.

Obtention de souches curées des souches évoluées ayant perdu le plasmide pJM105 : Pour obtenir des souches curées ayant perdu pJM105, les souches évoluées sont cultivées dans des conditions sans pression de sélection pour le maintien du plasmide à savoir en absence de tétracycline et absence de CH₃Cl. Une colonie isolée de chaque souche évoluée (issue d'un étalement sur milieu solide M3ClO avec BBT, tétracycline et incubé avec du CH₃Cl) sert à inoculer une culture de 3 mL de M3ClO liquide avec 4 mM de succinate. Après culture pendant 2 jours, un étalement à l'aide d'une oese est fait sur milieu solide M3ClO avec du succinate mais sans tétracycline afin d'obtenir des colonies isolées dont le phénotype de perte de croissance du plasmide est vérifié. Si aucune colonie ne présente le phénotype attendu, un second transfert est effectué dans du milieu frais à partir de la première culture. Pour cela, 1 µL de la culture est transféré dans un autre tube contenant le même volume du même milieu.

Afin de tester si certaines des colonies obtenues présentent un phénotype de perte du plasmide pJM105, elles sont repiquées avec des cure-dents stériles dans les puits d'une plaque 96 puits à fond plat Nunclon™ Thermo Scientific™ contenant 100 µL de milieu M3ClO liquide. En utilisant un hérisson, des répliques sont effectuées successivement sur 3 types de milieu solide : M3ClO avec succinate, M3ClO avec succinate et tétracycline, M3ClO avec BBT et incubé avec du CH₃Cl. Après une semaine d'incubation, deux colonies ayant pu se développer sur M3ClO et succinate mais pas en présence de tétracycline ou avec du CH₃Cl comme source de carbone sont réétalées à l'oese sur milieu solide M3ClO avec succinate. Par la suite, le phénotype (perte de croissance avec le CH₃Cl ou en présence de tétracycline) est confirmé par croissance en milieu liquide.

Chaque colonie est suspendue dans 200 µL de milieu M3ClO. De cette suspension, 40 µL servent à inoculer 3 mL de milieu M3ClO contenant soit 4 mM de succinate (vérification de la viabilité), soit le même milieu mais avec de la tétracycline (vérification la résistance à la tétracycline), soit 10 mM de CH₃Cl (test de capacité de croissance avec le CH₃Cl), et enfin soit sans aucune source de carbone (vérification de la non-contamination par une autre source de carbone). Après une incubation de 2 jours, la DO₆₀₀ des cultures en succinate est mesurée et au bout de 7 jours pour les autres (pour bien s'assurer qu'il n'y ait pas de croissance lente). Les souches curées ne se développent que dans le milieu avec uniquement du succinate. Pour leur conservation, 2 mL de ces cultures sont centrifugés 10 min à 10 000 rpm, le culot

suspendu dans 200 µL de surnageant et mélangé à 200 µL de glycérol 20% stérile dans un cryotube et stocké à -80°C.

Réintroduction du plasmide pJM105 dans les souches curées et test des capacités de croissance des transconjugants : la conjugaison tri-parentale est utilisée pour transférer le plasmide pJM105 ou le plasmide témoin pCM62 (non porteur des gènes *cmu*) dans les souches curées obtenues à partir des souches évoluées. Deux croisements indépendants sont réalisés pour chaque souche curée par plasmide. Une colonie TcR issue de chaque croisement sert à inoculer 3 mL de milieu M3Cl0 avec 15 mM de CH₃Cl afin de vérifier si la réintroduction du plasmide pJM105 conduit au rétablissement de la croissance avec le CH₃Cl. En parallèle, 3 mL de milieu M3Cl0 avec 4 mM de succinate et de la tétracycline sont aussi inoculés comme contrôle de viabilité. Après 2 jours d'incubation, 2 mL des cultures avec du succinate servent à un archivage à -80°C. Les cultures avec du CH₃Cl sont incubées 12 jours (car les cellules prennent plus de temps à croître avec du CH₃Cl avant que la DO₆₀₀ soit mesurée).

6.6. Méthodes analytiques in silico

6.6.1. Analyse des génomes et recherche des mutations des souches évoluées avec la plateforme MaGe de Microscope

Les outils disponibles sur la plateforme MaGe <https://mage.genoscope.cns.fr> sont utilisés pour l'analyse de génomes selon les instructions fournies.

Les données de séquençage Illumina de chaque souche évoluée mais aussi les données de séquençage des souches parentales sont analysées à l'aide du pipeline bioinformatique dédié de la plateforme MaGe de MicroScope (<https://mage.genoscope.cns.fr>) (Vallenet et al., 2020). Dans un premier temps, les lectures sont alignées sur les séquences génomiques des réplicons des souches ancestrales associées (*Methylobacterium extorquens* AM1 et DM4) à l'aide du package SSAHA2 (v.2.5.1) (Ning, 2001). Seuls les alignements uniques ayant un score d'alignement égal à au moins la moitié de leur longueur sont retenues comme graines pour un réalignement complet de Smith-Waterman (Smith et Waterman, 1981) avec une région étendue des deux côtés par cinq nucléotides du génome de référence. Tous les alignements calculés sont ensuite examinés pour détecter les différences entre les séquences des lectures

et la référence et un score basé sur la couverture, la fréquence des allèles, la qualité des bases et le biais de brin est calculé pour chaque événement détecté afin d'évaluer sa pertinence.

Enfin, via l'interface de MaGe (dans l'outil Variant profiling), les mutations (SNP et petits indels) trouvées dans les souches évoluées mais non retrouvées dans l'analyse du séquençage de la souche de départ avec un score supérieur à 0,8 et supportées par au moins 5 lectures sont conservées. L'effet des petites insertions et délétions dans la séquence codante des gènes a été déterminé avec Biopython (Cock et al. 2009) à partir des séquences d'ADN mutées ou sauvages. Un rapport généré à partir d'un script Python ad hoc (Domitille Jarrige https://github.com/JarrigeD/AIME_experimental_evolution_cmu) indique pour chaque gène muté des caractéristiques remarquables comme le premier acide aminé différant de la protéine sauvage, l'apparition d'un changement de cadre de lecture et la taille de la protéine chimérique résultante.

Les mutations dans pJM105 sont prédites pour chaque souche avec l'outil snippy v4.0 (<https://github.com/tseemann/snippy>) à partir de fichiers Genbank contenant tous les réplicons de chaque souche ainsi que du plasmide pJM105 (GenBank numéro d'accèsion KT950741.1).

6.6.2. Mise en évidence des grandes délétions au sein du génome des souches évoluées

Le nombre de lectures brutes alignées sur le génome pour chaque gène est normalisé pour chaque souche en Transcrits Par kilobase par Million (TPM) avec la formule :

$$TPM_i = \frac{RPK_i}{\sum_j RPK_j} \times 10^6$$

$$\text{Où } RPK_i = \frac{\text{reads mappés au gène } i}{\text{longueur du gène } i \text{ en kb}} \text{ et } RPK_j = \text{Ensemble des } RPK_i$$

sachant que RPK est en anglais « reads assigned per kilobase of target ». Le ratio entre les TPM des souches évoluées et des souches parentales (105A ou 105D) est calculé pour estimer les pertes potentielles de couverture pour chaque gène.

6.6.3. Recherche des liens paired-end entre le plasmide pJM105 et les réplicons des souches évoluées

Si le plasmide pJM105 n'est plus sous forme libre dans le génome des souches évoluées, mais intégré à un réplicon endogène, il devrait y avoir des fragments d'ADN chevauchant les régions du plasmide pJM105 et le réplicon du génome où a eu lieu l'intégration. De plus, des lectures correspondant aux extrémités (lectures paired-end) de ces fragments devraient être détectables dans l'ensemble des lectures du séquençage. Dans un premier temps, la séquence du plasmide pJM105 a été recherchée par homologie de séquence avec l'outil ncbi-blast+ v2.2.28 (Camacho et al. 2009) dans l'assemblage de la souche évoluée. Les lectures PE de la souche évoluée ont été mappées avec BWA v0.7.4 (<https://github.com/lh3/bwa>) sur les contigs de l'assemblage. Les liens PE entre le contig portant pJM105 et les autres contigs de l'assemblage ont été recherchés dans l'alignement (fichier bam) avec samtools view v1.11 (paramètre -F 2).

L'organisation des contigs ainsi obtenue a été comparée au génome de référence *M. extorquens* AM1 par blastN sur la plateforme MaGe (<https://mage.genoscope.cns.fr>).

6.6.4. Détermination de la couverture relative des lectures de chaque réplicon d'un génome par rapport à celle du chromosome

Les données Illumina de chaque variant ont été mappées sur les séquences de référence AM1 et DM4 issues de MaGe <https://mage.genoscope.cns.fr> (en plus du plasmide pJM105) avec BWA v0.7.4. L'outil coverage de samtools v1.11 (<https://github.com/samtools/samtools/releases/>) a permis d'obtenir la couverture moyenne (meandepth) de chaque réplicon. La couverture relative (cov) a été obtenue en divisant la couverture moyenne par celle du chromosome qui sert de référence (ce dernier étant présent en un exemplaire dans la cellule).

6.6.5. Prédiction de Structure de FtfL

Une prédiction de l'homotétramère de FtfL de *M. extorquens* est réalisée en générant un fichier multi-fasta contenant 4 fois la séquence protéique de *M. extorquens* DM4. La prédiction de structure est réalisée en utilisant Alphafold v2.2.3 (Jumper et al. 2021; Evans et al. 2021) en mode « multimer ». Les calculs sont effectués sur le cluster Core de l'Institut Français de Bioinformatique avec l'allocation de ressources suivante : utilisation d'une Graphics Processing Unit (GPU) complète (7g.40 gb), de 10 Central Processing Units (CPU), et

de 50G de mémoire. Le temps de calcul maximal a été porté à 24h. Les options « use_gpu_relax = true », « db_preset = full_dbs » sont utilisées. La prédiction de structure avec la meilleure vraisemblance est retenue. La visualisation du modèle est effectuée avec le logiciel UCSF ChimeraX (Goddard et al. 2018 ; Pettersen et al. 2021).

6.6.6. *Tests statistiques sur données de suivis de croissance ou tests de déshalogénation*

Les tests statistiques sur les données de croissance ou d'activité de déshalogénation sont réalisés avec des scripts Rmarkdown (écrits par Domitille Jarrige avec les conseils d'Adrien Borreca) avec la bibliothèque rstatix. Pour tester si les groupes d'échantillons sont issus de populations avec des distributions différentes, le test non paramétrique de Kruskal-Wallis suivi du test post-hoc de Dunn est utilisé dans un premier temps. Puis, si le test de Shapiro-Wilk effectué par la suite ne permet pas de rejeter que les données suivent une loi normale et que le test de Levene ne permet pas d'exclure l'homoscédasticité, un test ANOVA (analyse de variance) peut être réalisé, suivi d'un test post-hoc de Tukey pour tester si les moyennes des différents groupes d'échantillons sont significativement différentes entre elles. Le seuil de p-valeur pour l'ensemble des tests est fixé à 0.05. Les boîtes à moustache sont réalisées avec la bibliothèque ggplot2.

Bibliographie

Abe, K., Nomura, N. & Suzuki, S. Biofilms: hot spots of horizontal gene transfer (HGT) in aquatic environments, with a focus on a new HGT mechanism. *FEMS Microbiol. Ecol.* **96**, fiae031 (2020).

Acar Kirit, H., Lagator, M. & Bollback, J. P. Experimental determination of evolutionary barriers to horizontal gene transfer. *BMC Microbiology* **20**, 326 (2020).

Aluri, S., Rex, K. & Varshney, U. Simultaneous presence of *fhs* and *purT* genes is disadvantageous for the fitness of *Escherichia coli* growth. *FEMS Microbiol. Lett.* **362**, fmv101 (2015).

Amato, P. *et al.* A fate for organic acids, formaldehyde and methanol in cloud water: their biotransformation by micro-organisms. *Atmospheric Chem. Phys.* **7**, 4159–4169 (2007).

Anthamatten, D. & Hennecke, H. The regulatory status of the *fixL*- and *fixJ*-like genes in *Bradyrhizobium japonicum* may be different from that in *Rhizobium meliloti*. *Mol. Gen. Genet.* **225**, 38–48 (1991).

Anthony, C. *Biochemistry of Methylophages*. (Academic Press, 1982).

Antony, C. P. *et al.* *Methylophaga lonarensis* sp. nov., a moderately haloalkaliphilic methylophage isolated from the soda lake sediments of a meteorite impact crater. *Int. J. Syst. Evol. Microbiol.* **62**, 1613–1618 (2012).

Attieh, J. M., Hanson, A. D. & Saini, H. S. Purification and characterization of a novel methyltransferase responsible for biosynthesis of halomethanes and methanethiol in *Brassica oleracea*. *J. Biol. Chem.* **270**, 9250–9257 (1995).

Barrick, J. E. & Lenski, R. E. Genome dynamics during experimental evolution. *Nat. Rev. Genet.* **14**, 827–839 (2013).

Belkhef, S. *et al.* Continuous culture adaptation of *Methylobacterium extorquens* AM1 and TK 0001 to very high methanol concentrations. *Front. Microbiol.* **10**, 1313 (2019).

Belogurov, G. A. & Lahti, R. A lysine substitute for K⁺. A460K mutation eliminates K⁺ dependence in H⁺-pyrophosphatase of *Carboxydotherrnus hydrogenofomans*. *J. Biol. Chem.* **277**, 49651–49654 (2002).

Bershtein, S. *et al.* Protein homeostasis imposes a barrier on functional integration of horizontally transferred genes in bacteria. *PLoS Genet* **11**, e1005612 (2015).

Bertani, G. Studies on lysogenesis. I. The mode of phage liberation by lysogenic *Escherichia coli*. *J. Bacteriol.* **62**, 293–300 (1951).

Bhattacharya, S. *et al.* A ubiquitous platform for bacterial nanotube biogenesis. *Cell. Rep.* **27**, 334–342.e10 (2019).

- Bibi-Triki, S. *et al.* N-terminome and proteogenomic analysis of the *Methylobacterium extorquens* DM4 reference strain for dichloromethane utilization. *J Proteomics* **179**, 131–139 (2018).
- Bolger, A. M., Lohse, M. & Usadel, B. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* **30**, 2114–2120 (2014).
- Borodina, E., Cox, M. J., McDonald, I. R. & Murrell, J. C. Use of DNA-stable isotope probing and functional gene probes to investigate the diversity of methyl chloride-utilizing bacteria in soil. *Environ. Microbiol.* **7**, 1318–1328 (2005).
- Borodina, E., McDonald, I. R. & Murrell, J. C. Chloromethane-dependent expression of the *cmu* gene cluster of *Hyphomicrobium chloromethanicum*. *Appl. Environ. Microbiol.* **70**, 4177–4186 (2004).
- Bringel, F. *et al.* Methylo-trophs and methylo-troph populations for chloromethane degradation. *Curr. Issues Mol. Biol.* **33**, 149–172 (2019).
- Bringel, F. & Couée, I. Pivotal roles of phyllosphere microorganisms at the interface between plant functioning and atmospheric trace gas dynamics. *Front. Microbiol.* **6**, (2015).
- Bringel, F. *et al.* Genome sequence of the dichloromethane-degrading bacterium *Hyphomicrobium* sp. strain GJ21. *Genome Announc.* **5**, e00622-17 (2017).
- Brown, T. A. *Genomes. 2nd edition* (Wiley-Liss, 2002).
- Buddhuhs, N. *et al.* Complete genome sequence of the marine methyl-halide oxidizing *Leisingera methylohalidivorans* type strain (DSM 14336T), a representative of the Roseobacter clade. *Stand. Genomic. Sci.* **9**, 128–141 (2013).
- Camacho, C. *et al.* BLAST+: architecture and applications. *BMC Bioinform.* **10**, 421 (2009).
- Carpenter, L. J. *et al.* in *Scientific Assessment of Ozone Depletion* (ed. Ennis, C. A.) 21–125 (World Meteorological Organization, 2014).
- Carrillo, M. *et al.* Design and control of extrachromosomal elements in *Methylobacterium extorquens* AM1. *ACS Synth. Biol.* **8**, 2451–2456 (2019).
- Carroll, S. M., Lee, M.-C. & Marx, C. J. Sign epistasis limits evolutionary trade-offs at the confluence of single and multi-carbon metabolism in *Methylobacterium extorquens* AM1. *Evolution* **68**, 760–771 (2014).
- Carroll, S. M., Xue, K. S. & Marx, C. J. Laboratory divergence of *Methylobacterium extorquens* AM1 through unintended domestication and past selection for antibiotic resistance. *BMC Microbiol.* **14**, 2 (2014).
- Celeste, L. R. *et al.* Mechanism of N10-formyltetrahydrofolate synthetase derived from complexes with intermediates and inhibitors. *Protein Sci.* **21**, 219–228 (2012).

- Chaignaud, P. *et al.* Genomic and transcriptomic analysis of growth-supporting dehalogenation of chlorinated methanes in *Methylobacterium*. *Front. Microbiol.* **8**, 1600 (2017).
- Chaignaud, P. *et al.* Methanol consumption drives the bacterial chloromethane sink in a forest soil. *ISME J* **12**, 2681–2693 (2018).
- Chimileski, S., Dolas, K., Naor, A., Gophna, U. & Papke, R. T. Extracellular DNA metabolism in *Haloferax volcanii*. *Front. Microbiol.* **5**, 57 (2014).
- Chistoserdova, L. Modularity of methylotrophy, revisited. *Environ. Microbiol.* **13**, 2603–2622 (2011).
- Chistoserdova, L., Chen, S.-W., Lapidus, A. & Lidstrom, M. E. Methylotrophy in *Methylobacterium extorquens* AM1 from a genomic point of view. *J. Bacteriol.* **185**, 2980–2987 (2003).
- Chopra, I. & Roberts, M. Tetracycline antibiotics: mode of action, applications, molecular biology, and epidemiology of bacterial resistance. *Microbiol. Mol. Biol. Rev.* **65**, 232–260 (2001).
- Chou, H.-H., Berthet, J. & Marx, C. J. Fast growth increases the selective advantage of a mutation arising recurrently during evolution under metal limitation. *PLoS Genet.* **5**, e1000652 (2009).
- Chou, H.-H., Chiu, H.-C., Delaney, N. F., Segrè, D. & Marx, C. J. Diminishing returns epistasis among beneficial mutations decelerates adaptation. *Science* **332**, 1190–1192 (2011).
- Chou, H.-H. & Marx, C. J. Optimization of gene expression through divergent mutational paths. *Cell Rep* **1**, 133–140 (2012).
- Chubiz, L. M., Purswani, J., Carroll, S. M. & Marx, C. J. A novel pair of inducible expression vectors for use in *Methylobacterium extorquens*. *BMC Res. Notes* **6**, 183 (2013).
- Cock, P. J. A. *et al.* Biopython: freely available Python tools for computational molecular biology and bioinformatics. *Bioinformatics* **25**, 1422–1423 (2009).
- Conrad, T. M. *et al.* RNA polymerase mutants found through adaptive evolution reprogram *Escherichia coli* for optimal growth in minimal media. *Proc. Natl. Acad. Sci. U.S.A.* **107**, 20500–20505 (2010).
- Cox, M. J., Schäfer, H., Nightingale, P. D., McDonald, I. R. & Murrell, J. C. Diversity of methyl halide-degrading microorganisms in oceanic and coastal waters. *FEMS Microbiol. Lett.* **334**, 111–118 (2012).
- Crowther, G. J., Kosály, G. & Lidstrom, M. E. Formate as the main branch point for methylotrophic metabolism in *Methylobacterium extorquens* AM1. *J. Bacteriol.* **190**, 5057–5062 (2008).

- Cui, L.-Y. *et al.* Breeding of methanol-tolerant *Methylobacterium extorquens* AM1 by atmospheric and room temperature plasma mutagenesis combined with adaptive laboratory evolution. *Biotechnol. J.* **13**, 1700679 (2018).
- Dagert, M. & Ehrlich, S. D. Prolonged incubation in calcium chloride improves the competence of *Escherichia coli* cells. *Gene* **6**, 23–28 (1979).
- Dalldorf, C. *et al.* The hallmarks of a tradeoff in transcriptomes that balances stress and growth functions. *Res. Sq.* rs.3.rs-2729651 (2023).
- Davison, J. Genetic exchange between bacteria in the environment. *Plasmid* **42**, 73–91 (1999).
- Delaney, N. F. *et al.* Development of an optimized medium, strain and high-throughput culturing methods for *Methylobacterium extorquens*. *PLoS One* **8**, e62957 (2013).
- Doronina, N. V., Sokolov, A. P. & Trotsenko, Y. A. Isolation and initial characterization of aerobic chloromethane-utilizing bacteria. *FEMS Microbiol. Lett.* **142**, 179–183 (1996).
- Dubey, G. P. & Ben-Yehuda, S. Intercellular nanotubes mediate bacterial communication. *Cell* **144**, 590–600 (2011).
- Evans, G. J., Ferguson, G. P., Booth, I. R. & Vuilleumier, S. Growth inhibition of *Escherichia coli* by dichloromethane in cells expressing dichloromethane dehalogenase/glutathione S-transferase. *Microbiology (UK)* **146**, 2967–2975 (2000).
- Evans, R. *et al.* Protein complex prediction with AlphaFold-Multimer. *bioRxiv* 2021.10.04.463034 (2022).
- Falk-Krzesinski, H. J. & Wolfe, A. J. Genetic analysis of the *nuo* locus, which encodes the proton-translocating NADH dehydrogenase in *Escherichia coli*. *J. Bacteriol.* **180**, 1174–1184 (1998).
- Fan, H. & Chu, J.-Y. A brief review of short tandem repeat mutation. *Genomics Proteomics Bioinformatics* **5**, 7–14 (2007).
- Farhan Ul Haque, M., Nadalig, T., Bringel, F., Schaller, H. & Vuilleumier, S. Fluorescence-based bacterial bioreporter for specific detection of methyl halide emissions in the environment. *Appl. Environ. Microbiol.* **79**, 6561–6567 (2013).
- Figurski, D. H. & Helinski, D. R. Replication of an origin-containing derivative of plasmid RK2 dependent on a plasmid function provided *in trans*. *Proc. Natl. Acad. Sci. U.S.A.* **76**, 1648–1652 (1979).
- Freedman, D. L., Swamy, M., Bell, N. C. & Verce, M. F. Biodegradation of chloromethane by *Pseudomonas aeruginosa* strain NB1 under nitrate-reducing and aerobic conditions. *Appl. Environ. Microbiol.* **70**, 4629–4634 (2004).
- Fu, X. *et al.* DegP functions as a critical protease for bacterial acid resistance. *FEBS J.* **285**, 3525–3538 (2018).

- Fulton, G. L., Nunn, D. N. & Lidstrom, M. E. Molecular cloning of a methyl coenzyme A lyase gene from *Pseudomonas* sp. strain AM1, a facultative methylotroph. *J. Bacteriol.* **160**, 718–723 (1984).
- Galbally, I. E. & Kirstine, W. The production of methanol by flowering plants and the global cycle of methanol. *J. Atmos. Chem.* **43**, 195–229 (2002).
- Gay, P., Le Coq, D., Steinmetz, M., Ferrari, E. & Hoch, J. A. Cloning structural gene *sacB*, which codes for exoenzyme levansucrase of *Bacillus subtilis*: expression of the gene in *Escherichia coli*. *J. Bacteriol.* **153**, 1424–1431 (1983).
- Gay, P., Le Coq, D., Steinmetz, M., Berkelman, T. & Kado, C. I. Positive selection procedure for entrapment of insertion sequence elements in gram-negative bacteria. *J. Bacteriol.* **164**, 918–921 (1985).
- Gleizer, S. *et al.* Conversion of *Escherichia coli* to generate all biomass carbon from CO₂. *Cell* **179**, 1255–1263.e12 (2019).
- Goddard, T. D. *et al.* UCSF ChimeraX: meeting modern challenges in visualization and analysis. *Protein Sci.* **27**, 14–25 (2018).
- Gresham, D. & Dunham, M. J. The enduring utility of continuous culturing in experimental evolution. *Genomics* **104**, 399–405 (2014).
- Gruffaz, C. *et al.* Genes of the N-methylglutamate pathway are essential for growth of *Methylobacterium extorquens* DM4 with monomethylamine. *Appl. Environ. Microbiol.* **80**, 3541–3550 (2014).
- Guo, M., Ye, J., Gao, D., Xu, N. & Yang, J. *Agrobacterium*-mediated horizontal gene transfer: mechanism, biotechnological application, potential risk and forestalling strategy. *Biotechnol. Adv.* **37**, 259–270 (2019).
- Gyles, C. & Boerlin, P. Horizontally transferred genetic elements and their role in pathogenesis of bacterial disease. *Vet. Pathol.* **51**, 328–340 (2014).
- Hall, T. A. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symposium Series* (1999).
- Halsey, K. H., Carter, A. E. & Giovannoni, S. J. Synergistic metabolism of a broad range of C1 compounds in the marine methylotrophic bacterium HTCC2181. *Environ. Microbiol.* **14**, 630–640 (2012).
- Hamilton, J. T. G., McRoberts, W. C., Keppler, F., Kalin, R. M. & Harper, D. B. Chloride methylation by plant pectin: an efficient environmentally significant process. *Science* **301**, 206–209 (2003).
- Harper, D. B. Halomethane from halide ion—a highly efficient fungal conversion of environmental significance. *Nature* **315**, 55–57 (1985).

- Harper, D. B., Buswell, J. A., Kennedy, J. T. & Hamilton, J. T. G. Chloromethane, methyl donor in veratryl alcohol biosynthesis in *Phanerochaete chrysosporium* and other lignin-degrading fungi. *Appl. Environ. Microbiol.* **56**, 3450–3457 (1990).
- Hayoun, K. *et al.* Dichloromethane degradation pathway from unsequenced *Hyphomicrobium* sp. MC8b rapidly explored by pan-proteomics. *Microorganisms* **8**, 1876 (2020).
- Heikes, B. G. *et al.* Atmospheric methanol budget and ocean implication. *Glob. Biogeochem. Cycles* **16**, 80-1-80–13 (2002).
- Hu, B. *et al.* Comprehensive molecular characterization of *Methylobacterium extorquens* AM1 adapted for 1-butanol tolerance. *Biotechnol. Biofuels* **9**, 84 (2016).
- Hutchison, C. A. *et al.* Polar effects of transposon insertion into a minimal bacterial genome. *J. Bacteriol* **201**, e00185-19 (2019).
- Jaeger, N. *et al.* Chloromethane formation and degradation in the fern phyllosphere. *Sci. Total Environ.* **634**, 1278–1287 (2018).
- Janczarek, M. The Ros/MucR zinc-finger protein family in bacteria: structure and functions. *Int. J. Mol. Sci.* **23**, 15536 (2022).
- Johnson, C. M. & Grossman, A. D. Integrative and conjugative elements (ICEs): what they do and how they work. *Annu Rev Genet* **49**, 577–601 (2015).
- Johnston, C., Martin, B., Fichant, G., Polard, P. & Claverys, J.-P. Bacterial transformation: distribution, shared mechanisms and divergent control. *Nat. Rev. Microbiol.* **12**, 181–196 (2014).
- Jones, C. T., Susko, E. & Bielawski, J. P. Evolution of the connectivity and indispensability of a transferable gene: the simplicity hypothesis. *BMC Ecol. Evol.* **22**, 140 (2022).
- Jörg, G. & Bertau, M. Thiol-tolerant assay for quantitative colorimetric determination of chloride released from whole-cell biodehalogenations. *Anal. Biochem.* **328**, 22–28 (2004).
- Jumper, J. *et al.* Highly accurate protein structure prediction with AlphaFold. *Nature* **596**, 583–589 (2021).
- Kalyuzhnaya, M. G. & Lidstrom, M. E. QscR-mediated transcriptional activation of serine cycle genes in *Methylobacterium extorquens* AM1. *J. Bacteriol.* **187**, 7511–7517 (2005).
- Karalkar, N. B., Khare, K., Molt, R. & Benner, S. A. Tautomeric equilibria of isoguanine and related purine analogs. *Nucleosides Nucleotides Nucleic Acids* **36**, 256–274 (2017).
- Kato, Y., Asahara, M., Arai, D., Goto, K. & Yokota, A. Reclassification of *Methylobacterium chloromethanicum* and *Methylobacterium dichloromethanicum* as later subjective synonyms of *Methylobacterium extorquens* and of *Methylobacterium lusitanum* as a later subjective synonym of *Methylobacterium rhodesianum*. *J. Gen. Appl. Microbiol.* **51**, 287–299 (2005).

- Kayser, M. F. & Vuilleumier, S. Dehalogenation of dichloromethane by dichloromethane dehalogenase/glutathione S-transferase leads to formation of DNA adducts. *J. Bacteriol.* **183**, 5209–5212 (2001).
- Keppler, F., Kalin, R. M., Harper, D. B., McRoberts, W. C. & Hamilton, J. T. G. Carbon isotope anomaly in the major plant C1 pool and its global biogeochemical implications. *Biogeosci.* **1**, 123–131 (2004).
- Keppler, F., Harper, D. B., Röckmann, T., Moore, R. M. & Hamilton, J. T. G. New insight into the atmospheric chloromethane budget gained using stable carbon isotope ratios. *Atmos. Chem. Phys.* **5**, 2403–2411 (2005).
- Keppler, F. *et al.* Chlorine isotope fractionation of the major chloromethane degradation processes in the environment. *Environ Sci Technol* **54**, 1634–1645 (2020).
- Keppler, F. *et al.* Sources and sinks of chloromethane in a salt marsh ecosystem: constraints from concentration and stable isotope measurements of laboratory incubation experiments. *Environ. Sci.: Process Impacts* **22**, 627–641 (2020).
- Kim, H., Wu, K. & Lee, C. Stress-responsive periplasmic chaperones in bacteria. *Front. Mol. Biosci.* **8**, 678697 (2021).
- Kim, S., Lee, S. H., Seo, H. & Kim, K.-J. Biochemical properties and crystal structure of formate-tetrahydrofolate ligase from *Methylobacterium extorquens* CM4. *Biochem. Biophys. Res. Commun.* **528**, 426–431 (2020).
- Knief, C., Frances, L. & Vorholt, J. A. Competitiveness of diverse *Methylobacterium* strains in the phyllosphere of *Arabidopsis thaliana* and identification of representative models, including *M. extorquens* PA1. *Microb. Ecol.* **60**, 440–452 (2010).
- Knöppel, A., Lind, P. A., Lustig, U., Näsvall, J. & Andersson, D. I. Minor fitness costs in an experimental model of horizontal gene transfer in bacteria. *Mol Biol Evol* **31**, 1220–1227 (2014).
- Kolusu, S. R., Schlünzen, K. H., Grawe, D. & Seifert, R. Determination of chloromethane and dichloromethane in a tropical terrestrial mangrove forest in Brazil by measurements and modelling. *Atmos. Environ.* **173**, 185–197 (2018).
- Koo, C. W. & Rosenzweig, A. C. Biochemistry of aerobic biological methane oxidation. *Chem. Soc. Rev.* **50**, 3424–3436 (2021).
- Lang, A. S., Zhaxybayeva, O. & Beatty, J. T. Gene transfer agents: phage-like elements of genetic exchange. *Nat. Rev. Microbiol.* **10**, 472–482 (2012).
- Large, P. J., Peel, D. & Quayle, J. R. Microbial growth on C1 compounds. 2. Synthesis of cell constituents by methanol- and formate-grown *Pseudomonas* AM1, and methanol-grown *Hyphomicrobium vulgare*. *Biochem. J.* **81**, 470–480 (1961).
- Large, P. J. & Quayle, J. R. Microbial growth on C1 compounds. 5. Enzyme activities in extracts of *Pseudomonas* AM1. *Biochem. J.* **87**, 386–396 (1963).

- Lee, M.-C., Chou, H.-H. & Marx, C. J. Asymmetric, bimodal trade-offs during adaptation of *Methylobacterium* to distinct growth substrates. *Evolution* **63**, 2816–2830 (2009).
- Lee, M.-C. & Marx, C. J. Synchronous waves of failed soft sweeps in the laboratory: remarkably rampant clonal interference of alleles at a single locus. *Genetics* **193**, 943–952 (2013).
- Lenski, R. E. Revisiting the design of the long-term evolution experiment with *Escherichia coli*. *J. Mol. Evol.* **91**, 241–253 (2023).
- Leong, S. A., Ditta, G. S. & Helinski, D. R. Heme biosynthesis in *Rhizobium*. Identification of a cloned gene coding for delta-aminolevulinic acid synthetase from *Rhizobium meliloti*. *J. Biol. Chem.* **257**, 8724–8730 (1982).
- Lerminiaux, N. A. & Cameron, A. D. S. Horizontal transfer of antibiotic resistance genes in clinical environments. *Can. J. Microbiol.* **65**, 34–44 (2019).
- Li, H. Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. Preprint at <https://doi.org/10.48550/arXiv.1303.3997> (2013)
- López-Marqués, R. L., Pérez-Castiñeira, J. R., Losada, M. & Serrano, A. Differential regulation of soluble and membrane-bound inorganic pyrophosphatases in the photosynthetic bacterium *Rhodospirillum rubrum* provides insights into pyrophosphate-based stress bioenergetics. *J. Bacteriol.* **186**, 5418–5426 (2004).
- Lovell, C. R. & Leaphart, A. B. Community-level analysis: key genes of CO₂-reductive acetogenesis. *Meth. Enzymol.* **397**, 454–469 (2005).
- Luo, J. A new pathway for bacterial growth with chloromethane. (2021). Thèse de Doctorat, Université de Strasbourg, <https://www.theses.fr/2021STRAJ003>.
- Malhotra, A., Severinova, E. & Darst, S. A. Crystal structure of a sigma 70 subunit fragment from *E. coli* RNA polymerase. *Cell* **87**, 127–136 (1996).
- Manley, S. L., Wang, N.-Y., Walser, M. L. & Cicerone, R. J. Methyl halide emissions from greenhouse-grown mangroves. *Geophys. Res. Lett.* **34**, L01806 (2007).
- Marx, C. J. Development of a broad-host-range *sacB*-based vector for unmarked allelic exchange. *BMC Res. Notes* **1**, 1 (2008).
- Marx, C. J., Van Dien, S. J. & Lidstrom, M. E. Flux analysis uncovers key role of functional redundancy in formaldehyde metabolism. *PLoS Biol.* **3**, e16 (2005).
- Marx, C. J. Experimental evolution of *Methylobacterium*: 15 Years of Planned Experiments and Surprise Findings. *Curr. Issues Mol. Biol.* **33**, 249–266 (2019).
- Marx, C. J. *et al.* Complete genome sequences of six strains of the genus *Methylobacterium*. *J. Bacteriol.* **194**, 4746–4748 (2012).

Marx, C. J., Laukel, M., Vorholt, J. A. & Lidstrom, M. E. Purification of the formate-tetrahydrofolate ligase from *Methylobacterium extorquens* AM1 and demonstration of its requirement for methylotrophic growth. *J. Bacteriol.* **185**, 7169–7175 (2003).

Marx, C. J. & Lidstrom, M. E. Development of improved versatile broad-host-range vectors for use in methylotrophs and other Gram-negative bacteria. *Microbiology (UK)* **147**, 2065–2075 (2001).

Marx, C. J. & Lidstrom, M. E. Development of an insertional expression vector system for *Methylobacterium extorquens* AM1 and generation of null mutants lacking *mtdA* and/or *fch*. *Microbiology (UK)* **150**, 9–19 (2004).

Marx, C. J., O'Brien, B. N., Breezee, J. & Lidstrom, M. E. Novel methylotrophy genes of *Methylobacterium extorquens* AM1 identified by using transposon mutagenesis including a putative dihydromethanopterin reductase. *J. Bacteriol.* **185**, 669–673 (2003).

Maucourt, B., Roche, D., Chaignaud, P., Vuilleumier, S. & Bringel, F. Genome-wide transcription start sites mapping in *Methylobacterium* grown with dichloromethane and methanol. *Microorganisms* **10**, 1301 (2022).

McAnulla, C. *et al.* Chloromethane utilization gene cluster from *Hyphomicrobium chloromethanicum* strain CM2T and development of functional gene probes to detect halomethane-degrading bacteria. *Appl. Environ. Microbiol.* **67**, 307–316 (2001).

McDonald, I. R., Doronina, N. V., Trotsenko, Y. A., McAnulla, C. & Murrell, J. C. *Hyphomicrobium chloromethanicum* sp. nov. and *Methylobacterium chloromethanicum* sp. nov., chloromethane-utilizing bacteria isolated from a polluted environment. *Int. J. Syst. Evol. Microbiol.* **51**, 119–122 (2001).

McDonald, I. R. *et al.* *Aminobacter ciceronei* sp. nov. and *Aminobacter lissarensis* sp. nov., isolated from various terrestrial environments. *Int. J. Syst. Evol. Microbiol.* **55**, 1827–1832 (2005).

Metzger, L. C., Francez-Charlot, A. & Vorholt, J. A. Single-domain response regulator involved in the general stress response of *Methylobacterium extorquens*. *Microbiology (UK)* **159**, 1067–1076 (2013).

Michener, J. K., Camargo Neves, A. A., Vuilleumier, S., Bringel, F. & Marx, C. J. Effective use of a horizontally-transferred pathway for dichloromethane catabolism requires post-transfer refinement. *eLife* **3**, e04279 (2014).

Michener, J. K., Vuilleumier, S., Bringel, F. & Marx, C. J. Phylogeny poorly predicts the utility of a challenging horizontally transferred gene in *Methylobacterium* strains. *J. Bacteriol.* **196**, 2101–2107 (2014).

Michener, J. K., Vuilleumier, S., Bringel, F. & Marx, C. J. Transfer of a catabolic pathway for chloromethane in *Methylobacterium* strains highlights different limitations for growth with chloromethane or with dichloromethane. *Front. Microbiol.* **7**, 1116 (2016).

- Mo, X.-H. *et al.* Establishment of CRISPR interference in *Methylobacterium extorquens* and application of rapidly mining a new phytoene desaturase involved in carotenoid biosynthesis. *Appl. Microbiol. Biotechnol.* **104**, 4515–4532 (2020).
- Møller, T. S. B. *et al.* Relation between *tetR* and *tetA* expression in tetracycline resistant *Escherichia coli*. *BMC Microbiol.* **16**, (2016).
- Muller, E. E. L., Bringel, F. & Vuilleumier, S. Dichloromethane-degrading bacteria in the genomic age. *Res. Microbiol.* **162**, 869–876 (2011).
- Nadalig, T. *et al.* Detection and isolation of chloromethane-degrading bacteria from the *Arabidopsis thaliana* phyllosphere, and characterization of chloromethane utilization genes. *FEMS Microbiol. Ecol.* **77**, 438–448 (2011).
- Nadalig, T., Greule, M., Bringel, F., Keppler, F. & Vuilleumier, S. Probing the diversity of chloromethane-degrading bacteria by comparative genomics and isotopic fractionation. *Front. Microbiol.* **5**, 523 (2014).
- Nagatoshi, Y. & Nakamura, T. *Arabidopsis* HARMLESS TO OZONE LAYER protein methylates a glucosinolate breakdown product and functions in resistance to *Pseudomonas syringae* pv. *maculicola*. *J. Biol. Chem.* **284**, 19301–19309 (2009).
- Nagy, P. L., Marolewski, A., Benkovic, S. J. & Zalkin, H. Formyltetrahydrofolate hydrolase, a regulatory enzyme that functions to balance pools of tetrahydrofolate and one-carbon tetrahydrofolate adducts in *Escherichia coli*. *J. Bacteriol.* **177**, 1292–1298 (1995).
- Naor, A. & Gophna, U. Cell fusion and hybrids in archaea: prospects for genome shuffling and accelerated strain development for biotechnology. *Bioengineered* **4**, 126–129 (2013).
- Nayak, D. D. & Marx, C. J. Genetic and phenotypic comparison of facultative methylotrophy between *Methylobacterium extorquens* strains PA1 and AM1. *PLoS One* **9**, e107887 (2014).
- Nayak, D. D. & Marx, C. J. Methylamine utilization via the N-methylglutamate pathway in *Methylobacterium extorquens* PA1 involves a novel flow of carbon through C1 assimilation and dissimilation pathways. *J. Bacteriol.* **196**, 4130–4139 (2014).
- Nayak, D. D. & Marx, C. J. Experimental horizontal gene transfer of methylamine dehydrogenase mimics prevalent exchange in nature and overcomes the methylamine growth constraints posed by the sub-optimal N-methylglutamate pathway. *Microorganisms* **3**, 60–79 (2015).
- Nellen-Anthamatten, D. *et al.* *Bradyrhizobium japonicum* FixK2, a crucial distributor in the FixLJ-dependent regulatory cascade for control of genes inducible by low oxygen levels. *J. Bacteriol.* **180**, 5251–5255 (1998).
- Ning, Z., Cox, A. J. & Mullikin, J. C. SSAHA: a fast search method for large DNA databases. *Genome Res* **11**, 1725–1729 (2001).

Nygaard, P. & Smith, J. M. Evidence for a novel glycinamide ribonucleotide transformylase in *Escherichia coli*. *J. Bacteriol.* **175**, 3591–3597 (1993).

Ochsner, A. M. *et al.* Transposon sequencing uncovers an essential regulatory function of phosphoribulokinase for methylotrophy. *Curr. Biol.* **27**, 2579–2588.e6 (2017).

Ochsner, A. M. *et al.* Use of rare-earth elements in the phyllosphere colonizer *Methylobacterium extorquens* PA1. *Mol. Microbiol.* **111**, 1152–1166 (2019).

Ochsner, A. M., Sonntag, F., Buchhaupt, M., Schrader, J. & Vorholt, J. A. *Methylobacterium extorquens*: methylotrophy and biotechnological applications. *Appl Microbiol Biotechnol* **99**, 517–534 (2015).

Peel, D. & Quayle, J. R. Microbial growth on C1 compounds. 1. Isolation and characterization of *Pseudomonas* AM1. *Biochem. J.* **81**, 465–469 (1961).

Pettersen, E. F. *et al.* UCSF ChimeraX: Structure visualization for researchers, educators, and developers. *Protein Sci.* **30**, 70–82 (2021).

Peyraud, R. *et al.* Demonstration of the ethylmalonyl-CoA pathway by using ¹³C metabolomics. *Proc. Natl. Acad. Sci. U.S.A.* **106**, 4846–4851 (2009).

Peyraud, R. *et al.* Genome-scale reconstruction and system level investigation of the metabolic network of *Methylobacterium extorquens* AM1. *BMC Syst. Biol.* **5**, 189 (2011).

Phaneuf, P. V., Gosting, D., Palsson, B. O. & Feist, A. M. ALEdb 1.0: a database of mutations from adaptive laboratory evolution experimentation. *Nucl. Acids Res.* **47**, D1164–D1171 (2019).

Pomper, B. K., Vorholt, J. A., Chistoserdova, L., Lidstrom, M. E. & Thauer, R. K. A methenyl tetrahydromethanopterin cyclohydrolase and a methenyl tetrahydrofolate cyclohydrolase in *Methylobacterium extorquens* AM1. *European Journal of Biochemistry* **261**, 475–480 (1999).

Popa, O. & Dagan, T. Trends and barriers to lateral gene transfer in prokaryotes. *Curr. Opin. Microbiol.* **14**, 615–623 (2011).

Price, E. E., Román-Rodríguez, F. & Boyd, J. M. Bacterial approaches to sensing and responding to respiration and respiration metabolites. *Mol. Microbiol.* **116**, 1009–1021 (2021).

Rabinowitz, J. C. & Pricer, W. E. Formyltetrahydrofolate synthetase. I. Isolation and crystallization of the enzyme. *J. Biol. Chem.* **237**, 2898–2902 (1962).

Ragsdale, S. W. & Pierce, E. Acetogenesis and the Wood-Ljungdahl pathway of CO₂ fixation. *Biochim. Biophys. Acta.* **1784**, 1873–1898 (2008).

Rhew, R. C., Østergaard, L., Saltzman, E. S. & Yanofsky, M. F. Genetic control of methyl halide production in *Arabidopsis*. *Curr. Biol.* **13**, 1809–1813 (2003).

Rodionov, D. A., Vitreschak, A. G., Mironov, A. A. & Gelfand, M. S. Comparative genomics of the vitamin B12 metabolism and regulation in prokaryotes. *J. Biol. Chem.* **278**, 41148–41159 (2003).

Roselli, S. Génomique fonctionnelle de la dégradation microbienne du chlorométhane. Thèse de Doctorat, Université Louis-Pasteur (2009) <https://www.theses.fr/2009STRA6227>.

Roselli, S., Nadalig, T., Vuilleumier, S. & Bringel, F. The 380 kb pCMU01 plasmid encodes chloromethane utilization genes and redundant genes for vitamin B12- and tetrahydrofolate-dependent chloromethane metabolism in *Methylobacterium extorquens* CM4: a proteomic and bioinformatics study. *PLoS One* **8**, e56598 (2013).

Sampei, G. *et al.* Structures and reaction mechanisms of the two related enzymes, PurN and PurU. *J. Biochem.* **154**, 569–579 (2013).

Schada von Borzyskowski, L. *et al.* An engineered Calvin-Benson-Bassham cycle for carbon dioxide fixation in *Methylobacterium extorquens* AM1. *Metab. Eng.* **47**, 423–433 (2018).

Schada von Borzyskowski, L., Remus-Emsermann, M., Weishaupt, R., Vorholt, J. A. & Erb, T. J. A set of versatile brick vectors and promoters for the assembly, expression, and integration of synthetic operons in *Methylobacterium extorquens* AM1 and other alphaproteobacteria. *ACS Synth. Biol.* **4**, 430–443 (2015).

Schaefer, J. K., Goodwin, K. D., McDonald, I. R., Murrell, J. C. & Oremland, R. S. *Leisingera methylohalidivorans* gen. nov., sp. nov., a marine methylotroph that grows on methyl bromide. *Int. J. Syst. Evol. Microbiol.* **52**, 851–859 (2002).

Schäfer, H., McDonald, I. R., Nightingale, P. D. & Murrell, J. C. Evidence for the presence of a CmuA methyltransferase pathway in novel marine methyl halide-oxidizing bacteria. *Environ. Microbiol.* **7**, 839–852 (2005).

Schleper, C., Holz, I., Janekovic, D., Murphy, J. & Zillig, W. A multicopy plasmid of the extremely thermophilic archaeon *Sulfolobus* effects its transfer to recipients by mating. *J. Bacteriol.* **177**, 4417–4426 (1995).

Schneider, K. *et al.* The ethylmalonyl-CoA pathway is used in place of the glyoxylate cycle by *Methylobacterium extorquens* AM1 during growth on acetate. *J. Biol. Chem.* **287**, 757–766 (2012).

Scott, J. M. & Rabinowitz, J. C. The association-dissociation of formyltetrahydrofolate synthetase and its relation to monovalent cation activation of catalytic activity. *Biochem. Biophys. Res. Commun.* **29**, 418–423 (1967).

Singh, A., Müller, B., Fuxelius, H.-H. & Schnürer, A. AcetoBase: a functional gene repository and database for formyltetrahydrofolate synthetase sequences. *Database (Oxford)* **2019**, baz142 (2019).

Skovran, E. & Martinez-Gomez, N. C. Microbiology. Just add lanthanides. *Science* **348**, 862–863 (2015).

- Smith, T. F. & Waterman, M. S. Identification of common molecular subsequences. *J. Mol. Biol.* **147**, 195–197 (1981).
- Soler, N. & Forterre, P. Vesiduction: the fourth way of HGT. *Environ. Microbiol.* **22**, 2457–2460 (2020).
- Sorek, R. *et al.* Genome-wide experimental determination of barriers to horizontal gene transfer. *Science* **318**, 1449–1452 (2007).
- Soucy, S. M., Huang, J. & Gogarten, J. P. Horizontal gene transfer: building the web of life. *Nat. Rev. Genet.* **16**, 472–482 (2015).
- Southern, E. M. Detection of specific sequences among DNA fragments separated by gel electrophoresis. *J. Mol. Biol.* **98**, 503–517 (1975).
- Studer, A., Stupperich, E., Vuilleumier, S. & Leisinger, T. Chloromethane: tetrahydrofolate methyl transfer by two proteins from *Methylobacterium chloromethanicum* strain CM4. *Eur. J. Biochem.* **268**, 2931–2938 (2001).
- Studer, A., McAnulla, C., Büchele, R., Leisinger, T. & Vuilleumier, S. Chloromethane-induced genes define a third C1 utilization pathway in *Methylobacterium chloromethanicum* CM4. *J. Bacteriol.* **184**, 3476–3484 (2002).
- Studer, A., Vuilleumier, S. & Leisinger, T. Properties of the methylcobalamin:H4folate methyltransferase involved in chloromethane utilization by *Methylobacterium* sp. strain CM4. *Eur. J. Biochem.* **264**, 242–249 (1999).
- Studer, A. Aerobic microbial degradation of chloromethane. (2001). Thèse de Doctorat, École Polytechnique Fédérale de Zurich. <https://www.research-collection.ethz.ch/>
- Taylor, R. G., Walker, D. C. & McInnes, R. R. *E. coli* host strains significantly affect the quality of small scale plasmid DNA preparations used for sequencing. *Nucl. Acids Res.* **21**, 1677–1678 (1993).
- Toyama, H., Anthony, C. & Lidstrom, M. E. Construction of insertion and deletion *mxs* mutants of *Methylobacterium extorquens* AM1 by electroporation. *FEMS Microbiol Lett* **166**, 1–7 (1998).
- Traunecker, J., Preuß, A. & Diekert, G. Isolation and characterization of a methyl chloride utilizing, strictly anaerobic bacterium. *Arch. Microbiol.* **156**, 416–421 (1991).
- Tsai, W.-T. Fate of chloromethanes in the atmospheric environment: implications for human health, ozone formation and depletion, and global warming impacts. *Toxics* **5**, 23 (2017).
- Vallenet, D. *et al.* MicroScope: a platform for microbial genome annotation and comparative genomics. *Database (Oxford)* **2009**, bap021 (2009).
- Vallenet, D. *et al.* MicroScope: an integrated platform for the annotation and exploration of microbial gene functions through genomic, pangenomic and metabolic comparative analysis. *Nucleic Acids Res.* **48**, D579–D589 (2020).

Van den Bergh, B., Swings, T., Fauvart, M. & Michiels, J. Experimental design, population dynamics, and diversity in microbial experimental evolution. *Microbiol. Mol. Biol. Rev.* **82**, e00008-18 (2018).

Van Dien, S. J., Strovas, T. & Lidstrom, M. E. Quantification of central metabolic fluxes in the facultative methylotroph *Methylobacterium extorquens* AM1 using ¹³C-label tracing and mass spectrometry. *Biotechnol. Bioeng.* **84**, 45–55 (2003).

Vannelli, T., Messmer, M., Studer, A., Vuilleumier, S. & Leisinger, T. A corrinoid-dependent catabolic pathway for growth of a *Methylobacterium* strain with chloromethane. *Proc. Natl. Acad. Sci. U.S.A.* **96**, 4615–4620 (1999).

Vannelli, T., Studer, A., Kertesz, M. & Leisinger, T. Chloromethane metabolism by *Methylobacterium* sp. strain CM4. *Appl. Environ. Microbiol.* **64**, 1933–1936 (1998).

Varner, P. M. & Gunsch, C. K. Properties affecting transfer and expression of degradative plasmids for the purpose of bioremediation. *Biodegradation* **32**, 361–375 (2021).

Virolle, C., Goldlust, K., Djermoun, S., Bigot, S. & Lesterlin, C. Plasmid transfer by conjugation in gram-negative bacteria: from the cellular to the community level. *Genes (Basel)* **11**, 1239 (2020).

Vorholt, J. A. Cofactor-dependent pathways of formaldehyde oxidation in methylotrophic bacteria. *Arch Microbiol* **178**, 239–249 (2002).

Vorholt, J. A., Chistoserdova, L., Lidstrom, M. E. & Thauer, R. K. The NADP-dependent methylene tetrahydromethanopterin dehydrogenase in *Methylobacterium extorquens* AM1. *J. Bacteriol.* **180**, 5351–5356 (1998).

Vu, H. N. *et al.* Lanthanide-dependent regulation of methanol oxidation systems in *Methylobacterium extorquens* AM1 and their contribution to methanol growth. *J. Bacteriol.* **198**, 1250–1259 (2016).

Vuilleumier, S. *et al.* *Methylobacterium* genome sequences: a reference blueprint to investigate microbial metabolism of C1 compounds from natural and industrial sources. *PLoS One* **4**, e5584 (2009).

Vuilleumier, S., Ivoš, N., Dean, M. & Leisinger, T. Sequence variation in dichloromethane dehalogenases/glutathione S-transferases. *Microbiology (UK)* **147**, 611–619 (2001).

Vuilleumier, S. *et al.* Complete genome sequence of the chloromethane-degrading *Hyphomicrobium* sp. strain MC1. *J. Bacteriol.* **193**, 5035–5036 (2011).

Watford, S. & Warrington, S. J. Bacterial DNA Mutations. [Updated 2023 Apr 10]. In: StatPearls [Internet]. Treasure Island (FL): StatPearls Publishing (2023), <https://www.ncbi.nlm.nih.gov/books/NBK459274>.

Wellner, A. & Gophna, U. Neutrality of foreign complex subunits in an experimental model of lateral gene transfer. *Mol Biol Evol* **25**, 1835–1840 (2008).

- Wick, R. R., Judd, L. M., Gorrie, C. L. & Holt, K. E. Unicycler: Resolving bacterial genome assemblies from short and long sequencing reads. *PLOS Comput. Biol.* **13**, e1005595 (2017).
- Woodall, C. A., Warner, K. L., Oremland, R. S., Murrell, J. C. & McDonald, I. R. Identification of methyl halide-utilizing genes in the methyl bromide-utilizing bacterial strain IMB-1 suggests a high degree of conservation of methyl halide-specific genes in gram-negative bacteria. *Appl. Environ. Microbiol.* **67**, 1959–1963 (2001).
- Yanisch-Perron, C., Vieira, J. & Messing, J. Improved M13 phage cloning vectors and host strains: nucleotide sequences of the M13mp18 and pUC19 vectors. *Gene* **33**, 103–119 (1985).
- Yokouchi, Y., Ikeda, M., Inuzuka, Y. & Yukawa, T. Strong emission of methyl chloride from tropical plants. *Nature* **416**, 163–165 (2002).
- Yokouchi, Y., Saito, T., Ishigaki, C. & Aramoto, M. Identification of methyl chloride-emitting plants and atmospheric measurements on a subtropical island. *Chemosphere* **69**, 549–553 (2007).
- Yoshida, Y. *et al.* Inverse modeling of the global methyl chloride sources. *J. Geophys. Res.* **111**, (2006).
- Yurimoto, H., Shiraishi, K. & Sakai, Y. Physiology of methylotrophs living in the phyllosphere. *Microorganisms* **9**, 809 (2021).
- Zhang, G. & Darst, S. A. Structure of the *Escherichia coli* RNA polymerase alpha subunit amino-terminal domain. *Science* **281**, 262–266 (1998).
- Zhang, Y., Morar, M. & Ealick, S. E. Structural biology of the purine biosynthetic pathway. *Cell. Mol. Life Sci.* **65**, 3699–3724 (2008).

Communications scientifiques

(nom du présentateur souligné)

Poster (**Prix du meilleur poster**) Louis-François Mey, Stéphanie Fouteau, David Roche, Stéphane Vuilleumier, Françoise Bringel. **Experimental evolution of methylotrophic strains for growth with chloromethane. Microbial cycling of volatile organic compounds: Biogeochemistry to Biotechnology.** 25,26 mai 2022 - Norwich, Royaume-Uni

Présentation orale Louis-François Mey, Stéphanie Fouteau, David Roche, Stéphane Vuilleumier, Françoise Bringel. **Adaptation bactérienne à la croissance avec le chlorométhane après acquisition des gènes de la voie de dégradation *cmu*. Séminaire de Microbiologie de Strasbourg.** 24 mars 2022 - Illkirch-Graffenstaden, France

Présentation orale Louis-François Mey, Stéphane Vuilleumier, Françoise Bringel. **Laboratory-driven bacterial evolution for methylotrophic growth with chloromethane. Young Researchers in Life Science.** 16, 17, 18 juin 2021 – Paris, France (en ligne)

Présentation orale Louis-François Mey, Stéphane Vuilleumier, Françoise Bringel. **Bacterial adaptation to chloromethane, an ozone-depleting gas. Journées de l'Ecole Doctorale Vie et Santé de Strasbourg.** 21, 22 avril 2021 – Strasbourg, France

Poster Louis-François Mey, Yousra Louhichi, Stéphane Vuilleumier, Françoise Bringel. **Bacterial adaptation to the utilization of chloromethane.** Séminaire de Microbiologie de Strasbourg 28 mars 2019 – Strasbourg, France

Adaptations à la déshalogénation et à l'utilisation du chlorométhane par la voie *cmu* chez *Methylobacterium extorquens*

Le chlorométhane (CH_3Cl), un gaz organohalogéné principalement d'origine naturelle impliqué dans la dégradation de l'ozone dans l'atmosphère, est dégradé par des bactéries méthylotrophes capables d'utiliser des composés sans liaisons carbone-carbone tels que le méthanol comme seule source de carbone et d'énergie. Les Alpharotéobactéries méthylotrophes *Methylobacterium extorquens* CM4 et *Hyphomicrobium* sp. MC1 utilisent le CH_3Cl par la voie *cmu* initiée par transfert du carbone de CH_3Cl au tétrahydrofolate (H_4F) catalysé par CmuA et CmuB. Le méthyl- H_4F formé est ensuite dirigé vers la formation de biomasse ou oxydé en CO_2 via le formiate. Cette voie, acquise par transfert horizontal, ne suffit pas à conférer la croissance à une souche réceptrice. Les adaptations aux stress inhérents à la croissance avec le CH_3Cl sont peu connues.

Afin d'identifier de telles adaptations chez *M. extorquens*, le plasmide pJM105 contenant les gènes *cmu* de *Hyphomicrobium* sp. MC1 a été transféré à deux souches de *M. extorquens* (AM1 et DM4) incapables de se développer avec le CH_3Cl . Une approche d'évolution expérimentale sur 200 générations environ a permis d'améliorer la croissance de ces souches avec le CH_3Cl . Le séquençage du génome de 13 des 162 souches évoluées obtenues a révélé 64 mutations (SNP, délétion ou insertion d'un nucléotide, larges délétions) dans 45 gènes, ainsi que d'autres événements génétiques (réarrangement entre plasmides, grandes délétions). Des mutations inactivant le gène de la formyl-tétrahydrofolate ligase (FtfL) essentiel pour l'assimilation du carbone issu du méthanol ont été identifiées dans 8 des 13 souches séquencées. La construction et l'analyse de mutants de ce gène ont montré que l'inactivation de FtfL suffit à conférer la croissance avec le CH_3Cl par la voie *cmu* à *M. extorquens* DM4. Je fais l'hypothèse que cette enzyme interfère avec l'oxydation du CH_3Cl pour la production d'énergie tout en consommant le cofacteur H_4F essentiel à ce métabolisme, et que son activité s'oppose à celle de l'enzyme homologue PurU requise pour la croissance avec le CH_3Cl par la voie *cmu*. Parmi les autres mutations mises en évidence, une mutation dans le promoteur des gènes *nuo* pourrait faciliter l'expression de la NADH déshydrogénase impliquée dans la régénération du cofacteur NAD^+ . D'autres adaptations indépendantes de FtfL ont été détectées, qui diffèrent selon le contexte génétique de la souche receveuse des gènes de la voie *cmu*. Par exemple, les souches évoluées AM1 montrent toutes des mutations dans pJM105 (dans le gène *cmuA*, avec cointégration d'un plasmide endogène). Ces mutations n'ont cependant pas d'impact sur l'activité de déshalogénation du CH_3Cl dépendant de CmuA. L'obtention et la caractérisation de souches mutantes par évolution expérimentale au laboratoire représente un nouvel outil prometteur pour l'étude de la croissance avec le CH_3Cl par la voie *cmu*, notamment du fait des difficultés rencontrées jusqu'ici dans les approches de mutagenèse dirigée des bactéries dégradant le CH_3Cl .

Mots clés : bactéries, méthylotrophie, déshalogénation, chlorométhane, évolution expérimentale, adaptations métaboliques

Adaptations to dehalogenation and utilization of chloromethane by the *cmu* pathway in *Methylobacterium extorquens*

Chloromethane (CH_3Cl), a mainly naturally occurring organohalogen gas, is involved in ozone degradation in the atmosphere. CH_3Cl is degraded by methylotrophic bacteria capable of using compounds without carbon-carbon bonds, such as methanol, as their sole source of carbon and energy. Two such methylotrophic Alphaproteobacteria, *Methylobacterium extorquens* CM4 and *Hyphomicrobium* sp. MC1, utilize CH_3Cl via the *cmu* pathway initiated by carbon transfer from CH_3Cl to tetrahydrofolate (H_4F) catalyzed by CmuA and CmuB. Methyl- H_4F formed is then used for biomass production or oxidized to CO_2 via formate. The *cmu* pathway, often acquired by horizontal transfer, is not sufficient to confer growth to a recipient strain. However, little is known about the adaptations needed to overcome stresses inherent to growth with CH_3Cl .

To identify such adaptations in *M. extorquens*, the plasmid pJM105 containing the *cmu* genes from *Hyphomicrobium* sp. MC1 was transferred to two *M. extorquens* strains (AM1 and DM4) unable to grow with CH_3Cl . An experimental evolution approach, spanning some 200 generations, was performed to improve strain growth with CH_3Cl . Genome sequencing of 13 of the 162 evolved strains obtained revealed 64 mutations (SNPs, nucleotide deletions or insertions, large deletions) in 45 genes, as well as other genetic events (plasmid rearrangements, large deletions). Mutations inactivating the formyl-tetrahydrofolate ligase (FtfL) gene, essential for assimilation of carbon from methanol, were identified in 8 of the 13 sequenced strains. Construction and analysis of mutants of this gene showed that inactivation of FtfL was sufficient to confer growth with CH_3Cl via the *cmu* pathway to *M. extorquens* DM4. I hypothesize that this enzyme interferes with CH_3Cl oxidation for energy production while consuming the H_4F cofactor essential for this metabolism, and that its activity opposes that of the homologous PurU enzyme required for growth with CH_3Cl via the *cmu* pathway.

Among the other mutations identified, a mutation in the *nuo* gene promoter could facilitate expression of the NADH dehydrogenase involved in NAD^+ cofactor regeneration. Other FtfL-independent adaptations have also been detected, which differ according to the genetic context of the strain receiving the *cmu* pathway genes. For example, the evolved AM1 strains all show mutations in pJM105 (in the *cmuA* gene, with cointegration of an endogenous plasmid). However, these mutations have no impact on CmuA-dependent CH_3Cl dehalogenation activity. Obtaining mutant strains by experimental evolution in the laboratory is a promising new tool for studying growth with CH_3Cl via the *cmu* pathway, particularly given the difficulties encountered to date in directed mutagenesis approaches to CH_3Cl -degrading bacteria.

Key words: bacteria, methylotrophy, dehalogenation, chloromethane, experimental evolution, metabolic adaptations