

# Organohalide respiration in Firmicutes: energy metabolism and transcription regulation

Présentée le 16 juillet 2021

Faculté de l'environnement naturel, architectural et construit  
Laboratoire de biotechnologie environnementale  
Programme doctoral en génie civil et environnement

pour l'obtention du grade de Docteur ès Sciences

par

**Mathilde Stéphanie WILLEMIN**

Acceptée sur proposition du jury

Prof. K. Schirmer, présidente du jury  
Prof. C. Holliger, Dr J. Maillard, directeurs de thèse  
Prof. D. Kelly, rapporteur  
Prof. E. A. Edwards, rapporteuse  
Prof. R. Aebersold, rapporteur



# Acknowledgements

This manuscript is the result of more than four years of intense and sometimes truly exhausting work, which I would not have achieved without all the support and love that I got from my colleagues, my amazing friends and, most importantly, my family.

First, I would like to address my sincere gratitude to my thesis director, **Prof. Christof Holliger**, for the faith you have placed in me from day one, and for making me feel proud of what I could accomplish. Thank you for teaching me rigour and precision that we sometimes lack on this side of the “Graben”.

During this thesis, I was lucky enough to be doubly backed up. I would thus like to warmly thank my thesis co-director, **Dr. Julien Maillard**, for the countless hours you spent teaching/-explaining/discussing science (amongst other things) with me. During these few years by your side, I have learned that there is always a way to reach the target, though sometimes at the expense of a lot of sweat. I also thank you for curbing my “unbridled enthusiasm”, and for preventing me from rushing along while forgetting important steps on the way.

I would also like to take the opportunity to thank all the inspiring scientists who have composed the jury of this thesis: **Prof. Kristin Schirmer**, **Prof. Rudolf Aebersold**, **Prof. David Kelly**, and **Prof. Elizabeth Edwards**, for reviewing this document, providing me with helpful comments, and helping me to improve the quality of this manuscript.

Of course, my supervisors and jury members were not the only necessary people to successfully obtain this PhD degree. This achievement would not have been possible without the daily motivation to interact with the amazing (present and former) members of the LBE family. Here, I would especially like to thank the technical team for their incredible efficiency and for helping me in so many ways during these years. In particular, **Emmanuelle Rohrbach**, **Stéphane Marquis**, and **Marc Deront**, but also all the apprentices who, thanks to competent supervision from the former, became true “lab killers” (in the good sense). I was lucky to have several of these killers by my side for one or the other experiment presented in this thesis, and I thus owe them a special thanks: **Morgane Baumgartner**, **Vincent Jeannet**, **Romain Savary**, **Emmy Oppliger**, my dear friend **Marie Vingerhoets**, and, although not as much as what I would have liked (#covid), **Emylène Ostertag**. At the LBE, we can also constantly count on our “Mama Mena”, **Filomena Jacquier**, who I would like to thank as well for the numerous

comforting words and smiles she offered me when I needed it the most. And, last but not least, I also want to sincerely thank the people who have shared my daily pain, suffered my constant and loud complaints, and shared a beer when there was no other choice: the great passers-by **Dr. Nicolas Derlon**, **Marta Cerruti**, and **Dr. Alexandra Murray** for bringing your happy moods and some fresh air with you; the newly arrived **Dr. Laetitia Cardona** and **Natalia Rodilla Ramirez** for standing me during this last period, which was particularly stressful for me, and for your encouragements; **Dr. Aline Adler** for sharing parental duties of Muso, your office, but also precious pieces of advice with me; **Lorenzo Cimmino** for sharing all the pain of working with the anaerobic bacteria and for the countless and sincere good laughs in the office; **Arnaud Gelb** for sharing “pressure-release running sessions” with me but, above all, for being such an amazing friend who always puts others in front of your own needs. I hope you know how precious you are to the others and I could never thank you enough for your support during this thesis and also during this particularly difficult passed year. Finally, my former colleague **Dr. Géraldine Buttet** for your warm welcome at LBE; I am grateful that I have crossed your path at EPFL and that I can call you my dear friend today.

Secondly, having travelled from Neuchâtel to Lausanne by way of Fribourg and Bern has brought me so many amazing friends who I would like to thank here. I feel so lucky that life has put you along my path and I could never express in words how essential you are to me. For the sake of simplicity, since there are a lot of people to thank, I have decided to proceed by region so as to try to not to forget anybody.

So, a special thank goes to my friends who I met in Lausanne: the crazy people from our sister-lab (EML), especially **Dr. Niels Burzan**, **Dr. Cornelia List**, and **Dr. Solenne Marion** for showing me the way towards graduation; the members of the LUC badminton club, especially my dearest friends from the former LUC 4 and 3: **Valentine Barras**, **Aline Chauffard**, **Pilar Benavides**, and **Loan Dao** for the nice moments we shared on and off the courts; and, last but not least, the people who have been by my side during this last (weird) year, **Thomas Debarre** for his uncontrolled spontaneity and his unconditional friendship, **Dr. Eric Durandau** for being an innate BigDad with great listening skills and an even greater heart, **Timon Ganguillet** for being such a trusted friend who I cannot wait to see open his giant wings to fly over the world, and **Blaise Devaud** for always being captain smile and good laugh.

I would also like to thank my dearest friends and former colleagues from Bern. In particular, my former supervisor **Dr. Sandro Käser** and my friend **Dr. Moritz Niemann** for giving me the “taste for science” and for Berner rösti; as well as my two scientific women and amazing friends **Dr. Shikha Kumari** and **Dr. Mom. Anke Harsman** for the girls talk/wine events, for being such inspiring women, and for your friendship which is so precious to me.

During my time in Fribourg, I was lucky to meet some truly great friends who I would like to mention here. Many many thanks go to the members of the badminton club of Villars-sur-Glâne, and especially **Martin et Mahélie Benz-Hetzel**, **Antoine Cotting**, **Jérémie Stöckli**, and the crazy M&M’s girls team, with a special thanks to **Morgane Mondoux** and **Lucie Delley** for your sincere friendship, and, last but not least, to **Mylène Durwang** for all the good times

## Acknowledgements

---

spent together and for being the sweetest woman I know: I miss you and can't wait to have more time to visit you. In addition, I would like to acknowledge my friends from my bachelor studies, and particularly my dear friend **Nathalie Gremaud**, who I am so lucky to still have in Lausanne: to me you are the perfect balance between strength and softness and I admire you for being the most loyal and the nicest person in the world.

A special thanks also goes to **Dr. Sebastian Dobarco**; of course I could not mention Fribourg without mentioning you. I thank you for sharing my life during several years, for your love and support, and for making me grow by your side; you will always have a place in my heart. I also want to thank all your amazing friends and your family for welcoming me in their lives.

Lastly, I would like to thank my best and oldest friends from Neuchâtel; the ones with who I grew up with. It was sometimes hard to keep in touch while giving all I had to finish this thesis, and I thank and love you all for forgiving me when I was failing to do so. I especially thank my friend **Alan Vejzovic** for being an example of audacity and for showing me that you do not succeed unless you try. I am also proud to thank my warrior friend **Julie Botha**, who is a true example of strength: I could not thank you enough for your words, your love, and your support, especially during this particularly difficult last year. You taught me that you are only as strong as you think you are. Another inspiring woman who I want to mention here is **Marie Jacot**, who I am proud to have had by my side for almost twenty years now. Your capacity to be successful at basically everything you try will never cease to amaze me. Your latest achievement was to become the mom of the cutest girl, and once again you are doing it remarkably well. Finally, I would also like to address all my love and gratitude to **Lucie Turberg** for always being here for me, and for being that truly concerned friend with who I can share anything with, and expect nothing but sincerity and compassion in return. I am so glad to have you as one of my very closest friends despite the distance, and I treasure our friendship for its simplicity and its sincerity.

Finally, sometimes, there is a really thin line between friends and family. I think that you, **Pauline de Coulon**, have crossed this line by being the most faithful and precious friend I have. We have shared everything together and I could not wish for a better partner in crime than you. You are and will always be the person that I want to call first to talk, complain, laugh and cry. My life would not have been the same without you.

Speaking of family, they constitute the third and last sub-population of amazing people I want to thank here. Today, I am measuring how lucky I was to be surrounded by such a loving and supporting family.

I will start by thanking the person who I consider to be its newest member, my love and partner in life, **Dr. Thibault Asselborn**. I thank you for making me feel strong and proud of myself, for your unconditional support in any situation (even when I am such a pain), and for having been by my side during what might well be the most emotionally-intense year of my life. I am so thankful to have you in my life and still cannot believe how quickly you have become essential to me. You are like a little sparkle which I want carry everywhere in the palm of my

hand to stay warm, and on which I want to blow when it gets darker. You have made of my life what I wanted it to be, and I can't wait for our new adventures together. I also thank you for introducing me to your lovely family (especially **Sabine Goulon**, **François Asselborn**, but also **Sophie**, **Théo**, **Chloé**, and **Christian**, who I thank for their warm welcome and for making me feel that I have just found another home in theirs.

Special thanks also go to my big brother, **Thomas Willemin**, for standing by me literally from day one and sharing all the family events (good and bad). I am proud of the man you have become, and you can be proud of yourself too. Of course, the people around me know that another very important member of my family is my little sister, **Julie Willemin**, who I want to thank for her unconditional love, support, and for making me the proudest big sister on earth. Watching you grow up and become the sweet, beautiful, and strong woman you are today was the most satisfying event of my life so far.

I would also like to take the opportunity to thank all my cousins from both sides of my direct family (**Dhiraj**, **Vimala**, **Amandine**, **Jules**, **Axelle**, **Elliot** and **Pablo**); my aunts **Lulu**, **Valérie**, and **Stéphanie**, from who I proudly inherited my middle name; my uncles **Nicolas**, **Momo**, and **Yves**, and my grandparents **Gaby**, **Jean**, **Jean**, and **Germaine**. In addition, I am one of these people who are lucky enough to have an additional beautiful stepfamily that I have always considered to be mine, and I thank you (**Papi**, **Mamie**, **Velmira**, and **Bernard**, **Sam**, and **Mai**, **Roxane**, **Mathilde**, and **Margot**) for making me feel part of yours. All together, you form a very heterogeneous group of individuals who have all influenced me in a way or another, and I am forever grateful to you for always making me feel supported and loved.

Finally and most importantly, I want to thank my parents for making of me who I am today. Making you proud is my greatest reward from this success.

First, my dad, **Yves Willemin**, for being such a supportive father and such an example in life. I see myself in so many of your traits, and this makes me so proud. I love and respect you and I feel so grateful to have you as my dad. I also would like to take the opportunity to thank my father's wife, **Sophia Willemin**, for your love, genuine support, and for welcoming me in your life with open arms.

Last but not least, I thank my mom, **Sophie Wüthrich**, for always believing in me (sometimes more than myself), and for being such an example of courage, resourcefulness and strength. Despite all you can say, you have to know that you can be proud of how you have raised Thomas and I as a single mom. If ever you have a doubt about that, you just have to look at what we have become thanks to you.

# Abstract

Organohalides are a class of compounds often considered as persistent pollutants and harmful to environmental and human health. Some bacteria, among which are representatives from the Firmicutes phylum, are capable of using these compounds as terminal acceptors in a process called organohalide respiration (OHR). To do so, organohalide-respiring bacteria (OHRB) are equipped with enzymes called reductive dehalogenases which perform the terminal reduction in the OHR respiratory chain by replacing the halide by a hydrogen. OHR is not only interesting from a biological point of view but it also represents a powerful process for bioremediation purposes. The present work focuses on two different fundamental aspects of OHR which remain poorly understood.

The first part of the work focused on the transcriptional regulation of the reductive dehalogenase genes. OHRB can encode up to several dozens of different reductive dehalogenases in their genome which suggest the use of a broad number of organohalides, but this diversity remains largely unexplored. In Firmicutes OHRB, RdhK are transcriptional regulators specialised in the activation of reductive dehalogenase genes upon exposure of the cell to organohalides. Thus, the identification of RdhK binding partners (organohalide effector and DNA target sites) represents an interesting alternative to identify new OHR substrates while preventing the tedious and challenging characterisation of the complex membrane-associated and oxygen-sensitive reductive dehalogenases. Here, a strategy based on RdhK hybrid proteins was designed and optimised, and is proposed to serve to improve the efficiency of RdhK characterisation. The approach enabled a functional decoupling of the two domains of RdhK regulators targeting either the effector or the DNA target site. Therefore, it reduces the complexity of the screening procedure in the identification of RdhK binding partners. Further implementation of the hybrid strategy will increase the global comprehension of Firmicutes OHR regulatory networks and ultimately provide an indirect way to explore the reductive dehalogenase substrate diversity.

The second part of the thesis aimed to increase knowledge in the energy metabolism of Firmicutes OHRB. Quantitative comparative proteomics was applied to *Desulfitobacterium hafniense* strain DCB-2, which revealed the proteome adaptations to the growth on different substrates promoting either fermentation or respiration. In addition, possible roles of the

complex I-like enzyme was investigated using physiological and biochemical approaches. Respiratory complex I is the entry point of electrons produced by cytoplasmic metabolic activity in the respiratory chain and is composed of three modules. Most OHRB express a version of the complex which lacks the N-module responsible for the electron uptake from NADH, suggesting the use of alternative electron donors. While a few candidate partners for the OHR complex I-like enzyme were proposed based on proteomics results, the study of the physiological role of the enzyme helped in developing different respiratory models for strain DCB-2, which mainly differ in their dependence on the complex I-like enzyme.

Key words: Organohalide respiration, *Desulfitobacterium*, reductive dehalogenase, transcription regulation, RdhK regulators, hybrid protein design, comparative proteomics, energy metabolism, complex I, respiratory chain models.

# Résumé

Les composés organohalogénés (OH) sont souvent considérés comme polluants persistants et nocifs pour l'environnement et la santé humaine. Des bactéries, certaines appartenant au phylum des Firmicutes, peuvent utiliser ces composés comme accepteurs terminaux dans le processus de respiration des composés OH (ROH). Ces bactéries (abrégées BROH) possèdent des enzymes appelées déshalogénases réductrices (RdhA, en anglais) qui catalysent la réduction de l'halogénure et son remplacement par un atome d'hydrogène. La ROH n'est pas seulement intéressant d'un point de vue biologique, mais aussi à des fins de bioremédiation. Cette thèse aborde deux aspects fondamentaux de la ROH qui restent encore mal compris.

La première partie porte sur la régulation transcriptionnelle des gènes *rdh*. Les BROH contiennent plusieurs dizaines de RdhA dans leur génome, suggérant l'utilisation d'un grand nombre de composés OH, une diversité qui est encore méconnue. Chez les BROH parmi les Firmicutes, les protéines RdhK sont des régulateurs qui activent la transcription des gènes *rdh* en présence de composés OH. Ainsi, l'identification des partenaires des RdhK (le composé OH qui agit en tant qu'effecteur et les sites de liaison sur l'ADN) est une alternative intéressante pour découvrir des nouveaux substrats de la ROH, tout en évitant la caractérisation fastidieuse des enzymes RdhA (associées aux membranes et sensibles à l'oxygène). Une stratégie basée sur des RdhK hybrides a été conçue et optimisée pour accroître l'efficacité du processus de caractérisation des RdhK. L'approche permet de découpler les deux domaines fonctionnels des régulateurs qui ciblent l'effecteur ou l'ADN, et de faciliter l'identification des partenaires de liaison des RdhK. L'application de la stratégie d'hybrides permet une compréhension globale des réseaux de régulation chez les BROH et représente un moyen indirect d'explorer la diversité des substrats des enzymes RdhA.

La deuxième partie de la thèse s'intéresse au métabolisme énergétique des BROH parmi les Firmicutes. La souche DCB-2 de *Desulfitobacterium hafniense* a été soumise à une méthode de protéomique comparative et quantitative qui a révélé les adaptations du protéome de cette bactérie lorsqu'elle est cultivée sur différents substrats favorisant soit la fermentation soit la respiration. De plus, l'enzyme homologue au complexe I a été étudiée à l'aide d'approches physiologiques et biochimiques. Le complexe I respiratoire est composé de trois modules et sert

de point d'entrée dans la chaîne respiratoire aux électrons produits par l'activité métabolique du cytoplasme. La plupart des BROH expriment une version du complexe auquel manque le module N, qui reçoit les électrons du NADH, ce qui suggère l'utilisation d'autres donneurs d'électrons. À partir des données protéomiques, des candidats de donneur d'électrons pour le complexe ont été identifiés. De plus, différents modèles respiratoires pour la souche DCB-2, qui diffèrent principalement par leur dépendance à l'homologue du complexe I, sont proposés grâce à plusieurs expériences testant le rôle physiologique de l'enzyme.

Mots clefs : respiration des composés organohalogénés, *Desulfitobacterium*, déshalogénase réductrice, régulation de la transcription, régulateurs RdhK, conception de protéines hybrides, protéomique comparative, métabolisme énergétique, complexe I, modèles de chaînes respiratoires.

# Contents

<b>Acknowledgements</b>	<b>i</b>
<b>Abstract (English/Français)</b>	<b>v</b>
<b>List of figures</b>	<b>xiii</b>
<b>List of tables</b>	<b>xv</b>
<b>List of accronyms</b>	<b>xvii</b>
<b>1 General Introduction</b>	<b>1</b>
1.1 Organohalides and the environment . . . . .	1
1.2 Organohalide respiration and organohalide-respiring bacteria . . . . .	2
1.3 Diversity of reductive dehalogenases . . . . .	3
1.4 Composition and organisation of <i>rdh</i> gene clusters . . . . .	5
1.5 Energy metabolism of OHRB: physiological and biochemical considerations . .	6
1.6 Thesis objectives and outline . . . . .	10
<b>I Transcription regulation of organohalide respiration in Firmicutes</b>	<b>13</b>
<b>2 A review of transcription regulation in organohalide respiration</b>	<b>15</b>
2.1 Introduction . . . . .	15
2.2 Transcriptional regulators involved in organohalide respiration . . . . .	16
2.2.1 Two-component systems . . . . .	17
2.2.2 MarR-type regulators . . . . .	19
2.2.3 CRP/FNR-type transcriptional regulators . . . . .	20
<b>3 RdhK hybrid transcriptional regulators</b>	<b>29</b>
3.1 Introduction . . . . .	29
3.2 Material and Methods . . . . .	30
3.2.1 Plasmids construction and DNA manipulations . . . . .	30
3.2.2 Protein production and purification . . . . .	32
3.2.3 Electrophoretic mobility shift assay (EMSA) . . . . .	32
3.2.4 beta-galactosidase activity assay . . . . .	33

3.3	Results . . . . .	34
3.3.1	Domain definition for the design of RdhK hybrid proteins . . . . .	34
3.3.2	<i>In vitro</i> characterisation of RdhK hybrid proteins . . . . .	35
3.3.3	<i>In vivo</i> characterisation of RdhK hybrid proteins . . . . .	38
3.3.4	<i>In vivo</i> binding affinity of parental RdhK for DB8 and DB7 . . . . .	40
3.4	Discussion . . . . .	41
3.4.1	RdhK hybrid proteins are active regulatory proteins . . . . .	41
3.4.2	Hybrid B is more specific than Hybrid A . . . . .	41
3.4.3	RdhK6 recognises DB motifs located in other gene clusters . . . . .	42
3.5	Conclusion . . . . .	43
<b>4</b>	<b>Characterisation of new RdhK regulators</b>	<b>45</b>
4.1	Introduction . . . . .	45
4.2	Material and Methods . . . . .	46
4.2.1	Plasmids construction and DNA manipulations . . . . .	46
4.2.2	Protein production and purification . . . . .	47
4.2.3	Electrophoretic mobility shift assay (EMSA) . . . . .	47
4.2.4	Genome-scale pull-down assay . . . . .	47
4.3	Results and discussion . . . . .	49
4.3.1	<i>In vitro</i> identification of DNA motifs targeted by cross-species hybrid RdhK . . . . .	49
4.3.2	Identification at genome-scale of DNA sequences targeted by RdhK regulators . . . . .	51
4.4	Conclusion . . . . .	53
<b>II</b>	<b>Energy metabolism and complex I-like enzymes in organohalide-respiring Firmicutes</b>	<b>55</b>
<b>5</b>	<b>Metabolic versatility of Firmicutes OHRB and their complex I-like enzymes</b>	<b>57</b>
5.1	Comparative physiology of organohalide-respiring bacteria: special emphasis on the Firmicutes . . . . .	57
5.1.1	<i>Dehalobacter</i> spp. . . . .	58
5.1.2	<i>Desulfitobacterium</i> spp. . . . .	58
5.2	Complex I and complex I-like enzymes . . . . .	60
5.2.1	Respiratory complex I . . . . .	60
5.2.2	Evolution of complex I . . . . .	63
5.3	Complex I-like enzymes in OHRB . . . . .	65
<b>6</b>	<b>Comparative proteomics of <i>Desulfitobacterium hafniense</i> strain DCB-2</b>	<b>67</b>
6.1	Introduction . . . . .	67
6.2	Material and Methods . . . . .	68
6.2.1	Cultivation of <i>Desulfitobacterium hafniense</i> . . . . .	68
6.2.2	Cell harvesting and sample preparation . . . . .	69
6.2.3	Proteomic analysis . . . . .	69

## CONTENTS

---

6.3	Results and Discussion . . . . .	72
6.3.1	General presentation of the data . . . . .	72
6.3.2	Proteins up-regulated in fermentation with pyruvate . . . . .	73
6.3.3	Versatility of respiratory metabolisms . . . . .	82
6.4	Conclusions . . . . .	95
<b>7</b>	<b>Potential role of complex I-like enzymes in OHR metabolism</b>	<b>97</b>
7.1	Introduction . . . . .	97
7.2	Material and Methods . . . . .	98
7.2.1	Bacterial cultures and growth medium composition . . . . .	98
7.2.2	Rotenone-induced growth inhibition . . . . .	98
7.2.3	Rotenone effect on PCE dechlorination . . . . .	99
7.2.4	Cell fractionation and membrane protein extraction . . . . .	99
7.2.5	Two-steps purification attempt of the complex I-like enzymes . . . . .	100
7.2.6	LC-MS/MS analysis . . . . .	100
7.3	Results . . . . .	101
7.3.1	Comparison of the complex I-like enzymes from Firmicutes OHRB with the <i>E. coli</i> complex I . . . . .	101
7.3.2	Physiological roles of complex I-like enzymes in OHRB . . . . .	101
7.3.3	Tentative biochemical characterisation of native complex I-like enzymes in Firmicutes OHRB . . . . .	108
7.4	Discussion . . . . .	116
7.4.1	Identification of redox partners for complex I-like enzyme in OHR . . . . .	116
7.4.2	Physiological relevance of complex I-like enzymes in Firmicutes OHRB . . . . .	118
7.5	Conclusions . . . . .	120
<b>8</b>	<b>Concluding remarks and perspectives</b>	<b>121</b>
<b>A</b>	<b>Supplemental material to Chapter 3</b>	<b>125</b>
<b>B</b>	<b>Supplemental material to Chapter 4</b>	<b>131</b>
<b>C</b>	<b>Supplemental material to Chapter 6</b>	<b>133</b>
<b>D</b>	<b>Supplemental material to Chapter 7</b>	<b>139</b>
	<b>Bibliography</b>	<b>145</b>
	<b>Curriculum Vitae</b>	<b>161</b>



# List of Figures

1.1	Phylogenetic tree of organohalide-respiring bacteria. . . . .	3
1.2	Sequential reductive dechlorination of PCE to ethene by organohalide-respiring bacteria. . . . .	4
1.3	Representatives reductive dehalogenase <i>rdh</i> gene clusters. . . . .	5
1.4	Model of quinone-independent organohalide respiration in <i>D. mccartyi</i> . . . . .	7
1.5	Model of PCE reductive dechlorination by <i>Dehalobacter restrictus</i> . . . . .	8
2.1	DNA binding motifs of organohalide respiration transcriptional regulators. . .	18
2.2	Sequence likelihood analysis of RdhK proteins across <i>Dehalobacter</i> spp. and <i>Desulfitobacterium</i> spp. . . . .	27
3.1	Design of RdhK hybrid proteins. . . . .	34
3.2	SDS-PAGE analysis of the four RdhK proteins used in EMSA. . . . .	35
3.3	<i>In vitro</i> interactions of the RdhK parental and hybrid proteins with dehalobox DNA motifs. . . . .	37
3.4	<i>In vivo</i> beta-galactosidase reporter assay of RdhK parental and hybrid proteins. .	40
4.1	Characterisation of a cross-species RdhK hybrid protein towards the promoters of all <i>rdh</i> gene clusters from <i>D. restrictus</i> . . . . .	50
4.2	Co-immunoprecipitation of DB7 with RdhK6 and Hybrid A. . . . .	52
5.1	Schematic representation of bacterial respiratory complex I. . . . .	61
5.2	Genomic organisation of the <i>nuo</i> homologous gene clusters in organohalide-respiring bacteria. . . . .	66
6.1	Principle components analysis for DCB-2 proteome expression in different growth conditions. . . . .	73
6.2	Distribution of proteins up-regulated in Pyr-only conditions. . . . .	74
6.3	Proteome adaptations to the electron donor. . . . .	82
6.4	Wood-Ljungdahl pathway and butyrate metabolism in strain DCB-2. . . . .	87
6.5	Differential expression of strain DCB-2 hydrogenases in the six growth conditions. .	88
6.6	Pairwise comparisons of La/Fu vs Py/Fu or Pyr-only. . . . .	91
6.7	Proteome adaptations to the electron acceptor. . . . .	93
6.8	Rdh proteins up-regulated in CLOHPA conditions. . . . .	95

## LIST OF FIGURES

7.1	Complex I-like enzymes in Firmicutes organohalide-respiring bacteria. . . . .	102
7.2	Rotenone inhibition on the growth of strain DCB-2 in different conditions. . . .	104
7.3	Rotenone inhibition on the growth of <i>D. restrictus</i> . . . . .	105
7.4	Effect of rotenone inhibition on growing cultures of strain DCB-2. . . . .	106
7.5	H <sub>2</sub> protection effect on rotenone-treated La/Fu cultures. . . . .	107
7.6	Rotenone inhibition on the growth of <i>D. hafniense</i> strain TCE1 and on perchloroethylene (PCE) dechlorination activity. . . . .	108
7.7	Membrane extracts analysis on BN-PAGE. . . . .	110
7.8	Size exclusion chromatograms of membrane extracts initially purified by anion exchange chromatography. . . . .	112
7.9	Relative abundance of NADH:ubiquinone oxidoreductase (Nuo) homologous subunits in <i>D. hafniense</i> strain DCB-2 across different growth conditions. . . .	113
7.10	Identification of potential Nuo partners. . . . .	115
7.11	Potential redox partners for <i>D. hafniense</i> complex I-like enzyme. . . . .	116
7.12	Hypothetical models for the respiratory chain of <i>D. hafniense</i> strain DCB-2. . .	119
A.1	Design of <i>mgl</i> -DB chassis for EMSA analysis. . . . .	127
A.2	Design of the <i>cprBA</i> -based DB chassis for <i>in vivo</i> beta-galactosidase reporter assay.127	
A.3	Control EMSA experiments with the four RdhK proteins. . . . .	128
A.4	Preliminary <i>in vivo</i> beta-galactosidase assay for promoter-binding activity of RdhK proteins. . . . .	129
A.5	Sequence alignment of DB motifs recognised by RdhK6. . . . .	129
C.1	Comparison of the growth performance of strain DCB-2 in the six growth conditions. . . . .	133
C.2	Heat-map of the proteomic results and hierarchical clustering. . . . .	134
C.3	Sequence analysis of the unusual hydrogenase from strain DCB-2 . . . . .	135
C.4	The <i>dsr</i> gene cluster of <i>D. hafniense</i> strain DCB-2. . . . .	137
C.5	Comparison between the two putative electron bifurcating gene clusters of strains DCB-2 with that of <i>C. kluyveri</i> . . . . .	137
D.1	Purification attempt of the complex I-like enzyme of <i>D. hafniense</i> (1). . . . .	140
D.2	Purification attempt of the complex I-like enzyme of <i>D. hafniense</i> (2). . . . .	141
D.3	Purification attempt of the complex I-like enzyme of <i>D. restrictus</i> . . . . .	142
D.4	BN-PAGE with purified membrane extracts. . . . .	143

## List of Tables

3.1	List of plasmids used in the study. . . . .	31
4.1	List of plasmids used in the study. . . . .	46
4.2	List of primers used in the study. . . . .	47
6.1	Proteins up-regulated in Pyr-only conditions. . . . .	75
6.2	Selected proteins up-regulated in Pyr-only conditions in comparison to all respiratory conditions except La/CIOHPA. . . . .	79
6.3	Proteins showing up-regulation level of at least 10-fold in Pyr-only conditions when compared to Py/Fu conditions. . . . .	80
6.4	Proteome adaptations to the use of H <sub>2</sub> as electron donor. . . . .	83
7.1	Summary of the electron donor/acceptor combinations used for different OHRB strains. . . . .	98
A.1	Fusion PCR mix. . . . .	125
A.2	List of oligonucleotides used in the study. . . . .	126
B.1	List of primers used to amplify <i>rdhA</i> promoters in PERK-23. . . . .	132
C.1	Proteins up-regulated in fermentation metabolism in comparison to all respiratory metabolisms except La/CIOHPA. . . . .	136



# List of abbreviations

<b>ATP</b>	adenosine triphosphate
<b>DCP</b>	dichlorophenol
<b>DNA</b>	deoxyribonucleic acid
<b>cAMP</b>	cyclic adenosine monophosphate
<b>CISM</b>	complex iron-sulfur molybdoenzyme
<b>CIOHPA</b>	3-chloro-4-hydroxyphenylacetic acid
<b>CODH</b>	carbon monoxide dehydrogenase
<b>cpr</b>	chlorophenol respiration
<b>CRP</b>	cAMP receptor protein
<b>DB</b>	dehalobox
<b>DBD</b>	DNA-binding domain
<b>DDM</b>	n-dodecyl-beta-D-maltoside
<b>EMSA</b>	electrophoretic mobility shift assay
<b>ESI-MS</b>	electrospray ionisation mass spectrometry
<b>ETC</b>	electron transfer chain
<b>FeS</b>	iron-sulfur
<b>FMN</b>	flavin mononucleotide
<b>FNR</b>	fumarate nitrate regulatory proteins
<b>gDNA</b>	genomic DNA
<b>GFP</b>	green fluorescent protein

<b>HK</b>	histidine kinase
<b>HTH</b>	helix-turn-helix
<b>Hup</b>	uptake hydrogenase
<b>ITC</b>	isothermal titration calorimetry
<b>MarR</b>	multiple antibiotic resistance regulator
<b>MNG</b>	maltose-neopentyl glycol
<b>NADH</b>	nicotinamide adenine dinucleotide (reduced form)
<b>OHR</b>	organohalide respiration
<b>OHRB</b>	organohalide-respiring bacteria
<b>Nuo</b>	NADH:ubiquinone oxidoreductase
<b>PCBs</b>	polychlorinated biphenyls
<b>PCE</b>	perchloroethylene
<b>PCR</b>	polymeric chain reaction
<b>POPs</b>	persistent organic pollutants
<b>RDase</b>	reductive dehalogenase enzyme
<b>RDases</b>	reductive dehalogenase enzymes
<b>Rdh</b>	reductive dehalogenase homologous
<b>RdhA</b>	reductive dehalogenase (catalytic subunit)
<b>RNA</b>	ribonucleic acid
<b>RR</b>	response regulator
<b>Tat</b>	Twin-arginine translocation
<b>TCA</b>	tricarboxylic acid
<b>TCP</b>	trichlorophenol
<b>TCS</b>	two-component system
<b>TMT</b>	Tandem-Mass-Tag
<b>WLP</b>	Wood-Ljungdahl pathway

# 1 General Introduction

## 1.1 Organohalides and the environment

Organohalides are organic compounds with at least one covalently bound halogen atom which are extensively used in chemical industry and manufacture for different purposes like the synthesis of different types of biocides or as degreasing agents, among others. Many of them are considered as persistent organic pollutants (POPs) and as major threats for environmental and human health. As examples, some chlorophenols or aliphatic molecules like tetrachloroethene (perchloroethylene, PCE) are recognised as potential carcinogenic compounds with associated genotoxicity and mutagenicity effects on human and/or animals [1, 2].

Organohalide occurrence in nature was for long believed to be mostly the result of anthropogenic activities. Although most of polluted sites clearly appeared after improper industrial disposals, organohalides are also naturally present in the environment. Indeed, several biotic and abiotic processes are known as being natural sources of organohalides in both marine and terrestrial environments [3]. Moreover, organohalides were shown to play an important role in the natural cycles of halogens [4]. Microorganisms exposition to organohalides is thus dated way before the industrial development era which is at the origin of major polluted sites. The discovery of organisms that have adapted to either use those compounds in co-metabolic processes or use them to sustain growth either as carbon and energy source or as terminal electron acceptor is the witness of this long-term exposition [5]. The so-called organohalide-respiring bacteria (OHRB) are particularly interesting as they define a subgroup of bacteria capable of using organohalides as terminal electron acceptor in their respiratory chain while conserving energy [6, 7]. This type of organisms, in the centre of the present work, are studied for their potential application in the bioremediation of polluted sites as their metabolic activity eventually leads to the transformation of organohalides to halogen-free organics (most of the time much less or no more toxic). The environmental distribution of OHRB is not restricted to polluted sites (from where most of them were isolated) as different species were also successively identified in pristine environments (for more detailed review, see [4, 5]). Moreover, organohalide respiration (OHR) represents a fascinating and yet poorly

understood energy metabolism.

## 1.2 Organohalide respiration and organohalide-respiring bacteria

The discovery of organisms using organohalides in their metabolism started in the 1960's [8]. More than 20 years passed before the first evidences of a process coupling reductive dehalogenation and energy conservation was published [9–12]. More recently, the term "organohalide respiration" (OHR) was proposed to universally define this bacterial process which was by the time also sometimes named halorespiration or dehalorespiration [7, 8, 13]. Today, the number of strains considered as organohalide-respiring bacteria (OHRB) (i.e. capable of OHR) is substantial and they include a few bacterial genera from three major phyla: Chloroflexi, Firmicutes and Proteobacteria (gamma-, delta- and epsilon-) [5]. On top of their phylogenetic affiliation, another level of classification is often employed to define an OHRB. Indeed, a given organism can be obligate or facultative OHRB, by reference to its ability to perform alternative energy metabolisms. As an example, *Dehalobacter* spp. and *Desulfitobacterium* spp. are both part of the Firmicutes phylum. However, while the first one is restricted to the use of molecular hydrogen as electron donor and organohalides as electron acceptors, the latter can use sulfite (as indicated by its name) but also fumarate, nitrate and more as electron acceptors while using hydrogen, lactate, formate or others as electron source (for a detailed reviews of each species, see [14] and [15], respectively). Figure 1.1 present a phylogenetic tree of OHRB where the major species of each phylum is displayed with an indication of whether a specific species is part of the obligate or the facultative OHRB. As shown in this figure, Firmicutes, which are at the centre of this present thesis, is the only phylum with representatives displaying both obligate and facultative OHR metabolism. This feature makes Firmicutes OHRB particularly interesting to investigate the major differences between these two metabolic strategies. OHRB from the others phyla are either in the obligate OHRB (Chloroflexi) or the facultative category (Proteobacteria).

Among Firmicutes, one strain of *Desulfitobacterium hafniense*, strain DCB-2, was particularly studied during this thesis. The metabolic versatility of *Desulfitobacterium* spp. enables the application of experimental strategies that are not possible with *Dehalobacter* spp. For example, the variation of the cultivation conditions can lead to the identification of substrate-specific adaptations. Strain DCB-2 was the first member of its genus to be isolated and was obtained from a stable trichlorophenol-dechlorinating consortium selected from municipal sludge [16]. Since 3-chloro-4-hydroxyphenylacetic acid (CLOHPA) was shown to be an organohalide substrate for *D. dehalogenans* [17], CLOHPA, a highly soluble organohalide in comparison to chlorinated phenols, was used as a model compound to cultivate other *Desulfitobacterium* spp., among which strain DCB-2 [18]. Today, strain DCB-2 represents the most studied strain among *Desulfitobacterium* spp. on many aspects that are relevant for the present work.

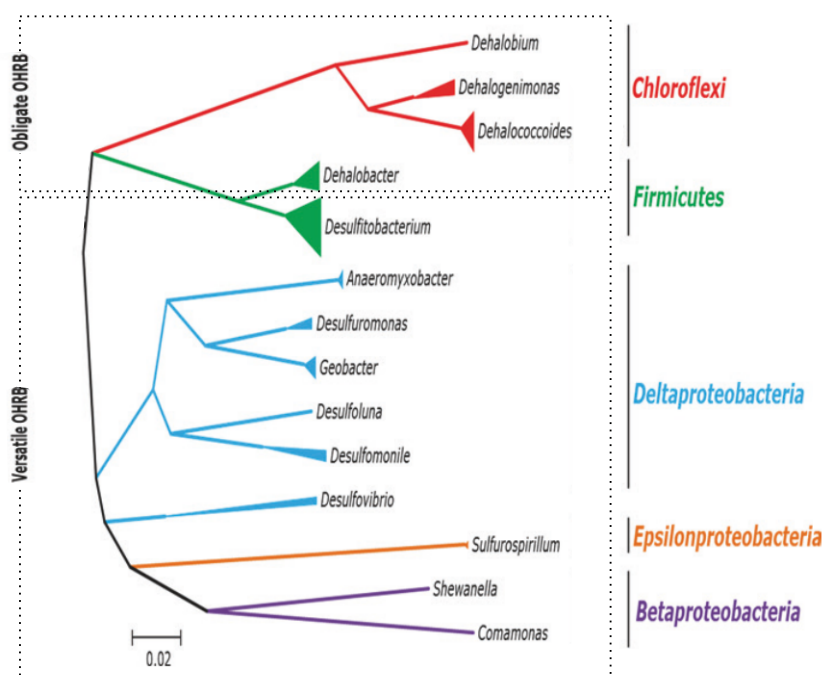


Figure 1.1 – Phylogenetic tree of organohalide-respiring bacteria. Adapted from [5].

The minimal gene set that defines putative OHRB is an operon constituted by one **reductive dehalogenase homologous** (*rdhA*) gene usually accompanied by the *rdhB* gene, both encoding the key catalytic enzyme in OHR (reductive dehalogenase, RdhA) and its putative membrane anchor (RdhB), respectively [7, 19, 20]. However, the presence of such an operon in the genome of an organism does not *per se* constitute the ultimate proof of a capacity to sustain growth by using organohalides as electron acceptors. Similarly, not every reductive dehalogenase (catalytic subunit) (RdhA) is necessarily involved in a respiration process (this will be further introduced in a later stage of the thesis). The physiological approach remains the most reliable way to identify a true OHRB.

### 1.3 Diversity of reductive dehalogenases

Not all OHRB have the same dehalogenation potential. One single OHRB can contain up to 36 *rdhA* copies [21]; a number that is generally higher in the obligate OHRB which reflect their dedication to this peculiar metabolic process. The scope of dehalogenation by one organism is determined by the number but also the nature of reductive dehalogenase enzymes (RDases) encoded in its genome. The reductive dehalogenation of PCE (see Figure 1.2, upper panel) is a good example to illustrate this concept. In this reduction reaction, the four chlorine atoms are sequentially removed from the organic backbone and replaced by a hydrogen. As indicated in the figure, some bacterial strains can perform the complete dechlorination of PCE, while the others are limited to one or more steps [22].

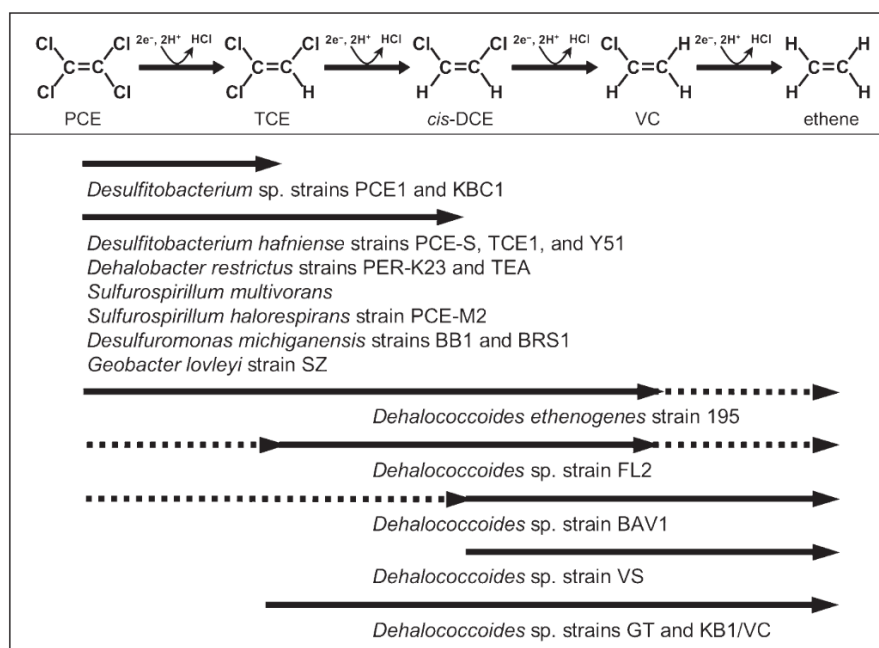


Figure 1.2 – Sequential reductive dechlorination of PCE to ethene by organohalide-respiring bacteria. The reactions is displayed at the top of the figure. The strains able to perform part or the complete reduction of PCE to ethene are indicated at the bottom (dash arrows indicate co-metabolic transformation). Figure from [22]

When the substrate for a given reductive dehalogenase enzyme (RDase) has been clearly identified, the general acronym "reductive dehalogenase homologous (Rdh)" is often replaced by one referring to the identified substrate (ex: *pceA* encodes for the catalytic subunit of the reductive dehalogenase dedicated to PCE). Despite the vast diversity of reductive dehalogenase genes recognised today [6, 19, 23], the number of biochemically characterised enzymes remains very limited. The sensitivity of RdhA to oxygen and the nature of the OHRB themselves (slow growing, strict anaerobes, genetically intractable) as well as the lack of sequence-substrate relationships are continuous challenges that have certainly contributed to slow down the progress in the OHR research topic [7]. For example, the structure of only one respiratory RdhA catalytic enzyme has been obtained so far [24].

Nevertheless, some common features obtained from sequence analysis and biochemical studies were established to define RdhA enzymes. First, RdhA enzymes are characterised by an N-terminal signal peptide for the Twin-arginine translocation (Tat) system. This latter transports RdhA enzymes in their already folded state across the cytoplasmic membrane [25]. Second, their active site is equipped by two types of redox cofactors: two iron-sulfur (FeS) centres and one cobamide cofactor, as shown in the crystal structure of *S. multivorans* PceA [24]. Third, as already mentioned, *rdhA* genes are usually co-transcribed (i.e. form an operon) with *rdhB* [20], a small protein displaying generally three trans-membrane helices and proposed to be the membrane anchor of RdhA (Note: recently, example of "orphan" RdhA

were reported [26, 27]). Finally, *rdhA* and *rdhB* are often part of larger gene clusters together with OHR-related accessory genes (so-called *rdh* gene clusters), which are found in restricted area of the genome [7, 21, 28, 29].

## 1.4 Composition and organisation of *rdh* gene clusters

The numerous *rdh* gene clusters present in the genomes of OHRB harbour different gene composition and organisation; a diversity that has been reviewed in the past (for recent work, see [7, 30]). In Figure 1.3 is displayed a selection of representative *rdh* gene clusters coming from different OHRB strains.

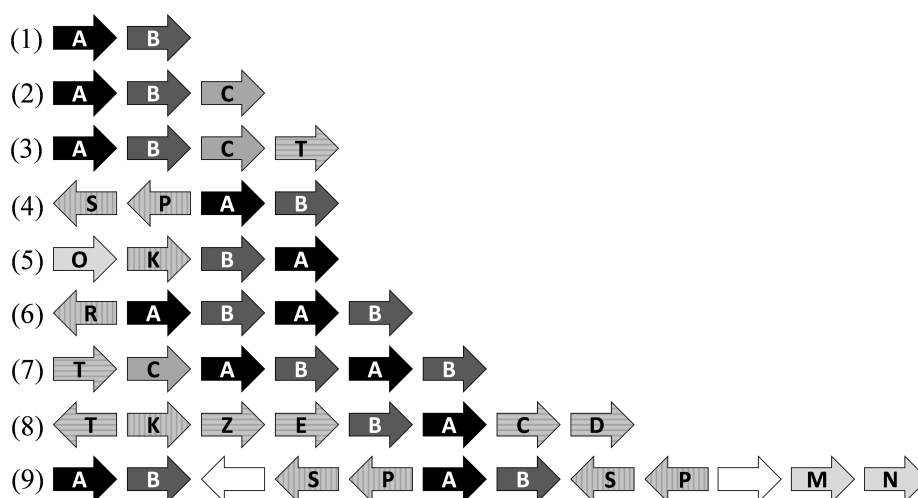


Figure 1.3 – Representatives reductive dehalogenase *rdh* gene clusters. Besides the minimal *rdh* gene cluster (1), the composition and organisation of *rdh* gene clusters vary significantly within and across OHRB phylogenetic groups. Examples for *rdh* gene clusters: (2) *vcrABC* of *Dehalococcoides mccartyi* strain VS [31]; (3) *pceABCT* of *Dehalobacter restrictus* [32]; (4) *pceAB* of *D. mccartyi* strain 195 [33]; (5) *rdh* cluster 1 of *Desulfitobacterium hafniense* strain DCB-2 [34]; (6) *cbdbA1452-6* of *D. mccartyi* strain CBDB1 [35]; (7) *pce* gene cluster of *Geobacter lovleyi* SZ [36]; (8) *cpr* gene cluster of *D. dehalogenans* [37]; (9) *pce* gene cluster of *Sulfurospirillum multivorans* [38]. The annotation of individual *rdh* genes follows the one proposed by Kruse *et al.* [39]. The (predicted) function of *rdh* genes is the following: (A) catalytic subunit; (B) membrane anchor for the catalytic subunit; (C) membrane-bound electron-transferring flavoprotein; (D) and (E) GroEL-type molecular chaperone; (K) CRP/FNR-type transcriptional regulator; (M) and (N) NapGH-like quinol dehydrogenase; (O) membrane-bound methyl-accepting chemotaxis sensor; (P) DNA-binding response regulator of a two-component system; (R) MarR-type transcriptional regulator; (S) histidine-kinase sensor of a two-component system; (T) trigger factor-like molecular chaperone; (Z) molecular chaperone. Taken from [7].

As shown in the figure, *rdhAB* gene clusters are often surrounded by a various number of accessory genes. A description of their associated function can be found in [30]. The *rdh* cluster dedicated to the *ortho*-substituted chlorophenols of *Desulfitobacterium dehalogenans*

strain JW/IU-DC1 constitutes a relevant example in the context of this thesis. It is composed by a total of eight genes forming the *chlorophenol respiration (cpr)TKZEBACD* gene cluster [37], which is conserved in other *Desulfitobacterium* strains [40]. Thus, beside the minimal set of reductive dehalogenase proteins (RdhA and B), this cluster encodes for proteins for which the predicted function is described below.

Briefly, RdhC proteins form a family of membrane-bound flavoproteins which show sequence similarities with the regulatory proteins NosR/NirI [37]. Despite this latter property and given the presence of a flavin mononucleotide (FMN)-binding domain, RdhC has recently been proposed as an electron-transferring protein to RdhA [41]. The RdhD and RdhE proteins have sequence similarities to GroEL-type molecular chaperones and were thus suggested to play a role in the maturation of the complex redox RdhA enzymes [30, 37]. RdhK proteins are defined as a subfamily of transcriptional regulators from the CRP/FNR superfamily [7, 42]. A whole section of this thesis is devoted to the transcriptional regulation in OHR with a particular focus on RdhK regulators and therefore will not be further discussed here (see Chapters 2 to 4). Finally, RdhT and RdhZ are both proposed to be active in the maturation of RdhA enzymes. RdhT proteins are described as trigger factor-like proteins that are binding to the Tat signal peptide of RdhA in order to delay its translocation across the membrane until the protein is properly folded [25, 43]. RdhZ is, for its part, not characterised, but was proposed to be a molecular chaperone based on the crystal structure of a homologous protein [30].

The multiplicity of *rdh* gene clusters in single OHRB are probably the result of a combination of adaptations to new substrates, horizontal gene acquisitions and duplication events [23, 44], maybe sometimes facilitated by the high plasticity of the genome region they occupy [21]. Despite the relatively large functional diversity of the proteins encoded in *rdh* gene clusters, only a few of them were proposed to be actively involved in the OHR electron transfer process. In fact, the biochemical characterisation of the global bacterial process remains poorly understood. The next section summarises the important findings and hypotheses published so far.

## 1.5 Energy metabolism of OHRB: physiological and biochemical considerations

Based on different genomic and proteomic studies available in the OHR field, the composition and the organisation of the OHR respiratory chain is believed to be variable between different bacterial genera which gave rise to several models [45, 46]. Although none of them is completely satisfying as they still contain a relatively high level of uncertainty, two main types of theoretical electron transport models can be distinguished: quinone-dependent and quinone-independent [45, 47, 48]. The latter is proposed for the Chloroflexi phylum and is first based on the fact that the genomes of organisms part of this phyla generally lack the quinone synthesis pathway and yet some were shown to grow in OHR conditions in quinone-depleted medium [49]. In addition, work performed with quinone analogues and/or antagonists on

different strains of *Dehalococcoides mccartyi* strongly supports this assumption (see detailed review in [50]). The current model for quinone-independent electron transport in *D. mccartyi* involves a supercomplex composed by the uptake hydrogenase (Hup), the RDase (RdhAB) and a member of the complex iron-sulfur molybdoenzyme (CISM) family [51, 52] (see Figure 1.4). This proposition was made after the results of recent biochemical cross-linking experiments and bolstered by the available proteomic data (for relevant reviews, see [45, 46, 50]).

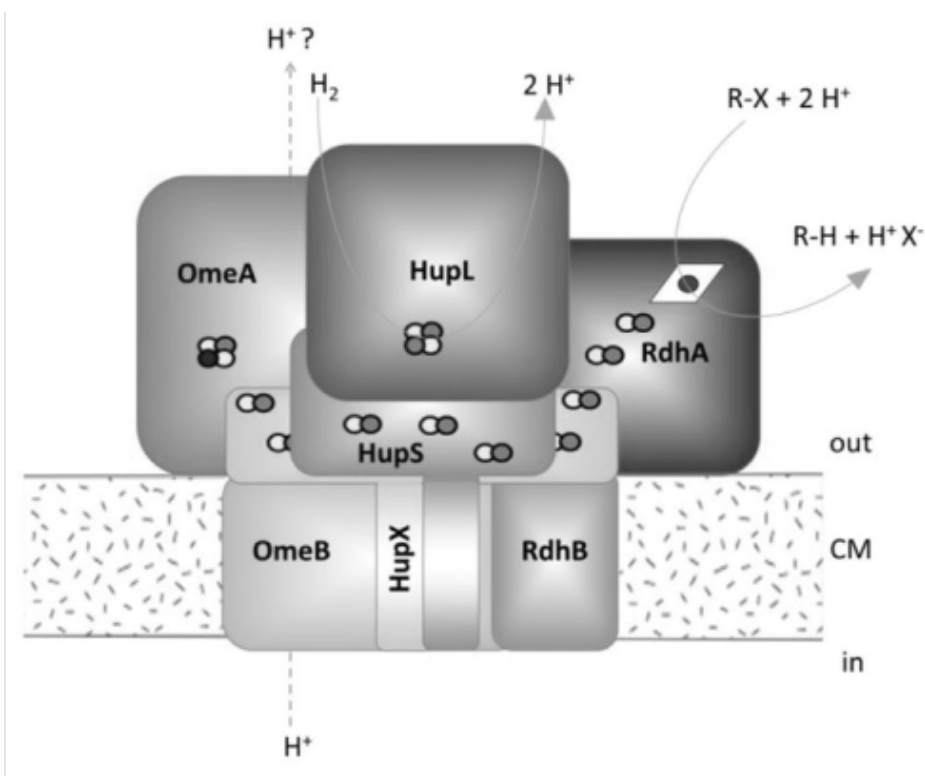


Figure 1.4 – Model of quinone-independent organohalide respiration in *D. mccartyi*. Adapted from [52].

Contrasting with the data obtained for some Chloroflexi OHRB representatives, the involvement of quinones as electron shuttle in the OHR respiratory metabolism was revealed in the earliest year by showing the inhibition of OHR by addition of quinone antagonists in *Dehalobacter restrictus* [12], *Desulfomonile tiedjei* [53] and *Sulfurospirillum multivorans* [45]. Furthermore, in *D. restrictus*, PCE dechlorination was shown to be induced after addition of quinone analogues and the quinone reduction was observed during electron transfer by spectrophotometry analysis [12]. As in other anaerobic respiratory metabolisms, menaquinones are most likely to be involved [54]. Indeed, this is supported by the presence of different types of menaquinones in *D. restrictus* [55] and *S. multivorans* [56]. Additionally, the identification of complete menaquinone synthesis pathways in several OHRB genomes, which, on the other hand, often lack the ubiquinone synthesis pathway, highly suggests that way [29, 34, 57]. The consensus model for quinone-dependent OHR electron transport chain thus show the electron donor machinery (which may vary depending on the nature of the electron donor in

use) linked to the terminal reductase by the menaquinone intermediate. As an example, in the model proposed for PCE respiration by *Dehalobacter restrictus*, displayed in Figure 1.5, the electrons are taken up from molecular hydrogen by the membrane bound Hup hydrogenase and serve to reduce the quinone to quinol. The quinol pool is shuttling the electrons further to the reductive dehalogenase which perform the terminal reduction of PCE [14].

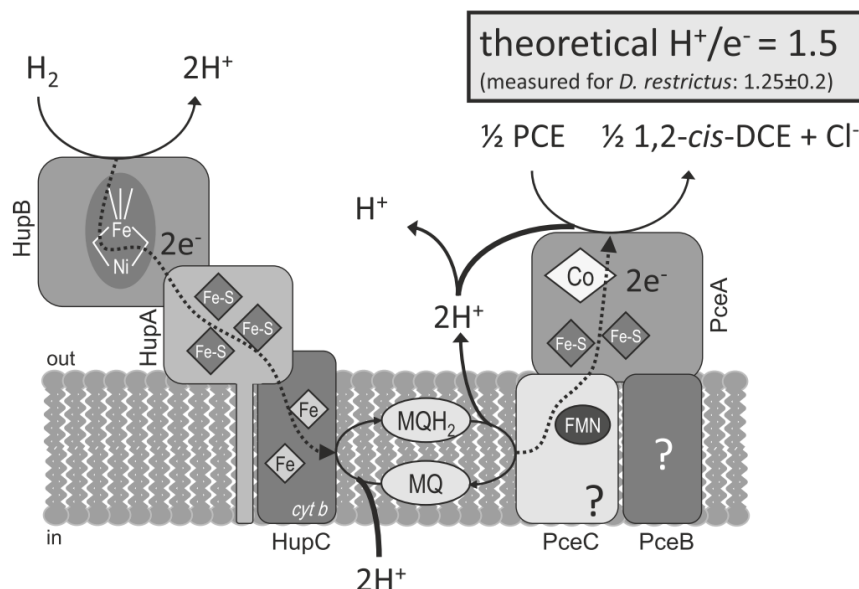


Figure 1.5 – Model of PCE reductive dechlorination by *Dehalobacter restrictus*. Example of quinone-dependent OHR electron transport in which menaquinone shuttle electron from the hydrogenase to the reductive dehalogenase. Figure from [14]

However, this model is challenged by a thermodynamic issue which has not been resolved so far. Indeed, the redox potential of the RDase corrinoid cofactor was measured for a few strains and the consensus value is situated around -370 mV [58–60]. In addition, the FeS clusters of *D. restrictus* were shown to be even lower in potential (-480 mV). Yet a direct transfer from menaquinol (-74 mV) [54] to the RDase redox cofactors is difficult to explain. Reverse electron flow, a mechanism which uses the proton motive force to lower the potential of electrons, was proposed to solve this problem [61]. More recently, with the discovery of electron bifurcation [62, 63] the involvement of additional redox players, like the FMN-binding protein PceC, was proposed [41].

Another surprising feature of OHR is the low growth yield achieved by OHRB in relation to the amount of free energy available. The relatively high redox potential of organohalides makes them very good electron acceptors [64]. The evidence for the establishment of a proton motive force (pmf) coupled to the RDase activity (i.e. to organohalide reduction) was obtained for a limited number of strains [11, 12, 65]. When the electrons are transferred from molecular hydrogen, the amount of energy from each electron transport is theoretically sufficient to translocate a minimum of three protons across the membrane [45, 54]. Yet,

from the above mentioned early studies, a relatively poor  $H^+/e^-$  ratio of 1 to 1.5 protons translocated per electron transferred was established. Taking the proposed model for PCE reductive dechlorination in *Dehalobacter restrictus* [14] (see Figure 1.5), such a ratio would correspond to one mole of adenosine triphosphate (ATP) produced per mole of released chloride, which fits with the published growth yields [14]. All this indicates that the OHR metabolism is under some constraints which prevents the whole available energy to be used for bacterial growth [45].

Except for the proton translocation that is likely coming from the quinone redox loop [48, 54], no actual active proton-pumping complex are known to take place in OHR metabolism. Anaerobic respiratory chains are known to be usually shorter than in aerobic respiration, conserving a lower amount of energy [54]. Then, OHR might represent a really good example to illustrate this phenomenon. Nevertheless, based on genomic and/or proteomic data, some potential redox intermediates have been proposed. As already mentioned, PceC was proposed to be delivering the electrons to the RDase through its FMN cofactor [41]. Similarly, a putative quinol dehydrogenase homologous to NapGH has been proposed as the missing link in *S. multivorans* [66]. Another example are the homologues of respiratory complex I which are found in most OHRB genomes (variable cluster compositions do exist, see details in Chapter 5) and detected in the proteome of cells cultivated in OHR conditions [33, 57, 66–68]. In aerobic respiration, complex I couples the transfer of electrons from nicotinamide adenine dinucleotide (reduced form) (NADH) to the quinone pool along with proton translocation [69]. The potential involvement of such proton-pumping machinery in OHR is at the centre of the second part of this thesis. Therefore, it will be further introduced and explored in Part II (Chapters 5 and 7).

## 1.6 Thesis objectives and outline

The objectives of this thesis concerned two main aspects of OHR in Firmicutes. Although some connections between the two aspects exist, it was decided to treat them separately and the thesis is thus divided in two main parts (Note: Chapters 1 and 8 correspond to the general introduction and to concluding remarks and perspectives, respectively, and are not listed here).

**Part I: Transcription regulation of organohalide respiration in Firmicutes.** The objective of this part was to develop a tool in order to use the transcriptional regulators RdhK as indirect way to explore the diversity of *rdh* gene clusters in Firmicutes and ultimately their dehalogenation potential. Indeed, the limited number of characterised clusters mainly come from the very nature of the OHRB. As strict anaerobes, slow growing and genetically intractable organisms, many biochemical approaches are challenging. RdhK represent an interesting alternative as they usually are activated upon binding to an organohalide compound and positively regulate the transcription of the genes involved in the respiration of this same compound through promoter binding. Thus, the goal was to develop a tool to efficiently screen the binding partners of yet uncharacterised RdhK regulators. Part I is divided in three chapters:

**Chapter 2** presents a literature review to introduce the OHR transcriptional regulation with a special emphasis on RdhK proteins. The content of this chapter was published in an extended form in 2019 [7].

In **Chapter 3**, a proof of concept of the RdhK hybrid proteins strategy is described, an approach which enables a more efficient characterisation of new RdhK proteins by the decoupling of the two domains of the regulators. The binding properties of two different versions of the RdhK hybrid composed by parts from two characterised RdhK proteins from strain DCB-2 (RdhK6 and RdhK1) were compared in order to define which design better serves the goal of this approach. The work presented in this chapter was published in 2020 [70].

In **Chapter 4**, different applications of the RdhK hybrid strategy are proposed. At first, the aim was to apply the RdhK hybrid strategy to the characterisation of new RdhK proteins from *D. restrictus*. Secondly, preliminary data on the adaptation of a genome-wide DNA-binding sites screening methodology for the RdhK hybrid system is presented in order to provide an efficient screening solution. Preliminary data obtained for both parts are discussed.

**Part II: Energy metabolism and complex I-like enzymes in organohalide-respiring Firmicutes.** The initial objective of the second part of the thesis was to further understand the OHR metabolic pathway with the characterisation of new players in the respiratory chain. Indeed, as presented in the general introduction (Chapter 1), the exact enzyme composition of the respiratory chain of Firmicutes OHRB remains elusive and the models available today fail to conciliate important thermodynamic aspects. In that context, a comparative proteomic study was conducted on *D. hafniense* strain DCB-2 in different growth conditions. As a facultative OHRB, the use of *D. hafniense* strain DCB-2 as model organism allowed to identify proteome adaptations specific to different types of energy metabolism, including OHR. This proteomic analysis gave access to a large dataset which was exploited to help a better understanding of the global metabolism of *D. hafniense* as well as to answer specific research questions such as the involvement of yet unconsidered enzymatic players in the different respiratory chains. In particular, the role of an 11-subunits complex-I like enzyme was investigated in strain DCB-2. This enzyme is expressed by most OHRB and the presence of such an important enzyme was often reported although its direct implication was never investigated. Part II is composed of three chapters:

**Chapter 5** presents a literature review to introduce the physiology and the metabolism of Firmicutes OHRB as well as a detailed introduction to respiratory complex I and complex I-like enzymes in OHRB.

In **Chapter 6** the results of the quantitative proteomic study performed on *D. hafniense* strain DCB-2 cultivated in six different growth conditions are presented. The growth conditions were chosen in order to highlight the proteome adaptations to different electron donors and acceptors through the comparison of cross-over electron donors/acceptors combinations. The results presented in this chapter form the basis for a manuscript in preparation.

Finally, **Chapter 7** presents the physiological and tentative biochemical characterisation of the complex I-like enzyme from strain DCB-2. A physiological approach was applied following the same different growth conditions as applied in Chapter 6 in order to reveal which specific energy metabolism relies on the activity of the complex I-like enzyme. Hypotheses formulated based on the physiological approach are being supported by the use of the proteomic dataset from Chapter 6. The data presented in Chapter 7 lay the foundation for a manuscript in preparation which specifically addresses the question of the role of complex I-like enzymes in the metabolism of OHRB.



## **Part I**

# **Transcription regulation of organohalide respiration in Firmicutes**



## 2 A review of transcription regulation in organohalide respiration

This chapter corresponds to a modified version of the following publication :

Maillard J. and **Willemín M. S.<sup>I</sup>**, "Regulation of organohalide respiration", *Advances in Microbial Physiology*, **2019**, 74, 191-238 [7]

### 2.1 Introduction

The study of regulatory mechanisms can be approached differently ranging from global proteomic and/or transcriptomic analysis in response to the presence of a given compound to the molecular characterisation of a specific regulatory protein. Each technique has its pros and cons and a combination of global and targeted approaches remains necessary for a complete understanding. Following the first mention of gene transcription regulation in the context of OHR research [37], many studies have been focusing on the cellular level of expression of *rdhA* genes after exposure to organohalide molecules using both targeted and broader methods (note that most of them were carried out in either *Dehalococcoides* or *Desulfitobacterium*, for more details, see [7]). These different studies have revealed that the transcription patterns of *rdhA* genes follow two general trends. In *Desulfitobacterium* spp. the transcription of individual *rdhA* genes were shown generally to respond strongly to specific organohalides [34, 37, 71–73] (note that some examples of non-regulated *rdhA* were reported in *Desulfitobacterium* spp., which were explained by the presence of an insertion sequence harbouring a strong promoter directly upstream of the *rdhA* genes [32, 74]), while in *Chloroflexi* (*Dehalococcoides* and *Dehalogenimonas* spp.), studies based on the presence/absence of transcripts highlighted seemingly unregulated steady transcription patterns for most of *rdhA* genes, independently of the organohalides tested [75–77]. Indeed, exposing *D. mccartyi* strain CBDB1 to different trichlorobenzenes or chlorinated dioxins revealed the transcription of most of (if not all) 32 *rdhA* genes present in that strain [76, 77]. Similarly, 19 out of 24 *rdhA* genes were detected on mRNA level in another obligate OHRB, *D. restrictus* strain PER-K23, growing on tetrachloroethene [57]. This general trend suggested that obligate OHRB, which

---

<sup>I</sup>Contribution of the PhD candidate: Elaboration and Writing of the parts selected for the present chapter.

strictly depends on the presence of organohalides for growth, developed a strategy of low and steady transcription level for many different *rdhA* genes in order to be ready upon exposure to any new organohalide substrate.

Overall, studies of OHR regulation highlighted the existence of various regulatory strategies depending on the type of OHRB. This can be partially explained by the fact that, depending on the organism, different types of regulators are involved in the OHR regulatory networks. The next sections aims to give an exhaustive overview of the different OHR regulation strategies.

## 2.2 Transcriptional regulators involved in organohalide respiration

Obligate OHRB encode up to three dozens of *rdh* gene clusters, suggesting that there is a need for an efficient regulatory system. Indeed, inducing gene expression of a particular cluster only when the corresponding substrate is available represents an important advantage in order to avoid producing unnecessary proteins and therefore to save energy. Likewise, facultative OHRB may use alternative metabolic pathways in absence of halogenated compounds or may repress *rdh* gene clusters in presence of more favourable electron acceptors. To do so, three main types of regulatory proteins have emerged in distinct OHRB phyla, and their distribution is not related to their energetic lifestyle. CRP/FNR-like regulators were found in the close vicinity of many *rdh* gene clusters of OHRB Firmicutes (i.e. *Dehalobacter* spp. and *Desulfotobacterium* spp.). In contrast, epsilon-Proteobacteria represented by *Sulfurospirillum* spp. seem to harbour regulators belonging to the family of two-component system (TCS). Finally, OHRB members of the Chloroflexi (*Dehalococcoides* spp. and *Dehalogenimonas* spp.) display a combination of TCS and MarR-type regulators. Even though our knowledge is limited only to a few studied regulators, experimental evidence has been obtained for the implication of MarR-type and CRP/FNR-type regulators. This section offers a survey of transcriptional regulators involved in OHR and addresses their mechanism of action.

Investigating the function of transcriptional regulators makes use of complementary techniques, the combinations of which have been applied to understand OHR regulation at the molecular level. While the use of *in vitro* analyses such as electrophoretic mobility shift assay (EMSA), isothermal titration calorimetry (ITC) or electrospray ionisation mass spectrometry (ESI-MS) gives valuable information on the affinity of regulators for their effector and deoxyribonucleic acid (DNA) partners, the production and purification of the regulatory proteins are a prerequisite and represent major bottlenecks in their characterisation. Alternatively, *in vivo* reporter assays were developed in *E. coli* to study OHR regulatory proteins offering the opportunity to screen more easily for possible binding partners. The nature of the different experimental approaches, however, requires a careful comparison and interpretation of the data obtained.

### 2.2.1 Two-component systems

Regulators of the two-component system (TCS) family are highly abundant in signalling pathways. They rely on the transfer of one phosphoryl group between two distinct proteins. The prototypical TCS is composed by a sensory histidine kinase (HK) component capable of auto-phosphorylation on a conserved histidine residue. The phosphoryl group is then transferred to a conserved asparagine residue on the response regulator (RR) component. The latter event will drive conformational change in the RR protein to activate the output response (for details on the mechanism, see [78]). HK components are sensing a large variety of signals (including small molecules and physico-chemical signals) via its sensor domain [79]. It was recently reported that a majority of HK proteins are membrane-associated which allow the systems to sense extracellular signals [78]. The signals can either induce or repress HK kinase activity and therefore control the level of RR phosphorylation via its kinase and phosphatase activities. The RR proteins are equipped with an effector domain in order to translate the stimuli into a response. The majority of RR components have a DNA-binding domain and thus act as transcriptional regulators (either as activator or repressor, or both). Since the targeted DNA sequences show only low conserved sequence motifs, experimental evidence is necessary for their identification [79].

#### Two-component systems in organohalide respiration

Until very recently, no experimental evidence was available showing the implication of TCS in *rdh* gene regulation. Only one TCS had been associated with the transcription activation of gene clusters responsible for the degradation of polychlorinated biphenyls (PCBs) in the non-OHR bacterium *Rhodococcus jostii* [80]. In 2020, the first complete characterisation of a TCS-type regulators was published [81].

In order to harmonise the nomenclature of TCS potentially involved in OHR, it was proposed to refer to HK and RR components as RdhS and RdhP, respectively [30]. TCS regulators were identified in the direct vicinity of *Rdh* gene clusters in *Sulfurospirillum* spp. [38, 82], *Dehalococcoides* spp. [33, 35, 77] and in *Dehalogenimonas lykanthroporepellens* [83]. In particular, *D. mccartyi* strains CBDB1 and 195 share a relatively high number of *Rdh* gene clusters in the proximity of which TCS encoding genes have been found (see cluster 4 in Figure 1.3, as example). Here, all the HK proteins are reported to be cytosolic [33, 35]. Additionally, several TCS representatives were detected at the transcription level in *D. lykanthroporepellens* along with their cognate *rdhA* genes [75]. *S. multivorans* [38] and *S. halorespirans* [82] and two strains of *Candidatus S. diekertiae* [41] display a highly conserved OHR core region. In all these genomes two sets of TCS are encoded in the vicinity of *pceAB* and *rdhAB* genes, respectively (1.3, cluster 9). In contrast to *Dehalococcoides* spp., HK components of *Sulfurospirillum* TCS are predicted to be membrane-associated [38]. On proteomic level, only the second TCS set (TCSII) located downstream of the *rdhAB* genes has been detected in cells cultivated with different electron donor/acceptor combinations [66, 84]. Recently, TCSII was shown to promote its own transcription as well as the transcription of the *pceA* in presence of PCE but

**(A)**

**(B)**

18

### 2.2.2 MarR-type regulators

MarR comprises a family of prokaryotic regulators which takes its name from *E. coli* multiple antibiotic resistance regulator (MarR), a regulator responsible for the transcription of an operon coding for a drug efflux pump. Most MarR proteins act as homodimeric repressors by binding specific palindromic DNA sequences usually overlapping with the -35 and/or -10 elements of the promoter region, thus preventing the ribonucleic acid (RNA) polymerase to recognise the promoter. MarR-type regulators are made of a single winged helix domain [88]. Usually, MarR encoding genes are located in opposite orientation to the genes under control. Additionally, many MarR family members show an autoregulation activity. Upon ligand binding, the affinity of the regulators for their targeted DNA sequence is lowered, the promoter region is consequently released and the transcription de-repressed. Typical ligands for MarR regulators are small phenolic compounds, small peptides or metal ions [89]. Even though most MarR-type regulators act as repressors, several examples exist where the regulators act as transcriptional activators. In this case, the exact position of the targeted DNA sequences in the promoter region determines if the regulator will activate or repress gene transcription [89].

#### MarR-type RdhR regulators in organohalide respiration

In addition to TCS encoding genes, several *MarR* homologous genes are in the vicinity of *Rdh* gene clusters in *Dehalococcoides* spp. [77] and *D. lykanthroporepellens* [83]. In *D. mccartyi* strain CBDB1, an increased transcription level was observed for two MarR-associated *rdhA* genes (cbdbA1453 and cbdbA1624) upon addition of either 1,2,3- or 1,2,4-trichlorobenzenes (TCBs), strongly suggesting the implication of the corresponding regulators [76]. Later, differential transcription levels for these two *rdhA* genes were also reported for cells cultivated in presence 2,3- and 1,3-dichlorodibenzo-*p*-dioxins which further supports the existence of a complex regulatory network [77]. In addition, several MarR-type regulators were also detected at the transcription level in *D. lykanthroporepellens* [75]. RdhR was proposed as a general term to designate MarR-type proteins involved in the regulation of OHR [30]. The diversity of putative RdhR regulators in the genomes of *Dehalococcoides* spp. has been presented previously [86].

The first RdhR member to be characterised is Rdh1R (cbdbA1625) from *D. mccartyi* strain CBDB1 which showed affinity for the intergenic region that separates it from its neighbour *rdhA* gene. EMSA analysis revealed a shift dependent on the Rdh1R concentration used in the assay only when the intergenic DNA fragment containing the putative promoter regions of *rdhA* and *rdh1R* (in opposite orientation) was included. This region showed a large palindromic sequence with two perfect 20-bp inverted repeats that overlaps the transcription starting sites of both *rdh1R* and *rdhA* genes (Figure 2.1A). Each of the 20-bp inverted repeats harbours a 14-bp palindromic sequence with 4-bp inverted repeats, as in typical MarR-type regulators [88]. Finally, evidence for Rdh1R transcription activity was obtained *in vivo* by using a promoter-LacZ fusion reporter assay. Inducing the expression of plasmid-borne *rdh1R* resulted in about

60% of the beta-galactosidase activity obtained with an empty plasmid, thus indicating the ability of Rdh1R to act as a repressor of its own transcription [77].

A second RdhR from *D. mccartyi* strain CBDB1 was characterised [86]. Rdh2R (cbdbA1456) is encoded in reverse orientation upstream of two *rdhAB* gene clusters (cbdbA1452-1455). The transcription of all five genes, including *rdh2R*, was induced by 1,2,3-TCB. *In vivo* reporter assay with the promoter regions of both *rdhA* and *rdh2R* genes showed some repressor activity of Rdh2R toward both *rdhA* promoters ( $P_{1453}$  and  $P_{1455}$ ) but not toward that of *rdh2R*. DNase I foot-printing experiments revealed regions of approximately 30 bp in  $P_{1453}$  and  $P_{1455}$  as target, helping defining the DNA binding site of Rdh2R. Rdh2R is likely to recognise a 17-bp palindromic sequence embedded in two 11-bp direct repeats. The palindromic sequence is made of 6-bp inverted repeats separated by 7 bp and is fully conserved in both targeted promoters (Figure 2.1A). EMSA analysis confirmed the interaction of Rdh2R towards the binding motifs identified in  $P_{1453}$ ,  $P_{1455}$ . In addition, repressor activity and DNA-binding activity of Rdh2R was also demonstrated toward the promoter region of another *rdhA* (cbdbA1598,  $P_{1598}$ ), which also harbours the conserved palindrome. This result is in line with the observed up-regulation of the gene products of *cbdbA1453*, *A1455* and *A1598* in the proteome of *D. mccartyi* strain CBDB1 cultivated on hexachlorobenzene [90], thus suggesting a certain degree of cross-talk between specific RdhR regulators and the conserved DNA binding motifs in distantly located *rdh* gene clusters [86]. Despite many trials, however, no specific organohalide effector could be identified for Rdh1R and Rdh2R regulators [77, 86].

### 2.2.3 CRP/FNR-type transcriptional regulators

The CRP/FNR superfamily of transcriptional regulators was named after the two firstly described members. The cAMP receptor protein (CRP) and the fumarate nitrate regulatory proteins (FNR) are homologous but are dedicated to the regulation of different metabolisms [91]. Indeed, while CRP controls the catabolic repression through cyclic adenosine monophosphate (cAMP) sensing, FNR is involved in monitoring the oxygen concentration through an FeS cluster [92–95]. CRP/FNR-type regulators are constituted by an N-terminal sensor domain which recognises the dedicated signal or effector molecule and a C-terminal DNA-binding domain which will positively (or, less frequently, negatively) influence gene transcription of the genes directly downstream of targeted DNA binding sites [96]. The consensus DNA motif recognised by the C-term helix-turn-helix (HTH) domain is generally palindromic (at least partially) and located at a precise position in the promoter region of the target genes [96]. It has been shown that the CRP/FNR-type regulators can act on the expression of a relatively high number of genes at the same time (for an example, see [97]). This family of regulators is widely distributed among bacteria and the list of CRP/FNR annotated sequences is exponentially growing since the first complete phylogenetic study conducted in 2003 [96]. A more recent study was reported in 2013 where more than 1500 sequences (against 369 in 2003) were considered for the analysis [42]. Today, more than 2'500 sequences are annotated as cAMP-binding proteins in UniProtKb/TrEMBL database and new experimental data are continuously

released on members of the CRP/FNR superfamily. Many studies report on regulator-specific features and are not applicable at the level of the superfamily (for some recent and relevant examples, see [98–100]). This reflects the extremely large functional diversity found in this protein superfamily.

### **RdhK family – CRP/FNR-type regulators dedicated to organohalide respiration**

Generally, RdhK are found in *rdh* gene clusters of OHRB belonging to the Firmicutes [30]. For example, six putative RdhK were identified in *Desulfitobacterium hafniense* strain DCB-2 [34] and up to twenty-five in *Dehalobacter restrictus* strain PER-K23 [57]. RdhK6 from *D. hafniense* strain DCB-2 (initially called CprK1, [40]), and the homologous CprK protein from *Desulfitobacterium dehalogenans* [37, 101] have been extensively studied. Both proteins were shown to fulfil the same function in their respective organism, namely activating the transcription of the chlorophenol respiration (*cpr*) gene cluster in presence of 3-chloro-4-hydroxyphenylacetic acid (ClOHPA). RdhK was proposed later as a general name to refer to this type of regulators involved in OHR since the effector molecules of most RdhK proteins are unknown [30]. However, for the sake of simplicity, we will refer to the characterised RdhK proteins by using their original names throughout this review.

Historically, CprK was first mentioned as a potential regulator of the *cpr* gene cluster by Smidt and co-authors in 2000 [37] (1.3, cluster 8). This assumption was based on the low but significant sequence similarity of CprK to CRP/FNR proteins as well as on the fact that it was constitutively transcribed at low level while the other *cpr* genes were up-regulated in cells respiring ClOHPA. Moreover, FNR box-like DNA binding motifs were identified in the promoter regions of several *cpr* genes suggesting the involvement of a FNR-like regulator [37]. To date, experimental evidence has been obtained for a limited number of RdhK regulators (CprK from *D. dehalogenans* strain JW/IU-DC1, CprK1, CprK2 and CprK4 from *D. hafniense* strain DCB-2).

### **Identification of RdhK binding partners**

Several studies have investigated the binding properties of a few RdhK proteins. Although today's knowledge on RdhK is limited to a few members of the family, the findings presented here are of general interest to understand the key features of RdhK regulators. Crystal structures of CprK and CprK1 proteins have been obtained and are of great help in identifying the nature of the interactions and the conformational changes upon binding of the effector [102, 103].

*Specificity of RdhK regulators to organohalide effectors.* Both CprK and CprK1 have a strong affinity for ClOHPA. In an *in vivo* approach based on a green fluorescent protein (GFP) reporter assay, the measured dissociation constant of CprK1 for ClOHPA was in the nanomolar range ( $K_D^{ClOHPA} = 19 \pm 2$  nM) [104]. This value is significantly smaller than the constant determined

via *in vitro* methods such as EMSA, intrinsic tryptophan fluorescence or ITC ( $K_D = 0.8$  to  $4.5 \mu\text{M}$ ) [101–103, 105]. It was suggested that the presence of the DNA in the reaction may stabilise the interaction with the effector and thus influence the apparent  $K_D$  value for the effector [101]. The  $K_D$  value obtained by EMSA, based on a ternary complex of protein, effector and DNA was of  $0.4 \mu\text{M}$ . It is therefore important to consider the nature of the complex being analysed (i.e. binary vs ternary) and the experimental approach used when comparing the levels of binding affinity.

CIOHPA was also used as an effector in a  $\beta$ -galactosidase based *in vivo* reporter assay conducted with CprK1, CprK2 and CprK4 of *D. hafniense* strain DCB-2 [87]. Not surprisingly, CprK1 showed an enhanced  $\beta$ -galactosidase activity in presence of CIOHPA. A similar result was observed with CprK2. In contrast, the activity of CprK4 was shown to be independent of the presence of CIOHPA. One of the explanations for this particular behaviour was that CprK4 might be activated by cAMP present naturally in *E. coli* but this was proved to be wrong since cAMP was not promoting any additional ternary complex formation in EMSA analysis [87, 101, 105]. Gábor and co-authors also hypothesised that CprK4 might be constitutively active and that its natural ligand might enhance its activity [87]. The activity of CprK4 was tested via EMSA in presence of several phenolic compounds including CIOHPA. The results obtained confirmed the peculiar behaviour of CprK4 which showed a relatively strong shift in absence of any effector. However, the amount of observed ternary complex formed was significantly higher in the presence of several other phenolic compounds displaying a maximum activity with 2,3-dichlorophenol (DCP) and 3,5-DCP, and to a lesser extent with 2,4,6-trichlorophenol (TCP). Four additional compounds were also acting as weak effectors. When these effectors were tested on CprK1, only 2,4-DCP and 2-bromo-4-chlorophenol could enhance binding of CprK1 to DNA at a similar level as CIOHPA [87]. These results are in agreement with data obtained by EMSA in another study [105]. Since some effectors activating CprK1 were not able to enhance CprK4 DNA-binding activity and vice versa, it was proposed that CprK4 and CprK1 have complementary functions. While CprK1 activity cannot be induced by phenols displaying a halogen substitution in the meta-position, CprK4 seems more tolerant, except for long side-chains in the para-position [87]. ESI-MS analysis has revealed 2,3-DCP as a weak effector for CprK1 while 2,4,5- and 2,4,6-TCP are promoting CprK1-DNA complex formation at low concentration. However, the relatively high effector concentrations used for ESI-MS have to be kept in mind when questioning the physiological relevance of these results and comparing with data obtained by EMSA analysis [105]. Nevertheless, the ESI-MS results follow the proposed hypothesis that meta-substituted effectors do not induce DNA-binding activity of CprK1 [87].

CprK and CprK1 have the capacity to discriminate between CIOHPA and OHPA, the corresponding dechlorination product. The dissociation constant of CprK1 for OHPA was reported in the millimolar range [102, 104]. The role of the chlorine atom in CIOHPA was proposed to be of both physical and chemical nature [102, 103, 105]. First, the chlorine atom provides an additional steric bulk that results in a proper positioning of the hydroxyl group of the phenolic compound inside the hydrophobic binding pocket [102]. Secondly, the chlorine is

responsible for lowering the pKa of the phenol moiety which will favour the stabilisation in its phenolate form and allows the establishment of bonds with several residues of the binding pocket, especially with Lys-133 (see below) [102, 103, 105]. Using many compounds similar to ClOHPA in the GFP reporter assay, the affinity of CprK1 for the effector is likely dependent to the radius and the electronegativity of the substituent at the ortho-position of the phenolic group. Furthermore, the substitution of the hydroxyl group by -NH<sub>2</sub> completely inhibits effector binding supporting the importance of the interaction with the phenolate moiety [104]. In the same study, the only potential effector molecule used that is able to compete with the activity observed with ClOHPA is 3-bromo-4-hydroxyphenylacetic acid (apparent  $K_D = 0.046 \pm 0.003 \mu\text{M}$ ), suggesting that the nature of the halogen atom has only a limited incidence on binding.

As described above, all the available biochemical data on RdhK proteins concerns phenol-activated regulators. Moreover, Lys-133, which interacts with the phenol moiety, is conserved in all the studied CprK regulators [87]. Thus, examples of RdhK regulators binding non-phenolic organohalides are lacking and still need to be identified.

*Identification of DNA binding motifs.* Smidt and co-authors have identified four FNR box-like palindromic sequences in the intergenic regions of the *cprTKZEBACD* gene cluster from *D. dehalogenans* [37]. The term "dehalobox (DB)" was proposed later to refer to these DNA motifs in the context of OHR [40]. The first evidence for the recognition of such a motif by CprK was obtained via EMSA analysis where the 193 bp DNA fragment including the *cprB* promoter region was shown to be retarded in presence of ClOHPA [101]. The amount of DNA shifted was dependent on the concentration of CprK with an apparent dissociation constant ( $K_D^{DB}$ ) of  $190 \pm 30 \text{ nM}$ . By DNase I foot-printing experiments, a stretch of 14 bp containing the nearly perfect 5-bp inverted repeats (TTAATACGCACTAA) was identified at -41.5 bp upstream of the transcription start site [101]. Inspection of the corresponding *cpr* gene cluster from *D. hafniense* strain DCB-2 revealed three putative DB sequences (DB5, DB6 and DB7, see Figure 2.1B) with the consensual palindromic motif TTAAT-N4-ATTAA. All three DB sequences are centred at -41.5 bp from the transcription start site of *cprT*, *cprZ* and *cprB* genes [40, 87]. Among them, only DB7 displays perfect 5-bp inverted repeats. A value of 90 nM could be determined for  $K_D^{DB7}$  by EMSA for CprK1 [40]. DB5 and to a lesser extent DB6, showing one or two mismatches in the inverted repeats, respectively, also serve as binding sites for CprK1, as demonstrated by the *in vivo* reporter assay [40, 87]. The study of additional *cpr* gene clusters in *D. hafniense* strain DCB-2 and their cognate CprK regulators led to the identification of additional putative DB sequences (Figure 2.1B). The reporter assay was performed with CprK2 and CprK4 and the DB-containing promoters found in their respective cluster. CprK2 showed more binding affinity toward DB1 and DB3 than with DB2 and DB4. In contrast, CprK4 only gave a positive answer with the promoter harbouring DB8 [87]. The possibility of cross-talk between the RdhK regulators and DB sequences identified outside of their respective gene clusters was not tested.

### Regulation mechanism of RdhK proteins

From the comparison of the crystal structures of CprK and CprK1 in combination with their binding partners, the conformational changes associated with RdhK activity was further explored. Since both regulators share 89 percent of sequence identity and display similar binding properties [40, 101], the structural and mechanistic features of both regulators are presented together here.

*Allosteric modulation of RdhK leading to DNA binding.* The comparison between effector-bound and effector-free CprK1 (PDB: 3E5X and 3E6B, respectively) has provided a detailed understanding of the structural rearrangements occurring upon binding of ClOHPA. The next section presents the main features that have been described in detail previously [102, 103].

Basically, effector binding leads to the reorganisation of the subunits in the homodimeric CprK1 to a conformation which is compatible with DNA binding. While the reduced effector-free protein has a rather flexible structure (for redox-dependent conformational changes, see below), the overall conformation of effector-bound CprK1 is more structured. The conformational reorganisation affects most of the protein regions with exception of a portion of the central  $\alpha$ -helix connecting the sensor domain and the DNA-binding domain. One of the major movements observed is a rigid hinge motion around the serine residue (Ser-108) located between the sensor  $\beta$ -barrel domain and the central  $\alpha$ -helix. This movement leads to the closing of the effector-binding pocket and is driven by the correct positioning of the chlorine atom of the effector by several protein-effector interactions. Consecutively, the N-terminal stretch (residues 1-18) gets ordered upon binding of ClOHPA and contributes to stabilise the DNA-binding conformation via direct interactions with the DNA-binding domain [103].

Residues binding to ClOHPA were identified from the available structures. The majority of them are involved in bond formation between residues of the effector-binding domain and either the phenol hydroxyl group (residues Tyr-76, Gly-85) or the acetic acid moiety of ClOHPA (Lys-86, Thr-90 and Thr-92). In addition, Lys-133 was shown to be crucial in forming a salt bridge to the deprotonated hydroxyl group of the phenolic compound. Indeed, a K133L variant of CprK1 lost *in vivo* activity completely, while the affinity of Tyr-76 and Gly-85 variants dropped by 50- to 4000-fold compared to wild-type CprK1 [104]. Since it was observed that Lys-133 is strictly conserved in all RdhK homologues of *D. hafniense* strain DCB-2, it was proposed that only halogenated phenolic compounds can be detected by this bacterium.

The CprK1 structure in complex with ClOHPA and the dehalobox containing DNA sequence (PDB: 3E6C) highlights several contacts between specific residues of the protein and nucleotides in the dehalobox, thus offering a way to understand the DNA-binding specificity. In addition to a network of polar contacts between the phosphate backbone of DNA and the protein, nucleotide-specific contacts are established by several amino acid residues. Indeed, water-mediated polar bounds were identified between Tyr-230 and adenine-10, as well as between Thr-193 and thymine-11 of the dehalobox sequence ( $T_1 T_2 A_3 A_4 T_5 - N_4 - A_{10} T_{11} T_{12} A_{13} A_{14}$ ). It is interesting to mention that these two nucleotides are not conserved in

all the dehaloboxes from *D. hafniense* strain DCB-2 [87], suggesting a possible way for different RdhK regulators to discriminate between DNA-binding sites. Another possible protein-DNA contact was proposed to involve His-191 and the first thymine in the dehalobox [103]. Also, a hydrophobic bond links Val-192 with thymine-12 in the CprK1 structure [103]. Val-192 replaces a glutamate residue in FNR which has been shown to interact with the FNR-box [93]. Interestingly, while CprK1 showed affinity for the typical FNR-box (TTGAT-N4-ATCAA, differences to the consensus dehalobox are underlined), a valine to glutamate (V192E) variant of CprK1 had no affinity anymore for the dehalobox but partial activity for the FNR-box. Thus, Val-192 is able to discriminate between the thymine and cytosine at position 12 of the dehalobox and FNR-box, respectively. Val-192 is conserved among the RdhK homologues of *D. hafniense* strain DCB-2 while most identified dehaloboxes display a thymine at position 12 [87].

*Redox sensing activity of RdhK proteins.* One of the first observations when looking at the sequence of CprK was that the protein shows an unusually high cysteine content [37]. Moreover, early *in vitro* experiments also revealed that the protein loses its regulatory activity in its oxidised form [101]. Since other members of the CRP/FNR superfamily display redox sensing activity [96, 106], the question of a possible redox switch in the control of CprK activity was addressed [101, 107, 108]. FeS cluster reconstitution procedure was applied to rule out their presence in CprK of *D. dehalogenans* [101]. An intra-molecular disulfide bond formation was identified between Cys-105 and Cys-111 of CprK [108]. However, Cys-111 is lacking in CprK1 of *D. hafniense* strain DCB-2. Moreover, a C105A variant of CprK kept the activity of the wild-type in binding DNA [101], thus ruling out the physiological relevance of this disulfide bridge. An intermolecular disulfide bond between Cys-11 of one monomer to Cys-200 of the second monomer reversibly inactivated DNA binding of CprK dimers [108]. A model was proposed in that a correct positioning of the HTH domain at the DNA is only possible after disrupting the Cys-11/Cys-200 disulfide bridge [102, 105]. Although contrasting interpretation was given for the implication of Cys-11 in redox inactivation of CprK [107], no clear evidence could be obtained from *in vivo* data for disulfide bond formation in the cytoplasm of aerobically cultivated *E. coli*, thus going against a possible physiological redox regulation in RdhK proteins [40]. This interpretation is also supported by sequence analysis of additional RdhK proteins showing that these particular cysteine residues are not conserved [40].

### **Tentative definition of the RdhK family of transcriptional regulators**

As mentioned earlier, the RdhK family of regulators still need to be clearly defined and, given the relative low numbers of characterised RdhK proteins, it is difficult to identify specific signature sequences to RdhK that could help distinguishing them from other CRP/FNR-type regulators. Here, we propose the following criteria to help defining the RdhK family: i) RdhK is a CRP/FNR-type regulator encoded in the direct vicinity of *Rdh* genes clusters; ii) RdhK effectors are organohalide molecules; and iii) RdhK has a restricted and specific spectrum of targeted genes (by opposition to other CRP/FNR proteins which can act on a large set of

genes). The two last criteria can only be investigated through experimental work. Nevertheless, a broader view of RdhK sequence diversity can be achieved using the first criterion. Indeed, in the context of the present chapter, a BLAST analysis was performed with CprK1 (RdhK6 of *D. hafniense* strain DCB-2) as initial sequence query against the 19 available genomes of *Dehalobacter* spp. and *Desulfitobacterium* spp. found in the NCBI database (September 2018). This resulted in a selection of 188 CRP/FNR unique non-redundant sequences. The genetic environment of each of these sequences was analysed in order to locate the nearest *rdhA* gene. A final list of 97 RdhK sequences was obtained when a distance smaller than 5 kb was considered between *rdhA*-*rdhK* gene pairs ([7]).

### Diversity of RdhK regulators in Firmicutes

The resulting selection of putative RdhK sequences was aligned together with 97 sequences of curated CRP/FNR-type regulators from the Swiss-Prot database (September 2018). From the sequence likelihood tree analysis (Figure 2.2), most of the RdhK sequences appear as a rather separate family despite a couple of outliers and the relatively deeply rooted position of some RdhK nodes. Interestingly, only 13 unique RdhK belong to the *Desulfitobacterium* genus, while *Dehalobacter* spp. harbour the majority of them (84 sequences). The four (at least partially) characterised RdhK regulators (CprK from *D. dehalogenans*, and CprK1, CprK2 and CprK4 from *D. hafniense*) belong to a node containing 48 unique RdhK sequences and are intertwined with sequences from *Dehalobacter* spp. This suggests that chlorophenol-sensing RdhK regulators (and the corresponding chlorophenol dechlorinating RdhA) may be widely distributed in *Dehalobacter* spp., contrasting with the only published report of one chlorophenol-respiring *Dehalobacter* strain [111]. Finally, none of the RdhK sequences was found in both *Desulfitobacterium* and *Dehalobacter* genera and the highest sequence identity score between two RdhK sequences coming from each genus peaks at 75 percent (CprK2 from *D. hafniense* [87] and RdhK09 (AHF09378.1) from *D. restrictus* [57]). Although sequence analysis is not replacing experimental evidence, the proposed selection of RdhK sequences invites to address questions on the relationships between yet uncharacterised RdhK regulators and their possible effector and DNA partners.

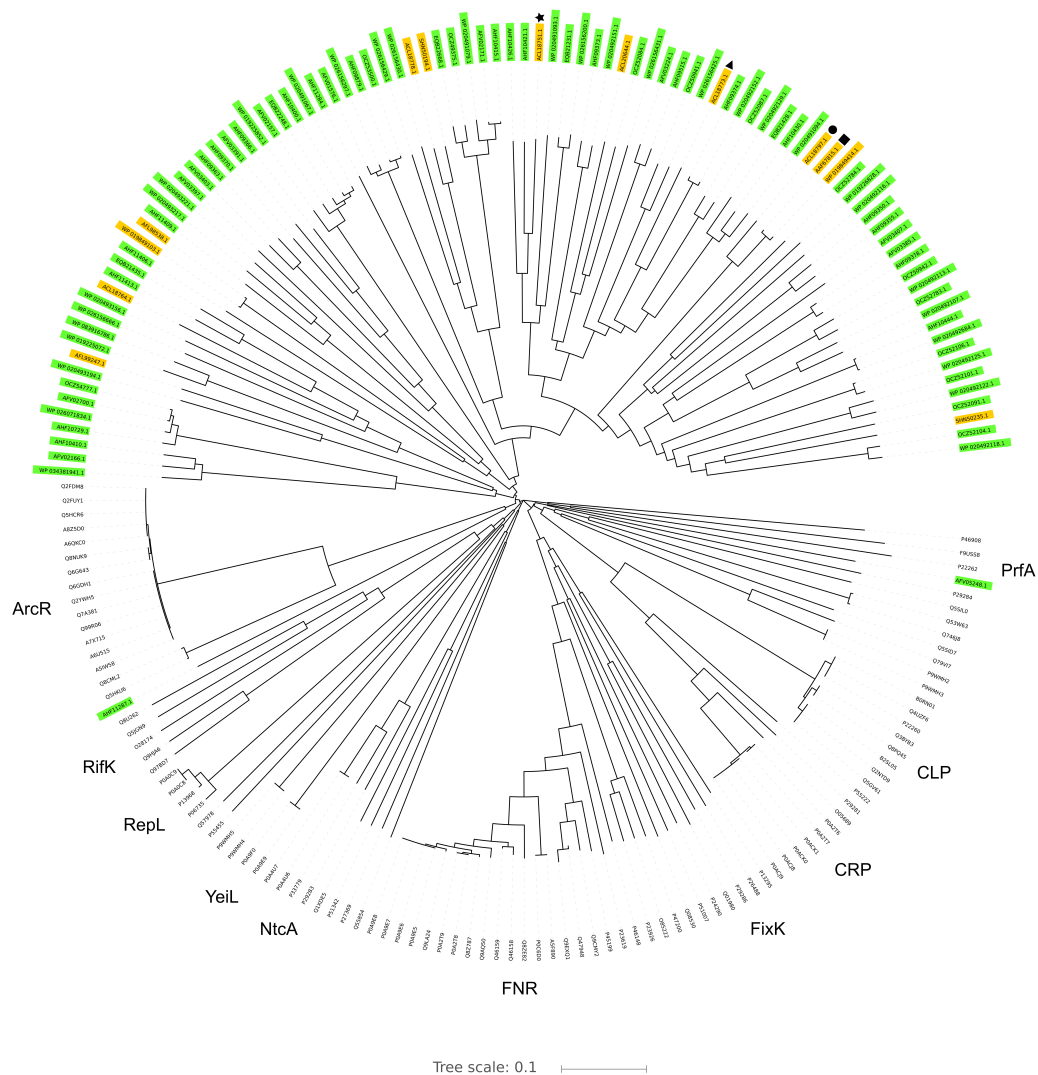


Figure 2.2 – Sequence likelihood analysis of RdhK proteins across *Dehalobacter* spp. and *Desulfitobacterium* spp. 97 non-redundant RdhK sequences were aligned with CRP/FNR entries from Swiss-Prot using ClustalX2.0 [109] and the tree was drawn using the online iTOL software [110]. RdhK belonging to *Dehalobacter* spp. and *Desulfitobacterium* spp. are highlighted in green and orange, respectively. The remaining sequences are Swiss-Prot entries for CRP/FNR-type regulators. The name of most important CRP/FNR regulators is indicated next to the tree. The characterised RdhK regulators are indicated by symbols: CprK from *D. dehalogenans* strain JW/IU-DC1 (square), *D. hafniense* strain DCB-2 CprK1 (circle), CprK2 (triangle) and CprK4 (star).



# 3 RdhK hybrid transcriptional regulators for the screening of target DNA motifs in organohalide-respiring bacteria

This chapter corresponds to a slightly modified version of the following publication :

**Willemin M. S.<sup>I</sup>**, Vingerhoets M., Holliger C. and Maillard J., "Hybrid Transcriptional Regulators for the Screening of Target DNA Motifs in Organohalide-Respiring Bacteria", *Frontiers in Microbiology*, **2020**, 11:310, 1-11 [70]

## 3.1 Introduction

Given the challenges of studying OHRB (slow growing, genetically intractable and strictly anaerobic bacteria) and the large diversity of *rdh* gene clusters encountered in most of the genomes, it is often difficult to associate a specific organohalide compound with a particular *rdh* gene cluster [7, 19, 112]. Therefore, the characterisation of RdhK proteins represents a great opportunity to explore the diversity of *rdh* gene clusters as they usually target the promoter regions of the genes responsible for the dehalogenation of their respective effector molecules. For each new RdhK protein, the identification of its preferred effector and DNA motifs represents an important challenge since both partners are interdependent in the formation of the ternary complex. In other words, the correct effector molecule is required to activate the RdhK protein of interest in order to screen for its DNA target sequences. Here we present a strategy where the complexity of the RdhK regulatory system is reduced by the use of hybrid proteins composed by the effector-binding domain (EBD) of an already characterised RdhK protein fused to the DNA-binding domain (DBD) of another member of the RdhK protein family. The advantage of this approach is to decouple the screening procedure for DNA motifs from the screening of the effector molecules. The present study

---

<sup>I</sup>Contribution of the PhD candidate: Design of the strategy, performing all the experimental assays and writing of the article.

aims at developing and validating the application of RdhK hybrid proteins for the screening of DNA motifs. To this respect, two different RdhK hybrid proteins composed by the EBD of RdhK6 (RdhK6E) that is activated by ClOHPA, and the DBD of RdhK1 (RdhK1D) that targets the dehalobox DB8, were designed and their binding properties were evaluated. Two versions of the RdhK6E-RdhK1D hybrid (in short RdhK6E1D), which differed in the position of the fusion site between the two domains, were tested both *in vitro* and *in vivo* for binding to selected organohalides and DNA motifs, and were compared to the results obtained with the native parental RdhK proteins.

## 3.2 Material and Methods

### Bacterial strains and growth conditions

For protein production, *Escherichia coli* strain BL21(DE3) carrying the expression plasmid of interest was cultivated overnight in 3 mL lysogeny broth (LB) [113] supplemented with kanamycin (30  $\mu$ M). The pre-culture was used to inoculate 500-mL of fresh medium following a 1:100 dilution. Large cultures were incubated at 37°C under agitation at 160 rpm until reaching approximately 0.6 of absorbance at 600 nm ( $A_{600}$ ). At that stage, to limit the formation of inclusion bodies, protein production was induced at 16°C by adding 0.1 mM of isopropyl- $\beta$ -D-1-thiogalactopyranoside (IPTG) and incubation was carried out for 3 h. After induction, cells were collected (7'000 x g, 4°C, and 15 min) and washed once with heparin binding buffer (see below for details) and the biomass was either stored at -80°C or directly used for protein purification.

In the context of the *in vivo*  $\beta$ -galactosidase reporter assay, *E. coli* strain JM109(DE3) carrying both the plasmid for RdhK protein production and the reporter plasmid were cultivated in 3 mL 1 $\times$  M9 medium (42 mM Na<sub>2</sub>HPO<sub>4</sub>, 22 mM KH<sub>2</sub>PO<sub>4</sub>, 8.5 mM NaCl, 18.7 mM NH<sub>4</sub>Cl supplemented with 2 mM MgSO<sub>4</sub>, 0.1 mM CaCl<sub>2</sub> and 0.2% glucose) with addition of kanamycin (30  $\mu$ g mL<sup>-1</sup>) and erythromycin (200  $\mu$ g mL<sup>-1</sup>) at 37 °C for approximately 24 h. This culture served as pre-culture for the inoculation of 20 mL cultures at a dilution of 1:100. Growth was carried out at 37°C until reaching an  $A_{600}$  value between 0.3 and 0.6. At this point, protein production was induced at 20 °C by adding 0.1 mM. When appropriate, 0.1 mM of chlorinated compounds (from a 100 mM stock solution in ddH<sub>2</sub>O) was added at the same point. Cells were then incubated for 12 h at 20 °C and 200 rpm before applying the  $\beta$ -galactosidase activity assay (see below).

### 3.2.1 Plasmids construction and DNA manipulations

#### Cloning of *rdhK* expression plasmids

All primers used in this study are given in supplementary Table A.2, while plasmids are described in Table 3.1. The sequence encoding *rdhK1* was cloned into the expression plasmid pET24d after polymeric chain reaction (PCR) amplification (with primers MW031 and MW032)

Plasmid	Description	References
pWUR176	pET24d plasmid for the expression of RdhK6	[40]
pMW021	pET24d plasmid for the expression of RdhK1	This study
pRDHK61A	pET24d plasmid for the expression of RdhK hybrid A	This study
pMW019	pET24d plasmid for the expression of RdhK hybrid B	This study
pWUR166	pAK80 plasmid for promoter fusion to <i>lacLM</i> genes	[40]
pMW032	pAK80 plasmid with DB07 containing original promoter	This study
pMW033	pAK80 plasmid with DB08 containing promoter	This study
pMW034	pAK80 plasmid with noDB containing promoter	This study

Table 3.1 – List of plasmids used in the study.

from genomic DNA (gDNA) of *Desulfitobacterium hafniense* strain DCB-2, resulting in the plasmid pMW021. The plasmid pWUR176 displaying the *rdhK6* gene of strain DCB-2 was obtained from Hauke Smidt (Wageningen University, the Netherlands). An *E. coli* codon-optimised version of the sequence encoding RdhK Hybrid A (*RdhK*<sub>61–148</sub> – *RdhK*<sub>145–228</sub>, see Figure 3.1) was first amplified using primers RdhK61A-F/R from a plasmid produced by Eurofins Scientific AG (Schönenwerd, Switzerland) and was inserted in pET24d resulting in plasmid pRDHK61A. The sequence encoding RdhK Hybrid B (*RdhK*<sub>61–188</sub> – *RdhK*<sub>185–228</sub>, see Figure 3.1) was obtained by fusion PCR. The sequences encoding the two protein domains were PCR amplified separately from gDNA of strain DCB-2 using primers MW033/MW034 and MW035/MW036, respectively and fused together in a second PCR reaction (for details, see Supplementary Information A.2). The resulting final sequence of 696 bp was cloned into linearised pET24d. The resulting plasmid was named pMW019.

### Design and construction of the DB chassis for *in vitro* analysis

For each DB of interest, a 79-bp oligonucleotide was obtained from Microsynth AG (Balgach, Switzerland). These oligonucleotides were designed as to replace the CRP-binding motif of *E. coli mgl* promoter by the 20-nt long DB motif (14-nt DB7, DB8 or a random and non-palindromic control sequence named noDB, flanked by 3 nt of the original DNA sequence) (Supplementary Figure A.1). The resulting DNA fragments, *mgl*-DB7, *mgl*-DB8 or *mgl*-noDB, were amplified by PCR using the primers DBC-F and DBC-R targeting the 5'- and 3'-end of the *P<sub>mgl</sub>* sequence, purified with the Qiagen PCR Purification Kit and quantified using the NanoDrop ND1000 apparatus.

### Cloning of reporter plasmids

All the reporter plasmids used in this study (pMW032, -033 and -034) were constructed based on pWUR166, which displays the DB7-containing promoter region of *cprBA* fused to the *lacLM* genes [40]. To generate the variants of the *cprB* promoter ( $P_{cpr}$ ) displaying alternative DB motifs, the sequence was divided in three distinct parts:  $P_{cpr}$ -5' (86 bp),  $P_{cpr}$ -3' (96 bp), both common to all promoters and flanking a central region of 50 bp harbouring the DB of interest (20-nt sequences, as above for the DB chassis) (Supplementary Figure A.2). First, the  $P_{cpr}$ -5' and -3' regions were amplified separately using genomic DNA from *D. hafniense* strain DCB-2 with the primer pairs BG1743/MW060 and MW061/BG1704, respectively. The 3'-end of MW060 and MW061 primers was designed in order to hybridise with the first and last 5 nt of the central part, respectively. The PCR fusion was performed by mixing the flanking PCR products with synthetic oligonucleotides carrying the different DB motifs in a 1:1:1 ratio, and by PCR amplification with primers BG1743 and BG1704. The resulting 222-bp DNA fragments were inserted in the linearised pAK80 vector as described previously [40].

### 3.2.2 Protein production and purification

Biomass pellets from cells producing the protein of interest were resuspended in heparin binding buffer (50 mM sodium phosphate buffer, pH = 7.2, (10 mM dithiothreitol (DTT) and (100 mM NaCl) supplemented with 1× SIGMAFAST protease inhibitor cocktail (Merck, Zug, Switzerland) and few crystals of DNase I (Merck), at a ratio of 10 mL per g of cells (wet weight) and lysed through 3 rounds of French press at 1000 psi. The soluble fraction was recovered by centrifugation (12'000 x g, 4°C, and 15 min) and loaded on a 5-mL heparin column for affinity purification (HiTrap™Heparin HP affinity column, GE Healthcare, Glattbrugg, Switzerland) attached to an ÄKTaprime™apparatus (GE Healthcare). Proteins were eluted from the column using a gradient of NaCl (up to 1 M) with heparin elution buffer and fractions were analysed by SDS-PAGE according to standard procedures. Protein fractions of interest were pooled and dialysed overnight against a phosphate buffer (50 mM sodium phosphate buffer, pH = 7.2, 10 mM NaCl and 1 mM DTT). To further increase protein purity, the dialysed fractions were concentrated (Amicon®Ultra 10 K cut-off, Merck) and purified through size-exclusion chromatography (SEC, Superose™12, 10/300 GL, GE Healthcare) in SEC buffer (50 mM sodium phosphate buffer, pH = 7.2, 100 mM NaCl and 1 mM DTT). Protein concentration was measured using the Qubit™Protein Assay kit (Fisher Scientific, Reinach, Switzerland). When needed, protein samples were further concentrated as described above.

### 3.2.3 Electrophoretic mobility shift assay (EMSA)

Aliquots of 100 ng of *mgl*-DB chassis DNA were mixed in a 15 µL reaction with 200 µmolL<sup>-1</sup> of chlorinated compound (from 2 mM aqueous solutions of CLOHPA or 3,5-DCP) and 2 µmolL<sup>-1</sup> of RdhK protein in 1× EMSA reaction buffer (100 mM Tris-HCl, pH = 8.5, 40% glycerol, 10 mM MgCl<sub>2</sub>, 500 mM NaCl, and 5 mM EDTA), which was freshly supplemented with 10 mM

DTT. The reaction was incubated for 30 min at room temperature. After incubation, 1.5  $\mu\text{L}$  of 10 $\times$  EMSA loading buffer (25% glycerol in 100 mM Tris-HCl, pH = 8.5, supplemented with a few crystals of bromophenol blue) were added to the reaction and 15  $\mu\text{L}$  of the mixture was loaded on 8% polyacrylamide gel (0.5 $\times$  Tris-borate-EDTA buffer (TBE) (45 mM Tris base, pH = 8.3, 45 mM boric acid, 1 mM EDTA), 8% acrylamide-bisacrylamide 37.5:1, 1.25% glycerol, and 1 mM EDTA) that was run for 30 min in 0.5 $\times$  TBE buffer prior to load the samples. After running the gel at 100 V for approximately 90 min, the gel was transferred in a 0.5 $\times$  TBE solution supplemented with 2  $\mu\text{g mL}^{-1}$  of ethidium bromide, and incubated for at least 30 min. UV signal was finally captured on the Universal Hood II Gel Doc System (Bio-Rad, Cressier, Switzerland), and quantified using the Image Lab Software (Bio-Rad).

### 3.2.4 $\beta$ -galactosidase activity assay

Following a 12 h period of RdhK protein induction (see the section on bacterial growth), 12 mL of the cultures were collected and centrifuged for 10 min at 4500  $\times$  g and room temperature. In order to strengthen the  $\beta$ -galactosidase activity, the cells were concentrated 20-fold by resuspending the biomass pellets in 600  $\mu\text{L}$  of growth medium. Aliquots of 100  $\mu\text{L}$  of the cell suspensions were used to measure the cell density at OD600 and the remaining samples were supplemented with 25  $\mu\text{L}$  of toluene for cell permeabilisation. The mixtures were vortexed for 30 s at full power and left on ice for 15 min before transferring 350  $\mu\text{L}$  into fresh tubes. Aliquots of 50  $\mu\text{L}$  of permeabilised cells were added to tubes containing 450  $\mu\text{L}$  of Z-buffer (per litre: 8.52 g  $\text{Na}_2\text{HPO}_4$ , 6.24 g  $\text{NaH}_2\text{PO}_4 \cdot 2 \text{H}_2\text{O}$ , 0.75 g KCl and 0.25 g  $\text{MgSO}_4 \cdot 7 \text{H}_2\text{O}$ ) freshly supplemented with 0.07%  $\beta$ -mercaptoethanol. After 10 min of temperature equilibration at 28  $^\circ$ , 100  $\mu\text{L}$  of ortho-nitrophenol- $\beta$ -galactoside (ONPG, 4  $\text{mg mL}^{-1}$  stock in Z-buffer) was added as substrate for the  $\beta$ -galactosidase. Reaction tubes were incubated at 28  $^\circ\text{C}$  until the development of faint yellow colour. At this point, reactions were stopped by the addition of 250  $\mu\text{L}$   $\text{Na}_2\text{CO}_3$  and the incubation time was recorded. Absorbance was measured at 420 nm. Miller units were calculated according to the following formula, where  $t$  is the reaction time in min and  $V$  is the volume in mL of concentrated cell suspension used in the assay:

$$\text{Activity [Miller unit]} = \frac{A_{420} \cdot 10^3}{t \cdot V \cdot \text{OD}_{600}}$$

Finally, linear regression based on least square fit method was used to evaluate the correlation between the presence of CLOHPA and an increase of  $\beta$ -galactosidase activity as well as its significance.

### 3.3 Results

#### 3.3.1 Domain definition for the design of RdhK hybrid proteins

Two different RdhK hybrid proteins were designed based on ligand-free and ligand-bound RdhK6 (CprK1) structures [103], as shown in Figure 3.1A. RdhK6 domain boundaries have been defined as follows: the EBD containing the  $\beta$ -barrel effector-binding pocket ends with residue F107 and is separated from the DBD by the central  $\alpha$ -helix region (S108-N148), which is partially affected by ligand binding. Consequently, the DBD begins with residue P149 [103]. Although structural analysis has revealed that the EBD and DBD act relatively independently from each other, both intra- and intermolecular interactions have been observed between the two domains upon dimer formation and ligand recognition [102, 103, 105]. To assess whether the domains as defined above can be completely decoupled, the EBD of RdhK6 and DBD of RdhK1 were fused at the residue N148 and P145 of the respective parent protein (see sequence alignment in Figure 3.1B), resulting in the first RdhK hybrid, named Hybrid A. In contrast, the two domains in Hybrid B were fused at the site corresponding to a conserved glycine residue in both parental proteins (G188 and G184 in RdhK6 and RdhK1, respectively). The corresponding C-terminal portion of RdhK6 (G188-Y232) is located after an amino acid stretch likely responsible for intramolecular inter-domain interactions. Moreover, residues strictly involved in DNA binding are located downstream of the conserved glycine [103].

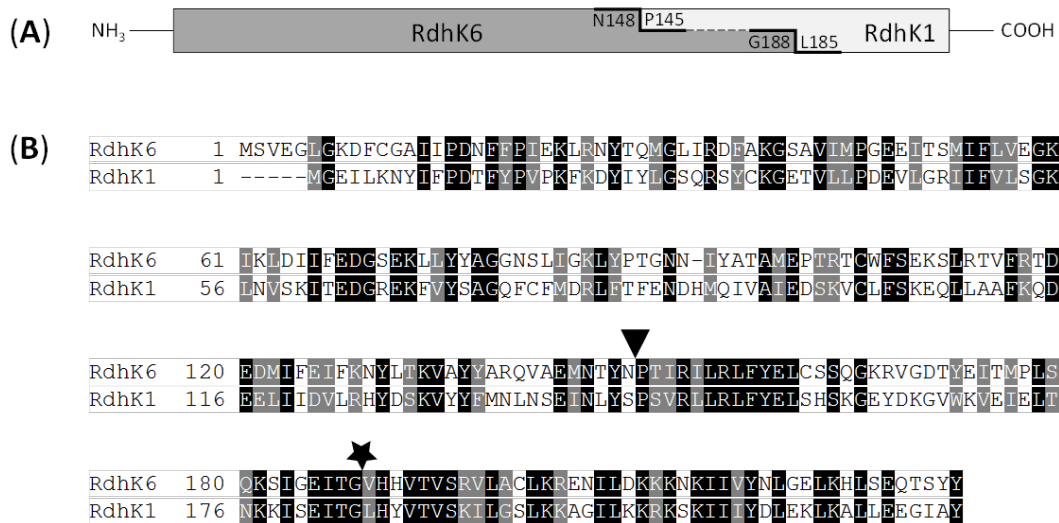


Figure 3.1 – Design of RdhK hybrid proteins. (A) Schematic representation of fusion sites. Hybrid A comprises the effector-binding domain of RdhK6 (dark grey, amino acids 1–148) and the DNA-binding domain of RdhK1 (light grey, amino acids 145–228). Hybrid B was fused further down along the RdhK6 sequence after residue G188 of RdhK6 with amino acids 185–228 of RdhK1. The residues at the fusion sites are indicated. (B) Sequence alignment of both parental RdhK proteins from *D. hafniense* strain DCB-2. The fusion points of both hybrid proteins are indicated by a triangle for Hybrid A and by a star for Hybrid B. The alignment was obtained with ClustalX2.0 and illustrated with Boxshade.

Hybrids A and B were recombinantly produced in and purified from *E. coli* along with RdhK6 and RdhK1. The binding activity of all four proteins, which were obtained in similar purity and comparable concentration (Figure 3.2), were first analysed *in vitro* using EMSA).

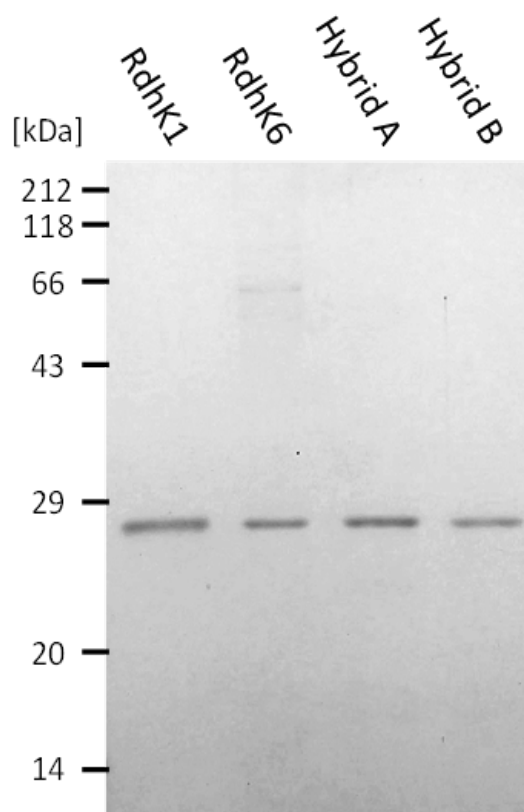


Figure 3.2 – SDS-PAGE analysis of the four RdhK proteins used in EMSA. All four proteins were recombinantly produced in *E. coli* and purified by heparin affinity chromatography followed by size exclusion chromatography. The figure shows all four proteins loaded at the same concentration as they were used in EMSA. Samples were run in a 14% acrylamide gel which was stained with Coomassie G250 following standard procedures.

### 3.3.2 *In vitro* characterisation of RdhK hybrid proteins

#### Experimental design and controls

Both RdhK hybrids along with RdhK6 and RdhK1 proteins were analysed by EMSA upon exposition to the effectors CIOHPA and 3,5-DCP, and to DB DNA motifs embedded in *E. coli mgl* promoter (Supplementary Figure A.1). To serve as negative control, a version of the *mgl* promoter was included which displayed a random and non-palindromic sequence, named noDB. No interaction was observed with this sequence for any of the four RdhK proteins, which confirmed that all observed ternary complexes described in the following sections depend on the presence of a genuine DB motif (Figure 3.3A). Additionally, a series of experiments was

also performed as control with a 1:1 mixture of the two effectors to show that no competition nor inhibition events are responsible for the absence of a ternary complex. Globally, the interactions pattern remained similar when both effectors were present (Supplementary Figure A.3).

### ***In vitro* binding of RdhK hybrids to DB8**

According to the design presented above, both hybrids were expected to bind specifically to DB8 in presence of ClOHPA. Yet, a first series of experiments was performed using DB8 as target DNA motif (Figure 3.3B). Upon exposure to ClOHPA, nearly 100% of DB8 was retained in a ternary complex by both hybrids. When ClOHPA was omitted or replaced by 3,5-DCP, no positive interaction could be observed with Hybrid B. The latter response agreed with what was expected with both hybrid proteins. In contrast, Hybrid A showed a slight binding to DB8 in presence of 3,5-DCP. Since a similar level of protein-DNA complex was also observed without effector, this most probably reflects a constitutive DNA-binding activity of this hybrid protein, which, however, was enhanced in the presence of ClOHPA. It was also noticed that the bands corresponding to DB8 in complex with Hybrid A did not appear as resolved as the ternary complexes observed with other proteins. Additionally, residual smears were observed on the top half of the lanes in EMSA experiments with Hybrid A. Consequently, the bands corresponding to protein-free DNA in these lanes were less intense even though the same initial DNA concentration was used for all reactions. These two observations may be explained by a possible aggregation of purified Hybrid A.

### ***In vitro* binding of parental RdhK to DB8**

As expected, nearly all DB8 was found in complex with RdhK1 when 3,5-DCP was used as effector molecule (Figure 3.3B). Also, in agreement with earlier work, a significant but lower amount of DB8 was retained in the complex without the addition of any effector [87]. The addition of ClOHPA did not increase the proportion of DB8 in complex with RdhK1 and indicated that this effector had likely no positive effect on RdhK1 for binding DB8. Gábor *et al.* proposed that a constitutive activation of the RdhK1 protein could explain these observations and may be considered of physiological relevance. They also suggested that only a ternary complex is able to recruit the RNA polymerase suggesting that effector-free RdhK1 protein can bind to the promoter but may not be sufficient to induce gene transcription [87], as it was also described for other members of the CRP/FNR superfamily [114]. Nevertheless, RdhK1 affinity for DB8 was enhanced by the addition of 3,5-DCP. RdhK6 was similarly tested for its interaction with DB8. Since this DNA motif is part of the *rdh-1* gene cluster, this combination has not been tested in earlier work and a positive interaction was unforeseen. However, nearly 100% of DB8 was retained in complex with RdhK6, when and only when ClOHPA was added to the reaction. This result suggested a potential cross-talk of RdhK6 with different *rdh* gene clusters, a phenomenon that was not considered so far.

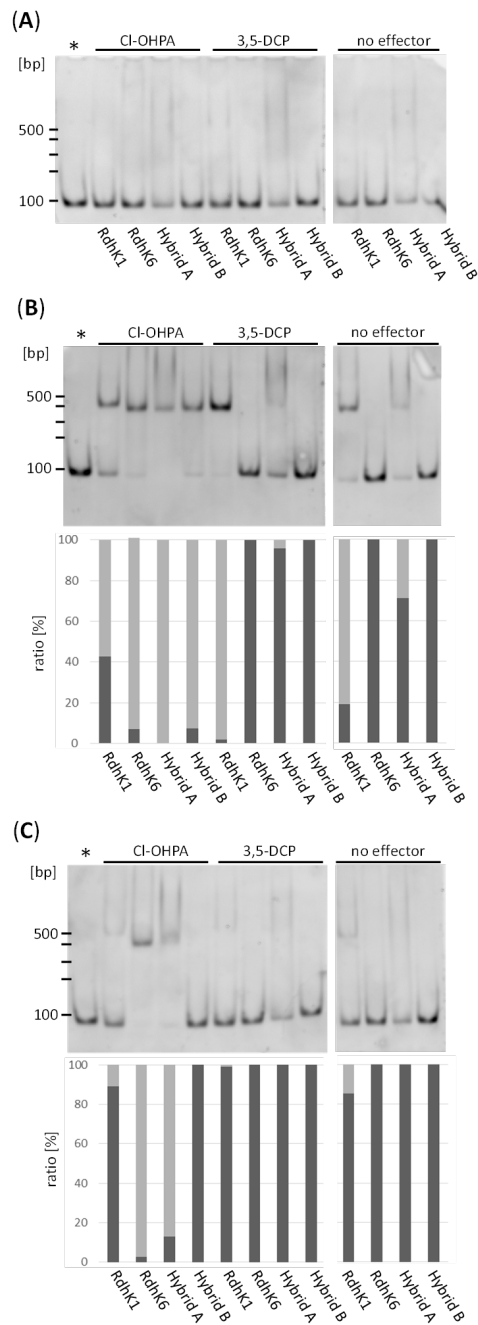


Figure 3.3 – *In vitro* interactions of the RdhK parental and hybrid proteins with dehalobox DNA motifs. Electrophoretic mobility shift assay was performed with the four RdhK proteins (RdhK1, RdhK6, Hybrid A and B) and the following three DNA sequence: no DB (A), DB8 (B) and DB7 (C). The different combinations were tested in presence and absence of two organohalide molecules, ClOHPA or 3,5-DCP. The top panel of the figure shows the gels from which DNA signals were quantified and expressed in the bottom panel (for B, C) as free DNA (dark grey) and protein/DNA complex (light grey). The asterisk indicates the position where free DNA migrated in each of the experiments.

### ***In vitro* binding of RdhK hybrids to DB7**

A second set of experiments was run with DB7 (Figure 3.3C). Since DB7 is the DNA motif targeted by RdhK6, the hybrids harbouring RdhK1 DBD were not expected to bind to this motif. No complex formation was observed with Hybrid B, even upon addition of CIOHPA. In contrast, approximately 90% of DB7 was retained in a ternary complex by Hybrid A in presence of CIOHPA. This result denoted a rather moderate DNA specificity of Hybrid A. Here again, the band corresponding to the ternary complex appeared less resolved and smears were observed in the top part of all lanes loaded with Hybrid A, as described in the DB8 experiment series. In contrast with the latter, the formation of complexes between Hybrid A and DB7 appeared to be strictly dependent on the presence of CIOHPA, as no corresponding bands were observed in its absence. This suggested that the constitutive binding activity of Hybrid A described above is only true for DB8.

### ***In vitro* binding of parental RdhK proteins to DB7**

In agreement with previously reported data, RdhK6 was able to retain almost 100% of DB7 when, and only when, CIOHPA was added to the reaction [37, 40, 87]. In addition, RdhK1 showed a weak binding affinity for DB7, independent of the presence of any effector molecule. Although this recalls the observations described with DB8, only 10% of DB7 was retained in complex by RdhK1 (against 80% of DB8, with no effector). Furthermore, unlike in DB8 series, the addition of 3,5-DCP did not enhance RdhK1 affinity for DB7. This suggests that the binding events observed between RdhK1 and DB7 are most probably the result of a residual activity of the protein while DB8 remains its true target motif.

The results presented in this section revealed that the *in vitro* DNA-binding activity of both hybrid proteins was enhanced by the presence of CIOHPA, which emphasises their ability to recognise this compound as effector and supports the production of hybrids actively binding DB motifs by fusing the EBD and DBD originating from different RdhK proteins. The results of EMSA experiments also highlighted that Hybrid A has only a limited DNA-binding specificity whereas Hybrid B is highly specific for the targeted DB as well as the effector. This shows a promising potential of the latter design for the investigation of uncharacterised RdhK proteins.

### **3.3.3 *In vivo* characterisation of RdhK hybrid proteins**

#### **Control and preliminary experiments**

To validate the data obtained *in vitro* and to further evaluate whether the two hybrids can act as genuine transcription activators beyond promoter binding, both hybrids were investigated with the *in vivo*  $\beta$ -galactosidase reporter assay as described earlier [40]. Hybrids A and B were expressed in *E. coli* cells carrying a reporter plasmid with the  $\beta$ -galactosidase gene placed under the control of promoters carrying either the DB7, DB8 or noDB DNA motif [40]. To focus on the specificity of the DB motifs exclusively, the native promoter region of *cprBA* (harbouring

DB7 or replaced, when applicable, by DB8 or noDB) from *D. hafniense* strain DCB-2 was used (Supplementary Figure A.2). A preliminary *in vivo* reporter assay was first performed in small volumes to screen for positive interactions between RdhK proteins and the DB motifs in presence of CIOHPA (Supplementary Figure A.4). With noDB, the correlation between the addition of CIOHPA and the increase of  $\beta$ -galactosidase activity was either negative or not significant (p-value > 0.05). Therefore, this control DNA motif was not further investigated. Experiments involving DB7 and DB8 were then performed in larger culture volumes to improve signal to noise ratio.

#### ***In vivo* binding affinity of RdhK hybrids for DB8 and DB7**

Upon addition of CIOHPA, a significant increase in  $\beta$ -galactosidase activity of 52- and 53-fold (p-value < 0.05) was observed with DB8 for Hybrid A and B, respectively (Figure 3.4). This result demonstrated that both hybrids can act as transcription regulator and recognise DB8 as target DNA motif. In the absence of CIOHPA, the  $\beta$ -galactosidase activity with both hybrids dropped to a level similar to the one obtained with cells carrying the empty pET24d vector instead of the hybrid expression plasmids. This confirmed that the enhancement of the  $\beta$ -galactosidase activity was dependent on the production of the corresponding hybrid proteins. When DB7 was used in combination with Hybrid B, the increase of  $\beta$ -galactosidase activity in presence of CIOHPA was only 2-fold, which likely suggested a poor affinity for this DNA motif in comparison to DB8. In contrast, the  $\beta$ -galactosidase activity measured with Hybrid A with DB7 increased about 250-fold upon addition of CIOHPA. The increase observed with DB7 was approximately five times more important than with DB8, which may indicate a higher affinity of Hybrid A for DB7 than for DB8. This trend was not observed *in vitro* and thus emphasises the need of using both *in vitro* and *in vivo* strategies to fully describe the behaviour of RdhK hybrid proteins. Globally, the results obtained from *in vivo* experiments with both hybrids were in good agreement with the conclusions drawn from the *in vitro* data. However, one aspect of the Hybrid A *in vitro* response did not have a visible counterpart in the *in vivo* results. Indeed, in the latter, Hybrid A activity remained dependent on the presence of CIOHPA at any time. Thus, the constitutive *in vitro* DNA-binding activity proposed for this hybrid is most likely an artefact resulting from the unstable behaviour of the purified protein rather than a genuine DNA-protein interaction responsible for gene transcription. Moreover, this observation could also be explained by the fact that a protein-DNA complex in absence of the effector may not be sufficient to properly induce gene transcription.

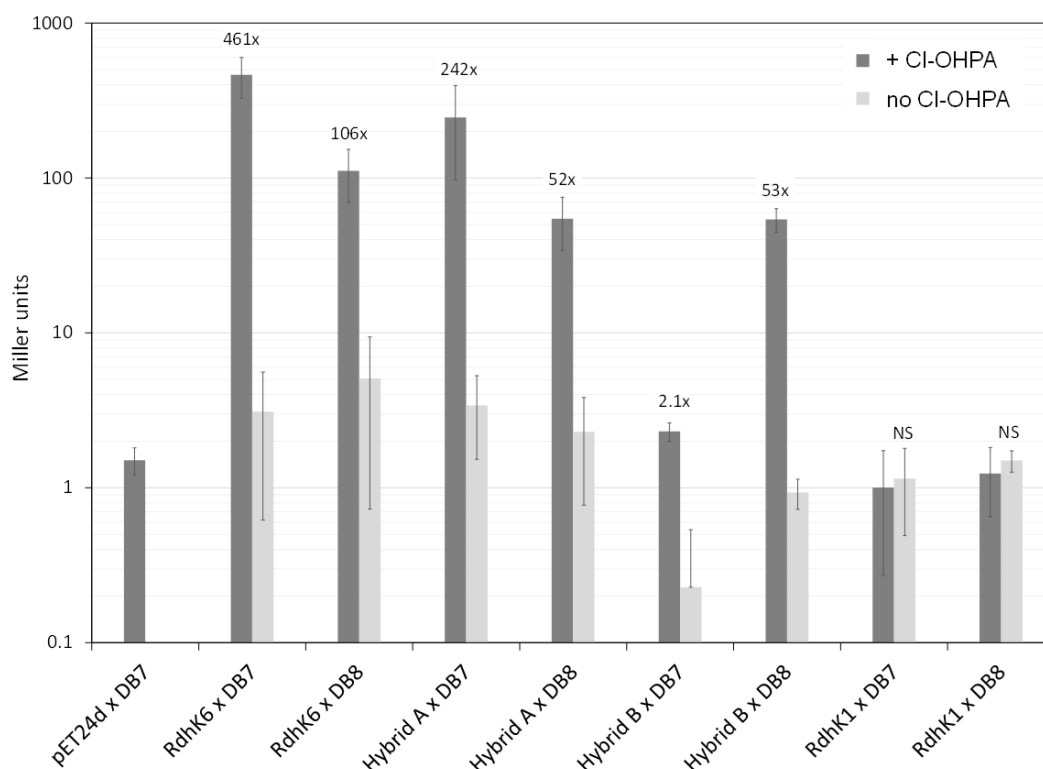


Figure 3.4 – *In vivo* beta-galactosidase reporter assay of RdhK parental and hybrid proteins. Each of the four proteins (RdhK6, RdhK1 and Hybrid A and B) was tested for their ability to promote  $\beta$ -galactosidase activity under promoters that display either the DB7 or DB8 motif. The resulting  $\beta$ -galactosidase activity was measured and is expressed in Miller units. ClOHPA was added as effector (dark grey bars). A control was included for each RdhK/DB combination, where no effector was added to the culture during protein expression (light grey bars). All experiments were performed in three biological replicates.

### 3.3.4 *In vivo* binding affinity of parental RdhK for DB8 and DB7

All the experiments described above were also performed with the two parental RdhK proteins. Despite many trials, no  $\beta$ -galactosidase activity could be measured with RdhK1. This observation is different from what has been described in the work by Gábor *et al.*, in which the activity of RdhK1 was significantly higher when it was combined with DB8 in comparison to other DB motifs [87]. An additional set of experiments was performed using 3,5-DCP as effector molecule for RdhK1 but resulted also in no  $\beta$ -galactosidase activity (data not shown). These negative results indicated that the experimental conditions applied here prevented the production of an active form of RdhK1 or that the design of the reporter platform based on *cprBA* intergenic region may not be compatible with RdhK1 recognition of DB8. In contrast, RdhK6 in combination with DB7 gave the highest  $\beta$ -galactosidase activity with a significant increase of 461-fold upon addition of ClOHPA. Also, in agreement with the *in vitro* results,

an increase in activity of 106-fold was measured when RdhK6 was combined with DB8. This result indicates again a potential physiological cross-talk between two *rdh* gene clusters in *D. hafniense* strain DCB-2.

### 3.4 Discussion

#### 3.4.1 RdhK hybrid proteins are active regulatory proteins

In this study, we show that the two domains (EBD and DBD) coming from different native RdhK proteins can be fused and the resulting hybrid proteins can be produced heterologously in *E. coli* as soluble and active proteins. Even though the results obtained with Hybrid A and Hybrid B are different in their binding properties, both proteins were active *in vitro* as well as *in vivo*. As expected, the hybrid proteins showed binding affinity for the DNA motif DB8 (initial target of RdhK1) which was enhanced in presence of CIOHPA (effector molecule of RdhK6). Thus, both hybrid designs proposed in this work gave rise to actual transcription regulators with binding properties that were neither equivalent to RdhK6 nor RdhK1. One aspect that Hybrid A and B had in common was the relatively low  $\beta$ -galactosidase activity. Indeed, the values obtained upon addition of CIOHPA for the hybrid proteins in combination to DB8 were always significant but remained relatively low in comparison to the activity measured with RdhK6 and DB7 (a mean value of 465 Miller units for RdhK6/DB7 against 55 for the hybrids/DB8). The RdhK6 activity is in good agreement with former reports [40, 87, 101], which indicates that the experimental set-up is likely not the reason for the lower activity measured with both hybrids. Pop *et al.* reported a higher activity upon increasing the final concentration of CIOHPA in the cultures [101]. Here, the addition of 20 mM CIOHPA did not result in an increase of the activity for the hybrid proteins (data not shown). A reduced expression rate or a reduced solubility of the hybrid proteins compared to RdhK6 might be the reason for the difference in activity. However, under the same experimental conditions, the unforeseen activity obtained with Hybrid A in combination with DB7 was much higher than with DB8 (a mean value of 246 Miller units) suggesting that an impaired protein expression may not be the only reason for the reduced activity. The values reported by Gabor *et al.* for their experiments involving RdhK1 and DB8 are in the same range as the ones we obtained with the hybrid proteins in combination with DB8 embedded in the *cprBA* promoter region [87]. Unfortunately, we could not confirm the values for RdhK1 activity but the fact that, globally, all the results obtained with DB8 gave lower activity (in absolute terms) suggests that the promoter sequence beyond the DB motif may also play a role. Moreover, DB8, unlike DB7, is only partially palindromic, therefore possibly explaining the lower activity.

#### 3.4.2 Hybrid B is more specific than Hybrid A

The comparison of the two hybrid proteins in the *in vivo* and *in vitro* approaches revealed a higher DNA-binding specificity of Hybrid B over Hybrid A, as the latter recognised both

DB8 and DB7. The *in vivo* experiments even showed a higher response when Hybrid A was combined with DB7 and may indicate a higher affinity for this motif. In addition, Hybrid A gave an unclear *in vitro* response that can likely be explained by a limited protein stability upon purification. Only Hybrid B showed a consistent ON/OFF response upon addition of CIOHPA, both *in vitro* and *in vivo*. The fact that Hybrid A is less specific in DNA-binding than Hybrid B remains difficult to explain since the former protein has a higher portion of RdhK1 C-terminal domain than the latter and should, consequently, be as stringent for DNA motifs as the parental RdhK1 protein if only the DBD is responsible for DNA recognition. This suggests that the protein-DNA interaction involves other residues than initially thought and may not be after all restricted to the very last region of the C-terminus. However, both hybrid proteins were able to recognise CIOHPA as effector, while 3,5-DCP did not activate them. Despite the fact that Hybrid A has a shorter portion of RdhK6 EBD, both hybrid proteins display all the residues that have been previously reported to be important for the interaction with CIOHPA [102–104]. Furthermore, since RdhK1 and RdhK6 are relatively close at the amino acid sequence level (89% identity), many of the residues highlighted in previous studies are essentially conserved in the two RdhK proteins and therefore also in the two hybrid proteins. The global conformation or the stability of both Hybrids may be partially responsible for the distinct binding patterns. Probably only a structural approach would allow us to clarify why Hybrid A is less specific in DNA binding. Nevertheless, our results showed that the strategy of domain fusion in Hybrid B created a higher specificity for the effector and for DNA than that of both parental proteins. Indeed, Hybrid B showed a higher effector dependency than RdhK1 and a higher DNA specificity than RdhK6.

### 3.4.3 RdhK6 recognises DB motifs located in other gene clusters

Beside the results obtained from the RdhK hybrid proteins, the different experiments performed in the context of this study revealed an interesting aspect of RdhK6. As expected, RdhK6 showed a high affinity for DB7 in presence of CIOHPA, confirming previously reported results [37, 40, 87, 101–103]. However, our dataset also highlighted the binding of RdhK6 to DB8 both *in vitro* and *in vivo*. In the latter case, CIOHPA-induced expression of the  $\beta$ -galactosidase was significantly lower for DB8 than for DB7, which may indicate a lower affinity of the protein for this DB motif and/or a weaker RNA polymerase activity when the DB8 motif replaced DB7 in the native promoter sequence. In any case, this observation points toward a new feature of RdhK proteins that has not been investigated in earlier work. Indeed, as DB8 is located in the *rdh-1* gene cluster (harbouring *rdhK1*), this result suggests a possible cross-talk of RdhK proteins with the different *rdh* gene clusters. Comparing the results of the present study with earlier studies, it appears that RdhK6 has affinity for all the DB tested so far, which include the three DB sequences identified in its own gene cluster [87], the modified version of the FNR box [40] and DB8 (this study). The alignment of these selected DB motifs shows that only the first two nucleotides (and the corresponding last two positions) are fully conserved in the palindrome (Supplementary Figure A.5), thus suggesting that RdhK6 regulation network might be broader than initially thought.

### 3.5 Conclusion

The present work shows the potential of using RdhK hybrid proteins to screen for DNA motifs targeted by yet uncharacterised RdhK proteins. Two different hybrid proteins were successfully produced in an active form and gave distinct response by *in vitro* and *in vivo* characterisation. While Hybrid A showed a mixed behaviour between the constitutive activation of RdhK1 and the relaxed DNA motif specificity of RdhK6, Hybrid B displayed a reliable and specific response both *in vitro* and *in vivo*. The latter appears as the design of choice to serve as a basis to create new RdhK hybrids for the investigation of uncharacterised RdhK DNA-binding domains by fusing them to the effector-binding domain of RdhK6. The proposed strategy based on the design of RdhK6E hybrid proteins could be applied to investigate the OHR regulation network in *Dehalobacter restrictus* strain PER-K23. This strain harbours up to 22 members of the RdhK protein subfamily, but is so far only known to dechlorinate tetra- and trichloroethene. Given the fact that many families of transcription regulators belong to the superfamily of CRP/FNR proteins [42], this strategy could provide a tool for the screening of DNA motifs for virtually any CRP/FNR protein.



## 4 Characterisation of new RdhK regulators using cross-species hybrid proteins and a genome-scale screening methodology

### 4.1 Introduction

RdhK proteins are dedicated to the transcriptional regulation of *rdh* gene clusters in the Firmicutes. Given the high number of *rdhK* genes homologues found sometimes in a single OHRB genome, there is a need for efficient screening methods to help their characterisation and to understand their associated regulatory networks. For example, the genome of *Dehalobacter restrictus* strain PER-K23 displays 24 *rdh* gene clusters and a total of 25 RdhK proteins are encoded in their direct vicinity (see Figure 4.1A) [57]. All but one *rdhA* genes are coming along with an *rdhB* subunit, which are located most of the time upstream of their cognate *rdhA*. Given the genome organisation of the OHR region in *D. restrictus*, it is likely that each of the 25 RdhK regulators is dedicated to the regulation of the *rdhA*, *rdhAB* or *rdhBA* present in its close vicinity. However, this hypothesis has never been tested experimentally and, as seen in the previous chapter, RdhK regulators do not seem to be necessarily restricted to the regulation of one single *rdh* gene cluster.

The RdhK hybrid approach presented in Chapter 3 is proposed to help in the characterisation of unknown RdhK proteins [70]. None of the 25 RdhK proteins of *D. restrictus* was characterised so far. Thus, this strain represents the perfect playground to put the RdhK hybrid strategy in application. To this purpose, a few RdhK hybrids were developed by fusing the effector-binding domain of RdhK6 from *D. hafniense* strain DCB-2 to the DNA-binding domain (DBD) of selected RdhK proteins of *D. restrictus* strain PER-K23. So far, the successful production of RdhK hybrids in a soluble form was only achieved for one candidate harbouring RdhK16 DBD. This hybrid protein was thus tested for *in vitro* interaction with the promoter regions located upstream of all the *rdhA* genes by EMSA. The results, presented in the first part of this chapter, show possible interactions with a limited number of promoter regions which invites to further investigate this particular RdhK hybrid and to apply this strategy with new regulators.

For a complete understanding of the OHR regulatory network involved in an OHRB, not only numerous regulators have to be characterised but each of them has to be assessed for interaction with a large number of possible DNA targets. The identification of the promoters targeted by a given RdhK can be identified using *in vitro* assays like EMSA. However, this type of approaches requires a prior knowledge of the expected DNA motifs which have to be amplified independently for the purpose of the assay. Thus, the number of DNA motifs that can be simultaneously tested is limited. Moreover, the purification of the RdhK protein (or RdhK hybrid protein) is time-consuming and represents a serious bottleneck given their relatively poor solubility when heterologously expressed. In the second part of this chapter, preliminary steps in modifying a method for a genome-wide screening of DNA binding sites for its application with the RdhK hybrid system are presented. The encouraging results obtained with RdhK6 and Hybrid A show the potential of this method for the identification of all possible targeted DNA regions in a genome by any yet uncharacterised RdhK regulator.

## 4.2 Material and Methods

### 4.2.1 Plasmids construction and DNA manipulations

#### Cloning of *rdhK* expression plasmids

See Chapter 3, Material and Methods and Table 4.1.

Plasmid	Description	Reference
pWUR176	pET24d plasmid for the expression of RdhK6	[40]
pRDHK61A	pET24d plasmid for the expression of RdhK hybrid B	[70]
pMW026	pET24d plasmid for the expression of $RdhK6_{1-188}^{Dha} :: RdhK16_{185-228}^{Dre}$	This study

Table 4.1 – List of plasmids used in the study.

#### Cloning of cross-species *rdhK* hybrid expression plasmids

The plasmid expressing  $RdhK6_E^{Dha}16_D^{Dre}$  (or  $RdhK6_{1-188}^{Dha} :: RdhK16_{185-228}^{Dre}$ ) was obtained by fusion PCR as described before (see Hybrid B cloning, in Chapter 3). This RdhK hybrid is composed by the same effector-binding domain (EBD) as Hybrid B but combined with the DNA-binding domain (DBD) belonging to RdhK16 of *D. restrictus* (*Dre*). The DBD was obtained by amplification using MW048/MW054 primers on *Dre* gDNA (Table 4.2). Except for this, the same cloning procedure was performed and resulted in a pET24d based plasmid expressing  $RdhK6_{1-188}^{Dha} :: RdhK16_{185-228}^{Dre}$  (see Table 4.1).

Primer	Sense	Sequence (5' → 3')	Target	Reference
MW048	+	CCGGTCTTCACTTTGTAAC GGTAAGTAAAATATTAGG	<i>RdhK16<sup>DRE</sup></i> (DBD)	This Study
MW054	-	GCGCGGATCCTCAGTAAGG TATGCCTTCCTCAAACAGTTC	<i>RdhK16<sup>Dre</sup></i> (DBD)	This Study
MW081	+	ATGACGGCTCTGCGCAAT	Endogenous DB7	This Study
MW082	-	TTCACCTCCCTTTAGCAAGA	Endogenous DB7	This Study

Table 4.2 – List of primers used in the study.

### Amplification of *rdhA1-24* promoter regions

The regions preceding the 24 *rdhA* genes (or *rdhB* genes in case of *rdhBA* configuration) genes were amplified from *D. restrictus* strain PER-K23 using the primers listed in supplementary Table B.1.

### 4.2.2 Protein production and purification

As described before. See Chapter 3, Material and Methods.

### 4.2.3 Electrophoretic mobility shift assay (EMSA)

As described before. See Chapter 3, Material and Methods.

### 4.2.4 Genome-scale pull-down assay

The method described here was adapted from the study published by Barlet *et al.* in 2017 [115]. The principle of this method relies on the interactions between an affinity-tagged version of the protein of interest that is immobilised on beads with fragmented genomic DNA. The DNA fragments targeted by the DNA-binding protein are thus recovered from the beads and identified using next-generation sequencing. In the modified version of the protocol that was applied in the present study, the anti-RdhK6 antibody was cross-linked to Protein G-coated beads, while ClOHPA was added to the mixture during the DNA-protein interaction step in order to bring RdhK6 in an active DNA-binding conformation.

### Preparation of the anti-RdhK6 associated beads

This procedure explains the preparation of the amount of beads necessary for one pull-down experiment. For one sample, 50 µL of Dynabeads™ Protein G (Thermo Fisher Scientific) was transferred in a 1.5-mL tube. The supernatant was removed using the DynaMag™ magnetic rack (Thermo Fisher Scientific) and discarded. Directly after that (preventing the drying of beads), the beads were incubated with 100 µL (i.e. roughly 5 µg) of purified antibody raised

against RdhK6 from *Desulfitobacterium hafniense* strain DCB-2 (R $\alpha$ RdhK6) for 30 min at 4°C on a overhead shaker. Following this incubation, the supernatant containing the unbound antibody was discarded and the beads were washed twice with 200  $\mu$ L of conjugation buffer (20 mM Na-phosphate buffer, 150 mM NaCl, pH = 7.2). To avoid recovering the bound antibody in the last wash or elution steps, the R $\alpha$ RdhK6-beads were treated with 250  $\mu$ L of fresh cross-linking solution (5 mM of Bis(sulfosuccinimidyl)suberate (BS3, Thermo Fisher Scientific) in conjugation buffer). After 30 min of cross-linking at room temperature (RT), the reaction was quenched by the addition of 12.5  $\mu$ L of Tris-buffer (1 M Tris-HCl, pH = 7.5). Quenching was carried out for 15 min at RT. Finally, the beads were conditioned by three consecutive wash steps in 200  $\mu$ L of PBS-T (1 X PBS with 0.02 % tween®-20).

### Antigen fixation on the beads

Soluble protein extracts of *E. coli* BL21 expressing RdhK6 or RdhK Hybrid B were prepared following the same procedure as in the protein purification protocol, except that the Heparin binding buffer was replaced by PBS supplemented with 1 $\times$  SIGMAFAST protease inhibitor cocktail (Merck, Zug, Switzerland) and a few crystals of DNase I (Merck) (see Chapter 3, Material and Methods). The soluble extracts were concentrated approximately 10-fold using Amicon® with 10 kDa-cutoff. After the last wash steps of the beads and removal of the supernatant, 400  $\mu$ L of the concentrated soluble protein extract were applied to the anti-RdhK6 associated beads, and incubated over night at 4°C. The next morning, the supernatant was collected and kept as flow-through (FT1) fraction. The beads were washed four times with 500  $\mu$ L of PBST (1X PBS, 0.005% Tween™-20) to completely remove the DNase which would compromise the subsequent work with DNA. Two more wash step were performed with 500  $\mu$ L of PBS only to remove the traces of the detergent which could affect the stability of the ternary complex.

### DNA preparation and pull-down assay

Aliquots of 2  $\mu$ g of gDNA from *Desulfitobacterium hafniense* strain DCB-2 were fragmented in PBS using a Covaris ultrasonicator (Covaris S220 with the following settings: average incident power = 17.5 W, peak incident power 175.0 W, duty factor = 10%, cycles burst = 200 counts, duration = 180 s). After sonication, the DNA fragments were recovered in 40  $\mu$ L of elution buffer (Qiagen AG, Hombrechtikon, Switzerland) using the Agencourt DNA extraction kit (AMPure XP kit, Beckman Coulter). The R $\alpha$ RdhK6-beads were stabilised in 40  $\mu$ L of PBS containing 10 mM of DTT. Besides, a 40  $\mu$ L mixture of fragmented DNA, ClOHPA and DTT in PBS (with final concentrations of 10 mM DTT, 800  $\mu$ M ClOHPA and approximately 10 ng of fragmented DNA) was prepared and applied to the beads. The tripartite binding reaction was carried out for 2 h at RT on the overhead shaker (at the minimum rotation rate to maximise the interactions). After incubation, the beads were collected on the magnetic rack, the supernatant was removed (fraction FT2) and the beads were washed four times in PBST. Fractions FT2, W1, W2, W3 and W4 were kept for analysis. Two more wash steps were performed in PBS-only. Upon the last

wash, the beads were transferred into a fresh tube for elution. Finally, 25  $\mu$ L of elution buffer was added to the beads which were incubated for 10 min at 95°C. The fractions were analysed by Western blot and PCR following standard protocols.

## 4.3 Results and discussion

### 4.3.1 *In vitro* identification of DNA motifs targeted by cross-species hybrid RdhK

Several RdhK hybrid proteins composed of the effector-binding domain (EBD) of RdhK6 from *Desulfitobacterium hafniense* DCB-2 (*Dha*) and the DNA-binding domain (DBD) of yet uncharacterised RdhK regulators from *D. restrictus* (*Dre*) were produced based on the design of Hybrid B (see Chapter 3). The gene fusions were cloned and heterologously expressed in *E. coli*. A representative example carrying the DBD of *RdhK16<sub>Dre</sub>* was successfully produced in a soluble form and partially purified for preliminary *in vitro* characterisation.

#### *RdhK6<sub>E<sup>Dha</sup></sub>16<sub>D<sup>Dre</sup></sub>* design and production

As for Hybrid B, the conserved glycine residue was chosen as fusion site and *RdhK16<sub>Dre</sub>* DBD, thus, consists of residues G185 to Y228. The resulting protein *RdhK6<sub>1-188<sup>Dha</sup></sub>::RdhK16<sub>185-228<sup>Dre</sup></sub>* (or to simplify, *RdhK6<sub>E<sup>Dha</sup></sub>16<sub>D<sup>Dre</sup></sub>*) was cloned, produce and purified as previously described [70]. The level of purity obtained for *RdhK6<sub>E<sup>Dha</sup></sub>16<sub>D<sup>Dre</sup></sub>* was lower (appr. 70% estimated visually from the Coomassie-stained gel) than what was achieved for the different RdhK proteins presented in Chapter 3. Nevertheless, the band corresponding to *RdhK6<sub>E<sup>Dha</sup></sub>16<sub>D<sup>Dre</sup></sub>* was dominating in the sample and was used for EMSA experiments.

#### *In vitro* screening of RdhK16 DNA targets

The 24 *rdh* gene clusters of *D. restrictus* are potentially regulated by RdhK16. In order to test this hypothesis, the genomic regions preceding each *rdhA*, *rdhAB* or *rdhBA* genes were amplified by PCR. As shown in Figure 4.1A, a total of 12 of these promoters contain a possible dehalobox DNA motif [116] which are listed in Figure 4.1B. The affinity of RdhK16 DBD for any of these 24 regions was tested *in vitro* by EMSA using purified *RdhK6<sub>E<sup>Dha</sup></sub>16<sub>D<sup>Dre</sup></sub>* (see previous section) using ClOHPA as effector molecule. Indeed, *RdhK6<sub>E<sup>Dha</sup></sub>16<sub>D<sup>Dre</sup></sub>* is displaying the same effector domain as Hybrid B and is therefore activated by ClOHPA. The formation of potential ternary complexes was observed for a few DNA targets as shown in Figure 4.1C. In particular, *RdhK6<sub>E<sup>Dha</sup></sub>16<sub>D<sup>Dre</sup></sub>* seemed to show some affinity for the promoter regions of *rdhA6*, *rdhBA10*, *rdhBA11*, *rdhBA16* and *rdhBA21*. Indeed in the lanes corresponding to these samples, one or several bands were observed in the upper part of the gel.

Important control experiments are still needed to draw any conclusion from this preliminary work. For example, all these experiments were done in presence of ClOHPA and should be repeated in its absence, to show the dependency of the observed ternary complex to

the presence of the effector. Furthermore, if the higher bands truly correspond to positive interactions, the proportion of DNA target retained in the complex remains relatively small in comparison to the band corresponding to free DNA. These interactions seemed to be significantly weaker than that from Hybrid B, as shown previously (Chapter 3 and [70]). It may be that the amount of active RdhK hybrid protein was too low (either because of partial purity or simply because of poor stability of the protein) to fully retain most of the DNA. A thorough experiment investigating the ratio of protein and DNA necessary to maximise the formation of ternary complexes would help to better understand the affinity of the hybrid protein for the DNA targets.

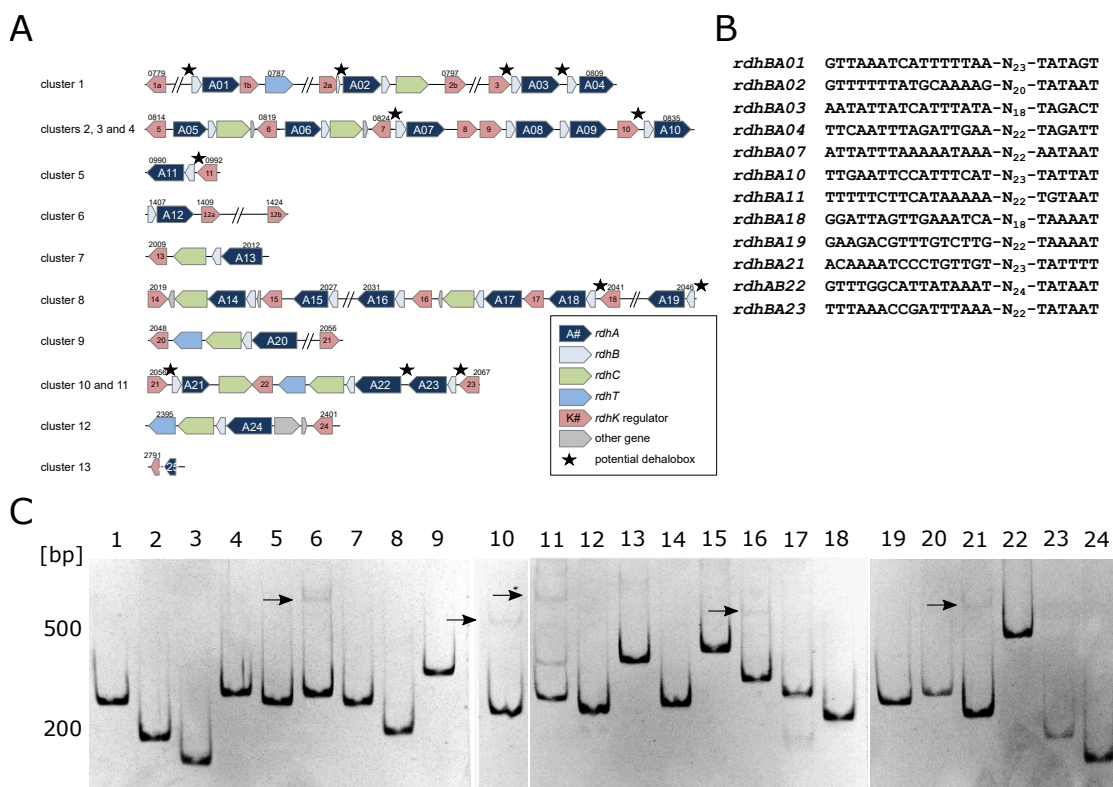


Figure 4.1 – Characterisation of a cross-species RdhK hybrid protein towards the promoters of all *rdh* gene clusters from *D. restrictus*. (A) Schematic representation of the 24 *rdh* gene clusters of *D. restrictus* strain PER-K23. The colour code corresponds to the different type of genes composing the clusters. In addition, the *rdhA* genes are numerated as well as their potentially-associated *rdhK* regulator. Finally, the promoter region containing a sequence resembling a dehalobox are indicated by a star. Adapted from [57]. (B) Sequences of the 12 potential dehalobox dehaloboxes detected in the promoters of *rdhAB* or *rdhBA* genes [116]. (C) *in vitro* characterisation of ternary complexes formed by  $RdhK6_E^{Dha}16_D^{Dre}$  with the 24 promoter regions of *D. restrictus* using EMSA. Promising signals indicating a potential ternary complex are indicated by a small arrow.

Despite the cytoplasmic nature of RdhK proteins, the solubility of recombinant RdhK proteins have been an issue that was recognised in early studies [87]. As for the production of native RdhK proteins, several attempts to produce the necessary RdhK hybrids often resulted in an insoluble form, preventing their purification and further use in *in vitro* experiments (data not shown). Therefore, efforts still need to be done to circumvent this difficulty or to find alternative methodologies to study them.

#### 4.3.2 Identification at genome-scale of DNA sequences targeted by RdhK regulators

##### Experimental strategy

In the following section, an adaptation of a method for the identification of sequences targeted by regulatory proteins described by Barlett *et al.* [115] to the RdhK system is described. In this approach, the genomic DNA of strain DCB-2 was fractionated in fragments of approximately 200 bp and exposed to RdhK6 or Hybrid A in presence of ClOHPA. Hybrid A was chosen over Hybrid B despite its weaker specificity to DNA motifs since it shares affinity for a DB7 with RdhK6. Both proteins were first bound to  $\alpha$ RdhK6-associated beads. Indeed, Hybrid A comprises a large portion of RdhK6 and is also recognised by the antibody raised against RdhK6. Using this approach, all possible genomic binding sites of RdhK6 (and any other RdhK DBD in case of hybrids) can be pulled-down using co-immunoprecipitation. The potential of this procedure was assessed using DB7 as positive target. Both RdhK6 and Hybrid A were previously shown to have *in vitro* and *in vivo* affinity for this dehalobox located in the promoter region of *rdhBA6* from strain DCB-2 [70]. Therefore, DB7 was expected to be recovered in the elution fraction of the co-immunoprecipitation experiments done with both RdhK proteins. These preliminary tests represent the initial step towards the identification of all possible genomic binding sites.

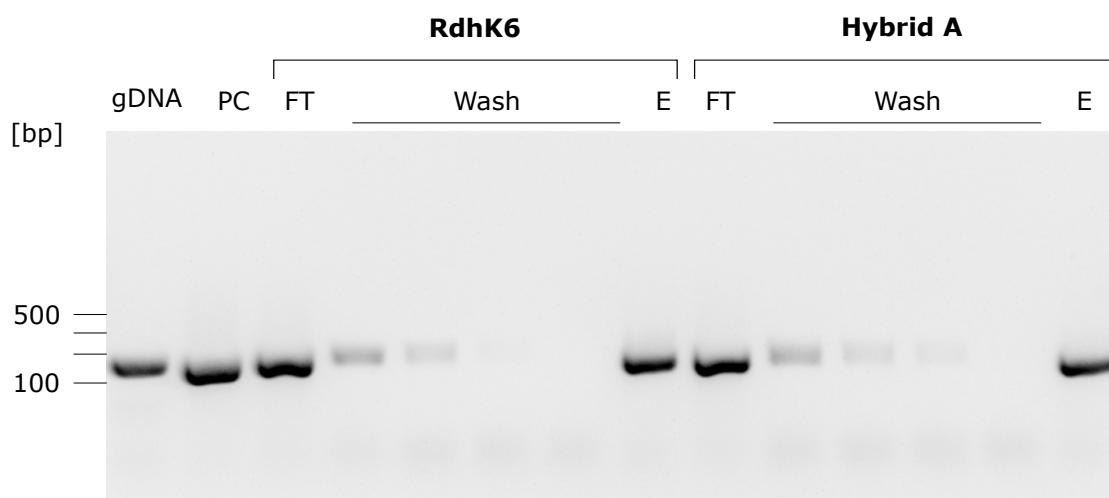


Figure 4.2 – Co-immunoprecipitation of DB7 with RdhK6 and Hybrid A. PCR amplification of DB7 was performed with all the fractions collected throughout the co-immunoprecipitation procedure applied to RdhK6 or Hybrid A, as indicated. Two positive controls were included: gDNA before and after fragmentation (lanes gDNA and PC, respectively). For each protein, the flow-through (FT), the four wash steps and the final elution (E) are shown.

#### Co-immunoprecipitation experiments of DB7 with RdhK6 and Hybrid A

Figure 4.2 displays the results obtained for the co-immunoprecipitation of DB7 with RdhK6 or Hybrid A in presence of CIOHPA. The gel shows the PCR amplification of DB7 from the fractions corresponding to different steps of the procedure. One can see that the results obtained for both proteins look similar which likely confirm the affinity of both proteins to DB7. These results show that both proteins are able to bind DB7 (lanes E) in this particular set-up, suggesting that they form a ternary complex even though the proteins are fixed on beads through their interaction with the antibody. While a substantial amount of DB7 could be amplified in the unbound fraction (FT), the four wash steps were successful in removing the excess DNA (4.2).

In summary, these results show that DB7, a DNA target of RdhK6 and Hybrid A can be efficiently pulled-down by co-immunoprecipitation in presence of the effector molecule. Moreover, the positive results obtained with Hybrid A, in which the part of RdhK6 sequence is shorter than in Hybrid B, indicate that the method is applicable to any hybrid protein based on RdhK6 EBD using the RdhK6 antibody. Here again, important controls should be conducted in order to draw robust conclusions concerning the elaboration of this approach. For example, random DNA sequences from strain DCB-2 genome could be used as target DNA for amplification in order to challenge the selectivity of the method. Also, the same experiments should be conducted in absence of CIOHPA in order to show the dependency of the system to the effector molecule.

The results obtained during this preliminary test are promising. However, it should be kept in

mind that, although the protein of interest is not required in a pure form, a significant amount of protein remains necessary in order to overcome the signal to noise ratio. Thus, recombinant protein (over)-expression in a at least partially soluble and active form still represents a fundamental challenge in the proposed methodology. This is not always achievable with any protein or may need some protocol adaptations. Nevertheless, the use of RdhK6 antibody to circumvent the purification of the RdhK hybrid of interest still constitutes a real advantage of this method as the purification RdhK proteins remains a true bottleneck for their *in vitro* characterisation.

## 4.4 Conclusion

In this chapter, preliminary results for two different approaches which aim to further characterise new RdhK regulators are presented. In the first approach, a RdhK hybrid protein displaying the DNA-binding domain of the yet uncharacterised RdhK16 from *D. restrictus* shed light on *in vitro* interactions with several promoters of *rdh* gene clusters from that same strain. Although important control experiments are necessary to validate these results,  $RdhK6_E^{Dha}16_D^{Dre}$  is the first example of a cross-species RdhK hybrid that could be obtained in a soluble form and showed *in vitro* binding activity. In the second approach, a DNA pull-down experiment was performed using *D. hafniense* RdhK6 and Hybrid A as baits. DB7, one of the DNA motifs targeted by both of these proteins was recovered in the elution fractions of both experiments. This result confirms that a DNA motif can be successfully captured and eluted through its interaction with the regulators that are immobilised on beads. Control experiments should be conducted to challenge the selectivity of the method.



## **Part II**

# **Energy metabolism and complex I-like enzymes in organohalide-respiring Firmicutes**



## **5 Metabolic versatility of Firmicutes organohalide-respiring bacteria and their complex I-like enzymes**

Over the years, a lot of efforts have been made to understand the global metabolism of OHRB. Many studies within the OHR field have focused on the characterisation of the terminal part of the respiratory chain (i.e. RDases). The electron donating moiety of the respiratory chain is usually deduced from the growing conditions (i.e. which electron donor is used by the cells), and sometimes combined with proteomic and/or transcriptomic approaches. This gave rise to metabolic models which are often proposed for one specific genus of OHRB and contain numerous question marks. Many thermodynamic and biochemical aspects are still not understood (see Chapter 1; Energy metabolism of OHRB). Also, the involvement of additional players in the respiratory chain was sometimes proposed [38, 41, 66], but strong evidences of their implication are still lacking. In the next sections, an non-exhaustive overview of the state of the art concerning the physiology of OHRB is proposed with a special emphasis on the Firmicutes.

### **5.1 Comparative physiology of organohalide-respiring bacteria: special emphasis on the Firmicutes**

In addition to their role as terminal reductases in OHR electron transport chains, RDases are also proposed to be associated with alternative metabolic activities (for example, electron sink under fermentative conditions or detoxification) [19, 48, 117]. Yet, the use of RDases as terminal reductases of the OHR electron transfer chain is the unifying property of OHRB. The identification of the physiological role of a new RdhA protein, the catalytic subunit of RDases, is thus a prerequisite to establish whether an organism is a true OHRB, especially when growing on fermentable compounds such as lactate and pyruvate [45].

As mentioned earlier, OHRB within the Firmicutes phylum are part of two bacterial genera: *Dehalobacter* spp. and *Desulfitobacterium* spp. Both genera show a close phylogenetic relationship, yet very different metabolic strategies: members of *Dehalobacter* spp. rely exclusively on OHR for growth, while *Desulfitobacterium* spp. are capable of performing alternative metabolisms such as the respiration of various organic and inorganic compounds and also fermentation [15]. Firmicutes is thus the only OHRB-containing phylum with both obligate and facultative representatives; a feature which invites to compare their OHR capabilities and their overall physiology.

### **5.1.1    *Dehalobacter* spp.**

All except one member of *Dehalobacter* spp. are limited to the use of molecular hydrogen as electron donor [14]. Most of the available literature on this genus reports its capacity to dechlorinate aliphatic compounds (example: PCE, TCE) but polychlorinated cyclic or aromatic molecules are also reported to be processed by different *Dehalobacter* sp. strains [14]. The capacity to use a broad range of organohalides as electron acceptors is reflected by the high number of *rdhA* genes present in the genomes; a property that *Dehalobacter* spp. share with other obligate OHRB. *Dehalobacter restrictus* is one of the most studied species of this genus with strain PER-K23 being one of the very first OHRB to be isolated [55]. As an example, the genome of strain PER-K23 encodes 25 RdhA homologous proteins [57]. The latter was also shown to display a quite large diversity of hydrogenases. In total, eight different multi-subunits hydrogenases were predicted including three membrane-bound NiFe uptake (Hup) and three cytosolic Fe-only hydrogenases [57]. In addition, a putative Wood-Ljungdahl pathway (WLP) for CO<sub>2</sub> fixation is encoded and expressed by strain PER-K23. On the other hand, *Dehalobacter* spp. cannot be cultivated without addition of acetate [14]. Actually, it has been established that *Dehalobacter* spp. use both autotrophic and heterotrophic carbon fixation with roughly one third of the assimilated carbon being from inorganic source. This metabolic feature is called mixotrophy and is shared with *Dehalococcoides* spp. [14].

### **5.1.2    *Desulfitobacterium* spp.**

*Desulfitobacterium* spp. are considered as the most abundant OHRB in nature, especially present in anaerobic terrestrial environments [15]. This probably reflects the wide metabolic diversity displayed by members of this genus and their capacity to adapt to different environments. As a consequence of their metabolic diversity, *Desulfitobacterium* spp. genomes are larger than those of *Dehalobacter* spp. as they encode several alternative metabolic pathways beside OHR. In return, being less dedicated to OHR than obligate OHRB, they generally display a smaller number of *rdh* gene clusters than the latter.

Despite the lower number of OHR dedicated enzymes, most *Desulfitobacterium* spp. have been isolated as OHRB and are capable of using a wide number of organohalides as electron acceptors ranging from chlorinated aliphatic and aromatic to iodine or bromine compounds

[15]. As already mentioned in Chapter 1, *D. hafniense* strain DCB-2 remains one of the most studied organism of this genus and is at the centre of the present work. Strain DCB-2 harbours the highest number of RDases among *Desulfitobacterium* spp., with five functional RdhA-encoding genes (note that two additional *rdhA* genes are found as pseudogenes in strain DCB-2). Among them, RdhA6 (previously annotated as CprA), involved in ClOHPA respiration, has been purified and characterised from strain DCB-2 [18] as well as other *Desulfitobacterium* spp. [15].

Representative strains from both *Desulfitobacterium hafniense* and *Desulfitobacterium dehalogenans* species have been shown to use "simple" electron donors like hydrogen or formate. However, in contrast to *Dehalobacter* spp., more complex organic molecules (for example pyruvate, lactate or vanillate) can also be used [15]. In the situation where a fermentable compound is available, it is relevant to consider the possibility of "facilitated fermentation" [45, 117]. In this case, the energy is conserved via substrate-level phosphorylation and the organohalide reduction via the RDase becomes a way to eliminate the excess of electrons (the organohalide as electron sink). A switch from true OHR to facilitated fermentation was proposed for *D. dehalogenans* when cultivated on pyruvate or lactate [117, 118]. In addition, *Desulfitobacterium* spp. can also grow purely fermentatively with pyruvate as sole carbon and energy source. Interestingly, the highest growth yield obtained with *D. dehalogenans* is for pyruvate-only cultures [118], suggesting a metabolic advantage for the cells in these conditions when compared to facilitated fermentation or pure OHR. To the best of my knowledge, the reasons for the differences in growth yields have so far never been addressed.

Apart from the "direct" energy metabolism, a complete set of genes encoding the WLP was also reported for *D. hafniense* strains DCB-2 and Y51 [15], suggesting an autotrophic carbon fixation which was experimentally confirmed for strain DCB-2 [34] as well as for *D. dehalogenans* [118]. In addition, the two *D. hafniense* strains seem to contain all but one of the enzymes involved in the tricarboxylic acid (TCA) cycle. Based on genomic analysis, this deficiency was proposed to be bypassed by the use of unconventional players [34]. However, no experimental data are available to confirm this hypothesis.

Interestingly, members from both *Dehalobacter* spp. and *Desulfitobacterium* spp. were shown to express an homologous enzyme of the respiratory complex I [38, 57, 66, 68, 119]. However, the role of the complex in the context of OHR and more generally in the metabolism of OHRB has never been addressed experimentally. Given the high importance of complex I in other respiratory pathways, the role of the homologous enzyme in OHRB and its potential involvement in the OHR respiratory chain appeared as a relevant research question. The next sections address functional, structural and evolutionary aspects of complex I and complex I-like enzymes to better understand the stakes of this research question.

## **5.2    Complex I and complex I-like enzymes**

### **5.2.1    Respiratory complex I**

In the context of oxidative phosphorylation, the respiratory complex I (or in short complex I) refers to the Nuo, a protein complex which serves as entry point in the electron transfer chain (ETC) for high energy electrons in the form of NADH generated in catabolic processes such as glycolysis, fatty acid oxidation or the TCA cycle [120, 121]. In bacteria, complex I is sometimes also named NADH dehydrogenase, or NDH-1, in opposition to the non-proton pumping version NDH-2. NDH-1 is composed of fourteen subunits which are common to the core-subunits of the mitochondrial respiratory complex I. The latter has acquired 32 additional subunits and form one of the biggest protein assemblies characterised so far [69]. Despite the difference in size, the main role of the respiratory complex I in prokaryotic or eukaryotic organisms remains similar: it couples the transfer of two electrons from NADH to the quinone pool along with the translocation of four protons across the membrane (i.e. either the inner mitochondrial membrane for eukaryotes or the cytoplasmic membrane in prokaryotes) to create a proton-motive force for ATP production. The fourteen subunits composing the bacterial enzyme are thus considered as sufficient to sustain its function and represent a good model to study the fundamental aspects of complex I.

Given its importance in the energy metabolism of organisms throughout the life kingdoms, it is not surprising that complex I has been extensively studied over the years. The next section gives an overview of the basic structural and functional information relevant for the present work, but more exhaustive reviews are available ([122–125]). Given the focus of this work and for the sake of simplicity, complex I will be used as a general term to designate the 14-subunit bacterial respiratory complex I (unless specified otherwise). In addition, *E. coli* Nuo nomenclature will be used to refer to the different subunits composing the complex.

#### **Structure and function of complex I**

Altogether, the fourteen subunits composing complex I form a membrane-bound L-shaped complex that consists of a peripheral arm exposed to the cytoplasm and a second arm spanning the membrane [120, 121, 126–128], as seen in the schematic representation shown in Figure 5.1. The enzyme is generally divided in three functional modules: N, Q and P.

The N-module is situated at the tip of the peripheral arm and consists of the NuoEFG subunits. It comprises the NADH-binding site where electrons are delivered to a non-covalently bound flavin mononucleotide (FMN) cofactor held in the Rossmann-type fold of NuoF [120, 129, 130]. From there, the electrons are tunnelled through a chain of seven FeS clusters to the quinone-reducing site located in the Q-module at the membrane side of the peripheral arm [131].

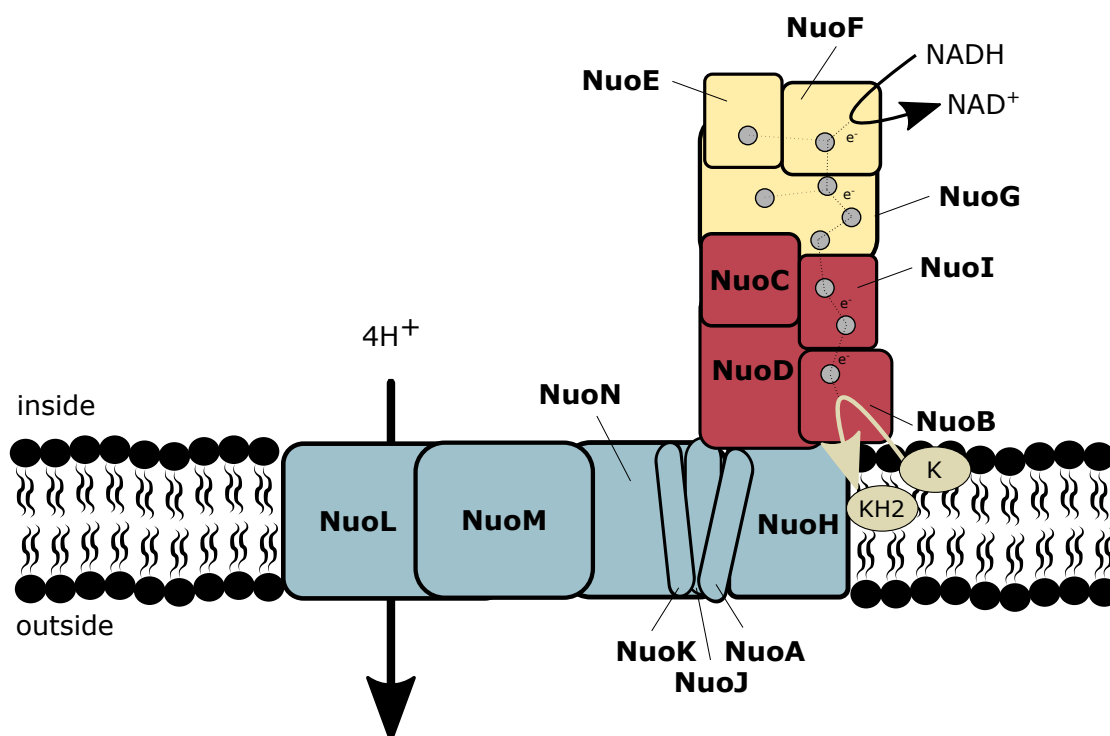


Figure 5.1 – Schematic representation of bacterial respiratory complex I. The fourteen subunits composing the classical respiratory complex I are indicated using *E. coli* nomenclature. The colour code corresponds to the three distinctive modules composing the L-shape complex: NADH-oxidising module (N-module) in yellow, the quinone-module (Q-module) in red and the proton-pumping module (P-module) in blue. FeS clusters are indicated by small grey circles in their respective subunits. The dotted line represents the path of the electrons from the NADH-oxidising to the quinone (K)-reducing sites (Note that two FeS clusters are not part of the main route and that the proton translocation is only schematically indicated and does not correspond to the exact location of the proton channels which are distributed along the P-module).

The Q-module is composed by four subunits (NuoBCDI) and displays three of the seven FeS clusters. Subunit NuoI is receiving the electrons from NuoG of the N-module. This passage between the two modules represents the longest distance of the overall transport chain and was therefore proposed to be the rate-limiting step [69, 129]. The last subunit involved in the electron tunnelling is NuoB which comprises the final redox centre that transfers the electrons to the quinone molecule [129, 130].

Thus, out of the seven subunits composing the peripheral arm, three are not directly involved in the redox transfer reaction. NuoE does not contribute to the main electron path although it contains an additional and conserved FeS cluster for which the physiological role remains unclear. It was recently proposed as a temporary storage unit for the second electron delivered to the FMN cofactor, in order to limit reactive oxygen species (ROS) production [69]. NuoCD (note that in *E. coli* these two proteins are fused) contributes to the formation of the quinone

cavity together with NuoB and two additional membrane subunits (NuoA and NuoH) [129, 132].

The quinone-binding site of complex I is unusually narrow [121, 133]. Nevertheless, in the consensus model, quinone molecules enter the channel all the way up to the close proximity of NuoB FeS cluster (note that some contradictory results were obtained with oversized synthetic quinone derivatives [133]). The unusual tightness of the cavity in comparison to other quinone-binding membrane proteins was proposed to help a tight regulation of the quinone reduction [69]. It was estimated that quinone could travel in and out the channel about 200 times per second [69, 124]. The quinone cavity was also proposed to welcome two quinone molecules at a time but this latter suggestion is still debated. Higher resolution structures of the quinone-bound enzyme are probably needed to conciliate all the data obtained. The quinone-binding site is also the target of long-established complex I inhibitors [134–136]. For example, rotenone is a complex I inhibitor found in the composition of some pesticides or insecticides.

In contrast to the peripheral subunits, the P-module is composed by seven membrane proteins (NuoAHJKLMN) that do not display any redox cofactor. The P-module, comprising of a total of sixty-four trans-membrane helices, is dedicated to proton translocation [132]. NuoLMN are ion antiporter subunits that form three potential routes for the proton to cross the membrane. A putative fourth channel is proposed to be formed by NuoHJK [129, 132, 137].

Overall, the structure of complex I seems very well adapted to fulfil its primary function : transferring electrons and pumping protons in order to create a proton-motive force used for ATP synthesis [120, 137]. However, the mechanism allowing complex I to couple the electron transfer reaction with the proton translocation is still not fully understood. A consensus tends to agree on the model of indirect coupling involving conformational changes rather than redox reaction intermediates [124, 138]. This is supported by the opposite location of the main antiporter subunits from the redox reaction centres which are separated by about 200 Å [124, 129]. The conserved hydrophilic residues found on the horizontal axis of the P-module are proposed to play a dominating role in this mechanism although the details are still under discussion [138].

Inherent to its primary function, another important aspect of complex I activity is the elimination of reducing equivalents and the recycling of  $\text{NAD}^+$  [126]. Under certain growth conditions, this function was even shown to prevail over the energy-conserving activity [120, 126, 139]. In addition, complex I is capable of reverse electron transfer (RET), for example in presence of high membrane potential, which leads to the reduction of  $\text{NAD}^+$ . The physiological implication of this mechanism is not fully understood yet but RET activity increases the release of detrimental ROS by complex I [69, 138]. Consequently, complex I dysfunction in mitochondria was associated with neurodegenerative and neuromuscular diseases [140]. This latter property certainly motivated many studies of the mitochondrial complex and its prokaryotic homologues.

The very-well defined structure of complex I which allows this giant structure to be so efficient

makes it a fascinating model to study evolution. Indeed, mitochondrial complex I comprises fourteen “core-subunits” that are homologous to those composing the bacterial enzyme and from which it evolved. An essence of what is known on the evolution of the complex is presented in next section.

### 5.2.2 Evolution of complex I

The consensus evolutionary model involves an ancestor, common to complex I and membrane-bound hydrogenases, that is composed by eleven subunits and that evolved to become the last common ancestor of the 14-subunit complex I family [123, 141, 142]. The 11-subunit enzyme, initially described in photosynthetic systems like chloroplasts and cyanobacteria [143], was shown to be present in the genome of representatives of various eubacterial phyla, including the Firmicutes [141]. This ancestor version of the enzyme does not have the N-module (i.e. lacks the subunits NuoEFG) but clearly differentiate from the membrane-bound hydrogenases. Sequence alignments suggested that the Q-module has the same function in the 11- or 14-subunit versions, which led to the proposition that the 11-subunit complex was the actual energy-coupling engine of the enzyme that may have acquired the N-module only later in the evolution [141].

The N-module is considered by Moparthy *et al.* as an electron donor apparatus, the acquisition of which allowed the respiratory chain to be connected with the TCA cycle [141]. Yet remains the question which electron donor is used by the 11-subunit enzyme. An attempt to find a common electron-donating protein that would be conserved in all bacteria harbouring the 11-subunit complex was not conclusive, although a few candidates have been proposed. Therefore, it was hypothesised that the 11-subunit version of the complex might work with several partners, acting as a docking platform for electron donors [141].

From the 11-subunit ancestor, the enzyme has acquired (in two steps: first NuoG, then NuoEF) the N-module resulting in the classical respiratory bacterial version, which represents the dominating version among eubacteria [141]. Indeed, roughly 50 % of the bacterial genomes display a full version of complex I [144]. Finally, the 14-subunit version sequentially acquired a total of thirty-two accessory proteins to become the largest enzyme of the mitochondrial respiratory chain [69, 120, 138, 145]. The accessory proteins of mitochondrial complex I mostly arrange around the central core of the enzyme and act as stabilisers or as sensors [69]. They do not improve the efficiency of the enzyme as both bacterial and mitochondrial complex I have the same reaction stoichiometry (four protons are pumped across the membrane for one NADH that is oxidised) [146]. Although they do not take part of the fundamental mechanism of complex I, they remain essential for the enzyme function, and mutations in accessory subunits were shown to be associated with severe mitochondria-related diseases [147]. Interestingly, the elaboration of a more sophisticated version of the enzyme seemed to have started prior to the endosymbiotic event as the presence of accessory subunits was shown in the  $\alpha$ -Proteobacteria, e.g. in *Paracoccus denitrificans* [138, 148].

Traces of evolution is contained in the bacterial genomes. Nevertheless, the 11-subunit version of the complex is surprisingly widely distributed among eubacteria [141]. As it lacks the NADH-oxidising module, this enzyme must rely on alternative electron donors. Based on this, it does no longer fall into the NADH oxidoreductase family. For nomenclature integrity, we propose here to refer to the 11-subunit version of complex I as compact 11-subunit complex I-like enzyme. Similarly, the term complex I-like enzymes will be used to define any protein assembly with homology to the conventional respiratory bacterial complex I but for which the electron donor is not necessarily NADH.

### **Examples of complex I-like enzymes**

As mentioned earlier, no universal electron donor could be identified for the compact 11-subunit complex I-like enzyme in bacteria [141]. Nevertheless, examples of peculiar complex I-like enzymes using alternative electron donors, mostly resulting from a need to adapt to the use of different substrates have been studied [142, 149]. Some selected cases are presented in the following paragraphs.

The  $\epsilon$ -Proteobacteria *Campylobacter jejuni* might illustrate the first step of the N-module acquisition as it does only harbour the NuoG subunit (note that it could also be the result of several gain and loss events, as suggested [141]). However, *nuoE* and *nuoF* are replaced in the gene cluster by two other genes. In the existing theoretical model, these two gene products were proposed to serve as adaptors for flavodoxins to feed electrons to NuoG and consequently in the complex I-like enzyme [150, 151]. Flavodoxins are generally small and acidic proteins that contains an FMN cofactor as redox centre [152]. As flavodoxins are widely distributed among several bacterial groups, they might serve as electron donors for complex I-like enzymes in other microorganisms as well.

The complex I-like enzymes identified in cyanobacterial genomes seem also devoid of the NuoEFG subunits [153]. The genes encoding for those enzymes are not necessarily arranged in one cluster (like mostly seen in bacteria) but can be found scattered in the genome and arranged in smaller operons [153]. Interestingly, cyanobacterial NuoM and NuoL homologues are generally present in several copies within the same genome. The sequence diversity of these multi-copy proteins gave raise to functionally distinct enzymes that may even co-occur within the same cell [153, 154]. Indeed, in cyanobacteria, complex I-like enzymes are not restricted to a role in respiration but are also involved in other cellular functions. In addition to the conventional subunits of compact 11-subunit complex I-like enzyme, the cyanobacterial complex I-like enzymes seem to contain four additional subunits which are specific to organisms performing oxygenic photosynthesis [154]. Only one of them was assigned with a function in carbon fixation but the functional roles of the others remain to be elucidated. Recently, a structural model was obtained for *Thermosynechococcus elongatus* [155]. This work has revealed the global structure of the enzyme which remains close to classical L-shape enzyme described earlier, but also has identified two more subunits with their position in the complex. In particular, a subunit named NdhS is important for electron

delivery from a ferredoxin that serves as intermediate between photosystem I and complex I. NdhS electrostatically attracts the ferredoxin close to the NuoI homologue and was proposed to help in the electron transfer toward the latter subunit by fine-tuning the ferredoxin binding angle [155, 156]. Ferredoxins are defined as acidic and redox active FeS cluster containing proteins [157] and, like flavodoxins, are commonly found in most organisms. They could also be considered more generally as possible electron donating proteins for complex I-like enzymes.

Another member of the complex I-like enzyme family was described from the methanogenic archaeon *Methanocarcina mazei* Gö1, which also lacks the N-module. Instead of NuoEFG, this version of the complex uses coenzyme F420 as electron donor. An additional subunit (FpoF), encoded in another part of the genome and homologous to a subunit of F420-reducing hydrogenases, was proposed as the electron input subunit [158].

### 5.3 Complex I-like enzymes in OHRB

As previously mentioned, *Dehalobacter* spp. and *Desulfitobacterium* spp. genomes encode complex I-like enzymes of the compact 11-subunit complex I-like type [34, 57, 67, 68]. Based on sequence similarity, three genes located distantly from the main operon were proposed to compose the N-module in *Desulfitobacterium* spp., which has sometimes led to consider the latter to display a classical 14-subunit complex I [15]. This, however, was never experimentally addressed. Some of the subunits composing the compact 11-subunit complex I-like enzymes were detected in proteomic analyses performed on representatives members of both genera. Three subunits (NuoBCD) were detected in the proteome of *D. restrictus* strain PER-K23 [57] and four subunits (NuoBCDI) in the proteome of *D. dehalogenans* growing on formate/fumarate or formate/CIOHPA [68].

The presence of complex I-like enzymes in OHRB is not limited to the Firmicutes. Indeed, a compact 11-subunit complex I-like enzyme was detected in the genome of several strains of *D. mccartyi* [33, 35]. Seven subunits (NuoBCDHIMN) were detected in the proteomic analysis of *D. mccartyi* strain CBDB1 cultivated on hexachlorobenzene [119]. As shown in Figure 5.2, the situation in *Sulfurospirillum* spp. seems different than in other OHRB as two different versions of the complex can be found [38]. On one hand, *Sulfurospirillum* spp. encodes a classical complex I with 14 subunits. On the other hand, the genome also contains an atypical version of the complex which seems to be formed by two operons (*nuoGHIJKLMN* and *nuoABCDEFGH*) found in opposite directions in the genome and separated by a gene of unknown function (see Figure 5.2). Based on proteomic data, both versions of the complex are expressed (with a small predominance of the classical 14-subunit complex encoded by *SMUL*<sub>0195–208</sub>), as observed in cells cultivated in various respiratory conditions (including OHR) [66].

So far, no biochemical characterisation nor evidence of the direct involvement of complex I or complex I-like enzymes in OHR metabolism have been reported. Nevertheless, the presence of the complex in the proteome of cells growing in OHR conditions strongly suggests so and

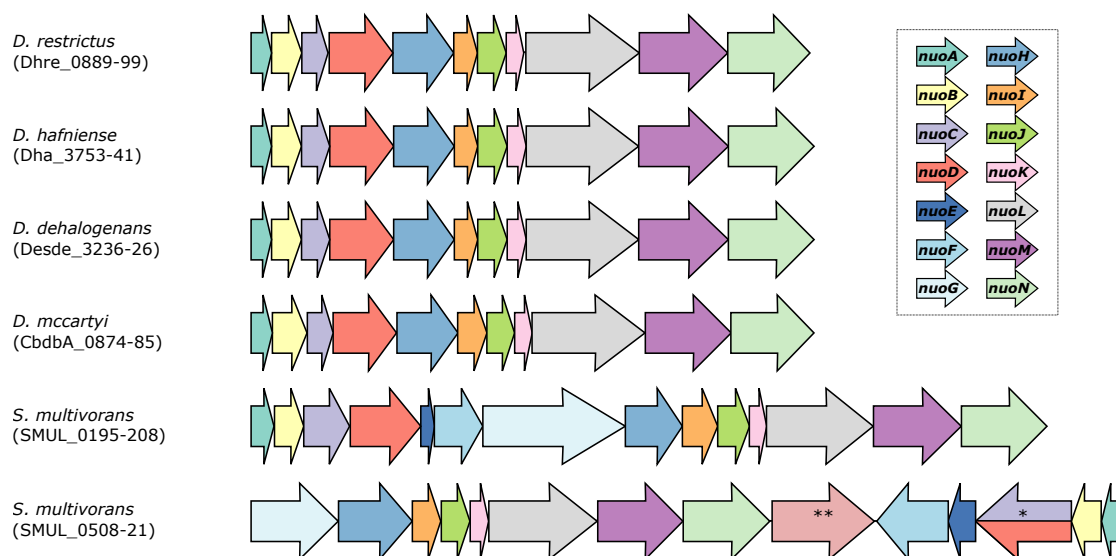


Figure 5.2 – Genomic organisation of the *nuo* homologous gene clusters in organohalide-respiring bacteria. The genome regions of four OHRB (*Dehalobacter restrictus* strain PER-K23, *Desulfitobacterium dehalogenans*, *Dehalococcoides mccartyi* strain CBDB1 and *Sulfurospirillum multivorans*) belonging to four different genera are schematically presented. The colour code indicates the Nuo homologous subunits as found in the classical respiratory complex I from *E. coli*. For *S. multivorans*, two different clusters are displayed as two versions of the complex are found. For the second version, two features are indicated: the genes encoding subunits C and D are fused (\*) and *SMUL*<sub>0516</sub> encodes an hypothetical protein (\*\*).

the participation of a complex I-like enzyme as a possible partner in the electron transfer chain was proposed in several hypothetical OHR models [47, 66]. Indeed, the expression of such a large protein complex with potential proton-pumping activity invites to consider it as part of the OHR energy metabolism. Otherwise, why would OHRB consume energy to produce this machinery without using it? Then, if complex I-like enzymes are part of the OHR respiratory chain, how do they fit in the existing models? In addition, the NADH-oxidising module is missing in most OHRB, which suggests the use of alternative electron donors for the complex I-like enzymes. So far, complex I-like enzymes were recognised as potential players in the OHR metabolism mostly due to its analogy to the respiratory complex I involved in the metabolism of aerobic respiration. At least two published hypothetical models include complex I-like enzymes in the OHR respiratory chain of *S. multivorans* [66] and *Dehalococcoides* spp. [47]. However, no evidence nor solid proposition of the potential role of the complex in OHR metabolism have been reported. Its repetitive occurrence in proteomic datasets from representative bacteria across the phylogeny of OHRB invites to investigate this research question and address its possible function in the metabolism of OHRB.

## 6 Comparative proteomics of *Desulfitobacterium hafniense* strain DCB-2

### 6.1 Introduction

Numerous studies focusing on differential gene expression in OHRB have been released over the years. Both transcriptomic and proteomic approaches contributed to today's comprehension of the OHR metabolism. Most of the global transcriptomic data available so far are based on the micro-array methodology which is limited in terms of sensitivity. More sensitive (but also more expensive) methods like RNA-seq was only performed on *D. mccartyi* [159]. In contrast, global proteome analyses were performed with representative bacteria from all OHRB phyla but the level of proteome coverage achieved vary significantly from one study to the other (for a review, see [160]). The best results in terms of proteome coverage were obtained with *D. mccartyi* spp. where up to 70 % of the proteins were detected (which represent a total of 1023 proteins) [90, 119]. Globally, the available datasets are largely dominated by studies targeting *D. mccartyi* sp. which somehow imbalance the global comprehension of OHR. Nevertheless, significant proteome coverage was achieved with *D. restrictus* strain PER-K23 [57]. As obligate OHRB, both *D. mccartyi* and *D. restrictus* do not enable to play with the growth conditions and thus to highlight proteome adaptations to different metabolisms in the same bacterium. In contrast, examples of studies comparing the proteomes of facultative OHRB such as *S. multivorans* [66] or *D. dehalogenans* [68] growing on different substrates were published. However, the proteome coverage was either globally limited or only high in one or the other condition impairing an extended proteome comparison.

In the present work, a Tandem-Mass-Tag (TMT) based quantitative proteomic approach was used to robustly compare the relative abundance of the proteins in *Desulfitobacterium hafniense* strain DCB-2 which was cultivated in six different growth conditions. The conditions were chosen in order to trigger different types of metabolism in the bacterium and highlight substrate-specific adaptations through the comparison of electron donor/acceptor combinations. On the one hand, respiration metabolisms were triggered by supplementing lactate or H<sub>2</sub> as electron donors in combination with either fumarate or ClOHPA as electron acceptors, and on the other hand fermentation was stimulated with addition of pyruvate in absence of

any electron acceptor. The data were analysed in order to target several research questions. At first, the protein specifically present in fermentative conditions when compared to respiration were analysed. Second, substrate-specific proteome adaptations were identified and analysed in order to understand their potential involvement in the metabolism of strain DCB-2.

## 6.2 Material and Methods

All chemicals used in the following protocols were purchased at Merck (Coinsins, Switzerland) unless specified otherwise.

### 6.2.1 Cultivation of *Desulfitobacterium hafniense*

*Desulfitobacterium hafniense* strain DCB-2 (DSM 10664) was cultivated anaerobically in 40-mL liquid batch cultures. The growth medium was based on a solution containing the following major elements:  $0.958 \text{ g L}^{-1}$  of  $\text{K}_2\text{HPO}_4 \cdot 3 \text{ H}_2\text{O}$ ,  $0.218 \text{ g L}^{-1}$  of  $\text{NaH}_2\text{PO}_4 \cdot 2 \text{ H}_2\text{O}$ ,  $0.1 \text{ g L}^{-1}$  of yeast extract and  $0.5 \text{ mg L}^{-1}$  of resazurin. This "basic" solution was complemented with two other solutions prepared separately which were both added at a 1:40 (v/v) ratio. The first one contained different important co-elements (such as vitamins, trace elements and co-factors). It was prepared by adding different preparation in ddH<sub>2</sub>O to reach a final concentration of  $20 \text{ mg L}^{-1}$  of EDTA (pH = 7.0),  $80 \text{ mg L}^{-1}$  of  $\text{FeCl}_2 \cdot 4 \text{ H}_2\text{O}$ ,  $4 \text{ mg L}^{-1}$  of  $\text{MnCl}_2 \cdot 4 \text{ H}_2\text{O}$ ,  $7.6 \text{ mg L}^{-1}$  of  $\text{CoCl}_2 \cdot 6 \text{ H}_2\text{O}$ ,  $2.8 \text{ mg L}^{-1}$  of  $\text{ZnCl}_2$ ,  $0.102 \text{ mg L}^{-1}$  of  $\text{CuCl}_2 \cdot 2 \text{ H}_2\text{O}$ ,  $0.2208 \text{ mg L}^{-1}$  of  $\text{AlCl}_3$ ,  $0.24 \text{ mg L}^{-1}$  of  $\text{H}_3\text{BO}_3$ ,  $1.656 \text{ mg L}^{-1}$  of  $\text{Na}_2\text{MoO}_4 \cdot 2 \text{ H}_2\text{O}$ ,  $0.96 \text{ mg L}^{-1}$  of  $\text{NiCl}_2 \cdot 6 \text{ H}_2\text{O}$ ,  $2 \text{ mg L}^{-1}$  of biotin,  $10 \text{ mg L}^{-1}$  of P-amidobenzoate,  $2 \text{ mg L}^{-1}$  of pantothenate,  $0.8 \text{ mg L}^{-1}$  of folic acid ( $\cdot 2 \text{ H}_2\text{O}$ ),  $2 \text{ mg L}^{-1}$  of lipoic acid,  $0.4 \text{ mg L}^{-1}$  of pyridoxine,  $22 \text{ mg L}^{-1}$  of nicotinic acid,  $4 \text{ mg L}^{-1}$  of thiamine-HCl,  $2 \text{ mg L}^{-1}$  of riboflavine and  $2 \text{ mg L}^{-1}$  of cyanocobalamin. The second one contained the minimal source of chlorides and was prepared by dissolving  $4.4 \text{ g L}^{-1}$  of  $\text{CaCl}_2 \cdot 2 \text{ H}_2\text{O}$  and  $4.06 \text{ g L}^{-1}$  of  $\text{MgCl}_2 \cdot 6 \text{ H}_2\text{O}$  in ddH<sub>2</sub>O. Finally, a third solution of carbonate/bicarbonate buffer (prepared by mixing  $9.01 \text{ g L}^{-1}$  of  $\text{NH}_4\text{HCO}_3$  and  $76.11 \text{ g L}^{-1}$   $\text{NaHCO}_3$ ) containing  $\text{Na}_2\text{S}$  ( $240.2 \text{ g L}^{-1}$  of  $\text{Na}_2\text{S} \cdot 9 \text{ H}_2\text{O}$  as reducing agent was added at 1:20 (v/v) ratio. All the solutions described were made anaerobic through 15 cycles of gas-exchange with N<sub>2</sub>, sterilized either by autoclave or filtration and stored under N<sub>2</sub> conditions. Note that for the basal medium and the carbonate/bicarbonate buffer, an additional step of boiling and cooling down while flushing with a mix of 80% N<sub>2</sub> / 20%CO<sub>2</sub> was done prior to sterilisation. Prior to inoculation, the medium was further completed with different combinations of substrates in order to trigger different types of metabolism. For fermentation, the medium was supplemented with a final concentration of 40 mM of pyruvate (Pyr-only). Alternatively, pyruvate was also used as electron donor at a final concentration of 20 mM in combination with 20 mM of fumarate as electron acceptor (Py/Fu). When applicable, pyruvate was replaced with 20 mM of lactate (La/Fu). To trigger organohalide respiration, 3-chloro-4-hydroxyphenylacetic acid (Cl-OHPA) was used as alternative electron acceptor at a final concentration 10 mM (from a 100 mM stock solution prepared in ddH<sub>2</sub>O) (La/CIOHPA). In addition, molecular hydrogen was used

as electron donor by replacing the gas phase with a mixture of 80% H<sub>2</sub> / 20% CO<sub>2</sub>, either in combination with fumarate (H<sub>2</sub>/Fu) or ClOHPA (H<sub>2</sub>/ClOHPA). In the case of H<sub>2</sub>, 2 mM acetate was added as carbon source. Once completed, the media were inoculated with 5% (v/v) of cultures prepared in the same conditions no longer than 3 days in advance. All cultures were incubated at 30°C under agitation (100 rpm).

### 6.2.2 Cell harvesting and sample preparation

The biomass was collected in the rather late growth phase to maximise the amount of biomass produced but clearly before reaching the stationary phase in order to analyse proteomes of physiologically active cells. The cultures were first transferred to 50-mL Falcon tubes and cells were pelleted for 15 min @ 4500x g and 4 °C. Biomass pellets were washed twice in ice-cold Tris-HCl buffer (50 mM, pH = 7.5) and stored in 1.5-mL Eppendorf tubes at -80°C until further use. All biomass samples used for proteomic analysis were treated at the same time. First, the pellets were resuspended in 100 µL of 50 mM HEPES buffer (pH = 8.0) supplemented with SigmaFast™ protease inhibitor cocktail (Sigma Aldrich, Buchs, Switzerland) and a few crystals of DNase I (Roche, Basel, Switzerland). Five rounds of 5 s (5 x 1 s) sonication at 60% were applied to the cell suspensions directly in the 1.5-mL tubes (Fisherbrand™ Q500 sonicator, Thermo Fisher Scientific, Basel, Switzerland). Samples were kept on ice in-between rounds. After a mild centrifugation (500x g, 5 min and 4°C) to remove unbroken cells, proteins were solubilized for 15 min at RT in 2% SDS (from a stock solution at 4% prepared in HEPES buffer). Protein concentration was determined using Pierce™ BCA protein assay kit (Fisher Scientific) and protein samples were diluted with 2% SDS solution and aliquoted to 30 µg in 30 µL.

### 6.2.3 Proteomic analysis

This part of the work was performed by Romain Hamelin and Florence Armand of the Proteomic Core Facility of EPFL.

#### Reduction/alkylation of the samples

Samples were trypsin digested using the Filter-Aided Sample preparation (FASP) protocol with minor modifications [161]. Quantified lysates containing a protein amount of 30 µg were resuspended in 200 µL of 8M Urea; 100 mM Tris-HCl and deposited on top of Microcon®-30K devices (Merck). Samples were centrifuged at 9391x g, at 20°C for 30 min. All subsequent centrifugation steps were performed using the same conditions. An additional 200 µL of 8 M Urea; 100 mM Tris-HCl was added and devices were centrifuged again. Reduction was performed by adding 100 µL of 10mM TCEP in 8 M Urea; 100 mM Tris-HCl on top of filters followed by 60 min incubation time at 37°C with gentle shaking and light protection. Alkylation solution was removed by centrifugation and filters were washed with 200 µL of 8 M Urea; 100 mM Tris-HCl. After removal of washing solution by centrifugation, alkylation was performed

by adding 100  $\mu$ L of 40 mM chloroacetamide in 8 M Urea; 100 mM Tris-HCl and incubating the filters at 37°C for 45 min with gentle shaking and protection from light. The alkylation solution was removed by centrifugation and another washing/centrifugation step with 200  $\mu$ L of 8 M Urea; 100 mM Tris-HCl was performed. This last Urea buffer washing step was repeated twice followed by three additional washing steps with 100  $\mu$ L of 5 mM Tris-HCl. Proteolytic digestion was performed overnight at 37°C by adding on top of filters 100  $\mu$ L of a combined solution of Endoproteinase Lys-C and Trypsin Gold in an enzyme/protein ratio of 1:50 (w/w) supplemented with 10 mM  $\text{CaCl}_2$ . Resulting peptides were recovered by centrifugation. The devices were then rinsed with 50  $\mu$ L of 4% trifluoroacetic acid and centrifuged. This step was repeated three times and peptides were finally desalted on SDB-RPS StageTips [162] and dried by vacuum centrifugation. A mixture of each biological samples was spiked as two additional channels and used as bridge channels.

### TMT-labelling

For TMT labelling, dried peptides were first reconstituted in 8  $\mu$ L 100 mM HEPES pH 8 and 3  $\mu$ L of TMT solution (40  $\mu\text{g}\mu\text{L}^{-1}$  in pure acetonitrile) was then added. TMT Labelling was performed at room temperature for 90 min and reactions were quenched with hydroxylamine to a final concentration of 0.4% (v/v) for 15 min. TMT-labeled samples were then pooled at a 1:1 ratio across all samples. A single shot control LC-MS run was performed to ensure similar peptide mixing across each TMT channel to avoid the need of further excessive normalisation. Quantities of each TMT-labeled sample were adjusted according to the control run. The combined samples were then desalted using a 100 mg SEP-PAK C18 cartridge (Waters) and vacuum centrifuged. Pooled samples were fractionated into 12 fractions using an Agilent OFF-Gel 3100 system following the manufacturer's instructions. Resulting fractions were dried by vacuum centrifugation and again desalted on C18 StageTips.

### LC-MS/MS analysis

Each individual fraction was resuspended in 2% Acetonitrile; 0.1% Formic acid and nano-flow separations were performed on a Dionex Ultimate 3000 RSLC nano UPLC system on-line connected with a Lumos Fusion Orbitrap Mass Spectrometer. A capillary pre-column (Acclaim Pepmap C18 ; 3  $\mu\text{m}$ -100 Å ; 2cm x 75  $\mu\text{m}$  ID) was used for sample trapping and cleaning. Analytical separations were performed at 250 nl/min over a 150 min. biphasic gradients on a 50 cm long in-house packed capillary column (75  $\mu\text{m}$  ID; ReproSil-Pur C18-AQ 1.9  $\mu\text{m}$  silica beads; Dr. Maisch). Acquisitions were performed through Top Speed Data-Dependent acquisition mode using 3 s cycle time. First MS scans were acquired at a resolution of 120'000 (at 200 m/z) and the most intense parent ions were selected and fragmented by High energy Collision Dissociation (HCD) with a Normalised Collision Energy (NCE) of 37.5% using an isolation window of 0.7 m/z. Fragmented ions scans were acquired with a resolution of 50'000 (at 200 m/z) and selected ions were then excluded for the following 120 s.

### Data analysis

Raw data were processed using SEQUEST, Mascot, MS Amanda [162] and MS Fragger [163] in Proteome Discoverer v.2.4 against the *Desulfitobacterium hafniense* strain DCB-2 database (4851 sequences). Enzyme specificity was set to trypsin and a minimum of six amino acids was required for peptide identification. Up to two missed cleavages were allowed and a 1% FDR cut-off was applied both at peptide and protein identification levels. For the database search, carbamidomethylation (C), TMT tags (K and Peptide N termini) were set as fixed modifications whereas oxidation (M) was considered as a variable one.

Resulting text files were processed through in-house written R scripts (version 3.6.3) [164]. A first normalisation step was applied according to the Sample Loading normalisation [165]. Assuming that the total protein abundances were equal across the TMT channels, the reporter ion intensities of all spectra were summed and each channel was scaled according to this sum, so that the sum of reporter ion signals per channel equals the average of the signals across samples. The multiplexing design of the experiment required a second normalisation step to correct variations between the 2 TMT experiments. Internal Reference Scaling (IRS) process was here applied [166]. For each TMT run, the mean of reporter ion signals in the 2 corresponding bridge channels was calculated for each protein. The normalisation factor, scaling the 2 bridge channels means between the 2 TMT experiments, was used to scale protein intensities. A Trimmed M-Mean normalisation step was also applied using the package EdgeR (version 3.26.8) [167]. Assuming that samples contain a majority of non-differentially expressed proteins, this third step calculates normalisation factors according to these presumed unchanged protein abundances. Proteins with high or low abundances and proteins with larger or smaller fold-changes are not considered. Differential protein expression analysis was performed using R bioconductor package limma (version 3.34.9, 2018-02-22) [168], followed by Benjamini-Hochberg multiple-testing method [169]. Normalised and standardised protein abundances were used for the following data treatments. A heat map was processed to visualise unsupervised hierarchical clustering of protein abundances using the Euclidean distance and complete linkage method for both, sample clustering and protein clustering. Additionally, as an attempt to highlight new putative Nuo candidates by comparing their global behaviour, the Euclidean distance was calculated between the cluster 4 protein abundances and known Nuo complex profiles defined with the median of the Nuo complex protein abundances.

### 6.3 Results and Discussion

The relative abundance of proteins from *Desulfitobacterium hafniense* strain DCB-2 was compared from cells cultivated in six different growth conditions. The conditions chosen were triggering the metabolic processes fermentation and respiration including OHR. The different growth conditions and corresponding metabolisms are summarised in Figure 6.1A. As the proteomes obtained from different growth conditions are compared to each others, it is important to keep in mind that growth performances of strain DCB-2 varies substantially depending on the substrates in use, as shown in Figure C.1. In order to have statistical information, each of the six conditions was performed in triplicates which resulted in a total of eighteen batch cultures. After extraction, the proteins of each biological samples were quantified and analysed by a tandem-mass-tag (TMT) proteomic workflow. As the number of samples was higher than the number of isobaric labels available, the analysis was performed in two batches. A mixture of all samples was also included in each batch to correct for potential intra- and inter-batch effects (for more details about normalisation, see Material and Methods section).

#### 6.3.1 General presentation of the data

A total of 2796 proteins were quantified across all of the 18 samples (divided in six growth conditions) which represent 58 % of the theoretical proteome of *D. hafniense* strain DCB-2 [34]. Thus, the present study resulted in the most extended proteomic dataset so far obtained in OHRB across six growth conditions run in parallel; going beyond the 1023 proteins quantified in Schiffmann *et al.* [90]. After several steps of normalisation (see details in materials and methods), the data displayed in Figure 6.1B revealed a good reproducibility between the different replicates series. Indeed, the three replicates of each condition co-localised in the PCA (see Figure 6.1B). The Py/Fu replicates appeared slightly more dispersed than the replicates of the other conditions but remained grouped. In addition, this analysis also revealed a co-localisation of the mixed samples which reflects the efficiency of the different normalisation steps (see 6.1B, black series).

The PCA revealed two subgroups of growth conditions with Py/Fu and La/Fu on one side and the four other conditions in the other side of the major axis (PC1). This was confirmed by the hierarchical clustering of the conditions from the normalised abundances of the proteins in the different samples (see supplemental data C.2. The branch formed by the conditions Pyr-only, La/CIOHPA, H<sub>2</sub>/Fu and H<sub>2</sub>/CIOHPA appeared more heterogeneous with two subgroups formed by Pyr-only and La/CIOHPA on one hand and the two conditions involving H<sub>2</sub> on the other. A hierarchical clustering was also applied to the proteins abundances across the conditions in order to group the proteins following a similar trend. A number of fifteen clusters was chosen to divide the dataset, as it resulted in a coherent separation of the protein patterns.

The comparative proteomic analysis gave rise to a significant amount of data which are explored in the next sections. At first, a comparison between the fermentative and the respiratory

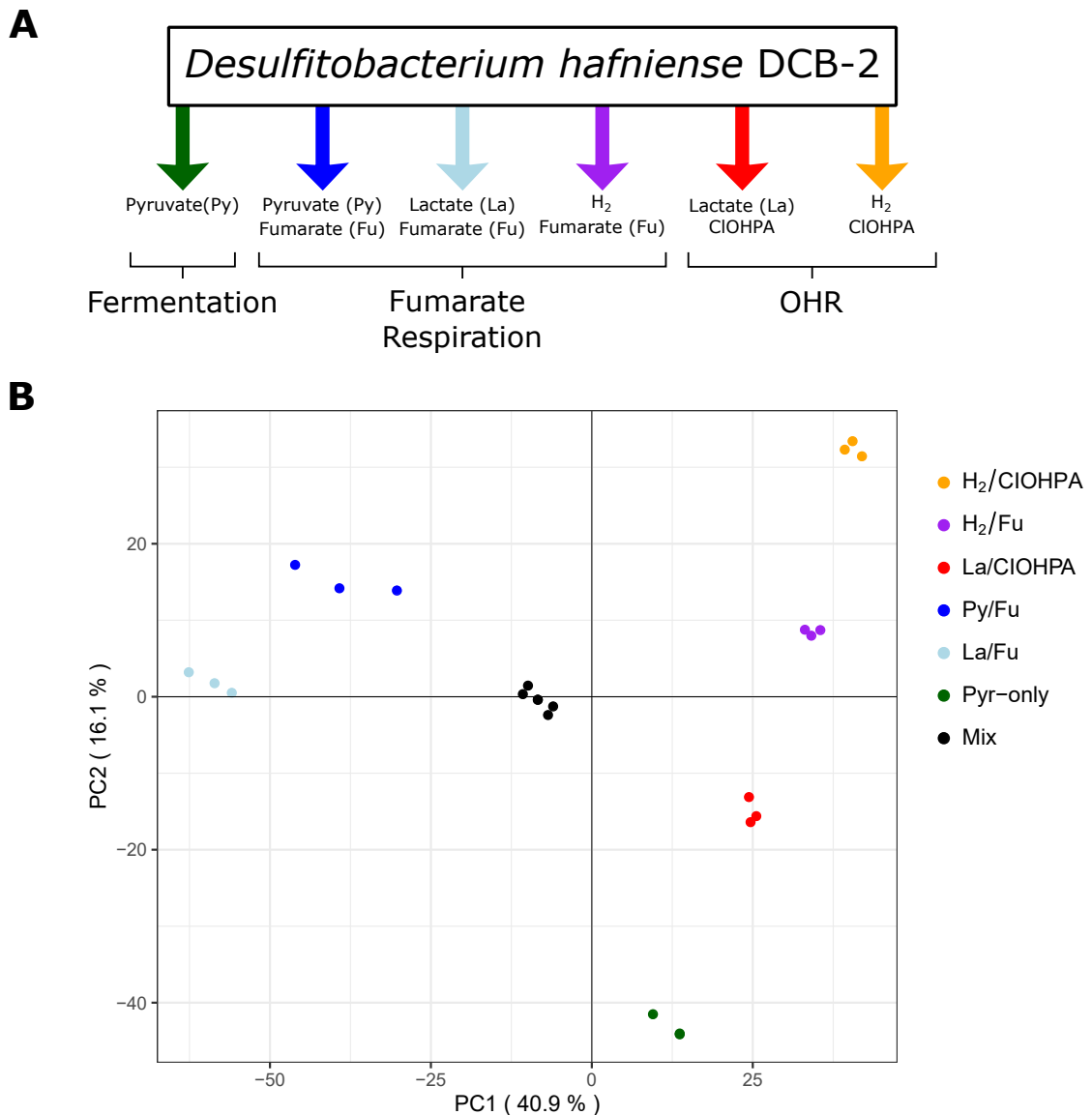


Figure 6.1 – Principle components analysis for DCB-2 proteome expression in different growth conditions.

metabolisms was made. Then, the dataset was analysed according to the different electron donors and acceptors in order to reveal substrate-specific proteomic adaptations of strain DCB-2.

### 6.3.2 Proteins up-regulated in fermentation with pyruvate

The proteomic dataset was first analysed with emphasis on the proteins showing significant up-regulation levels during fermentation of pyruvate (Pyr-only) in comparison to all respiratory conditions (combinations of electron donors and acceptors) (Table S1, excel document).

Figure 6.2 shows in a Venn diagram the distribution of 463 proteins displaying at least a 2-fold up-regulation during fermentation against any of the other five conditions. A set of 19 proteins were found at the very centre of the diagram, thus representing the proteins that are likely important for the adaptation of strain DCB-2 to fermentative growth (Table 6.1).

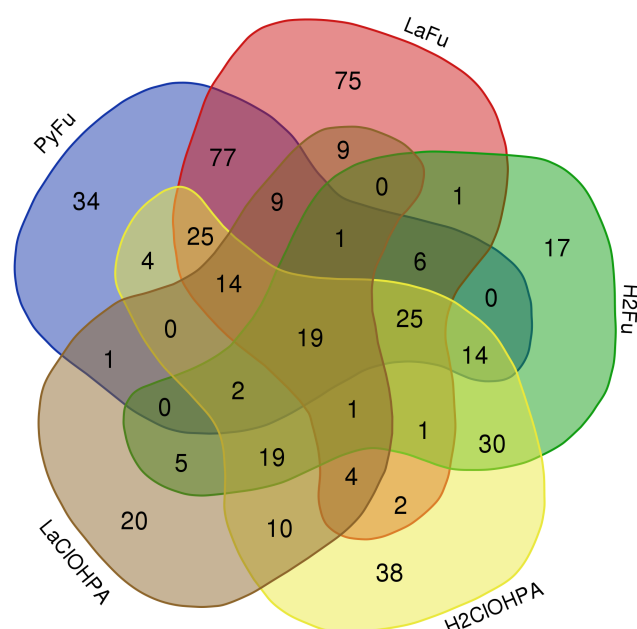


Figure 6.2 – Distribution of proteins up-regulated in Pyr-only conditions. The Venn diagram represents the number and distribution of proteins that were up-regulated in fermentation in comparison to all five respiratory conditions taken individually.

Proteins from a five-gene cluster (ACL21108-12) were generally up-regulated with high level (between 5 and 30-fold, depending on the protein and conditions). On the one hand, ACL21108, which is likely encoded by a monocistronic gene, is annotated as a NAD-dependent decarboxylating malate dehydrogenase (dMDH, E.C. 1.1.1.38). In contrast to the classical malate dehydrogenase (E.C. 1.1.1.37) that produces oxaloacetate in the TCA cycle, dMDH is usually involved in anaplerotic pathways and produces pyruvate from malate. While this activity is not likely to be required in presence of pyruvate in the medium, ACL21108 may function in the reverse direction and produces malate from pyruvate, CO<sub>2</sub> and NADH, thus replenishing TCA cycle intermediates. On the other hand, ACL21109-12 correspond to a predicted operon (in reverse orientation) coding for proteins with sequence homology to enzymes involved in citrate and succinate metabolism. ACL21110 and ACL21112 show 40% and 35% sequence identity with the small ( $\alpha$ ) and large ( $\beta$ ) subunits, respectively, of the succinyl-CoA synthetase (SucC and SucD, SCS) from *E. coli*, which catalyses the conversion of succinyl-CoA to succinate in the TCA cycle. The lack of the  $\alpha$ -ketoglutarate dehydrogenase in strain DCB-2, thus interrupting the TCA cycle, requires an alternative source of succinyl-CoA for anabolism. While no function could be assigned to ACL21111, ACL21109 displays some sequence homology to the  $\beta$  subunit of the citryl-CoA lyase (CitE) from *E. coli* or the malonyl-CoA thioesterase from

Protein ref.	Predicted function	Fold change in Pyr-only vs					
		Py/Fu	La/Fu	H2/Fu	H2/Cl	La/Cl	
ACL18433	Glycyl radical enzyme of unknown function (GUF)	8	5	4	9	3	
ACL18746	Succinate:quinone oxidoreductase, flavoprotein	7	4	9	9	4	
ACL19672	Transcriptional regulator, CdaR family	2	2	2	2	2	
ACL20074	NADP-oxidoreductase/NADPH-reducing hydrogenase	8	4	6	6	3	
ACL20535	Succinate:quinone oxidoreductase, flavoprotein	17	5	9	11	4	
ACL21007	Succinate:quinone oxidoreductase, flavoprotein	6	5	5	3	2	
ACL21108	NAD-dependent decarboxylating malate DH	30	8	10	9	9	
ACL21109	Citryl-CoA lyase, $\beta$ subunit	22	7	14	11	6	
ACL21110	Succinyl-CoA synthetase, $\alpha$ subunit	13	6	11	10	7	
ACL21111	Hypothetical protein	11	5	9	9	5	
ACL21112	Succinyl-CoA synthetase, $\beta$ subunit	15	5	11	9	7	
ACL21807	4Fe-4S ferredoxin	3	4	16	16	12	
ACL21808	CODH, clade F, catalytic subunit	3	5	22	20	15	
ACL21809	Cobyrinic acid a,c-diamide synthase	3	3	3	3	3	
ACL22391	Transcriptional regulator, IclR family	3	3	3	2	3	
ACL22517	Methyl-accepting chemotaxis sensory transducer	5	4	5	3	7	
ACL22518	Protein of unknown function (DUF224)	8	4	8	5	17	
ACL22519	FAD-linked oxidase protein	9	5	8	5	25	
ACL22584	ABC transporter	5	4	2	3	4	

Table 6.1 – Proteins up-regulated in Pyr-only conditions.

*Rhodobacter sphaeroides* [170]. While the former catalyses together with CitD and CitF the citrate lyase activity in *E. coli* that is involved in citrate fermentation [171, 172], the latter has been proposed to participate in the assimilation of acetyl-CoA via the ethylmalonyl-CoA pathway and glyoxylate. The exact substrates of ACL21109 and the ACL21110/2 pair of enzymes cannot be inferred from sequence homology. However, taken together, this five-gene cluster is likely to play a role in anaplerotic reactions in the context of the TCA cycle.

A glycyl radical enzyme (GRE) (ACL18433) was found up-regulated between 3- and 9-fold in Pyr-only conditions. GREs are diverse in function but can be classified into two groups based on the generation of aldehydes. GREs generating aldehydes require the establishment of bacterial micro-compartments (BMC), protein-based organelles that protect the cytoplasm from toxic reaction intermediates [173]. Zarzycki *et al.* have identified ACL18433 as a GRE of unknown function (GUF) based on the peculiar gene composition and organisation [174], suggesting that no functional BMC is produced around ACL18433. Nevertheless, it is likely that the GRE encoding gene is part of a five-gene operon (ACL18432-6), from which three proteins were detected as up-regulated during fermentation. While ACL18432 is a transcriptional regulator, ACL18434 is the GRE-associated activating enzyme [175]. ACL18435 belongs to the family of tripartite ATP-independent periplasmic (TRAP) transporters [176] that is composed of two subunits which are homologous to DctP and to a fusion of DctQ and DctM [177]. No clear function, however, could be associated with ACL18433 in the fermentation metabolism of strain DCB-2. Interestingly, the GRE and associated activating enzyme likely involved as pyruvate formate-lyase in strain DCB-2 (PFL, ACL19303-4) did not show any up-regulation pattern throughout the proteomic dataset, thus reflecting the central role of PFL in carbon metabolism [178].

In Pyr-only conditions, the catalytic flavoprotein subunit of three members of the family of succinate:quinone oxidoreductases were also up-regulated (ACL18746, ACL20535 and ACL21007, Table 6.1). This large family of membrane-bound oxidoreductases encompasses both subfamilies of succinate:quinone reductases (SQR) and quinol:fumarate reductases (QFR) [179]. While the paradigmatic SQR member is the succinate dehydrogenase (SdhABCD) involved as complex II in aerobic respiration that delivers electrons to the quinone pool and contributes to the TCA cycle, QFR is represented by the terminal fumarate reductase (FrdABC) in anaerobic fumarate respiration [180]. According to Kim *et al.*, 19 paralogues of this type of flavoproteins were identified in strain DCB-2 [34]. However, 31 of them were obtained by BlastP analysis (data not shown), suggesting that strain DCB-2 has the capacity to transfer electrons via an extended network of redox enzymes. The three flavoproteins identified in Pyr-only conditions display diverse genetic environments (none of them are part of the classical 4- or 3-gene operons as for SdhABCD and FrdABC, respectively), but all of them harbour a signal peptide for the Twin-arginine translocation (Tat) system, indicating that these flavoproteins are exported across the cytoplasmic membrane. Among them, only ACL20535 displays a possible membrane-bound partner (ACL20534). So far, no substrate can be associated with any of the three flavoproteins.

When comparing the large protein ACL20074 with the UniProt database, the major part of the N-terminal half of ACL20074 looks like an NADPH-dependent oxidoreductase (with SfrB of *Geobacter sulfurreducens* as best match [181]), and the C-terminal half resembles to the HndD subunit of the NADP-reducing hydrogenase from *Desulfovibrio fructosovorans* [182]. While SfrAB is likely involved in acetate metabolism with a possible function in the NADPH/NADP homeostasis, the HndABCD complex has been proposed recently to be a ferredoxin- and NAD (and not NADP)-dependent flavin-based electron-bifurcating hydrogenase [183]. The sequence alignment of ACL20074 with the characterised [FeFe] electron-bifurcating hydrogenases (taken from Figure S4 in Kpebe, 2018) revealed that the major sequence features are conserved in ACL20074 (Figure C.3). Moreover, it appears that the first 110 amino acids of ACL20074, that comprise the conserved [2Fe2S] cluster, match with the N-terminal sequence of the electron-bifurcating hydrogenases. The sequence homology then stops in the middle of the first [4Fe4S] cluster and starts again around position 580 of ACL20074. It is the intermediate region of ACL20074 that shows homology to SfrB from *G. sulfurreducens* (Figure C.3). The lack of the other Hnd subunits in the genetic environment of ACL20074, and most importantly HndC which harbours the flavin- and NAD-binding sites and is responsible for electron bifurcation, is possibly replaced by the intermediate region in ACL20074 that displays possible binding sites for FAD and NAD(P)H, according to [184]. Despite the clear sequence homology to the electron-bifurcating hydrogenases, no clear function can be assigned yet to this unusual composite hydrogenase. However, it seems to be widely distributed in the Firmicutes and invites further physiological and biochemical investigations.

One among several carbon monoxide dehydrogenase (CODH) encoded in the genome of strain DCB-2 was up-regulated in Pyr-only conditions. ACL21808, the CODH catalytic subunit (CooS), is part of a predicted 3-gene operon (ACL21807-9) that also displays a typical ferredoxin (CooF) and a CODH maturation protein. Sequence analysis revealed that ACL21808 belongs to the clade F of CODH enzymes [185]. Not up-regulated, ACL21806 (CooA) is a member of the CRP/FNR regulatory proteins and is likely regulating the transcription of the CODH encoding operon. Interestingly, the catalytic enzymes (ACL21807-8) are strongly up-regulated and mainly when compared to H<sub>2</sub> conditions and to La/CIOHPA conditions. Finally, the proteins encoded by a series of three consecutive genes (ACL22517-9) showed up-regulation in Pyr-only conditions. Although no clear function could be predicted, the sequence of ACL22518 and ACL22519 indicate that they are redox enzymes, suggesting a possible involvement in the energy metabolism of fermentatively growing strain DCB-2 cells.

Figure 6.2 also shows subsets of 25 and 14 proteins that are up-regulated in Pyr-only conditions in comparison to all other conditions except La/CIOHPA or H<sub>2</sub>/Fu, respectively. While the latter subset mainly consists of proteins harbouring radical SAM domain and other cofactors likely involved in anabolism, Table 6.2 lists a selection of the former subset (given in full in Table C.1). Here two clusters of proteins are noteworthy: a six-gene cluster (ACL21175-80) encoding a member of the complex iron-sulfur molybdoenzyme (CISM) superfamily and two subunits of a putative NiFe hydrogenase five-gene cluster (ACL21187-91). Representatives of the CISM superfamily in bacteria are extremely diverse in function and harbour

various gene composition [186]. Interestingly, a genomic survey has identified *D. hafniense* to have the largest set of molybdoenzymes, most of them belonging to the diverse DMSO reductase subfamily [187]. A thorough sequence analysis suggests that the ACL21175-80 gene cluster forms an operon, which displays, beside the typical CISM subunits A (the catalytic molybdoenzyme) and B (a ferredoxin), another copy of the B subunit and an enzyme showing low sequence homology to the NAD(P)H-nitrite reductase (NirB) or the CoA/CoA-disulfide reductase (CoADR). Therefore, this composite cluster of redox enzymes may indicate that, beside Pyr-only conditions, strain DCB-2 exploits additional sources of energy from minor components of the medium. The NiFe hydrogenase, that is encoded in ACL21287-91, contains, in addition to the typical large subunit and Tat-dependent small subunit, a Tat-dependent *c*-type cytochrome (ACL21287), and two proteins with unknown function. The replacement of *b*-type cytochrome by a *c*-type cytochrome in NiFe H<sub>2</sub>-uptake hydrogenases has been reported for most  $\delta$ -Proteobacteria [188]. The *c*-type cytochrome, which was also found up-regulated in fermentation (however only when compared to Py/Fu and La/Fu conditions), belongs to the family of split-Soret diheme *c*-type cytochromes (Ssc) [189] and may well receive electrons from the associated NiFe hydrogenase as it was proposed in the metabolism of *Desulfovibrio desulfuricans* [190]. The fact that the up-regulation level of this hydrogenase is substantially higher when comparing to Py/Fu and La/Fu conditions than to H<sub>2</sub> conditions may reflect the common need of hydrogenases in fermentative metabolism (likely producing H<sub>2</sub>) and in respiration with H<sub>2</sub> as electron donor. Proteome adaptations to La/CIOHPA in comparison to all other conditions indicate that it shares some metabolic pathways with fermentation, suggesting that cells growing with lactate and CIOHPA display a mixed energy metabolism with both respiratory and fermentative aspects. It is likely that CIOHPA is reduced from the oxidation of lactate to pyruvate via respiration, but also possible that the resulting pyruvate is then fermented.

The direct comparison of Pyr-only with Py/Fu respiration highlighted about 30 proteins with an up-regulation level higher than 10-fold (Table 6.3). Beside the proteins already discussed in Tables 6.1 and 6.2 (shaded in grey in Table 6.3), some proteins are noteworthy. A lactate transporter (ACL21424) was highly up-regulated in the pairwise comparisons of Pyr-only with Py/Fu, H<sub>2</sub>/Fu and H<sub>2</sub>/CIOHPA. On the contrary, it was not up-regulated when compared to lactate conditions. This suggests that ACL21424 is involved in both lactate uptake and secretion. The uptake function occurs when lactate is used as electron donor, while lactate needs to be secreted as one of the end products of Pyr-only conditions (see section 6.3.3 for more information about lactate metabolism).

Seven proteins belonging to a possible 14-gene operon (ACL18316-29) were highly up-regulated in Pyr-only conditions in comparison to Py/Fu (Figure C.4, A). Sequence analysis revealed that many proteins encoded in this cluster are homologous to enzymes involved in dissimilatory sulfite reduction (DSR), as in *Desulfovibrio* sp. [191, 192]. Whether all the 14 genes belong to the same operon is not clear. Indeed, no clear predicted function in dissimilatory sulfite reduction was found for the first four genes (ACL18316-9). Moreover, ACL18319 was not detected at all in the proteomic analysis. Nevertheless, if only considering the proteins

Protein ref.	Predicted function	Fold change in Pyr-only vs					
		Py/Fu	La/Fu	H2/Fu	H2/Cl	La/Cl	
ACL21175	4Fe-4S ferredoxin binding protein, CISM subunit B	8	13	6	7	<2	
ACL21176	Molybdopterine oxidoreductase, CISM subunit A	5	6	3	4	<2	
ACL21177	FAD-dependent NAD(P)H-disulfide oxidoreductase	40	31	15	28	<2	
ACL21178	4Fe-4S ferredoxin binding protein, DmsB-like	3	3	3	3	<2	
ACL21179	Protein of unknown function DUF81	73	83	28	84	<2	
ACL21180	Hypothetical protein	6	6	4	4	<2	
ACL21287	Tat-dependent <i>c</i> -type cytochrome	14	8	<2	<2	<2	
ACL21288	NiFe-dependent hydrogenase, large subunit	17	6	2	2	<2	
ACL21289	NiFe-dependent hydrogenase, small subunit	12	5	2	2	<2	

Table 6.2 – Selected proteins up-regulated in Pyr-only conditions in comparison to all respiratory conditions except La/CIOHPA.

Protein ref. 1	Predicted function	Fold change in Pyr-only vs				
		Py/Fu	La/Fu	H2/Fu	H2/Cl	La/Cl
ACL21179	Protein of unknown function DUf81	73	83	28	84	-
ACL21177	FAD-dependent NAD(P)H-disulfide oxidoreductase (CoA-disulfide reductase?)	40	31	15	28	-
ACL21424	L-lactate transporter	68	-	66	32	-
ACL21087	Hypothetical protein	59	-	14	8	-
ACL18321	Dissimilatory sulfite reductase, $\beta$ subunit (DsrB)	40	6	-	-	-
ACL18320	Dissimilatory sulfite reductase, $\alpha$ subunit (DsrA)	23	5	-	-	-
ACL18322	Dissimilatory sulfite reductase, unknown function (DsrD)	19	4	-	-	-
ACL18317	Protein of unknown function (ferredoxin?)	18	4	-	-	-
ACL18316	Nitrate reductase, $\gamma$ subunit (DsrM-like)	17	4	-	-	-
ACL18325	Dissimilatory sulfite reductase, 4Fe-4S ferredoxin (DsrK)	15	4	-	-	-
ACL18324	Dissimilatory sulfite reductase (DsrM)	13	4	-	-	-
ACL18314	Hypothetical protein	32	9	-	4	-
ACL18315	Hypothetical protein	22	5	-	3	-
ACL21108	NAD-dependent decarboxylating malate DH	30	8	10	9	9
ACL21109	Citryl-CoA lyase, $\beta$ subunit	22	7	14	11	6
ACL21110	Succinyl-CoA synthetase, $\alpha$ subunit	13	6	11	10	7
ACL21112	Succinyl-CoA synthetase, $\beta$ subunit	15	5	11	9	7
ACL21111	Hypothetical protein	11	5	9	9	5
ACL20535	Succinate:quinone oxidoreductase, flavoprotein	17	5	9	11	4
ACL21288	NiFe-dependent hydrogenase, large subunit	17	6	2	2	-
ACL21287	Tat-dependent c-type cytochrome	14	8	-	-	-
ACL21289	NiFe-dependent hydrogenase, small subunit	12	5	2	2	-
ACL21182	Hypothetical protein	16	15	8	12	-
ACL21183	Protein of unknown function DUf81	13	12	7	8	-
ACL20816	5-methyl-THF corrinoid/FeS methyltransferase	13	10	-	-	3
ACL20817	CODH/ACS, $\delta$ subunit	11	8	-	-	3
ACL20820	CODH/ACS, $\delta$ subunit	10	7	-	-	3
ACL22682	Hypothetical protein	13	4	-	-	-
ACL18084	Hypothetical protein	10	19	-	2	-
ACL18955	UDP-glucose/GDP-mannose dehydrogenase	10	10	-	-	-
ACL19345	Molybdopterin oxidoreductase, CISM subunit A	10	8	-	-	-

Table 6.3 – Proteins showing up-regulation level of at least 10-fold in Pyr-only conditions when compared to Py/Fu conditions. <sup>1</sup>Proteins are ordered in the table according to the value of fold change in Pyr-only vs Py/Fu and then directly followed by the proteins belonging to the same gene cluster (separated by dashed lines).

ACL18320-9, it is clear that DSR is induced in Pyr-only conditions (Table 6.3). One explanation for this observation is the addition of sodium sulfide as a reducing agent (at a final concentration of 1 mM) in the anaerobic medium preparation. It is proposed that sulfide is at least partially oxidised abiotically to a variety of sulfur compounds in the medium [193], thus likely producing sulfur or sulfite which can be used by *Desulfitobacterium* spp. as electron acceptor [17, 34]. Although the concentration of sulfite in the medium remains probably low in the culture medium, it would be enough to induce the expression of the *dsr* gene cluster and possibly also to account for some respiratory metabolism, or at least for an improved energy conservation during fermentation. The behaviour of ACL18320-29 proteins in the pairwise comparison across all six growth conditions suggests that DSR metabolism is also strongly triggered during respiration with H<sub>2</sub> as electron donor and in La/CIOHPA conditions (Figure C.4, B). The induction of alternative energetic pathways such as DSR is possibly correlated with the lower growth yields observed for these conditions in comparison to Py/Fu and La/Fu (see Figure C.1). Similarly, *D. hafniense* strain TCE1 has shown a higher growth yield in La/Fu conditions compared to H<sub>2</sub>/PCE [194].

Table 6.3 also reports on several proteins that are highly up-regulated in Pyr-only conditions but do not display a clear function. Noteworthy, however, are the proteins ACL20816, -17 and -20, which belong the gene cluster ACL20816-23 that is mainly responsible for the Methyl branch of the Wood-Ljungdahl pathway (WLP). These proteins are clearly up-regulated in Pyr-only conditions, but also in respiration with H<sub>2</sub> as electron donor (see section 6.3.3 for a detailed discussion on the WLP).

Taken altogether, the proteomic dataset highlighting adaptations to the metabolism of fermentation is complex. The complexity in the response of the proteome from cells fermenting pyruvate with that of cells respiring various electron donors and acceptors is reflected by the distribution of all up-regulated proteins during fermentation in comparison to any combination of electron donor and acceptor, and by the small number of proteins commonly up-regulated across all pairwise comparisons. Nevertheless, this analysis mainly revealed proteins catalysing reactions involved in the central carbon metabolism (TCA-related activity, CO<sub>2</sub> assimilation, lactate transport, ...) and in the energy metabolism (H<sub>2</sub> metabolism, various redox enzymes, ...). Interestingly, an unexpected outcome was the induction of the dissimilatory sulfite reduction pathway in fermentation but also in respiratory conditions associated with relatively low growth yields. This reflects the high versatility and flexibility of the energy metabolism in *D. hafniense* and invites us to consider that partial sulfite respiration was likely occurring in the medium composition used for fermentation. Also, a new type of cytoplasmic hydrogenase showing sequence homology to electron-bifurcating NAD-dependent hydrogenases is possibly involved in the NAD<sup>+</sup>/NADH homeostasis, a function that appears to be specially required during fermentation.

### 6.3.3 Versatility of respiratory metabolisms

#### Proteome adaptations to electron donors

In this section, both electron donors lactate and molecular hydrogen ( $H_2$ ) were compared. The respective proteome adaptations were highlighted by confronting two pairwise comparisons of growth conditions. Indeed, by comparing  $H_2$ /CLOHPA vs La/CLOHPA and  $H_2$ /Fu vs La/Fu, the proteins specific to  $H_2$  (i.e. which were up-regulated in both pairwise comparisons) or lactate (down-regulated in both pairwise comparisons) were identified. These results are displayed in Figure 6.3.

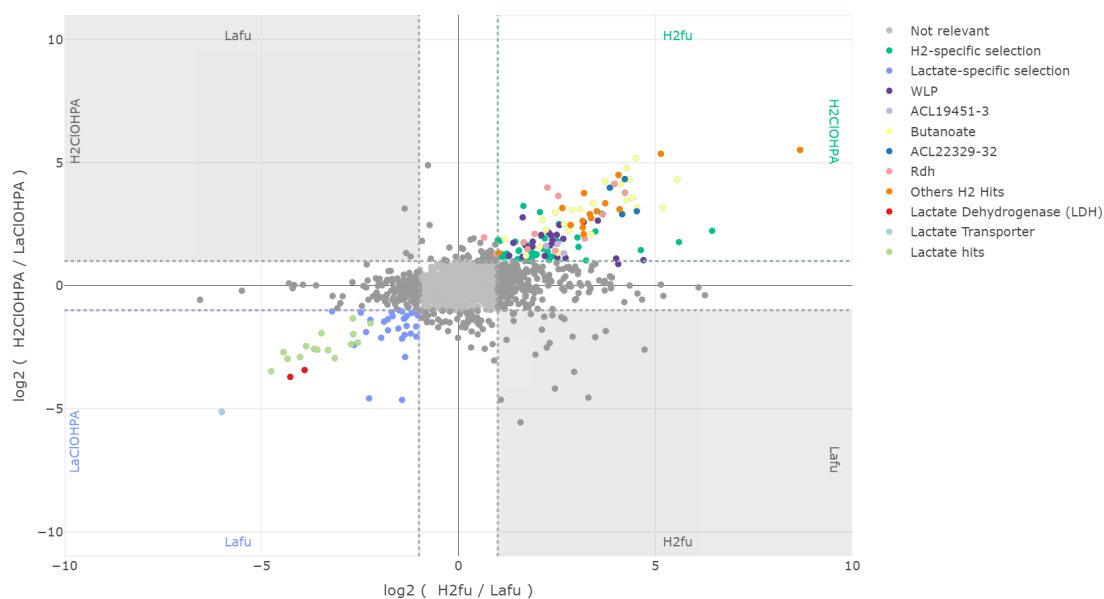


Figure 6.3 – Proteome adaptations to the electron donor.

#### $H_2$ -specific proteome adaptations

The top right panel in Figure 6.3 contains all the proteins which were up-regulated in both conditions with  $H_2$  as electron donor when compared to the corresponding conditions with lactate. A total of 123 proteins were observed in this selection which was larger than any of the other substrate-specific selections that are presented in the later sections. This high number of  $H_2$ -specific proteins can be partially explained by the fact that many of them are encoded by genes that are part of predicted large operons. The most relevant proteins of that selection are displayed in Table 6.4.

Table 6.4 – Proteome adaptations to the use of H<sub>2</sub> as electron donor. Selection of the most relevant proteins up-regulated with H<sub>2</sub> as electron donor. The color code indicates the different metabolic pathways (as on the scatter plot of Figure 6.3).

Group	Accession	Description
	ACL18146	Methylene-THF reductase
	ACL18217	Formate-THF ligase
	ACL18622	Formate-THF ligase
	ACL18623	Formiminotransferase-cyclodeaminase
	ACL18624	Methylene-THF dehydrogenase (NADP(+))
	ACL18944	4Fe-4S ferredoxin iron-sulfur protein
	ACL18945	CODH, catalytic subunit
	ACL18946	FAD-dependent NAD(P)-disulfide oxidoreductase
	ACL20816	5-methyl-TFH corrinoid/FeS protein methyltransferase
	ACL20817	CODH/ACS, $\delta$ subunit
	ACL20819	Ferredoxin
	ACL20820	CODH/ACS, $\delta$ subunit
	ACL20821	CODH/ACS, $\beta$ subunit
	ACL20822	CODH, catalytic subunit
	ACL21899	5-methyl-TFH corrinoid/FeS protein methyltransferase
	ACL21900	Hypothetical protein
	ACL21901	Hypothetical protein
	ACL21902	Cobalamin-binding domain protein
	ACL22046	Glycine dehydrogenase (decarboxylating)
	ACL22047	Glycine dehydrogenase (decarboxylating)
	ACL22273	Formate dehydrogenase, formation protein FdhE
	ACL22274	Formate dehydrogenase, $\gamma$ subunit
	ACL22275	4Fe-4S ferredoxin domain protein
	ACL22276	Formate dehydrogenase, $\alpha$ subunit
	ACL19451	5-formyl-THF cyclo-ligase
	ACL19452	NADH dehydrogenase, NuoE-like subunit
	ACL19453	NADH dehydrogenase, NuoF-like subunit
	ACL20879	Enoyl-CoA hydratase/isomerase
	ACL20880	Acetyl-CoA acetyltransferase
	ACL20881	Acetyl-CoA hydrolase/transferase
	ACL20882	Hypothetical protein
	ACL20883	Unknown function DUF224 cysteine-rich region domain protein
	ACL20884	Electron transfer flavoprotein, $\alpha$ subunit
	ACL20885	Electron transfer flavoprotein, $\beta$ subunit
	ACL20886	Acyl-CoA dehydrogenase domain protein
	ACL20887	3-hydroxybutyryl-CoA dehydrogenase
Continued on next page		

Table 6.4 – continued from previous page

Group	Accession	Description
	ACL20888	Acyl-CoA dehydrogenase domain protein
	ACL20889	Sigma54 specific transcriptional regulator, Fis family
	ACL21561	Phosphate butyryltransferase
	ACL21563	Acetyl-CoA acetyltransferase
	ACL22520	GntR domain protein
	ACL22529	Unknown function DUF224 cysteine-rich region domain protein
	ACL22530	Acetyl-CoA hydrolase/transferase
	ACL22531	Acyl-CoA dehydrogenase domain protein
	ACL22532	Acetyl-CoA acetyltransferase
	ACL22533	3-hydroxyacyl-CoA NAD-binding dehydrogenase
	ACL22535	Electron transfer flavoprotein, $\alpha$ subunit
	ACL22536	Electron transfer flavoprotein, $\beta$ subunit
	ACL22537	NADH:flavin oxidoreductase/NADH oxidase
	ACL22538	Enoyl-CoA hydratase/isomerase
	ACL22329	Betaine transporter, BCCT
	ACL22330	Methyltransferase, cognate corrinoid protein
	ACL22331	Glycine betaine methyltransferase, MtgB
	ACL22332	Methylcobalamin:THF methyltransferase, MtrH
	ACL18082	Pyridoxamine 5-phosphate oxidase-related FMN-binding
	ACL18534	Acetate/CoA ligase
	ACL18581	Fe-containing alcohol dehydrogenase
	ACL19102	ABC transporter
	ACL19103	Extracellular solute-binding protein
	ACL20568	Hypothetical protein
	ACL20673	Short-chain dehydrogenase/reductase, SDR
	ACL20674	Molybdopterin oxidoreductase
	ACL20850	Hypothetical protein
	ACL20994	Sigma54 specific transcriptional regulator, Fis family
	ACL21054	Hypothetical protein
	ACL21793	4Fe-4S ferredoxin
	ACL22411	Transcriptional regulator, MarR family
	ACL22699	Beta-lactamase domain protein

Interestingly, the H<sub>2</sub>-specific protein selection is composed mostly by proteins involved in the carbon metabolism rather than in the energy metabolism. This suggests that many observed changes in protein expression is a consequence of the use of acetate as carbon source as opposed to lactate rather than due to the direct use of H<sub>2</sub> as electron donor. In particular, many proteins of the selection were either directly or indirectly linked to the Wood-Ljungdahl pathway (WLP). Indeed, 6 out of 8 proteins encoded by the CO dehydrogenase/acetyl-CoA

synthase (CODH/ACS) operon (ACL20816-22) were up-regulated in H<sub>2</sub> conditions. Additionally, the three subunits composing the formate dehydrogenase (ACL22274-76) and its associated maturation factor (ACL22273) were also identified in the selection, supporting the WLP pathway as previously proposed by Kim *et al.* [34]. The H<sub>2</sub>-specific protein selection also contains one or more protein homologues for each step of the Methyl branch of the WLP (see table 6.4 and the metabolic map in Figure 6.4). In addition, an alternative way to produce 5,10-methenyl-THF is provided by ACL19451, which was also up-regulated. Interestingly, the two proteins directly following ACL19451, initially annotated as NuoE (ACL19452) and NuoF (ACL19453), were also up-regulated with a similar behaviour. Although not up-regulated in H<sub>2</sub>/CLOHPA vs La/CLOHPA, ACL19454, a molybdoenzyme of the CISM family, is likely part of the same operon (ACL19451-4). ACL19452-4 is likely to form an NAD<sup>+</sup>-dependent formate dehydrogenase, similarly to the FdsABG complex identified in *Rhodobacter capsulatus* [195]. This suggests that the NuoE- and NuoF-like proteins could be indirectly associated with the WLP rather than forming an incomplete NADH-oxidising module of complex I as previously proposed [117] (more details about complex I-related proteomics are discussed in Chapter 7). Finally, a four-genes cluster (ACL22329-32) appear to be relatively highly up-regulated. As an example, ACL22330 is found 32-times more expressed in H<sub>2</sub>/CLOHPA compared to La/CLOHPA. Interestingly, this cluster was previously identified as a putative demethylase operon [196]. The methyltransferase subunit (ACL22331) encoded by this operon is homologous to MtgB in *D. hafniense* strain Y51, which has been shown to use glycine betaine as methyl donor to methylate tetrahydrofolate (THF), which can be oxidised to CO<sub>2</sub> and serve as alternative electron donor in respiration [197]. The involvement of an O-demethylase in the use of phenyl methyl ethers such as vanillate to follow the Methyl branch of WLP in the reverse direction was also proposed for strains DCB-2 and Y51 in the context of growth on vanillate [34]. Alternatively, this route represents also a way to by-pass the Methyl branch of the WLP by producing methyl-THF which can enter the carbonyl branch directly leading to the production of acetate. The growth medium applied here, however, does not contain any source of glycine betaine. Therefore, the physiological role of this gene cluster in the adaptation of the proteome toward the use of H<sub>2</sub> remains to be elucidated.

The higher expression level of numerous proteins directly or indirectly linked with the WLP in the H<sub>2</sub>-conditions may be a consequence of the co-occurrence of H<sub>2</sub> and CO<sub>2</sub> which could trigger the expression of the pathway. This could also suggest that the bacteria rely on the WLP to produce acetyl-CoA from CO<sub>2</sub> even though acetate was added as organic carbon source. The first step in the assimilation of acetate, catalysed by the acetate kinase (ACL21842), is an energy-consuming process and may not be the preferred route in bacteria growing on H<sub>2</sub> and harbouring the WLP. The down-regulation of ACL21842 in H<sub>2</sub> compared to La conditions may corroborate this hypothesis. Dedicated experiments should be performed to determine the nature of the carbon source in strain DCB-2 growing on H<sub>2</sub>. A mixotrophic growth (combination of autotrophic and heterotrophic carbon assimilation mechanisms) has already been proposed for other OHRB [6, 55, 198].

Besides the WLP-related proteins, two large clusters (ACL20879-89 and ACL22530-38) and a

small one (ACL21561-3) seem to contain at least one protein that can be mapped to the butyrate metabolism. Both of the large clusters include predicted flavin-based electron bifurcating proteins, a mechanism actually discovered in butyrate-forming anaerobes [63]. Moreover, the gene cluster ACL20879-89, and to a lesser extent the ACL22530-38 cluster, show a similar genetic organisation as the typical electron-bifurcating gene cluster of *Clostridium kluyveri* that was shown to be involved in butyrate metabolism [199–201] (Figure C.5). This strongly suggested that strain DCB-2 is able to produce butyrate from acetyl-CoA. Indeed, at first sight, strain DCB-2 does not have the possibility to produce butyrate as ACL20886, proposed to perform the conversion from butyryl-CoA to crotonyl-CoA, does not perform the reverse reduction reaction. To the best of our knowledge, strain DCB-2 is not known as a butyrate-producing organism. But, as no butyrate is present in the growth medium and a large amount of acetyl-CoA might result from the use of WLP, it is more likely that strain DCB-2 uses the butyrate synthesis pathway (which would become possible by the bifurcation mechanism) as an additional route for substrate-level phosphorylation. This hypothesis should be verified by assessing the production of butyrate in cultures of strain DCB-2. All the proteins of the selection that could be mapped to the butyrate pathway are displayed on the metabolic map in Figure 6.4, together with the WLP pathway.

On top of those two main metabolic pathways, some proteins encoded by the reductive dehalogenase clusters (Rdh) were also appearing in the H<sub>2</sub> selection, a surprising observation which will be discussed in the section dedicated to ClOHPA-specific hits. Similarly, the putative fumarate reductase (ACL18807-9) was found up-regulated in the H<sub>2</sub>-specific selection but the reason for a higher expression of this enzyme in H<sub>2</sub> conditions is not yet understood. Nevertheless, these proteins will be more extensively discussed in the fumarate-specific section. In addition, ACL18581, annotated as alcohol dehydrogenase and ACL18082, a protein potentially involved in the vitamin B6 synthesis, appeared with the most extreme level of up-regulation (up to 60 times more expressed in H<sub>2</sub>-conditions). The annotation of the first one, can be misleading as it represents a quite general term and the fact to find no other protein associated or with a similarly extreme behaviour does not help to discuss the potential function of this protein. Similarly, ACL18082 is not supported by other proteins encoded in its close vicinity. The over-expression of vitamin biosynthesis might indicate a poor energetic state of the cells which maybe try to compensate for important missing intermediates. This same reason might explain the presence of two other proteins (ACL19102-03) which seem to be involved in the active transport of solutes. Arguments to explain the presence of the latter examples in the H<sub>2</sub>-specific selection still need to be further developed.

Other hits are worth a notice as they represent interesting candidates for important metabolic processes. For example, the predicted acetate-CoA ligase (ACL18534) could be involved in the active carbon fixation from acetate as already proposed by [34]. Indeed, ACL18534 was already proposed to perform the conversion from acetate to acetyl-CoA with the use of ATP. In the present conditions, acetate is the only organic source of carbon and, as mentioned before, carbon assimilation is thus expected to happen in addition to autotrophic CO<sub>2</sub> fixation through the WLP. Additionally, the ferredoxin (ACL21793), also present in the H<sub>2</sub>-specific adaptation,

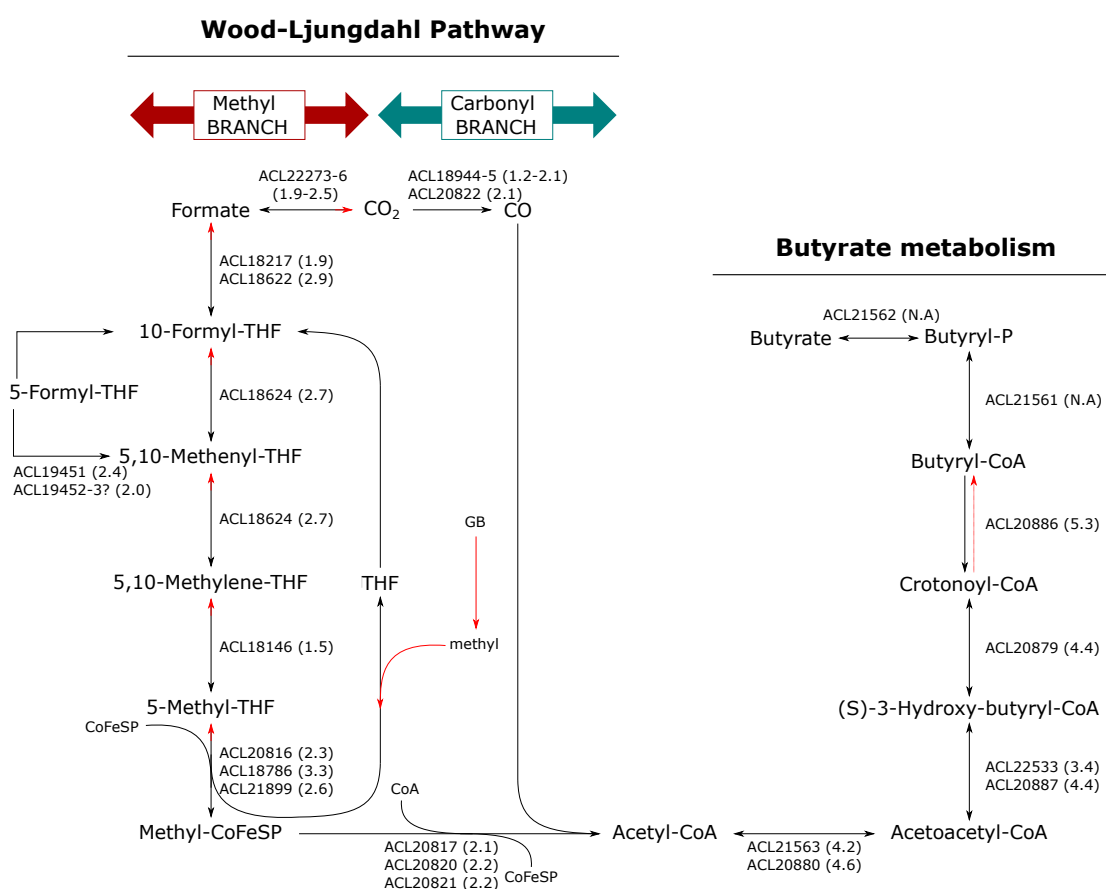


Figure 6.4 – Wood-Ljungdahl pathway and butyrate metabolism in strain DCB-2. Metabolic map displaying both pathways linked to each other by acetyl-CoA. Each arrow represent a reaction that is proposed to be performed by one of the proteins that were over-expressed in  $H_2$  conditions as compared to lactate. The fold changes of  $H_2/Fu$  vs  $La/Fu$  and  $H_2/CIOHPA$  vs  $La/Fu$  were averaged and are indicated in brackets for each protein. (Note: ACL21561-62 were not included in the  $H_2$ -selection)

could represent an important alternative electron shuttle in the  $H_2$ -related energy metabolism. In any case, additional experiments would be needed to confirm those hypotheses. Several other proteins are listed in Table 6.4 as they showed high up-regulation ratios (at least 5-times in the two pairwise comparisons). However, a clear function could not be assessed to them and therefore they will not be further discussed.

As already discussed, all of the above mentioned proteins probably reflect the need to find alternative ways of energy conservation and/or carbon fixation when growing in  $H_2$  and acetate conditions. In comparison, the use of lactate which can undergo two rounds of oxidation (to pyruvate and then to acetyl-CoA) is probably a significant physiological advantage as it provides energy and carbon at the same time. This might explain the much higher growth performances of strain DCB-2 in lactate conditions (Note that  $La/CIOHPA$  is a particular case, as already discussed in section 6.3.2). None of the proteins of the  $H_2$  selection could

be linked to the direct oxidation of molecular hydrogen. Yet, the genome of DCB-2 encodes eight different hydrogenases but none of them seems to be specialised for the growth in  $H_2$  conditions. In order to better understand the use of hydrogenases across the tested conditions, the eight hydrogenases were plotted according to their relative abundance in the 18 samples as expressed by a Z-score. The resulting graph, displayed in Figure 6.5, revealed that one of the four NiFe hydrogenase (ACL20766-9) seems to dominate in  $H_2$ /C10HPA conditions. Although the same is not observed in  $H_2$ /Fu in which the hydrogenase appears at the same level as in Py/Fu or La/C10HPA, it remains a serious candidate as uptake hydrogenase delivering electrons to menaquinones in the electron transfer chain. Indeed, the encoded proteins correspond to a three-subunit hydrogenase complex (HydABC) along with its associated maturation factor (HydD). Similar membrane-bound complexes have already been mentioned in the context of OHR in other strains [68] and belong to the group 1 hydrogenases [202]. By contrast, the relative abundance of the multi-subunits NiFe-hydrogenase belonging to the energy-converting hydrogenases (Ech)-type hydrogenases (ACL22288-93) of group 4 [202] appeared less abundant in the two conditions involving  $H_2$  as electron donor in comparison to all the other tested conditions. In acetoclastic methanogens, the Ech-type hydrogenase was proposed to work in both directions (proton reduction or  $H_2$  oxidation) using a cytoplasmic ferredoxin as redox partner while transferring protons through the membrane [203]. If the enzyme has a similar function in strain DCB-2, one could imagine that its use is limited in  $H_2$  conditions because the electron transfer chain does not involve any cytoplasmic intermediate.

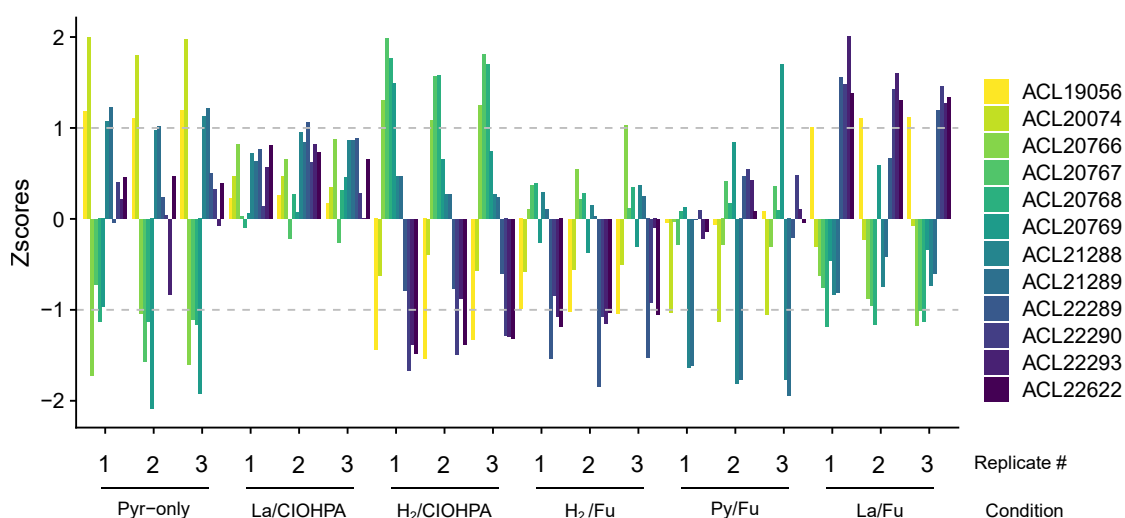


Figure 6.5 – Differential expression of strain DCB-2 hydrogenases in the six growth conditions. The relative abundance of the subunits composing the selected enzymes in each of the 18 samples is expressed as Z-scores in order to evaluate the amplitude of variation of each protein across the proteomic dataset.

### Lactate-specific proteome adaptations

In the bottom left panel of the scatter plot displayed in Figure 6.3 are shown the proteins up-regulated when lactate was used as electron donor in comparison to H<sub>2</sub>. This selection is composed of 48 proteins in total which, in contrast to H<sub>2</sub> selection, were rarely encoded by large gene clusters. Exceptions are the eight proteins involved in the UMP synthesis pathway (ACL22022-9) and six proteins forming a complex iron-sulfur molybdoenzyme (CISM) with two tandem-domain rhodanese proteins (ACL19560-65). A similar cluster was already described in *D. hafniense* strain TCE1 where it was proposed to be involved in sulfur metabolism [204]. At the time, only one copy of the rhodanese subunits was detected with a 4-fold ratio in La/PCE in comparison to La/Fu. In contrast, the present analysis identified the complete CISM cluster up-regulated in lactate conditions. Up-regulation was also observed when pyruvate was used as electron donor, or fumarate as electron acceptor, thus suggesting that this cluster is rather down-regulated in H<sub>2</sub> conditions and during Pyr-only conditions. A member of the CISM family has been proposed to be part of the quinone-independent OHR respiratory chain in *D. mccartyi* [51], which bolstered the interest for this type of membrane-bound complexes in the metabolism of OHRB. However, it is still unclear whether the CISM complex is an intermediate in the electron transfer chain or represents an alternative route for the electrons in presence of organic electron donors in strain DCB-2.

In fact, a striking feature is noteworthy as the majority of the proteins in the lactate-specific selection were also up-regulated in Py/Fu or Pyr-only conditions when compared to H<sub>2</sub> conditions. Thus, their presence in the lactate-specific corner of the scatter plot may be the result of the comparison with H<sub>2</sub> conditions mainly. In other terms, the proteins composing this selection probably rather reflect a down-regulation when H<sub>2</sub> is used as electron donor than proteins truly up-regulated due to the presence of lactate.

To reveal the true lactate hits, the lactate-specific selection revealed by the comparison to H<sub>2</sub> was highlighted in separate volcano plots comparing La/Fu to Py/Fu and to Pyr-only conditions (Figure 6.6A and B, respectively). The idea was to see which of the proteins composing the selection was also significantly up-regulated when comparing to these two latter conditions and thus might be considered as true lactate specific hits. As shown in Figure 6.6A and B, most of the proteins of the lactate-specific selection appeared in the centre of the volcano plots meaning that only a few of them were also significantly up-regulated in Py/Fu and/or Pyr-only (with logFC values between -1 and 1). The proteins of the lactate-specific selection that were found with a minimum of 2-fold ratio when comparing La/Fu with either Py/Fu or Pyr-only are listed in Figure 6.6 C and D, respectively. None of the proteins highlighted there was up-regulated both against Pyr-only and against Py/Fu, revealing that none of them can be affiliated to lactate alone. These overlapping results can most probably be explained by the fact that lactate is oxidised to pyruvate and thus cells in La/Fu conditions must also express the proteins responsible for pyruvate utilisation. Although no true lactate-only hit could be identified, a few candidates of the lactate-specific proteins display interesting properties which are detailed below.

At first, ACL21424 which occupies the most extreme position in the scatter plot was identified as the lactate transporter. Secondly, two proteins (ACL21106-7) with unknown function are encoded by a small gene cluster and also appear with relatively high logFC values. Protein domain prediction indicated that they were both harbouring a lactate utilisation domain (LUD) and therefore caught our interest. Sequence analysis with BlastP, revealed that these two proteins display significant similarity to LutBC of *Bacillus subtilis* [205]. In contrast, while being also part of the La vs H<sub>2</sub> selection, no evidence allowed us to assign ACL21105 as a LutA homologue. Nevertheless, we propose ACL21106-7 to form the lactate dehydrogenase (LDH) active in the respiratory chain of strain DCB-2. Interestingly, these two proteins also appeared as up-regulated in La/Fu vs Py/Fu (Figure 6.6A and C) but not vs Pyr-only (Figure 6.6B and D). The fact that the LDH would be present in fermentative conditions at similar level as when lactate is the electron donor can be explained by the production of lactate from pyruvate in the lactic acid fermentation pathway, which could induce the expression of the LDH. The same situation is observed also for the lactate transporter (ACL21424), which is up-regulated in La/Fu vs Py/Fu but not vs Pyr-only. This could also indicate that in the Py/Fu conditions, all the pyruvate is converted to acetyl-CoA and thus there is no need for lactate-specific enzymes as there is no lactate production.

Except for these three proteins linked to lactate utilisation, the proteome observed in La/Fu conditions remains globally closer to that in Py/Fu than to Pyr-only conditions, which is translated by a higher number of differently expressed proteins in La/Fu vs Pyr-only than in La/Fu vs Py/Fu. Interestingly, the already mentioned CISM cluster seem to be restricted to the two respiratory systems which may indicate its involvement as an intermediate in the respiratory chain.

In conclusion, the comparison between the use of H<sub>2</sub> and lactate as electron donor mainly revealed differences in the carbon and/or central metabolism and in the direct uptake machinery. It highlighted the importance of alternative energy production pathways in cells growing in H<sub>2</sub> conditions with acetate as a carbon source. Besides a few proteins, no strong indications for the involvement of substrate-specific redox enzymes in the respiratory chain could be obtained. This suggests that the electron flow from the electron donor via the electron-donating moiety of the respiratory chain (either the H<sub>2</sub>-uptake hydrogenase or the lactate dehydrogenase) to the electron-acceptor moiety (see below) via redox enzymes that do not differ substantially with the different electron donors and acceptors used. This is in agreement with the existing OHR model proposing that the uptake hydrogenase transfers the electrons directly to menaquinone which transfers them further to the final reductase. This model matches with respiratory systems based on H<sub>2</sub> as electron donor. In contrast, when lactate is used, a cytosolic electron shuttle is expected to transfer electrons from the LDH complex to the menaquinone pool. The fact that no additional redox enzyme could be identified does not exclude their presence at a steady-state level in all growth conditions.

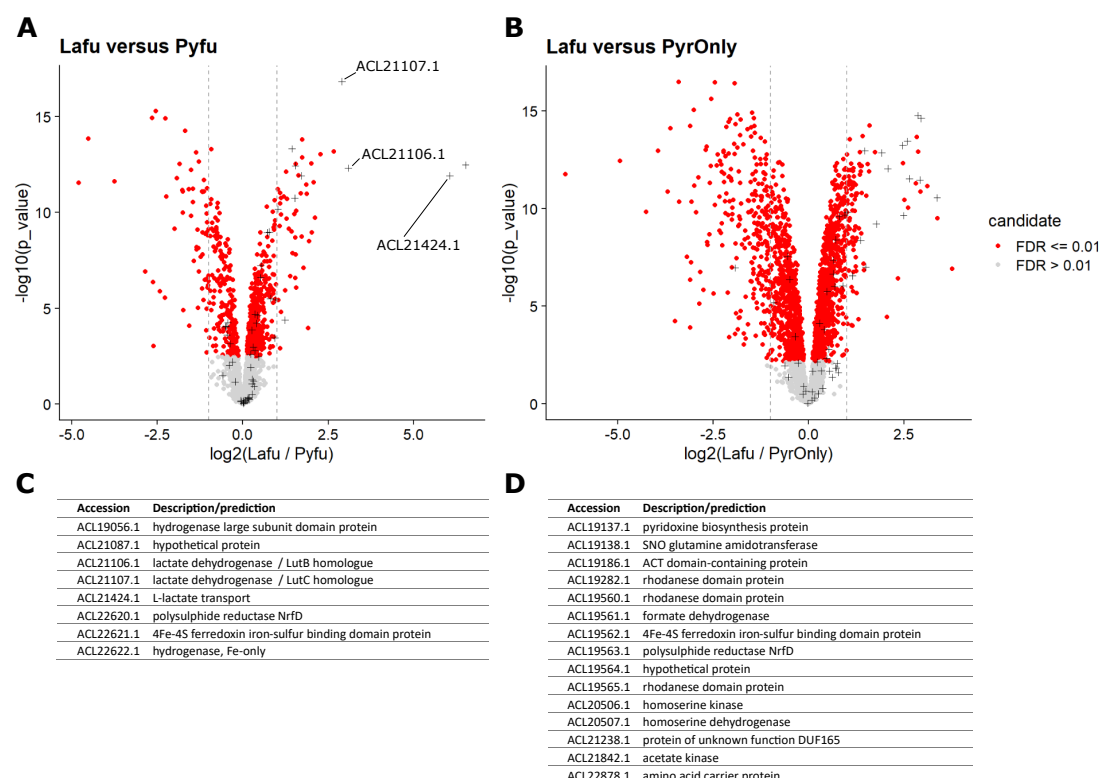


Figure 6.6 – Pairwise comparisons of La/Fu vs Py/Fu or Pyr-only. Volcano plots representing the relative abundance of all the proteins of the dataset in La/Fu conditions compared to Py/Fu (A) or to Pyr-only (B). The fold-change ratio associated to each proteins corresponds to the x-axis and is expressed in  $\log_2$ . The  $p$ -value of each results is displayed on the y-axis and expressed in  $\log_{10}$ . The red dots are representing the significant values ( $FDR \leq 0.01$ ) as opposed to the grey dots which are considered as not significant ( $FDR > 0.01$ ). The proteins of the lactate-specific adaptation as revealed by comparing lactate to  $H_2$  (found in the left down corner of Figure 6.3) are marked by black crosses. The proteins from the lactate-specific selection which were at least 2-times more abundant in Py/Fu or Pyr-only when opposed to La/Fu are listed in Tables (C) and (D), respectively.

### Proteome adaptations to electron acceptors

The same type of analyses were performed with a focus on the change in protein expression triggered by the change in the electron acceptor. Similarly, by combining the two pairwise comparisons La/Fu vs La/CIOHPA and  $H_2$ /Fu vs  $H_2$ /CIOHPA, the fumarate- and CIOHPA-specific proteins could be highlighted using a scatter plot representation, as displayed in Figure 6.7. As in previous section, the results are presented in separated subsections relative to either fumarate or CIOHPA.

### Fumarate specific proteome adaptations

The fumarate specific proteins are displayed in the down left corner of Figure 6.7. What is striking in this panel is that there are very few proteins. Indeed, only five proteins were identified with a logFC value above 1 in cells exposed to fumarate instead of CIOHPA, and, as it can be seen in the figure, all of them are only slightly exceeding a fold-change ratio of 2. This indicated that they were only weakly up-regulated in the fumarate conditions. Furthermore these proteins do not have a clear predicted function which could justify their presence in this selection. Although one could imagine that the nature of the final electron acceptor does not influence the composition of proteins involved in electron transfer chain, an up-regulation of the terminal fumarate reductase was expected. While 15 out of the 19 proteins identified as the flavoprotein subunit belonging to the succinate:quinone oxidoreductase family (with the quinol:fumarate reductases (QFR) among them) in the genome of strain DCB-2 were successfully quantified in the analysis. Although none of them has been characterised yet, these flavoproteins have been proposed to play a relatively large variety of functions in strain DCB-2 [34]. The fact that none of them is showing up in the fumarate-specific selection was quite surprising and it did not help identifying the terminal reductase involved in fumarate respiration. As briefly mentioned in the section dedicated to the H<sub>2</sub>-specific adaptations, ACL18807-9 represent the most promising QFR candidate. Indeed, the proteins encoded in this three-gene cluster show similarities with the FrdABC complex described in *Wolinella succinogenes* [180]. Among the 19 flavoproteins, only ACL18808 is surrounded by genes predicted to encode both a *b*-type cytochrome (ACL18807) and an FeS protein (ACL18809), similarly to the gene composition in *W. succinogenes*. When looking at the relative expression level (expressed in Z-scores), none of the flavoproteins (including ACL18808) was found specifically abundant in the fumarate conditions across the 18 samples. The proteins of the gene cluster ACL18807-9 were even more abundant in H<sub>2</sub> conditions which might indicate that ACL18807-9 were rather expressed at steady state level in all tested conditions and thus might be involved in other more general metabolic pathways beside dissimilatory fumarate reduction. As for example, strain DCB-2 is predicted to have an incomplete TCA cycle where the reduction of fumarate to succinate is known to occur.

### OHR and CIOHPA-specific proteome adaptations

Contrasting with the small number of proteins found in the fumarate conditions, 39 proteins appeared significantly up-regulated in the CIOHPA conditions when opposed to fumarate. The CIOHPA-specific proteome, displayed in the top right corner of the scatter plot of Figure 6.7, is composed mostly by proteins that are part of or directly adjacent to reductive dehalogenase (*rdh*) gene clusters (Figure 6.8A). In total, proteins from five out of the seven different *rdh* gene clusters were identified in this selection. Most of the proteins encoded in gene clusters 5 (ACL18776-83) and 6 (ACL18796-804) were successfully detected and quantified allowing for an almost complete coverage on these two clusters. As expected, proteins from cluster 6, known as dedicated to the respiration of CIOHPA [37, 40], were showing the highest up-regulation factors with RdhA6 (ACL18801) being up to 305 times more expressed in La/CIOHPA

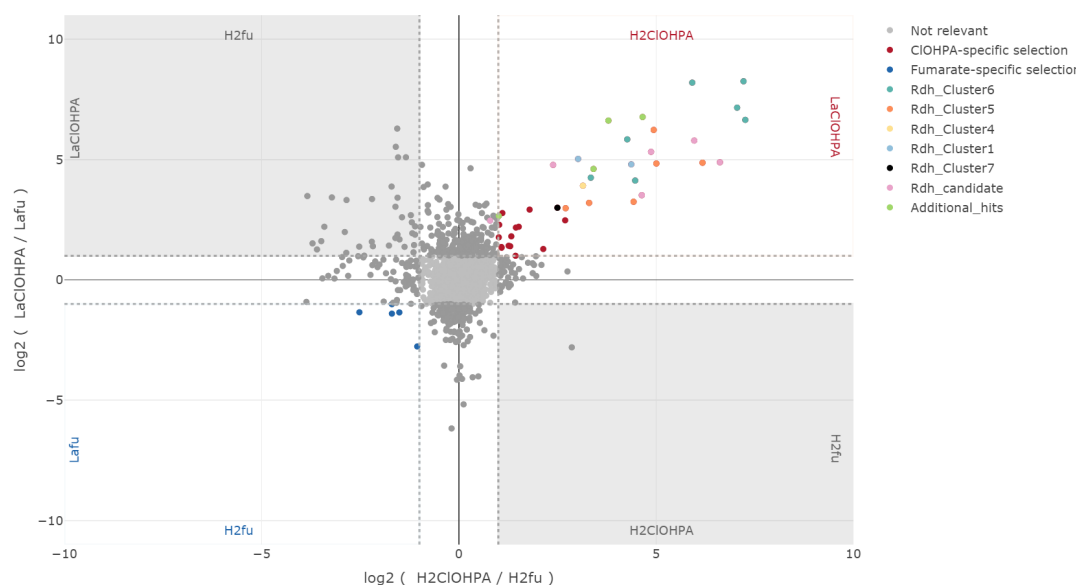


Figure 6.7 – Proteome adaptations to the electron acceptor.

in comparison to La/Fu and 149 times in H<sub>2</sub>/CIOHPA vs H<sub>2</sub>/Fu. More surprisingly, all the detected proteins of cluster 5 were showing high expression ratios between 7 and 75 times in CIOHPA conditions compared to fumarate conditions. Interestingly, four proteins encoded by genes that are directly downstream to cluster 5 (ACL18784-7) appeared with similar up-regulation factors as the latter. Although the role of these proteins remains to be elucidated, we propose that they might be considered as potential new members of the Rdh protein family.

In addition to *rdh* gene clusters 5 and 6, isolated Rdh proteins were also up-regulated in CIOHPA conditions: the reductive dehalogenase RdhA4 (ACL18775), a chemotaxis sensory transducer (RdhO1, ACL18750), and the CRP/FNR transcriptional regulators RdhK1 (ACL18751) and RdhK7 (ACL20644). One protein, encoded by a gene preceding the cluster 4 (ACL18768) followed the Rdh trend and thus is also proposed to be considered as such. However, this proposition remains to be taken carefully as the gene is not directly adjacent to the *rdh* gene cluster 4 and its function remains to be identified.

The induction by CIOHPA of the transcription of the gene cluster in *D. dehalogenans*, that is homologous to *rdh* gene cluster 6 of strain DCB-2, was shown already 20 years ago through Northern blot analysis [37] and, more recently, confirmed by proteomic analysis of the same OHRB [68]. In addition, transcriptomic analysis performed on *D. hafniense* strain DCB-2 already revealed the induction of the transcription of *rdhA4*, *rdhA5* and *rdhA6* genes upon addition of 3-chloro-4-hydroxybenzoate (CLOHBA), a compound highly similar to CIOHPA [34], which is coherent with the present data. However, the level of up-regulation obtained by Kim *et al.* were significantly lower than the one obtained in the present work. Kim *et al.* have also reported RdhA3 to be up-regulated by the presence of CLOHBA although to a lesser extent than in presence of 3,5-dichlorophenol [34]. Despite the high protein coverage obtained in

the present analysis, no proteins of the *rdh* gene cluster 3 was detected in the dataset which does not match with the proposition made by Kim *et al.* This later mentioned discrepancy can be partially explained that the published data were obtained by Northern blot targeting only the reductive dehalogenase (*rdhA*) at the transcript level which complicate the comparison to our proteomic results. Also, it cannot be excluded that the difference in the compounds used to induce gene expression in the two studies (CLOHPA vs. CLOHBA) may trigger a slightly different induction.

Taken all together, our data provide significant and robust evidences for the involvement of many of the Rdh proteins in CLOHPA respiration; an observation that reflects the relatively permissive *rdh* response already reported upon addition of this particular compound and other organohalides [72, 87]. Whether this involvement is direct or indirect remains to be tested, but the high level of coverage of the proteins encoded by the *rdh* gene clusters gives a new dimension to the understanding as so far only RdhA proteins were mainly detected [34, 72]. For example, the fact that the regulators from four different *rdh* gene clusters (RdhK1, RdhK5, RdhK6 and RdhK7) were up-regulated in the presence of a single organohalide compound further supports the existence a cross-talk within the OHR regulation network [7, 70]. However, it does not help understanding which RdhK regulator targets which *rdh* gene clusters, reinforcing the need to apply more direct methods (such as *in vitro* RdhK protein characterisation) in combination with global proteomic analysis as the one presented here. Interestingly, RdhK4 (i.e. CprK2), previously reported to be activated in presence of CLOHPA in an *in vivo* reporter assay [87] does not seem to be up-regulated in presence of the latter; thus, recalling the importance to apply physiologically relevant approaches in complement of the direct regulator characterisation. Additional experiments are necessary for a complete understanding of the regulatory networks involved in OHR. In this optic, more efforts should be made towards the development of efficient methods to characterise the regulators, like the ones proposed in Chapter 3 and 4.

An interesting observation, yet still hard to understand, arose from the pairwise comparison of the two CLOHPA-conditions. Indeed, most of the Rdh proteins were, in proportion, more expressed in H<sub>2</sub>/CLOHPA compared to La/CLOHPA. This phenomenon, which was less pronounced for the proteins encoded in cluster 6, can be visualised on the graph in Figure 6.8B. This suggests that the bacteria in H<sub>2</sub>/CLOHPA expressed more Rdh proteins and at a higher level than in La/CLOHPA with the exception of cluster 6. Why the same proteins were also appearing slightly more expressed in the H<sub>2</sub>/Fu conditions in relation to La/Fu is even less understood. However, it might highlight the more restricted energy metabolism of strain DCB-2 when using H<sub>2</sub> as electron donor (in comparison to organic sources of energy), similarly to obligate OHRB, such as *Dehalococcoides* spp., which express many *rdhA* genes at steady state level [7].

Besides the 22 Rdh proteins discussed above, the CLOHPA-specific selection comprised additional 17 more isolated hits. Most of them showed relatively low ratios and none of them had an obvious link to OHR. The relevance of their occurrence in the CLOHPA-specific adaptation

requires further investigation which goes beyond the scope of this thesis.

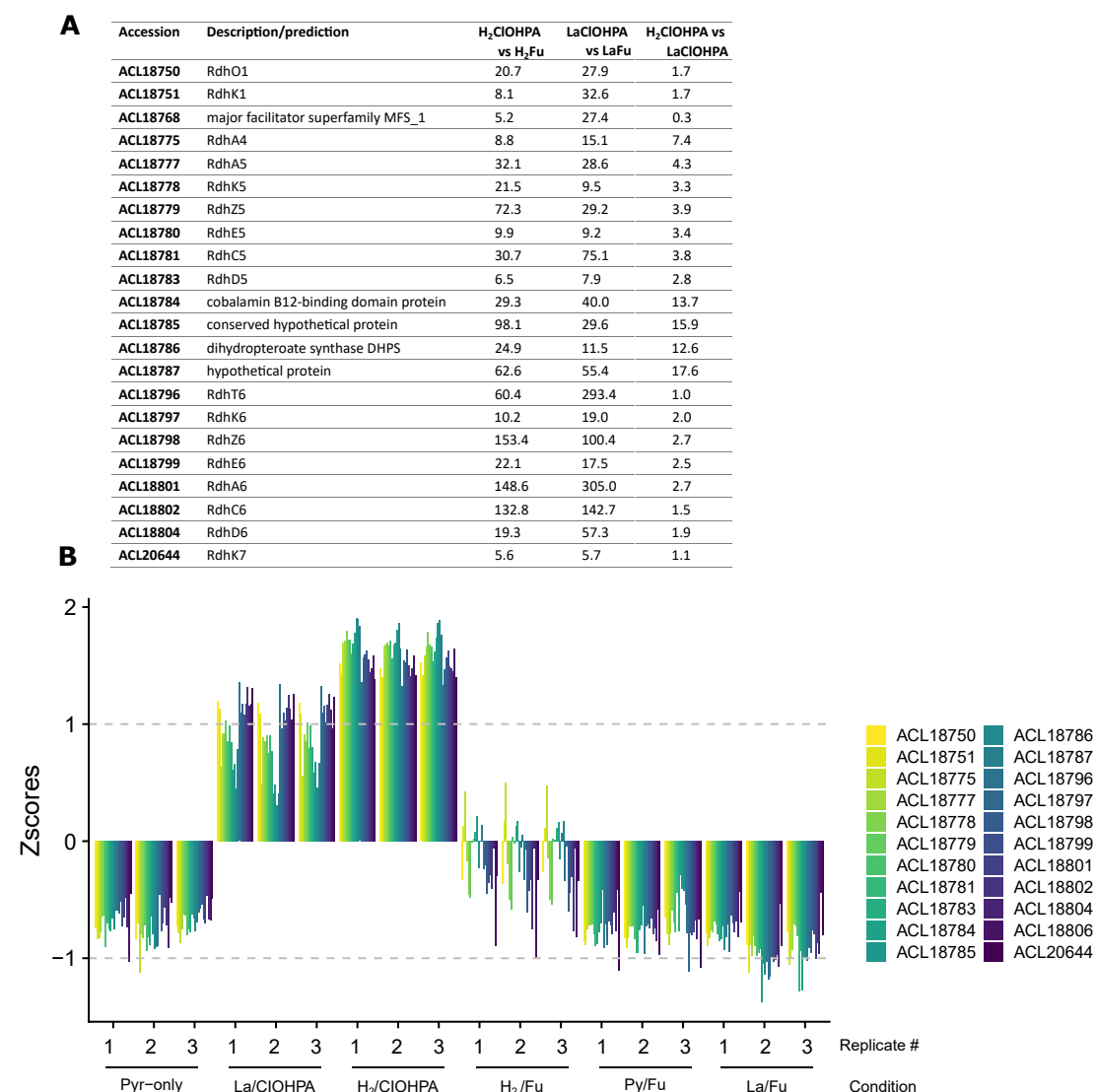


Figure 6.8 – Rdh proteins up-regulated in CIOHPA conditions. (A) Table with the Rdh proteins and candidates found in the CIOHPA growth conditions. Their respective fold-change are displayed for three different pairwise comparisons (H<sub>2</sub>/CIOHPA vs H<sub>2</sub>/Fu, La/CIOHPA vs La/Fu and H<sub>2</sub>/CIOHPA vs La/CIOHPA). (B) Relative abundance of the Rdh proteins over the 18 samples analysed in this study.

## 6.4 Conclusions

This chapter presents an extended proteomic study of *D. hafniense* strain DCB-2 that gave access to a robust comparison of nearly 3000 proteins which represent close to 60% of the theoretical proteome. This work is the first proteomic analysis performed on an OHRB which

resulted in such a level of proteome coverage across six different growth conditions [84].

Overall, the data highlighted many interesting features which, when put in perspective, offers a better understanding of *D. hafniense* proteome adaptations to the different metabolisms. Although it remains complicated to draw strong conclusions only based on the up-regulation of a protein in given conditions, several proteins were highlighted for their role in the central carbon or energy metabolisms, more specifically in fermentation conditions. Based on this dataset, dissimilatory sulfite reduction (possibly as alternative respiratory metabolism) was suggested to co-occur with fermentation, probably from sulfite produced abiotically in the culture medium. In addition, the presence of hydrogenases raised the question of a potential H<sub>2</sub> production during fermentation metabolism with pyruvate by strain DCB-2.

Concerning the respiratory conditions, this analysis shed light on important proteins involved in the corresponding metabolic pathways but also on proteins which are not directly linked to the applied substrates. For example, the complete Wood-Ljungdahl and the butyrate pathways were up-regulated in H<sub>2</sub> conditions where acetate is used as carbon source, in contrast to the conditions with lactate as electron and carbon source; revealing the importance of alternative carbon assimilation pathways when using acetate. The lactate conditions helped identifying the so far unknown lactate dehydrogenase and lactate transporter. Except for the proteins directly involved in lactate utilisation, no other lactate-specific pathways were identified, which probably reflects a certain similarity between the lactate- and the pyruvate-based metabolisms. The data concerning the fumarate conditions did not help in identifying the respiratory fumarate reductase. Nevertheless, an interesting candidate was shown to be expressed at steady-state level. Finally, when the dataset was used in order to identify proteins involved in the respiration of CIOHPA, proteins from 5 out of 7 *rdh* gene clusters were significantly up-regulated. The high coverage obtained in the proteomic analysis allowed us to get access to quantitative information on 17 Rdh proteins that were already identified on gene level, but also on other proteins encoded in the close vicinity of the *rdh* gene clusters which could be additional, yet unknown, Rdh proteins and need further investigation.

The data obtained with the analysis presented in this chapter provide an almost infinite amount of information which is sometimes difficult to put in perspective. Information on the substrate consumption and product formation should help proposing a better interpretation of the proteomic data. While first experiments were recently performed, these physiological data still need to be carefully analysed, and should certainly help addressing certain metabolic pathways that were identified in the present chapter (acetate assimilation, lactate production in fermentative conditions, ...). In addition, additional dedicated experiments will be performed in order to investigate other questions that were raised during the proteomic data analysis (for example, the production of butyrate and the dissimilatory sulfite reduction). Finally, the high quality dataset produced in this analysis represent a solid basis that can be further exploited for present and future research questions.

## 7 Potential role of complex I-like enzymes in the metabolism of organohalide respiration

### 7.1 Introduction

Respiratory complex I, or NADH:quinone oxidoreductase, is mostly known for playing a central role in aerobic respiration as it generates a proton motive force by coupling the transfer of two electrons from NADH to the quinone pool while pumping four protons across the cytoplasmic membrane. Most OHRB genomes encode a compact 11-subunit complex I-like enzyme lacking the NADH module and thus probably functioning with another electron source. The role of this complex I-like enzymes in the OHR metabolism was never really addressed. Yet the presence of such a potentially proton-pumping machinery in the proteomic data of representative strains from all OHRB phyla have raised the interest of the OHR research community. So far, only speculative information is available on the potential role of this enzyme in OHRB and in particular OHR metabolism. In some existing respiratory chain models, the complex I-like enzyme is actually proposed to be part of the OHR pathway [66, 117] (see more details in Chapter 5).

In this chapter, physiological and biochemical approaches were applied to two Firmicutes OHRB in order to start elucidating this research question. A long-established complex I inhibitor was used to see the importance of an active complex I-like enzyme for the growth of *D. hafniense* strains DCB-2 and TCE1, and *D. restrictus* strain PER-K23, and for electron transfer processes in *D. hafniense* strain TCE1. When possible, several growth conditions were tested in parallel in order to see the effect of complex I-like enzyme inhibition while promoting different types of metabolism. In addition, the purification and characterisation of the native complex I-like enzyme from both *D. hafniense* strain DCB-2 and *D. restrictus* PER-K23 was attempted. The results obtained are presented together with the limitations and difficulties encountered during the process. Finally, the quantitative proteomic dataset obtained for *D. hafniense* strain DCB-2 and presented in Chapter 6 was used to compare the

relative abundance of complex I-like enzyme in the different growth conditions.

## 7.2 Material and Methods

### 7.2.1 Bacterial cultures and growth medium composition

*D. hafniense* strain DCB-2 was cultivated in 200-mL anaerobic flasks following the media preparation described in section 6.2.1. Small variations from that recipe were made to adapt the growth medium to *D. hafniense* strain TCE1 (DSM 12704) or *Dehalobacter restrictus* strain PER-K23 (DSM 9455). In that cases, 0.1 g/L of peptone was used instead of yeast extract in the basal solution. In addition, tetrachloroethene (PCE) was used as electron acceptor for the latter strains but each one of them required a different preparation. PCE stock solutions were prepared in hexadecane and were added after inoculation on top of the aqueous phase resulting in a bi-phasic system to sustain a constant and slow diffusion of PCE to the aqueous phase and avoid any toxic effect on growth of PCE in the direct environment of cells. For *D. restrictus*, 4% v/v of 500 mM PCE in hexadecane was added to the culture flask while 1% v/v of 2 M was used for *D. hafniense* strain TCE1, giving in both cases a nominal concentration of 20 mM. H<sub>2</sub> was used for both strains as electron donor (see section 6.2.1) (H<sub>2</sub>/PCE). When applicable, 20 mM lactate was alternatively used as electron donor for TCE1 cultures (La/PCE). For a summary of the different culture conditions, see table 7.1.

Organism	<i>D. hafniense</i>							<i>D. restrictus</i>	
Strain	DCB-2						TCE1		PER-K23
Condition	Pyr-only	Py/Fu	La/Fu	La/CIOHPA	H <sub>2</sub> /Fu	H <sub>2</sub> /CIOHPA	H <sub>2</sub> /PCE	La/PCE	H <sub>2</sub> /PCE
Pyruvate	40 mM	20 mM	-	-	-	-	-	-	-
Fumarate	-	20 mM	20 mM	-	20 mM	-	-	-	-
Lactate	-	20 mM	20 mM	20 mM	-	-	-	20 mM	-
CIOHPA	-	-	-	10 mM	-	10 mM	-	-	-
H <sub>2</sub>	-	-	-	-	80%	80%	80%	-	80%
Acetate	-	-	-	-	2 mM	2 mM	2 mM	-	2 mM
PCE	-	-	-	-	-	-	20 mM	-	20 mM

Table 7.1 – Summary of the electron donor/acceptor combinations used for different OHRB strains. The short names describing the different growth conditions relevant in this study are indicated as well as the final concentration of the corresponding medium components.

### 7.2.2 Rotenone-induced growth inhibition

When applicable, the cultures were treated with 10 µM of rotenone (stock solution of 10 mM prepared in absolute ethanol, Merck, Zug, Switzerland). Absolute ethanol was added in control cultures. Growth was followed spectrophotometrically by measuring the absorbance of culture samples at 600 nm ( $A_{600}$ ) as proxy for the cell density.

### 7.2.3 Rotenone effect on PCE dechlorination

Cells from 400 mL of culture were collected anaerobically. The cultures flasks were opened in the anaerobic chamber, the aqueous phase was separated from the organic phase containing the PCE using a decanter and was transferred to collection bags to maintain the anaerobic conditions during centrifugation. Cells were pelleted (30 min at 7'000 x g and 4°C) and washed twice in 1 mL of washing buffer (100 mM Tris-HCl, 100 mM NaCl, pH = 7.5, N<sub>2</sub>-flushed) in the anaerobic chamber. The cells were finally resuspended in a small volume of washing buffer and kept anaerobic. Reactions were carried out individually in 10-mL anaerobic flasks containing 1 mL of cell suspension and either lactate at a final concentration of 20 mM or H<sub>2</sub> added in the gas phase to an over-pressure of 0.8 bar. The reactions were started by the addition of 20 µM of PCE. In addition, and when applicable, rotenone was added as inhibitor to a final concentration of 10 µM (unless notified otherwise). Alternatively, the same volume of ethanol was added as control.

### 7.2.4 Cell fractionation and membrane protein extraction

The protocol described here was used for the purification of protein complexes. Approximately 3 g of biomass (wet weight) from *D. hafniense* strain DCB-2 cultivated in Pyr-only conditions or 2.5 g of *D. restrictus* were first resuspended in 20 mL of sucrose buffer (10 mM Tris-HCl, 150 mM KCl, 250 mM sucrose, pH 7.4) supplemented with a few DNase I crystals and protease inhibitor cocktail. Cells were lysed through 3 rounds of French press at 1000 psi. Unbroken cells and residual cell debris were eliminated by centrifuging the lysed cell suspensions 15 min at 3'000 x g and 4°C. The cell-free extract was fractionated into soluble and membrane fractions by ultra-centrifugation for 1 h at 175'000 x g and 4°C (fixed-angle rotor 45Ti, Beckman-Coulter, Nyon, Switzerland). After careful elimination of the supernatant, the membrane pellet was resuspended in 6 mL of sucrose buffer by gentle stirring overnight on ice. The protein concentration of the membrane suspension was determined by the BCA assay and the sample was diluted with sucrose buffer to reach 5-10 mg/mL of proteins. Membrane proteins were solubilized for 20 min on ice using a final concentration of 1 % (m/v) n-dodecyl-beta-D-maltoside (DDM). The membrane extracts were cleared from insoluble material by centrifugation (85'000 x g, 7 min, 4°C) and further filtered at 0.45 µm.

**Alternative protocol.** When membrane extracts were prepared for Blue Native (BN)-PAGE analysis, lower amounts of cells (200-600 mg) were used as starting material. Sucrose buffer was replaced by Tris buffer (50 mM Tris-HCl, 50 mM NaCl, pH 7.5) and some modifications were brought to the downstream procedure. Ultra-centrifugation was performed 185'000 x g (fixed-angle rotor TLA-100.1, Beckman-Coulter). The totality of the membrane pellet was resuspended in 500 µL of MES buffer (50 mM MES-NaOH, 50 mM NaCl, 5 mM MgCl<sub>2</sub>, 0.005 % (w/v) maltose-neopentyl glycol (MNG), pH = 6.0) using an homogenizer and by 1 h of gentle stirring on ice. Extraction was performed with 1 % MNG with gentle stirring for 30 min at 4°C followed by an additional 10 min at RT. Extraction was followed by 15 min

ultra-centrifugation at 215'000 x g and 4°C (fixed-angle rotor Ti 50.2, Beckman-Coulter) and filtration through 22 µm filters to clear out the sample before sucrose gradient separation since the presence of particles in the sample disrupt the regularity of the gradient. At this point, the sample was either diluted 2-fold in MES buffer and concentrated again on a 300 kDa cut-off Amicon® column in order to reduce the concentration of detergent prior to load the samples on native PAGE gels. Samples for sucrose gradient separation were directly loaded without the dilution/concentration steps as described elsewhere [206].

### 7.2.5 Two-steps purification attempt of the complex I-like enzymes

Membrane extracts prepared as described in the first protocol were loaded on an anion-exchange column (HiTrap™QFF sepharose, 5-mL) attached to an AKTA™system using a peristaltic pump. The column was equilibrated in buffer A (20 mM Tris-HCl, 2 mM EDTA, 10% v/v ethylene glycol, 0.2 % w/v DDM, pH = 7.55) freshly supplemented with 0.02 % of asolectin (stock solution of 2 % prepared in 2 % CHAPS) and filtrated at 0.22 µM. The elution was carried out by a 150-mL gradient with buffer B (buffer A supplemented with 1 M NaCl) and 1-mL fractions were collected. During the elution process, the absorbance at 280 nm and 420 nm were constantly measured to help in the selection of fractions for downstream analysis. Fractions of interest were pooled together and concentrated using a 100-kDa cut-off Amicon® column. The protein complexes present in 500 µL of the concentrated eluate were then separated by size-exclusion chromatography (SEC, Superose 6 INCREASE 10/300 gl, 25-mL) equilibrated in buffer C (20 mM Tris-HCl, 200 mM NaCl, 10 % glycerol, 0.005 % w/v DDM, pH = 7.55). Fractions of 0.4 mL from an isocratic elution were collected and monitored at 280 nm and 420 nm.

### 7.2.6 LC-MS/MS analysis

For solid (gel bands) or liquid samples analysis by LC-MS/MS, see standard operating procedure described in [207]. For in-gel digestion, small variations to the original protocol were applied. After digestion, the peptides were resuspended in buffer A (2% acetonitrile and 0,1% trifluoroacetic acid in ddH<sub>2</sub>O).

## 7.3 Results

### 7.3.1 Comparison of the complex I-like enzymes from Firmicutes OHRB with the *E. coli* complex I

As mentioned earlier, unlike in *E. coli* (*nuo* operon), the complex I-like enzymes encoded in the genomes of *D. hafniense* and *D. restrictus* lack the *nuoEFG* homologous genes. Except for this difference, the organisation of the rest of the operon is conserved in the three species, as shown in Figure 7.1A. In contrast, the genes found in the direct vicinity of the *nuo* homologous operon are species-specific and of various predicted function. Interestingly, a protein annotated as desulfoferredoxin (Dehre\_0888) is encoded directly upstream of the operon and seems conserved among *D. restrictus* strains. Sequence alignments of Nuo protein subunits from *E. coli* strain K12 with the homologous proteins from *D. hafniense* strain DCB-2 and *D. restrictus* strain PER-K23 were performed. The sequence identity and similarity scores obtained from this analysis are indicated in Figure 7.1B. As expected, *D. hafniense* and *D. restrictus* Nuo homologous subunits are significantly closer to each other in comparison to those of *E. coli*. Nevertheless, given that *D. restrictus* and *D. hafniense* are phylogenetically close to each other, the level of sequence homology remains relatively low, suggesting a possible different origin of the complex in the two OHRB.

### 7.3.2 Physiological roles of complex I-like enzymes in OHRB

In this section different experiments performed in presence or absence of rotenone are presented. Rotenone is a known inhibitor for complex I activity as it binds to the quinone-binding pocket [208]. Here the inhibitor was used in order to see if an active complex I-like enzyme is needed for OHRB to grow. The effect of rotenone on growth was first tested with different OHRB strains that were cultivated in different culture conditions. Subsequently, the effect of rotenone on the electron transfer chain involved in OHR was investigated using concentrated cell suspensions of *D. hafniense* strain TCE1.

#### Rotenone-induced growth inhibition

The effect of complex I inhibition by rotenone on the growth of strain DCB-2 was tested for the six different growth conditions already applied in Chapter 6 (i.e. Pyr-only, Py/Fu, La/Fu, La/CIOHPA, H<sub>2</sub>/CIOHPA and H<sub>2</sub>/Fu) and following two different experimental approaches. In the first approach, rotenone was added to the culture at the time of inoculation and growth was followed by measuring cell density at 600 nm over time. As rotenone was prepared in ethanol, control experiments were included in which the same volume of ethanol alone was added to the cultures. In addition, another series of cultures was prepared without any treatment to serve as positive controls. Both treatment and control series were performed in triplicates.

The results of this experiment showed different responses of the bacteria to the presence of

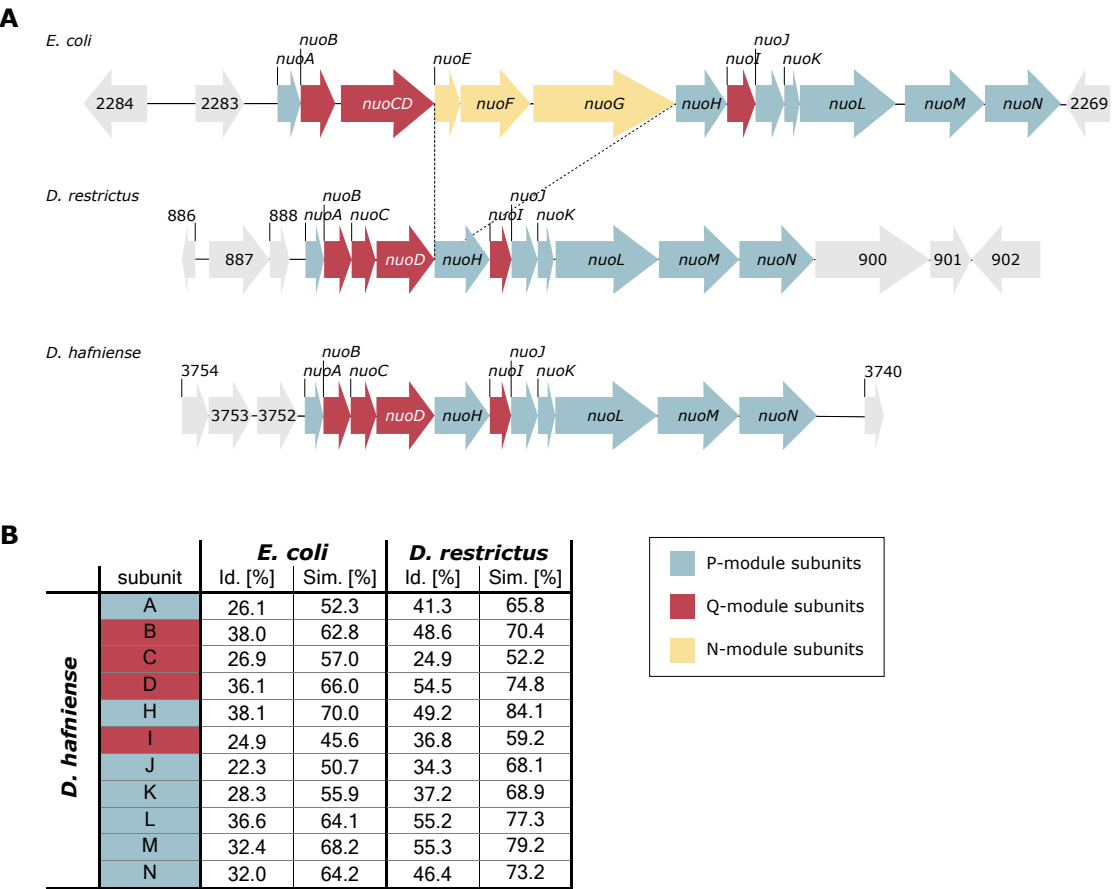


Figure 7.1 – Complex I-like enzymes in Firmicutes organohalide-respiring bacteria. **A** Schematic representation of the *nuo* operon in *E. coli* strain K12, and homologous gene clusters from *D. restrictus* strain PER-K23 and *D. hafniense* strain DCB-2. All the *nuo* subunits are indicated along with the genes directly flanking the operons. Numbers correspond to the respective gene loci (Eck\_, Dehre\_ or Dha\_). **B** Homologous sequences were aligned using LALIGN from the Swiss Institute of Bioinformatics (available here: [https://embnet.vital-it.ch/software/LALIGN\\_form.html](https://embnet.vital-it.ch/software/LALIGN_form.html)). Global alignment parameters (with default setups) were chosen as the size is relatively conserved within homologous proteins. An exception was made when comparing NuoC- or NuoD-homologous proteins from OHRB to the fused NuoCD of *E. coli* where the alignment was performed using local parameters to overcome the size discrepancy between the individual or fused gene products. The identity and similarity scores obtained are expressed in percent and are given for each homologous protein comparison between *D. hafniense* and *E. coli* or *D. restrictus*. The different subunits are coloured according to the structural module they belong to: proton-pumping module (P-module), quinone-reducing module (Q-module) or NADH-oxidising module (N-module).

rotenone depending on the growth conditions. The conditions Pyr-only, Py/Fu, La/CIOHPA and La/Fu (Figure 7.2 A to D, respectively) showed a clear growth defect. In this subset of conditions, Py/Fu and La/Fu were more extremely affected by the presence of rotenone as the bacteria did not grow at all while the two corresponding control series reached a maximum

$OD_{600}$  of almost 0.6 and 0.3, respectively. In the La/CIOHPA condition, the response was similar as in the two latter conditions except that a small rescue of growth was observed after 20 h of rotenone treatment which finally reached a cell density corresponding to 30% ( $OD_{600} = 0.055$ ) of the density observed in the control series ( $OD_{600} = 0.170$ ). In contrast, after approximately 30 h, the treated cultures in Pyr-only conditions started to grow and with a similar rate as in the control series. The treated cultures reached 70% of the cell density observed with the controls (+Rot :  $OD_{600} = 0.520$  and controls:  $OD_{600} = 0.730$ ). Interestingly, the two growth conditions using  $H_2$  as electron donor were not or only slightly affected by the addition of rotenone (Figure 7.2 E and F). Indeed, a minor reduction of the growth rate was observed for  $H_2$ /CIOHPA but the same cell density was reached in the treated cultures as in the controls ( $OD_{600} = 0.1$ ). No rotenone-induced growth defect was noticed with cultures in  $H_2$ /Fu conditions.

In addition to the growth inhibition experiments done with *D. hafniense*, rotenone was added similarly to *D. restrictus* strain PER-K23 in the only condition possible for this bacterium (i.e.  $H_2$ /PCE). The results are displayed in Figure 7.3 and showed no growth defect in presence of rotenone which is consistent with what was observed with strain DCB-2 in  $H_2$ -dependent growth conditions.

In all the above-mentioned experiments, the ethanol treated cells did not show any difference in their growth in comparison to the untreated positive controls. On this basis, it was decided to use the ethanol treatment only as control for further experiments.

As a complement of the first series of experiments on strain DCB-2, other cultures were prepared in the same conditions but instead of adding the rotenone directly in the growth medium at the time of inoculation, the same amount of rotenone (or ethanol for controls) was spiked in the cultures after growth has clearly started. The same six growth conditions as before were tested in the spiked series and the results are displayed in Figure 7.4. As already mentioned, strain DCB-2 does not grow at the same rate in each growth condition nor to the same extent. The timing for spiking the cultures (indicated by a red arrow in the graphs) was therefore adapted for each condition. The spikes were performed after the start of the exponential growth phase to clearly capture the phenotype of growth inhibition due to rotenone and avoid the confusion with the entry into the stationary phase.

The results obtained here reflected the ones obtained during the first experimental approach. After being treated with rotenone, Pyr-only, Py/Fu, La/Fu and La/CIOHPA cultures either completely stopped growing or show an important reduction in growth rate and extent in comparison to the positive controls (Figure 7.4 A to D). Again the growth of the cultures in  $H_2$ /CIOHPA and  $H_2$ /Fu conditions were not perturbed by the addition of rotenone (Figure 7.4 E and G).

These results suggested that an active complex I-like enzyme is required for strain DCB-2 in most of the tested growth conditions but the bacteria seem to be insensitive to the rotenone when the conditions involved  $H_2$  as electron donor.

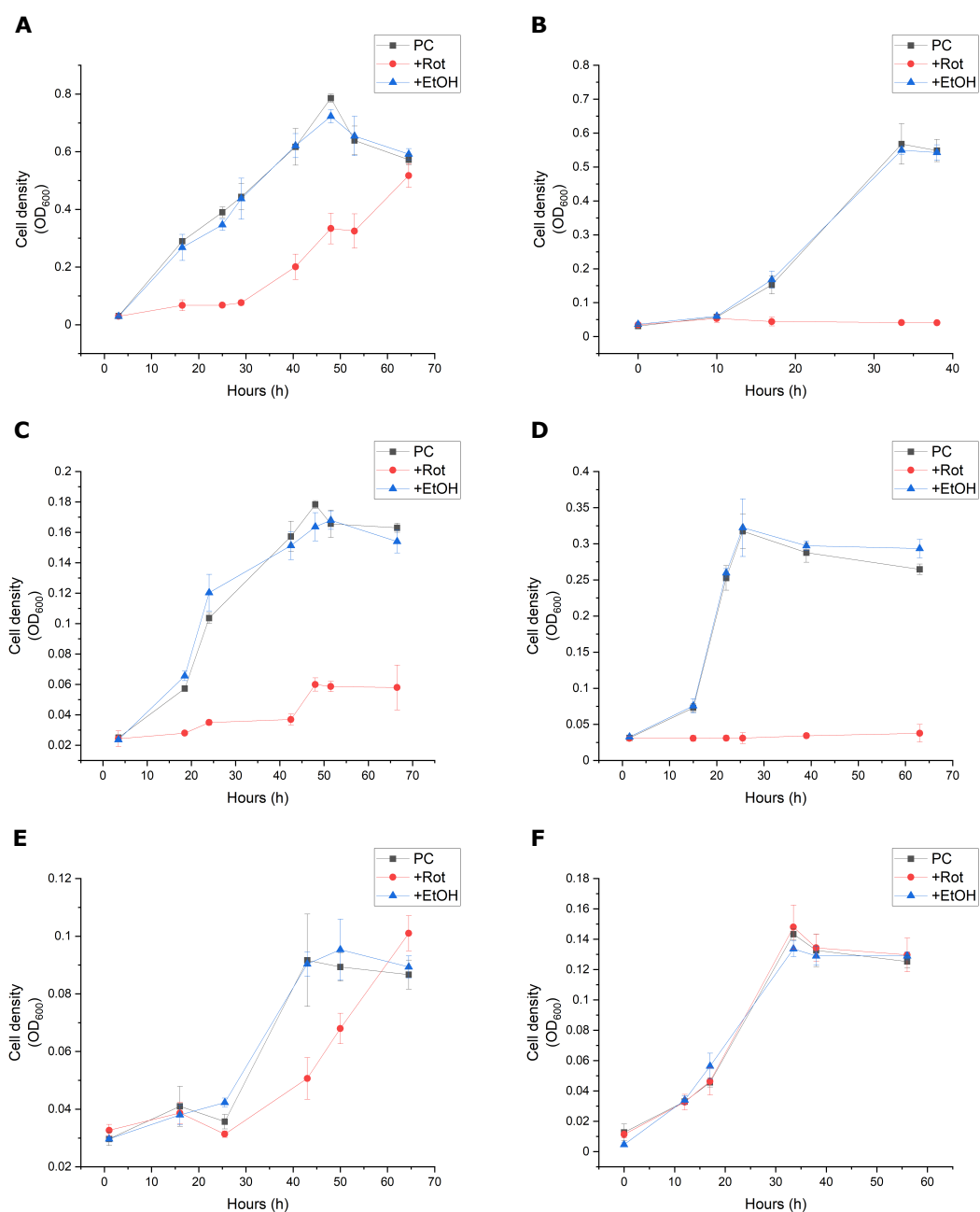


Figure 7.2 – Rotenone inhibition on the growth of strain DCB-2 in different conditions. Each graph displayed in this figure corresponds to a different condition: **A** Pyr-only, **B** Py/Fu, **C** La/CIOHPA, **D** La/Fu, **E** H<sub>2</sub>/CIOHPA and **F** H<sub>2</sub>/Fu. For each condition, different triplicates of cultures were either untreated (growth positive control, PC), treated with rotenone (+Rot) or with the same volume of ethanol (+EtOH). Growth was followed by measuring cell density at 600 nm over time.

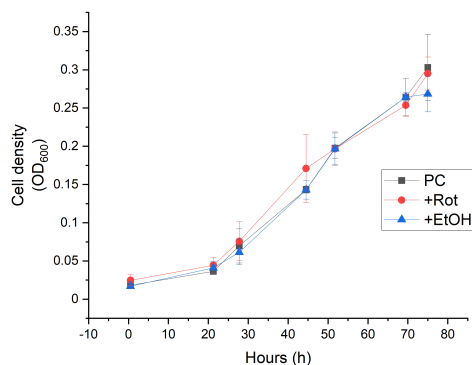


Figure 7.3 – Rotenone inhibition on the growth of *D. restrictus*. Three different culture series of strain PER-K23 growing in H<sub>2</sub>/PCE condition were performed, each in triplicates: untreated cells (PC), cells treated with rotenone (+Rot) or with the same volume of ethanol (+EtOH). Growth was followed by measuring cell density at 600 nm over time.

To further explore the "protection" effect of H<sub>2</sub> with respect to the rotenone inhibition, two other growth experiments were performed. For both of them La/Fu condition was used as it has shown the strongest growth inhibition phenotype in response to rotenone (Figures 7.2 and 7.4 D). In the first experiment, rotenone and hydrogen were included in the culture before inoculation while in the second, rotenone was first spiked in the already growing cultures and hydrogen was spiked after the growth of treated cultures started to decline. The results obtained here are displayed in the two graphs of Figure 7.5 and look relatively similar. In both experiments, the H<sub>2</sub> addition was apparently not sufficient to let the growth recover completely from the rotenone treatment. However, the rotenone-treated series supplemented with H<sub>2</sub> in the two experiments (be it from the inoculation time or from spike-time) showed a significant increase in the extent of growth.

As presented earlier through the different experimental approaches, the complex I-like enzyme seems to play an important role for OHRB in most of the tested metabolisms. However, these results are not informative on the role of the complex in these bacteria. To see if the complex is directly involved in the electron transfer chain involved in OHR, rotenone was applied to cell suspensions in the enzymatic assay described below.

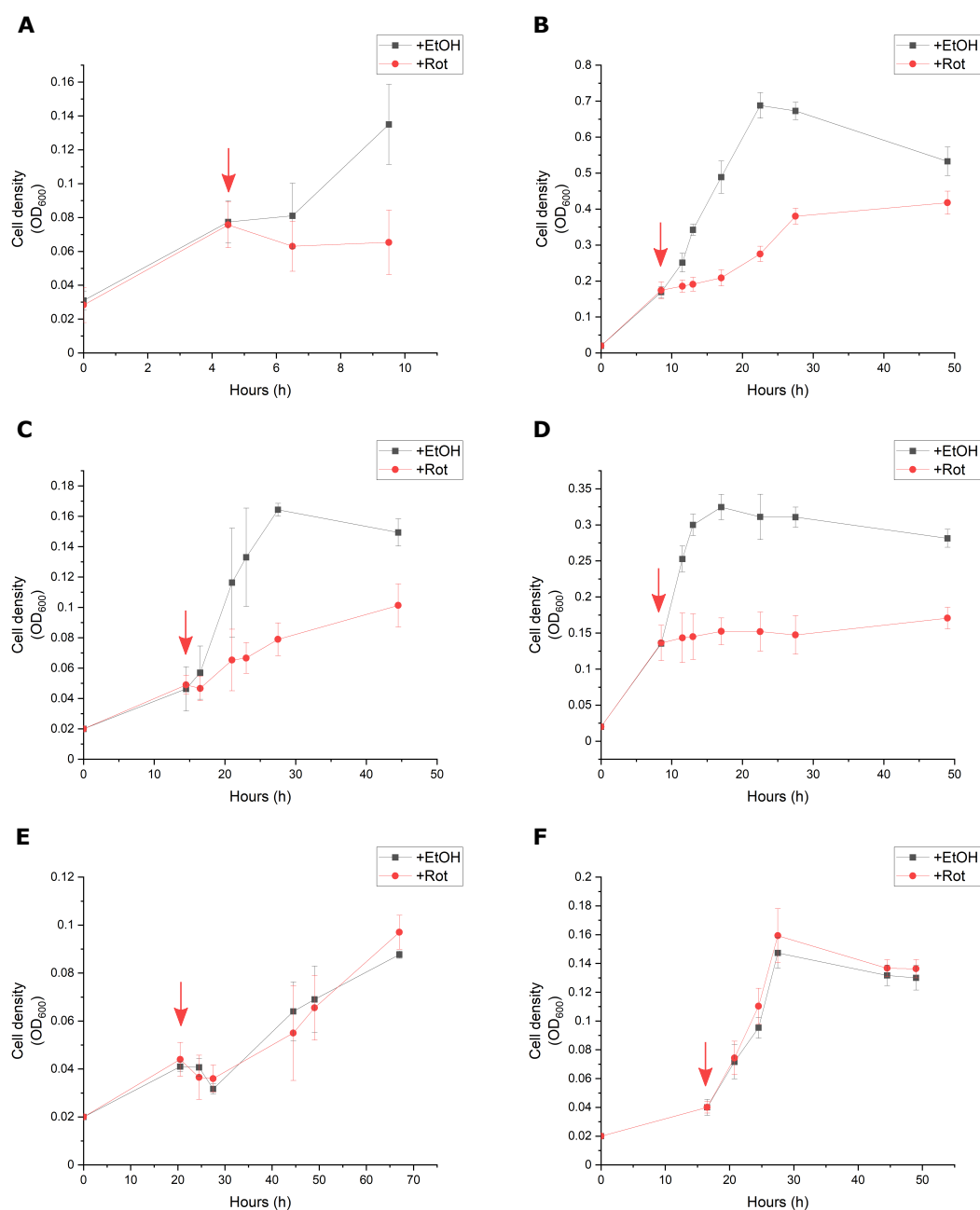


Figure 7.4 – Effect of rotenone inhibition on growing cultures of strain DCB-2. Each graph displayed in this figure corresponds to a different condition: **A** Pyr-only, **B** Py/Fu, **C** La/CIOHPA, **D** La/Fu, **E** H<sub>2</sub>/CIOHPA and **F** H<sub>2</sub>/Fu. Two different series of triplicates were performed for each growth condition and were spiked at the time-point indicated by the red arrow. In one series, the cultures were treated with rotenone (+Rot) and in the second, with the same volume of ethanol (+EtOH) to serve as control. Growth was followed by measuring cell density at 600 nm over time.

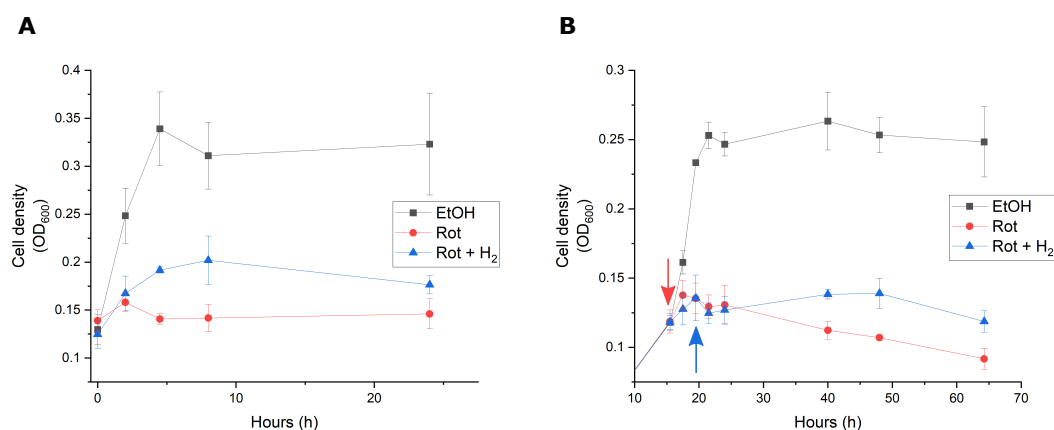


Figure 7.5 – H<sub>2</sub> protection effect on rotenone-treated La/Fu cultures. The growth of strain DCB-2 in La/Fu conditions measuring cell density at 600 nm over time. Two different experimental strategies were applied for which three different series of triplicates were performed. On the left panel **A** the cultures were supplemented with H<sub>2</sub> and/or rotenone right after inoculation. On the right panel **B** the rotenone (Rot) or ethanol (EtOH) was spiked in the cultures on the time-point indicated by the red arrow. When applicable (Rot+H<sub>2</sub>), an additional spike of 20 mL of H<sub>2</sub> was performed at the time-point indicated by the blue arrow.

### Rotenone effect on PCE dechlorination

In this experiment, the efficiency of the electron transfer in OHR metabolism from the electron donor (H<sub>2</sub> or lactate) to PCE was assessed by measuring the PCE dechlorination in presence/absence of rotenone. PCE dechlorination to *cis*-DCE can be easily monitored over time from the gas phase of cell suspensions by GC analysis, in contrast to ClOHPA which requires liquid samples for analysis. Therefore, *D. hafniense* strain TCE1 was used in this experiment.

Before running the actual enzymatic assay, the effect of rotenone on the growth of strain TCE1 was tested in La/PCE and H<sub>2</sub>/PCE conditions and the results are shown in Figure 7.6 A and B. As described for strain DCB-2 in La/ClOHPA conditions, strain TCE1 showed a strong growth defect in presence of rotenone from the start but growth slightly recovered after 30 h of incubation. Surprisingly, however, the growth of strain TCE1 growth was also partially perturbed by rotenone in H<sub>2</sub>/PCE conditions. This latter observation mitigates the conclusion on the protective effect of H<sub>2</sub> observed on other strains.

Globally, this experiment revealed that, although the growth of strain TCE1 was significantly affected by the addition of rotenone in La/PCE conditions and also partially perturbed in H<sub>2</sub>/PCE conditions, the inhibition of the complex I-like enzyme did not influence the PCE dechlorination activity of the bacteria using lactate nor H<sub>2</sub> as electron donor. This observation suggests that the complex I-like enzyme is not playing a direct role in the electron transport chain involved in OHR in strain TCE1 in both tested experimental conditions. Thus, the growth phenotype observed in most of the tested conditions is resulting from another essential

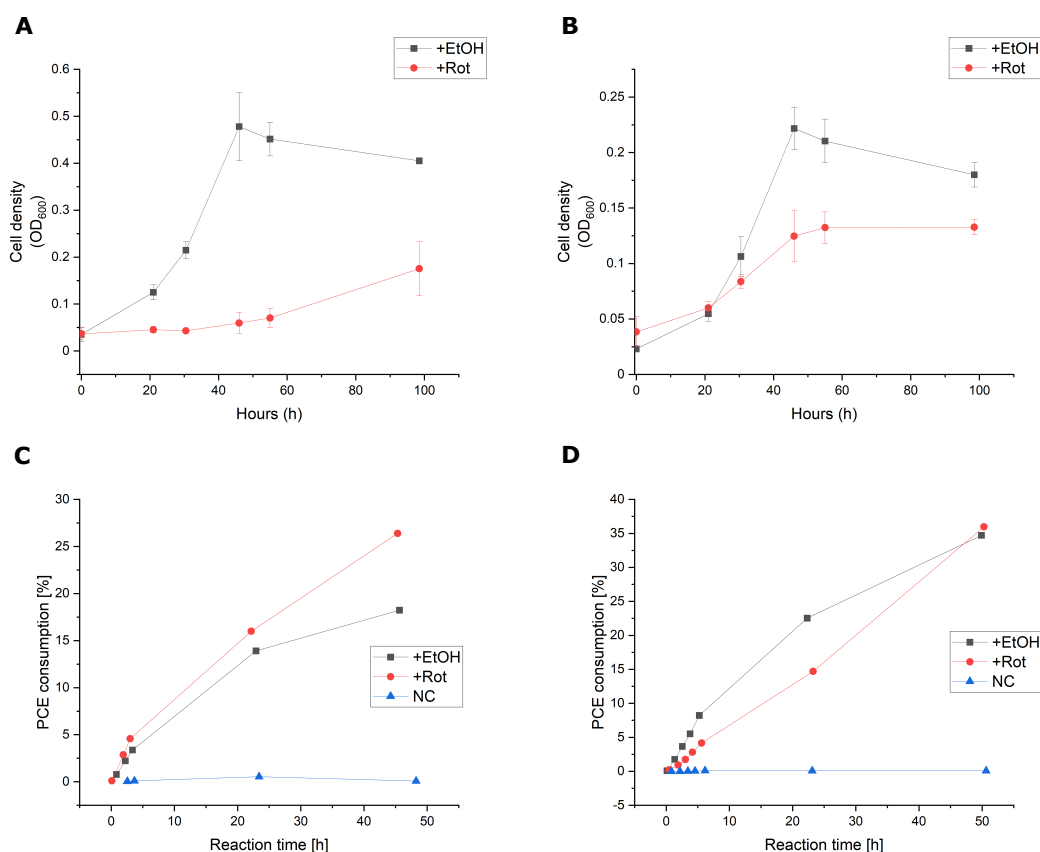


Figure 7.6 – Rotenone inhibition on the growth of *D. hafniense* strain TCE1 and on PCE dechlorination activity. The growth of strain TCE1 in either La/PCE or H<sub>2</sub>/PCE conditions were followed by measuring cell density at 600 nm over time and results are displayed in **A** and **B**, respectively. For each condition, two different series of cultures were performed in triplicates and were treated either with ethanol (+EtOH) or rotenone (+Rot). In parallel, PCE dechlorination activity of strain TCE1 was measured on concentrated cell suspension with cells cultivated in La/PCE and H<sub>2</sub>/PCE conditions and the results are displayed in **C** and **D**, respectively. The results are expressed as percentile of dechlorinated PCE to the initial concentration over time.

function provided by the complex I-like enzyme.

### 7.3.3 Tentative biochemical characterisation of native complex I-like enzymes in Firmicutes OHRB

#### Blue-Native PAGE analysis of membrane extracts from Firmicutes OHRB

The separation of the native complex I-like enzyme by BN-PAGE serves several purposes: (i) determine its subunits composition, (ii) identify potential additional subunits associated to the complex, and (iii) eventually get access to a possible in-gel enzymatic activity assay

to test potential electron donors. Altogether, it would enable to solve important structural and functional aspects of the complex. In this optic, membrane fractions were prepared from *D. hafniense* strain DCB-2 and *D. restrictus* strain PER-K23 cells. The proteins and protein complexes from the two preparations were solubilized from the membranes using 1% MNG and these extracts were either directly loaded on BN-PAGE, or, some part of *D. restrictus* membrane extracts were further separated on sucrose gradient and the fractions were analysed by native gel. The fractions showing a similar pattern on BN-PAGE were pooled together, concentrated and run as one sample in BN-PAGE. As shown in Figure 7.7A, several protein complexes were observed in the samples and were properly separated by the gradient. All the samples obtained with *D. restrictus* (i.e. membrane extracts and sucrose gradient fractions) showed a higher diversity of complexes in comparison to *D. hafniense* membrane extracts where only one clear band could be observed around the size of 66 kDa. All the *D. restrictus* samples did not show strong differences in their profile, fraction 16 was thus chosen as a representative sample for *D. restrictus* and three bands were excised from the BN-PAGE gel profile (the first at around 480 kDa, the second between 66 and 480 kDa and the third lower than 66 kDa. Note that for the last two bands, double bands were excised together). These three bands and the one of *D. hafniense* as indicated by red boxes on Figure 7.7A were analysed by LC-MS/MS, along with the initial membrane extracts from both OHRB strains.

The proteomic analysis identified most Nuo homologous subunits in the total membrane extracts obtained from *D. restrictus* and *D. hafniense* (7 and 10 out of 11 subunits, respectively), suggesting an efficient extraction procedure of the complex I-like enzymes in the condition applied (see Figure 7.7B). Furthermore, for *D. restrictus* five subunits were detected in the gel band excised around 480 kDa. However, Nuo homologous proteins were detected among many other proteins and the subunits of the complex were not dominating the dataset which probably indicates that complex I-like enzymes migrate around this area of the gel but do not account for the signal observed in the gel. Except for this band, none of the Nuo homologous subunits were detected in any other gel bands. Detecting the complex I-like enzyme in the higher molecular range of the gel represents an encouraging result as it suggests that the complex was not or only partially disrupted by the extraction and analysis procedure.

In parallel, an in-gel enzymatic assay using NADH as electron donor was performed using purified complex I from *E. coli* as positive control. No signal was obtained in the lanes corresponding to samples from *D. hafniense* or *D. restrictus* suggesting that none of the protein complexes were able to oxidise NADH (data not shown).

### Purification attempt of complex I-like enzymes

Enrichment of the complex I-like enzymes was attempted using a protocol that was developed for the purification of respiratory complex I [209]. This protocol was applied on concentrated membrane extracts from *D. hafniense* from cells cultivated in Pyr-only conditions and from *D. restrictus* cells obtained in H<sub>2</sub>/PCE conditions. For this procedure, the membrane extracts were first separated on an anion exchange column with monitoring the elution at both 280

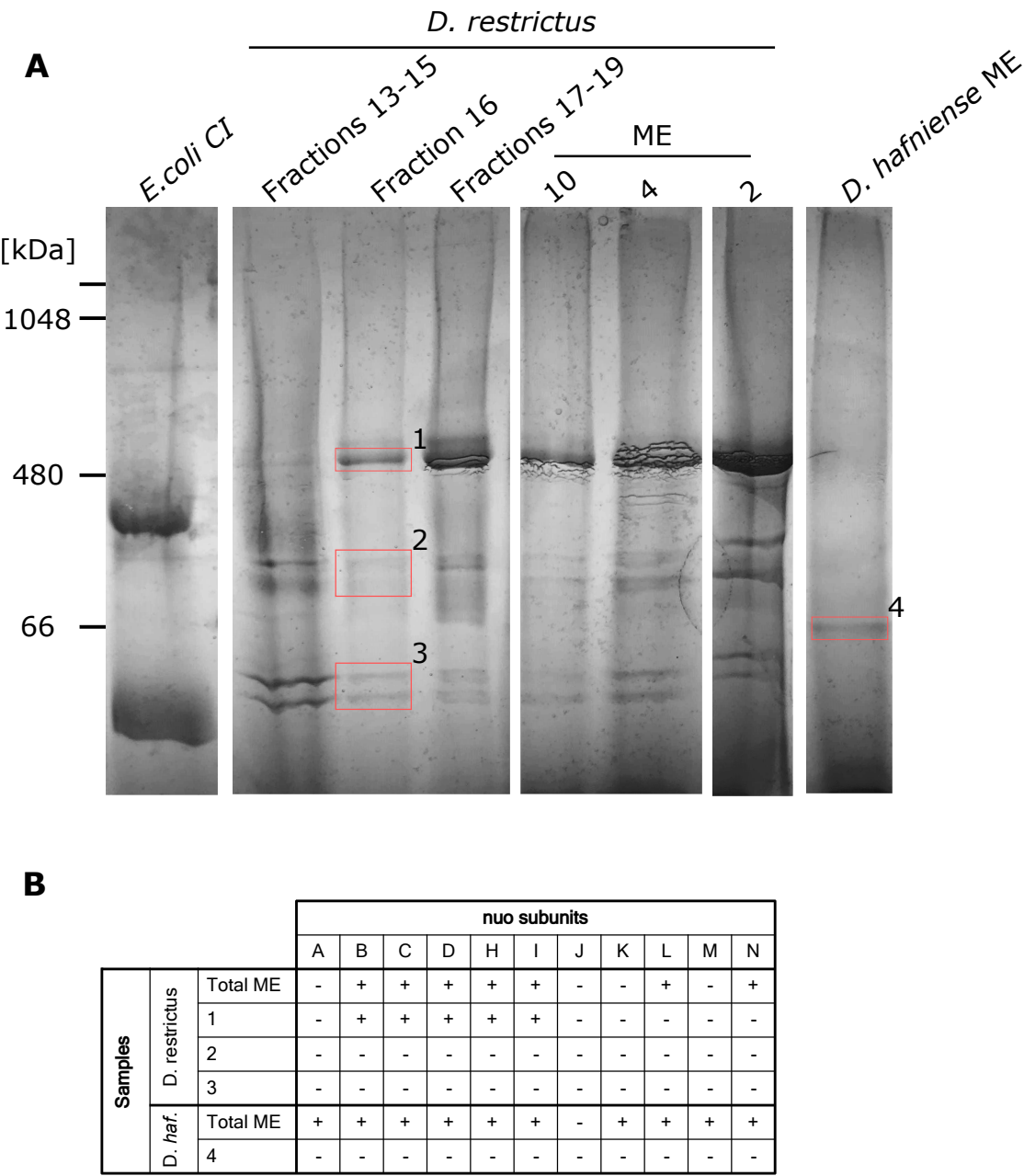


Figure 7.7 – Membrane extracts analysis on BN-PAGE. *E. coli* purified complex I was used as control. Four different samples of *D. restrictus* were loaded: pooled fractions 13-15, fraction 16 and pooled fractions 17-19 from sucrose gradient and total membrane extract (ME) in three different dilutions (10x, 4x or 2x). In addition, total membrane extracts (ME) from *D. hafniense* were also analysed. The red box indicates the bands that were excised for LC-MS/MS analysis. **B** Summary of LC-MS/MS analysis on total membrane extracts (ME) and gel bands from both *D. restrictus* and *D. hafniense* (*D. haf.*). The presence of each *nuo* homologous subunits is reported for each sample. When a subunit was detected in a given sample, it is indicated in the table by a (+) in contrast to (-).

nm and 420 nm wavelengths. As complex I-like enzymes are predicted to harbour multiple FeS clusters, only fractions absorbing at 420 nm were considered, pooled, concentrated, and further separated by size exclusion chromatography. The different chromatograms obtained and the summary of the different steps operated are shown in supplemental Figure D.1 for *D. hafniense* and Figure D.3 for *D. restrictus*.

The two anion exchange chromatograms from both OHRB looked similar with an important peak absorbing at 420 nm. When analysed on SEC, the pooled and concentrated fractions resulted in one main peak centred around 16 mL of elution volume (Note: for *D. hafniense* the peaks appeared more resolved when analysing an intermediate samples taken during the concentration process; the signals obtained for the final concentrated sample were saturated, see Figure D.1). Some minor peaks were also observed directly before and after the major peak but without clear separation. As shown in Figure 7.8, when the fractions obtained by SEC were analysed by SDS-PAGE, one main profile was obtained for each bacterial species which followed the peak signal. Although a relatively high diversity of proteins were present in the corresponding samples (given the fact that they resulted from two rounds of purification), the two signal profiles were dominated by a few proteins.

To obtain a better resolution of the sub-populations present in *D. hafniense* sample, another round of SEC was performed with three fractions obtained in the first round. As shown in Figure D.2, one single and symmetrical peak could be observed with fractions C7. Unfortunately, the attempt to monitor the complex I-like enzymes during the purification process using Western blot and an antibody raised against *E. coli* NuoI failed, which prevented identify the peak observed in C7 fraction as the purified complex. This was probably due to the poor affinity of the antibody toward the NuoI homologous protein from OHRB (Figure 7.1B) and the seemingly low abundance of the complex I-like enzymes in the membrane extracts. Alternatively, in order to see if the peak observed for sample C7 was enriched with complex I-like subunits, three fractions of the C7 elution (corresponding to before, during and after the peak) were analysed by LC-MS/MS. Unfortunately, although all but one of the eleven subunits (exc. NuoK) were detected in the C7 peak (with some being significantly enriched), the proteomic analysis also revealed the presence of a relatively high protein diversity suggesting that the peak observed in the chromatogram corresponded to several co-eluting complexes. This was confirmed by the analysis of the fractions on BN-PAGE. The gel, displayed in Figure D.4, showed that both *D. hafniense* and *D. restrictus* final purified samples still contained several protein complexes and no specific enrichment around 400 kDa (size where the complex I-like enzyme was expected).

To summarise, the relatively low abundance of the complex, the high diversity of the membrane proteome and the impossibility to efficiently visualise the complex impaired the biochemical characterisation of the complex I-like enzymes. No clear conclusion could be drawn about the complex composition nor the identification of potential partners.

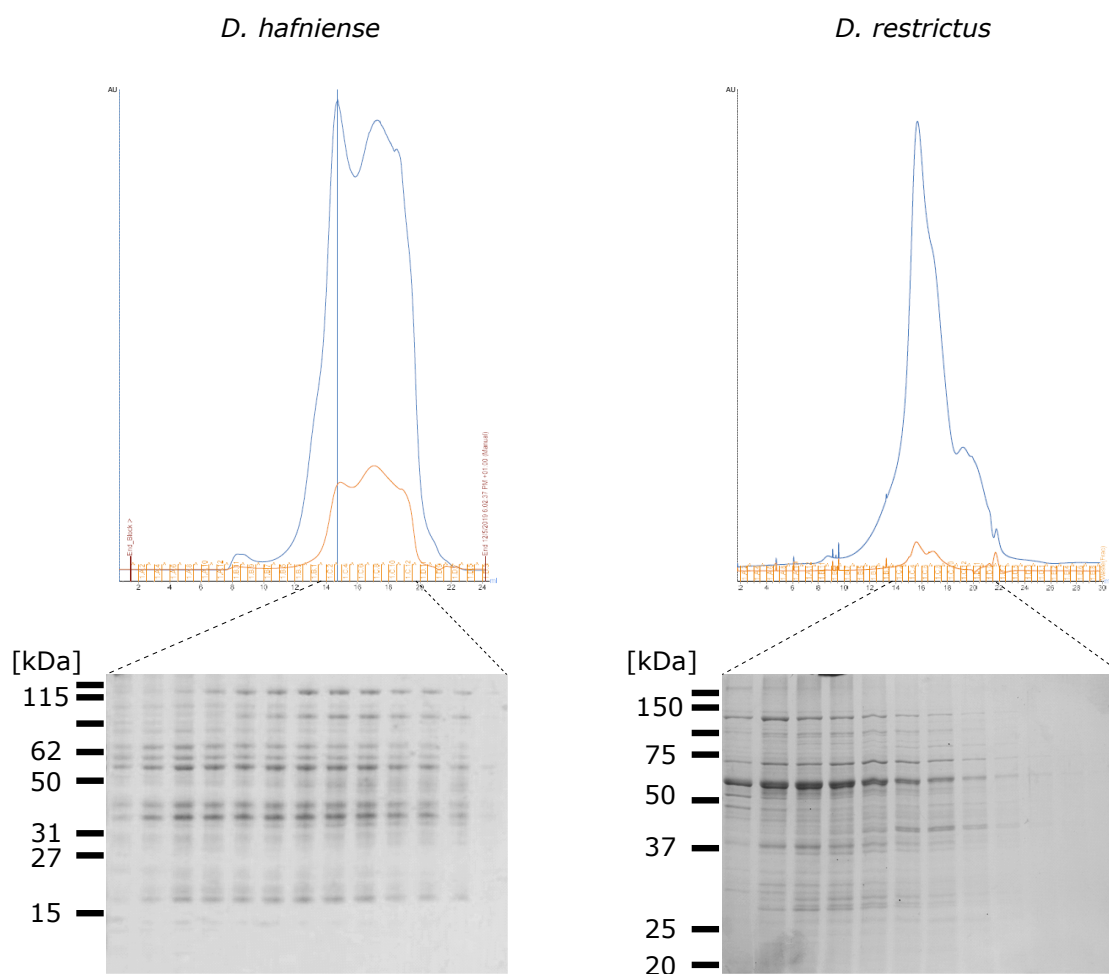


Figure 7.8 – Size exclusion chromatograms of membrane extracts initially purified by anion exchange chromatography. The chromatograms were obtained with the membrane extracts of *D. hafniense* and *D. restrictus* after a first round of purification on an anion exchange column. The fractions corresponding to the peak were analysed by SDS-PAGE and the gels corresponding to each purification procedure are displayed at the bottom of the figure.

### Complex I-like enzyme of strain DCB-2 in the comparative proteomic dataset

As presented in the previous chapter, the comparative proteomic approach gave access to a rather complete dataset of high quality, in which the relative quantification of ~ 60% of the total proteome of *D. hafniense* strain DCB-2 was obtained across 6 different growth conditions.

The proteomic dataset was used to compare the relative abundance of the Nuo homologous subunits across the tested conditions, since a predominance of the complex I-like enzyme in one or the other growth condition may reflect its relevance. In Figure 7.9 are shown the Z-scores of each detected subunit (i.e. all of 11 subunits except the NuoK homologue) in the 18 samples. The Z-scores indicate if the protein is more (positive scores) or less (negative scores) expressed across the different samples in comparison to its mean abundance. This result

indicated that the abundance of the complex I-like enzyme did not vary drastically among all the samples as all the scores were comprised between 1 to -1.5. Nevertheless, as shown on the graph, a clear trend can be drawn from the common behaviour of all the Nuo homologous subunits. Indeed, except for the two subunits which showed a more noisy pattern in z-scores (NuoA and NuoJ homologues) all the other components of the complex were relatively more expressed in the Pyr-only, Py/Fu and La/Fu growth conditions and less expressed in H<sub>2</sub>/ClOHPA condition. In La/ClOHPA and H<sub>2</sub>/Fu conditions the Z-scores were slightly below 0, indicating that in these conditions the level of expression of the complex was close to the average Z-score. Given the fact that the Nuo homologous subunits are following the same

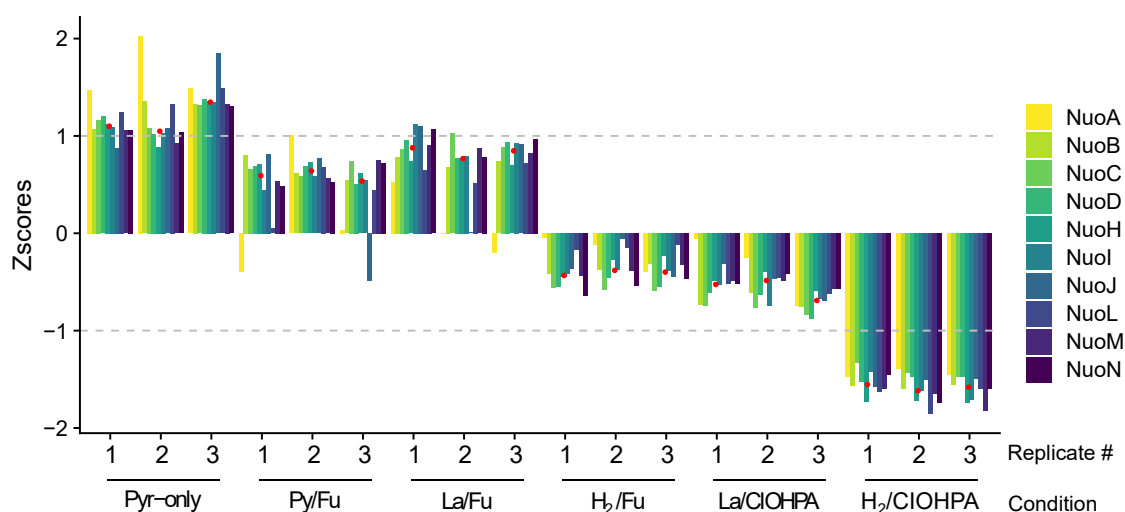


Figure 7.9 – Relative abundance of Nuo homologous subunits in *D. hafniense* strain DCB-2 across different growth conditions. The Z-scores for each of the Nuo homologous proteins were calculated across the different growth conditions. They are displayed together and follow a similar trend over the dataset with the median being indicated by the red dots.

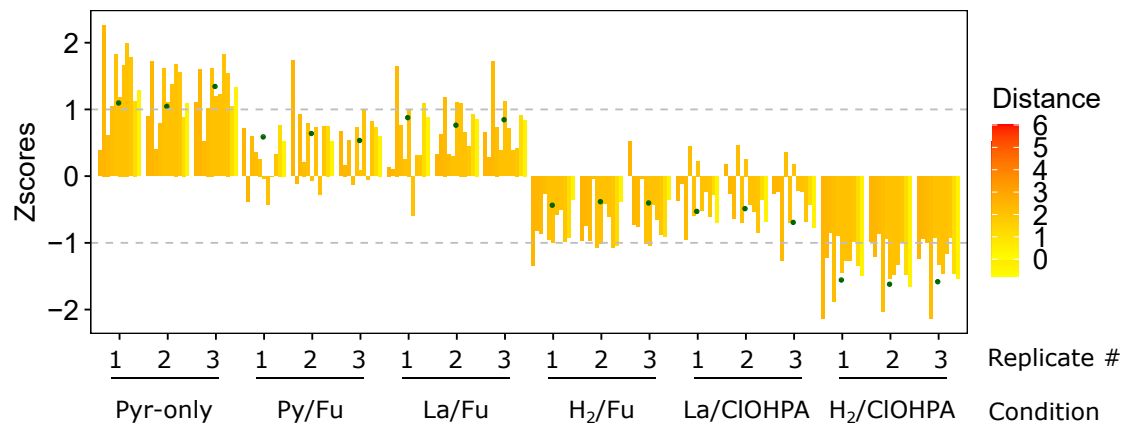
and clear trend, a median representing the typical behaviour of a Nuo protein was calculated (displayed in red on Figure 7.9). This median was used to identify potential complex I-like partner proteins, in particular those feeding electrons to it, by screening all the proteins part of the cluster 4 (cluster containing the Nuo homologous subunits) showing a similar behaviour. Proteins were then ranked according to the euclidean distance to the median; proteins with the smallest distance were considered as the most promising partners. To evaluate their possible involvement in electron transfer to the complex, their predicted function and sequences were analysed for redox activity (BlastP, [210]), FeS content (MetalPredator [211]) and association to the membrane (TMHMM sever 2.0).

The results are displayed in Figure 7.10A and showed that most of the Nuo homologous subunits are comprised within an euclidean distance between 0.42 and 0.91. One protein (ACL21016) was also comprised in this distance range. ACL21016 is a short protein of unknown function of 66 amino acids (7.7 kDa) with one predicted trans-membrane helix and no redox centre. Whether it plays a role in the activity of the complex I-like enzyme cannot be answered.

As already mentioned, NuoJ and NuoA homologues showed a more noisy behaviour than the rest of the detected subunits. They therefore appeared with a higher euclidean distance of 1.49 and 2.21, respectively. Nevertheless they globally appeared to follow the same pattern and do not influence the median which is not sensitive to outliers. An arbitrary maximum of euclidean distance was thus set at 2.5 for screening. The resulting profile of the selected proteins is shown in Figure 7.10A. when the threshold was increased to 3, only six more proteins were included, which did not seem to be of particular interest. Beyond that threshold, a more heterogeneous profile started to emerge (data not shown).

With a 2.5 threshold, eight candidates were identified from the cluster 4 for a total of nineteen proteins including the Nuo homologous proteins. As shown in 7.10B, the selected putative partner proteins are of various predicted functions. Four of them were predicted as cytoplasmic proteins equipped with an FeS cluster (ACL21807, ACL21808, ACL20267 and ACL20268) and two proteins as membrane-associated but without any redox centre (ACL21570 and ACL18521). Finally, two proteins that fell into the 2.5 threshold criterion were neither predicted as membrane nor redox proteins, and were therefore excluded from the selection.

When extending this search to the whole proteomic dataset (outside cluster 4), more proteins were obtained with an euclidean distance  $<2.5$ . Indeed, the clusters close to cluster 4 (mainly 2, 3 and 5) were also containing proteins following a similar trend to complex I-like enzyme. Those candidates have to be considered carefully since they were not initially clustered with the complex. Nevertheless, interesting proteins were identified in this selection and were also included in the list (7.10B). In particular, four of these proteins were predicted to contain a redox cofactor (ACL18123, ACL19056, ACL21378 and ACL20933).

**A****B**

Dist	Accession	Description/Prediction	Cluster	TMD	FeS
0.40	ACL21016	Hypothetical protein	4	1	-
0.42	ACL21758	<b>NuoM</b>	4	14	-
0.43	ACL21764	<b>NuoD</b>	4	-	-
0.45	ACL21762	<b>NuoI</b>	4	-	+
0.47	ACL21763	<b>NuoH</b>	4	8	-
0.48	ACL21757	<b>NuoN</b>	4	14	-
0.50	ACL21766	<b>NuoB</b>	4	-	+
0.68	ACL21765	<b>NuoC</b>	4	-	-
0.91	ACL21759	<b>NuoL</b>	4	16	-
1.20	ACL18123	Pyruvate:ferredoxin oxidoreductase	5	-	+
1.49	ACL21761	<b>NuoJ</b>	4	5	-
1.96	ACL21807	4Fe-4S ferredoxin	4	-	+
2.07	ACL20267	Rubredoxin Fe(Cys) <sub>4</sub>	4	-	+
2.09	ACL21570	Hypothetical protein, type-II secretion system	4	1	-
2.10	ACL19056	Fe-Fe hydrogenase	5	-	+
2.10	ACL21808	CODH	4	-	+
2.21	ACL21767	<b>NuoA</b>	4	3	-
2.21	ACL21378	4Fe-4S ferredoxin	3	-	+
2.35	ACL20933	Lactate utilization protein	2	-	+
2.45	ACL20268	Rubrerithrin	4	-	+
2.47	ACL18521	ABC transporter permease	4	6	-

Figure 7.10 – Identification of potential Nuo partners. (A) Z-scores of possible Nuo protein partners selected from the overall proteins of the dataset using an euclidean distance smaller than 2.5 from the Nuo homologues median (as represented by the green dots). (B) Summary table with the different Nuo homologues and interesting candidates calculated with an euclidean distance <2.5. For each protein, the euclidean distance (Dist.) is indicated as well as the number of trans-membrane domain (TMD) predicted for each entry and whether or not it is predicted to display an FeS cluster. The candidates which were not initially found in cluster 4 (like the Nuo homologues) are highlighted in grey.

## 7.4 Discussion

### 7.4.1 Identification of redox partners for complex I-like enzyme in OHR

Respiratory complex I is especially known for its function as a central player in *E. coli* aerobic respiration, although it is also predicted to be involved in various bacterial physiological pathways [144]. Based on the similar genetic organisation of the region encoding the *nuo* homologous genes, it is likely that complex I-like enzymes adopt a similar conformation in OHRB as in other bacteria and thus might play a similar role. Indeed, except for the NADH-oxidising module, which may be replaced by an alternative electron donor, the important features for electrons channelling to the quinone reduction site and the formation of the proton pumping module are conserved.

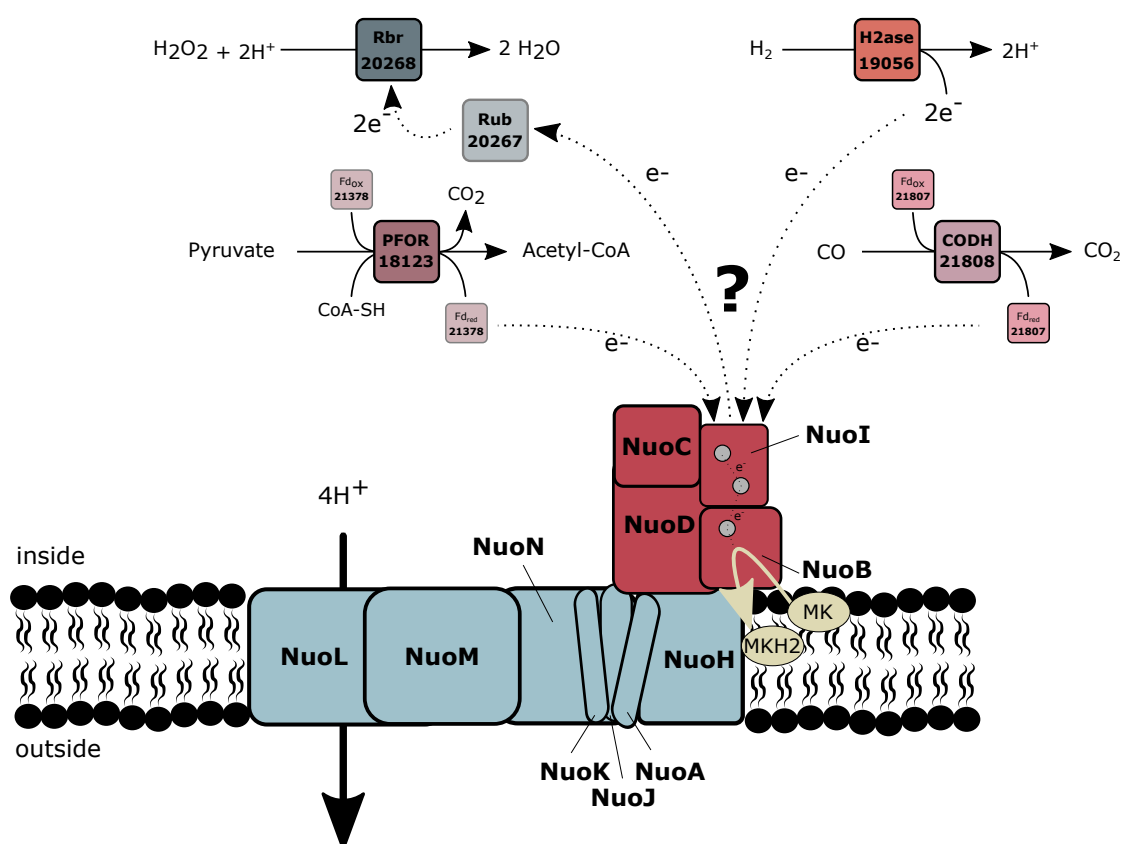


Figure 7.11 – Potential redox partners for *D. hafniense* complex I-like enzyme. This model proposes four putative protein partners to the complex I-like enzyme that were shown to follow the similar trend as the Nuo homologous proteins in the proteomic analysis.

In this work, the proteomic data already discussed in Chapter 6 were used to identify potential electron donor partners for the complex-I enzyme of *D. hafniense* strain DCB-2. The proteins initially annotated as NuoE (ACL19452) or NuoF (ACL19453) and sometimes considered as part of the complex I-like enzyme in *D. hafniense*, were found within the proteins that are the most distant from the median calculated for Nuo homologous proteins,

thus strongly suggesting that they are not likely partners of the complex. However, among the proteins which showed an expression pattern very close to the Nuo homologues across the 18 samples analysed by proteomic, four putative redox partners were identified. As shown in Figure 7.11, these proteins were either constituted by one single protein (ACL19056) or by a set of two proteins which are likely to interact with each others (ACL21807-8, ACL20267-8 and ACL18123/ACL21378).

ACL19056 corresponds to an Fe-only hydrogenase. Interestingly, the NuoG subunit composing the N-module in *E. coli* was shown to display similarities with molybdopterin-containing proteins, like Fe-only hydrogenases [141]. Of course, the fact alone that ACL19056 follows a similar trend as other Nuo homologues does not represent a proof in itself that it directly interacts with the complex. Nevertheless, it represents an interesting candidate and we could imagine the involvement of other proteins to stabilise the interaction between the hydrogenase and the NuoI subunit which displays the first redox cofactor of the Q-module that is usually accepting electrons from the N-module. Another proposition is that the complex-I like enzyme may receive electrons from a reduced ferredoxin, as already described in cyanobacteria (for details, see Chapter 5) [153, 154]. Indeed both ACL21378 and ACL21807 correspond to ferredoxins which could be reduced by the pyruvate:ferredoxin oxidoreductase (PFOR, ACL18173) or the carbon monoxide dehydrogenase (CODH, ACL21808), respectively. Unlike for ACL21807 which is encoded directly upstream of CODH, ACL21378 is not found directly adjacent to PFOR. However, no other ferredoxin encoding gene is found in the direct vicinity of PFOR in the genome. The cyanobacterial complex-I like enzyme display additional subunits for an efficient delivery of the electrons from the ferredoxin [154]. The presence of additional subunits playing a similar role in OHRB is not excluded but could only be verified by a thorough biochemical characterisation of the enzyme. Pyruvate is present in many of the tested growth conditions and thus the activity of PFOR is easy to imagine. In contrast, carbon monoxide was not directly added to the culture medium. Thus, the oxidation of CO to CO<sub>2</sub> coupled with the reduction of a ferredoxin may seem rather unlikely to occur. Yet, ACL21808 and its associated ferredoxin show a high sequence identity with CODH I (CooSI) and CooF3 (70% and 50 %, respectively) of *Carboxydotherrmus hydrogenoformans*. In the latter bacterium, the electrons supplied by CODH I are used for the translocation of protons through the cytoplasmic membrane by a complex I-related hydrogenase [212]. CODH activity thus remains an interesting candidate for feeding electrons to the complex I-like enzyme in strain DCB-2. The last potential redox partner for the complex I-like enzyme highlighted in the proteomic data is a set of proteins composed by a rubrerythrin and a rubredoxin (ACL20267-8). In *Desulfovibrio vulgaris*, homologous proteins are found to be involved in a protective mechanism against oxidative stress [213]. Rubredoxin is receiving electrons from an undefined source which, in our case, could be the complex I-like enzyme, and transfers them on the rubrerythrin moiety which performs the reduction of H<sub>2</sub>O<sub>2</sub> to H<sub>2</sub>O [213]. This implies that the complex I-like enzyme works in the reverse direction using the proton gradient to transfer electrons from reduced menaquinones to the rubredoxin. Consuming energy for the protection against oxidative stress might be a necessity in strict anaerobes like OHRB. Furthermore, this proposition somehow recalls

the mechanism of ROS generation by mitochondrial complex I where, in contrast to our hypothesis, electrons delivered by the complex I represents a source of oxidative stress.

#### 7.4.2 Physiological relevance of complex I-like enzymes in Firmicutes OHRB

Globally, the results obtained from the growth inhibition experiments with rotenone highlighted what seems like a fundamental difference between conditions involving  $H_2$  as electron donor and the other conditions. Indeed, bacterial growth was totally inhibited by the addition of rotenone when *D. hafniense* strains DCB-2 and TCE1 were cultivated in respiratory or fermentative conditions with lactate or pyruvate as primary carbon and energy source. In contrast, the inhibition of the complex I-like enzyme had no effect on the growth of *D. hafniense* strain DCB-2 in  $H_2$ /CLOHPA or  $H_2$ /Fu, neither on *D. restrictus* growing in  $H_2$ /PCE. Although *D. hafniense* strain TCE1 in  $H_2$ /PCE showed a slight growth defect, it seems that OHRB mostly rely on the use of complex I-like enzymes for growth except when using  $H_2$  as electron donor.

From that observation, we propose that the complex I-like enzyme might be necessary for the electrons to enter the respiratory chain only when the source of electrons is located in the cytoplasm. In contrast, in the respiration where the uptake hydrogenase is the electron donating moiety, the electron donor ( $H_2$ ) delivers them directly to the membrane and thus there is no need of a cytoplasmic entry point. This proposition is illustrated for *D. hafniense* strain DCB-2 in the Figure 7.12 and consists of two different models: a complex I-like enzyme-dependent model 7.12A or -independent one 7.12B. In the first case, we propose that the electrons are delivered to the complex I-like enzyme by a reduced ferredoxin resulting from PFOR activity, as described in the previous section. The interactions of the complex I-like enzyme with several electron donors, however, is not excluded.

These two models globally conciliate most of the results obtained with rotenone for growth experiments in respiration conditions (a small reserve is made for *D. hafniense* TCE1 in  $H_2$ /PCE conditions which seemed more slightly sensitive to rotenone). However, rotenone had no effect on the PCE dechlorination by cell suspensions of *D. hafniense* strain TCE1, no matter if lactate or  $H_2$  was used as electron donor. This suggested that an active complex I-like enzyme was not directly involved in the electron transfer chain between  $H_2$  or lactate and PCE, a statement that goes against the proposed models. It has to be kept in mind that in the experiment using strain TCE1 cell suspension, the efficiency of the electron transfer between the primary electron donor and the terminal acceptor is decoupled from growth. The effective PCE dechlorination by cell suspensions does not reflect the need of the generation of the proton motive force, as it has to in the case of growing cells.

Another aspect which remains to be discussed is the special case of fermentation. Indeed, when tested, *D. hafniense* strain DCB-2 growing in Pyr-only showed a clear growth defect in response to the rotenone treatment. But, interestingly, a resurgence of growth was observed after 30 h of incubation. This phenomenon remains difficult to explain but could be linked

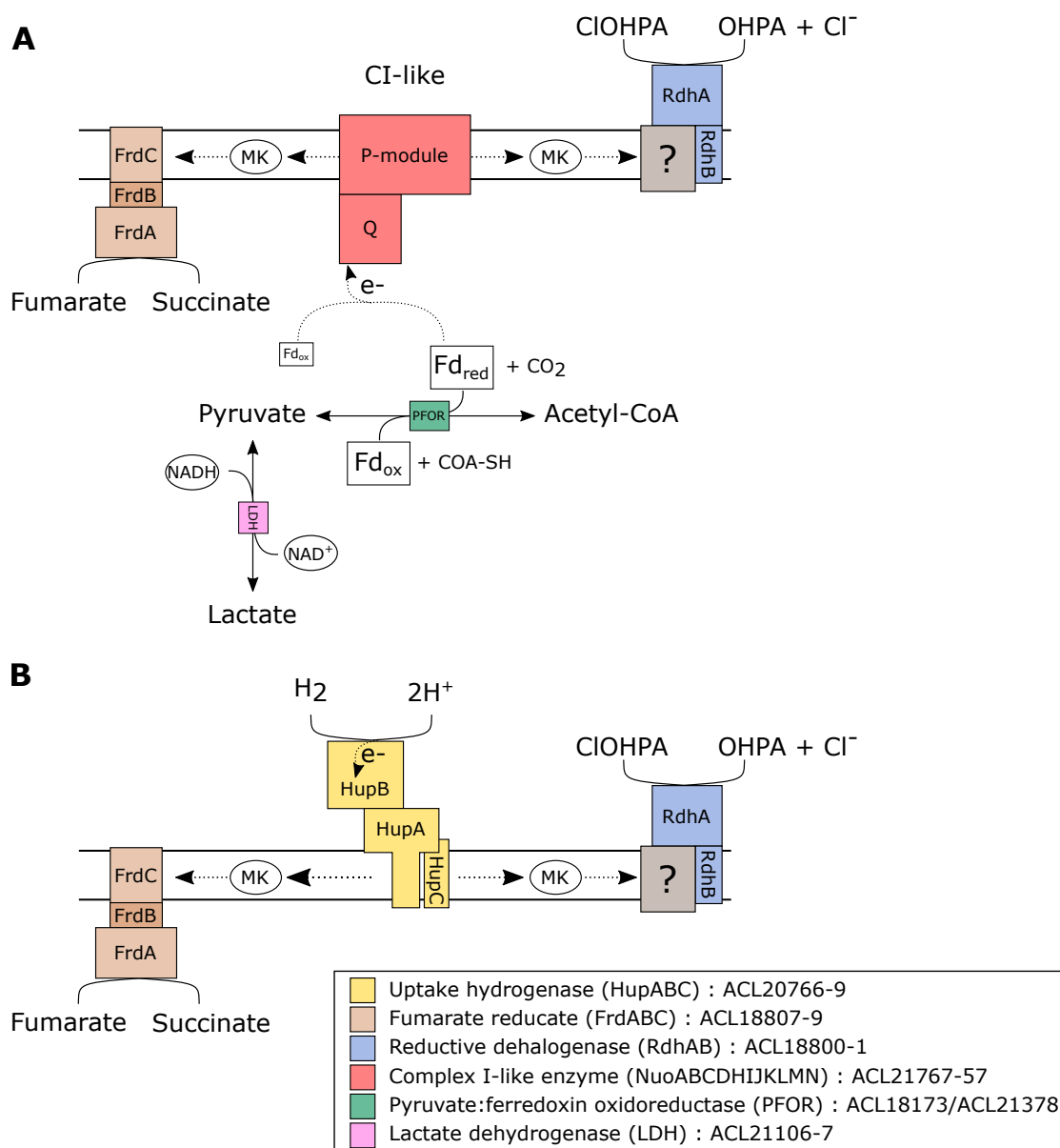


Figure 7.12 – Hypothetical models for the respiratory chain of *D. hafniense* strain DCB-2. Two different models are proposed for fumarate or ClOHPA respiration. (A) Complex I-like dependent model when lactate or pyruvate are used as electron donors and (B) an independent model when molecular hydrogen is the electron donor.

with a potential production of  $H_2$  in fermentative conditions. Indeed  $H_2$  is a common product of fermentation and a slight  $H_2$  "protective" effect was observed in cultures treated with rotenone. Still, this does not explain why the complex I-like enzyme is necessary in absence of any terminal electron acceptor. One explanation could be that the electrons are still transferred to an electron acceptor which is so far not considered as such. This proposition would also account for the protective effect of  $H_2$  which, according to our model, enables to bypass the

complex I-like enzyme.

## 7.5 Conclusions

In this chapter, the first study of the physiological implication of complex I-like enzymes in OHRB is presented. The different experimental approaches revealed that the metabolism of Firmicutes OHRB substantially varies in their dependence to this enzyme. An active complex I-like enzyme seem to be necessary unless the bacteria use molecular hydrogen as electron donor. This observation gave rise to two different respiratory models for *D. hafniense* strain DCB-2; one dependent and one independent of the complex I-like enzyme. The reason why an active complex I-like enzyme is needed for growth remains unclear and requires further investigation. In parallel, redox protein partners, potentially replacing the missing NADH-oxidising module, were proposed based on their co-occurrence in the proteomic data with the subunits of the complex I-like enzyme. In total, four different protein partners were proposed: two ferredoxins potentially reduced through the activity of the pyruvate:ferredoxin oxidoreductase or one specific carbon monoxide dehydrogenase, an Fe-only hydrogenase, and a couple of rubredoxin/rubrerhythrin proteins. The possible implication of these different partners in the activity of the complex I-like enzyme needs to be confirmed by further experiments.

## 8 Concluding remarks and perspectives

### **Towards the complete characterisation of new RdhK regulators and a global understanding of the OHR regulatory network in the Firmicutes**

The very high diversity of *rdh* gene clusters found in OHRB (in particular obligate OHRB) suggests that one single organism has the potential to use a broad range of organohalides in their metabolism. However, only a few *rdh* gene clusters were assigned to the respiration of specific organohalide compounds and the dehalogenation potential remains largely unexplored. This can be explained by the numerous difficulties encountered when working with OHRB as model organisms (genetically intractable anaerobes with low growth yield) and specifically with the reductive dehalogenases that are oxygen-sensitive complex redox enzymes.

The transcriptional regulators, such as members of the RdhK family, represent an interesting alternative for the study of the functional diversity of *rdh* gene clusters in the Firmicutes as they activate the transcription of the genes responsible for the dehalogenation of the organohalide compounds that are recognised as effector molecule. As developed in Chapters 3 and 4 of the thesis, the RdhK hybrid approach represents a promising strategy to characterise the activity of new RdhK regulators. In this work, the methodology was shown to enable the screening of the DNA motifs recognised by a given RdhK protein without prior knowledge of its effector. The results presented in Chapter 3 suggest the existence of possible cross-talks of a specific RdhK protein towards different *rdh* gene clusters. From these successful results, efforts need to be made to adapt available high-throughput methodologies for the screening of DNA target sites to the RdhK hybrid system in order to identify all the potential binding sites on the genomic DNA for a given RdhK regulator. Thus, a genome-wide approach is the only way to get a complete understanding of the RdhK-based regulatory networks. In that optic, a method relying on co-immunoprecipitation of fragmented genomic DNA together with the RdhK (hybrid) protein of interest followed by next-generation sequencing represents a very promising approach. Encouraging preliminary results were obtained during the first application of this method. Further experiments using quantitative PCR will be performed to check the selective enrichment of expected DNA targets in the elution fractions.

Ultimately, the sequencing of the captured DNA fragments should lead to the identification of all the DNA motifs targeted by a given RdhK regulator. Sequence analysis of these motifs should help defining the minimal conserved sequence (or dehalobox) recognised by this particular regulator. Moreover, the method could be easily applied to any RdhK hybrid protein made of the effector-binding domain from RdhK6 without the need for the development of additional antibodies. Another advantage is that it does not require a pure RdhK protein which circumvents the tedious and challenging work of protein purification. However, in order to have a satisfying signal to noise ratio, cell extracts with a significantly enriched fraction of soluble and active RdhK (or RdhK hybrid) proteins of interest should be used. This implies the over-expression of the proteins in a genetically tractable host organism; a process which often resulted in the formation of inclusion bodies in *E. coli*. To overcome this solubility issue, alternative expression systems and/or host organisms could be tested.

The identification of the target DNA sequence is the first step in the characterisation of new RdhK proteins. Once a DNA motif was assigned to the DNA-binding domain of a yet uncharacterised RdhK regulator fused to the effector-binding domain of *D. hafniense* RdhK6, the screening of the effector molecule is relatively easy as it can be performed directly with the native RdhK protein in combination with the newly identified DNA sequence by applying simple methods such as EMSA. At that stage, the chemical nature of the organohalides could cause some technical issues as they are often hydrophobic molecules and sometimes also acidic, therefore they might not be available for interaction with the RdhK proteins or interfere with the downstream method, respectively. Ultimately, the assignment of an organohalide as effector molecule for a specific RdhK protein targeting one or several *rdh* gene clusters constitutes a strong indication for the use of this particular organohalide as substrate by the cognate reductive dehalogenase. This can be later verified experimentally for example by spiking the given organohalide in a growing culture and measure the induction of the expression of the corresponding genes/proteins. Ultimately, this organohalide can be added as sole electron acceptor to evaluate whether it sustains growth by OHR.

#### **Development of a metabolic model for *D. hafniense* strain DCB-2: a tool for present and future research questions**

The facultative OHRB *D. hafniense* strain DCB-2 is known to be capable of diverse metabolisms including fermentation with pyruvate and different types of anaerobic respiration, among which OHR is the focus of this thesis. This bacterium shows drastically different growth performances depending on the conditions it is cultivated in. Surprisingly, the highest growth yield was obtained with the bacteria cultivated in fermentative conditions (Pyr-only). Indeed, fermentation is generally considered as energetically less favourable than respiration for the cells. This discrepancy could be explained by an underlying and partial sulfite respiration revealed by the proteomic data (Chapter 6) which could likely compensate for energy conservation. Our results showed that, in terms of growth yield, Pyr-only condition is followed by Py/Fu, La/Fu, La/CIOHPA, H<sub>2</sub>/Fu and H<sub>2</sub>/CIOHPA (in that order). Between the highest and the lowest growth yield a difference of more than 5-fold was observed. The comparative proteomic approach performed on strain DCB-2 and the downstream data analysis achieved in the course

of the thesis (presented in Chapters 6 and 7) highlighted numerous proteome adaptations to the substrates used. Despite the detailed analysis of the results, the interpretation of the proteomic data in terms of metabolic adaptations is still in its infancy and requires further analytical but also experimental work. Therefore, to better exploit the proteomic dataset and to improve the understanding of strain DCB-2 metabolism, the information collected here on the relative protein abundances will be used in combination with physiological data to constrain a genome-scale metabolic model of strain DCB-2 and to run flux balance analysis.

To this purpose, a collaboration was established towards the end of the thesis with the laboratory of Prof. Vassily Hatzimanikatis (EPFL, Lausanne). The idea is to develop a genome-scale metabolic model of *D. hafniense* strain DCB-2, whose foundation is the functional annotation of the genome of the organism. Genome-scale metabolic models are organism-specific biochemical databases which enable the simulation of cellular metabolism under different conditions. Substantial experimental work has already been done by cultivating strain DCB-2 in the different growth conditions and monitoring substrate consumption and product formation as well as biomass formation. Evangelia Vayena, a PhD student in Prof. Hatzimanikatis' laboratory, is currently developing and refining the metabolic model for strain DCB-2 based on its theoretical proteome. The draft metabolic reconstruction was acquired from the CarveMe collection [214] and was manually curated based on a list of growth-related metabolic tasks adapted from Tymoshenko *et al.* [215] and on our experimental observations. After the manual curation, the model accounts for 810 unique metabolites across the cytoplasm and extracellular space, 1009 biotransformations, 382 transport reactions and 167 boundary reactions. Thermodynamic curation of the network was also performed using method described by Salvay *et al.* [216] resulting in the addition of thermodynamic constraints for 69 % of the reactions. The substrate consumption and product formation data for the six different culture conditions were integrated in the refined version of strain DCB-2 metabolic model in order to reconstruct condition-specific metabolic networks. In the next steps, the proteomic data will be further incorporated into the model in a similar way to what was already described by Hadadi *et al.* [217]. When ready, this tool will help identifying alternative states for the energy and carbon metabolisms of the cells in the different conditions, that can explain the experimental measurements, by identifying the active pathways in each one of the culture media.

In addition, given the high quality dataset and the good level of proteome coverage obtained, the comparative proteomic data and the derived integrative metabolic model provide a solid basis for addressing present and future research questions. For example, the complex I-like enzyme-dependent or -independent models proposed in Chapter 7 could be tested using the metabolic model. Similarly, the implication of the potential redox protein partners for the complex I-like enzyme could be addressed through the integrative metabolic model. So far, it has to be kept in mind that those candidates for redox partners were proposed exclusively based on the similar expression pattern across the proteomic data as that of the complex I-like enzyme. Indeed, no conclusive result could be obtained with the biochemical approach applied to the complex I-like enzymes of *D. hafniense* and *D. restrictus*, thus not supporting the involvement of any of the proposed protein partner candidates. The poor results obtained

here can be explained by the fact that the enzyme, which was not highly abundant, was diluted in the large protein diversity found in the membrane proteomes. Additional identification of the complex I-like enzyme with BN-PAGE or new purification attempts should be performed from membrane extracts subjected to an enrichment procedure. As the tagging of the native complex is impossible in OHRB since they are genetically intractable, an enrichment of the complex using co-immunoprecipitation remains a possible option but requires the development of an antibody specific to one of the subunits of the complex I-like enzyme. A proper isolation of the complex should be helpful to determine its exact subunit composition and could be coupled with a cross-linking approach to improve the chance of isolating possibly transient interacting protein partners.

## A Supplemental material to Chapter 3

Two primer pairs were used in Fusion PCR: F1, R1, F2 and R2. Primers F1 and R2 were designed as primers targeting the beginning and the end of the overall sequence to amplify. When needed, they include a restriction site to subsequently insert the sequence in the expression plasmid. Primers R1 and F2 are targeting the end of the first domain and the beginning of the second domain, respectively. They include a small tail of 5 nucleotides complementary to the fusion site. In a first round of PCR, the coding sequence of the two domains were amplified separately using the pair of primers F1/R1 and F2/R2, respectively. The two PCR products are then mixed in a 1:1 ratio (typically 50 ng of each products were completed with ddH<sub>2</sub>O to a total of 2 µL) to serve as template for a second round of PCR with primers F1 and R2 to generate the final PCR product.

Reactant	Stock Concentration	Round-1		Round-2 Fusion PCR
		1 <sup>st</sup> domain	2 <sup>nd</sup> domain	
Hi-Fi Buffer*	5x	10 µL	10 µL	10 µL
DMSO	100 %	1.5 µL	1.5 µL	-
MgCl <sub>2</sub>	50 mM	2 µL	2 µL	2 µL
dNTPs	10 mM	5 µL	5 µL	5 µL
Genomic DNA	>5 ng/uL	2 µL	2 µL	-
Mix of 1 <sup>st</sup> and 2 <sup>nd</sup> domains	50 ng each	-	-	2 µL
Primer F1	10 uM	2.5 µL	-	1.25 µL
Primer R1	10 uM	2.5 µL	-	-
Primer F2	10 uM	-	2.5 µL	-
Primer R2	10 uM	-	2.5 µL	1.25 µL
ddH <sub>2</sub> O	100 %	23.5 uL	23.5 uL	27.5 µL
VELOCITY DNA Polymerase*	-	1 µL	1 µL	1 µL
Final volume	-	50 µL	50 µL	50 µL

Table A.1 – Fusion PCR mix. \*Bioline, Labgene scientific, Châtel-St-Denis, Switzerland

Primer	Sense	Sequence (5' → 3')	Gene target / plasmid	Reference
MW031	+	gaatCCATGGgagaattttaaaattatattttc	<i>rdhkl</i> / PET24d	This study
MW032	-	gcgcAAGCTTtcaataagctatccccttcgagaag	<i>rdhkl</i> / PET24d	This study
<i>rdhkl61A-F</i>	+	gaatCCATGGgcgttaggggttaggaaaagac	<i>rdhkl61A</i> / PET24d	This study
<i>rdhkl61A-R</i>	-	gcgcAAGCTTtcaataacgcaatgcctcttcacg	<i>rdhkl61A</i> / PET24d	This study
MW033	+	gaatCCATGGcgttgaagggtttggg	<i>rdhkl61B</i> (EBD) / PET24d	This study
MW034	-	tgaagaccggtatctcccatactctt	<i>rdhkl61B</i> (EBD) / PET24d	This study
MW035	+	ccggtcttcattatgtcaccgtaagcagata	<i>rdhkl61B</i> (DBD) / PET24d	This study
MW036	-	gcgcCTCGAGtcaataagctatccccttcgagaag	<i>rdhkl61B</i> (DBD) / PET24d	This study
DBC-F	+	gaatttcagggcgatgtaac	<i>mgI</i> -DB chassis	This study
DBC-R	-	cggtgcgaatcatagctatc	<i>mgI</i> -DB chassis	This study
<i>mgI</i> -DB7	+	gaatttcagggcgatgtaacgcgttcaagttataacaca-	DB7 internal primer /	This study
		ttaatacttaacaatgtgatagctatgattgcacgcg	<i>mgI</i> -DB chassis	
<i>mgI</i> -DB8	+	gaatttcagggcgatgtaacgcgttcaaatgtatatacgc-	DB8 internal primer /	This study
		taataatacttaacaatgtgatagctatgattgcacgcg	<i>mgI</i> -DB chassis	
<i>mgI</i> -noDB	+	gaatttcagggcgatgtaacgcgttcaagcctciagttgct-	noDB internal primer /	This study
		ggcttgatcttaacaatgtgatagctatgattgcacgcg	<i>mgI</i> -DB chassis	
BG1743	+	gcgcAAGCTTggccaaggctcgccatgacg	<i>cprB4</i> promoter / PAK80	((40))
BG1704	-	gcGATCCgctataaaaaataaaatgatacc	<i>cprB4</i> promoter / PAK80	((40))
MW060	-	tctcaacccttccttcgtagt	<i>cpr</i> -DB7 promoter / PAK80	This study
MW061	+	gatacctattacaataatacccaactactt	<i>cpr</i> -DB7 promoter / PAK80	This study
MW062	+	tgagaatcaggttaaatgtatatacgtaaatattgcgta-	<i>cpr</i> -DB8 internal primer /	This study
		tgtagatcc	PAK80	
MW063	+	tgagaatcaggttagcctciagttggctgggttgattgcgta-	<i>cpr</i> -noDB internal primer /	This study
		tgtagatcc	PAK80	

Table A.2 – List of oligonucleotides used in the study.

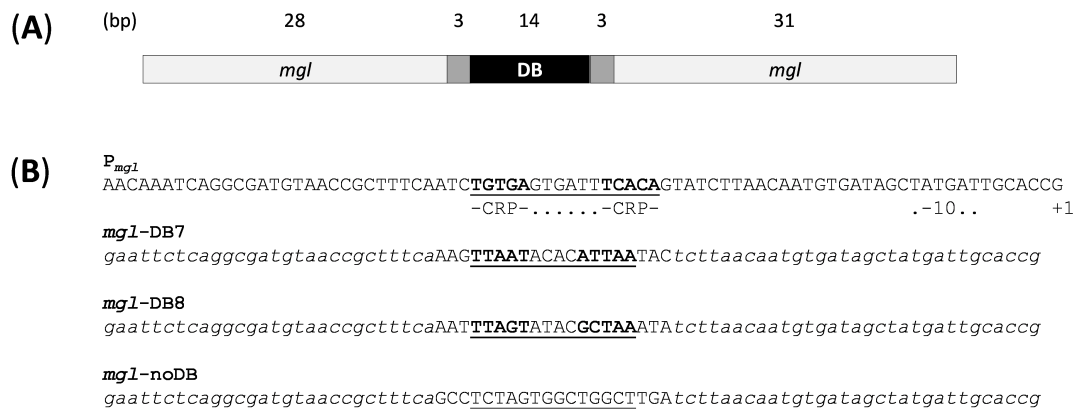


Figure A.1 – Design of *mgl*-DB chassis for EMSA analysis. (A) Schematic representation of the *mgl* promoter containing the DB motif. Each *mgl*-DB DNA fragment is composed by a unique 20-nt long sequence corresponding to the DB sequence (14 nt) flanked by 3 nt on each side from the original *rdh* promoter, and embedded in a common DNA chassis derived from the *mglBAC* promoter (*P<sub>mgl</sub>*) from *E. coli*. This design was chosen to avoid experimental incompatibility as the *mgl* promoter shares structural features with promoters targeted by RdhK proteins and is known to be the target of CRP [218]. (B) For each DB of interest, a 79-nt long oligonucleotide was designed and used as template for PCR amplification with a pair of primers targeting the 5'- and 3'-ends of *P<sub>mgl</sub>*.

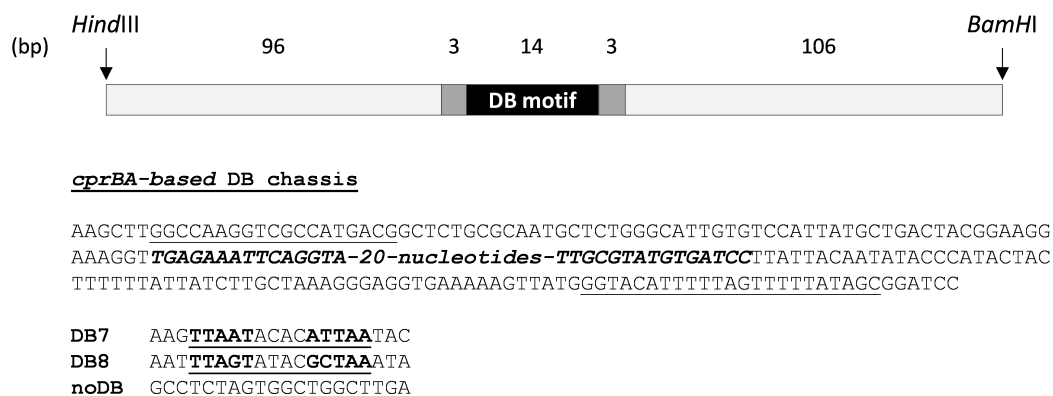


Figure A.2 – Design of the *cprBA*-based DB chassis for *in vivo* beta-galactosidase reporter assay. Schematic representation of the cloning strategy for the DB chassis used in fusion with the reporter gene. The DB7 motif and three flanking nucleotides on each side (20 bp in total) was replaced by DB8 or the non-palindromic noDB sequence. Using DB motifs embedded in the native *cprBA* promoter was chosen to address the question of the specificity of the DB motif exclusively. The 222-bp fragments were produced by fusion PCR, and cloned into pAK80 carrying the beta-galactosidase gene.

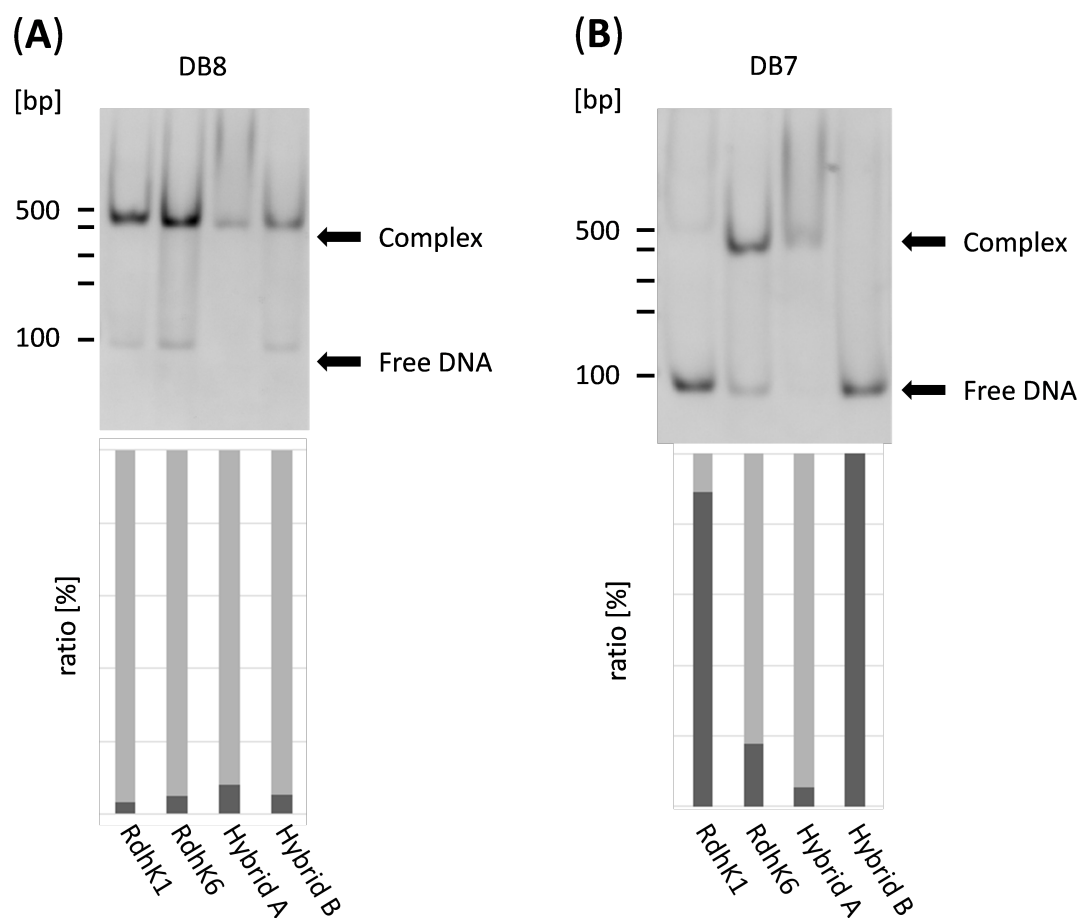


Figure A.3 – Control EMSA experiments with the four RdhK proteins, DB8 (A) or DB7 (B), and the combination of both effectors. Both effectors (CLOHPA and 3,5-DCP) were added in a 1:1 ratio in the reactions. This experiment was used to exclude the influence of potential inhibitory or competitive events between the two compounds.

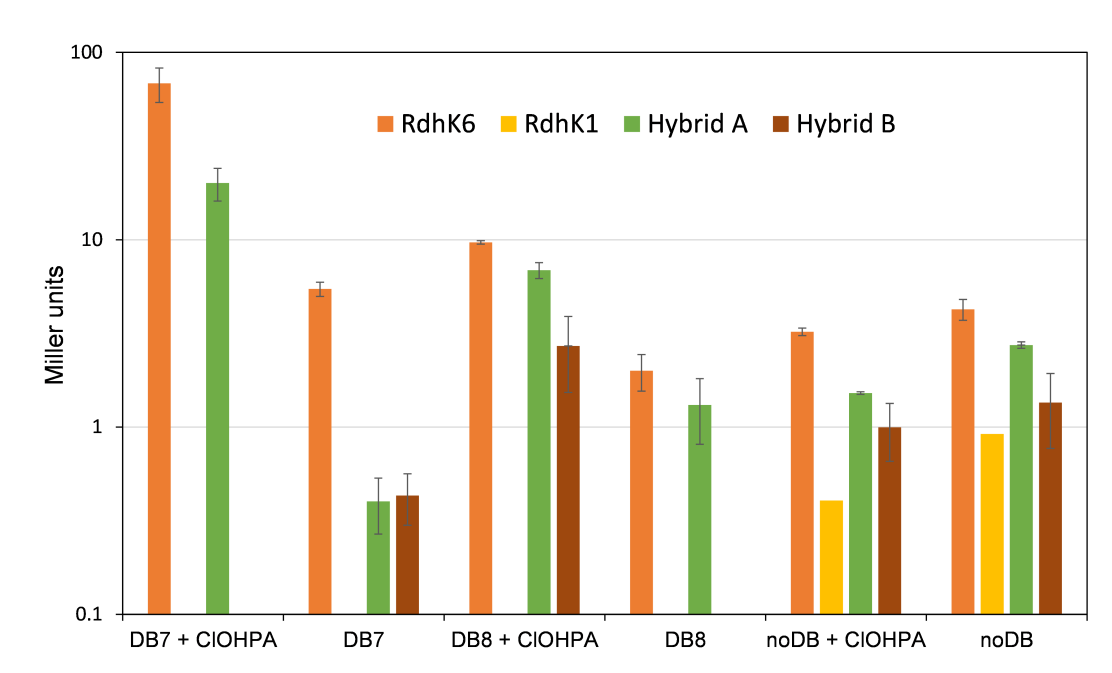


Figure A.4 – Preliminary *in vivo* beta-galactosidase assay for promoter-binding activity of RdhK proteins. The preliminary experiment was performed at small scale with all four proteins (RdhK6, RdhK1, Hybrid A and B) in combination with promoters carrying either DB7, DB8 or noDB motif in presence/absence of CIOHPA as effector. Each experiment was run in triplicates and the protocol was adapted to be run in a screening mode using a 96-well reader plate. The most interesting RdhK/DB combinations were further tested at a larger scale in order to strengthen the signal-to-noise ratio.

DB5	TTATTgcacATTAA
DB6	TTAGTgcacTCTAA
DB7	TTAATacacATTAA
DB8	TTAGTatacGCTAA
FNR	TTGATacacATCAA

Figure A.5 – Sequence alignment of DB motifs recognised by RdhK6. The DB5, DB6 and DB7 motifs are part of the *rdh-6* gene cluster, DB8 belongs to the *rdh-1* gene cluster, while FNR motif represents the paradigmatic palindrome recognised by FNR.



## **B Supplemental material to Chapter 4**

Primer	Sense	Sequence (5' → 3')	Target	Ref.
ARP050F	+	CATGTGCATACCTCCTTGTTAAATC	PR <i>rdhBA1</i>	[116]
ARP051R	-	CATAATTTACCTCCCTACCTCA	PR <i>rdhBA1</i>	[116]
ARP052F	+	AGTTTTTATTGGGGAACCTTACCTTC	IR <i>rdhK2-rdhBA2</i>	[116]
ARP053R	-	CATAACCAAACCTCCTTTTTTCT	IR <i>rdhK2-rdhBA2</i>	[116]
ARP054F	+	CATTGTGAAAATATTATCA	IR <i>rdhK3-rdhBA3</i>	[116]
ARP055R	-	CATCATTTTCACCTCCTCTTTAG	IR <i>rdhK3-rdhBA3</i>	[116]
ARP056F	+	GCATATCCAACGCTTAATATGGA	IR <i>rdhA3-rdhBA4</i>	[116]
ARP057R	-	CATTTTCACCTCCTCTCTCAGC	IR <i>rdhA3-rdhBA4</i>	[116]
ARP058F	+	AAAAATCAAACCTCCCTAATATTGATTAT	IR <i>rdhK5-rdhAB5</i>	[116]
ARP059R	-	CATTCCGACTGAGGTTTAAAGAAA	IR <i>rdhK5-rdhAB5</i>	[116]
ARP060F	+	AAAAATCAAACCTCCCTTTCGATAGATT	IR <i>rdhK6-rdhAB6</i>	[116]
ARP061R	-	CATACCTTATCTCCTCCTTAAATAT	IR <i>rdhK6-rdhAB6</i>	[116]
ARP062F	+	AAAAATCAAACCTCTATATCGTGGATT	IR <i>rdhK7-rdhBA7</i>	[116]
ARP063R	-	CATTGTGTTTCCCTCACCTCCTTAATTA	IR <i>rdhK7-rdhBA7</i>	[116]
MW026	+	TGAACGACGCGGAAAGACCATT	IR <i>rdhK9-rdhBA8</i>	This study
APR023R	-	CATTGTTTTACACCTCCTATCTTC	IR <i>rdhK9-rdhBA8</i>	[116]
ARP103F	+	ATAAGTGGTGGCTGGAATGG	IR <i>rdhK10-rdhBA10</i>	[116]
ARP104R	-	GGCTCTGGGTTTGACGAAAAG	IR <i>rdhK10-rdhBA10</i>	[116]
ARP066F	+	AATTAAAGTTTGACAATTCATTTGTACAA	IR <i>rdhK11-rdhBA11</i>	[116]
ARP067R	-	CATCATTTTCACCTCCCTTTTGATAATT	IR <i>rdhK11-rdhBA11</i>	[116]
ARP105F	+	CATTAAATACCGCCGAAAGG	PR <i>rdhBA12</i>	[116]
ARP106R	-	AACAAATGAAATCCCCATGC	PR <i>rdhBA12</i>	[116]
ARP072F	+	GAAAACACGCTCCCAAGTGCA	PR <i>rdhBA13</i>	[116]
ARP098R	-	CATCTGTTCACCTCCATTCCAATT	PR <i>rdhBA13</i>	[116]
ARP074F	+	CTTTGGATTGTTGATGGTAAAAAAT	IR <i>rdhK15-rdhBA14</i>	[116]
ARP075R	-	CACGTCTCCCTTCTTCTACTTT	IR <i>rdhK15-rdhBA14</i>	[116]
ARP024F	+	TAGAACTTCCATCTGTTATCTGAAA	PR <i>rdhBA15</i>	[116]
APR025R	-	CATATTTCTCCTCCTTTATTTTTATTAGC	PR <i>rdhBA15</i>	[116]
ARP026F	+	AAGGTAAAAATCAAAATCCCTGCTTC	IR <i>rdhK16-rdhBA16</i>	[116]
APR027R	-	TTATCTTTTCACCTCCCTTTTATGCG	IR <i>rdhK16-rdhBA16</i>	[116]
ARP107F	+	TCAATAGGGTAACCGTCGCA	IR <i>rdhK17-rdhBA17</i>	[116]
ARP108R	-	TGGAACCAATTCTGGTGCT	IR <i>rdhK17-rdhBA17</i>	[116]
ARP028F	+	TATAGTTATAAAAACTGCTTGCTCAG	IR <i>rdhK18-rdhBA18</i>	[116]
APR029R	-	CATTATATTTACCTCCTTCTACTC	IR <i>rdhK18-rdhBA18</i>	[116]
ARP116F	+	CGGTGACAAGGTATATGATATT	PR <i>rdhBA19</i>	[116]
ARP117R	-	CATAGCTTTTTACCTCCCTTATA	PR <i>rdhBA19</i>	[116]
ARP030F	+	CAATAAAAACCAAGCTTATGATTACC	PR <i>rdhBA20</i>	[116]
APR031R	-	CATGTAGTCTCTCCTCCCTTAA	PR <i>rdhBA20</i>	[116]
ARP080F	+	CTACGAAAAAATTTATCCACAAAGCTT	IR <i>rdhK21-rdhBA21</i>	[116]
ARP081R	-	CATGATTACGGACACCTAATC	IR <i>rdhK21-rdhBA21</i>	[116]
ARP101F	+	AACAAATCGAGAAGTACAAGTGGT	IR <i>rdhKA23-rdhAB22</i>	[116]
ARP102R	-	TTCCCAGTGATTTTAGAAACATCATCGTCC	IR <i>rdhKA23-rdhAB22</i>	[116]
ARP084F	+	TTGAAAAATAGAAAACAACAAAATATAGTT	IR <i>rdhK23-rdhBA23</i>	[116]
ARP085R	-	CATATTTTTCACCTCCTTCCCAGTTAA	IR <i>rdhK23-rdhBA23</i>	[116]
ARP086F	+	CTGTTGGTGTGCTCATATATTGTAT	PR <i>rdhBA24</i>	[116]
ARP086R	-	CATTGCAATCTTCCTCCTTAATAAT	PR <i>rdhBA24</i>	[116]

Table B.1 – List of primers used to amplify *rdhA* promoters in PERK-23. IR = intergenic region, PR = promoter region.

## C Supplemental material to Chapter 6

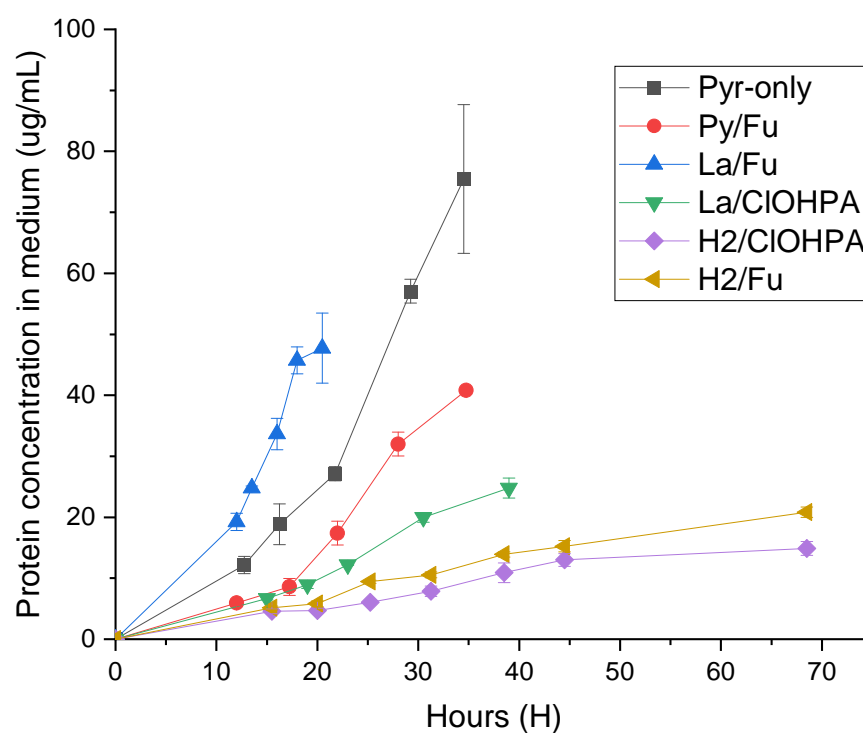


Figure C.1 – Comparison of the growth performance of strain DCB-2 in the six growth conditions. The protein concentration in 1 mL of growth media was followed overtime and used as a proxy for comparing the growth performances of the bacteria in the six growth conditions relevant in Chapter 6.

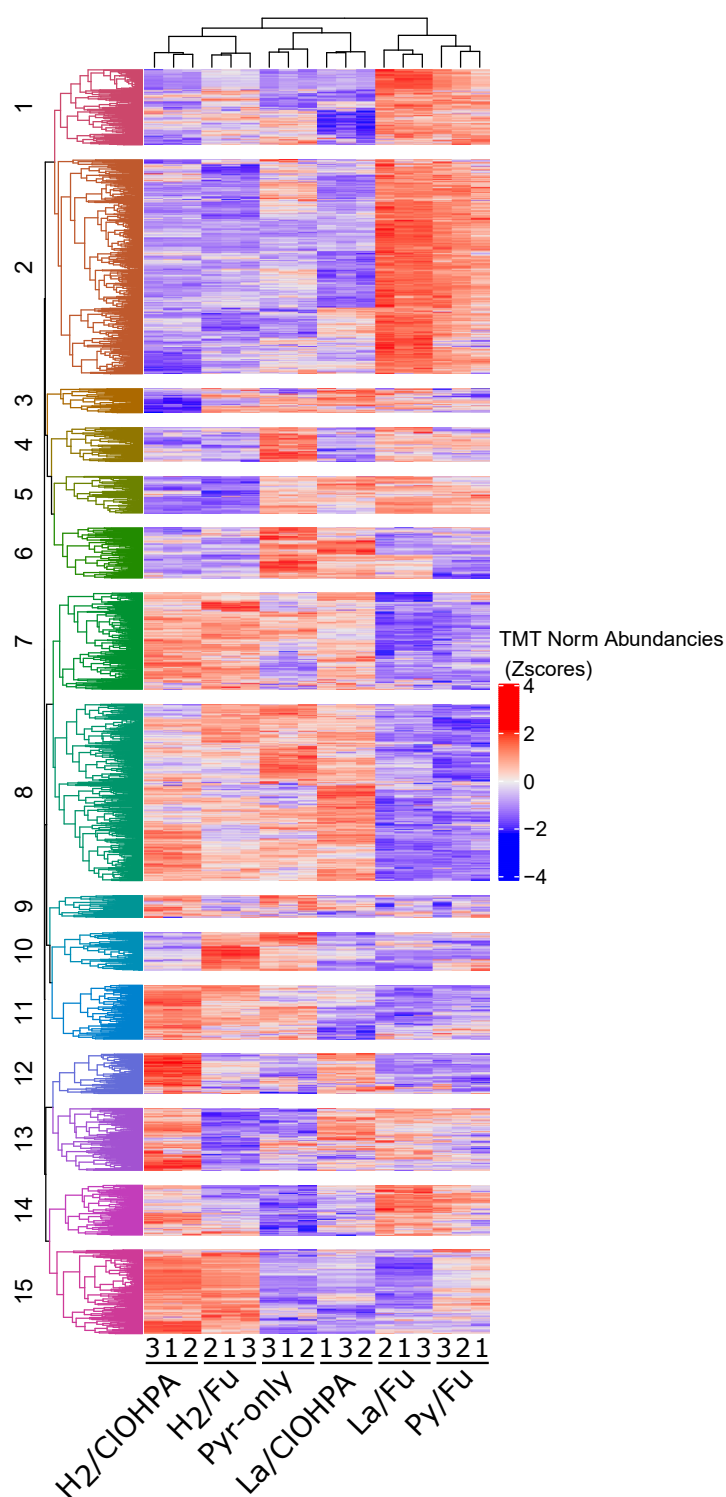
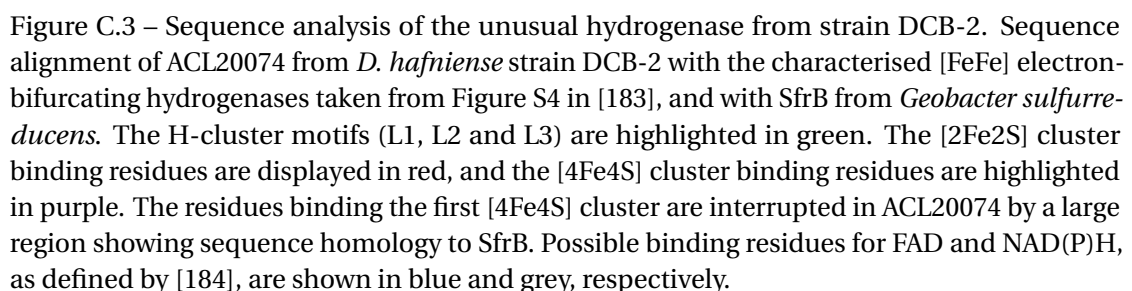


Figure C.2 – Heat-map of the proteomic results with hierarchical clustering. Heat-map representing the abundance (expressed in Z-scores) of the 2796 proteins across the 18 samples. The proteins were distributed vertically in 15 different clusters according to their behaviour over the dataset. In addition, the samples were grouped and arranged by similarity on the vertical axes.



Protein ref.	Predicted function	Fold change in Pyr-only vs				
		Py/Fu	La/Fu	H2/Fu	H2/Cl	La/Cl
ACL18239	Arylsulfotransferase	5.82	3.75	2.99	3.48	1.69
ACL18309	Thioredoxin domain protein	3.94	2.18	2.15	2.93	1.60
ACL18432	PAS modulated sigma54 transcriptional regulator	4.86	2.39	3.42	4.73	1.83
ACL18435	TRAP dicarboxylate transporter, DctP subunit	9.45	3.86	3.63	6.05	1.89
ACL18837	Succinate dehydrogenase	3.45	2.22	2.22	2.56	1.76
ACL18870	Hypothetical protein	5.13	3.51	3.25	3.26	1.67
ACL18923	FAD-dependent pyridine nucleotide-disulfide oxidoreductase	6.40	2.71	3.16	3.89	1.38
ACL19889; ACL19995	Dimethylsulfoxide reductase, chain B	7.55	4.87	3.46	2.13	0.54
ACL21058	Hypothetical protein	2.63	2.17	3.66	3.53	1.88
ACL21175	4Fe-4S ferredoxin binding protein, DmsB-like	7.85	12.89	5.53	6.68	1.45
ACL21176	Molybdopterin oxidoreductase, DmsA-like	5.41	5.90	3.46	3.68	1.36
ACL21177	FAD-dependent NAD(P)H-disulfide oxidoreductase, (CoA-disulfide reductase?)	39.73	30.64	14.51	27.76	1.11
ACL21178	4Fe-4S ferredoxin binding protein, DmsB-like	3.25	3.46	2.74	2.77	1.18
ACL21179	Unknown function DUF81	73.30	82.56	27.68	83.91	1.79
ACL21180	Hypothetical protein	6.42	6.01	3.68	3.79	1.97
ACL21182	Hypothetical protein	15.92	15.32	8.17	11.89	1.45
ACL21183	Unknown function DUF81	13.36	12.21	6.69	7.74	1.49
ACL21288	NiFe-dependent hydrogenase, large subunit	17.37	6.46	2.24	2.10	1.25
ACL21289	NiFe-dependent hydrogenase, small subunit	12.46	4.56	2.41	2.04	1.38
ACL21751	Sodium/sulfate symporter	5.79	3.03	4.01	3.39	1.67
ACL21798	FMN-binding domain protein	5.78	3.46	3.53	3.84	1.50
ACL22160	Flagellar basal-body rod protein FlgC	2.69	2.31	2.19	2.23	1.54
ACL22386	Cupin 2 conserved barrel domain protein	3.55	2.65	2.41	2.28	1.23
ACL22616	Transcriptional regulator, LysR family	3.63	2.09	2.37	2.31	1.34
ACL22695	Molybdopterin oxidoreductase	8.60	3.54	2.78	3.74	1.62

Table C.1 – Proteins up-regulated in fermentation metabolism in comparison to all respiratory metabolisms except La/ClOHPA.

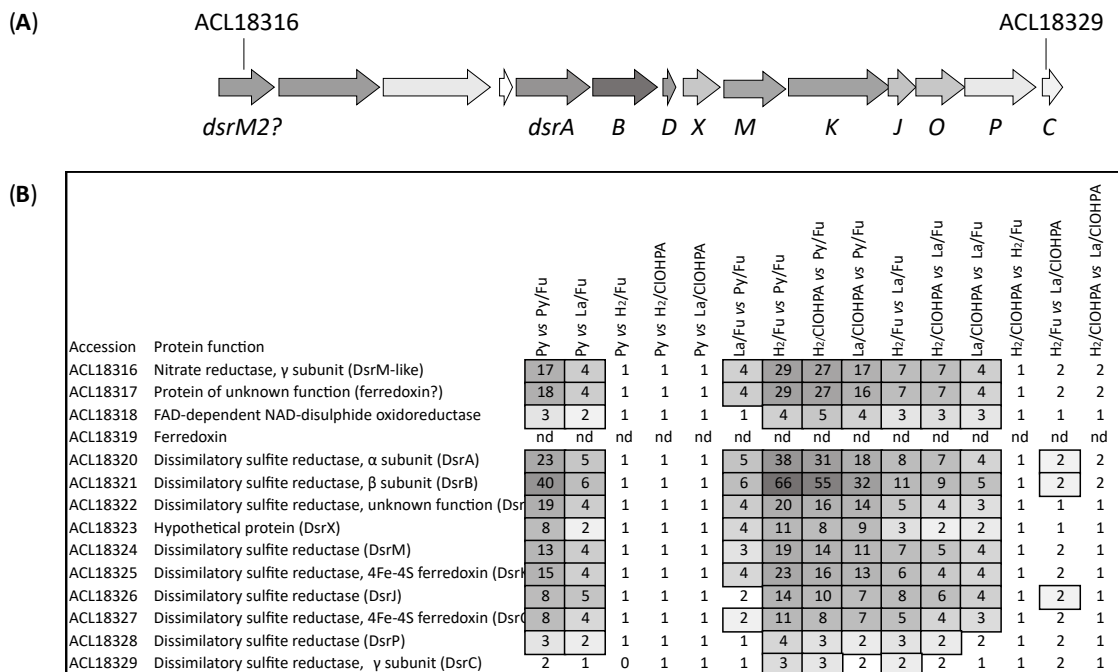


Figure C.4 – The *dsr* gene cluster of *D. hafniense* strain DCB-2. (A) Genetic organisation of *dsr* genes in strain DCB-2. No annotation is given for the genes without a clear function. The intensity of grey shading corresponds to the level of up-regulation of the different proteins in fermentation metabolism as compared to pyruvate/fumarate respiration. (B) Up-regulation fold-changes of Dsr proteins in pairwise comparisons across the six growth conditions. (nd: not detected in the proteomic analysis)

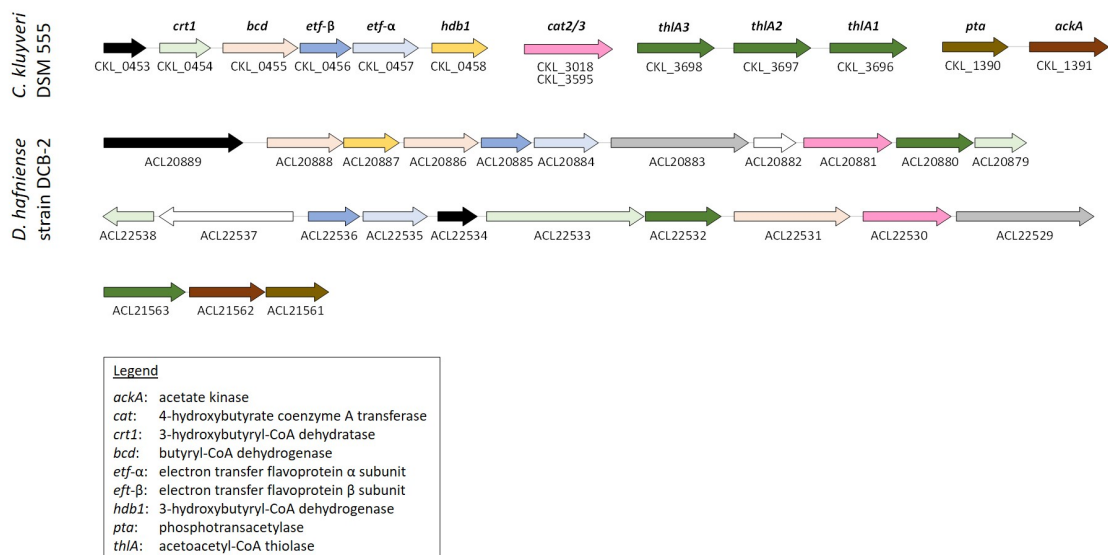


Figure C.5 – Comparison between the two putative electron bifurcating gene clusters of strains DCB-2 with that of *Clostridium kluyveri* described in [199].



## **D Supplemental material to Chapter 7**

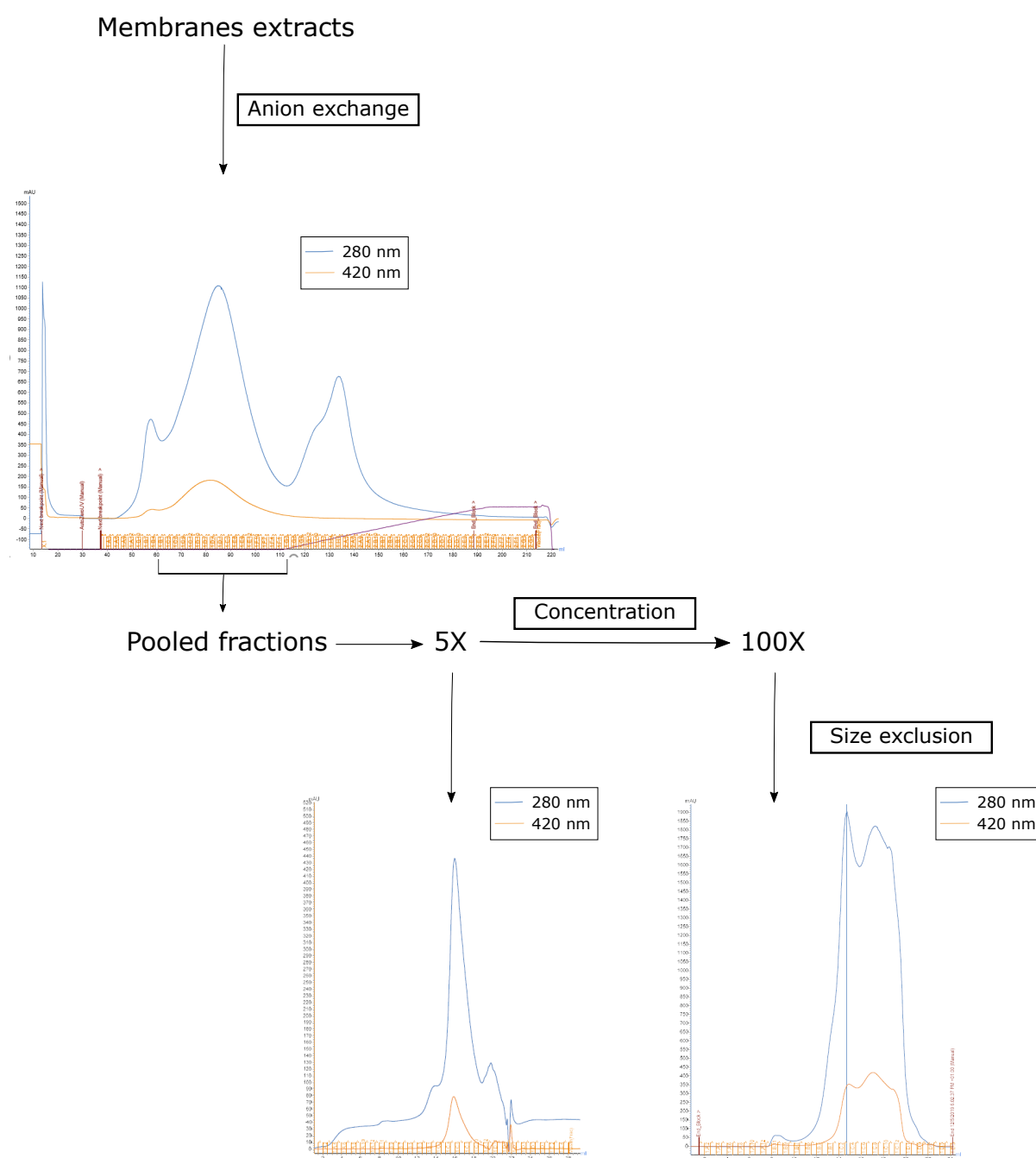


Figure D.1 – Purification attempt of the complex I-like enzyme of *D. hafniense* (1). The membrane extracts were first purified by anion exchange chromatography. The chromatogram obtained is displayed at the top of the figure. The fractions showing absorbance at 420 nm were pooled together and concentrated to be further processed by size exclusion chromatography (SEC). Two different concentrations of the fractions were loaded on SEC (either 5x or 100x) and the resulting chromatograms are shown at the bottom of the figure. The resolution obtained for the 100X concentrated sample was poor probably due to a too high protein concentration in the sample and to the overloading of the column, thus preventing an efficient purification. This is supported by the saturation of the 280 nm signal and by the fact that only one protein pattern is observed by SDS-PAGE (See Figure 7.8).

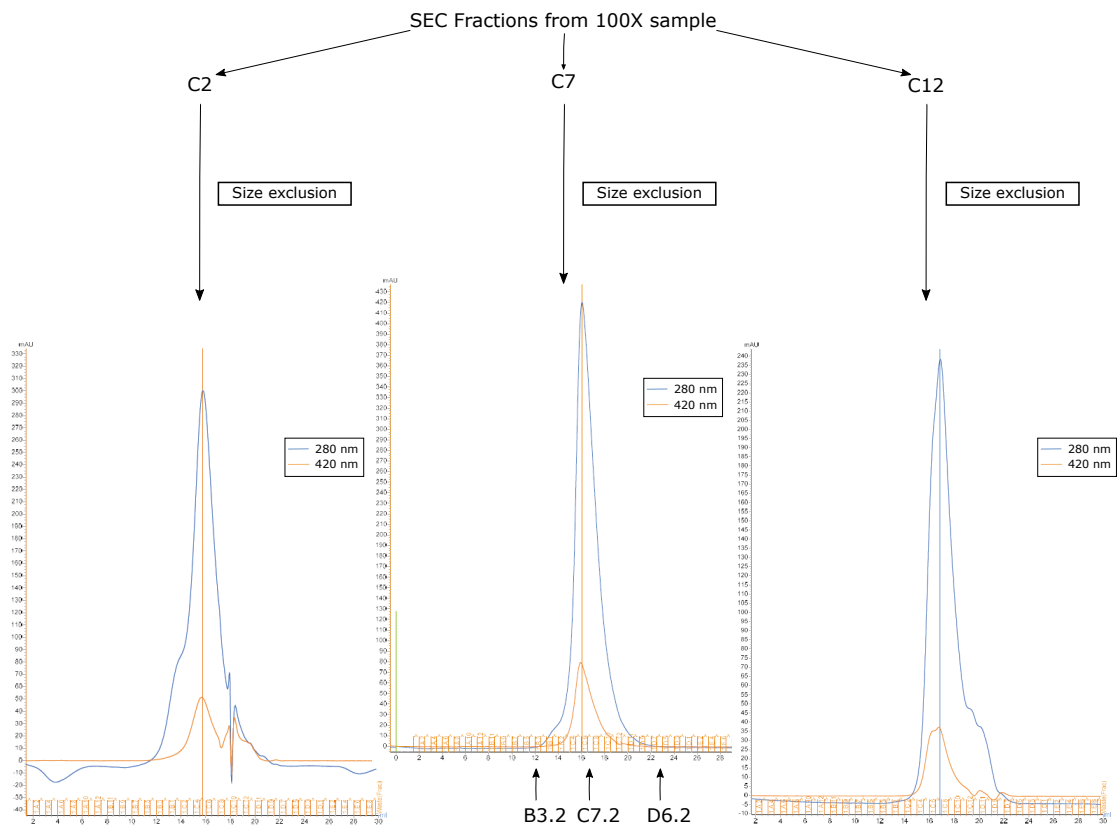


Figure D.2 – Purification attempt of the complex I-like enzyme of *D. hafniense* (2). Three different fractions (C2, C7 and C12) of the 100X concentrated sample obtained from the anion exchange pooled fractions and showing slightly different protein profiles on SDS-PAGE gels were re-loaded for a second round of SEC. In fractions C2 and C12, some co-eluting signals were observed before and after the main peak, respectively. In contrast, one single and symmetrical peak was observed centred at the elution volume of 16 mL with fraction C7. As the chromatogram of fraction C7 (shown in the middle) displayed a well defined protein pattern, fractions B3.2, C7.2 and D6.2 were analysed by LC-MS/MS to determine the protein composition enriched in fraction C7.2.

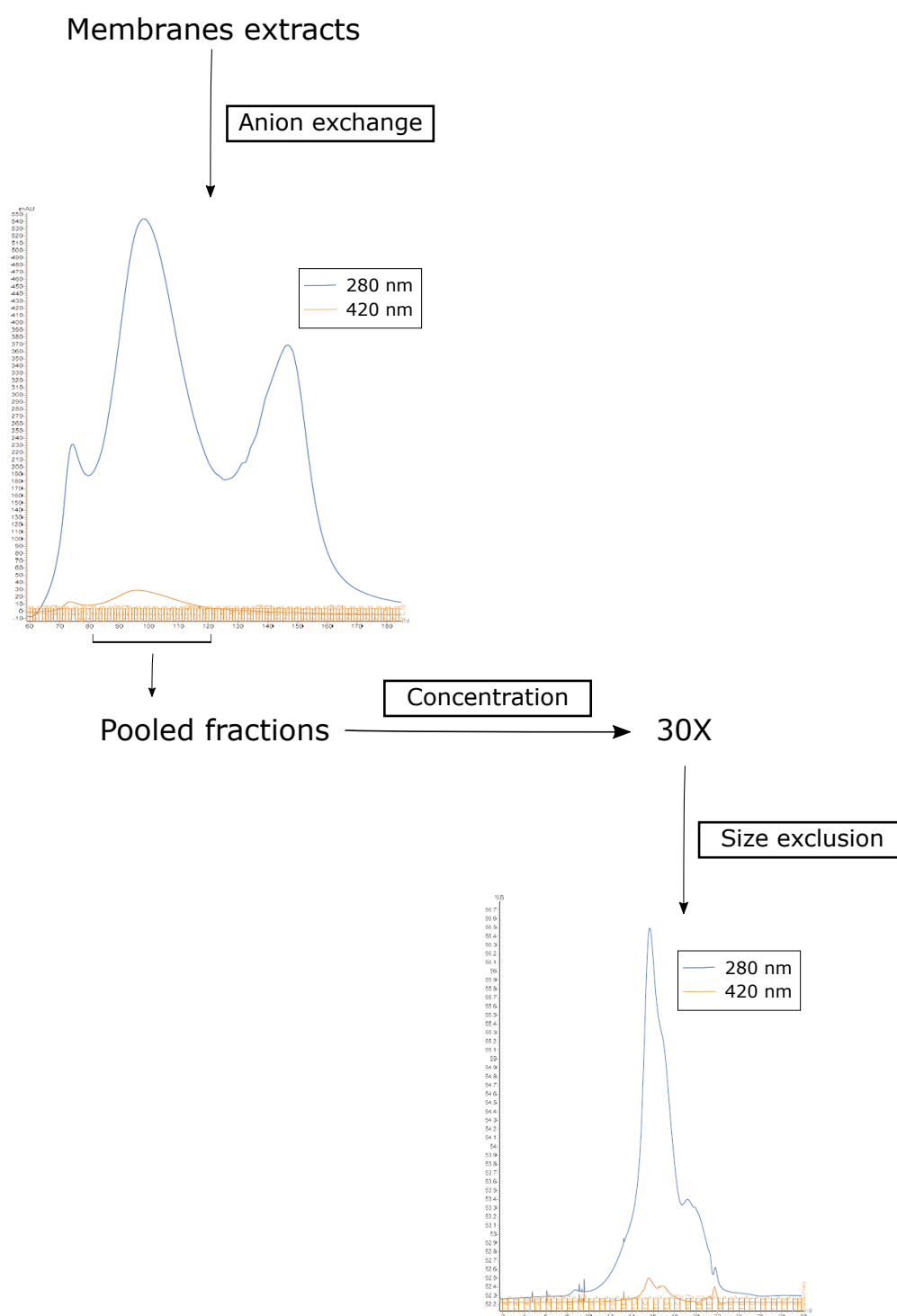


Figure D.3 – Purification attempt of the complex I-like enzyme of *D. restrictus*. The membrane extracts were first purified by anion exchange chromatography. The chromatogram obtained is displayed at the top of the figure. The fractions showing absorbance at 420 nm were pooled together and concentrated 30x to be further processed on size exclusion chromatography as shown at the bottom of the figure.

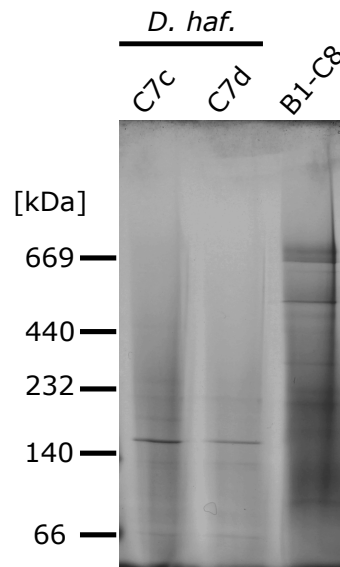


Figure D.4 – BN-PAGE with purified membrane extracts. The gel was loaded with the two fractions corresponding to the centre of the peak obtained after the first round of SEC with both *D. hafniense* samples (5x = C7<sub>d</sub> or 100x = C7<sub>c</sub>). These results confirmed the LC-MS/MS analysis which suggested that several protein complexes were present in the main peak observed at the elution volume of 16 mL in the SEC chromatograms. Indeed, corresponding lanes display several protein bands with one dominating signal slightly higher than 140 kDa. No signal showed a potential enrichment of a complex at the size expected for complex I-like enzymes. In addition to *D. hafniense* samples, the fractions resulting from *D. restrictus* SEC analysis corresponding to the main peak (fraction B1 to C8) were pooled together and concentrated prior to load on the gel in order to have sufficient material. The lane corresponding to the sample (B1-C8) was of poorer resolution but still revealed many proteins likely corresponding to several different complexes with widely distributed molecular weights.



## Bibliography

1. Oluwasanu, A. A. Fate and Toxicity of Chlorinated Phenols of Environmental Implications: A Review. *Medicinal & Analytical Chemistry International Journal* **2**, 1–8 (2018).
2. World Health Organization (WHO). Tetrachloroethene in Drinking-water Background. *WHO Guidelines for Drinking-water Quality* **2**, 1–9 (2003).
3. Field, J. A. in *Organohalide-Respiring Bacteria* 7–29 (2016).
4. Atashgahi, S., Häggblom, M. M. & Smidt, H. Organohalide respiration in pristine environments: implications for the natural halogen cycle. *Environmental Microbiology* **20**, 934–948 (2018).
5. Atashgahi, S., Lu, Y. & Smidt, H. in *Organohalide-Respiring Bacteria* (eds Adrian, L. & Löffler, F. E.) 63–105 (Springer, 2016).
6. Adrian, L. & Löffler, F. E. *Organohalide-Respiring Bacteria* (eds Adrian, L. & Löffler, F. E.) 632 (Springer, 2016).
7. Maillard, J. & Willemin, M. S. in *Adv Microb Physiol* (ed Poole, R. K.) 191–238 (Elsevier, 2019).
8. McCarty, P. L. in *Organohalide-Respiring Bacteria* (eds Adrian, L. & Löffler, F. E.) 51–62 (Springer, 2016).
9. Dolfig, J. Reductive dechlorination of 3-chlorobenzoate is coupled to ATP production and growth in an anaerobic bacterium, strain DCB-1. *Microbiology* **153**, 264–266 (1990).
10. Mohn, W. W. & Tiedje, J. M. Strain DCB-1 conserves energy for growth from reductive dechlorination coupled to formate oxidation. *Archives of Microbiology* **153**, 267–271 (1990).
11. Mackiewicz, M. & Wiegel, J. Comparison of energy and growth yields for *Desulfitobacterium dehalogenans* during utilization of chlorophenol and various traditional electron acceptors. *Applied and Environmental Microbiology* **64**, 352–355 (1998).
12. Schumacher, W. & Holliger, C. The proton/electron ratio of the menaquinone-dependent electron transport from dihydrogen to tetrachloroethene in "*Dehalobacter restrictus*". *Journal of Bacteriology* **178**, 2328–2333 (1996).

13. Leys, D., Adrian, L. & Smidt, H. Organohalide respiration: Microbes breathing chlorinated molecules. *Philosophical Transactions of the Royal Society B: Biological Sciences* **368**, 8–10 (2013).
14. Maillard, J. & Holliger, C. in *Organohalide-Respiring Bacteria* (eds Adrian, L. & Löffler, F. E.) 153–171 (Springer, 2016).
15. Futagami, T. & Furukawa, K. in *Organohalide-Respiring Bacteria* (eds Adrian, L. & Löffler, F. E.) 173–206 (Springer, 2016).
16. Madsen, T. & Licht, D. Isolation and characterization of an anaerobic chlorophenol-transforming bacterium. *Applied and Environmental Microbiology* **58**, 2874–2878 (1992).
17. Utkin, I., Woese, C. & Wiegel, J. Isolation and Characterization of *Desulfitobacterium dehalogenans* gen. nov., sp. nov., an Anaerobic Bacterium Which Reductively Dechlorinates Chlorophenolic Compounds. *International Journal of Systematic Bacteriology* **44**, 612–619 (1994).
18. Christiansen, N., Ahring, B. K., Wohlfarth, G. & Diekert, G. Purification and characterization of the 3-chloro-4-hydroxy-phenylacetate reductive dehalogenase of *Desulfitobacterium hafniense*. *FEBS letters* **436**, 159–162 (1998).
19. Hug, L. A. & Edwards, E. A. Diversity of reductive dehalogenase genes from environmental samples and enrichment cultures identified with degenerate primer PCR screens. *Frontiers in Microbiology* **4**, 1–16 (2013).
20. Neumann, A., Wohlfarth, G. & Diekert, G. Tetrachloroethene dehalogenase from *Dehalospirillum multivorans*: Cloning, sequencing of the encoding genes, and expression of the *pceA* gene in *Escherichia coli*. *Journal of Bacteriology* **180**, 4140–4145 (1998).
21. McMurdie, P. J. *et al.* Localized plasticity in the streamlined genomes of vinyl chloride respiring *Dehalococcoides*. *PLoS Genetics* **5**, 1–10 (2009).
22. Futagami, T., Goto, M. & Furukawa, K. Biochemical and genetic bases of dehalorespiration. *Chemical Record* **8**, 1–12 (2008).
23. Molenda, O. *et al.* Insights into origins and function of the unexplored majority of the reductive dehalogenase gene family as a result of genome assembly and ortholog group classification. *Environmental Science: Processes and Impacts* **22**, 663–678 (2020).
24. Bommer, M. *et al.* Structural basis for organohalide respiration. *Science* **346**, 455–458 (2014).
25. Maillard, J., Genevaux, P. & Holliger, C. Redundancy and specificity of multiple trigger factor chaperones in *Desulfitobacterium*. *Microbiology* **157**, 2410–2421 (2011).
26. Low, A. *et al.* A comparative genomics and reductive dehalogenase gene transcription study of two chloroethene-respiring bacteria, *Dehalococcoides mccartyi* strains MB and 11a. *Scientific Reports* **5**, 1–12 (2015).
27. Key, T. A. *et al.* Genome sequence of the organohalide-respiring *Dehalogenimonas alkenigignens* type strain (IP3-3T). *Standards in Genomic Sciences* **11**, 1–11 (2016).

28. Tang, S., Wang, P. H., Higgins, S. A. & Löffler, F. E. Sister *Dehalobacter* Genomes Reveal Specialization in Organohalide Respiration and Recent Strain Differentiation Likely Driven by Chlorinated Substrates. *Frontiers in Microbiology* **7**, 1–14 (2016).
29. Kruse, T. *et al.* Complete genome sequence of *Dehalobacter restrictus*. *Standards in Genomic Sciences* **8**, 375–388 (2013).
30. Kruse, T., Smidt, H. & Lechner, U. in *Organohalide-Respiring Bacteria* (eds Adrian, L. & Löffler, F. E.) 345–376 (Springer, 2016).
31. Müller, J. A. *et al.* Molecular identification of the catabolic vinyl chloride reductase from *Dehalococcoides* sp. strain VS and its environmental distribution. *Applied and Environmental Microbiology* **70**, 4880–4888 (2004).
32. Maillard, J., Regeard, C. & Holliger, C. Isolation and characterization of Tn-Dha1, a transposon containing the tetrachloroethene reductive dehalogenase of *Desulfotobacterium hafniense* strain TCE1. *Environmental Microbiology* **7**, 107–117 (2005).
33. Seshadri, R. *et al.* Genome sequence of the PCE-dechlorinating bacterium *Dehalococcoides ethenogenes*. *Science* **307**, 105–108 (2005).
34. Kim, S.-H. *et al.* Genome sequence of *Desulfotobacterium hafniense* DCB-2, a Gram-positive anaerobe capable of dehalogenation and metal reduction. *BMC Microbiology* **12**, 1–20 (2012).
35. Kube, M. *et al.* Genome sequence of the chlorinated compound-respiring bacterium *Dehalococcoides* species strain CBDB1. *Nature Biotechnology* **23**, 1269–1273 (2005).
36. Wagner, D. D. *et al.* Genomic determinants of organohalide-respiration in *Geobacter lovleyi*, an unusual member of the Geobacteraceae. *BMC Genomics* **13**, 1–17 (2012).
37. Smidt, H., Van Leest, M., Van der Oost, J. & De Vos, W. M. Transcriptional regulation of the *cpr* gene cluster in ortho-chlorophenol-respiring *Desulfotobacterium dehalogenans*. *Journal of Bacteriology* **182**, 5683–5691 (2000).
38. Goris, T. *et al.* Insights into organohalide respiration and the versatile catabolism of *Sulfurospirillum multivorans* gained from comparative genomics and physiological studies. *Environmental Microbiology* **16**, 3562–3580 (2014).
39. Kruse, T. *et al.* Comparative genomics of the genus *Desulfotobacterium*. *FEMS microbiology ecology* **93**, 1–22 (2017).
40. Gábor, K. *et al.* Characterization of CprK1, a CRP / FNR-Type Transcriptional Regulator of Halorespiration from *Desulfotobacterium hafniense*. *Journal of Bacteriology* **188**, 2604–2613 (2006).
41. Buttet, G. F., Willemin, M. S., Hamelin, R., Rupakula, A. & Maillard, J. The membrane-bound C subunit of reductive dehalogenases: Topology analysis and reconstitution of the FMN-binding domain of pceC. *Frontiers in Microbiology* **9**, 1–15 (2018).
42. Matsui, M., Tomita, M. & Kanai, A. Comprehensive computational analysis of bacterial CRP/FNR super family and its target motifs reveals stepwise evolution of transcriptional networks. *Genome Biology and Evolution* **5**, 267–282 (2013).

43. Morita, Y., Futagami, T., Goto, M. & Furukawa, K. Functional characterization of the trigger factor protein PceT of tetrachloroethene-dechlorinating *Desulfotobacterium hafniense* Y51. *Applied Microbiology and Biotechnology* **83**, 775–781 (2009).
44. Hug, L. A. in *Organohalide-Respiring Bacteria* (eds Adrian, L. & Löffler, F. E.) 377–392 (Springer, 2016).
45. Schubert, T., Adrian, L., Sawers, R. G. & Diekert, G. Organohalide respiratory chains : composition , topology and key enzymes. *FEMS microbiology ecology* **94**, 1–17 (2018).
46. Fincker, M. & Spormann, A. M. Biochemistry of Catabolic Reductive Dehalogenation. *Annu. Rev. Biochem* **86**, 357–86 (2017).
47. Mayer-Blackwell, K., Holly, S., Fincker, M. & Spormann, A. M. in *Organohalide-Respiring Bacteria* (eds Adrian, L. & Löffler, F. E.) 259–280 (Springer, 2016).
48. Schubert, T. & Diekert, G. in *Organohalide-Respiring Bacteria* (eds Adrian, L. & Löffler, F. E.) 397–428 (Springer, 2016).
49. Schipp, C. J., Marco-Urrea, E., Kublik, A., Seifert, J. & Adrian, L. Organic cofactors in the metabolism of *Dehalococcoides mccartyi* strains. *Philosophical Transactions of the Royal Society B: Biological Sciences* **368**, 1–12 (2013).
50. Zinder, S. H. in *Organohalide-Respiring Bacteria* (eds Adrian, L. & Löffler, F. E.) 107–137 (Springer, 2016).
51. Kublik, A. *et al.* Identification of a multi-protein reductive dehalogenase complex in *Dehalococcoides mccartyi* strain CBDB1 suggests a protein-dependent respiratory electron transport chain obviating quinone involvement. *Environmental Microbiology* **18**, 3044–3056 (2016).
52. Seidel, K., Kühnert, J. & Adrian, L. The complexome of *Dehalococcoides mccartyi* reveals its organohalide respiration-complex is modular. *Frontiers in Microbiology* **9**, 1–14 (2018).
53. Louie, T. M. & Mohn, W. W. Evidence for a chemiosmotic model of dehalorespiration in *Desulfomonile tiedjei* DCB-1. *Journal of Bacteriology* **181**, 40–46 (1999).
54. Simon, J., van Spanning, R. J. M. & Richardson, D. J. The organisation of proton motive and non-proton motive redox loops in prokaryotic respiratory systems. *Biochimica et Biophysica Acta - Bioenergetics* **1777**, 1480–1490 (2008).
55. Holliger, C. *et al.* *Dehalobacter restrictus* gen. nov. and sp. nov., a strictly anaerobic bacterium that reductively dechlorinates tetra- and trichloroethene in an anaerobic respiration. *Archives of Microbiology* **169**, 313–321 (1998).
56. Scholz-Muramatsu, H., Neumann, A., Meßmer, M., Moore, E. & Diekert, G. Isolation and characterization of *Dehalospirillum multivorans* gen. nov., sp. nov., a tetrachloroethene-utilizing, strictly anaerobic bacterium. *Archives of Microbiology* **163**, 48–56 (1995).
57. Rupakula, A. *et al.* The restricted metabolism of the obligate organohalide respiring bacterium *Dehalobacter restrictus*: lessons from tiered functional genomics. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences* **368**, 1–16 (2013).

58. Van De Pas, B. A. *et al.* Purification and molecular characterization of ortho-chlorophenol reductive dehalogenase, a key enzyme of halo-respiration in *Desulfitobacterium dehalogenans*. *Journal of Biological Chemistry* **274**, 20287–92 (1999).
59. Schumacher, W., Holliger, C., Zehnder, A. J. & Hagen, W. R. Redox chemistry of cobalamin and iron-sulfur cofactors in the tetrachloroethene reductase of *Dehalobacter restrictus*. *FEBS Letters* **409**, 421–425 (1997).
60. Kräutler, B. *et al.* The Cofactor of Tetrachloroethene Reductive Dehalogenase of *Dehalospirillum multivorans* Is Norpseudob12, a New Type of a Natural Corrinoid. *Helvetica Chimica Acta* **86**, 3698–3716 (2003).
61. Miller, E., Wohlfarth, G. & Diekert, G. Studies on tetrachloroethene respiration in *Dehalospirillum multivorans*. *Archives of Microbiology* **166**, 379–387 (1996).
62. Buckel, W. & Thauer, R. K. Energy conservation via electron bifurcating ferredoxin reduction and proton/Na<sup>+</sup> translocating ferredoxin oxidation. *Biochimica et Biophysica Acta - Bioenergetics* **1827**, 94–113 (2013).
63. Buckel, W. & Thauer, R. K. Flavin-Based Electron Bifurcation, A New Mechanism of Biological Energy Coupling. *Chemical Reviews* **118**, 3862–3886 (2018).
64. Dolfing, J. & Novak, I. The Gibbs free energy of formation of halogenated benzenes, benzoates and phenols and their potential role as electron acceptors in anaerobic environments. *Biodegradation* **26**, 15–27 (2015).
65. Mohn, W. W. & Tiedje, J. M. Microbial reductive dehalogenation. *Microbiological Reviews* **56**, 482–507 (1992).
66. Goris, T. *et al.* Proteomics of the organohalide- respiring Epsilonproteobacterium *Sulfurospirillum multivorans* adapted to tetrachloroethene and other energy substrates. *Nature Publishing Group* **5**, 1–13 (2015).
67. Nonaka, H. *et al.* Complete genome sequence of the dehalorespiring bacterium *Desulfitobacterium hafniense* Y51 and comparison with *Dehalococcoides ethenogene* 195. *Journal of Bacteriology* **188**, 2262–2274 (2006).
68. Kruse, T. *et al.* Genomic, proteomic, and biochemical analysis of the organohalide respiratory pathway in *Desulfitobacterium dehalogenans*. *Journal of Bacteriology* **197**, 893–904 (2015).
69. Sazanov, L. A. in *Mechanisms of Primary Energy Transduction in Biology* (ed Wikström, M.) 25–59 (Royal Society of Chemistry, 2017).
70. Willemin, M. S., Vingerhoets, M., Holliger, C. & Maillard, J. Hybrid Transcriptional Regulators for the Screening of Target DNA Motifs in Organohalide-Respiring Bacteria. *Frontiers in Microbiology* **11**, 1–11 (2020).
71. Bisailon, A., Beaudet, R., Lépine, F. & Villemur, R. Quantitative Analysis of the Relative Transcript Levels of Four Chlorophenol Reductive Dehalogenase Genes in *Desulfitobacterium hafniense* PCP-1 Exposed to Chlorophenols. *Applied and Environmental Microbiology* **77**, 6261–6264 (2011).

72. Mac Nelly, A., Kai, M., Svatoš, A., Schubert, T. & Diekert, G. Functional Heterologous Production of Reductive Dehalogenases from *Desulfitobacterium hafniense* Strains. *Applied and Environmental Microbiology* **80**, 4313–4322 (2014).
73. Tsukagoshi, N., Satoshi, Uenaka, T., Suzuki, N. & Kurane, R. Isolation and transcriptional analysis of novel tetrachloroethene reductive dehalogenase gene from *Desulfitobacterium* sp. strain KCB1. *Applied Genetics and Molecular Biotechnology* **69**, 543–553 (2006).
74. Futagami, T., Tsuboi, Y., Suyama, A., Goto, M. & Furukawa, K. Emergence of two types of nondechlorinating variants in the tetrachloroethene-halo-respiring *Desulfitobacterium* sp. strain Y51. *Applied Microbiology and Biotechnology* **70**, 720–728 (2006).
75. Mukherjee, K. *et al.* *Dehalogenimonas lykanthroporepellens* BL-DC-9T simultaneously transcribes many *rdhA* genes during organohalide respiration with 1,2-DCA, 1,2-DCP, and 1,2,3-TCP as electron acceptors. *FEMS Microbiology Letters* **354**, 111–118 (2014).
76. Wagner, A., Adrian, L., Kleinstaub, S., Andreesen, J. R. & Lechner, U. Transcription analysis of genes encoding homologues of reductive dehalogenases in "*Dehalococcoides*" sp. Strain CBDB1 by using terminal restriction fragment length polymorphism and quantitative PCR. *Applied and Environmental Microbiology* **75**, 1876–1884 (2009).
77. Wagner, A. *et al.* Regulation of reductive dehalogenase gene transcription in *Dehalococcoides mccartyi*. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences* **368**, 1–10 (2013).
78. Jacob-Dubuisson, F., Mechaly, A., Betton, J.-M. & Antoine, R. Structural insights into the signalling mechanisms of two-component systems. *Nature Reviews Microbiology* **16**, 585–593 (2018).
79. Gao, R. & Stock, A. M. Biological Insights from Structures of Two-Component Proteins. *Annual Review of Microbiology* **63**, 133–154 (2009).
80. Takeda, H. *et al.* Dual two-component regulatory systems are involved in aromatic compound degradation in a polychlorinated-biphenyl degrader, *Rhodococcus jostii* RHA1. *Journal of Bacteriology* **192**, 4741–4751 (2010).
81. Esken, J. *et al.* Tetrachloroethene respiration in *Sulfurospirillum* species is regulated by a two-component system as unraveled by comparative genomics, transcriptomics, and regulator binding studies. *Microbiology Open* **9**, 1–31 (2020).
82. Goris, T. *et al.* The complete genome of the tetrachloroethene-respiring Epsilonproteobacterium *Sulfurospirillum halorespirans*. *Journal of Biotechnology* **255**, 33–36 (2017).
83. Siddaramappa, S. *et al.* Complete genome sequence of *Dehalogenimonas lykanthroporepellens* type strain (BL-DC-9T) and comparison to *Dehalococcoides* strains. *Standards in Genomic Sciences* **6**, 251–264 (2012).
84. Türkowsky, D. *et al.* Thermal proteome profiling allows quantitative assessment of interactions between tetrachloroethene reductive dehalogenase and trichloroethene. *Journal of Proteomics* **192**, 10–17 (2018).

85. John, M. *et al.* Retentive memory of bacteria: Long-term regulation of dehalorespiration in *Sulfurospirillum multivorans*. *Journal of Bacteriology* **191**, 1650–1655 (2009).
86. Krasper, L. *et al.* The MarR-type regulator RdhR1456 regulates *rdh* gene transcription in *Dehalococcoides mccartyi* strain CBDB1. *Journal of Bacteriology* **198**, 3130–3141 (2016).
87. Gábor, K., Sene, K. H., Smidt, H., de Vos, W. M. & van der Oost, J. Divergent roles of CprK paralogues from *Desulfitobacterium hafniense* in activating gene expression. *Microbiology* **154**, 3686–3695 (2008).
88. Wilkinson, S. P. & Grove, A. Ligand-responsive transcriptional regulation by members of the MarR family of winged helix proteins. *Current Issues in Molecular Biology* **8**, 51–62 (2006).
89. Grove, A. MarR family transcription factors. *Current Biology* **23**, R142–R143 (2013).
90. Schiffmann, C. L. *et al.* Proteomic dataset of the organohalide-respiring bacterium *Dehalococcoides mccartyi* strain CBDB1 grown on hexachlorobenzene as electron acceptor. *Data in Brief* **7**, 253–256 (2016).
91. Shaw, D. J., Rice, D. W. & Guest, J. R. cAMP receptor protein (CRP) and fumarate and nitrate reduction regulatory protein (FNR) are homologues. *Journal of Molecular Biology* **166**, 241–247 (1983).
92. Busby, S. & Ebright, R. H. Transcription activation by catabolite activator protein (CAP). *Journal of Molecular Biology* **293**, 199–213 (1999).
93. Green, J., Scott, C. & Guest, J. R. Functional versatility in the CRP-FNR superfamily of transcription factors: FNR and FLP. *Adv Microb Physiol* **44**, 1–34 (2001).
94. Guest, J. R., Green, J., Irvine, A. S. & Spiro, S. in *Regulation of Gene Expression in Escherichia coli* 317–342 (Springer US, Boston, MA, 1996).
95. Kolb, A., Busby, S., Buc, H., Garges, S. & Adhya, S. Transcriptional Regulation by cAMP and its Receptor Protein. *Annual Review of Biochemistry* **62**, 749–797 (1993).
96. Körner, H., Sofia, H. J. & Zumft, W. G. Phylogeny of the bacterial superfamily of Crp-Fnr transcription regulators : exploiting the metabolic spectrum by controlling alternative gene programs. *FEMS Microbiology Reviews* **27**, 559–592 (2003).
97. Shimada, T., Fujita, N., Yamamoto, K. & Ishihama, A. Novel roles of camp receptor protein (CRP) in regulation of transport and metabolism of carbon sources. *PLoS ONE* **6**, 1–11 (2011).
98. Ebert, M. *et al.* Heme and nitric oxide binding by the transcriptional regulator DnrF from the marine bacterium *Dinoroseobacter shibae* increases *napD* promoter affinity. *Journal of Biological Chemistry* **292**, 15468–15480 (2017).
99. McLaughlin, P. T., Bhardwaj, V., Feeley, B. E. & Higgs, P. I. MrpC, a CRP/Fnr homolog, functions as a negative autoregulator during the *Myxococcus xanthus* multicellular developmental program. *Molecular Microbiology* **109**, 245–261 (2018).

100. Ranganathan, S. *et al.* Characterization of a cAMP responsive transcription factor, Cmr (Rv1675c), in TB complex mycobacteria reveals overlap with the DosR (DevR) dormancy regulon. *Nucleic Acids Research* **44**, 134–151 (2016).
101. Pop, S. M., Kolarik, R. J. & Ragsdale, S. W. Regulation of anaerobic dehalorespiration by the transcriptional activator CprK. *Journal of Biological Chemistry* **279**, 49910–49918 (2004).
102. Joyce, M. G. *et al.* CprK Crystal structures reveal mechanism for transcriptional control of halo-respiration. *Journal of Biological Chemistry* **281**, 28318–28325 (2006).
103. Levy, C. *et al.* Molecular basis of halo-respiration control by CprK, a CRP-FNR type transcriptional regulator. *Molecular Microbiology* **70**, 151–167 (2008).
104. Kemp, L. R., Dunstan, M. S., Fisher, K., Warwicker, J. & Leys, D. The transcriptional regulator CprK detects chlorination by combining direct and indirect readout mechanisms. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences* **368**, 20120323 (2013).
105. Mazon, H. *et al.* Transcriptional activation by CprK1 is regulated by protein structural changes induced by effector binding and redox state. *Journal of Biological Chemistry* **282**, 11281–11290 (2007).
106. Kiley, P. J. & Beinert, H. Oxygen sensing by the global regulator , FNR : the role of the iron-sulfur cluster. *FEMS Microbiology Reviews* **22**, 341–352 (1999).
107. Gupta, N. & Ragsdale, S. W. Dual roles of an essential cysteine residue in activity of a redox-regulated bacterial transcriptional activator. *Journal of Biological Chemistry* **283**, 28721–28728 (2008).
108. Pop, S. M., Gupta, N., Raza, A. S. & Ragsdale, S. W. Transcriptional activation of dehalorespiration: Identification of redox-active cysteines regulating dimerization and DNA binding. *Journal of Biological Chemistry* **281**, 26382–26390 (2006).
109. Larkin, M. A. *et al.* Clustal W and Clustal X version 2 . 0. *Bioinformatics* **23**, 2947–2948 (2007).
110. Letunic, I. & Bork, P. Interactive tree of life ( iTOL ) v3 : an online tool for the display and annotation of phylogenetic and other trees. *Nucleic Acids Research* **44**, 242–245 (2016).
111. Wang, S., Zhang, W., Yang, K.-l. & He, J. Isolation and characterization of a novel *Dehalobacter* species strain TCP1 that reductively dechlorinates. *Biodegradation* **25**, 313–323 (2014).
112. Maphosa, F., Vos, W. M. D. & Smidt, H. Exploiting the ecogenomics toolbox for environmental diagnostics of organohalide-respiring bacteria. *Trends in Biotechnology* **28**, 308–316 (2010).
113. Bertani, G. Studies on lysogenesis. I . The Mode of Phage Liberation by Lysogenic *Escherichia Coli*. *Journal of Bacteriology* **62**, 293–300 (1951).
114. Diaz, E. & Prieto, M. A. Bacterial promoters triggering biodegradation of aromatic pollutants. *Current Opinion in Biotechnology* **11**, 467–475 (2000).

115. Bartlett, A. *et al.* Mapping genome-wide transcription-factor binding sites using DAP-seq. *Nature Protocols* **12**, 1659–1672 (2017).
116. Rupakula, A. *Understanding the metabolism of tetrachloroethene- respiring Dehalobacter restrictus : from genome analysis , corrinoid cofactor biosynthesis to regulation of reductive dehalogenases* PhD thesis (EPFL, 2015).
117. Mayer-Blackwell, K., Sewell, H., Fincker, M. & Spormann, A. M. in *Organohalide-Respiring Bacteria* (eds Adrian, L. & Löffler, F. E.) 259–280 (Springer, 2016).
118. Van De Pas, B. A. *et al.* Energy Yield of Respiration on Chloroaromatic Compounds in *Desulfitobacterium dehalogenans*. *Applied and Environmental Microbiology* **67**, 3958–3963 (2001).
119. Schiffmann, C. L. *et al.* Proteome profile and proteogenomics of the organohalide-respiring bacterium *Dehalococcoides mccartyi* strain CBDB1 grown on hexachlorobenzene as electron acceptor. *Journal of Proteomics* **98**, 59–64 (2014).
120. Brandt, U. Energy Converting NADH: Quinone Oxidoreductase (Complex I). *Annual Review of Biochemistry* **75**, 69–92 (2006).
121. Berrisford, J. M., Baradaran, R. & Sazanov, L. A. Structure of bacterial respiratory complex I. *Biochimica et Biophysica Acta (BBA) - Bioenergetics* **1857**, 892–901 (2016).
122. Zickermann, V. *et al.* Mechanistic insight from the crystal structure of mitochondrial complex I. *Science* **347**, 44–49 (2015).
123. Parey, K., Wirth, C., Vonck, J. & Zickermann, V. Respiratory complex I — structure, mechanism and evolution. *Current Opinion in Structural Biology* **63**, 1–9 (2020).
124. Sazanov, L. A. A giant molecular proton pump: structure and mechanism of respiratory complex I. *Nature Reviews Molecular Cell Biology* **16**, 375–388 (2015).
125. Wirth, C., Brandt, U., Hunte, C. & Zickermann, V. Structure and function of mitochondrial complex i. *Biochimica et Biophysica Acta - Bioenergetics* **1857**, 902–914 (2016).
126. Friedrich, T. The NADH:ubiquinone oxidoreductase (complex I) from *Escherichia coli*. *Biochimica et Biophysica Acta - Bioenergetics* **1364**, 134–146 (1998).
127. Peng, G. *et al.* Isolation, characterization and electron microscopic single particle analysis of the NADH:ubiquinone oxidoreductase (complex I) from the *Hyperthermophilic eubacterium* Aquifex aeolicus. *Biochemistry* **42**, 3032–3039 (2003).
128. Yagi, T. & Matsuno-Yagi, A. The proton-translocating NADH-quinone oxidoreductase in the respiratory chain: The secret unlocked. *Biochemistry* **42**, 2266–2274 (2003).
129. Friedrich, T. On the mechanism of respiratory complex I. *Journal of Bioenergetics and Biomembranes* **46**, 255–268 (2014).
130. Gnannt, E., Dörner, K., Strampaad, M. F., de Vries, S. & Friedrich, T. The multitude of iron–sulfur clusters in respiratory complex I. *Biochimica et Biophysica Acta - Bioenergetics* **1857**, 1068–1072 (2016).

131. Hayashi, T. & Stuchebrukhov, A. A. Electron tunneling in respiratory complex I. *PNAS* **107**, 19157–19162 (2010).
132. Baradaran, R., Berrisford, J. M., Minhas, G. S. & Sazanov, L. A. Crystal structure of the entire respiratory complex i. *Nature* **494**, 443–448 (2013).
133. Uno, S., Kimura, H., Murai, M. & Miyoshi, H. Exploring the quinone / inhibitor-binding pocket in mitochondrial respiratory complex I by chemical biology approaches. *American Society for Biochemistry and Molecular Biology* **294**, 679–696 (2019).
134. Singer, T. P. Biomembranes Part F: Bioenergetics: Oxidative Phosphorylation. *Methods in Enzymology* **55**, 454–462 (1979).
135. Hu, P., Lv, J., Fu, P. & Hualing, M. Enzymatic characterization of an active NDH complex from *Thermosynechococcus elongatus*. *FEBS Letters* **587**, 2340–2345 (2013).
136. Esposti, M. D. Genome analysis of structure-function relationships in respiratory complex i, an ancient bioenergetic enzyme. *Genome Biology and Evolution* **8**, 126–147 (2016).
137. Sazanov, L. A. Respiratory complex I: Mechanistic and structural insights provided by the crystal structure of the hydrophilic domain. *Biochemistry* **46**, 2275–2288 (2007).
138. Galemou Yoga, E., Angerer, H., Parey, K. & Zickermann, V. Respiratory complex I – Mechanistic insights and advances in structure determination. *Biochimica et Biophysica Acta - Bioenergetics* **1861**, 148153 (2020).
139. Calhoun, M. W., Oden, K. L., Gennis, R. B., Mattos, M. J. T. D. & Neijssel, O. M. Energetic Efficiency of *Escherichia coli* : Effects of Mutations in Components of the Aerobic Respiratory Chain. *Journal of Bacteriology* **175**, 3020–3025 (1993).
140. Rodenburg, R. J. Mitochondrial complex I-linked disease. *Biochimica et Biophysica Acta - Bioenergetics* **1857**, 938–945 (2016).
141. Moparthi, V. K. & Hägerhäll, C. The Evolution of Respiratory Chain Complex I from a Smaller Last Common Ancestor Consisting of 11 Protein Subunits. *Journal of Molecular evolution* **72**, 484–497 (2011).
142. Brandt, U. Adaptations of an ancient modular machine. *Science* **363**, 230–231 (2019).
143. Friedrich, T. & Scheide, D. The respiratory complex I of bacteria, archaea and eukarya and its module common with membrane-bound multisubunit hydrogenases. *FEBS Letters* **479**, 1–5 (2000).
144. Spero, M. A., Aylward, F. O., Currie, C. R. & Donohue, T. J. Phylogenomic Analysis and Predicted Physiological Role of the Proton-Translocating NADH:Quinone Oxidoreductase (Complex I) Across Bacteria. *mBio* **6**, 00389–15 (2015).
145. Ohnishi, T., Ohnishi, S. T. & Salerno, J. C. Five decades of research on mitochondrial NADH-quinone oxidoreductase (complex I). *Biological Chemistry* **399**, 1249–1264 (2018).

146. Jones, A. J., Blaza, J. N., Varghese, F. & Hirst, J. Respiratory complex i in *bos taurus* and *paracoccus denitrificans* pumps four protons across the membrane for every NADH oxidized. *Journal of Biological Chemistry* **292**, 4987–4995 (2017).
147. Fassone, E. & Rahman, S. Complex I deficiency: Clinical features, biochemistry and molecular genetics. *Journal of Medical Genetics* **49**, 578–590 (2012).
148. Yip, C. Y., Harbour, M. E., Jayawardena, K., Fearnley, I. M. & Sazanov, L. A. Evolution of respiratory complex I "Supernumerary" subunits are present in the  $\alpha$ -proteobacterial enzyme. *Journal of Biological Chemistry* **286**, 5023–5033 (2011).
149. Schut, G. J. *et al.* The role of geochemistry and energetics in the evolution of modern respiratory complexes from a proton-reducing ancestor. *Biochimica et Biophysica Acta - Bioenergetics* **1857**, 958–970 (2016).
150. Weerakoon, D. R. & Olson, J. W. The *Campylobacter jejuni* NADH:Ubiquinone oxidoreductase (complex I) utilizes flavodoxin rather than NADH. *Journal of Bacteriology* **190**, 915–925 (2008).
151. Taylor, A. J. & Kelly, D. J. The function, biogenesis and regulation of the electron transport chains in *Campylobacter jejuni*: New insights into the bioenergetics of a major food-borne pathogen. *Advances in Microbial Physiology* **74**, 239–329 (2019).
152. Sancho, J. Flavodoxins: Sequence, folding, binding, function and beyond. *Cellular and Molecular Life Sciences* **63**, 855–864 (2006).
153. Battchikova, N., Eisenhut, M. & Aro, E.-m. Cyanobacterial NDH-1 complexes : Novel insights and remaining puzzles. *Biochimica et Biophysica Acta - Bioenergetics* **1807**, 935–944 (2011).
154. Nowaczyk, M. M. *et al.* NdhP and NdhQ: Two novel small subunits of the cyanobacterial NDH-1 complex. *Biochemistry* **50**, 1121–1124 (2012).
155. Schuller, J. M. *et al.* Structural adaptations of photosynthetic complex I enable ferredoxin-dependent electron transfer. *Science* **363**, 257–260 (2019).
156. Yamamoto, H., Peng, L., Fukao, Y. & Shikanai, T. An Src homology 3 domain-like fold protein forms a ferredoxin binding site for the chloroplast NADH dehydrogenase-like complex in *Arabidopsis*. *Plant Cell* **23**, 1480–1493 (2011).
157. Bruschi, M. & Guerlesquin, F. Structure , function and evolution of bacterial ferredoxins. *FEMS Microbiology Reviews* **54**, 155–176 (1988).
158. Bäumer, S. *et al.* The F420 H<sub>2</sub> Dehydrogenase from *Methanosarcina mazei* Is a Redox-driven Proton Pump Closely Related to NADH Dehydrogenases. *The Journal of Biological Chemistry* **275**, 17968–17973 (2000).
159. Men, Y. *et al.* Metagenomic and Metatranscriptomic Analyses Reveal the Structure and Dynamics of a Dechlorinating Community Containing *Dehalococcoides mccartyi* and Corrinoid-Providing Microorganisms under Cobalamin-Limited Conditions. *Applied and Environmental Microbiology* **83**, 1–14 (2017).

160. Türkowsky, D. *et al.* An integrative overview of genomic, transcriptomic and proteomic analyses in organohalide respiration research. *FEMS Microbiology Ecology* **94**, 1–21 (2018).
161. Wiśniewski, J. R., Zougman, A., Nagaraj, N. & Mann, M. Universal sample preparation method for proteome analysis. *Nature Methods* **6**, 359–362 (2009).
162. Kulak, N. A., Pichler, G., Paron, I., Nagaraj, N. & Mann, M. Minimal, encapsulated proteomic-sample processing applied to copy-number estimation in eukaryotic cells. *Nature Methods* **11**, 319–324 (2014).
163. Dorfer, S. V. *et al.* MS Amanda, a universal identification algorithm optimized for high accuracy tandem mass spectra. *Journal of Proteome Research* **13**, 3679–3684 (2014).
164. Kong, A. T., Leprevost, F. V., Avtonomov, D. M., Mellacheruvu, D. & Nesvizhskii, A. I. MSFragger: Ultrafast and comprehensive peptide identification in mass spectrometry-based proteomics. *Nature Methods* **14**, 513–520 (2017).
165. Team, R. C. *R: A language and environment for statistical computing*. R Foundation for Statistical Computing 2018.
166. Plubell, D. L. *et al.* Extended multiplexing of tandem mass tags (TMT) labeling reveals age and high fat diet specific proteome changes in mouse epididymal adipose tissue. *Molecular and Cellular Proteomics* **16**, 873–890 (2017).
167. Robinson, M. D., McCarthy, D. J. & Smyth, G. K. edgeR: A Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* **26**, 139–140 (2009).
168. Ritchie, M. E. *et al.* Limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Research* **43**, e47 (2015).
169. Benjamini, Y. & Hochberg, Y. Controlling the False Discovery Rate : A Practical and Powerful Approach to Multiple Testing. *Journal of Royal Statistical Society* **57**, 289–300 (1995).
170. Erb, T. J., Frerichs-Revermann, L., Fuchs, G. & Alber, B. E. The apparent malate synthase activity of *Rhodobacter sphaeroides* is due to two paralogous enzymes, (3S)-malyl-coenzyme A (CoA)/ $\beta$ -methylmalyl-CoA lyase and (3S)-malyl-CoA thioesterase. *Journal of Bacteriology* **192**, 1249–1258 (2010).
171. Bott, M. & Dimroth, P. *Klebsiella pneumoniae* genes for citrate lyase and citrate lyase ligase: localization, sequencing, and expression. *Molecular Microbiology* **14**, 347–356 (1994).
172. Jiang, F. *et al.* Citrate utilization under anaerobic environment in *Escherichia coli* is under direct control of Fnr and indirect control of ArcA and Fnr via CitA-CitB system. *Environmental Microbiology* **00** (2020).
173. Kerfeld, C. A., Aussignargues, C., Zarzycki, J., Cai, F. & Sutter, M. Bacterial microcompartments. *Nature Reviews Microbiology* **16**, 277–290 (2018).

174. Zarzycki, J., Erbilgin, O. & Kerfeld, C. A. Bioinformatic characterization of glycy radical enzyme-associated bacterial microcompartments. *Applied and Environmental Microbiology* **81**, 8315–8329 (2015).
175. Shisler, K. A. & Broderick, J. B. Glycyl radical activating enzymes: Structure, mechanism, and substrate interactions. *Archives of Biochemistry and Biophysics* **546**, 64–71 (2014).
176. Forward, J. A., Behrendt, M. C., Wyborn, N. R., Cross, R. & Kelly, D. J. TRAP transporters: A new family of periplasmic solute transport systems encoded by the dctPQM genes of *Rhodobacter capsulatus* and by homologs in diverse gram-negative bacteria. *Journal of Bacteriology* **179**, 5482–5493 (1997).
177. Mulligan, C., Fischer, M. & Thomas, G. H. Tripartite ATP-independent periplasmic (TRAP) transporters in bacteria and archaea. *FEMS Microbiology Reviews* **35**, 68–86 (2010).
178. Backman, L. R., Funk, M. A., Dawson, C. D. & Drennan, C. L. New tricks for the glycyl radical enzyme family. *Critical Reviews in Biochemistry and Molecular Biology* **52**, 674–695 (2017).
179. Hägerhäll, C. Succinate : quinone oxidoreductases Variations on a conserved theme. *Biochimica et Biophysica Acta* **1320**, 107–141 (1997).
180. Kröger, A. *et al.* Fumarate respiration of *Wolinella succinogenes*: Enzymology, energetics and coupling mechanism. *Biochimica et Biophysica Acta - Bioenergetics* **1553**, 23–38 (2002).
181. Coppi, M. V. *et al.* Involvement of *Geobacter sulfurreducens* SfrAB in acetate metabolism rather than intracellular, respiration-linked Fe(III) citrate reduction. *Microbiology* **153**, 3572–3585 (2007).
182. Malki, S. *et al.* Characterization of an operon encoding an NADP-reducing hydrogenase in *Desulfovibrio fructosovorans*. *Journal of Bacteriology* **177**, 2628–2636 (1995).
183. Kpebe, A. *et al.* A new mechanistic model for an O<sub>2</sub>-protected electron-bifurcating hydrogenase, Hnd from *Desulfovibrio fructosovorans*. *Biochimica et Biophysica Acta - Bioenergetics* **1859**, 1302–1312 (2018).
184. Kaufmann, F. & Lovley, D. R. Isolation and characterization of a soluble NADPH-dependent Fe(III) reductase from *Geobacter sulfurreducens*. *Journal of Bacteriology* **183**, 4468–4476 (2001).
185. Inoue, M. *et al.* Structural and phylogenetic diversity of anaerobic carbon-monoxide dehydrogenases. *Frontiers in Microbiology* **10**, 1–14 (2019).
186. Rothery, R. A., Workun, G. J. & Weiner, J. H. The prokaryotic complex iron-sulfur molybdoenzyme family. *Biochimica et Biophysica Acta - Biomembranes* **1778**, 1897–1929 (2008).
187. Zhang, Y. & Gladyshev, V. N. General trends in trace element utilization revealed by comparative genomic analyses of Co, Cu, Mo, Ni, and Se. *Journal of Biological Chemistry* **285**, 3393–3405 (2010).

188. Greening, C. *et al.* Genomic and metagenomic surveys of hydrogenase distribution indicate H<sub>2</sub> is a widely utilised energy source for microbial growth and survival. *ISME Journal* **10**, 761–777 (2016).
189. Abreu, I. A. *et al.* A novel iron centre in the split-Soret cytochrome c from *Desulfovibrio desulfuricans* ATCC 27774. *Journal of Biological Inorganic Chemistry* **8**, 360–370 (2003).
190. Da Silva, S. M., Pacheco, I. & Pereira, I. A. Electron transfer between periplasmic formate dehydrogenase and cytochromes c in *Desulfovibrio desulfuricans* ATCC 27774. *Journal of Biological Inorganic Chemistry* **17**, 831–838 (2012).
191. Oliveira, T. F. *et al.* The crystal structure of *Desulfovibrio vulgaris* dissimilatory sulfite reductase bound to DsrC provides novel insights into the mechanism of sulfate respiration. *Journal of Biological Chemistry* **283**, 34141–34149 (2008).
192. Pires, R. H. *et al.* Characterization of the *Desulfovibrio desulfuricans* ATCC 27774 DsrMK-JOP complex - A membrane-bound redox complex involved in the sulfate respiratory pathway. *Biochemistry* **45**, 249–262 (2006).
193. Findlay, A. J. & Kamyshny, A. Turnover rates of intermediate sulfur species ( $Sx_2^-$ ,  $S_0$ ,  $S_2O_{32}^-$ ,  $S_4O_{62}^-$ ,  $SO_{32}^-$ ) in anoxic freshwater and sediments. *Frontiers in Microbiology* **8** (2017).
194. Gerritse, J. *et al.* Influence of different electron donors and acceptors on dehalorespiration of tetrachloroethene by *Desulfitobacterium frappieri* TCE1. *Applied and Environmental Microbiology* **65**, 5212–5221 (1999).
195. Hartmann, T. & Leimkühler, S. The oxygen-tolerant and NAD<sup>+</sup>-dependent formate dehydrogenase from *Rhodobacter capsulatus* is able to catalyze the reduction of CO<sub>2</sub> to formate. *FEBS Journal* **280**, 6083–6096 (2013).
196. Studenik, S., Vogel, M. & Diekert, G. Characterization of an O-Demethylase of *Desulfitobacterium hafniense* DCB-2. *Journal of Bacteriology* **194**, 3317–3326 (2012).
197. Ticak, T., Kountz, D. J., Girosky, K. E., Krzycki, J. A. & Ferguson, D. J. A nonpyrrolysine member of the widely distributed trimethylamine methyltransferase family is a glycine betaine methyltransferase. *PNAS* **111**, E4668–E4676 (2014).
198. Holliger, C., Schraa, G., Stams, A. J. M. & Zehnder, A. J. B. A Highly Purified Enrichment Culture Couples the Reductive Dechlorination of Tetrachloroethene to Growth. *Applied and Environmental Microbiology* **59**, 2991–2997 (1993).
199. Boynton, Z. L., Bennett, G. N. & Rudolph, F. B. Cloning, sequencing, and expression of clustered genes encoding  $\beta$ -hydroxybutyryl-coenzyme A (CoA) dehydrogenase, crotonase, and butyryl-CoA dehydrogenase from *Clostridium acetobutylicum* ATCC 824. *Journal of Bacteriology* **178**, 3015–3024 (1996).
200. Garcia Costas, A. M. *et al.* Defining Electron Bifurcation in the Electron-Transferring Flavoprotein Family. *Journal of Bacteriology* **199**, 1–14 (2017).

201. Seedorf, H. *et al.* The genome of *Clostridium kluyveri*, a strict anaerobe with unique metabolic features. *Proceedings of the National Academy of Sciences of the United States of America* **105**, 2128–2133 (2008).
202. Peters, J. W. *et al.* [FeFe]- and [NiFe]-hydrogenase diversity, mechanism, and maturation. *Biochimica et Biophysica Acta - Molecular Cell Research* **1853**, 1350–1369 (2015).
203. Welte, C. & Deppenmeier, U. Bioenergetics and anaerobic respiratory chains of acetlastic methanogens. *Biochimica et Biophysica Acta - Bioenergetics* **1837**, 1130–1147 (2014).
204. Prat, L., Maillard, J., Rohrbach-Brandt, E. & Holliger, C. An unusual tandem-domain rhodanese harbouring two active sites identified in *Desulfitobacterium hafniense*. *FEBS Journal* **279**, 2754–2767 (2012).
205. Chai, Y., Kolter, R. & Losick, R. A widely conserved gene cluster required for lactate utilization in *Bacillus subtilis* and its involvement in biofilm formation. *Journal of Bacteriology* **191**, 2423–2430 (2009).
206. Burschel, S. *et al.* Iron-sulfur cluster carrier proteins involved in the assembly of *Escherichia coli* NADH:ubiquinone oxidoreductase (complex I). *Molecular Microbiology* **111**, 31–45 (2019).
207. Rudinskiy, N. *et al.* Amyloid-beta oligomerization is associated with the generation of a typical peptide fragment fingerprint. *Alzheimer's and Dementia* **12**, 996–1013 (2016).
208. Murai, M. & Miyoshi, H. Current topics on inhibitors of respiratory complex i. *Biochimica et Biophysica Acta - Bioenergetics* **1857**, 884–891 (2016).
209. Jones, A. J. *et al.* A Self-Assembled Respiratory Chain that Catalyzes NADH Oxidation by Ubiquinone-10 Cycling between Complex I and the Alternative Oxidase. *Angewandte Chemie - International Edition* **55**, 728–731 (2016).
210. Altschul, S. F., Gish, W., Miller, W., Myers, E. W. & Lipman, D. J. Basic local alignment search tool. *Journal of Molecular Biology* **215**, 403–410 (1990).
211. Valasatava, Y., Rosato, A., Banci, L. & Andreini, C. MetalPredator: A web server to predict iron-sulfur cluster binding proteomes. *Bioinformatics* **32**, 2850–2852 (2016).
212. Svetlitchnyi, V., Peschel, C., Acker, G. & Meyer, O. Two membrane-associated NiFeS-carbon monoxide dehydrogenases from the anaerobic carbon-monoxide-utilizing eubacterium *Carboxydotherrmus hydrogenoformans*. *Journal of Bacteriology* **183**, 5134–5144 (2001).
213. Lumpio, H. L., Shenvi, N. V., Summers, A. O., Voordouw, G. & Kurtz, D. M. J. Rubrerythrin and rubredoxin oxidoreductase in *Desulfovibrio vulgaris*: A novel oxidative stress protection system. *Journal of Bacteriology* **183**, 101–108 (2001).
214. Machado, D., Andrejev, S., Tramontano, M. & Patil, K. R. Fast automated reconstruction of genome-scale metabolic models for microbial species and communities. *Nucleic Acids Research* **46**, 7542–7553 (2018).

215. Tymoshenko, S. *et al.* Metabolic Needs and Capabilities of *Toxoplasma gondii* through Combined Computational and Experimental Analysis. *PLoS Computational Biology* **11**, 1–28 (2015).
216. Salvy, P. *et al.* PyTFA and matTFA: A Python package and a Matlab toolbox for Thermodynamics-based Flux Analysis. *Bioinformatics* **35**, 167–169 (2019).
217. Hadadi, N. *et al.* Mechanistic insights into bacterial metabolic reprogramming from omics-integrated genome-scale models. *npj Systems Biology and Applications* **6** (2020).
218. Zheng, D., Constantinidou, C., Hobman, J. L. & Minchin, S. D. Identification of the CRP regulon using *in vitro* and *in vivo* transcriptional profiling. *Nucleic Acids Research* **32**, 5874–5893 (2004).

## Mathilde Stéphanie Willemin

Ch. Bois de la Fontaine 6, CH-1007 Lausanne, Switzerland

+41 79 964 50 86

mathilde.willemin@epfl.ch

### Education

---

Lausanne, Switzerland	<b>Swiss Federal Institutes of Technology, EPFL</b>
2021	<i>PhD, Laboratory for Environmental Biotechnology</i> <i>Doctoral program in Chemistry and Chemical Engineering</i>
Bern, Switzerland	<b>Bern University</b>
2016	<i>MSc, Molecular life sciences, magna cum laude</i> <i>Special qualification in Biochemistry and Chemical biology</i>
Fribourg, Switzerland	<b>Fribourg University</b>
2014	<i>BSc, Biochemistry, bilingual german/french</i> <i>Minor in environmental sciences</i>

### Research Experience

---

2016-present	<b>PhD researcher   Swiss Federal Institutes of Technology</b> Biochemical and functional characterization of proteins involved in organohalide respiration, <u>Prof. Christof Holliger</u>
2016	<b>Research assistant   Bern University</b> Finalization of MSc research, <u>Prof. André Schneider</u>
2014-2016	<b>MSc researcher   Bern University</b> Characterization of novel components involved in mitochondrial DNA segregation in <i>Trypanosoma Brucei</i> , <u>Prof. André Schneider</u>

### Publications

---

- Willemin, M. S.**, Vingerhoets, M., Holliger, C., & Maillard, J. (2020). Hybrid Transcriptional Regulators for the Screening of Target DNA Motifs in Organohalide-Respiring Bacteria. *Frontiers in Microbiology*, 11 (March), 1–11. <https://doi.org/10.3389/fmicb.2020.00310>
- Maillard, J., & **Willemin, M. S.** (2019). Regulation of organohalide respiration. In R. K. Poole (Ed.), *Adv Microb Physiol*, Vol 74 (pp. 191–238). Elsevier. <https://doi.org/10.1016/bs.ampbs.2019.02.002>
- Buttet, G. F., **Willemin, M. S.**, Hamelin, R., Rupakula, A., & Maillard, J. (2018). The Membrane-Bound C Subunit of Reductive Dehalogenases : Topology Analysis and Reconstitution of the FMN-Binding Domain of PceC. *Frontiers in Microbiology*, 9 (April), 1–15. <https://doi.org/10.3389/fmicb.2018.00755>
- Käser, S., **Willemin, M.**, Schnarwiler, F., Schimanski, B., Poveda-Huertes, D., Oeljeklau, S., Warscheid, B., Meisinger, C., Schneider, A. (2017). Biogenesis of a mitochondrial DNA inheritance machinery in the mitochondrial outer membrane of *Trypanosoma brucei*. *PLOS Pathogen*, 13(12), e1006808. <https://doi.org/10.1101/190751>

## Contributions to Scientific Meetings

---

2017/2018/2019	Swiss Society of Microbiology – Annual Congress, Zurich/Lausanne/Basel Switzerland (Oral, poster)
2019	85 <sup>th</sup> Harden Conference: Dynamic Membrane Complexes: Respiration And Transport, Bonn, Germany (Flash Poster)
2019	7 <sup>th</sup> Swiss Microbial Ecology Meeting, Lausanne, Switzerland (Poster)
2018	European Federation of Biotechnology- Electron transport processes and their regulation, Saint-Tropez, France (Flash Poster)
2017	DehaloCon II – a Conference on Anaerobic Reductive Dehalogenation, Leipzig, Germany (Oral)
2016	Swiss Trypanosomatid meeting, Leysin Switzerland (Oral)

## Teaching Experience (assistant)

---

BSc3 Environmental Chemistry: exercise sessions, EPFL, Switzerland (2019)  
 BSc2 Biology: exercise sessions, EPFL, Switzerland (2018-present)  
 BSc4 Microbiology for Engineers: practical and exercise sessions, EPFL, Switzerland (2016-present)  
 BSc3 Biochemistry lab course: *In vitro* protein import assay, University of Bern, Switzerland (2016)

## Grad Level Courses

---

Academic Writing for Doctoral Students – Language Center, EPFL, Lausanne, Switzerland (2019)  
 EMBO Practical Course: Targeted Proteomics – Experimental Design and Data analysis, Prof. Ruedi Aebersold, Center for Genomic Regulation, Barcelona, Spain (2018)  
 Chemical Biology – Prof. Christian Heinis, EPFL, Lausanne, Switzerland (2018)  
 Lecturing/Presenting in Engineering – Teaching Support Center, EPFL, Lausanne, Switzerland (2018)  
 Design of Experiment – Prof. Jean-Marie Fuerbringer, EPFL, Lausanne, Switzerland (2017)

## Additional Experience

---

2018-2019 Lab technicians training and supervision

## References

---

Prof. Christof Holliger EPFL, Lausanne christof.holliger@epfl.ch	Dr. Julien Maillard EPFL, Lausanne julien.maillard@epfl.ch	Prof. André Schneider University of Bern andre.schneider@dcb.unibe.ch
--	--	---

## Language

---

French <i>Mother tongue</i>	English <i>Fluent</i>	German <i>Basic skills</i>
--------------------------------	--------------------------	-------------------------------

## Hobbies

---

Sports ( <i>Marathon/Badminton</i> )	Brewing	Birdwatching
--------------------------------------	---------	--------------