University of Nottingham
UK | CHINA | MALAYSIA

Study of Glyceraldehyde-3-Phosphate Oxidoreductases from Archaea and Bacteria

Amaury Montarnal, Ing.

Thesis submitted to the University of Nottingham for the degree of

Doctor of Philosophy

April 2022

Declaration

Unless otherwise acknowledged, the work presented in this thesis is my own. No part has been submitted for another degree in the University of Nottingham or any other institute of learning.

Amaury Montarnal

April 2022

Abstract

Carbon monoxide is a greenhouse gas mainly produced by anthropogenic sources, but the electrons liberated for CO oxidation toCO_2 have a very low potential of -520 mV. Interestingly, CO can be produced in the syngas by the gasification of renewable sources. Hence, to exploit these low potential electrons and carbon from renewable sources, the general objective of this project was to engineer *Escherichia coli* to use carbon monoxide as the sole energy and carbon source. One of the engineering points was the heterologous expression in *E. coli* of the glyceraldehyde-3-phosphate oxidoreductase (GAPOR) to function in the gluconeogenic direction to fix the low potential electrons from the carbon monoxide.

The already characterised GAPOR from the mesophilic *Methanococcus maripaludis* has been produced in a mutant *E. coli* strain optimised for crude extract enzyme assay. Three genes were deleted from the *E. coli* Rosetta-gami 2(DE3): *iscR* to improve the availability of the [4Fe-4S] cofactor in the cell; *selA* and *hypF* to diminish the benzyl viologen reduction background activity observed in the preliminary test. The glyceraldehyde-3-phosphate (G3P) oxidation activity of GAPOR in cell crude extract and affinity purified has been tested, but no activity was observed. In parallel, the enzymes from the *M. maripaludis* molybdenum cofactor biosynthesis pathway were co-expressed with GAPOR to improve the G3P-dependent activity. However, no enzymatic activity has been detected. A new LC-MS method of molybdopterin cofactor identification has been used to detect the cofactor, but so far, no cofactor was observed in the protein extract. The expression of the newly discovered bacterial homologous genes of GAPOR: GOR-SL from the mesophilic bacterium *Geosporobacter ferrireducens*, was tentatively expressed in *E. coli*, but no soluble enzyme has been detected. The expression of both proteins was tested in aldehyde ferredoxin oxidoreductase expressing bacteria *Clostridium autoethanogenum* and *Clostridium acetobutylicum* under the control of different inducible and constitutive promoters. GAPOR and GOR expression

has been observed in the soluble fraction of *C. acetobutylicum*. NoG3P-dependent activity has been detected during the enzyme assay on the purified enzyme and no impact on the cell solvent production in batch culture was either seen.

This study did not allow the production of active enzymes; nevertheless, it opened a new path of research in the archaeal molybdenum cofactor biosynthesis pathway and the pursuit of the investigation in GAPOR and GOR-SL-like protein in *C. acetobutylicum.*

Acknowledgements

I would like to start by thanking my supervisors. I will always be grateful to Prof. Philippe Soucaille for the opportunity to work on this project with him and for all the advice, guidance, supervision, and knowledge he provided me to complete this work. I would like to thank Dr Naglis Malys for all the advice and for always having left his door open when I had questions. Both of you have helped me to grow as a scientist and a person. I know that I will be able to use all the experience and knowledge I have gained throughout my career and life.

I would like to sincerely thank Dr Minyeong Yoo for her help, advice, and tips at the bench; she has kindly answered to all my random questions. I want to thank her for her friendship and all the discussion about food, restaurant, or Korean fried chicken. It was an absolute pleasure to work with her. I want to thank Dr David Tooth who made mass spectrometry analysis possible and shared great protein analysis discussions with me. I also want to thank Dr Céline Foulquier for performing the last enzyme assay in Toulouse that I could not do. I want to thank Prof. Isabelle Meynial-Salles who allowed me to go to her lab in Toulouse to perform an experiment that I could not have done in Nottingham; my gratitude also goes to all the EAD3 team members who made the four months I spent there at the end of the first lockdown absolutely memorable.

I also would like to thank my teammate Ruud Hendricks, as well as Claudio Tomi-Andrino, François Seys, Margaux Poulalier-Delavelle, Barbara Bourgade, Marcel Te Vrügt for their friendship: thanks to them, I have spent a wonderful time in and out of the lab and shared great discussions about my work. I would like to thank all my other colleagues in B27/B39 who made my time in the lab unforgettable. I also want to thank my teammates at the Nottingham Casuals RFC for giving me the opportunity to think about other things than science. Thank you too to my friends from INSA: Aline, Fabien, Joa, Manon, Marine, Melina and Sylouan, for their valuable friendship during all these years.

Finally, I want to thank my family in France for all their unconditional support during this journey. Without them, I would never have been able to become who I am now.

This PhD was funded by SBRC and the School of Life Sciences.

Table of Contents

List of Figures

List of Tables

List of Abbreviations

°C	Degree Celsius
1,3-BPG	1,3-bisphosphoglycerate
3PG	3-phosphoglycerate
aa	Amino acid
ACS	Acetyl-CoA synthase
Amp	Ampicillin
AOR	Aldehyde: ferredoxin oxidoreductase
Arc	Aerobic respiration control
ATCC	American Type Culture Collection
BCR	benzoyl-CoA reductase
Blast	Basic Local Alignment Search Tool
BSA	Bovine serum albumin
BV	Benzyl viologen
Cam	Chloramphenicol
Car	Carbenicillin
CatP	Chloramphenicol thiamphenicol marker
CBB	Calvin-Benson-Bassham
CMP	Cytosine monophosphate
CO	Carbon monoxide
CO_2	Carbon Dioxide
CODH	CO dehydrogenase
CODH/ACS	CO dehydrogenase/acetyl-CoA synthase
cPMP	cyclic pyranopterin monophosphate
CTP	Cytosine triphosphate
DHAP	dihydroxyacetone phosphate
DMSO	Dimethyl sulfoxide
DMSOR	dimethyl sulfoxide reductase
DSMZ	Deutsche Sammlung von Mikroorganismen und Zellkulturen
DTT	Dithiothreitol
EMP	Embden-Meyerhof-Parnas
EPR	electronic paramagnetic resonance
ErmB	Erythromycin marker
EU	European Union
FAD	flavin adenine dinucleotide
$FADH_2$	Reduced flavine adenine dinucleotide
FDH	Formate dehydrogenase
Fd_{ox}	Oxidised ferredoxin
Fd_{red}	Reduced ferredoxin
Fdx	*E. coli* ferredoxin
Fdx_{red} Fdx_{ox}	Reduce or oxidised from of ferredoxin
FeMoco	Iron-based molybdenum cofactor
FHL	Formate hydrogen lyase

FLP	Flippase
FNR	fumarate nitrate regulation
FOR	Formaldehyde: ferredoxin oxidoreductase
FRT	FLP recognition target
Fur	Ferric Uptake Regulator
G3P	glyceraldehyde-3-phosphate
GTP	Guanine triphosphate
H_2	dihydrogen
H_2O	Water
H_2S	hydrogen sulphide
HPLC	High pressure liquid chromatography
ICP-MS	Ion coupled plasma-mass spectrometry
IHF	integration host factor
IPCC	Intergovernmental Panel on Climate Change
IPTG	Isopropyl β-d-1-thiogalactopyranoside
Kan	Kanamycin
k_{cat}	Catalytic constant
k_d	Dissociation constant
k_m	Michaelis constant
LC-MS	Liquide chromatography-Mass spectrometry
Mo/W-bisMPT	Mo/W bis-molybdopterin
Mo-bisMGD	Mo bis molybdopterin guanin dinucleotide
Moco	Molybdenum cofactor
Mo-MCD	Mo-molybdopterin cytosine dinucleotide
MPT	molybdopterin
MS	Mass spectrometry
mV	Milli volte
MWCO	Molecular weight cut off
NADH	Nicotinamide adenine dinucleotide
NADPH	Nicotinamide adenine dinucleotide phosphate
O_3	Ozone
OD	Optic density
PBS	Phosphate Buffer Saline
PGK	phosphoglycerate kinase
ppm	Part per million
RBS	Ribosome Binding Site
RT°	Room temperature
RT-qPCR	Reverse trans
Ru5P	ribulose-1,5-biphosphate
RuBisCO	ribulose 1,5-bisphosphate carboxylase/oxygenase
SAM	S-adenyl methionine
SDS-PAGE	Sodium dodecyl sulphate-Polyacrylamide gel electrophoresis
SDT	Sodium dithionite
SO	Sulphite oxidase
TAE	Tris, acetate, EDTA buffer
Thm	Thiamphenicol

Tm	Melting temperature
US EPA	United State Environmental Protection Agency
V_{max}	Maximal reaction rate
W-bisMPT	tungsten bis molybdopterin
Wco	Tungsten cofactor
WOR	Tungsten-containing: ferredoxin oxidoreductase
WT	Wild type
XO	Xanthine oxidase

To my grandfathers Louis Dancy and Philippe Montarnal

To my grandmother Rose Marie Dancy

To my parents Françoise and Jean-Michel Montarnal

To my sister Bérénice Montarnal

« En essayant continuellement on finit par réussir.

Donc : plus ça rate, plus on a de chance que ça marche »

Jacques Rouxel

So far away, we wait for the day

For the lives all so wasted and gone

We feel the pain of a lifetime lost in a thousand days

Through the fire and the flames, we carry on

Through the Fire and the Flames

Dragon Force

Chapter I:

Introduction

1. Introduction

In its recent report, the Intergovernmental Panel on Climate Change (IPCC) (IPCC, 2021a) presented that the Global surface temperature in 2020 increased by 1.09°C since 1850 (IPCC, 2021b). It was stated in the report that it was very likely that greenhouse gases caused by human activities were responsible for this increase (IPCC, 2021b). Atmospheric carbon monoxide impacts the three main greenhouse gases: ozone (O_3), methane and carbon dioxide (CO_2). Indeed, carbon monoxide reacts with hydroxyl radicals influencing the tropospheric oxidation potential and methane. Moreover, CO is a source of O_3 and CO_2 (Shindell et al., 2006). CO is a tasteless, odourless, and colourless gas that is toxic for many organisms due to its propensity to bind metalloenzymes such as haemoglobin or myoglobin in humans. The CO is naturally present in the atmosphere at a concentration varying between 40 ppb to 120 ppb on average, with a lifetime in the atmosphere of 1 to 4 months (IPCC, 2021a; Szopa et al., 2021). It is a product of the incomplete combustion of organic compounds in an oxygen-limited atmosphere. The principal sources of atmospheric CO in the United States of America (*ca.* 66%) are industrial and human activity. About 66% of these emission originated from transport, mainly automobile (Figure I-1) (US EPA, 2021b). The remaining part is naturally produced by volcanic activity and forest fires (Diender, Stams, & Sousa, 2015). To limit the increase of temperature, it is necessary to reduce the production of greenhouse gases (IPCC, 2021b). One of the solutions is to use more renewable raw materials to produce energy or chemicals (Daniell, Köpke, & Simpson, 2012; Stern & Stern, 2007).

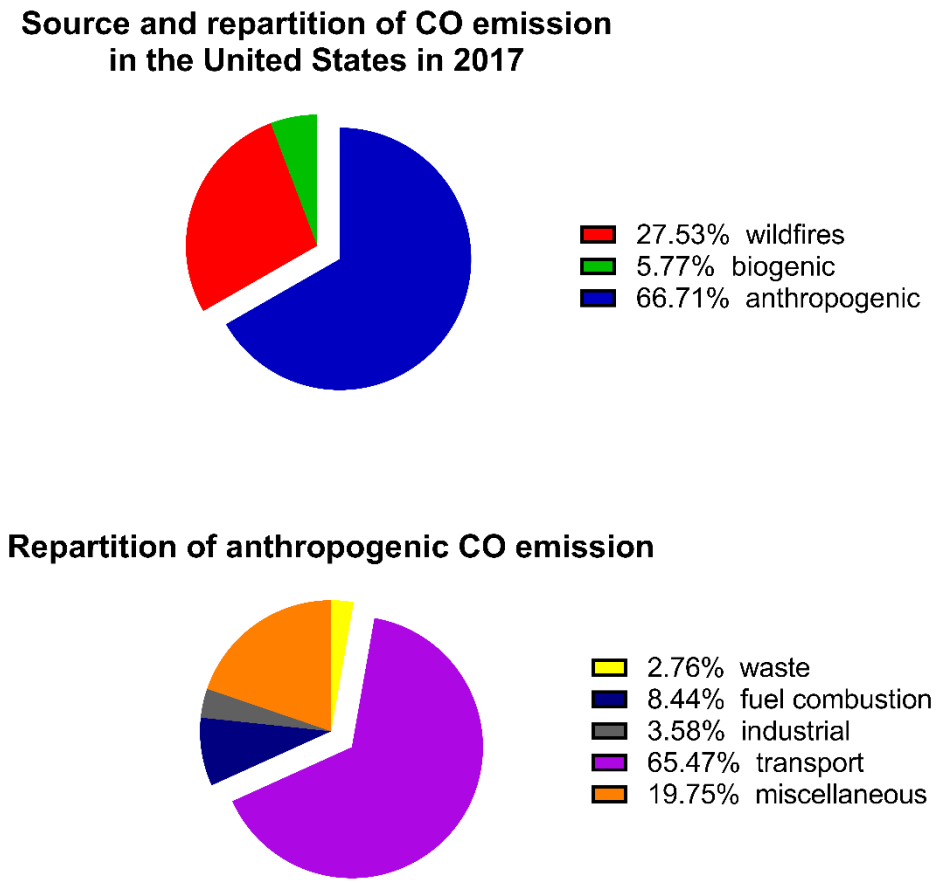


Figure I-1 Source of atmospheric CO in the United States according to US Environmental Protection Agency 2017 report updated in January 2021 (US EPA, 2021b). Ca. 66% of the CO emissions come from human activities, and about 66% of the human emissions come from transport. The data were collected from the U.S. EPA website (US EPA, 2021a).

CO can be co-produced with dihydrogen (H_2) and other minor components in the synthesis gas (syngas). The syngas, composed mainly of CO, H_2 and CO_2 with different ratios depending on the production process (Molino, Chianese, & Musmarra, 2016), is produced by gasification of carbon-containing material such as biomass. Biomass is defined by 2009/28/EC European directives as: "*The biodegradable fraction of products, waste and residues from biological origin from agriculture, including vegetal and animal substances, from forestry and related industries, fisheries and aquaculture, as well as the biodegradable fraction of waste, including industrial and municipal waste of biological origin*"(EU, 2009) and still according to the European Union: "*Using biomass as a fuel is deemed carbon neutral as carbon was trapped from the atmosphere during the biomass life cycle (it grows)*"(Eurostat, 2019).

The gasification is completed in three steps: 1) drying raw materials, 2) heating between 300°C and 500°C without oxidising agents to achieve pyrolysis and produce gases, vaporised tars, oil, and solid char. Step three is the gasification of pyrolysis products with an oxidising agent (Bridgwater, 2003). Syngas can be produced from renewable resources such as lignocellulosic biomass, municipal waste solid or liquid (Diender et al., 2015; Heijstra, Leang, & Juminaga, 2017; Luo, Zhou, & Yi, 2012). In Europe currently, several power plants use biomass gasification from different sources to produce syngas to produce energy (heat and power) by combustion of syngas, biomethane, liquid-biofuel or hydrogen (Molino et al., 2016). The production of biofuels could be either by thermochemical or biochemical conversion (Sun, Atiyeh, Huhnke, & Tanner, 2019). The production of syngas allows its conversion to valuable chemical components by microorganisms from materials that are difficult to metabolise, such as straw or wood (Henstra, Sipma, Rinzema, & Stams, 2007). Furthermore, patents from universities and industrial companies (LanzaTech, *i.e. (Koepke & Chen, 2018)*, Genomatica,…) to produce ethanol, isoprene, 3-hydroxypropionate (3-HP) from syngas from gas-fermentation, as well as improved fermenters for industrial purpose, were reported by Sun (Sun et al., 2019).

CO is naturally used as a substrate by anaerobic mesophilic or thermophilic acetogenic, methanogenic and hydrogenogenic archaea and bacteria to produce acids (acetate, butyrate, formate), solvent (ethanol and butanol), methane, hydrogen sulphide (H_2S) and H_2 (Henstra et al., 2007). CO was also observed to be used by aerobic carboxydotrophic and carboxydovore bacteria, with some of them using CO as sole carbon and energy source. They are fixing the carbon through the Calvin-Benson-Bassham (CBB) cycle and using oxygen, sulphate or quinone as electron acceptor (Diender et al., 2015; G. M. King & Weber, 2007; Meyer, Jacobitz, & Krüger, 1986; Oelgeschläger & Rother, 2008). Some of these aerobic CO-oxidisers could not use CO above a concentration of 1000 ppm. Interestingly, it was estimated that these bacteria, found in a wide variety of biomes, were consuming about 20 % of the annual CO emission (G.

M. King & Weber, 2007). Aerobic and anaerobic bacteria with CO metabolism use carbon monoxide dehydrogenase to oxidise the CO to CO_2 (Diender et al., 2015). This reaction possesses a very low redox potential (-520 mV), naturally used by hydrogenogenic microorganisms to produce H_2 from protons (Diender et al., 2015). However, only a few organisms can use CO as sole energy and carbon source (Oelgeschläger & Rother, 2008), such as *Eubacterium limosum* (Genthner & Bryant, 1982) and *Clostridium ljungdahlii* (Tanner, Miller, & Yang, 1993). A reason might be the sensitivity of hydrogenase toward CO (Bertsch & Muller, 2015).

The Calvin-Benson-Bassham (CBB) cycle (Figure I-2) is the principal carbon dioxide (CO_2) fixing pathway in the Biosphere. The ribulose-1,5-bisphosphate carboxylase/oxygenase (RuBisCO) fixes three molecules of CO_2 by catalysing the synthesis of six molecules of 3-phosphoglycerate (3PG) with three molecules of ribulose-1,5-biphosphate (Ru5P). Then, the phosphoglycerate kinase (PGK) catalyses the adenosine triphosphate (ATP) dependent synthesis of 1,3-bisphosphoglycerate (1,3-BPG) followed by the nicotinamide adenine dinucleotide phosphate (NADPH) dependent production of six glyceraldehyde-3-phosphate (G3P). To save the consumption of these 6 ATP and 6 NADPH, we aim to replace the glyceraldehyde-3-phospahte dehydrogenase (GAPDH) and the PGK step by a single one driven by the glyceraldehyde-3-phosphate: ferredoxin oxidoreductase (GAPOR) (*cf* Figure I-21) (N. Antonovsky, S. Gleizer, & R. Milo, 2017). The other principal CO_2 fixation pathway is the Wood-Ljungdahl involving the CO dehydrogenase/acetyl-CoA synthase (CODH/ACS) (Jeoung, Martins, & Dobbek, 2019; Ragsdale, 2008).

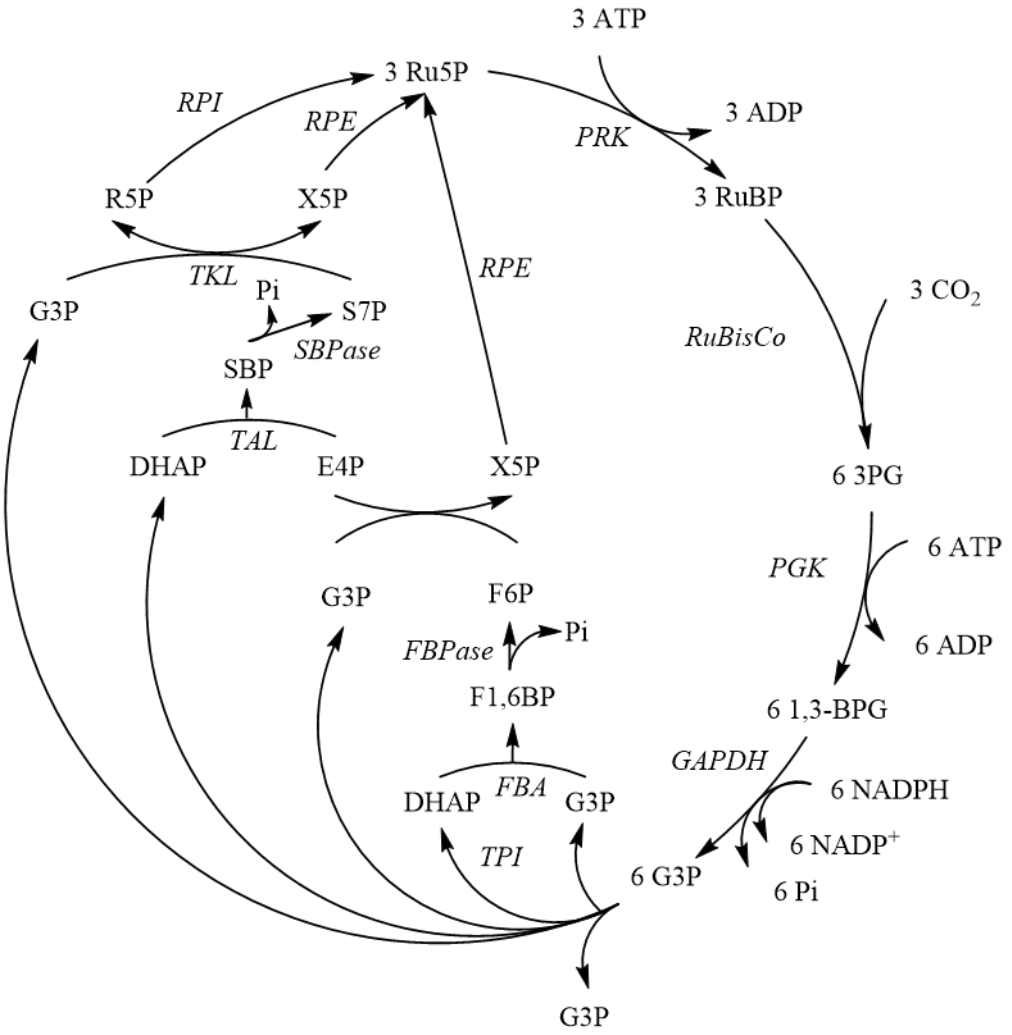


Figure I-2 Scheme of Calvin-Benson-Bassham cycle. ATP: adenosine triphosphate, ADP: adenosine diphosphate, Pi: monophosphate, NADPH: reduced nicotinamide adenine dinucleotide phosphate, NADP: nicotinamide adenine dinucleotide phosphate, 3PG: 3-phosphoglycerate, 1,3BPG: 1,3-bisphosphoglycerate, G3P: glyceraldehyde-3-phosphate, DHAP: dihydroxyacetone phosphate, F1,6BP: fructose-1,6-bisphosphate, F6P: fructose-6-phosphate, E4P: erythrose-4-phosphate, X5P: xylulose-5-phosphate, R5P: ribose-5-phosphate, Ru5P: ribulose-5-phosphate, RuBP: ribulose-1,5-bisphosphate, SBP: sedoheptulose-1,7-bisphosphate, S7P: sedoheptulose-7-phosphate; RuBisCo: Ribulose-1,5-bisphosphate carboxylase-oxygenase, PGK: phosphoglycerate kinase, GAPDH: glyceraldehyde-3-phosphate dehydrogenase, TPI: triosephosphate isomerase, FBA: fructose-bisphosphate aldolase, SBPase: sedoheptulose-1,7-bisphosphatase, FBPase: fructose bisphosphatase, TKL: transketolase, RPI: ribose-5-phosphate isomerase, RPE: Ribulose-5-phosphate epimerase, PRK: phosphoribulokinase.(N. Antonovsky et al., 2017)

Carbon monoxide oxidation has a very low redox potential. Other oxidoreductase enzymes in the synthetic Calvin cycle could use this redox potential to catalyse a reaction in the reverse way to produce metabolites or to save energy in the cells. This energy could be used to produce heterologous metabolite. The development of a novel *Escherichia coli* able to grow on CO as sole carbon and energy source would build a platform for the development or the production of chemicals from one of the best-known bacteria.

2. CODH

To to use the carbon and the electrons from the CO oxidation in our hypothetical system. In prokaryotes, CO and CO_2 can be interconverted by one enzyme, the carbon monoxide dehydrogenase (CODH). This is enzyme sable to catalyse the water-gas shift which is the interconversion of CO to CO_2 with water (H_2O) and H_2 as describe by (1) (Oelgeschläger & Rother, 2008).

$$CO + H_2O = CO_2 + 2H^+ + 2e^- \qquad (E^0 = -520\ \mathrm{mV}) \qquad (1)$$

The CODH is a metallo-enzyme. Two classes of CODH were observed: Cu,Mo-CODH are found in aerobic organisms, and Ni,Fe-CODH are found in anaerobic archaea and bacteria (Jeoung et al., 2019).

The most studied, O_2 tolerant, Cu,Mo-CODH from *Oligotropha carboxydovorans* (Meyer & Schlegel, 1978) is a $(\alpha\beta\gamma)_2$ hexamer which contains a Mo-molybdopterin cytosine dinucleotide (Mo-MCD) with Cu^I thiolate bounded to the Mo centre. Two [2Fe-2S] clusters and a flavin adenine dinucleotide (FAD) form the remaining electron transfer pathway. It carries out CO oxidation to CO_2, oxidising H_2 and transferring the electron to membrane-bound quinones. (Jeoung et al., 2019; Wilcoxen & Hille, 2013).

The oxygen-sensitive Ni,Fe_CODH is a more diverse group of enzymes that can be divided into four classes (Lindahl, 2002). Classes I and II correspond to acetyl-CoA decarbonylase/synthase. These are 2 MDa pentameric ($\alpha\beta\gamma\Delta\varepsilon$) bifunctional complex with CODH and acetyl-CoA synthase (ACS) and are mainly found in methanogenic archaea. Class III enzymes are two independent enzymes $\alpha_2\beta_2$ CODH/ACS which form a complex with the heterodimer corrinoid iron-CoFeSP carrying the methyl to the ACS subunit for the synthesis of acetyl-CoA (Fast & Papoutsakis, 2018). The Class III enzymes are mainly found in acetogenic bacteria and are used during growth on CO_2 and H_2 by the Wood-Ljungdahl pathway using ferredoxin as an electron donor (Fast & Papoutsakis, 2018; Lindahl, 2002; Ragsdale & Pierce, 2008). The class IV enzymes are homodimer CODH found in CO utilising bacteria as both carbon and energy source. These enzymes are monofunctional (Jeoung et al., 2019). The CODH

homodimer sub-complex contains five iron-sulfur clusters: one [4Fe-4S] is at the interface of the two subunits, another [4Fe-4S] in the middle of each subunit and a last one [Ni-4Fe-4S] cluster is in the active site (Jeoung et al., 2019).

As said, CODH and CODH/ACS are part of the CO fixation pathway in acetogenic bacteria primarily by the Wood-Ljungdahl pathway (WLP) to produce acetyl-CoA (Figure I-3). The pyruvate ferredoxin oxidoreductase (PFOR) can use acetyl-CoA with reduced ferredoxin and CO_2 to form pyruvate. This pyruvate can be used for gluconeogenesis or to produce acids and solvents (Köpke et al., 2010a; Liew et al., 2017; Norman, Millat, Winzer, Minton, & Hodgman, 2018; Zhu et al., 2020). The CODH uses ferredoxin or ferredoxin-like proteins as electron acceptors or donors (Oelgeschläger & Rother, 2008). The redox potential of the couple CO/CO_2 is very low (-520 mV). Indeed, the redox potential of the reduced ferredoxin is very low (*Cf* following section)(Carlson & Papoutsakis, 2017). The reduction of CO_2 requires electrons from ferredoxin and $2H^+$. For the reverse reaction, CODH needs ferredoxin to collect the low-potential electrons released by the reaction. Liew and co-worker reported that ferredoxin reduced by CODH could contribute to other enzymatic reactions, such as the production acetaldehyde in *Clostridium autoethanogenum* by the aldehyde: ferredoxin oxidoreductase (AOR) (*cf* section 5.1.1) (Liew et al., 2017).

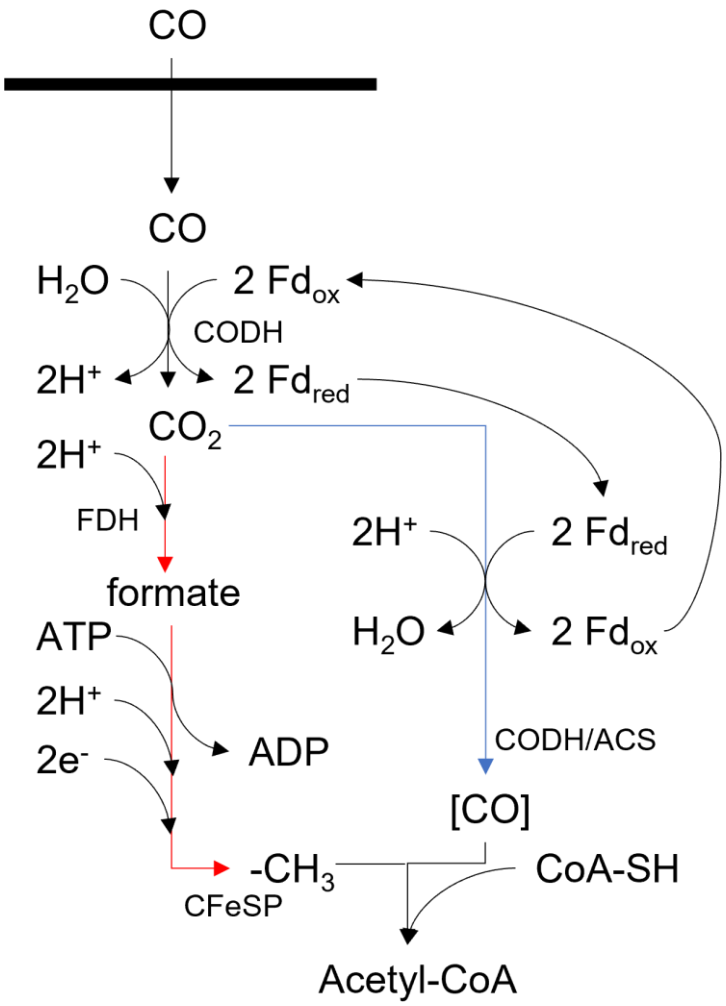


Figure I-3 Wood-Ljungdahl pathway for CO fixation. The pathway was adapted from (Fast & Papoutsakis, 2018; Köpke et al., 2010a; Norman et al., 2018; Ragsdale & Pierce, 2008). $2e^-$ represent a reductive power unknown in C. autoethanogenum and C. ljungdahlii it was supposed to be either ferredoxin or NADH/ferredoxin electron bifurcation. In red, the methyl branch and in blue the carbonyl branch of the acetyl-CoA synthase. FDH: formate dehydrogenase, CODH: carbon monoxide dehydrogenase, ACS: acetyl-CoA synthase, Fd_{red}: reduce ferredoxin, Fd_{ox}: oxidised ferredoxin, CoA-SH: Coenzyme A, CFeSP: corrinoid iron-sulfur protein, ATP: adenosine triphosphate, ADP: adenosine diphosphate.

Carlson & Papoutsakis (2017) demonstrated that the CODH subunit from *Clostridium carboxidovorans* expressed in *Clostridium acetobutilicum*, which did not fix CO_2, was active with or without the co-expression of the ACS subunit. *C. acetobutilicum* expressing the CODH subunit can oxidise CO up to 5% (vol/vol) CO. Moreover, they showed that CODH needs a protein to insert the Ni to activate the enzyme, except when additional Ni is added to the medium. (Carlson & Papoutsakis, 2017).

3. Ferredoxin

To transport the reduced electrons produced by the CO oxidation, ferredoxin is required. The ferredoxin is one of the smallest metalloproteins, to our knowledge. Its sequence is between 50 to 120 amino acids for a molecular weight between 5 to 15 kDa. Ferredoxins contains one or two iron-sulfur centres ([2Fe-2S], [3Fe-4S], [4Fe-4S], 2[4Fe-4S], [7Fe-8S]) (Figure I-4) (Nzuza et al., 2021). These proteins have been found in archaea, bacteria, and eukaryotes to perform the transport of electrons from different oxidases, which produce low potential electrons, to enzymes requiring them (Bruschi & Guerlesquin, 1988; Buckel & Thauer, 2013; Campbell, Bennett, & Silberg, 2019). It was proposed that [4Fe-4S] ferredoxin was the first to appear in life and evolved toward [2Fe-2S] cluster (Sousa et al., 2013). In fact, It was found that [4Fe-4S] cofactor is more sensitive to O_2 and is mostly found in anaerobic organisms, while [2Fe-2S] are mostly found in aerobic organisms and eukaryotes (Imlay, 2006). Moreover, they have the property to carry very high or very low redox potential electrons (300 mV, - 660 mV) (Atkinson, Campbell, Bennett, & Silberg, 2016; Buckel & Thauer, 2013; Guerrini et al., 2008).

The iron-sulfur clusters are coordinated by four cysteine residues binding to the Fe atom. It was observed that each type of cofactor has a specific binding motif (Table I-1).

Table I-1 Iron-sulfur cluster types and consensus binding motif and their literature low-potential (Atkinson et al., 2016; Nzuza et al., 2021)

Type of cofactor	**motif**	**E^0 (mV)**
[2Fe-2S]	$CX_{5/4}CX_{2/n}CX_{n/3}C$	-455 to -152
[3Fe-4S]	CX_5CX_nCP	-340 to -31[a]
[4Fe-4S]	$CX_{2/5}CX_{2/3}CX_nC(P)$	-453 to -280
2[4Fe-4S]	$CX_2CX_2CX_3CX_nCX_2CX_2CX_3(CX_3)C$	-660 to -390
[7Fe-8S] ([3Fe-4S]/ [4Fe-4S])	$CX_7CX_3CPX_nCX_2CX_2CX_3CP$	-647 to -280

[a] (Brereton, Verhagen, Zhou, & Adams, 1998; Child et al., 2018)

The *E. coli* ferredoxin (Fdx) was discovered by Knoell and Knappe (Knoell & Knappe, 1974). This ferredoxin is encoded by the *fdx* gene inside the *isc* operon, which is responsible for the housekeeping production of protein for the iron-sulfur biogenesis. This ferredoxin contains a single [2Fe-2S] cluster (Figure I-4) with a redox potential of about -380 mV (Knoell & Knappe, 1974). The understanding of the metabolic role of Fdx in *E. coli* and its place in the *isc* operon took time. Kim *et al*. found that it provided an electron to the IscS during the conversion of a cysteine to an alanine (Kim, Frederick, Reinen, Troupis, & Markley, 2013). Moreover, in the pathogenic *E. coli* K1 strain, the Fdx was found to be involved in the secretion of cytotoxic necrotizing factor 1 responsible of the invasion by the bacterium of the human brain cells (H. Yu & Kim, 2010). In *P. furiosus*, the ferredoxin involved in the interaction with the formaldehyde ferredoxin oxidoreductase was a single [4Fe-4S] cluster ferredoxin (Zhou & Adams, 1997) with a low potential of -400 mV and a high potential of +350 mV. In contrast, *M. maripaludis* possesses plenty of ferredoxins or polyferredoxins with one or more [4Fe-4S] centre (Hendrickson, Kaul, Zhou, Bovee, Chapman, Chung, De Macario, et al., 2004).

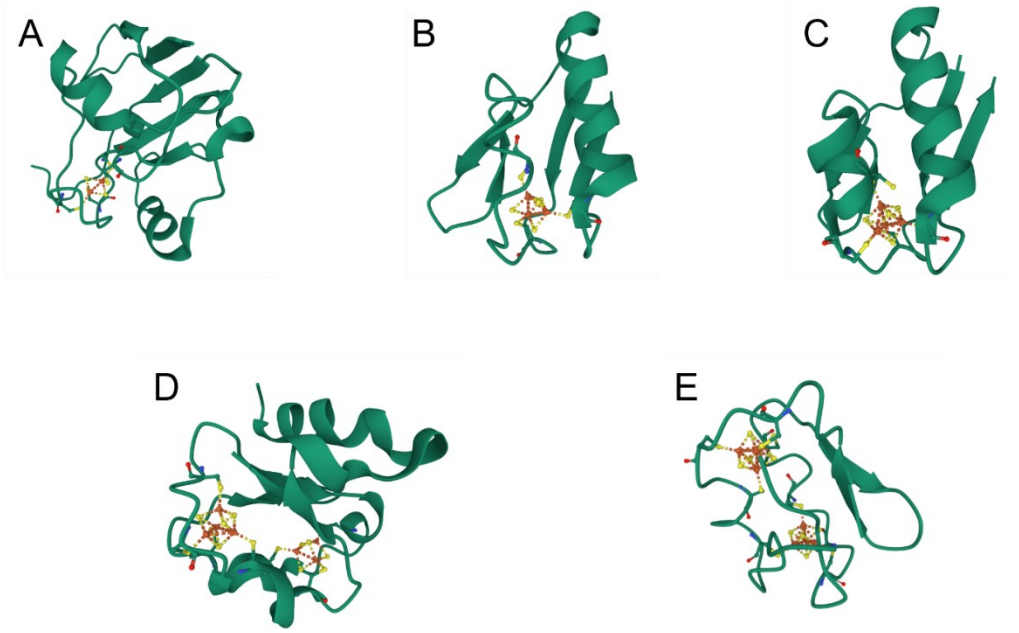


Figure I-4 Crystal stuctures of different types of ferredoxin found in prokaryotes. A) E. coli [2Fe-2S] ferredoxin 1I7H (Kakuta, Horio, Takahashi, & Fukuyama, 2001) B) and C) P. furiosus [3Fe-4S] 3PNI and [4Fe-4S] 2Z8Q variant of D14C ferredoxin (Løvgreen, Martic, Windahl, Christensen, & Harris, 2011); D) Thermus thermophilus [7Fe-8S] ferredoxin 1H98 (Macedo-Ribeiro et al., 2001) E) Clostridium pasteurianum 2[4Fe-4S] ferredoxin 1CLF (Bertini et al., 1995).

4. Iron-sulfur cluster main synthesis pathways and regulation in *E. coli*

In *E. coli,* the Fe-S cluster's assembly and incorporation proteins are encoded by the *Iron-sulfur cluster (isc)* operon under non-stress conditions(D. C. Johnson, Dean, Smith, & Johnson, 2005). It is the primary formation and incorporation system. The operon is composed of five genes, *iscRSUAX,* incorporated in a more extensive system with *hscBA fdx* (Figure I-5) and *cyaY* (Akhtar & Jones, 2008; Barras, Loiseau, & Py, 2005).

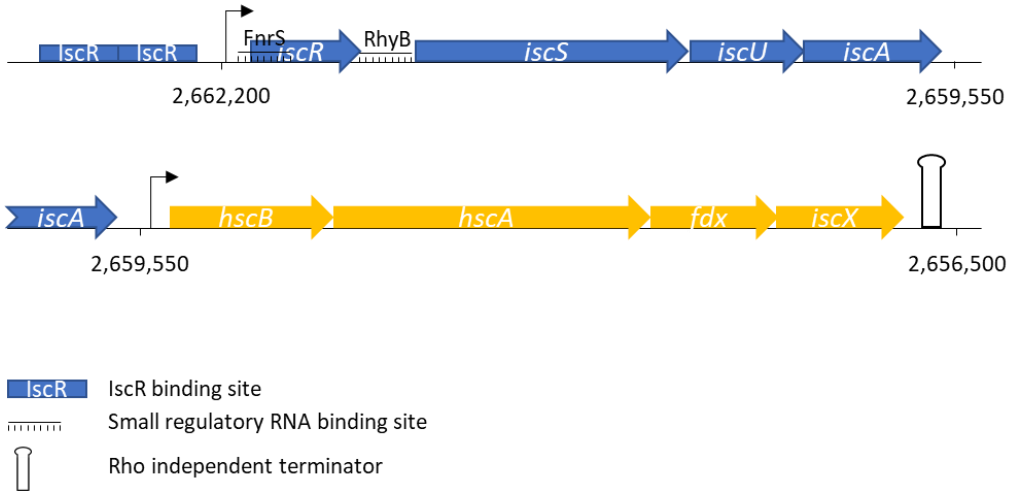


Figure I-5 *iscRSUAX hscBA fdx* operon in *E. coli* with transcription starts, regulators binding site and terminator.

iscS codes for a cysteine desulfurase (IscS) which provides sulfur for the iron-sulfur cluster (Schwartz et al., 2001). It also provides sulfur for the sulfuration of MoaD in the molybdenum cofactor biosynthesis pathway (A. Magalon & Mendel, 2015). The second gene of this cluster (*iscU*) expresses a homodimer (IscU) which can form a heterotetrameric complex with two IscS. IscU is a scaffold for the biosynthesis of labile $[2Fe\text{-}2S]^{2+}$ and $[4Fe\text{-}4S]^{2+}$ clusters (Markley et al., 2013). This biosynthesis of the $[4Fe\text{-}4S]^{2+}$ cluster, formed sequentially from two $[2Fe\text{-}2S]^{2+}$, is directed by the sulfur transfer from IscS (Agar et al., 2000). The formation mechanism of the [4Fe-4S] cluster (Figure I-6) has been resolved with native MS (Lin, McCabe, Russell, & Barondeau, 2020). They found that the cofactor formation was performed sequentially when IscS and IscU were in complex. The formation starts with the incorporation of an iron atom and the supply of an electron from a ferredoxin, followed by the desulfuration of an L-Cys to an L-Ala. The same sequence is repeated until the formation of [2Fe-2S]. Once the cofactor is formed, IscU is liberated from the complex (Lin et al., 2020). The IscU is then recruiting an apo-IscU to form an asymmetric dimer which seems to be preferentially used by the HscA and HscB chaperone protein to transfer the cofactor to the targeted protein (Chandramouli & Johnson, 2006). On the other side, if another [2Fe-2S]-IscU is recruited, the transfer of two electrons from Fdx for a reducing coupling of the two [2Fe-2S] forms the [4Fe-4S] cluster (Chandramouli et al.,

2007). The cofactor is then transferred to the chaperone or to target proteins (Cai, Frederick, & Markley, 2020; Unciuleac et al., 2007). The last gene of this operon is *iscA*. It has been shown by Wang et al. that under aerobic condition IscA may act as an iron chaperone for the iron-sulfur cluster biosynthesis (Wang et al., 2010). Moreover, IscA could also play the role of chaperone protein inserting of the Fe-S cluster in the apo-protein (Barras et al., 2005). The two other proteins produced from this cluster of genes are the chaperone proteins HscB and HscA, which have a size of 20 and 66 kDa. These two chaperone proteins are essential in the iron-sulfur protein biosynthesis by interacting with the scaffold protein IscU and by improving the transfer of the iron-sulfur cluster to the apo-protein, which is ATP-dependent (Chandramouli & Johnson, 2006). Under oxidative stress and low iron conditions, *sufABCDSE* take over the production of the iron-sulfur cluster (Py, Moreau, & Barras, 2011). This system works with two subcomplexes: $SufBC_2D$ and SufSE. It was found that the consumption of an ATP started the synthesis by the SufC ATPase subunit bringing conformational change to the complex and exposing the binding site to the surface. The mobilisation of SufSE follows it to interact with the complex. SufS catalyses the desulfuration of L-Cys to persulfurate Cys364. SufE Cys51 allows transpersulfuration of this cysteine. The persulphide is then transferred to SufB. The S atom is then sequentially transferred to the active site. It is supposed that reduced flavine adenine dinucleotide ($FADH_2$) is used to reduce iron and form the [2Fe-2S] cofactor. When the cofactor is formed, SufSE is released, allowing the transfer of the cofactor to SufA, which can then maturate the cofactor to [4Fe-4S] (Pérard & Ollagnier de Choudens, 2018). Another system, *nif,* is only responsible for incorporating the Mo-Fe-S cluster in the nitrogenase (Tokumoto, Kitamura, Fukuyama, & Takahashi, 2004).

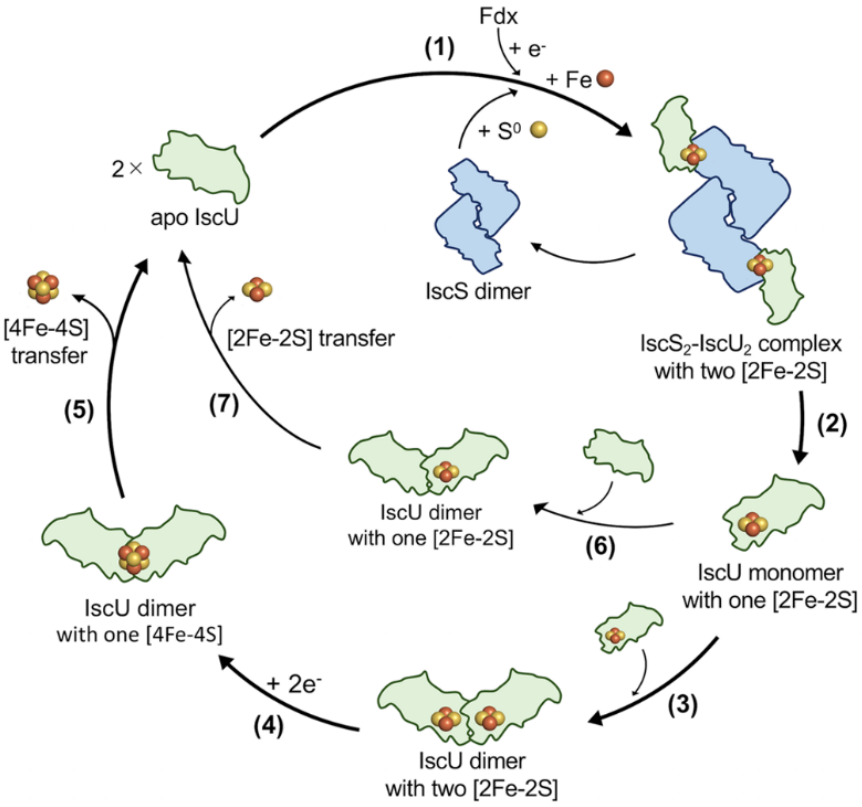


Figure I-6 Scheme of the [4Fe-4S] and [2Fe-2S] formation in E. coli. (Kunichika, Nakamura, Fujishiro, & Takahashi, 2021), Fdx: Ferredoxin, IscS: iron sulfur cluster cysteine desulfurase, IscU: iron sulfur cluster scaffold protein.

The expression regulation of the *isc* operon and the switch to the *suf* operon expression is directed by IscR. In addition to these two operons, IscR is involved in the regulation of 40 genes in *E. coli* (Giel, Rodionov, Liu, Blattner, & Kiley, 2006). IscR is a [2Fe-2S]-cluster containing regulator. In normal conditions, IscR is co-transcribed with IscSUA and maturated to form the Holo-IscR. Apo-IscR competes with the other apo-cluster protein. In its Holo-form, IscR binds the DNA in the P_{IscR} promoter and represses the expression of the *isc* operon (Schwartz et al., 2001). Under anoxic conditions, the demand for the Fe-S cluster is lower as the iron-sulfur-containing enzymes are less damaged by the presence of dioxygen (Giel et al., 2013). Thus, there is less competition for apo-IscR to incorporate the Heme conducting to the inhibition of the *isc* operon. Inversely, under oxic conditions, there this a higher number of clusters containing protein, increasing the demand for the iron-sulfur cluster. Hence, less holo-IscR are matured lifting the expression inhibition to cope with the needs (Giel et al., 2013). The switch from the *isc* operon to the *suf* operon is

driven by the low iron conditions response mechanism. Indeed, the oxidative stress response regulator OxyR induces the expression of the Ferric Uptake Regulator (Fur) (M. Zheng, Doan, Schneider, & Storz, 1999). Both Fur and OxyR combined with the Integration host factor (IHF), upregulate the *suf* operon (Lee, Yeo, & Roe, 2008). Furthermore, Giel et al. showed that apo-IscR is a transcriptional activator of *suf* (Giel et al., 2006). To increase the amount of apo-IscR in the cell and decrease the activity of the *isc* system, the small regulatory RNA *RhyB* is produced under Fur regulation. This sRNA binds the *isc* mRNA in the intergenic sequence between *iscR* and *iscS* and promotes RNA degradation by the degradosome of *iscSUA* (Desnoyers, Morissette, Prévost, & Massé, 2009). Thus, oxidative stress and iron starvation lead to the switch from the *isc* to the *suf* genes expression(Py & Barras, 2010).

Another layer of regulation of the *isc* operon is given by IscX and CyaY. It was found that *in vivo,* they had a positive cumulative effect on cluster maturation (Roche, Huguenot, Barras, & Py, 2015). IscX expression is under IscR regulation, while CyaY expression is not. It was found that IscX and CyaY are competing for the same binding site on IscS (Kim, Bothe, Frederick, Holder, & Markley, 2014). However, it was recently found by Adinolfi et al. that IscX has a second binding site (Adinolfi et al., 2018). They also suggested that the CyaY has an inhibitory effect on IscS by blocking the catalytic loop conducting to inhibitthe cluster formation. Under low iron concentration, CyaY has a lower affinity with IscS than IscXpreventing the binding of CyaY to the complex and allowing a normal cluster synthesis. However, CyaY affinity for the IscS increases along with the iron concentration. Thus, at high iron concentrations, CyaY has a higher affinity for IscS than IscX, starting to inhibit the cluster synthesis. Thus, the cell regulates the synthesis at enzyme level by inhibition counter-inhibition based on iron concentration using the CyaY as an Iron sensor. Regulation of the expression is applied to the two genes *iscX* and *cyaY* as both proteins have sub-stoichiometric concentrations compared to IscS under normal conditions. This concentration can go up by 6 to 8 fold under different conditions (Adinolfi et al., 2018).

5. GAPOR an AOR with a specific substrate

The last step of the theoretical pathway involves the utilisation of the glyceraldehyde-3-phosphate: ferredoxin oxidoreductase (GAPOR). The enzyme was supposed to use the carbon and the electrons from the CO oxidation. The GAPOR is part of the large enzyme superfamily.

5.1. Aldehyde: ferredoxin oxidoreductase superfamily

The Aldehyde: ferredoxin oxidoreductase (AOR) family is composed of enzyme driving the reversible ferredoxin-dependent reaction of oxidoreduction of an aldehyde to the corresponding carboxylate following the reaction (2) (Roy, Dhawan, Johnson, Rees, & Adams, 2011) (Fd: oxidised and reduced form of ferredoxin with only one [4Fe-4S] cluster).

$$RCOH + H_2O + 2Fd_{ox} = RCOOH + 2\,H^+ + 2Fd_{red} \quad E'_0 \approx -560\,mV \ (2)$$

The enzymes of this family are characterised by their tungsten (W) atom coordinated by the dithiolene of two molybdopterins (MPT) (Figure I-7) forming a tungsten bis molybdopterin W-bisMPT and a [4Fe-4S] cluster per subunit coordinated by four cysteines (Chan, Mukund, Kletzin, Adams, & Rees 1995).

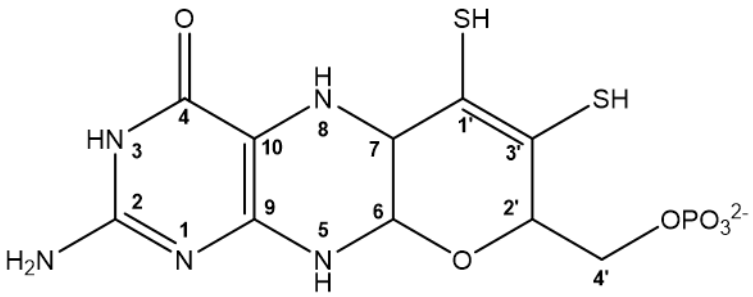


Figure I-7 Tautomeric form of molybdopterin moieties inserted in AOR family enzymes. The tungsten is coordinated by the thiol group in 1' and 3' (Roy et al., 2011).

After phylogenetic analysis of more than 4,000 protein sequences from the AOR family, Adams and co-workers have shown that the superfamily was composed of 92 distinct clades (Schut et al., 2021; Scott et al., 2019). Among them, only eight had one of their enzymes purified: the Aldehyde: ferredoxin oxidoreductase (AOR) (EC 1.2.7.5) (Arndt et al., 2019; Roy, Menon, & Adams,

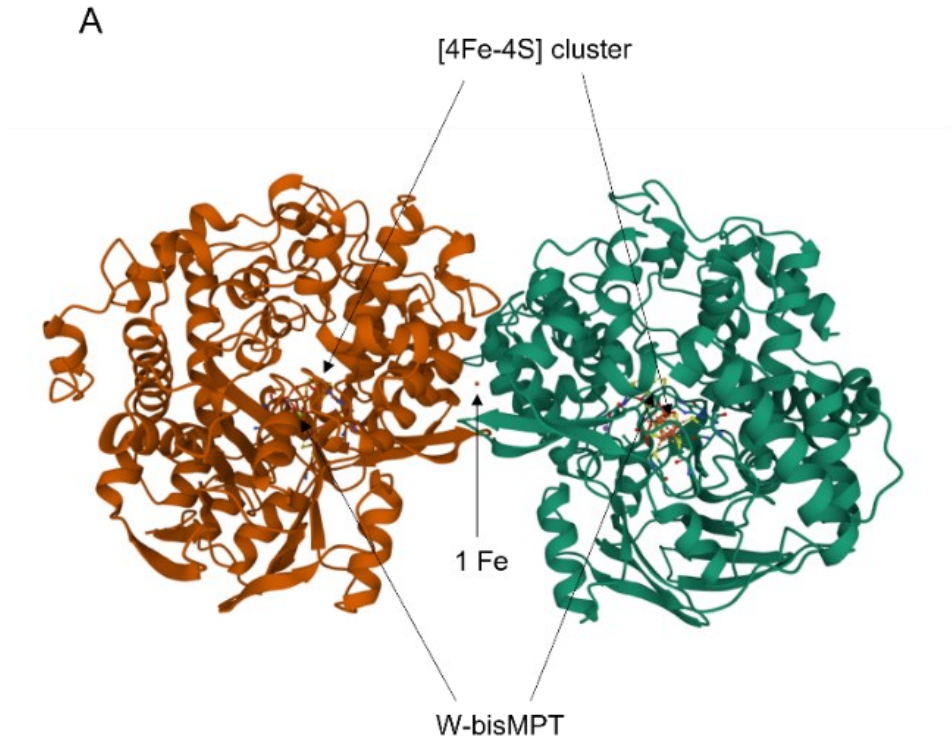
2001), Formaldehyde: ferredoxin oxidoreductase (FOR) (EC 1.2.7.B2) (Roy et al., 2001; R. Roy et al., 1999) and the Glyceraldehyde-3-phosphate: ferredoxin oxidoreductase (GAPOR) (EC 1.2.7.6) (Mukund & Adams, 1995; Park, Mizutani, & Jones, 2007), GOR (bacterial GAPOR) (Scott et al., 2019) and tungsten-containing: ferredoxin oxidoreductase WOR5 (EC 1.2.7.B3) (L. E. Bevers, Bol, Hagedoorn, & Hagen, 2005). Another Clade of enzymes corresponds to the BamB, the active site subunit of class II benzoyl-CoA reductase (BCR) (Weinert et al., 2015). Three other clades of characterised enzymes with unknown activities were also identified WOR4 (Roy & Adams Michael, 2002), AOR1 and YdhV (Reschke et al., 2019), with the latter being the only clades only composed of molybdenum-containing AOR (Seelmann, Willistein, Heider, & Boll, 2020). Moreover, their physiological substrates are only known for AOR, GAPOR, GOR and BCR.

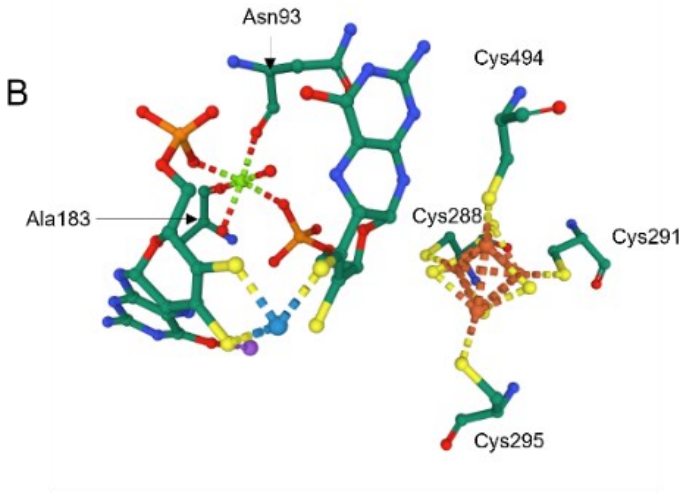
The enzymes from this "superfamily" were mainly characterised in hyperthermophiles archaeon such as *Pyrococcus furiosus* (Mukund & Adams, 1991; Roy et al., 2001)*, Pyrobaculum aerophilum* (Hagedoorn et al., 2005), *Thermococcus litoralis* (Heider, Ma, & Adams, 1995; J. L. Johnson, Rajagopalan, Mukund, & Adams, 1993). Enzymes from this super family were also characterised in mesophilic archaea like *Methanococcus maripaludis* (Park et al., 2007) and bacteria from the Clostridial class like *Eubacterium acidaminophilum* (Rauh, Graentzdoerffer, Granderath, Andreesen, & Pich, 2004) or *Clostridium formicoaceticum* (C. Huber, Caldeira, Jongejan, & Simon, 1994) and other bacteria such as *Aromatoleum aromaticum* EbN1 (Arndt et al., 2019) and *E. coli* with the recently characterised YdhV with yet unknown activity (Reschke et al., 2019). So far, only the crystal structure of AOR (PDB: 1AOR (Chan et al., 1995)), FOR (PDB: 1B25 and 1B4N (Hu, Faham, Roy, Adams, & Rees, 1999)) and very recently WOR5 (PDB: 6X1O and 6X6U not yet published) from *P. furiosus* have been solved. The crystallographic study of AOR from *P. furiosus* helped to confirm the tricyclic structure of the pterin cofactor (Figure I-8 C) and the W/Mo coordination mode (Chan et al., 1995). The AORs were found to be very abundant proteins in this microorganism.

It was found that the W cofactor could be replaced by Molybdenum (Mo) in the characterised enzymes of *P. furiosus,* but the enzymes were then inactivated (Mukund & Adams, 1996; Sevcenco et al., 2010). Only a few characterised enzymes of this family were found to be Mo-dependent: one AOR from *Proteus vulgaris* (Trautwein, Krauss, Lottspeich, & Simon, 1994), the GAPOR from *Methanococcus maripaludis* (Park et al., 2007) and one of the Mo-AOR from *C. formicoaceticum* (C. Huber et al., 1994). It was found that the tungsten centre oxidation level was not homogeneous, and the three oxidative states $W^{IV}/W^{V}/W^{VI}$ could be found in purified enzymes.

5.1.1. AOR

AOR *stricto sensus* isoenzymes are the most studied enzyme of the "superfamily" (Seelmann et al., 2020). Chan and co-workers resolved the crystal structure of the *P. furiosus* AOR (2.3 Å) for the first time in 1995. This work shows that the main structure of the characterised AOR isoenzymes from the thermophilic organism is a homodimer, each containing a W-bisMPT, a [4Fe-4S] cluster and a single tetrahedral iron atom coordinated by the two subunits. The Fe atom was found to be coordinated by Glu332 and His383, forming an EXXH motif. They also found an additional Mg2+ coordinated by the phosphate group of the MPT cofactor with residues Asn93 and Ala183 and two H2O molecules. Additionally, a sodium atom coordinated by O-4 (see Figure I-7 for atom numbers) of the pterin was found (Figure I-8) (Chan et al., 1995).


A
[4Fe-4S] cluster
1 Fe
W-bisMPT


B
Asn93
Cys494
Ala183
Cys288
Cys291
Cys295

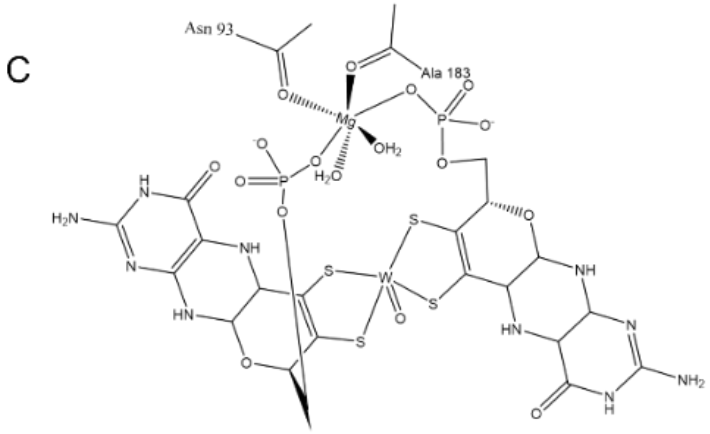

C
Asn 93
Ala 183
Mg
W

Figure I-8 Crystal structure of P. furisosus AOR (PDB: 1AOR).A) crystal structure of the homodimer with the W cofactor, [4Fe-4S] cluster and the Fe atom at the interface of the two subunits. B) crystal structure of the two cofactors within a subunit with the residue coordinating Mg (Asn93 and Ala183) and the four cysteines coordinating the iron-sulfur cluster (Cys288, Cys291, Cys295 and Cys494). C) representation of the complete W-bisMPT present in AOR. W in light blue, Mg in light green, C in dark green, O in red, N in blue, S in yellow, P in orange, Fe in brown and Na in purple (Chan et al., 1995). The modelisation was rendered using mol*viewer (Sehnal et al., 2021)The AOR subunit is composed of three domains. Both cofactor binding sites are located at the interface of these three domains. Domain I (residues 1-210) is composed of 12 β-sheet with a saddle-like shape on which the W-bisMPT sits. Domain II (residues 210-417) and domain III (residues 418-605) close the cage around the cofactors and provide residues to stabilise the cofactors. The tungsten cofactor is bound to the protein with residues from two similar motifs, Asp-X-X-Gly-Leu-Cys/Asp-X, present in domain II (residues 338 to 344) and III (residues 489 to 495). Asp, Leu and Gly interact with their respective pterin through the amino group of the first ring, while the Asp/Cys bind the N-8 of the second pterin ring. Hydrogen bonds are formed between the second nitrogen ring of the pterin and Thr344 and Leu495. The [4-Fe-4S] cluster is positioned ~10 Å from the pterin and is coordinated by three cysteine from domain II (Cys288, Cys291, Cys295) and one from the third domain (Cys494) Figure I-8 B). The cluster is also placed 6 Å below the protein surface, suggesting it plays a role in the electron transfer from the W to the ferredoxin. Cys494 interacts with cofactors and Arg76, suggesting that the pterin might play a role in the redox reaction. Moreover, a 15 Å hydrophobic channel was found between the surface of the protein to the tungsten cavity, which seems to be able to accommodate a wide range of substrates. The exact mechanism of the AOR is not well known (Chan et al., 1995; Roy et al., 2011).

Interestingly, AOR with an $\alpha_2\beta_2\gamma_2$ were found in *Aromatoleum aromaticum* EbN1 (Arndt et al., 2019) and *M. thermoacetica* (Claudia Huber, Skopan, Feicht, White, & Simon, 1995). The two additional subunits contain, respectively, four [4Fe-4S] cluster and a FAD cofactor, allowing to use different electron acceptors such as NAD^+ (Arndt et al., 2019).

The AORs were found to use a wide range of substrates, such as short-chain and branched-chain aliphatic aldehyde and aromatic aldehydes, as well as there their corresponding carboxylate (Napora-Wijata, Strohmeier, & Winkler, 2014; Roy et al., 2001). The redox potential of the carboxylate/aldehyde ($E'_0 \approx$ -560 mV) is very low, making the reaction thermodynamically very favourable toward the aldehyde oxidation (Mukund & Adams, 1991). However, AOR can catalyse the reverse reaction using reduced ferredoxin as electron donor. Hence, physiologically the AOR are involved in pathways where carboxylates or aldehydes are involved. In wild-type *P. furiosus*, it is supposed that AOR is involved in the cell detoxification from aldehydes produced from 2-ketoacids during sugar and peptide fermentation despite not being experimentally proven (Heider et al., 1995). However, in *Thermoanaerobacter* strain X514, Basen et al. found that the native AOR could convert acetate to acetaldehyde. In contrast, when *adhA*, a heterologous gene coding for an alcohol

dehydrogenase, was introduced, acetaldehyde was converted to ethanol (>20 mM). Moreover, when a CODH-encoding gene from *Thermococcus onnurineus* was heterologously expressed in the strain, they showed that the CO could be used as the sole source of reducing power to produce alcohol from diverse organic acids, in cell-free extract, with the native AOR and heterologous AdhA (Basen et al., 2014). The same kind of pathways were shown to happen natively in acetogenic clostridia, such as *Clostridium autoethanogenum*. This strict anaerobic gram+ mesophilic bacteria can grow on CO, CO_2 + H_2 or fructose and possesses two putative AORs which were confirmed to be involved in the reduction of organic acid to the corresponding alcohol using *aor1* and *aor2* mutant strain (Liew et al., 2017). Moreover, Marcellin et al. demonstrated that under autotrophic condition the reduction of acetate to acetaldehyde by AOR in the ethanol production could be coupled to the CO oxidation by CODH (Marcellin et al., 2016). The same kind of reaction involving an AOR was found in the closely related strain *Clostridium ljungdahlii* (Whitham Jason et al., 2015). Moreover, one of the three putative AORs of this strain is more expressed and is supposed to use molybdenum instead of tungsten contrary to the other two (Köpke et al., 2010a; Whitham Jason et al., 2015).

5.1.1. FOR

The FOR was characterised after the AOR and GAPOR in *P. furiosus*. The enzyme was named after discovering its ability to oxidise formaldehyde as well as other short-chain aldehydes (C2-C_4) with lower affinity ($K_m \geq 25$ mM) than AOR. However, the FOR was also found to use glutaric dialdehyde (which does not belong to any pathway in *P. furiosus*) with a higher affinity (K_m = 800 µM). Moreover, the FOR was found to be activated *in vitro* by incubation with sulphide. It was then suggested that the FOR *in vivo* actual substrate was some C_4 to C_6 aldehyde or semialdehyde produced during Arg, Lys and Pro metabolism, implying that FOR might play a role during the growth on peptides (Roy et al., 2001; R. Roy et al., 1999).

FOR was characterised as a tetramer of four identical subunits of ~70 kDa. The metallic content of each subunit is identical to AOR except for Ca^{2+} replacing

Na^+ bounded to O-4 of one of the MPT moieties. Also, the FOR sequence lacks the EXXH motif indicating that no Fe atom was creating a bridge between the two subunits. These were confirmed by the crystal structure. The crystal structure of the *P. furiosus* FOR was the second to be solved (1.85 Å) (Figure I-9 A) and showed a tetrameric structure composed of identical subunits with high homology with AOR subunits. The subunits interact with each other by two interfaces: I) a ~750 Å² between the domain I of subunit A and D, B and C, II) ~700 Å² between the domain II of subunit A and B, C and D. Interestingly the FOR was also crystalised in complex with glutarate (the oxidation product of glutaric aldehyde a known inhibiter of the enzyme (Roy et al., 2001)) and ferredoxin (Hu et al., 1999).

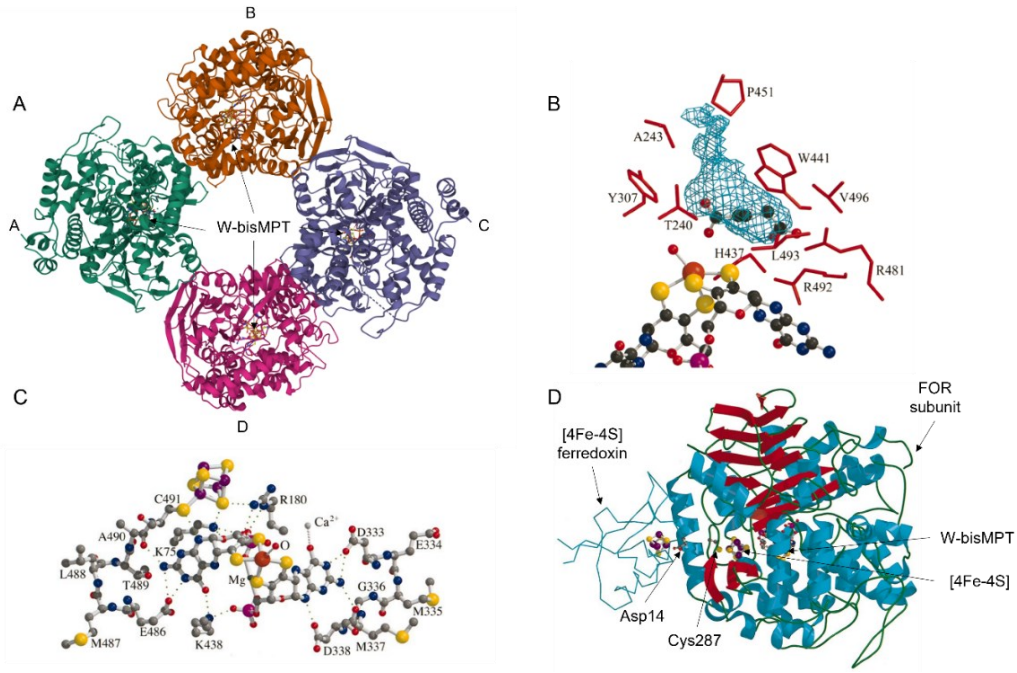


Figure I-9 FOR crystal structures. A) The overall structure of P. furiosus FOR (PBD 1B25) the image was taken from PDB using Mol*Viewer. B) Active site of the FOR subunits with glutarate mimicking the FOR substrate next to the W-bis MPT cofactor as defined using VIDOO.C) Representation of the cofactor with the surrounding residues. D) Complex of a FOR subunit and P. furiosus ferredoxin (backbone representation). The residues from each protein in van der Waals contact Asp14 (Fd) and Cys287 (FOR), Which are supposed to help couple the two iron-sulfur centres. Figures B), C), and D) were taken from Hu et al. article (Hu et al., 1999).

Like the *P. furiosus* AOR, the protein is composed of three domains. Domain I (residues 1-208) forms the base of the enzymes where the W-bisMPT sits with domains II (residues 209-406) and III (residues 407-619), closing the protein over the two cofactors. The first domain has the highest similarities with AOR

domain I (54%). It decreases with the domains (II: 31.4%), while the third domain has the lowest sequence similarity (III: 26%) (Guerrini, 2007).

The tungsto-pterin cofactor is identical to the AOR cofactor apart from Ca, replaced by Na and an oxygen atom coordinating W in addition to the four pterin sulfurs. However, it is still ambiguous if the W is coordinated with an oxo or a hydroxyl group. Moreover, in AOR, the pterin is bound to the protein by two almost identical motifs in domains II and III (DXXGL[C/D]X).However, in domain III of the FOR, it is replaced by the sequence EMLTAC (starts with Glu486), leading to an unusual *cis* peptide bond configuration in place of an α-helix. The first sequence motif in domain II is conserved in FOR and starts with Asp333 (Figure I-9 C). As mentioned above, a Calcium atom is coordinated by the O-4 of one of the pterin, three protein residues (Gly 179, Glu304 and Asp306) and two water molecules. In addition, one of the water molecules has a hydrogen bond with the phosphate group of the second pterin. The second cofactor is coordinated by four cysteines with the same pattern as in AOR: three in the domain II CXXCXXXCG (Cys284, Cys287, Cys291) and one in the domain III (Cys491). Similarly to the AOR, three residues are linked by hydrogen (Cys491) bond and salt bridges (Lys75 and Arg180) between the iron-sulfur centre and the pterin. This suggests that the pterin should play a role in the electron transfer (Figure I-9 C) from the oxidised W to the iron-sulfur cluster (Hu et al., 1999).

The FOR was crystallised in a complex with its physiological electron acceptor (apparent K_m = 100 µM) (Figure I-9 D). In addition, it was measured that the W is 13 Å away from the [4Fe-4S] cluster. The FOR interface with the Fd is a 350 Å² shallow depression rather hydrophobic without any charged amino acid centred around Tyr286 and Cys287. The latter is one of the coordinating residues of the [4Fe-4S] cluster. On the other side, the Fd interface mainly consists of Asp14, Ala15 and Ile16. Those three residues are part of the iron-sulfur cluster binding motif, with Asp14 coordinating one of the sulfurs. It was found that Asp14 (Fd) and Cys287 (FOR) are in van der Waals contact and should help to couple the two iron-sulfur clusters for electron transfer. Hence,

it was suggested that the electrons were transferred from the oxidised W centre to the 13 Å away FOR iron-sulfur centre with the help of either or both hydrogen bounds and salt bridges. The redox centre transfers the electrons to the ferredoxin with the help of the two residues in van der Waals contact (Hu et al., 1999).

The higher resolution of the X-Ray structure helped to have a better understanding of the catalytic cavity (Hu et al., 1999). It has a volume of 1500 $Å^3$ with a hydrophobic channel leading toward the surface of the domain and a large chamber where the W-bisMPT lies at the bottom. In the surrounding of the W, the surface of the cavity is layered with side chains of charged and polar amino acids. Among these amino acids, Thr240, Asp320, Tyr307, Glu308, Tyr416 and His437 conserved a configuration around the W centre very similar to the AOR and other AOR-like proteins. The interaction of a substrate with the residue and the W cofactor was studied thanks to the FOR-glutarate complex. It was observed that one of the carboxylates of the glutarate was forming a hydrogen bond with the Glu308 side chain, and the second carboxylate had electrostatic interaction with Arg481 and Arg492. Moreover, the first carboxylate group of the substrate was in the correct distance to create hydrogen bound with Glu308, Tyr416 and His437 side chains.

By comparison with the Mo aldehyde oxidoreductase from *D. gigas,* Hu and co-workers proposed a catalytic mechanism for the FOR and, by extension, a hint for the AOR mechanism as both catalytic sites are pretty similar (Hu et al., 1999). The proposed mechanism (Figure I-10 pathway A) involved a nucleophilic attack on the α-carbon of the aldehyde by the water bounded tungsten continued by a hybrid transfer to the oxo-ligand of the tungsten. The Glu308 has the role of activating a water molecule to attack the bound substrate. A more recent computational study proposed a reaction coherent with the experimental apparent V_{max} (Figure I-10 pathway B). The substrate would bind directly to the W centre, followed by a nucleophilic attack on the carbonyl carbon atom of the substrate by the W tungsten oxo group. Then

simultaneously, Glu308 abstracts a proton from the intermediate substrate and two electrons reduce the W atom (Liao, Yu, & Himo, 2011).

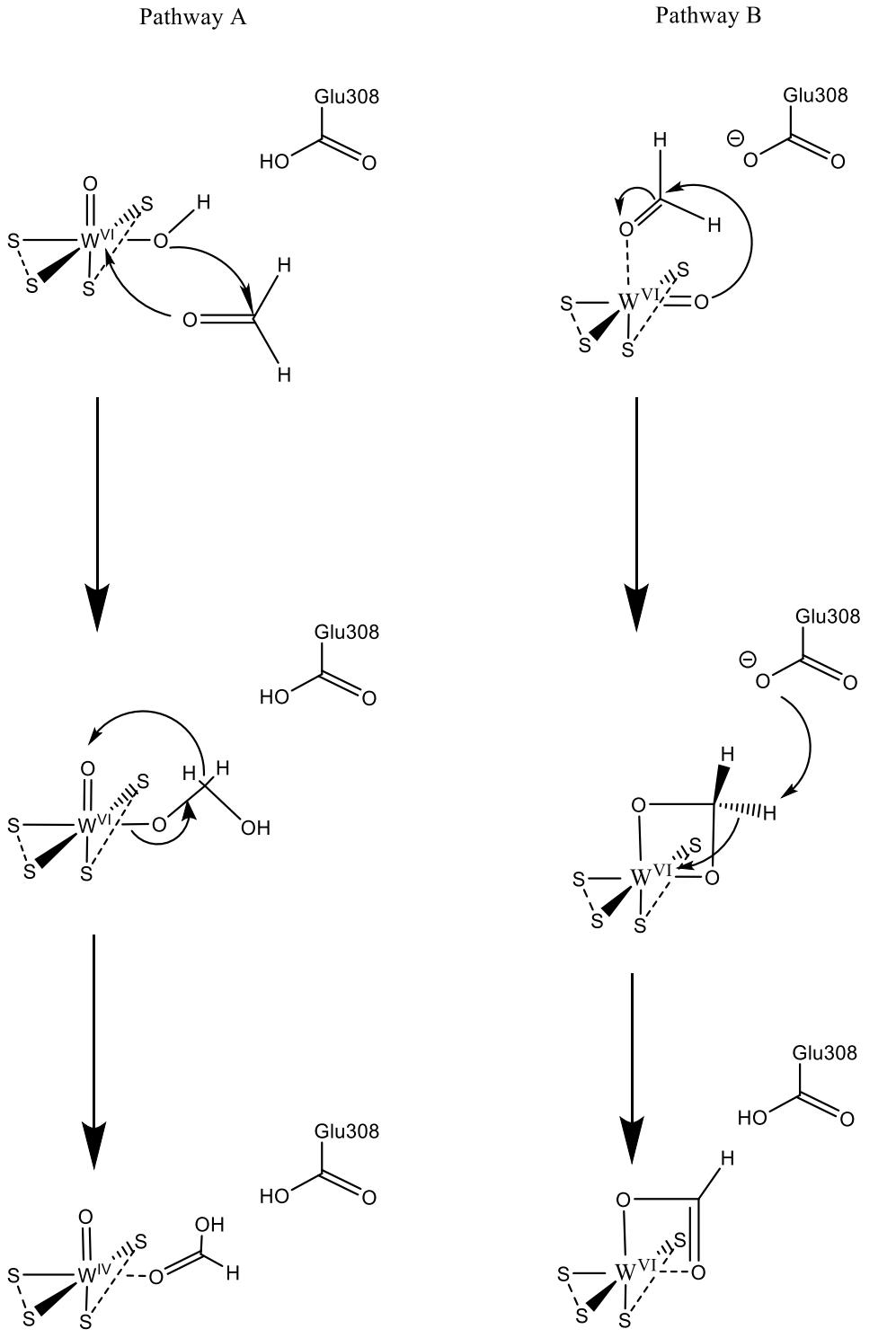


Figure I-10 Theoretical reaction mechanism of FOR with formaldehyde. Pathway A is the theoretical mechanism from Adams and co-worker and pathway B was the computationally estimated proposed by Liao at al. The reaction schemes were redrawn from Liao et al. article for visibility (Liao et al., 2011).

5.1.2. WOR and other

Two other tungsten-containing aldehyde ferredoxin oxidoreductases, WOR4 and WOR5, were also found in *P. furiosus* (R. Roy et al., 1999). WOR4 was identified as a homodimer (ca. 69 kDa subunit) containing one W, one Ca atom, and a [3Fe-4S]. However no substrates were found for this enzyme. WOR5 was characterised as homodimer with a large 65 kDa subunit containing the tungsten cofactor and one [4Fe-4S] cluster per subunit (L. E. Bevers et al., 2005). During this study, the adjacent gene in the *wor5* operon, was characterised as coding for a 19 kDa protein with 16 cysteines. However, this protein was not observed after purification. When the first bacterial GAPOR (GOR) was discovered (Scott et al., 2019), it was found that this gene was homologous to the *gorS* coding the small subunits 4[4Fe-4S] of GOR, which led to propose that WOR5 also had an undetected 4[4Fe-4S] containing subunit. Very recently, the X-ray crystal structure of WOR5 (1.944 Å) was published on PDB in complex with taurine (PDB: 6X6U,Figure I-11) (Lanzilotta & Mathew, To be published). The structure was resolved as heterotetrameric in $\alpha_2\beta_2$, with the already studied 65 kDa large subunit and a small subunit of 166 amino acids (~19 kDa). In the crystal structure no Ca was detected and it was replaced by Mg. WOR5 was found to be able to use the same range of aldehydes as AOR with a lower apparent affinity but had its highest catalytic efficiency with hexanal (V_{max} = 15.6 U/mg, K_m = 0.18 mM). Ferredoxin was supposed to be its physiological electron acceptor (L. E. Bevers et al., 2005); however, its role in *P. furiosus* remains unknown.

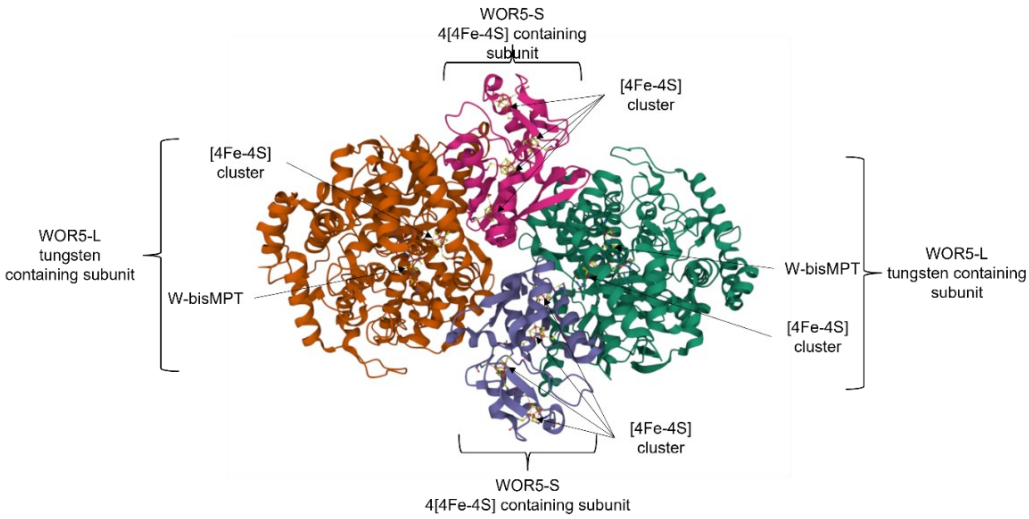


Figure I-11 Structure of the P. furiosus WOR5 (6X6U). The large 65 kDa subunit (orange and green) contains one W-bisMPT and a single [4Fe-4S] cluster. The small ~19 kDa subunit contains four [4Fe-4S]

Recently another member of the AOR superfamily was identified: YdhV (Reschke et al., 2019) from *E. coli*, a 700 amino acids protein found to incorporate a bisMPT cofactor without nucleotide substitution, a preference for Mo instead of W and a predicted presence of a [4Fe-4S] redox centre. The *ydhV* is part of the *ydhYVWXUT* expressed under anaerobic conditions. YdhY and YdhX were identified as ferredoxin-like protein. YdhY is predicted to contain four iron-sulfur clusters susceptible to correspond to the small subunit of WOR5 or GOR and YdhX to be periplasmic due to the presence of a TAT sequence. YdhU was predicted to be an integral membrane protein containing a cytochrome *b*. The role and activity of YdhW and YdhT are still unknown (Partridge et al., 2008). No activity was detected for YdhV, probably due to a potentially disassembled iron-sulfur cluster. The cluster might be sensitive due to a "hyperactive" cysteine residue (Reschke et al., 2019).

The Adams group also purified two new tungsten-containing enzymes during a study on tungsten-containing oxidoreductase's potential role in detoxifying human guts (Schut et al., 2021). The enzymes were isolated from *Eubacterium limosum* and *Acetomicrobium mobile*. Interestingly both WOR purified from *A. mobile* were found to have a large tungsten-containing subunit and a small 4[4Fe-4S] clusters containing subunit. Moreover, AmWOR1 is the first electron bifurcation WOR able to couple the aldehyde oxidation with the reduction of

Fd and production of NADH. This enzyme contains three more subunits, containing Fe-S centres and a flavine cofactor (Figure I-12) (Schut et al., 2021). All the purified enzymes were able to reduce-food related toxic aldehydes. However, their actual physiological role is still unknown.

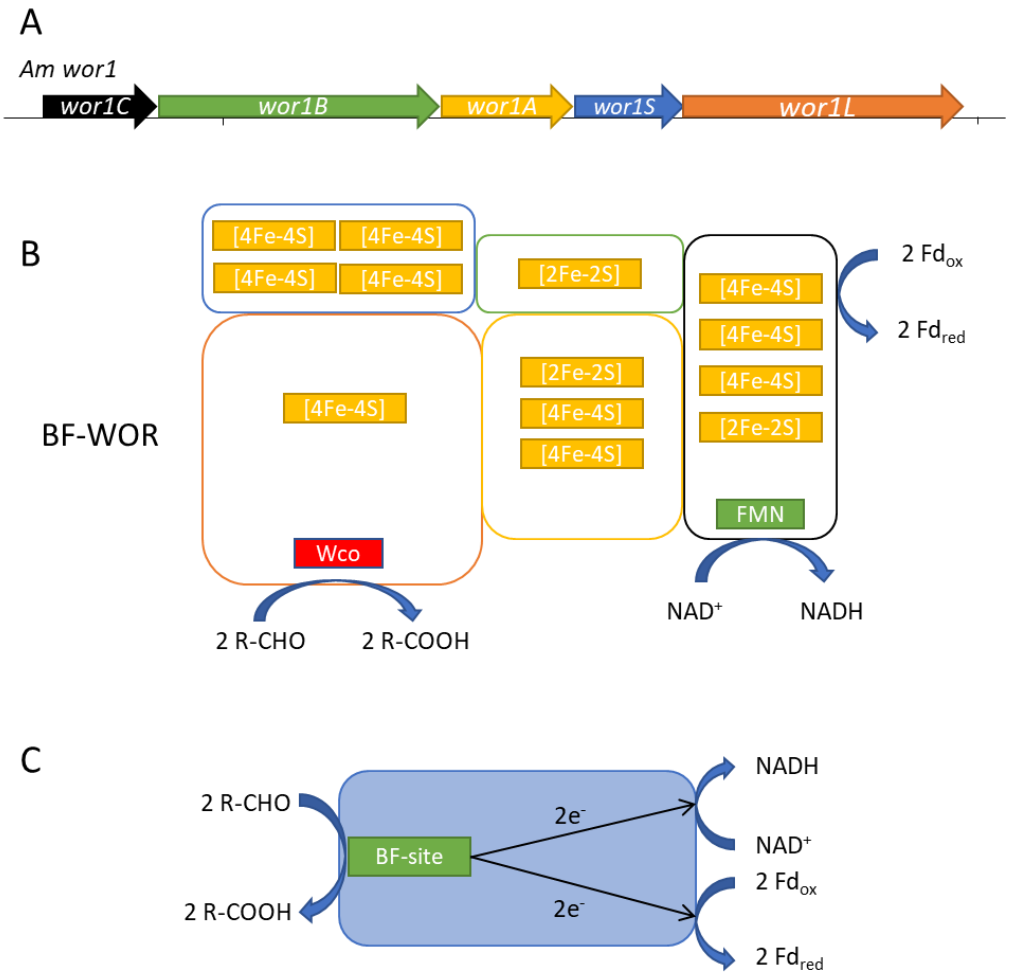


Figure I-12 Bifurcation tungsten ferredoxin oxidoreductase 1 (BF-WOR1) from *Aacetomicrobium mobile*. A) scheme of the AmWOR1 operon. B) Representation of the BF-WOR with the cofactor of each subunit and the supposed role. C) schematic representation of the aldehyde oxidation electron bifurcation reaction of BF-WOR. All the figures were reproduced for visibility from the Adams group article (Schut et al., 2021).

These recent findings lead to believe that the AOR superfamily is rather important (Figure I-13), more than only in archaea. Future studies will bring more understanding on their role in bacterial metabolism and their potential role in the human gut's microbiome.

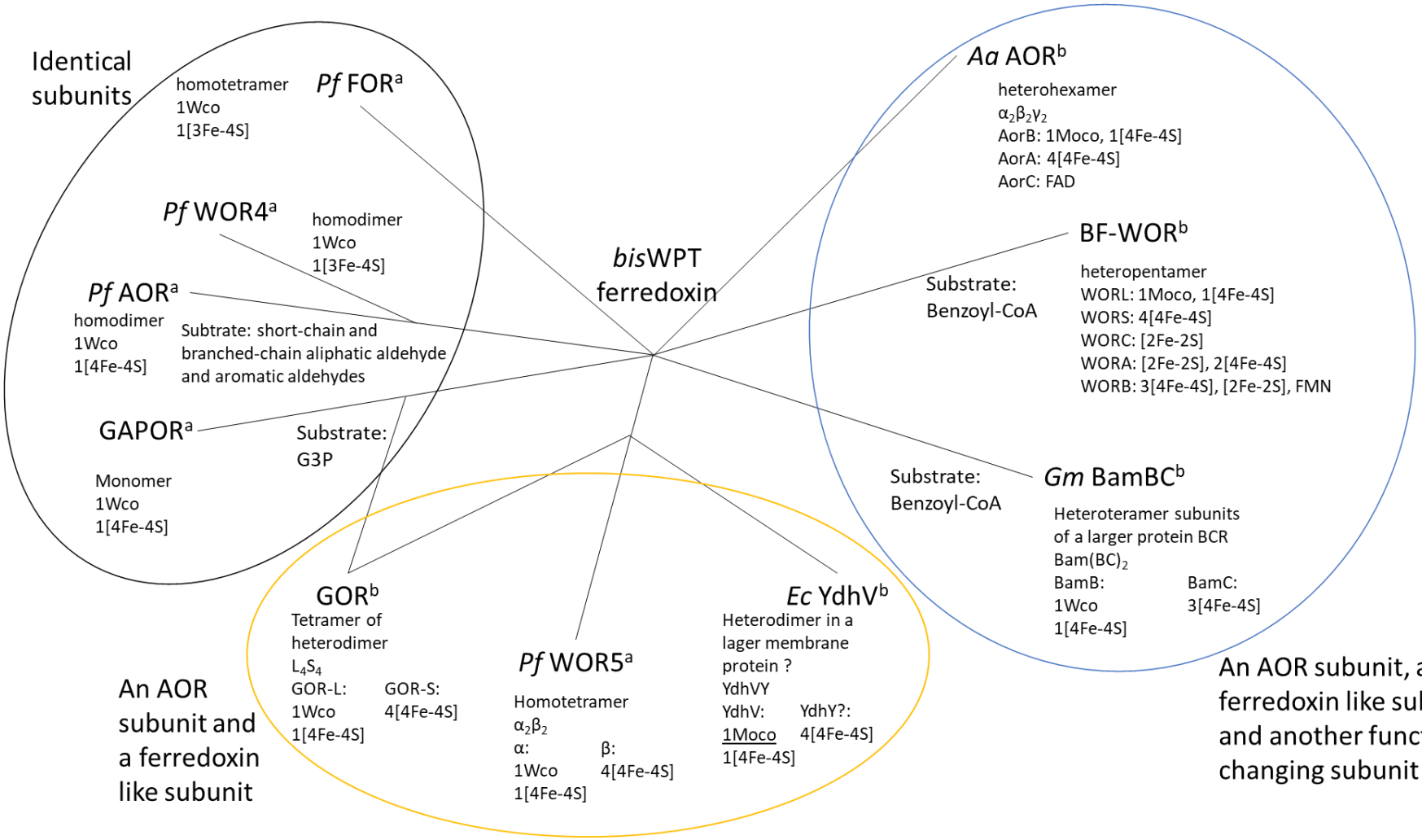

Identical subunits
homotetramer
1Wco
1[3Fe-4S]
Pf FOR[a]
Pf WOR4[a]
homodimer
1Wco
1[3Fe-4S]
Pf AOR[a]
homodimer
1Wco
1[4Fe-4S]
Subtrate: short-chain and branched-chain aliphatic aldehyde and aromatic aldehydes
GAPOR[a]
Substrate: G3P
Monomer
1Wco
1[4Fe-4S]
bisWPT ferredoxin
Aa AOR[b]
heterohexamer
α2β2γ2
AorB: 1Moco, 1[4Fe-4S]
AorA: 4[4Fe-4S]
AorC: FAD
BF-WOR[b]
Substrate: Benzoyl-CoA
heteropentamer
WORL: 1Moco, 1[4Fe-4S]
WORS: 4[4Fe-4S]
WORC: [2Fe-2S]
WORA: [2Fe-2S], 2[4Fe-4S]
WORB: 3[4Fe-4S], [2Fe-2S], FMN
Substrate: Benzoyl-CoA
Gm BamBC[b]
Heteroteramer subunits of a larger protein BCR
Bam(BC)2
BamB: 1Wco 1[4Fe-4S]
BamC: 3[4Fe-4S]
GOR[b]
Tetramer of heterodimer
L4S4
GOR-L: 1Wco 1[4Fe-4S]
GOR-S: 4[4Fe-4S]
Pf WOR5[a]
Homotetramer
α2β2
α: 1Wco 1[4Fe-4S]
β: 4[4Fe-4S]
Ec YdhV[b]
Heterodimer in a lager membrane protein ?
YdhVY
YdhV: 1Moco 1[4Fe-4S]
YdhY?: 4[4Fe-4S]
An AOR subunit and a ferredoxin like subunit
An AOR subunit, a ferredoxin like subunit, and another function changing subunit

Figure I-13 Summary scheme of the model enzymes of the AOR superfamily.In this figure are presented the best-known enzymes among the more than 4, 000 identified protein of the superfamily. 92 clades were identified by Schut and Scott. The enzymes are grouped depending on their subunit's composition and lines represent their similarities. The first group circled in black comprises enzymes with identical subunits containing one tungsten cofactor and an iron-sulfur centre, mainly [4Fe-4S]. This subunit is considered the basic AOR subunit containing the catalytic centre. The second group, circled in yellow, is composed of heteromeric enzymes. The primary subgroup is a heterodimer of an AOR subunit and a ferredoxin-like subunit. Two or more subgroups compose the enzyme. The third group, circled in blue, comprises enzymes with the yellow group heterodimer and other subunits to change the function of the enzyme, electron bifurcation, part of a larger enzyme complex. The cofactor compositions of the subunits are described in the scheme and the known physiological substrate. *Aa* AOR substrates are identical to *Pf* AOR. *Pf*: *Pyrococcus furiosus*, *Aa*: *Aromatoleum aromaticum* EbN1, *Ec: Escherichia coli*, FOR: Formaldehyde: ferredoxin oxidoreductase, WOR: tungsten- containing aldehyde ferredoxin oxidoreductases ,AOR: aldehyde: ferredoxin oxidoreductases, GAPOR: archaeal Glyceraldehyde-3-phosphate: ferredoxin oxidoreductase, GOR: bacterial Glyceraldehyde-3-phosphate: ferredoxin oxidoreductase, BF-WOR: electron bifurcation-WOR, BCR: benzoyl–coenzyme A reductases, Wco: tungsten cofactor here tungsten-bis-molybdopterin, Moco: molybdenum cofactor, here Mo-bis-molybdopterin, BCR: benzoyl–coenzyme A reductases, FAD: flavin adenine dinucleotide, FMN: Flavin mononucleotide. [a] enzyme found in archaea, [b] enzyme found in bacteria (Arndt et al., 2019; Mukund & Adams, 1991, 1995; Reschke et al., 2019; Roy, Mukund, Dunn, Weiss, & Adams, 1999; Schut et al., 2021; Scott et al., 2019; Weinert et al., 2015).

5.2. Glyceraldehyde-3-phosphate: ferredoxin oxidoreductase

5.2.1. Structure

The Glyceraldehyde-3-phosphate: ferredoxin oxidoreductase (GAPOR) (EC 1.2.7.6) was first described in *Pyrococcus furiosus* as an enzyme catalysing the ferredoxin-dependent oxidation of the glyceraldehyde-3-phosphate (G3P) in 3-phosphoglycerate (3PG) without the intermediary step of 1,3-diphosphoglycerate (1,3PG) (Mukund & Adams, 1995). The gene was identified by Oost et al. as a 1956 bp (Van Der Oost et al., 1998) coding for a 73.9 kDa (apo-enzyme) monomeric enzyme contrary to all the other purified AOR that were all either homomultimeric or heteromeric. GAPOR polypeptide sequence has a 50% similarity with AOR/FOR sequences, andcoth cofactor binding sequences are conserved in the same overall position in the protein sequence. In addition, with the elemental analysis of the purified enzyme from *P. furiosus,* GAPOR was found to contain W-bisMPT cofactor withan Mg atom. AOR's Na and FOR's Ca were apparently replaced by 2 Zn atoms. However, in the absence of structural information, the Zn atoms' position is unknown.

In the absence of a crystal structure, Guerrini created during its thesis a homology model (Figure I-14) of the GAPOR from *Methanococcus janaschii* (Van Der Oost et al., 1998) (57% homology and 42% identity with *P. furiosus* GAPOR) thanks to the crystal structure of *P. furiosus* AOR and FOR. He found that the GAPOR sequence was also divided into three domains: domain I (residues 1 to 220) has the highest secondary structure homology with the domain I of the two other AOR and plays the same role of base for the pterin cofactor. Domain II (residues 201 to 440) has high sequence and structure conservation in the cofactor binding motif and around the reaction centre. However, the homology is lower than the domain I. The third and last domain (residue 441 to 622) has the lowest homology with the comparison sequence. Several gaps were observed in the alignment. However, the cofactor binding sites were still conserved. The same type of variability was observed when solving the FOR structure (Hu et al., 1999). Hence, it was supposed that this

variability could play a role in the enzyme's substrate specificity. The same cavity and channel toward the surface were observed. Two highly conserved Tyr352 and Glu353 (Glu308 in FOR) were observed in the cavity along with W-bisMPT. This led to supposed that the reaction mechanism of the GAPOR is similar to the already characterised AOR (Guerrini, 2007).

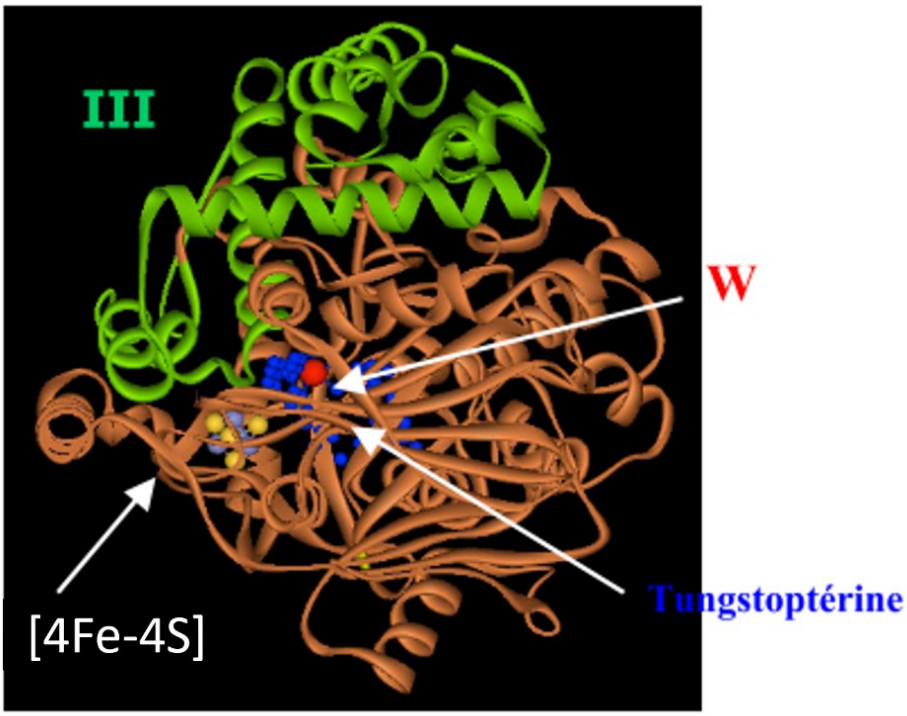


Figure I-14 Homology model of the M. janashii GAPOR using P. furiosus AOR and FOR resolved structures as templates (Guerrini, 2007). Domain III is visualised in green, and the position of the two cofactors in the model are marked with arrows.

GAPORs have been found in different archaea such as *Archaeoglobus fulgidus* (Labes & Schönheit, 2001), *M. maripaludis* (Costa, Lie, Jacobs, & Leigh, 2013; Guerrini, 2007; Park et al., 2007), *Aeropyrum pernix* and *Pyrobaculum aerophilum* (Reher, Gebhard, & Schönheit, 2007). The last two organisms have GAPOR, but only *P. aerophilum* has exhibited GAPOR activity. The most studied GAPOR is from *P. furiosus* (Mukund & Adams, 1995). Moreover, GAPOR from *Methanoccoccus jannashii* has been studied by Guerrini. However, no G3P oxidation activity was observed after its heterologous expression in *E. coli* and *C. acetobutylicum* (Guerrini, 2007) in the presence of W excess in the medium. Putative GAPOR encoding genes have also been identified in *Pyrococcus*

Horikoshi, and Pyrococcus abyssi (Guerrini, 2007). Interestingly, the GAPOR from the mesophilic methanogen *M. maripaludis* was shown to be Mo-dependent and inactive in the presence of W (Park et al., 2007).

5.2.2. Catalytic activity

GAPOR, in contrast to the other characterised AORs, which use an extensive range of substrates, only use glyceraldehyde-3-phosphate (G3P). The enzyme catalyses the oxidation of G3P to 3-phosphoglycerate (3PG) without the intermediary state of 1,3-diphosphoglycerate (1,3PG) following the reaction (3) (Mukund & Adams, 1995).

$$RCOH + H_2O + 2Fd_{ox} = RCOOH + 2\,H^+ + 2Fd_{red} \quad E_0' \approx -560\,mV \ (2)$$

Furthermore, it was demonstrated that GAPOR is enantioselective toward D-G3P, but the DL-G3P mixture was not seen to have an inhibitory effect on GAPOR activity, making it suitable for activity assays (P. L. Hagedoorn, 2019). *In vitro* benzyl viologen (BV) or methyl viologen was used as the electron acceptor and indicator in the colorimetric enzyme assay with *P. furiosus.* GAPOR exhibited a K_m of 0.43 µM for BV. *P. furiosus* GAPOR had an apparent K_m for G3P of 30 µM and an apparent V_{max} of 350 U/mg at 70°C using BV (range between 0.01 mM and 0.5 mM) (Mukund & Adams, 1995). Above 0.5 mM G3P, a significant substrate inhibition was observed. This inhibition was recently studied by Hagedoorn and was found to be partial substrate inhibition with a K_i at least 10-fold smaller than K_m in all the tested conditions. He suggested that the GAPOR probably has an inhibitory binding site with a high affinity for G3P (P. L. Hagedoorn, 2019). This inhibition was found to be reduced in the presence of sodium chloride (NaCl) by probably reducing the affinity of G3P for the inhibitory binding site. Hence GAPOR was shown to have higher activity at higher substrate despite increasing the K_m for G3P. After global fitting of the steady-state kinetic with the different NaCl concentrations tested, he reported a V_{max} at 60°C with D-G3P of 470 U/mg (P. L. Hagedoorn, 2019). Other substances were found to stimulate the activity *i.e.* potassium phosphate, sodium arsenate, potassium chloride, sodium citrate or sodium sulphate (Mukund & Adams, 1995). In addition to the substrate inhibition, the *M.*

maripaludis GAPOR was found to be inhibited by ATP, which was suggested to be a post-translational regulator of the enzyme (Park et al., 2007). In the Hagedoorn study, GAPOR was found to have an optimal pH of 9, suggesting that the negatively charged phosphate group of the G3P is required for substrate recognition as not activity was detected with glyceraldehyde or methylglyoxal. Moreover, the optimal temperature was observed around 80°C, but it was noted that G3P is unstable at *ca*. 100°C.

It has been shown that GAPOR uses oxidised ferredoxin (Fd_{ox}) as an electron acceptor (K_m = 6 µM for ferredoxin from *P. furiosus at 70°C*) and does not accept NAD^+ or $NADP^+$ as electron acceptors. The GAPOR with it physiological electron acceptor has an apparent V_{max} of 90 U/mg (for assays with 0.1 to 50 µM of ferredoxin) (Mukund & Adams, 1995).

The electronic paramagnetic resonance (EPR) study of the GAPOR found two species of tungsten. The active one, W^{V1+}, is an intermediary state in the catalytic cycles of two subsequent 1 e^- reductions of the W from W^{IV} to W^{VI} with two midpoint potentials $E_m(W^{IV/V})$ = -507 mV and $E_m(W^{V/IV})$ = - 419 mV. The inactivate one with a higher potential $E_m(W^{IV/V})$ = -329 mV. This type of tungsten was found in 10 % of the GAPOR (Hagedoorn Peter, Freije, & Hagen Wilfred, 1999).

The EPR analysis of $[4Fe\text{-}4S]^+$ cluster showed a usual spin S= 3/2 and an unusual S = 1/2 with very similar potential $E_m([4Fe\text{-}4S]^{2+/1+})$ = -333 mV and -336 mV respectively. The first spin of the cluster was also found for the redox centre of the *P. furiosus* AOR. It was suggested that the S = 1/2 was reflecting the interaction of the active W^{VI+} with the [4Fe-4S] cluster (Hagedoorn Peter et al., 1999).

The results of the substrate/product (GAP/3PG) redox titration with G3P, GAPOR and ferredoxin showed that the electrons released by the oxidation of the G3P are transferred to the W centre before going to the [4Fe-4S] cluster and finally reduce the iron-sulfur cluster of the Fd_{ox}. The pterin moieties were

not observed to play a role in this reduction chain (Hagedoorn Peter et al., 1999).

Interestingly, Guerrini during its heterologous expression of GAPOR from *M. jannashii*, *P. furiosus* and *M. maripaludis* in *E. coli* and *C. acetobutylicum,* the purified enzyme exhibited hydrogenase-like activity in the presence of H_2 (Guerrini, 2007).

5.2.3. Physiological role

The GAPOR reaction was found to be part of a modified Embden-Meyerhof-Parnas (EMP) pathway along the glyceraldehyde dehydrogenase (GAPDH) and phosphoglycerate kinase (PGK) (Mukund & Adams, 1995; Park et al., 2007; Reher et al., 2007). It was shown that GAPOR was only catalysing the reaction in the glycolytic direction, and the GAPDH/PGK were observed to only work in the gluconeogenic direction. This type of modified EMP leads to no or very low ATP production. In *P. furiosus,* the GAPOR Fd reduction was shown to be linked with a membrane-bound hydrogenase (MBH) and a proton gradient with an ATP synthase (Figure I-15) (Sapra, Bagramyan, & Adams, 2003). Moreover, in a recent study, a Δ*gapn P. furiosus* strain expressing a primary NADPH-dependent alcohol dehydrogenase encoding gene (*adhA*) from a thermophilic *Caldanaerobacter subterraneus* was able to produce ethanol at 80°C at a significant amount (8:1 ethanol:acetate). The AOR pathway mentioned above uses the Fd_{red} produced by the GAPOR to produce the NADPH needed for AdhA (Straub et al., 2020).

In *M. maripaludis* which only use CO_2 and formate/H_2 to produce methane (Goyal, Zhou, & Karimi, 2016). It has been proposed that GAPOR, GAPDH and PGK can form a cycle (Figure I-15), despite leading to a loss of energy and the consumption of NADPH (Park et al., 2007). The latter can be produced by $F_{420}H2$:$NADP^+$ oxidoreductase following the synthesis of CO_2 from formate by the formate dehydrogenase (Costa et al., 2013). In *M. maripaludis,* the *gorS2* gene is constitutively expressed. Moreover, in this organism, GAPOR was inhibited by G3P and ATP and it has been concluded by Park et al. that this inhibition might play a regulatory role. Moreover, it appears that an

overproduction of GAPOR allows growth on formate without dihydrogen and CO in a *M. maripaludis* mutant deficient in hydrogenases. Therefore, it was concluded that GAPOR was most likely used during non-optimal growth conditions. In addition, as GAPOR is the only enzyme able to oxidise G3P, it seems to play a key role in the glycogen pathway regulation.

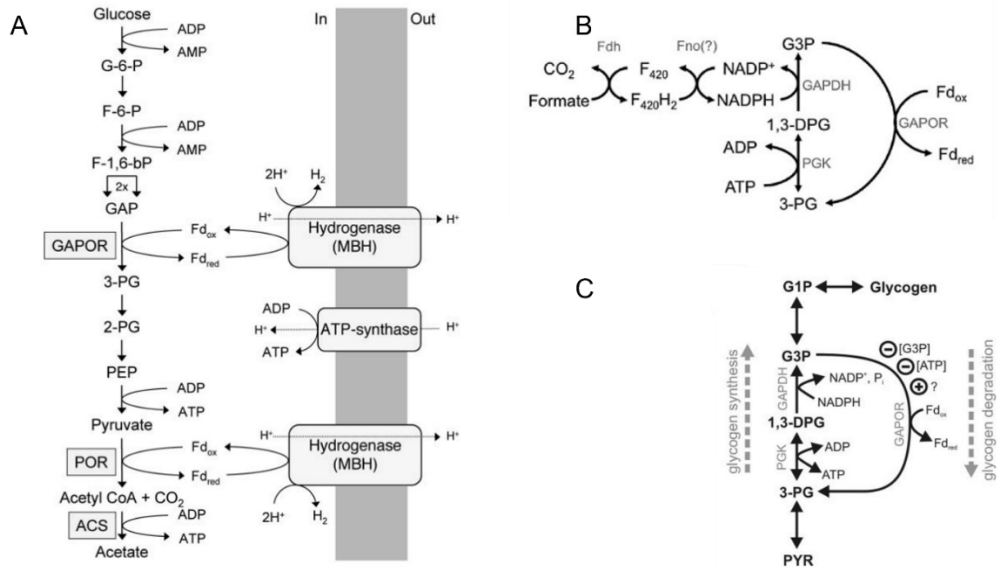


Figure I-15 Modified Embden-Meyerhof-Parnas (EMP) pathway in P. furiosus and M. maripaludis. A) P. furiosus modified EMP pathway in the glycolysis direction, coupling Fd reduction with hydrogenase creating a proton gradient to create ATP with ATP-synthase (Sapra et al., 2003). B) Proposed complementation of the GAPOR-GAPDH cycle with the input of formate dehydrogenase and $F_{420}H_2$:$NADP^+$ oxidoreductase (Costa et al., 2013). C) Proposed GAPOR-GAPDH cycle with the different GAPOR inhibitors for the proposed glycogen pathway (Park et al., 2007). GAPOR: glyceraldehyde-3-phosphate: ferredoxin oxidoreductase, POR: pyruvate: ferredoxin oxidoreductase, ACS: acetyl-CoA synthase, GAPDH: glyceraldehyde-3-phosphate dehydrogenase, PGK: phosphoglycerate kinase, Fdh: formate dehydrogenase, Fno: $F_{420}H_2$:$NADP^+$ oxidoreductase, Fd: ferredoxin (ox: oxidised, red: reduced), G1P: glucose-1-phosphate, G-6-P: glucose-6-phosphate, F-6-P: Fructose-6-phosphate, F-1,6-bp: fructose-1,6-biphosphate, G3P/ GAP: glyceraldehyde-3-phosphate,1,3-DPG: 1,3 diphosphoglycerate, 3-PG: 3-phosphoglycerate, PEP: phosphoenolpyruvate, PYR: pyruvate.

5.2.4. GOR: a bacterial GAPOR

Very recently, the first GAPOR homologous protein was characterised from the extremely thermophilic strict anaerobic cellulolytic bacterium *Caldicellulosiruptor bescii* by the Adams group (Scott et al., 2019). The enzyme was found to be only able to use G3P as the substrate to produce 3PG. However, the phylogenetic study showed that this enzyme belongs to another clade than the archaeal GAPOR and was hence named GOR. Putative isoenzymes were present in 44 other mostly anaerobic bacteria. GOR is a heterodimer like WOR5 with a 65,916 Da large subunit, GOR-L (UniProt:

B9MQI2), coded by Athe_0821 and a 14,618 Da small subunit, GOR-S (Uniprot B9MQI1), coded by Athe_0820. The heterodimer was proposed to form a tetramer of homodimer (L_4S_4) holoenzyme due to the presence of two peaks on size exclusion chromatography with a similar specific activity of apparent size of 315 kDa and 73 kDa. The metal content (Fe/W) and the amino acid sequence analysis suggest that GOR-L contains the same cofactor as GAPOR, and GOR-S contains four [4Fe-4S] clusters like the WOR5 small subunit whose structure was recently solved. The actual structure of GOR is yet to be solved.

When GOR was tested with 1 mM BV and different concentrations of G3P, a linear increase of activity was observed with no saturation up to 20 mM with a k_{cat}/K_m of $4.7x10^4$ $s^{-1}.M^{-1}$ at 70°C. It is interesting to remind that *P. furiosus* GAPOR activity was found to be inhibited above 5 mM G3P. Hence, they proposed that the lack of saturation was due to a low affinity toward G3P and a high turnover rate. They determined a K_m for *C. bescii* ferredoxin was 38.9 ± 10.7 µM with a k_{cat} 5.57 s^{-1} at 70°C with 5 mM G3P as a kinetic saturation was observed making it consistent with being the physiological electron acceptor. *C. bescii* GOR was found to be inhibited by dihydroxyacetone phosphate (DHAP).

In *C. bescii,* GOR is part of a secondary modified EMP pathway and is expressed along with reversible GAPDH and PGK carrying the conventional EMP. Interestingly the GAPDH has an identical unsaturated linear kinetic activity up to 20 mM using 1 mM NAD^+ at 70°C. Moreover, the triose phosphate isomerase was found to be strongly biased toward DHAP displaying a different way of regulating the glycolysis. It was supposed that this alternative pathway might be used by the cell to regulate the carbon and electron fluxes during excess carbon loading. GOR was suggested to be coupled with an Ech hydrogenase creating proton motive force and ATP production with an ATP synthase. This pathway seems to be used by the cell to keep high H_2 production in high NADH or H_2 concentration by carrying ferredoxin-only glycolysis, avoiding NADH excess.

6. Molybdenum Cofactor biosynthesis pathway

As presented in the section before, the GAPOR contains a W/Mo cofactor. This cofactor is different from the final product of the *E. coli* molybdenum cofactor biosynthesis pathway. Hence, the pathway might be a path of activity modification.

6.1. Introduction

Archaea, bacteria, and eukaryotes widely use molybdenum (Mo) and tungsten (W) as transition elements. They are very versatile compounds able of transitions between VI, V and IV oxidation states, which cover reactions involving one or two electrons exchange and a large range of redox potentials (Hille, Schulzke, & Kirk, 2017). Hence, thanks to this wide range of oxidation states, Mo and W can couple proton transfer with electron transfer. Mo/W are involved in the active site of the molybdo/tungsto-enzymes. These enzymes are part of carbon, sulfur and nitrogen metabolism and mainly catalyse oxo-transfer reactions as hydroxylation of carbon centre but also C-H cleavages (Hille, Hall, & Basu, 2014). The electron-transfer through the Mo/W atoms is mediated by other redox centres such as iron-sulfur clusters, cytochromes or FAD/FMN (Arndt et al., 2019; Iobbi-Nivol & Leimkühler, 2013; Schut et al., 2021; Yokoyama & Leimkühler, 2015) or ferredoxins (Mukund & Adams, 1995; Park et al., 2007).

To be incorporated at the correct position in the protein active site, the Mo/W must be chelated by a pyranopterin called a molybdopterin (MPT) to form the molybdenum/tungsten cofactor (Moco/Wco). The structure of the cofactor has been characterised by Rajagopalan as tricyclic pyranopterin with a dithiolene group coordinating an Mo/W atom (Amy & Rajagopalan, 1979; Rajagopalan & Johnson, 1992). The MPT might also be substituted with nucleotides such as guanine (Mo-bisMGD) or cytosine Mo-MCD) (J. L. Johnson, Bastian, & Rajagopalan, 1990). Also, in some nitrogenases, molybdenum is coordinated by the Fe-S cluster to form an iron-based molybdenum cofactor (FeMoco) (Ribbe, Hu, Hodgson, & Hedman, 2014; Schwarz, Mendel, & Ribbe, 2009).

Thermophilic and hyperthermophilic archaea (represented by *P. furiosus*) incorporates W instead of Mo. The tungsten possesses the same oxidation states as Mo. Mo and W are taken by the cells from the environment under their oxidized soluble form of molybdate or tungstate (MoO_4^{2-}, WO_4^{2-}) with two different active transport systems ModABC , WtpABC and TupABC (Hagen, 2011). It has been found that exchange of Mo and W could produce active enzymes as dimethylsulfoxide (DMSO) reductase from *Rhodobacter sphaeroides* expressed in *E. coli* (Neumann, 2008). However, only some enzymes were found to be active with the other cofactor.

Mo and W can be incorporated in heterologous enzymes expressed in *E. coli* thanks to a similar ion radius, discrimination is talked about later. However, this inversion from Mo to W would lead to an inactive human enzyme (human sulfite oxidase) or an active dimethyl sulfoxide (DMSO) reductase from *Rhodobacter sphaeroides* (Neumann & Leimkühler, 2008). Also, it has been shown that when Mo was used to replace W in *P. furiosus* AOR, the enzymes were inactivated (Sevcenco et al., 2010).

The molybdenum cofactor biosynthesis pathway is ubiquitous and the *E. coli* pathway is the best studied. The tungsto-pterin cofactor biosynthesis pathway is assumed to be very similar (Seelmann et al., 2020). The molybdopterin biosynthesis pathway has been fully solved (Mendel & Leimkühler, 2015; Rajagopalan, 1996) and is composed of four general steps using guanine triphosphate (GTP) as the precursor and four intermediary molecules (Figure I-20): (i) cyclic pyranopterin monophosphate (cPMP), (ii) formation of the molybdopterin (MPT) with the addition of two sulfur, (iii) insertion of the Mo atoms on an activated MPT (MPT-AMP) and (iv) synthesis of the Mo-bis-molybdopterin from two MPT. Further modification can be done as substitution with nucleotides or maturation by chaperone protein (Hille, 1996). The enzymes containing Mo/W cofactor were proposed to be classified in function of Mo/W cofactor (Hille et al., 2014; Leimkühler, 2020; Axel Magalon, Fedor, Walburger, & Weiner, 2011; Zupok, Iobbi-Nivol, Méjean, & Leimkühler, 2019):

- Family I: composed of Mo/W-BisMPT cofactor-containing enzyme usually linked to the one [4Fe-4S] cluster, *i.e.* archaeal and bacterial AOR (Hagen, 2011; Reschke et al., 2019).
- Family II: Xanthine dehydrogenase and aldehyde oxidase family, also known as xanthine oxidase (XO) family. They are characterised by a Mo-MPT/MCD with the oxo/hydroxo and sulfur-bounded Mo^{VI} centre. The XO enzymes were found to carry one or more [2Fe-2S] in a second subunit, along with a FAD subunit.
- Family III: dimethyl sulfoxide reductase (DMSOR) family and are exclusively found in prokaryotes, bound a Mo-BisMGD typically with the Mo centre bound with two groups: a sulfur or an oxygen ligand and either a hydroxo or amino acid (Ser, Cys, Sec, and Asp). The enzymes were found to bind different types of iron-sulfur cofactors in the Moco binding subunit and in a second subunit, usually three to four.
- Family IV: Sulfite oxidase (SO) family. They are characterised by containing Mo-MPT with the Mo^{VI} binding a dioxo group as well as a cysteine. In the recent review no SO family members were observed to carry iron-sulfur cofactor.

Moco biosynthesis in *E. coli* involves five operons coding for 15 proteins, including the one coding for the Mo transport: *moaABCDE*, *mobAB*, *modABC*, *modEF*, *moeAB* and *mogA* (Figure I-16) (Rajagopalan*, 1996; Shanmugam et al., 1992)*.

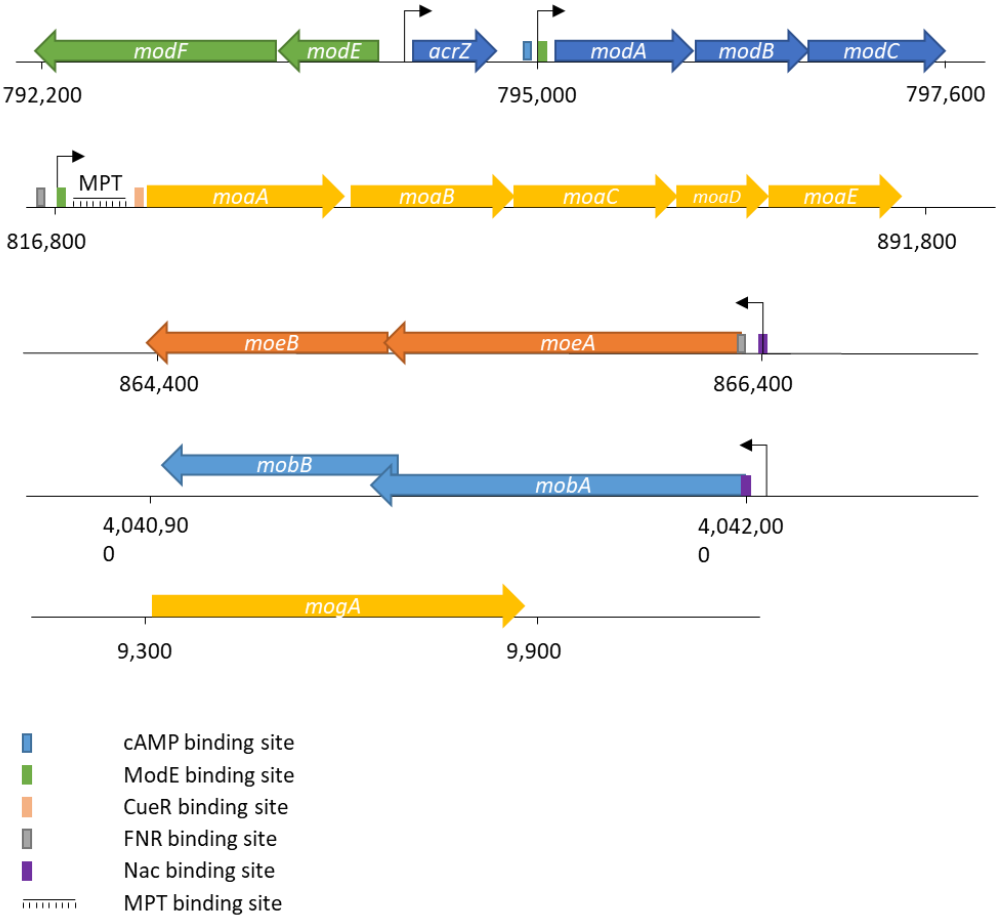


Figure I-16 *moaABCDE*, *mobAB*, *modABC*, *modEF*, *moeAB* and *mogA* operons from *E. coli* str.K-12 substr. MG1655

6.2. Mo/W uptake and regulation

Mo and W enter the cell as molybdate (MoO_4^{2-}) or tungstate (WO_4^{2-}) mainly via the molybdate ABC transporter with the consumption of ATP: Mod (Grunden & Shanmugam, 1997), Wtp (Bevers Loes, Hagedoorn, Krijger Gerard, & Hagen Wilfred, 2006), Tup (Makdessi, Andreesen, & Pich, 2001). The three AB_2C_2 periplasmic active transporters use ATP to translocate the oxyanion (Aguilar-Barajas, Díaz-Pérez, Ramírez-Díaz, Riveros-Rosas, & Cervantes, 2011; Hagen, 2011). Their A subunit is the oxyanion binding and discriminating part of the system in the periplasm. ModA, found in diverse organisms, was shown to have the same high affinity toward Mo and W (Imperial, Hadi, & Amy, 1998). WtpA discovered in *P. furiosus* can also transport W and Mo with the highest affinity for tungsten and a similar affinity for Mo as ModA. TupA, first observed in *Eubacterium acidaminophilum,* can only bind W. ModABC and TupABC are

widely observed in archaea and bacteria, contrary to Wtp, which is only found in archaea which does not have the two others (Seelmann et al., 2020). The B subunits pair are the transmembrane part of the transport system bound to the C subunits pair which contains the ATP-binding cassette (Hagen, 2011).

The expression of *modABC* is regulated by the protein ModE coded by *modEF* (Lisa A. Anderson et al., 1997). ModE is a homodimer binding indistinctively molybdate or Tungstate with a high affinity (K_d = 0.8 µM) (Grunden, Ray, Rosentel, Healy, & Shanmugam, 1996). When bound to the oxyanion, ModE has an increased affinity toward the DNA target sequence (ATCGCTATATA-N6/7-TATATAACGAT *E. coli* consensus sequence), leading to a down-regulation of the transporter expression in the presence of Mo/W in the cell (Lisa A. Anderson et al., 1997; McNicholas, Rech, & Gunsalus, 1997). Homologous protein was found in *Rhodobacter capsulatus* MopA, and MopB with *mopA* coexpressed with *modABC* and *modB* constitutively expressed (Zupok, Iobbi-Nivol, et al., 2019). On the other hand, *modABC* expression is upregulated by the cyclic-AMP (cAMP) receptor protein (Crp) (D. Zheng, Constantinidou, Hobman, & Minchin, 2004). This catabolic transcription initiation regulator control over 100 genes mainly involved in the non-glucose sugar catabolism. It forms a complex with cAMP to form Crp-cAMP to bind this consensus sequence 5'-AAATGTGATCTAGATCACATTT-3' of the gene promoter (Kolb, Busby, Buc, Garges, & Adhya, 1993).

6.3. The biosynthesis

The *E. coli* molybdenum cofactor biosynthesis pathway was best studied in *E. coli* and is used as a model (Leimkühler, Wuebbens, & Rajagopalan, 2011). Hence it is the *E. coli* pathway will be presented here (Leimkühler, 2020).

6.3.1. 5'-GTP to cPMP

As mentioned before, the precursor molecule is a 5'-GTP. This nucleotide is converted in cPMP (6-alkyl pterin substituted on C2' and C4' atom with phosphate groups, Figure I-17) by two enzymes, MoaA (UniProt *E. coli* P30745) and MoaC (UniProt *E. coli* P0A738), from the *moa* operon (Wuebbens &

Rajagopalan, 1995). Thanks to the enzyme from Staphylococcus aureus, both enzymes and the reaction mechanism were studied (Hover, Lilla, & Yokoyama, 2015; Hover, Loksztejn, Ribeiro, & Yokoyama, 2013). MoaA, GTP 3',8'-cyclase (EC 4.1.99.22), is an homodimer with two [4Fe-4S] clusters per subunit belonging to the S-adenyl methionine (SAM) dependent radical enzyme. It is responsible for the conversion of the GTP to (8S)-30,8-cyclo-7,8-dihydroguanosine 50-triphosphate (3',8cH_2GTP) by the insertion of the C8 of the GTP between C2' and C3' of the ribose. It is preceded by the cleavage of the SAM radical to 5'-deoxyadesnosine-H and methionine with the transfer of the H from the C3' of the ribose (Hover et al., 2013). This intermediate compound is used as substrate by the cyclic pyranopterin monophosphate synthase MoaC (EC 4.6.1.17) catalysing the formation of the pterin ring by using phosphodiester bond cleavage as driving force. This reaction allows the formation of cPMP and the release of a pyrophosphate (Hover et al., 2015). MoaC is a trimer of homodimer (Kanaujia et al., 2010).

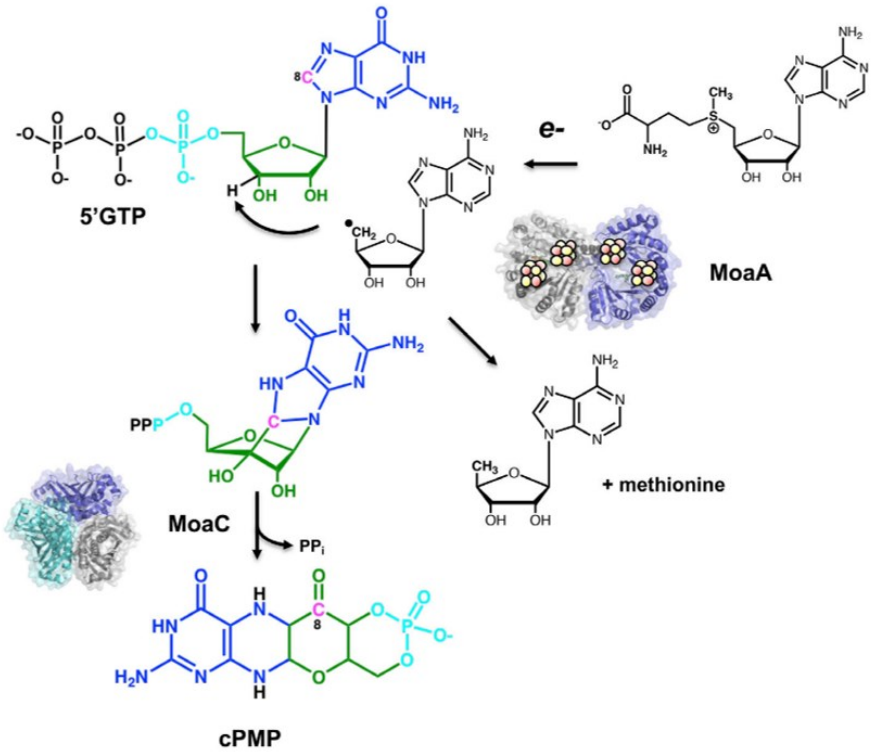


Figure I-17 Scheme summarising the conversion of 5'-GTP to cPMP catalysed by MoaA and MoaC by Leimkühler (Leimkühler, 2020). 5'-GTP: 5'-guanosine triphosphate, cPMP: cyclic pyranopterin monophosphate, e^-: electron, PPi: pyrophosphate, MoaA: (GTP 3',8'-cyclase, MoaC: cyclic pyranopterin monophosphate synthase.

6.3.2. cPMP to MPT

The MPT is synthesised from cPMP with a heterotetrameric enzymatic complex, the MPT synthase (EC 2.8.1.12) composed of two large subunits MoaE (~16,850 kDa, UniProt *E. coli* P30749) and two small subunits MoaD (~8,750 kDa, UniProt *E. coli* P30748) with the two MoaE in the centre and two thiocarboxylated MoaD binding on this centre (Pitterle & Rajagopalan, 1989). The complex possesses one active site per MoaE subunit. MoaD plays the role of the sulfur carrier thanks to its Gly-Gly C-terminal tail, which is buried in the MoaE active site (Pitterle & Rajagopalan, 1993; Schmitz, Wuebbens, Rajagopalan, & Leimkühler, 2007). The role of this complex is the catalysis of the cPMP sulfuration. The two sulfuration reactions are achieved sequentially: first in C2' and then in C1' (Hille et al., 2014). To complete the disulfuration, MoaD, which gave it sulfur in C2', comes off from the MPT synthase complex and is replaced by a thiocarboxylated MoaD. The disulfuration is followed by a dehydration to finalise the cycle formation (Figure I-18) (Daniels, Wuebbens, Rajagopalan, & Schindelin, 2008; Wuebbens & Rajagopalan, 2003).

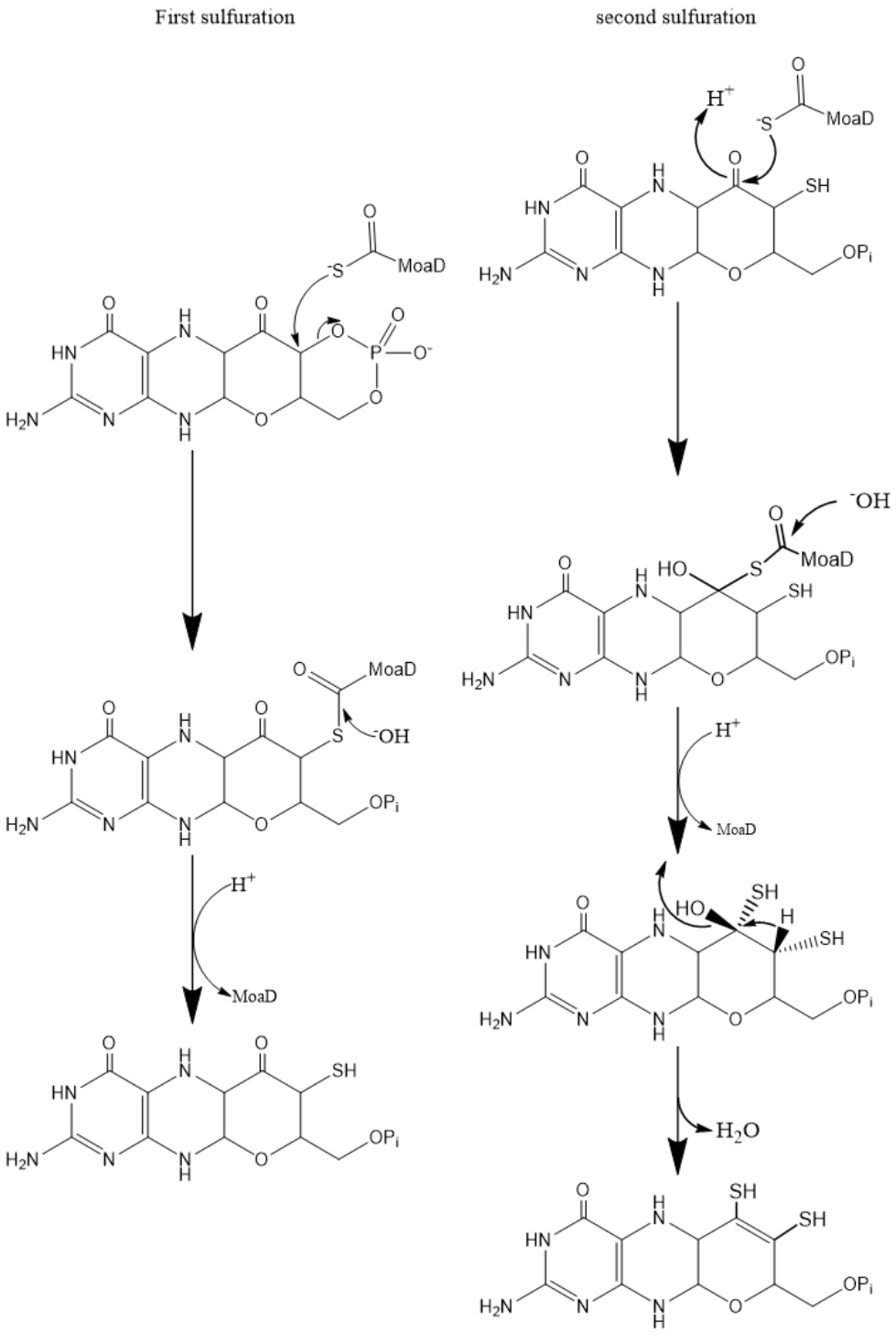


Figure I-18 synthesis mechanism of MPT from cPMP by the MPT synthase complex EC 2.8.1.12. "In the second sulfuration step (bottom), the stereochemistry of the pyran ring is inferred from the spatial disposition of cPMP relative to the C-terminus of MoaD". The figure was reproduce for visibility from Hille's article(Hille et al., 2014).

The desulfurated MoaD needs to be regenerated. The thiocarboxylation of MoaD is catalysed by MoeB (UniProt *E. coli* P12282) in two steps in a MoaD-MoeB complex (Leimkühler, Wuebbens, & Rajagopalan, 2001). This complex forms another complex with an identical dimer to form a heterotetramer $(\alpha\beta)_2$.

The first reaction is the activation (EC 2.7.7.80) by adenylation of the C-terminal Gly-Gly end of MoaD, and this reaction is dependent on Mg^{2+}·ATP. The second is the sulfur transfer (EC 2.8.1.11) to the activated MoaD by direct interaction with the sulfur transferase within the MoaD-MoeB complex leading to the release of the AMP and MoeB. MoaD-SH can then bind to the MoaE (Schmitz et al., 2007; Tong, Wuebbens, Rajagopalan, & Fitzgerald, 2005). Iobbi-Nivol and Leimkühler proposed that the persulphide from the sulfur provider attacks the activated MoaD protein-AMP bond to form perthiocarboxylate intermediate with the sulfur transferase to form the MoaD-SH by reduction cleavage (Iobbi-Nivol & Leimkühler, 2013). The sulfur is provided by the cysteine desulfurase IscS in *E. coli* which uses L-cysteine to produce L-alanine (W. Zhang et al., 2010). It was shown that at least two other proteins were involved in the sulfur transfer, TusA and YnjE (J.-U. Dahl et al., 2011; J. U. Dahl et al., 2013). TusA forms a sulfur relay between IscS and MoaD-MoeB complex with a formation of the persulphide on TusA, which will transfer it to MoaD. The rhodanese-like protein, YnjE, seems to play a similar role as TusA regarding the Moco biosynthesis pathway, but its exact role needs further investigation. TusA is not expressed under anaerobic conditions, and the YnjE expression is increased under this condition. It places YnjE as the first choice to interact with IscS without oxygen. Interestingly it was shown that SufS could not replace IscS as sulfur provider for the Moco biosynthesis pathway due to its inability to interact with TusA (Leimkühler, Bühning, & Beilschmidt, 2017). Then, the sulfurated MoaD is released from the complex to bind MoaE and reform an active MPT synthase complex. It is important to note that the role of YnjE as a potential replacement for TusA was not clearly mentioned in the following article/reviews of the Leimkühler group (Leimkühler, 2020; Mendel, Hercher, Zupok, Hasnat, & Leimkühler, 2020).

6.3.3. MPT adenylation and Mo/W incorporation

A couple of enzymes catalyses the insertion of the Molybdenum atom: the molybdopterin adenylyltransferase MogA (EC 2.7.7.75, UniProt *E. coli* P0AF03) (Liu, Wuebbens, Rajagopalan, & Schindelin, 2000) and the molybdopterin

molybdotransferase MoeA (EC 2.10.1.1, UniProt *E. coli* P12281) (J. Nichols & Rajagopalan, 2002). MogA forms a homotrimeric complex and is responsible for the MPT's activation by adding the AMP to the phosphate moiety. This reaction is Mg^{2+}·ATP dependent (Kuper, Llamas, Hecht, Mendel, & Schwarz, 2004). Then the Mo atom is added to the MPT-AMP from cytoplasmic molybdate by MoeA with the release of an AMP (Jason D Nichols & Rajagopalan, 2005), and MoeA forms a homodimer (Xiang, Nichols, Rajagopalan, & Schindelin, 2001). A MoeA monomer is divided into four domains, and the dimer's crystallographic structure (Figure I-19 shows that the domain II of one monomer faces the domains III and IV of the other monomers. The putative active site is positioned in a cleft between domain II of the first subunit and domains III/IV of the second subunit (Xiang et al., 2001). The adenylation of the MPT is useful when the concentration of the Molybdate/tungstate is physiological (1 – 10 µM). At higher concentrations, the MoeA alone can incorporate the metal (Hille et al., 2014). Over 1 mM, neither enzyme was required to produce the molybdenum cofactor (Neumann & Leimkühler, 2008).

As presented above, the enzymes from the SO family incorporate the end-product of the MoeA, which is then coordinated by a Cys to form the SO protein as an MPT-$Mo^{VI}O_2$-Cys (Hille et al., 2014). In *E. coli,* MsrP is the only Mo-MPT-containing SO enzyme yet characterised (Gennaris et al., 2015; Leimkühler, 2020).

It is important to note that, due to its significant instability, Moco must be kept bound to the enzymes until the incorporation of Molybdopterin in the Mo-enzymes. It has been shown that there are some interactions between all the enzymes involved in the incorporation of Mo and the final maturation, MogA-MoeA and MoeA-MobA. These enzymes are involved in subsequent modifications of the Mo-MPT and forms a big complex (A. Magalon, Frixon, Pommier, Giordano, & Blasco, 2002; Vergnes, Gouffi-Belhabich, Blasco, Giordano, & Magalon, 2004).

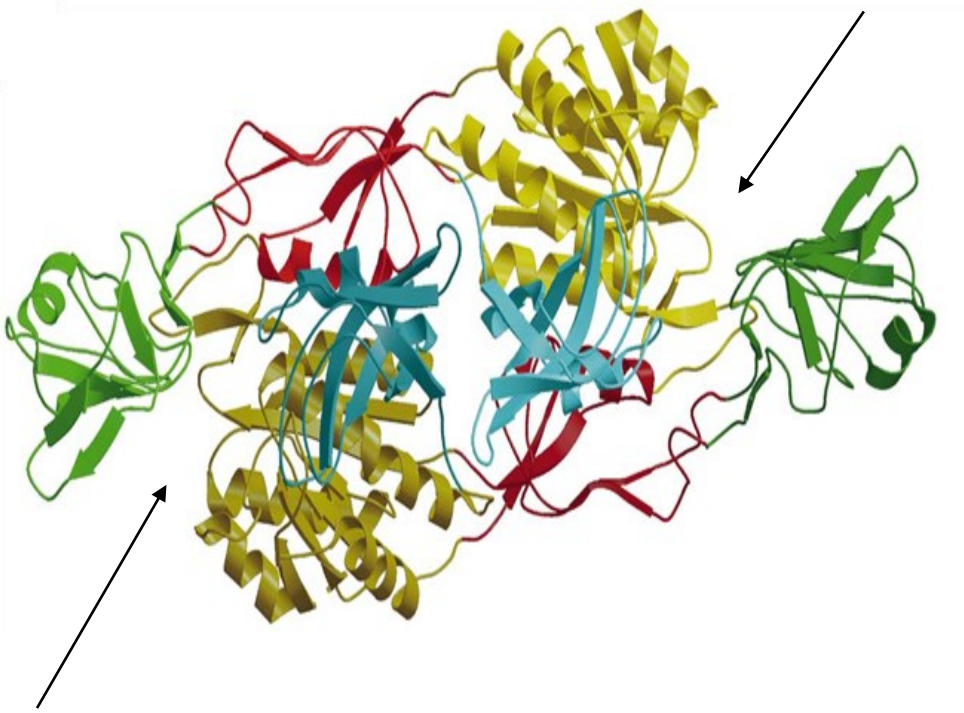


Figure I-19 E. coli MoeA dimer , domain I (red) (residues 27-53, 140-175) , domain II (green) (residues 53-140), domain III (yellow) (residues 7-27, 176-323) and domain IV (bleu) (324-408) (Xiang et al., 2001). The putative active sites are pointed with arrows.

6.3.4. Nucleotide addition and maturation

After the Mo/W transfer on the MPT cofactor, more modifications are needed so that it can be used by most of the Mo-dependent proteins in *E. coli* (Leimkühler, 2020).

The formation of the Mo-MGD is catalysed by the molybdenum cofactor guanylyltransferase MobA (EC 2.7.7.77, UniProt *E. coli* P32173) and MobB (UniProt *E. coli* P32125), a GTP binding protein (McLuskey, Harrison, Schüttelkopf, Boxer, & Hunter, 2003). The latter is not essential for the formation of Mo-bisMGD, but its presence enhances the activity of MobA. It has been shown that the dimeric MobB possesses a GTP binding site, but it has also been demonstrated that MobA can bind GTP by itself (Temple & Rajagopalan, 2000). Indeed, thanks to the crystal structure of MobA, two conserved binding sites have been observed: one probably for the pyranopterin and the other for Mg-GTP, as suggest the Rossman-fold in N-terminals. Moreover, the docking modeling between MobA and MobB shows a shared binding site for GTP (Reschke et al., 2013). The reaction catalysed by

MobA happens in two steps. The first is the formation of the intermediary Mo-*bis*MPT from two Mo-MPT or one Mo-MPT and one MPT. If it is the first hypothesis, a Molybdate is released. It has been suggested that both Mo-MPT/MPT bind together on a monomeric MobA in the respective GTP and MPT binding sites. In fact, the apo-cofactor might probably bind the nucleotide binding region, GTP playing the role of precursor. The second step is the addition of the GMP on both phosphate moieties of the Mo-bisMPT. These GTPs come from cytoplasmic Mg^{2+}·GTP carried by MobB. The nucleotide addition is followed by the release of two pyrophosphates. However, the catalysis mechanism of Mo-bisMPT formation is not completely solved so far (Mendel & Leimkühler, 2015; Reschke et al., 2013). Without further modification, this cofactor can be directly inserted in some DMSO family proteins.

As presented in the previous section, AORs contain a specific W-bisMPT cofactor. This type of cofactor, as observed by Reschke and co-workers, is an intermediate compound used by MobA during the synthesis of Mo-bisMGD (Reschke et al., 2013). Recently a protein incorporating a similar cofactor was found in *E. coli*: YdhV (Reschke et al., 2019). The biosynthesis pathway of the Mo-bisMPT cofactor is unknown in *E. coli* (Leimkühler, 2020).

The formation of Mo-MCD is catalysed by the molybdenum cofactor cytidylyltransferase MocA (EC 2.7.7.76, UniProt *E. coli* Q46810). The CMP addition from Mg-CTP on the phosphate moiety of the Mo-MPT is identical to the reaction catalysed by MobA but with a high affinity for CTP. Indeed, this protein has some amino acid sequence similarities to MobA with 22% of identities but cannot to guide the formation of Mo-bisMCD as MobA with Mo-bisMGD (Neumann, Mittelstadt, Seduk, Iobbi-Nivol, & Leimkuhler, 2009).

Some enzymes from the DMSO reductase family and all the enzymes from the XO family require sulfurated molybdenum cofactor (Hille et al., 2014). The Mo-bisMGD are sulfurated by molecular chaperone enzymes such as FdhD by binding the Mo-bisMGD from MobA and uses the previously seen L-cysteine desulfurase IscS to provide sulfur before inserting the sulfurated cofactor in the

apo-protein (Arnoux et al., 2015). The mechanism is almost the same for the sulfuration of Mo-PCD. It was found that PaoD from the XdhC-family of chaperones (Iobbi-Nivol & Leimkühler, 2013) was binding the Mo-MCD to protect it from the oxidation while helping for sulfuration before inserting the cofactor in the apo-protein (Neumann, Stocklein, & Leimkuhler, 2007; Neumann, Stöcklein, Walburger, Magalon, & Leimkühler, 2007).

Some molybdenum enzymes need the help of chaperone proteins to insert the molybdenum cofactor into the apo-molybdo-enzymes (Axel Magalon et al., 2011). In *E. coli,* it was found that the majority of DMSOR family members have their own chaperone protein (NarJ, NarW) (Francis Blasco et al., 1998; F. Blasco, Pommier, Augier, Chippaux, & Giordano, 1992) which as presented above can also carry the Moco maturation (NapD, TorD) (Iobbi-Nivol & Leimkühler, 2013). On the contrary, several proteins use FdhD and DsmD as chaperone and cofactor sulfuration enzymes (Ray, Oates, Turner, & Robinson, 2003; Thomé et al., 2012). Moreover, many Mo-enzymes are membrane proteins that must be led to the membrane. Those molecular chaperones usually belong to the TorD family and protect the MoCo from oxidation (Genest, Méjean, & Iobbi-Nivol, 2009). For instance, TorD prevents the translocation of TorA before the insertion of the cofactor into the enzyme. However, MobA can carry the molybdenum cofactor to the enzymes without the help of a chaperone, in *Rhodobacter sphaeroides,* for instance (Hille et al., 2014).

6.4. Regulation

The ModE was presented earlier as a repressor of the ModABC transporter expression when molybdates bound ModE, which increased its affinity for the operator DNA (Grunden et al., 1996; Grunden & Shanmugam, 1997). The binding of ModE on other operators also activates the expression of many operons coding for molybdenum cofactor synthesis, notably *moaABCDE,* which code for the first step of MoCo biosynthesis. Hence, the presence of Mo in the cell increases the production of molybdopterin (L. A. Anderson, McNairn, Lubke, Pau, & Boxer, 2000). Hence *moa* is regulated by the availability of molybdate in the cell (Zupok, Iobbi-Nivol, et al., 2019).

Other regulators are also involved in the regulation of genes involved in molybdenum cofactor synthesis (Zupok, Iobbi-Nivol, et al., 2019).

The *moaABCDE* transcription is upregulated by the fumarate nitrate regulator (FNR), a dual transcriptional regulator possessing a [4Fe-4S] cluster. This cluster function as an oxygen sensor: when intact (under anaerobiosis), the FNR can bind to the promoter and upregulates the transcription of genes involved in anaerobic respiration and downregulates genes involved in aerobic respiration. In presence of oxygen, the iron-sulfur cluster is degraded in two steps: [3Fe-4S], which improves hydrogen peroxide production, accelerating the degradation to [2Fe-2S] state leading to the loss of DNA binding activity. FNR up regulate the expression of *moaABCDE* (L. A. Anderson et al., 2000). Moreover, it was observed that the *moaA* was repressed by copper thanks to CueR and that *moaB* has a copper inducible transcription promoter (Yamamoto & Ishihama, 2005)

Two posttranslational regulations were observed (Patterson-Fortin, Vakulskas, Yakhnin, Babitzke, & Romeo, 2013; Regulski et al., 2008). It was shown that CsrA, the carbon storage regulator, a global transcription regulator repressing stationary phase metabolism and activating the central carbon metabolism (Sabnis, Yang, & Romeo, 1995), was activating the transcription of *moa* operon (Patterson-Fortin et al., 2013). This mechanism would ensure that the cofactor is produced under high-demand conditions (Leimkühler, 2020; Zupok, Iobbi-Nivol, et al., 2019). On the contrary, the molybdenum cofactor is down-regulating the moa operon's expression, preventing the cofactor's overexpression (L. A. Anderson et al., 2000). The Moco-sensing riboswitch represses the translation of MobA when Moco binds the mRNA. In *E. coli,* a 138 bp RNA sequence forms five stem loops preventing the RBS from being accessible by the ribosome (Regulski et al., 2008). This prevents cofactor overproduction by stopping the translation of the first enzyme involved in the cofactor biosynthesis pathway. However the repression is lifted when Wco was produced instead of Moco, suggesting that this riboswitch is specific toward Moco (Leimkühler, 2020; Regulski et al., 2008; Zupok, Iobbi-Nivol, et al., 2019).

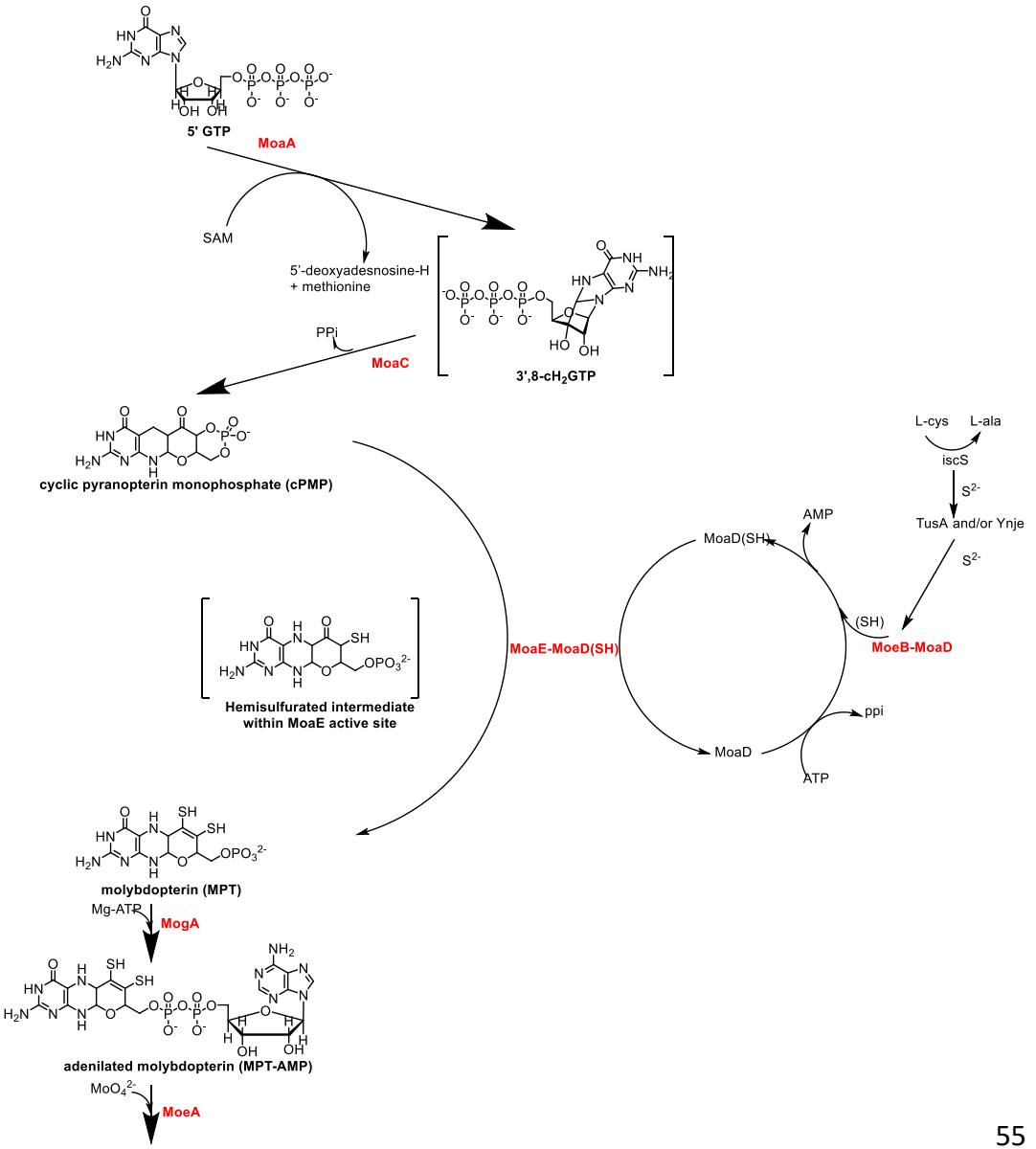
MoeAB basal expression is regulated by two global regulators, FNR and ArcA, with the latter activating the expression and the former repressing the activity (Hasona, Self, & Shanmugam, 2001). ArcA is the cytoplasmic regulatory component of the aerobic respiration control system, with ArcB being a transmembrane sensor detecting anaerobiosis/microaerobiosis. When ArcB senses these conditions, it is autophosphorylated and activate ArcA by phosphorylation, then playing its regulation role (S. Iuchi & Lin, 1988; S Iuchi & Lin, 1992a, 1992b). Under this condition it was found that ArcA binds the promoter region of *moeAB* and upregulates the gene expression (Hasona et al., 2001). Moreover, the formerly presented FNR is, here, downregulating the expression of the *moe* operon (L. A. Anderson et al., 2000; Hasona et al., 2001). Moreover, both regulators increase the expression of *arcA* under anaerobic conditions (Compan & Touati, 1994). One last upregulation is with NarL. This protein is part of the nitrate respiration regulation system, and coupled with NarX, can sense NO_2 and/or NO_3, which leads to the phosphorylation of NarL to activate it (Rabin & Stewart, 1993). In the presence of nitrate, the expression of the cofactor is increased as the enzyme involved in the nitrate respiration contains the cofactor (Hasona et al., 2001). It was observed that FNR was repressing the NarL expression (Myers et al., 2013). According to Hasona and co-workers, the expression of *moeAB* is controlled in order to express these genes at three different levels: a low expression level under aerobic conditions when the expression is not upregulated; an intermediary level modulated by ArcA and FNR; and a high expression level under anaerobic conditions and during nitrate respiration under the control of ArcA-P and NarL-P (Hasona et al., 2001).

In addition, the cofactor biosynthesis also seems to depend on iron availability. Indeed, several proteins involved in the biosynthesis contain [4Fe-4S] cluster (MoaA, FNR) or are dependent on cell iron availability for its expression, like IscS. Under iron-limiting conditions, the cellular molybdenum cofactor content was decreased due to inactive MoaA, no upregulation with active FNR and lack of expressed IscS (Zupok, Gorka, Siemiatkowska, Skirycz, & Leimkühler, 2019).

Overall, in *E. coli,* the molybdenum cofactor biosynthesis pathway is regulated by the Mo cellular content (ModE), the anaerobiosis (FNR, ArcA), the iron availability (FNR, MoaA, IscS), the presence of nitrate for nitrate respiration (NapL) and central carbon metabolism regulation (CsrA) (Leimkühler, 2020; Zupok, Iobbi-Nivol, et al., 2019).

6.5. Mo/W discrimination

It is still unsure how Mo or W are selected by the cell or proteins. Some organisms only use W, such as *P. furiosus*, others only Mo, like *E. coli*. However, it was found that some organisms, *E. limosum* and *A. mobile*, were strongly biased toward W, despite containing Mo enzymes and ModABC transporter in addition to TupABC (Schut et al., 2021). Other organisms, such as *A. aromaticum* or *Geobacter metallireducens,* contain both Mo and W cofactor enzymes (Arndt et al., 2019; Wischgoll et al., 2005). In nature, Mo is more available than W. indeed seawater Mo concentration is 500,000-fold higher than W (ca. 9 to 13.5 ppm for ca. $2x10^{-3}$ ppm), in soil 3 to 10-fold higher and in freshwater 2 to 5-fold higher where the W concentration does not exceed 0.5 nM (Kietzin & Adams, 1996). Several hypotheses were raised to understand how the organisms could discriminate between the two elements (Seelmann et al., 2020). The first hypothesis is the selection only from the uptake, thanks to the higher specificity of the Tup and Wtp transporter toward W. In comparison, the Mod transporter is not as specific toward Mo (Hagen, 2011). The second hypothesis is a selection at the Mo/W insertion level with MoeAs. In tungsten-specific or mixed organisms, it was observed that more than one MoeA was produced (Seelmann et al., 2020). It was observed that some MoeAs encoding genes were placed close to genes i) involved in W transport, ii) W-containing enzyme or iii) involved in the biosynthesis pathway, but it does not appear to be systematic. It has been proposed that some MoeA could be specific to molybdate or tungstate. It seems that MobA is not specific to one or the other but that the formation of the MoeA-MobA complex would give the specificity. However, more experiments are needed to confirm these hypotheses (Seelmann et al., 2020).


5' GTP
MoaA
SAM
5'-deoxyadesnosine-H
+ methionine
PPi
MoaC
3',8-cH2GTP
cyclic pyranopterin monophosphate (cPMP)
L-cys
L-ala
iscS
S2-
TusA and/or Ynje
S2-
AMP
MoaD(SH)
(SH)
MoeB-MoaD
MoaE-MoaD(SH)
Hemisulfurated intermediate
within MoaE active site
MoaD
ppi
ATP
molybdopterin (MPT)
Mg-ATP
MogA
adenilated molybdopterin (MPT-AMP)
MoeA

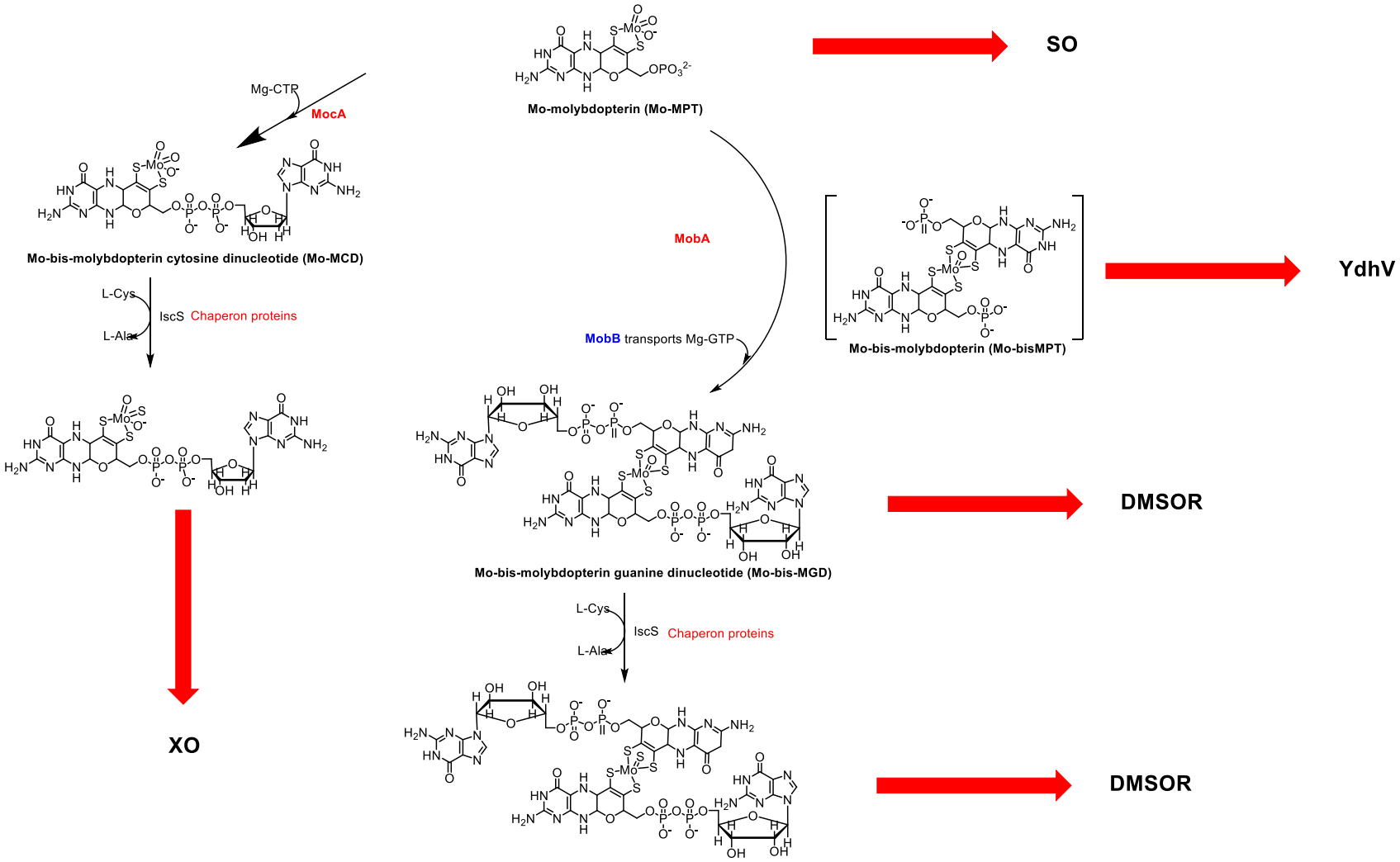

Mg-CTP
MocA
Mo-bis-molybdopterin cytosine dinucleotide (Mo-MCD)
L-Cys
L-Ala
IscS
Chaperon proteins
XO
Mo-molybdopterin (Mo-MPT)
SO
MobA
MobB transports Mg-GTP
Mo-bis-molybdopterin (Mo-bisMPT)
YdhV
Mo-bis-molybdopterin guanine dinucleotide (Mo-bis-MGD)
DMSOR
L-Cys
L-Ala
IscS
Chaperon proteins
DMSOR

Figure I-20 Scheme of the molybdenum cofactor biosynthesis pathway in E. coli. In red enzymes involved in the formation and maturation.MoaA: GTP 3',8'-cyclase (EC 4.1.99.22, UniProt P30745) , MoaC: cyclic pyranopterin monophosphate synthase (EC 4.6.1.17, UniProt P0A738), MoaE: molybdopterin synthase catalytic subunit (EC 2.8.1.12, UniProt P30749), MoaD: molybdopterin synthase sulfur carrier subunit (EC 2.8.1.12, UniProt P30748), MoeB: molybdopterin-synthase adenylyltransferase (EC 2.7.7.80, UniProt P12282), MogA: molybdopterin adenylyltransferase (EC 2.7.7.75, UniProt P0AF03), MoeA: molybdopterin molybdotransferase (EC 2.10.1.1, UniProt P12281), MobA: molybdenum cofactor guanylyltransferase (EC 2.7.7.77, UniProt P32173), MobB: GTP binding protein UniProt P32125), IscS: cysteine desulfurase (EC 2.8.1.7, UniProt P0A6B7), TusA: sulfur transfer protein (EC 2.8.1.-, UniProt P0A890) MocA: molybdenum cofactor cytidylyltransferase (EC 2.7.7.76, UniProt Q46810) , SO: Sulfite oxidase , YdhV: newly discovered Mo-bis PPT containing membrane enzyme putative oxidoreductase (UniProt P76192) , DMSOR: dimethylsulfoxide oxidoreductase, XO: xanthine oxidase .

7. Research project

There is an increasing interest in producing chemicals from cheap substrates like CO. As presented above, CO can be produced by gasification from a wide variety of biomass sources including municipal wastes. CO is used by a wide range of aerobic and anaerobic archaea and bacteria as a source of carbon and energy. These organisms can convert CO to CO_2 and low potential electrons (E^0 ~-520 mV carried by ferredoxin) using the CODH enzyme. These organisms are fixing CO_2 either through the CBB cycle or the Wood-Ljungdahl pathway.

It was hypothesised that the heterologous co-expression of CODH, GAPOR, and ferredoxin genes could allow *E. coli* to grow on CO as the sole carbon and energy source by fixing the CO2 produced with a synthetic CBB cycle (Figure I-21) (Niv Antonovsky, Shmuel Gleizer, & Ron Milo, 2017). CO would be oxidised with a heterologous CODH (Carlson & Papoutsakis, 2017) producing the production CO_2 and low potential reduced ferredoxin. The CO_2 would be fixed by a synthetic CBB cycle (Niv Antonovsky et al., 2017; Gleizer et al., 2019). In this cycle, it was planned to replace the GAPDH and PGK by the GAPOR from *M. maripaludis* (Park et al., 2007) functioning in the reverse direction thanks to the low potential reduced ferredoxin. If this pathway is functional, it would theoretically (4) save 6 ATP and 6 NADPH compared to the normal CBB reaction (5).

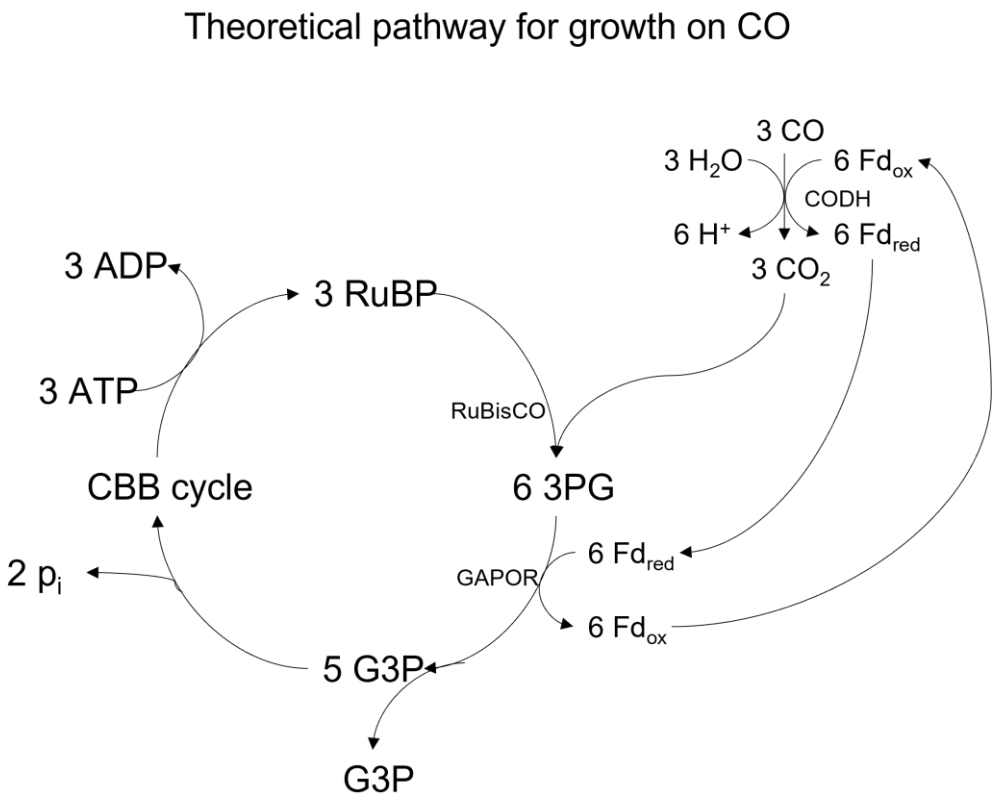


Figure I-21 Theoretical pathway for CO fixation in a modified Calvin-Benson-Bassham (CBB) cycle. RuBP: ribulose-1,5-bisphosphate, G3P: glyceraldehyde-3-phosphate, 3PG: 3-phosphoglycereate, Fd: ferredoxin in both reduced and oxidised form, CODH: carbon monoxide dehydrogenase, GAPOR: glyceraldehyde-3-phosphate: ferredoxin oxidoreductase, RuBisCo: Ribulose-1,5-bisphosphate carboxylase-oxygenase. The detailed CBB cycle is presented on Figure I-2.

Equation of CBB cycle (Niv Antonovsky et al., 2017; Dijkhuizen & Harder, 1984):

$$3\,CO_2 + 9\,ATP + 6\,NADPH \rightarrow G3P + 9\,ADP + NAD + 8\,p_i \qquad (4)$$

Theoretical equation of the modified CBB cycle from CO:

$$3CO + 3H_2O + 3\,ATP \rightarrow G3P + 3ADP + 2\,p_i + 6H^+ \qquad (5)$$

Herein are presented the results of the first step of this ambitious project: Is it possible to express the gorS2 gene from M. maripaludis to produce a functional GAPOR in E. coli by reproducing the Park experiment? Also, with the objective of evaluating the reverse reaction when low-potential ferredoxin is generated from the oxidation of CO by a CODH. The GAPOR contain a specific cofactor, and after encountering problems, it was supposed that expressing the correct cofactor could help to obtain an active enzyme. Hence, it was decided to characterise the key genes involved in the

molybdopterin biosynthesis pathway of M. maripaludis by an in vivo complementation approach in E. coli. As this approach was also unsuccessful, it was hypothesised that expressing the gene in AOR-expressing bacteria would help produce active enzymes. Hence, it was finally decided to express both the GAPOR and the GOR (a recently discovered new enzyme from bacteria that catalyse the GAPOR reaction) encoding genes in two anaerobic bacteria C. autoethanogenum and C. acetobutylicum known to produce the right molybdenum cofactor needed to heterologously produce functional GAPOR and a GOR.

Chapter II:

Materials and Methods

1. Strains and plasmids

The strains and plasmids used in this study are listed in Table II-1 and Table II-2.

Table II-1. Bacterial strains used in this study

***E. coli* Strain**	**Characteristics**	**Origin**
DH5α	F^- endA1 glnV44 thi-1 recA1 relA1 gyrA96 deoR nupG purB20 ϕ80dlacZΔM15 Δ(lacZYA-argF)U169, hsdR17($r_K^-m_K^+$), λ^-	NEB
TOP10	F- mcrA Δ(mrr-hsdRMS-mcrBC) ϕ80lacZΔM15 ΔlacX74 nupG recA1 araD139 Δ(ara-leu)7697 galE15 galK16 rpsL(Str^R) endA1 λ^-	Invitrogen
RG_WT	E. coli rosetta-gami 2(DE3): Δ(ara-leu)7697 ΔlacX74 ΔphoA Pvu II phoR araD139 ahpC gale galK rspL (DE3) F'[lac^+ $lacI^q$ pro] gor522::Tn 10 trxB pRARE2(Cam^R, Str^R, Tet^R)	Novagen
RG	*E. coli* rosetta-gami 2(DE3) pET28a(+) *gorS2* (Kan^R)	(Guerrini, 2007)[a]
RGP	*E. coli* rosetta-gami 2(DE3) pET28a(+) (Kan^R)	This study
RGΔmobAB	*E. coli* rosetta-gami 2(DE3) Δ*mobAB::kan*^{R} pET28a(+) *gorS2* (*Amp*^{R})	(Guerrini, 2007)[a]
RGΔiscR	*E. coli* rosetta-gami 2(DE3) Δ*iscR*	This study
RGΔiscR_gapor	*E. coli* rosetta-gami 2(DE3) Δ*iscR* pET28a(+) *gorS2* (Kan^R)	This study
RG_AM020::kan	E. coli rosetta-gami 2(DE3) ΔiscRΔhypF::kan^R	This study
RG_AM021	*E. coli* rosetta-gami 2(DE3) Δ*iscRΔmoeA*	This study
RG_AM022::kan	*E. coli* rosetta-gami 2(DE3) ΔiscR ΔmoeA *ΔhypF::kan*^{R}	This study
RG_AM023	E. coli rosetta-gami 2(DE3) *ΔiscRΔhypFΔselA*	This study
RG_AM023gapor	RG_AM023 pET28a(+) *gorS2* (*Amp*^{R})	This study
RG_AM024	E. coli rosetta-gami 2(DE3) *ΔiscR ΔmoeA ΔhypF ΔselA*	This study
RG_AM024gapor	RG_AM024 pET28a(+) *gorS2* (*Amp*^{R})	This study

RG_AM025	*E. coli* rosetta-gami 2(DE3) *ΔiscR ΔhypF ΔselA ΔmobAB::kan*^{R}	This study
RG_AM025gapor	RG_AM025 pET28a(+) *gorS2* (*Amp*^{R})	This study
RG_AM026	E. coli rosetta-gami 2(DE3) ΔiscRΔmoeAΔhypFΔselAΔmobAB::kan^R	This study
RG_AM025gapor	RG_AM026 pET28a(+) *gorS2* (*Amp*^{R})	This study
sExpress	NEB Express strain [*fhuA2 [lon] ompTgal sulA11 R(mcr- 73:miniTn10–TetS)2 [dcm] R(zgb-210:Tn10–TetS) endA1 D(mcrC-mrr*)114:IS10] transformed with R702 conjugal transfer plasmid from E. coli CA434 (Cam^R)	(Woods et al., 2019)
Pan2	*E. coli* TOP10 transformed with pAN2 *C. acetobutylicum* methylation plasmid (tet^R)	(Heap, Pennington, Cartman, Carter, & Minton, 2007)
***E. coli* mini-library strains**		
RG_AM001	RG_AM025 pAM_001	This study
RG_AM002	RG_AM025 pAM_002	This study
RG_AM003	RG_AM024 pAM_003	This study
RG_AM004	RG_AM024 pAM_004	This study
RG_AM005	RG_AM024 pAM_005	This study
RG_AM006	RG_AM024 pAM_006	This study
RG_AM007	RG_AM025 pAM_007	This study
RG_AM008	RG_AM026 pAM_008	This study
RG_AM009	RG_AM026 pAM_009	This study
RG_AM010	RG_AM026 pAM_010	This study
RG_AM011	RG_AM026 pAM_011	This study
RG_AM012	RG_AM026 pAM_012	This study
RG_AM013	RG_AM026 pAM_013	This study
RG_AM014	RG_AM026 pAM_014	This study

RG_AM015	RG_AM026 pAM_015	This study
RG_AM016	RG_AM026 pAM_016	This study
RG_AM017	RG_AM026 pAM_017	This study
RG_AM018	RG_AM026 pAM_018	This study
RG_AM019	RG_AM026 pAM_019	This study
Clostridium autoethanogenum strains		
C. autoethanogenum DSM10061	Wild type	(Abrini, Naveau, & Nyns, 1994)
C24	*C. autoethanogenum* DSM10061 with *bgaR*-P_{gaL}::*tcdR* inserted at *pyrE* loccus	(Woods et al., 2022)
CA_AM001	*C. autoethanogenum* DSM10061 pMTL83151_AM001	This study
CA_AM002	*C. autoethanogenum* DSM10061 pMTL83151_AM002	This study
CA_AM003	*C. autoethanogenum* DSM10061 pMTL83151_AM003	This study
CA_AM004	C. autoethanogenum C24 pMTL83151_AM004	This study
CA_AM005	*C. autoethanogenum* DSM10061 pMTL83151_AM005	This study
CA_AM005.1	*C. autoethanogenum* DSM10061 pMTL83151_AM005.1	This study
CA_AM006	*C. autoethanogenum* DSM10061 pMTL83151_AM006	This study
CA_AM007	C. autoethanogenum C24 pMTL83151_AM007	This study
Clostridium acetobutylicum strains		
MGCΔcac1502 Δupp	*C. acetobutylicum ATCC 824* WT *ΔCA_C1502 Δupp* this strain was named ATCC 824 in the report as actual WT was not used	(Croux et al., 2016)
C. acetobutylicum tcdr+	*C. acetobutylicum* ATCC824 WT with *bgaR*-P_{gaL}::*tcdR* inserted at *pyrE* loccus	Omorotionmwan, 2022, unpublished/ under review
CAC_AM001	C. acetobutylicum MGCΔcac1502Δupp pMTL83151_AM001	This study
CAC_AM003	C. acetobutylicum MGCΔcac1502Δupp pMTL83151_AM003	This study

CAC_AM004	C. acetobutylicum *tcdR*$^{+}$ pMTL83151_AM004	This study
CAC_AM005	C. acetobutylicum MGCΔcac1502Δupp pMTL83151_AM005	This study
CAC_AM005.1	C. acetobutylicum MGCΔcac1502Δupp pMTL83151_AM005.1	This study
CAC_AM006	C. acetobutylicum MGCΔcac1502Δupp pMTL83151_AM006	This study
CAC_AM007	C. acetobutylicum *tcdR*$^{+}$ pMTL83151_AM007	This study

Table II-2 Plasmids used in this study

Plasmids	**Characteristics**	**Origin**
pET28a(+)	T7 promoter *lacO* T7 transcription start His-tag T7-tag MCS T7 terminator *lacI* pBR322 *ori* Kan^{R} f1 ori	Novagen
pET28a(+) gorS2 (Kan^{R})	pET28a(+) with *gorS2* insert, N-term his-tag, Kan^{R}	Soucaille, unpublished [a]
pET28a(+) gorS2 (Amp^{R})	pET28a(+) with *gorS2*, N-term His-tag, Amp^{R} inserted within Kan^{R}	Soucaille, unpublished [a]
pCP20	FLP^{+}, λ ci857^{+}, λ p_R Rep^{ts} , Amp^{R}, Cam^{R}	(Cherepanov & Wackernagel, 1995)
pRARE2	*E. coli* rare tRNAs Cam^{R}, Str^{R}, Tet^{R}	Novagen
2_Pfdx_Rb3_pUC57	pUC57 backbone carrying the synthesised P*fdx_Rb3*, Amp^{R}	GenScript Biothech (Netherlands) B.V.
4_PtcdB_pUC57	pUC57 backbone carrying the synthesised P*tcdB*, Amp^{R}	GenScript Biothech (Netherlands) B.V.
20ACIQFP_gorS-6hisL_for_synth_pMK-T	Plasmids carrying the synthesised *gorSL* genes with the his-tag sequence uptstream	Geneart, Thermofisher
Plasmid mini library for *gorS2 (GAPOR gene)* and *M. maripaludis S2* molybdenum cofactor biosynthesis enzyme expression		
pAM_001	pET28a (+) *gor*S2*mobA* (Amp^{R})	This study

pAM_002	pET28a (+) *gor*S2 *mobB* (Amp^R)	This study
pAM_003	pET28a (+) *gor*S2*moeA1* (Amp^R)	This study
pAM_004	pET28a (+) *gor*S2 *moeA2* (Amp^R)	This study
pAM_005	pET28a (+) *gor*S2 *moeA3* (Amp^R)	This study
pAM_006	pET28a (+) *gor*S2 *moeA4* (Amp^R)	This study
pAM_007	pET28a (+) *gor*S2*mobA mobB* (Amp^R)	This study
pAM_008	pET28a (+) *gor*S2 *mobA moeA1* (Amp^R)	This study
pAM_009	pET28a (+) *gor*S2 *mobA moeA2* (Amp^R)	This study
pAM_010	pET28a (+) *gor*S2 *mobA moeA3* (Amp^R)	This study
pAM_011	pET28a (+) *gor*S2 *mobA moeA4* (Amp^R)	This study
pAM_012	pET28a (+) *gor*S2 *mobB moeA1* (Amp^R)	This study
pAM_013	pET28a (+) *gor*S2 *mobB moeA2* (Amp^R)	This study
pAM_014	pET28a (+) *gor*S2 *mobB moeA3* (Amp^R)	This study
pAM_015	pET28a (+) *gor*S2 *mobB moeA4* (Amp^R)	This study
pAM_016	pET28a (+) gorS2 *mobA mobB moeA1* (Amp^R)	This study
pAM_017	pET28a (+) gorS2 *mobA mobB moeA2* (Amp^R)	This study
pAM_018	pET28a (+) gorS2 *mobA mobB moeA3* (Amp^R)	This study
pAM_019	pET28a (+) gorS2 *mobA mobB moeA4* (Amp^R)	This study
Plasmid for *Clostridium* expression		
pMTL83151	Modular plasmid for Clostridium expression pCB102 *catP* ColE1 + tra MCS	(Heap, Pennington, Cartman, & Minton, 2009)
pMTL83151_AM001	Derived from pMTL83151 P*hydA hydA-RBS gorS2*	This study
pMTL83151_AM002	Derived from pMTL83151 P*thl thl-RBS gorS2*	This study
pMTL83151_AM003	Derived from pMTL83151 P*fdx* synthetic riboswitch[b] *gorS2*	This study

pMTL83151_AM004	Derived from pMTL83151 P*tcdB tcdB-RBS*[c] *gorS2*	This study
pMTL83151_AM005	Derived from pMTL83151 P*hydA hydARBS gor*	This study
pMTL83151_AM005.1	Derived from pMTL83151_AM005 *gorS-RBS*[d]	This study
pMTL83151_AM006	Derived from pMTL83151 P*fdx* synthetic riboswitch[b] *gorS2*	This study
pMTL83151_AM007	Derived from pMTL83151 P*tcdB tcdB-RBS*[c] *gorS2*	This study

[a] These strains and plasmids were given by Philippe Soucaille, INSA Toulouse, France
[b] synthetic RBS F from (Cañadas, Groothuis, Zygouropoulou, Rodrigues, & Minton, 2019)
[c] from *C. difficile* (Dupuy & Matamouros, 2006)
[d] from G. ferrireducens

2. Bacterial cultures and media

2.1. *E. coli*

2.1.1. Routine culture and Storage

E. coli strains were routinely cultured in LB media at 37°C. Antibiotics were appropriately used with the following concentration: Kanamycin 50 µg/ml, carbenicillin 100 µg/ml, and Chloramphenicol 30 µg/ml.

E. coli cells were stored at -80°C in 20 % glycerol in screw-caped tubes.

2.1.2. Culture for protein production

For GAPOR/GOR expression in *E. coli* RG strains, the cells were revived in 5 mL MAC media overnight at 37°C complemented with carbenicillin 100 µg/ml and Chloramphenicol 30 µg/ml and 100 µM sodium molybdate ($NaMoO_4$).

For 300 ml protein expression culture, a pre-culture was performed in Flask in 100 ml MAC media at 37°C. The cultures were shaken at 180 rpm overnight.

The protein expression cultures were performed under microaerobic conditions. The cultures were performed in 50 ml MAC in a 120 ml serum bottle and 300 ml MAC in a 600 ml serum bottle. The media was complemented with 100 µM molybdate. The cultures were inoculated at 0.3 $OD_{600\ nm}$ from either reviving culture for 50 ml or flask preculture for 300 ml culture. The flasks were

sealed with rubber stoppers to ensure no gas exchange with the outside. After inoculation the cultures were incubated for 8h at either room temperature (RT°) or 37°C. The cultures were induced with 100 µM Isopropyl β-d-1-thiogalactopyranoside (IPTG) and incubated for 16h at RT°, testes were also performed at 37°C.

Antibiotics were appropriately used with the following concentration: Carbenicillin 100 µg/ml, Chloramphenicol 30 µg/ml.

2.1.3. Media

Luria-Bertani (LB) Broth, miller:

Tryptone	10 g/L
Yeast extract	5 g/L
NaCl	10 g/L

The media was prepared by the lab technician from a pre-mixed powder. 15 g/L of technical Agar n°2 (Oxoid) was added by the lab technician to prepare solid media.

SOC media:

Tryptone	20 g/L
Yeast extract	5 g/L
NaCl	10 mM
KCl	2.5 mM
$MgCl_2$	10 mM
$MgSO_4$	10 mM
Glucose	20 mM

The SOC media was used for *E. coli* transformation recovery was commercial media from NEB (1X SOC outgrowth media) and Invitrogen.

MAC media:

Glycerol	20 g/L
Tryptone	10 g/L
Yeast extract	5 g/L
HEPES	23 g/L
$FeSO_4$	50 mg/L
K_2HPO_4	0.5 g/L
NaCl	2 g/L
NTA	200 mg/L

The pH was adjusted to 7.4 with concentrated ammonia.

M9 media:

NH_4Cl	1g/L
NaCl	0.5 g/L
$Na_2HPO_4{\cdot}2H_2O$	7.5 g/L
KH_2PO_4	3 g/L
$MgSO_4{\cdot}2H_2O$	2 mM
$CaCl_2$	0.2 mM
Thiamine HCl	0.3 µM
$FeSO_4{\cdot}7H_2O$	10 mg/L
$ZnSO_4{\cdot}7H_2O$	1.8 mg/L
$CaCl_2{\cdot}6H_2O$	1.8 mg/L
$MnSO_4{\cdot}H_2O$	1.2 mg/L
$CoCl_2{\cdot}2H_2O$	1.8 mg/L
Glucose	2 g/L
L-leucine	40 mg/L

The pH was checked and adjusted at 7.4, if necessary, with NaOH or HCl.

2.2. *Clostridium autoethanogenum*

2.2.1. Routine culture and Storage

C. autoethanogenum was stored at -80°C in 15 % dimethyl sulfoxide (DMSO) in screw-capped tube in either 1 ml long-term stock or 300 µl one-shot stock.

The cells were routinely revived in 2 ml pre-reduced YTF media with D-cycloserine incubated 48 h to 72 h at 37°C in the anaerobic cabinet.

Subculture of the inoculum in 10 mL YTF with D-cycloserine at $OD_{600\ nm}$ ≈ 0.05 overnight at 37°C in the anaerobic cabinet. The culture could then be used to reprepare cryo-stock.

When appropriate, the cultures were supplemented with thiamphenicol (15 µg/ml) and D-cycloserine (250 µg/ml).

2.2.2. Culture for protein production

The cultures were prepared routinely as above. The 10 ml cultures were sub-cultured at $OD_{600\ nm}$ ≈ 0.05-0.1 in 60 ml YTF in 120 ml seal serum bottle. The cells were incubated overnight at 37°C in the anaerobic cabinet. The cultures were complemented with 100 µM $NaMoO_4$.

The cultures with bacteria harbouring a riboswitch were supplemented with 5 mM theophylline. When appropriate, the cultures were supplemented with thiamphenicol (15 µg/ml) and D-cycloserine (250 µg/ml).

2.2.3. Media

YTF:

Tryptone	16 g/L
Yeast extract	10 g/L
Fructose	10 g/L
NaCl	0.2 g/L
MES	20 g/L
Acidic 1000X trace element solution	1 mL/L
Basic 1000X trance element solution	1 mL/L
Vitamin 1000X stock solution	1 mL/L

The pH was adjusted at 5.8 with HCl. 15 g/L of technical agar n°2 (Oxioid) was added to prepare solid media. YTF serum bottle was sparged with N_2 for 15-30 min. The media was complemented with D-cycloserine and thiamphenicol.

Acidic 1000X trace element solution:

HCl	50 mM
H_3BO_3	100 mg/L
$MnCl_4 \cdot 4H_2O$	230 mg/L
$FeCl_2 \cdot 4H_2O$	780 mg/L
$CoCl_2 \cdot 6H_2O$	103 mg/L
$NiCl_2 \cdot 6H_2O$	602 mg/L
$ZnCl_2$	78 mg/L
$CuSO_4 \cdot 5H_2O$	50 mg/L
$AlK(SO_4)_2 \cdot 12H_2O$	50 mg/L

Basic 1000X trace element solution:

NaOH	10 mM
Na_2SeO_3	58 mg/L
Na_2WO_4	53 mg/L
$Na_2MbO_4 \cdot 2H_2O$	52 mg/L

Vitamin 1000X stock solution:

p-aminobenzoate	114 mg/L
riboflavin	104 mg/L
thiamine	200 mg/L
nicotinate	206 mg/L
pyridoxin	510 mg/L
Ca D-(+)-pantothenate	104 mg/L
Cyanocobalamin	78 mg/L
d-biotin	22 mg/L
folate	48 mg/L
lipoate/thioctic acid	50 mg/L

The trace elements and vitamin stock solutions were prepared for collective use and stored at 4°C.

2.3. *Clostridium acetobutylicum*

2.3.1. Routine culture and Storage

The cells were stored at -80°C in 20 % glycerol in screw-capped tubes. The spores were stored in 30 ml MS media in sealed serum bottles at 4°C and -20°C in screw-capped tube aliquots.

The cells were revived on pre-reduced CGM plates supplemented with the correct antibiotics from a glycerol stock or frozen spore aliquot. The streaked plates were incubated in the anaerobic cabinet for three days.

The colonies were then used to inoculate liquid media (1 to 10 ml) of 2xYTG or CGM and incubated overnight at 37°C. The culture was then used to prepare cryo-stock or inoculate expression culture.

2.3.2. Spore preparation

The cells were streaked on pre-reduced MS plates with the corresponding antibiotics at 37°C in the anaerobic cabinet for three days.

1 ml MS media was inoculated with colonies from the plates and incubated in the anaerobic cabinet overnight at 37°C.

30 ml MS media in serum bottle was inoculated with all the 1 ml pre-culture and incubated at 37°C for 3 to 5 days without agitation. The serum bottle was degassed when needed, andspore formation was checked under the microscope. The serum bottle was stored, and aliquots were prepared in the anaerobic cabinet.

2.3.3. Culture for protein production

The cells were revived identically on CGM plates, and a 5 ml CGM preculture were performed. The GAPOR and GOR-SL production cultures were performed in 60 mL CGM media in a 120 ml sealed serum bottle. The media was inoculated at 0D ≈ 0.1 and incubated at 37°C for 7 h to 8 h. The gene expression was

induced with 10 mM D-lactose at OD between 0.6 and 0.8 and incubated for 5 h at 37°C. The media was complemented with 100 µM $NaMoO_4$.

Small preparation cultures were performed in 10 ml after pre-cultures were induced with lactose at the inoculation time.

The cultures were supplemented with thiamphenicol (15 µg/ml) when required.

2.3.4. Media

2x YTG:

Tryptone	16 g/L
Yest extract	10 g/L
Glucose	30 g/L
NaCl	5 g/L

The pH was adjusted at 5.2. 2xYTG was used for *C. acetobutylicum* transformation and routine culture.

CGM:

Yeast extract	6.25 g/L
KH_2PO_4	0.9375 g/L
K_2HPO_4	0.9375 g/L
$MgSO_4 \cdot 7H_2O$	0.5 g/L
$MnSO_4 \cdot H_2O$	0.0125 g/L
NaCl	1.25 g/L
Asparagine	2.5 g/L
$(NH_4)_2SO_4$	2.5 g/L
MES	20 g/L

The pH was adjusted at 6.4 with concentrated ammonia.

$FeSO_4 \cdot 7H_2O$	0.0125 g/l

The $FeSO_4$ was added just before sparging from a 100X stock solution. The media was complemented at the same time with Na_2MoO_4 at a final concentration of 100 µM.

Before use, degassed and autoclaved glucose was added at a final concentration of 60 g/L. Degassed cysteine/HCl at 0.5 g/L was added to the media to reduce the trace of O_2.

Solid media was prepared by the lab technician with a similar composition with 15 g/L agar.

MS (sporulation medium):

Glucose	60 g/L
KH_2PO_4	0.55 g/L
K_2HPO_4	0.55 g/L
$MgSO_4 \cdot 7H_2O$	0.22 g/L
Acetic acid	2.3 ml/L
$FeSO_4 \cdot 7H_2O$	0.011 g/L
Para-aminobenzoic acid	8 mg/L
Biotin	0.08 mg/L

The pH was adjusted at 6.4 with concentrated ammonia. MS glucose base synthetic media was used to prepare *C. acetobutylicum* spore stock.

3. Cell harvesting

3.1. *E. coli*

E. coli cells were routinely harvested for cell protein extraction from 50 ml to 3 L of culture in serum bottle. For enzyme assays, the serum bottles were introduced in the anaerobic tent. The cultures were dispatched in 50 ml falcon tubes and harvested by centrifugation at 7,193 ×g for 20 minutes at 4°C. For cultures larger than 50 ml, the supernatant was discarded, and more culture was poured over the pellet. This cycle was repeated until the complete harvesting of the cells. The cell pellets were kept on ice, resuspended in 25 ml

phosphate buffer (10 mM, pH 8.0) and gathered in three tubes. Then the resuspension was centrifuged at 7,193 ×g for 15 min at 4°C.

The phosphate buffer was sparged with N_2 for 15-30 min and placed in the cabinet overnight to complete the reduction.

The same protocol was applied for aerobic harvesting.

The pellet could be kept at -80°C after resuspension in the appropriate volume of phosphate buffer for lysis in a falcon tube or a sealed serum bottle for anaerobic experiments.

3.2. *C. autoethanogenum*

The cell culture was transferred in two 50 ml pre-equilibrated falcon tubes tightly closed and spun at 7,193 ×g for 20 minutes at 4°C. The tubes were reintroduced in the anaerobic cabinet, the supernatants were discarded, the pellets were resuspended in 7.5 ml pre-reduced phosphate buffer and the resuspended pellets were pooled. The solution was then centrifuged again. The pellet was resuspended in 0.425 µl phosphate buffer and transferred to an Eppendorf 1.5 ml tube. The pellets were either kept at -80°C or directly lysed. The pellet was kept on ice.

3.3. *C. acetobutylicum*

Under an anaerobic atmosphere, the cell culture was transferred in two 50 ml pre-equilibrated falcon tubes tightly closed and spun at 7,193 ×g for 20 minutes at 4°C.Cell lysis. The supernatant was discarded, and the pellets resuspended in 10 ml PBS and pooled together. The resuspended cells were centrifuged, and the supernatant was discarded. The pellets were resuspended in 1 ml 10 mM phosphate buffer pH 8 and stored at -80°C in a sealed serum bottle (pre-reduced in the cabinet). For enzyme assay, the pellet was washed and resuspended in 3 ml 0.1 M Tri-HCl, 10 µl glycerol, and 2 mM DTT. The resuspended pellet was kept at -80°C until lysis, purification, and enzyme assay.

4. Cell lysis

4.1. *E. coli*

4.1.1. Chemical lysis

The harvested cells were lysed with 1 ml BugBuster (1.8 ml 10 mM phosphate buffer pH 8.4, with 200 µl of 10x BugBuster plus benzonase nuclease, Novagen) per 100 ml of culture in the 50 ml tube and incubated for 1 hour on ice. The treated samples were then centrifuged at 7,193xg for 15 min, and supernatant was collected. During the crude extract preparation, all the manipulations were performed under the COY anaerobic tent, and the phosphate buffer solution was sparged with N_2 for 20 to 30 min.

4.1.2. Sonication

The cells were resuspended in 2.5 ml to 15 ml depending on the volume of culture harvested. The resuspended cells were treated with 6 cycles at 30% amplitude for 30 sec on and 90 sec off, kept on ice. For the larger volume, the resuspended cells were split in two and treated with 12 cycles with the same conditions. The treated samples were then centrifuged at 7,193xg for 15 min, and the supernatant was collected.

4.2. *C. autoethanogenum*

The harvested cells were lysed with 1 ml of BugBuster lysozyme solution (425 µl per 100 ml of culture, 50 µl 10x BugBuster, 25 µl of 75 mg/ml lysozyme from chicken egg white, 0.5 µl Benzonase nuclease, Novagen). The resuspended cells were incubated on ice for 1 h. The treated samples were then centrifuged at 7,193xg for 15 min, the supernatant was collected, and the pellet was kept. All the experiments were performed in anaerobic cabinet except the centrifugation.

4.3. *C. acetobutylicum*

1 ml of resuspended cells were lysed by sonication in a 2 ml centrifugation tube. The cells were sonicated for 15 sec and rested for 60 sec on ice at 30 % for 10

cycles. The tubes were centrifuged for 15 min at 18,000 xg, 4°C. The supernatant was recovered.

For enzyme assay all the experiments were performed in an anaerobic chambre and all the solution was pre reduced with N_2 prior being entered in the anaerobic chamber. The pellet was thaw on ice and lysed by sonication with 5 cycles of 30 sec on 2 min off on ice at 30 % power. The cells were centrifuged for 10 min at 4°C at 7,193xg.

5. Molecular biology

5.1. Preparation of competent cells

5.1.1. Preparation of *E. coli* competent cells

(a) Chemically competent cells

E. coli competent cells were for heat shock transformation of plasmids from the library in the mutated *E. coli* cells.

100 ml of LB medium in a flask was inoculated with 1 ml *E. coli* overnight culture. The 100 ml culture was incubated at 37°C with agitation up to an $OD_{600\ nm} \approx 0.25 – 0.3$. The culture was chilled on ice for 15 min and centrifuged for 10 min at 3,300 xg at 4°C. The supernatant was discarded. The pellet was resuspended in 30 ml of ice-cold 0.1 M $CaCl_2$ buffer and incubated for 30 min on ice. The cells were centrifuged for 10 min at 3,300 xg at 4°C. The pellet was resuspended in 1 ml 0.1 M $CaCl_2$ + 15 % glycerol. 100 µl of the resuspended competent cells were dispatched in 1 ml Eppendorf tubes. The tubes were stored at -80°C

(b) Electrocompetent cells

Electrocompetent *E. coli* cells were used to prepare *E. coli* TOP10 containing pAN2 plasmids for plasmid methylation for transformation in *C. acetobutylicum* (Heap et al., 2007).

50 ml of LB medium was inoculated with 500 µl of overnight culture in 250 ml flask. The culture was incubated at 37°C with agitation until reaching an $OD_{600\ nm} \approx 0.5 – 0.8$ ml. The culture was chilled on ice for 15 min. The cells were

centrifuged at 5,000 xg for 15 min at 4°C. The supernatant was discarded, and the pellet was resuspended in sterile ice-cold distilled water. The cells were centrifuged as before. The supernatant was discarded, and the pellet was washed with 2 ml 10 % glycerol. The cells were centrifuges again. The supernatant was discarded, and the pellet was finally resuspended in 300 µl 10 % glycerol. 40 µl of electrocompetent cells were aliquoted in Eppendorf 1.5 ml tubes and stored at -80°C.

5.1.2. Preparation of electrocompetent *C. acetobutylicum*

Inoculation of 10 ml of 2xYTG from cells revived on plates. Then serial dilution down to 10^{-4} in 20 ml sample tubes were performed. The cultures were incubated in the anaerobic cabinet. The serial dilution overnight culture with an $OD_{600\ nm} \approx 0.6 – 0.8$ or the least grown culture was used to inoculate 100 ml of pre-reduced 2xYTG in a flask at a starting OD of 0.05. The culture was incubated at 37°C in the anaerobic cabinet until reaching an OD between 0.6 – 0.8. The cells were centrifuged at 4,000 xg for 10 min at 4°C. The centrifuged cells were kept on ice and returned to the cabinet. The supernatant was discarded, and the pellet was resuspended with 10 ml of pre-reduced sterile ice-cold electroporation buffer (EPB: 284 mM sucrose, 5 mM $NaHPO_4$, pH 7.4). The cells were centrifuged in the same condition. The supernatant was discarded. The pellet was resuspended in 2 ml of ice-cold EPB and kept on ice. The competent was directly used for electroporation.

5.2. Transfer of plasmid in bacteria

5.2.1. *E. coli* heat shock transformation

Heat shock transformation was the principal method used to transform *E. coli* with plasmid DNA. When *E. coli* DH5α chemically competent cells (NEB) or *E. coli* TOP10 chemically competent cells (Invitrogen) were used, the manufacturer protocol was followed. The lab-made *E. coli* chemically competent cells were transformed with a standard protocol. The plasmid DNA was added to 50-100 µl competent cells and incubated on ice for 30 min. The cells were then incubated for 30 sec at 42°C then transferred on ice for 2 min. 950-900 µl of SOC media was added to the cells and incubated for 1 h at 37°C.

All or part of the cells were spread on LB plates complemented with the corresponding antibiotics. The plates were incubated overnight at 37°C

5.2.2. *C. autoethanogenum* conjugation

2 ml of pre-reduced YTF was inoculated from *C. autoethanogenum* DMSO stock and incubated at 37°C in the anaerobic cabinet for 48 to 72 h. When growth was observed in the pre-culture, the plasmid was transformed in *E. coli* sExpress as above, and *C. autoethanogenum* preculture was subculture into 1 ml YTF using different ratios between 1:50 to 1:100 for an approximate starting OD ≈ 0.02 and incubate overnight.

The day after, one colony from the *E. coli* sExpress transformation plate was inoculated in 5 ml LB medium with the correct antibiotic (kanamycin and chloramphenicol) and incubated until reaching OD between 0.2 and 0.4. Once the OD was reached 1 ml of *E. coli* sExpress culture was aliquoted and centrifuged at 3,000xg for 3 min. The supernatant was discarded, the pellet was resuspended with PBS by flickering, and then centrifuged. The supernatant was discarded carefully. The pellet was dried as much as possible, and the tube was introduced into the anaerobic cabinet. The pellet was resuspended in 200 µl overnight *C. autoethanogenum* culture by flicking the tube. Then straight after, the resuspension was poured onto a pre-reduced YTF plate without selection. Incubate for 16 to 24 h at 37°C in the anaerobic cabinet.

The following day, the plate was flooded with 600 µl pre-reduced PBS and the spots were scraped with a spreader to resuspend them in the PBS. 50 µl and the remaining mating slurry was spread on pre-reduced YTF selection plates supplemented with D-cycloserine (250 µg/ml) and selection antibiotics. The plates were incubated for 3 to 4 days at 37°C in the anaerobic cabinet.

When transconjugant was observed on plates, the colonies were picked and patched on a fresh pre-reduced selective plate with the same antibiotics. The colonies could then be tested by colony PCR and inoculated in small-volume liquid media for cryo-stock.

5.2.3. *E. coli* electroporation

Plasmids were transformed in *E. coli* pAN2 electrocompetent cells for methylation. The plasmids were put in contact with 40 µl of competent cells on ice for 5 min in the electroporation cuvette 2 mm. The cells were then shocked (using Biorad Gene Pulser Xcell Microbial System with preset Eco2) and 250 µl of SOC was added directly to the cuvette and transferred to anEppendorf tube. The cells were incubated at 37°C for 1 to 2 h. All or part of the cells were spread on LB plates complemented with the corresponding antibiotics and tetracycline (10 µg/ml for pAN2 selection). The plates were incubated overnight at 37°C

5.2.4. C. acetobutylicum electroporation

In the anaerobic cabinet, 570 µl freshly made competent cells were put in contact with 10 to 30 µl (containing 5 to 10 µg of plasmid DNA) plasmids solution and kept on ice. For *C. acetobutylicum tcdR*$^{+}$ strain, methylated DNA had to be used and was prepared as presented (Chapter II:5.2.3). The cells and plasmids were transferred on 4 mm electroporation cuvette chilled at -20°C and kept on ice. The cells were shocked at 1 Ω, 2 kV and 25 µF using Gene Pulser Xcell Microbial System (BioRad). Immediately after the shock, the content of the cuvette was used to inoculate 10 ml of pre-reduced 2xYTG in 50 ml falcon tubes and incubated at 37°C in the anaerobic cabinet for 3 to 4 h. After recovery, the cells were centrifuged at 4,000 xg for 10 min at 4°C and returned to the anaerobic cabinet. The supernatant was discarded, and the pellet was resuspended in 1 ml 2xYTG. Then the resuspended cells were spread on CGM plates complemented with the correct antibiotics and incubated at 37°C in the anaerobic cabinet.

5.3. Plasmid extraction

Plasmids were extracted from overnight *E. coli* culture in 1 to 4 ml LB. The plasmids were extracted using the QIAprep Spin Miniprep kit (Qiagen) or Monarch® Plasmid miniprep kit (New England Biolabs). One modification was brought to the kits' protocol by incubating the tubes on ice for 10 min after adding of the neutralisation buffer.

For the plasmids extraction from *C. autoethanogenum* and *C. acetobutylicum,* the harvested cells were treated with 200 µl lysozyme $2.115x10^6$ unit/ml and incubated for 30 min at 37°C prior to the addition of kits lysis buffer. 5 µl of the plasmid DNA preparation was transformed by heat shock in *E. coli* DH5α chemically competent cells (NEB) for plasmids amplification and plasmid preparation. For those transformations, 5 ml of LB media was complemented with antibiotics after the outgrowth phase.

5.4. P1 transduction

P1 transduction was used to delete genes in *E. coli* by transferring selective mutation, including a selection marker using a P1 phage (Thomason, Costantino, & Court, 2007).

5.4.1. Preparation of phage lysate

5 ml of LB supplemented with 0.2 % glucose, and 5 mM $CaCl_2$ was inoculated with 50 µl of overnight donor strain culture. The cells were incubated at 37°C for 1 h with agitation, and 50 µl of phage lysate was added. The culture was then incubated for 4 h in the same condition until cell lysis. 50 µl of chloroform was added to the culture and vortexed. The cells were then centrifuged at 4,500 xg for 10 min at 4°C. The supernatant was recovered and 50 µl of chloroform was added to it. The P1 phage lysate was kept on ice for direct utilisation or stored in 200 µl aliquot in sterile PCR tubes at 4°C.

5.4.2. Transduction

2 ml of the overnight culture of the target strain was centrifuged for 2 min at 16,000 xg. Then the pellet was resuspended in 1 ml of sterile 10 mM $MgSO_4$, 5 mM $CaCl_2$ solution and kept on ice. 5 different conditions were always tested and prepared as follows:

Tube number	Cells resuspension	P1 Phage lysate	Sterile water
1	100 µl	/	100 µl
2	100 µl	10 µl	90 µl
3	100 µl	50 µl	50 µl
4	100 µl	100 µl	/
5	/	100 µl	100 µl

The tubes were incubated for 30 min at 30°C without shaking. 100 µl of 1 M sodium citrate was added to each tube and mixed. Then, 1 ml LB medium was added, and the cultures were incubated at 37°C for 1 h. Following the incubation, the cells were centrifuged with the same condition to recover the pellets. The pellets were resuspended in 200 µl LB medium and spread on selection plates supplemented with Kan for incubation at 37°C.

5.4.3. FLP recombination

After P1 transduction, to remove the kanamycin marker, the marker sequence was flanked by flippase (FLP) recognition target (FRT) sites (Baba et al., 2006; Datsenko & Wanner, 2000). The kanamycin marker removal was performed using a flippase (*flp*) placed on pCP20 under the control of a temperature-sensitive promoter and a temperature-sensitive replicon.

Chemically competent cells of the new mutant strain were prepared following the standard protocol (Chapter II:1.1.1(a)), and the pCP20 plasmid was transformed in the cells by heat shock (Chapter II:5.2.1). The transformants were selected on LB plates complemented with kanamycin and carbenicillin and incubated at 30°C.

A Few colonies were picked from the selection plate, inoculated on 5 ml LB without antibiotic and grown overnight at 43°C to induce FLP recombinase and lose the plasmid.

Serial dilution 1:100 up to 10^{-6} of the overnight culture and 50 µl of the last dilution was spread on an LB plate without selection for overnight incubation at 30°C. Colonies from the overnight plates were picked and patched on three

different plates, LB + kan, LB + Amp and LB, in that order and incubated at 37°C for LB + Kan and LB plates and 30°C for LB + Amp. The colonies growing on LB plates and showing sensitivity to both antibiotics could be grown on liquid LB media to prepare cryo-stock.

The cells could then be transformed with the plasmids, including pRARE2 prepped from wild-type *E. coli* Rosetta-gami2(DE3).

5.5. DNA manipulation

5.5.1. Polymerase Chain Reaction

DNA fragments were amplified by polymerase chain reaction (PCR). 1 ng to 1 µg of genomic DNA or 1 pg to 10 ng of plasmid DNA was used as templates. For Colony PCR, a fragment of the colonies was picked from the plates and resuspended in 100 µl of sterile water for *E. coli*, 300 µl for *C. autoethanogenum* and *C. acetobutylicum*. These cells' resuspensions were used as templates and PCR mix water. The PCR mixture was prepared following the polymerase manufacturer template in a final volume between 25 µl and 50 µl. The PCR was performed following the polymerase manufacturer protocols. All the polymerase mix was in a 2X master mix containing all the dNTPs. The PCRs were run for 25 to 35 cycles composed as below (the temperature and times represent a summary of the used enzymes conditions)

(a) Colony PCR and control PCR

Colony PCR was performed using OneTaq Quick-Load 2X Master (NEB, Table II-3) to check plasmid transformation and gene deletion. The same polymerase was used to check the cloning of genes.

Table II-3 Cycle parameters for PCR amplification with OneTaq Quick-Load 2x master mix

Step	Temperature	time
Initial denaturation	95°C	30 sec to 10 min
30 cycles	95°C -5°C lower than the Tm to Tm 68°C	15-30 sec 15-60 sec 1 min/kb
Final extension	68°C	5 min
Hold	15°C	

(b) Fragment amplification for cloning

Fragments amplification for cloning and plasmids construction was performed using Q5 High Fidelity 2X Master Mix (NEB, Table II-4). The Template was either purified genomic DNA or plasmid DNA. The non-master mix version of the polymerase was used in Toulouse.

Table II-4 Cycle parameters for PCR amplification with Q5 High fidelity 2x master mix

Step	Temperature	time
Initial denaturation	98°C	30 sec
25-35 cycles	98°C Tm to +3°C higher than the Tm 72°C	5-10 sec 10-30 sec 20-30 sec/kb
Final extension	72°C	2 min
Hold	15°C	

(c) PCR primers

The complete list of primers used in this project are presented in the appendix, Hifi assembly: Table S 1, control of hifi assembly cloning and sequencing: Table S 2, *gor* plasmid construction: Table S 3, Primers for *gapor* and *gor* expression plasmids construction for *C. autoethanogenum* and *C. acetobutylicum:* Table S 4*,* Primers for the control of the construction of the pMTL based plasmid: Table S 5, Primers for the control of the gene deletion in E. coli: Table S 6, Primers for 16s ribosomal DNA for contamination control: Table S 7

5.5.2. DNA synthesis

(a) Oligonucleotide synthesis

The oligonucleotides were synthesised by Sigma-aldrich.

(b) Gene synthesis

P*fdx*_RB3 and P*tcdB* promoter DNA sequences were ordered to GenScript Biotech (Netherlands) B.V. and cloned on pUC57. The *gorSL* genes from *Geosporobacter ferrireducens* modified with the insertion of the 6-histag and thrombin cleavage site sequence were ordered using Geneart, thermoFisher. Sequences are available in Table S 8.

5.5.3. Agarose gel electrophoresis

All the DNA electrophoresis were performed on 0.8% Agarose in Tris, acetate, EDTA (TAE) buffer and run at 100 V for 50 to 55 min. Quick-Load Purple 1 kb plus DNA ladder (New England Biolabs) was used as standard. The samples were mixed with the Gel Loading Dye, Purple (6X).

5.5.4. DNA purification

The PCR product and digested plasmid were extracted from the agarose gel with the Zymoclean Gel DNA recovery (Zymo Research) and the Monarch DNA gel extraction Kit. The PCR products were typically purified using DNA Clean & Concentrator (Zymo Research).

5.5.5. Construction of *gorS2* and *M. maripaludis* molybdenum cofactor biosynthesis pathway enzymes plasmids mini library by Hifi assembly

The assemblies were achieved with pET28a (+) *gor*S2 (Amp^R) digested with BamHI and BlpI as the backbone (1 µg of DNA, 0.5 µl of BamHI, 1 µl of BlpI, 5 µl of CutSmart (10X) buffer and sufficient volume of water up to 50 µl; Enzymes and buffer came from New England Biolabs). The insert *mobA, mobB, moeA1, moeA2, moeA3* and *moeA4* were amplified as explained above, using *M. maripaludis* S2 commercial genomic DNA (DSM-14266, DSMZ GmbH); The primers were designed to have 25 base pair (bp) overlapping ends. A specific synthetic ribosome binding site (RBS) has been attributed to each gene (Table II-5).

Table II-5 synthetic ribosome Binding site allocated to the cloned genes (Zelcbuch et al., 2013)

Genes	RBS	RBS name
mobA	AAGAGGTTTGGA	C
mobB	TTCGCAGGGGGAAG	D
moeAs	AAGTTAAGAGGCAAGA	A’

The HiFi Assembly were performed as advised in the NEBuilder HiFi Assembly Kit in 20 µl and the plasmids we assembled as presented in Table II-6. 5 µl of the assembly mix were transformed in commercial *E. coli* 5α chemically competent cells as presented above and selected on LB-Amp plates. Colonies were control by PCR. The plasmids were validated by sequencing.

Table II-6 HiFi assembly primers for genes PCR amplification. Primer sequences are described in Table S 1

Plasmid	Primer pairs		
	Gene 1	Gene 2	Gene 3
pAM_001	Hifi_1 /Hifi_2	-	-
pAM_002	Hifi_9.1 /Hifi_5	-	-
pAM_003	Hifi_10.1 /Hifi_8	-	-
pAM_004	Hifi_11.2 /Hifi_12	-	-
pAM_005	Hifi_13.1 /Hifi_14	-	-
pAM_006	Hifi_15.1 /Hifi_16	-	-
pAM_007	Hifi_1 /Hifi_3.1	Hifi_4.1 /Hifi_5	-
pAM_008	Hifi_1 /Hifi_25.1	Hifi_18.1 /Hifi_8	-
pAM_009	Hifi_1 /Hifi_26.1	Hifi_20.2 /Hifi_12	-
pAM_010	Hifi_1 /Hifi_27.1	Hifi_22.1 /Hifi_14	-
pAM_011	Hifi_1 /Hifi_28.1	Hifi_24.1 /Hifi_16	-
pAM_012	Hifi_9.1 /Hifi_17.1	Hifi_18.1 /Hifi_8	-
pAM_013	Hifi_9.1 /Hifi_19.1	Hifi_20.2 /Hifi_12	-
pAM_014	Hifi_9.1 /Hifi_21.1	Hifi_22.1 /Hifi_14	-
pAM_015	Hifi_9.1 /Hifi_23.1	Hifi_24.1 /Hifi_16	-
pAM_016	Hifi_1 /Hifi_3.1	Hifi_4.1 /Hifi_29.1	Hifi_18.1 /Hifi_8
pAM_017	Hifi_1 /Hifi_3.1	Hifi_4.1 /Hifi_30.1	Hifi_20.2 /Hifi_12
pAM_018	Hifi_1 /Hifi_3.1	Hifi_4.1 /Hifi_31.1	Hifi_22.1 /Hifi_14
pAM_019	Hifi_1 /Hifi_3.1	Hifi_4.1 /Hifi_32.1	Hifi_24.1 /Hifi_16

5.5.6. Construction of *gorS2* and *gorSL* expression plasmids for *C. autoethanogenum* and *C. acetobutylicum*

GAPOR and GOR encoding genes were expressed in clostridia strains with the pMTL83151 (Heap et al., 2009).

P*hydA* promoter fragment was amplified by PCR using pHydA-LL-Ctag as template. The PCR fragment ends and pET28a(+) gorS2(amp^R) were digested by restriction enzymes SgrAI and NcoI to remove T7 promoter upstream of *gorS2* before cloning P*hydA* upstream of *gorS2* using T4 DNA ligase (NEB). P*hydA-gorS2* was extracted from the plasmids by restriction digestion (NheI, NotI) before being cloned in pMTL83151 MCS using T4 DNA ligase (pMTL83151_AM001). The *gorSL* genes and P*hydA* prompter were amplified by PCR, and the vector was linearised with restriction enzymes NheI and NotI. The three fragments were assembled using NEBuilder HiFi Assembly Kit (NEB) (pMTL83151_AM005). The same vector was prepared identically to replace the *hydA* Ribosome Binding Site (RBS) with the native *gorS* RBS. P*hydA* and *gor* fragments were amplified by PCR with primers containing native RBS. The fragments and linearised vector were assembled using GeneArt Seamless Cloning and assembly kit (Invitrogen) pMTL83151_AM005.1)

P*thl* DNA fragment was amplified from primer dimers by PCR amplification without a double-strand DNA template. P*hydA* was excised from pMTL83151_AM001 by restriction digestion with NotI and NcoI. P*thl* fragment and linearised pMTL83151_AM001 were assembled using NEBuilder HiFi Assembly Kit (NEB).

The P*fdx*_RB3 was extracted from synthetic plasmid 2_Pfdx_RB3_pUC57 by digestion with the restriction enzymes NotI and NdeI. P*hydA* was excised from pMTL83151_AM002 using the same restriction enzyme. The promoter fragment and the linearised vector were gel purified and ligated using T4 DNA ligase (NEB) (pMTL83151_AM003). The *gorS2* gene was excised from pMTL83151_AM003 to linearise the vector with restriction enzymes NdeI and NheI. The *gor* fragment was amplified by PCR using a synthetic *gor*-containing

plasmid. The fragment and vector were assembled using GeneArt Seamless Cloning and assembly kit (Invitrogen) (pMTL83151_AM006).

pMTL83151_AM004 was constructed identically to pMTL83151_AM003 extracting P*tcdB* from 4_PtcdB_pUC57. pMTL83151_AM007 was constructed identically to pMTL83151_AM006 using a primer with a different overhang 5' end for annealing in PtcdB. *gorS2* was excised from pMTL83151_AM004 with the same pair of restriction enzyme.

All the cloning was validated by sequencing (Eurofins) and plasmid restriction map. All the enzymes and kits were used following the manufacturer's protocols.

The plasmids were transferred in *C. autoethanogenum* and *C. acetobutylicum* following the protocol presented above.

5.5.7. Construction of the gor expression plasmid for *E. coli*

The *gorS2* fragment was excised from pET28a(+) *gorS2* (*Amp*^{R}) and linearised the vector. The *gor* genes fragments were amplified by PCR using the synthesised *gor* plasmid as a template. The two were assembled using GeneArt Seamless Cloning and assembly kit (Invitrogen) (pAM020).

5.5.8. DNA quantitation

Double-stranded DNA was typically quantified using a NanoDrop Lite spectrophotometer (ThermoFisher Scientific).

5.5.9. Plasmid design and sequencing alignment

Plasmids design, primers and DNA sequencing alignment were performed using the online tool Benchling (https://benchling.com/signin/welcome).

5.5.10. Promoters and RBS initial transcription rate calculation

The initiation transcription rate of promoter and RBS translation rate estimation was performed using the online modelisation tool using De Novo DNA Operon calculator (https://salislab.net/software/predict_operon_calculator) (Cetnar & Salis, 2020; Farasat et al., 2014; Halper, Hossain, & Salis, 2020; Ng, Farasat, Maranas, & Salis, 2015; Reis & Salis, 2020). DNA and RNA

secondary structures were modelled using Mfold online tool (http://www.unafold.org/) (Zuker, 2003).

6. Protein purification

6.1. Affinity

6.1.1. GAPOR produced in *E. coli*

Affinity purification was performed with a HisTrap FF crude 1 ml (GE Healthcare). The crude extract was diluted 1:1 with binding buffer (20 mM phosphate buffer, 500 mM NaCl, 20 mM imidazole, pH 7.4). The purification was conducted by following the manufacturer's protocol.

(a) Gradient elution

For gradient elution, the purification was performed under aerobic conditions using an Äkta pure (GE Healthcare). The column was equilibrated with 5ml binding buffer (1 ml/min). The diluted sample was loaded in the column using a 2 ml loop on several rounds at 0.5 ml/min. The column was washed with 15 ml of 4% buffer B (20 mM sodium phosphate buffer, 500 mM NaCl, 500 mM imidazole, pH 8.4). The proteins were eluted with a 4-100% buffer B gradient on 20 ml at 1 ml/min (20 mM to 500 mM imidazole). The elution was collected in 0.5 ml fractions.

(b) Step elution

In Nottingham, in the anaerobic COY tent, all the injections, washes and elution were achieved with a 10 ml syringe, and a syringe pump was used to a apply steady flow rate. In aerobic conditions, the purification was performed using an Äkta pure, and a 2 ml loop was used for sample loading. The diluted sample was loaded on the column in two injections at 0.5 ml/min. A first elution at 75 mM imidazole 1ml/min (20 mM phosphate buffer, 500 mM NaCl, 75 mM imidazole, pH 7.4) was applied to remove a part of the contaminant before the GAPOR elution at 125 mM imidazole elution buffer 1 ml/min (20 mM phosphate buffer, 500 mM NaCl, 125 mM imidazole, pH 7.4).

In Toulouse an identical protocol was applied using an Äkta pure in the anaerobic chamber, but a sample pump was used to load the larger diluted sample on the column at 0.3 ml/min.

Äkta methods conception and piloting were done using Unicorn 7 software. The protein flow were monitored using an inline UV spectrophotometer (280 nm) and conductometer.

To remove the imidazole and to concentrate the enzyme solution, the 5 ml GAPOR solution in elution buffer was applied on a Vivaspin 10,000 MWCO centrifugal concentrator (Sartorius) by following the manufacturer's protocol to recover a final volume of 200 µl in 10 mM sodium phosphate buffer pH 8. PD-10 desalting columns (GE Healthcare) were also used following protein elution when an Äkta pure apparatus was used. In Toulouse, the centrifugal protocol was followed by loading 2.5 ml on the desalting column, and 3.5 ml was recovered after desalting. The gravitational protocol was used during protein preparation for cofactor analysis.

For the purification prior to enzyme assays, all the solutions were sparged with N_2 for 20 to 30 min, and the manipulations were completed under anaerobic conditions.

6.1.2. GAPOR produced in *C. acetobutylicum*

GAPOR was purified following an identical protocol to *E. coli*-produced GAPOR purification.

For enzyme assay, 2.4 ml of supernatant was recovered after lysis and was diluted with 600 µl of binding buffer. 500 µl was aliquoted for latter SDS-PAGE and Western blot analysis. The remaining was desalted using a PD-10 desalting column (GE Healthcare). The proteins were eluted in a 3.5 ml binding buffer. 500 µl was aliquoted for further purification. 3 ml was loaded on 1 ml HisTrap Nickel affinity purification column, and the chromatography was performed following the protocol presented above. 5 fractions of 1 ml were collected. 50 µl of each was aliquoted for SDS-PAGE and western blot analysis. The first two fractions were pooled and desalted on a PD-10 desalting column, then eluted

with 3.5 ml of 20 mM sodium phosphate buffer, 0.024 mM sodium dithionite, pH 8.4. Buffer A (Binding buffer): 20 mM sodium phosphate buffer, 500 mM NaCl, 0.024 mM sodium dithionite, pH 8.4. Buffer B: 20 mM sodium phosphate buffer, 500 mM NaCl, 0.024 mM sodium dithionite, 500 mM imidazole, pH 8.4. All the manipulation was performed under anaerobic condition. For the purification prior to enzyme assays, all the solutions were sparged with N_2 for 20 to 30 min, and the manipulations were completed under anaerobic conditions.

6.1.3. GOR produced in *C. acetobutylicum*

GOR was purified using an identical protocol, but no second wash at 75 mM was performed. The proteins were eluted with 200 mM imidazole.

The enzyme assay was performed as presented in the section above. The three first fractions were pooled and desalted for enzyme assay. For the purification prior to enzyme assays, all the solutions were sparged with N_2 for 20 to 30 min, and the manipulations were completed under anaerobic conditions.

6.2. Gel filtration

The gel filtration chromatography was performed by Dr Philip Bardelang on a Superdex 200 10/300 (GE Healthcare) with an Äkta pure. The injection volume was 200 µl, and the elution flow rate was 0.5 ml/min. 0.5 ml fractions were collected from 2.8 ml to 22.5 ml (run 1) and 2.7 ml to 21.7 ml (run 2) of eluent (phosphate buffer saline). Elution fractions containing GAPOR were collected in one pool and subjected to concentration in the same buffer with a Vivaspin 30,000 MWCO centrifugal concentrator (Sartorius)

7. Enzymatic activity assays

7.1. GAPOR assay on crude extract

All the GAPOR enzyme assays were performed under the Coy anaerobic tent in Nottingham or inToulouse TBI EAD3 anaerobic chamber, and the absorbances were measured with Jenway 7205 spectrophotometer with a 37°C heated cuvette holder in Nottingham and HP/Agilent 8453 UV/Vis Spectrophotometer

with a heated 8 slot cuvette passer (Agilent Technology, Hewlett-Packard). All the solutions except the GAPOR solution and the concentrated G3P were sparged with N_2 for 20 to 30 min. 500 µl of 2x reaction mix (6 mM Benzyl viologen, 112 µM Na_2MoO_4, and 100 mM EPPS were pre-mixed, G3P was added just before the dilution to a concentration of 60 µM) was pre-heated at 37°C. Typically 250 µl, 100 µl, 125 µl and 60 µl of crude extract were added to 100 µl of 38.6 µl sodium dithionite (SDT) solution and sufficient volume of deionised water for 1 ml total volume pre-heated at 37°C. The reaction was started with the diluted enzyme solution in Nottingham experiments; in Toulouse, the reaction was started with the addition of 100 µl of 100 µM G3P. The kinetics were typically followed at 600 nm (ε_{BV} =7,400 $M^{-1}.cm^{-1}$) for 5 to 15 min with a measurement every 7 seconds. For the test of the GAPOR reactivation, 10 µl of 100 mM L-cysteine-HCL was added to 480 µl of pre-heated solution and 10 µl of purified GAPOR (Park et al., 2007) 1 min before the start of the assays.

7.2. GAPOR assay on purified enzymes

For enzyme assay on purified GAPOR, the protocol is relatively similar. The 500 µl of reaction mix without G3P was added to the cuvette with typically 300 µl of deionised water and 100 µl, 75 µl and 50 µl of purified desalted enzymes. The reaction was started with 100 µl of 100 µM G3P. The GAPOR solution was put in contact with 63 mM L-cysteine-HCl for 10 min in a tentative to reactivate the GAPOR. For aerobically expressed GAPOR, the cysteine was dissolved in the 2X EPPS buffer and the pH was roughly adjusted with NaOH (sparged with N_2) at pH 8 with pH paper to compensate for the pH drop. The reaction was monitored with HP/Agilent 8453 UV/Vis Spectrophotometer with a heated 8-slot cuvette passer (Agilent Technology, Hewlett-Packard). A complete visible spectrum was measured for each time point. The reaction monitoring was started before the addition of enzymes for 25 min.

7.3. GAPOR and GOR produced in *C. acetobutylicum* enzyme activity

Both enzymatic activity assays were performed identically. The 500 µl of reaction mix without G3P was added to the cuvette with typically 300 µl of deionised water and incubated at 37°C. The monitoring was started, and different volume of purified and desalted enzyme solution was added to the cuvette (50 µl only for GAPOR, 100 µl, 200 µl, and 400 µl for both GAPOR and GOR). The reaction was started with 100 µl of 100 mM G3P. The reaction was monitored with HP/Agilent 8453 UV/Vis Spectrophotometer with a heated 8-slot cuvette passer (Agilent Technology, Hewlett-Packard). A complete visible spectrum was measured for each time point.

8. Protein detection and titration

8.1. SDS-PAGE

The sodium dodecyl sulphate polyacrylamide gel electrophoresis (SDS-PAGE) was typically performed using 4%-12% NuPAGE Bis-Tris Midi or Mini Gels (Invitrogen) following the manufacturer's recommendation. The running buffer was the MES-glycine (20x) (Invitrogen). The midi gels were run at 200 V for 40 or 45 min, while the mini gels were run at 200 V for 35 min. NuPAGE LDS Sample Buffer (4X) (Invitrogen) was used as loading buffer and 0.5 M Dithiothreitol (DTT) was used as a 10X reducing agent. To prepare the sample mix from the lysis pellet: the pellet was resuspended in the sample mix (loading buffer, DTT and water), then incubated at 4°C, heated more than 10 min at 100°C and cooled on ice. The cooled solution was centrifuged at 10,000 ×g for 1 min. The supernatant was pipetted in clean tubes and loaded on the gels. Gels were stained using blue silver staining (Candiano et al., 2004)

For increased band separation between 40 kDa and 10 kDa, Novex 16 %, tricine, 1.0 mm, Mini Protein Gel, 10 wells (Invitrogen) were used. For 20 µl SDS sample, 8 µl of protein solution or 2 µl of resuspended pellets were mixed with 10 µl Novex tricine SDS sample buffer 2X (Invitrogen), 2 µl 10x reducing agent, deionised water was added for pellet sample. The sample was heated for 5 min

at 75°C and spun 1 min at 10,000 xg. The gel was run for 90 min at 125 V using Tricine SDS Running Buffer (10X) (Invitrogen).

8.2. Blue silver straining

The proteins were fixed on the SDS-PAGE gel by incubation with a fixative solution (40 % v/v methanol, 10 % v/v glacial acetic acid in deionised water) for 20 min to 16 h. The gel was then rinsed with water. The gel was then stained for 1 h to 16 h using Blue-Silver stain (reagent A: 14.75 % v/v of 85 % phosphoric acid 125 g/L of ammonium acetate in deionised water; reagent B: 6 g/L of Coomassie Blue G-250 in methanol; reagent were mixed 4:1 (A:B) just before used). The gel was then rinsed with deionised water for 1 h to 16 h to remove the background strain. (Candiano et al., 2004).

8.3. Western Blot

The proteins were transferred to the PVDF membrane using Trans-Blot Turbo Transfer System and Trans-Blot Turbo Mini and Midi PVDF Transfer Packs (Bio-Rad) with the mixed MW program. The His-tagged GAPOR were revealed using anti-Histag HRP conjugated antibody (HISP12-HRP Alpha diagnostic international) and 3,3',5,5'-Tetramethylbenzidine (TMB) (Sigma-Aldrich) as colourimetric substrate (Ni, Xu, & Gallagher, 2017).

8.4. Bradford titration

The protein solution concentration was measured following the Bio-Rad micro-assay protocol, and the 0.125 -2 mg/ml bovine serum albumin standard kit was used as a standard on microplates. The plate was read with the Promega GloMax®-Multi+ Microplate Multimode Reader.

8.5. BCA titration

The protein solution concentration was also measured using Pierce™ BCA Protein Assay Kit, ThermoFisher Scientific. The plate was read with the Promega GloMax®-Multi+ Microplate Multimode Reader.

9. Molybdenum cofactor characterisation

9.1. FORM A fluorescence titration

Xanthin oxidase cofactor extraction was performed following Meckenstock and co-workers' protocol (Meckenstock, Krieger, Ensign, Kroneck, & Schink, 2001). 50 µl of (6 mg/ml) Xanthine Oxidase from bovine milk, Grade III, ammonium sulfate suspension, >=0.8 units/mg protein (Sigma-Aldrich) was diluted in 200 µl of 55 mM $KMnO_4$ (in 100 mM NaOH) and heated for 15 min at 100°C. 200 µl of 66 % absolute ethanol and centrifuge for 15 min at 15,000 xg. The supernatant was recovered and dispatched in the 96-well microplate (Greiner 96 Flat Bottom Transparent Polystyrol) as well as a 1/200 and 1/500 sample dilution. Pterin-6-carboxylic acid was used as standard, solubilised in 55 mM $KMnO_4$ solution, and was treated like the sample. 0 µM, 0.78 µM, 1.56 µM, 3.125 µM, 6.25 µM, 12.5 µM, 25 µM and 50 µM Pterin-6-carboxylic acid solutions was for the standard curves. The plate was read with a Tecan Spark plate reader with an excitation wavelength of 380 nm and an emission wavelength of 450 nm.

9.2. Molybdenum cofactor liquid chromatography-mass spectrometry

9.2.1. Cofactor extraction

The GAPOR cofactor was extracted from purified GAPOR produced with *E. coli* RG_AM023gapor by acidic treatment and ultrafiltration. 2 ml of purified enzymes were used for each condition. The 30,000 MWCO Vivaspin 2 was used as a 2 ml ultrafiltration apparatus, the apparatuses were spun following the manufacturer's recommendations. The membrane was washed with 6 ml ultrapure Elga water. The filters were then washed with 4 ml of 100 mM ammonium acetate buffered at pH 3 (acetic acid) and pH 8 (concentrated ammonia) for control. The filtrates for both conditions were kept as chemical background control. 2 ml of purified GAPOR was filtrated, and the filtrate was kept as second background control. The protein was then washed with 6 ml of ammonium acetate buffer up to recover typically 400 µl of the retentate. The

6 ml filtrate was collected for later LC-MS analysis. All the solutions were kept at -20°C until analysis.

9.2.2. LC-MS

Buffered ultrafiltered extracts were extracted using reversed-phase solid-phase-extraction (International Sorbent Technologies; 50mg, 1mL, ENV+) cartridges. Recovered eluted components were dried to residue using vacuum centrifugation, resuspended (0.1mL water 0.1%[v/v] formic acid) and analysed (10μL) using high-pressure-reversed-phase liquid-chromatography with a superficially porous (Waters; CORTECS T3; 2.1x50mm) column chromatographed with a linear gradient of 2-80% (v/v) acetonitrile in 0.1% (v/v) formic acid at 0.2mL.min^{-1}. The Eluent was directed to a hybrid quadrupole-time-of-flight mass spectrometer (Waters; QToF Premier) equipped with an electrospray ion-source operated in positive-ion mode and scanning spectra (30-1600 m/z) were calibrated using a lock-mass to achieve an accurate mass of typically 20ppm. This experiment was performed by Dr David Tooth.

Chapter III:

Heterologous Expression of the Glyceraldehyde-3-Phosphate: Ferredoxin Oxidoreductase from *Methanococcus maripaludis* in *Escherichia coli*

1. Introduction

The goal of this study, as presented in Chapter I:7, was to modify *E. coli* to use CO as the sole carbon and energy source. It was hypothesised that expressing a combination of CODH, ferredoxin and GAPOR heterologous genes might solve this problem. There were different candidates for GAPOR among the known functioning GAPOR. The best studied is the W-dependent GAPOR from the thermophilic archaea *Pyrococcus furiosus* (P.-L. Hagedoorn, 2019; Mukund & Adams, 1995). The heterologous production of GAPOR from the thermophilic *Methanocaldococcus janaschii* and the mesophilic *Methanococcus maripaludis* in *E. coli* was performed by Guerrini during his PhD works. However, no G3P dependent activity of the GAPOR was observed in W-enriched cultures despite observing a hydrogenase-like Benzyl viologen activity in the presence of hydrogen (Guerrini, 2007). Interestingly, the same year, the GAPOR from the mesophilic *M. maripaludis* was found to be Mo-dependent (Park et al., 2007) enzyme. In *M. maripaludis* S2, the GAPOR is a 613 amino acids monomeric protein encoded by the 1,842-bp-gene MMP_RS04900 (*gorS2)* (Costa et al., 2013; Park et al., 2007; Van Der Oost et al., 1998) with a molecular weight of 70.9 kDa. Interestingly, the *M. maripaludis* GAPOR is a strictly Mo-dependent enzyme as it did not exhibit any activity when the cells were grown in a medium supplemented with W, even in the presence of Mo (Park et al., 2007). Moreover, during a heterologous expression of *gorS2* in *Escherichia coli,* only one-fifth of the GAPOR proteins contained an atom of Mo. This low incorporation rate could be explained by the inability of *E. coli* to synthesise the specific MoCo (Park et al., 2007). Moreover, the heterologous production of another GAPOR in *E. coli* from *M. jannaschii* did not show any activity despite the insertion of the W atom (Guerrini, 2007).

To date, only one study has been published about the heterologous production, purification and activity assay of *M. maripaludis* GAPOR (Park et al., 2007). GAPOR has been produced in two different *E. coli* strains (BL21 (DE3) and Rosetta-gami 2 (DE3)). The specific activity of these two strains' GAPOR was 14 U/mg of protein and 120 U/mg of protein for the oxidation of G3P with benzyl

viologen as the electron acceptor, respectively. In those experiments, the maximum activity has been observed at 0.1 mM G3P, and a strong inhibition was observed at higher concentrations (Park et al., 2007). In this study, the proteins from BL21(DE3) and Rosetta-gami2 (DE3) contained 0.17 and 0.74 g-atoms of Mo per mol of protein, respectively. A GAPOR purified from *P. furiosus* contained 1 g-atoms W/mol of proteins instead. The *P. furiosus* GAPOR is enzymatically able to function in the gluconeogenesis direction using Fd_{red} as an electron donor. However, the reduction of 3PG in G3P by the *M. maripaludis* GAPOR has not been observed *in vitro* using reduced benzyl viologen as an electron donor (BV) (Park et al., 2007).

The alignment of the amino acids sequence of the GAPORs from *P. furiosus*, *M. maripaludis* and *M. jannaschii* (Figure III-1) presented numerous sequence similarities. The Four cysteines (C292, C296, C300, C516), which interacted with the cubane-type cluster near the active site, were conserved in the *M. maripaludis* GAPOR. Moreover, the residues from the active site, close to the Mo/W atom, were conserved. Furthermore, the residues interacting with the two organic moieties of the Mo-*bis*pyranopterin (Guerrini, 2007) were also conserved. Except the use of a different metallic atom, all the important GAPOR function seems to be conserved in the *M. maripaludis* GAPOR. (Park et al., 2007). Thus, to produce a GAPOR in *Escherichia coli* to construct our new strain, the GAPOR from the mesophilic archaea *Methanococcus maripaludis* was chosen. In fact, it is the only one found to have G3P dependent activity after heterologous expression in *E. coli.* Moreover, *E. coli* produces Mo enzymes only (Leimkühler, 2020; A. Magalon & Mendel, 2015) and does not possess any of the tungsten-specific transporters, unlike *P. furiosus* (Bevers Loes et al., 2006). Park and co-workers demonstrated that *gorS2* was best expressed in *E. coli Rosetta-gami2(DE3)* with a specific activity of 120 U/mg (Park et al., 2007). According to Park et al., both aerobic and anaerobic conditions could lead to the production of active GAPOR after reactivation with cysteine-HCl. Hence, it was then decided to use the *gorS2* from *M. maripaludis* expressed in *E. coli* Rosetta-gami2(DE3).

To know if the GAPOR could catalyse the gluconeogenesis reaction, the GAPOR from *M. maripaludis* should be used in combination with heterologously expressed *codh* and ferredoxin. Hence, it was first important to characterise the GAPOR produced in our strain using Park et al. experimental conditions.

To achieve this, a strain producing the GAPOR, already built by Prof. Philippe Soucaille's lab in Toulouse, was tested. The strain was also optimised to improve the expression and to perform crude extract enzyme assay to observe the potential *in vivo* as well as the characterisation of the specific activity.

```
CLUSTAL O(1.2.4) multiple sequence alignment
GAPOR P. furiosus     MKFSVLKLDVGKREVEAQEIEREDIFGVVDYGIMRHNELRTYEVDPYDPRNIVIFGIGPF        60
GAPOR M. maripaludis  -MNIL--IDGSRQN--YEE--LEESEFPISFGINLHTKQETWKYDAFDEKNLFCFGKGIL        53
GAPOR M. Janashii     MKNAL--INATTKK--FEI--IEKTVLPITWGLYWHNKFETWKYDAYDEKNVFCFGSGVL        54
                            :  ::   ::   :    *.    : :*:  *.: .*:: * :* :*:. ** * :
GAPOR P. furiosus     AGSVLPGSHRLVFFFRSPLYGGLFPSTMGGAGYQFKNVGVDFVEIHGKAEKPTVIILKND       120
GAPOR M. maripaludis  P---IIGGHRLIFSFRSPLWDGFHFSAMGGAGYTFKDTGIQNVAITGKCEVPTVIVLNGE       110
GAPOR M. Janashii     P---VIGGHRLIFSFRSPLWDGFYFSSMGGAGYQFKSTGLNNVAIIGRCENPSILVIEN-       110
                          : *.***:* *****:.*:. *:****** **..*:: * * *:.* *::::::.
GAPOR P. furiosus     GEKLSVDFYEIELEKLLDVWKEYKGEEGVYALTQYLLDNLASVFEGMEFRIAVVGPAALN       180
GAPOR M. maripaludis  EDKLKIEFMPFTEEI-----------TDIYEFNDKI----IDLFKEKNYRAFLVGPASKT       155
GAPOR M. Janashii     DGQLRIDFIEVKEEL-----------KTVYEVSKYI----LELYKDKNLRSVVVGEAAKR       155
                        :* ::*  .  *             :* ... :     .:::  : *  :** *:
GAPOR P. furiosus     TNMGAIFSQALRNGKRAVGSEDWAARGGPGSVLLRAHNVVAIAFGGKKRKREFPGEDISD       240
GAPOR M. maripaludis  TNMGGIYSQTIRNGKIVEGSEDWAARGGGGSVLYQAHNVLGVVFFGKKTPEKN-------       208
GAPOR M. Janashii     TNMGGLFSQTVRNGKFVEGSEDWAARGGGGSVLYRAHNIMGIVFFGDEKEDKE------E       209
                      ****.::**::**** . ********** **** :***::.:.* *.:   :
GAPOR P. furiosus     VKVAKRVVEGIHKKAQRDVINESTVKYRYNPKLNTGGTFGGNYPAEGDLVPVLNWQMPYI       300
GAPOR M. maripaludis  ---LKEIVEEHYNKPYTKVVLEHTEKYRYSEEKKTGGTFGNNYHVTMELTPVFNWRMPFI       265
GAPOR M. Janashii     KEKAKKIIESYYKKPMSKVVLEHTKKYRYDEETKTGGTFGNNWLLYKEKVPIFNWRMPYI       269
                          *.::*  ::*   .*: * * ****. : :******.*:    : .*::**:**:*
GAPOR P. furiosus     PKEERIKIHELIMKYYWEPFNKESIQPKNWTTCGEPCPVVCKKHRKGHHVEYEPYEANGP       360
GAPOR M. maripaludis  DKNKRMKLHKKIIEYFVNRFDEEAIETKNWTNCGEPCPVVCKKYRKGLHVDYEPYEANGP       325
GAPOR M. Janashii     DKEDRKKILEKILKFYLEIFNKETIEPKRWANCGEPCPVLCKKYRNKNKVDYEPYASNGT       329
                       *:.* *: : *:::: : *::*:*: *.*:.*******:***:*:  :*:**** :**
GAPOR P. furiosus     LSGSIYLYASDISVHAVDAMGFDAIEFGGTAAWVLELVHKGLLKPAEVGISDV----PEF       416
GAPOR M. maripaludis  CIGVFDIYAADKVVHTIDKLGFDAIEFGNLCSWTFELLDNGMLKPEEVGIEKPVFDISNF       385
GAPOR M. Janashii     LLGIFDLYEADRVVKTADALGFDAIEIGNLTAWVFELLDVGLLKEEELNIKKPIFDYKKI       389
                        * : :* :*  *:: * :******:*.  :*.:**:. *:**  *:.*..      ::
GAPOR P. furiosus     TKDDLITKPVEASEKNAKLVAELAHSIAFGKTEVARIIGMGKRKASKILDEKFKDRLS-Y       475
GAPOR M. maripaludis  END---EDILKNSMHNAEQAVKLAEIIAFQTNEFGKICKSGTRRAGKILNEKYPDRIK--       440
GAPOR M. Janashii     TNDD-DEEIREISKHNAEQAIKFMHNLAENSNDLYKILSLGKRKAAKILNERFKSRVNKI       448
                       :*    .  : * :**: . :: . :*  ..:. :*   *.*:*.***:*:: .*:.
GAPOR P. furiosus     GESFKDYGVYTPLGDDGEINPTMYWAIGNFIPLPIQGRYWTFYQFGVFLEPEELAQKIVS       535
GAPOR M. maripaludis  DKKFEDFGVYDSFGERGQISPTMYWAIGNFMPYLIQGKYLTHYQCGVFLEPEELAELSVK       500
GAPOR M. Janashii     GKKFNDFAVYVPFGDWGEIAPNLYWTPGFFMPFVIQGRYLTYYK-PEFNEPEKLAELVVE       507
                      .:.*:*:.**  :*: *:* *.:**: * *:*  ***:* *.*:   * ***:**:  *.
GAPOR P. furiosus     SALWEFWYDNVGWCRFHRGWMKKVLKALFMEAYGVSIDMEEHAKKQIRKLIDYLKKAGYE       595
GAPOR M. maripaludis  NSIEEITLENLGICRFHRKWVTPIIEKLVKEMSDV-N-LNEESMELFKKIAKYDSNIGCP       558
GAPOR M. Janashii     SIKLELPIENLGICRFHRKWLKPVLKELVKELLGIED-IVEDSINLYREICEYNKKIGYP       566
                      .   *:  :*:* ***** *:. ::: *. *  .:   : *.: :  ::: .* .: *
GAPOR P. furiosus     PVFWDSMRVIDLVAKGSEEFGNENWAKKFKEDKIGTAKEYLKRVLDAYSQLIGTEWTL--       653
GAPOR M. maripaludis  -E-MESERVKELIIAGAFEFENEKWSKEFEN---GNFDEYIKRVLEKYSELLEIDWKLKE       613
GAPOR M. Janashii     -AKIESERVKDLIIAMAKEFGNEEWTKKFEN--KENVDEYVKRVLNKYSELLGIDWRIS-       622
                          :* ** :*:   : ** **:*:*:*::    . .**:****: **:*:  :* :
```

Figure III-1 Alignment of the GAPOR amino acids sequence from *P. furiosus* (AAL80588.1), *M. maripaludis* (WP_011170889.1) and *M. janashii* (AAB99186.1) with Clustal omega version 1.2.4. The residues interacting with the [4Fe-4S] centre were highlighted in yellow, the residue of the active site close to the Mo/W atom were highlighted in blue and the residues, which interact with the Molybdenum cofactor in green (Guerrini, 2007).

2. GAPOR gene heterologous expression in *Escherichia coli*

2.1. Aerobic expression

Before performing any enzyme assay, it was important to check if the *E. coli* rosetta-gami2(DE3) pET28a(+)gorS2 (amp^R) (Figure S 1) send from Prof. Philippe Soucaille Toulouse lab was properly producing GAPOR following the Park et al. culture protocol on a modified M9 media.

The expression of the gene at a lower temperature improves the protein solubility (Kataeva et al.) by avoiding inclusion body formation (Francis & Page, 2010; Rosano & Ceccarelli, 2014). Hence, to potentially improve the solubility and avoid inclusion bodies formation, the cultures were performed at lower temperatures (22°C to 25°C) than the normal 37°C. The GAPOR production was induced by the addition of 0.1mM IPTG to the *E. coli* GAPOR-producing strain.

The growth of *E. coli* Rosetta-gami 2 (DE3) pET28a (+) *gor*S2 (Kan^R) (RG) was studied for 48 hours after inoculation from a 24h grown preculture (performed at room temperature as well). At the end of the growth of the preculture, the $OD_{600\ nm}$ was only about 0.5-0.6. Also, a lag phase of 24h was observed after inoculation at 1.5%. The OD of induction (0.4 to 0.6) was reached after 30h. The specific growth rate was calculated to be $\mu = 0.2295 \pm 0.03964\ h^{-1}$ which was about two-fold less than the specific growth rate of *E. coli NC81* at 23°C in glucose minimal medium, ~0.4 h^{-1} (Herendeen, VanBogelen, & Neidhardt, 1979). Hence, the growth of the *E. coli* RG strain is much slower due to the presence of the two plasmids pRARE2 and pET28a (+) *gor*S2 and not due to GAPOR production. However, after induction, it seems that the growth slowdown.

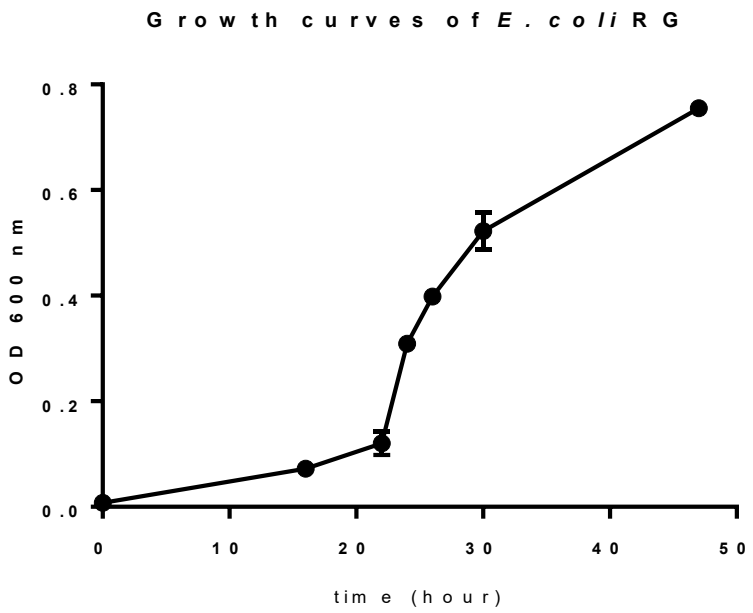


Figure III-2 Growth curve of *E. coli* RG. OD at 600 nm of the culture of *E. coli* Rosetta-gami 2 (DE3) pET28a (+) *gorS2* (*Kan*^{R}) (RG) inoculated with 24h grown preculture at 1% or 1.5% at room temperature in 25 ml of M9 medium is presented. The induction was carried out after 30-31h with 1 mM IPTG. The exponential regression curve (from t=0 to T=26) was calculated with GraphPad Prism 7: k = 0.2295 ± 0.03964 h^{-1} with a doubling time of about 3 hours.

To confirm the production of soluble GAPOR under these conditions: the cells were 10-fold concentrated and then lysed with BugBuster. Two different supernatant samples were loaded on SDS-PAGE: the culture supernatant to control an unwanted excretion of the target protein and the lysate supernatant. The *M. maripaludis* GAPOR is a 70.9 kDa (Park et al., 2007). The GAPOR is a soluble protein which possesses two cofactors. If the GAPOR is present in the soluble fraction, the chances of obtaining active enzymes for *in vivo* activity and *in vitro* enzyme assay are increased. Also, for purification and enzyme activity characterisation purposes, to have soluble protein would avoid the issue of protein re-solubilisation, refolding, and cofactor insertion.

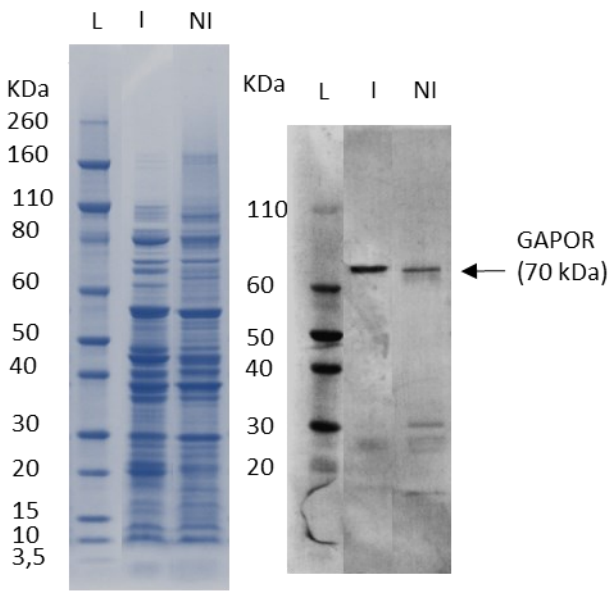


Figure III-3 Detection of soluble GAPOR production from *E. coli* Rosetta-gami 2 (DE3) pET28a (+) *gorS2* (*Kan*^{R}) (RG) aerobic culture on M9.The cultures were performed in 25 ml M9 media and inoculated at 1.5%. The cultures were incubated at room temperature (*ca.* 22°C) for 30 h to be induced with 1 mM IPTG and further incubated at room temperature for 24h. The Control culture was not induced (NI). The cells were harvested following the routine protocol. GAPOR was observed in both due to low temperature IPTG promoter leakage at lower temperature. Left: Coomassie blue silver stain SDS-PAGE (False colour). Right: Westernblot anti-histag. The wells order is identical for both gels: I: Lysate supernatant of IPTG-induced RG culture; NI: lysate supernatant of non-induced control RG culture; L: ladder.

As shown in Figure III-3, bands were observed at 70 kDa on the Western blot and the Coomassie blue SDS-PAGE. This band certainly corresponded to GAPOR as it was normally the only His-Tagged protein produced by these strains. Therefore, it was concluded that the cells were able to produce a soluble GAPOR. It was concluded that when GAPOR is produced, the growth is slower than a normal *E. coli* strain, even at low temperatures. This temperature was chosen to protect the GAPOR during the long induction periods (Park et al., 2007), avoid a premature degradation of the enzyme, and increase the proportion of soluble GAPOR. However, it could lead to some promoter leakage as observed on the Figure III-3. Also, minimal media utilisation did not seem relevant for the future experiments, as no precise knowledge and control of the media composition was required. Therefore, it was decided to abandon the minimal media to use the rich MAC media designed to improve *E. coli* growth under anaerobic conditions (Guerrini, 2007).

2.2. Anaerobic experiments

The GAPOR is an oxygen-sensitive enzyme (Park et al., 2007): the presence of molecular oxygen could lead to spontaneous formation of a disulphide bond between the cysteine coordinating the [4Fe-S] cluster of the apo-GAPORs (Messens & Collet, 2006). So, the absence of oxygen during, at least, the induction of the protein production would prevent a part of the soluble GAPOR from aggregation. *M. maripaludis* is an anaerobic archaeon (Goyal et al., 2016), and all the published AORs were found in strictly anaerobic bacteria (Nissen & Basen, 2019), including the recently found YdhV from *E. coli*, produced only during anaerobic conditions (Reschke et al., 2019). Although Park and co-workers showed that the anaerobic expression of the *gorS2* was not mandatory to produce active GAPOR if cysteine-HCL was added to the enzyme assay reaction mix (Park et al., 2007), it seemed to us that the production of GAPOR had to be performed anaerobically to reduce the risk of apo-GAPOR aggregation and preserve potential active GAPOR.

2.2.1. Anaerobic culture.

After our first culture experiments presented above, we observed slow growth and a long lag phase during minimal media aerobic culture at room temperate. Moreover, the OD600nm observed after 48h incubation did not reach 1 Figure III-2. We supposed that to achieve our objectives, a strictly controlled media composition was not required for the enzyme assay, as no specific nutrient regulation was expected. Furthermore, we supposed that the *E. coli* growth would have been slower under anaerobic conditions. Thus, we chose to use the rich, anaerobic specific media: MAC medium (Guerrini, 2007). This culture broth was already used in similar conditions by Guerrini. The anaerobic cultures were then performed in sealed serum bottles. The medium was not pre-reduced with nitrogen before inoculation. It allowed *E. coli* to use the oxygen in the sealed serum bottle to start its growth, helping to reach the induction OD faster. Around induction, the cells switched to anaerobiosis for GAPOR production In the culture on rich medium (Figure III-4) even though the culture turned into anaerobic, the exponential phase started after few hours, between

6 to 10 hours. While on M9 minimal medium the exponential phase started between 16 to 22 h but ended between 31 to 48 h Figure III-2, whereas the anaerobic rich medium cultures, which ended the exponential phase after 48 h. However, the growth under anaerobic conditions in 25 ml is very close to the growth under aerobic conditions.

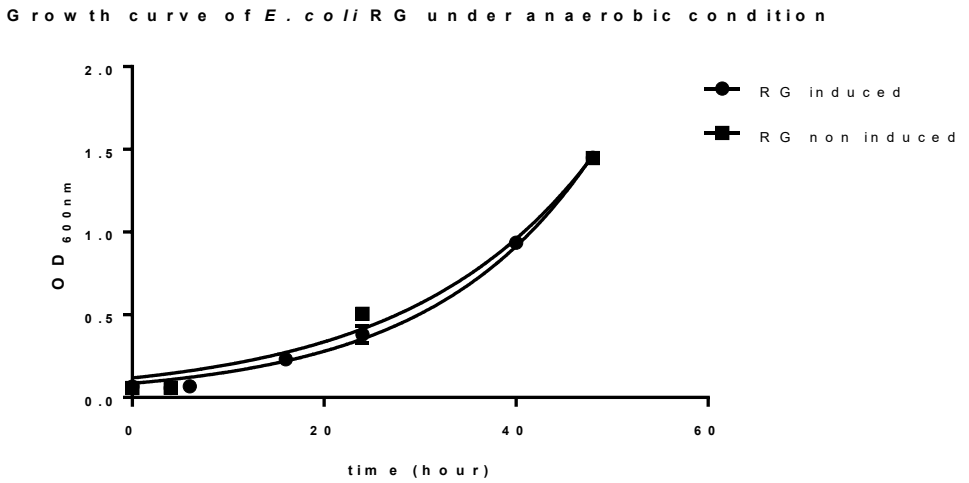


Figure III-4: Growth curve of *E. coli* RG under anaerobic conditions. OD at 600 nm of the culture of *E. coli* Rosetta-gami 2 (DE3) pET28a (+) *gorS2* (Kan^R) (RG) at room temperature in 25 ml of MAC medium under anaerobic conditions inoculated after 24 hours preculture under the same conditions. The culture was induced with 1 mM IPTG after 24 h of growth, and no IPTG was added to the control culture.

Furthermore, the induction with IPTG did not seem to impact on the cells' growth, contrary to the aerobic growth, where the growth slowdown after the induction. However, this slowdown could have been caused by nutrient limitation in the liquid medium, which could have occurred later in the rich medium. The OD600nm was measured to monitor the growth and not to draw the growth curves. Hence, even though those data gave information about the strain's growth, they are not suitable for calculating a relevant growth rate.

2.2.2. GAPOR production

We previously showed that soluble GAPOR was produced under aerobic conditions (Figure III-3). However, it was supposed that the GAPOR was prone to form insoluble aggregate due to the cysteine coordinating the iron-sulfur cluster. These insoluble aggregates are recovered in the lysis pellet with other insoluble proteins and cell debris. Thus, to observe the proportion of insoluble

GAPOR produced, the proteins from the lysed pellet were also run on SDS-PAGE for Coomassie straining and anti-his-tag western blot.

If necessary, the protein produced in inclusion bodies can be recovered, with a treatment of the pellet with chaotropic agents such as urea, or Guanidinium (Kaur, Singh, Panda, & Lal, 2021). However, GAPOR contains two cofactors (molybdenum cofactor and [4Fe-4S] cluster). Although it was known that the FeS cluster could be reconstituted in the presence of an excess of both ions (Guerrini et al., 2008), it was unsure that the protein would be able to be refolded with both cofactors. Consequently, it was decided that the inclusion bodies would not be used in our experiment. To later investigate the kinetic activity of the GAPOR, it was essential to understand which condition would allow us to gather the most soluble GAPOR. Indeed, later, to measure the specific activity of the enzyme of interest, it is necessary to purify an active enzyme. Therefore, the advantage of having more protein in the supernatant was to avoid the protein's solubilisation and refolding step. Hence, to understand which proportion of the produced enzyme was lost in the inclusion body, both pellet and supernatant of anaerobic and aerobic cultures were loaded respectively on an acrylamide gel for SDS-PAGE electrophoresis for comparison (Figure III-5). During the experiment shown in Figure III-5, two different methods of lysis, chemical (BugBuster and B-PER) and physical (sonication), were also tested to understand which of them would be the most relevant for the study.

In Figure III-5, the GAPOR was detected in all the tested conditions (marked with the red rectangle). An equivalent amount of soluble proteins from lysed cells for both conditions is visible in the silver-blue Coomassie straining SDS-PAGE. However, on the anti-his-tag western blot, more GAPOR was observed in the supernatant from the anaerobic condition than that from the aerobic condition. Moreover, more GAPOR inclusion bodies seem to be present in the insoluble phase of the lysed cells. This may lead to concluding that less GAPOR was produced as inclusion bodies under anaerobic conditions, conducting to more soluble enzymes. One of our hypotheses could explain this: the iron-

sulfur cluster could be oxidised and increase the possibility of inclusion bodies formation (Messens & Collet, 2006).

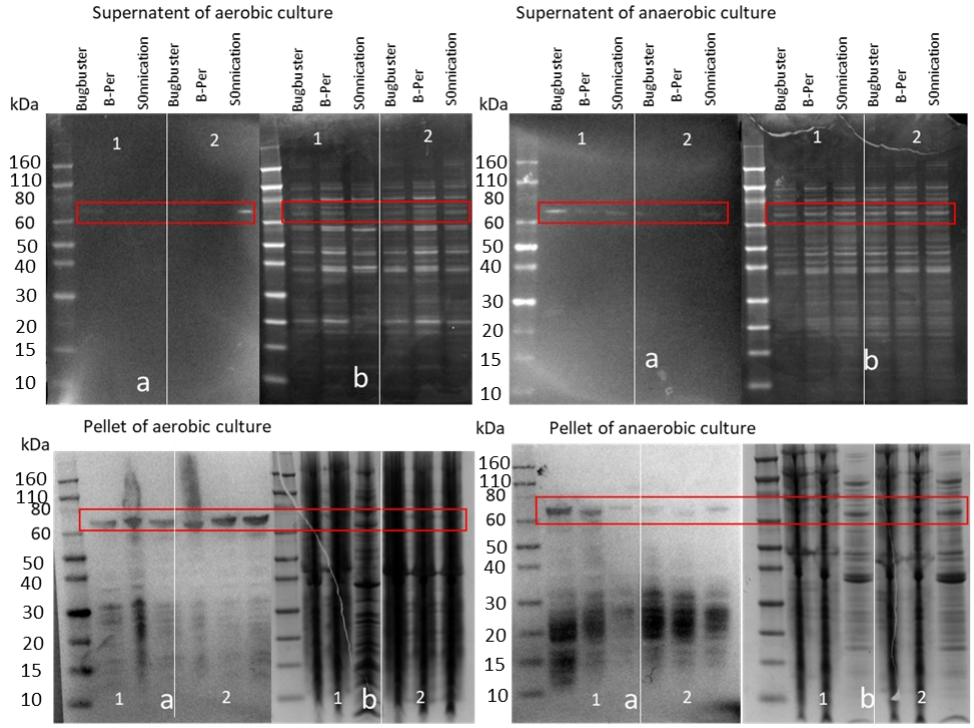


Figure III-5 Comparison of GAPOR production and extraction from aerobic and anaerobic culture using three methods of lysis.Western blot using anti-Histag HRP conjugated antibody (HISP12-HRP Alpha diagnostic international) and 3,3',5,5'-Tetramethylbenzidine (TMB) (Sigma-Aldrich) as colourimetric substrate (a) and Coomassie blue-staining SDS-PAGE (b) of supernatants and pellets of aerobic and anaerobic cultures at room temperature of *E. coli* Rosetta-gami 2 (DE3) pET28a (+)-*gorS2* (*Kan*^{R}) (1) and *E. coli* Rosetta-gami 2 (DE3) *ΔmobAB* pET28a (+)-gorS2 (*Amp*^{R}) (2). Three different lysis methods, namely two chemicals, BugBuster and B-Per, and sonication, were used on the same cultures. The pattern for the wells is the same for each gel. The bands which correspond to GAPOR are squared in red. The colours of supernatant strained SDS-PAGE gels, and Western blot membranes have been inverted. Moreover, the Coomassie blue gel has not been colourised in blue informatically.

Finally, the overexpression of *gorS2* under micro-aerobiosis seemed to be a better option. Indeed, it seemed to be the condition allowing the best production of soluble GAPOR. It is noteworthy that, there were no major differences between the three different lysis methods. Thus, the choice has been made to use BugBuster which seemed to be the most convenient in the anaerobic tent, due to the volume used at the time

To enhance protein overproduction in the soluble phase, the decrease of the culture temperature to 16°C, was tried. Thus, *gorS2* under room temperature conditions was probably expressed without induction. So, due to the time

taken by the cell growth at 16°C ($OD_{600\ nm}$ of 0.492 to 0.772 after 64 hours, measures were done with the Jenway spec.) it was decided to do no induction in the culture with IPTG. To control the background due to the absence of induction an *E. coli* Rosetta-gami 2 (DE3) pET28a (+) (Kan^R) (RGP) has been constructed.

The transcription system used is directed by LacI, the *lac* operon repressor. The *lac* operon comprises genes for lactose uptake and metabolism. LacI allows a very low constitutive expression of the *lac* operon genes to save energy for the cell while keeping the ability to incorporate and metabolise the sugar. LacI is a homotetramer that binds the *lacO* binding sequence to repress the expression. Lactose or allolactose molecules, such as IPTG, can bind the protein to lift the repression by decreasing the protein affinity of LacI for *the lacO* DNA sequence, leading to a full expression. The basal constitutive is often called leakiness(Gatti-Lafranconi, Dijkman, Devenish, & Hollfelder, 2013; Wilson, Zhan, Swint-Kruse, & Matthews, 2007). The repressor leakage was visible on the Western blot (Figure III-6) for both temperatures (RT° and 16°C). However, due to the low OD of the culture at 16°C the quantity of GAPOR produced was too low to be noticeable on the Coomassie blue stained SDS-PAGE. In 48h of growth at room temperature (24h after the induction), RG culture reached OD_{600nm} 3.3 compared to 1.8 after 72h of incubation at 16°C (both times were the harvesting point). (OD measured with the Jenway spectrophotometer). On the other hand, for the room temperature cultures induced and non-induced, GAPOR corresponding bands were visible on the membrane as well as on the gel. Although the GAPOR was constitutively produced without induction, the addition of 0.1 mM of IPTG helped the overproduction of the target enzyme. Moreover, unlike the condition used in the experiment presented in Figure III-5, the cells were resuspended in 0.5 ml of BugBuster instead of 2.5 ml. Thus, the proteins were more concentrated, and more easily observed on the gel (His-Tagged protein bands were visible around 30 kDa). It has been supposed that those bands (marked with an arrow) corresponded to degraded protein containing the N-terminal Histag.

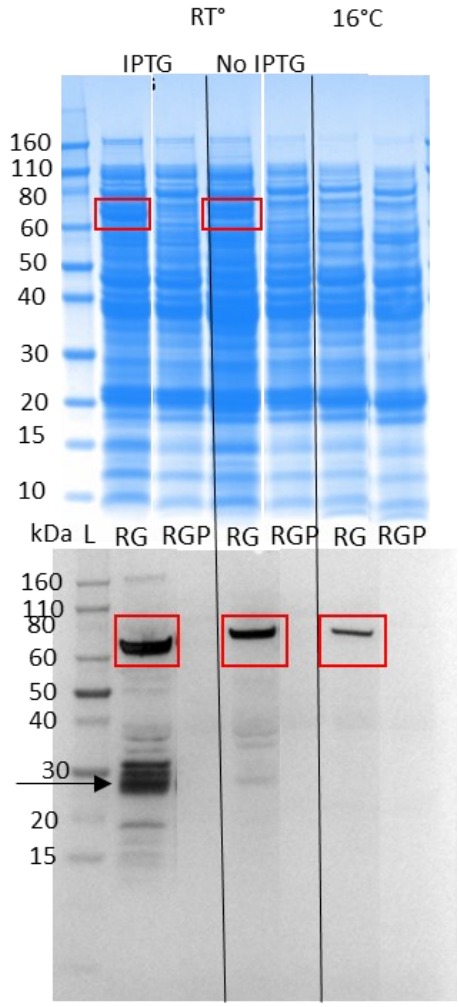


Figure III-6 Comparison of GAPOR production at room temperature and 16°C under anaerobic conditions.The cultures were performed in 25 ml of MAC medium in sealed serum bottles. At room temperature, the cultures were incubated for 24h after inoculation and then induced with 0.1 mM IPTG. The control cultures were not induced. Due to the time taken by the cell growth at 16°C (OD600 nm of 0.492 to 0.772 after 64 hours), it was decided to do no induction in the culture with IPTG. Coomassie blue stain SDS-PAGE (top) and Westernblot anti-Histag of *E. coli* Rosetta-gami 2 (DE3) pET28a (+)-*gorS2* (*Kan*^{R}) (RG), *E. coli* Rosetta-gami 2 (DE3) *ΔmobAB* pET28a (+)-*gorS2* (*Amp*^{R}) (RGΔ) and *E. coli* Rosetta-gami 2 (DE3) pET28a (+)- (*Kan*^{R}) (RGP) lysed with BugBuster. The arrow points to 6-histag containing protein bands that were supposed to be degraded GAPOR.

The overproduction of GAPOR has been observed at 16°C. However, due to the slow growth, it was impossible to proceed with this condition. Thus, the GAPOR overproduction at room temperature with 0.1mM IPTG induction was in better condition despite a potentially higher proportion of inclusion bodies.

2.3. Mutation to improve the GAPOR production

2.3.1. Deletion of *iscR* to improve the insertion of the Iron-sulfur cluster in the GAPOR

There is still a non-negligible fraction of the produced GAPOR in the insoluble fraction of the lysate. One possible reasons of the insoluble GAPOR is the formation of protein aggregates due improperly folded protein. Indeed, the presence four cysteines in the GAPOR for coordinating of the iron-sulfur cluster could create disulphide bonds between apo-GAPORs (Messens & Collet, 2006). To reduce aggregate formation, increasing the availability of the iron-sulfur cluster might help to improve the incorporation in the GAPOR, preventing the cysteine from interacting.

It has been shown that the deletion of the *iscR* gene in *E. coli* encoding the repressor of the *isc* operon helped the production of active recombinant hydrogenase (Akhtar & Jones, 2008).

Hence, *iscR* was deleted by P1 transduction in the *E. coli* Rosetta-gami2(DE3), the strain used for the GAPOR overproduction, as shown in Figure III-7 a,b. To be able to use the same plasmid to compare the evolution, the kanamycin selection marker in the strain had to be removed by FLP recombination (Figure III-7c).

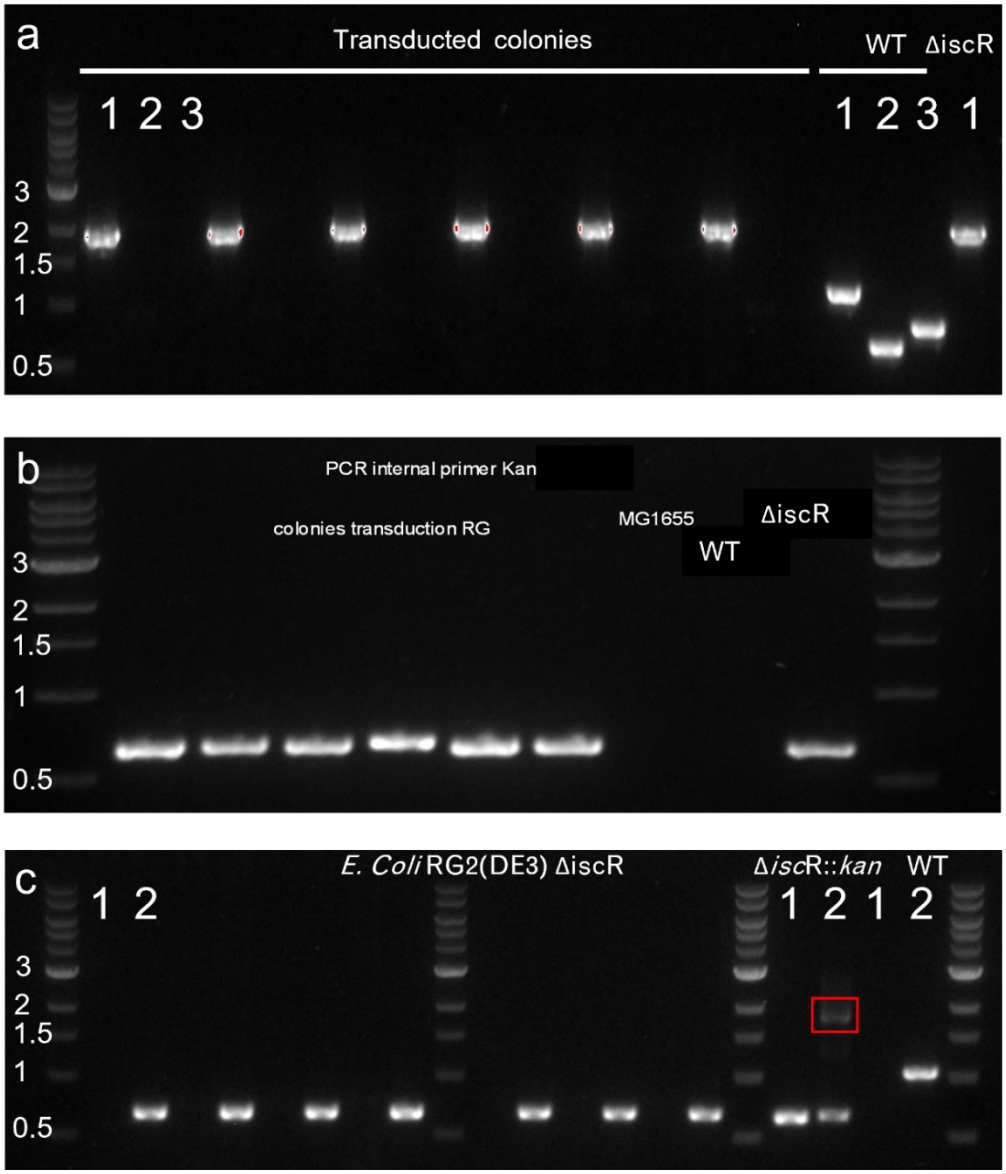


Figure III-7 Deletion of *iscR* in *E. coli* Rosetta-gami2(DE3) by P1 transduction using *E. coli* BW25113 *ΔiscR*::*kan*.(a) Colony PCR of *E. coli* Rosetta-gami2(DE3) *ΔiscR*::*kan* growing on LB kanamycin plates after P1 transduction , WT: *E. coli* Rosetta-gami2(DE3), and *ΔiscR*: *E. coli* BW25113 *ΔiscR*::*kan* using primer pairs of 1: external primers; 2: external forward, internal reverse; 3: internal forward, external reverse, the expected fragment size for the WT control was 1044 bp, 810 bp and 742 bp respectively. (b) Colony PCR of transducted colonies using two kanamycin internal primers to confirm the presence of the kanamycin marker. MG1655: *E. coli* K-12 MG1655, WT: *E. coli* Rosetta-gami2(DE3), and *ΔiscR*: *E. coli* BW25113 *ΔiscR*::*kan*; the expected fragment size was 638 bp. (c) Colony PCR of *E. coli* Rosetta-gami2(DE3) *ΔiscR,* which grew solely on LB plates without antibiotics after the kanamycin cassette removal by FLP recombination. 1: internal kanamycin primers; 2: external primers of *iscR*. WT: *E. coli* Rosetta-gami2(DE3), and *ΔiscR*::*Kan*: *E. coli* BW25113 *ΔiscR*::*kan* were used as control. The expected fragment size for *ΔiscR*::*Kan* was 638 bp which is identical to the band on gel (a) (band in the red square); while a fragment size of 1044 bp was expected for WT. The absence of amplification for the reaction 1 and the smaller fragment size compared to the controls confirmed the deletion of *iscR*.

Once the *E. coli* Rosetta-gami 2(DE3) Δ*iscR* pET28a(+)*gorS2* (kan^R) (named RGΔ*isc*R_gapor) strain was constructed, production assays were conducted (Figure III-8). To be able to determine the potential changes in the soluble

GAPOR production level, the crude extracts of *E. coli* RG_gapor and RGΔ*isc*R_gapor were purified using 1 ml Histrap column and using syringes to operate the column. Both strains were cultured in the same condition and harvested at similar OD. 600 ml of culture split into 6 of 100ml cultures in sealed serum bottles. The results of the purification are shown in the figure below (Figure III-8 a).

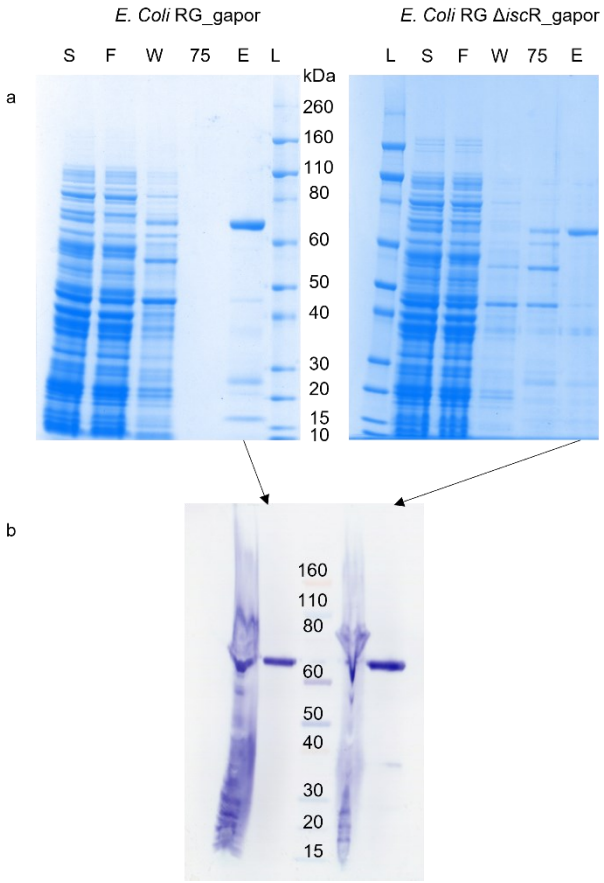


Figure III-8 GAPOR purification from newly constructed strain *E. coli* Rosetta-gami 2 (DE3) *ΔiscR* pET28a (+)-*gorS2* (*Kan*^{R}) (*E. coli RGΔiscR*_gapor) and *E. coli* Rosetta-gami 2 (DE3) pET28a (+)-*gorS2* (*Kan*^{R}) (*E. coli* RG_gapor).The proteins were produced from 6 100 ml MAC media complemented with 100 µM molybdate cultures in sealed serum bottles. The cultures were incubated at room temperature with agitation for 10 h, induced with 0.1 mM IPTG and incubated overnight. The cells were harvested and lysed following the routine protocol. The GAPOR was purified from both cultures following our standard affinity purification protocol. (a) SDS-PAGE blue-silver Coomassie stained of S: lysis supernatant mixed 1:1 with 20 mM imidazole purification buffer; F: flowthrough (12 ml); W: column wash with 20 mM imidazole (15 ml); 75: column wash with 75 mM imidazole (5 ml); E: GAPOR elution in 5ml of 125 mM imidazole and concentrated down to ca. 200 µl using a viva-spin2- 10000 MWCO; L: Novex sharp pre-stained Protein ladder. The purified GAPOR band is clearly visible at ca. 70 kDa for both E lanes. (b) Western-blot anti-His-tag HRP-conjugated revealed with TMB for membrane of both concentrated purified GAPOR solution (marked with arrows).

Along with the supernatant samples, resuspended pellet samples were run on SDS-PAGE and analysed on an anti-histag western blot. No major observable

differences in GAPOR protein concentration between RGΔiscR_gapor and RG_gapor strains were observed on the Western Blot anti-His-tag (Figure III-8). The Bradford assays performed on the purified and concentrated GAPOR elution solution from RGΔiscR_gapor culture affinity purification showed a slight increase in the protein concentration (103.096 µg/ml for RG and 134.349 µg/ml for RGΔ*iscR*). The normalised concentration with the average final optical density ($0.92 \pm 7.45\times10^{-3}$ for RGΔiscR_gapor and $0.85 \pm 4.6\times10^{-2}$ for RG_gapor) confirmed this slight protein concentration increase to 146.03 µg/ml/OD_{600nm} for RGΔiscR_gapor compared to 121.29 µg/ml/OD_{600nm} for RG_gapor.

The quantity of protein produced under anaerobic conditions was low, about 27 µg of GAPOR, both degraded and nondegraded forms from 600 ml of culture. This low yield could be explained by the final OD_{600nm} of the production culture, which was rarely above 1.1. Moreover, the OD_{600nm} increase during the induction phase (16 h) was also very low, from 0.8 to 1.1.

To understand why the growth was slow, it was decided to perform a growth curve experiment under aerobic conditions. Indeed, by removing the main stress condition, it was easier to check if there were any problems with the strains. It appeared that 4 hours after the induction, the three strains, including the control strain RGP, which contains the pET28a(+) empty plasmid, exhibited slowdowns in growth by when the inducer was added (Figure III-9). Therefore, the IPTG seems to slow down the growth regardless of GAPOR production. Hence, in addition to the slow growth due to the anaerobic conditions, the growth inhibition by the IPTG is clearly a reason for the low GAPOR production yield.

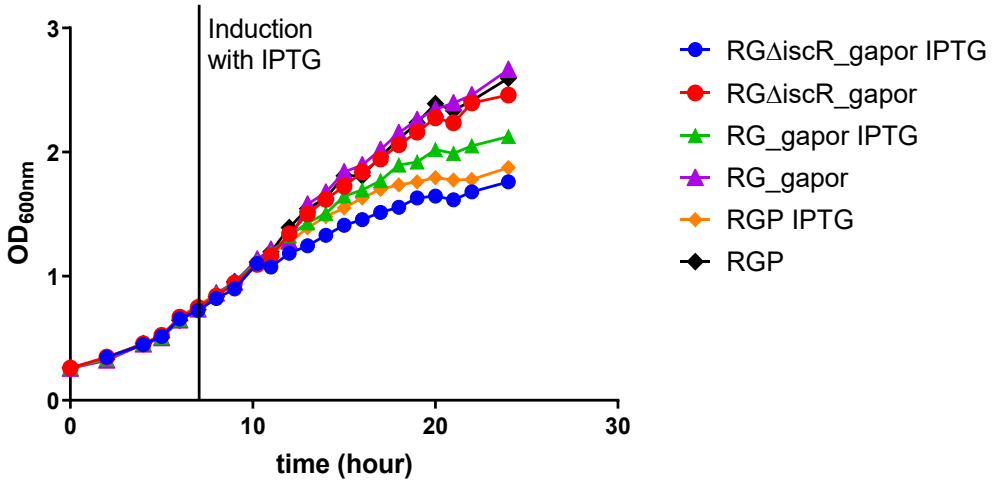


Figure III-9. Growth curves of *E. coli RGΔiscR*, Rosetta-gami 2(DE3) *ΔiscR* pET28a(+) *gorS2* (kan^R) ; RG , Rosetta-gami 2(DE3) pET28a(+) *gorS2* (kan^R); RGP, Rosetta-gami 2(DE3) pET28a(+) (kan^R) with and without IPTG induction. Samples were taken every hour for 24h. In the case of induction, 0.1 mM IPTG was used after 6h incubation.

The deletion of *iscR* was a good way to increase the amount of GAPOR in the soluble phase of the lysate, even though the increase is weak. Also, under aerobic growth, no difference was observed in the growth curves between *ΔiscR* mutant strains producing the GAPOR and the *E. coli* rosetta-gami2(DE3) wild type with or without the GAPOR gene present on the pET28a(+) plasmid. There is no observable difference in the growth between the *iscR* mutant and the WT during the GAPOR production after IPTG induction. Hence the increase of soluble GAPOR does not seem to disturb the *iscR* mutant growth significantly. Thus, it can be supposed that the increase of soluble GAPOR amount in the mutant strain can be mainly due to a "transfer" of protein from insoluble to the soluble phase compared to the WT strains rather than an increase of the protein production, which could have slow down the growth. The GAPOR production with Δ*iscR* mutant seems to be sufficient to perform enzyme assay on crude extract to analyse if the recombinant GAPOR is active.

3. Enzyme assay with all the mutations in *E. coli* to reduce unspecific BV reduction in crude extract

3.1. Preliminary assay for the detection of GAPOR activity in *E. coli* crude extract

Enzyme assay on lysed crude extract was supposed to be a relativelyr fast way to control if the GAPOR produced in our strain had a G3P-dependent activity. It was also supposed that this type of assay would allow getting insight into the *in vivo* activity of the GAPOR. Moreover, it was important to test the possibility of observing any potential G3P-dependent GAPOR activity in the crude extract with the objective of testing a small library of molybdenum cofactor biosynthesis pathway mutants.

In these preliminary GAPOR assays, the reaction tested was the oxidation of G3P to 3PG. This kinetic assay was designed to use benzyl viologen (BV) as GAPOR's final electron acceptor instead of ferredoxin, the natural electron acceptor (Figure III-10). Indeed, the reduced BV turns blue from colourless. This colour change can be followed by a spectrophotometer at 600 nm and used to monitor the GAPOR catalytic reaction (Park et al., 2007). These preliminary assays were performed with the crude cell lysate supernatant. The assay on purified enzyme would be performed on the strain that would be found active.

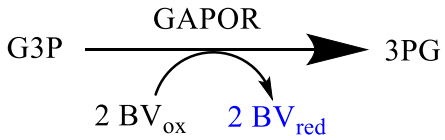


Figure III-10 Scheme of the reaction of G3P oxidation to 3PG catalysed by the GAPOR using oxidised benzyl viologen (BV_{ox}). The reaction is followed at 600 nm by following the reduction of the BV, which turn blue when reduced.

Figure III-11 shows the kinetic curves of the BV reduction enzyme assay that was performed with 10 µl crude extract for RG and RGP strain as a negative control. The control with no crude extract did not display any reduction of BV, implying that there is no self-reduction or reduction by the reaction mix. Also, the BV is similarly reduced by both crude extracts, whether the GAPOR is produced or not. Hence, the GAPOR had no significant impact on the BV

reduction. However, the reduction of BV by crude extracts of the negative control strain indicates that there is a strong background which could potentially mask the GAPOR activity.

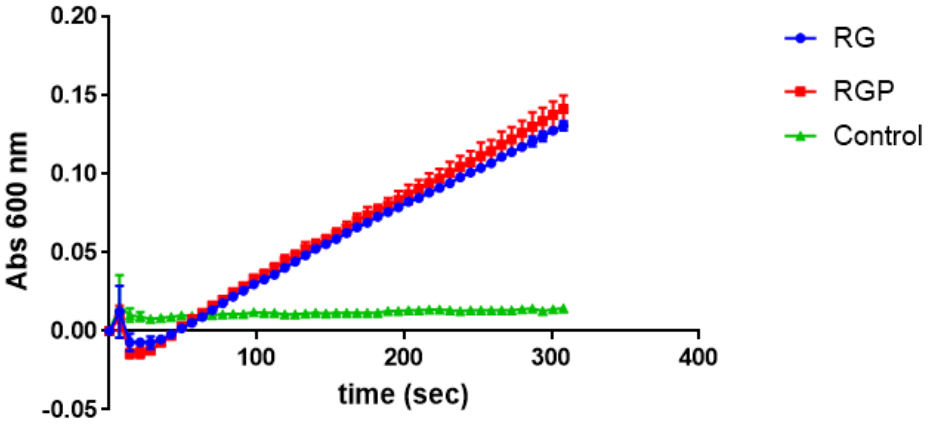


Figure III-11. Kinetic measurement of the BV reduction enzyme assay on crude extractof *E. coli* rosetta-gami 2(DE3) pET28a(+) *gorS2* (*kan*^{R}) (RG, bleu) and *E. coli* rosetta-gami 2(DE3) pET28a(+) (*kan*^{R}) (RGP, red) and the control (green) for what the crude extract was replaced by deionised water. The kinetics have been followed with the reduction of Benzyl viologen at 600 nm for 5 min. The cells were grown anaerobically, and the cell-free extracts have been used to start the reaction. 10 µl of crude extract was added to 500 µl of reaction mix and 490 µl of deionised water for a total volume of 1 ml (60 µM G3P, 3 mM BV, 56 µM Na_2MoO_4, 50 mM EPPS, pH8.4).

3.2. Strain construction to diminish the BV reducing background

As shown by the preliminary enzyme assays (Figure III-11), there was a strong background of BV reduction activity in *E. coli*. This background activity prevented the detection of the GAPOR activity via the utilisation of BV as the final electron acceptor. Two main BV-reducing enzymes were identified in *E. coli*: 1) hydrogenase and 2) formate dehydrogenase. First, there are three hydrogenases in *E. coli*, which all belong to the [NiFe] family (Blokesch et al., 2004). HypF is a carbamoyltransferase coded by *hypF* and is responsible for the maturation of the three *E. coli* hydrogenases (Hyd1-3) (Petkun et al., 2011). It has been demonstrated that when *hypF* is mutated, the strain shows a hydrogenase$^-$ phenotype (Maier, Binder, & Böck, 1996; Paschos, Bauer, Zimmermann, Zehelein, & Bock, 2002) and the loss of *[hydrogenase]* phenotype was confirmed by Guerrini (Guerrini, 2007). The second is the formate dehydrogenase (FDH), a membrane-bound, molybdo and

selenocysteine-dependent enzyme (Mihara et al., 2008). It has been demonstrated by Mandrand-Berthelot that the FDH_H, in combination with Hyd3 in the formate hydrogen lyase (FHL) reduces the BV *in vivo* (Mandrand-Berthelot, Wee, & Haddock, 1978; Wu & Mandrand-Berthelot, 1987). It has been demonstrated by (Mihara et al., 2008) and Mandrand-Berthelot (1978) the deletion of *selA*, formerly called *fdhA*, coding for the selenocysteine synthase was disabling the FDH_H activity. Thus Δ*hypFselA E. coli* mutant strains were constructed sequentially by P1 transduction from Keio collection strains in *E. coli* Rosetta-gami 2(DE3) *ΔiscR* strain. The kanamycin cassettes were both removed Figure III-12.

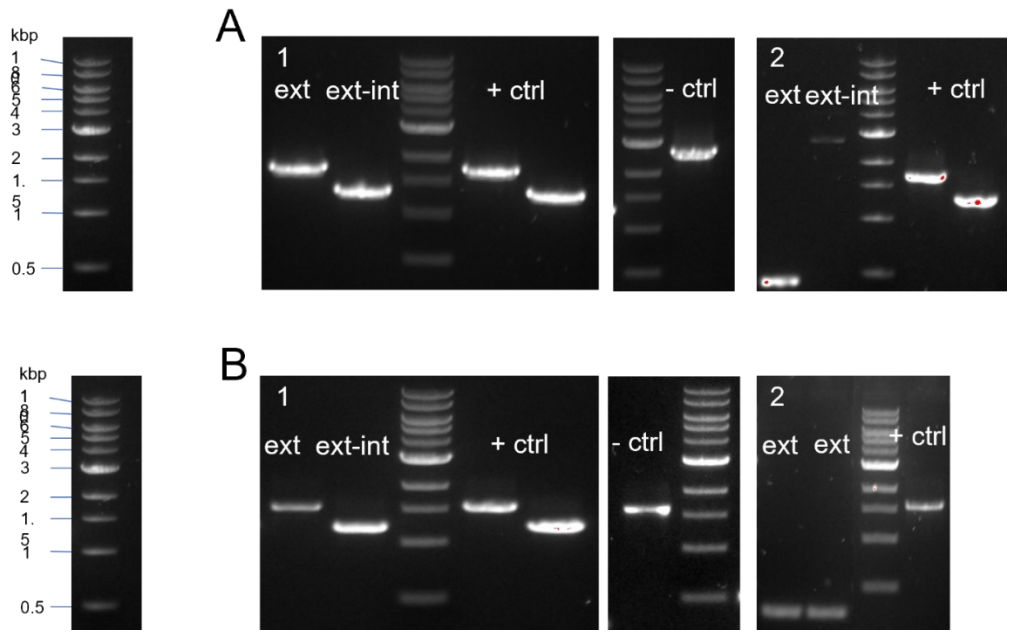


Figure III-12 Colony PCR of the deletion of *hypF* and *selA* in *E. coli* RGcoli RG*ΔiscR* by P1 transduction. A: 1: Deletion of *hypF* in *E. coli* RG*ΔiscR* ext: colony PCR with two external primers, the fragment size for an *hypF*$^+$ strain is 2590 bp (-ctrl), *E. coli* BW25113 *ΔhypF*::*kan* (Keio collection) was used as a positive control for the deletion and insertion of the kanamycin cassette (+ctrl). 2: the same two pairs of primers were used to control the removal of the kanamycin marker by FLP recombination. The antibiotic marker removal was visible thanks to the DNA fragment in the external-primers amplification. The band observed in the internal-external amplification is assumed to be an unspecific amplification. B: Control of the deletion of *selA* in *E. coli RGΔiscRΔhypF* by P1 transduction from *E. coli* BW25113 *ΔselA*::*kan*. 1: colony PCR of a transducted colony using two pairs of primers similar to the PCR performed in A, *E. coli* BW25113 *ΔselA*::*kan* was used as deletion positive control (+ctrl); The expected fragment size of the *selA*$^+$ strain used as a negative control was 1606 bp (-ctrl). 2: The removal of the kanamycin was only controlled by PCR amplification with external primers. The fragment obtained for the tested colonies was lower than both controls. To note, the colonies controlled by PCR amplification were chosen for both experiments after displaying growth only on non-selective plates.

3.3. Enzyme assay of GAPOR produced in [BV-reducing]⁻ *E. coli* strains.

3.3.1. Verification of the background reduction

To test the loss of the BV-reducing phenotype, new BV-reducing assays were performed using crude extract (Figure III-13).

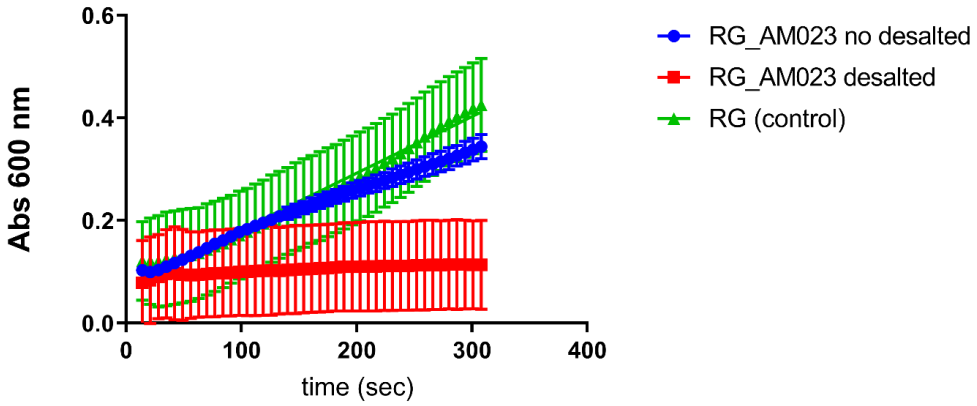


Figure III-13 Kinetic curves of BV reduction assay after deletion of the *hypF* and *selA*. Kinetic curves of BV reduction in presence of 100 µl of crude extract of *E. coli* Rosetta-gami2 (DE3) *ΔiscRΔhypFΔselA* (RG_AM023). *E. coli* RG *ΔiscR* pET28a(+) *gorS2* (RG) was used as a negative control. The mutant strain crude extract was also assayed after desalting with zebra spin desalting column 7KMWCO 2 ml, and the control was not desalted.

As shown in Figure III-13, there was no significant modification of the GAPOR-independent background BV-reducing activity after the deletion of the two genes compared to the original strain RG. However, the most important result of this experiment was that most of the BV-reduction activity came from the presence of residual metabolites and media traces. Indeed, after desalting the lysate supernatant, no absorbance change was measured, and no change of colour was visible even after 5 min. Hence, even with the presence of H_2 in the tent atmosphere, the hydrogenase did not seems to impact as much than the other reactions.

Even if the deletion performed were potentially not mandatory, an expression strain was constructed to perform GAPOR enzyme assay on crude extract. Moreover, this experiment gave the condition to test the GAPOR activity on crude extract for the next experiments and mainly for testing of the molybdenum cofactor biosynthesis pathway library.

3.3.2. G3P-dependent activity assay

After determinating of the condition to reduce the BV-reducing background activity, GAPOR was produced in *E. coli* RG_AM023. The same type of enzyme assay on crude extract was performed (Figure III-14). However, the sampling time was extended. Indeed, it was supposed that the reaction rate would be very low as there was not even a slight perturbation of the BV-reduction rate. The reactions were started by adding of the crude extract instead of the substrate (G3P at 30 µM) as it was more convenient for the reaction solution homogenisation.

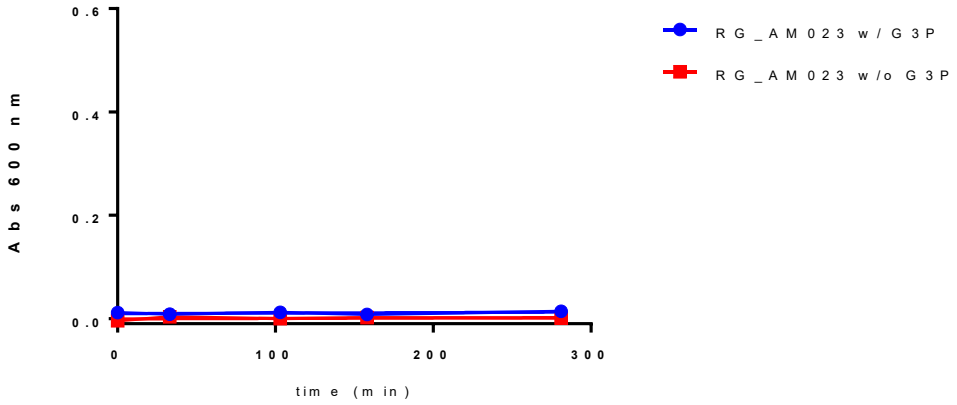


Figure III-14 G3P oxidation by GAPOR produced by *E. coli* Rosetta-gami2 (DE3) *ΔiscRΔhypFΔselA* pET28a(+) *gorS2* (RG_AM023_gapor) followed by the reduction of BV. The crude extract samples were prepared identically to the desalted sample presented in Figure III-13. No changes were brought to the reaction mix for this experiment (60 µM G3P, 3 mM BV, 56 µM Na_2MoO_4, 50 mM EPPS, pH8.4), though the data were collected on a longer time frame. The GAPOR activity was tested in the presence and absence of 30 µM G3P in the reaction mix buffer.

The results of these experiments show an absence of BV reduction activity, consequently, no GAPOR activity was detected, whether G3P was in the reaction mix or not. Moreover, Guerrini showed that GAPOR from *P. furiosus, M. janshii and M. maripaludis* had a hydrogenase-like activity in the presence of H_2, shown by the reduction of BV (Guerrini, 2007). Although the atmosphere of the Coy anaerobic tent is composed of 2%-3% of H_2, no BV-reducing were observed. Hence, even if the deletion of *iscR* increased the quantity of soluble GAPOR, the GAPOR produced were not active with either of the known activities.

The first modifications on the host *E. coli* strains were more or less successful improving the recombinant gene expression and removing background BV reduction activity on crude extract assay. All the enzyme assays were carried out in the Coy anaerobic tent in Nottingham. However, there was concerns about too high oxygen concentration in the tent atmosphere. It was the principal reason why the assays were performed on crude extract and not with purified GAPOR, even though the protein was already tagged with a poly-histidine marker. The presence of oxygen might have to effect: i) inactivation of the oxygen-sensitive GAPOR by the oxidation of the [4Fe-4S] cluster or the Moco (Roy et al., 2001), ii) re-oxidation of BV during the assay. To ensure that oxygen would not be a problem during the later assay, they were performed in TBI, INSA Toulouse in their anaerobic cabinet.

The assay was performed with the same GAPOR-producing strain, and a plasmid-less strain was used as a negative control. Note that in Toulouse the lysis was performed by sonication, not BugBuster. Three different volumes of crude extract were tested 250 µL, 125 µL and 60 µL. As shown in Figure III-15, the BV reduction activity can be observed for both strains (Figure III-15A). This meant that the activity observed was not GAPOR dependent or the GAPOR enzymatic activity was not strong enough compared to the background activity. Despite the BV-activity, the hydrogenase-like activity was not considered as the cabinet atmosphere was composed of 100% N_2, and the solutions were sparged with N_2 before entering the cabinet.

The BV reduction activity of the crude extract proteins was calculated (Figure III-15 C) to observe a potentially higher reduction rate after normalisation with the quantity of protein added to the reaction. Despite normalisation, no significant difference between the GAPOR producing strain and the control can be observed.

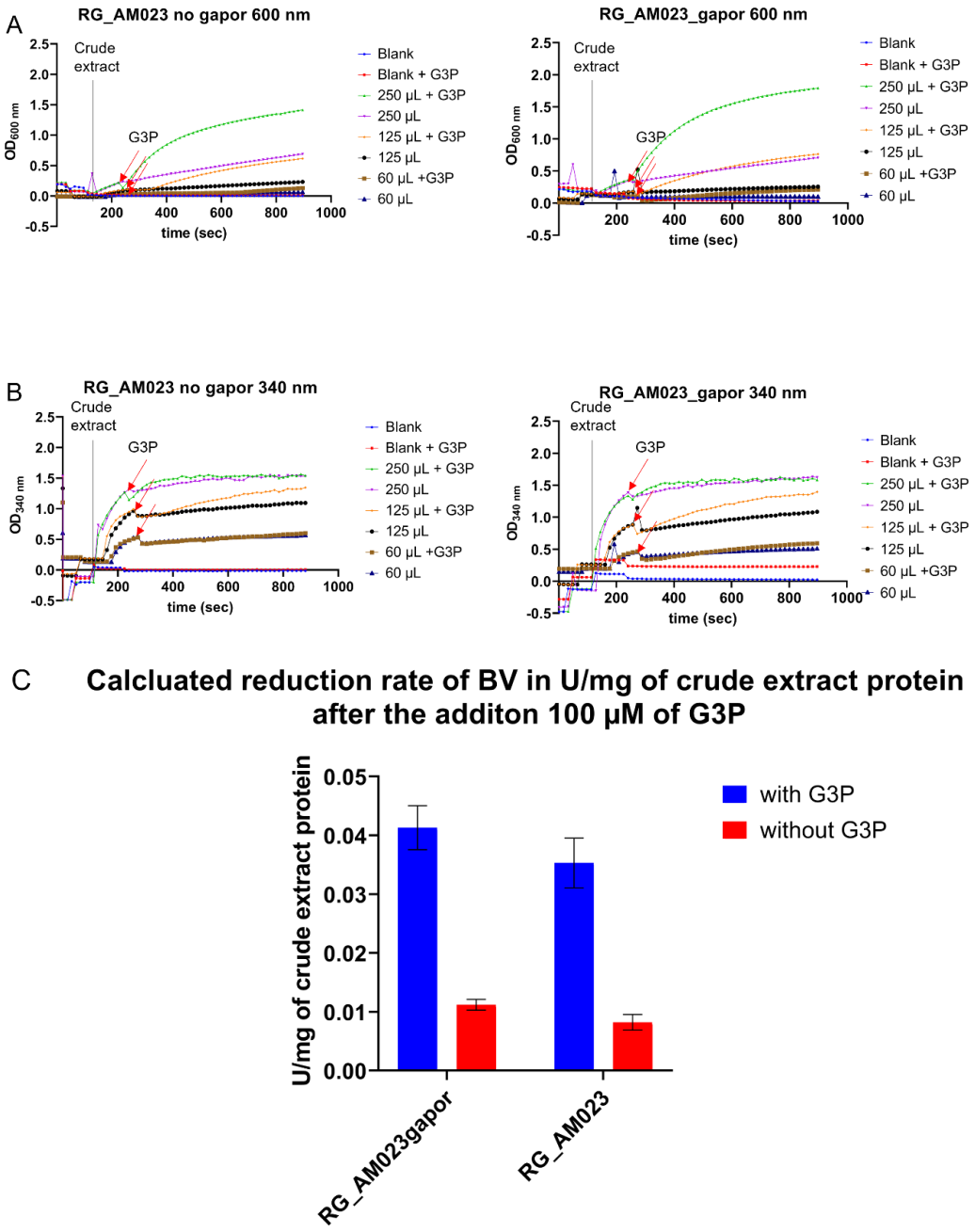


Figure III-15 Kinetic curves following the BV reduction at 600 nm and 340 nm in the presence of the crude extract of *E. coli* Rosetta-gami2 (DE3) *ΔiscRΔhypFΔselA* (RG_AM023) expressing or not the *gorS2* (pET28a(+) *gorS2* : RG_AM023_gapor) and G3P reduction rate per mg of crude extract protein. 500 µl of reaction mix (50 mM EPPS, 3 mM Benzyl viologen and 56 µM Na_2MoO_4) were complemented with 100 µl of sodium dithionite; 50 µl, 175 µl and 220 µl of water for the test with 250 µl, 125 µl and 60 µL of crude extract respectively. The measurement was started when 250 µl, 125 µl and 60 µL of crude extract was added (black vertical bar). For the blank, the crude extract was replaced with water. The reactions were started by adding 100 µl of 1 mM G3P (final concentration 100 µM) or water for the control without G3P. the reaction was monitored for 15 min, and a full spectrum was measured every 16 sec. All the conditions were measured at the same time. A: Kinetic curves following the reduction of BV at 600 nm. B: Monitoring of the reactions at 340 nm following the evolution of NAH concentration. C: Calculated reduction rate of BV in U/mg of crude extract protein. The ΔOD was measured by the software controlling the spectrophotometer using the data collected in Figure III-15 A. The total protein concentration was measured by BCA titration of the crude extract RG_AM023gapor 2.559 ± 0.162 mg.ml-1 and RG_AM023 2.692 ± 0.135 mg.ml-1.

The BV-reduction activity seems to be G3P dependent. Indeed, an acceleration in the reduction of BV can be observed after adding 100 µM of G3P (the final concentration in the reaction solution) (Figure III-15). The spectrophotometer used in Toulouse measured the complete spectrum for each point contrary to the apparatus used in Nottingham. It allowed to follow the reaction at 340 nm. This wavelength permitted us to observe the production of NAD(P)H. On these graphs (Figure III-15 B), a strong production of NAD(P)H is visible after the addition of G3P. The measurement at 340 nm (NAD(P)H) for the 250 µl of crude extract experiment saturates rapidly after the addition of G3P. However, no saturation or few is visible for the measurement at 600 nm (following the BV reduction). With this observation and the absence of saturation for a lower volume of crude extract, it does not seem that the plateau was due to the complete consumption of NAD^+ but to the saturation of the spectrophotometer. It can be supposed that the G3P is oxidised by the glyceraldehyde-3-phosphate dehydrogenase using $NAD(P)^+$ and phosphate that is still present in the extract.

$$G3P + PO_4 + NAD^+ \leftrightarrow 1{,}3BPG + NADH + H^+$$

Furthermore, the increase of the NADH concentration can be observed directly after the addition of the crude extract. This led to the conclusion that desalting was not optimal during the experiment, and enough G3P and NAD^+ remained in the desalted crude extract leading to the BV observed in the figure below. This activity was not visible when the experiment was performed in Nottingham and only appeared in a better-controlled anoxic atmosphere. Indeed, if the reaction rate of BV reduction by this NAD^+-dependent activity was lower than the BV oxidation rate by a trace of O_2 in the Nottingham anaerobic tent, it was not possible to monitor it.

Despite the deletion of genes meant to reduce the background activity all the potential GAPOR-unspecific activity observed in Nottingham, some remained, as the hypothetic reaction presented formerly. Moreover, with a deficiency in removing of all metabolites, it was not possible to observe any GAPOR and G3P dependent BV reduction activity free from background activity. Furthermore,

no significant difference in the activity per mg of total protein was observed, ruling definitely out any potential GAPOR-dependent activity in this experience. Hence it can be concluded that the GAPOR produced in this condition is not active. To test if, without the surrounding reactions, a G3P-dependent activity can be observed, GAPOR was purified and tested.

4. Purification of GAPOR

To observe the GAPOR-specific G3P-dependent BV reduction activity and consequently calculate the kinetic constant of the studied enzyme, it was needed to purify the enzyme. With that in mind, the 6-histidine-tag sequence was introduced at the N-terminal end of the GAPOR and spaced by a thrombin cleavage site (Figure III-16). This tag allowed the purification of the target protein by affinity binding of the 6-histidine tag to the nickel in the nickel affinity chromatography column.

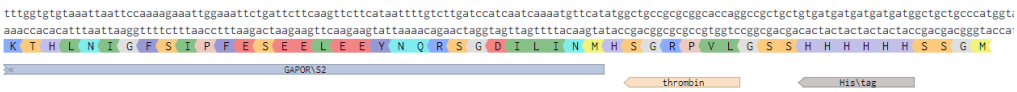


Figure III-16 DNA sequence of the N-terminal end of *gorS2* expressing the fusion protein 6-his-GAPOR.

4.1. Affinity purification

First purification experiments were performed using 1 ml Histrap column (GE Healthcare, nickel affinity column). For those first experiments, all the solutions were manually applied on the column using syringes. SDS-PAGE and WB of purification from anaerobic culture were presented in the section above in Figure III-8. On these figures, bands at lower molecular weight are visible. These bands are also observed on western-blot anti-his-tag. Hence, as previously stated those bands were supposed to be protein degradation containing an intact N-terminal end coeluted with the GAPOR. With future enzyme assays and or cofactor analysis in mind, it was tried to remove those contaminants. Thus, gel-filtration chromatography was tested as a second step of purification.

In this instance, the gel filtration chromatography was only conducted on aerobically produced GAPOR preparation. Indeed, it was the easiest way to

produce a maximum of GAPOR at the time of the experiment. Thus, two cultures of twice 600 ml of the RG*ΔiscR*_gapor strain were grown aerobically to produce about 600 µg of protein. As visible on the SDS-PAGE and Western Blot anti-His-tag (Figure III-17), the GAPOR production on both 1.2 L batches was equivalent. However, a lot of contaminant proteins are visible on the Blue Silver stained SDS-PAGE. Among those contaminations the Western Blot results indicate that most of these contaminations are his-tagged degraded GAPOR retained during the Nickel affinity chromatography (Figure III-17).

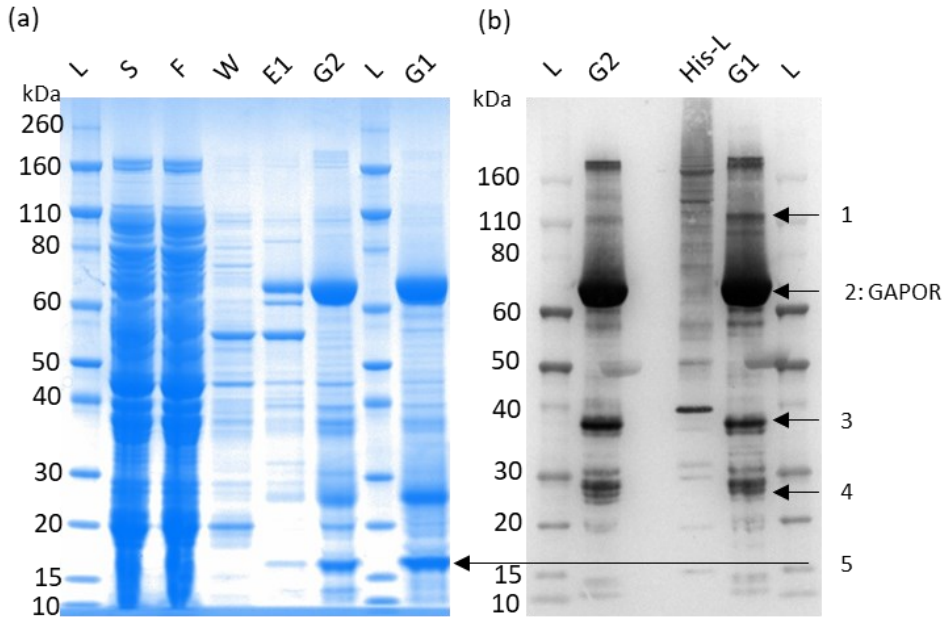


Figure III-17. SDS-PAGE and Western blot GAPOR nickel affinity purification.(a) Blue Silver stained SDS-PAGE of the purification of aerobically produced GAPOR in 600 ml culture of *E. coli* rosetta-gami2(DE3) *ΔiscR* pET28a(+) *gorS2* (*kan*^{R}); (b) Western Blot anti-Histag . L: Novex-sharp protein pre-strained protein ladder (Invitrogen); His-L: His-tagged standard BenchMark his-tagged ladder (Invitrogen) S: cell-free extract; F: flowthrough; W: wash; E1: elution with 75 mM imidazole; G1: GAPOR elution with 125 mM imidazole batch 1; G2 GAPOR elution with 125 mM imidazole batch 2. GAPOR bands are visible at 70 kDa on both (a) and (b). Band 1 might correspond to GAPOR dimer (110 kDa), 3 and 4 to GAPOR degradation (37 kDa and 25 kDa) and 5 to a contaminant protein, which was coeluted with GAPOR (17 kDa).

These degradation bands (Figure III-17) were also observed by Guerrini during his thesis at similar molecular weights (45, 23 and 21 kDa). These degradation bands were observed despite the addition of anti-protease (Guerrini, 2007).

4.2. Gel filtration

To remove these contamination bands, affinity-purified protein solutions were applied to size exclusion chromatography. GAPOR from two different

production batches were separated on a Superdex 200 10/300 on an Äkta pure for size exclusion chromatography. The two runs were completed in a row with a column wash between the runs. First, fractions of GAPOR were eluted from the column after approximately 13 ml of running buffer was passed through the column (Figure III-18). The first chromatogram shows two main peaks. According to the SDS-PAGE, the first peak seems to correspond to the intact GAPOR, whereas the second peak corresponds to the 17 kDa protein signal visible on the control (Figure III-18 lane 1). The last contamination observed in the control lane is a His-tagged GAPOR degradation (25 kDa) poorly separated between the two blocs. Although chromatogram 2 is different, the SDS-PAGEs of fractions are close. The principal difference between the two runs is the small peak near 8 ml after injection. One of the hypotheses was a complex between two GAPOR. However, the protein concentration might have been too low to be detected with the colloidal Blue Silver staining (Figure III-18).

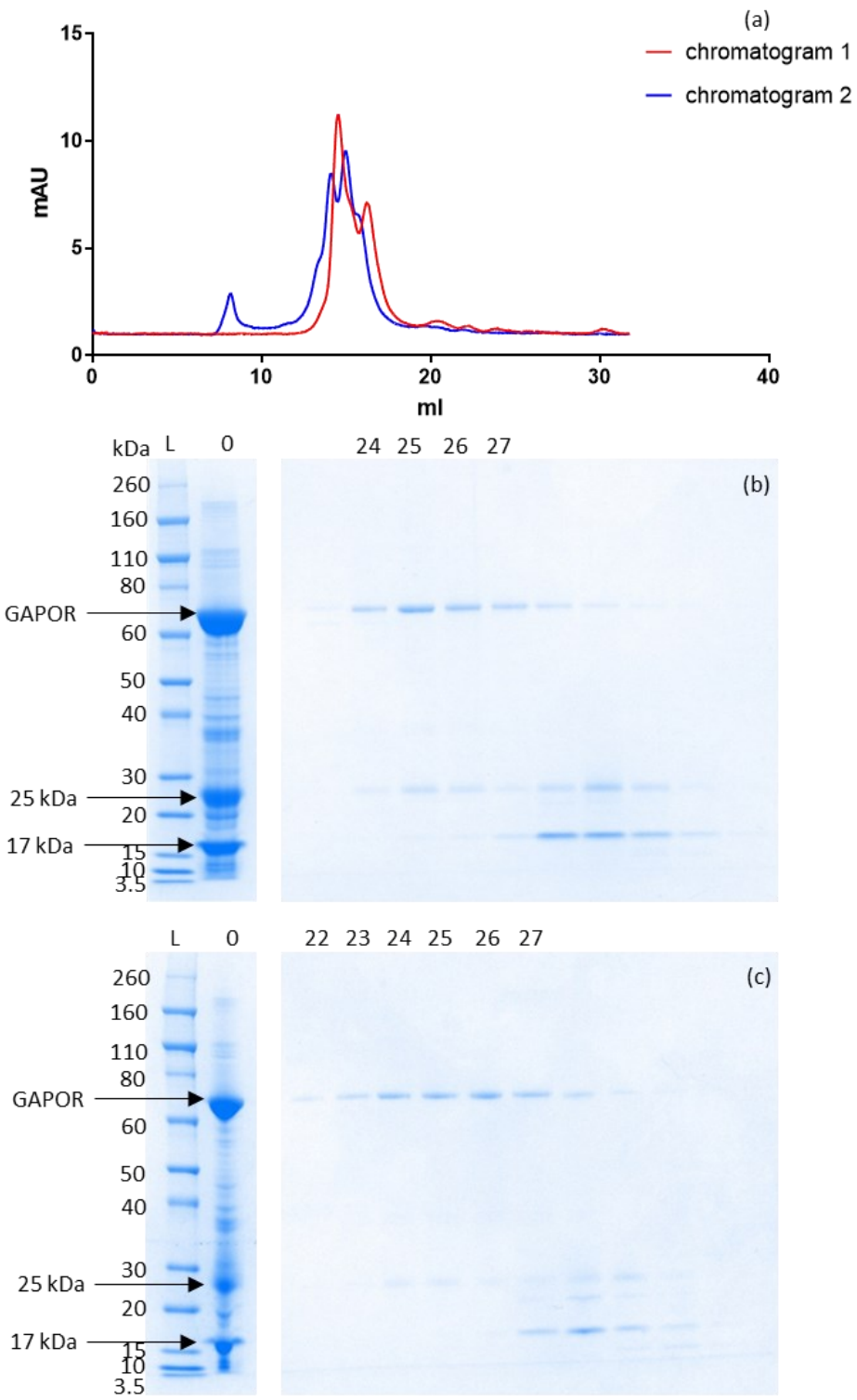


Figure III-18. Size exclusion chromatography of affinity purified GAPOR. (a) chromatogram of the two runs of 200 µl GAPOR solution on Superdex200 10/300, fractions of 0.5 ml ha been collected from respectively 2.8 ml to 22.5 ml and 2.7 ml to 21.7 ml of eluent. (b) (c) Coomassie blue silver strained SDS-PAGE of the collected fractions from both runs (not all data shown). Fractions 24, 25, 26, 27 and 22, 23, 24, 25, 26, 27, respectively, run 1 and 2 were kept and gathered for later concentration. L: Novex-sharp protein pre-strained protein ladder (Invitrogen); 0: Mix of the two solutions injected for chromatography. GAPOR bands are visible at 70 kDa. The main contaminant protein bands are visible at 25 and 17 kDa.

Following this purification all the fractions containing GAPOR were collected in one pool of 5 ml and concentrated 25 times with a 30,000 MWCO Vivaspin 2 ml centrifugal concentrator. Contrary to the first concentration step after the affinity chromatography, where a Vivaspin was used to retain all the protein and remove the imidazole, the molecular weight cut-off is higher to retain only the GAPOR and remove all the low molecular weight protein contaminants. As shown in Figure III-19, the contaminants were still present in the final concentrated solution. Moreover, the quantity of protein is drastically lower. Indeed, the recovery yield after these steps of purification-concentration is about 10%, 200 µl at 0.28 µg/µl. During a purification process, the loss of the product of interest is always important, but, in this case, further investigation need to be conducted to improve this last purification step. The final overall yield of the purification process is 23 µg of GAPOR per litre of culture. To perform certain analysis, such as ICP-MS, hundreds of micrograms of GAPOR are required, implying that under these production and purification conditions from, 20 to 200 L of anaerobic culture. Hence, modification of the expression level, as well as purification technics, need to be performed.

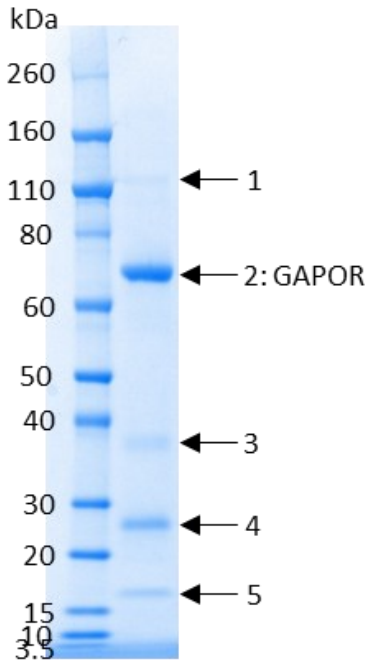


Figure III-19. Coomassie blue silver stained SDS-PAGE of the concentrated fraction with 30,000 MWCO Vivaspin centrifugal concentrator.GAPOR band (noted 2 in the figure) is visible at 70 kDa. At 115, 38, and 25 kDa, noted 1, 3, and 4, contaminant protein bands (assumed to be GAPOR

degradation or aggregate) correspond to bands on WB (Figure III-17 (b)). The band noted 5 at 17 kDa band of non-His-tagged contaminant protein.

After these experiments the gel filtration step is not efficient enough to remove the contaminant. Also, at this stage of the study, where no GAPOR activity was observed, this step was superfluous and potentially increased protein degradation by increasing the purification process time. Hence, no gel filtration was applied to affinity-purified GAPOR for enzyme assay.

4.3. GAPOR affinity purification using an Äkta apparatus

4.3.1. Elution with a gradient elution

As presented above, the first experiment was performed manually, that was preventing to perform a proper gradient elution. Therefore, before going to Toulouse and using the Äkta purifier present in their anaerobic cabinet, a gradient purification was performed to confirm the protocol use and see if any optimisation of the purification protocol could be brought. The purification was performed on 600 ml GAPOR production culture. The elution was performed from 20 mM imidazole to 500 mM imidazole. The chromatogram of this experiment can be observed in Figure III-20. Most of the GAPOR was eluted between fractions 10 to 16 (Figure III-21). These fractions correspond to an elution range between 100 mM to 165 mM of imidazole.

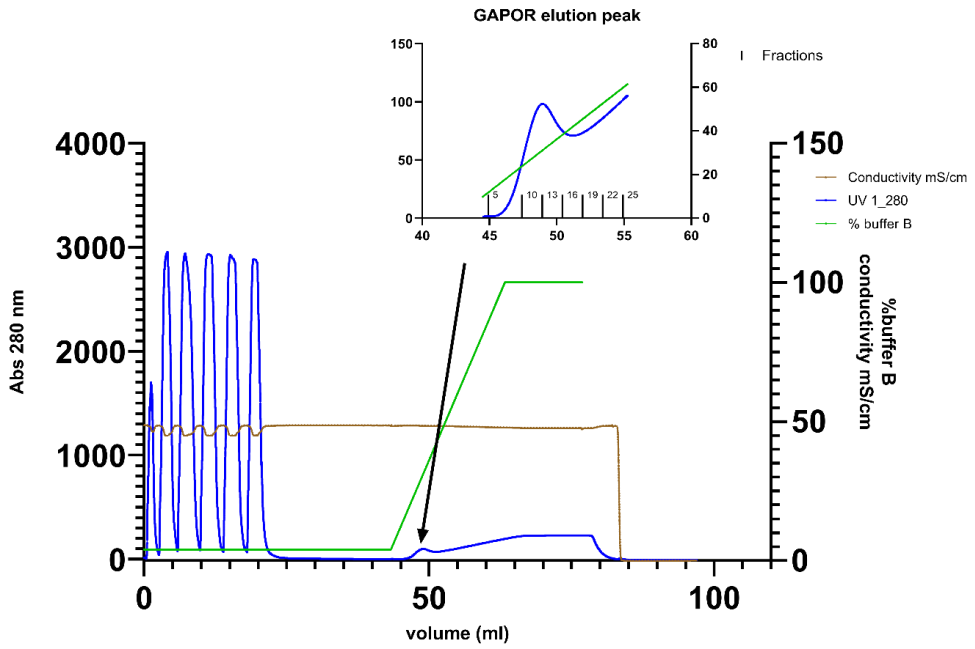


Figure III-20 Gradient elution GAPOR affinity purification with an Äkta pure. 600 ml of cultures of *E. coli* rosetta-gami2(DE3) Δ*iscRhypFselA* pET28a(+) *gorS2*-(*ampR*) were used for this purification. The cultures were performed following the routine protocol. After lysis 6 ml of crude extract was recovered and mixed 1:1 with 20 mM imidazole binding buffer (20 mM sodium phosphate buffer, 500 mM NaCl, pH 8.4) and loaded on the column with a 2 ml loop. The column was washed with 15 ml of 4% buffer B (20 mM sodium phosphate buffer, 500 mM NaCl, 500 mM imidazole, pH 8.4). The proteins were eluted with a 4-100% of buffer B gradient (20 mM to 500 mM imidazole) green curve. The elution was collected in 0.5 ml fraction. The blue curve corresponds to the monitoring of the absorbance at 280 nm to detect the protein leaving the column. In brown: the conductivity of the solution exiting the column. The GAPOR elution peak was magnified where the fractions were run on the gel (Figure III-21) are also indicated.

An important contamination band is visible in fraction 10 (Figure III-21, marked by the red arrow). This contaminant was also visible in the 75 mM imidazole wash in Figure III-17 (a) lane E1). The figure below shows that the 38 kDa and 25 kDa contaminants were eluted simultaneously with the GAPOR. contrary to the 17 kDa contaminant protein that was more retained on the column.

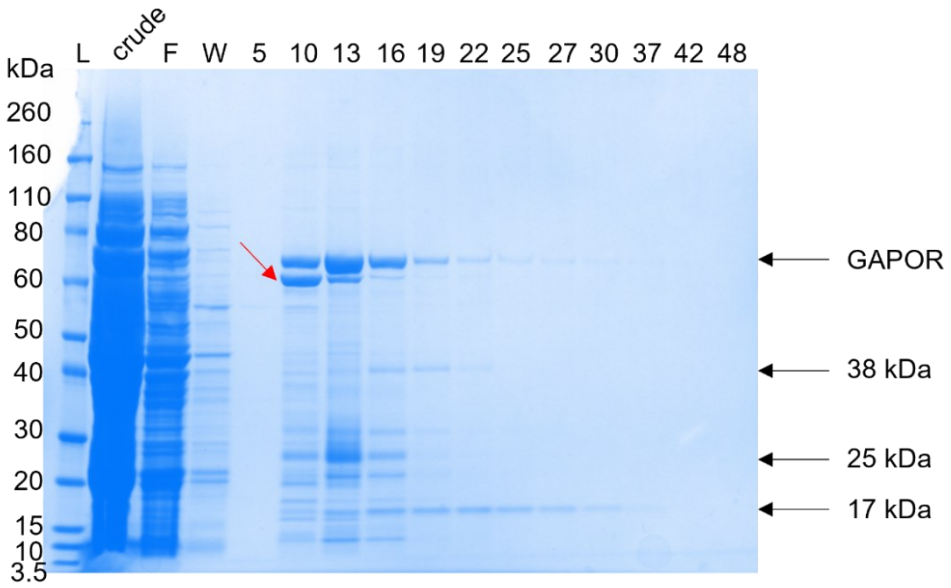


Figure III-21 Coomassie blue-silver stained SDS-PAGE of the sample recovered from the GAPOR gradient elution affinity purification. SDS-PAGE on NuPAGE 4-12% Bis-Tris-acrylamide (Invitrogen) 200 V, 35 min. The crude extract (crude); F: flowthrough (15 ml total); W: column wash; fractions 5, 10, 13, 16, 19, 22, 25, 27, 30, 37, 42, and 48 were also run on this gel (Figure I-20). The GAPOR, as well as the mains contaminants are indicated with arrows. The band indicated with a red arrow correspond to another big contaminant also visible in the 75 mM imidazole wash in Figure III-17 (a) lane E1.

4.3.2. Step elution

With the objective to reduce the contaminant but still collect a maximum of GAPOR in mind, it was decided to keep a similar imidazole concentration for the elution step as previously used. Indeed, as previously shown, with the syringe purification experiment (Figure III-8), the 75 mM imidazole 5 ml wash was sufficient to remove most of the 60 kDa contaminant. Furthermore, 125 mM imidazole was kept for the elution step. Even though increasing the concentration would probably have allowed recovering more GAPOR, more 17 kDa contaminant would have also been collected. Keeping this concentration allowed the possibility of recovering the maximum GAPOR and limiting the contamination by the 17 kDa protein. Also, the amount of protein lost with a higher elution concentration is very low, as visible in Figure III-22.

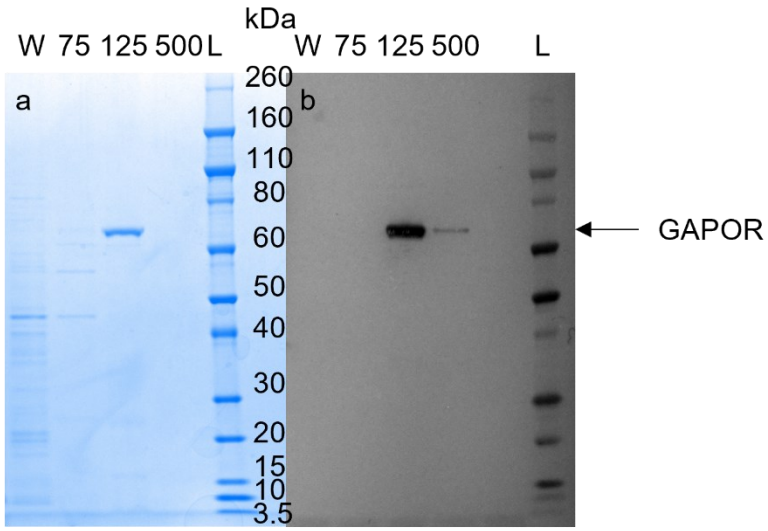


Figure III-22 SDS-PAGE and western blot of GAPOR purification elution process.This figure presents samples of the four main steps of the GAPOR elution process. W: the wash of the column with 20 mM imidazole binding buffer; 75: 5 ml of 75 mM imidazole buffer; 125: The GAPOR was eluted with 5 ml of 125 mM imidazole buffer; 500: wash of the column with 5 ml of 500 mM imidazole. The purification was performed using a 1 ml Histrap column (GE Healthcare); syringes were operated with a syringe-pusher apparatus to apply a steady flow rate. The GAPOR was produced, harvested and lysed with an anaerobic culture of *E. coli* rosetta-gami2(DE3) Δ*iscRhypFselA* pET28a(+) *gorS2*-(*ampR*) following our routine protein production protocol.

The chromatogram of the affinity purification of GAPOR produced with *E. coli* RG_AM023gapor and recovered by a step-elution is presented in Figure III-23. The GAPOR elution peak is clearly visible on the chromatogram despite being rather small compared to the peak observed at 75 mM imidazole wash. Using densitometry measurement, it was possible to estimate that about 52% of the GAPOR was recovered in the elution (269 µg from 517 µg of GAPOR). The remaining was lost during the binding (about 10%) and washes (about 23%). The remaining 15% is supposed to be recovered at the 500 mM imidazole step. The amount of GAPOR lost during the purification was important, but most of the degradation products were removed. In addition, to these results, the possibility to use an Äkta apparatus in Nottingham (aerobic) and Toulouse (anaerobic) helped to get better reliability of the process with controlled flow rate and absorbance monitoring. Moreover, the possibility to work with higher volumes, mainly thanks to the sample pump in Toulouse, where a higher volume (30 ml) was loaded on the column while keeping the crude extract on ice.

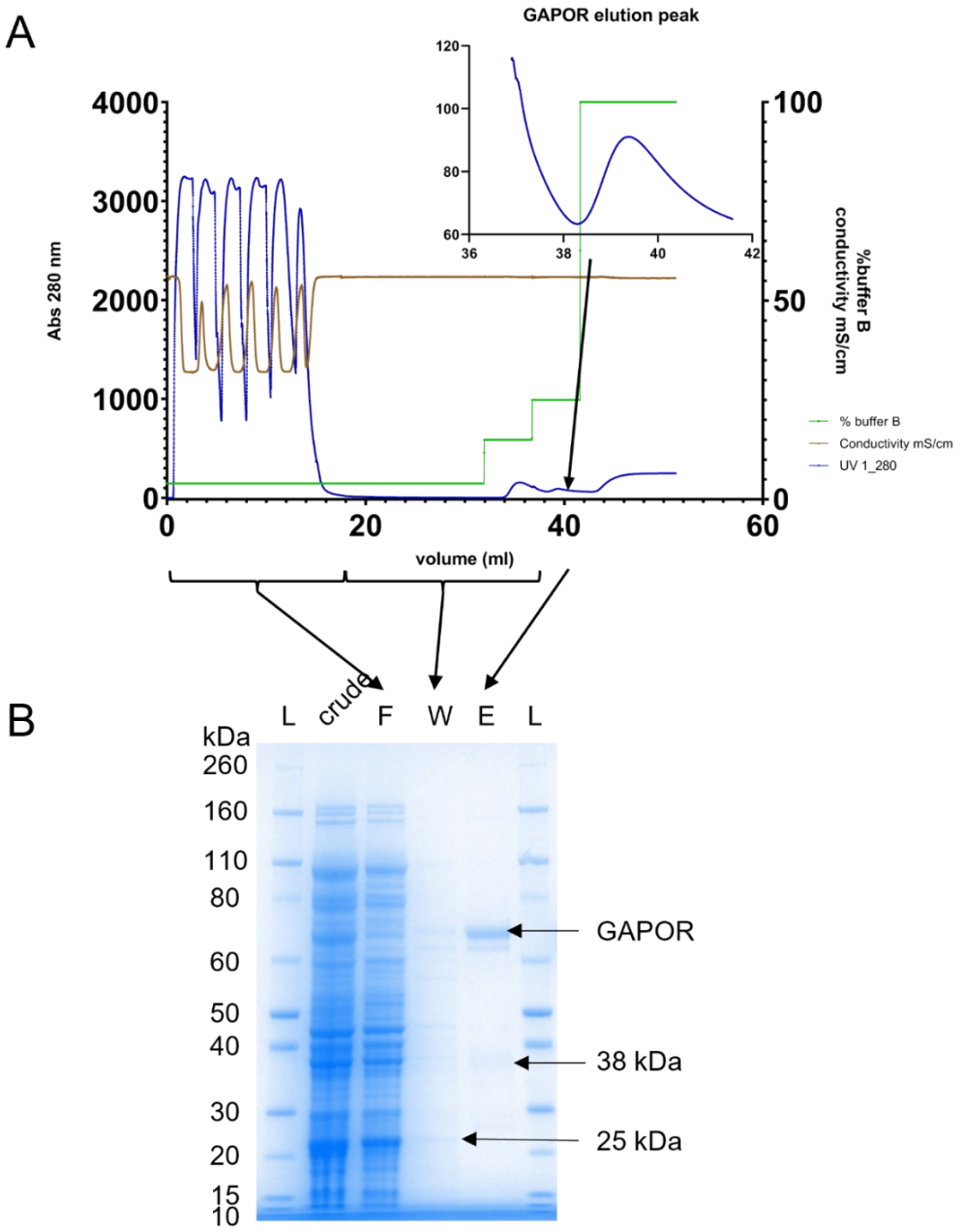


Figure III-23 GAPOR affinity purification from 600 ml of *E. coli* rosetta-gami2(DE3) Δ*iscRhypFselA* pET28a(+) *gorS2*-(*ampR*) (*E. coli* RG_AM023gapor) using an Äkta pure. A) After lysis, 600 ml of crude extract were recovered and mixed 1:1 with 20 mM imidazole binding buffer (20 mM sodium phosphate buffer, 500 mM NaCl, pH 8.4 mixed with 4% of elution buffer, 20 mM sodium phosphate buffer, 500 mM NaCl, 500 mM imidazole, pH 8.4) and loaded on the column using a 2 ml loading loop. The column was washed with 15 ml of 20 mM imidazole binding buffer and 5 ml of 75 mM imidazole buffer (15% elution buffer). The GAPOR was eluted with 5 ml of 125 mM imidazole buffer (25 % elution buffer) and collected. The elution peak is magnified in the top right corner (only the absorbance curve). The column was then washed with 500 mM imidazole (100% elution buffer). B) SDS-PAGE Blue-silver stained ran at 200 V for 35 min with MOPS running buffer (Invitrogen). 10 µl of the sample were loaded on the gel. L: Novex-sharp protein pre-strained protein ladder (Invitrogen); Crude: crude extract loaded on the column after 1:1 dilution; F: flowthrough recover during the loading of the column; W: column washes with 20 mM and 75 mM imidazole pooled together; E: GAPOR elution. GAPOR is indicated with an arrow as well as degradation at 38 kDa in the elution and 25 kDa in the W lane. To note that the lysis was performed in Nottingham by sonication under aerobic conditions following an anaerobic growth. Densitometry analysis was performed using ImageLab 6.1 (BioRad). The standard used was the concentrated eluate (not shown in the picture)

which was titrated by BCA titration at 0.292 mg/ml. About 1.9 µg of protein was loaded on the gel. GAPOR in the tested lane represented 53% of the total detected protein, and the band used represented 14.9% (the same band as the one below GAPOR). After recalculating the protein concentration by densitometry of this standard lane as control gave a concentration of 0.298 mg/ml. Thus, the protein concentration of each sample was calculated as follows W: 0.034 mg/ml, E: 0.067 mg/ml; giving a quantity of protein in each solution of W: 676 µg of protein and E: 335 µg. With the same software, the quantity of GAPOR in each sample was calculated, giving: crude: 517 µg of GAPOR, Flowthrough: 55 µg, Washes: 120 µg and elution 269 µg

5. Enzymatic activity on purified GAPOR

5.1. GAPOR activity assay without addition of L-cysteine

As formerly presented (Figure III-15), no GAPOR G3P-dependent activity was observed during the crude extract enzyme assay due to strong background activity. We then concluded that the activity assay had to be performed on purified enzyme, which would allow for measuring the specific activity at the same time. Hence, once a correct protocol was set up to purify the protein by affinity purification using an Äkta protein chromatography apparatus, it was possible to perform the GAPOR kinetic assay on the purified protein under anaerobic conditions in Toulouse. The Äkta apparatus present in the anaerobic chamber was equipped with a sample pump that allowed purifying a larger volume of crude extract (30 mL). The experiment was carried out in Toulouse with 3 L of *E. coli* RG_AM023gapor anaerobic culture. The purification was performed under anaerobic conditions on the Äkta pure present in their anaerobic cabinet. The 15 ml sample recovered from this culture was mixed with 1:1 with the binding buffer. The 30 ml were then loaded on the column with a sample pump. The purification was conducted as in Figure III-23. The GAPOR eluted was desalted with a PD-10 to remove the 500 mM NaCl and 125 mM imidazole. After desalting, about 427 µg of GAPOR was recovered (122 ± 2 µg/ml).

The purified GAPOR activity was then assayed with a protocol like the one used for the crude extract assays. The kinetic curves from this assay are presented below (Figure III-24a). No BV reduction was observed after 15 min of the assay, which was confirmed after calculating the specific activity from the kinetic curve slope. As shown in Figure III-24 b, the order of magnitude of specific activity is 10^{-5}, with the highest specific activity for the 75 µl negative control

at $4.72x10^{-5} \pm 7.23x10^{-7}$ U/mg of GAPOR. In comparison, the activity measured by Park et al. for the same enzyme was 120 U/mg of GAPOR (Park et al., 2007).

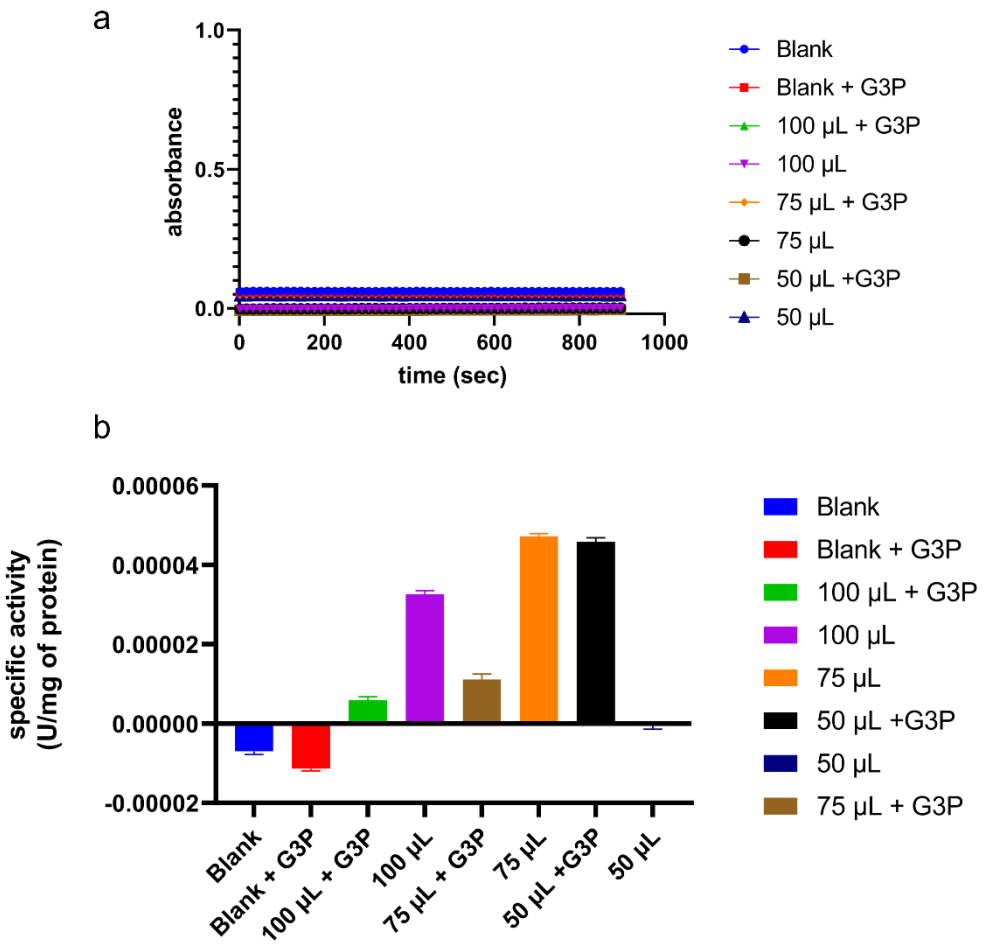


Figure III-24 Kinetic curves of purified GAPOR BV reduction in the presence of G3P. (a) 3 L of *E. coli* rosetta-gami2(DE3) Δ*iscRhypFselA* pET28a(+) *gorS2*-(*ampR*) (*E. coli* RG_AM023gapor) anaerobic culture were purified using a 1 ml Histrap nickel column and an Äkta apparatus in the anaerobic cabinet in Toulouse. The eluate was desalted using PD-10 following the gravitational protocol. Three different volumes of desalted eluate were tested: 100 µl, 75 µl, and 50 µl. The protein solution was added to 500 µl reaction mix (3 mM Benzyl viologen, 56 µM Na_2MoO_4, 50 mM EPPS), 300 µl, 325 µl, and 350 µl of water in depending on the volume of the protein solution tested. The monitoring was started after the addition of the GAPOR solution. The reactions were started by adding 100 µl G3P from a 1 mM stock solution. Water was used as a control. The GAPOR solution was replaced in the blank with 10 mM phosphate buffer. The reactions were monitored for 15 min. All the solutions were sparged with N_2 for at least 30 min before being brought into the anaerobic cabinet. (b) specific activity calculated from the data presented on (a). The enzymatic specific activity unite is defined as µmol of G3P consumed/mg of GAPOR/min. To compare all the different conditions, a coefficient was applied to compensate for the dilution factor: 1 for blank and 100 µl, 1.33 for 75 µl and 2 for 50 µl.

It can be concluded that in the tested condition presented above, the GAPOR is not active. In their paper Park et al. stated that purified GAPOR exposed to air are reversibly inactivated but was reactivated by the addition of L-cysteine. They also said that aerobically produced and purified GAPOR has the same yield

of GAPOR activity compared to the anaerobically prepared enzyme as long as L-cysteine is added for reactivation (Park et al., 2007). It was then supposed that then the addition of L-cysteine according to their protocol and results would help to observe GAPOR-specific activity.

5.2. GAPOR "reactivation"

In accordance with our hypothesis and the results of Park et al. presented above, the addition of L-cysteine during the enzyme assay needed to be tested to potentially observe the GAPOR activity. The addition of L-cysteine is supposed to ensure the reduction of the molybdenum and the Glu353 in the active site as well as the Fe-S cluster for the oxidation of G3P. In theory, the Mo/W was supposed to be in a Mo^{VI} oxidation state with the OH group of the Glu 353 near the Mo centre supposed to be reduced at the beginning of the reaction (*cf* Chapter I:5.1.1). In addition, it should ensure that no trace of oxygen could mitigate the reduction of BV, as observed in the previous experiments. Hence, it was decided to test the reactivation of GAPOR with L-cysteine with anaerobically produced and aerobically produced GAPOR. Nevertheless, the GAPOR purification was performed under anoxic conditions. A new purification of 3 L cultures was done, and 3.5 ml of protein solution was recovered after desalting at 774 ± 36 µg/ml. During this assay, the GAPOR solution was put in contact with 63 mM L-cysteine-HCl for 10 min in a tentative to reactivate the GAPOR. As can be visible in Figure III-25, BV reduction only happened when L-cysteine was present in the reaction mix. Moreover, the BV reduction was not G3P-dependant. The GAPOR-specific activity reported by Park et al. was 120 U/mg of GAPOR (Park et al., 2007). In Figure III-25, it is clear that BV is only chemically reduced by L-cysteine Hence, no GAPOR activity after reactivation were detected.

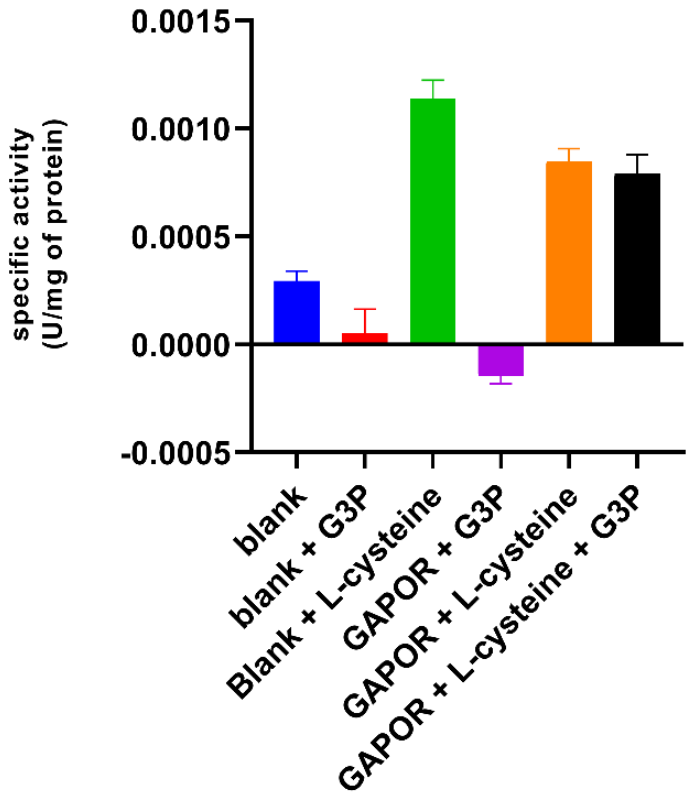


Figure III-25 Specific activities calculated after anaerobic GAPOR enzyme assay reactivated with L-cysteine. All the activity were calculated using the concentration of GAPOR measured by BCA titration (774 ± 36 µg/ml). After purification, the 480 µl of GAPOR was incubated 10 min with 20 µl of 1.43 M L-cysteine (final concentration 63 mM), 100 µl of the reduced GAPOR solution was used for the assay, the same volume of 10 mM phosphate buffer was used for the control. The reaction was started using 100 µl of G3P stock solution for a final concentration of 100 µM. The reaction was monitored for 15 min. The protein were produced with *E. coli* rosetta-gami2(DE3) Δ*iscRhypFselA* pET28a(+) *gorS2*-(*ampR*) (*E. coli* RG_AM023gapor) in 3L MAC medium culture under anaerobic conditions.

As the reactivation of an anaerobically produced GAPOR did not work, it was decided to test the reactivation of an aerobically produced GAPOR. The precultures were carried out the same way as the anaerobic GAPOR production. The culture conditions were modified as the expression under aerobic conditions has the advantage of producing more GAPOR in a smaller culture volume. The aerobic cultures were performed in 300 ml of MAC media. The 300 ml cultures were inoculated at 0.1 OD and incubated at 37°C for 7 to 8 hours (about 4.5 to 5). However, the GAPOR, for the same initial reasons (protection of the protein from aggregation and denaturation), were performed at room temperature. In Toulouse, a cooled incubator was

available, and 22°C was chosen as room temperature. The cultures were chilled to 22°C, induced with 100 µM of IPTG and incubated overnight at the same temperature. The cells were harvested at an OD_{600nm} around 6.3 and washed aerobically with the same conditions. The lysis was also performed under aerobic conditions. The purification with the standard protocol was performed under an anaerobic atmosphere. For the experiment presented in Figure III-27, 3.5 ml of desalted GAPOR solution at 829 ± 22 µg/mL was recovered which correspond to about 2.9 mg of protein (Figure III-26). The SDS-PAGE protein separation profile of aerobic-produced affinity-purified GAPOR (Figure III-26) was similar to the anaerobically produced GAPOR. The same degradation was indeed observed.

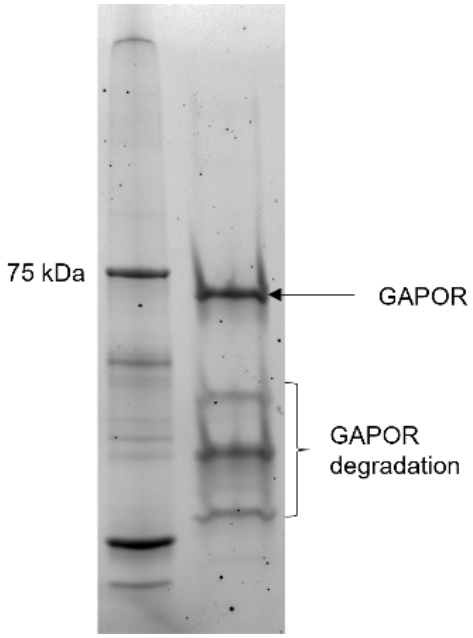


Figure III-26 SDS-PAGE of aerobically produced GAPOR after purification desalting. 12.5 µl of the sample (5 µl of GAPOR solution in 25 µl loading buffer, 5min at 95°C) was loaded on 12% mini-protean TGX stain-free protein gel (no Coomassie staining). The gel was run for 35 min at 200 V. The ladder used was Precision Plus Protein standard unstained. The assumed GAPOR degradations bands were, unfortunately, not extracted from the gel for sequencing, which would have permitted the validation of protein origins

The reactivation protocol was modified for optimisation, leading to a modification of the enzyme assay protocol. For this experiment, the cysteine was dissolved in the 2X EPPS buffer and to compensate for the pH drop, the pH was roughly adjusted with NaOH (sparged with N_2) at pH 8 with pH paper. All the solutions were sparged prior to entering the anaerobic chamber.

To test the modification brought to the reactivation step, an enzyme assay was performed with L-cysteine solution only. This experiment was done to estimate the BV reduction background activity and test the state of reduction of the reaction solution. Hence, the same experiment was run twice: just after introducing the pre-reduced solution and after overnight equilibration in the anaerobic chamber. This test of BV reduction by cysteine was performed as follows: 500 µL of buffered cysteine solution was mixed with 400 µL of water, and the reactions were started with 100 µL of 10x BV solution (3 mM). It gave an equivalent activity of 0.03 (for 800 µg/mL). This equivalent specific activity is higher by 10 to 100-fold than the specific activity calculated in the previous experiment (Figure III-25). After this overnight equilibration in the anaerobic chamber, when the reaction was performed under the same conditions, the activity without enzymes and G3P was 0.168 ± 0.03 U/mg. This 10-fold increase in BV reduction activity seems to indicate that sparging with N_2 was not sufficient to remove all the oxygen present in the solutions.

As observed before, L-cysteine is reducing BV, and thus the reaction had to be started not by the substrate but by the last electron acceptor. Performing the experiment this way allowed the enzyme to be incubated for 10 min in the L-cysteine EPPS buffer by reducing the [4-fe-4S] cluster and the molybdenum cofactor. The reduction of BV in the presence of GAPOR and G3P gave an activity of 0.190 ± 0.045 U/mg, and without G3P, 0.166 ± 0.047 U/mg (Figure III-27). Even though the GAPOR-G3P experiment has a higher activity of 0.024 U/mg, this difference is not significant compared to the standard deviation (*P* value = 0.7149). Hence, no GAPOR G3P-dependent activity was observed after the “reactivation” of the enzyme by the addition and incubation with L-cysteine.

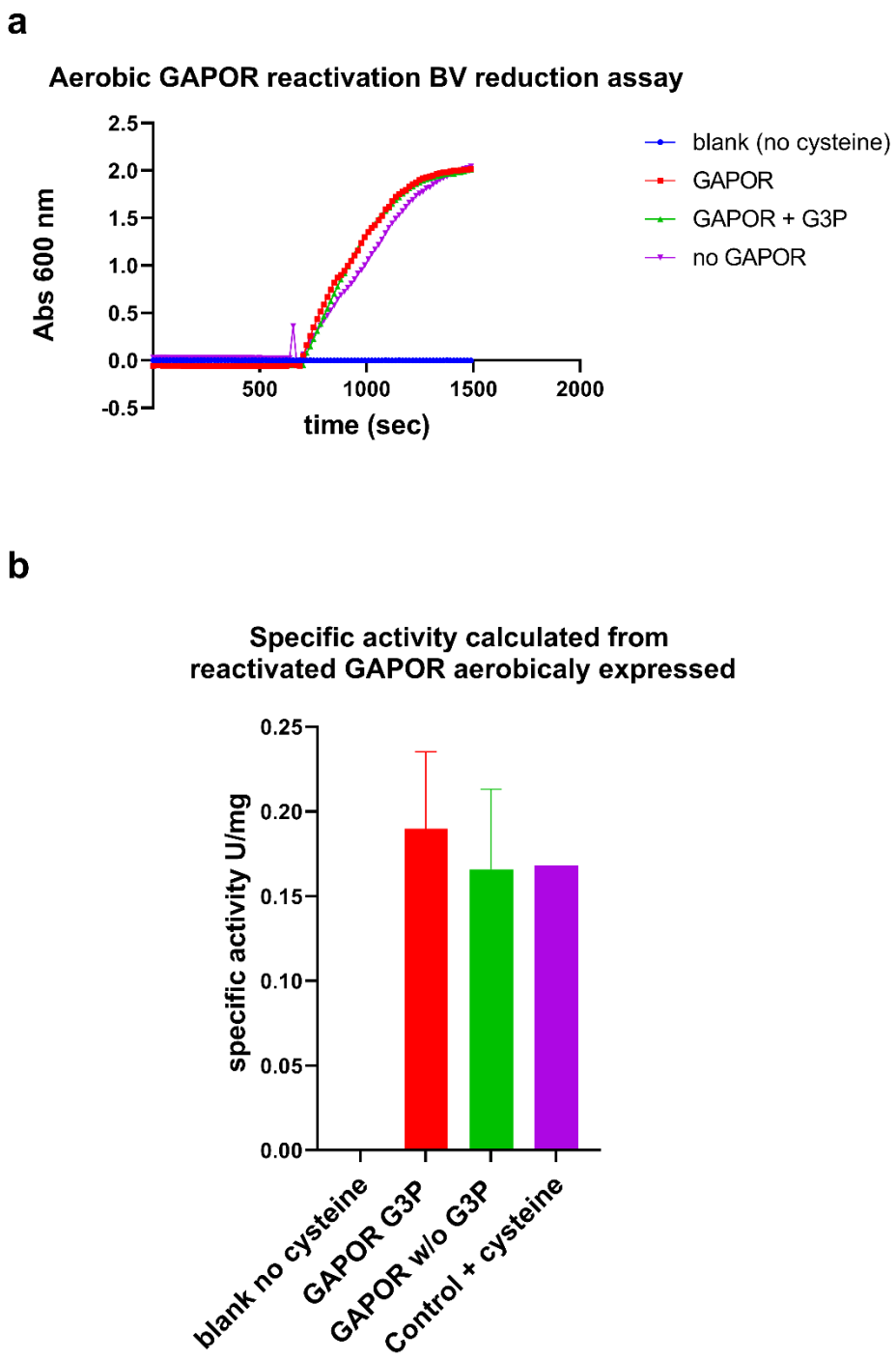


Figure III-27 Specific activity calculated from reactivated GAPOR aerobically produced. For this experiment, 300 ml of aerobic culture of *E. coli* rosetta-gami2(DE3) Δ*iscRhypFselA* pET28a(+) *gorS2*-(*ampR*) (*E. coli* RG_AM023gapor) was used. The culture was inoculated at 0.1 OD600 nm, incubated for 7 to 8 hours at 37°C (OD600nm 4.51), cooled down to 22°C and induced with 100 µM IPTG. The culture was then incubated at 22°C overnight. The cells were harvested and lysed aerobically. The lysate supernatant was then purified following the standard protocol under an anoxic atmosphere. After desalting, 2.9 mg of GAPOR was recovered in 3.5 ml of 10 mM sodium phosphate buffer (0.829 mg/ml). (a) Example of kinetic curve collected from the GAPOR reactivation with cysteine. In this experiment, the cysteine was added to 500 µl of EPPS buffer (50 mM EPPS, 56 µM Na2MoO4, pH 8.4), and the pH was roughly adjusted back to pH 8.4 with pre-reduced NaOH. The final L-cysteine concentration was 63 mM. 100 µl of purified and desalted GAPOR solution was added to the cuvette and incubated for 10 min in the cysteine EPPS buffer. 100 µl of G3P was then added to the solution.

In this case, the reaction was started by the addition of BV solution (final concentration 3 mM). Water was used to replace the tested solution in the control assay. The assay was monitored from the addition of the protein solution for 25 min. (b) The specific activity was calculated from the ΔOD/min from GAPOR reactivation with L-cysteine during the first 200 sec after the addition of BV. After these two tentatives of reactivation, no GAPOR G3P-dependent activity was observed and even less enzyme-dependent activity. All the BV-reduction activity was dependent on L-cysteine addition. The reaction mixture was supplemented with molybdate to provide an excess of molybdenum; as done by Park et al., iron sulphate may have been added to the mixture for the same purpose.

In conclusion, the reduction of the reaction mix solution, the enzymes' active site and metallic centre by L-cysteine did not successfully restored a G3P-dependent BV reduction by the GAPOR. It was supposed that all the GAPOR after purification was correctly folded. However, GAPOR degradations were visible in the crude extract (Figure III-6 western blot) but also after purification (Figure III-26). The pattern of degradation was vastly conserved, but it was not determined when the proteolysis happened and what was causing the proteolysis.

6. Discussion and Conclusion

The experiments presented in this section aimed to produce and characterise the GAPOR activity with the ultimate objective of using it in the 3-PG to G3P direction to conserve the electrons given by the CODH. The GAPOR from *M. maripaludis* was chosen because of its characterisation by Park and co-workers (Park et al., 2007). They have shown that GAPOR from this mesophilic archaeon was active after a heterologous *gorS2* expression in *E. coli*. This activity is Mo dependent contrary to all the other studied GAPOR and AOR from archaea, which are W-dependent.

To improve the expression and maximise the probability of observing GAPOR activity, *iscR* was deleted. This aim was to improve the [4Fe-4S] cluster production in order to increase the incorporation of this cluster into the GAPOR and potentially reduce the amount of inclusion bodies. A minor improvement was observed, and it was then supposed that increasing the cytosolic availability of the [4Fe-4S] cluster helped increase the cofactor incorporation (Chapter III:2.3.1) However, it would have been interesting to be able to quantify the amount of cofactor incorporated by subunit by Ion coupled

plasma-mass spectrometry (ICP-MS) or electronic paramagnetic resonance (EPR).

To evaluate the functionality of the GAPOR, it was decided to measure it in the crude extract (Chapter III:3). After the first essay to measure GAPOR activity in Nottingham, strong non-G3P-dependent background activities were observed. To reduce these activities, *hyp*F and *sel*A genes were deleted. These were very effective when it was tested in Nottingham. However, when the same experiment was reproduced in Toulouse, a G3P-dependent activity was observed. This activity seemed to lead to both the reduction of $NAD(P)^+$ and BV. When the activity was calculated following the experiment with and without GAPOR, no significant differences were observed in the BV-reduction activity that could indicate a GAPOR, G3P-dependent, BV reduction. Hence on crude extract, no GAPOR and G3P-dependent activity was observed. The experiment in Toulouse emphasised the oxygen level in the Nottingham enzymology anaerobic cabinet. Indeed, the G3P-dependent activity of *E. coli* RG_AM023gapor was tested in Nottingham, and no BV-reduction activity was observed. Contrary to the anaerobic cabinet used in Toulouse or for microbiology in Nottingham, where the anaerobic atmosphere is automatically monitored and regulated, the one in the enzymology anaerobic cabinet is not. After analysis of the gathered data, it seems that I was not able to bring, in Nottingham, the anoxic level required for the enzymology experiments. The oxygen perturbation could have come from the permeation through the vinyl wall of the cabinet, being higher than the reduction rate by the catalyst, an insufficient degassing of the solution used for the assay or, more basically, a problem with the utilisation of the cabinet. As the work presented in the next chapter was conducted in parallel, it was impossible to detect the background activity observed in Toulouse on time. Moreover, the mistake of not performing cleaner desalting prevented us from having a background-free kinetic assay. Using the standard gravity protocol could have helped us remove more metabolites than the centrifugal methods. Thus, if active GAPOR was

present in the cuvette, weaker GAPOR-dependent activity might have been observed.

To observe any GAPOR activity, the assays were conducted on affinity-purified GAPOR (Chapter III:4). The enzyme was successfully purified thanks to its N-terminal 6-histidine tag. The GAPOR band observed on the different band was once sequenced by MS with a good coverage (data not presented in this report) However, degraded proteins were eluted along with the GAPOR. Protein degradation was also detected in the crude extract by western blot against a his-tag in N-terminal. More degradation probably occurred from N-terminal, but it was not possible to observe them due to the absence of a tag in the C-terminal. In addition, to the potential degradation during the lysis and the purification, some of these degradations were also observed in protein crude extract. They could be due to non-complete translation of the enzymes, or transcription of non-complete mRNA. To verify this hypothesis, RT-qPCR could have been performed on samples. With correctly chosen primers depending on the degradation observed, it could have been possible to estimate the proportion of incomplete mRNA fragments. However, when and what caused the protein degradation was not studied. The first trial to characterise the purified enzyme did not permit the observation of any kind of BV reduction (Chapter III:5.1). To palliate potential oxidation of the two GAPOR cofactors prohibiting any enzymatic activity, the anaerobically produced GAPOR was exposed to L-cysteine (Chapter III:5.2). This exposure to L-cysteine was supposed to help get all the cofactors in their correct oxidation state, including the Glu residue near the Mo centre. In this first experiment, all the BV reduction measured was cysteine dependent, no difference of activity was observed between cysteine alone, GAPOR alone and GAPOR with G3P. A similar experiment was conducted on aerobically produced enzymes. In this occurrence, purified GAPOR was exposed to L-cysteine for a longer time, and the addition of the last electron acceptor, BV, started the reaction. No GAPOR-dependent and G3P-dependent BV reduction was noted either. Hence, overall, no oxidation of G3P to 3-PG was detected. It was not possible to reproduce the

experiments presented by Park et al. (Park et al., 2007), even with the addition of L-cysteine. In addition, it would have been interesting to prepare the assay with natively produced GAPOR from either *M. maripaludis* or *P. furiosus.* This could have been used as a positive control to validate the assay beforehand. Moreover, the possibility of using the exact *E. coli* strain presented by Park in his article might have been a good addition to the control to determine if the modification brought to the strain genome successfully expressed an active GAPOR.

The first hypothesis was about the correct folding of the enzyme. It was not possible to observe the GAPOR activity in crude extract due to background activities. It is then difficult to conclude regarding the proper folding of the enzyme during these experiments. However, inclusion bodies were still present in the pellet after lysis despite the improvement with the deletion of *iscR*. Indeed, by releasing the repression of the *isc* operon, the [4Fe-4S] availability was improved (Akhtar & Jones, 2008), and as a consequence, it facilitated the molybdenum cofactor biosynthesis. In fact, it was recently demonstrated that iron-sulfur cluster availability was impacting the molybdenum cofactor biosynthesis pathway at different stages, as presented in Chapter I.6.4 (Zupok, Gorka, et al., 2019). Hence, if denaturation occurred, it is most likely to have happened during the purification and/or during the desalting. Furthermore, it was shown that traces of oxygen in the solution could decrease the BV reduction rate. However, when all the solution and columns were introduced in the anaerobic chamber overnight, no activity was observed despite the "reactivation" by cysteine. Reschcke et al. have shown during their study of YdhV in *E. coli* that under aerobic conditions, the cofactor was inserted into the protein and kept under an oxidised form within the enzyme (Reschke et al., 2019). Hence, it seems that even in the case of contact with oxygen, the molybdenum cofactor was not released from the protein. The addition of iron and sulfur in excess in either a "refolding" step prior to L-cysteine treatment, as previously shown to work for ferredoxin (Guerrini et al., 2008), would potentially have helped to ensure the proper cofactor composition of the

enzyme. Again, as explained above, titration of the iron content might have helped to understand if the limitation could have originated from the denaturation of the protein. The ratio then would have been lower than 4:1 as observed in the literature (Park et al., 2007) for a correctly folded protein.

In molybdenum-sufficient background such as here, the cofactor production is low (Anderson Lisa, McNairn, Leubke, Pau Richard, & Boxer David, 2000). In normal conditions, the molybdate availability was driving the upregulation of *modE* four- to five-fold under aerobic conditions (Anderson Lisa et al., 2000). However, this was tested with 10 mM molybdate as the final concentration (Anderson Lisa et al., 2000) when100-fold less was used in our media (Park et al., 2007). Hence, it is possible that despite the supplementation, no significant *modE* strong upregulation occurred. Moreover, as mentioned above (Chapter I:6.4) it can be supposed that the *moaA, moaE* and *moaD* genes should be upregulated by the deletion of *iscR* and the supplementation of media with 330 mM $FeSO_4$. Despite this upregulation, *moaA* translation was shown to be downregulated by unbounded molybdenum cofactor due to a riboswitch upstream of *moaA*. It can then be supposed that no upregulation of the cofactor happened. When tested by Park and co-workers only a fifth of the enzyme had both cofactor incorporated and iron and in Rosetta-gami 2(DE3) strain 0.74 mol of Mo/mol of protein. The cofactor of GAPOR and other AORs is known to differ from the main molybdenum cofactor produced in *E. coli* (*cf* Chapter I:5). However, Reschke et al. recently found that YdhV, an *E. coli* periplasmic enzyme with an unknown activity, was containing an identical cofactor as GAPOR (Reschke et al., 2019). However, neither MobA, the enzyme responsible for the formation of the bis-MGD using bis-Mo-MPT as an intermediary, nor MocA, the equivalent for the formation of MCD, are involved in the formation of YdhV cofactor (Reschke et al., 2019). They also suggest that two other proteins encoded by the *ydh* operon, YdhW and YdhT, might be involved in the cofactor formation. At the beginning of this study, in regards to the cofactor analysis of Park and co-workers, it was decided to design experiments to improve the production and integration of the proper cofactor.

After the results presented in this chapter, the goal of the following experiment was modified to try to incorporate the correct cofactor and potentially restore GAPOR activity.

Chapter IV:

Methanococcus maripaludis Molybdenum Cofactor Biosynthesis Pathway

1. Introduction

In the previous chapter, the heterologous production of an active GAPOR from *M. maripaludis* was not observed. One of the hypotheses was a problem with the insertion of the right molybdenum cofactors in the enzyme. This hypothesis was reinforced by the fact that when the GAPOR form *Methanococcus maripaludis* was produced in *Escherichia coli*, only a fifth of the enzyme had both cofactor incorporated (Park et al., 2007). As presented in the literature review (Chapter I:6), the molybdenum cofactor biosynthesis pathway is ubiquitous. However, it presents some differences depending on the strain. The AORs of *archaea* are known to harbour a W coordinated by dithiolene groups of two molybdopterins (Mo/W-bisMPT) (*cf* Chapter I:5) while the GAPOR from *M. maripaludis* was found to be Mo dependent (Park et al., 2007). In *E. coli,* the final molybdenum cofactors are Mo-MPT, MCD or bis-MGD with different maturation steps. However, The *E. coli* enzymeYdhV was recently found to contain the same cofactor as *M. maripaludis* GAPOR (Reschke et al., 2019).

Thus, in parallel with the works presented in the previous chapter, a mini plasmids library and strains for the expression *gorS2,* along with 6 genes from the *M. maripaludis* molybdenum cofactor biosynthesis pathway, was constructed. This library aimed to improve the GAPOR activity after reproducing the Park et al. experiment by providing the native molybdenum cofactor to the GAPOR. However, after the first preliminary results obtained in Nottingham (Figure III-14) and their confirmation in Toulouse (Figure III-15, Figure III-24, Figure III-25, Figure III-27) the cofactor insertion became a way to express an active GAPOR.

Herein are presented the results of the study of the GAPOR complementation with molybdenum cofactor from *M. maripaludis*. First is presented the choice of the molybdenum cofactor biosynthesis pathway encoding genes to be expressed in *E. coli* along with *gorS2* and the construction of the plasmids and strain library. Their expressions were tested and tentatively analysed by SDS-

PAGE. GAPOR kinetic assays and test to identify the cofactor incorporated in the target enzyme were done from the mini-library of strains described below.

2. The *M. maripaludis* molybdenum cofactor biosynthesis pathway compares to *E. coli* pathway

First, the *M. maripaludis* molybdenum cofactor biosynthesis pathway was compared to the *E. coli* pathway to select the genes that will be co-expressed with the GAPOR gene (*gorS2*) to produce the native cofactor and supposedly a functional GAPOR. In *M. maripaludis,* 11 genes were annotated in the genome assembly (*Methanococcus maripaludis S2* taxid: 267 377, Figure IV-1) (Hendrickson, Kaul, Zhou, Bovee, Chapman, Chung, Conway de Macario, et al., 2004) as coding for an enzyme from the molybdenum cofactor biosynthesis pathway. In comparison, in *E. coli,* ten genes are responsible for the cofactor production, excluding the regulation.

The enzymes from the *E. coli* MoCo biosynthesis pathway have been aligned using Blast protein (https://blast.ncbi.nlm.nih.gov/Blast.cgi) (Altschul, Gish, Miller, Myers, & Lipman, 1990) with the *M. maripaludis* enzymes proteome (Table IV-1). All *E. coli* protein has a counterpart in *M. maripaludis* except MoaD and MogA.

In *M. maripaludis* MoaB and MogA have 32 % identity and 56% homology, and 29%, 53%, respectively, with the same 161 amino acid protein annotated as MoaB (Table IV-1). In fact, MoaB is a MogA-like protein that could catalyse the adenylyltransferase activity. However, in *E. coli* its activity is unknown (Leimkühler, 2020). In archaeons such as *Pyrococcus furiosus* and other archaea only MoaB, is produced (Loes E. Bevers et al., 2008b). Hence, it could be supposed that MoaB plays the role of MogA in *M. maripaludis*. It could also be supposed that MoaB might be an ancestral protein of MogA conserved in *E. coli*. It is to note that the molybdenum transfer in molybdate excess can be performed by MoeA alone (Hille et al., 2014). When the 81 amino-acid protein *E. coli,* MoaD, was blasted against the archaea strain proteome, none matched with a *M. maripaludis* protein. To keep the same comparison, MoaD

Table IV-1 Results of the Blastp between the protein of the *E. coli* molybdenum cofactor biosynthesis pathway against the *M. maripaludis* proteome. (https://blast.ncbi.nlm.nih.gov/Blast.cgi).

Protein in E. coli	***Length (number of aa)***	***Ref code NCBI of M. maripaludis S2 equivalent***	***% of Identity***	***% of homology***	***Length (number of aa)***	***Name in M. maripaludis***
MoaA	329	WP_011170515.1	29%	49%	297	MoaA
MoaB	170	WP_011171429.1	32%	56%	161	MoaB
MoaC	161	WP_011171010.1	43%	62%	153	MoaC
MoaD	81					
MoaE	150	WP_011171179.1	36%	50%	141	MoaE
MogA	195	WP_011171429.1	29%	53%	161	
MoeA	411	WP_011171563.1	27%	47%	402	MoeA1
		WP_011170489.1	31%	48%	616	MoeA2
		WP_011170457.1	34%	52%	373	MoeA3
		WP_011170145.1	29%	49%	375	MoeA4
MoeB	249	WP_011171178.1	40%	55%	239	MoeB
MobA	194	WP_011170517.1	31%	44%	214	MobA
MobB	175	WP_011171385.1	28%	48%	229	MobB

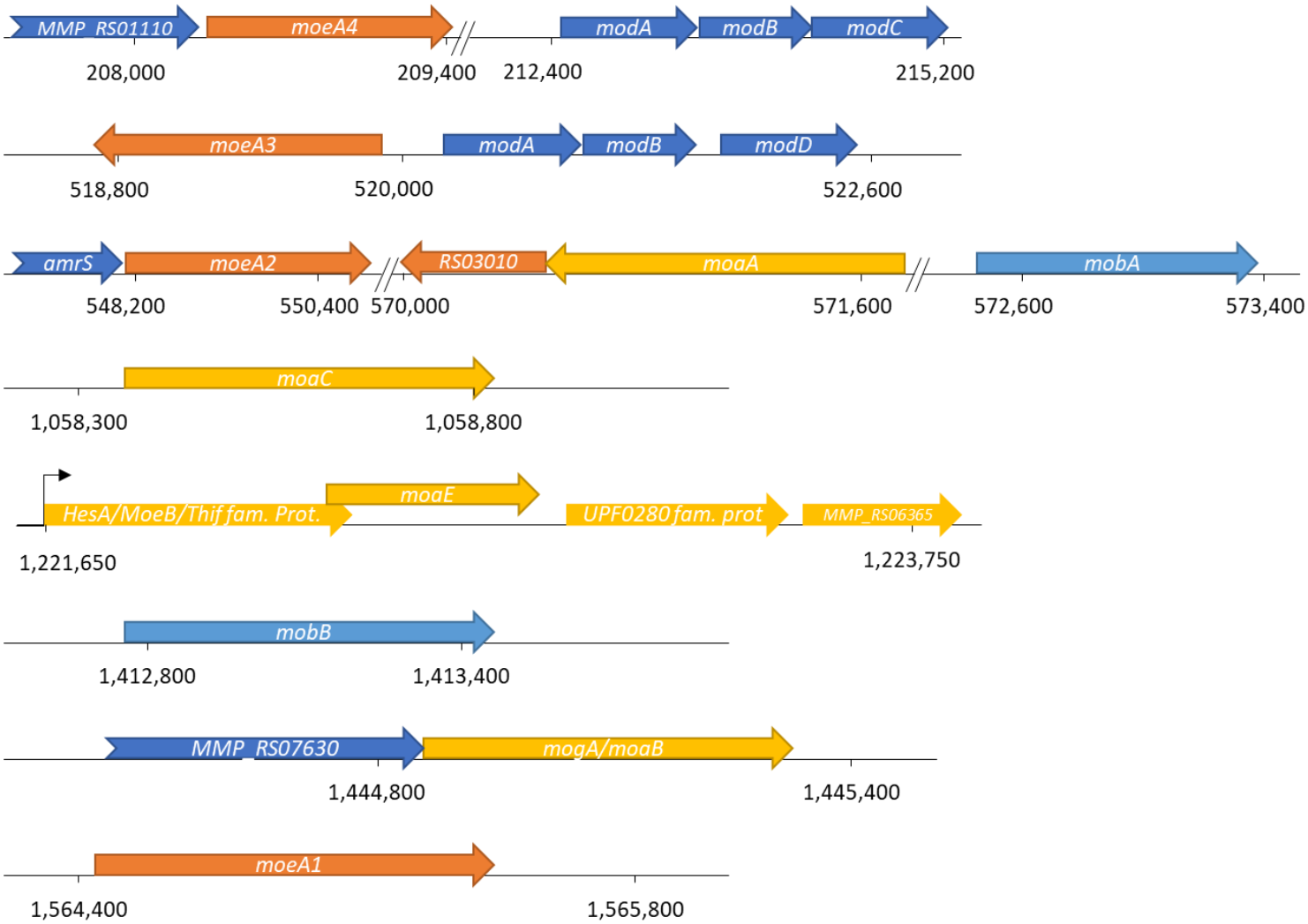

MMP_RS01110
moeA4
208,000
209,400
212,400
modA
modB
modC
215,200
moeA3
518,800
520,000
modA
modB
modD
522,600
amrS
moeA2
548,200
550,400
570,000
RS03010
moaA
571,600
mobA
572,600
573,400
moaC
1,058,300
1,058,800
HesA/MoeB/ThiF fam. Prot.
moaE
UPF0280 fam. prot
MMP_RS06365
1,221,650
1,223,750
mobB
1,412,800
1,413,400
MMP_RS07630
mogA/moaB
1,444,800
1,445,400
moeA1
1,564,400
1,565,800

Figure IV-1 Schemes of the molybdenum cofactor biosynthesis pathway genes in *Methanococcus maripaludis S2.*

was aligned against *P. furiosus* proteome, and it matched with an 82 amino-acid protein annotated as MoaD with 29% identities and 42% homology. In *E. coli,* this enzyme is involved in a complex with MoaE responsible for the double sulfuration of cPMP to form the PPT. MoaD plays the role of the sulfur carrier (Hille et al., 2014). Thus, another enzyme is probably responsible for the disulfuration of the cPMP to proceed with the synthesis of the cofactor. So, in *M. maripaludis,* MoaE might be supposed to also be responsible for carrying the sulfur group. Thus, *E. coli* MoaD was aligned against *M. maripaludis* MoaE from both strains using Clustal omega (https://www.ebi.ac.uk/Tools/msa/clustalo/) (Sievers et al., 2011). Both MoaE showed a similar identity percentage with *E. coli* MoaD (23% for *E. coli* and 25% for *M. maripaludis*). This lead to suppose that *M. maripaludis* MoaE has a similar activity. Another candidate is MoeB which is part of the complex sulfurating MoaD along with IscS, TusA and YnjE (*cf* Chapter I:6.2). However, when *E. coli* MoaD was aligned with *M. maripaludis* MoeB, it showed 22 % identity but more importantly mainly aligned itself in the middle of the archaeal MoeB. The sulfur is carried by MoaD on its C-terminal tail thanks to the Gly-Gly end, which insert deeply into MoaE (J.-U. Dahl et al., 2011). *M. maripaludis* MoeB does not seem to present that kind of C-terminal tail. Hence, it is unknown so far how the two sulfur atoms are brought to the cofactor. Furthermore, the homology and identity percentage presented in Table IV-1 between the protein from *E. coli* and *M. maripaludis* for this protein is one of the highest of this pathway and, without MoaD, the role of MoeB, which is to sulfurate MoaD, is unknown. It is noteworthy that all the genes of the MoCo biosynthesis pathway are in different parts of the *M. maripaludis* genome except *moeB* and *moaE,* which form an operon (Figure IV-1). The genes are positioned at 1 221 650… 1 222 790 in the *M. maripaludis* S2 genome (NC_005791.1 NCBI (Hendrickson, Kaul, Zhou, Bovee, Chapman, Chung, Conway de Macario, et al., 2004)) with *moeB* in the first position and *moaD* start codon embedded in *moeB.* This operon seems to be composed of 2 more

predicted genes MMP_RS06360, a UPF0280 family protein, and MMP_RS06365, a hypothetical protein. A terminator sequence seems to be present in the 128 nucleotides intergenic region with the biotin synthase BioB. In *E. coli,* MoaD is involved in the interaction between MoeB and MoaE (J.-U. Dahl et al., 2011). In the absence of MoaD, the fact that *moeB* and *moaE* are in the same operon tends to push toward the hypothesis of MoeB replacing or partially replacing MoaD as the sulfur carrier. However, no studies were found to look into this interaction.

Moreover, in *M. maripaludis,* four genes in different places inside its genome have been annotated as MoeA and align with their *E. coli* counterpart. As mentioned in the literature review (Chapter I:6.5), there are hypotheses that MoeAs could be involved in the discrimination between Mo and W in organism using both elements. All of them have a percentage of homology, *ca.* 50 % and a percentage of identity, *ca.* 30 % (Table IV-1). Three of them are shorter than the *E. coli* MoeA, but one of them (MoeA2) has a 162 amino acids tail at the C-terminal end. To continue investigating the absence of MoaD in the archaea, *E. coli* MoaD was aligned against MoeA2. Some homologies were observed in N-terminal, but the Gly-Gly C-terminal tail motif was not present in the MoeA2 sequence. Therefore, MoeA2 was supposed not to be involved in the sulfuration. Then, to understand if this tail was present in other organism the tail sequence was aligned using Blastp, excluding the *Methanococcus maripaludis.* The tail aligned with other MoeA from *Methanococci* including *Methanocaldococcus jannaschii* (Q58296 UniProt). When the two putative proteins were aligned (Clustal omega) they had 61% identities. This archaeon is also known to express a GAPOR (Guerrini, 2007).

Besides that, according to the proteomics report, only three of them, MoeA1, MoeA2 and MoeA4, are produced by *M. maripaludis* (Xia et al., 2006). Two MoeA have been found in *P. furiosus.* Bevers et al. proposed these two enzymes would be involved in incorporation of either Mo or W, as formerly mentioned (Loes E. Bevers et al., 2008b). Hence, following this hypothesis, one of those four MoeAs could be supposed to be used by the cell to incorporate

W along with Mo. The alignment between the MoeA from *E. coli* and the four *M. maripaludis* MoeA shows amino acids sequence homology (Figure IV-2). For instance, the highly conserved amino acids in domains III and IV in the cleft, where the active site has been supposed (Jason D. Nichols, Xiang, Schindelin, & Rajagopalan, 2007; Schrag et al., 2001).

Moreover, the results of a proteomics study conducted on *M. maripaludis* by Xia et al. indicated that *mobA* was transcribed into RNA, but no translation was observed (Xia et al., 2006). Thus, if MobA is not really produced, this tail could be used by this archaeon to produce the Mo-bisPPT required by GAPOR. So, MobA was aligned against MoeA2 but none of the identified active site amino acids (Lake, Temple, Rajagopalan, & Schindelin, 2000) were present in the MoeA2 tail.

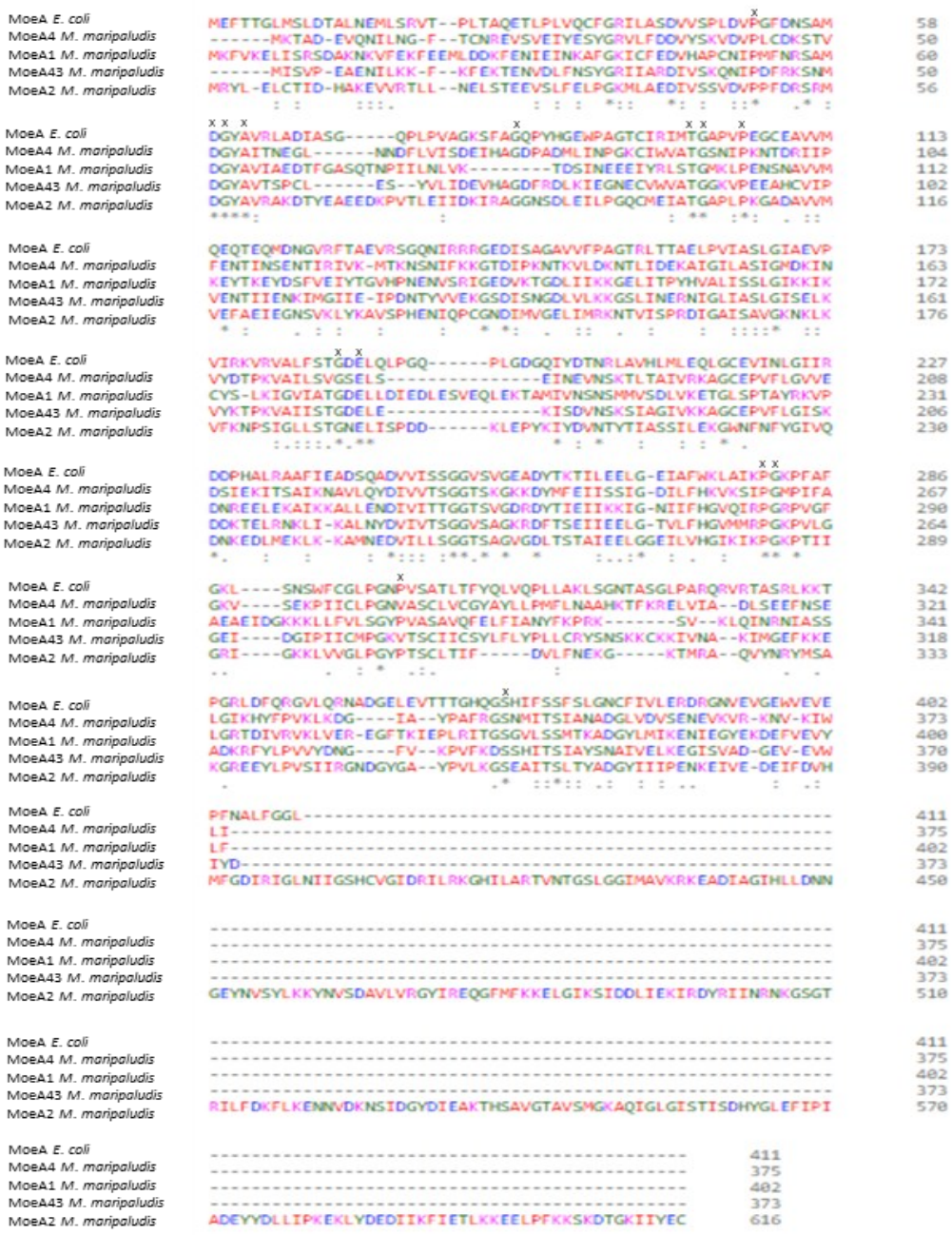


Figure IV-2 Protein alignment between MoeA from *E. coli* and MoeAs from *M. maripaludis*. The alignment has been done with Clustal omega version 1.2.4. The strictly conserved residues found by Xiang S., et al. 2001 are shown by an "x". the strictly conserved residues in the particular protein alignment are shown by the software with "*", highly similar and similar residues are respectively shown with ":" and ".".

In *E. coli*, some Molybdo-enzymes, such as DMSO reductase (Leimkühler, 2020; A. Magalon & Mendel, 2015), need the presence of a molecular chaperone to incorporate the MoCo. No molecular chaperone has been identified so far in *Methanococcus maripaludis* for the insertion of the Mo-bisMPT into the GAPOR protein.

Hence, after the comparison of the *E. coli* molybdenum cofactor biosynthesis pathway with *M. maripaludis*, two enzymes were not present in the archaea. It is still being determined how the cells are replacing them for the sulfuration the MPT. However, it is supposed that MobA might plays the MogA role of MPT adenylyltransferase in *M. maripaludis*. MoeA, MobA and MobB are responsible in *E. coli* for the incorporation of the molybdenum atom and the formation of the Mo-bisMPT/MGD. As presented above, when multiple MoeAs isozymes are coded in the genome, it was hypothesised to be for the discrimination between Mo and W. Therefore, it was supposed that one of the four MoeA might be involved in GAPOR's cofactor synthesis. It was then decided to study them in *E. coli*. In addition, MobA is known to be the enzyme catalysing the formation of the Mo-bisMGD with an intermediary of Mo-bisMPT and MobB is supposed to be involved in carrying the GTP to MobA. Thus, to investigate if these enzymes might be involved, it was decided to also study them. Hence, the genes encoding the four MoeAs, MobA and MobB from *M. maripaludis* were chosen to be cloned along with *gorS2* for a co-expression in *E. coli*.

3. Construction of the mini library

These constructions were planned to be a way to improve the GAPOR-specific activity and not to obtain some activity. Consequently, the plan was to express different combinations of the selected *M. maripaludis* molybdenum cofactor biosynthesis pathway along with the GAPOR in a same operon in *mobAB* and *moeA* mutants *E. coli.* To observe the potential impact of the archaeal genes expression on the GAPOR activity, *E. coli moeA* and *mobAB* genes had to be deleted.

3.1. Deletion of *E. coli mobAB* and *moeA* genes in Δ*isc*R*hyp*F*sel*A strain *E. coli* RG_AM023

In parallel to the plasmid construction, *moeA* and *mobAB* genes were deleted in the *E. coli* RG_AM023 strain. Three different mutant strains were constructed Δ*moeA,* Δ*mobAB,* Δ*moeAmobAB*, *E. coli* RG_AM024, *E. coli* RG_AM025 and *E.* RG_AM026, respectively. The deletions were performed by

P1 transduction (Figure IV-3). The deletion of *moeA* was performed first as the kanamycin marker was flanked by two Flippase (FLP) recognition target (FRT) sites as the phage lysate was prepared with *E. coli* BW25113 Δ*moeA*::*kan*. The phage lysate for the deletion of *mobAB* was prepared with *E. coli* Rosetta-gami2(DE3) Δ*mob*AB::kan received from Toulouse. In *E. coli,* as described in the literature review (Figure I-16), *mobAB* forms an operon. It would have been more complicated to knock-out the first gene separately while conserving the second functional. Hence, the entire operon was knocked out. However, the kanamycin marker of this strain was not flanked by FRT sites and it was more difficult to remove it. Thus, it was decided to keep it.

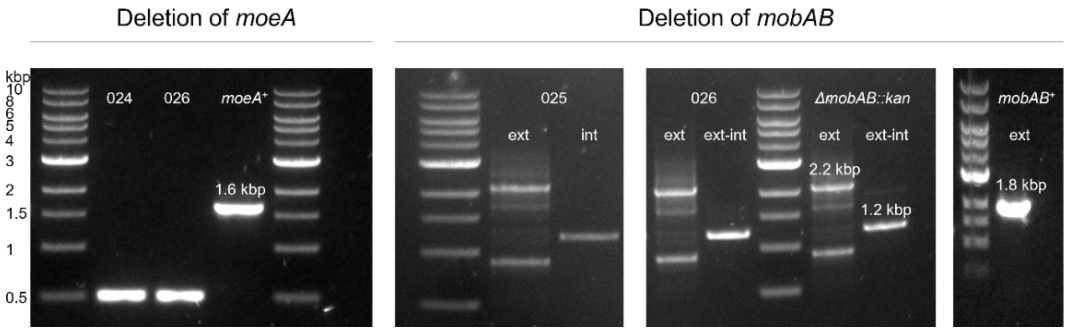


Figure IV-3 Control of *moeA* and *mobAB* deletion by P1 transduction.The deletion of both genes was performed by P1 transduction in *E. coli* rosetta-gami2(DE3) Δ*iscRhypFselA* (*E. coli* RG_AM023).The phage lysates were prepared with *E. coli* BW25113 Δ*moeA*::*kan* and *E. coli* Rosetta-gami2(DE3) Δ*mobAB*::*kan*. The kanamycin marker was removed after *moeA* deletion using a flip-recombinase. Here the final strain *E. coli* RG_AM024 (*E. coli* rosetta-gami2(DE3) Δ*iscRhypFselAmoeA)*, *E. coli* RG_AM025 (*E. coli* rosetta-gami2(DE3) Δ*iscRhypFselAmobAB::*Kan) , and *E. coli* RG_AM026 (*E. coli* rosetta-gami2(DE3) Δ*iscRhypFselAmoeAmobAB::*Kan) colony PCR are presented along with *moeA*$^{+}$ *E. coli* strain (*E. coli* Rosetta-gami2(DE3) *ΔmobAB::kan*). External primers were used for this control PCR. The expected fragment size for *moeA*$^{+}$ was 1.6 kbp. After the deletion of *mobAB,* the kanamycin marker was not removed. It was found that *E. coli* Rosetta-gami2(DE3) *ΔmobAB::kan* was constructed as Palmer et al. (Palmer et al., 1996). Two couples of primers were used: one external with an expected size for *mobAB*$^{+}$ (*E. coli* RG_AM023) was 1.8 kbp and 2.2 kbp for *ΔmobAB::kan* (*E. coli* Rosetta-gami2(DE3) *ΔmobAB::kan*) and an external, internal binding within the kanamycin gene with an expected size of 1.2 kbp. Several unspecific fragments were also amplified with the external primers, even in the positive control *ΔmobAB::kan*.

3.2. Plasmid library construction

To achieve the production of active GAPOR and determine the *M. maripaludis* enzymes responsible for the biosynthesis of the Mo-bispyranopterin, *mobA*, *mobB,* and *moeA*1-4 genes must be cloned in the plasmid used for the expression of GAPOR encoding gene (*gor*S2). These genes were cloned in the pET28a (+) *gor*S2 (Amp^{R}) (7,944 bp)

These genes were cloned to be expressed as an operon under the *lac* operator and a T7 promotor. Nineteen different operons of three types were designed (Figure IV-4). *M. maripaludis* genes were organised in those operons to be expressed individually to observe the impact of each gene on the GAPOR activity. Then the operons were designed to express Moco biosynthesis genes along with two and three genes to observe if there was any synergy between the different proteins. All 19 different combinations are summarised in Table IV-2. In those operons, the first gene was *gor*S2, and the *moeA* gene was always in the last position. Furthermore, *mobA* was always in the second position when cloned along with other *M. maripaludis* genes. (Figure IV-4).

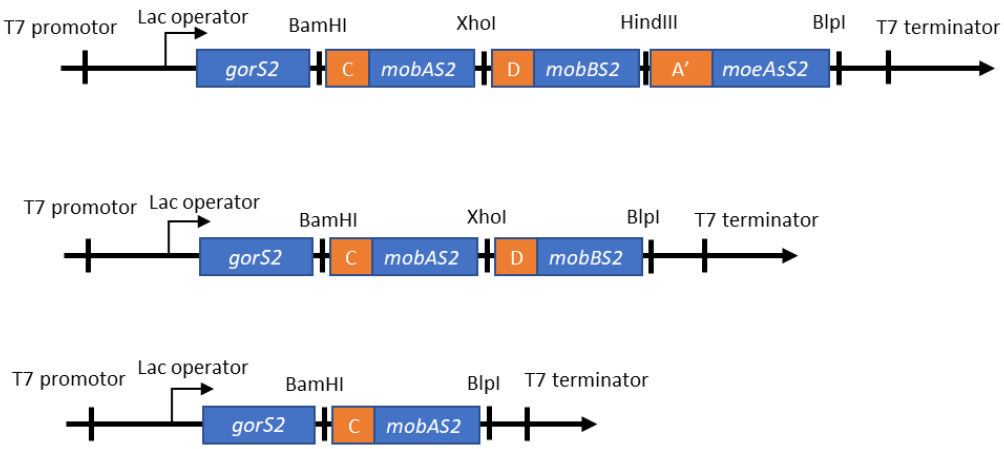


Figure IV-4 The Operon schemes representing the three types of assembly with their respective RBS and the introduced restriction sites. The RBSs were chosen to have a similar translation rate, and each gene was assigned a particular one C (AAGTTAAGAGGCAAGA) for *mobA*, D (TTCGCAGGGGGAAG) for *mobB* and A' (AAGAGGTTTGGA) for *moeA*s (Zelcbuch et al., 2013).

Moreover, a specific synthetic RBS was assigned to each type of gene (Figure IV-4). RBS C (AAGTTAAGAGGCAAGA) and D (TTCGCAGGGGGAAG) had medium efficiency in expressing the *mob* genes. RBS A' (AAGAGGTTTGGA) was a high-efficiency RBS which was modified (A2 instead of G2) to decrease its efficiency (Zelcbuch et al., 2013) as it is placed at the end of the mRNA sequence. They have been chosen to translate *M. maripaludis* genes with almost the same efficiency. This was checked by the prediction of translation initiation rate, as shown in Figure IV-5. These synthetics RBS cloned with the *M. maripaludis* genes are predicted to have a translation initiation rate 100-fold to 1800-fold the theoretical translation rate of *gorS2* (60.66). To have actual translation

higher than the *gorS2* gene would ensure that the cofactor biosynthesis would not limit the obtention of a functional GAPOR. It is important to note that no codons optimisation was performed as the *E. coli* rosetta-gami2(DE3) has a second plasmid pRARE2 which expresses *E. coli* rare tRNA.

The plasmids were constructed using Hifi assembly (NEB) following the combination presented in Table IV-2 and with the strategy presented in Figure IV-6. It was chosen to use Hifi Assembly methods as it allows cloning several genes at once. Moreover, it is a well-known technic in the group. The plasmids, once constructed, were then transformed into *E. coli* DH5α and validated by sequencing.

Once the library and the mutant strains were validated, the plasmids were transformed in the corresponding mutant Table IV-2. The transformed strains were named according to the plasmid it was hosting, *i.e.* pAM_001 transformed in *E. coli* RG_AM024 was named *E. coli* RG_AM001 *etc.* The mutant strains were also transformed with the original plasmid pET28a(+) *gorS2* (Amp^R) to be used as controls. All the strains were validated a second time by sequencing and restriction map.

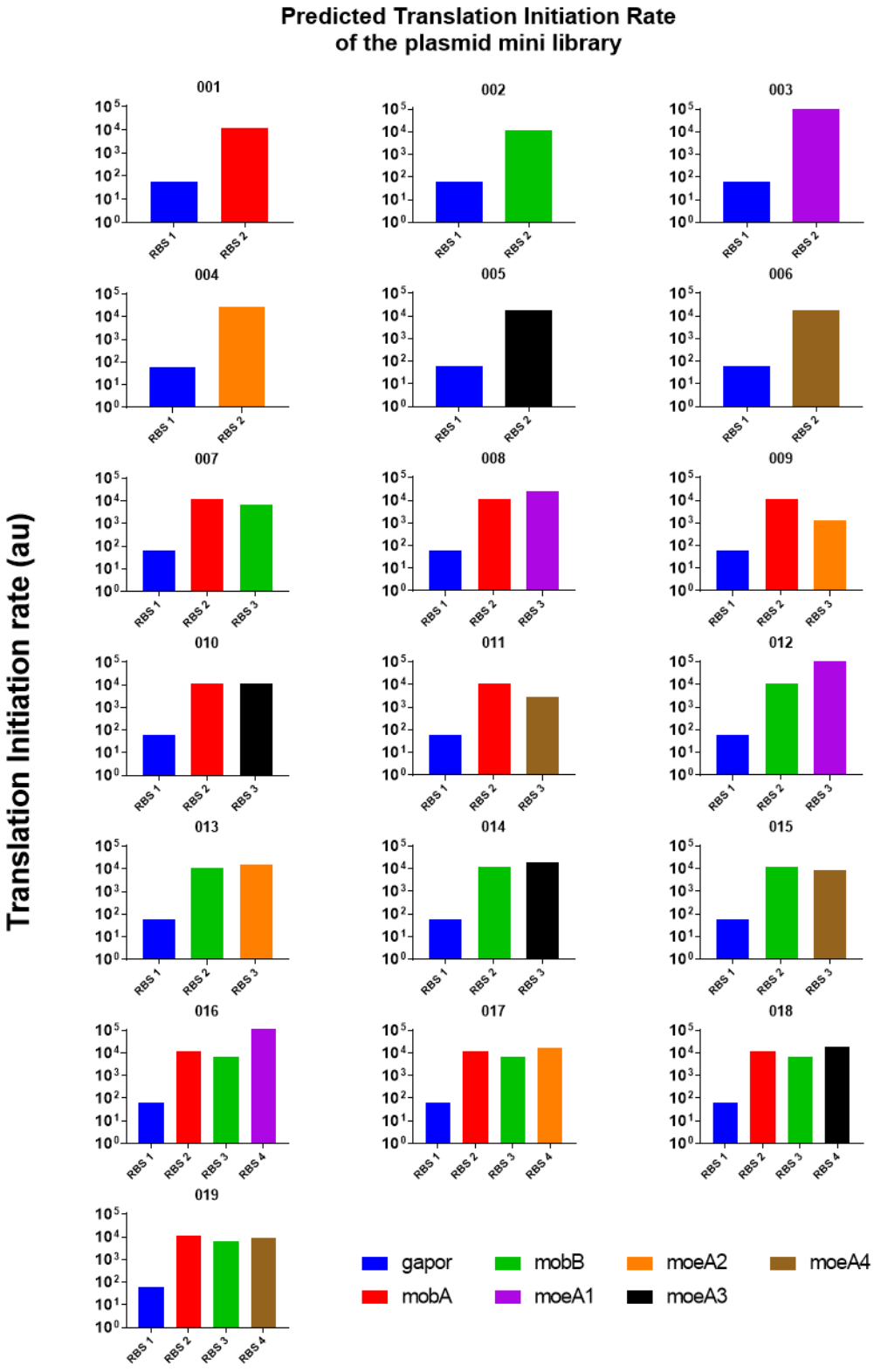
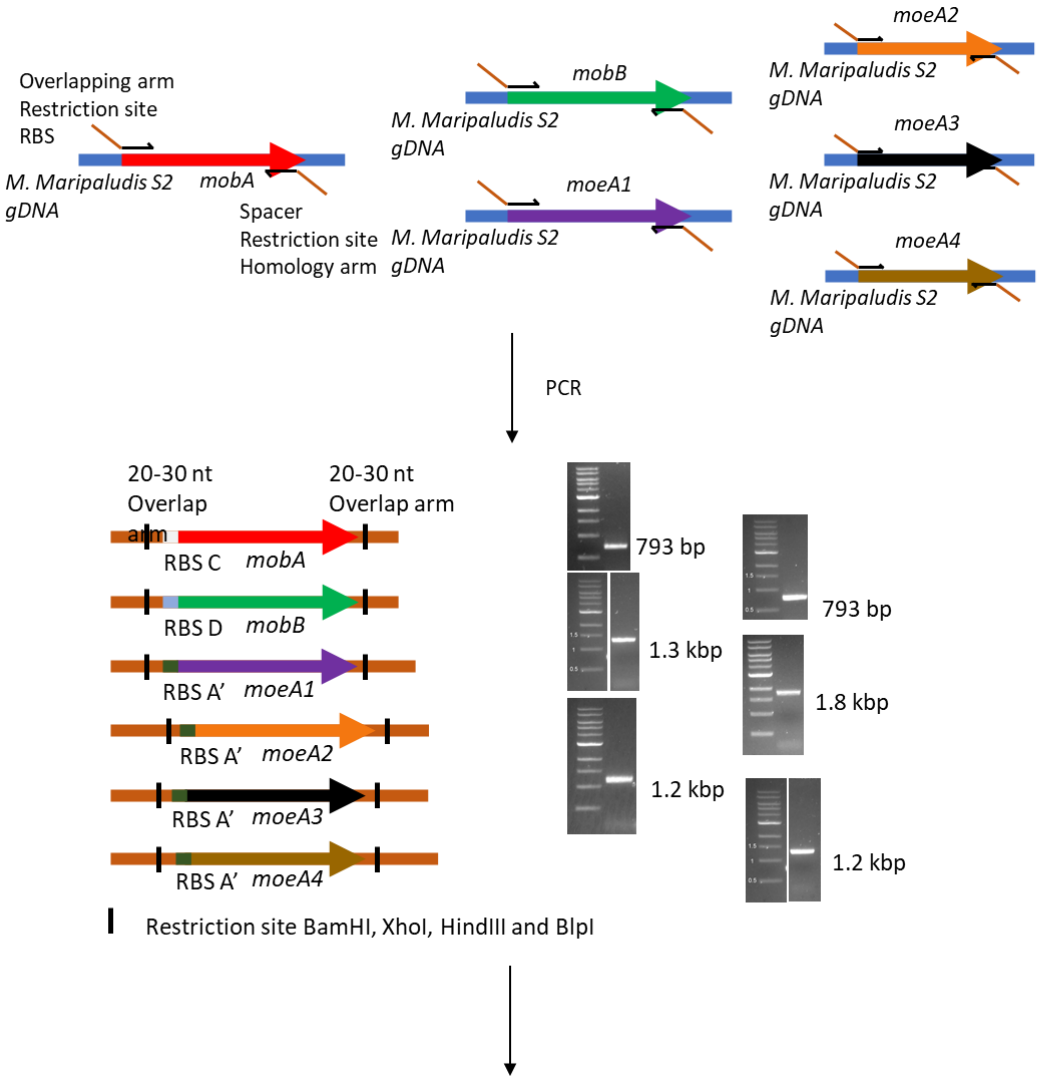


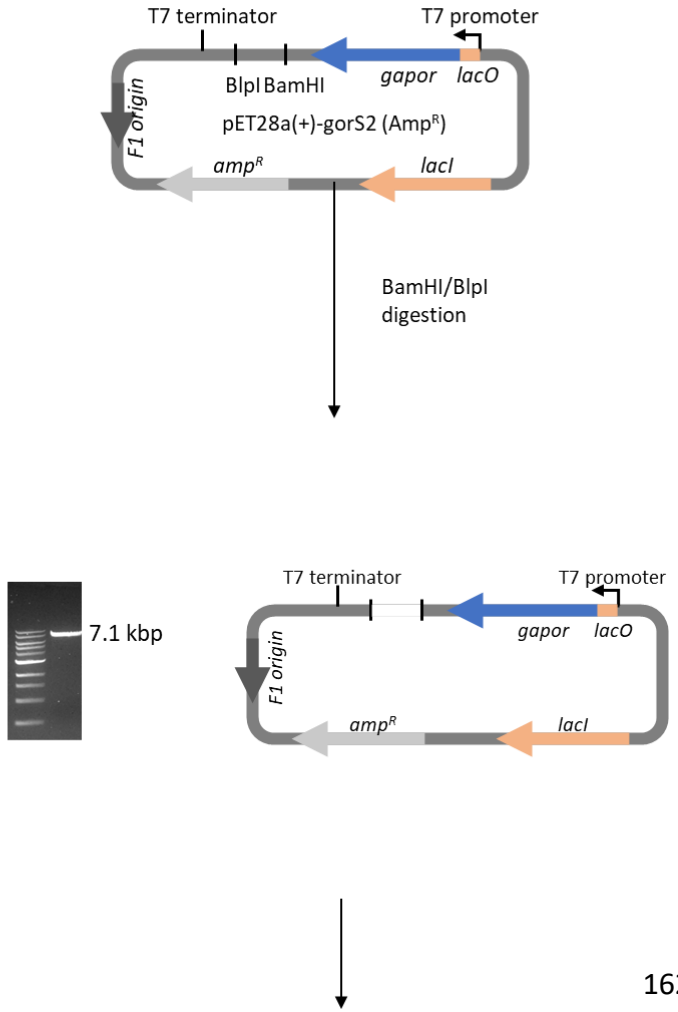
Figure IV-5 Predicted translation initiation rate of plasmids mini library. The transcription initiation rates were calculated for each operon in *E. coli* str. k-12 substr. MG1655 using De Novo DNA Operon calculator (https://salislab.net/software/predict_operon_calculator) (Cetnar & Salis, 2020; Farasat et al., 2014; Halper et al., 2020; Ng et al., 2015; Reis & Salis, 2020). *Gapor,* here referring *gorS2,* was predicted to have a translation initiation rate of 60.6. The Y-axis is in log10.

Table IV-2 HiFi Assembly genes combination and plasmids host strains

Plasmid	***mobA***	***mobB***	***moeA1***	***moA2***	***moeA3***	***moeA4***
pAM_001	+					
pAM_002		+				
pAM_003			+			
pAM_004				+		
pAM_005					+	
pAM_006						+
pAM_007	+	+				
pAM_008	+		+			
pAM_009	+			+		
pAM_010	+				+	
pAM_011	+					+
pAM_012		+	+			
pAM_013		+		+		
pAM_014		+			+	
pAM_015		+				+
pAM_016	+	+	+			
pAM_017	+	+		+		
pAM_018	+	+			+	
pAM_019	+	+				+

Mutant strain	Transformed plasmids
E. coli RG_AM024	pAM_003; pAM_004, pAM_005, pAM_006
E. coli RG_AM025	pAM_001, pAM_002, pAM_007
E. coli RG_AM026	pAM_008, pAM_009, pAM_010, pAM_011, pAM_012, pAM_013, pAM_014, pAM_015, pAM_016, pAM_017, pAM_018, pAM_0119


Overlapping arm
Restriction site
RBS
M. Maripaludis S2 gDNA
mobA
Spacer
Restriction site
Homology arm
mobB
M. Maripaludis S2 gDNA
moeA1
M. Maripaludis S2 gDNA
moeA2
M. Maripaludis S2 gDNA
moeA3
M. Maripaludis S2 gDNA
moeA4
M. Maripaludis S2 gDNA
PCR
20-30 nt Overlap arm
20-30 nt Overlap arm
RBS C mobA
RBS D mobB
RBS A' moeA1
RBS A' moeA2
RBS A' moeA3
RBS A' moeA4
793 bp
1.3 kbp
1.2 kbp
793 bp
1.8 kbp
1.2 kbp
Restriction site BamHI, XhoI, HindIII and BlpI


T7 terminator
T7 promoter
BlpI BamHI
gapor
lacO
F1 origin
pET28a(+)-gorS2 (AmpR)
ampR
lacI
BamHI/BlpI digestion
7.1 kbp
T7 terminator
T7 promoter
gapor
lacO
F1 origin
ampR
lacI

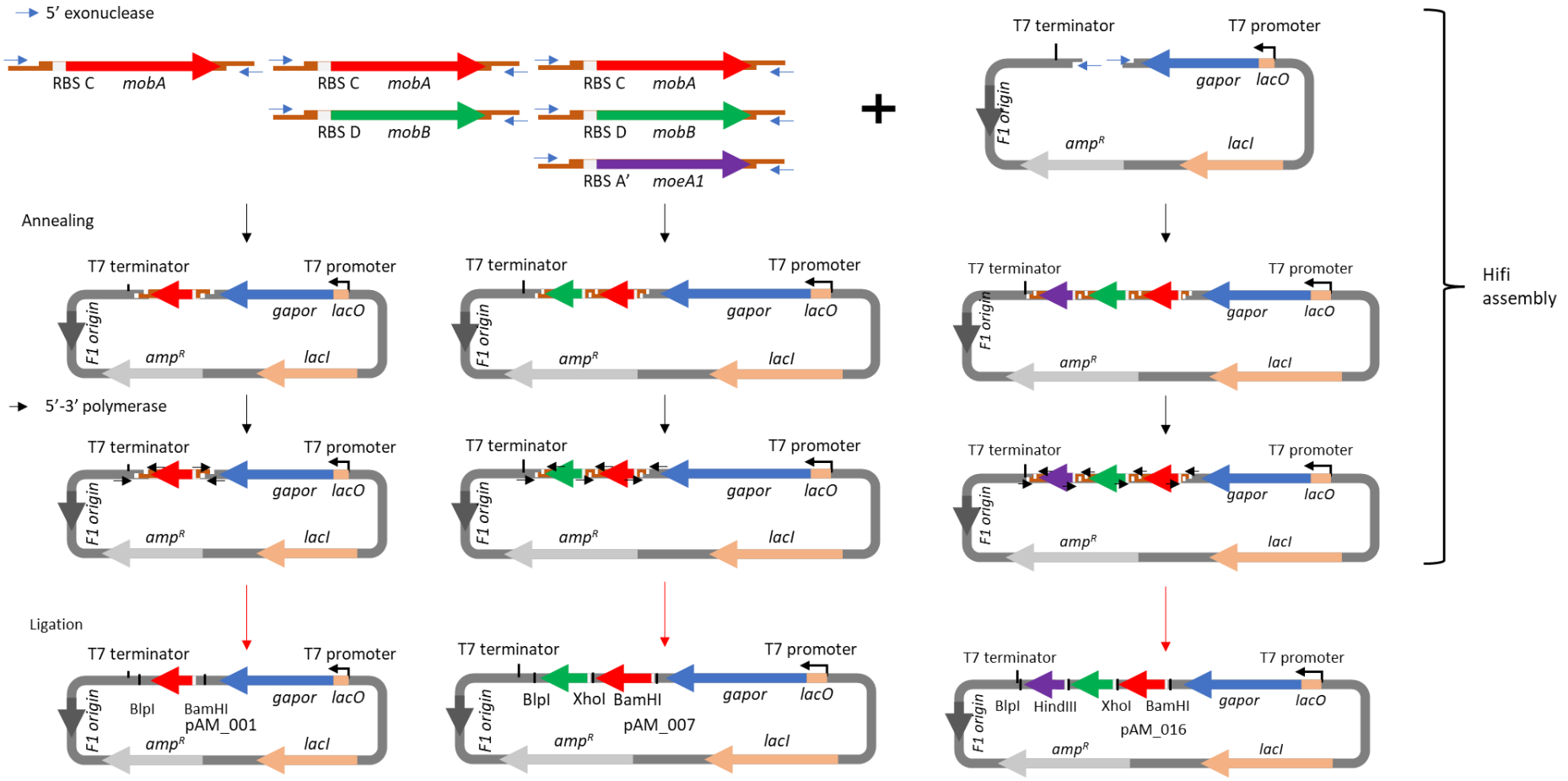

5' exonuclease
RBS C
mobA
RBS D
mobB
RBS A'
moeA1
+
T7 terminator
T7 promoter
gapor
lacO
F1 origin
amp^R
lacI
Annealing
5'-3' polymerase
Ligation
BlpI
BamHI
pAM_001
XhoI
pAM_007
HindIII
pAM_016
Hifi assembly

Figure IV-6 Scheme of cloning strategy of *M. maripaludis* molybdenum cofactor genes along *gorS2*.The construction of *mobA*, *mobAmobB* and *mobAmobBmoeA1*-containing plasmids is presented here as examples. The *M. maripaludis* genes were amplified by PCR using primers with an overhang sequence. These overhang sequences containied 20-30 nt overlapping sequence for annealing, the new RBS for the forward primers, a spacer for the reverse primers, anda restriction site. The overlapping sequence needed to have a Tm above or equal to 48°C. Thenwere adapted in function of the surroundings and the possibility to have the best conditions for the annealing. The pET28a(+) *gorS2* (Amp^R) vector were digested with BlpI and BamHI. After purification, the fragments were mixed according to Table IV-2. The mix was then transformed into *E. coli* DH5α. The plasmids were then validated by sequencing.

4. Control of the molybdenum cofactor enzyme production

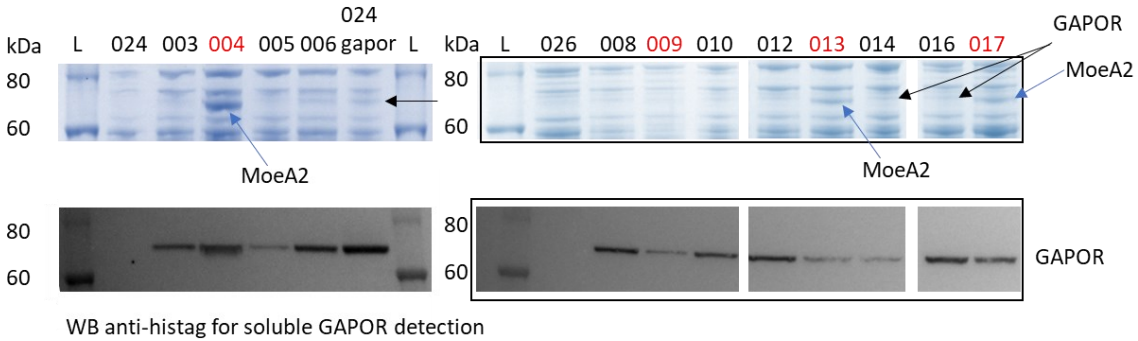
Once the strains were constructed and validated, a first protein production assay was performed. The six theoretical sizes of the archaeal protein were calculated using Expasy compute pI/Mw tool (https://web.expasy.org/compute_pi/). MobA was calculated to be a 24.8 kDa protein, MobB a 25.1 kDa, MoeA1 a 45.1 kDa, MoeA2 a 68.2 kDa, MoeA3 a 41.1 kDa and MoeA4 a 41.1 kDa.

MoeA2 was the easiest to visualise from a cell-free extract (Figure IV-7). The MoeA2 encoding gene was expressed in *E. coli* RG_AM004-009-013 and 017. On the SDS-PAGE gel, both MoeA2 and GAPOR were well separated. MoeA2 was predicted to have a translation initiation rate 22-fold higher than GAPOR in RG_AM009, 260-fold higher in RG_AM0013 and RG_AM0017 and up to 450-fold higher in RG_AM004 (Figure IV-5). Those ratios seemed correct as the MoeA2 band is more visible in RG_AM004 (Figure IV-7 A and B) than for the three other strains. It is important to note that the method used to reveal the Western Blot using TMB for membrane is as sensitive as Blue Silver staining. So, in Figure IV-7B, where the western blot and the Coomassie-stained gel pictures are merged, it is possible to see the two bands together and the relative difference in expression between the two genes. Furthermore, on the opposite side, the GAPOR was poorly produced during the RG_AM009 culture, which led to poor production of MoeA2, coupled with a poor extraction, MoeA2 is not visible on the SDS-PAGE gel in Figure IV-7. GAPOR is only visible on the western blot, as only the encoding gene was expressed along with *gorS2*, and by extension, that RBS A' was functional, and the modification brought to this RBS did not abolish its function.

MoeA2 expression

A

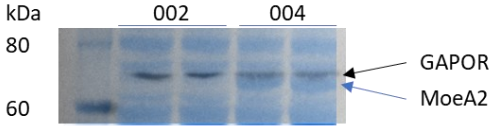
SDS-PAGE silver-blue Coomassie staining



WB anti-histag for soluble GAPOR detection

B

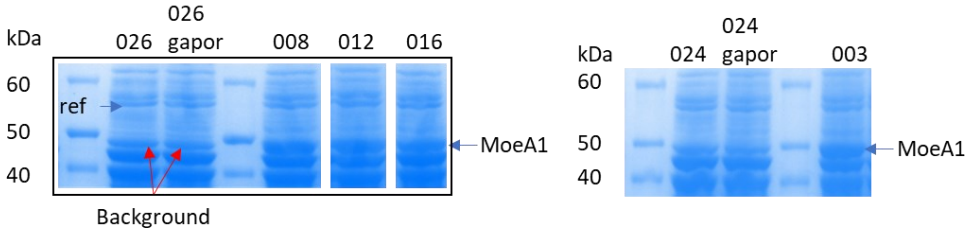
Superposition of Coomassie stained and WB anti-histag



MoeA1 expression

C

SDS-PAGE silver-blue Coomassie staining



D

Estimation of the relative quantity of MoeA1 against a reference band

E. coli strain	026	026gapor	008	012	016
	Estimated Mol. Wt. (KDa)	Relative Quantity compared to reference band			
Control band	55.61	1 (ref)	1.13	1.12	0.91	0.83
MoeA1 band	47.79	1.48	1.28	3.01	2.80	0.85

Figure IV-7 MoeA1 and MoeA2, the two detected proteins expressed from the ***E. coli*** Rosetta-gami2(DE3) mini library.Are presented here, the SDS-PAGE silver-blue Coomassie stained and western blot anti-histag of the lysis supernatant of 50 ml MAC anaerobic culture in 120 ml sealed serum bottle complemented with molybdate at room temperature of *E. coli* Rosetta-gami2(DE3) Δ*iscRhypFselA* RG_AM0XX) mutant mini-library deficient in either *moeA, mobAB,* or both. The strains were transformed with pAM0XX plasmids containing *gorS2* and one, two or three *mobA*, *mobB* or *moeA* as presented in Table IV-2 (pET28a(+) *gorS2* (AmpR)). Data presented in this figure show the evidence of soluble expression of MoeA1 and MoeA2. No evidence of the production of MobA, MobB, MoeA3 and MoeA4 in the lysate soluble fraction due to background proteins. A) Silver-Blue Coomassie stained 4%-12% Bis tris SDS-PAGE gel of *E. coli* RG_AM0XX strains crude extract, and the western blot anti-histag of the same gel revealed with TMB for membrane. In red are marked the strain expressing *moeA2* RG_AM004 (*ΔmoeA, gorS2moeA2*), RG_AM009 (*ΔmoeAmobAB, gorS2moeA2*), RG_AM013 (*ΔmoeAmobAB, gorS2moeA2*) and RG_AM017 (*ΔmoeAmobAB, gorS2mobAmobBmoeA2*). The Black arrows mark the GAPOR Bands (70 kDa), and the blue arrows mark the MoeA2 bands (68 kDa). RG_AM024 (*ΔmoeA*), RG_AM024gapor (*ΔmoeA, gorS2*), RG_AM003 (*ΔmoeA, gorS2moeA1*), RG_AM005 (*ΔmoeA, gorS2moeA3*), RG_AM006 (*ΔmoeA, gorS2moeA4*), RG_AM026 (*ΔmoeAmobAB*), RG_AM008 (*ΔmoeAmobAB, gorS2mobAmoeA1*), RG_AM010 (*ΔmoeAmobAB, gorS2mobAmoeA3*), RG_AM012 (*ΔmoeAmobAB, gorS2mobBmoeA1*), RG_AM014 (*ΔmoeAmobAB, gorS2mobBmoeA3*), RG_AM016 (*ΔmoeAmobAB, gorS2mobAmobBmoeA1*) were used as negative controls. B) Identical procedures were followed to obtain these results. Here, only the lysis supernatant of duplicates cultures of RG_AM004 (*ΔmoeA, gorS2moeA2*) are presented, with RG_AM002 (*ΔmobAB, gorS2mobB*) being used as a non-*moeA2-expressing* control strain. The Silver-Blue Coomassie-stained SDS-PAGE and the revealed western blot picture were merged to show the difference between the GAPOR (black arrow) and the MoeA2 (blue arrow) bands. C) Observation of MoeA1 production in RG_AM008, RG_AM012, RG_AM016 and RG_AM003. RG_AM026 and RG_AM024 with and without the *gorS2* expression plasmids were used as control. The Constitutive *E. coli* background proteins band, marked with a red arrow, is observable at the same theoretical size as MoeA1 (45 kDa). *moeA1* was predicted to have the highest initiation translation rate (Figure IV-5**)**. MoeA1 bands are indicated with blue arrows. Hence, comparing the intensity difference between MoeA1-producing strains and background protein natively expressed by *E. coli* even in control. So, to estimate the amount of protein loaded on the gel, a band present in all the lanes was used as a reference (marked ref in the figure) at 55 kDa. The results of the relative quantities are presented in table D. D) The relative amount of protein of the MoeA1 size band and the 55 kDa reference band in each lane of the RG_AM026 gel (C) was calculated by densitometry using ImageLab 6.1 (BioRad) and RG_AM026 55 KDa band as reference. The software estimated the MoeA1 band at 48 kDa instead of the theoretical 45 kDa.

This led to assume that the other MoeAs encoding genes were also expressed, as the same RBS drove their translation . In addition, the predicted translation initiation rate of *moeA1* was predicted to be 6 to 20-fold higher than *moeA2*. The expression of *moeA1* was controlled on SDS-PAGE by comparison of RG_AM008, RG_AM012, and RG_AM016 against the control RG_AM026 and RG_AM026gapor around 45 kDa (Figure IV-7 C). A band was visible in the controls' lanes and in the *moeA1* expressing strain lane. The band in the test lanes seemed bigger and brighter. To verify that a similar amount of proteins was loaded in the test lane compared to the control, a consistent protein with a 55 kDa MW was chosen as a reference. The relative amount was then measured by densitometry with ImageLab 6.1 (BioRad) (Figure IV-7 D). Data shows that an equivalent amount of the 55 kDa protein was present in all wells (between 0.85 to 1.05 in relative intensity). The same was observed for

RG_AM024 and RG_AM003. The intensity of the 45 kDa band was 1.48-fold higher than the intensity of the 55 kDa reference in the control lane RG_AM026. In RG_AM008 and RG_AM012, the intensities were about 3-fold higher than the reference band once normalised. Therefore, it could be concluded that MoeA1 seemed to be produced in all the strains. However, there is no direct proof of expression of the two other *moeAs* encoding genes. In Figure IV-7 C a very thick band is present in all the lanes below the MoeA band, and no difference in intensity were observed on the densitometry curve compared to the control.

The same experiment was performed, but the protein extracts were separated on 16% tricine gel (Novex^{TM}) to improve the protein separation between 40 kDa and 10 kDa. Despite improving the protein separation in the desired range, bands were observed at the target molecular weight, including in the lane of the control strain. Moreover, no change in intensity was detected at those molecular weights. 16% tricine gel was not the most appropriate gel as it performs better in separating protein below 10 kDa.

A lower signal was expected for these proteins as less Coomassie brilliant blue could bind to them, a higher amount would be required to observe them at the same level as GAPOR or MoeA1/2. However, as shown in Figure IV-5, most of the constructs should have a similar level of expression as *moeA2*. The only gene with a significantly higher predicted translation initiation rate was *moeA1*, which protein could be detected on gel. Nevertheless, the remaining genes had a similar predicted translation initiation rate as *moeA2*, but the background signal was low enough to allow the visualisation of the protein on a gel.

Therefore, only two of the *M. maripaludis* genes were expressed in a soluble form, and the gene product was visualised. But these two observations helped us to conclude that the chosen RBS was functional and that *moeA3* and *moeA4* might have also been expressed. Moreover, *moeAs* were the last genes of the operons (Figure IV-4) when the genes were present. Thus, it was concluded that the transcription of the operons was complete and that neither *mobA* nor *mobB* presence were affecting the transcription of the following genes. To be

able to detect the MobA, MobB, MoeA3 and MoeA4, two methods were identified. The first one would have been introducing a tag in the protein sequence (for example, a Strep-tag) to allow its detection by Western blot. Also, a combination of tags could have been used to differentiate the protein with similar molecular weight, like MobA and MobB (about 25 kDa), when encoding genes were expressed together (RG_AM007 and RG_AM0016 to RG_AM019). The second could have been to compare the peptide sequence of the band supposed to contain one of the target proteins against the same band from a control strain. It is supposed that it would be possible to identify the targeted peptides against the background thus. Nevertheless, the protein bands assumed to correspond to MoeA1 and MoeA2 should have been sequenced for validation.

To go further, the complete transcription of the operon could have been confirmed by RT-qPCR. It was not performed as, as stated above, the last gene of the operons was translated meaning that there was no disturbance of the operon mRNA transcription.

Overall, only two of the six enzymes were spotted on gels. It was strongly supposed that the two other *moeAs* encoding genes were expressed as their RBS was confirmed to be functional. Also, the mRNA seemed to be completely transcribed. The RBSs were assumed to work as published by Zelcbuch et al. for the *moeA1* and *moeA2* (Zelcbuch et al., 2013). Hence it was decided to go ahead with the experiment.

In fact, when the experiment was initially designed, the goal was to improve the GAPOR specific activity of an already active GAPOR as published by Park et al. (Park et al., 2007). The aim was to improve the incorporation of the cofactor to increase the GAPOR specific activity in the ultimate target of the combination of GAPOR, ferredoxin and CODH pathway rather thanto get an active GAPOR. Thus, if the gene coding for the cofactor synthesis pathway were not expressed in the soluble fraction – actively expressed, it would not have improved the activity and could not have been used for complementation. These experiments and the following experiments were performed in parallel

with Chapter III:3.3.2 experiments, so to evaluate which enzyme could improve the GAPOR activity, it was decided to go ahead with the experiment.

5. Enzyme assay

After the library's construction, the GAPOR assay was ready to be performed. The results of our experiments with heterologous *gorS2* expression using the Park et al. protocol to measure G3P-dependent GAPOR activity was negative. Therefore, the analysis of the new data was focused on the detection of activity.

5.1. Assays done in Nottingham

A first round of assay on crude extract was performed. This series of assays was planned as a first screening of the strains. The strain showing an activity would be re-tested for confirmation before moving forward with purification and calculating their kinetic constant. The cultures were grown at room temperature in 50 ml MAC medium in sealed serum bottle. In this first round, all the strains were induced with IPTG. The cells were harvested, lysed, and crude extracts were desalted the same way *E. coli* RG_AM023gapor presented in Chapter III:3.3.2. In the preceding chapter no G3P-dependent activity was observed when RG_AM023gapor was tested. Thus, these experiments, which were planned to check the possible improvement of the GAPOR activity with different cofactor biosynthesis enzymes, became a way to observe any potential GAPOR activity.

All the strains were tested for 5 minutes with G3P. During this first round of tests, 8 strains showed some activity ranging from 0.0219 mU per mg of total protein (RG_AM016) to 0.846 mU per mg of total protein (RG_AM0018). Additionally, the order of magnitude of the activity is 10^{-7}-fold lower than the measurement presented by Park et al. (Park et al., 2007). The difference might mainly be explained by the fact that they performed their assay on purified GAPOR compared to the work presented here. Interestingly, except for RG_AM002, all the strain that showed some reduction of the benzyl viologen expressed the *mobB* gene. Noteworthy that the assays were performed in

batches of three or four strains, so the absence of activity for a batch might have had an origin in the status of the tent when the experiment was performed.

All the strains displaying some BV-reducing activity were tested a second time in biological duplicate. The assays were performed with and without G3P. Moreover, the original RG strain crude extract was used to control if oxygen were present in the tent atmosphere and solutions in a too high amount. As presented in Chapter III:3.1, this strain displayed an intense background activity. Therefore, if BV-reduction did not occur during the less strong than usual, it signalled that oxygen was present in an excessive amount at the time of the assay. The activity calculated from those assays are presented in Figure IV-8.

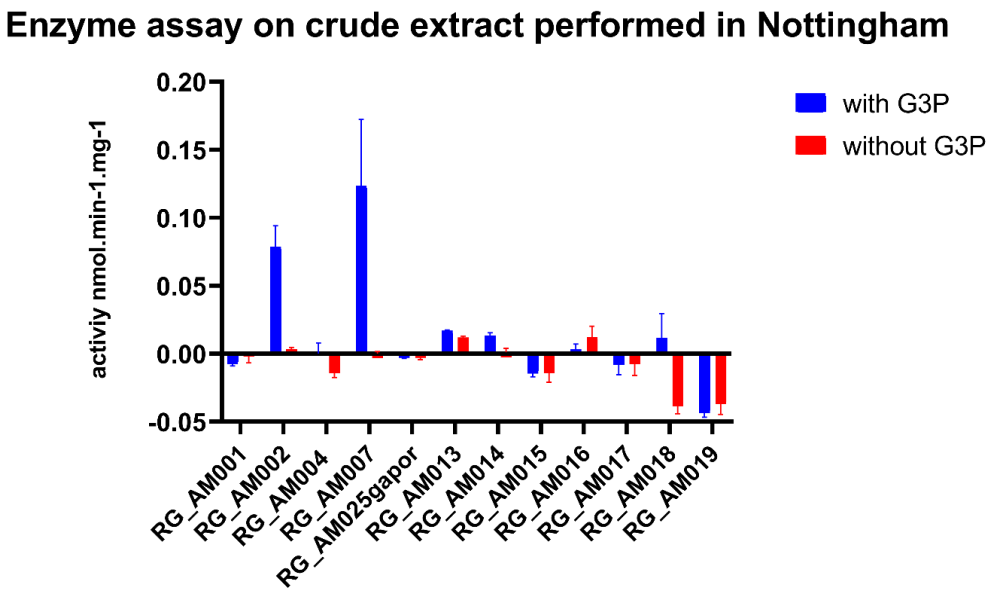


Figure IV-8 Enzyme assay on crude extract performed in Nottingham. The strains were tested separately with biological duplicates, and *the E. coli* RG strain (*E. coli* Rosetta-gami2(DE3) Δ*iscRhypFselA* pET28a(+) *gorS2* (AmpR)) was used as oxygen control during the assay. If no BV reduction was observed during the control strain crude extract assay, the tent was considered unfit to perform the assay. Thus, the assay was re-attempted later. Single points measurements at 600 nm were taken every 30 min up to 300 min. The reaction solution: 500 µl of 2x reaction mix (30 µl of DL-G3P, 3 mM BV, 56 µM Na2MoO4, 50 mM EPPS pH8, final concentration), 400 µL of 10 mM potassium phosphate buffer, the assay was started with 100 µl of desalted crude extract. The error bars represent the standard deviation calculated by GraphPad Prism 9. *E. coli* Rosetta-gami2(DE3) Δ*iscRhypFselA* RG_AM0XX) mutant mini-library efficient in either *moeA, mobAB,* or both accordingly to the plasmids content. The strains were transformed with pAM0XX plasmids containing *gorS2* and one, two or three *mobA*, *mobB* or *moeA* as presented in Table IV-2 (pET28a(+) *gorS2* (AmpR)). RG_AM001 (*ΔmobAB, gorS2mobA*), RG_AM002 (*ΔmobAB, gorS2mobB)*, RG_AM004 (*ΔmoeA, gorS2moeA2*), RG_AM007 (*ΔmobAB, gorS2mobAmobB*), RG_AM025gapor (*ΔmobAB, gorS2*), RG_AM013 (*ΔmobABmoeA, gorS2mobBmoeA2*), RG_AM014 (*ΔmobABmoeA, gorS2mobBmoeA3*),

RG_AM015 (*ΔmobABmoeA, gorS2mobBmoeA4*), RG_AM016 (*ΔmobABmoeA, gorS2mobBmoeA1*), RG_AM017 (*ΔmobABmoeA, gorS2mobBmoeA2*), RG_AM018 (*ΔmobABmoeA, gorS2mobBmoeA3*), RG_AM019 (*ΔmobABmoeA, gorS2mobBmoeA4*).

As observed during the first round, the activities of BV reduction were low. Thus the absorbance measurement procedure was modified. So, instead of automatically taking a measurement every 7 seconds for 5 min, a single measurement was taken every 30 min, and then all the assays were run simultaneously. Hence, it prevented potential protein degradation and mainly reduced the overall experiment time as the measurements took up to 300 min. Additionally, the cabinet temperature was risen to 37°C during the assay to prevent a fall of the cuvette temperature during the assay mimicking the first round. It can be observed in Figure IV-8 that only two assays displayed a G3P dependent BV reduction. RG_AM002 and RG_AM007 showed an enzymatic activity of 0.0548 mU per mg of protein and 0.089 mU per mg of total protein. After the calculation of the t-test between the two, there was no significant difference (P value of 0.3383). None of the other strains displayed any BV reduction contrary to the first round's observation. The activity's order of magnitude was very low. This brought the question of the repeatability of the experiments. Indeed, there was a difference of 10-fold between the activity measured. Also, the way the experiment was performed was adding some variability to the measurement due to the potential modification of the cuvette positioning in the single cuvette holder, contrary to the first round of assay where the cuvette was not moved during the entirety of the monitoring.

G3P-dependent activities were detected in two strains, RG_AM002 expressing *moeA2* and RG_AM007 expressing *mobA mobB*. However, RG_AM017, which should express the three genes, does not present any activity. These enzymes impact different stages of the molybdenum cofactor biosynthesis pathway, MoeA2, the insertion of the Mo and MobAMobB, the formation of the MGD (Chapter I:6.3). It was then difficult to understand which reaction was key in synthesising the cofactor when both individually studied seemed to impact the GAPOR activity positively, and the combination of the three did not reduce BV.

The variability of the results observed between the two rounds might be explained by the change in the measurement procedure, which would explain why the activities measured were 10-fold lower. The daily utilisation of the tent might potentially bring some changes to the anaerobic atmosphere condition despite the care taken to maintain it the best possible without an accurate monitor.

At this stage of the experiments, two strains presented more activity than RG_AM023gapor, the original test strain. Nevertheless, there were many doubts regarding this set of experiments as it was not understood why the BV reduction activity were not confirmed in the second assay.

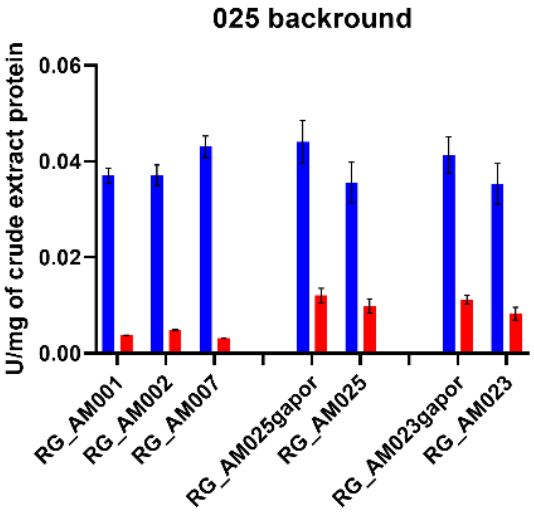
Due to the absence of activity observed in RG_AM023gapor, it was already decided to go to Toulouse to perform this experiment in a more controlled environment. Consequently, it was agreed to reperform the library crude extract experiment at the same time. In Toulouse, as stated in the preceding chapter (Chapter III:3.3.2 and Chapter III:5), the spectrophotometer is placed in an anaerobic chamber, and the cuvette holder contains 8 positions to run several conditions simultaneously. The principal concern was the presence of traces of oxygen in the tent, the chamber in Toulouse had proper monitoring of the oxygen, and it was maintained below 10 ppm.

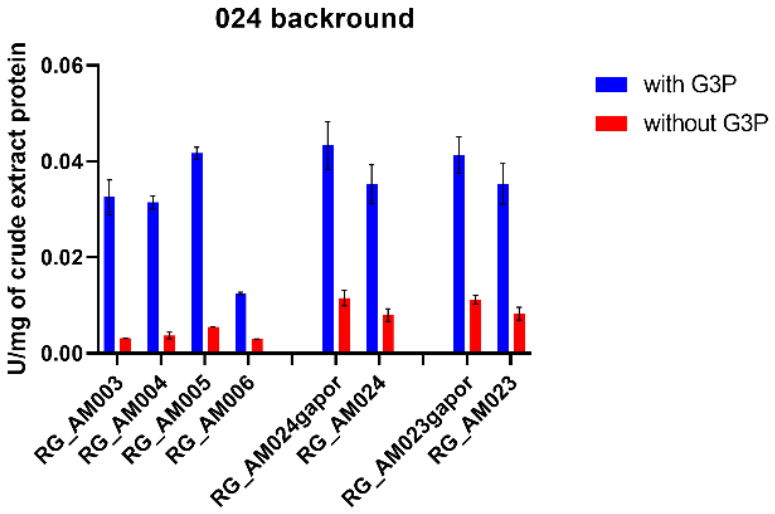
Hence, at this stage, it was concluded that activities could be measured but that better experimental conditions were required to confirm or invalidate these first observations.

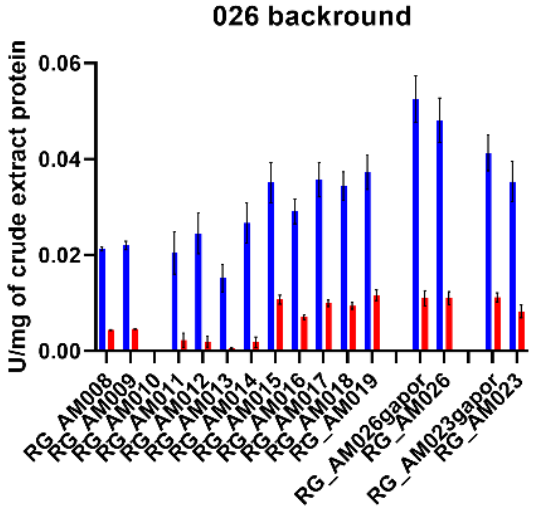
5.2. Assays performed in Toulouse

After the first two rounds of experiments, a new complete round was performed in the Toulouse laboratory along with the test of *E. coli* RG_AM023gapor presented in the Chapter III:3.3.2. There, the 19 strains of the library were tested, as well as the three control strains with and without the standard *gapor* containing plasmid pET28a(+) gorS2 (Amp^R). The cultures were performed the same way as in Nottingham (Chapter IV:5.1); however, the lysis methods were modified to suit the product available in the lab. Thus, the

harvested cells were resuspended in 2.5 ml phosphate buffer and lysed by sonication under anaerobic conditions. The volume of crude extract added to the reaction solution was proportionally increased to 250 µl to compensate for the increased crude extract dilution. Moreover, the concentration of G3P was also increased, like for RG_AM023gapor assay, to 100 µM final to fit the GAPOR optimum. Finally, to ensure that the remaining traces of oxygens were removed from the cuvette, 100 µl of sodium dithionite at 13.5 µM final was added. No reduction of BV was observed in the blanks.


025 backround
U/mg of crude extract protein
0.06
0.04
0.02
0.00
RG_AM001
RG_AM002
RG_AM007
RG_AM025gapor
RG_AM025
RG_AM023gapor
RG_AM023


024 backround
U/mg of crude extract protein
0.06
0.04
0.02
0.00
RG_AM003
RG_AM004
RG_AM005
RG_AM006
RG_AM024gapor
RG_AM024
RG_AM023gapor
RG_AM023
with G3P
without G3P


026 backround
U/mg of crude extract protein
0.06
0.04
0.02
0.00
RG_AM008
RG_AM009
RG_AM010
RG_AM011
RG_AM012
RG_AM013
RG_AM014
RG_AM015
RG_AM016
RG_AM017
RG_AM018
RG_AM019
RG_AM026gapor
RG_AM026
RG_AM023gapor
RG_AM023

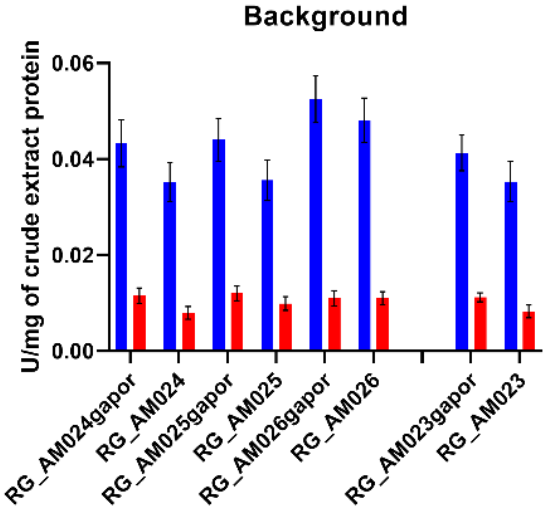

Background
U/mg of crude extract protein
0.06
0.04
0.02
0.00
RG_AM024gapor
RG_AM024
RG_AM025gapor
RG_AM025
RG_AM026gapor
RG_AM026
RG_AM023gapor
RG_AM023

Figure IV-9 Results of the crude extract assay performed in Toulouse The cultures were performed identically to those in Nottingham. The cells were harvested and collected the same way, but the lysis was performed by sonication in 2.5 ml phosphate buffer. 500 µl of reaction mix (50 mM EPPS, 3 mM Benzyl viologen and 56 µM Na2MoO4) were complemented with 100 µl of sodium dithionite, 50 µl, 175 µl and 220 µl of water for the test with 250 µl, 125 µl and 60 µL of crude extract respectively. The measurement was started, and 250 µl, 125 µl and 60 µL of crude extract were added (black vertical bar). For the blank, the crude extract was replaced with water. The reactions were started by adding 100 µl of 1 mM G3P (final concentration 100 µM) or water for the control without G3P. The reaction was monitored for 15 min. A full spectrum was measured every 16 sec. The experiment was performed in batches of 3 to 5: RG_AM001/002/007 (*gorS2mobA*, *gorS2mobB*, *gorS2mobAmobB*), RG_AM003/004/005 (*gorS2moeA1, gorS2moeA2, gorS2moeA3*), RG_AM006/008/009/010 (*gorS2moeA4, gorS2mobAmoeA1, gorS2mobAmoeA2, gorS2mobAmoeA3*), RG_AM011/012/013/014 (*gorS2mobAmoeA4, gorS2mobBmoeA1, gorS2mobBmoeA2, gorS2mobBmoeA3*), RG_AM015/016/017/018/019/023gapor (*gorS2mobBmoeA4, gorS2mobAmobBmoeA1, gorS2mobAmobBmoeA2, gorS2mobAmobBmoeA3, gorS2mobAmobBmoeA4, gorS2*), RG_AM024gapor/025gapor/026gapor (*ΔmoeA gorS2, ΔmobAB gorS2, ΔmobABmoeA gorS2*), and RG_AM023/024/025/026 (*ΔmoeA, ΔmobAB, ΔmobABmoeA*). No oxygen control was performed as the atmosphere of the anaerobic chamber was precisely monitored. The error bars represent the linear regression error on the measurement.

The calculated activities are presented in Figure IV-9. All the strained displayed G3P-dependent BV reduction activity. However, these activities are not GAPOR dependent. Indeed, as observed for RG_AM023 (with and without GAPOR expressing plasmid), there was a strong G3P-dependent background activity. The observed activity in the GAPOR-less strains was at the same level as GAPOR expressing strain. With the control strains (mutant strains transformed with *gorS2*-only plasmid), the *gorS2* expressing strains seemed to show a stronger activity than the *gorS2*-less counterparts. However, the differences were not significant (within the measurement error). Additionally, the G3P independent activity within the control strains was also lower for all the *gorS2*-less strains.

Hence, the variation of activity between the two were probably not due to the GAPOR but more surely by measure bias and testing conditions. The highest activity measured was from RG_AM026, and for all the other tested strains, G3P-dependent BV reduction activities were like the background *gorS2*-less strains. Again, to optimize the available time in the chamber, the assays were performed in batches of three to 5. It seems that two batches lead to lower activities (RG_AM006/8/9/10 and RG_AM011/12/13/14) than the last batch (RG_AM015/16/17/18/19). Thus, the assay conditions seem to have more impact than the presence of GAPOR and other heterologous-produced enzymes. Hence, some solutions for these two specific batches were supposed

not to have been prepared as correctly as the other batches. Therefore, no significant differences were observed between the tested strain.

The observed G3P-dependent activity was described in Chapter III:3.3.2, and NAD+ was supposed to still be present in the crude extract solution after the PD-10 desalting step. Better metabolite removal would have increased the chance of observing GAPOR-dependent BV reduction activity. It is important to remember that other enzymes use G3P as their substrate and that $NAD(P)^+$ dependent reactions were the only ones observed. In fact, some pyruvate was most probably still present in the assays' solution diverting G3P from GAPOR. In addition, BV was also diverted from GAPOR to complete the G3P oxidation. Hence the presence of other metabolites in significant amount was very detrimental toward the studied reaction. Unfortunately, the issue regarding the remaining NAD(P)+ was only observed at the end of Toulouse's first work period, leaving no time to reperform the experiments in suitable conditions.

These experiments emphasised the problem of oxygen in Nottingham's anaerobic tent, where no BV reduction was observed for most of the strain. The G3P-dependent enzyme reaction was most probably not oxygen sensitive, however the reaction rate was probably slower than the oxidation rate of BV with the oxygen trace. It was then impossible to anticipate this background activity before going to Toulouse. Also, the BV-reduction activity observed in the Nottingham assays was most certainly coming from less oxygen in the chamber and the solution.

Hence, no GAPOR activity was observed during the crude extract assay due to this G3P-dependent enzyme activity.

Due to these unexpected reactions, it was then decided to move completely to purified enzyme assays. However, as shown in (Chapter III:5), no GAPOR activity observed after purification. So, when back in Nottingham, as no more enzyme assay on purified enzyme would be possible under correct anaerobic conditions, it was decided to move to the investigation of the cofactor inserted in the GAPOR produced in the basic strain *E. coli* RG_AM023gapor.

6. Cofactor analysis

This experiment was planned to try to understand the reason for the absence of activity of the heterologously expressed *gorS2*. In *P. furiosus,* the cofactor identified in the GAPOR was a tungsten cofactor which is a tungsten atom coordinated by two molybdopterins (Loes E. Bevers, Hagedoorn, & Hagen, 2009; Chan et al., 1995). The bioinformatic study by Guerrini showed that *M. maripaludis* GAPOR was similar to *M. jannaschii* and *P. furiosus* GAPOR. As a consequence, the *M. maripaludis* GAPOR was supposed to have a similar type of cofactor. However, he showed that this GAPOR was not active in the presence of W (Guerrini, 2007). Park and co-workers later confirmed that *M. maripaludis* GAPOR was Mo-dependent (Park et al., 2007). It was then assumed that the cofactor was composed of a Mo atom coordinated by two molybdopterins. Hence it was necessary to analyse the cofactor incorporated in the studied GAPOR.

6.1. Presentation of the literature cofactor analysis

6.1.1. Detection and quantification of the molybdenum atom

The literature shows several methods for molybdenum cofactor analysis. One of them is measuring the number of metal atoms per enzyme molecule. Detecting the metal atom in the molybdenum cofactor is mainly done by inductively coupled plasma-mass spectrometry (ICP-MS). This is the titration method used by Guerrini and Park to determine the metal contained in the protein (Guerrini, 2007; Park et al., 2007). Another method based on ICP is the inductively coupled plasma-optical emission spectroscopy (Neumann, Mittelstadt, Iobbi-Nivol, et al., 2009). Both of these methods are quantitative and can detect trace elements. The total reflection X-ray fluorescence analysis used to detect the molybdenum in a Mo-MPT (Reschke et al., 2013) is an ultra-trace elements analysis tool used with very small samples (Wobrauschek, 2007). This method has also been used to quantify the Phosphorus atom in the molybdenum cofactor (Reschke et al., 2019; Reschke et al., 2013).

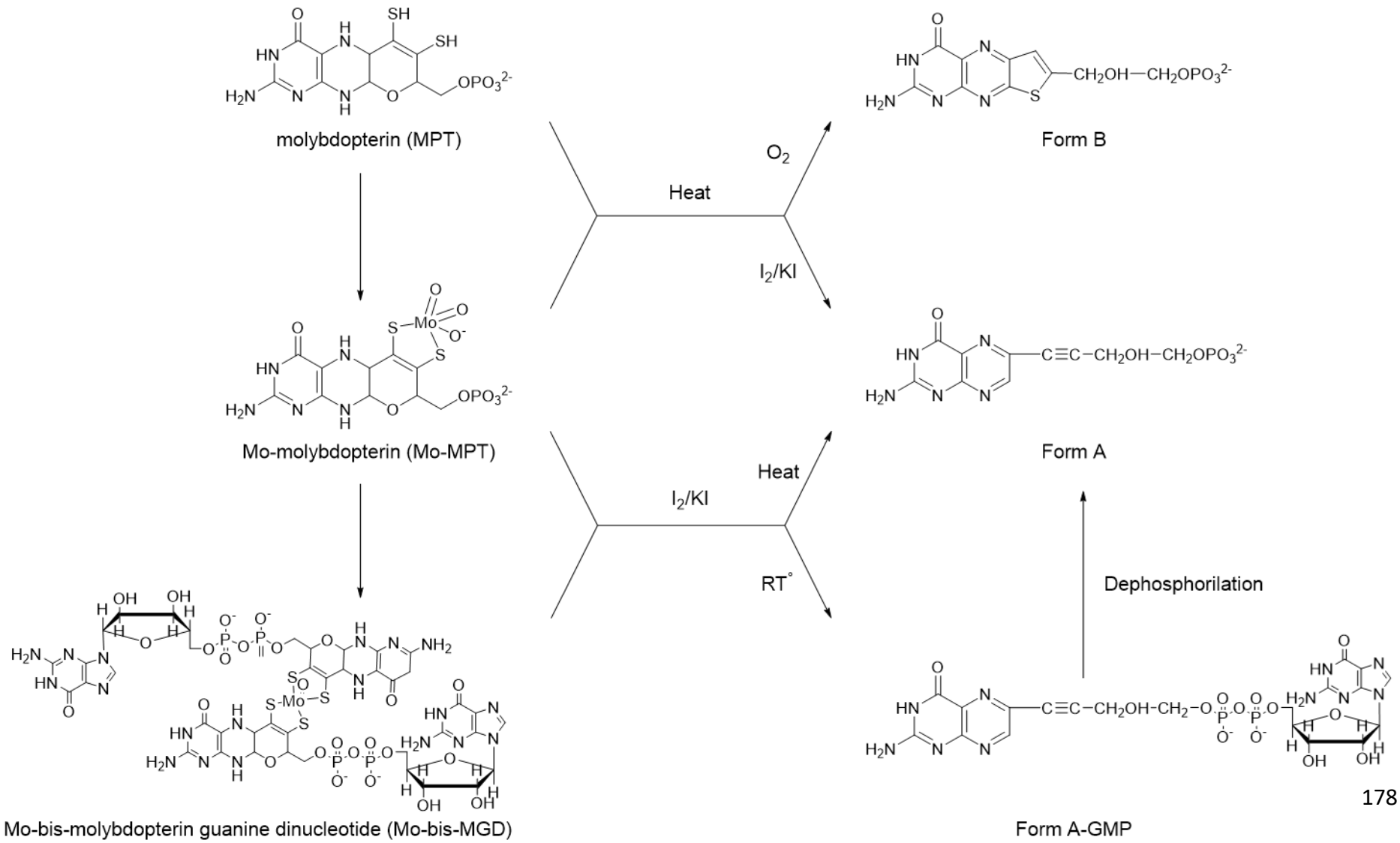

molybdopterin (MPT)
Mo-molybdopterin (Mo-MPT)
Mo-bis-molybdopterin guanine dinucleotide (Mo-bis-MGD)
Heat
O2
I2/KI
Form B
Form A
Heat
I2/KI
RT°
Dephosphorilation
Form A-GMP

Figure IV-10 Scheme of the obtention of the three most studied fluorescent molybdenum cofactor derivatives from MPT, Mo-MPT and Mo-*bis*MGD The structures were reproduced form Leimkühler et al. and Jon(J. L. Johnson, Indermaur, & Rajagopalan, 1991; Leimkühler et al., 2011)

6.1.1. Quantification of the molybdopterin

Associated with metal atom quantitation, the most used method in the literature to quantify pterin is to measure the fluorescence of a derivative compound. The molybdopterin with or without a molybdenum atom is very unstable (J. L. Johnson et al., 1990; A. Magalon & Mendel, 2015); hence, the "form A" (Figure IV-10) is produced by heating the molybdo-enzyme at pH 2.5 for 20 min at 100°C with 1% I_2/ 2% KI (J. L. Johnson & Rajagopalan, 1982).In the absence of the iodine oxidant, form B is obtain (J. L. Johnson, Hainline, Rajagopalan, & Arison, 1984). It is the degradation product obtained after air exposure once extracted from the protein (Leimkühler et al., 2011). An equivalent to this form exists for the Molybdopterin Guanin dinucleotide (Mo-MGD), the form A-GMP or with Molybdopterin cytosine dinucleotide (Mo-MCD) when no heat was applied during the treatment with I_2/KI (J. L. Johnson et al., 1991). This derivative product will have a lower fluorescence intensity due to the GMP's presence, which quenches it. During the reaction, a dephosphorylated form, the "form A (dephospho)", may be produced. The heating step can be replaced by overnight incubation at room temperature, to avoid dephosphorylation. Hence, all the "form A (dephospho)" collected may come from dephosphorylated molybdopterin (J. L. Johnson et al., 1991). These three compounds can be separated on a strong anion exchange HPLC column. The latter is eluted by 10 mM acetic acid, the form A by 10 mM HCl and the form A-GMP by 50 mM HCl (J. L. Johnson et al., 1991). Neumann *et al*. eluted those molecules in only two steps 10 mM acetic acid for form A and 50 mM for form A-GMP (M. Neumann et al., 2007). After converting the form A-GMP to form A by a dephosphorylation step, the fractions are analysed by C-18 reverse phase HPLC and fluorescence detection. Indeed, the form A has a more intense fluorescence than form A-GMP. The conversion is done by overnight digestion by pyrophosphatase. The extinction coefficient of form A at 380 nm is 13,200 $cm^{-1}.M^{-1}$ (J. L. Johnson et al., 1984). The important point is that form A no longer coordinates a molybdenum atom.

In combination with the Molybdenum quantification, this method also permits the determination of the nature of the Molybdenum cofactor in the Molybdo-enzyme. Indeed, for the Molybdo-pyranopterin, which is a Molybdenum atom coordinated by a single molybdopterin moiety, the ratio is 1, for the Mo bispyranopterin, two molybdopterin moieties, the theoretical ratio is 2 and for the Mo-bisMGD the theoretical ratio with the form A-GMP is also two (Reschke et al., 2013).

The combination of the first two presented methods was used for instance to characterise the cofactor structure (Leimkühler et al., 2011) and the role of MobA in the *E. coli* molybdenum cofactor biosynthesis pathway (Reschke et al., 2013).

6.1.2. Phosphate quantification

In addition to the total reflection X-ray fluorescence analysis, a chemical method can be used to titrate the phosphate group (Ames, 1966) in the molybdopterin. The Mo-MPT contains one phosphate per Mo, MGD contains four per Mo. Hence, the phosphate titration can be coupled with the Mo quantification to characterise the ratio (Hilton & Rajagopalan, 1996; Reschke et al., 2013).

6.1.3. GMP quantitation

GMPs are released from the molybdopterin by incubation of the molybdoenzyme in an acidic solution, *e.g.* with 5% sulfuric acid for 15 min at room temperature (Neumann, Mittelstadt, Iobbi-Nivol, et al., 2009) or with 0.1 N HCl for 1h at 100°C (Hilton & Rajagopalan, 1996). Then for both methods, the nucleotide is purified with C-18 column reversed phase HPLC. The quantitation was done inline by absorbance at 280 nm using guanine as a standard (Hilton & Rajagopalan, 1996). Then, with the ratio GMP:MO, the type of molybdenum cofactor can be determined. Indeed, with a ratio of around 2, it figures out an MGD.

Mass spectrometry is used as a characterisation method, as well as NMR or IR spectroscopy, for the structure and the composition of synthetic molybdopterin-like (Burgmayer et al., 2007; Ghosh, Samuel, & Schulzke, 2017).

6.2. Presentation of our cofactor analysis strategy

At the beginning of the study, it was planned to use the most used methods in the literature ICP-MS for the Mo and P quantitation and us the fluorescence assay to quantify the different types of molybdopterin potentially inserted in the GAPOR. However, contrary to the literature, our analytical lab was not equipped with any inline or offline fluorimeter. The only fluorimeter in the lab was for microplates which reduced the measurement accuracy due to the reduced light path.

Preliminary tests were conducted using a microplate fluorimeter to determine the range of concentration detectable by the equipment by establishing a linear standard curve. The standard molecule chosen was pterin-6-carboxylic acid which has the same fluorescence spectra as Form A (Meckenstock et al., 2001). Park and colleagues used this molecule as a standard (Park et al., 2007). This linear range was found between 0.8 µM and 50 µM of standard. To evaluate the method, we used as a positive control the Xanthine oxidase from bovine milk that contains a Mo-MPT (Meckenstock et al., 2001; Park et al., 2007; Rauh et al., 2004). Following the protocol used by Park et al. and set up by Meckenstock et al., a test using the positive control was performed. The protein was oxidised by permanganate and boiled for 15 min to extract the cofactor and form “Form A”. Theoretically, a 10.88 µM was expected in the tested sample; however, the measured concentration was 1.072 ± 0.089 µM. This represents about 10% of the expected amount of extracted cofactor. The problem encountered here probably came from the ammonium sulphate present in the xanthine oxidase resuspension.

Due to the absence of GAPOR activity observed with all the strains, there was a concern regarding the concentration of pterin, Mo and P that could be measured. The sensitivity of the measurement of the pterin concentration was

a limitation. In fact, the lowest concentration of the standards requires a complete extraction of the cofactor from 21 µg of GAPOR in 50 µl if all the GAPOR was inserted with the Mo-bisMPT. Also, the team running the ICP-MS within the University of Nottingham in Sutton Bonington was contacted, and they advised to provide samples containing 10 to 100 µg/L of Mo. At the time of the first experiment, only 0.558 µg was obtained with a production yield of 232.5 ng/L of culture. The two methods originally planned were requiring a certain amount of GAPOR that was considered too high at the time specially when Park et al used 0.5 mg of GAPOR for their published experiments.

Therefore, other analytical methods were looked for in collaboration with Dr David Tooth, who is responsible for the group's analytical platform (HPLC and LC-MS). First was considered to measure the intact protein mass of the cofactor-containing GAPOR of purified protein on LC-MS to identify the cofactor by its mass and quantify the protein simultaneously. However, a similar method was tested by Guerrini during his thesis but was unsuccessful(Guerrini, 2007). Hence, the cofactor identification had to be performed on extracted cofactor. The enzymes were then produced under micro-anaerobic conditions, identically as *gorS2* expression in Toulouse, to maximise the GAPOR production (Chapter III:5). However, the protein would be purified under aerobic conditions using a 1 ml Histrap affinity column on the Äkta apparatus. The aerobic conditions were not considered detrimental to the experiment, as Reschke and co-workers demonstrated that the molybdoenzymes YdhV stayed conformed after aerobic purification if there was no denaturation of the enzyme (Reschke et al., 2019). The sample would be treated at pH 3 on an ultrafiltration spin column (Vivaspin 2) 30 kDa molecular weight cut off To extract the cofactor (Figure IV-11) (Martín-Tornero, Gómez, Durán-Merás, & Espinosa-Mansilla, 2016). Control would be run at pH 8 for comparison purposes. All the filtra would then be run on MS and compared to the pterin-6-carboxylique acid. Xanthine oxidase would be used as a test and positive control.

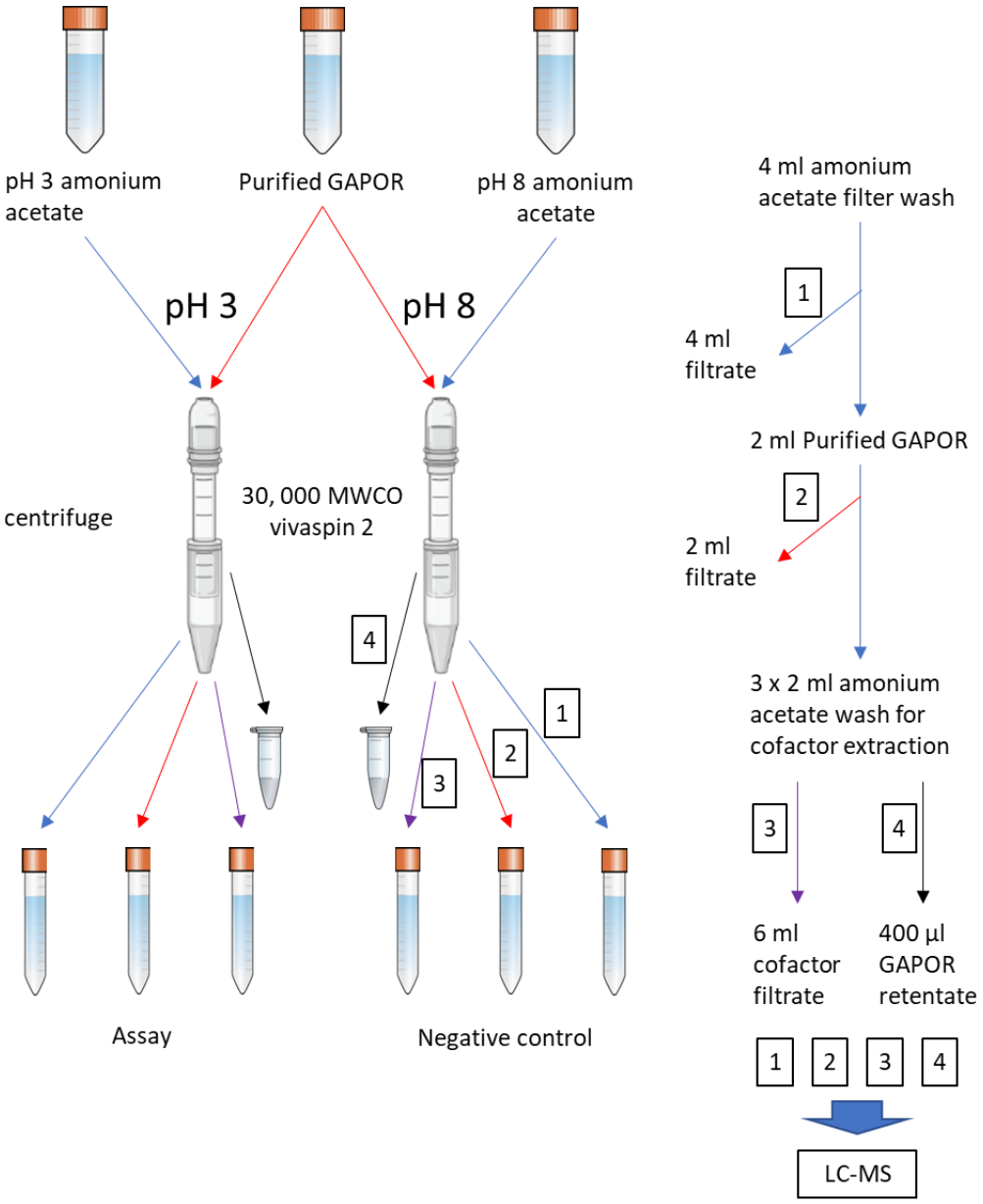


Figure IV-11 Scheme of the cofactor extraction process for LC-MS analysis.Scheme of the cofactor extraction process for LC-MS analysis. This scheme describes the process of acid (pH3) molybdenum cofactor extraction from GAPOR. The enzyme was produced with E. coli Rosetta-gami2(DE3) ΔiscRhypFselA pET28a(+) gorS2 (AmpR) (RG_AM023gapor) 600 ml MAC supplemented with 100 µM molybdate culture in sealed serum bottles. The cultures were incubated from the induction with 100 µM IPTG at room temperature and lysed by sonication. The GAPOR was purified under aerobic conditions using a 1 ml Histrap affinity column on the Äkta apparatus. Before applying the protein to the 30 000 molecular weight control Vivaspin 2 centrifugal ultrafiltration apparatus, the membrane was washed with 6 ml ultrapure Elga water (discarded) and 4 ml of 100 mM ammonium acetate at the corresponding pH. The filtrate (1) was used as chemical background control. 2 ml of purified GAPOR was applied to the filter. The filtrate (2) was kept to control if the cofactor was already realised at this stage. The protein was washed with 6 ml of ammonium acetate at the correct pH to extract the cofactor in the filtrate (3). The protein retentate was also recovered (4). The cofactor extraction was performed at pH 3, and control was run simultaneously at pH 8. The collected samples were given to Dr Dave Tooth for LC-MS analysis. All the solution was prepared with ultrapure Elga water and chemical MS-grade.

The above method must be coupled with a very accurate GAPOR titration method to determine which type of molybdenum cofactor was inserted in the enzyme (Chapter IV:6.1.1). Also, after purification, on SDS-PAGE and western blot, unspecific bands were observed along the GAPOR band and assumed as GAPOR degradation. Thus GAPOR quantitation was performed using MS, and three GAPOR-specific standard peptides from the N-terminal end, middle and C-terminal end of the GAPOR (G. Zhang et al., 2010). Three peptides would help estimate the proportion of intact GAPOR after purification. Unfortunately, this method could not be satisfactorily completed before writing this thesis.

6.3. Mass spectrometry analysis

6.3.1. Cofactor analysis

As explained above, the GAPOR was produced in 600 ml culture under micro-aerobic conditions. The culture was inoculated under aerobic condition and then sealed. The *gorS2* expression was then induced under micro aerobic conditions after about 8h of culture at 37°C and was further incubated for 16h at room temperature. Then, the cells were harvested, lysed and the protein purified under aerobic conditions. The cofactor was then putatively extracted from the enzyme by several washes with ammonium acetate pH 3 on an ultrafiltration membrane. The same experiment was conducted with ammonium acetate pH 8 as a control. The two pH were used to allow a comparison between the two extracts. In fact, by comparing the two sets of LC-MS data, it would be possible to spot the differences and control if these signals could correspond to molybdopterin or a degradation of it. No intact or degraded GAPOR was observed on SDS-PAGE (Figure IV-12), even after pH neutralisation. The hypothesis was a denaturation of the protein and precipitation on the filter. It was assumed that the cofactor was released from the protein before precipitation. This hypothesis was backed by the absence of brown precipitate on the membrane, suggesting that the iron-sulfur cluster was potentially liberated from the enzyme. Native-PAGE was tried to control the denaturation of the protein and estimate the part of denatured enzymes in pH 8, but no consistent results were obtained.

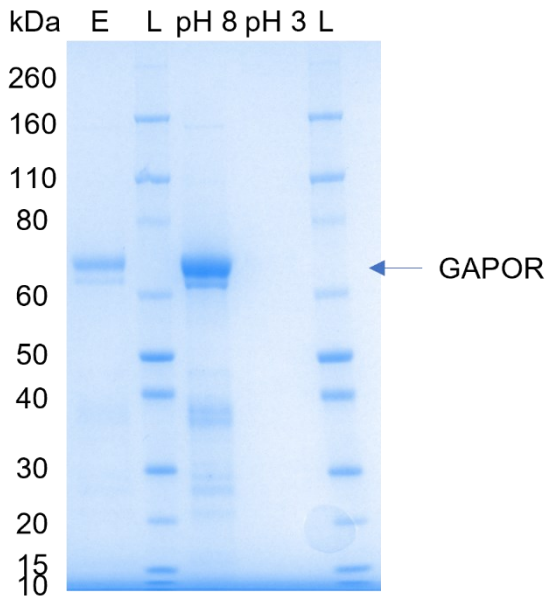


Figure IV-12 Blue-silver Coomassie-stained SDS-PAGE of purified GAPOR after cofactor extraction concentration. E: GAPOR elution from *E. coli* RG_AM023gapor affinity chromatography. pH8: control of cofactor extraction and protein concentration. 2 ml of elution solution was introduced in ultrafiltration apparatus (Vivaspin2 30000 MWCO) and washed with 6 ml ammonium acetate pH 8, final volume ca. 500 µl. pH3: cofactor extraction and protein concentration at pH3, same protocol as pH 8. L: Novex sharp pre-stained standard (Invitrogen).

By extracting the cofactor by acidic treatment at room temperature in the presence of oxygen the most suspected derivative of the cofactor is Form B (Hageman & Rajagopalan, 1986; J. L. Johnson et al., 1984; Leimkühler et al., 2011). As the sample was not heat treated, it reduced the chance of cleavage of the GMP off the air oxidised cofactor. Thus some Form B-GMP might be observed as well. Other potential degradation of the cofactor might be observed. The degradations that were looked for are summarised in the table below.

Table IV-3 List of the compounds which $[M+1H]^+$ were searched in MS spectra.

Name	description	formula	Probability to observe it if inserted	reference
Form B	Air oxidised molybdopterin	$C_{10}H_8N_5O_6PS$	High	(J. L. Johnson et al., 1984)
Form B-GMP	Air oxidised MGD	$C_{20}H_{20}N_{10}O_{13}P_2S$	High	(Leimkühler et al., 2011)
Form B-CMP	Air oxidised MCD	$C_{19}H_{20}N_8O_{13}P_2S$	High	(Leimkühler et al., 2011)
MPT	Un-oxidised molybdopterin	$C_{10}H_{14}N_5O_6PS_2$	Low	
	Pterin-6-carboxylic acid	$C_7H_5N_5O_3$	Low	
H2NPT	dihydroneopterin	$C_9H_{13}N_5O_4$	Very low	(Martín-Tornero et al., 2016)
H2 NPT-P3	dihydroneopterin triphosphate	$C_9H_{16}N_5O_{13}P_3$	Very low	(Martín-Tornero et al., 2016)
P-H4 -Pt	6-pyruvoyltetrahydropterin	$C_9H_{11}N_5O_3$	Very low	(Martín-Tornero et al., 2016)
H4BPt	tetrahydrobiopterin	$C_9H_{15}N_5O_3$	Very low	(Martín-Tornero et al., 2016)
H4BPt-Glu	biopterin glucoside	$C_{15}H_{21}N_5O_8$	Very low	(Martín-Tornero et al., 2016)
H2MPT	dihydromonapterin	$C_9H_{12}N_5O_{13}P_4$	Very low	(Martín-Tornero et al., 2016)
H4MPT	tetrahydromonapterin	$C_9H_{15}N_5O_4$	Very low	(Martín-Tornero et al., 2016)
2'OMet-H4MPT	2'-O methyltetrahydromonapterin		Very low	(Martín-Tornero et al., 2016)
6-HMP	6-hydroxymethyl dihydropterin	$C_7H_9N_5O_2$	Very low	(Martín-Tornero et al., 2016)
H4-CyPt	tetrahydrocyanopterin	$C_7H_8N_6O$	Very low	(Martín-Tornero et al., 2016)

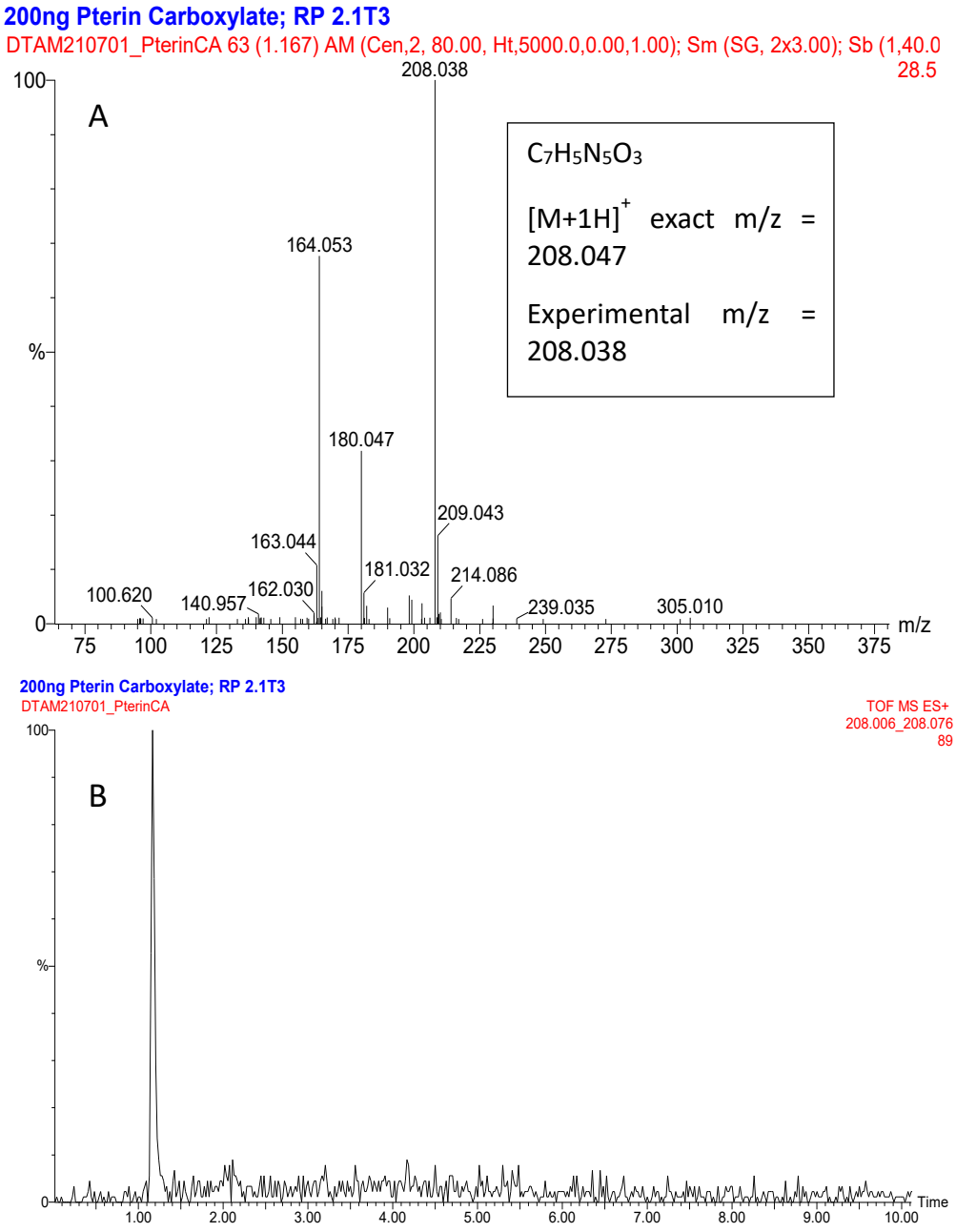


Figure IV-13 LC-MS spectrum and chromatogram of the pterin-6-carboxylic acid. A) Positive ion mass spectrum of 1.1minute eluting components, including pterin-6-carboxylic acid. Note the high experimental mass accuracy observed and high signal:noise with a theoretical on-column load of 100pmol recovered using the employed extraction protocol. B) Extracted ion chromatogram of the 208.04 m/z component. Note the high signal:noise 1.1 minute peak obtained with a theoretical on-column load of 100pmol Pterin-6-carboxylate recovered using the employed extraction protocol

When the method was tested, the pterin-6-carboxylic acid was used as a positive control. A high signal:noise spectra (Figure IV-13 A) and a high peak on extracted chromatograms (Figure IV-13 B) were observed when as little as 200ng (1.0 nmol) was extracted and analysed, with an on-column load of theoretically 20ng (100 pmol). Extracted ion chromatograms for all candidate

target compounds (Table IV-3) were generated, and nothing was detected (data not shown). The same results were observed with xanthin oxidase.

There are three possible explanations for the absence of observed cofactor. The first hypothesis is the absence of incorporation of the cofactor in the strain due to either an issue with the cofactor production or with its incorporation in the protein if a chaperon is involved, like some *E. coli* DMSOR enzymes (Francis Blasco et al., 1998; Iobbi-Nivol & Leimkühler, 2013; A. Magalon & Mendel, 2015). The second explanation is associated with the cofactor extraction itself. A hypothesis is that the GAPOR was precipitated by the pH 3 treatment without losing its cofactor. Regarding this point, the protocol could be modified with some knowledge from the literature: the form-A extraction, might be improved by including a boiling step before the ultra-filtration step (Rajagopalan & Johnson, 1992). A softer method would also be to add some chaotropic agent such as urea or guanidine hydrochloride. Those kinds of modifications could help the cofactor extraction.

The third explanation is that the amount of cofactor is too low to be detected by the analytical equipment. 2 ml of GAPOR eluate at 0.142 ± 0.015 µg/µl was used for each condition corresponding to 2 µmol of GAPOR. Theoretically, if all the enzymes contained an Mo-MGD and the extraction was performed with 100 % yield, 4 µmol of pterin moieties should be recovered. Park and co-worker measured 0.74 mol of Mo per mol of GAPOR. Therefore, 2.96 µmol of pterin would be expected to be extracted in these conditions. As presented above, as low as 100 pmol of extracted cofactor could be detected on the column. Moreover, the protocol (Martín-Tornero et al., 2016) suggested that pterin would be detected if the target components were present over 10 pmol on the column. This would represent 0.1 nmol in solid-phase extraction extract, similar to 10 nM in the filtrate. In conclusion, no pterin was present in the protein extraction solution after ultra-filtration and solid-phase extraction.

Further experiments are required to improve this analysis protocol, mainly the extraction part. Importantly, it would be very interesting to be able to use the original *E. coli* GAPOR strain presented in the Park et al. article as a control (Park

et al., 2007). This control would validate if the lack of cofactor came from the strain used. Another control could be the overexpressed YdhV encoding gene from *E. coli* (Reschke et al., 2019)bearing potentially the same Mo-bisMPT cofactor as *M. maripaludis* GAPOR. Hence, it is difficult to conclude regarding the type of cofactor contained in the enzyme and even more if a cofactor was incorporated. At least a test should have been done to extract the cofactor in "Form A" and try to titrate the pterin within the base strain the *E. coli* RG_AM023gapor. Especially with the improvement of the expression protocol for *E. coli* RG_AM023gapor strain, the production yield after purification was about 710 µg of GAPOR recovered per litre of culture. Running these two experiments would have permitted us to know more about the incorporation of the molybdenum cofactor in the recombinant GAPOR.

7. Discussion and Conclusion

This study aimed to improve the GAPOR activity by increasing the insertion of the molybdenum cofactor. The expression of the native GAPOR cofactor Mo-bisMPT instead of the Mo-MGD (Moco) produced by *E. coli* was supposed to improve its insertion into the target protein.

The *M. maripaludis* molybdenum cofactor biosynthesis pathway was compared to the *E. coli* pathway, and several differences were observed. The absence of MogA isoenzyme in the archaea strain was supposed to be compensated by MoaB, which role is unknown in *E. coli* but can catalyse adenylyl transfer (Loes E. Bevers et al., 2008a). Other archaea also express only MoaB encoding genes. It was hence supposed that MoaB could be an ancestral MogA. Despite hypotheses, the absence of MoaD is still not understood. Further study of the archaeal molybdenum cofactor biosynthesis pathway to investigate the differences with the well-studied *E. coli* pathway. Moreover, four MoeAs were found in *M. maripaludis*. In the literature there are hypothesis about MoeA being involved in the discrimination between Mo and W insertion(Loes E. Bevers et al., 2009). Hence to observe if the GAPOR activity was modified, six genes from the *M. maripaludis* Moco pathway were expressed alongside *gorS2*. Only two proteins, MoeA1 and MoeA2, were observed in the soluble

fraction on the SDS-PAGE gel. Other methods should have been implemented to detect the four other proteins, such as small tags or polyclonal antibodies. Despite those issues, the experiments were performed with the justification that finding the right combination of proteins involved in the molybdenum cofactor biosynthesis would help to improve the GAPOR activity. However, unfortunately, no combination was found to measure a GAPOR and G3P-dependent activity for BV reduction unambiguously. Similarly to the results presented in the previous chapter, the background G3P-dependent activity was present in all the tested strains, including the controls without *gorS2* expressed. Moreover, the results did not show any higher reduction activity from the tested strains' crude extract than the controls. Therefore, it can be concluded that no active GAPOR was produced in our *E. coli* strains despite the complementation with the *M. maripaludis* genes potentially encoding the enzymes for the molybdenum cofactor biosynthesis. The experimental parameters for the crude extract assay have to be modified to reduce the background activity if the same kind of experiment has to be done again.

The GAPOR content was tentatively tested with a new LC-MS method. The lack of an appropriate fluorimeter to perform the literature methods led us to test this new method. However, no cofactor was detected with the new method, and further developments are needed. It would be interesting to validate the method with a known cofactor containing protein such as YdhV from *E. coli* seemingly carrying the same cofactor, as the homologous overexpression of *ydhV* and purification of the protein was already published (Reschke et al., 2019).

In conclusion, it was not possible to produce an active GAPOR and reproduce the results presented by Park and co-workers in their publication (Park et al., 2007). Despite trying to produce Mo-bisMPT cofactor by expressing *M. maripaludis* genes encoding the enzymes supposed to be involved in the cofactor biosynthesis pathway, no improvement of the GAPOR activity was unambiguously demonstrated. Moreover, due to a non-optimised LC-MS

cofactor analysis method, it was not possible to characterise the cofactor composition of the GAPOR produced in the *E. coli* RG_AM023gor.

With the aim of heterologously express an active GAPOR, it was then decided to express *gorS2* in two strains, *C. autoethanogenum* and *C. acetobutylicum,* known to produce active AOR with molybdenum cofactor. It was hypothesised that those organisms could heterologously express *gorS2* to produce an active GAPOR by inserting the correct cofactor. in addition, it was decided to express the *gorSL* genes recently discovered to encode a bacterial GAPOR (Scott et al., 2019).

Chapter V:

Heterologous Expression of Bacterial and Archaeal Genes Encoding GAPOR in *Clostridium autoethanogenum* and *C. acetobutylicum*

1. Introduction

Early in the study, an *in vivo* experiment was planned to palliate the potential impossibility of producing active GAPOR heterologously. In the first instance, *Clostridium autoethanogenum* was chosen. This acetogen is a strict anaerobic mesophile able to grow on sole CO or CO_2 + H_2 isolated from rabbit faeces in 1994 (Abrini et al., 1994). It is closely related to *C. ljungdahlii* (Humphreys et al., 2015). This *Clostridii* expressed an aldehyde: ferredoxin oxidoreductase involved in the first specific step for ethanol production, *i.e*. the acetate to acetaldehyde conversion using reduced ferredoxin as an electron donor (Brown et al., 2014; Liew et al., 2017; Liew et al., 2016). The *C. autoethanogenum* AOR has not been purified and characterised, but its activity and role in the metabolism were investigated (Liew et al., 2017). Thus, it was supposed that, because another AOR was expressed, it might help to express an active GAPOR, as the cofactor was supposed to be identical all the other AOR. Moreover, the *in vivo* experiment was based on the hypothesis that GAPOR could replace the RNF ferredoxin oxidation in *C. autoethanogenum* or *C. ljungdahlii* RNF^- strain and allow the growth on sole CO (Tremblay, Zhang, Dar Shabir, Leang, & Lovley Derek, 2012; Zhu et al., 2020). *C. ljungdahlii* is a strict bacterium use to produce chemical at an industrial scale (Köpke et al., 2010a; Zhang, Zhao, Jia, Jiang, & Gu, 2020), notably by Lanzatech. The possibility to produce chemical from CO without diverting electron to the RNF system for H_2 accumulation. In addition, another well-studied Clostridium expressing an AOR is *Clostridium acetobutylicum*. The role of the protein encoded by CAC2018 is still unknown, but was supposed to be involved in the acid production from aldehydes. The protein was found to have 30 % homology with *P. furiosus* AOR (Jang et al., 2012). Hence, it was also decided to try to express the genes in *C. acetobutylicum* as the plasmids constructed for the expression in *C. autoethanogenum* could also be used in *C. acetobutylicum.* Moreover, *C. acetobutylicum* is a model organism biological hydrogen production and other solvent (Yoo, Nguyen, & Soucaille, 2020). Thus, starting

to investigate a novel manner to increase the hydrogen and other metabolites might be promising.

During the thesis, the first active bacterial GAPOR was reported in 2019 by Scott et al. and was designated as GOR (Chapter I.5.2.4). This enzyme was discovered in *Caldicellulosiruptor bescii*, an extremely thermophilic bacteria (optimum growth at 78°C) which can ferment high concentration of untreated plant biomass (Scott et al., 2019). The enzyme catalyses the same reaction as GAPOR with k_{cat}/k_m 4.7 $\times 10^4$ $s^{-1}M^{-1}$ at 70°C, in comparison *P. furiosus* GAPOR has an apparent V_{max} and k_m of 350 U/mg and 30 µM respectively at 70°C (Roy et al., 2001). *C. bescii* enzyme is composed of 2 subunits a large GOR-L of 65 kDa and a one GOR-S of 15 kDa. This enzyme is one of the few heteromeric AOR, such as the recently discovered YdhV (Reschke et al., 2019). The large subunit was found to harbour a tungstopterin and a [4Fe-4S] cluster, the small subunit is a ferredoxin binding subunit with four [4Fe-4S] cluster binding domains (Figure V-1). Interestingly, they purified the enzyme by size exclusion chromatography in two different forms, the tetramer of the heterodimer and the heterodimer solely (Scott et al., 2019). The two forms displayed the same specific activity. Our study was already focused on the expression of a gene encoding a mesophilic homologue of *P. furiosus* GAPOR. It was then decided to express GOR-SL encoding genes in both Clostridial strains as well.

The objective of this chapter was to study the expression of GAPOR and GOR encoding genes in both *C. autoethanogenum* and *C. acetobutylicum.* Furthermore, the G3P-dependent activity of the two heterologously produced enzymes will be tested.

Herein are presented our choices for the source of mesophilic GOR and the construction of plasmids and strain for the gene expression. The results of our production study of GAPOR in the Clostridial strains and GOR in *E. coli* and both Clostridial strains will also be presented as well as the potential impact of the protein expression on the metabolism of the cells. The enzymes assays results will then be presented.

2. Bioinformatics comparison of GOR-SL

C. bescii GOR, as presented in Chapter I.5.2.4, is tungsten dependent protein with a cofactor identical to GAPOR or AOR from *P. furiosus* or *M. maripaludis* (Scott et al., 2019) (Mukund & Adams, 1995). Scott et al. showed that most of the GOR-containing microorganisms were thermophilic bacteria, but some mesophilic bacteria were also found. Following our directive line to use a protein homolog from a mesophilic bacterium, *C. bescii* GOR-L and GOR-S were blasted. Among the mesophilic bacteria, *Geosporobacter ferrireducens* GOR was chosen. *G. ferrireducens* was isolated in oil-contaminated soil and had an optimal growth temperature between 25°C and 45°C. It is a strict anaerobe alkaliphilic and iron-reducing bacteria (Hong et al., 2015). *G. ferrireducens* GOR-L (AOT72683) has 54 % identities, 72 % homologies, and GOR-S (AOT72682) has 49 % identities and 68% homologies with *C. bescii* GOR. GOR heterodimer is coded by two collocated genes a putative 429 nt (coding for the (Fe-S) cluster domains): *gor-S* (Gferi_25880) and a putative 1,791 nt (coding for the aldehyde: ferredoxin oxidoreductase) *gor-L* (Gferi_25885) with *gor-L* RBS and start codon embedded in *gor-S* (Figure V-1 C). *gor-S* is coding for a 142 amino acids protein GOR-S with a theoretical molecular weight of 15.5 kDa. The protein, like *C. beskii* small subunit, contains four [4Fe-4S] cluster domains (Figure V-1 B). The large subunit GOR-L is a 596 amino acid protein with the same binding domain for a pterin and a [4Fe-4S] cluster domain with a theoretical molecular weight of 66.5 kDa (Figure V-1 A).

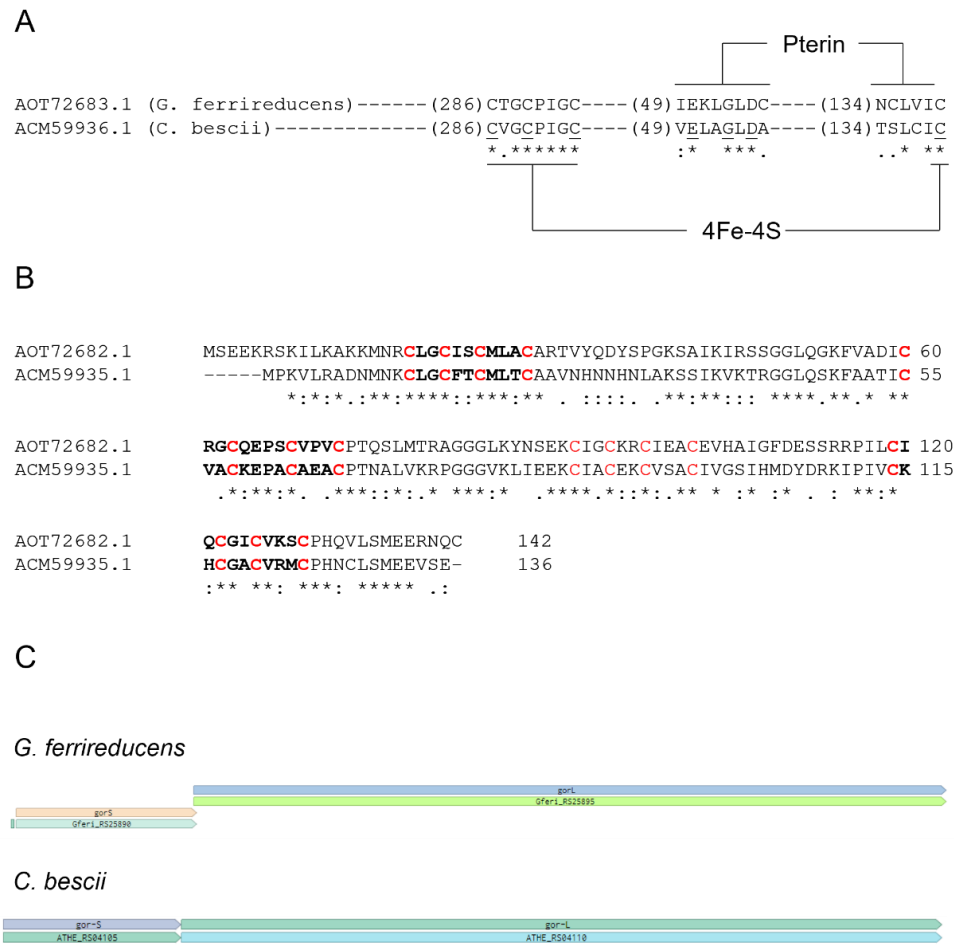


Figure V-1 Comparison of *G. ferrireducens* GOR (*gor-S*: AOT72682.1, *gor*-L :AOT72653.1) and *C. bescii* GOR (*gor*-S: ACM59935.1, *gorL*: ACM59936.1). A) Identification of the conserved binding motif in *G. ferrireducens* for GOR-L pterin and [4Fe-4S], B) [4Fe-4S] cysteine binding motif in GOR-S in bold and red. Alignment made with Clustal Omega. C) Schematic representation of both GOR operon from both strains, and the start codons of the *gorL* is embedded in *gorS*. The figure was made using the Benchling visualisation tool.

3. GOR-SL expression in *E. coli*

In their study, Scott and co-workers expressed in *C. bescii gorS* and a *gorL* gene encoding an N-terminal tagged (histag) GOR-L for affinity purification. Their enzyme was still active after purification in *C. bescii* (Scott et al., 2019). Thus, it was decided to reproduce their construction. However, as explained above (Figure V-2), the first 7 nucleotides of the large subunit are shared with the *gor-S* sequence. Hence the his-tagged coding sequence was placed between L2 and N3. We used the same his-tag-thrombin cleavage site sequence used previously for *gorS2*. Moreover, to extend the GOR-L sequence the least possible, the Ochre stop codon of *gor-S* was changed to Ambre to accommodate the

insertion of the glycine codon Figure V-2. The complete sequence of the *gor-SL* operon was synthesised externally (Geneart, Thermofisher).

Figure V-2 Histag sequence insertion in 5' of *gorL*. The top figure shows the native position of the 5'-end of *gor*-L embedded within the 3'-end of *gor-S*. The bottom figure presents how the six-histidine-tag and the thrombin cleavage sequence were cloned between the native Ochre start codon and the new Ambre start codon. The figures were made with the Benchling visualisation tool.

Following the initial work with GAPOR, the *gorSL* genes were first expressed in *E. coli* RG_AM023. In addition to the fact this strain is now well known to us, the expression in the mutant *E. coli* Rosetta-gami2(DE3) avoids potential translational issues with rare codons as the strain contains a pRARE2 plasmid expressing *E. coli* rare tRNA. Furthermore, it would allow a comparison with GAPOR during crude extract enzyme assay. It is important to note that the strain was constructed while the mini-library "crude extract enzyme assay", presented in the previous chapter, was underway. Then, it was decided to clone the synthesised genes in pET28a(+) gorS2 (AmpR) in place of *gorS2*. The genes are expressed under the control of the lac operator and T7 promoter, and the RBS of *gorS2* was used for *gorS*. The molecular weight of GOR-L with the addition of the tag was calculated at 68.7 kDa. The plasmid was constructed using a Geneart assembly kit (Invitrogen), and the plasmid was validated by sequencing. Then the plasmid was transformed in *E. coli* RG_AM023 to construct *E. coli* RG_AM020.

gorSL expression was performed following the *standards gorS2* expression protocol. As only GOR-L was tagged only the large subunit would be revealed by Western blot. It can be observed in Figure V-3 that no soluble GOR-L was

recovered in the insoluble fraction. Moreover, no band was observed at 15 kDa even though GOR-S had a theoretical initiation translation rate 30-fold higher than GOR-L.

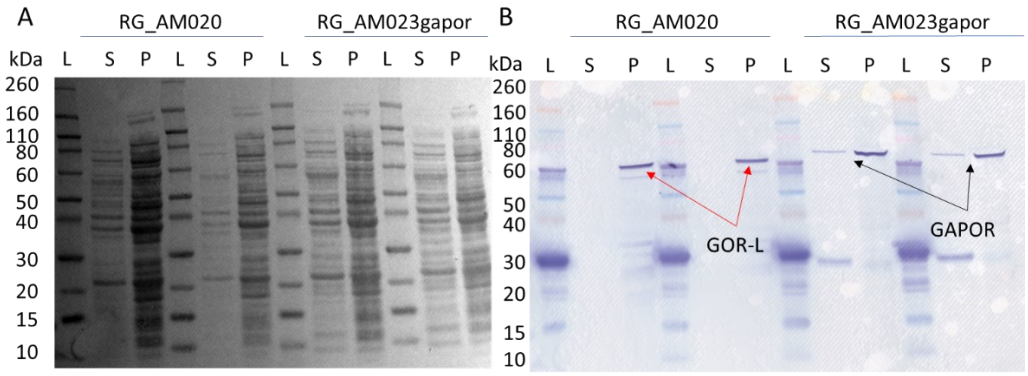


Figure V-3 GOR-SL production in *E. coli* RG_AM020. A) Ponceau red-stained WB membrane B) WB revealed with TMB for membrane. *gor*-SL was expressed in *E. coli* RG_AM020. (*E. coli* Rosetta-gami2(DE3) Δ*iscRhypFselA* pET28a(+) *gorSL*(AmpR)) E. coli RG_AM023gapor (*E. coli* Rosetta-gami2(DE3) Δ*iscRhypFselA* pET28a(+) *gorS2* (AmpR)) was used as a control. S: supernatant, P: pellet, L: Novex Sharp pre-strain protein standards. GAPOR was visible at 70 kDa and GOR-L at 68 kDa. The protein production cultures were performed under anaerobic conditions in 50 ml MAC media complemented with 100 µM molybdate in 120 ml sealed serum bottles. The protein production was induced after 8 h with 100 µM IPTG and incubated at room temperature overnight. The cells were routinely harvested and lysed.

Hence, no soluble GOR was produced under the tested condition in the *E. coli* RG_AM020 strain despite GOR-L being produced in an insoluble form.

Due to those negative preliminary results, *gorSL* expression was not pursued in *E. coli,* and it was chosen to focus on the expression in the two Clostridia.

4. *gorS2* and *gorSL* expression in C*lostridia*

4.1. AOR in C. acetobutylicum and C. acetobutylicum

As presented above, both Clostridial strains produced aldehyde: ferredoxin oxidoreductase. *C. autoethanogenum* possesses two isoforms of the enzyme coded by *aor1* (CLAU_0081) and *aor2* (CLAU_0099) (Humphreys et al., 2015). The two genes are not expressed at the same level, but the expression level is increased under autotrophic growth conditions, 4-fold and 5.3-fold higher under CO_2+ H_2 than CO and fructose fermentation, respectively (Liew et al., 2017; Marcellin et al., 2016; Mock et al., 2015). AORs play an important role in ethanol production by reducing acetic acid to acetaldehyde using reduced ferredoxin as an electron donor (Köpke et al., 2010b; Mock et al., 2015). Also,

it was found that this reaction could be coupled with CO oxidation by CODH, producing the reduced ferredoxin needed by AOR (Marcellin et al., 2016). It was found that the two isoenzymes had different roles in *C. autoethanogenum* metabolism, and the deletion of *aor2* increased ethanol production under CO and fructose fermentation (Liew et al., 2017). *C. acetobutylicum* also possesses a gene coding for an AOR (CAC2018). However the enzyme has yet to be characterised (Jang et al.).

It was supposed because of their AOR expression, the heterologous GAPOR and GOR soluble expression might be facilitated. A potential chaperon or unknown protein folding mechanism may be present in these bacteria, helping the heterologous enzymes correct conformation. Noteworthily, no codon optimisation of the DNA sequence was performed.

4.2. Enzymes expression in C. autoethanogenum and C. acetobutylicum

4.2.1. Choice of promoters and plasmids

C. autoethanogenum is a bacterium well-studied in the laboratory. The plasmid transfer in *C. autoethanogenum* is performed by conjugation, using *E. coli* sExpress (Woods et al., 2019) as a shuttle host (Minton et al., 2016). Due to this, the plasmids needed to contain both a Gram+ and a Gram- replicon. In the laboratory, the primary vectors used in *Clostridia* are the modular plasmids pMTL80000 vector (Heap et al., 2009). This modular vector combination was pCB102 as Gram+ replicon, *catP* as a selection marker, ColE1 as Gram- replicon and a multiple cloning site (MCS) for gene cloning. At first, *ErmB* was used as an antibiotic resistance marker. Nevertheless, during the first test of *gorS2* expression, it was observed that Clarithromycin (a more stable analogue of erythromycin) was not inhibitory or stable enough and quickly led to plasmid loss. So, the selection marker was replaced by *catP* and thiamphenicol was used as a selective pressure in *C. autoethanogenum.*

Constitutive and inducible promoters were tested to control the *gorS2* transcription (Table V-1). One of the issues that arose in the *gorS2* expression

in *E. coli* was the low expression level. Hence two relatively strong constitutive promoters were tested. *C. acetobutylicum hydA* promoter and RBS (Girbal et al., 2003; Girbal et al., 2005) were cloned in front of *gorS2* for the first expression test. The second tested promoter, P_{thl} (for thiolase promoter), was shown to be 14-fold weaker than P_{hydA} (Girbal et al., 2003). However, the expression level, when tested in *C. autoethanogenum* by other laboratory members (unpublished data), has shown a strong constitutive expression. Thus, the *C. acetobutylicum* thiolase promoter and RBS were also tested (Girbal et al., 2003; Heap et al., 2009).

In addition, two inducible promoters were tested (Table V-1): P_{tcdB} and P_{fdx}_RB3. The first is a two parts inducible system from the toxin B system of *Clostridium difficile* (Cartman Stephen & Minton Nigel, 2010; Dupuy & Matamouros, 2006). This system used a lactose-inducible *tcdR*$^{+}$ mutant strain to express genes controlled by P_{tcdB} on the plasmid. TcdR is toxin specific sigma factor specific to P_{tcdB} (Dupuy & Matamouros, 2006). In this system, the expression was induced by IPTG/lactose to express TcdR, which induced the expression of *gorS2* under the P_{tcdB} control (Minton et al., 2016). Hence the plasmids had to be conjugated in *C. autoethanogenum* C24, the *tcdR*$^{+}$ strain (Woods et al., 2022). The same plasmids were also transformed in a *tcdR*$^{+}$ mutant strain of *C. acetobutylicum*. The second inducible system was based on a *Clostridium sporogenes* ferredoxin promoter coupled with a riboswitch designed to control the translation of *cas9* for CRISPR/Cas9 genome edition (Cañadas et al., 2019). The protein translation was induced by the addition of theophylline, relaxing the stem-loop structure on the RBS and allowing the binding of the ribosome to the Shine-Delgado sequence (Cañadas et al., 2019). These two inducible systems were also formerly tested in the *C. autoethanogenum* (unpublished data) and showed that the expression level due to the leakage of the P_{tcdB} promoter was equivalent to the induced level of the riboswitch. The latter designed, to control *cas9* translation, is very tightly controlled.

Table V-1 Plasmids for the expression of *gorS2* and *gorSL* in *C. autoethanogenum*

plasmid name	promoter	RBS	studied gene	C. autoethanogenum strain	strain name	C. acetobutylicum strain	Strains name
pMTL83151_AM001	P_{hydA}	*hydA*[1]	*gorS2*	*C. autoethanogenum* DSM10061	*CA_AM001*	*C. acetobutylicum ATCC824*	*CAC_AM001*
pMTL83151_AM002	P_{thl}	*thl*[1]	*gorS2*	*C. autoethanogenum* DSM10061	*CA_AM002*	*C. acetobutylicum ATCC824*	-
pMTL83151_AM003	P_{fdx}	Riboswitch synthetic[2]	*gorS2*	*C. autoethanogenum* DSM10061	*CA_AM003*	*C. acetobutylicum ATCC824*	*CAC_AM003*
pMTL83151_AM004	P_{tcdB}	*tcdB*[3]	*gorS2*	*C. autoethanogenum* C24[5]	*CA_AM004*	*C. acetobutylicum tcdR*$^{+}$	*CAC_AM004*
pMTL83151_AM005	P_{hydA}	*hydA*	*gorSL*	*C. autoethanogenum* DSM10061	*CA_AM005*	*C. acetobutylicum ATCC824*	*CAC_AM005*
pMTL83151_AM005.1	P_{hydA}	*gor-S*[4]	*gorSL*	*C. autoethanogenum* DSM10061	*CA_AM005.1*	*C. acetobutylicum ATCC824*	*CAC_AM005.1*
pMTL83151_AM006	P_{fdx}	Riboswitch synthetic	*gorSL*	*C. autoethanogenum* DSM10061	*CA_AM006*	*C. acetobutylicum ATCC824*	*CAC_AM006*
pMTL83151_AM007	P_{tcdB}	*tcdB*	*gorSL*	*C. autoethanogenum* C24	*CA_AM007*	*C. acetobutylicum tcdR*$^{+}$	*CAC_AM007*

[1] from *C. acetobutylicum* ATCC 824 (Girbal et al., 2003)
[2] synthetic RBS F from (Cañadas et al., 2019)
[3] from *C. difficile* (Dupuy & Matamouros, 2006)
[4] from *G. ferrireducens*
[5] (Woods et al., 2022)

It is important to note that the *tcdR*$^{+}$ *C. acetobutylicum* mutant strain is deleted from all the restriction enzyme encoding genes. Hence the pMTL83151_AM004 and pMTL83151_AM007 plasmids had to be methylated before electroporation by *E. coli* harbouring pAN2 (Heap et al., 2007).

P_{hydA}, P_{fdx}_RB3 and P_{tcdB} were also used to control the expression of *gorSL* in *C. autoethanogenum*. The plasmids were constructed using HiFi assembly (NEB) or GeneArt assembly kits (Invitrogen). All the steps for the plasmid construction were followed by DNA sequencing for validation. After the transfer of the plasmid into the host strain, two plasmids displayed mutations and were not tested for protein expression. The mutations were observed only in *C. autoethanogenum* for pMTL83151_AM002 and pMTL83151_AM006. For the former, mutations were observed in the sequence after cloning and transformation in the *E. coli* sExpress. The thiolase promoter was known in the laboratory to be functional in both Gram+ and Gram- bacteria and with a high level of expression in *E. coli*. Hence, one of the mutated plasmids was chosen with the insertion of a G in the -35 sequence (ttGgata) of the promoter, supposing that it would decrease enough the strength of the promoter to allow the conjugation with *C. autoethanogenum*. Even with the mutation in the promoter, after conjugation and plasmid rescue, sequencing revealed that a one-base deletion introduced a frameshift in *gorS2* sequence. Hence, the *gorS2* expression in *E. coli* and *C. autoethanogenum* could be toxic for the bacteria. It was unknown if the toxicity was due to the GAPOR activity or the overexpression. It was then decided not to use it to express *gorSL*. pMTL83151_AM006 was transferred into both strains. Despite no problem with the electroporation of the plasmid in *C. acetobutylicum*, the conjugated plasmid in *C. autoethanogenum* displayed a 1kb insertion of *E. coli* genomic DNA within *gor-L*. The insertion was supposed to occur during either the *E. coli* TOP10 or sExpress incubation. The *C. acetobutylicum* strain hosting pMTL83151_AM006 was tested for the *gorS2* expression, but the results will not be presented here due to some doubt about the strain validation. Finally, pMTL83151_AM005.1 conjugated *C. autoethanogenum* strain was not tested

for *gorSL* expression either. In this plasmid, the *gor-S* translation was placed under the control of the native *gor-S* RBS instead of *C. acetobutylicum hydA*. When the plasmid was checked after the conjugation, the complete *gor-S* sequence was deleted. The RBS functioning in *C. autoethanogenum* was supposed to be too strong, leading to the gene's deletion. Here as well, it is not known if the deletion was due to a too-strong expression or due to the potential toxicity of the gene

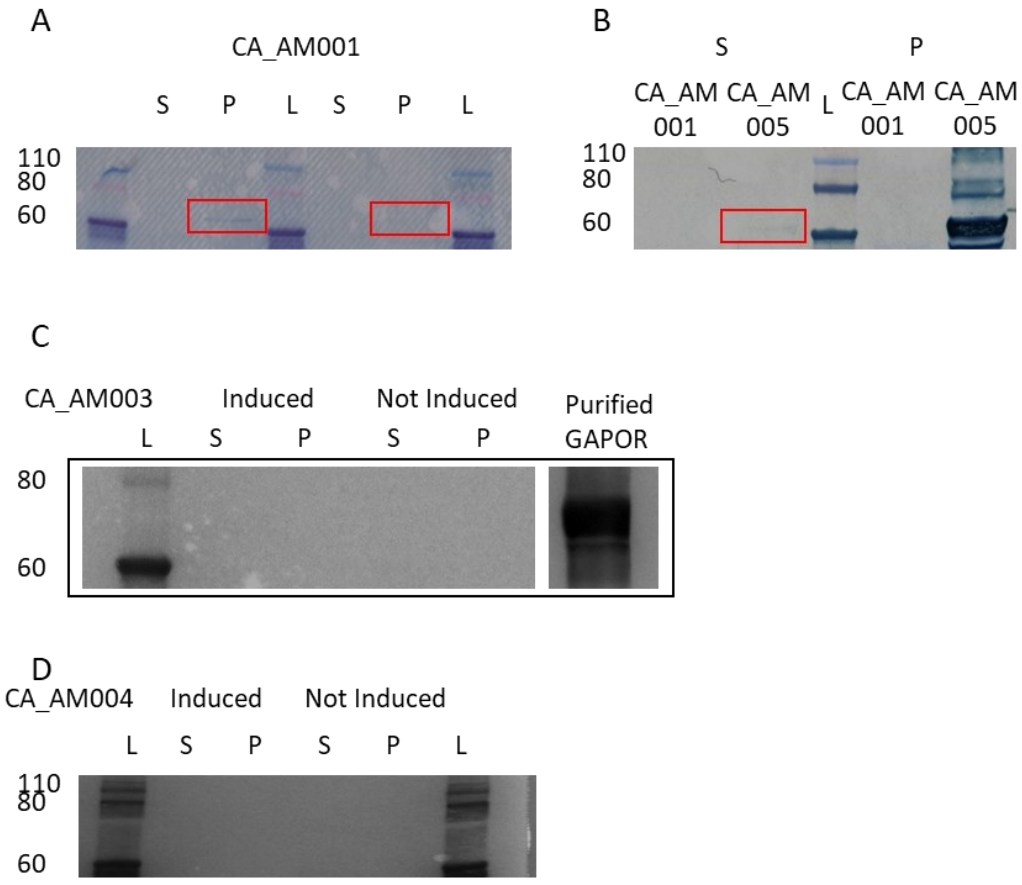
4.2.1. Protein production in Clostridia

Initially, these expression experiments aimed to express *gorS* and *gorSL* in *C. autoethanogenum* to later try an *in vivo* experiment in *C. ljungdahlii* RNF$^-$ strain. The hypothesis was that GAPOR/GOR would, if functional, in the opposite of the known reaction, like *Clostridial* AOR, allow growth of the deficient strain on CO by producing oxidised ferredoxin for the CODH (Chapter I.5.1.1).

The first expression assays were performed with CA_AM001 and CA_AM005 (Figure V-4 A&B). No soluble GAPOR was detected on the western blot and only fainted band were observed in the insoluble fraction. A faint signal of soluble GOR-L was seen in the cell-free supernatant, but an important signal was noted in the insoluble fraction. However, many other bands were revealed by the HRP conjugated antibodies (not shown in the figure). Those bands were supposed to be either degraded GOR-L or aggregates without or with the tag on the protein as none of those bands were visible for the CA_AM001 control pellet revealed by western blot. GOR was shown by Scott et al. to form tetra-heterodimers (Scott et al., 2019), hence as hypothesised before for the *E. coli* experiments, it might be possible that GOR-L could be more prone to form aggregates. The results obtained with *gorL* expression showed that the *hydA* promoter was functional in *C. autoethanogenum*,.However, the RBS used to control the translation of *gapor* and *gor-S* was weak in *C. autoethanogenum* or not inserted with the right spacing upstream of the genes. This seems to correlate with the deletion of the promoter and *gorS* in pMTL83151_AM005.1 presented in the previous section. Hence, it could be supposed that the level

of expression of *gorL* was acceptable for the cells, especially because most of it was insoluble. However, the expression of the *gorS* either because of the activity or the level of expression of the small subunit was toxic for the cells leading to the deletion.

C. autoethanogenum



C. acetobutylicum

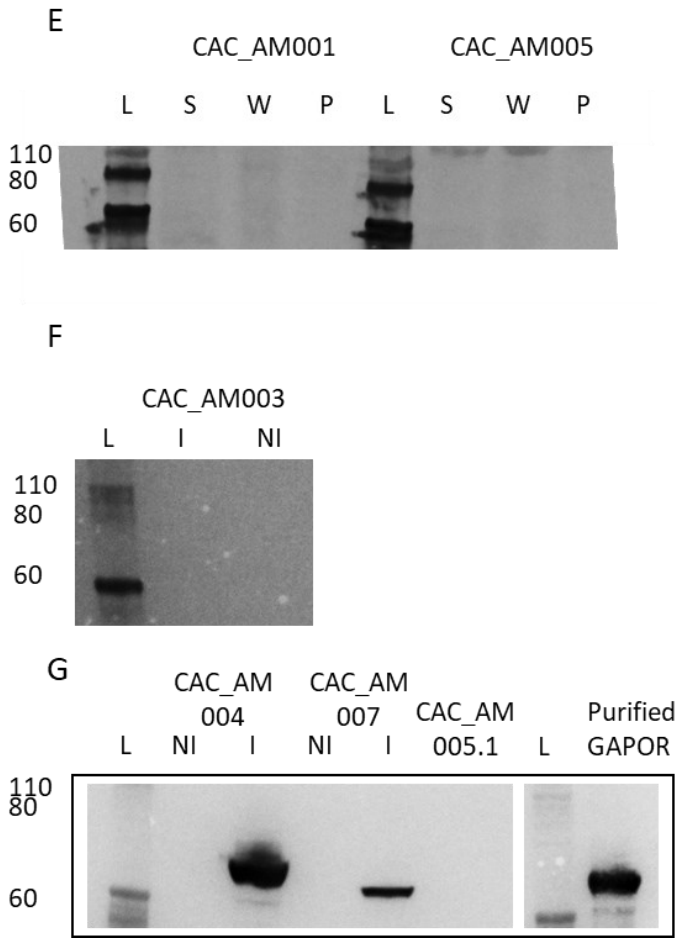


Figure V-4 *gorS2* and *gorSL* in *C. autoethanogenum* and *C. acetobutylicum*. Western Blot anti-histag revealed with TMB for membrane. A) GAPOR under the control of P_{hydA} from CA_AM001, S: supernatant, P: pellet, fainted band of insoluble GAPOR are indicated by red rectangle, fake colour was applied; B) *gorS2* and *gorSL* under the control of P_{hydA} from CA_AM001 and CA_AM005, S: supernatant, P: pellet, fainted bands of soluble GOR-L is marked by a red rectangle ; C) *gorS2* under the control of P_{fdx}_RB3 (riboswitch) from CA_AM003, purified GAPOR from E. coli RG_AM025gapor was used as a positive control, induction with 5 mM theophylline; D) *gorS2* under the control of P_{tcdB} in lactose inducible *tcdR*$^{+}$ *C. autoethanogenum* C24 from CA_AM004, S: supernatant, P: pellet, induction with 100 µM IPTG; E) *gorS2* and *gor* expression under the control of P_{hydA} from CAC_AM001 and CAC_AM005, S: supernatant, W: supernatant after the wash of the pellet with lysis buffer, P: pellet; F) *gorS2* expression under the control of Pfdx_RB3 from CAC_AM003, I: induced with 5 mM theophylline, NI: no induction; G) *gorS2* and *gorSL* expression under the control of P_{tcdB} in *C. acetobutylicum* lactose inducible *tcdR* CAC_AM004 and CAC_AM007 and *gorSL* under the control of P_{hydA} with *gor*S native RBS. Purified GAPOR form E. coli RG_AM026gapor was used as positive control, I: induction with 50 mM lactose, NI: no induction. L: Novex sharp pre-stained protein standard, on display on the pictures only 60 kDa, 80 kDa and 110 kDa. The contrast on the image was increased to allow a more straightforward observation of the faint band. *C. autoethanogenum* cultures were performed in 60 ml YTF media in 120 ml sealed serum bottles under strict anaerobic conditions at 37°C. The media was complemented with 100 µM molybdate. The cells were incubated overnight after induction if induction was required. *C. acetobutylicum* cultures were performed in 60 ml CGM media complemented with 100 µM molybdate in sealed serum bottles under strict anaerobic conditions. The cells were incubated for 7 to 8 h after induction at *ca.* 0.1 OD_{600nm}.

In *C. acetobutylicum,* no *gorS2* or *gorSL* expression was observed neither in the supernatant, pellet wash, nor pellets. The promoter and RBS sequence were cloned from *C. acetobutylicum* (Girbal et al., 2003); hence it was assumed that the promoter and RBS would be functional. Therefore, it could be supposed that the RBS was too close or too far away from the start codon leading to a poor translation of *gorS2* or *gorS*. Moreover, the *gorL* RBS did not seem to be functional in *C. acetobutylicum* despite being functional in *C. autoethanogenum.* The absence of soluble or insoluble protein could come from culture or harvesting conditions.

The two inducible promoters tested in *C. autoethanogenum* did not permit expression or at least enough expression of *gorS2* for the protein to be visible on Western blot. Both promoters and RBSs were shown to work in this strain by other laboratory members. The absence of soluble or insoluble protein could be explained by how the gene was cloned with the RBS and promoter. Another hypothesis to explain the absence of expression could be the methods of culture induction that did not suit the protein expression. The induction was inspired by the methods used for *E. coli*, but the induction time might needed to be shorter, leading to protein degradation. The same results were observed for *gorS2* expression with P_{fdx}_RB3 in *C. acetobutylicum.* It might be possible

that the P*fdx* riboswitch was improperly constructed to allow gene expression from this plasmid.

However, in *C. acetobutylicum,* P_{tcdB} controlled expression of *gorS2* and *gorSL* led soluble production of GAPOR and GOR-L after induction with 50 mM lactose (Figure V-5 G). GAPOR and GOR-L were also produced as inclusion bodies. At a molecular weight corresponding to GOR-S (15 kDa), an increased signal was observed between induced and not induced conditions, suggesting that the complex between GOR-S and GOR-L is probably formed. It might explain why Gor-L was more soluble than in *C. autoethanogenum*.

GAPOR was shown to be poorly produced in *C. autoethanogenum,* while GOR-L was mainly produced in an insoluble form. It was then concluded that P_{hydA} was functional, but the *hydA* RBS was either driving a weak expression or a problem with the respective position of the RBS and the start codon. However, *gor*L RBS was functional. In *C. autoethanogenum,* none of the tested inducible promoters led to any protein production. Our principal hypothesis regarding the absence of expression would be a problem with the induction protocol. It would be interesting to develop a better protocol for heterologous protein expression in *C. autoethanogenum*. The impossibility at the time of testing *gorS2* expression with the inducible promoters in *C. autoethanogenum* did not help to understand where our problem of expression was. It was demonstrated that the *gorL* RBS was functional in the strain. Therefore, if GOR-L was detected, this implied that the promoters were working correctly, and the RBS area designed to express *gorS* required revision. P_{hydA} and P_{fdx}_RB3 did not allow either to express genes in *C. acetobutylicum*. *gorL* expressed from P_{tcdB} showed that the native RBS was also functional in this strain Thus, the absence of expression with the *C. acetobutylicum* P*hydA* promoter was probably due to a culture issue. The cells might have been incubated for too long before conducting the cell lysis. Finally, *C. acetobutylicum* CAC_AM004 and CAC_AM007 showed GAPOR, GOR-L and GOR-S production after lactose induction. The presence of soluble GOR-L indicates a potential correctly conformed protein. Hence, it was decided to pursue the study by testing the

purification of both enzymes expressed in these strains. As none of the studied enzymes was expressed in *C. autoethanogenum*, the *in vivo* experiment was abandoned. It was decided to send the strain to Dr Céline Foulquier in Toulouse to perform the assays on purified enzymes. Harvested cells had to be sonicated to extract the soluble protein from *C. acetobutylicum*. In Nottingham, it was impossible to perform the complete lysis, purification, and assay under strictly controlled anaerobic conditions.

4.1. Enzymes purification

C. acetobutylicum possesses three hydrogenases one [NiFe]-Hydrogenase (Nölling et al., 2001) and 2 [FeFe]-hydrogenases (Gorwa, Croux, & Soucaille, 1996; P. W. King, Posewitz, Ghirardi, & Seibert, 2006) which catalyse the reversible reaction $2H^+ + 2e^- = H_2$ using ferredoxin as electron donor or acceptor. In these study methyl viologen, and by consequence, Benzyl viologen, could be used as electron acceptor to follow the hydrogen uptake activity of the enzyme (Demuez, Cournac, Guerrini, Soucaille, & Girbal, 2007; P. W. King et al., 2006). Hence, using another anaerobic chamber in Nottingham to test the GAPOR and GOR activities in crude extract was impossible, as all the microbiology cabinets contain H_2 in their atmosphere. Moreover, in Nottingham, contrary to Toulouse, it was not possible to lyse the cells by sonication under correct anaerobic conditions. Although, crude extract assays help detect activity, it is impossible to characterise an enzyme's kinetic parameters fully. Therefore, it was required to determine the parameter for the purification.

After validation, the purification of GAPOR produced in *C. acetobutylicum* behaved exactly the same way as the purification of GAPOR produced from *E. coli* Figure S 2.

The *gorSL* operon was designed with a his-tag in the N-terminal of the large subunit as it was shown by Scott et al. that this design allowed the purification of the heterodimeric GOR as well as tetraheterodimers (Scott et al., 2019). The GOR was found to be eluted with a higher imidazole concentration (200 mM instead of 125 mM). Analysis of the SDS-PAGE (Figure V-5) showed that along

with GOR-L, three other bands were visible on the gel of concentrated GOR (Figure V-5). The GOR bands were given to the analytical team of the group. However, they were not able to get any results other than the protease (porcine-trypsin) and keratin which are both background contamination. Hence, the expression of a complete GOR could not be validated.

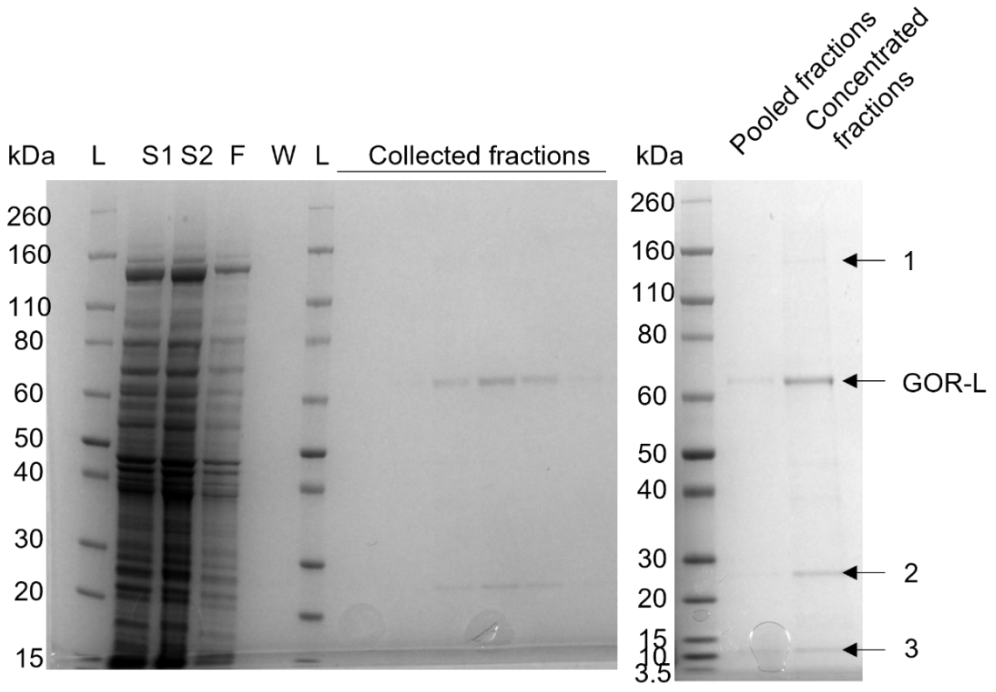


Figure V-5 GOR from *C. acetobutylicum* CAC_AM007 affinity purification SDS-PAGE. Blue-Silver Coomassie strained SDS-PAGE (MOPS buffer) after nickel affinity purification and gradient elution, 0.5 ml fraction were collected during the elution (fractions 6, 8, 10, 12, 14, 16 are presented in the picture). S: Lysis supernatant of two 50 mM lactose induced cultures. F: sample flowthrough. W: column wash. Fractions 8 to 16 were pooled, and 200 µl of the pooled fractions were acetone precipitated. The precipitated protein pellet was resuspended in a 20 µl SDS-PAGE sample mix and was fully loaded on the gel. An arrow marks the GOR-L band. Bands 1, 2 and 3 were assumed to be degradation or aggretes of GOR-L and were sent for sequencing by mass spectrometry in addition to the GOR-L band, However, no results were obtained. Comparison of GAPOR and GOR can be found in Figure S 2

Gene expressions were optimised by reducing the concentration of the inducer to 10 mM to limit the impact on the growth, but no difference in the expression level was observed. After affinity purification, 205 ± 8 µg of GAPOR and 295 ± 41 µg of GOR were recovered in 5 ml from about 120 ml cultures (two 60 ml cultures). After desalting with a PD-10 of half of the elution (2.5 ml), 77 ± 12 µg of GAPOR and 94.5 ± 12 µg of GOR were recovered in 3.5 ml. In comparison, *ca.* 2.7 mg of GAPOR was recovered from 3 L anaerobic *E. coli* culture (Chapter III:5.2). For the same volume, it would be expected to recover *ca.* 5 mg of

GAPOR and 7.3 mg of GOR. Therefore, the conditions (promoters, RBS, and strain) in *C. acetobutylicum* are better for expressing the enzymes, at least regarding the amount of protein.

4.2. Effect of the expression of *gorS2* and *gorSL* on the *C. acetobutylicum* metabolism

The production of active GAPOR and-or GOR in *C. acetobutylicum* during glucose fermentation was hypothesised to increase the pool of reduced ferredoxin and NAD^+ is both enzymes are considered to have similar activity with GAPDH. Reduced ferredoxin could be used to produce hydrogen, NADH and NADPH with the help of a hydrogenase, a ferredoxin-NAD^+ reductase and a ferredoxin-$NADP^+$ reductase, respectively (Yoo et al., 2015). Also, the production level of H_2, NADH and NADPH change according to the fermentation conditions (solventogenic, alcohologenic and acidogenic) (Yoo et al., 2015). Consequently, increasing the pool of reduced ferredoxin could potentially increase the production of H_2 or NAD(P)H. The solvent production pathway requires the consumption of NAD(P)H, and this redox balance has been key to engineer the solvent production pathway in *C. acetobutylicum* (Yoo et al., 2015). Accordingly, increasing reduced ferredoxin production by *gorS2* or *gorSL* expression could increase solvent production. In the case, that GAPOR or GOR expression improve the solvent production, this study could open a different way to produce chemicals from renewable source. Hence, batch cultures of the *gorS2* or *gorSL* expressing strain *C. acetobutylicum* CAC_AM004 and CAC_AM007 were performed to examine this potential.

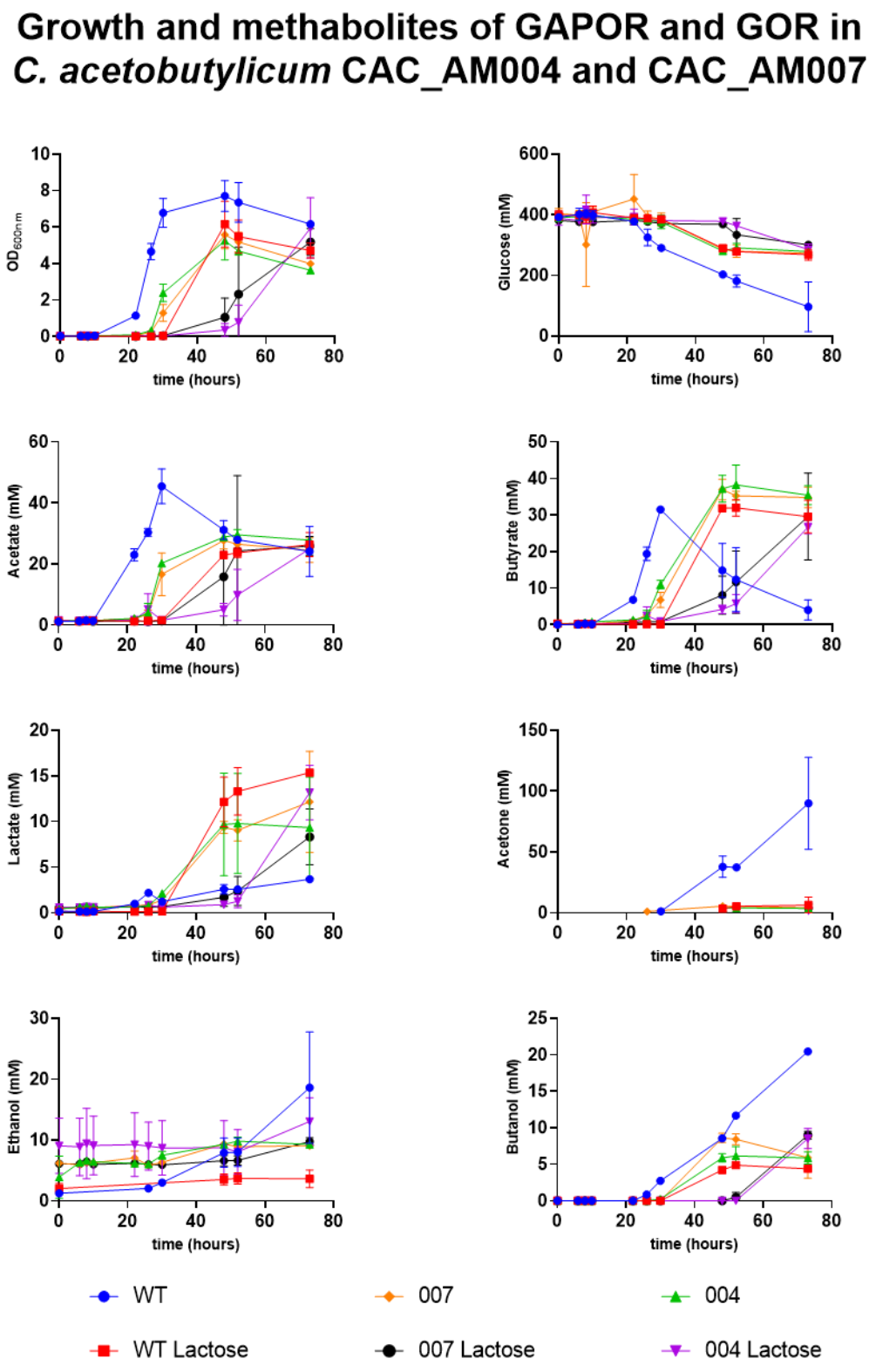


Figure V-6 Growth and metabolites during *gorS2* and *gorSL* expression in *C. acetobutylicum tcdR*$^{+}$ under the control of P_{tcdB} CAC_AM004 and CAC_AM007. The cultures were performed in duplicate in 60 ml CGM in 120 ml sealed serum bottles, and the *gorS2* and *gorSL* expressions were induced with 50 mM D-lactose at t = 0 h. The cultures were incubated for 73 h at 37°C and samples for OD and HPLC analysis at 0 h, 6 h, 8 h, 10 h, 22 h, 26 h 30, 30 h, 48 h, 52 h and 73 h. *gorS2* was expressed using CAC_AM004 (004) and *gorSL* was expressed using CAC_AM007 (007), C. acetobutylicum *tcdR*$^{+}$ strain (WT) was used as control. Cultures without induction was used to control the potential impact

of lactose on growth other than protein expression induction. The samples' supernatants were analysed by HPLC by the group analytical team and the six main metabolites of *C. acetobutylicum* as well as glucose and lactose (not shown), were looked for. The results of each strain were plotted per metabolite, wild-type *C. acetobutylicum tcdR*$^{+}$ (WT) in blue, WT induced with lactose in red. *C. acetobutylicum tcdR*$^{+}$ transformed with *gorSL*-containing plasmid non-induced (orange) and induced (black), *C. acetobutylicum tcdR*$^{+}$ transformed with *gorS2*-containing plasmid non-induced (green) and induced (purple).

The impact of GAPOR and GOR expression on the cells' metabolism was studied during a 73 h-batch fermentation with induction of the protein at t = 0 h (Figure V-6). The first observation is the difference in the lag phase with the lactose-complemented cultures, which was longer than the non-induced cultures. It was observed that the lag phase of the induced culture was 15 h to 20 h longer than the non-induced cultures of the same strain. Moreover, none of the strains reached glucose limitation. Interestingly, the lactose concentration stayed stable in the cultures during the 73 h of the experiment, indicating that there was catabolic repression of the use of this substrate by the cells (Y. Yu, Tangney, Aass Hans, & Mitchell Wilfrid, 2007).

The WT strain cultures start producing acetate and butyrate with an increase of the concentration in the medium following the growth until a sharp fall of the concentration corresponding to the acids reassimilation concomitant with an augmentation of acetone, lactate, ethanol and butanol concentration. In Comparison, no acid reassimilation was observed in the WT culture complemented with lactose. Consequently, low solvent production was detected. Similar behaviour was observed for the two non-induced mutant strains.

Except for the delay, the *gorS2* or *gorSL* expressing strains had similar metabolite production, and differences between induced and non-induced cultures were not detected. However, the addition of lactose in the culture had an impact on the growth. This could be explained by unknown inhibition by a too high lactose concentration (50 mM). However, a simpler hypothesis was the presence of traces of oxygen in the lactose solution despite the stock solution being equilibrated in the anaerobic cabinet at least overnight. The oxygen could have also been brought into the serum bottle by the syringe and needle used to add a liquid to the reduced media. The increased lag time could

be explained by one more injection than the non-induced culture. Our last hypothesis was a problem with the starting OD, which would have impacted the lag time.

Hence, the expression of *gorS2* or *gorSL* did not impact the metabolism of the cells, as the main difference between induced and non-induced was for the WT. Unfortunately, the H_2 production was not measured, so it is difficult to conclude a total absence of target enzyme activity. However, with the current results, it might be possible that both the GAPOR and the GOR were not active. Different hypotheses could explain the lack of activity. The first is an incorrect folding of the protein in the cells, but according to the signal observed on Western blot (0) it is assumed that at least a part of the proteins was correctly folded. Nonetheless, the expression of soluble GOR-S could unfortunately not be confirmed by MS sequencing. The Adams group showed that Δ*gor-L* and *Δgor-S* mutants of *C. bescii* had similar growth behaviour (Scott et al., 2019). There is a high probability that GOR-L alone is not active. Hence, the lack of effect *gor-SL* of expression might be explained by an absence of either the production or a correct folding of the small subunit. A second hypothesis is a too low production of the enzymes, despite a strong induction, to impact the metabolites profile. Finally, the last hypothesis is the cofactor insertion. Even though *C. acetobutylicum* express an AOR, the enzyme was not characterised, and its cofactor is unknown. It is supposed to be identical to the *P. furiosus* AOR as the same cofactor was found in other bacteria and firmicutes (*i.e. C. bescii*). In the future, if the study is pursued, it would be interesting to test the same experiment but with complementation of the media with W instead of Mo.

4.3. Enzyme assay

An enzyme assay on purified enzyme had to be performed to understand if GAPOR and GOR produced in *C. acetobutylicum* were active. As presented in this chapter's former section, it was impossible to perform the complete growth, sonication lysis, purification, and enzyme assay in correct anaerobic conditions in Nottingham. Therefore, it was decided to send spores of the *C. acetobutylicum* CAC_AM004 and CAC_AM007 to the Toulouse laboratory,

where the enzyme assays on the GAPOR produced in *E. coli* were performed. Unfortunately, I could not go there, so Dr Céline Foulquier performed the assays following my protocols.

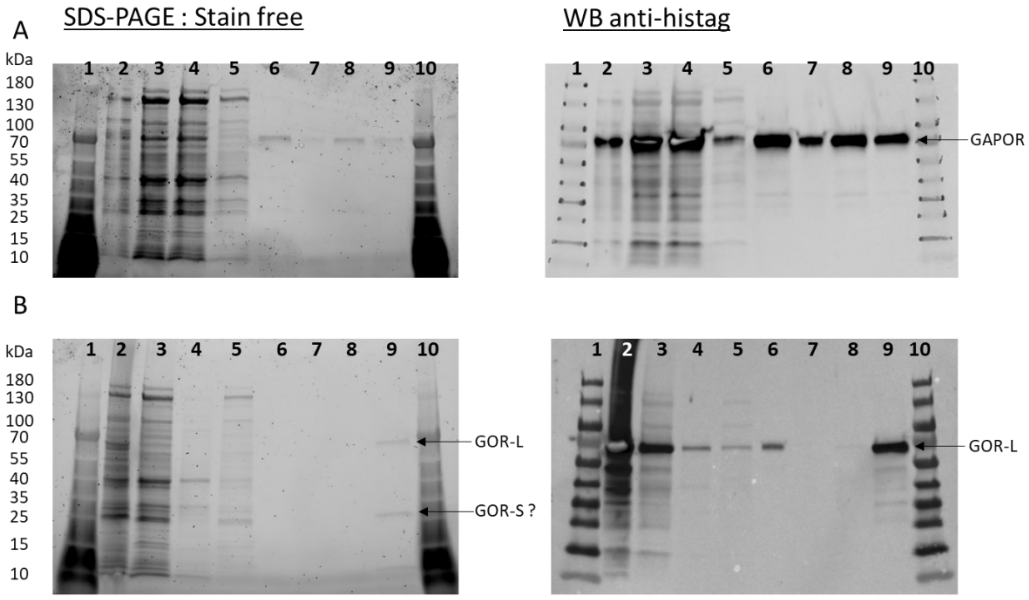


Figure V-7 SDS-PAGE and Western Blot anti-histag of the affinity purification of GAPOR and GOR produced in *C. acetobutylicum tcdR*$^{+}$ CAC_AM004 (*gorS2*) and CAC_AM007 (*gorSL*), respectively. Proteins were expressed in 120 ml batch cultures induced at OD 0.2 with 10 mM imidazole. The GAPOR and GOR present in the supernatants were purified by affinity chromatography in the anaerobic chamber. The purification buffer was supplemented with 0.024 mM sodium dithionate. SDS-PAGE was performed using 4 – 15 % Mini-PROTEAN TGX Stain-free Protein Gels, 10 wells (Bio-Rad), the ladder used was PageRuler Prestained Protein ladder, 10 to 180 kDa (Thermo Scientific) A) GAPOR purification 1 and 10: ladder; 2: lysis pellet;3: lysis supernatant; 4: buffer exchanged lysis supernatant; 5: flowthrough; 6: 75 mM imidazole wash; 7: desalted pooled elution fractions; 8: elution fraction 1 (1 ml); 9: elution fraction 2 (1 ml). GAPOR bands are marked with an arrow. B) GOR purification 1 and 10: ladder; 2: lysis pellet; 3: buffer exchanged lysis supernatant; 4: flowthrough; 5: wash 20 mM imidazole; 6: desalted pooled elution fractions; 7: elution fraction 1 (1 ml); 8: elution fraction 2 (1 ml); 9: elution fraction 2 (1 ml). GOR-L and putative GOR-S bands are marked with an arrow.

Unfortunately, no GAPOR or GOR activities were observed during the enzyme assay on purified enzymes, flowthrough, or desalted crude extract. As observed on the SDS-PAGE and western blot 'Figure V-7), the amount desalted purified GAPOR and GOR was drastically reduced compared to the crude extract. The actual causes for the lower amount of recovered protein are unknown, and it was impossible to reperform the experiments on time. In comparison to the preliminary experiment, the cells were harvested at similar OD or slightly lower for CAC_AM007, 0.8 instead of 1.2, which might explain the lower amount of GOR collected compared to GAPOR. A second cause might be the loss of GAPOR during the 75 mM imidazole wash. The GAPOR loss at this step was known but

was tolerated to remove a contaminant. However, this particular contaminant was not observed on SDS-PAGE or Western Blot, in this experiment. Hence, the second wash might not have been mandatory in this condition, and its removal would have probably allowed for recovering more enzymes.

Regarding the GAPOR assays, 17.3 µg of protein was collected after desalting, representing 8-fold less proteins collected compared to preliminary purification works. Nonetheless, the protein solution with a concentration of 5 µg/ml allowed an amount of enzyme in the cuvette between 250 ng and 2 µg under the different conditions tested. Park and co-workers reported a range of GAPOR-specific activity between 14 U/mg of GAPOR and 120 U/mg of GAPOR depending on the *E. coli* strain used (Park et al., 2007). Hence, if the GAPOR produced here was active, kinetic curves with a slope ranging from $5.2x10^{-2}$ ΔOD/min to 3.6 ΔOD/min should have been obtained. Therefore, in combination with the results of the batch culture experiments, it is highly probable that no active GAPOR was produced.

The GOR recovery was even lower as 10-fold less protein was recovered, 3.5 µg. Scott et al. reported a GOR specific activity around 50 U/mg of enzyme after nickel affinity purification (Scott et al., 2019). Hence, if the GOR produced here was active, kinetic curves with a slope ranging from $7.4x10^{-2}$ ΔOD/min to 0.3 ΔOD/min should have been obtained. Therefore, an identical conclusion to GAPOR could be drawn: no active GOR were expressed.

The lack of activity could have several explanations outside the low amount of protein. One hypothesis could be the incorporation of the wrong metallic atom. GAPOR from *M. maripaludis*, as formerly presented, was characterised as Mo-dependent when expressed in *E. coli* (Park et al., 2007). However, GOR from the mesophilic strain produced in this study was not characterised before. It is then possible that despite being from a mesophilic bacterium, the GOR might be W-dependent. In the future, it would then be interesting to test the expression of the encoding genes and the activity of both enzymes in media supplemented with tungstate instead of molybdate. It would also be

interesting to see if the level of soluble protein would be increased or lowered, as well as the impact on hydrogen, alcohol, and solvent production.

5. Discussion and conclusion

The project presented in this chapter aimed to produce active GAPOR and its newly discovered bacterial counterpart GOR in two Clostridial strains known to express AOR: *C. autoethanogenum* and *C. acetobutylicum*. The secondary objectives were to investigate the possibility of growing *the C. ljundahlii* RNF$^-$ strain on the sole CO and the possibility of using GAPOR or GOR to improve the metabolite production in *C. acetobutylicum*.

Encoding genes of the GOR characterised by Scott and co-workers were expressed in *C. bescii,* a thermophilic bacterium. As our aim was to use this enzyme in a metabolically engineered mesophilic bacteria, we decided to study the expression of GOR-encoding genes from a mesophilic bacterium. GOR from *G. ferrireducens* was picked among the other mesophilic bacteria potentially possessing GOR-encoding genes. Moreover, these mesophilic bacteria were not extensively studied, so it was difficult to estimate if the GOR-encoding genes were functional and expressed in their original strains. Hence, it would have been interesting to test the expression of potential GOR-encoding genes from other bacteria to maximise the odds of producing an active enzyme. Based on the work from Scott et al. (Scott et al., 2019) and then increasing the probability to produce an active GOR, the gene was synthesised with a his-tag sequence upstream *gorL.*

The heterologous production of soluble GAPOR and GOR was only detected in *C. acetobutylicum* under the control of the P*tcdB* inducible promoter. Moreover, mutation and deletion in the promoter and the genes were observed, indicating potential toxicity of the protein in both strains. It is unknown whether the toxicity was due to the enzyme activity or overexpression. The combination of the in vivo study with the enzyme assay strongly supports that both the GAPOR and GOR were produced but in an inactive form in *C. acetobutylicum*. Then if the study had to be continued, it

would be extremely interesting to test the *gorS2* and *gorSL* expression in a tungstate-supplemented medium following the recent findings of the Adams' group (Schut et al., 2021), which demonstrated the presence of active tungsten-containing AOR in mesophilic Mo-dependent bacteria. Moreover, As mentioned above, it would be interesting to test GOR encoding genes from different mesophilic bacteria with both Mo and W. This could help to find active GOR during heterologous production.

In conclusion, GAPOR and GOR-L were successfully expressed in *C. acetobutylicum,* but the GOR-S production could not be validated; none of them were found active. Hence, active enzymes were not produced despite the use of AOR-producing strains as a host for *gorS2* and *gorSL* expression. When the method for the molybdenum cofactor analysis is correctly set up and validated, it would be interesting to determine the type of cofactor incorporated in the enzyme, if any cofactor is incorporated, during the expression of *gorS2* and *gorSL* in *C. acetobutylicum*.

Chapter VI:

Conclusion and Perspectives

This research project was built with the objective of engineering *E. coli* to make it able to grow on CO as the sole carbon and energy source. To do so, it was conceptualised to use a CODH to produce CO_2 and reduced ferredoxin carrying very low potential electrons. The CBB cycle would then be used to fix CO_2 with the RuBisCo. The gluconeogenesis part of the cycle was planned to be replaced by a GAPOR functioning in the reverse direction using the 3PG produced by the RuBisCo as substrate and the reduced ferredoxin produced by the CODH, bringing the low potential electrons. To start, it was decided to study the expression of *gorS2* encoding the *M. maripaludis* GAPOR in the chosen *E. coli* strain.

The study of the *gorS2* expression in *E. coli* showed an absence of glyceraldehyde-3-phosphate oxidation activity in our test. The absence of activity prevented us from pursuing further studying the reverse reaction activity. Even though we have reproduced the *gorS2* expression presented in the Park *et al.* article (Park et al., 2007), it was not the case for the GAPOR activity. There are several hypotheses regarding the lack of GAPOR activity.

Firstly, the conditions used for the growth and production of the GAPOR. There were some differences between our protocol and the Park *et al*. protocol: the growth was in relatively small batches using rich media under micro-aerobic conditions, while they used litre scale baches in minimal media under strictly anaerobic conditions (Park et al., 2007). The low level of *gors2* expression was a struggle all around the project that, for instance, prevented us from performing crucial experiment such as ICP-MS. Nevertheless, it would be surprising if these modifications to their protocol was the cause of the lack of activity. Indeed, they stated that when the cells were grown aerobically and purified in the presence of oxygen, the GAPOR was active as long as L-cysteine was added to the enzyme assay reaction mix. Surprisingly, we found that L-cysteine, with the concentration used in Park *et al.* article, was reducing BV at a very strong rate. It was difficult for us to understand how they could measure GAPOR activity with such a high background due to the chemical reduction of BV by L-cysteine.

Secondly, the total amount of GAPOR produced and purified might be too low. Despite the production of enzyme performed in large volume (3 L) to compensate for it, no activity was observed either. Hence, if any active GAPOR was produced, the ratio of active GAPOR:total produced GAPOR was too low to be detected during the assay.

Another hypothesis, as explained in the thesis, was an issue with the incorporation of the cofactor. It is known that the insertion of a complex cofactor, such as the Mo-bisMPT, could be a blocking point during heterologous protein production. During the course of the project, the public knowledge regarding the molybdenum cofactor synthesis increased, mainly regarding the Mo-bisMPT in *E. coli* (Leimkühler, 2020; Reschke et al., 2019). Moreover, the complementation of the *E. coli* strain by the genes from *M. maripaludis* encoding the enzymes involved in the molybdenum cofactor synthesis was planned from the beginning of the project. Indeed,increasing the 0.74 mol of Mo per mol of protein measured by Park and co-workers (Park et al., 2007) was supposed to be a way to improve the GAPOR specific activity. However, these experiments did not either allow the production of an active GAPOR. The unambiguous production of only two enzymes of the Moco biosynthesis pathway was fine with the objective of improving the GAPOR activity. However, it drastically reduced the chance of increasing the level of the correct molybdenum cofactor in the enzyme.

The principal issue of the cofactor study was the impossibility of setting up a robust method to identify the cofactor. It was The MS methods was suitable to observe the pterin cofactor. However, the cofactor extraction method needs more work to be fully validated. Incorporating different cofactor-containing enzymes and other extraction methods used for Mo-cofactors extraction would be useful to validate the method entirely. The absence of this kind of analytical method did not allow us to validate or rule out our hypothesis nor to make progress on the understanding of the cofactor biosynthesis in *M. maripaludis*. In the cofactor analysis, it would have been necessary to produce and purify enough GAPOR for Mo quantitation with the *E. coli* RG_AM023gapor

by ICP-MS. Several milligrams of GAPOR were indeed collected for activity assay in Toulouse but never used for metal content quantitation. This result would have given an estimation of the basal cofactor incorporation yield. Another less-discussed aspect here was the potential second cofactor synthesis pathway in *E. coli* responsible for the cofactor synthesis and insertion in YdhV (Scott et al., 2019). In our study, no GAPOR activity was observed above the background in the strains Δ*moeAmobAB*. Hence, it could be supposed that the YdhV cofactor pathway might be very specific to this enzyme. Investigating this pathway could help produce an active GAPOR in *E. coli*. In fact, the MS cofactor analysis methods should have been validated with already published homologously expressed, such as *P. furiosus* GAPOR, before testing with the studied enzyme.

Overall, the results obtained with the production of the GAPOR from *M. maripaludis* in *E. coli* raised more questions than they brought answers mainly due to the impossibility to reproduce what was planned to be the basal level of active GAPOR production to allow the study of the opposite enzymatic activity.

The study of the GAPOR and its bacterial homologue, GOR, production and G3P oxidation activity in *C. autoethanogenum* and *C. acetobutylicum* gave promising results. Both enzymes were shown to be produced in *C. acetobutylicum* and extracted in the soluble fraction of the lysate with a higher yield than in *E. coli*. Despite being produced in an AOR-expressing strain, no G3P oxidation activity was measured during the enzyme assay neither was a modification of the bacterial metabolism observed. It can then be concluded that no active enzymes were purified nor produced in *C. acetobutylicum.*

In the future, when the cofactor analysis methods are finalised, the type of cofactor inserted into the GAPOR or GOR produced in the Clostridia must be determined. Comparing the results in *E. coli* would give us insight into the cofactor expression in these strains. Both clostridia strains are better candidates than *E. coli* to produce active GAPOR and GOR in terms of the level of production of the protein as well as for the type of Mo/W-cofactor that is

produced. Finally, it will be interesting to test if tungstate complementation, for instance, will allow the production of an active GAPOR or GOR in both studied Clostridia.

In conclusion, the absence of reproducibility of the GAPOR oxidoreduction activity was a major drawback in the project's initial plan, which prevented us from further investigating the 3PG reduction activity of the enzymes and, as a consequence, the study of the CODH:ferredoxin:GAPOR pathway. However, the main failure of the study was the absence of a reliable analytical method for determining of the enzymatic cofactor content. Due to delays caused by the unavailability of the LC-MS apparatus, the first results were received too late in the project, not allowing us to validate the method. Nevertheless, the LC-MS method was shown to be accurate in detecting amount of pterin-6-carboxylic acid as low as 100 pmol. Hence, with the test of different cofactor extraction methods, it should be possible to set up robust cofactor identification methods. The next step would be to implement standard molecules to be able to couple the cofactor identification with accurate titration, which would, in complementation of ICP-MS measurement and accurate protein titration, allow a very accurate cofactor identification/quantification.

The lack of a correct analysis method prevented us from taking full advantage of the investigation of the *M. maripaludis* cofactor biosynthesis pathway. Even though the experiments were planned to improve the GAPOR activity, it was a way to investigate the Mo-bisMPT synthesis pathway in *M. maripaludis*. Therefore, the lack of activity and a validated cofactor analysis method kept us from gathering more insight into the archaeal pathway. Nevertheless, the comparison between *E. coli* and *M. maripaludis* molybdenum cofactor pathway gave the opportunity to observe the difference of this archaeal pathway and the most studied one. These differences opened new research paths to understand how the Mo-bisMPT cofactor is formed and acquire new understanding of the archaeal pathway.

It seems interesting to pursue the development of the GAPOR and GOR enzyme production in clostridia. Indeed, GAPOR and GOR production in

C. autoethanogenum requires more fine-tuning. However, the potential toxicity observed during the *gorSL* expression suggests that active GAPOR and GOR could be produced from this strain with the correct level of expression and evolution. The same remarks could be made for *C. acetobutylicum*, as producing an active form of these enzymes could open new ways to improve or modify solvent and hydrogen production in *C. acetobutylicum* from renewable sources.

References

Abrini, J., Naveau, H., & Nyns, E.-J. (1994). Clostridium autoethanogenum, sp. nov., an anaerobic bacterium that produces ethanol from carbon monoxide. *Archives of Microbiology, 161*(4), 345-351. doi:10.1007/BF00303591

Adinolfi, S., Puglisi, R., Crack, J. C., Iannuzzi, C., Dal Piaz, F., Konarev, P. V., . . . Pastore, A. (2018). The Molecular Bases of the Dual Regulation of Bacterial Iron Sulfur Cluster Biogenesis by CyaY and IscX. *Frontiers in Molecular Biosciences, 4*, 97. Retrieved from https://www.frontiersin.org/article/10.3389/fmolb.2017.00097

Agar, J. N., Krebs, C., Frazzon, J., Huynh, B. H., Dean, D. R., & Johnson, M. K. (2000). IscU as a Scaffold for Iron−Sulfur Cluster Biosynthesis: Sequential Assembly of [2Fe-2S] and [4Fe-4S] Clusters in IscU. *Biochemistry, 39*(27), 7856-7862. doi:10.1021/bi000931n

Aguilar-Barajas, E., Díaz-Pérez, C., Ramírez-Díaz, M. I., Riveros-Rosas, H., & Cervantes, C. (2011). Bacterial transport of sulfate, molybdate, and related oxyanions. *BioMetals, 24*(4), 687-707. doi:10.1007/s10534-011-9421-x

Akhtar, M. K., & Jones, P. R. (2008). Deletion of iscR stimulates recombinant clostridial Fe–Fe hydrogenase activity and H2-accumulation in Escherichia coli BL21(DE3). *Applied Microbiology and Biotechnology, 78*(5), 853-862. doi:10.1007/s00253-008-1377-6

Altschul, S. F., Gish, W., Miller, W., Myers, E. W., & Lipman, D. J. (1990). Basic local alignment search tool. *J Mol Biol, 215*(3), 403-410. doi:10.1016/s0022-2836(05)80360-2

Ames, B. N. (1966). [10] Assay of inorganic phosphate, total phosphate and phosphatases. In *Methods in Enzymology* (Vol. 8, pp. 115-118): Academic Press.

Amy, N. K., & Rajagopalan, K. V. (1979). Characterization of molybdenum cofactor from Escherichia coli. *Journal of Bacteriology, 140*(1), 114-124. doi:10.1128/jb.140.1.114-124.1979

Anderson, L. A., McNairn, E., Lubke, T., Pau, R. N., & Boxer, D. H. (2000). ModE-dependent molybdate regulation of the molybdenum cofactor operon moa in Escherichia coli. *J Bacteriol, 182*(24), 7035-7043.

Anderson, L. A., Palmer, T., Price, N. C., Bornemann, S., Boxer, D. H., & Pau, R. N. (1997). Characterisation of the Molybdenum-Responsive ModE Regulatory Protein and its Binding to the Promoter Region of the modABCD (Molybdenum Transport) Operon of Escherichia Coli. *European Journal of Biochemistry, 246*(1), 119-126. doi:https://doi.org/10.1111/j.1432-1033.1997.00119.x

Anderson Lisa, A., McNairn, E., Leubke, T., Pau Richard, N., & Boxer David, H. (2000). ModE-Dependent Molybdate Regulation of the Molybdenum Cofactor Operon moa in Escherichia coli. *Journal of Bacteriology, 182*(24), 7035-7043. doi:10.1128/JB.182.24.7035-7043.2000

Antonovsky, N., Gleizer, S., & Milo, R. (2017). Engineering carbon fixation in E . coli : from heterologous RuBisCO expression to the Calvin – Benson – Bassham cycle. *Current Opinion in Biotechnology, 47*, 83-91. doi:10.1016/j.copbio.2017.06.006

Antonovsky, N., Gleizer, S., & Milo, R. (2017). Engineering carbon fixation in E. coli: from heterologous RuBisCO expression to the Calvin–Benson–Bassham cycle. *Current Opinion in Biotechnology, 47*, 83-91. doi:10.1016/j.copbio.2017.06.006

Arndt, F., Schmitt, G., Winiarska, A., Saft, M., Seubert, A., Kahnt, J., & Heider, J. (2019). Characterization of an Aldehyde Oxidoreductase From the Mesophilic Bacterium Aromatoleum aromaticum EbN1, a Member of a New Subfamily of Tungsten-Containing Enzymes. *Frontiers in Microbiology, 10*. doi:10.3389/fmicb.2019.00071

Arnoux, P., Ruppelt, C., Oudouhou, F., Lavergne, J., Siponen, M. I., Toci, R., . . . Walburger, A. (2015). Sulphur shuttling across a chaperone during molybdenum cofactor maturation. *Nat Commun, 6*(1), 6148. doi:10.1038/ncomms7148

Atkinson, J. T., Campbell, I., Bennett, G. N., & Silberg, J. J. (2016). Cellular Assays for Ferredoxins: A Strategy for Understanding Electron Flow through Protein Carriers That Link Metabolic Pathways. *Biochemistry, 55*(51), 7047-7064. doi:10.1021/acs.biochem.6b00831

Baba, T., Ara, T., Hasegawa, M., Takai, Y., Okumura, Y., Baba, M., . . . Mori, H. (2006). Construction of Escherichia coli K-12 in-frame, single-gene knockout mutants: the Keio collection. *Molecular Systems Biology, 2*, 2006.0008-2006.0008. doi:10.1038/msb4100050

Barras, F., Loiseau, L., & Py, B. (2005). How Escherichia coli and Saccharomyces cerevisiae Build Fe/S Proteins. In R. K. Poole (Ed.), *Advances in Microbial Physiology* (Vol. 50, pp. 41-101): Academic Press.

Basen, M., Schut, G. J., Nguyen, D. M., Lipscomb, G. L., Benn, R. A., Prybol, C. J., . . . Adams, M. W. W. (2014). Single gene insertion drives bioalcohol production by a thermophilic archaeon. *Proceedings of the National Academy of Sciences, 111*(49), 17618. doi:10.1073/pnas.1413789111

Bertini, I., Donaire, A., Feinberg, B. A., Luchinat, C., Picciolp, M., & Yuan, H. (1995). Solution Structure of the Oxidized 2[4Fe-4S] Ferredoxin from Clostridium Pasteurianum. *European Journal of Biochemistry, 232*(1), 192-205. doi:https://doi.org/10.1111/j.1432-1033.1995.tb20799.x

Bertsch, J., & Muller, V. (2015). CO Metabolism in the Acetogen Acetobacterium woodii. *Appl Environ Microbiol, 81*(17), 5949-5956. doi:10.1128/aem.01772-15

Bevers, L. E., Bol, E., Hagedoorn, P. L., & Hagen, W. R. (2005). WOR5, a novel tungsten-containing aldehyde oxidoreductase from Pyrococcus furiosus with a broad substrate specificity. *Journal of Bacteriology, 187*(20), 7056-7061. doi:10.1128/JB.187.20.7056-7061.2005

Bevers, L. E., Hagedoorn, P.-L., & Hagen, W. R. (2009). The bioinorganic chemistry of tungsten. *Coordination Chemistry Reviews, 253*(3), 269-290. doi:https://doi.org/10.1016/j.ccr.2008.01.017

Bevers, L. E., Hagedoorn, P.-L., Santamaria-Araujo, J. A., Magalon, A., Hagen, W. R., & Schwarz, G. (2008a). Function of MoaB Proteins in the Biosynthesis of the Molybdenum and Tungsten Cofactors. *Biochemistry, 47*(3), 949-956. doi:10.1021/bi7020487

Bevers, L. E., Hagedoorn, P. L., Santamaria-Araujo, J. A., Magalon, A., Hagen, W. R., & Schwarz, G. (2008b). Function of MoaB proteins in the biosynthesis of the molybdenum and tungsten cofactors. *Biochemistry, 47*(3), 949-956. doi:10.1021/bi7020487

Bevers Loes, E., Hagedoorn, P.-L., Krijger Gerard, C., & Hagen Wilfred, R. (2006). Tungsten Transport Protein A (WtpA) in Pyrococcus furiosus: the First Member of a New Class of Tungstate and Molybdate Transporters. *Journal of Bacteriology, 188*(18), 6498-6505. doi:10.1128/JB.00548-06

Blasco, F., Dos Santos, J.-P., Magalon, A., Frixon, C., Guigliarelli, B., Santini, C.-L., & Giordano, G. (1998). NarJ is a specific chaperone required for molybdenum cofactor assembly in nitrate reductase A of Escherichia coli. *Molecular Microbiology, 28*(3), 435-447. doi:https://doi.org/10.1046/j.1365-2958.1998.00795.x

Blasco, F., Pommier, J., Augier, V., Chippaux, M., & Giordano, G. (1992). Involvement of the narJ or narW gene product in the formation of active nitrate reductase in Escherichia coli. *Molecular Microbiology, 6*(2), 221-230. doi:https://doi.org/10.1111/j.1365-2958.1992.tb02003.x

Blokesch, M., Paschos, A., Bauer, A., Reissmann, S., Drapal, N., & Böck, A. (2004). Analysis of the transcarbamoylation-dehydration reaction catalyzed by the hydrogenase maturation proteins HypF and HypE. *European Journal of Biochemistry, 271*(16), 3428-3436. doi:10.1111/j.1432-1033.2004.04280.x

Brereton, P. S., Verhagen, M. F. J. M., Zhou, Z. H., & Adams, M. W. W. (1998). Effect of Iron-Sulfur Cluster Environment in Modulating the Thermodynamic Properties and Biological Function of Ferredoxin from Pyrococcus furiosus. *Biochemistry, 37*(20), 7351-7362. doi:10.1021/bi972864b

Bridgwater, A. V. (2003). Renewable fuels and chemicals by thermal processing of biomass. *Chemical Engineering Journal, 91*(2), 87-102. doi:https://doi.org/10.1016/S1385-8947(02)00142-0

Brown, S. D., Nagaraju, S., Utturkar, S., De Tissera, S., Segovia, S., Mitchell, W., . . . Köpke, M. (2014). Comparison of single-molecule sequencing and hybrid approaches for finishing the genome of Clostridium autoethanogenum and analysis of CRISPR systems in industrial relevant Clostridia. *Biotechnology for Biofuels, 7*(1), 40. doi:10.1186/1754-6834-7-40

Bruschi, M., & Guerlesquin, F. (1988). Structure, function and evolution of bacterial ferredoxins. *FEMS Microbiology Letters, 54*(2), 155-175. doi:10.1016/0378-1097(88)90175-9

Buckel, W., & Thauer, R. K. (2013). Energy conservation via electron bifurcating ferredoxin reduction and proton/Na+ translocating ferredoxin oxidation. *Biochimica et Biophysica Acta (BBA) - Bioenergetics, 1827*(2), 94-113. doi:https://doi.org/10.1016/j.bbabio.2012.07.002

Burgmayer, S. J. N., Kim, M., Petit, R., Rothkopf, A., Kim, A., BelHamdounia, S., . . . Kirk, M. L. (2007). Synthesis, Characterization, and Spectroscopy of Model Molybdopterin Complexes(). *Journal of inorganic biochemistry, 101*(11-12), 1601-1616. doi:10.1016/j.jinorgbio.2007.07.012

Cai, K., Frederick, R. O., & Markley, J. L. (2020). ISCU interacts with NFU1, and ISCU[4Fe-4S] transfers its Fe-S cluster to NFU1 leading to the production of holo-NFU1. *Journal of Structural Biology, 210*(2), 107491. doi:https://doi.org/10.1016/j.jsb.2020.107491

Campbell, I. J., Bennett, G. N., & Silberg, J. J. (2019). Evolutionary Relationships Between Low Potential Ferredoxin and Flavodoxin Electron Carriers. *Frontiers in Energy Research, 7*. doi:10.3389/fenrg.2019.00079

Cañadas, I. C., Groothuis, D., Zygouropoulou, M., Rodrigues, R., & Minton, N. P. (2019). RiboCas: A Universal CRISPR-Based Editing Tool for Clostridium. *ACS Synthetic Biology, 8*(6), 1379-1390. doi:10.1021/acssynbio.9b00075

Candiano, G., Bruschi, M., Musante, L., Santucci, L., Ghiggeri Gian, M., Carnemolla, B., . . . Righetti Pier, G. (2004). Blue silver: A very sensitive colloidal Coomassie G-250 staining for proteome analysis. *ELECTROPHORESIS, 25*(9), 1327-1333. doi:10.1002/elps.200305844

Carlson, E. D., & Papoutsakis, E. T. (2017). Heterologous Expression of the Clostridium carboxidivorans CO Dehydrogenase Alone or Together with the Acetyl Coenzyme A Synthase Enables both Reduction of CO2 and Oxidation of CO by Clostridium acetobutylicum. *Appl Environ Microbiol, 83*(16). doi:10.1128/aem.00829-17

Cartman Stephen, T., & Minton Nigel, P. (2010). A mariner-Based Transposon System for In Vivo Random Mutagenesis of Clostridium difficile. *Applied and Environmental Microbiology, 76*(4), 1103-1109. doi:10.1128/AEM.02525-09

Cetnar, D. P., & Salis, H. M. (2020). Systematic Quantification of Sequence and Structural Determinants Controlling mRNA stability in Bacterial Operons. *bioRxiv*, 2020.2007.2022.216051. doi:10.1101/2020.07.22.216051

Chan, M. K., Mukund, S., Kletzin, A., Adams, M. W. W., & Rees , D. C. (1995). Structure of a Hyperthermophilic Tungstopterin Enzyme, Aldehyde Ferredoxin Oxidoreductase. *Science, 267*(5203), 1463-1469. doi:10.1126/science.7878465

Chandramouli, K., & Johnson, M. K. (2006). HscA and HscB Stimulate [2Fe-2S] Cluster Transfer from IscU to Apoferredoxin in an ATP-Dependent Reaction. *Biochemistry, 45*(37), 11087-11095. doi:10.1021/bi061237w

Chandramouli, K., Unciuleac, M.-C., Naik, S., Dean, D. R., Huynh, B. H., & Johnson, M. K. (2007). Formation and Properties of [4Fe-4S] Clusters on the IscU Scaffold Protein. *Biochemistry, 46*(23), 6804-6811. doi:10.1021/bi6026659

Cherepanov, P. P., & Wackernagel, W. (1995). Gene disruption in Escherichia coli: TcR and KmR cassettes with the option of Flp-catalyzed excision of the antibiotic-resistance determinant. *Gene, 158*(1), 9-14. doi:https://doi.org/10.1016/0378-1119(95)00193-A

Child, S. A., Bradley, J. M., Pukala, T. L., Svistunenko, D. A., Le Brun, N. E., & Bell, S. G. (2018). Electron transfer ferredoxins with unusual cluster binding motifs support secondary metabolism in many bacteria. *Chemical science, 9*(41), 7948-7957.

Compan, I., & Touati, D. (1994). Anaerobic activation of arcA transcription in Escherichia coli: roles of Fnr and ArcA. *Molecular Microbiology, 11*(5), 955-964. doi:https://doi.org/10.1111/j.1365-2958.1994.tb00374.x

Costa, K. C., Lie, T. J., Jacobs, M. A., & Leigh, J. A. (2013). H2-independent growth of the hydrogenotrophic methanogen Methanococcus maripaludis. *MBio, 4*(2), 00062-00013.

Croux, C., Nguyen, N.-P.-T., Lee, J., Raynaud, C., Saint-Prix, F., Gonzalez-Pajuelo, M., . . . Soucaille, P. (2016). Construction of a restriction-less, marker-less mutant useful for functional genomic and metabolic engineering of the biofuel producer Clostridium acetobutylicum. *Biotechnology for Biofuels, 9*(1), 23. doi:10.1186/s13068-016-0432-2

Dahl, J.-U., Urban, A., Bolte, A., Sriyabhaya, P., Donahue, J. L., Nimtz, M., . . . Leimkühler, S. (2011). The Identification of a Novel Protein Involved in Molybdenum Cofactor Biosynthesis in Escherichia coli. *The Journal of Biological Chemistry, 286*(41), 35801-35812. doi:10.1074/jbc.M111.282368

Dahl, J. U., Radon, C., Bühning, M., Nimtz, M., Leichert, L. I., Denis, Y., . . . Leimkühler, S. (2013). The sulfur carrier protein tusa has a pleiotropic role in escherichia coli that also affects molybdenum cofactor biosynthesis. *Journal of Biological Chemistry, 288*(8), 5426-5442. doi:10.1074/jbc.M112.431569

Daniell, J., Köpke, M., & Simpson, S. D. (2012). Commercial biomass syngas fermentation. *Energies, 5*(12), 5372-5417. doi:10.3390/en5125372

Daniels, J. N., Wuebbens, M. M., Rajagopalan, K. V., & Schindelin, H. (2008). Crystal Structure of a Molybdopterin Synthase−Precursor Z Complex: Insight into Its Sulfur Transfer Mechanism and Its Role in Molybdenum Cofactor Deficiency. *Biochemistry, 47*(2), 615-626. doi:10.1021/bi701734g

Datsenko, K. A., & Wanner, B. L. (2000). One-step inactivation of chromosomal genes in Escherichia coli K-12 using PCR products. *Proceedings of the National Academy of Sciences of the United States of America, 97*(12), 6640-6645. doi:10.1073/pnas.120163297

Demuez, M., Cournac, L., Guerrini, O., Soucaille, P., & Girbal, L. (2007). Complete activity profile of Clostridium acetobutylicum [FeFe]-hydrogenase and kinetic parameters for endogenous redox partners. *FEMS Microbiology Letters, 275*(1), 113-121. doi:10.1111/j.1574-6968.2007.00868.x

Desnoyers, G., Morissette, A., Prévost, K., & Massé, E. (2009). Small RNA-induced differential degradation of the polycistronic mRNA iscRSUA. *The EMBO Journal, 28*(11), 1551-1561. doi:10.1038/emboj.2009.116

Diender, M., Stams, A. J. M., & Sousa, D. Z. (2015). Pathways and Bioenergetics of Anaerobic Carbon Monoxide Fermentation. *Frontiers in Microbiology, 6*, 1275. doi:10.3389/fmicb.2015.01275

Dijkhuizen, L., & Harder, W. (1984). Current views on the regulation of autotrophic carbon dioxide fixation via the Calvin cycle in bacteria. *Antonie van Leeuwenhoek, 50*(5), 473-487. doi:10.1007/BF02386221

Dupuy, B., & Matamouros, S. (2006). Regulation of toxin and bacteriocin synthesis in Clostridium species by a new subgroup of RNA polymerase σ-factors. *Research in Microbiology, 157*(3), 201-205. doi:https://doi.org/10.1016/j.resmic.2005.11.004

Directive 2009/28/EC of the European Parliament and of the Council of 23 April 2009 on the promotion of the use of energy from renewable sources and amending and subsequently repealing Directives 2001/77/EC and 2003/30/EC, L 140/16 C.F.R. (2009).

Eurostat. (2019). Glossary:Biomass Retrieved from https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Glossary:Biomass

Farasat, I., Kushwaha, M., Collens, J., Easterbrook, M., Guido, M., & Salis, H. M. (2014). Efficient search, mapping, and optimization of multi-protein genetic systems in diverse bacteria. *Molecular Systems Biology, 10*(6), 731. doi:https://doi.org/10.15252/msb.20134955

Fast, A., G., & Papoutsakis, E., T. (2018). Functional Expression of the Clostridium ljungdahlii Acetyl-Coenzyme A Synthase in Clostridium acetobutylicum as Demonstrated by a Novel In Vivo CO Exchange Activity En Route to Heterologous Installation of a Functional Wood-Ljungdahl Pathway. *Applied and Environmental Microbiology, 84*(7), e02307-02317. doi:10.1128/AEM.02307-17

Francis, D. M., & Page, R. (2010). Strategies to optimize protein expression in E. coli. In *Current Protocols in Protein Science*.

Gatti-Lafranconi, P., Dijkman, W. P., Devenish, S. R. A., & Hollfelder, F. (2013). A single mutation in the core domain of the lac repressor reduces leakiness. *Microbial Cell Factories, 12*(1), 67. doi:10.1186/1475-2859-12-67

Genest, O., Méjean, V., & Iobbi-Nivol, C. (2009). Multiple roles of TorD-like chaperones in the biogenesis of molybdoenzymes. *FEMS Microbiology Letters, 297*(1), 1-9. doi:10.1111/j.1574-6968.2009.01660.x

Gennaris, A., Ezraty, B., Henry, C., Agrebi, R., Vergnes, A., Oheix, E., . . . Barras, F. (2015). Repairing oxidized proteins in the bacterial envelope using respiratory chain electrons. *Nature, 528*(7582), 409-412. doi:10.1038/nature15764

Genthner, B. R., & Bryant, M. P. (1982). Growth of Eubacterium limosum with Carbon Monoxide as the Energy Source. *Appl Environ Microbiol, 43*(1), 70-74.

Ghosh, A. C., Samuel, P. P., & Schulzke, C. (2017). Synthesis, characterization and oxygen atom transfer reactIVity of a pair of Mo(IV)O- and Mo(VI)O2-enedithiolate complexes-a look at both ends of the catalytic transformation. *Dalton Transactions, 46*(23), 7523-7533. doi:10.1039/c7dt01470h

Giel, J. L., Nesbit, A. D., Mettert, E. L., Fleischhacker, A. S., Wanta, B. T., & Kiley, P. J. (2013). Regulation of iron–sulphur cluster homeostasis through

transcriptional control of the Isc pathway by [2Fe–2S]–IscR in Escherichia coli. *Molecular Microbiology, 87*(3), 478-492. doi:10.1111/mmi.12052

Giel, J. L., Rodionov, D., Liu, M., Blattner, F. R., & Kiley, P. J. (2006). IscR-dependent gene expression links iron-sulphur cluster assembly to the control of O2-regulated genes in Escherichia coli. *Molecular Microbiology, 60*(4), 1058-1075. doi:10.1111/j.1365-2958.2006.05160.x

Girbal, L., Mortier-Barrière, I., Raynaud, F., Rouanet, C., Croux, C., & Soucaille, P. (2003). Development of a Sensitive Gene Expression Reporter System and an Inducible Promoter-Repressor System for Clostridium acetobutylicum</em&gt. *Applied and Environmental Microbiology, 69*(8), 4985. doi:10.1128/AEM.69.8.4985-4988.2003

Girbal, L., von Abendroth, G., Winkler, M., Benton, P. M. C., Meynial-Salles, I., Croux, C., . . . Soucaille, P. (2005). Homologous and Heterologous Overexpression in Clostridium acetobutylicum and Characterization of Purified Clostridial and Algal Fe-Only Hydrogenases with High Specific Activities. *Applied and Environmental Microbiology, 71*(5), 2777. doi:10.1128/AEM.71.5.2777-2781.2005

Gleizer, S., Ben-Nissan, R., Bar-On, Y. M., Antonovsky, N., Noor, E., Zohar, Y., . . . Milo, R. (2019). Conversion of Escherichia coli to Generate All Biomass Carbon from CO2. *Cell, 179*(6), 1255-1263.e1212. doi:https://doi.org/10.1016/j.cell.2019.11.009

Gorwa, M. F., Croux, C., & Soucaille, P. (1996). Molecular characterization and transcriptional analysis of the putative hydrogenase gene of Clostridium acetobutylicum ATCC 824. *Journal of Bacteriology, 178*(9), 2668-2675. doi:10.1128/jb.178.9.2668-2675.1996

Goyal, N., Zhou, Z., & Karimi, I. A. (2016). Metabolic processes of Methanococcus maripaludis and potential applications. *Microbial Cell Factories, 15*, 107. doi:10.1186/s12934-016-0500-0

Grunden, A. M., Ray, R. M., Rosentel, J. K., Healy, F. G., & Shanmugam, K. T. (1996). Repression of the Escherichia coli modABCD (molybdate transport) operon by ModE. *Journal of Bacteriology, 178*(3), 735-744. doi:10.1128/jb.178.3.735-744.1996

Grunden, A. M., & Shanmugam, K. T. (1997). Molybdate transport and regulation in bacteria. *Archives of Microbiology, 168*(5), 345-354. doi:10.1007/s002030050508

Guerrini, O. (2007). *Etude des relations structure-fonction de la Glycéraldéhyde 3-phosphate oxydoréductase et Ingénierie métabolique de Clostridium acetobutylicum pour la production de Biohydrogène*. Retrieved from http://www.theses.fr/2007ISAT0039/document Available from http://www.theses.fr/2007ISAT0039

Guerrini, O., Burlat, B., Léger, C., Guigliarelli, B., Soucaille, P., & Girbal, L. (2008). Characterization of Two 2[4Fe4S] Ferredoxins from Clostridium acetobutylicum. *Current Microbiology, 56*(3), 261-267. doi:10.1007/s00284-007-9072-x

Hagedoorn, P.-L. (2019). Steady-state kinetics of the tungsten containing aldehyde: Ferredoxin oxidoreductases from the hyperthermophilic archaeon Pyrococcus furiosus. In *Journal of Biotechnology* (Vol. 306, pp. 142-148). Department of Biotechnology, Delft University of Technology, Van der Maasweg 9, Delft, 2629HZ, Netherlands.

Hagedoorn Peter, L., Freije, J. R., & Hagen Wilfred, R. (1999). Pyrococcus furiosus glyceraldehyde 3-phosphate oxidoreductase has comparable W6+/5+ and W5+/4+ reduction potentials and unusual [4Fe-4S] EPR properties. *FEBS Letters, 462*(1-2), 66-70. doi:10.1016/S0014-5793(99)01511-2

Hagedoorn, P. L. (2019). Steady-state kinetics of the tungsten containing aldehyde: Ferredoxin oxidoreductases from the hyperthermophilic archaeon Pyrococcus furiosus. *Journal of Biotechnology, 306*, 142-148. doi:10.1016/j.jbiotec.2019.10.005

Hagedoorn, P. L., Chen, T., Schröder, I., Piersma, S. R., Vries, S. d., & Hagen, W. R. (2005). Purification and characterization of the tungsten enzyme aldehyde:ferredoxin oxidoreductase from the hyperthermophilic denitrifier Pyrobaculum aerophilum. *JBIC Journal of Biological Inorganic Chemistry, 10*(3), 259-269. doi:10.1007/s00775-005-0637-5

Hageman, R. V., & Rajagopalan, K. V. (1986). [62] Assay and detection of the molybdenum cofactor. In *Methods in Enzymology* (Vol. 122, pp. 399-412): Academic Press.

Hagen, W. R. (2011). Cellular uptake of molybdenum and tungsten. *Coordination Chemistry Reviews, 255*(9), 1117-1128. doi:https://doi.org/10.1016/j.ccr.2011.02.009

Halper, S. M., Hossain, A., & Salis, H. M. (2020). Synthesis Success Calculator: Predicting the Rapid Synthesis of DNA Fragments with Machine Learning. *ACS Synthetic Biology, 9*(7), 1563-1571. doi:10.1021/acssynbio.9b00460

Hasona, A., Self, W. T., & Shanmugam, K. T. (2001). Transcriptional regulation of the moe (molybdate metabolism) operon of Escherichia coli. *Archives of Microbiology, 175*(3), 178-188. doi:10.1007/s002030100252

Heap, J. T., Pennington, O. J., Cartman, S. T., Carter, G. P., & Minton, N. P. (2007). The ClosTron: A universal gene knock-out system for the genus Clostridium. *Journal of Microbiological Methods, 70*(3), 452-464. doi:https://doi.org/10.1016/j.mimet.2007.05.021

Heap, J. T., Pennington, O. J., Cartman, S. T., & Minton, N. P. (2009). A modular system for Clostridium shuttle plasmids. *Journal of Microbiological Methods, 78*(1), 79-85. doi:https://doi.org/10.1016/j.mimet.2009.05.004

Heider, J., Ma, K., & Adams, M. W. (1995). Purification, characterization, and metabolic function of tungsten-containing aldehyde ferredoxin oxidoreductase from the hyperthermophilic and proteolytic archaeon Thermococcus strain ES-1. *Journal of Bacteriology, 177*(16), 4757-4764. doi:10.1128/jb.177.16.4757-4764.1995

Heijstra, B. D., Leang, C., & Juminaga, A. (2017). Gas fermentation: cellular engineering possibilities and scale up. *Microbial Cell Factories, 16*, 60. doi:10.1186/s12934-017-0676-y

Hendrickson, E. L., Kaul, R., Zhou, Y., Bovee, D., Chapman, P., Chung, J., . . . Leigh, J. A. (2004). Complete genome sequence of the genetically tractable hydrogenotrophic methanogen Methanococcus maripaludis. *J Bacteriol, 186*(20), 6956-6969. doi:10.1128/jb.186.20.6956-6969.2004

Hendrickson, E. L., Kaul, R., Zhou, Y., Bovee, D., Chapman, P., Chung, J., . . . Leigh, J. A. (2004). Complete genome sequence of the genetically tractable hydrogenotrophic methanogen Methanococcus maripaludis. *Journal of Bacteriology, 186*(20), 6956-6969. doi:10.1128/JB.186.20.6956-6969.2004

Henstra, A. M., Sipma, J., Rinzema, A., & Stams, A. J. M. (2007). Microbiology of synthesis gas fermentation for biofuel production. *Current Opinion in Biotechnology, 18*(3), 200-206. doi:https://doi.org/10.1016/j.copbio.2007.03.008

Herendeen, S. L., VanBogelen, R. A., & Neidhardt, F. C. (1979). Levels of major proteins of Escherichia coli during growth at different temperatures. *Journal of Bacteriology, 139*(1), 185. Retrieved from http://jb.asm.org/content/139/1/185.abstract

Hille, R., Hall, J., & Basu, P. (2014). The mononuclear molybdenum enzymes. *Chemical Reviews, 114*(7), 3963-4038. doi:10.1021/cr400443z

Hille, R., Schulzke, C., & Kirk, M. L. (2017). Molybdenum and Tungsten Enzymes. In (pp. P001-348): The Royal Society of Chemistry.

Hilton, J. C., & Rajagopalan, K. V. (1996). Identification of the Molybdenum Cofactor of Dimethyl Sulfoxide Reductase fromRhodobacter sphaeroidesf. sp.denitrificansas Bis(molybdopterin guanine dinucleotide)molybdenum. *Archives of Biochemistry and Biophysics, 325*(1), 139-143. doi:https://doi.org/10.1006/abbi.1996.0017

Hong, H., Kim, S.-J., Min, U.-G., Lee, Y.-J., Kim, S.-G., Jung, M.-Y., . . . Rhee, S.-K. (2015). Geosporobacter ferrireducens sp. nov., an anaerobic iron-reducing bacterium isolated from an oil-contaminated site. *Antonie van Leeuwenhoek, 107*(4), 971-977. doi:10.1007/s10482-015-0389-3

Hover, B. M., Lilla, E. A., & Yokoyama, K. (2015). Mechanistic Investigation of cPMP Synthase in Molybdenum Cofactor Biosynthesis Using an Uncleavable Substrate Analogue. *Biochemistry, 54*(49), 7229-7236. doi:10.1021/acs.biochem.5b00857

Hover, B. M., Loksztejn, A., Ribeiro, A. A., & Yokoyama, K. (2013). Identification of a Cyclic Nucleotide as a Cryptic Intermediate in Molybdenum Cofactor Biosynthesis. *Journal of the American Chemical Society, 135*(18), 7019-7032. doi:10.1021/ja401781t

Hu, Y., Faham, S., Roy, R., Adams, M. W. W., & Rees, D. C. (1999). Formaldehyde ferredoxin oxidoreductase from Pyrococcus furiosus: the 1.85 Å resolution crystal structure and its mechanistic implications11Edited by I. A. Wilson. *Journal of Molecular Biology, 286*(3), 899-914. doi:https://doi.org/10.1006/jmbi.1998.2488

Huber, C., Caldeira, J., Jongejan, J. A., & Simon, H. (1994). Further characterization of two different, reversible aldehyde oxidoreductases from Clostridium formicoaceticum, one containing tungsten and the other molybdenum. *Archives of Microbiology, 162*(5), 303-309. doi:10.1007/BF00263776

Huber, C., Skopan, H., Feicht, R., White, H., & Simon, H. (1995). Pterin cofactor, substrate specificity, and observations on the kinetics of the reversible tungsten-containing aldehyde oxidoreductase fromClostridium thermoaceticum. *Archives of Microbiology, 164*(2), 110-118. doi:10.1007/BF02525316

Humphreys, C. M., McLean, S., Schatschneider, S., Millat, T., Henstra, A. M., Annan, F. J., . . . Minton, N. P. (2015). Whole genome sequence and manual annotation of Clostridium autoethanogenum, an industrially relevant bacterium. *BMC genomics, 16*, 1085-1085. doi:10.1186/s12864-015-2287-5

Imlay, J. A. (2006). Iron-sulphur clusters and the problem with oxygen. *Molecular Microbiology, 59*(4), 1073-1082. doi:https://doi.org/10.1111/j.1365-2958.2006.05028.x

Imperial, J., Hadi, M., & Amy, N. K. (1998). Molybdate binding by ModA, the periplasmic component of the Escherichia coli mod molybdate transport system. *Biochimica et Biophysica Acta (BBA) - Biomembranes, 1370*(2), 337-346. doi:https://doi.org/10.1016/S0005-2736(98)00003-0

Iobbi-Nivol, C., & Leimkühler, S. (2013). Molybdenum enzymes, their maturation and molybdenum cofactor biosynthesis in Escherichia coli. *Biochimica et Biophysica Acta (BBA) - Bioenergetics, 1827*(8), 1086-1101. doi:https://doi.org/10.1016/j.bbabio.2012.11.007

IPCC. (2021a). *Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change [Masson-Delmotte, V., P. Zhai, A. Pirani, S.L. Connors, C. Péan, S. Berger, N. Caud, Y. Chen, L. Goldfarb, M.I. Gomis, M. Huang, K. Leitzell, E. Lonnoy, J.B.R. Matthews, T.K. Maycock, T. Waterfield, O. Yelekçi, R. Yu, and B. Zhou (eds.)].* Retrieved from https://www.ipcc.ch/report/ar6/wg1/#FullReport

IPCC. (2021b). Summary for Policymakers. In *Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change [Masson-Delmotte, V., P. Zhai, A. Pirani, S. L. Connors, C. Péan, S. Berger, N. Caud, Y. Chen, L. Goldfarb, M. I. Gomis, M. Huang, K. Leitzell, E. Lonnoy, J.B.R. Matthews, T. K. Maycock, T. Waterfield, O. Yelekçi, R. Yu and B. Zhou (eds.)]*: Cambridge University Press, In Press.

Iuchi, S., & Lin, E. C. (1988). arcA (dye), a global regulatory gene in Escherichia coli mediating repression of enzymes in aerobic pathways. *Proceedings of the National Academy of Sciences, 85*(6), 1888-1892. doi:10.1073/pnas.85.6.1888

Iuchi, S., & Lin, E. C. (1992a). Mutational analysis of signal transduction by ArcB, a membrane sensor protein responsible for anaerobic repression of operons involved in the central aerobic pathways in Escherichia coli.

Journal of Bacteriology, 174(12), 3972-3980. doi:doi:10.1128/jb.174.12.3972-3980.1992

Iuchi, S., & Lin, E. C. (1992b). Purification and phosphorylation of the Arc regulatory components of Escherichia coli. *Journal of Bacteriology, 174*(17), 5617-5623. doi:doi:10.1128/jb.174.17.5617-5623.1992

Jang, Y.-S., Lee Jin, Y., Lee, J., Park Jin, H., Im Jung, A., Eom, M.-H., . . . Lovley Derek, R. (2012). Enhanced Butanol Production Obtained by Reinforcing the Direct Butanol-Forming Route in Clostridium acetobutylicum. *MBio, 3*(5), e00314-00312. doi:10.1128/mBio.00314-12

Jeoung, J.-H., Martins, B. M., & Dobbek, H. (2019). Carbon Monoxide Dehydrogenases. In Y. Hu (Ed.), *Metalloproteins: Methods and Protocols* (pp. 37-54). New York, NY: Springer New York.

Johnson, D. C., Dean, D. R., Smith, A. D., & Johnson, M. K. (2005). STRUCTURE, FUNCTION, AND FORMATION OF BIOLOGICAL IRON-SULFUR CLUSTERS. *Annual Review of Biochemistry, 74*(1), 247-281. doi:10.1146/annurev.biochem.74.082803.133518

Johnson, J. L., Bastian, N. R., & Rajagopalan, K. V. (1990). Molybdopterin guanine dinucleotide: a modified form of molybdopterin identified in the molybdenum cofactor of dimethyl sulfoxide reductase from Rhodobacter sphaeroides forma specialis denitrificans. *Proceedings of the National Academy of Sciences of the United States of America, 87*(8), 3190-3194. Retrieved from http://www.ncbi.nlm.nih.gov/pmc/articles/PMC53861/

Johnson, J. L., Hainline, B. E., Rajagopalan, K. V., & Arison, B. H. (1984). The pterin component of the molybdenum cofactor. Structural characterization of two fluorescent derivatives. *J Biol Chem, 259*(9), 5414-5422.

Johnson, J. L., Indermaur, L. W., & Rajagopalan, K. V. (1991). Molybdenum cofactor biosynthesis in Escherichia coli. Requirement of the chlB gene product for the formation of molybdopterin guanine dinucleotide. *J Biol Chem, 266*(19), 12140-12145.

Johnson, J. L., & Rajagopalan, K. V. (1982). Structural and metabolic relationship between the molybdenum cofactor and urothione. *Proceedings of the National Academy of Sciences of the United States of America, 79*(22), 6856-6860. Retrieved from http://www.ncbi.nlm.nih.gov/pmc/articles/PMC347232/

Johnson, J. L., Rajagopalan, K. V., Mukund, S., & Adams, M. W. (1993). Identification of molybdopterin as the organic component of the tungsten cofactor in four enzymes from hyperthermophilic Archaea. *Journal of Biological Chemistry, 268*(7), 4848-4852. doi:https://doi.org/10.1016/S0021-9258(18)53474-8

Kakuta, Y., Horio, T., Takahashi, Y., & Fukuyama, K. (2001). Crystal Structure of Escherichia coli Fdx, an Adrenodoxin-Type Ferredoxin Involved in the Assembly of Iron−Sulfur Clusters. *Biochemistry, 40*(37), 11007-11012. doi:10.1021/bi010544t

Kanaujia, S. P., Jeyakanthan, J., Nakagawa, N., Balasubramaniam, S., Shinkai, A., Kuramitsu, S., . . . Sekar, K. (2010). Structures of apo and GTP-bound

molybdenum cofactor biosynthesis protein MoaC from Thermus thermophilus HB8. *Acta Crystallographica Section D: Biological Crystallography, 66*(7), 821-833.

Kataeva, I., Chang, J., Xu, H., Luan, C. H., Zhou, J., Uversky, V. N., . . . Wang, B. C. (2005). Improving solubility of Shewanella oneidensis MR-1 and Clostridium thermocellum JW-20 proteins expressed into Escherichia coli. In *Journal of Proteome Research* (Vol. 4, pp. 1942-1951).

Kaur, J., Singh, A., Panda, A. K., & Lal, R. (2021). Protocol for in-vitro purification and refolding of hexachlorocyclohexane degrading enzyme haloalkane dehalogenase LinB from inclusion bodies. *Enzyme and Microbial Technology, 146*, 109760. doi:https://doi.org/10.1016/j.enzmictec.2021.109760

Kietzin, A., & Adams, M. W. W. (1996). Tungsten in biological systems. *FEMS Microbiology Reviews, 18*(1), 5-63. doi:10.1016/0168-6445(95)00025-9

Kim, J. H., Bothe, J. R., Frederick, R. O., Holder, J. C., & Markley, J. L. (2014). Role of IscX in Iron–Sulfur Cluster Biogenesis in Escherichia coli. *Journal of the American Chemical Society, 136*(22), 7933-7942. doi:10.1021/ja501260h

Kim, J. H., Frederick, R. O., Reinen, N. M., Troupis, A. T., & Markley, J. L. (2013). [2Fe-2S]-Ferredoxin Binds Directly to Cysteine Desulfurase and Supplies an Electron for Iron–Sulfur Cluster Assembly but Is Displaced by the Scaffold Protein or Bacterial Frataxin. *Journal of the American Chemical Society, 135*(22), 8117-8120. doi:10.1021/ja401950a

King, G. M., & Weber, C. F. (2007). Distribution, diversity and ecology of aerobic CO-oxidizing bacteria. *Nature Reviews Microbiology, 5*(2), 107-118. doi:10.1038/nrmicro1595

King, P. W., Posewitz, M. C., Ghirardi, M. L., & Seibert, M. (2006). Functional Studies of [FeFe] Hydrogenase Maturation in an Escherichia coli Biosynthetic System. *Journal of Bacteriology, 188*(6), 2163-2172. doi:10.1128/JB.188.6.2163-2172.2006

Knoell, H. E., & Knappe, J. (1974). Escherichia coli Ferredoxin, an Iron-Sulfur Protein of the Adrenodoxin Type. *European Journal of Biochemistry, 50*(1), 245-252. doi:10.1111/j.1432-1033.1974.tb03893.x

Koepke, M., & Chen, M. Y. (2018). World Wide Patent No.: L. N. Inc.

Kolb, A., Busby, S., Buc, H., Garges, S., & Adhya, S. (1993). Transcriptional regulation by cAMP and its receptor protein. *Annual Review of Biochemistry, 62*(1), 749-797.

Köpke, M., Held, C., Hujer, S., Liesegang, H., Wiezer, A., Wollherr, A., . . . Dürre, P. (2010a). *Clostridium ljungdahlii* represents a microbial production platform based on syngas. *Proceedings of the National Academy of Sciences, 107*(29), 13087. doi:10.1073/pnas.1004716107

Köpke, M., Held, C., Hujer, S., Liesegang, H., Wiezer, A., Wollherr, A., . . . Dürre, P. (2010b). Clostridium ljungdahlii represents a microbial production platform based on syngas. *Proceedings of the National Academy of Sciences, 107*(29), 13087. doi:10.1073/pnas.1004716107

Kunichika, K., Nakamura, R., Fujishiro, T., & Takahashi, Y. (2021). The Structure of the Dimeric State of IscU Harboring Two Adjacent [2Fe–2S] Clusters Provides Mechanistic Insights into Cluster Conversion to [4Fe–4S]. *Biochemistry, 60*(20), 1569-1572. doi:10.1021/acs.biochem.1c00112

Kuper, J., Llamas, A., Hecht, H.-J., Mendel, R. R., & Schwarz, G. (2004). Structure of the molybdopterin-bound Cnx1G domain links molybdenum and copper metabolism. *Nature, 430*(7001), 803-806. doi:10.1038/nature02681

Labes, A., & Schönheit, P. (2001). Sugar utilization in the hyperthermophilic, sulfate-reducing archaeon Archaeoglobus fulgidus strain 7324: starch degradation to acetate and CO2 via a modified Embden-Meyerhof pathway and acetyl-CoA synthetase (ADP-forming). *Archives of Microbiology, 176*(5), 329-338. doi:10.1007/s002030100330

Lake, M. W., Temple, C. A., Rajagopalan, K. V., & Schindelin, H. (2000). The Crystal Structure of the Escherichia coliMobA Protein Provides Insight into Molybdopterin Guanine Dinucleotide Biosynthesis *. *Journal of Biological Chemistry, 275*(51), 40211-40217. doi:10.1074/jbc.M007406200

Lanzilotta, W. N., & Mathew, L. G. (To be published). *A Novel Function For WOR5 Revealed Through in Situ Crystallography and Biochemical Analysis*.

Lee, K.-C., Yeo, W.-S., & Roe, J.-H. (2008). Oxidant-Responsive Induction of the suf Operon, Encoding a Fe-S Assembly System, through Fur and IscR in Escherichia coli. *Journal of Bacteriology, 190*(24), 8244. doi:10.1128/JB.01161-08

Leimkühler, S. (2020). The biosynthesis of the molybdenum cofactors in Escherichia coli. *Environmental Microbiology, n/a*(n/a). doi:10.1111/1462-2920.15003

Leimkühler, S., Bühning, M., & Beilschmidt, L. (2017). Shared Sulfur Mobilization Routes for tRNA Thiolation and Molybdenum Cofactor Biosynthesis in Prokaryotes and Eukaryotes. *Biomolecules, 7*(1), 5. Retrieved from https://www.mdpi.com/2218-273X/7/1/5

Leimkühler, S., Wuebbens, M. M., & Rajagopalan, K. V. (2001). Characterization of Escherichia coli MoeB and Its Involvement in the Activation of Molybdopterin Synthase for the Biosynthesis of the Molybdenum Cofactor *. *Journal of Biological Chemistry, 276*(37), 34695-34701. doi:10.1074/jbc.M102787200

Leimkühler, S., Wuebbens, M. M., & Rajagopalan, K. V. (2011). The history of the discovery of the molybdenum cofactor and novel aspects of its biosynthesis in bacteria. *Coordination Chemistry Reviews, 255*(9), 1129-1144. doi:https://doi.org/10.1016/j.ccr.2010.12.003

Liao, R.-Z., Yu, J.-G., & Himo, F. (2011). Tungsten-dependent formaldehyde ferredoxin oxidoreductase: Reaction mechanism from quantum chemical calculations. *Journal of inorganic biochemistry, 105*(7), 927-936. doi:https://doi.org/10.1016/j.jinorgbio.2011.03.020

Liew, F., Henstra, A. M., Köpke, M., Winzer, K., Simpson, S. D., & Minton, N. P. (2017). Metabolic engineering of Clostridium autoethanogenum for

selective alcohol production. *Metabolic Engineering, 40*, 104-114. doi:https://doi.org/10.1016/j.ymben.2017.01.007

Liew, F., Henstra, A. M., Winzer, K., Köpke, M., Simpson, S. D., & Minton, N. P. (2016). Insights into CO₂ Fixation Pathway of Clostridium autoethanogenum by Targeted Mutagenesis. *MBio, 7*(3), e00427-00416. doi:10.1128/mBio.00427-16

Lin, C.-W., McCabe, J. W., Russell, D. H., & Barondeau, D. P. (2020). Molecular Mechanism of ISC Iron–Sulfur Cluster Biogenesis Revealed by High-Resolution Native Mass Spectrometry. *Journal of the American Chemical Society, 142*(13), 6018-6029. doi:10.1021/jacs.9b11454

Lindahl, P. A. (2002). The Ni-Containing Carbon Monoxide Dehydrogenase Family: Light at the End of the Tunnel? *Biochemistry, 41*(7), 2097-2105. doi:10.1021/bi015932+

Liu, M. T. W., Wuebbens, M. M., Rajagopalan, K. V., & Schindelin, H. (2000). Crystal Structure of the Gephyrin-related Molybdenum Cofactor Biosynthesis Protein MogA from Escherichia coli *. *Journal of Biological Chemistry, 275*(3), 1814-1822. doi:10.1074/jbc.275.3.1814

Løvgreen, M. N., Martic, M., Windahl, M. S., Christensen, H. E. M., & Harris, P. (2011). Crystal structures of the all-cysteinyl-coordinated D14C variant of Pyrococcus furiosus ferredoxin: [4Fe–4S] ↔ [3Fe–4S] cluster conversion. *JBIC Journal of Biological Inorganic Chemistry, 16*(5), 763-775. doi:10.1007/s00775-011-0778-7

Luo, S., Zhou, Y., & Yi, C. (2012). Syngas production by catalytic steam gasification of municipal solid waste in fixed-bed reactor. *Energy, 44*(1), 391-395. doi:10.1016/j.energy.2012.06.016

Macedo-Ribeiro, S., Martins, B. M., Barbosa Pereira, P., Buse, G., Huber, R., & Soulimane, T. (2001). New insights into the thermostability of bacterial ferredoxins: high-resolution crystal structure of the seven-iron ferredoxin from Thermus thermophilus. *JBIC Journal of Biological Inorganic Chemistry, 6*(7), 663-674. doi:10.1007/s007750100243

Magalon, A., Fedor, J. G., Walburger, A., & Weiner, J. H. (2011). Molybdenum enzymes in bacteria and their maturation. *Coordination Chemistry Reviews, 255*(9-10), 1159-1178. doi:10.1016/j.ccr.2010.12.031

Magalon, A., Frixon, C., Pommier, J., Giordano, G., & Blasco, F. (2002). In vivo interactions between gene products involved in the final stages of molybdenum cofactor biosynthesis in Escherichia coli. *J Biol Chem, 277*(50), 48199-48204. doi:10.1074/jbc.M205806200

Magalon, A., & Mendel, R. R. (2015). Biosynthesis and Insertion of the Molybdenum Cofactor. *EcoSal Plus, 6*(2), 0006-2013.

Maier, T., Binder, U., & Böck, A. (1996). Analysis of the hydA locus of Escherichia coli: two genes (hydN and hypF) involved in formate and hydrogen metabolism. *Archives of Microbiology, 165*(5), 333-341.

Makdessi, K., Andreesen, J. R., & Pich, A. (2001). Tungstate Uptake by a Highly Specific ABC Transporter inEubacterium acidaminophilum *. *Journal of Biological Chemistry, 276*(27), 24557-24564. doi:10.1074/jbc.M101293200

Mandrand-Berthelot, M.-A., Wee, M. Y. K., & Haddock, B. A. (1978). An improved method for the identification and characterization of mutants of Escherichia coli deficient in formate dehydrogenase activity. *FEMS Microbiology Letters, 4*(1), 37-40. doi:10.1111/j.1574-6968.1978.tb02841.x

Marcellin, E., Behrendorff, J. B., Nagaraju, S., DeTissera, S., Segovia, S., Palfreyman, R. W., . . . Speight, R. (2016). Low carbon fuels and commodity chemicals from waste gases–systematic approach to understand energy metabolism in a model acetogen. *Green Chemistry, 18*(10), 3020-3028.

Markley, J. L., Kim, J. H., Dai, Z., Bothe, J. R., Cai, K., Frederick, R. O., & Tonelli, M. (2013). Metamorphic protein IscU alternates conformations in the course of its role as the scaffold protein for iron–sulfur cluster biosynthesis and delivery. *FEBS Letters, 587*(8), 1172-1179. doi:10.1016/j.febslet.2013.01.003

Martín-Tornero, E., Gómez, D. G., Durán-Merás, I., & Espinosa-Mansilla, A. (2016). Development of an HPLC-MS method for the determination of natural pteridines in tomato samples. *Analytical Methods, 8*(34), 6404-6414. doi:10.1039/C6AY01519K

McLuskey, K., Harrison, J. A., Schüttelkopf, A. W., Boxer, D. H., & Hunter, W. N. (2003). Insight into the role of Escherichia coli MobB in molybdenum cofactor biosynthesis based on the high resolution crystal structure. *Journal of Biological Chemistry, 278*(26), 23706-23713. doi:10.1074/jbc.M301485200

McNicholas, P. M., Rech, S. A., & Gunsalus, R. P. (1997). Characterization of the ModE DNA-binding sites in the control regions of modABCD and moaABCDE of Escherichia coli. *Molecular Microbiology, 23*(3), 515-524. doi:https://doi.org/10.1046/j.1365-2958.1997.d01-1864.x

Meckenstock, R. U., Krieger, R., Ensign, S., Kroneck, P. M. H., & Schink, B. (2001). Acetylene hydratase of Pelobacter acetylenicus. *European Journal of Biochemistry, 264*(1), 176-182. doi:10.1046/j.1432-1327.1999.00600.x

Mendel, R. R., Hercher, T. W., Zupok, A., Hasnat, M. A., & Leimkühler, S. (2020). The Requirement of Inorganic Fe-S Clusters for the Biosynthesis of the Organometallic Molybdenum Cofactor. *Inorganics, 8*(7), 43. Retrieved from https://www.mdpi.com/2304-6740/8/7/43

Mendel, R. R., & Leimkühler, S. (2015). The biosynthesis of the molybdenum cofactors. *Journal of Biological Inorganic Chemistry, 20*(2), 337-347. doi:10.1007/s00775-014-1173-y

Messens, J., & Collet, J.-F. (2006). Pathways of disulfide bond formation in Escherichia coli. *The International Journal of Biochemistry & Cell Biology, 38*(7), 1050-1062. doi:https://doi.org/10.1016/j.biocel.2005.12.011

Meyer, O., Jacobitz, S., & Krüger, B. (1986). Biochemistry and physiology of aerobic carbon monoxide-utilizing bacteria. *FEMS Microbiology Reviews, 2*(3), 161-179. doi:10.1111/j.1574-6968.1986.tb01858.x

Meyer, O., & Schlegel, H. G. (1978). Reisolation of the carbon monoxide utilizing hydrogen bacterium Pseudomonas carboxydovorans (Kistner) comb. nov. *Archives of Microbiology, 118*(1), 35-43. doi:10.1007/BF00406071

Mihara, H., Abe, K., Inagaki, K., Takahata, M., Esaki, N., Nakano, R., . . . Tamura, T. (2008). Selenite Assimilation into Formate Dehydrogenase H Depends on Thioredoxin Reductase in Escherichia coli. *The Journal of Biochemistry, 143*(4), 467-473. doi:10.1093/jb/mvm247

Minton, N. P., Ehsaan, M., Humphreys, C. M., Little, G. T., Baker, J., Henstra, A. M., . . . Zhang, Y. (2016). A roadmap for gene system development in Clostridium. *Anaerobe, 41*, 104-112. doi:https://doi.org/10.1016/j.anaerobe.2016.05.011

Mock, J., Zheng, Y., Mueller Alexander, P., Ly, S., Tran, L., Segovia, S., . . . Metcalf, W. W. (2015). Energy Conservation Associated with Ethanol Formation from H2 and CO2 in Clostridium autoethanogenum Involving Electron Bifurcation. *Journal of Bacteriology, 197*(18), 2965-2980. doi:10.1128/JB.00399-15

Molino, A., Chianese, S., & Musmarra, D. (2016). Biomass gasification technology: The state of the art overview. *Journal of Energy Chemistry, 25*(1), 10-25. doi:https://doi.org/10.1016/j.jechem.2015.11.005

Mukund, S., & Adams, M. W. (1991). The novel tungsten-iron-sulfur protein of the hyperthermophilic archaebacterium, Pyrococcus furiosus, is an aldehyde ferredoxin oxidoreductase. Evidence for its participation in a unique glycolytic pathway. *Journal of Biological Chemistry, 266*(22), 14208-14216. doi:https://doi.org/10.1016/S0021-9258(18)98669-2

Mukund, S., & Adams, M. W. (1995). Glyceraldehyde-3-phosphate ferredoxin oxidoreductase, a novel tungsten-containing enzyme with a potential glycolytic role in the hyperthermophilic archaeon Pyrococcus furiosus. *J Biol Chem, 270*(15), 8389-8392.

Mukund, S., & Adams, M. W. (1996). Molybdenum and vanadium do not replace tungsten in the catalytically active forms of the three tungstoenzymes in the hyperthermophilic archaeon Pyrococcus furiosus. *Journal of Bacteriology, 178*(1), 163-167. doi:10.1128/jb.178.1.163-167.1996

Myers, K. S., Yan, H., Ong, I. M., Chung, D., Liang, K., Tran, F., . . . Kiley, P. J. (2013). Genome-scale Analysis of Escherichia coli FNR Reveals Complex Features of Transcription Factor Binding. *PLOS Genetics, 9*(6), e1003565. doi:10.1371/journal.pgen.1003565

Napora-Wijata, K., Strohmeier, G. A., & Winkler, M. (2014). Biocatalytic reduction of carboxylic acids. *Biotechnology Journal, 9*(6), 822-843. doi:https://doi.org/10.1002/biot.201400012

Neumann, M., & Leimkühler, S. (2008). Heavy metal ions inhibit molybdoenzyme activity by binding to the dithiolene moiety of molybdopterin in Escherichia coli. *Febs j, 275*(22), 5678-5689. doi:10.1111/j.1742-4658.2008.06694.x

Neumann, M., Mittelstadt, G., Iobbi-Nivol, C., Saggu, M., Lendzian, F., Hildebrandt, P., & Leimkuhler, S. (2009). A periplasmic aldehyde oxidoreductase represents the first molybdopterin cytosine

dinucleotide cofactor containing molybdo-flavoenzyme from Escherichia coli. *Febs j, 276*(10), 2762-2774. doi:10.1111/j.1742-4658.2009.07000.x

Neumann, M., Mittelstadt, G., Seduk, F., Iobbi-Nivol, C., & Leimkuhler, S. (2009). MocA is a specific cytidylyltransferase involved in molybdopterin cytosine dinucleotide biosynthesis in Escherichia coli. *J Biol Chem, 284*(33), 21891-21898. doi:10.1074/jbc.M109.008565

Neumann, M., Stocklein, W., & Leimkuhler, S. (2007). Transfer of the molybdenum cofactor synthesized by Rhodobacter capsulatus MoeA to XdhC and MobA. *J Biol Chem, 282*(39), 28493-28500. doi:10.1074/jbc.M704020200

Neumann, M., Stöcklein, W., Walburger, A., Magalon, A., & Leimkühler, S. (2007). Identification of a Rhodobacter capsulatus l-Cysteine Desulfurase That Sulfurates the Molybdenum Cofactor When Bound to XdhC and before Its Insertion into Xanthine Dehydrogenase. *Biochemistry, 46*(33), 9586-9595. doi:10.1021/bi700630p

Ng, C. Y., Farasat, I., Maranas, C. D., & Salis, H. M. (2015). Rational design of a synthetic Entner–Doudoroff pathway for improved and controllable NADPH regeneration. *Metabolic Engineering, 29*, 86-96. doi:https://doi.org/10.1016/j.ymben.2015.03.001

Ni, D., Xu, P., & Gallagher, S. (2017). Immunoblotting and Immunodetection. *Current Protocols in Protein Science, 88*(1), 10.10.11-10.10.37. doi:10.1002/cpps.32

Nichols, J., & Rajagopalan, K. V. (2002). *Escherichia coli* MoeA and MogA: Function in metal incorporation step of molybdenum cofactor biosynthesis *Journal of Biological Chemistry, 277*(28), 24995-25000. doi:10.1074/jbc.M203238200

Nichols, J. D., & Rajagopalan, K. V. (2005). In vitro molybdenum ligation to molybdopterin using purified components. In *Journal of Biological Chemistry* (Vol. 280, pp. 7817-7822).

Nichols, J. D., Xiang, S., Schindelin, H., & Rajagopalan, K. V. (2007). Mutational analysis of Escherichia coli MoeA: Two functional activities map to the active site cleft. *Biochemistry, 46*(1), 78-86. doi:10.1021/bi061551q

Nissen, L. S., & Basen, M. (2019). The emerging role of aldehyde:ferredoxin oxidoreductases in microbially-catalyzed alcohol production. In *Journal of Biotechnology* (Vol. 306, pp. 105-117): Elsevier.

Nölling, J., Breton, G., Omelchenko Marina, V., Makarova Kira, S., Zeng, Q., Gibson, R., . . . Smith Douglas, R. (2001). Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium Clostridium acetobutylicum. *Journal of Bacteriology, 183*(16), 4823-4838. doi:10.1128/JB.183.16.4823-4838.2001

Norman, R. O. J., Millat, T., Winzer, K., Minton, N. P., & Hodgman, C. (2018). Progress towards platform chemical production using Clostridium autoethanogenum. *Biochemical Society Transactions, 46*(3), 523-535. doi:10.1042/BST20170259

Nzuza, N., Padayachee, T., Chen, W., Gront, D., Nelson, D. R., & Syed, K. (2021). Diversification of Ferredoxins across Living Organisms. *Current Issues in*

Molecular Biology, 43(3), 1374-1390. Retrieved from https://www.mdpi.com/1467-3045/43/3/98

Oelgeschläger, E., & Rother, M. (2008). Carbon monoxide-dependent energy metabolism in anaerobic bacteria and archaea. *Archives of Microbiology, 190*(3), 257-269. doi:10.1007/s00203-008-0382-6

Palmer, T., Santini, C.-L., Iobbi-Nivol, C., Eaves, D. J., Boxer, D. H., & Giordano, G. (1996). Involvement of the narJ and mob gene products in distinct steps in the biosynthesis of the molybdoenzyme nitrate reductase in Escherichia coli. *Molecular Microbiology, 20*(4), 875-884. doi:10.1111/j.1365-2958.1996.tb02525.x

Park, M. O., Mizutani, T., & Jones, P. R. (2007). Glyceraldehyde-3-phosphate ferredoxin oxidoreductase from Methanococcus maripaludis. *J Bacteriol, 189*(20), 7281-7289.

Partridge, J. D., Browning, D. F., Xu, M., Newnham, L. J., Scott, C., Roberts, R. E., . . . Green, J. (2008). Characterization of the Escherichia coli K-12 ydhYVWXUT operon: regulation by FNR, NarL and NarP. *Microbiology, 154*(2), 608-618. doi:https://doi.org/10.1099/mic.0.2007/012146-0

Paschos, A., Bauer, A., Zimmermann, A., Zehelein, E., & Bock, A. (2002). HypF, a carbamoyl phosphate-converting enzyme involved in [NiFe] hydrogenase maturation. *J Biol Chem, 277*(51), 49945-49951. doi:10.1074/jbc.M204601200

Patterson-Fortin, L. M., Vakulskas, C. A., Yakhnin, H., Babitzke, P., & Romeo, T. (2013). Dual Posttranscriptional Regulation via a Cofactor-Responsive mRNA Leader. *Journal of Molecular Biology, 425*(19), 3662-3677. doi:https://doi.org/10.1016/j.jmb.2012.12.010

Pérard, J., & Ollagnier de Choudens, S. (2018). Iron–sulfur clusters biogenesis by the SUF machinery: close to the molecular mechanism understanding. *JBIC Journal of Biological Inorganic Chemistry, 23*(4), 581-596. doi:10.1007/s00775-017-1527-3

Petkun, S., Shi, R., Li, Y., Asinas, A., Munger, C., Zhang, L., . . . Cygler, M. (2011). Structure of Hydrogenase Maturation Protein HypF with Reaction Intermediates Shows Two Active Sites. *Structure, 19*(12), 1773-1783. doi:10.1016/j.str.2011.09.023

Pitterle, D. M., & Rajagopalan, K. V. (1989). Two proteins encoded at the chlA locus constitute the converting factor of Escherichia coli chlA1. *Journal of Bacteriology, 171*(6), 3373-3378. doi:10.1128/jb.171.6.3373-3378.1989

Pitterle, D. M., & Rajagopalan, K. V. (1993). The biosynthesis of molybdopterin in Escherichia coli. Purification and characterization of the converting factor. *Journal of Biological Chemistry, 268*(18), 13499-13505. doi:https://doi.org/10.1016/S0021-9258(19)38677-6

Py, B., & Barras, F. (2010). Building Fe–S proteins: bacterial strategies. *Nature Reviews Microbiology, 8*(6), 436-446. doi:10.1038/nrmicro2356

Py, B., Moreau, P. L., & Barras, F. (2011). Fe–S clusters, fragile sentinels of the cell. *Current Opinion in Microbiology, 14*(2), 218-223. doi:https://doi.org/10.1016/j.mib.2011.01.004

Rabin, R. S., & Stewart, V. (1993). Dual response regulators (NarL and NarP) interact with dual sensors (NarX and NarQ) to control nitrate- and nitrite-regulated gene expression in Escherichia coli K-12. *Journal of Bacteriology, 175*(11), 3259-3268. doi:doi:10.1128/jb.175.11.3259-3268.1993

Ragsdale, S. W. (2008). Enzymology of the Wood–Ljungdahl Pathway of Acetogenesis. *Annals of the New York Academy of Sciences, 1125*(1), 129-136. doi:https://doi.org/10.1196/annals.1419.015

Ragsdale, S. W., & Pierce, E. (2008). Acetogenesis and the Wood–Ljungdahl pathway of CO2 fixation. *Biochimica et Biophysica Acta (BBA) - Proteins and Proteomics, 1784*(12), 1873-1898. doi:https://doi.org/10.1016/j.bbapap.2008.08.012

Rajagopalan, K. (1996). Biosynthesis of the molybdenum cofactor. *Escherichia coli and Salmonella typhimurium*, 674-679.

Rajagopalan, K., & Johnson, J. L. (1992). The pterin molybdenum cofactors. *Journal of Biological Chemistry, 267*(15), 10199-10202.

Rauh, D., Graentzdoerffer, A., Granderath, K., Andreesen, J. R., & Pich, A. (2004). Tungsten-containing aldehyde oxidoreductase of Eubacterium acidaminophilum. *European Journal of Biochemistry, 271*(1), 212-219. doi:https://doi.org/10.1111/j.1432-1033.2004.03922.x

Ray, N., Oates, J., Turner, R. J., & Robinson, C. (2003). DmsD is required for the biogenesis of DMSO reductase in Escherichia coli but not for the interaction of the DmsA signal peptide with the Tat apparatus. *FEBS Letters, 534*(1), 156-160. doi:https://doi.org/10.1016/S0014-5793(02)03839-5

Regulski, E. E., Moy, R. H., Weinberg, Z., Barrick, J. E., Yao, Z., Ruzzo, W. L., & Breaker, R. R. (2008). A widespread riboswitch candidate that controls bacterial genes involved in molybdenum cofactor and tungsten cofactor metabolism. *Molecular Microbiology, 68*(4), 918-932. doi:https://doi.org/10.1111/j.1365-2958.2008.06208.x

Reher, M., Gebhard, S., & Schönheit, P. (2007). Glyceraldehyde-3-phosphate ferredoxin oxidoreductase (GAPOR) and nonphosphorylating glyceraldehyde-3-phosphate dehydrogenase (GAPN), key enzymes of the respective modified Embden-Meyerhof pathways in the hyperthermophilic crenarchaeota Pyrobaculum aerophi. *FEMS Microbiology Letters, 273*(2), 196-205. doi:10.1111/j.1574-6968.2007.00787.x

Reis, A. C., & Salis, H. M. (2020). An Automated Model Test System for Systematic Development and Improvement of Gene Expression Models. *ACS Synthetic Biology, 9*(11), 3145-3156. doi:10.1021/acssynbio.0c00394

Reschke, S., Duffus, B. R., Schrapers, P., Mebs, S., Teutloff, C., Dau, H., . . . Leimkühler, S. (2019). Identification of YdhV as the First Molybdoenzyme Binding a Bis-Mo-MPT Cofactor in Escherichia coli. *Biochemistry, 58*(17), 2228-2242. doi:10.1021/acs.biochem.9b00078

Reschke, S., Sigfridsson, K. G. V., Kaufmann, P., Leidel, N., Horn, S., Gast, K., . . . Leimkühler, S. (2013). Identification of a bis-molybdopterin

intermediate in molybdenum cofactor biosynthesis in escherichia coli. *Journal of Biological Chemistry, 288*(41), 29736-29745. doi:10.1074/jbc.M113.497453

Ribbe, M. W., Hu, Y., Hodgson, K. O., & Hedman, B. (2014). Biosynthesis of Nitrogenase Metalloclusters. *Chemical Reviews, 114*(8), 4063-4080. doi:10.1021/cr400463x

Roche, B., Huguenot, A., Barras, F., & Py, B. (2015). The iron-binding CyaY and IscX proteins assist the ISC-catalyzed Fe-S biogenesis in Escherichia coli. *Molecular Microbiology, 95*(4), 605-623. doi:10.1111/mmi.12888

Rosano, G. L., & Ceccarelli, E. A. (2014). Recombinant protein expression in Escherichia coli: Advances and challenges. *Frontiers in Microbiology, 5*(APR), 1-17. doi:10.3389/fmicb.2014.00172

Roy, R., & Adams Michael, W. W. (2002). Characterization of a Fourth Tungsten-Containing Enzyme from the Hyperthermophilic Archaeon Pyrococcus furiosus. *Journal of Bacteriology, 184*(24), 6952-6956. doi:10.1128/JB.184.24.6952-6956.2002

Roy, R., Dhawan, I. K., Johnson, M. K., Rees, D. C., & Adams, M. W. W. (2011). Aldehyde Ferredoxin Oxidoreductase. In *Encyclopedia of Inorganic and Bioinorganic Chemistry* (pp. 1-13).

Roy, R., Menon, A. L., & Adams, M. W. W. (2001). [11] Aldehyde Oxidoreductases from Pyrococcus furiosus. In *Methods in Enzymology* (Vol. 331, pp. 132-144): Academic Press.

Roy, R., Mukund, S., Dunn, D. M., Weiss, R., & Adams, M. W. W. (1999). Purification and Molecular Characterization of the Tungsten-Containing Formaldehyde Ferredoxin Oxidoreductase from the Hyperthermophilic Archaeon. In (Vol. 181, pp. 1171-1180).

Roy, R., Mukund, S., Schut, G. J., Dunn, D. M., Weiss, R., & Adams, M. W. (1999). Purification and molecular characterization of the tungsten-containing formaldehyde ferredoxin oxidoreductase from the hyperthermophilic archaeon Pyrococcus furiosus: the third of a putative five-member tungstoenzyme family. *J Bacteriol, 181*(4), 1171-1180.

Sabnis, N. A., Yang, H., & Romeo, T. (1995). Pleiotropic Regulation of Central Carbohydrate Metabolism in Escherichia coli via the Gene csrA(*). *Journal of Biological Chemistry, 270*(49), 29096-29104. doi:10.1074/jbc.270.49.29096

Sapra, R., Bagramyan, K., & Adams, M. W. W. (2003). A simple energy-conserving system: Proton reduction coupled to proton translocation. *Proceedings of the National Academy of Sciences, 100*(13), 7545. doi:10.1073/pnas.1331436100

Schmitz, J., Wuebbens, M. M., Rajagopalan, K. V., & Leimkühler, S. (2007). Role of the C-Terminal Gly-Gly Motif of Escherichia Coli MoaD, a Molybdenum Cofactor Biosynthesis Protein with a Ubiquitin Fold. *Biochemistry, 46*(3), 909-916. doi:10.1021/bi062011w

Schrag, J. D., Huang, W., Sivaraman, J., Smith, C., Plamondon, J., Larocque, R., . . . Cygler, M. (2001). The crystal structure of Escherichia coli MoeA, a protein from the molybdopterin synthesis pathway11Edited by D. Rees.

Journal of Molecular Biology, 310(2), 419-431. doi:https://doi.org/10.1006/jmbi.2001.4771

Schut, G. J., Thorgersen, M. P., Poole, F. L., Haja, D. K., Putumbaka, S., & Adams, M. W. W. (2021). Tungsten enzymes play a role in detoxifying food and antimicrobial aldehydes in the human gut microbiome. *Proceedings of the National Academy of Sciences, 118*(43), e2109008118. doi:10.1073/pnas.2109008118

Schwartz, C. J., Giel, J. L., Patschkowski, T., Luther, C., Ruzicka, F. J., Beinert, H., & Kiley, P. J. (2001). IscR, an Fe-S cluster-containing transcription factor, represses expression of Escherichia coli genes encoding Fe-S cluster assembly proteins. *Proceedings of the National Academy of Sciences, 98*(26), 14895. doi:10.1073/pnas.251550898

Schwarz, G., Mendel, R. R., & Ribbe, M. W. (2009). Molybdenum cofactors, enzymes and pathways. *Nature, 460*(7257), 839-847. doi:10.1038/nature08302

Scott, I. M., Rubinstein, G. M., Poole, F. L., 2nd, Lipscomb, G. L., Schut, G. J., Williams-Rhaesa, A. M., . . . Adams, M. W. W. (2019). The thermophilic biomass-degrading bacterium Caldicellulosiruptor bescii utilizes two enzymes to oxidize glyceraldehyde 3-phosphate during glycolysis. *J Biol Chem, 294*(25), 9995-10005. doi:10.1074/jbc.RA118.007120

Seelmann, C. S., Willistein, M., Heider, J., & Boll, M. (2020). Tungstoenzymes: Occurrence, Catalytic Diversity and Cofactor Synthesis. *Inorganics, 8*(8). doi:10.3390/inorganics8080044

Sehnal, D., Bittrich, S., Deshpande, M., Svobodová, R., Berka, K., Bazgier, V., . . . Rose, A. S. (2021). Mol* Viewer: modern web app for 3D visualization and analysis of large biomolecular structures. *Nucleic Acids Research, 49*(W1), W431-W437. doi:10.1093/nar/gkab314

Sevcenco, A. M., Bevers, L. E., Pinkse, M. W. H., Krijger, G. C., Wolterbeek, H. T., Verhaert, P. D. E. M., . . . Hagedoorn, P. L. (2010). Molybdenum incorporation in tungsten aldehyde oxidoreductase enzymes from Pyrococcus furiosus. *Journal of Bacteriology, 192*(16), 4143-4152. doi:10.1128/JB.00270-10

Shanmugam, K. T., Stewart, V., Gunsalus, R. P., Boxer, D. H., Cole, J. A., Chippaux, M., . . . Rajagopalan, K. V. (1992). Proposed nomenclature for the genes involved in molybdenum metabolism in Escherichia coli and Saimonella typhimurium. *Molecular Microbiology, 6*(22), 3452-3454. doi:https://doi.org/10.1111/j.1365-2958.1992.tb02215.x

Shindell, D. T., Faluvegi, G., Stevenson, D. S., Krol, M. C., Emmons, L. K., Lamarque, J. F., . . . Schultz, M. G. (2006). Multimodel simulations of carbon monoxide: Comparison with observations and projected near-future changes. *Journal of Geophysical Research: Atmospheres, 111*(D19).

Sievers, F., Wilm, A., Dineen, D., Gibson, T. J., Karplus, K., Li, W., . . . Higgins, D. G. (2011). Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Molecular Systems Biology, 7*(1), 539. doi:https://doi.org/10.1038/msb.2011.75

Sousa, F. L., Thiergart, T., Landan, G., Nelson-Sathi, S., Pereira, I. A. C., Allen, J. F., . . . Martin, W. F. (2013). Early bioenergetic evolution. *Philosophical Transactions of the Royal Society B: Biological Sciences, 368*(1622), 20130088. doi:doi:10.1098/rstb.2013.0088

Stern, N., & Stern, N. H. (2007). *The economics of climate change: the Stern review*: cambridge University press.

Straub, C. T., Schut, G., Otten, J. K., Keller, L. M., Adams, M. W. W., & Kelly, R. M. (2020). Modification of the glycolytic pathway in Pyrococcus furiosus and the implications for metabolic engineering. In *Extremophiles*.

Sun, X., Atiyeh, H. K., Huhnke, R. L., & Tanner, R. S. (2019). Syngas fermentation process development for production of biofuels and chemicals: A review. *Bioresource Technology Reports, 7*, 100279. doi:https://doi.org/10.1016/j.biteb.2019.100279

Szopa, S., Naik, V., Adhikary, B., Artaxo, P., Berntsen, T., Collins, W. D., . . . Zanis, P. (2021). Short-Lived Climate Forcers. In *Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change [Masson-Delmotte, V., P. Zhai, A. Pirani, S.L. Connors, C. Péan, S. Berger, N. Caud, Y. Chen, L. Goldfarb, M.I. Gomis, M. Huang, K. Leitzell, E. Lonnoy, J.B.R. Matthews, T.K. Maycock, T. Waterfield, O. Yelekçi, R.Yu, and B. Zhou (eds.)].* Cambridge University Press. In Press.

Tanner, R. S., Miller, L. M., & Yang, D. (1993). Clostridium ljungdahlii sp. nov., an acetogenic species in clostridial rRNA homology group I. *International Journal of Systematic and Evolutionary Microbiology, 43*(2), 232-236.

Temple, C. A., & Rajagopalan, K. V. (2000). Mechanism of assembly of the Bis(Molybdopterin guanine dinucleotide)molybdenum cofactor in Rhodobacter sphaeroides dimethyl sulfoxide reductase. *J Biol Chem, 275*(51), 40202-40210. doi:10.1074/jbc.M007407200

Thomason, L. C., Costantino, N., & Court, D. L. (2007). E. coli Genome Manipulation by P1 Transduction. *Current Protocols in Molecular Biology, 79*(1), 1.17.11-11.17.18. doi:10.1002/0471142727.mb0117s79

Thomé, R., Gust, A., Toci, R., Mendel, R., Bittner, F., Magalon, A., & Walburger, A. (2012). A Sulfurtransferase Is Essential for Activity of Formate Dehydrogenases in Escherichia coli*. *Journal of Biological Chemistry, 287*(7), 4671-4678. doi:10.1074/jbc.M111.327122

Tokumoto, U., Kitamura, S., Fukuyama, K., & Takahashi, Y. (2004). Interchangeability and Distinct Properties of Bacterial Fe-S Cluster Assembly Systems: Functional Replacement of the iscand suf Operons in Escherichia coli with the nifSU-Like Operon from Helicobacter pylori. *The Journal of Biochemistry, 136*(2), 199-209. doi:10.1093/jb/mvh104

Tong, Y., Wuebbens, M. M., Rajagopalan, K. V., & Fitzgerald, M. C. (2005). Thermodynamic Analysis of Subunit Interactions in Escherichia coli Molybdopterin Synthase. *Biochemistry, 44*(7), 2595-2601. doi:10.1021/bi047762h

Trautwein, T., Krauss, F., Lottspeich, F., & Simon, H. (1994). The (2R)-hydroxycarboxylate-viologen-oxidoreductase from Proteus vulgaris is a molybdenum-containing iron-sulphur protein. *European Journal of Biochemistry, 222*(3), 1025-1032. doi:https://doi.org/10.1111/j.1432-1033.1994.tb18954.x

Tremblay, P.-L., Zhang, T., Dar Shabir, A., Leang, C., & Lovley Derek, R. (2012). The Rnf Complex of Clostridium ljungdahlii Is a Proton-Translocating Ferredoxin:NAD+ Oxidoreductase Essential for Autotrophic Growth. *MBio, 4*(1), e00406-00412. doi:10.1128/mBio.00406-12

Unciuleac, M.-C., Chandramouli, K., Naik, S., Mayer, S., Huynh, B. H., Johnson, M. K., & Dean, D. R. (2007). In Vitro Activation of Apo-Aconitase Using a [4Fe-4S] Cluster-Loaded Form of the IscU [Fe−S] Cluster Scaffolding Protein. *Biochemistry, 46*(23), 6812-6821. doi:10.1021/bi6026665

US EPA. (2021a). 2017 National Emissions Inventory (NEI) Data. Retrieved from https://www.epa.gov/air-emissions-inventories/2017-national-emissions-inventory-nei-data

US EPA. (2021b). 2017 National Emissions Inventory: January 2021 Updated Release, Technical Support Document. (January). Retrieved from https://www.epa.gov/sites/production/files/2021-02/documents/nei2017_tsd_full_jan2021.pdf

Van Der Oost, J., Schut, G., Kengen, S. W. M., Hagen, W. R., Thomm, M., & De Vos, W. M. (1998). The ferredoxin-dependent conversion of glyceraldehyde-3-phosphate in the hyperthermophilic archaeon Pyrococcus furiosus represents a novel site of glycolytic regulation. *Journal of Biological Chemistry, 273*(43), 28149-28154. doi:10.1074/jbc.273.43.28149

Vergnes, A., Gouffi-Belhabich, K., Blasco, F., Giordano, G., & Magalon, A. (2004). Involvement of the molybdenum cofactor biosynthetic machinery in the maturation of the Escherichia coli nitrate reductase A. *J Biol Chem, 279*(40), 41398-41403.

Wang, W., Huang, H., Tan, G., Si, F., Liu, M., Landry, A. P., . . . Ding, H. (2010). In vivo evidence for the iron-binding activity of an iron-sulfur cluster assembly protein IscA in Escherichia coli. *The Biochemical journal, 432*(3), 429-436. doi:10.1042/BJ20101507

Weinert, T., Huwiler, S. G., Kung, J. W., Weidenweber, S., Hellwig, P., Stärk, H.-J., . . . Boll, M. (2015). Structural basis of enzymatic benzene ring reduction. *Nature Chemical Biology, 11*(8), 586-591. doi:10.1038/nchembio.1849

Whitham Jason, M., Tirado-Acevedo, O., Chinn Mari, S., Pawlak Joel, J., Grunden Amy, M., & Parales, R. E. (2015). Metabolic Response of Clostridium ljungdahlii to Oxygen Exposure. *Applied and Environmental Microbiology, 81*(24), 8379-8391. doi:10.1128/AEM.02491-15

Wilcoxen, J., & Hille, R. (2013). The Hydrogenase Activity of the Molybdenum/Copper-containing Carbon Monoxide Dehydrogenase of Oligotropha carboxidovorans*. *Journal of Biological Chemistry, 288*(50), 36052-36060. doi:10.1074/jbc.M113.522441

Wilson, C. J., Zhan, H., Swint-Kruse, L., & Matthews, K. S. (2007). The lactose repressor system: paradigms for regulation, allosteric behavior and protein folding. *Cellular and Molecular Life Sciences, 64*(1), 3-16. doi:10.1007/s00018-006-6296-z

Wischgoll, S., Heintz, D., Peters, F., Erxleben, A., Sarnighausen, E., Reski, R., . . . Boll, M. (2005). Gene clusters involved in anaerobic benzoate degradation of Geobacter metallireducens. *Molecular Microbiology, 58*(5), 1238-1252. doi:https://doi.org/10.1111/j.1365-2958.2005.04909.x

Wobrauschek, P. (2007). Total reflection x-ray fluorescence analysis—a review. *X-Ray Spectrometry, 36*(5), 289-300. doi:10.1002/xrs.985

Woods, C., Humphreys, C. M., Rodrigues, R. M., Ingle, P., Rowe, P., Henstra, A. M., . . . Minton, N. P. (2019). A novel conjugal donor strain for improved DNA transfer into Clostridium spp. *Anaerobe, 59*, 184-191. doi:https://doi.org/10.1016/j.anaerobe.2019.06.020

Woods, C., Humphreys, C. M., Tomi-Andrino, C., Henstra, A. M., Köpke, M., Simpson, S. D., . . . Nojiri, H. (2022). Required Gene Set for Autotrophic Growth of Clostridium autoethanogenum. *Applied and Environmental Microbiology, 88*(7), e02479-02421. doi:doi:10.1128/aem.02479-21

Wu, L. F., & Mandrand-Berthelot, M.-A. (1987). Regulation of the fdhF gene encoding the selenopolypeptide for benzyl viologen-linked formate dehydrogenase in Escherichia coli. *Molecular and General Genetics MGG, 209*(1), 129-134.

Wuebbens, M. M., & Rajagopalan, K. V. (1995). Investigation of the early steps of molybdopterin biosynthesis in Escherichia coli through the use of in vivo labeling studies. *J Biol Chem, 270*(3), 1082-1087.

Wuebbens, M. M., & Rajagopalan, K. V. (2003). Mechanistic and mutational studies of Escherichia coli molybdopterin synthase clarify the final step of molybdopterin biosynthesis. *J Biol Chem, 278*(16), 14523-14532.

Xia, Q., Hendrickson, E. L., Zhang, Y., Wang, T., Taub, F., Moore, B. C., . . . Leigh, J. A. (2006). Quantitative proteomics of the archaeon Methanococcus maripaludis validated by microarray analysis and real time PCR. *Mol Cell Proteomics, 5*(5), 868-881. doi:10.1074/mcp.M500369-MCP200

Xiang, S., Nichols, J., Rajagopalan, K. V., & Schindelin, H. (2001). The Crystal Structure of Escherichia coli MoeA and Its Relationship to the Multifunctional Protein Gephyrin. *Structure, 9*(4), 299-310. doi:https://doi.org/10.1016/S0969-2126(01)00588-3

Yamamoto, K., & Ishihama, A. (2005). Transcriptional response of Escherichia coli to external copper. *Molecular Microbiology, 56*(1), 215-227. doi:https://doi.org/10.1111/j.1365-2958.2005.04532.x

Yokoyama, K., & Leimkühler, S. (2015). The role of FeS clusters for molybdenum cofactor biosynthesis and molybdoenzymes in bacteria. *Biochimica et Biophysica Acta (BBA) - Molecular Cell Research, 1853*(6), 1335-1349. doi:https://doi.org/10.1016/j.bbamcr.2014.09.021

Yoo, M., Bestel-Corre, G., Croux, C., Riviere, A., Meynial-Salles, I., & Soucaille, P. (2015). A Quantitative System-Scale Characterization of the

Metabolism of Clostridium acetobutylicum. In E. L. Papoutsakis Sang Yup (Ed.), *MBio* (Vol. 6, pp. e01808-01815).

Yoo, M., Nguyen, N.-P.-T., & Soucaille, P. (2020). Trends in Systems Biology for the Analysis and Engineering of Clostridium acetobutylicum Metabolism. *Trends in Microbiology, 28*(2), 118-140. doi:https://doi.org/10.1016/j.tim.2019.09.003

Yu, H., & Kim, K. S. (2010). Ferredoxin is involved in secretion of cytotoxic necrotizing factor 1 across the cytoplasmic membrane in Escherichia coli K1. *Infection and Immunity, 78*(2), 838-844. doi:10.1128/IAI.00674-09

Yu, Y., Tangney, M., Aass Hans, C., & Mitchell Wilfrid, J. (2007). Analysis of the Mechanism and Regulation of Lactose Transport and Metabolism in Clostridium acetobutylicum ATCC 824. *Applied and Environmental Microbiology, 73*(6), 1842-1850. doi:10.1128/AEM.02082-06

Zelcbuch, L., Antonovsky, N., Bar-Even, A., Levin-Karp, A., Barenholz, U., Dayagi, M., . . . Milo, R. (2013). Spanning high-dimensional expression space using ribosome-binding site combinatorics. *Nucleic Acids Research, 41*(9), e98-e98. doi:10.1093/nar/gkt151

Zhang, G., Ueberheide, B. M., Waldemarson, S., Myung, S., Molloy, K., Eriksson, J., . . . Fenyö, D. (2010). Protein Quantitation Using Mass Spectrometry. In D. Fenyö (Ed.), *Computational Biology* (pp. 211-222). Totowa, NJ: Humana Press.

Zhang, L., Zhao, R., Jia, D., Jiang, W., & Gu, Y. (2020). Engineering Clostridium ljungdahlii as the gas-fermenting cell factory for the production of biofuels and biochemicals. *Current Opinion in Chemical Biology, 59*, 54-61. doi:https://doi.org/10.1016/j.cbpa.2020.04.010

Zhang, W., Urban, A., Mihara, H., Leimkühler, S., Kurihara, T., & Esaki, N. (2010). IscS Functions as a Primary Sulfur-donating Enzyme by Interacting Specifically with MoeB and MoaD in the Biosynthesis of Molybdopterin in Escherichia coli*. *Journal of Biological Chemistry, 285*(4), 2302-2308. doi:10.1074/jbc.M109.082172

Zheng, D., Constantinidou, C., Hobman, J. L., & Minchin, S. D. (2004). Identification of the CRP regulon using in vitro and in vivo transcriptional profiling. *Nucleic Acids Research, 32*(19), 5874-5893. doi:10.1093/nar/gkh908

Zheng, M., Doan, B., Schneider, T. D., & Storz, G. (1999). OxyR and SoxRS Regulation offur. *Journal of Bacteriology, 181*(15), 4639. doi:10.1128/JB.181.15.4639-4643.1999

Zhou, Z. H., & Adams, M. W. W. (1997). Site-directed mutations of the 4Fe-ferredoxin from the hyperthermophilic archaeon Pyrococcus furiosus: Role of the cluster-coordinating aspartate in physiological electron transfer reactions. *Biochemistry, 36*(36), 10892-10900. doi:10.1021/bi9708141

Zhu, H.-F., Liu, Z.-Y., Zhou, X., Yi, J.-H., Lun, Z.-M., Wang, S.-N., . . . Li, F.-L. (2020). Energy Conservation and Carbon Flux Distribution During Fermentation of CO or H2/CO2 by Clostridium ljungdahlii. *Frontiers in Microbiology,*

11, 416. Retrieved from https://www.frontiersin.org/article/10.3389/fmicb.2020.00416

Zuker, M. (2003). Mfold web server for nucleic acid folding and hybridization prediction. *Nucleic Acids Res, 31*(13), 3406-3415. doi:10.1093/nar/gkg595

Zupok, A., Gorka, M., Siemiatkowska, B., Skirycz, A., & Leimkühler, S. (2019). Iron-Dependent Regulation of Molybdenum Cofactor Biosynthesis Genes in *Eschericha coli*. *Journal of Bacteriology*, JB.00382-00319. doi:10.1128/JB.00382-19

Zupok, A., Iobbi-Nivol, C., Méjean, V., & Leimkühler, S. (2019). The regulation of Moco biosynthesis and molybdoenzyme gene expression by molybdenum and iron in bacteria. *Metallomics*. doi:10.1039/C9MT00186G

Supplementary Materials

Table S 1 Primers sequence for the HiFi assembly PCR. Primers combinations are presented in Table II-6

Names	Sequence
Hifi_1	5' ACTGGTGGACAGCAAATGGGTCGCGGATCCTAATAGAAATAATTTTGTTTAACTTTAAAGTTAAGAGGCAAGAATGATATCAGCAATTATTTTATCCGGTGGAAAGGC 3'
Hifi_2	5' CCAAGGGGTTATGCTAGTTATTGCTGAGCTTATTATCTGTCAAAACACGCGTTTTCAAGTTCC 3'
Hifi_3	5' CTTCCCCCTGCGAATAAAGTTAAACAAAATTATTTCTATTACTCGAGTTATTATCTGTCAAAACACGCGTTTTCAAGTTCC 3
Hifi_3.1	5' GTCAAAACACGCGTTTTCAAGTTCCTTTAATG 3'
Hifi_4	5' GTTTAACTTTATTCGCAGGGGGAAGATGATAAGAGTAATAGGTGTAATCGGAAGAAAGGATACTG 3'
Hifi_4.1	5' GGAACTTGAAAACGCGTGTTTTGACAGATAATAACTCGAGTAATAGAAATAATTTTGTTTAACTTTATTCGCAGGGGGAAGATGATAAGAGTAATAGGTGTAATCGGAAGAAAGGATACT 3'
Hifi_5	5' CCAAGGGGTTATGCTAGTTATTGCTGAGCTTATTATTTTTTTTCAATTTTGATTGAAATTGTTTCAGGGTTTTCCG 3'
Hifi_6	5' CGCCGCCGGTCCTGCTTATAAAGTTAAACAAAATTATTTCTATTAAAGCTTTTATTATTTTTTTTCAATTTTGATTGAAATTGTTTCAGGGTTTTCCG 3'
Hifi_6.1	5' GTTCTTTTACAAATTTCATCGCCGCCGGTCCTGCTTATAAAGTTAAACAAAATTATTTCTATTAAAGCTTTTATTATTTTTTTTCAATTTTGATTGAAATTGTTTCAGGGTTTTCCG 3'
Hifi_7	5' AACTTTATAAGCAGGACCGGCGGCGATGAAATTTGTAAAAGAACTGATTAGTCGCAGTGATG 3'
Hifi_7.1	5' GCGGCGATGAAATTTGTAAAAGAACTGATTAGTCGCAGTGATG 3'
Hifi_8	5' CCAAGGGGTTATGCTAGTTATTGCTGAGCTTATTAAAACAAGTAAACTTCAACGAATTCGTCCTTTTCA 3'
Hifi_9.1	5' ACTGGTGGACAGCAAATGGGTCGCGGATCCTAATAGAAATAATTTTGTTTAACTTTATTCGCAGGGGGAAGATGATAAGAGTAATAGGTGTAATCGGAAGAAAGGATACTG 3'

Hifi_10.1	5' ACTGGTGGACAGCAAATGGGTCGCGGATCCTAATAGAAATAAT TTTGTTTAACTTTAAAGAGGTTTGGAATGAAATTTGTAAAAGAACT GATTAGTCGCAGTGATG 3'
Hifi_11.1	5' ACTGGACAGCAAATGGGTCGCGGATCCTAATAGAAATAATTTT GTTTAACTTTAAAGTTAAGAGGCAAGAATGCTTGCAGAAGATATA GTTTCAAGTGTAGATGTAC 3'
Hifi_11.2	5' ACTGGTGGACAGCAAATGGGTCGCGGATCCTAATAGAAATAAT TTTGTTTAACTTTAAAGAGGTTTGGAATGAGATACCTAGAACTTTG TACAATTGATCATGCTAAAGA 3'
Hifi_12	5' CCAAGGGGTTATGCTAGTTATTGCTGAGCTTATTAACATTCATA AATTATTTTTCCAGTATCTTTTGACTTTTTAAATGGAAGTTC 3'
Hifi_13.1	5' ACTGGTGGACAGCAAATGGGTCGCGGATCCTAATAGAAATAAT TTTGTTTAACTTTAAAGAGGTTTGGAATGATATCGGTACCCGAAGC AGAAAATATTTTAAAAAAATTTAAAT 3'
Hifi_14	5' CCAAGGGGTTATGCTAGTTATTGCTGAGCTTATTAATCATATAT CCAGACTTCAACTTCTCCATCTGCA 3'
Hifi_15.1	5' ACTGGTGGACAGCAAATGGGTCGCGGATCCTAATAGAAATAAT TTTGTTTAACTTTAAAGAGGTTTGGAATGAAAACTGCTGATGAAG TTCAAAATATTTTAAATGGATTTACA 3'
Hifi_16	5' CCAAGGGGTTATGCTAGTTATTGCTGAGCTTATTAAATCAGCCA GATTTTAACATTTTTTCTGACTTTTACCTC 3'
Hifi_17.1	5' CAGTTCTTTTACAAATTTCATTCCAAACCTCTTTAAAGTTAAACA AAATTATTTCTATTACTCGAGTTATTATTTTTTTTCAATTTTGATTGA AATTGTTTCAGGGTTTTCCG 3'
Hifi_18.1	5' TGGAATGAAATTTGTAAAAGAACTGATTAGTCGCAGTGATG 3'
Hifi_19.1	5' ATCTTCTGCAAGCATTCCAAACCTCTTTAAAGTTAAACAAAATTA TTTCTATTACTCGAGTTATTATTTTTTTTCAATTTTGATTGAAATTGT TTCAGGGTTTTCCG 3'
Hifi_19.2	5' TTCTAGGTATCTCATTCCAAACCTCTTTAAAGTTAAACAAAATTA TTTCTATTACTCGAGTTATTATTTTTTTTCAATTTTGATTGAAATTGT TTCAGGGTTTTCCG 3'
Hifi_20.1	5' GAGGTTTGGAATGCTTGCAGAAGATATAGTTTCAAGTGTAGAT GTAC 3'
Hifi_20.2	5' GAGGTTTGGAATGAGATACCTAGAACTTTGTACAATTGATCAT GCTAAAGA 3'
Hifi_21.1	5' CTGCTTCGGGTACCGATATCATTCCAAACCTCTTTAAAGTTAAA CAAAATTATTTCTATTACTCGAGTTATTATTTTTTTTCAATTTTGATT GAAATTGTTTCAGGGTTTTCCG 3'
Hifi_22.1	5' GGAATGATATCGGTACCCGAAGCAGAAAATATTTTAAAAAAAT TTAAAT 3'

Hifi_23.1	5' GAACTTCATCAGCAGTTTTCATTCCAAACCTCTTTAAAGTTAAAC AAAATTATTTCTATTACTCGAGTTATTATTTTTTTTCAATTTTGATTG AAATTGTTTCAGGGTTTTCCG 3'
Hifi_24.1	5' GGAATGAAAACTGCTGATGAAGTTCAAAATATTTTAAATGGATT TACA 3'
Hifi_25.1	5' CAGTTCTTTTACAAATTTCATTCCAAACCTCTTTAAAGTTAAACA AAATTATTTCTATTACTCGAGTTATTATCTGTCAAAACACGCGTTTT CAAGTTCC 3'
Hifi_26.1	5' ATCTTCTGCAAGCATTCCAAACCTCTTTAAAGTTAAACAAAATTA TTTCTATTACTCGAGTTATTATCTGTCAAAACACGCGTTTTCAAGTT CC 3'
Hifi_26.2	5' TTCTAGGTATCTCATTCCAAACCTCTTTAAAGTTAAACAAAATTA TTTCTATTACTCGAGTTATTATCTGTCAAAACACGCGTTTTCAAGTT CC 3'
Hifi_27.1	5' CTGCTTCGGGTACCGATATCATTCCAAACCTCTTTAAAGTTAAA CAAAATTATTTCTATTACTCGAGTTATTATCTGTCAAAACACGCGTT TTCAAGTTCC 3'
Hifi_28.1	5' GAACTTCATCAGCAGTTTTCATTCCAAACCTCTTTAAAGTTAAAC AAAATTATTTCTATTACTCGAGTTATTATCTGTCAAAACACGCGTTT TCAAGTTCC 3'
Hifi_29.1	5' CAGTTCTTTTACAAATTTCATTCCAAACCTCTTTAAAGTTAAACA AAATTATTTCTATTAAAGCTTTTATTATTTTTTTTCAATTTTGATTGA AATTGTTTCAGGGTTTTCCG 3'
Hifi_30.1	5' TTCTAGGTATCTCATTCCAAACCTCTTTAAAGTTAAACAAAATTA TTTCTATTAAAGCTTTTATTATTTTTTTTCAATTTTGATTGAAATTGT TTCAGGGTTTTCCG 3'
Hifi_31.1	5' CTGCTTCGGGTACCGATATCATTCCAAACCTCTTTAAAGTTAAA CAAAATTATTTCTATTAAAGCTTTTATTATTTTTTTTCAATTTTGATT GAAATTGTTTCAGGGTTTTCCG 3'
Hifi_32.1	5' GAACTTCATCAGCAGTTTTCATTCCAAACCTCTTTAAAGTTAAAC AAAATTATTTCTATTAAAGCTTTTATTATTTTTTTTCAATTTTGATTG AAATTGTTTCAGGGTTTTCCG 3'

Table S 2 Primers for the control of the hifi assembly cloning and sequencing

Name	Sequence	Utilisation
Fwd ext gor 1	ATCCGGATATAGTTCCTCCTTTCAGCAAAAAAC	Sequencing of *gorS2* gene in pET28a
Fwd int gor 2	CCCGCACTGGTAGTGTGTTAAATATTTTCCTT	Sequencing of *gorS2* gene in pET28a
Fwd int gor 3	TCCACAATTAGTCCAGTTTTTAGTTTCAATTGCTTCTT	Sequencing of *gorS2* gene in pET28a
Fwd int gor 4	TCTTCTTCCCCATTAAGTACAATTACAGTTGGAACTT	Sequencing of *gorS2* gene in pET28a
Rev int gor 1	GCCAACAATGTACTGGGCAATTGGAAAT	Sequencing of *gorS2* gene in pET28a
Rev int gor 2	CACCGGTTTTTAACTGGAGAATGCCTTTTATTG	sequencing of *gorS2* gene in pET28a
Rev int gor 3	ATGGATTTCACTTTTCTGCAATGGGCG	sequencing of *gorS2* gene in pET28a
Rev int gor 4	GGCGTAGAGGATCGAGATCTCGATCC	sequencing of *gorS2* gene in pET28a
reseq GAPOR 3 rev	ATACATTTAAAGATACGGGAATTCAAACGTTGCAAT	sequencing of *gorS2* gene in pET28a
reseq GAPOR 3 fwd	CATCAAACCCAAGTTTGTCGATTGTATGGACC	sequencing of *gorS2* gene in pET28a

reseq GAPOR 4 rev	AAACAAGCGCTCATGAGCCCGA	sequencing of *gorS2* gene in pET28a
reseq GAPOR 4 rev	GAAGCAGGTCCGACTAAAAATGCCCT	sequencing of *gorS2* gene in pET28a
mobB assembly seq	CGCAGTGACCGCAGTTTAGATTATAGGA	Sequencing assembly primer binding in *mobB*
moeA1 assembly seq	TCTGAAACCATCATGGAATTTGAATTAACGATCATTG	Sequencing assembly primer binding in *moeA1*
moeA2 assembly seq	AATGTCGCCAAACATGTGGACATCAAATATC	Sequencing assembly primer binding in *moeA2*
moeA2 seq 2	TTCATAAGTATCCTTTGCTCTCACG	Sequencing assembly primer binding in *moeA2* 5'end
moeA3 sequencing	TCATCTATTAAAACGTAGGATTCTA	Sequencing assembly primer binding in *moeA3*
moeA4 sequencing	TCTGCAGGGTCTCCAGCGTGAATTT	Sequencing assembly primer binding in *moeA4*

Table S 3 Primers for *gor* plasmid construction for expression in *E. coli*

Name	Sequence	Utilisation
fwd *gorL* recomb	TGTCCACCAGTCATGCTAGCTTAGACCATTCTTCCCTGGCTC	Amplification of *gor* for cloning in pET28a(+)
rev *gorS* recomb	AAGAAGGAGATATACCATGTCAGAGGAGAAGC	Amplification of *gor* for cloning in pET28a(+)

Table S 4 Primers for *gapor* and *gor* expression plasmids construction for *C. autoethanogenum* and *C. acetobutylicum*.

Name	Sequence	Utilisation
promHydA fwd	ATATATCACCGGTGGCGGCCGCTTTACTTAAATGTAACCGATTTGTG	Amplification of PhydA and RBS (forward) pMTL83151_AM001
promHydA rev 2	ATACCATGGTTATCCTCCCAAAATGTAAAATATAATTAAAATATATTAATAAA	Amplification of PhydA and RBS (reverse) pMTL83151_AM001
prim thl fwd	TTTATCAGGAAACAGCTATGACCGCGGCCGCTTTTTAACAAAATATATTGATAAAAATAATAATAGTGGGTATAATTAAGTTGTTAGAGAAAACGTATAAATTAGGGATAAACTATGGAA	Amplification of P*thl* and RBS by primer dimer PCR pMTL83151_AM002
prim thl rev	TGATGATGATGATGGCTGCTGCCCATATGAACTAACCTCCTAAATTTTGATACGGGGTAACAGATAAACCATTTCAATCTATTTCATAAGTTCCATAGTTTATCCCTAATTTATACGTTT	Amplification of P*thl* and RBS by primer dimer PCR pMTL83151_AM002
PhydA hifi fwd	TATCAGGAAACAGCTATGACCGCGGCCGCTTTACTTAAATGTAACCGATTTGTG	Amplification of P*hydA* and RBS for Hifi assembly with *gor* pMTL83151_AM005 (forward)

PhydA rev hyfi gorSL	GGTTATCCTCCCAAAATGTAAAATATAATTAA AATATATTAATAAACTTCGTTAAAAAATTAAC GTTTAATCAACGTAAATATTACGTACATT	Amplification of P*hydA* and RBS for Hifi assembly with *gor* pMTL83151_A M005 (reverse)
PhydA rev gorS RBS	CACCTGCTAAAATGTAAAATATAATTAAAATA TATTAATAAACTTCGTTAAAAAATTAACGTTT AATCAACGTAAATATTACGTACAT	Amplification of P*hydA* and native *gorS* RBS for Hifi assembly with *gor* pMTL83151_A M005.1 (reverse)
fwd gorS pHydA	TATTTTACATTTTGGGAGGATAACCATGTCAG AGGAGAAGCGCTCAAAAATATTGA	Amplification of *gor* for assembly of pMTL83151_A M005 (forward)
gorS fwd g-art PfdX	GGAGGTAACAACAACATGTCAGAGGAGAAG CGCTCAAAAATATTG	Amplification of *gor* for assembly of pMTL83151_A M006 (forward)
gorS fwd g-art PtcdB	AAAGGAGAAAATCACATGTCAGAGGAGAAG CGCTCAAAAATATTGA	Amplification of *gor* for assembly of pMTL83151_A M007 (forward)
rev gorL pMTL	TGCAGGCTTCTTATTTTTATGCTAGCTTAGACC ATTCTTCCCTGGCTCTCC	Amplification of *gor* for cloning in pMTL plasmids

Table S 5 Primers for the control of the pMTL based plasmids construction

Name	Sequence	Utilisation
ColE1+tra-F2	CCATCAAGAAGAGCGAC	Control PCR and sequencing upstream of the MCS
pCB102-R1	CTGTTATGCCTTTTGACTATC	Control PCR and sequencing downstream of the MCS
pMTLsec1	GGATAAAAAAATTGTAGATAAATTTTATAAAATAGTTTTA	Sequencing pMTL83151_AM001 and *gorS2* containing pMTL
pMTLsec2	AGAAAAGATTTTTTTCATCAAATGCATCGTATTTCCATGT	Sequencing pMTL83151_AM001 and *gorS2* containing pMTL
pMTLsec3	CAATGGGCGGTGCTGGATATACATTTAAAGATACGGGAAT	Sequencing pMTL83151_AM001 and *gorS2* containing pMTL
pMTLsec4	TAATTCCATTGTAACGTGGTAATTATTTCCAAATGTTCCG	Sequencing pMTL83151_AM001 and *gorS2* containing pMTL
pMTLsec5	TTTTTAACTGGAGAATGCCTTTTATTGATAAAAATAAAAG	Sequencing pMTL83151_AM001 and *gorS2* containing pMTL
pMTLsec6	CCAAAAGAATCGTAAACTCCAAAGTCTTCGAATTTTTAT	Sequencing pMTL83151_AM001 and *gorS2* containing pMTL
pMTLsec7	GGAGAACGGGGCCAAATTTCGCCAACAATGTACTGGGCAA	Sequencing pMTL83151_AM001 and *gorS2*

		containing pMTL
pMTLsec 8	CTATAATTTTCTTTTCTGTAAATTTCTTTCTATTC AGCAC	Sequencing pMTL83151_A M001 and *gorS2* containing pMTL
pMTLsec 9	AAGTATGAAATCATAAATAAAGTTTAATTTTGA AGTTATT	Sequencing pMTL83151_A M001 and *gorS2* containing pMTL
pMTLsec 10	ATGAGTGATAACACTGCGGCCAACTTACTT	Sequencing pMTL83151_A M001 and *gorS2* containing pMTL
pMTLsec 11	CTTAATTACAAATTTTTAGCATCTAATTTAACTT CAATTC	Sequencing pMTL83151_A M001 and *gorS2* containing pMTL
pMTLsec 12	CGCCACATAGCAGAACTTTAAAAGTGCTC	Sequencing pMTL83151_A M001 and *gorS2* containing pMTL
seq fwd gorS	ACGCTGTATAGAAGCTTGTGAAG	Sequencing of *gor* containing pMTL
seq fwd gorL1	CCACTATGGGGTTTAGATACAGA	Sequencing of *gor* containing pMTL
seq fwd gorL2	TTATTGACTGCATCGAGAAGCTT	Sequencing of *gor* containing pMTL

Table S 6 Primers for the control of the gene deletion in *E. coli*

Names	Sequence	Utilisation
Delta mobA Fwd	AGGTTAAGGTCTGGCTCTTTTTGCCAGG	Control of *mobA* deletion in *E. coli*
Delta mobA Rev	AGCTCATAGCTATCTTTGCCTGGCTTATCAACATC	Control of *mobA* deletion in *E. coli*
Delta mobB Fwd	GCAAGCAGGAGAACGCCGG	Control of *mobB* deletion in *E. coli*
Delta mobB Rev	CCCGCCATAAACTGCCAGGCATC	Control of *mobB* deletion in *E. coli*
KanR int rev ctrl DmobAB	GTGAGAATGGCAAAAGTTTATGCATTTCTTTCCA	Control of the Kan^R maker after P1 transduction
KanR Fwd ext ctrl DmobAB	GCGTATTTCGTCTCGCTCAGGC	Control of the Kan^R maker after P1 transduction
KanR ext rev ctrl DmobAB	CCGAACTCAGAAGTGAAACGCCGTA	Control of the Kan^R maker after P1 transduction
KanR int Fwd ctrl DmobAB	GCGTATTTCGTCTCGCTCAGGC	Control of the Kan^R maker after P1 transduction
fwd EcmoeA cntrl PCR	ATGACTGCTTTGCCCATAGTATTCGTCC	Control of the deletion of *E. coli* moeA

rev EcmoeA cntrl PCR	GTCGAAGTCGAGCAGCGTCAGGTTA	Control of the deletion of *E. coli* moeA
iscR del fwd ext 2	AAACGCCACGATAAAAAAATGGCACTGAA	Control of the deletion of *E. coli iscR*
iscR del fwd int 2	GTTTTCCCGTCTGCGTAAAAATGGTCTG	Control of the deletion of *E. coli iscR*
iscR del rev ext	TTTCTGATAAAAGTTGGCTGCACCTTTGATC	Control of the deletion of *E. coli iscR*
iscR del rev int	GACCAGACACATCCAGCACTTCCT	Control of the deletion of *E. coli iscR*
DhypF fwd	ACAGAACTGTTCTGGCATCAAAGGAATTCC	Control of the deletion of *E. coli hypF*
DhypF rev	CCTTTGTGGTCAGTGAGGATCGGATGA	Control of the deletion of *E. coli hypF*
DselA fwd	ATGTCCTCCTGACCCATCTCACGT	Control of the deletion of *E. coli selA*
DselA rev	GGCATAGCCGAGATCGATGGTCA	Control of the deletion of *E. coli selA*

Table S 7 Primers for 16s ribosomal DNA for contamination control

Names	Sequence	Utilisation
27F	AGAGTTTGATCCTGGCTCAG	Amplification of 16s ribosomal DNA for strain identification
1492R	GGTTACCTTGTTACGACTT	Amplification of 16s ribosomal DNA for strain identification

Table S 8 Synthesised DNA fragments

Name	DNA sequences
P_{fdx}_RB3	CAGGAAACAGCTATGACCGCGGCCGCGTGTAGTAGCCTGTGAA ATAAGTAAGGAAAAAAAAGAAGTAAGTGTTATATATGATGATTA TTTTGTAGATGTAGATAGGATAATAGAATCCATAGAAAATATAG GTTATACAGTTATATAAAAATTACTTTAAAAATTAATAAAAACAT GGTAAAATATAAATCGGTACCAATACGACTCACTATAGGTTCCG GTGATACCAGCATCGTCTTGATGCCCTTGGCAGCACCCTGCTAA GGAGGTAACAACAACATGGGCAGCAGCCATCATCATCATCATCA CACAGCGGCCTGGTGCCGCGCGGCAGCCATATGAACATTTTGAT TGATGGATCAAGACAA
P_{tcdB}	CAGGAAACAGCTATGACCGCGGCCGCGGTTAATGAATTTAAAG AAATATTTACAATAGAAATCAAATTTTAGAATTAACTTTATTGTA AAATCAATAACTTAATCTAAGAATATCTTAATTTTTATATTTTATA TAGAACAAAGTTTACATATTTATTTCAGACAACGTCTTTATTCAAT CGAAGAGCAAATTAATCAACTGAGTGTCTTCAATTTAAAATGTTA GGAAGTGAATGTATATGAAAACCTAAGTAGATATTAGTATATTT TATAAATAGAAAGGAGGATATATAAAAGAGTTTTAGCATTTAGA TGTAAAAATATTCAATAAAAATATTATAGTAAAGGAGAAAATCA CATGGGCAGCAGCCATCATCATCATCATCACAGCAGCGGCCTGG TGCCGCGCGGCAGCCATATGAACATTTTGATTGATGGATCAAGA
gorSL	ATTCTAGAAATAATTTTGTTTAACTTTAAGAAGGAGATATACCAT GTCAGAGGAGAAGCGCTCAAAAATATTGAAAGCTAAAAAAATG AATCGGTGTTTAGGTTGTATTTCATGTATGCTGGCCTGTGCAAG AACTGTTTATCAGGATTACTCTCCCGGGAAAAGTGCTATAAAAA TCAGAAGCAGTGGGGGCTTGCAGGGAAAGTTTGTTGCAGATAT CTGCAGGGGGTGTCAGGAACCTTCTTGTGTTCCAGTTTGTCCAA CACAGTCATTGATGACTAGGGCAGGGGGAGGCTTGAAATATAA CAGCGAAAAGTGTATCGGCTGTAAACGCTGTATAGAAGCTTGTG AAGTACATGCCATTGGTTTTGATGAGAGCAGTAGGAGACCAATC CTGTGCATACAGTGCGGAATATGTGTAAAATCTTGTCCGCATCA GGTTCTGTCAATGGAGGAACGAAATCAATGTTAGGTAGCAGCC ATCATCATCATCATCACAGCAGCGGCCTGGTGCCGCGCGGCAGC CATGGTTTAAATAAGGAAGGAATTCGAATTTTATATATAGACTT GACTACGGAAGAAGTTGAAAAAATAGAACGGCAAGATTTGAAA GATTACTTAGGGGGTGTAGGTGTTGCGATCAAGTTATTTGATGA GTATAAACAACCAGATTTAGACGTTTTTCATGAAAAACAACCAAT CATTTTTTCTATTGGACCTTTGGAAACGATTTATCCTGTGGTAAC GAAAACTGTTGCTGTTTTTCGATCTCCCTTAACTGGGGAACTAGG AGAAAGTTATGCTGGATTAAGATTGGGAATGGCCATTCGTTTCG CTGGATACGATGCGATCCTGATAACTGGTAAAGCGAAAGGGCC AATATATCTTAGTATTAACAGTCAAAGTGTAGAATTAAAAAAAG CACAACCACTATGGGGTTTAGATACAGAAGAAAGCGGAAGATT GCTGAGAGAAATGGAGTCAGGCAGTGGTTTACGTTCAATTCTGC GAATTGGTCAAGGAGGAGAAAGAGGAATAAGGTATGCTGGAG

TTAATGTAGATACCTATAGACATTTCGGAAGATTGGGTTTAGGG
TCTTTATTCGGAAGTAAGAATTTAAAAGGAATGGTGATTACAGG
GGATAGTACAGTACCGATAGAAGGTAAAAAAGAATATCAAAAA
ACGTATGCTGAAATTTATAAAAAAGTTACGGAAACCGATGTGAT
GGAAAAATATCATGGCTTAGGAACAACGGTCAATGTAATTCCAT
TAAATGATATGAATGGACTGCCTACAAGAAATTTAAAGCAAAGT
TATTTTGAAGCTGCAGATGAAGTTAGTGGTGAAGCATTTGCTAA
AGAAAAACTTATTAGAAAAATTGCATGCACGGGTTGTCCAATAG
GGTGTATACATATTGGACTTTACAGAAATCAATTTCAAGAACCA
ATGGAATATGAATATTCCCCGGTTGCCTACGATCATGAATTGGTT
TTTGCAGTAGGAACATTTATTGGTATAAAAGATAAAAATAAAAT
TTTAGCAATTATTGACTGCATCGAGAAGCTTGGATTAGATTGTAT
TACAACGGGAGTACTATTGGGATGGATAACGGAAGCGTATAAG
AGCGGTATCATAACTGAAAAAGATTTAGATCTACAGGTAGATTT
CGGAAATGAAAGTGCGTACATACAGATCATTGAAAACATTGTTA
ACGAAAAAAATAGTTTTTATAAAGATTTAGCATTTGGAACAGGA
TATGCTTCAAAAAAATATGGCGGAGAGGATTTCGCAGCTGTGCT
TGGCGGTCATGAAATCGCAGGATATCATACGGGTTATGGCAATC
TATTAGGGATGGCAGTGGGGGCTAGACATTCCCACTTAGACAAT
GCAGGTTATTCCATTGATCAAGCAACAGATCCAGCAGATCAAGA
TCAAATAGTAGATAAGTTAATTGAAGAGGAAATAACAAGAAAC
ATCTTAAACTGTTTAGTCATTTGCTTATTTGCCAGAAATATTTATG
ACTTAGGTACTGTAATAAAAGCACTGGAAACAATCGGTATCAGC
TGGAAAGAAGATGAATTAAAAAAACTTGGAAAAAAAATATTTTA
CCTTAAAATACAGGTTAAAAAAACATTAGGCTTTGATTATAAAA
AAATAGCATTTCCGAAGCGTTTCTTTGAAACAGAGACACTTCATG
GAAAGATGGATCAGGACATAGCGATGGAGTTATTAAGGCTATA
CAATAAAAAAATTGATCAAATTTGTTCTGAGGCAACAACGCTATT
TGCTACAACGGAGAGCCAGGGAAGAATGGTCTAAGCTAGCATA

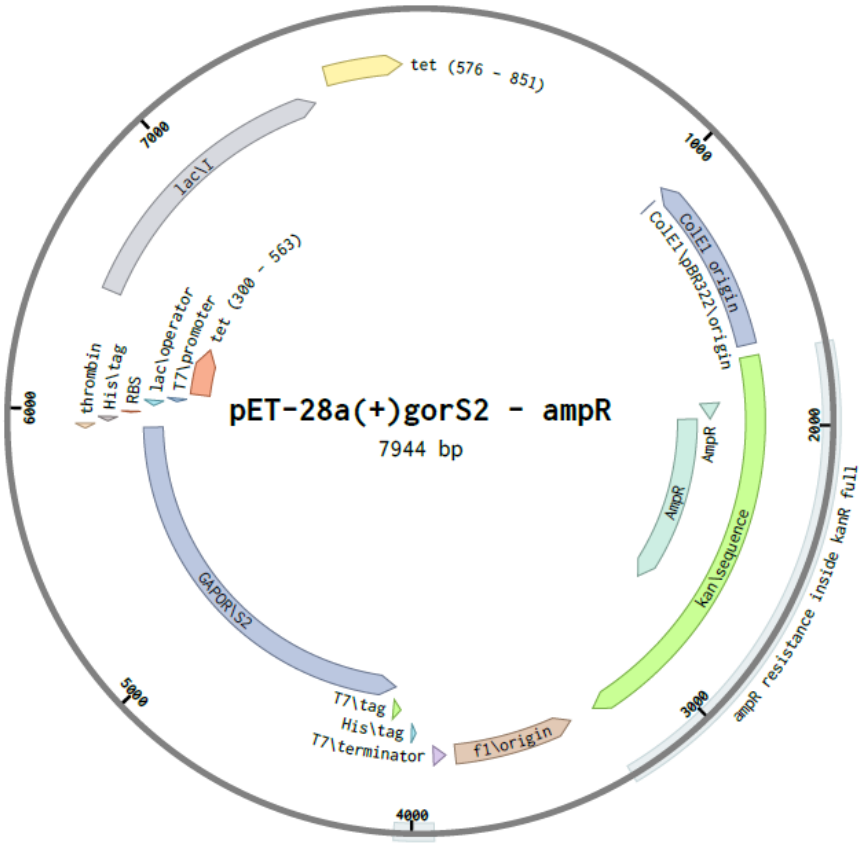


Figure S 1 pET-28a(+)*gorS2* (*Amp^R)* plasmid map.

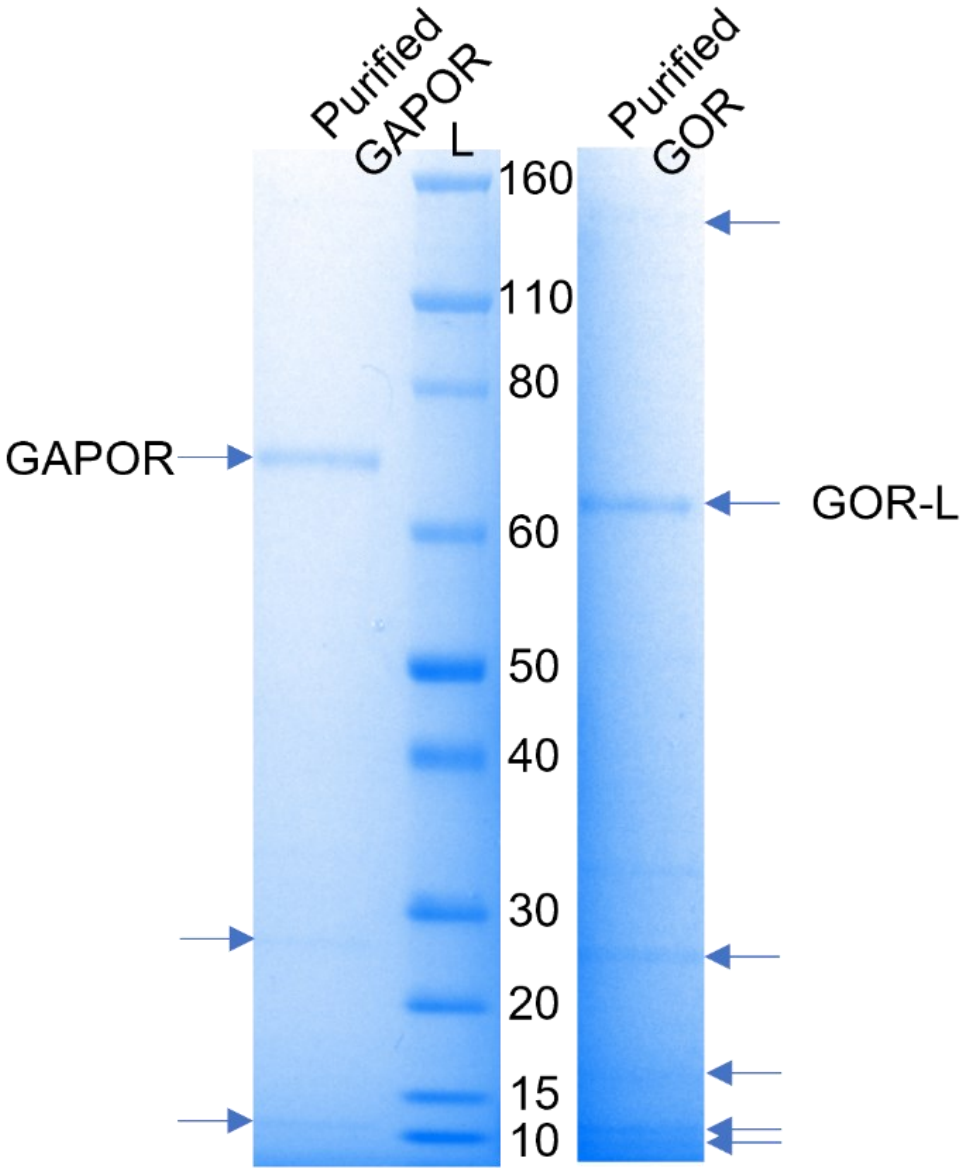


Figure S 2 Blue Silver Coomassie stained SDS PAGE Gel of affinity purified GAPOR and GOR-SL produced in CAC_AM004 and CAC_AM007 GAPOR and GOR-L bands are marked. Contaminant bands are marked with arrows allowing to observe contaminant common to both purification and extra contaminant in GOR-SL purified protein.