NOVA SCHOOL OF SCIENCE & TECHNOLOGY

DEPARTMENT OF LIFE SCIENCES

Structural and functional insights on the electrifying pathways of *Geobacter sulfurreducens*

Tomás Monteiro Fernandes

Master in Biochemistry

DOCTORATE IN BIOTECHNOLOGY

NOVA University Lisbon

September, 2024

NOVA SCHOOL OF SCIENCE & TECHNOLOGY

DEPARTMENT OF LIFE SCIENCES

Structural and functional insights on the electrifying pathways of *Geobacter sulfurreducens*

Tomás Monteiro Fernandes

Master in Biochemistry

Adviser: Maria Leonor Carvalho Morgado
Assistant Investigator, NOVA School of Science and Technology, NOVA University Lisbon

Co-adviser: Carlos Alberto Gomes Salgueiro
Full Professor, NOVA School of Science and Technology, NOVA University Lisbon

Examination Committee:

Chair: Maria João Lobo de Reis Madeira Crispim Romão,
Full Professor, NOVA University Lisbon

Rapporteurs: Irene Díaz-Moreno,
Full Professor, University of Seville
Ricardo Saraiva Loureiro Oliveira Louro,
Associate Professor, NOVA University Lisbon

Adviser: Maria Leonor Carvalho Morgado,
Assistant Professor, NOVA University Lisbon

Members: Thomas Boesen,
Associate Professor, Aarhus University
Maria João Lobo de Reis Madeira Crispim Romão,
Full Professor, NOVA University Lisbon

DOCTORATE IN BIOTECHNOLOGY

NOVA University Lisbon
September, 2024

Structural and functional insights on the electrifying pathways of *Geobacter sulfurreducens*

Acknowledgments

The writing and conceptualization of a PhD thesis is perhaps the greatest academic challenge and achievement there is. How can one accordingly convey thousands of hours of blood, sweat and tears (literally, in my case – you only realize how sharp a cracked NMR tube can be once you accidentally pierce one of your fingers with one) in a singular document? No matter the effort, sheets of paper will never grasp the sheer passion, dedication and care needed to finalize this stage of a scientist's life. This remarkable journey, which I would say started long before the actual period of time I dedicated to this degree, would not have been possible without the unmeasurable help of a long list of people, which I will soon start to mention and acknowledge.

Before we start getting all emotional, I would like to use this segment as a thesis preface. Those who are brave enough to get past this section and read my actual work will realize that while it regards fundamental science, it might one day be part of the solutions required to develop more sustainable technologies. These will hopefully help mitigate the impacts of climate change, including the rise of global temperatures, extreme weather events and loss of biodiversity, contributing for the definitive establishment of renewable energy sources and efficient management of our planet's resources. Therefore, while my Thesis contains a considerable volume of text, figures and references, which will necessarily result in environmentally disastrous printed hardcopies, I have dedicated some time to making it as much user-friendly and navigable as possible for electronic readers. The navigation panel will allow you to select the specific section you wish to read, while the different hyperlinks I have placed in all figures, tables, equations and sections mentions will help you connect the different segments of the text. Once you access a specific hyperlink, you can always switch between your views by using the command "Alt + left/right arrow". This will not only allow you to navigate the document smoothly, but hopefully give you second thoughts about getting a printed version. Moreover, I would like to state that while a thesis is an individual work, I believe that mine is the result of the support and collaboration of a great group of people. Therefore, I chose to use "we" throughout this document to acknowledge the collective effort and invaluable contributions of my peers.

Now that this small but important message is conveyed, I ask you to get your tissues out, especially if you are the kind of person to get emotional reading thesis' acknowledgments.

First, I would like to thank my supervisor, Doctor Leonor Morgado. I solemnly hope that none of these words come as a surprise to you and that I have been able to show my gratitude accordingly throughout these 7 years of work and friendship. Your endless patience, availability, work ethic, dedication, knowledge, expertise and perfectionism always strove me to be the best possible version of myself. Thank you for guiding me through the most difficult and frustrating times, and for keeping my feet on the ground whenever I entered my "Icarus moods", and wanted to change the world with an NMR tube. Science has a lot to do with passion, but a rather responsible one. You have taught me to ask questions and seek answers, but also to never take them as definitive. You are the main reason I am the scientist that I am today, and for that I am deeply and forever grateful to you. Thank you for paving the way for all the experiences and opportunities I was able to live through, and for always believing in me. I know that I have a great friend in you and that our journey is far from over.

This is not a goodbye, but rather a "see you soon". I surely hope so, and wish you all the best in all aspects of your life. You deserve it.

Secondly, I would like to thank my co-supervisor and long term "boss", Professor Carlos Salgueiro, without whom I would not be writing this Thesis. Professor, I will never forget everything you did for me. I remember the day you accepted me to do my bachelor final project almost 10 years ago. At the time, my interest arose not necessarily because of the topic, but very much because of you. Your posture and simple taking on science. Your clear explanations (using humor to simplify what others would complicate), personality and amicability, were the great catalyzers to what I think was and still is a very successful journey and partnership. I hope you feel the same way. I have a tremendous admiration for your professionalism, dedication and work ethic. I will never forget our discussions about our Sporting Clube de Portugal, a great passion shared beyond science and that brought us great happiness in recent times. I hope it goes on for a long time, for both our sakes. Thank you for everything you taught me and for always giving me the tools I needed to be successful in science and happy outside of it. As I said to Dr. Leonor, I hope this is not a goodbye. You can always count on me for anything you need.

Moreover, I would like to acknowledge Dr. Aldino Viegas for not only having accepted to be part of my thesis advisory committee, but also for all the indispensable help with the acquisition, processing and analysis of the 3D NMR experiments required to determine the solution structure of ExtJ. Thank you for your valuable insights, patience and expertise. It was a pleasure to learn from you and I only wish we would have had more time, so that I could learn even more.

To Dr. Arménio Barbosa, thank you for introducing me to the "magic" world of molecular dynamics simulations. I am grateful for your guidance and expertise. For the patience to hear me debate about all my theories, which, for the most part, were unfortunately later proven wrong. I guess that is what science is. Also, for your kindness and attention to detail. I really enjoyed learning from you, and rest assured, I will carry this knowledge anywhere I go. I would also like to acknowledge Prof. Cecília Roque for receiving me in her lab and for allowing me to learn from Dr. Arménio. Thank you.

To Dr. Cristiano Mota, thank you for your availability to help me uncover the mystery behind the homodimerization of ExtJ. Our last-minute collaboration was intense, efficient and enriching. I found it really easy to work with you, mostly because of your incredible knowledge and proficiency. Thanks to you, I now have a clearer picture of the information we can obtain from SAXS. Considering how successful and simple it was to unravel a crucial result of my Thesis, it is now definitely on my list of "go-to" techniques. Thank you, Dr. Cristiano. For your knowledge, kindness and efficiency. I would also like to acknowledge Prof. Maria João Romão for allowing my ExtJ samples to be sent to the ESRF in Grenoble, and to learn from Dr. Cristiano. Thank you.

To Dr. Filipe Folgosa, thank you for receiving me in ITQB and for taking the time to introduce me to EPR. I really enjoyed learning from you and always felt that we would work great as a team. GSU0105 was quite a challenging system, and your expertise was a key contribution for us to unravel some of its secrets. I would also like to acknowledge Prof. Miguel Teixeira not only for receiving me in his lab, but also for his expertise, crucial for the publication of the GSU0105's work.

To Prof. David Turner, for the assistance with the modified sequential model used to fit the redox titration curves of GSU0105 and for all the valuable suggestions for that work.

To Prof. Oliver Einsle, it was an honour and pleasure to have you as member of my thesis advisory committee. Thank you for introducing me to AlphaFold and for showing me how to easily predict the models of all my target proteins. The introduction of this tool was paramount for this Thesis, and for that, I am really grateful. I would also like to thank you for your suggestions and advice over the years, which significantly enriched my research and academic growth. Vielen Dank für alles.

To Elisabete Ferreira, for all the help with the acquisition of the CD, DSC and MST data. For your expertise and advice. For your patience whenever I asked for the smooth versions of the CD data sets. For always being kind and having a smile on your face. You have likely contributed to dozens of BSc projects, MSc and PhD thesis. I really hope you have been mentioned in all of them, because you really deserve it. Thank you for everything and let's hope you can continue to be mentioned in thesis such as this one for many more years.

To all the lab members of the Biochemistry and Bioenergetics of Heme Proteins group, the 611, I will never forget anyone who crossed my path in this beautiful journey. It was a pleasure to learn and share knowledge with all of you. To share victories and defeats. To work alongside you when the sun was at its peak, and when winter was at its coldest. During times of happiness, and times of sadness. I really look back to us as a family, with all the flaws and issues that families have, but which in the end are forgotten, and replaced by the great memories we have cherished together. A huge thank you to Dr. Marta Silva, Dr. Ana Fernandes, Dr. Joana Dantas, Dr. Liliana Teixeira, Dr. Marisa Ferreira, soon to be Dr. Pilar Portela, Bruno Mendes, Jorge Antunes, João Martinho, Alexandre Almeida, and Mafalda Fernandes. You do not realize how important your simple presence, energy and chatting was for me throughout the years. Thank you for everything.

To all the BSc students who passed through this great lab while I was here, thanks to all of you. You ignite the spark of our lab every year and are part of a crucial cycle of different scientific generations. Now that I am leaving, I sincerely hope that you are fortunate enough to have a path as happy and as successful as I was able to have. In particular, I would like to thank Dalila Pedro for trusting me to be her supervisor in her final BSc project. It was a pleasure to teach you and to push you. You never stopped fighting and I was really proud of the final result. A considerable part of the ExtJ chapter is there to show thanks to your work, as well. Thank you and good luck pursuing all your dreams.

I would also like to acknowledge the NOVA School of Science and Technology and its professors, assistants and other staff that played a role in my academic path (which started back in 2013), for all the opportunities and experiences that contributed to my education as a man and as a scientist. To all my BSc, MSc and PhD colleagues. You were very important for this journey. All in all, thank you very much, glorious FCT.

Now, to my friends. I am fortunate enough to be surrounded by inspiring people that enrich me every day with their knowledge, grit, courage, experience and love. I cannot thank you enough Tomás Ribeiro, Marcelo Domingues, Carolina Costa, André Silva, Maria Sena-Esteves, Renato Domingues,

João Bernardo, Tomás Candeias, Beatriz Batista and Fábio Teles. Thank you for all the great moments and for always being there whenever I needed.

A special thank you to Tomás Ribeiro, Marcelo Domingues and André Silva for getting me out of the toughest moments of this journey. You guys are the best. Please know that you can count on me for anything.

The list of friends whom I must acknowledge continues – Andreia Faustino, Ana Rita Bello, Ana Rita Ayala, Ana Teresa, António Seixas, Beatriz Ramalho, Francisco Henriques, Guilherme Ilunga, Inês Franco, Paulo Santos, and Rúben Rodrigues. Look how far we have come. Doctors, engineers, nurses, consultants, parents. You guys are family and while we are not together as often as we would like, it always feels that we are still hanging out every day when I am with you, just like we did in highschool. Thank you for allowing me to have this sort of friendship, which I will nurture for the rest of my life. I love you all.

To my direct family, which I proudly say it is really big. I want to thank all my grandparents, uncles, aunts and cousins. My grandmothers, Ana Paula, Fernanda, Maria Ivone, and Luísa. My grandfathers, Sebastião, António, and José. My aunts, Inês, Paula, Maria, Marta, Karline, Cláudia, Sofia, and Ana. My uncles, Nuno Rosinha, Nuno Pires, Nuno Martins, Nuno Figueira, Henrique, Luís, Pedro Carvalho and Pedro Magarreiro. My cousins Gonçalo, Afonso, Rita, Maria, Diogo, Mariana Pires, Mariana Carvalho, Bernardo, Luís, Ana Clara, Francisca, Leonor, and Lourenço. My godfather, António. My godmother, Paula. To Paulo Félix and Beta. To Gui and Rafa. To Manuel Negrelli, who will always be part of my family. To its son, my friend, Hugo. To Paulo Cidra, part of our family and a friend. To his son, Rodrigo. To Tommy, Pipa, Lobo, Mike and Tita. To all of you, and your loved ones, I love you and wish you all the best, always. Thank you for being part of my life.

In particular, I want to thank my little brother and sister. Dear Constança and Salvador. I love you more than words can convey and I wish that I was 20 years younger, so that I could share my childhood with you. It breaks my heart every time you guys ask me to play a bit longer, to spend a bit more time with you, knowing that I have to go to work. But I do so, so that one day I can have a better future. A future in which I will be able to help you through anything you need. There is not a day I do not think about you. Thank you for being a source of inspiration and happiness in my life. May you always be happy, with those beautiful smiles and shiny blue eyes. I love you very much.

To Gonçalo. You know how much you mean to me. You are my oldest companion and friend. I was, am and will always be incredibly lucky to have you as a brother. You are the strongest pilar of my life, while also being my greatest weakness. I do not remember a day of my life in which I have not loved you, nor one in which I was not loved back. Our bond, beyond our shared blood, will forever be tied by thousands of shared experiences. Our happiest conquests and also our greatest defeats. May it be so forever. I would not change a thing. May you always be the happiest person in this world. I love you.

To my father. Warrior, companion, idol, friend. Thank you for everything you have always done for me. I know of all the sacrifices you did and do to ensure that your children have everything they need to be successful and happy. Thank you for allowing me to pursue my dreams. I have had the happiest of lifes, thanks to your efforts. Thank you for being the man you are. For all our shared moments of

happiness and sadness. For you care. For your presence. For your calls. You never failed me. I hope I have never failed you, as well. A paragraph of mere words can never fully measure the love I have for you. I hope one day I can repay you half of what you have given me. And that I can be for my children, the father you were to me. I love you, Pai.

To my mother, whose unwavering love and support have been the foundation of everything I have achieved in my life. From a small baby to a boy who loved to play the numbers with you, to now a man that loves you infinitely. Your endless care and love have carried me through every challenge. Your bravery and resilience have been a source of endless inspiration and admiration. From the smallest acts of kindness to the greatest sacrifices, you have always put my needs before your own. I thank you deeply for always allowing me to pursue my dreams. I will never forget your bed time stories. Your meals. Your touch. Your perfume. Your care. Your patience. Everything you have given me. I love you. And I thank you, Mãe.

To Sara. My love. The absolute most important person of this journey. The one who endured the most difficult times and with which I shared my biggest frustrations. Thank you for being an endless source of love, care, patience, comfort and friendship. Thank you for taking care of me when I was sick. For being my source of energy when I was tired. For fulfilling me when I was frustrated. For making me happy when I was the saddest. For being my guiding light. Your endless encouragement, patience, and belief carried me through every challenge. You are, by far and unquestionably, the best thing I have taken from this long academic journey. I will never forget the day we met. And I will never stop loving you. You are the one. Now that this journey is about to end, I hope we can finally start the rest of our lives. I am the happiest man in the world when I am with you. You are all I need. Thank you for everything. Amo-te muito.

As a general message, I hope I have made all the people I mentioned proud. This Thesis is the result of thousands of hours of work, and all of you had a say in it. Thank you, once again.

To Fundação para a Ciência e Tecnologia (FCT), for my PhD Scholarships (SRFH/BD/145039/2019 and COVID/BD/153449/2023) and the following research & development projects, attributed to:

Prof. Carlos Salgueiro – PTDC/BIA-BQM/31981/2017 and PTDC/BIA-BQM/4967/2020
Dr. Leonor Morgado – EXPL/BIA-BQM/0770/2021
UCIBIO – UIDP/04378/2020 and UIDB/04378/2020
i4HB – LA/P/0140/2020

The NMR spectrometers are part of the National NMR Network and are supported through FCT (ROTEIRO/0031/2013 and PINFRA/22161/2016) and cofounded by FEDER through COMPETE 2020, POCI, PORL and FCT through PIDDAC.

Abstract

Electroactive microorganisms possess the remarkable ability to transfer electrons extracellularly, enabling innovative applications in microbial electrochemical technologies (MET). Despite their widely accepted potential, the application and efficiency of these technologies is scarce, largely due to their low-current production. Subsequently, understanding the fundamental mechanisms that govern extracellular electron transfer (EET) in electroactive microorganisms, from a structural and functional perspective, is a critical step towards the optimization of these processes.

Geobacter sulfurreducens is a model electroactive Gram-negative bacterium that is associated with the highest current densities observed in MET. Genetic and proteomic studies have identified some of the main players in this bacterium's EET pathways, but the precise mechanisms underlying this complex and dynamic process are still under scrutiny. The work developed in this Thesis uses a combination of biophysical techniques and computational methods to provide insights on several players of the EET routes of *G. sulfurreducens*.

AlphaFold 2 model predictions of the porin-cytochrome complexes of this bacterium provided insights on heme axial coordination and spatial disposition, fold arrangement, function and overall assembly of these multiprotein complexes. This information was explored to define new protein production strategies.

NMR and SAXS studies showed that the small periplasmic protein ExtJ, part of the ExtHIJKL porin-cytochrome complex, forms homodimers by establishing intramolecular disulfide bridges through a cysteine residue located in its flexible C-terminal domain. The solution structure of this protein is homologous to those of flavin-binding proteins, however, the protein does not bind FMN. AlphaFold-Multimer predictions suggest that ExtJ might interact with the ExtI β-barrel porin. We postulate that the protein homodimerization is controlled by cellular mechanisms of disulfide bridge formation and cleavage, which concomitantly regulate the formation of the transient ExtI-ExtJ complex and the flow of small molecules through this porin.

The extracellular cytochrome PgcA possesses a fuzzy structure with three cytochrome domains linked by unstructured stretches. Biophysical studies performed on the individual domains and the full-length cytochrome showed that the protein is able to establish a flexible electron chain, adopting multiple conformations that promote electron transfer at variable distances. This tethered diffusion mechanism, unprecedented in biological systems, might be advantageous during metal reduction.

Finally, the periplasmic triheme cytochrome GSU0105 was characterized by EPR, NMR and UV-visible spectroscopies, which showed that the protein contains one heme group that undergoes a redox state dependent low- to high-spin interconversion. Moreover, the reduction potential of this heme is modulated by the solution pH, which indirectly controls the spin state interconversion by promoting different structural conformations.

Overall, these results constitute an important contribute to the current understanding of the EET mechanisms of *G. sulfurreducens* and towards the development of efficient MET.

Keywords: *Geobacter*, Extracellular electron transfer, Cytochromes, AlphaFold, NMR

Resumo

Os microorganismos electroactivos são capazes de transferir electrões para o meio extracelular, o que permite que sejam utilizados em tecnologias electroquímicas microbianas (TEM). No entanto, apesar do seu potencial, a aplicação e eficiência destas tecnologias é ainda limitada, principalmente devido aos baixos níveis de corrente produzidos. Consequentemente, é fundamental compreender os mecanismos que governam o fenómeno de transferência electrónica extracelular (TEE) de uma perspectiva estrutural e funcional, de forma a que estes processos possam ser optimizados.

Geobacter sulfurreducens é uma bactéria Gram-negativa electroactiva modelo, sendo responsável pelas maiores densidades de corrente observadas em TEM. Os componentes mais relevantes das vias de TEE desta bactéria já foram identificados por estudos genómicos e proteómicos, no entanto, os mecanismos subjacentes a este processo não foram clarificados. O trabalho desenvolvido durante esta Tese resulta da aplicação de uma combinação de técnicas biofísicas e métodos computacionais que permitiram o estudo de vários componentes das vias de TEE de *G. sulfurreducens*.

Os modelos dos complexos porina-citocromo desta bactéria foram previstos pelo AlphaFold 2, o que permitiu obter informação sobre a coordenação axial e disposição espacial dos hemos, a função, o enrolamento e a organização geral destes complexos. Esta informação serviu para definir novas estratégias de produção destas proteínas.

Estudos de RMN e SAXS demonstraram que a proteína periplasmática ExtJ, parte do complexo porina-citocromo ExtHIJKL, forma homodímeros através de pontes dissulfureto intramoleculares, utilizando um resíduo de cisteína localizado no domínio flexível do C-terminal. A sua estrutura em solução indica que a proteína é homóloga a proteínas que ligam moléculas de flavina, no entanto, a proteína não interage com FMN. Previsões do AlphaFold-Multimer sugerem que a ExtJ poderá interagir com a porina ExtI. Com base nestes resultados, postula-se que o processo de dimerização possa ser controlado por mecanismos de quebra e formação de pontes dissulfureto, que por sua vez controlam a formação do complexo ExtI-ExtJ e a passagem de moléculas através desta porina.

O citocromo extracelular PgcA possui uma estrutura difusa com três domínios citocromo ligados por secções desestruturadas. Estudos biofísicos dos domínios individuais e do citocromo completo demonstram que a proteína estabelece uma cadeia electrónica flexível, adoptando múltiplas conformações que promovem transferência electrónica a várias distâncias. O mecanismo descrito, sem precedentes em sistemas biológicos, poderá ser vantajoso durante a redução de metais.

Finalmente, o citocromo trihémico periplasmático GSU0105 foi caracterizado por espectroscopias de RPE, RMN e UV-visível, tendo estas demonstrado que a proteína contém um grupo hemo que sofre uma conversão de spin-baixo para spin-alto dependente do estado redox. Ademais, o potencial de redução deste hemo é modulado pelo pH da solução, que controla indirectamente a conversão do estado de spin, ao promover diferentes conformações estruturais.

Globalmente, estes resultados constituem um contributo importante para o nosso conhecimento actual dos mecanismos de TEE de *G. sulfurreducens* e para o desenvolvimento de TEM mais eficientes.

Palavras-chave: *Geobacter*, Transferência electrónica extracelular, Citocromos, AlphaFold, RMN

Table of contents

Acknowledgments vii
Abstract xiii
Resumo xv
Table of contents xvii
List of figures xxi
List of tables xxv
Abbreviations, symbols and constants xxvii

Introduction

1.1. Microbes as the pioneers of Life's evolution 35
1.2. Electroactive microorganisms 37
1.2.1. Extracellular electron transfer in bacteria 40
1.3. Microbial electrochemical technologies 41
1.4. *Geobacter* bacteria 45
1.4.1. *Geobacter sulfurreducens* 46
1.4.2. Extracellular electron transfer pathways 47
1.4.3. Electron harvesting pathways 57
1.5. Cytochromes – The chemistry of the key players in extracellular electron transfer 59
1.6. Protein engineering of multiheme cytochromes 62
1.7. Objectives and thesis outline 66
1.8. References 67

Unfolding the outer gates of extracellular electron transfer in *Geobacter* with AlphaFold

2.1. Summary 93
2.2. Introduction 95
2.2.1. The protein folding problem 95
2.2.2. The impact and applications of AlphaFold 96
2.2.3. Structural insights on porin-cytochrome complexes 97
2.3. Results and discussion 99
2.3.1. OmabcB and OmabcC 99
2.3.2. ExtABCD 105
2.3.3. ExtEFG 110
2.3.4. ExtHIJKL 114
2.4. Using structural models to optimize protein production 122
2.4.1. "Blind" protein-cloning 122
2.4.2. AlphaFold-guided protein cloning 124
2.4.3. Preliminary expression tests of target proteins 126
2.5. Conclusions 128

2.6. Materials and methods 131
2.6.1. Protein targets and putative porin-cytochrome complexes model prediction 131
2.6.2. RF cloning of the target proteins 131
2.6.3. Insertion of affinity-tags in cytochrome constructs 135
2.6.4. Production of the truncated ExtA and ExtK constructs 136
2.6.5. Protein expression tests 137
2.7. References 138

Dissecting a porin-cytochrome complex: Structural characterization of ExtJ

3.1. Summary 155
3.2. Introduction 157
3.3. Results and discussion 158
3.3.1. Amino acid sequence analysis 158
3.3.2. ExtJ forms homodimers in vitro 159
3.3.3. Preliminary structural characterization of ExtJ 160
3.3.4. NMR solution structure and relaxation experiments of ExtJ 169
3.3.5. ExtJ homodimerization is driven solely by an intermolecular disulfide bridge 173
3.3.6. Narrowing down the function of ExtJ in *Geobacter sulfurreducens* 183
3.4. Conclusions 191
3.5. Materials and methods 193
3.5.1. DNA manipulation 193
3.5.2. Protein sequence analysis and AlphaFold predictions 193
3.5.3. Protein expression and purification 194
3.5.4. NMR spectroscopy 195
3.5.5. CD spectroscopy 199
3.5.6. Molecular dynamics simulations 199
3.5.7. Microscale thermophoresis 200
3.5.8. SAXS experiments 201
3.6. References 202

Disentangling a tangled cytochrome: Functional mechanisms of PgcA

4.1. Summary 217
4.2. Introduction 219
4.3. Results and discussion 220
4.3.1. PgcA and homologs establish a new class of cytochromes 220
4.3.2. PgcA's cytochrome domains are composed of α-helical structures 223
4.3.3. PgcA contains an intramolecular redox chain 227
4.3.4. NMR features of PgcA's cytochrome domains 228
4.3.5. PgcA's cytochrome domains exchange electrons promiscuously 235

4.3.6. The cytochrome domains of PgcA reduce several electron acceptors 239
4.3.7. Proline-threonine stretches are not responsible for Fe(III) oxides binding in PgcA 240
4.3.8. Domain 3 is intramolecularly recharged for continuous Fe(III) oxides reduction 242
4.4. Conclusions 244
4.5. Materials and methods 247
4.5.1. DNA manipulation 247
4.5.2. Protein sequence and structural homology analysis 247
4.5.3. Protein expression and purification 248
4.5.4. CD spectroscopy 249
4.5.5. DSC experiments 249
4.5.6. Redox titrations followed by visible spectroscopy 250
4.5.7. Electron transfer experiments with terminal electron acceptors 250
4.5.8. Fe(III) oxides binding experiments 251
4.5.9. NMR spectroscopy 251
4.6. References 254

Depicting a unique triheme cytochrome: Biophysical studies of GSU0105

5.1. Summary 263
5.2. Introduction 265
5.3. Results and discussion 266
5.3.1. Amino acid sequence analysis of GSU0105 266
5.3.2. Spectroscopic characterization of GSU0105 268
5.3.3. Redox properties of GSU0105 275
5.4. Conclusions 279
5.5. Materials and methods 281
5.5.1. DNA manipulation 281
5.5.2. Protein sequence analysis and AlphaFold predictions 281
5.5.3. Protein expression and purification 282
5.5.4. CD spectroscopy 283
5.5.5. EPR spectroscopy 283
5.5.6. NMR spectroscopy 284
5.5.7. Redox titrations followed by visible spectroscopy 284
5.6. References 287

Conclusions and future perspectives

6.1. References 300

Appendix

7.1. AlphaFold models with pLDDT score coloring 305
7.2. Supplementary tables 310
7.3. MD simulations – Additional plots 327
7.4. PyMol script for insertion of *c*-type hemes in AlphaFold models 328
7.5. References 330

List of figures

Introduction

Figure 1.1. Current view of the tree of life. 36
Figure 1.2. Diversity of electroactive microorganisms. 38
Figure 1.3. Morphology of cable bacteria. 39
Figure 1.4. Respiratory chains of Gram-negative bacteria. 40
Figure 1.5. Protein nanowire device and electric output from ambient humidity. 42
Figure 1.6. Examples of microbial electrochemical technologies. 43
Figure 1.7. Engineering of a synthetic electron conduit in living cells. 44
Figure 1.8. Structural features of cytochrome *c* peroxidases CcpA and MacA. 46
Figure 1.9. Extracellular electron transfer pathways of *G. sulfurreducens*. 48
Figure 1.10. AlphaFold 2 models of the quinol dehydrogenases ImcH, CbcL and CbcBA. 49
Figure 1.11. Structures of periplasmic cytochromes of *G. sulfurreducens*. 50
Figure 1.12. Structure of the entire HDH assembly of *K. stuttgartiensis*. 51
Figure 1.13. Cryo-EM structures of the cytochrome nanowires of *G. sulfurreducens*. 52
Figure 1.14. Electron nanowire forms the central spine of the HDCR filament. 53
Figure 1.15. Structural comparison of the four-heme arrays of different cytochrome nanowires. 54
Figure 1.16. Structural properties of the PilA-N-C pili. 56
Figure 1.17. Schematic models of PilA and cytochrome nanowire secretion. 57
Figure 1.18. Periplasmic cytochromes crucial for the electron harvesting pathways. 58
Figure 1.19. Chemical structures of different heme groups. 60
Figure 1.20. Rational design of PpcA mutants. 63
Figure 1.21. Scheme of the preparation of *G. sulfurreducens* strains with mutated cytochromes. 64
Figure 1.22. Simplified *Geobacter* EET pathways towards electrodes. 65

Unfolding the outer gates of extracellular electron transfer in *Geobacter* with AlphaFold

Figure 2.1. Model architecture of AlphaFold 2. 96
Figure 2.2. Bacterial porin-cytochromes. 98
Figure 2.3. Structural features of cytochromes OmaB/C. 100
Figure 2.4. Relative arrangement of OmaB/C domains predicted by AlphaFold. 101
Figure 2.5. Structural features of the OmbB/C transmembrane β-barrel porins. 101
Figure 2.6. Structural features of the OmcB and OmcC dodecaheme *c*-type cytochromes. 102
Figure 2.7. Structural homology between Ccnr and OmcB/C. 103
Figure 2.8. AlphaFold-Multimer models of OmabcB/C complexes. 104
Figure 2.9. Structural features of ExtA. 105
Figure 2.10. Structural conservation of tetraheme domains of ExtA. 106
Figure 2.11. Structural features of ExtB transmembrane β-barrel porin. 107
Figure 2.12. Structural features of ExtC. 107

Figure 2.13. Extended and compacted models of the ExtD hexaheme cytochrome. 108
Figure 2.14. AlphaFold-Multimer models of the ExtBCD complex. 109
Figure 2.15. Structural features of ExtE transmembrane β-barrel porin. 110
Figure 2.16. ExtC and ExtF are structural homologs. 111
Figure 2.17. ExtG contains multiple non-canonical heme binding motifs. 112
Figure 2.18. AlphaFold-Multimer models of the ExtEFG complex. 113
Figure 2.19. ExtH is a group II rhodanese-like protein. 115
Figure 2.20. Structural features of the ExtI β-barrel porin. 116
Figure 2.21. Structural homology of ExtJ with riboflavin synthase-like β-barrel proteins. 117
Figure 2.22. Structural features of the cytochrome ExtK and the hypothetical protein ExtL. 118
Figure 2.23. Structure of the ExtKL tandem cytochrome. 119
Figure 2.24. Structural homologs of ExtKL. 120
Figure 2.25. AlphaFold-Multimer model of the ExtHIJKL complex. 121
Figure 2.26. AlphaFold model of PgcA. 130

Dissecting a porin-cytochrome complex: Structural characterization of ExtJ

Figure 3.1. Amino acid sequence of the pET-28a(+)-ExtJ construct. 157
Figure 3.2. Amino acid sequence alignment of ExtJ with homolog sequences. 158
Figure 3.3. ExtJ homodimer formation. 160
Figure 3.4. 1D ^{1}H-NMR spectrum of ExtJ WT. 161
Figure 3.5. 2D ^{1}H, ^{15}N-HSQC spectrum of ExtJ WT. 162
Figure 3.6. NMR sequential assignment and structural features of ExtJ. 163
Figure 3.7. Structural models of ExtJ. 164
Figure 3.8. Structural impact of the C81* mutation and the N-terminal affinity tag. 165
Figure 3.9. Far-UV CD spectral features of ExtJ. 167
Figure 3.10. Solution structure of ExtJ. 170
Figure 3.11. ^{15}N relaxation parameters of ExtJ. 172
Figure 3.12. 2D ^{1}H, ^{15}N-HSQC spectra of ExtJ in the presence and absence of 1 mM DTT. 174
Figure 3.13. Plot of rotational correlation time as a function of protein molecular weight. 175
Figure 3.14. Electrostatic surface of the homodimer. 176
Figure 3.15. MD simulation of the ExtJ homodimer in explicit water. 176
Figure 3.16. MST measurements of the monomer-dimer equilibrium of ExtJ. 178
Figure 3.17. SAXS studies of ExtJ. 180
Figure 3.18. Relative orientation of the riboflavin synthase domains and the ExtJ homodimer. 184
Figure 3.19. Homodimer structures of the N-terminal domain of riboflavin synthase. 184
Figure 3.20. Binding pocket conservation in riboflavin synthase and homologs. 185
Figure 3.21. Structural conservation of the riboflavin binding pocket. 186
Figure 3.22. Monitorization of ExtJ-FMN interaction by NMR. 187
Figure 3.23. Monitorization of ExtJ-PpcA interaction by 1D ^{1}H-NMR. 188

Figure 3.24. Monitorization of ExtJ-PpcA interaction by 2D ^{1}H, ^{15}N-HSQC. 189
Figure 3.25. Structural contextualization of ExtJ in the ExtHIJKL complex. 190
Figure 3.26. AlphaFold 3 model predictions of putative ExtJ-FAD complexes. 192

Disentangling a tangled cytochrome: Functional mechanisms of PgcA

Figure 4.1. AlphaFold models of PgcA and homologs. 220
Figure 4.2. Structural features of PgcA. 221
Figure 4.3. Structural depiction of a model triheme cytochrome. 222
Figure 4.4. Beads on a string like arrangement of PgcA. 223
Figure 4.5. Biochemical and spectroscopic features of the cytochrome domains of PgcA. 224
Figure 4.6. Structural homology between OmcF and the cytochrome domains of PgcA. 225
Figure 4.7. CD spectral deconvolution and unfolding profiles of PgcA. 226
Figure 4.8. 1D ^{1}H-NMR spectra of the cytochrome domains of PgcA. 229
Figure 4.9. Temperature dependence of the heme methyl chemical shifts.... 230
Figure 4.10. Diagram of a *c*-type heme. 231
Figure 4.11. 2D ^{1}H, ^{13}C-HMQC spectra of the cytochrome domains of PgcA. 232
Figure 4.12. 2D ^{1}H-EXSY NMR spectra of cytochrome domains 1 and 2 of PgcA. 234
Figure 4.13. 2D ^{1}H-EXSY NMR spectrum of cytochrome domain 3 of PgcA. 235
Figure 4.14. Chemical shift perturbation studies. 236
Figure 4.15. Electron transfer reactions monitored by NMR. 237
Figure 4.16. Role of PgcA in the extracellular electron transfer pathways of *G. sulfurreducens*. 238
Figure 4.17. Electron transfer experiments with electron acceptors. 240
Figure 4.18. Binding experiments between PgcA and Fe(III) oxides. 241
Figure 4.19. Electrostatic surface of the cytochrome domains of PgcA and akageneite. 242
Figure 4.20. Monitorization of Fe(III) oxides reduction by full-length PgcA. 243
Figure 4.21. Proposed model of extracellular electron transfer in *G. sulfurreducens*. 245

Depicting a unique triheme cytochrome: Biophysical studies of GSU0105

Figure 5.1. Amino acid sequence analysis of GSU0105. 266
Figure 5.2. Dendogram of representative families of periplasmic multiheme cytochromes. 268
Figure 5.3. Far-UV CD spectral features of oxidized GSU0105. 269
Figure 5.4. Structure of PpcD from *G. sulfurreducens*. 270
Figure 5.5. Temperature-induced molar ellipticity variation. 271
Figure 5.6. UV-visible spectral features of cytochrome GSU0105. 272
Figure 5.7. EPR and NMR features of GSU0105. 273
Figure 5.8. Redox titrations of GSU0105 and PpcA-family cytochromes. 276
Figure 5.9. AlphaFold models of GSU0105. 280
Figure 5.10. Electronic distribution scheme for a triheme cytochrome. 285

Conclusions and future perspectives

Figure 6.1. Updated view of the EET pathways of *G. sulfurreducens*. 299

Appendix

Figure 7.1. AlphaFold models represented in Chapter 1. 305
Figure 7.2. AlphaFold models of Cyc2 and PioAB. 305
Figure 7.3. AlphaFold models of OmabcB/C porin-cytochrome complexes. 306
Figure 7.4. AlphaFold models of ExtABCD porin-cytochrome complex. 306
Figure 7.5. AlphaFold models of ExtEFG porin-cytochrome complex. 307
Figure 7.6. AlphaFold models of ExtHIJKL porin-cytochrome complex 307
Figure 7.7. AlphaFold models of the ExtJ homodimer, ExtIJ and ExtJKL complexes. 308
Figure 7.8. AlphaFold 3 models of monomer and homodimer ExtJ with FAD. 308
Figure 7.9. AlphaFold 2 and 3 models of GSU0105. 308
Figure 7.10. AlphaFold models of different quinol hydrogenases of *G. sulfurreducens*. 309
Figure 7.11. Physical parameter plots for the MD simulation of the ExtJ homodimer. 327

List of tables

Unfolding the outer gates of extracellular electron transfer in *Geobacter* with AlphaFold

Table 2.1. List of residues of the target proteins to clone. 123
Table 2.2. List of AlphaFold-guided constructs..... 125
Table 2.3. Optimized expression conditions for the cytochrome targets and ExtJ. 127
Table 2.4. List of constructs available in the laboratory..... 129
Table 2.5. Sequences of the RF cloning primers used to produce the vectors. 132
Table 2.6. PCR cycling conditions for the RF cloning of the target proteins. 133
Table 2.7. Optimization of the RF cloning protocol conditions for each target protein. 134
Table 2.8. Sequences of the primers used to produce the vectors containing affinity tags. 135
Table 2.9. Sequences of the primers used to produce the truncated ExtA and ExtK vectors. 136

Dissecting a porin-cytochrome complex: Structural characterization of ExtJ

Table 3.1. Secondary structural elements of ExtJ determined by BeStSel and AlphaFold. 168
Table 3.2. NMR and refinement statistics of ExtJ monomer structure..... 171
Table 3.3. SAXS data processing. 182
Table 3.4. Sequences of the primers used to produce the ExtJ C81* vector. 193

Disentangling a tangled cytochrome: Functional mechanisms of PgcA

Table 4.1. Secondary structural elements of the cytochrome domains of PgcA. 227
Table 4.2. Sequences of the primers used to produce the PgcA cytochrome domains vectors. 247

Depicting a unique triheme cytochrome: Biophysical studies of GSU0105

Table 5.1. Secondary structural elements of GSU0105 and the PpcA-family cytochromes. 270
Table 5.2. Comparison of the E_{app} values of GSU0105 and the PpcA-family cytochromes. 277
Table 5.3. Sequences of the primers used to produce the GSU0105 WT and mutant vectors. 281

Appendix

Table 7.1. Assignment of the ^{1}H, ^{13}C and ^{15}N resonances of ExtJ. 310
Table 7.2. SAXS data collection. 323
Table 7.3. NMR signals of PgcA's heme substituents and axial ligands in the reduced state. 323
Table 7.4. NMR signals of PgcA's heme substituents and axial ligands in the oxidized state. 324
Table 7.5. Redox-active windows and E_{app} values of cytochromes from *G. sulfurreducens*. 325
Table 7.6. Sequences of the primers used to produce the additional GSU0105 mutants. 326

Abbreviations, symbols and constants

1D – One dimensional

2D – Two dimensional

2xYT – 2x yeast extract-tryptone medium

3D – Three dimensional

ADP – Adenosine diphosphate

AF(2/3) – AlphaFold (2/3)

AMPS – AMBER-based portal server

ARTINA – Artificial intelligence for nuclear magnetic resonance applications

ATP – Adenosine triphosphate

AvECN – *Archaeoglobus veneficus* extracellular cytochrome nanowire

BeStSel – Beta structure selection

BLAST – Basic local alignment search tool

BMRB – Biological magnetic resonance data bank

CASP – Critical assessment of protein structure prediction

Ccm – Cytochrome *c* maturation

Ccnr – Cytochrome *c* nitrite reductase

CCO – Cytochrome *c* oxidase

CD – Circular dichroism

CF(1/2) – ColabFold (1/2)

CFT – Crystal field theory

CFU – Colony forming units

COSY – Correlation spectroscopy

CPMG – Carr-Purcell-Meiboom-Gill

CRISPR-Cas9 – Clustered regularly interspaced palindromic repeats – CRISPR-associated protein 9

Cryo-EM – Cryogenic electron microscopy

CSI – Chemical shift index

C_{Target} – Concentration of target

D – Axial component of the zero-field splitting

DB – AlphaFold protein structure database

D_{max} – Maximum dimension of the particle in solution

DMSO – Dimethyl sulfoxide

DNA – Deoxyribonucleic acid

dNTPs – Deoxynucleotides

DpnI – *Diplococcus pneumoniae* enzyme I

DSC – Differential scanning calorimetry

DSS – 4,4-dimethyl-4-silapentane-1-sulfonic acid

DSSP – Dictionary of secondary structures of proteins

DTT – Dithiothreitol

E – Rhombic component of the zero-field splitting

E/D – Rhombicity of the zero-field splitting

E_{app} – Apparent midpoint reduction potential

EDTA – Ethylenediamine tetraacetic acid

EET – Extracellular electron transfer

EPR – Electron paramagnetic resonance

ESMF/ESMFold – Evolutionary scale modeling fold

ESRF – European synchrotron radiation facility

EXSY – Exchange spectroscopy

FAD – Flavin adenine dinucleotide

F_{cold} – Relative fluorescence in the "cold" region

F_{hot} – Relative fluorescence in the "hot" region

FMN – Flavin mononucleotide

F_{norm} – Normalized fluorescence

F_{rel} – Relative fluorescence

fw – Forward

g – Electron g-factor

g_{max} – Maximum electron g-factor

GROMACS – Groningen machine for chemical simulations

HDCR – Hydrogen-dependent carbon dioxide reductase

HetNOE – Heteronuclear nuclear Overhauser effect

HMQC – Heteronuclear multiple quantum coherence

HP – High-potential

HS – High salt concentration

HSQC – Heteronuclear single quantum coherence

HT – His-tag

ICP-AES – Inductively coupled plasma – atomic emission spectroscopy

Ig – Immunoglobulin

IM – Inner membrane

IMC – Inner membrane cytochrome

IPTG – Isopropyl β-D-thiogalactopyranoside

I_{sat} – Measured peak intensity in the presence of ^{1}H saturation

IUB – International union of biochemistry

I_{unsat} – Measured peak intensity in the absence of ^{1}H saturation

IUPAC – International union of pure and applied chemistry

K_D – Dissociation constant

KEGG – Kyoto encyclopedia of genes and genomes

LB – Luria-Bertani medium

LED – Light-emitting diode

Lgt – Prolipoprotein diacylglyceryl transferase

Lnt – Apolipoprotein N-acyltransferase

LP – Low-potential

LS – Low salt concentration

LspA/SpII – Lipoprotein-specific signal peptidase

MALDI-TOF-MS – Matrix-assisted laser desorption ionization – time-of-flight – mass spectrometry

MCS – Multiple cloning site

MD – Molecular dynamics

MET – Microbial electrochemical technologies

MFC – Microbial fuel cell(s)

MQ – Menaquinone

MQH_2 – Menaquinol

MSA – Multiple sequence alignment

MST – Microscale thermophoresis

MW/MM – Molecular-weight / Molecular mass

MWCO – Molecular-weight cut-off

NAD^+ – Nicotinamide adenine dinucleotide (oxidized form)

NADH – Nicotinamide adenine dinucleotide (reduced form)

NADHD – NADH dehydrogenase

Ni-NTA – Nickel – nitriloacetic acid

NMR – Nuclear magnetic resonance

NOE – Nuclear Overhauser effect

NOESY – Nuclear Overhauser effect spectroscopy

NRMSD – Normalized root mean square deviation

OD_{600} – Optical density at 600 nm

OM – Outer membrane

OMC – Outer membrane cytochrome

OzpA – Outer membrane cytochrome Z protease

PAE – Predicted alignment error

PBS – Phosphate-buffered saline

PCC – Porin-cytochrome complex

PcECN – *Pyrobaculum calidifontis* extracellular cytochrome nanowire

PCR – Polymerase chain reaction

PDB – Protein data bank

PgcA D1/2/3 – PgcA cytochrome domain 1/2/3

pLDDT – Predicted local distance difference test

PME – Particle-Mesh Ewald

PMSF – Phenylmethylsulphonyl fluoride

PPC – Periplasmic cytochrome

ppm – Parts per million

p (r) – Distance distribution function

R^2 – Coefficient of determination

RB – Riboflavin

RCCS – Random coil chemical shifts

RF – Reduced fraction

RF cloning – Restriction-free cloning

R_g – Radius of gyration

RMSD – Root mean square deviation

ROS – Reactive oxygen species

RNA – Ribonucleic acid

rRNA – Ribosomal ribonucleic acid

rv – Reverse

S^2 – Order parameter

SAXS – Small-angle X-ray scattering

SDH – Succinate dehydrogenase

SDS-PAGE – Sodium dodecyl sulfate – polyacrylamide gel electrophoresis

SEC – Size-exclusion chromatography

SECIS – Selenocysteine insertion sequence

SEM – Scanning electron microscopy

SHE – Standard hydrogen electrode

ST – Strep-tag

STC – Small tetraheme cytochrome

T_1 – Longitudinal relaxation

T_2 – Transverse relaxation

Taq DNA polymerase – *Thermus aquaticus* thermostable DNA polymerase

TCI – Triple-resonance cryoprobe

TEM – Transmission electron microscopy

tet – Tetracycline

TEV – Tobacco Etch Virus

T_m – Melting temperature

TMBZ – 3,3',5,5'-tetramethylbenzidine

TOCSY – Total correlation spectroscopy

TR – Thrombin

Tris – Tris(hydroxymethyl)aminomethane

tRNA – Transfer ribonucleic acid

UCSF – University of California, San Francisco

UV – Ultraviolet

VMD – Visual molecular dynamics

V_p – Porod volume

W-bisPGD – Tungsten-bis-pyranopterin guanine dinucleotide

$W_{1/2/3}$ – Weighted contributions of hemes 1/2/3

WT – Wild-type

δ – Chemical shift

Δδ – Chemical shift variation

ΔH – Enthalpy of unfolding

$|\pm 1/2\rangle$ – Ground Kramers doublet

γ^1H – ^1H gyromagnetic ratio

$\gamma^{15}N$ – ^{15}N gyromagnetic ratio

a – Hydrodynamic radius

[θ] – Mean residue ellipticity

$[\theta]_F$ – Mean residue ellipticity in the folded state

$[\theta]_U$ – Mean residue ellipticity in the unfolded state

ε – Molar extinction coefficient

υ_N – ^{15}N nuclear frequency

τ_c – Rotational correlation time

ΔNOE – Uncertainty of the heteronuclear nuclear Overhauser effect

λ – Wavelength

k – Boltzmann's constant (1.38×10^{-23} kg m^2 s^{-2} K^{-1})

F – Faraday constant (96485 C mol^{-1})

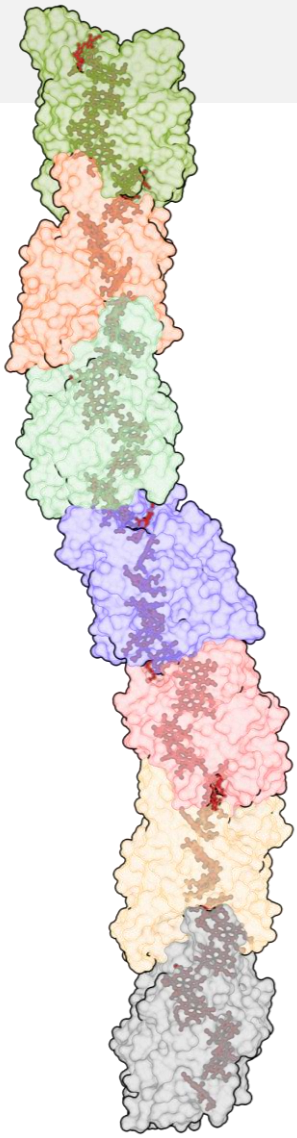
R – Molar gas constant (8.314 J K^{-1} mol^{-1})

η – Viscosity of water (0.89×10^{-3} kg m^{-1} s^{-1})

Amino acid abbreviations		
Alanine	Ala	A
Arginine	Arg	R
Asparagine	Asn	N
Aspartate	Asp	D
Cysteine	Cys	C
Glutamate	Glu	E
Glutamine	Gln	Q
Glycine	Gly	G
Histidine	His	H
Isoleucine	Ile	I
Leucine	Leu	L
Lysine	Lys	K
Methionine	Met	M
Phenylalanine	Phe	F
Proline	Pro	P
Selenocysteine	Sec	U
Serine	Ser	S
Threonine	Thr	T
Tryptophan	Trp	W
Tyrosine	Tyr	Y
Valine	Val	V

1

Introduction



This chapter was partially reproduced from (i) **T. M. Fernandes**, L. Morgado, D. L. Turner and C. A. Salgueiro, Protein engineering of electron transfer components from electroactive *Geobacter* bacteria, Antioxidants, 10 (2021) 844, (ii) C. A. Salgueiro, L. Morgado, M. A. Silva, M. R. Ferreira, **T. M. Fernandes** and P. C. Portela, From iron to bacterial electroconductive filaments: Exploring cytochrome diversity using *Geobacter* bacteria, Coordin. Chem. Rev., 452 (2022) 214284, and (iii) L. R. Teixeira, **T. M. Fernandes**, M. A. Silva, L. Morgado and C. A. Salgueiro, Characterization of a novel cytochrome involved in *Geobacter sulfurreducens*' electron harvesting pathways, Chem. Eur. J., 28 (2022) e202202333.

List of contents

Introduction 33

1.1. Microbes as the pioneers of Life's evolution 35

1.2. Electroactive microorganisms 37

1.2.1. Extracellular electron transfer in bacteria 40

1.3. Microbial electrochemical technologies 41

1.4. *Geobacter* bacteria 45

1.4.1. *Geobacter sulfurreducens* 46

1.4.2. Extracellular electron transfer pathways 47

1.4.3. Electron harvesting pathways 57

1.5. Cytochromes – The chemistry of the key players in extracellular electron transfer 59

1.6. Protein engineering of multiheme cytochromes 62

1.7. Objectives and thesis outline 66

1.8. References 67

1.1. Microbes as the pioneers of Life's evolution

The origin of Life is a captivating event that has fascinated humanity for centuries. Humankind continually strives to comprehend the unlikelihood of this cosmic event, seeking not only to clarify our understanding of the Universe, but also to define our place and purpose within it. Planet Earth was formed around 4.6 billion years ago as the result of countless astronomical collisions between dust particles, asteroids and gases that dwelled around a young Sun [1]. Given the vast number of stars in the Universe and the abundance of planetary systems discovered so far, one can assume that planet formation is a relatively common phenomenon. The emergence of Life, however, is a much more complex and rarer event that depends not only on planetary habitability [2] but also on a process of abiogenesis, which refers to the natural development of life from non-living matter, in a primordial soup, according to the Oparin-Haldane theory [3].

Earth's first living beings, simple anaerobic heterotrophic bacteria, appeared in primitive sulfurous oceans nearly 4 billion years ago, under a poisonous reducing atmosphere full of carbon dioxide, nitrogen, hydrogen, methane and water vapor [4]. These microorganisms harvested electrons from abundant mineral sources near active volcanic sites to reduce specific compounds such as nitrate and sulphate, in order to fix carbon oxides (such as CO and CO_2) while generating organic matter [5]. Upon millions of years and due to oxygenic phototrophs [6] that drove the rise of atmospheric oxygen in the Great Oxidation Event [7], microorganisms developed mechanisms to efficiently perform aerobic respiration and to obtain, transfer and store substrates of variable nature. This evolutionary trigger played a crucial role in the transition from single-celled organisms to the vast array of species observed today [8], a testament of the remarkable journey of Life on Earth.

Over billions of years, through the processes of evolution [9] and natural selection [10], Life has undergone diversification and complexification. Single-celled organisms gradually gave rise to multicellular life forms, unlocking new avenues for specialization and cooperation [11]. This allowed for the development of diverse body plans, physiological adaptations and ecological strategies. Environmental changes, such as shifting climates, geological events, and interactions among organisms, played crucial roles in driving speciation and fostering the proliferation of Life's intricacies [12]. The accumulation of genetic diversity, genetic drift, and the emergence of new traits fueled the branching of lineages, leading to the multitude of species that inhabit our planet today. This ongoing evolutionary journey continues to shape and redefine Life's extraordinary tapestry, highlighting the remarkable resilience and adaptive capacity of organisms across the tree of life (Figure 1.1) [13].

Our planet, which has once inhabitable, is now predicted to host more than 8.5 million species [14], spread across all types of environments. Amongst microorganisms, which today are crucial for a multitude of purposes [15, 16], electroactive bacteria are a remarkable branch of evolution that has been under tight scrutiny over the last 30 years [17].

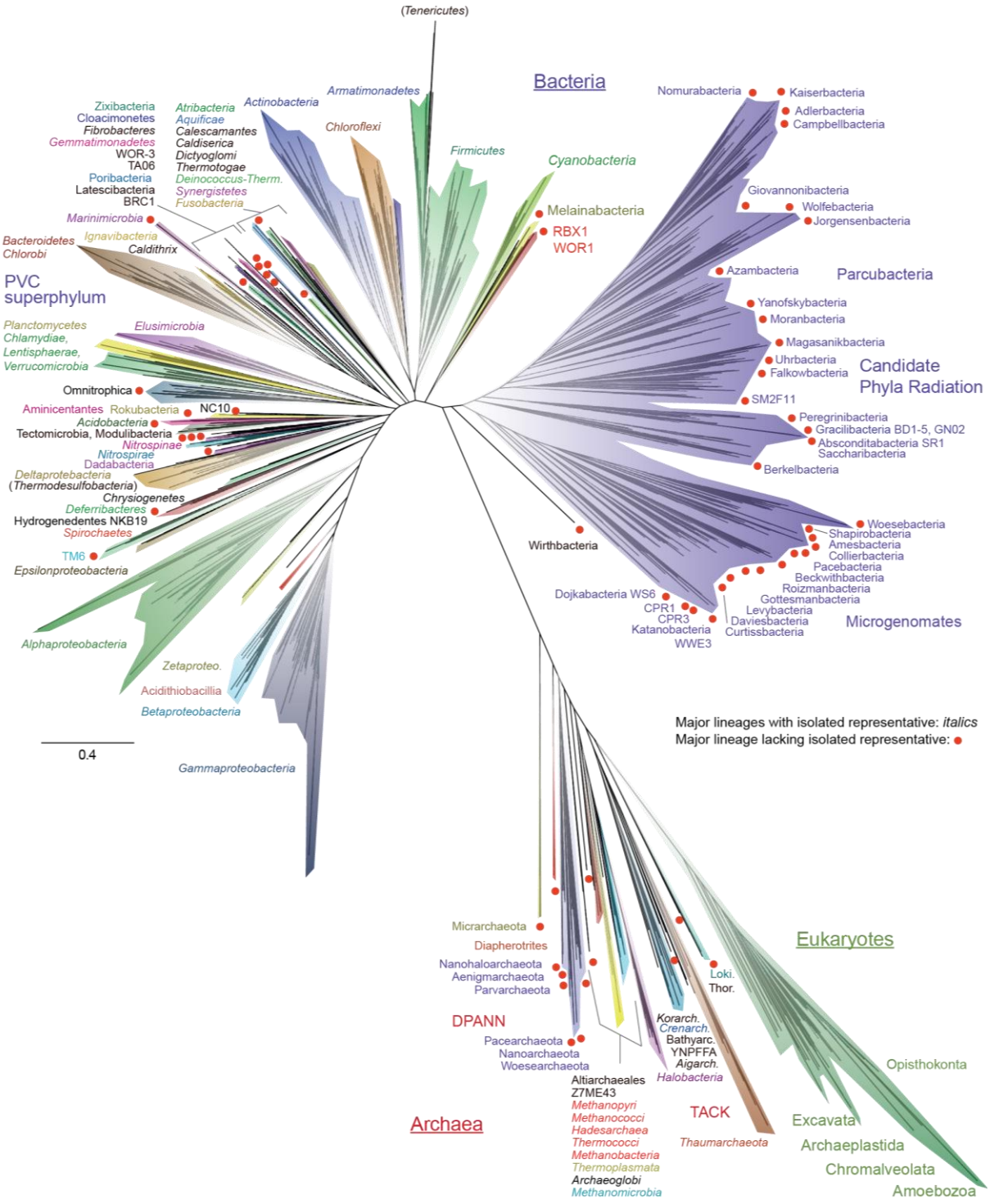


Figure 1.1. Current view of the tree of life. Major lineages are assigned arbitrary colors and named, with well-characterized lineage names, in italics. Lineages lacking an isolated representative are highlighted with non-italicized names and red dots. The names Tenericutes and Thermodesulfobacteria are bracketed to indicate that these lineages branch within the Firmicutes and the Deltaproteobacteria, respectively. Eukaryotic supergroups are noted, but not otherwise delineated due to the low resolution of these lineages. The Candidate Phyla Radiation are assigned a single colour as they are composed entirely of organisms without any isolated representatives and are still in the process of definition at lower taxonomic levels. This figure was reproduced from [13].

1.2. Electroactive microorganisms

The first descriptions of microscopic organisms emerged during the 17th century, credited to the pioneering work of Robert Hooke and Antoni van Leeuwenhoek [18]. Using simple microscopes, they made remarkable observations that laid the foundations for the field of Microbiology, which was slowly developed during more than two centuries. During this time, dedicated researches diligently worked in the development of techniques to isolate and culture microorganisms, leading to the breakthrough establishment of the Germ Theory of Disease [19]. These pivotal advancements were spearheaded by Louis Pasteur and Robert Koch [20], and revolutionized our understanding of microorganisms and their profound impact on human health and environment.

Shortly thereafter, in 1911, Michael Potter reported an intriguing phenomenon involving sedimentary bacteria, later identified as species of *Escherichia coli* and *Saccharomyces cerevisiae*, that demonstrated the ability to carry out electrolysis and produce gas when exposed to an electric current [21]. This early observation marked the first description of bacteria capable of engaging in electron exchange with an electrode, a feature that would later be recognized as a defining trait of electroactive microorganisms. Nonetheless, the significance of this physiological feature and its biotechnological potential went largely unnoticed until the latter half of the 20th century, when the groundbreaking work of Kenneth H. Nealson [22] and Derek R. Lovley [23] showed that bacteria such as *Shewanella* and *Geobacter* can use extracellular metal oxides as terminal electron acceptors. These findings resulted in the development of bioelectrochemical systems without the use of toxic and expensive exogeneous redox mediators [24-26] and propelled the field of Electromicrobiology [27].

Electroactive microorganisms span all three domains of Life [28, 29] and are capable of transferring (exoelectrogens, Figure 1.2A) or receiving (electrotrophs, Figure 1.2B) electrons from or to different electrode surfaces in a multitude of bioelectrochemical devices [30]. This is only possible due to their unique morphological, physiological and metabolic features, developed over millions of years to cope with extreme environments in which electron donors and acceptors are scarce [31]. These include the secretion of electron shuttles [32-35], such as flavins, and, more commonly, cytochrome-dependent extracellular electron transfer (EET) mechanisms [36, 37].

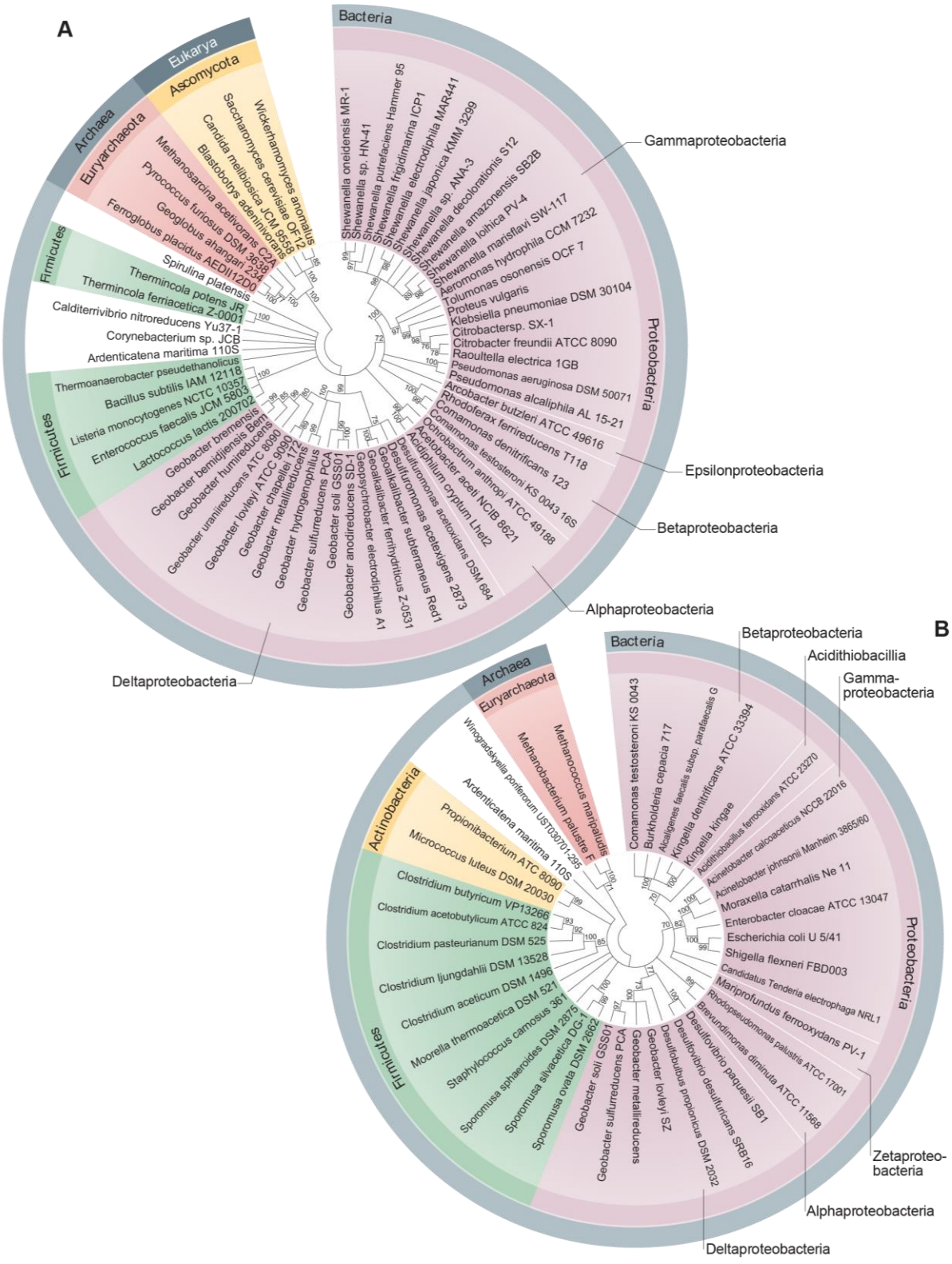


Figure 1.2. Diversity of electroactive microorganisms. The schemes show a molecular phylogenetic analysis of 16S (Archaea and Bacteria) and 18S (Eukarya) rRNA gene sequences organized using the maximum likelihood for selected exoelectrogenic (A) and electrotrophic (B) communities. The outer circle classifies the species based on kingdom, whereas the inner circles show the phylum-level classification for groups of two or more species. The percentage of trees (≥70%) in which the associated taxa clustered together is shown next to the branch nodes, based on 1000 tree iterations. This figure was reproduced from [17].

Electroactive archaea are usually methanogenic and obligate anaerobes that can generate electricity at high temperatures [38], while being capable of reducing iron and oxidizing acetate [39]. They are suggested to use cytochrome-dependent EET [40] rather than mediated electron transfer, however, their preferable mechanisms are still under debate [41, 42].

Electroactive bacteria are the most diverse and studied type of electroactive microorganisms, spreading across a range of taxonomic groups that host both Gram-negative (proteobacteria, acidobacteria and cyanobacteria) and Gram-positive (firmicutes and actinobacteria) bacteria [43]. These microorganisms thrive in diverse environments, using both electron shuttles and cytochrome networks to reduce an extensive list of electron acceptors. One of the most remarkable types of electroactive bacteria discovered to date are cable bacteria of the family *Desulfobulbaceae*, which grow as electrically conductive multicellular filaments (also designated cables, Figure 1.3A) that can reach centimeters in length, enabling long-range electron transport (Figure 1.3B) [44, 45].

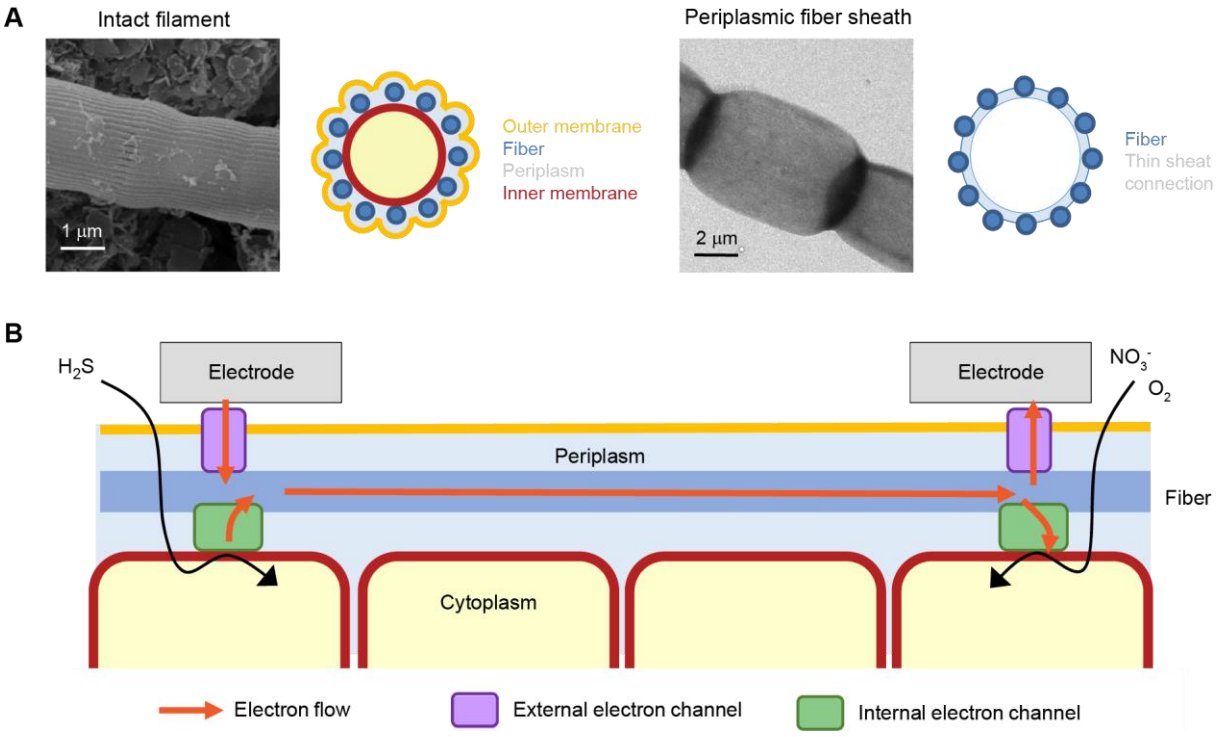


Figure 1.3. Morphology of cable bacteria. (A) SEM and TEM images of an intact cable bacterium filament (left) and an extracted cable bacterium filament retaining the fiber sheath (right), respectively. For both cases, a schematic of the cross-sections is shown. (B) Proposed mechanism of electron transport through cable bacterium filaments. Filamentous multicellular cable bacteria perform centimeter-scale electron transport in a process that couples the oxidation of an electron donor (H_2S) in deeper sediment to the reduction of an electron acceptor (most commonly O_2, but also NO_3^-) near the surface. This figure was adapted from [46].

Remarkably, while all other known biological electron transport metalloproteins contain redox-active iron or copper metalloproteins [47], long-range electron transport in cable bacteria is believed to involve unique sulfur-ligated nickel cofactors, located in the periplasmic bacterial fibers [48-50].

Finally, the most notable electroactive eukaryotes are yeasts, part of the *Fungus* kingdom, with *S. cerevisiae* holding the spotlight [51]. Yeasts transfer electrons to diverse substrates in both aerobic and anaerobic conditions, by secreting endogenous mediators [52] or by using specialized protein appendages [53].

1.2.1. Extracellular electron transfer in bacteria

Electroactive bacteria share an untraditional organization of their electron transport chains. Using Gram-negative bacteria as an example, bacterial electron transport chains are typically composed by a set of membrane-bound electron carriers organized from low to high reduction potentials, thus allowing a seamless flow of electrons to a specific terminal electron acceptor (Figure 1.4A).

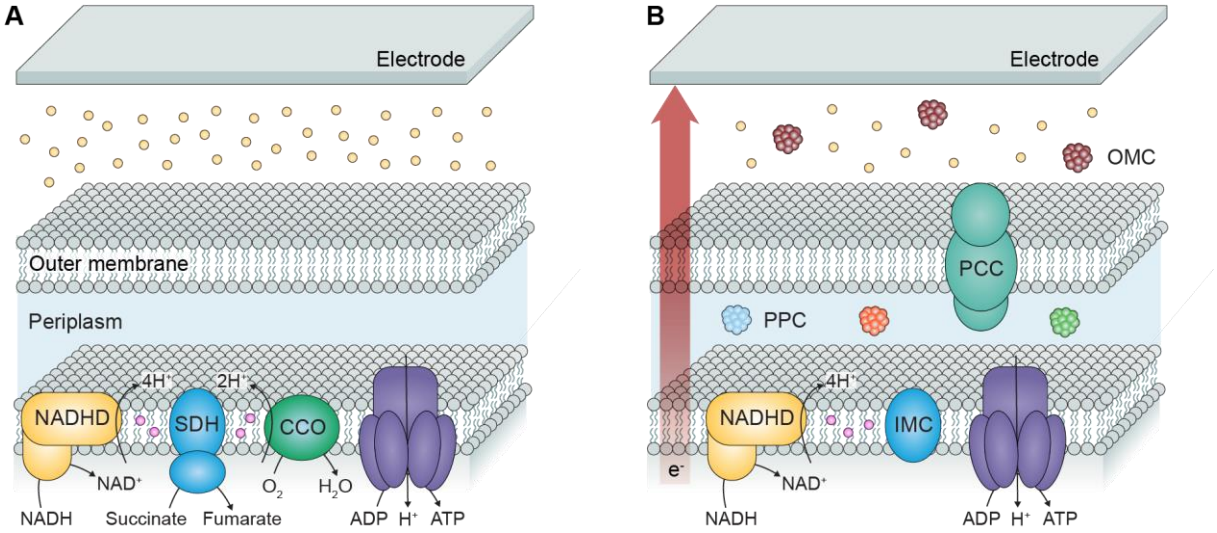


Figure 1.4. Respiratory chains of Gram-negative bacteria. (A) Electron transport chain for the cytoplasmic reduction of O_2 and fumarate in *E. coli*. SDH refers to succinate dehydrogenase and CCO to cytochrome *c* oxidase. (B) Electron transport chain for a putative exoelectrogen. IMC refers to inner membrane cytochrome, PPC to periplasmic cytochrome, PCC to porin-cytochrome complex and OMC to outer membrane cytochrome. In both panels, NADHD refers to NADH dehydrogenase. The rose spheres located within the inner membrane represent liposoluble electron-shuttles. The yellow spheres in the extracellular environment represent soluble electron shuttles. The ATP synthase is represented in purple. For simplicity, the menaquinone pool (MQH_2/MQ) is not represented. This figure was adapted from [54, 55].

These electron carriers diffuse laterally in the membrane and interact transiently to promote a fast exchange of electrons [56]. Alternatively, liposoluble electron-shuttles, such as quinones, can be used to electronically connect these protein complexes [57]. The electron transfer along the transport chain is accompanied by the generation of a transmembrane ion gradient that creates a chemical imbalance in the form of a proton motive force. This force is used by the ATP synthase complex to drive the synthesis of ATP [55]. Nevertheless, a traditional electron chain lacks the reach needed to transport electrons to the cell exterior, and is not sufficient to warrant the survival of microorganisms that dwell in anaerobic environments in which the available electron acceptors are often scarce and insoluble. Hence, microorganisms that perform EET use dedicated electron transfer

chains that bridge the gap between the inner membrane and the cell exterior (Figure 1.4B). In these specialized metabolic pathways, an inner membrane associated NADH dehydrogenase catalyzes the electron transfer from NADH to a menaquinone pool. Then, specific menaquinol oxidoreductases receive electrons from the menaquinone pool and transfer them to periplasmic cytochromes. In Gram-negative bacteria, electrons are transferred to cytochromes located in the outer membrane, which in turn are responsible for the reduction of the extracellular electron acceptors [58]. In Gram-positive bacteria, the EET pathways are similar to those in Gram-negative bacteria, with the final steps of EET being warranted by cytochromes embedded in the peptidoglycan layer [59]. Alternatively, microorganisms lacking the required cellular organization to perform direct EET can use electron shuttles, such as phenazines, quinones, flavins and even endogenous antibiotics [60], to bridge the gap between the cell and the final electron acceptor, which can be specific substrates or electrodes (Figure 1.4A). These redox mediators can be endogenous or exogeneous and diffuse freely between cells and substrates in biofilms to maximize the efficiency of electron transfer processes [61]. Certain Gram-negative bacteria, such as *Geobacter* and *Shewanella*, are capable of conjugating both mediator- and cytochrome-dependent EET (Figure 1.4B).

1.3. Microbial electrochemical technologies

The unique metabolic features of electroactive microorganisms allow them to thrive in extreme environments, using organic and inorganic compounds as terminal electron acceptors, some of which are toxic or radioactive [62]. These features, as well as the biological components that enable them, make interesting targets for microbial electrochemical technologies (MET), which include bioenergy production in microbial fuel cells (MFC) [63-65], bioremediation stations for wastewater treatment [62, 66], sustainable biorefineries for microbial electrosynthesis of green chemicals [67], bioelectronic sensors [68-70] and several bionanotechnologies [71]. A remarkable example of these technologies is the application of protein nanowires from electroactive bacteria in a recently designed device that generates power from ambient humidity (Figure 1.5) [72]. This thin-film device made from nanometer-scale protein wires harvested from the *Geobacter sulfurreducens* bacterium, which are deposited between two gold electrodes, generates continuous electric power in an ambient environment. The device produces a sustained voltage of around 0.5 V across a 7 μm-thick film, with a current density of around 17 $\mu A\ cm^{-2}$. For reference, this current density would only be sufficient to power microsensors, low-power LED indicators or other small-scale electronic devices. The driving force behind this energy generation is a self-maintained moisture gradient that forms within the film, when the film is exposed to the humidity that is naturally present in air. Additionally, by connecting several devices linearly, one may scale up the voltage and current to power electronics [72]. This method offers the promise of clean power for self-sustained systems by harvesting energy from the environment. The exact mechanism for energy harvesting in the protein nanowire film has not been elucidated completely, however, it is believed that water molecules in air either naturally comprise ionized species [73-75], or are ionized when absorbed on the nanowire surface [72]. The ionized clusters donate charge to the nanowire, supplying the closed-loop current flow driven by the

voltage resulting from the moisture gradient. Finally, a dynamic absorption-desorption exchange of water molecules at the interface provides a large reservoir for this continuous exchange of water molecules to generate a sustained electric output.

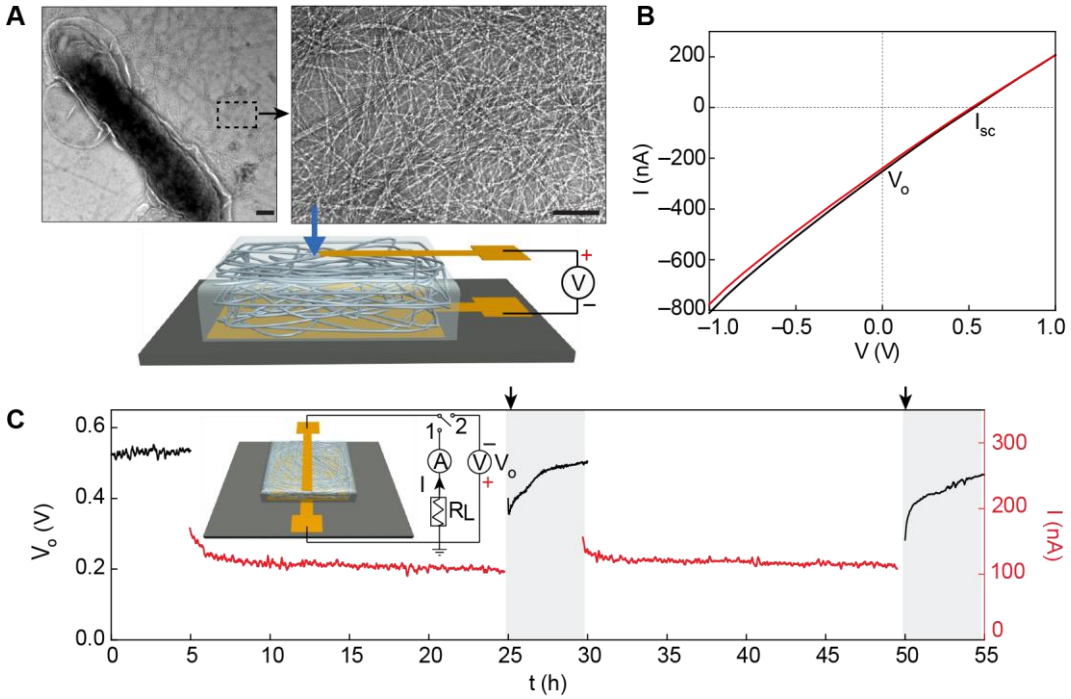


Figure 1.5. Protein nanowire device and electric output from ambient humidity. (A) TEM images of the purified network (right panel) produced by the microorganism *G. sulfurreducens* (dark shape in the left panel). Scale bars (100 nm). Below the TEM images, a diagram of the device is shown. (B) Typical current (I)-voltage (V) curves from the device with ambient lighting (red) and in the dark (black), at an ambient relative humidity of 50%. (C) Evolution of I (red) and V_0 (black) from a nanowire device in the ambient environment (at a relative humidity of 50%). The device initially had a V_0 of approximately 0.52 V (0-5 h). Connecting to a load resistor (R_L (resistor/inductor) = 2 MΩ) yielded a continuous and gradually stabilized I of about 110 nA for 20 h (red curve, 5-25 h). Then, R_L was disconnected and V_0 was recorded (indicated by the arrow at t = 25 h). V_0 gradually increased to the initial value of 0.5 V (25-30 h), showing a self-recharging process. Reconnecting to R_L yielded a repeated continuous powering to the R_L (I is approximately 115 nA) – red curve (30-50 h). Disconnecting the R_L yielded a second self-recharging process (50-55 h), which brought V_0 back to 0.5 V again. The inset shows the circuit diagram, in which connections to terminals 1 and 2 correspond to I and V_0 measurements, respectively. This figure was reproduced from [72].

In other MET, living microorganisms can interact with electrodes to perform substrate oxidation (anodic) or substrate reduction (cathodic) reactions, depending if they are electrotrophic or exoelectrogenic. Exoelectrogens oxidize organic matter, releasing electrons that travel through an external circuit, while generating electricity (Figure 1.6, MFC). Electrotrophs utilize provided electrical current to drive the reduction of carbon dioxide or other chemicals into biofuels and other useful compounds (Figure 1.6, microbial electrosynthesis cell). These configurations are the basis for all MET, whose applications result from the intricate interplay between microbial physiology,

electrochemistry and environmental engineering. The addition of different substrates in the anodic or cathodic chambers can be used to tune the products of such technologies (Figure 1.6, microbial electrolysis cell), while the presence of additional chambers creates an internal potential that may drive reactions such as water desalination (Figure 1.6, microbial desalination cell).

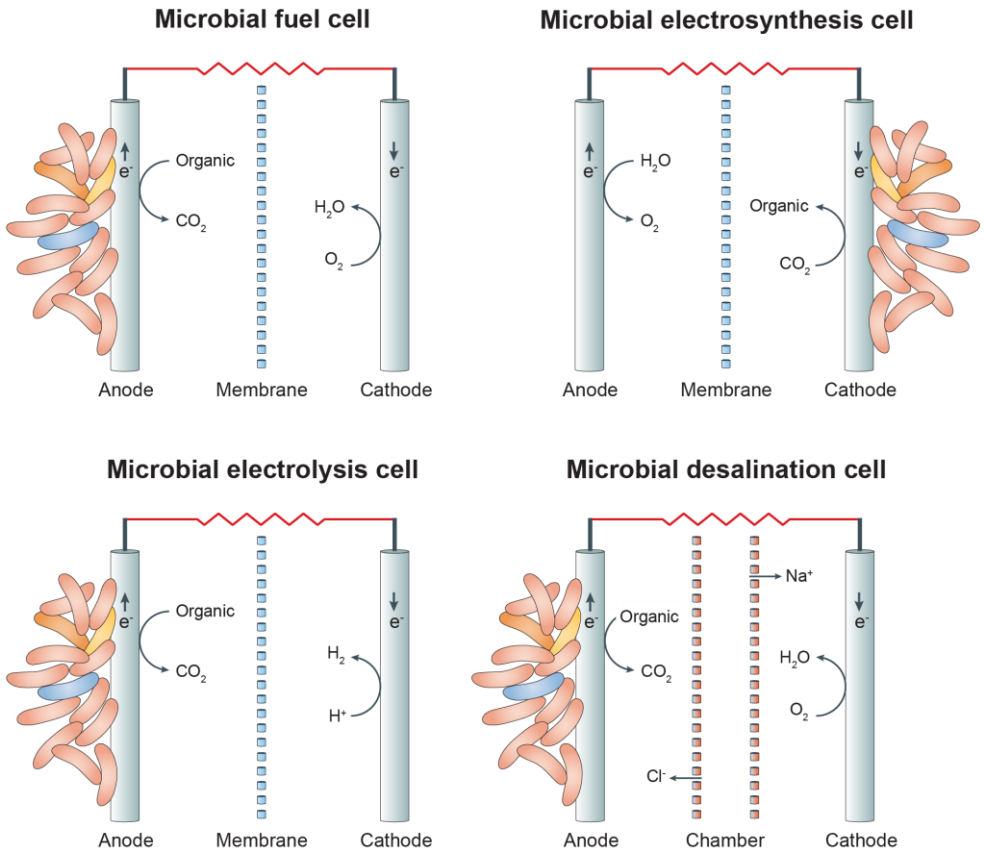


Figure 1.6. Examples of microbial electrochemical technologies. Electroactive microorganisms can colonize either the anode (exoelectrogenic) or the cathode (electrotrophic) to perform substrate oxidation or reduction, respectively. Substrate oxidation releases electrons that travel through an external circuit, which can be used to either produce electricity or to promote additional reactions. Substrate reduction, which occurs after an external power source provides electrons, can be used to turn carbon dioxide or other chemicals into biofuels and other useful compounds. In MET, the cathodic and anodic chambers are usually separated by one or multiple membranes, which allow the passage of protons and ions resulting from microbial metabolism, thus creating ionic gradients that can be used for multiple purposes. This figure was adapted from [54].

Despite their undeniable potential and versatility, MET are mostly limited to lab-scale setups, except for a few pilot-scale MFC [76] and U(VI) bioremediation stations [77]. Nevertheless, some companies, such as Aquacycl (https://aquacycl.com/), iMetland (http://imetland.iot4water.com/) and Nanoelectra (https://nanoelectra.com/), are starting to work on the development of sustainable industrial-scale MET with applications in wastewater treatment, biosensing and production of added-value products. Additionally, a peculiar example of a MFC setup that surpassed the lab-scale is the one applied at the 2019 Glastonbury Music Festival [78]. This system was developed to continuously power public toilet lighting using solely the urine of festival goers, with no harvesting

circuitry or energy storage, showing the potential of these technologies as energy sources in decentralized areas, such as refugee camps.

The development of different MET depends on a better understanding of their abiotic and biotic components, which include the nature of the electrode materials [79, 80], the reactor design and operating conditions [81], the medium or electrolyte composition [82, 83] and the type of microbial catalyst [84]. From the biotic point of view, the efficacy of the electron transfer processes of the selected microbial catalysts can be optimized through several strategies that include selective adaptation [85], synthetic biology [86-92], and protein engineering (particularly of multiheme cytochromes [93-95]), among others. For example, Yi and co-workers showed that appropriate selective pressures in these microorganisms could lead to the development of new strains with enhanced capability for current production in MFC [85]. In their work, a wild-type (WT) strain of *G. sulfurreducens* was inoculated in a system in which a graphite anode was poised at −400 mV (vs Ag/AgCl) for 5 months, from which an isolate, designated strain KN400, was later recovered. This strain possessed several phenotypic changes in the outer surface of the cell, namely a higher abundance of protein nanowires, pili and flagella, which resulted in higher current and power densities than the WT strain [85]. Jensen and co-workers used a synthetic-biology approach to engineer a microorganism to increase its EET properties [96]. By inserting the porin-cytochrome outer membrane MtrCAB complex of *Shewanella oneidensis* MR-1 into *E. coli*, they were able to significantly increase the capacity of the latter to reduce metal ions and solid metal oxides (Figure 1.7) [96].

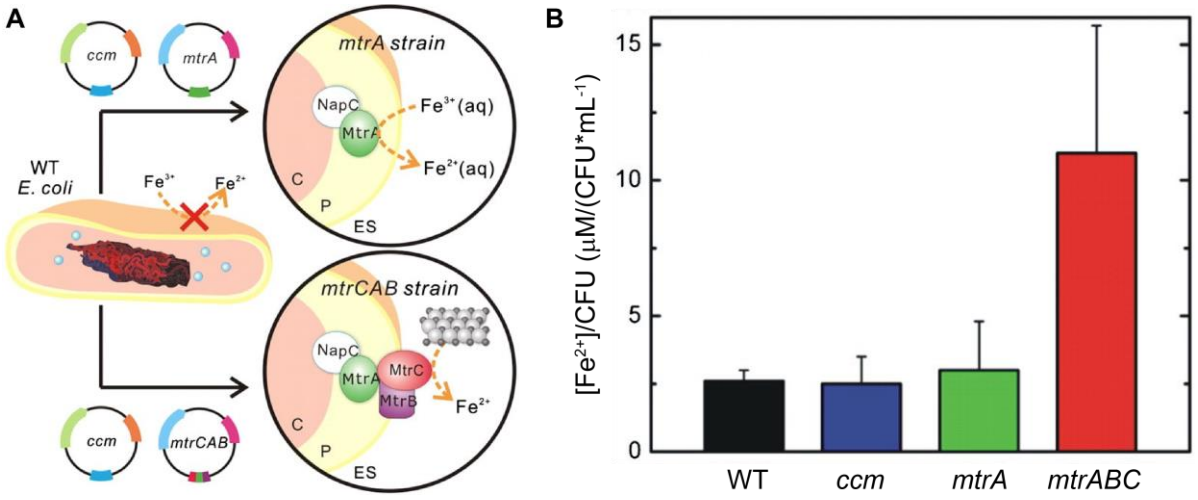


Figure 1.7. Engineering of a synthetic electron conduit in living cells. (A) Schematic of the engineered *E. coli mtrA* and *mtrCAB* strains for soluble and extracellular metal reduction, respectively. NapC is a native *E. coli* inner membrane tetraheme cytochrome, shown to be able to reduce MtrA [97, 98]. (B) Concentration of bulk α-Fe_2O_3 reduced by WT, *ccm*, *mtrA*, and *mtrCAB* strains, normalized by colony forming units (CFU). Error bars represent the standard deviation between triplicates from separate starting cultures. The *ccm* control strain corresponds to a WT strain carrying only the pEC86 plasmid, required for *c*-type cytochrome maturation. This figure was reproduced from [96].

The *mtrCAB* genes from *S. oneidensis* MR-1 encode for the periplasmic decaheme cytochrome MtrA, the 28 strand β-barrel protein MtrB, and the outer membrane decaheme cytochrome MtrC, thus establishing a porin-cytochrome complex capable of performing EET (Figure 1.7A) [99]. These genes were selected as a potential minimal set required to create a synthetic electron conduit to allow *E. coli* to reduce insoluble Fe(III) oxides extracellularly. Additionally, in order to dissect the electron transfer paths of this heterologous pathway and to separately investigate the role of MtrA and MtrC in Fe(III) reduction, MtrA was also expressed by itself (Figure 1.7A). In both cases, since the *E. coli* cytochrome *c* maturation (*ccmABCDEFGH*) genes required for heme insertion are not transcribed under aerobic conditions, the *mtrA* and *mtrCAB* plasmids were co-transformed with pEC86, a plasmid containing *ccmA-H* under the constitutive *tet* promoter (Figure 1.7A) [100, 101]. As expected, the results showed that the *mtrCAB* strain was the most efficient at reducing extracellular iron oxides (Figure 1.7B) and proved that synthetic biology approaches may be used in different genetically tractable, easily manipulated bacteria to create artificial electroactive microorganisms.

Other emerging technologies, such as the CRISPR-Cas9 system [102, 103], are expected to greatly contribute to the development of engineered microorganisms with similar approaches, using those with promising natural features as a chassis [104-106]. The highest current densities recorded to date in MET come from mixed cultures that are usually dominated by δ-proteobacteria of the genus *Geobacter* [17], making them the most often selected microbial catalyst for such applications.

1.4. *Geobacter* bacteria

Geobacter species are one of the primary agents in Nature for coupling the oxidation of organic compounds to the reduction of insoluble Fe(III) and Mn(IV) oxides in soils and sediments, a process of global biogeochemical significance that allows these bacteria to fill important niches in a variety of environments [107, 108]. These bacteria are part of the *Geobacteraceae* family, which was previously considered to be part of the *Desulfuromonadales* order in the δ-subclass of the Proteobacteria. This order was phylogenetically branched between the *Desulfobacterales* and *Desulfovibrionales* orders, and consisted of the genus *Geobacter* and the sole species *Pelobacter propionicus* [17]. However, a recent study proposed a reclassification of the δ-proteobacteria and *Oligoflexia* proteobacterial subclasses. This placed the *Geobacteraceae* family within a new order termed *Geobacterales*, now part of the *Desulfuromonadia* class [109]. Presently, there are 21 classified species of *Geobacter* bacteria [110], 12 of which have their genomes sequenced [110-117].

The first member of the *Geobacter* species to be isolated was *Geobacter metallireducens*, found in 1987 in the Potomac River Estuary, Maryland, USA, following the efforts of Derek R. Lovley and Elizabeth J. Phillips [118-120]. This microorganism was the first bacterium shown to be capable of coupling the oxidation of organic matter to the reduction of insoluble compounds in anaerobic conditions, including iron, uranium, nitrate, gold, silver, mercury and chromate [23]. In 1994, a new *Geobacter* strain capable of reducing iron, sulfur and carbon, while using either acetate or hydrogen as electron donor, was isolated from sediments of a hydrocarbon-contaminated ditch near Norma, Oklahoma, USA, and denominated *G. sulfurreducens* [121]. This bacterium was the first *Geobacter* with

a fully sequenced genome [111], and for which a genetic manipulation system was developed [122]. These tools opened the venue for functional genomic and proteomic studies designed to understand the bacterium's metabolism, gene regulation and EET mechanisms, establishing it as the representative model organism of this family of bacteria.

1.4.1. *Geobacter sulfurreducens*

G. sulfurreducens is mainly found in water sediments, under anaerobic conditions, and was initially classified as a strict anaerobe microorganism [121]. However, the presence of several genes encoding proteins typically involved in aerobic respiration or in protective mechanisms against oxidative stress, such as multiple superoxide dismutases, hydrogenases, peroxiredoxins, rubrerythrins, cytochrome *c* peroxidases and a catalase, hinted towards the existence of oxygen toleration mechanisms in this bacterium [111, 123-125]. This observation was supported by additional studies that not only showed that the bacterium can tolerate O_2 exposure up to 24 h, but that it can also use this molecule as an electron acceptor under microaerobic conditions [126, 127]. These features convey the bacterium with a competitive advantage, since in water sediments near the oxic-anoxic interface, the Fe(II) resulting from microbial extracellular reduction can be reoxidized to Fe(III) and pledge continuous EET processes [127]. Moreover, it has been shown that *G. sulfurreducens* activates different mechanisms depending on the level of oxygen exposure [126, 128]. These include the overexpression of pili to increase motility and allow an escape from oxygen-contaminated areas, or of genes involved in cell encapsulation and biofilm production to decrease O_2 exposure. Out of the proteins involved in potential oxygen toleration mechanisms, two diheme cytochrome *c* peroxidases have been studied in detail: CcpA (gene ID number GSU2813) and MacA (GSU0466). The structures of these proteins, which catalyze the reduction of hydrogen peroxide to water, have been determined by X-ray crystallography (Figure 1.8).

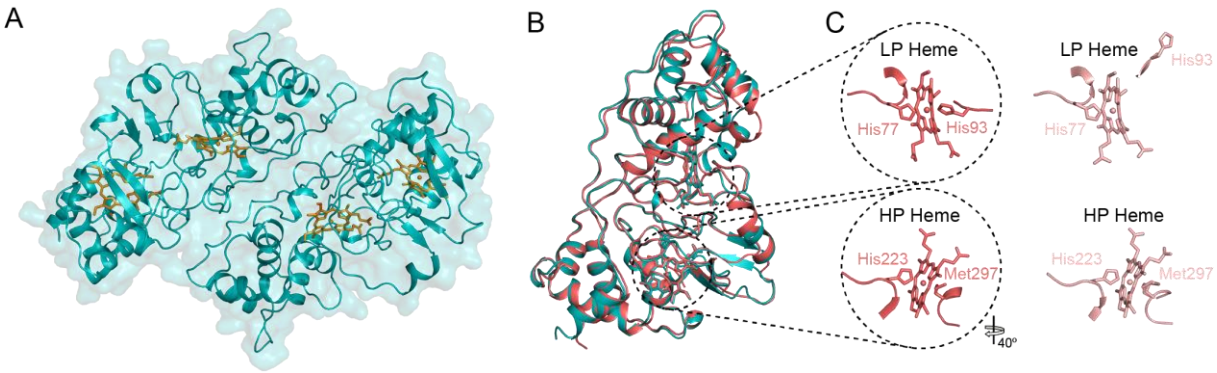


Figure 1.8. Structural features of cytochrome *c* peroxidases CcpA and MacA. (A) Surface representation of the CcpA homodimer in the oxidized form (PDB: 3HQ6 [129]). (B) Overlay of CcpA (blue) and MacA (pink, PDB: 4AAL [130]) oxidized monomers. (C) Heme axial ligands of MacA in the oxidized (pink) and reduced (light pink, PDB: 4AAN [130]) forms. In the reduced form, the LP distal ligand His^{93} is detached. LP and HP stand for low- and high-potential heme, respectively. This figure was reproduced from [131].

As commonly found in other cytochrome *c* peroxidases, CcpA (Figure 1.8A) and MacA (Figure 1.8B) are homodimers organized in two domains, each containing a heme group. These hemes possess different axial coordination and reduction potential values. The heme located at the N-terminal is the catalytical heme and is coordinated by two histidine residues in the oxidized form, having a low midpoint reduction potential (LP heme). The heme located at the C-terminal is His-Met coordinated and has a high midpoint reduction potential (HP heme). During catalytical activity, the reduction of the HP heme leads to a structural rearrangement that triggers the dissociation of the LP heme distal ligand (His^{93}), leaving the protein in a mixed valence state, while making the active site accessible for the binding of hydrogen peroxide (Figure 1.8C).

Besides cytochrome *c* peroxidases, *G. sulfurreducens* codes for an unprecedent number of other *c*-type cytochromes [111], superior to those observed in *Desulfovibrio* and *Shewanella* bacteria, also known for relying on different cytochromes for respiration [132, 133]. In fact, out of the 3466 predicted protein-encoding open reading frames of the bacterium's genome, 111 contain at least one match to the *c*-type cytochrome motif that identifies heme groups (CXXCH, where X corresponds to any amino acid) [111]. An independent analysis of the bacterium's genome shows that the number of cytochromes may be higher (at least 132 genes contain one CXXCH motif), even if contracted and extended motifs observed in other *c*-type cytochromes are not considered [134-148]. If these types of motifs are considered, the number can go up to 180, meaning that *c*-type cytochromes comprise 3-5% of the bacterium's genome. Additionally, more than 65% of these cytochromes contain more than one heme group. Interestingly, due to the great number of heme-containing proteins expressed, *G. sulfurreducens* cells have a pink to reddish color.

The incredible number of cytochromes coded by the genome of *G. sulfurreducens* highlights the importance of electron transport for this microorganism and suggests that this bacterium possesses electron transfer networks with high flexibility and apparent redundancy, which allows the reduction of diverse acceptors in natural environments. In fact, no single gene deletion on the bacterium's genome was found to eliminate electron transfer to all electron acceptors, thus confirming the complexity of its electron transfer networks [58].

1.4.2. Extracellular electron transfer pathways

The ability of *G. sulfurreducens* to reduce or oxidize extracellular insoluble compounds depends on the strategic localization of multiple cytochromes in the different cell compartments, connecting the cytoplasmic metabolism with the cell exterior. These cytochromes form complex networks with different players, which will vary depending on the environmental conditions, working either in the export or import of electrons [149].

Considering a model in which the bacterium is using acetate as an energy source and ATP generation is completely dependent on EET [150], this carbon substrate is initially converted into acetyl-CoA and then oxidized to CO_2 through the Krebs cycle, while reducing NAD^+ to NADH [151, 152]. The generated NADH molecule transfers electrons to the menaquinone pool via a proton pumping NADH dehydrogenase [153], which in turn might reduce different quinol dehydrogenases,

depending on the reduction potential of the final electron acceptor [154]. The exact mechanisms that lead to the expression of specific quinol dehydrogenases in response to this stimuli are still unknown, however, they likely involve environmental sensing and signal transduction pathways that affect gene regulation and protein synthesis, such as those promoted by the GSU0582 and GSU0935 heme-containing signal transducers [155].

Quinol dehydrogenases link the central metabolism with periplasmic cytochromes

The genome of *G. sulfurreducens* encodes six putative quinol dehydrogenases in the bacterial inner membrane, either in complex with other proteins or as single cytochromes. Out of these, ImcH, CbcL and CbcBA have been shown to be required for the reduction of specific electron acceptors (Figure 1.9).

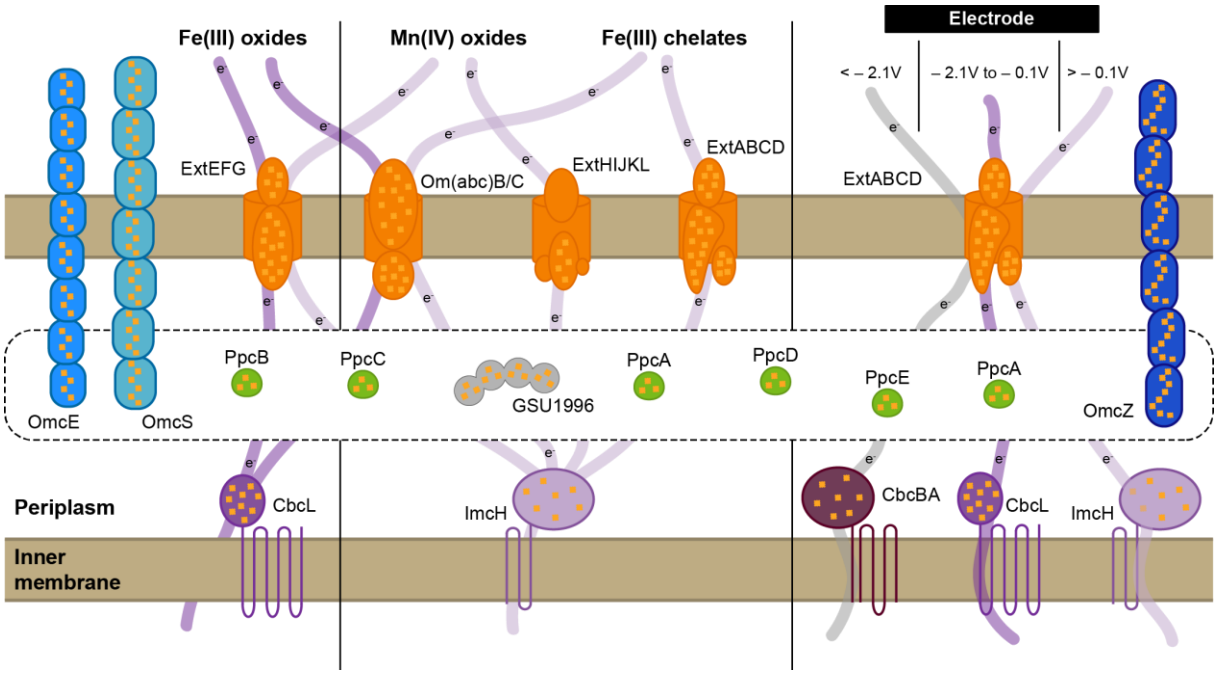


Figure 1.9. Extracellular electron transfer pathways of *G. sulfurreducens*. The different *c*-type cytochromes involved in the electron transfer pathways towards Fe(III) oxides, Mn(IV) oxides, Fe(III) chelates and electrodes are represented, together with their putative electron flows. The *c*-type heme groups are represented in light orange. The tricarboxylic acid (Krebs) cycle, the NADH dehydrogenase, the menaquinone pool and the ATP synthase are all omitted for simplification. This figure was partially reproduced and adapted from [131].

The ImcH cytochrome contains 7 *c*-type heme groups and is homologous to the NapC/NirT/NrfH family, possessing a conserved menaquinone-binding site (Figure 1.10) [156]. A mutated strain in which the gene that codes for ImcH was deleted was unable to reduce Fe(III) citrate, Fe(III)-EDTA, and Mn(IV) oxides, all of which contain a reduction potential above –100 mV [157]. In contrast, this strain was able to reduce Fe(III) oxides with reduction potentials below this value. The cytochrome CbcL, also known as Cbc1 or CbcY, is expected to hold 2 *b*-type heme groups in a transmembrane domain composed by six α-helices, and 9 *c*-type heme groups in its periplasmic domain (Figure 1.10) [158]. This *bc* cytochrome complements the phenotype of ImcH, since a mutated strain in which the gene

that codes for CbcL was deleted was unable to reduce low-potential electron acceptors, such as Fe(III) oxides and electrodes poised at −100 mV [159]. Recently, a third mutated strain, lacking the CbcBA heterodimer, formed by CbcA and CbcB, ceased Fe(III) reduction at −210 mV and was unable to transfer electrons to electrodes poised in a redox window between −210 and −280 mV [160].

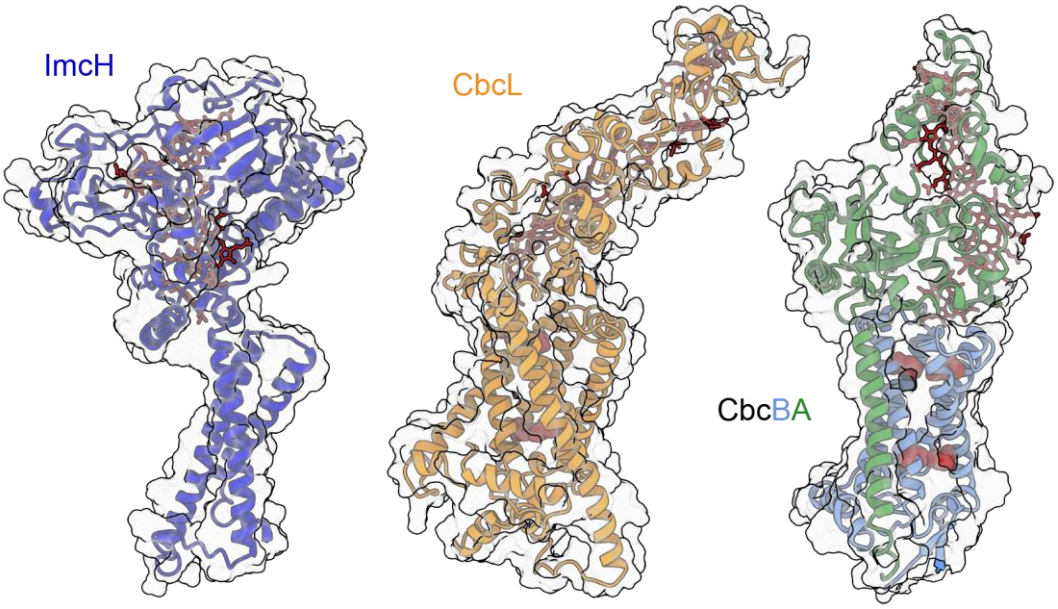


Figure 1.10. AlphaFold 2 models of the quinol dehydrogenases ImcH, CbcL and CbcBA. The surfaces of each model are represented in light gray. The *c*-type heme groups are represented as red sticks. The putative binding sites for the *b*-type hemes in CbcL and CbcBA are represented by red spheres, corresponding to the putative axial histidine residues. The pLDDT scores of each model can be found in Figure 7.1.

The CbcBA heterodimer forms a chimeric *bc* cytochrome, with CbcA holding 7 *c*-type heme groups, facing the periplasm, while CbcB is a transmembrane protein, composed by 4 α-helices and expected to hold 2 *b*-type heme groups (Figure 1.10). This heterodimer is part of the Cbc5 complex, composed by the CbcABCDE proteins [161].The remaining quinol dehydrogenase containing complexes – Cbc3 (CbcVWX), Cbc4 (CbcSTU) and Cbc6 (CbcMNOPQR) – have not been associated with particular working reduction potential ranges, but they likely provide complementary pathways for EET [161].

G. sulfurreducens has been shown to grow on anodes poised between +0.4 and −0.3 V vs the standard hydrogen electrode (SHE) with similar biomass growth, independently of the applied potential [162]. These observations contrast with the those made on a similar study performed with *Shewanella putrefaciens*, in which biomass growth was dependent on the anodic potential applied [163, 164]. This shows that *G. sulfurreducens* is able to maintain a constant growth under different reduction potential ranges by adjusting its metabolic pathways to balance energy conservation [165]. Remarkably, the bacterium is able to regulate the number and nature of inner membrane-bound respiratory proteins under different thermodynamic conditions, by creating invaginations of the inner membrane, designated intracytoplasmic membrane structures, which could increase the extracellular electron flux [166].

Altogether, the quinol dehydrogenases of *G. sulfurreducens* are expected to not only warrant the electron transfer from the menaquinone pool to the periplasmic cytochromes, but also to be involved in the generation of the proton gradient that ultimately drives the synthesis of ATP by ATP synthase [58, 116].

Periplasmic electron transfer is promiscuous and redundant

The periplasm of *G. sulfurreducens* is dominated by the highly conserved and multifunctional PpcA-family of triheme cytochromes (Figure 1.11), which seamlessly links the electron transfer between the inner and outer membrane components (Figure 1.9) [131, 167].

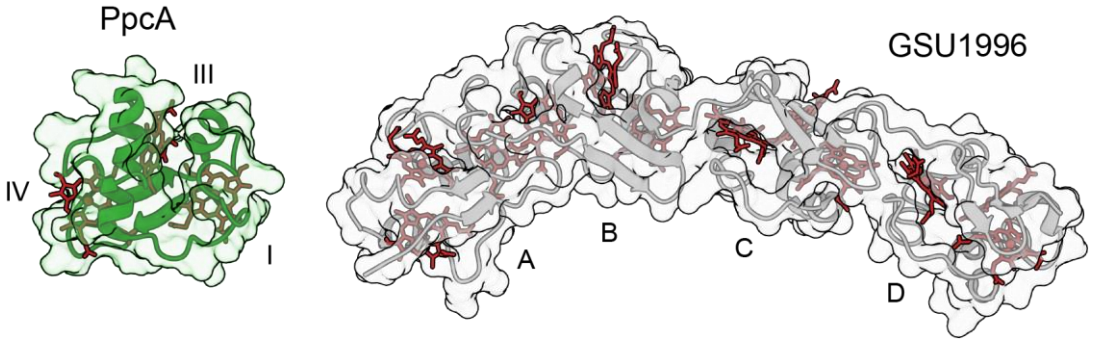


Figure 1.11. Structures of periplasmic cytochromes of *G. sulfurreducens*. The solution structure of PpcA (PDB: 2LDO [168]) and the crystal structure of GSU1996 (PDB: 3OV0 [169]) are represented as ribbon and surface, in green and gray, respectively. The GSU1996 cytochrome has a 12 nm long modular structure with the hemes distributed along four covalently linked domains (designated A-D). Each domain contains three hemes with a disposition similar to those of the PpcA-family. The hemes in these cytochromes are numbered I, III and IV, a designation that derives from the superimposition of the hemes in cytochromes c_7 with those of the structurally homologous tetraheme cytochromes c_3 [170].

These five cytochromes (PpcA-E) exhibit promiscuous interactions both among themselves and with various redox partners [171], thus participating in several EET pathways [172-175], in contrast with those of the inner and outer membranes, which can be associated with specific routes of electron acceptor reduction. Additional periplasmic proteins may also play a role in this process, while those containing a large number of *c*-type heme groups, such as GSU1996 (Figure 1.11) [169], are expected to work as electron biocapacitors, contributing to the enhancement of the electron-storage capacity of the periplasm, thus assuring a continuous flow of electrons and concomitant production of metabolic energy [176, 177]. Other cytochromes that might work as biocapacitors in the periplasm of *G. sulfurreducens* are GSU0702 (35 hemes), GSU2210 (27 hemes), GSU2494 (15 hemes), GSU2495 (26 hemes), GSU2882 (18 hemes), GSU2883 (24 hemes), GSU2884 (27 hemes), GSU2899 (23 hemes), GSU2912 (27 hemes) and GSU3218 (15 hemes), all of which apparently present a linear modular structure, similar to that of GSU1996 (Figure 1.11), according to their predicted AlphaFold 2 models deposited in the AlphaFold protein structure database [178, 179].

While the genome of *G. sulfurreducens* encodes for an unprecedent number of multiheme cytochromes (more than 70), some of which containing a remarkably high number of *c*-type hemes

(up to 35), a survey of prokaryotic genomes conducted in 2010 identified putative *c*-type cytochromes capable of binding up to 46 hemes [180]. Six years later, Badalamenti and co-workers sequenced the genome of the anaerobic Gram-negative bacterium *Desulfuromonas soudanensis* WTL and found a gene coding a putative *c*-type cytochrome (DSOUD_0664) able to bind 69 hemes [181]. If one considers rare extended heme binding motifs, such as CX_3CH, CX_4CH and CX_5CH, the DSOUD_0664 cytochrome may be able to bind up to 77 hemes [181]. The current record holder cytochrome, with 113 predicted heme-binding motifs, was found in an anaerobic methanotrophic archaea [182]. These impressive numbers are only surpassed when supramolecular assemblies are considered, such as that of the hydrazine dehydrogenase (HDH) of the anammox organism *Kuenenia stuttgartiensis*, which assembles into a 1.7 MDa multiprotein complex containing an extended electron transfer network of 192 *c*-type heme groups (Figure 1.12) [183].

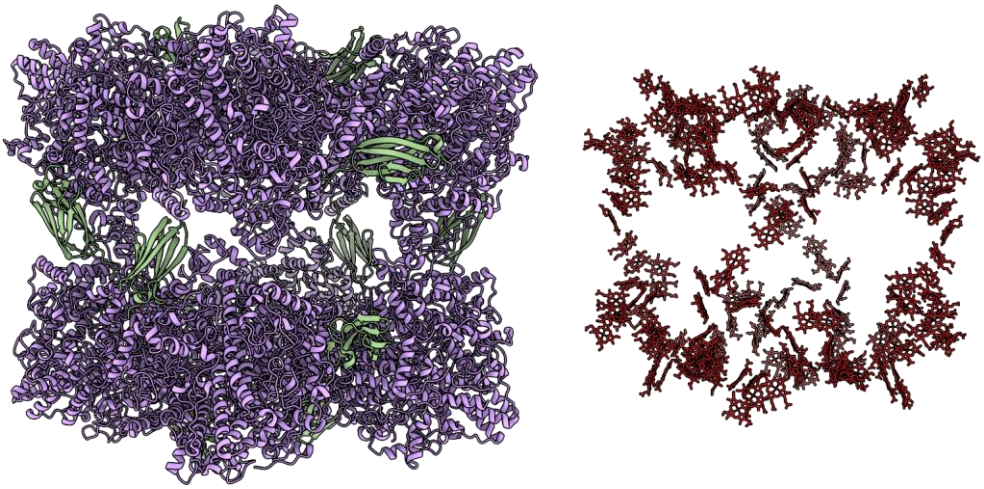


Figure 1.12. Structure of the entire HDH assembly of *K. stuttgartiensis*. The supramolecular assembly (PDB: 6HIF [183]) contains 24 copies of the HDH octaheme monomer arranged as an octamer of trimers $(\alpha_3)_8$ (in purple), together with 12 copies of Kustc1130, which acts as an assembly factor (in green). This results in a 1.7 MDa $(\alpha_3)_8\beta_{12}$ complex with a cube shape of 150 x 150 x 150 $Å^3$. The 192 *c*-type heme network of HDH is represented on the right.

Extracellular reduction of electron acceptors is dependent on different types of cytochromes

Moving back to the current model for EET in *G. sulfurreducens*, once electrons reach the outer membrane, electron transfer is ensured by porin-cytochrome complexes, cytochrome nanowires and freely diffusing cytochromes, which can directly reduce extracellular electron acceptors (Figure 1.9). A thorough description of the properties of the five porin-cytochrome complexes of *G. sulfurreducens* is provided in Chapter 2, while those of freely diffusing cytochromes are addressed in Chapter 4.

Cytochrome nanowires are electrically conductive filaments made of polymeric assemblies of cytochrome subunits and were first discovered in 2019, when the structure of the OmcS nanowire from *G. sulfurreducens*, which can be reduced by all the PpcA-family cytochromes [184], was revealed by cryogenic electron microscopy (cryo-EM, Figure 1.13) [185, 186].

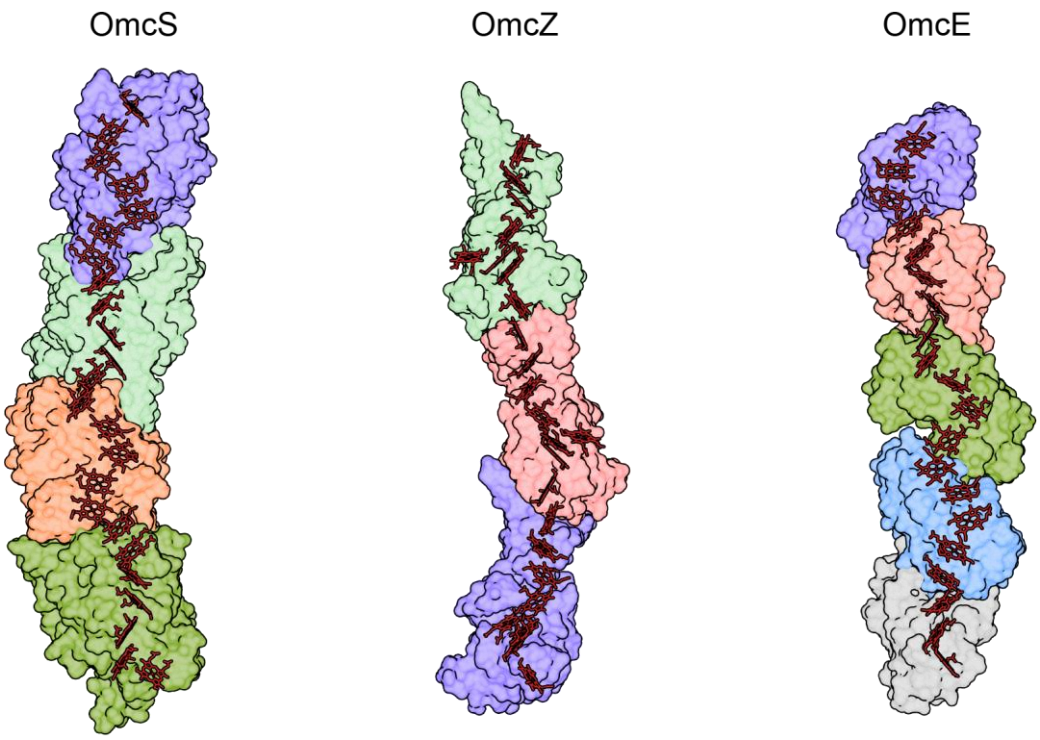


Figure 1.13. Cryo-EM structures of the cytochrome nanowires of *G. sulfurreducens*. The electrically conductive filaments of OmcS (4 hexaheme subunits, PDB: 6EF8 [185]), OmcZ (3 octaheme subunits, PDB: 7LQ5 [145]) and OmcE (5 tetraheme subunits, PDB: 7TFS [187]) are represented as surface, with their heme groups in red.

On one hand, prior to this discovery, polymerization of *c*-type cytochromes had only been observed in horse heart cytochrome solutions in the presence of ethanol, due to successive domain-swapping of the C-terminal α-helix [188, 189]. On the other hand, filament formation by non-cytoskeletal enzymes had been known for decades, having a wide-spread role in enzyme regulation and biology [190]. An example of a bacterial enzyme that undergoes filamentation is the hydrogen-dependent CO_2 reductase (HDCR, Figure 1.14) [191], which directly converts H_2 and CO_2 into formic acid, with a higher activity than any other known biological or chemical catalyst [192, 193]. The minimum repeating unit of this oligomer is a hexamer consisting of two hydrogenases and a formate dehydrogenase, bound around a central core of hydrogenase Fe-S subunits and two small bacterial polyferredoxin-like proteins, whose C-terminal α-helices are responsible for the filament formation (Figure 1.14) [191].

Recently, the structures of the OmcE and OmcZ nanowires have also been determined [145, 148, 187], expanding the range of cytochromes which naturally polymerize in *G. sulfurreducens* (Figure 1.13).

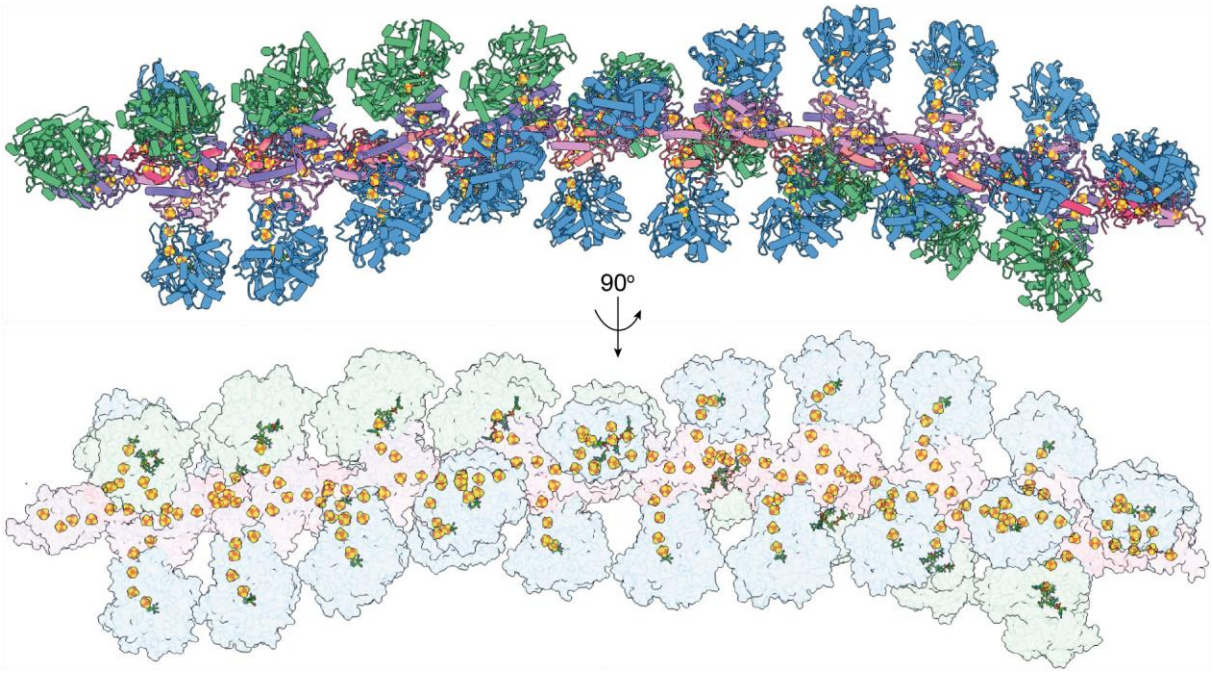


Figure 1.14. Electron nanowire forms the central spine of the HDCR filament. The model of a dodecameric HDCR filament is shown in ribbon (top) and surface (bottom) representations, with each subunit colored differently. The orange and yellow spheres represent the 228 [4Fe4S]-clusters, showcasing the central electron wire of the HDCR filament. The W-bisPGD cofactors are represented as green sticks. This figure was reproduced from [191].

A structural survey of the deposited AlphaFold 2 models of the multiheme cytochromes of this bacterium performed using the Dali [194], Foldseek [195] and PDBeFold [196] servers indicates that (i) the GSU1787 and OmcP (GSU2913) tetraheme cytochromes likely assemble into OmcE-like nanowires, (ii) the OmcJ (GSU0701), OmcT (GSU2503) and OmcU (GSU2501) hexaheme cytochromes likely assemble into OmcS-like nanowires and that (iii) the GSU1334 octaheme cytochrome likely assembles into OmcZ-like nanowires. This observation further supports the view that cytochrome nanowires are essential for EET in *G. sulfurreducens* and that additional types of nanowires, with distinct arrangements and properties, may also assemble from other cytochromes in this bacterium.

The list of naturally occurring cytochrome nanowires is expanding towards other branches of Life, and it is now believed that these filaments are ubiquitous in prokaryotes, establishing a wide-spread mechanism for long-range EET [197]. Remarkably, genes coding for putative cytochrome nanowires have been found in Borg extrachromosomal elements of archaea [198, 199], which are giant genetic elements (up to 1.1 million base pairs) that expand the metabolic capacity of these organisms [200].

While the five structurally characterized cytochrome nanowires have no folding similarity, they contain a conserved heme packing, with alternating antiparallel and T-shaped arrangements (Figure 1.15). Moreover, these filaments present a low percentage of secondary structural elements, a common feature that can be mainly attributed to the pressure of evolutionary selection, which focuses on maintaining closely stacked heme arrangements for efficient electron transfer, rather than on a conserved protein fold [197].

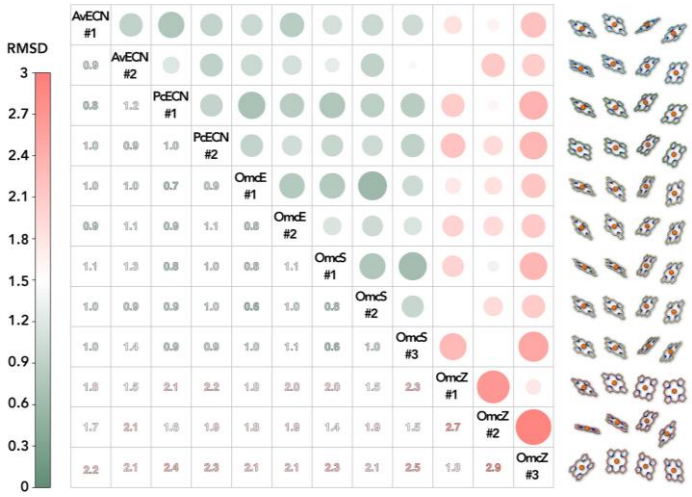


Figure 1.15. Structural comparison of the four-heme arrays of different cytochrome nanowires. The matrix is based on the pairwise RMSD comparison of 20 atoms between two arrays (five atoms per heme, including the iron and the four surrounding nitrogen atoms). The resulting RMSD is shown in color, as indicated in the left bar. Extreme RMSD values are highlighted by larger dot sizes to emphasize either very similar or very different heme arrangements. The tetraheme cytochrome nanowires OmcE, AvECN (*Archaeoglobus veneficus* extracellular cytochrome nanowire) and PcECN (*Pyrobaculum calidifontis* extracellular cytochrome nanowire), possess two four-heme arrays along the filament structure, while the respective OmcS and OmcZ hexaheme and octaheme cytochrome nanowires possess three different arrays. This figure was reproduced from [197].

The three cytochrome nanowires of *G. sulfurreducens* are distinct in their putative function, size and morphology (Figure 1.13). Deletion of the genes encoding the OmcE and OmcS cytochromes affected the reduction of Fe(III) oxides, but not of soluble Fe(III) forms [161, 173, 201-208], whereas OmcZ has been shown to be essential for the formation of conductive electrode biofilms and for the cathodic corrosion of stainless steel (Fe^0) [161, 209-215].

The OmcE nanowire has a rise per subunit of 34 Å with 6.1 subunits per turn, while OmcS has a rise per subunit of 46.7 Å with 4.3 subunits per turn. OmcE filaments' thickness ranges between 3-6 nm, whereas that of OmcS ranges between 5-7.5 nm. These two nanowire cytochromes differ in overall appearance, thickness, internal protein fold, helical pitch and primary sequence, but their heme motifs are superimposable (Figure 1.15). An additional feature shared by OmcE and OmcS is the coordination of a heme in one subunit by a histidine residue from the following subunit, which stabilizes the subunit's interfaces and places all heme groups within 3-5 Å, in a tightly packed arrangement for efficient electron transfer [185-187].

The octaheme OmcZ cytochrome is initially produced by cells as a 50 kDa precursor, which is then cleaved by the serine protease OzpA (OmcZ protease) into a nanowire-forming version of 30 kDa [145]. This protease is suggested to act as a molecular switch that regulates the assembly of OmcZ nanowires. The resulting 3-5 nm filaments contain a linear chain of hemes, with the exception of the highly exposed heme VI, that is likely crucial for electron transfer events and provides OmcZ with a distinctive z shape in cryo-EM images [145, 148]. Unlike the OmcE and OmcS nanowires, there is no heme coordination between different subunits in OmcZ filaments, whose interfaces are alternatively

stabilized by heme-heme and heme-aromatic stacking, as well as by hydrophobic and electrostatic interactions [145].

Overall, the discovery of naturally occurring cytochrome filaments opened the venue for the production of biobased, non-toxic electronics [216-218]. Consequently, the mechanisms underlying their long-range conductivity, which might include coherent or non-coherent transport, such as electron delocalization or electron hopping, have been under tight scrutiny [219-225]. However, most conductivity studies on bacterial filaments are conducted on air-dried sample preparations, assuming that their atomic structure is unchanged from solution, which cannot be the case, since their native folding can only be maintained in a fully hydrated state [219, 220, 226-229]. Therefore, Baquero and co-workers have questioned the physiological relevance of these conductivity measurements [197]. Moreover, these authors stressed that conductivity measurements must be performed with an experimental apparatus that perfectly aligns electrodes such that exposed hemes at both ends of a hypothetical filament are in touch with the measuring system, which is highly unlikely [148, 197]. For example, OmcZ filaments have been stated to be 1000x-fold more conductive than OmcS, despite their similar heme-heme packing [220]. Supposedly, the branched hemes in OmcZ, with the solvent-exposed heme VI in each subunit, accounted for the higher-measured conductivity [145, 148], but these differences might be exacerbated by the different heme-electrode surface distances of the two cytochrome nanowires [148]. An additional issue raised by Baquero and co-workers relates with the fact that the samples used in these measurements are often contaminated with 3 nm filaments of extracellular DNA [187], previously reported to be conductive [230], which might thwart their real conductivity [197]. Ultimately, the techniques currently used to measure the conductivity of individual fully hydrated filaments have several technical limitations, and there is an urgent need for the development of appropriate experimental tools for this purpose.

The pili-nanowire paradigm

The view that *Geobacter* bacteria require extracellular conductive appendages to perform long-range electron transport was established almost 20 years ago [231]. These conductive extracellular filaments were originally thought to be type IV pili, based on atomic force microscopy low-resolution observations of dried samples and other additional studies [231-235]. In fact, mutations disrupting type IV pili production disrupted *Geobacter*'s electron transport [231]. These thin filaments are 6-8 nm in diameter and have lengths of μm, thus being able to connect bacteria with different cells, surfaces and electron acceptors [236]. Their conserved aromatic residues have been argued to be the main players behind their conductivity, by performing electron hopping via π-π stacking [237]. While this hypothesis contrasts with the general view that proteins are poor conductors [238], strategic mutations in specific aromatic residues altered the apparent conductivity of these filaments [239]. Altogether, this view was established in the scientific community as the e-pili hypothesis [240]. The preliminary computational models of the e-pili, together with an experimentally determined nuclear magnetic resonance (NMR) structure of detergent-solubilized monomeric pilin, assumed that these were built entirely from the PilA-N gene product, forming a continuous α-helix [237, 241, 242].

However, a recently determined cryo-EM structure of the type IV pili of *G. sulfurreducens* (Figure 1.16A) shows features incompatible with these assumptions and with the model for electron transport in e-pili (Figure 1.16B) [243].

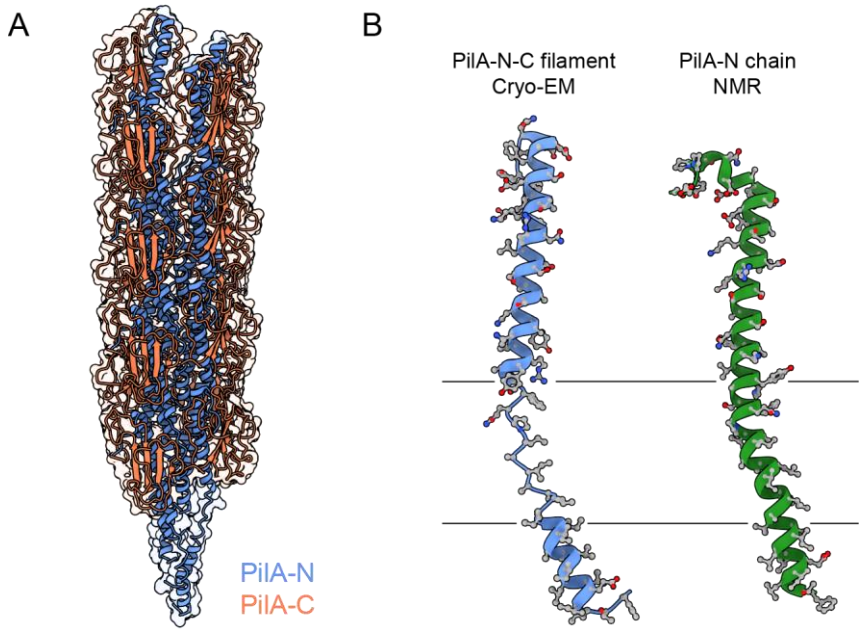


Figure 1.16. Structural properties of the PilA-N-C pili. (A) Cryo-EM structure of the PilA-N-C filament (PDB: 6VK9 [243]). The ribbon and surface representation of the structure show the PilA-C chains (orange) positioned along the PilA-N subunits (blue) to form the PilA-N-C heterodimer. (B) Partial melting of the PilA-N α-helix. Comparison of the experimental structures obtained for a PilA-N chain within a PilA-N-C filament (blue) and for a detergent-solubilized PilA-N chain (green, PDB: 2M7G [241]). The α-helix of the structure determined by cryo-EM (green) is melted between the black lines, preventing the stacking of aromatic residues that has been believed to be the basis for the pili filaments conductivity.

The PilA-N and PilA-C proteins assemble into a heterodimeric pili, in which the C-terminal residues of PilA-N stabilize the copolymerization with PilA-C, positioning this protein along the outer surface of the filament through electrostatic and hydrophobic interactions [243]. Importantly, as seen for all other type IV pili whose structures have been reported to date [244-248], the N-terminal α-helix of *Geobacter*'s filaments is partially melted, positioning the conserved aromatic residues such that electron hopping is not possible (Figure 1.16B). Alternatively, it is believed that these pili assemble and participate in a nanowire translocation system similar to the one encountered in type II secretion pseudopili (or endopili) [246, 249] (Figure 1.17), as suggested by previous studies [214, 250, 251].

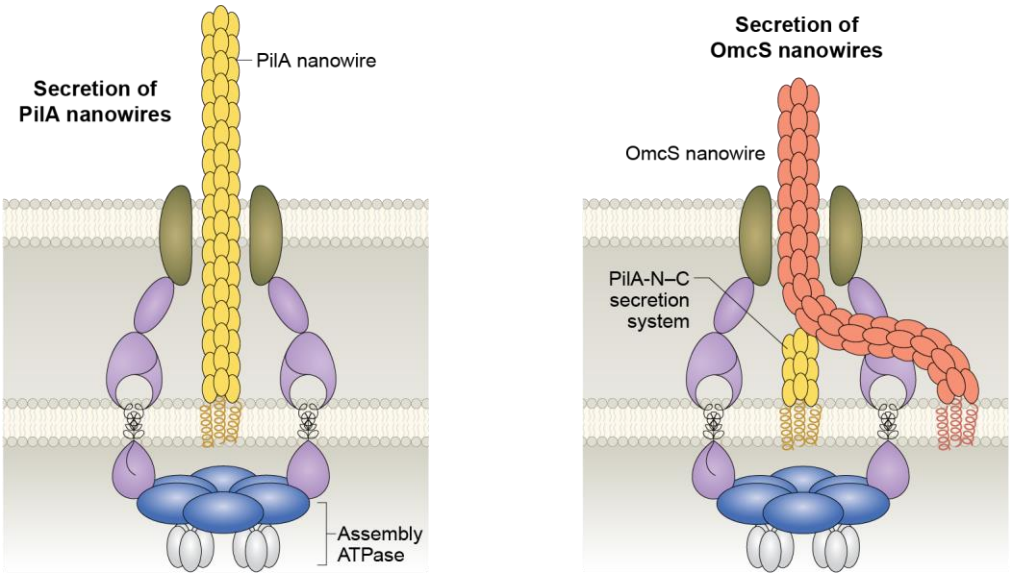


Figure 1.17. Schematic models of PilA and cytochrome nanowire secretion. On the left, a past concept, in which PilA nanowires (e-pili), composed of the monomeric PilA-N protein, are exported through an ATP-driven pilus secretion machinery spanning the inner and outer membranes, is presented (see [252] for a Review). On the right, a present model of pilus-facilitated OmcS nanowire secretion is presented [243]. In this model, the PilA-N-C filaments do not conduct electrons, rather participating in the secretion of OmcS. Importantly, this model implies that cytochrome nanowires might also play a role in periplasmic EET [184]. This figure was reproduced from [253].

This model correlates the expression of the type IV pili with the cellular localization of *Geobacter*'s cytochrome nanowires [187, 243], and explains many prior observations that mistakenly linked these filaments with the bacterium's conductive properties. Consequently, the pili-nanowire paradigm is currently under strong debate [250, 254-263]. Undeniably, the functions of the OmcE, OmcS, OmcZ and PilA-N-C filaments are deeply intertwined, as further illustrated by the fact that deletion of the GSU1771 gene, which encodes a transcriptional regulator, triggered a 5- to 100-fold change in the expression levels of all five of these genes [264, 265].

1.4.3. Electron harvesting pathways

The *G. sulfurreducens* bacterium is capable of exchanging electrons with electrode surfaces in both directions [27, 107, 266]. The mechanisms responsible for the electron transfer from the cell interior to the cell exterior are generally established, but those of the opposite pathway are poorly understood. This pathway should involve cytochromes placed at the same cell compartments to allow electrons to reach the cytoplasm, alter the cytoplasmic NADH/NAD^+ ratios and drive the bacterial metabolic processes towards the production of biofuel or desired chemical compounds [67, 267, 268]. Nevertheless, to warrant directionality to the electron transfer in an opposite direction, the pathway may not necessarily involve the same proteins. In fact, genomic and proteomic studies carried out in current-consuming vs current-producing biofilms revealed that there are specific electron transfer proteins for each direction [269-272]. For example, the OmcZ

nanowire, which was showed to be essential for electron transfer to electrode surfaces, is not crucial when electrons are supplied to cells or to drive the metabolic conversion of the terminal electron acceptor fumarate to succinate [271]. An opposite observation was made for the GSU2515 and PccH monoheme periplasmic cytochromes, which showed the two highest increased expression levels in studies performed in electron-uptake biofilms [271]. Additional genes encoding redox-active proteins showing higher transcript abundance in current-consuming cells were found, but in less abundant amounts. Additionally, their cellular localization could not be predicted with confidence [271].

The PccH cytochrome contains a unique fold with one His-Met coordinated *c*-type heme (Figure 1.18A) [273-275], and was suggested to be an electron transfer partner of GSU2515 in the periplasm of the bacterium [276].

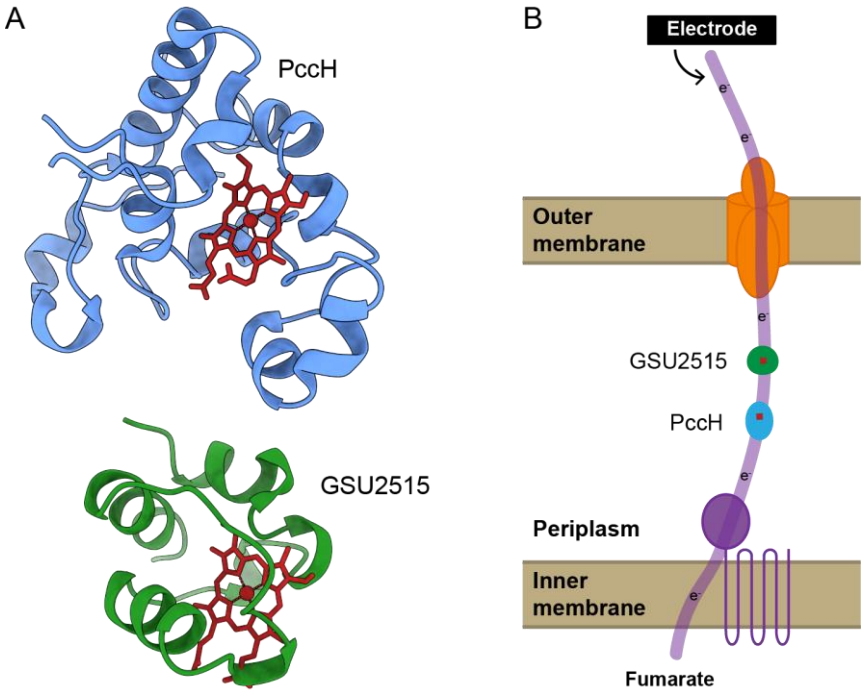


Figure 1.18. Periplasmic cytochromes crucial for the electron harvesting pathways of *G. sulfurreducens*. (A) Ribbon representations of the crystal structure of PccH (PDB: 4RLR [274], blue) and the AlphaFold 2 model of GSU2515 (green). The *c*-type heme groups are shown as red sticks. For simplicity, the disordered N-terminal domain of GSU2515 is omitted. The pLDDT score of the GSU2515 model can be found in Figure 7.1. (B) Schematic representation of the putative electron transfer pathways involving GSU2515 (green) and PccH (blue) in biofilms grown on graphite cathodes, using fumarate as terminal electron acceptor. The *c*-type heme groups are represented as red squares. This figure was partially reproduced from [277].

This was later confirmed in a study in which GSU2515 was shown to transfer electrons to PccH in a thermodynamically favorable reaction (Figure 1.18B) [277]. According to its AlphaFold 2 model, the global fold of the GSU2515 cytochrome resembles those of the cytochromes c_6 [277], known for having their heme groups placed in a highly hydrophobic core, thus possessing a relatively low heme exposed surface area (Figure 1.18A) [278]. These results clarified which are the mechanisms of

electron transfer in the periplasm of the bacterium in electron harvesting conditions, but additional studies are required to identify the key players of this process in the inner and outer membranes of the bacterium (Figure 1.18B).

1.5. Cytochromes – The chemistry of the key players in extracellular electron transfer

Primitive anaerobic bacteria transferred electrons from abundant mineral sources, including the highly available soluble ferrous iron near active volcanic sites, to terminal acceptors, leading to the fixation of carbon oxides, while generating organic matter [5]. However, with the increase of oxygen concentration, and the concomitant decrease of the soluble iron availability, organisms had to evolve not only to adapt to aerobic respiration, but also to develop strategies to obtain, transfer and store iron in an efficient manner [7, 279, 280].

Presently, iron is the fourth most abundant element and the second most abundant metal (behind aluminum) in the Earth's crust [281]. The natural abundance of iron and its chemical versatility are the basis for its natural selection for incorporation in biomolecules throughout Life's evolutionary stages [282]. Overall, organisms evolved to incorporate and utilize iron in their cellular pathways in distinct ways, such as proteins' cofactors, electron donors or electron acceptors. However, excess of iron can lead to numerous conditions, mostly due to the formation of reactive oxygen species (ROS), that damage membranes, proteins and nucleic acids, which can result in cell death [283-285]. Consequently, most aerobic organisms have evolved to safely acquire, transport and store iron [286], while efficiently acting on ROS to prevent cellular damage [287].

Iron, with an atomic number of 26, is part of the first transition series of the Periodic Table, characterized by having non-filled *d* orbitals, variable oxidation states, as well as a rich coordination and organometallic chemistry [288]. This transition metal has been observed in ten different oxidation states (from −4 to +7), although most of these are found in particular complexes and/or under specific conditions [289]. The two most common iron states found at physiological conditions are the ferrous (Fe^{2+}) and the ferric (Fe^{3+}) states, in which iron acts as an electron donor or acceptor, respectively. The reduction potential of the Fe^{2+}/Fe^{3+} pair is tuned by several factors and ranges from −700 to +700 mV vs SHE, thus covering the entire biologically relevant range of reduction potentials [290, 291]. These features make iron the ideal redox mediator for biochemical processes in living organisms.

In heme proteins, the iron atom is covalently bound to four nitrogen atoms of a protoporphyrin IX molecule (Figure 1.19). Porphyrins are heterocyclic macrocycle organic compounds composed of four pyrrole subunits connected via methine bridges. These compounds contain 26 π-electrons, 18 of which form a large, planar, circular conjugated system [292, 293]. Consequently, porphyrins have a strong absorbance in the visible region of the electromagnetic spectrum, resulting in a strong, purple red color in solution.

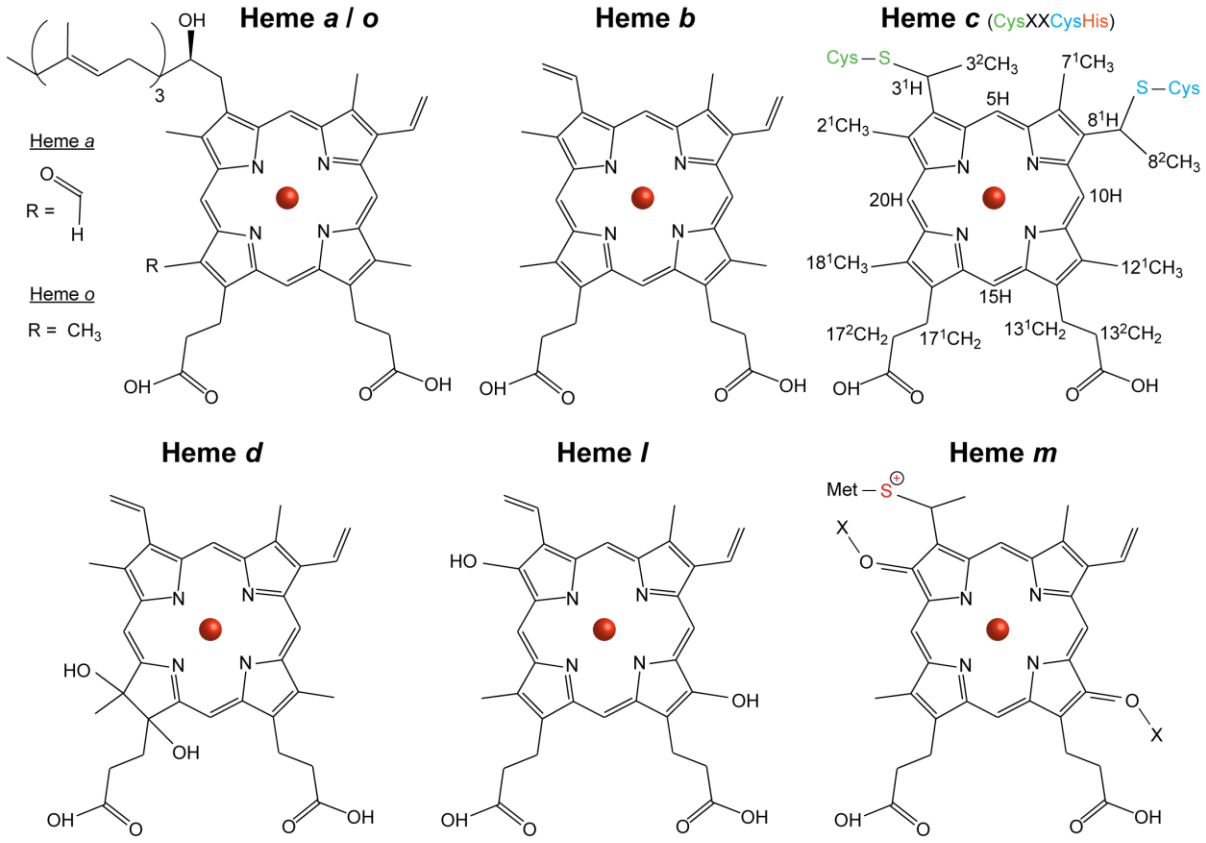


Figure 1.19. Chemical structures of different heme groups. The structures for the *a*, *b*, *c*, *d*, *l*, *m* and *o*-type hemes are represented using condensed structural formulas. The iron cation is represented by an orange sphere. In the *c*-type heme structure, the IUPAC-IUB nomenclature for tetrapyrroles is represented [294]. The two covalently attached cysteines, part of the CXXCH heme binding motif, and their respective sulfur atom, are colored green and blue. The proximal axial ligand of the *c*-type heme (His residue) is colored in orange. In the *m*-type heme structure, which is present in myeloperoxidases [295], the heme moiety is attached to three amino acids, one of which forms a vinyl-sulfonium ion linkage with a methionine side-chain (red sulfur atom). X stands for any amino acid. This figure was reproduced from [131].

Besides the covalent bonds formed with the four nitrogen atoms of the porphyrin, the iron atom can also bind one or two axial ligands. These axial ligands are usually amino acids from the protein and their nature greatly modulates the properties of the iron atom, together with the porphyrin molecule substituents and the local heme environment. In fact, different kinds of porphyrins associated with iron have distinct properties and classifications, resulting in three major types of hemes, named *a*, *b* or *c* (Figure 1.19). The different porphyrin substituents, and hence the different structural features of each type of heme, yield different visible spectroscopic signatures, with α-bands at 605, 560 and 550 nm for hemes *a*, *b* and *c*, respectively, a feature that has been used to classify different cytochromes [296]. Type *a* hemes contain a long hydrophobic tail of isoprene units attached to the porphyrin and a formyl group instead of one of the methyl substituents (Figure 1.19). Type *b* and *c* hemes have a similar structure, with the only difference being that in the latter the vinyl groups form covalent thioether bonds with cysteine residues of the protein, in the conserved binding motif CXXCH (Figure 1.19).

Variations of the hemes described above were also identified but are much less frequent and include hemes *d* [297], *l* [298], *m* [295] and *o* [299] (Figure 1.19).

Porphyrins are strong field ligands that place the energies of iron orbitals close enough, so that small energy contributions from the axial ligands can generate sufficient ligand field strength to alter the iron's spin-state. The iron atom contains five 3*d* orbitals: two orbitals of higher energy (d_{xy}^2 and d_z^2, also known as the e_g set of orbitals in Crystal Field Theory (CFT)) and three of lower energy (d_{xy}, d_{xz} and d_{yz}, also known as the t_{2g} set of orbitals in CFT) and is naturally high-spin outside an octahedral coordination complex because the energies of these orbitals are degenerate [300]. However, in an octahedral coordination complex, iron axial ligands with strong field (certain protein amino acid residues) will result in a low-spin configuration, whereas weak field ligands (usually small molecules) will result in a high-spin configuration. Thus, the nature of the axial ligands plays a crucial role in the definition of the iron spin-state.

Porphyrin is a weak dibasic acid which loses two pyrrole N-protons upon coordination of iron, thus becoming a divalent anion. This complex, in which iron is coordinated by a negatively charged ligand, significantly stabilizes its oxidized state and the reduction potential value of the Fe^{2+}/Fe^{3+} pair is shifted to −115 mV from the value of +771 mV for the aqueous species [301]. Compared to the nature of the porphyrin substituents found in the different types of hemes, the impact of the axial ligands' nature on the reduction potential values of the heme-containing proteins is substantially higher [302-305]. Most heme-containing proteins are axially coordinated by two histidine residues, but mixed coordination with histidine and other amino acids (such as cysteine, methionine, asparagine, tyrosine and lysine) is also common. In addition, one of the heme axial positions can be transiently vacant or occupied by small ligands [306, 307].

Histidine is a good electron donor and stabilizes the hemes' oxidized state, resulting in lower heme reduction potential values [308]. The opposite effect is observed when the iron is coordinated by a methionine residue in the same position because the sulfur atom in the methionine side-chain is a good electron acceptor and stabilizes the reduced state of the iron [309, 310]. Thus, typically, hemes with His-His axial coordination have lower reduction potential values compared to those with His-Met coordination.

In addition to the nature of the axial ligands, the heme reduction potential values can also be modulated by other intrinsic and extrinsic factors, such as the heme puckering [311] and exposure to the dielectric environment of the solvent [312], as well as the distribution of polar and charged residues near the heme [313-315]. The unique coordination chemistry of iron in heme proteins evolved throughout Life in such a way that the iron properties can be modulated by selective changes in the polypeptide chain and heme axial ligands to achieve maximum effectiveness in the proteins' functions.

The flexibility of cytochromes' functional properties has made them ideal and crucial components of several biological processes, ranging from electron transfer to signal transduction events across all kingdoms of Life, being found in mitochondria, chloroplasts, endoplasmic reticulum and bacterial redox chains [316]. Cytochromes are commonly known as electron transport proteins that mediate

the electron transfer between components in diverse respiratory chains. The different types of cytochromes are named according to their heme type letter, which can be followed by a number. This number depends on some of its properties, such as the number of heme groups, optical or functional properties, and the nature of its axial ligand. Furthermore, cytochromes can perform enzymatic reactions with a multitude of purposes [317-319].

The combination of two or more heme groups in the same polypeptide chain or in multiple domains has made cytochromes even more efficient and versatile. This is well illustrated by cytochromes with enzymatic activity, such as peroxidase or fumarate reductase, in which a catalytic center is readily supplied with electron(s) from heme(s) located in its vicinity [320, 321]. Multiheme cytochromes with no catalytic activity, exclusively dedicated to electron transfer, are also advantageous compared to monoheme cytochromes, particularly in the case of microorganisms subjected to considerable environmental fluctuations. Multiheme cytochromes typically extend the proteins' functional redox potential range as a result of the different contributions of each of the heme reduction potential values and heme-heme interactions between neighboring redox centers [322]. These cytochromes generally have multiple bis-histidinyl *c*-type heme groups and can receive or donate multiple electrons in a cooperative way, depending on the intrinsic properties of the neighboring hemes or of heme-surrounding protonable centers. Nevertheless, there are examples of multiheme cytochromes with mixed coordination *c*-type hemes, such as the mentioned GSU1996 cytochrome [169], or with more than one type of heme, such as the mentioned CbcL quinol dehydrogenase [159, 323]. Finally, multiheme cytochromes in which the heme groups seamlessly adopt a linear topology (such as cytochrome nanowires), endow the microorganisms with the ability to perform long range electron transfer, without the need for successive binding events, providing them with a competitive advantage over other microorganisms [145, 185, 187].

1.6. Protein engineering of multiheme cytochromes

The abundance and relevance of multiheme cytochromes in electroactive bacteria, particularly in *Geobacter*, make them the logical targets for rational protein engineering focusing on the tuning of their redox properties, resulting in improved forms of these electron transfer components. For this, several techniques have been developed to characterize multiheme cytochromes, both functionally and structurally [324-330]. These strategies have been useful to reveal the functional mechanisms and key residues involved in the redox reactions of these proteins. The data obtained for the WT cytochromes can then be conjugated with structural information and explored to design mutated forms with optimal functional properties.

The main goal of multiheme cytochrome engineering is to improve the current production output of electrogenic bacteria by either increasing the bacterium's biomass formation or by optimizing the bacterium's EET mechanisms [93]. In the first case, cytochrome mutants with enhanced e^-/H^+ mechanisms, contributing to increased membrane potential and ATP production, can be envisaged. In the second case, the design of proteins with enhanced electron-transfer driving force for either upstream or downstream partners, will contribute to the creation of bacterial cells with improved

electron transfer capabilities. Since the redox properties of the heme groups of cytochromes can be modulated considering different structural aspects, mutations can be designed to explore different chemical properties, including the heme's solvent exposure and network of interactions, such as surrounding hydrogen bonds and ionizable residues. Therefore, depending on the location and nature of the targeted residues, different strategies can be applied.

Over the last 15 years, several studies performed on PpcA mutants have shown that the replacement of residues that alter the heme solvent exposure might be explored to modify the functional properties of the protein, however, the most promising substitutions involve charge alterations in the vicinity of the heme groups [331-336]. A summary of these results is presented in Figure 1.20.

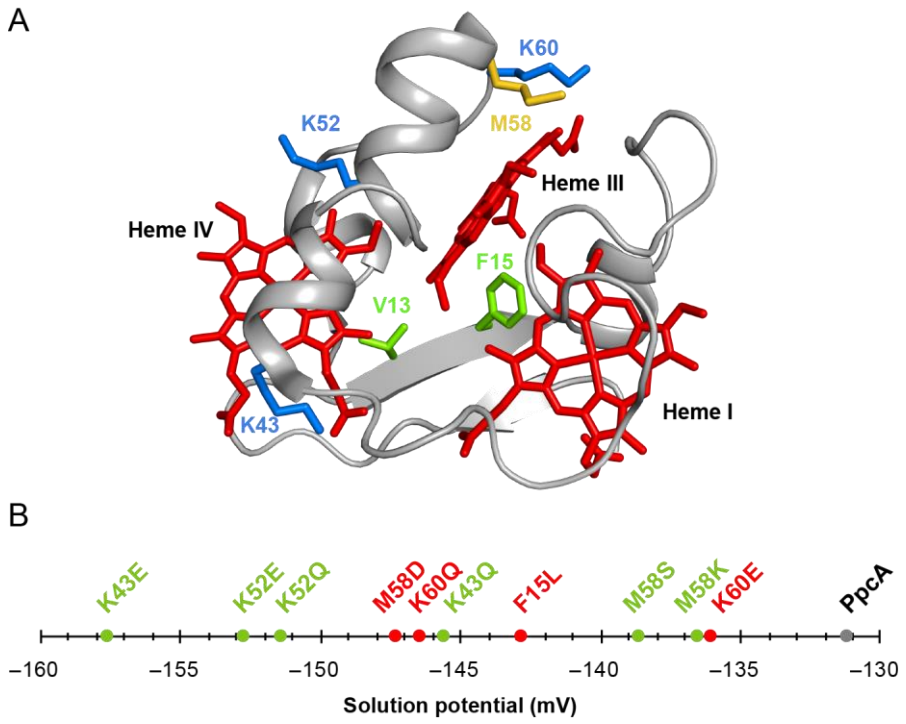


Figure 1.20. Rational design of PpcA mutants. (A) Spatial location of the residues mutated in the PpcA solution structure. The PpcA polypeptide chain (PDB: 2MZ9 [337]) is shown as a Cα ribbon in gray, with the heme groups in red. The side-chains of V13 and F15 are represented in green (conserved residues found in the PpcA-family cytochromes); K43, K52 and K60 in blue (positively charged lysine residues in the vicinity of the hemes); and M58 in yellow (residue that controls heme III's solvent accessibility), all in stick drawings. (B) Apparent midpoint reduction potentials of PpcA mutants. The mutants that retained their e^-/H^+ transfer capabilities are labeled in green, while mutants for which the preferential pathway is disrupted are labeled in red. WT PpcA is labeled in gray. These panels were reproduced from [93].

The rational principles behind these mutations are transverse and can be applied to other proteins [338], constituting the first step towards the engineering of optimized *Geobacter*-mutated strains, with optimal applicability in MET. The second step is to realistically evaluate the effect of these mutated proteins in the bacterium's electron transfer capabilities when incorporated into cells. In 2003, Lloyd and co-workers used a *G. sulfurreducens* strain with the *ppcA* gene knocked out to show

that the bacterial growth with fumarate as an electron acceptor is not affected, while the growth rate significantly decreased when Fe(III) citrate was used as an electron acceptor [339]. The ability of the bacterium to grow with Fe(III) citrate was restored when PpcA was expressed in *trans* using a complementation plasmid inserted by electroporation [339]. A similar strategy (Figure 1.21) can be envisaged to probe the effects of the PpcA mutants, using either the *G. sulfurreducens* strain with the *ppcA* gene knocked out [339] or, ideally, a *G. sulfurreducens* strain with all five cytochromes from the PpcA-family knocked out [174], to allow a more straightforward interpretation of results caused by the expression of the mutated forms of PpcA.

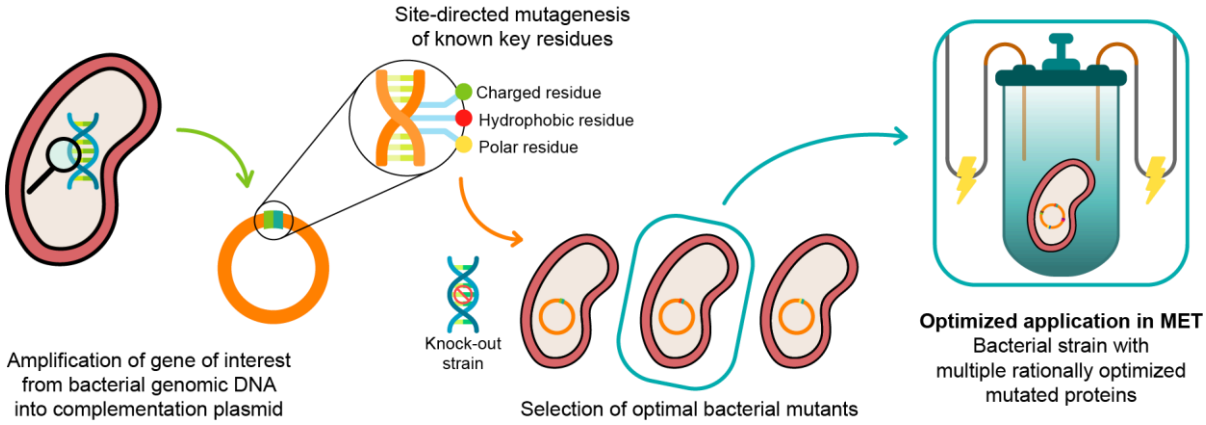


Figure 1.21. Schematic picture of the preparation of *G. sulfurreducens* strains with mutated cytochromes. After the bacterial genome is extracted, the specific target gene is amplified and further inserted into a complementation plasmid. The gene is mutated in key residues through site-directed mutagenesis, and the resultant mutants are inserted into a bacterial knock-out strain. The EET capabilities of the engineered bacteria are then tested in media with different electron acceptors. This process is applied to different key EET components of the bacterium, and the optimized mutants are selected and conjugated. The resultant strain, with higher electron transfer driving force and current production, is finally applied in MET. This figure was reproduced from [93].

After the production of the mutated strains, their viability must be evaluated by analyzing the growth curves in the recommended standard bacterial media for *G. sulfurreducens* (NBAF [122]), so they can later be studied in media with different electron acceptors and their mutations evaluated in terms of electron transfer efficiency and current production. However, a single mutation in a single cytochrome should not cause significant changes *per se* in the electron transfer efficiency of the bacterium. Therefore, after the impact and viability of a certain mutation are evaluated, other mutated forms of cytochromes belonging to the same EET pathway must be inserted using a similar methodology to improve the directionality of the electron transfer. For that purpose, the creation of a "stripped-down" strain, trimming the redundant pathways of *G. sulfurreducens*, would constitute an excellent vehicle for the implementation of this strategy [58, 93]. Based on our current knowledge of the EET pathways, a simplified, controlled and clearly defined electron transfer chain at each subcellular location can be envisioned (Figure 1.22).

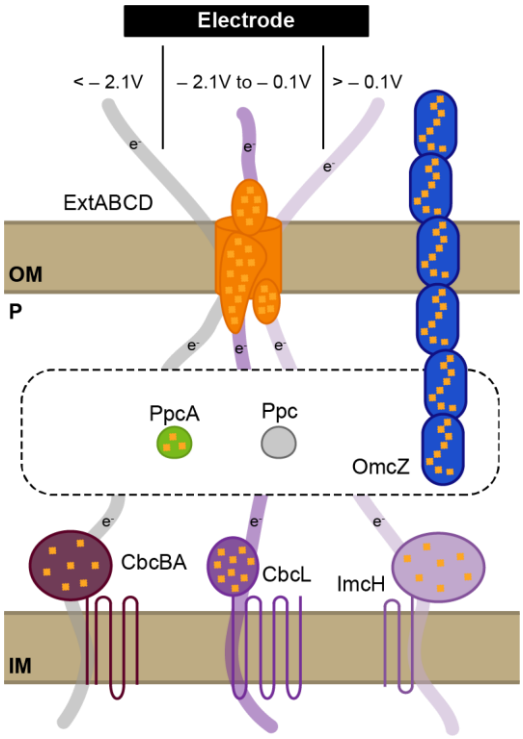


Figure 1.22. Simplified *Geobacter* EET pathways towards electrodes. The quinol dehydrogenase required for a specific MET would depend on the applied electrode potential, while the periplasmic electron transfer could be solely dependent on the highly abundant and versatile triheme cytochrome PpcA. Alternatively, other suitable periplasmic cytochrome could also be selected, since the PpcA-family cytochromes have been shown not to be essential for EET towards electrodes [175]. The final electron transfer steps would be dependent either on the ExtABCD porin-cytochrome complex (see section 2.3.2) or on the OmcZ nanowire, which might also participate in the periplasmic electron transfer. IM, OM and P stand for inner membrane, outer membrane and periplasm, respectively. This figure was adapted from [131].

This approach would be labor-intensive and time-consuming due to the considerable number of cytochromes expressed by *G. sulfurreducens* and substantial redundancy of the EET pathways [58]. Alternatively, as shown previously, one might use genetically tractable bacteria with no or few cytochromes, such as *E. coli*, to engineer heterologous electron transport chains [96, 340].

Importantly, the development of optimized *Geobacter*-mutated strains through rational protein engineering of multiheme cytochromes relies on the principle that the multiple protein components of a certain EET pathway are not limiting electron transfer. While this has not been exactly confirmed directly in *G. sulfurreducens* [341, 342], experimental values tied with a few theoretical assumptions indicate that electron transfer through the MtrABC complex from *S. oneidensis*, which possesses a similar organization of its EET chains, is not limiting [343-348].

1.7. Objectives and thesis outline

The structural and functional characterization of the individual electron transfer components of the EET pathways of *Geobacter* bacteria is a fundamental step towards the sustainable development of improved mutated strains, with increased respiratory rates and optimal application in MET. The work performed in this Thesis links the fundamental biophysical aspects of multiheme cytochromes and other proteins linked to EET, allowing their rational modulation for different biotechnological applications.

The goal of this Thesis was to biochemically characterize specific components of the EET pathways of *G. sulfurreducens*, thereby contributing to the fundamental understanding of their properties and the bacterium's physiology. These targets were selected based either on their unique features or on their relevant role in the EET pathways of the bacterium.

To properly engage the reader into the context and background of our research, this initial chapter provided a state-of-the-art overview of electroactive microorganisms, their EET mechanisms and how their components can be modulated to design optimal biotechnological applications.

In the second chapter, we make use of the AlphaFold 2 algorithm to gain insights into the structural features and organization of the five porin-cytochrome complexes of *G. sulfurreducens*, establishing guidelines that might be used for optimized cloning, expression and purification of these proteins, while updating our general understanding of the function and architecture of these complexes.

The third chapter is focused on the characterization of ExtJ, a small periplasmic protein from the putative porin-cytochrome complex ExtHIJKL, with no predicted cofactors, but yet linked to an EET-dedicated protein complex.

Biophysical insights into the features of PgcA, a freely-diffusing cytochrome particularly relevant for extracellular Fe(III) oxides reduction, are presented in detail in the fourth chapter. The unique features of PgcA suggest that this protein forms a new class of cytochromes – microbial heme-tethered redox strings.

In the fifth chapter, we characterized GSU0105, a triheme cytochrome with unprecedent structural and functional features, which might work as a complementary player to the PpcA-family triheme cytochromes in the promiscuous periplasmic electron transfer network.

The final chapter of this work contextualizes our findings, clarifies the ongoing studies and establishes the foundations for the following steps towards developing sustainable MET.

1.8. References

[1] G.B. Dalrymple, The age of the Earth in the twentieth century: A problem (mostly) solved, Geol. Soc. S. P., 190 (2001) 205-221.
[2] C.P. McKay, Requirements and limits for life in the context of exoplanets, Proc. Natl. Acad. Sci., 111 (2014) 12628-12633.
[3] A. Pross, R. Pascal, The origin of Life: What we know, what we can know and what we will never know, Open Biol., 3 (2013) 120190.
[4] K. Zahnle, L. Schaefer, B. Fegley, Earth's earliest atmospheres, Cold Spring Harb. Perspect. Biol., 2 (2010) a004895.
[5] D.E. Canfield, M.T. Rosing, C. Bjerrum, Early anaerobic metabolisms, Philos. T. R. Soc. B: Biol. Sci., 361 (2006) 1819-1836.
[6] P. Sánchez-Baracaldo, T. Cardona, On the origin of oxygenic photosynthesis and cyanobacteria, New Phytol., 225 (2020) 1440-1446.
[7] H.D. Holland, The oxygenation of the atmosphere and oceans, Philos. T. R. Soc. B: Biol. Sci., 361 (2006) 903-915.
[8] T.W. Lyons, C.T. Reinhard, N.J. Planavsky, The rise of oxygen in Earth's early ocean and atmosphere, Nature, 506 (2014) 307-315.
[9] K.N. Laland, T. Uller, M.W. Feldman, K. Sterelny, G.B. Müller, A. Moczek, E. Jablonka, J. Odling-Smee, The extended evolutionary synthesis: Its structure, assumptions and predictions, Proc. R. Soc. B: Biol. Sci., 282 (2015) 20151019.
[10] A. Booth, C. Mariscal, W.F. Doolittle, The modern synthesis in the light of microbial genomics, Annu. Rev. Microbiol., 70 (2016) 279-297.
[11] M.A. Nowak, Five rules for the evolution of cooperation, Science, 314 (2006) 1560-1563.
[12] D.H. Erwin, Climate as a driver of evolutionary change, Curr. Biol., 19 (2009) R575-R583.
[13] L.A. Hug, B.J. Baker, K. Anantharaman, C.T. Brown, A.J. Probst, C.J. Castelle, C.N. Butterfield, A.W. Hernsdorf, Y. Amano, K. Ise, Y. Suzuki, N. Dudek, D.A. Relman, K.M. Finstad, R. Amundson, B.C. Thomas, J.F. Banfield, A new view of the tree of life, Nat. Microbiol., 1 (2016) 16048.
[14] C. Mora, D.P. Tittensor, S. Adl, A.G.B. Simpson, B. Worm, How many species are there on Earth and in the ocean?, PLoS Biol., 9 (2011) e1001127.
[15] E.B. Graham, J.E. Knelman, A. Schindlbacher, S. Siciliano, M. Breulmann, A. Yannarell, J.M. Beman, G. Abell, L. Philippot, J. Prosser, A. Foulquier, J.C. Yuste, H.C. Glanville, D.L. Jones, R. Angel, J. Salminen, R.J. Newton, H. Bürgmann, L.J. Ingram, U. Hamer, H.M.P. Siljanen, K. Peltoniemi, K. Potthast, L. Bañeras, M. Hartmann, S. Banerjee, R.-Q. Yu, G. Nogaro, A. Richter, M. Koranda, S.C. Castle, M. Goberna, B. Song, A. Chatterjee, O.C. Nunes, A.R. Lopes, Y. Cao, A. Kaisermann, S. Hallin, M.S. Strickland, J. Garcia-Pausas, J. Barba, H. Kang, K. Isobe, S. Papaspyrou, R. Pastorelli, A. Lagomarsino, E.S. Lindström, N. Basiliko, D.R. Nemergut, Microbes as engines of ecosystem function: When does community structure enhance predictions of ecosystem processes?, Front. Microbiol., 7 (2016) 214.

[16] K. Timmis, V. de Lorenzo, W. Verstraete, J.L. Ramos, A. Danchin, H. Brüssow, B.K. Singh, J.K. Timmis, The contribution of microbial biotechnology to economic growth and employment creation, Microb. Biotechnol., 10 (2017) 1137-1144.
[17] B.E. Logan, R. Rossi, A. Ragab, P.E. Saikaly, Electroactive microorganisms in bioelectrochemical systems, Nat. Rev. Microbiol., 17 (2019) 307-319.
[18] H. Gest, The discovery of microorganisms by Robert Hooke and Antoni van Leeuwenhoek, Fellows of The Royal Society, Notes Rec. R. Soc. Lond., 58 (2004) 187-201.
[19] M.X. Byndloss, A.J. Bäumler, The germ-organ theory of non-communicable diseases, Nat. Rev. Microbiol., 16 (2018) 103-110.
[20] A. Ullmann, Pasteur-Koch: Distinctive ways of thinking about infectious diseases, Microbe, 2 (2007) 383-387.
[21] M.C. Potter, Electrical effects accompanying the decomposition of organic compounds, Proc. R. Soc. B: Biol. Sci., 84 (1911) 260-276.
[22] C.R. Myers, K.H. Nealson, Bacterial manganese reduction and growth with manganese oxide as the sole electron acceptor, Science, 240 (1988) 1319-1321.
[23] D.R. Lovley, S.J. Giovannoni, D.C. White, J.E. Champine, E.J. Phillips, Y.A. Gorby, S. Goodwin, *Geobacter metallireducens* gen. nov. sp. nov., a microorganism capable of coupling the complete oxidation of organic compounds to the reduction of iron and other metals, Arch. Microbiol., 159 (1993) 336-344.
[24] B.H. Kim, H.J. Kim, M.S. Hyun, D.H. Park, Direct electrode reaction of Fe(III)-reducing bacterium, *Shewanella putrefaciens*, J. Microbiol. Biotechnol., 9 (1999) 127-131.
[25] H.J. Kim, M.S. Hyun, I.S. Chang, B.H. Kim, A microbial fuel cell type lactate biosensor using a metal-reducing bacterium, *Shewanella putrefaciens*, J. Microbiol. Biotechnol., 9 (1999) 365-367.
[26] B.H. Kim, T. Ikeda, H.S. Park, H.J. Kim, M.S. Hyun, K. Kano, K. Takagi, H. Tatsumi, Electrochemical activity of an Fe(III)-reducing bacterium, *Shewanella putrefaciens* IR-1, in the presence of alternative electron acceptors, Biotechnol. Tech., 13 (1999) 475-478.
[27] D.R. Lovley, Electromicrobiology, Annu. Rev. Microbiol., 66 (2012) 391-409.
[28] C. Koch, F. Harnisch, Is there a specific ecological niche for electroactive microorganisms?, ChemElectroChem, 3 (2016) 1282-1295.
[29] L. Digel, R. Bonné, K. Aiyer, Are all microbes electroactive?, Cell Rep. Phys. Sci., 5 (2024) 102200.
[30] T. Zheng, J. Li, Y. Ji, W. Zhang, Y. Fang, F. Xin, W. Dong, P. Wei, J. Ma, M. Jiang, Progress and prospects of bioelectrochemical systems: Electron transfer and its applications in the microbial metabolism, Front. Bioeng. Biotechnol., 8 (2020) 10.
[31] A.D. Goldman, J.M. Weber, D.E. LaRowe, L.M. Barge, Electron transport chains as a window into the earliest stages of evolution, Proc. Natl. Acad. Sci., 120 (2023) e2210924120.
[32] S.H. Light, L. Su, R. Rivera-Lugo, J.A. Cornejo, A. Louie, A.T. Iavarone, C.M. Ajo-Franklin, D.A. Portnoy, A flavin-based extracellular electron transfer mechanism in diverse Gram-positive bacteria, Nature, 562 (2018) 140-144.

[33] E. Marsili, D.B. Baron, I.D. Shikhare, D. Coursolle, J.A. Gralnick, D.R. Bond, *Shewanella* secretes flavins that mediate extracellular electron transfer, Proc. Natl. Acad. Sci., 105 (2008) 3968-3973.
[34] A. Okamoto, K. Saito, K. Inoue, K.H. Nealson, K. Hashimoto, R. Nakamura, Uptake of self-secreted flavins as bound cofactors for extracellular electron transfer in *Geobacter* species, Energy Environ. Sci., 7 (2014) 1357-1361.
[35] H. von Canstein, J. Ogawa, S. Shimizu, R. Lloyd Jonathan, Secretion of flavins by *Shewanella* species and their role in extracellular electron transfer, Appl. Environ. Microbiol., 74 (2008) 615-623.
[36] C.M. Paquete, L. Morgado, C.A. Salgueiro, R.O. Louro, Molecular mechanisms of microbial extracellular electron transfer: The importance of multiheme cytochromes, Front. Biosci. L., 27 (2022) 174.
[37] L. Shi, H. Dong, G. Reguera, H. Beyenal, A. Lu, J. Liu, H.-Q. Yu, J.K. Fredrickson, Extracellular electron transfer mechanisms between microorganisms and minerals, Nat. Rev. Microbiol., 14 (2016) 651.
[38] N. Sekar, C.-H. Wu, M.W.W. Adams, R.P. Ramasamy, Electricity generation by *Pyrococcus furiosus* in microbial fuel cells operated at 90 °C, Biotechnol. Bioeng., 114 (2017) 1419-1427.
[39] Y.D. Yilmazel, X. Zhu, K.-Y. Kim, D.E. Holmes, B.E. Logan, Electrical current generation in microbial electrolysis cells by hyperthermophilic archaea *Ferroglobus placidus* and *Geoglobus ahangari*, Bioelectrochemistry, 119 (2018) 142-149.
[40] D. Gupta, K. Chen, S.J. Elliott, D.D. Nayak, MmcA is an electron conduit that facilitates both intracellular and extracellular electron transport in *Methanosarcina acetivorans*, Nat. Commun., 15 (2024) 3300.
[41] K. Gao, Y. Lu, Putative extracellular electron transfer in methanogenic archaea, Front. Microbiol., 12 (2021) 611739.
[42] X. Zhang, G.H. Joyce, A.O. Leu, J. Zhao, H. Rabiee, B. Virdis, G.W. Tyson, Z. Yuan, S.J. McIlroy, S. Hu, Multi-heme cytochrome-mediated extracellular electron transfer by the anaerobic methanotroph '*Candidatus* Methanoperedens nitroreducens', Nat. Commun., 14 (2023) 6118.
[43] P. Mani, F. V T, K. Bowman, C. T S, T. Keshavarz, G. Kyazze, Development of an electroactive aerobic biocathode for microbial fuel cell applications, Env. Microbiol. Rep., 12 (2020) 607-612.
[44] L.P. Nielsen, N. Risgaard-Petersen, Rethinking sediment biogeochemistry after the discovery of electric currents, Annu. Rev. Mar. Sci., 7 (2015) 425-442.
[45] C. Pfeffer, S. Larsen, J. Song, M. Dong, F. Besenbacher, R.L. Meyer, K.U. Kjeldsen, L. Schreiber, Y.A. Gorby, M.Y. El-Naggar, K.M. Leung, A. Schramm, N. Risgaard-Petersen, L.P. Nielsen, Filamentous bacteria transport electrons over centimetre distances, Nature, 491 (2012) 218-221.
[46] F.J.R. Meysman, R. Cornelissen, S. Trashin, R. Bonné, S.H. Martinez, J. van der Veen, C.J. Blom, C. Karman, J.-L. Hou, R.T. Eachambadi, J.S. Geelhoed, K.D. Wael, H.J.E. Beaumont, B. Cleuren, R. Valcke, H.S.J. van der Zant, H.T.S. Boschker, J.V. Manca, A highly conductive fibre network enables centimetre-scale electron transport in multicellular cable bacteria, Nat. Commun., 10 (2019) 4120.
[47] J. Liu, S. Chakraborty, P. Hosseinzadeh, Y. Yu, S. Tian, I. Petrik, A. Bhagi, Y. Lu, Metalloproteins containing cytochrome, iron-sulfur, or copper redox centers, Chem. Rev., 114 (2014) 4366-4469.

[48] H.T.S. Boschker, P.L.M. Cook, L. Polerecky, R.T. Eachambadi, H. Lozano, S. Hidalgo-Martinez, D. Khalenkow, V. Spampinato, N. Claes, P. Kundu, D. Wang, S. Bals, K.K. Sand, F. Cavezza, T. Hauffman, J.T. Bjerg, A.G. Skirtach, K. Kochan, M. McKee, B. Wood, D. Bedolla, A. Gianoncelli, N.M.J. Geerlings, N. Van Gerven, H. Remaut, J.S. Geelhoed, R. Millan-Solsona, L. Fumagalli, L.P. Nielsen, A. Franquet, J.V. Manca, G. Gomila, F.J.R. Meysman, Efficient long-range conduction in cable bacteria through nickel protein wires, Nat. Commun., 12 (2021) 3996.
[49] A. Hiralal, J.S. Geelhoed, S. Neukirchen, F.J.R. Meysman, Comparative genomic analysis of nickel homeostasis in cable bacteria, BMC Genomics, 25 (2024) 692.
[50] D. Pankratov, S. Hidalgo Martinez, C. Karman, A. Gerzhik, G. Gomila, S. Trashin, H.T.S. Boschker, J.S. Geelhoed, D. Mayer, K. De Wael, F. J.R. Meysman, The organo-metal-like nature of long-range conduction in cable bacteria, Bioelectrochemistry, 157 (2024) 108675.
[51] S.V. Raghavulu, R.K. Goud, P.N. Sarma, S.V. Mohan, *Saccharomyces cerevisiae* as anodic biocatalyst for power generation in biofuel cell: Influence of redox condition and substrate load, Bioresour. Technol., 102 (2011) 2751-2757.
[52] S. Wu, Y. Xiao, L. Wang, Y. Zheng, K. Chang, Z. Zheng, Z. Yang, J.R. Varcoe, F. Zhao, Extracellular electron transfer mediated by flavins in Gram-positive *Bacillus* sp. WS-XY1 and yeast *Pichia stipitis*, Electrochim. Acta, 146 (2014) 564-567.
[53] Y. Hubenova, M. Mitov, Extracellular electron transfer in yeast-based biofuel cells: A review, Bioelectrochemistry, 106 (2015) 177-185.
[54] A. Kumar, L.H.-H. Hsu, P. Kavanagh, F. Barrière, P.N.L. Lens, L. Lapinsonnière, J.H. Lienhard V, U. Schröder, X. Jiang, D. Leech, The ins and outs of microorganism-electrode electron transfer reactions, Nat. Rev. Chem., 1 (2017) 0024.
[55] G. Reguera, Biological electron transport goes the extra mile, Proc. Natl. Acad. Sci., 115 (2018) 5632-5634.
[56] S. Gupte, E.S. Wu, L. Hoechli, M. Hoechli, K. Jacobson, A.E. Sowers, C.R. Hackenbrock, Relationship between lateral diffusion, collision frequency, and electron transfer of mitochondrial inner membrane oxidation-reduction components, Proc. Natl. Acad. Sci., 81 (1984) 2606-2610.
[57] G. Unden, J. Bongaerts, Alternative respiratory pathways of *Escherichia coli*: Energetics and transcriptional regulation in response to electron acceptors, Biochim. Biophys. Acta - Bioenergetics, 1320 (1997) 217-234.
[58] T. Ueki, Cytochromes in extracellular electron transfer in *Geobacter*, Appl. Environ. Microbiol., 87 (2021) e03109-03120.
[59] O.M. Gomaa, N.L. Costa, C.M. Paquete, Electron transfer in Gram-positive bacteria: Enhancement strategies for bioelectrochemical applications, World J. Microb. Biot., 38 (2022) 83.
[60] Q.Z. Wu, L.W. Cao, Z.H. Liu, W.Z. Du, C.C. Wang, W.W. Li, Endogenous antibiotics as extracellular electron shuttles for *Microbacterium* spp., ACS Electrochem., (2024).
[61] X. Liu, L. Shi, J.-D. Gu, Microbial electrocatalysis: Redox mediators responsible for extracellular electron transfer, Biotechnol. Adv., 36 (2018) 1815-1827.

[62] X. Wang, F. Aulenta, S. Puig, A. Esteve-Núñez, Y. He, Y. Mu, K. Rabaey, Microbial electrochemistry for bioremediation, Environ. Sci. Ecotech., 1 (2020) 100013.
[63] J.B.A. Arends, W. Verstraete, 100 years of microbial electricity production: Three concepts for the future, Microb. Biotechnol., 5 (2012) 333-346.
[64] D.R. Lovley, Microbial fuel cells: Novel microbial physiologies and engineering approaches, Curr. Opin. Biotechnol., 17 (2006) 327-332.
[65] B.E. Logan, Exoelectrogenic bacteria that power microbial fuel cells, Nat. Rev. Microbiol., 7 (2009) 375-381.
[66] D.R. Lovley, Cleaning up with genomics: Applying molecular biology to bioremediation, Nat. Rev. Microbiol., 1 (2003) 35-44.
[67] P. Dessì, L. Rovira-Alsina, C. Sánchez, G.K. Dinesh, W. Tong, P. Chatterjee, M. Tedesco, P. Farràs, H.M.V. Hamelers, S. Puig, Microbial electrosynthesis: Towards sustainable biorefineries for production of green chemicals from CO_2 emissions, Biotechnol. Adv., 46 (2021) 107675.
[68] D.R. Lovley, e-Biologics: Fabrication of sustainable electronics with "green" biological materials, mBio, 8 (2017) e00695.
[69] N.L. Ing, M.Y. El-Naggar, A.I. Hochbaum, Going the distance: Long-range conductivity in protein and peptide bioelectronic materials, J. Phys. Chem. B, 122 (2018) 10403-10423.
[70] M.J. Guberman-Pfeffer, N.-M. Dorval Courchesne, D.R. Lovley, Microbial nanowires for sustainable electronics, Nat. Rev. Bioeng., 2 (2024) 869-886.
[71] A.F. Smith, X. Liu, T.L. Woodard, T. Fu, T. Emrick, J.M. Jiménez, D.R. Lovley, J. Yao, Bioelectronic protein nanowire sensors for ammonia detection, Nano Res., 13 (2020) 1479-1484.
[72] X. Liu, H. Gao, J.E. Ward, X. Liu, B. Yin, T. Fu, J. Chen, D.R. Lovley, J. Yao, Power generation from ambient humidity using protein nanowires, Nature, 578 (2020) 550-554.
[73] L.C. Soares, S. Bertazzo, T.A.L. Burgo, V. Baldim, F. Galembeck, A new mechanism for the electrostatic charge build-up and dissipation in dielectrics, J. Braz. Chem. Soc., 19 (2008) 277-286.
[74] R.F. Gouveia, F. Galembeck, Electrostatic charging of hydrophilic particles due to water adsorption, J. Am. Chem. Soc., 131 (2009) 11381-11386.
[75] T.R.D. Ducati, L.H. Simões, F. Galembeck, Charge partitioning at gas-solid interfaces: Humidity causes electricity buildup on metals, Langmuir, 26 (2010) 13763-13766.
[76] S. Wu, H. Li, X. Zhou, P. Liang, X. Zhang, Y. Jiang, X. Huang, A novel pilot-scale stacked microbial fuel cell for efficient electricity generation and wastewater treatment, Water Res., 98 (2016) 396-403.
[77] R.T. Anderson, H.A. Vrionis, I. Ortiz-Bernad, C.T. Resch, P.E. Long, R. Dayvault, K. Karp, S. Marutzky, D.R. Metzler, A. Peacock, D.C. White, M. Lowe, D.R. Lovley, Stimulating the *in situ* activity of *Geobacter* species to remove uranium from the groundwater of a uranium-contaminated aquifer, Appl. Environ. Microbiol., 69 (2003) 5884-5891.
[78] X.A. Walter, J. You, J. Winfield, U. Bajarunas, J. Greenman, I.A. Ieropoulos, From the lab to the field: Self-stratifying microbial fuel cells stacks directly powering lights, Appl. Energy, 277 (2020) 115514.

[79] A.S. Mathuriya, D.A. Jadhav, M.M. Ghangrekar, Architectural adaptations of microbial fuel cells, Appl. Microbiol. Biot., 102 (2018) 9419-9432.
[80] S. Grover, L.E. Doyle, Advanced electrode materials for microbial extracellular electron transfer, Trends Chem., 6 (2024) 144-158.
[81] M. Zhou, M. Chi, J. Luo, H. He, T. Jin, An overview of electrode materials in microbial fuel cells, J. Power Sources, 196 (2011) 4427-4435.
[82] D. Jiang, B. Li, W. Jia, Y. Lei, Effect of inoculum types on bacterial adhesion and power production in microbial fuel cells, Appl. Biochem. Biotech., 160 (2009) 182.
[83] F. Li, Y. Sharma, Y. Lei, B. Li, Q. Zhou, Microbial fuel cells: The effects of configurations, electrolyte solutions, and electrode materials on power generation, Appl. Biochem. Biotech., 160 (2009) 168.
[84] O. Schaetzle, F. Barrière, K. Baronian, Bacteria and yeasts as catalysts in microbial fuel cells: Electron transfer from micro-organisms to electrodes for green electricity, Energy Environ. Sci., 1 (2008) 607-620.
[85] H. Yi, K.P. Nevin, B.-C. Kim, A.E. Franks, A. Klimes, L.M. Tender, D.R. Lovley, Selection of a variant of *Geobacter sulfurreducens* with enhanced capacity for current production in microbial fuel cells, Biosens. Bioelectron., 24 (2009) 3498-3503.
[86] A. Hirose, T. Kasai, R. Koga, Y. Suzuki, A. Kouzuma, K. Watanabe, Understanding and engineering electrochemically active bacteria for sustainable biotechnology, Bioresour. Bioprocess., 6 (2019) 10.
[87] J. Zhao, F. Li, Y. Cao, X. Zhang, T. Chen, H. Song, Z. Wang, Microbial extracellular electron transfer and strategies for engineering electroactive microorganisms, Biotechnol. Adv., 53 (2020) 107682.
[88] J. Zhang, F. Li, D. Liu, Q. Liu, H. Song, Engineering extracellular electron transfer pathways of electroactive microorganisms by synthetic biology for energy and chemicals production, Chem. Soc. Rev., 53 (2024) 1375-1446.
[89] W. Tu, I.P. Thompson, W.E. Huang, Engineering bionanoreactor in bacteria for efficient hydrogen production, Proc. Natl. Acad. Sci., 121 (2024) e2404958121.
[90] F. Li, H. Yu, B. Zhang, C. Hu, F. Lan, Y. Wang, Z. You, Q. Liu, R. Tang, J. Zhang, C. Li, L. Shi, W.-W. Li, K.H. Nealson, Z. Liu, H. Song, Engineered cell elongation promotes extracellular electron transfer of *Shewanella oneidensis*, Adv. Sci., 11 (2024) 2403067.
[91] M. Mouhib, M. Reggente, L. Li, N. Schuergers, A.A. Boghossian, Extracellular electron transfer pathways to enhance the electroactivity of modified *Escherichia coli*, Joule, 7 (2023) 2092-2106.
[92] B. Qin, G. Yang, X. Chen, X. Wu, Y. Fang, X. Quan, L. Zhuang, Specific interaction of resorufin to outer-membrane cytochrome OmcE of *Geobacter sulfurreducens*: A new insight on artificial electron mediators in promoting extracellular electron transfer, Water Res., 266 (2024) 122403.
[93] T.M. Fernandes, L. Morgado, D.L. Turner, C.A. Salgueiro, Protein engineering of electron transfer components from electroactive *Geobacter* bacteria, Antioxidants, 10 (2021) 844.
[94] A.S. Alves, N.L. Costa, M. Tien, R.O. Louro, C.M. Paquete, Modulation of the reactivity of multiheme cytochromes by site-directed mutagenesis: moving towards the optimization of microbial electrochemical technologies, J. Biol. Inorg. Chem., 22 (2017) 87-97.

[95] J. Zhang, Z. You, D. Liu, R. Tang, C. Zhao, Y. Cao, F. Li, H. Song, Conductive proteins-based extracellular electron transfer of electroactive microorganisms, Quant. Biol., 11 (2023) 405-420.
[96] H.M. Jensen, A.E. Albers, K.R. Malley, Y.Y. Londer, B.E. Cohen, B.A. Helms, P. Weigele, J.T. Groves, C.M. Ajo-Franklin, Engineering of a synthetic electron conduit in living cells, Proc. Natl. Acad. Sci., 107 (2010) 19213.
[97] J.S. Gescher, C.D. Cordova, A.M. Spormann, Dissimilatory iron reduction in *Escherichia coli*: Identification of CymA of *Shewanella oneidensis* and NapC of *E. coli* as ferric reductases, Mol. Microbiol., 68 (2008) 706-719.
[98] K.E. Pitts, P.S. Dobbin, F. Reyes-Ramirez, A.J. Thomson, D.J. Richardson, H.E. Seward, Characterization of the *Shewanella oneidensis* MR-1 decaheme cytochrome MtrA: Expression in *Escherichia coli* confers the ability to reduce soluble Fe(III) chelates, J. Biol. Chem., 278 (2003) 27758-27765.
[99] M.J. Edwards, G.F. White, J.N. Butt, D.J. Richardson, T.A. Clarke, The crystal structure of a biological insulated transmembrane molecular wire, Cell, 181 (2020) 665-673.e610.
[100] L. Thöny-Meyer, Biogenesis of respiratory cytochromes in bacteria, Microbiol. Mol. Biol. Rev., 61 (1997) 337-376.
[101] L. Thöny-Meyer, F. Fischer, P. Kunzler, D. Ritz, H. Hennecke, *Escherichia coli* genes required for cytochrome *c* maturation, J. Bacteriol., 177 (1995) 4321-4326.
[102] M. Jinek, K. Chylinski, I. Fonfara, M. Hauer, J.A. Doudna, E. Charpentier, A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity, Science, 337 (2012) 816-821.
[103] J.A. Doudna, E. Charpentier, The new frontier of genome engineering with CRISPR-Cas9, Science, 346 (2014) 1258096.
[104] S. Shanmugam, H.-H. Ngo, Y.-R. Wu, Advanced CRISPR/Cas-based genome editing tools for microbial biofuels production: A review, Renew. Energ., 149 (2020) 1107-1119.
[105] W. Ding, Y. Zhang, S. Shi, Development and application of CRISPR/Cas in microbial biotechnology, Front. Bioeng. Biotechnol., 8 (2020) 711.
[106] W.-Q. Lin, Z.-H. Cheng, Q.-Z. Wu, J.-Q. Liu, D.-F. Liu, G.-P. Sheng, Efficient enhancement of extracellular electron transfer in *Shewanella oneidensis* MR-1 via CRISPR-mediated transposase technology, ACS Synth. Biol., 13 (2024) 1941-1951.
[107] D.R. Lovley, T. Ueki, T. Zhang, N.S. Malvankar, P.M. Shrestha, K.A. Flanagan, M. Aklujkar, J.E. Butler, L. Giloteaux, A.E. Rotaru, D.E. Holmes, A.E. Franks, R. Orellana, C. Risso, K.P. Nevin, *Geobacter*: The microbe electric's physiology, ecology, and practical applications, Adv. Microb. Physiol., 59 (2011) 1-100.
[108] T. Li, Q. Zhou, The key role of *Geobacter* in regulating emissions and biogeochemical cycling of soil-derived greenhouse gases, Environ. Pollut., 266 (2020) 115135.
[109] D.W. Waite, M. Chuvochina, C. Pelikan, D.H. Parks, P. Yilmaz, M. Wagner, A. Loy, T. Naganuma, R. Nakai, W.B. Whitman, M.W. Hahn, J. Kuever, P. Hugenholtz, Proposal to reclassify the proteobacterial classes *Deltaproteobacteria* and *Oligoflexia*, and the phylum *Thermodesulfobacteria*

into four phyla reflecting major functional capabilities, Int. J. Syst. Evol. Microbiol., 70 (2020) 5972-6016.

[110] D. Sun, X. Wan, W. Liu, X. Xia, F. Huang, A. Wang, J.A. Smith, Y. Dang, D.E. Holmes, Characterization of the genome from *Geobacter anodireducens*, a strain with enhanced current production in bioelectrochemical systems, RSC Advances, 9 (2019) 25890-25899.

[111] B.A. Methé, K.E. Nelson, J.A. Eisen, I.T. Paulsen, W. Nelson, J.F. Heidelberg, D. Wu, M. Wu, N. Ward, M.J. Beanan, R.J. Dodson, R. Madupu, L.M. Brinkac, S.C. Daugherty, R.T. DeBoy, A.S. Durkin, M. Gwinn, J.F. Kolonay, S.A. Sullivan, D.H. Haft, J. Selengut, T.M. Davidsen, N. Zafar, O. White, B. Tran, C. Romero, H.A. Forberger, J. Weidman, H. Khouri, T.V. Feldblyum, T.R. Utterback, S.E. Van Aken, D.R. Lovley, C.M. Fraser, Genome of *Geobacter sulfurreducens*: Metal reduction in subsurface environments, Science, 302 (2003) 1967-1969.

[112] M. Aklujkar, J. Krushkal, G. DiBartolo, A. Lapidus, M.L. Land, D.R. Lovley, The genome sequence of *Geobacter metallireducens*: Features of metabolism, physiology and regulation common and dissimilar to *Geobacter sulfurreducens*, BMC Microbiol., 9 (2009) 109.

[113] M. Aklujkar, N.D. Young, D. Holmes, M. Chavan, C. Risso, H.E. Kiss, C.S. Han, M.L. Land, D.R. Lovley, The genome of *Geobacter bemidjiensis*, exemplar for the subsurface clade of *Geobacter* species that predominate in Fe(III)-reducing subsurface environments, BMC Genomics, 11 (2010) 490.

[114] T. Aoyagi, H. Koike, T. Morita, Y. Sato, H. Habe, T. Hori, Draft genome sequence of *Geobacter pelophilus* strain Dfr2, a ferric iron-reducing bacterium, Genome Announc., 5 (2017) e00517-00537.

[115] J.P. Badalamenti, D.R. Bond, Complete genome of *Geobacter pickeringii* G13^{T}, a metal-reducing isolate from sedimentary kaolin deposits, Genome Announc., 3 (2015) e00015-e00038.

[116] J.E. Butler, N.D. Young, D.R. Lovley, Evolution of electron transfer out of the cell: Comparative genomics of six *Geobacter* genomes, BMC Genomics, 11 (2010) 40.

[117] K. Inoue, Y. Ogura, Y. Kawano, T. Hayashi, Complete genome sequence of *Geobacter sulfurreducens* strain YM18, isolated from river sediment in Japan, Genome Announc., 6 (2018) e00318-00352.

[118] D.R. Lovley, E.J. Phillips, Novel mode of microbial energy metabolism: Organic carbon oxidation coupled to dissimilatory reduction of iron or manganese, Appl. Environ. Microbiol., 54 (1988) 1472-1480.

[119] D.R. Lovley, E.J.P. Phillips, Organic matter mineralization with reduction of ferric iron in anaerobic sediments, Appl. Environ. Microb., 51 (1986) 683-689.

[120] D.R. Lovley, J.F. Stolz, G.L. Nord Jr, E.J.P. Phillips, Anaerobic production of magnetite by a dissimilatory iron-reducing microorganism, Nature, 330 (1987) 252-254.

[121] F. Caccavo, Jr., D.J. Lonergan, D.R. Lovley, M. Davis, J.F. Stolz, M.J. McInerney, *Geobacter sulfurreducens* sp. nov., a hydrogen- and acetate-oxidizing dissimilatory metal-reducing microorganism, Appl. Environ. Microbiol., 60 (1994) 3752-3759.

[122] M.V. Coppi, C. Leang, S.J. Sandler, D.R. Lovley, Development of a genetic system for *Geobacter sulfurreducens*, Appl. Environ. Microbiol., 67 (2001) 3180-3187.

[123] M.V. Coppi, The hydrogenases of *Geobacter sulfurreducens*: A comparative genomic perspective, Microbiology, 151 (2005) 1239-1254.
[124] P.-L. Tremblay, D.R. Lovley, Role of the NiFe hydrogenase Hya in oxidative stress defense in *Geobacter sulfurreducens*, J. Bacteriol., 194 (2012) 2248-2253.
[125] M.V. Coppi, R.A. O'Neil, D.R. Lovley, Identification of an uptake hydrogenase required for hydrogen-dependent reduction of Fe(III) and other electron acceptors by *Geobacter sulfurreducens*, J. Bacteriol., 186 (2004) 3022-3028.
[126] C.E.A. Engel, D. Vorländer, R. Biedendieck, R. Krull, K. Dohnt, Quantification of microaerobic growth of *Geobacter sulfurreducens*, PLoS One, 15 (2020) e0215341.
[127] W.C. Lin, M.V. Coppi, D.R. Lovley, *Geobacter sulfurreducens* can grow with oxygen as a terminal electron acceptor, Appl. Environ. Microbiol., 70 (2004) 2525-2528.
[128] A.M. Speers, G. Reguera, Competitive advantage of oxygen-tolerant bioanodes of *Geobacter sulfurreducens* in bioelectrochemical systems, Biofilm, 3 (2021) 100052.
[129] M. Hoffmann, J. Seidel, O. Einsle, CcpA from *Geobacter sulfurreducens* is a basic di-heme cytochrome *c* peroxidase, J. Mol. Biol., 393 (2009) 951-965.
[130] J. Seidel, M. Hoffmann, K.E. Ellis, A. Seidel, T. Spatzal, S. Gerhardt, S.J. Elliott, O. Einsle, MacA is a second cytochrome *c* peroxidase of *Geobacter sulfurreducens*, Biochemistry, 51 (2012) 2747-2756.
[131] C.A. Salgueiro, L. Morgado, M.A. Silva, M.R. Ferreira, T.M. Fernandes, P.C. Portela, From iron to bacterial electroconductive filaments: Exploring cytochrome diversity using *Geobacter* bacteria, Coordin. Chem. Rev., 452 (2022) 214284.
[132] L. Shi, T.C. Squier, J.M. Zachara, J.K. Fredrickson, Respiration of metal (hydr)oxides by *Shewanella* and *Geobacter*: A key role for multihaem *c*-type cytochromes, Mol. Microbiol., 65 (2007) 12-20.
[133] W.B.R. Pollock, G. Voordouw, Molecular biology of *c*-type cytochromes from *Desulfovibrio vulgaris* Hildenborough, Biochimie, 76 (1994) 554-560.
[134] O. Einsle, A. Messerschmidt, P. Stach, G.P. Bourenkov, H.D. Bartunik, R. Huber, P.M.H. Kroneck, Structure of cytochrome *c* nitrite reductase, Nature, 400 (1999) 476-480.
[135] A. Belbelazi, R. Neish, M. Carr, J.C. Mottram, M.L. Ginger, Divergent cytochrome *c* maturation system in kinetoplastid protists, mBio, 12 (2021) e00166.
[136] M.J. Edwards, D.J. Richardson, C.M. Paquete, T.A. Clarke, Role of multiheme cytochromes involved in extracellular anaerobic respiration in bacteria, Protein Sci., 29 (2020) 830-842.
[137] S.J. Ferguson, J.M. Stevens, J.W.A. Allen, I.B. Robertson, Cytochrome *c* assembly: A tale of ever increasing variation and mystery?, Biochim. Biophys. Acta - Bioenergetics, 1777 (2008) 980-984.
[138] J.M. Stevens, O. Daltrop, J.W.A. Allen, S.J. Ferguson, *c*-type cytochrome formation: Chemical and biological enigmas, Accounts Chem. Res., 37 (2004) 999-1007.
[139] J.W.A. Allen, M.L. Ginger, S.J. Ferguson, Maturation of the unusual single-cysteine (XXXCH) mitochondrial *c*-type cytochromes found in trypanosomatids must occur through a novel biogenesis pathway, Biochem. J., 383 (2004) 537-542.

[140] V. Fülöp, K.A. Sam, S.J. Ferguson, M.L. Ginger, J.W.A. Allen, Structure of a trypanosomatid mitochondrial cytochrome *c* with heme attached via only one thioether bond and implications for the substrate recognition requirements of heme lyase, FEBS J., 276 (2009) 2822-2832.
[141] C. Ferousi, S. Lindhoud, F. Baymann, E.R. Hester, J. Reimann, B. Kartal, Discovery of a functional, contracted heme-binding motif within a multiheme cytochrome, J. Biol. Chem., 294 (2019) 16953-16965.
[142] R. Pisa, T. Stein, R. Eichler, R. Gross, J. Simon, The *nrfI* gene is essential for the attachment of the active site haem group of *Wolinella succinogenes* cytochrome *c* nitrite reductase, Mol. Microbiol., 43 (2002) 763-770.
[143] D. Aragão, C. Frazão, L. Sieker, G.M. Sheldrick, J. LeGall, M.A. Carrondo, Structure of dimeric cytochrome c_3 from *Desulfovibrio gigas* at 1.2 Å resolution, Acta Crystallogr. D, 59 (2003) 644-653.
[144] P. Simões, P.M. Matias, J. Morais, K. Wilson, Z. Dauter, M.A. Carrondo, Refinement of the three-dimensional structures of cytochrome c_3 from *Desulfovibrio vulgaris* Hildenborough at 1.67 Å resolution and from *Desulfovibrio desulfuricans* ATCC 27774 at 1.6 Å resolution, Inorg. Chim. Acta, 273 (1998) 213-224.
[145] Y. Gu, M.J. Guberman-Pfeffer, V. Srikanth, C. Shen, F. Giska, K. Gupta, Y. Londer, F.A. Samatey, V.S. Batista, N.S. Malvankar, Structure of *Geobacter* cytochrome OmcZ identifies mechanism of nanowire assembly and conductivity, Nat. Microbiol., 8 (2023) 284-298.
[146] B. Hermann, M. Kern, L. La Pietra, J. Simon, O. Einsle, The octahaem MccA is a haem *c*-copper sulfite reductase, Nature, 520 (2015) 706-709.
[147] J.W.A. Allen, E.B. Sawyer, M.L. Ginger, P.D. Barker, S.J. Ferguson, Variant *c*-type cytochromes as probes of the substrate specificity of the *E. coli* cytochrome *c* maturation (Ccm) apparatus, Biochem. J., 419 (2009) 177-186.
[148] F. Wang, C.H. Chan, V. Suciu, K. Mustafa, M. Ammend, D. Si, A.I. Hochbaum, E.H. Egelman, D.R. Bond, Structure of *Geobacter* OmcZ filaments suggests extracellular cytochrome polymers evolved independently multiple times, eLife, 11 (2022) e81551.
[149] N. Heidary, N. Kornienko, S. Kalathil, X. Fang, K.H. Ly, H.F. Greer, E. Reisner, Disparity of cytochrome utilization in anodic and cathodic extracellular electron transfer pathways of *Geobacter sulfurreducens* biofilms, J. Am. Chem. Soc., 142 (2020) 5194-5203.
[150] J. Meng, Z. Xu, J. Guo, Y. Yue, X. Sun, Analysis of enhanced current-generating mechanism of *Geobacter sulfurreducens* strain via model-driven metabolism simulation, PLoS One, 8 (2013) e73907.
[151] A.S. Galushko, B. Schink, Oxidation of acetate through reactions of the citric acid cycle by *Geobacter sulfurreducens* in pure culture and in syntrophic coculture, Arch. Microbiol., 174 (2000) 314-321.
[152] D. Segura, R. Mahadevan, K. Juárez, D.R. Lovley, Computational and experimental analysis of redundancy in the central metabolism of *Geobacter sulfurreducens*, Plos Comput. Biol., 4 (2008) e36.
[153] M. Mollaei, P.H.A. Timmers, M. Suarez-Diez, S. Boeren, A.H. van Gelder, A.J.M. Stams, C.M. Plugge, Comparative proteomics of *Geobacter sulfurreducens* PCA^T in response to acetate, formate and/or hydrogen as electron donor, Environ. Microbiol., 23 (2021) 299-315.

[154] C.E. Levar, C.L. Hoffman, A.J. Dunshee, B.M. Toner, D.R. Bond, Redox potential as a master variable controlling pathways of metal reduction by *Geobacter sulfurreducens*, ISME J., 11 (2017) 741-752.
[155] M.A. Silva, C.A. Salgueiro, Multistep signaling in Nature: A close-up of *Geobacter* chemotaxis sensing, Int. J. Mol. Sci., 22 (2021) 9034.
[156] A.I. Pimenta, C.M. Paquete, L. Morgado, M.J. Edwards, T.A. Clarke, C.A. Salgueiro, I.A.C. Pereira, A.G. Duarte, Characterization of the inner membrane cytochrome ImcH from *Geobacter* reveals its importance for extracellular electron transfer and energy conservation, Protein Sci., 32 (2023) e4796.
[157] C.E. Levar, C.H. Chan, M.G. Mehta-Kolte, D.R. Bond, An inner membrane cytochrome required only for reduction of high redox potential extracellular electron acceptors, mBio, 5 (2014) e02034.
[158] J.M.A. Antunes, M.A. Silva, C.A. Salgueiro, L. Morgado, Electron flow from the inner membrane towards the cell exterior in *Geobacter sulfurreducens*: Biochemical characterization of cytochrome CbcL, Front. Microbiol., 13 (2022) 898015.
[159] L. Zacharoff, C.H. Chan, D.R. Bond, Reduction of low potential electron acceptors requires the CbcL inner membrane cytochrome of *Geobacter sulfurreducens*, Bioelectrochemistry, 107 (2016) 7-13.
[160] K. Joshi, C.H. Chan, D.R. Bond, *Geobacter sulfurreducens* inner membrane cytochrome CbcBA controls electron transfer and growth yield near the energetic limit of respiration, Mol. Microbiol., 116 (2021) 1124-1139.
[161] M. Aklujkar, M.V. Coppi, C. Leang, B.C. Kim, M.A. Chavan, L.A. Perpetua, L. Giloteaux, A. Liu, D.E. Holmes, Proteins involved in electron transfer to Fe(III) and Mn(IV) oxides by *Geobacter sulfurreducens* and *Geobacter uraniireducens*, Microbiology, 159 (2013) 515-535.
[162] F. Scarabotti, L. Rago, K. Bühler, F. Harnisch, The electrode potential determines the yield coefficients of early-stage *Geobacter sulfurreducens* biofilm anodes, Bioelectrochemistry, 140 (2021) 107752.
[163] C. Grobbler, B. Virdis, A. Nouwens, F. Harnisch, K. Rabaey, P.L. Bond, Effect of the anode potential on the physiology and proteome of *Shewanella oneidensis* MR-1, Bioelectrochemistry, 119 (2018) 172-179.
[164] A.A. Carmona-Martínez, F. Harnisch, U. Kuhlicke, T.R. Neu, U. Schröder, Electron transfer and biofilm formation of *Shewanella putrefaciens* as function of anode potential, Bioelectrochemistry, 93 (2013) 23-29.
[165] E. Howley, R. Krajmalnik-Brown, C.I. Torres, Cytochrome gene expression shifts in *Geobacter sulfurreducens* to maximize energy conservation in response to changes in redox conditions, Biosens. Bioelectron., 237 (2023) 115524.
[166] E. Howley, A. Mangus, D. Williams, C.I. Torres, Intracytoplasmic membranes develop in *Geobacter sulfurreducens* under thermodynamically limiting conditions, NPJ Biofilms Microbiomes, 9 (2023) 18.

[167] L. Morgado, M. Bruix, M. Pessanha, Y.Y. Londer, C.A. Salgueiro, Thermodynamic characterization of a triheme cytochrome family from *Geobacter sulfurreducens* reveals mechanistic and functional diversity, Biophys. J., 99 (2010) 293-301.

[168] L. Morgado, V.B. Paixão, M. Schiffer, P.R. Pokkuluri, M. Bruix, C.A. Salgueiro, Revealing the structural origin of the redox-Bohr effect: The first solution structure of a cytochrome from *Geobacter sulfurreducens*, Biochem. J., 441 (2012) 179-187.

[169] P.R. Pokkuluri, Y.Y. Londer, N.E. Duke, M. Pessanha, X. Yang, V. Orshonsky, L. Orshonsky, J. Erickson, Y. Zagyansky, C.A. Salgueiro, M. Schiffer, Structure of a novel dodecaheme cytochrome *c* from *Geobacter sulfurreducens* reveals an extended 12 nm protein with interacting hemes, J. Struct. Biol., 174 (2011) 223-233.

[170] D.L. Turner, H.S. Costa, I.B. Coutinho, J. Legall, A.V. Xavier, Assignment of the ligand geometry and redox potentials of the trihaem ferricytochrome c_3 from *Desulfuromonas acetoxidans*, Eur. J. Biochem, 243 (1997) 474-481.

[171] M.R. Ferreira, L. Morgado, C.A. Salgueiro, Periplasmic electron transfer network in *Geobacter sulfurreducens* revealed by biomolecular interaction studies, Protein Sci., 33 (2024) e5082.

[172] Y.H. Ding, K.K. Hixson, C.S. Giometti, A. Stanley, A. Esteve-Núñez, T. Khare, S.L. Tollaksen, W. Zhu, J.N. Adkins, M.S. Lipton, R.D. Smith, T. Mester, D.R. Lovley, The proteome of dissimilatory metal-reducing microorganism *Geobacter sulfurreducens* under various growth conditions, Biochim. Biophys. Acta, 1764 (2006) 1198-1206.

[173] E.S. Shelobolina, M.V. Coppi, A.A. Korenevsky, L.N. DiDonato, S.A. Sullivan, H. Konishi, H. Xu, C. Leang, J.E. Butler, B.C. Kim, D.R. Lovley, Importance of *c*-type cytochromes for U(VI) reduction by *Geobacter sulfurreducens*, BMC Microbiol., 7 (2007) 16.

[174] T. Ueki, L.N. DiDonato, D.R. Lovley, Toward establishing minimum requirements for extracellular electron transfer in *Geobacter sulfurreducens*, FEMS Microbiol. Lett., 364 (2017) fnx093.

[175] S. Choi, C.H. Chan, D.R. Bond, Lack of specificity in *Geobacter* periplasmic electron transfer, J. Bacteriol., 204 (2022) e00322.

[176] A. Esteve-Núñez, J. Sosnik, P. Visconti, D.R. Lovley, Fluorescent properties of *c*-type cytochromes reveal their potential role as an extracytoplasmic electron sink in *Geobacter sulfurreducens*, Environ. Microbiol., 10 (2008) 497-505.

[177] N.S. Malvankar, T. Mester, M.T. Tuominen, D.R. Lovley, Supercapacitors based on *c*-type cytochromes using conductive nanostructured networks of living bacteria, ChemPhysChem, 13 (2012) 463-468.

[178] J. Jumper, R. Evans, A. Pritzel, T. Green, M. Figurnov, O. Ronneberger, K. Tunyasuvunakool, R. Bates, A. Žídek, A. Potapenko, A. Bridgland, C. Meyer, S.A.A. Kohl, A.J. Ballard, A. Cowie, B. Romera-Paredes, S. Nikolov, R. Jain, J. Adler, T. Back, S. Petersen, D. Reiman, E. Clancy, M. Zielinski, M. Steinegger, M. Pacholska, T. Berghammer, S. Bodenstein, D. Silver, O. Vinyals, A.W. Senior, K. Kavukcuoglu, P. Kohli, D. Hassabis, Highly accurate protein structure prediction with AlphaFold, Nature, 596 (2021) 583-589.

[179] M. Varadi, S. Anyango, M. Deshpande, S. Nair, C. Natassia, G. Yordanova, D. Yuan, O. Stroe, G. Wood, A. Laydon, A. Žídek, T. Green, K. Tunyasuvunakool, S. Petersen, J. Jumper, E. Clancy, R. Green, A. Vora, M. Lutfi, M. Figurnov, A. Cowie, N. Hobbs, P. Kohli, G. Kleywegt, E. Birney, D. Hassabis, S. Velankar, AlphaFold protein structure database: Massively expanding the structural coverage of protein-sequence space with high-accuracy models, Nucleic Acids Res., 50 (2022) D439-D444.
[180] S. Sharma, G. Cavallaro, A. Rosato, A systematic investigation of multiheme *c*-type cytochromes in prokaryotes, J. Biol. Inorg. Chem., 15 (2010) 559-571.
[181] J.P. Badalamenti, Z.M. Summers, C.H. Chan, J.A. Gralnick, D.R. Bond, Isolation and genomic characterization of *Desulfuromonas soudanensis* WTL, a metal- and electrode-respiring bacterium from anoxic deep subsurface brine, Front. Microbiol., 7 (2016) 913.
[182] A.O. Leu, C. Cai, S.J. McIlroy, G. Southam, V.J. Orphan, Z. Yuan, S. Hu, G.W. Tyson, Anaerobic methane oxidation coupled to manganese reduction by members of the *Methanoperedenaceae*, ISME J., 14 (2020) 1030-1041.
[183] M. Akram, A. Dietl, U. Mersdorf, S. Prinz, W. Maalcke, J. Keltjens, C. Ferousi, N.M. de Almeida, J. Reimann, B. Kartal, M.S.M. Jetten, K. Parey, T.R.M. Barends, A 192-heme electron transfer network in the hydrazine dehydrogenase complex, Sci. Adv., 5 (2019) eaav4310.
[184] P.C. Portela, C.C. Shipps, C. Shen, V. Srikanth, C.A. Salgueiro, N.S. Malvankar, Widespread extracellular electron transfer pathways for charging microbial cytochrome OmcS nanowires via periplasmic cytochromes PpcABCDE, Nat. Commun., 15 (2024) 2434.
[185] F. Wang, Y. Gu, J.P. O'Brien, S.M. Yi, S.E. Yalcin, V. Srikanth, C. Shen, D. Vu, N.L. Ing, A.I. Hochbaum, E.H. Egelman, N.S. Malvankar, Structure of microbial nanowires reveals stacked hemes that transport electrons over micrometers, Cell, 177 (2019) 361-369.
[186] D.J. Filman, S.F. Marino, J.E. Ward, L. Yang, Z. Mester, E. Bullitt, D.R. Lovley, M. Strauss, Cryo-EM reveals the structural basis of long-range electron transport in a cytochrome-based bacterial nanowire, Commun. Biol., 2 (2019) 219.
[187] F. Wang, K. Mustafa, V. Suciu, K. Joshi, C.H. Chan, S. Choi, Z. Su, D. Si, A.I. Hochbaum, E.H. Egelman, D.R. Bond, Cryo-EM structure of an extracellular *Geobacter* OmcE cytochrome filament reveals tetrahaem packing, Nat. Microbiol., 7 (2022) 1291-1300.
[188] E. Margoliash, J. Lustgarten, Interconversion of horse heart cytochrome *c* monomer and polymers, J. Biol. Chem., 237 (1962) 3397-3405.
[189] S. Hirota, Y. Hattori, S. Nagao, M. Taketa, H. Komori, H. Kamikubo, Z. Wang, I. Takahashi, S. Negi, Y. Sugiura, M. Kataoka, Y. Higuchi, Cytochrome *c* polymerization by successive domain swapping at the C-terminal helix, Proc. Natl. Acad. Sci., 107 (2010) 12854-12859.
[190] C.K. Park, N.C. Horton, Structures, functions, and mechanisms of filament forming enzymes: A renaissance of enzyme filamentation, Biophys. Rev., 11 (2019) 927-994.
[191] H.M. Dietrich, R.D. Righetto, A. Kumar, W. Wietrzynski, R. Trischler, S.K. Schuller, J. Wagner, F.M. Schwarz, B.D. Engel, V. Müller, J.M. Schuller, Membrane-anchored HDCR nanowires drive hydrogen-powered CO_2 fixation, Nature, 607 (2022) 823-830.

[192] K. Sordakis, C. Tang, L.K. Vogt, H. Junge, P.J. Dyson, M. Beller, G. Laurenczy, Homogeneous catalysis for sustainable hydrogen storage in formic acid and alcohols, Chem. Rev., 118 (2018) 372-433.
[193] V. Müller, New horizons in acetogenic conversion of one-carbon substrates and biological hydrogen storage, Trends Biotechnol., 37 (2019) 1344-1354.
[194] L. Holm, Dali server: Structural unification of protein families, Nucleic Acids Res., 50 (2022) W210-W215.
[195] M. van Kempen, S.S. Kim, C. Tumescheit, M. Mirdita, J. Lee, C.L.M. Gilchrist, J. Söding, M. Steinegger, Fast and accurate protein structure search with Foldseek, Nat. Biotechnol., 42 (2023) 243-246.
[196] M. Varadi, J. Berrisford, M. Deshpande, S.S. Nair, A. Gutmanas, D. Armstrong, L. Pravda, B. Al-Lazikani, S. Anyango, G.J. Barton, K. Berka, T. Blundell, N. Borkakoti, J. Dana, S. Das, S. Dey, P. Di Micco, F. Fraternali, T. Gibson, M. Helmer-Citterich, D. Hoksza, L.C. Huang, R. Jain, H. Jubb, C. Kannas, N. Kannan, J. Koca, R. Krivak, M. Kumar, E.D. Levy, F. Madeira, M.S. Madhusudhan, H.J. Martell, S. MacGowan, J.E. McGreig, S. Mir, A. Mukhopadhyay, L. Parca, T. Paysan-Lafosse, L. Radusky, A. Ribeiro, L. Serrano, I. Sillitoe, G. Singh, P. Skoda, R. Svodobova, J. Tyzack, A. Valencia, E.V. Fernandez, W. Vranken, M. Wass, J.M. Thornton, M. Sternberg, C. Orengo, S. Velankar, PDBe-KB: A community-driven resource for structural and functional annotations, Nucleic Acids Res., 48 (2020) D344-D353.
[197] D.P. Baquero, V. Cvirkaite-Krupovic, S.S. Hu, J.L. Fields, X. Liu, C. Rensing, E.H. Egelman, M. Krupovic, F. Wang, Extracellular cytochrome nanowires appear to be ubiquitous in prokaryotes, Cell, 186 (2023) 2853-2864.
[198] L.-D. Shi, J. West-Roberts, M.C. Schoelmerich, P.I. Penev, L. Chen, Y. Amano, S. Lei, R. Sachdeva, J.F. Banfield, *Methanotrophic methanoperedens* archaea host diverse and interacting extrachromosomal elements, Nat. Microbiol., 9 (2024) 2422-2433.
[199] M.C. Schoelmerich, L. Ly, J. West-Roberts, L.-D. Shi, C. Shen, N.S. Malvankar, N. Taib, S. Gribaldo, B.J. Woodcroft, C.W. Schadt, B. Al-Shayeb, X. Dai, C. Mozsary, S. Hickey, C. He, J. Beaulaurier, S. Juul, R. Sachdeva, J.F. Banfield, Borg extrachromosomal elements of methane-oxidizing archaea have conserved and expressed genetic repertoires, Nat. Commun., 15 (2024) 5414.
[200] B. Al-Shayeb, M.C. Schoelmerich, J. West-Roberts, L.E. Valentin-Alvarado, R. Sachdeva, S. Mullen, A. Crits-Christoph, M.J. Wilkins, K.H. Williams, J.A. Doudna, J.F. Banfield, Borgs are giant genetic elements with potential to expand metabolic capacity, Nature, 610 (2022) 731-736.
[201] T. Mehta, M.V. Coppi, S.E. Childers, D.R. Lovley, Outer membrane *c*-type cytochromes required for Fe(III) and Mn(IV) oxide reduction in *Geobacter sulfurreducens*, Appl. Environ. Microbiol., 71 (2005) 8634-8641.
[202] Y.H. Ding, K.K. Hixson, M.A. Aklujkar, M.S. Lipton, R.D. Smith, D.R. Lovley, T. Mester, Proteome of *Geobacter sulfurreducens* grown with Fe(III) oxide or Fe(III) citrate as the electron acceptor, Biochim. Biophys. Acta - Proteins Proteom., 1784 (2008) 1935-1941.

[203] D.E. Holmes, S.K. Chaudhuri, K.P. Nevin, T. Mehta, B.A. Methé, A. Liu, J.E. Ward, T.L. Woodard, J. Webster, D.R. Lovley, Microarray and genetic analysis of electron transfer to electrodes in *Geobacter sulfurreducens*, Environ. Microbiol., 8 (2006) 1805-1815.

[204] C.A. Kellenberger, S.C. Wilson, S.F. Hickey, T.L. Gonzalez, Y. Su, Z.F. Hallberg, T.F. Brewer, A.T. Iavarone, H.K. Carlson, Y.-F. Hsieh, M.C. Hammond, GEMM-I riboswitches from *Geobacter* sense the bacterial second messenger cyclic AMP-GMP, Proc. Natl. Acad. Sci., 112 (2015) 5383-5388.

[205] B.C. Kim, B.L. Postier, R.J. DiDonato, S.K. Chaudhuri, K.P. Nevin, D.R. Lovley, Insights into genes involved in electricity generation in *Geobacter sulfurreducens* via whole genome microarray analysis of the OmcF-deficient mutant, Bioelectrochemistry, 73 (2008) 70-75.

[206] C. Leang, X. Qian, T. Mester, D.R. Lovley, Alignment of the *c*-type cytochrome OmcS along pili of *Geobacter sulfurreducens*, Appl. Environ. Microbiol., 76 (2010) 4080-4084.

[207] P.M. Shrestha, A.E. Rotaru, M. Aklujkar, F. Liu, M. Shrestha, Z.M. Summers, N. Malvankar, D.C. Flores, D.R. Lovley, Syntrophic growth with direct interspecies electron transfer as the primary mechanism for energy exchange, Environ. Microbiol. Rep., 5 (2013) 904-910.

[208] P.M. Shrestha, A.E. Rotaru, Z.M. Summers, M. Shrestha, F. Liu, D.R. Lovley, Transcriptomic and genetic analysis of direct interspecies electron transfer, Appl. Environ. Microbiol., 79 (2013) 2397-2404.

[209] K.P. Nevin, D.R. Lovley, Mechanisms for Fe(III) oxide reduction in sedimentary environments, Geomicrobiol. J., 19 (2002) 141-159.

[210] K.P. Nevin, B.C. Kim, R.H. Glaven, J.P. Johnson, T.L. Woodard, B.A. Methe, R.J. Didonato, S.F. Covalla, A.E. Franks, A. Liu, D.R. Lovley, Anode biofilm transcriptomics reveals outer surface components essential for high density current production in *Geobacter sulfurreducens* fuel cells, PLoS One, 4 (2009) e5628.

[211] A.E. Franks, R.H. Glaven, D.R. Lovley, Real-time spatial gene expression analysis within current-producing biofilms, ChemSusChem, 5 (2012) 1092-1098.

[212] K. Inoue, C. Leang, A.E. Franks, T.L. Woodard, K.P. Nevin, D.R. Lovley, Specific localization of the *c*-type cytochrome OmcZ at the anode surface in current-producing biofilms of *Geobacter sulfurreducens*, Environ. Microbiol. Rep., 3 (2011) 211-217.

[213] L. Peng, Y. Zhang, Cytochrome OmcZ is essential for the current generation by *Geobacter sulfurreducens* under low electrode potential, Electrochim. Acta, 228 (2017) 447-452.

[214] H. Richter, K.P. Nevin, H. Jia, D.A. Lowy, D.R. Lovley, L.M. Tender, Cyclic voltammetry of biofilms of wild type and mutant *Geobacter sulfurreducens* on fuel cell anodes indicates possible roles of OmcB, OmcZ, type IV pili, and protons in extracellular electron transfer, Energ. Environ. Sci., 2 (2009) 506-516.

[215] H.-Y. Tang, D.E. Holmes, T. Ueki, P.A. Palacios, D.R. Lovley, Iron corrosion via direct metal-microbe electron transfer, mBio, 10 (2019) e00303.

[216] D.R. Lovley, J. Yao, Intrinsically conductive microbial nanowires for 'green' electronics with novel functions, Trends Biotechnol., 39 (2021) 940-952.

[217] J. Neu, C.C. Shipps, M.J. Guberman-Pfeffer, C. Shen, V. Srikanth, J.A. Spies, N.D. Kirchhofer, S.E. Yalcin, G.W. Brudvig, V.S. Batista, N.S. Malvankar, Microbial biofilms as living photoconductors due to ultrafast electron transfer in cytochrome OmcS nanowires, Nat. Commun., 13 (2022) 5150.
[218] T. Lin, W. Ding, D. Zhang, Z. You, Y. Yang, F. Li, D. Xu, D.R. Lovley, H. Song, Expression of filaments of the *Geobacter* extracellular cytochrome OmcS in *Shewanella oneidensis*, Biotechnol. Bioeng., 121 (2024) 2002-2012.
[219] P.J. Dahl, S.M. Yi, Y. Gu, A. Acharya, C. Shipps, J. Neu, J.P. O'Brien, U.N. Morzan, S. Chaudhuri, M.J. Guberman-Pfeffer, D. Vu, S.E. Yalcin, V.S. Batista, N.S. Malvankar, A 300-fold conductivity increase in microbial cytochrome nanowires due to temperature-induced restructuring of hydrogen bonding networks, Sci. Adv., 8 (2022) eabm7193.
[220] S.E. Yalcin, J.P. O'Brien, Y. Gu, K. Reiss, S.M. Yi, R. Jain, V. Srikanth, P.J. Dahl, W. Huynh, D. Vu, A. Acharya, S. Chaudhuri, T. Varga, V.S. Batista, N.S. Malvankar, Electric field stimulates production of highly conductive microbial OmcZ nanowires, Nat. Chem. Biol., 16 (2020) 1136-1142.
[221] M.J. Guberman-Pfeffer, Assessing thermal response of redox conduction for anti-Arrhenius kinetics in a microbial cytochrome nanowire, J. Phys. Chem. B, 126 (2022) 10083-10097.
[222] M.J. Guberman-Pfeffer, Structural determinants of redox conduction favor robustness over tunability in microbial cytochrome nanowires, J. Phys. Chem. B, 127 (2023) 7148-7161.
[223] Y. Eshel, U. Peskin, N. Amdursky, Coherence-assisted electron diffusion across the multi-heme protein-based bacterial nanowire, Nanotechnology, 31 (2020) 314002.
[224] Y. Agam, R. Nandi, A. Kaushansky, U. Peskin, N. Amdursky, The porphyrin ring rather than the metal ion dictates long-range electron transport across proteins suggesting coherence-assisted mechanism, Proc. Natl. Acad. Sci., 117 (2020) 32260-32266.
[225] E. Papp, G. Vattay, Computation of biological conductance with Liouville quantum master equation, Sci. Rep., 14 (2024) 19571.
[226] E. Szmuc, D.J.F. Walker, D. Kireev, D. Akinwande, D.R. Lovley, B. Keitz, A. Ellington, Engineering *Geobacter* pili to produce metal:organic filaments, Biosens. Bioelectron., 222 (2023) 114993.
[227] J.D. Bernal, I. Fankuchen, M.A.X. Perutz, An X-ray study of chymotrypsin and haemoglobin, Nature, 141 (1938) 523-524.
[228] M.-C. Bellissent-Funel, A. Hassanali, M. Havenith, R. Henchman, P. Pohl, F. Sterpone, D. van der Spoel, Y. Xu, A.E. Garcia, Water determines the structure and dynamics of proteins, Chem. Rev., 116 (2016) 7673-7697.
[229] M. Tarek, D.J. Tobias, The dynamics of protein hydration water: A quantitative comparison of molecular dynamics simulations and neutron-scattering experiments, Biophys. J., 79 (2000) 3244-3257.
[230] S.H. Saunders, E.C.M. Tse, M.D. Yates, F.J. Otero, S.A. Trammell, E.D.A. Stemp, J.K. Barton, L.M. Tender, D.K. Newman, Extracellular DNA promotes efficient extracellular electron transfer by pyocyanin in *Pseudomonas aeruginosa* biofilms, Cell, 182 (2020) 919-932.
[231] G. Reguera, K.D. McCarthy, T. Mehta, J.S. Nicoll, M.T. Tuominen, D.R. Lovley, Extracellular electron transfer via microbial nanowires, Nature, 435 (2005) 1098-1101.

[232] N.S. Malvankar, M.T. Tuominen, D.R. Lovley, Lack of cytochrome involvement in long-range electron transport through conductive biofilms and nanowires of *Geobacter sulfurreducens*, Energy Environ. Sci., 5 (2012) 8651-8659.

[233] N.S. Malvankar, D.R. Lovley, Microbial nanowires for bioenergy applications, Curr. Opin. Biotechnol., 27 (2014) 88-95.

[234] N.S. Malvankar, D.R. Lovley, Microbial nanowires: A new paradigm for biological electron transfer and bioelectronics, ChemSusChem, 5 (2012) 1039-1046.

[235] Derek R. Lovley, Long-range electron transport to Fe(III) oxide via pili with metallic-like conductivity, Biochem. Soc. Trans., 40 (2012) 1186-1190.

[236] L. Craig, K.T. Forest, B. Maier, Type IV pili: Dynamics, biophysics and functional consequences, Nat. Rev. Microbiol., 17 (2019) 429-440.

[237] N.S. Malvankar, M. Vargas, K. Nevin, P.-L. Tremblay, K. Evans-Lutterodt, D. Nykypanchuk, E. Martz, M.T. Tuominen, D.R. Lovley, Structural basis for metallic-like conductivity in microbial nanowires, mBio, 6 (2015) e00084.

[238] L.A. Zotti, J.C. Cuevas, Electron transport through homopeptides: Are they really good conductors?, ACS Omega, 3 (2018) 3778-3785.

[239] R.Y. Adhikari, N.S. Malvankar, M.T. Tuominen, D.R. Lovley, Conductivity of individual *Geobacter* pili, RSC Advances, 6 (2016) 8354-8357.

[240] D.R. Lovley, Electrically conductive pili: Biological function and potential applications in electronics, Curr. Opin. Electrochem., 4 (2017) 190-198.

[241] P.N. Reardon, K.T. Mueller, Structure of the type IVa major pilin from the electrically conductive bacterial nanowires of *Geobacter sulfurreducens*, J. Biol. Chem., 288 (2013) 29260-29266.

[242] N. Lebedev, S. Mahmud, I. Griva, A. Blom, L.M. Tender, On the electron transfer through *Geobacter sulfurreducens* PilA protein, J. Polym. Sci. Pol. Phys., 53 (2015) 1706-1717.

[243] Y. Gu, V. Srikanth, A.I. Salazar-Morales, R. Jain, J.P. O'Brien, S.M. Yi, R.K. Soni, F.A. Samatey, S.E. Yalcin, N.S. Malvankar, Structure of *Geobacter* pili reveals secretory rather than nanowire behaviour, Nature, 597 (2021) 430-434.

[244] F. Wang, M. Coureuil, T. Osinski, A. Orlova, T. Altindal, G. Gesbert, X. Nassif, E.H. Egelman, L. Craig, Cryoelectron microscopy reconstructions of the *Pseudomonas aeruginosa* and *Neisseria gonorrhoeae* type IV pili at sub-nanometer resolution, Structure, 25 (2017) 1423-1435.

[245] S. Kolappan, M. Coureuil, X. Yu, X. Nassif, E.H. Egelman, L. Craig, Structure of the *Neisseria meningitidis* type IV pilus, Nat. Commun., 7 (2016) 13015.

[246] A. López-Castilla, J.-L. Thomassin, B. Bardiaux, W. Zheng, M. Nivaskumar, X. Yu, M. Nilges, E.H. Egelman, N. Izadi-Pruneyre, O. Francetic, Structure of the calcium-dependent type 2 secretion pseudopilus, Nat. Microbiol., 2 (2017) 1686-1695.

[247] B. Bardiaux, G.C. de Amorim, A. Luna Rico, W. Zheng, I. Guilvout, C. Jollivet, M. Nilges, E.H. Egelman, N. Izadi-Pruneyre, O. Francetic, Structure and assembly of the enterohemorrhagic *Escherichia coli* type 4 pilus, Structure, 27 (2019) 1082-1093.

[248] A. Neuhaus, M. Selvaraj, R. Salzer, J.D. Langer, K. Kruse, L. Kirchner, K. Sanders, B. Daum, B. Averhoff, V.A.M. Gold, Cryo-electron microscopy reveals two distinct type IV pili assembled by the same bacterium, Nat. Commun., 11 (2020) 2231.
[249] K.V. Korotkov, M. Sandkvist, W.G.J. Hol, The type II secretion system: Biogenesis, molecular architecture and mechanism, Nat. Rev. Microbiol., 10 (2012) 336-351.
[250] Y. Ye, X. Liu, K.H. Nealson, C. Rensing, S. Qin, S. Zhou, Dissecting the structural and conductive functions of nanowires in *Geobacter sulfurreducens* electroactive biofilms, mBio, 13 (2022) e03822.
[251] X. Liu, S. Zhuo, C. Rensing, S. Zhou, Syntrophic growth with direct interspecies electron transfer between pili-free *Geobacter* species, ISME J., 12 (2018) 2142-2151.
[252] G. Reguera, Harnessing the power of microbial nanowires, Microb. Biotechnol., 11 (2018) 979-994.
[253] T. Boesen, L.P. Nielsen, A. Schramm, Pili for nanowires, Nat. Microbiol., 6 (2021) 1347-1348.
[254] F. Wang, L. Craig, X. Liu, C. Rensing, E.H. Egelman, Microbial nanowires: Type IV pili or cytochrome filaments?, Trends Microbiol., 31 (2023) 384-392.
[255] D.R. Lovley, Response to Wang et al.: Evidence contradicting the cytochrome-only model, Trends Microbiol., 31 (2023) 548-549.
[256] F. Wang, L. Craig, X. Liu, C. Rensing, E.H. Egelman, Models are useful until high-resolution structures are available, Trends Microbiol., 31 (2023) 550-551.
[257] D.R. Lovley, D.J.F. Walker, *Geobacter* protein nanowires, Front. Microbiol., 10 (2019).
[258] D.R. Lovley, On the existence of pilin-based microbial nanowires, Front. Microbiol., 13 (2022) 872610.
[259] D.R. Lovley, Microbial nanowires, Curr. Biol., 32 (2022) R110-R112.
[260] D.R. Lovley, Untangling *Geobacter sulfurreducens* nanowires, mBio, 13 (2022) e00850.
[261] X. Liu, K.H. Nealson, S. Zhou, C. Rensing, Reply to Lovley, "Untangling *Geobacter sulfurreducens* nanowires", mBio, 13 (2022) e01041.
[262] M.J. Guberman-Pfeffer, To be or not to be a cytochrome: Electrical characterizations are inconsistent with *Geobacter* cytochrome 'nanowires', Front. Microbiol., 15 (2024) 1397124.
[263] I.A. Schwarz, B. Alsaqri, Y. Lekbach, K. Henry, S. Gorman, T. Woodard, L. Dion, L. Real, D.E. Holmes, J.A. Smith, D.R. Lovley, Lack of physiological evidence for cytochrome filaments functioning as conduits for extracellular electron transfer, mBio, 15 (2024) e00690.
[264] A. Hernández-Eligio, G.A. Huerta-Miranda, S. Martínez-Bahena, D. Castrejón-López, M. Miranda-Hernández, K. Juárez, GSU1771 regulates extracellular electron transfer and electroactive biofilm formation in *Geobacter sulfurreducens*: Genetic and electrochemical characterization, Bioelectrochemistry, 145 (2022) 108101.
[265] P.L. Tremblay, Z.M. Summers, R.H. Glaven, K.P. Nevin, K. Zengler, C.L. Barrett, Y. Qiu, B.O. Palsson, D.R. Lovley, A *c*-type cytochrome and a transcriptional regulator responsible for enhanced extracellular electron transfer in *Geobacter sulfurreducens* revealed by adaptive evolution, Environ. Microbiol., 13 (2011) 13-23.

[266] K.B. Gregory, D.R. Bond, D.R. Lovley, Graphite electrodes as electron donors for anaerobic respiration, Environ. Microbiol., 6 (2004) 596-604.
[267] C.S. Butler, D.R. Lovley, How to sustainably feed a microbe: Strategies for biological production of carbon-based commodities with renewable electricity, Front. Microbiol., 7 (2016) 1879.
[268] O. Choi, B.I. Sang, Extracellular electron transfer from cathode to microbes: Application for biofuel production, Biotechnol. Biofuels, 9 (2016) 11.
[269] E. Marsili, J. Sun, D.R. Bond, Voltammetry and growth physiology of *Geobacter sulfurreducens* biofilms as a function of growth stage and imposed electrode potential, Electroanalysis, 22 (2010) 865-874.
[270] M. Rimboud, E. Desmond-Le Quemener, B. Erable, T. Bouchez, A. Bergel, Multi-system Nernst-Michaelis-Menten model applied to bioanodes formed from sewage sludge, Bioresour. Technol., 195 (2015) 162-169.
[271] S.M. Strycharz, R.H. Glaven, M.V. Coppi, S.M. Gannon, L.A. Perpetua, A. Liu, K.P. Nevin, D.R. Lovley, Gene expression and deletion analysis of mechanisms for electron transfer from electrodes to *Geobacter sulfurreducens*, Bioelectrochemistry, 80 (2011) 142-150.
[272] R.A. Yoho, S.C. Popat, C.I. Torres, Dynamic potential-dependent electron transport pathway shifts in anode biofilms of *Geobacter sulfurreducens*, ChemSusChem, 7 (2014) 3413-3419.
[273] J.M. Dantas, D.M. Tomaz, L. Morgado, C.A. Salgueiro, Functional characterization of PccH, a key cytochrome for electron transfer from electrodes to the bacterium *Geobacter sulfurreducens*, FEBS Lett., 587 (2013) 2662-2668.
[274] J.M. Dantas, L. Campelo, N. Duke, C.A. Salgueiro, P.R. Pokkuluri, The structure of PccH from *Geobacter sulfurreducens*: A novel low reduction potential monoheme cytochrome essential for accepting electrons from an electrode, FEBS J., 282 (2015) 2215-2231.
[275] T.C. Santos, A.R. de Oliveira, J.M. Dantas, C.A. Salgueiro, C.M. Cordas, Thermodynamic and kinetic characterization of PccH, a key protein in microbial electrosynthesis processes in *Geobacter sulfurreducens*, Biochim. Biophys. Acta - Bioenergetics, 1847 (2015) 1113-1118.
[276] L.R. Teixeira, P.C. Portela, L. Morgado, D. Pantoja-Uceda, M. Bruix, C.A. Salgueiro, Backbone assignment of cytochrome PccH, a crucial protein for microbial electrosynthesis in *Geobacter sulfurreducens*, Biomol. NMR Assign., 13 (2019) 321-326.
[277] L.R. Teixeira, T.M. Fernandes, M.A. Silva, L. Morgado, C.A. Salgueiro, Characterization of a novel cytochrome involved in *Geobacter sulfurreducens*' electron harvesting pathways, Chem. Eur. J., 28 (2022) e202202333.
[278] P.R. Pokkuluri, Y.Y. Londer, S.J. Wood, N.E. Duke, L. Morgado, C.A. Salgueiro, M. Schiffer, Outer membrane cytochrome *c*, OmcF, from *Geobacter sulfurreducens*: High structural similarity to an algal cytochrome c_6, Proteins, 74 (2009) 266-270.
[279] S.C. Andrews, Iron storage in bacteria, Adv. Microb. Physiol., 40 (1998) 281-351.
[280] K.M. Sendra, A. Barwinska-Sendra, E.S. Mackenzie, A. Baslé, T.E. Kehl-Fie, K.J. Waldron, An ancient metalloenzyme evolves through metal preference modulation, Nat. Ecol. Evol., 7 (2023) 732-744.

[281] P.M. Williams, M.G. Edmunds, The abundances of the elements, Sci. Prog. Oxf. , 61 (1974) 323-347.

[282] H. Beinert, R.H. Holm, E. Münck, Iron-sulfur clusters: Nature's modular, multipurpose structures, Science, 277 (1997) 653.

[283] Y. Christen, Oxidative stress and Alzheimer disease, Am. J. Clin. Nutr., 71 (2000) 621S-629S.

[284] R.R. Crichton, R.J. Ward, Iron species in iron homeostasis and toxicity, Analyst, 120 (1995) 693-697.

[285] M. Valko, C.J. Rhodes, J. Moncol, M. Izakovic, M. Mazur, Free radicals, metals and antioxidants in oxidative stress-induced cancer, Chem. Biol. Interact., 160 (2006) 1-40.

[286] S.C. Andrews, A.K. Robinson, F. Rodríguez-Quiñones, Bacterial iron homeostasis, FEMS Microbiol. Rev., 27 (2003) 215-237.

[287] J.A. Imlay, Cellular defenses against superoxide and hydrogen peroxide, Annu. Rev. Biochem., 77 (2008) 755-776.

[288] M. Sánchez, L. Sabio, N. Gálvez, M. Capdevila, J.M. Dominguez-Vera, Iron chemistry at the service of life, IUBMB Life, 69 (2017) 382-388.

[289] F.W. Outten, E.C. Theil, Iron-based redox switches in biology, Antioxid. Redox Sign., 11 (2008) 1029-1046.

[290] D. Alvarez-Paggi, L. Hannibal, M.A. Castro, S. Oviedo-Rouco, V. Demicheli, V. Tórtora, F. Tomasina, R. Radi, D.H. Murgida, Multifunctional cytochrome *c*: Learning new tricks from an old dog, Chem. Rev., 117 (2017) 13382-13460.

[291] Z. Zheng, M.R. Gunner, Analysis of the electrochemistry of hemes with E_ms spanning 800 mV, Proteins, 75 (2009) 719-734.

[292] A.S. Ivanov, A.I. Boldyrev, Deciphering aromaticity in porphyrinoids via adaptive natural density partitioning, Org. Biomol. Chem., 12 (2014) 6145-6150.

[293] T.D. Lash, Origin of aromatic character in porphyrinoid systems, J. Porphyr. Phthalocya., 15 (2011) 1093-1115.

[294] G.P. Moss, Nomenclature of tetrapyrroles: Recommendations of the 1986 IUPAC-IUB joint commission on biochemical nomenclature (JCBN), Eur. J. Biochem., 178 (1988) 277-328.

[295] T.J. Fiedler, C.A. Davey, R.E. Fenna, X-ray crystal structure and characterization of halide-binding sites of human myeloperoxidase at 1.8 Å resolution, J. Biol. Chem., 275 (2000) 11964-11971.

[296] D. Keilin, On cytochrome, a respiratory pigment, common to animals, yeast, and higher plants, Proc. R. Soc. B: Biol. Sci., 98 (1925) 312-339.

[297] G.N. Murshudov, A.I. Grebenko, V. Barynin, Z. Dauter, K.S. Wilson, B.K. Vainshtein, W. Melik-Adamyan, J. Bravo, J.M. Ferrán, J.C. Ferrer, J. Switala, P.C. Loewen, I. Fita, Structure of the heme *d* of *Penicillium vitale* and *Escherichia coli* catalases, J. Biol. Chem., 271 (1996) 8863-8868.

[298] T.D. Rae, H.M. Goff, The heme prosthetic group of lactoperoxidase: Structural characteristics of heme *l* and heme *l*-peptides, J. Biol. Chem., 273 (1998) 27968-27977.

[299] M.R. Cheesman, V.S. Oganesyan, N.J. Watmough, C.S. Butler, A.J. Thomson, The nature of the exchange coupling between high-spin Fe(III) heme o_3 and Cu_B(II) in *Escherichia coli* quinol oxidase, cytochrome bo_3: MCD and EPR studies, J. Am. Chem. Soc., 126 (2004) 4157-4166.
[300] J.H. van Vleck, Theory of the variations in paramagnetic anisotropy among different salts of the iron group, Phys. Rev., 41 (1932) 208-215.
[301] J. Shack, W.M. Clark, Metalloporphyrins: Cycles of changes in systems containing heme, J. Biol. Chem., 171 (1947) 143-187.
[302] G. Battistuzzi, M. Borsari, J.A. Cowan, A. Ranieri, M. Sola, Control of cytochrome *c* redox potential: Axial ligation and protein environment effects, J. Am. Chem. Soc., 124 (2002) 5315-5324.
[303] G. Liu, W. Shao, S. Zhu, W. Tang, Effects of axial ligand replacement on the redox potential of cytochrome *c*, J. Inorg. Biochem., 60 (1995) 123-131.
[304] P. Rydberg, E. Sigfridsson, U. Ryde, On the role of the axial ligand in heme proteins: A theoretical study, J. Biol. Inorg. Chem., 9 (2004) 203-223.
[305] G.R. Moore, G.W. Pettigrew, Cytochromes *c*: Evolutionary, structural and physicochemical aspects, Springer-Verlag Heidelberg, Berlin, 1990.
[306] L.J. Smith, A. Kahraman, J.M. Thornton, Heme proteins-diversity in structural characteristics, function, and folding, Proteins, 78 (2010) 2349-2368.
[307] I. Bertini, G. Cavallaro, A. Rosato, Cytochrome *c*: Occurrence and functions, Chem. Rev., 106 (2006) 90-115.
[308] P. Hosseinzadeh, Y. Lu, Design and fine-tuning redox potentials of metalloproteins involved in electron transfer in bioenergetics, Biochim. Biophys. Acta, 1857 (2016) 557-581.
[309] A.L. Raphael, H.B. Gray, Axial ligand replacement in horse heart cytochrome *c* by semisynthesis, Proteins, 6 (1989) 338-340.
[310] Y. Takayama, N.D. Werbeck, H. Komori, K. Morita, K. Ozawa, Y. Higuchi, H. Akutsu, Strategic roles of axial histidines in structure formation and redox regulation of tetraheme cytochrome c_3, Biochemistry, 47 (2008) 9405-9415.
[311] C. Olea, Jr., J. Kuriyan, M.A. Marletta, Modulating heme redox potential through protein-induced porphyrin distortion, J. Am. Chem. Soc., 132 (2010) 12794-12795.
[312] E. Stellwagen, Haem exposure as the determinate of oxidation-reduction potential of haem proteins, Nature, 275 (1978) 73-74.
[313] R.J. Kassner, Theoretical model for the effects of local nonpolar heme environments on the redox potentials in cytochromes, J. Am. Chem. Soc., 95 (1973) 2674-2677.
[314] J. Mao, K. Hauser, M.R. Gunner, How cytochromes with different folds control heme redox potentials, Biochemistry, 42 (2003) 9829-9840.
[315] P. Voigt, E.-W. Knapp, Tuning heme redox potentials in the cytochrome *c* subunit of photosynthetic reaction centers, J. Biol. Chem., 278 (2003) 51993-52001.
[316] F. Scott Mathews, The structure, function and evolution of cytochromes, Prog. Biophys. Mol. Bio., 45 (1985) 1-56.

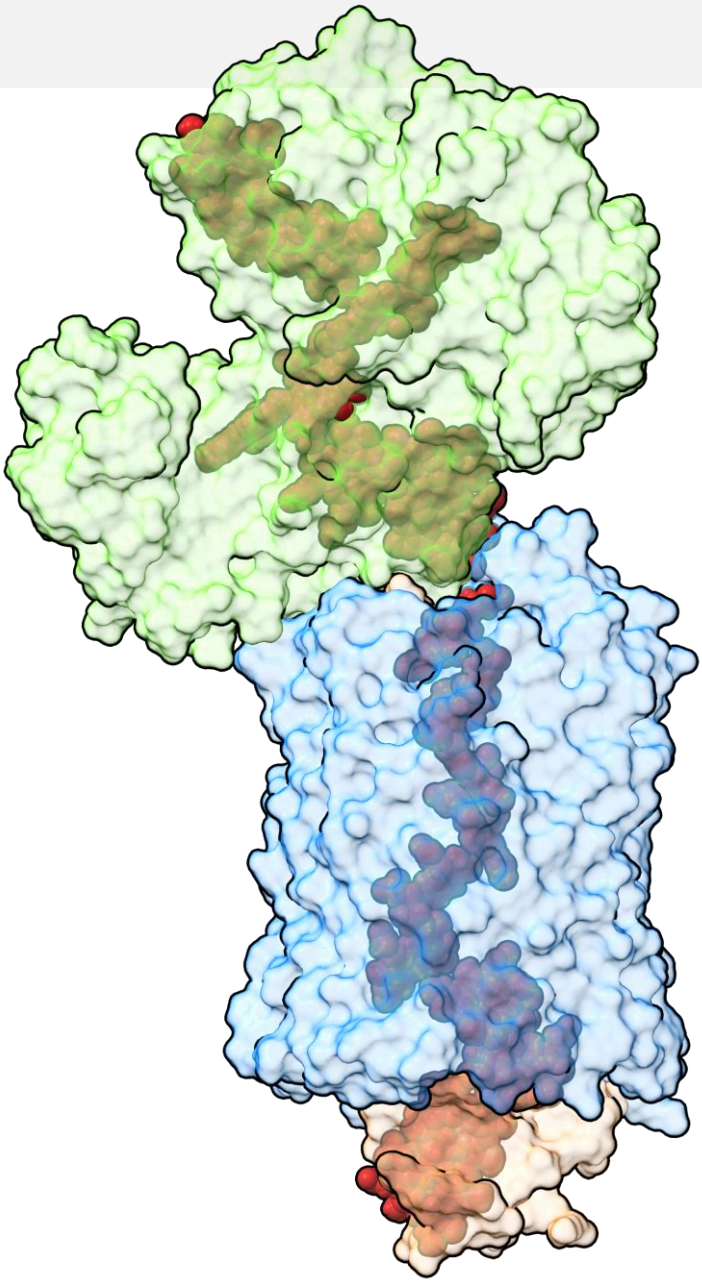
[317] F. Malatesta, G. Antonini, P. Sarti, M. Brunori, Structure and function of a molecular machine: Cytochrome *c* oxidase, Biophys. Chem., 54 (1995) 1-33.
[318] M.J. Coon, Cytochrome P450: Nature's most versatile biological catalyst, Annu. Rev. Pharmacol. Toxicol., 45 (2004) 1-25.
[319] S. Ma, R. Ludwig, Direct electron transfer of enzymes facilitated by cytochromes, ChemElectroChem, 6 (2019) 958-975.
[320] P. Taylor, S.L. Pealing, G.A. Reid, S.K. Chapman, M.D. Walkinshaw, Structural and mechanistic mapping of a unique fumarate reductase, Nat. Struct. Biol., 6 (1999) 1108-1112.
[321] S. Shin, V.L. Davidson, MauG, a diheme enzyme that catalyzes tryptophan tryptophylquinone biosynthesis by remote catalysis, Arch. Biochem. Biophys., 544 (2014) 112-118.
[322] B.M. Fonseca, C.M. Paquete, C.A. Salgueiro, R.O. Louro, The role of intramolecular interactions in the functional control of multiheme cytochromes *c*, FEBS Lett., 586 (2012) 504-509.
[323] V. Fulop, J.W. Moir, S.J. Ferguson, J. Hajdu, The anatomy of a bifunctional enzyme: Structural basis for reduction of oxygen to water and synthesis of nitric oxide by cytochrome cd_1, Cell, 81 (1995) 369-377.
[324] M. Assfalg, L. Banci, I. Bertini, M. Bruschi, P. Turano, 800 MHz ^{1}H NMR solution structure refinement of oxidized cytochrome c_7 from *Desulfuromonas acetoxidans*, Eur. J. Biochem., 256 (1998) 261-270.
[325] A.P. Fernandes, I. Couto, L. Morgado, Y.Y. Londer, C.A. Salgueiro, Isotopic labeling of *c*-type multiheme cytochromes overexpressed in *E. coli*, Prot. Expr. Purif., 59 (2008) 182-188.
[326] A.C. Messias, D.H. Kastrau, H.S. Costa, J. LeGall, D.L. Turner, H. Santos, A.V. Xavier, Solution structure of *Desulfovibrio vulgaris* (Hildenborough) ferrocytochrome c_3: Structural basis for functional cooperativity, J. Mol. Biol., 281 (1998) 719-739.
[327] L. Morgado, A.P. Fernandes, Y.Y. Londer, M. Bruix, C.A. Salgueiro, One simple step in the identification of the cofactors signals, one giant leap for the solution structure determination of multiheme proteins, Biochem. Biophys. Res. Commun., 393 (2010) 466-470.
[328] H. Santos, D.L. Turner, A.V. Xavier, J. Le Gall, Two-dimensional NMR studies of electron transfer in cytochrome c_3, J. Magn. Reson., 59 (1984) 177-180.
[329] D.L. Turner, C.A. Salgueiro, P. Schenkels, J. Le Gall, A.V. Xavier, Carbon-13 NMR studies of the influence of axial ligand orientation on haem electronic structure, Biochim. Biophys. Acta, 1246 (1995) 24-28.
[330] L. Morgado, C.A. Salgueiro, Elucidation of complex respiratory chains: A straightforward strategy to monitor electron transfer between cytochromes, Metallomics, 14 (2022) mfac012.
[331] J.M. Dantas, L. Morgado, M. Aklujkar, M. Bruix, Y.Y. Londer, M. Schiffer, P.R. Pokkuluri, C.A. Salgueiro, Rational engineering of *Geobacter sulfurreducens* electron transfer components: A foundation for building improved *Geobacter*-based bioelectrochemical technologies, Front. Microbiol., 6 (2015) 752.

[332] J.M. Dantas, P.C. Portela, A.P. Fernandes, Y.Y. Londer, X. Yang, N.E.C. Duke, M. Schiffer, P.R. Pokkuluri, C.A. Salgueiro, Structural and functional relevance of the conserved residue V13 in the triheme cytochrome PpcA from *Geobacter sulfurreducens*, J. Phys. Chem. B, 123 (2019) 3050-3060.
[333] J.M. Dantas, L. Morgado, Y.Y. Londer, A.P. Fernandes, R.O. Louro, P.R. Pokkuluri, M. Schiffer, C.A. Salgueiro, Pivotal role of the strictly conserved aromatic residue F15 in the cytochrome c_7 family, J. Biol. Inorg. Chem., 17 (2012) 11-24.
[334] J.M. Dantas, T. Simões, L. Morgado, C. Caciones, A.P. Fernandes, M.A. Silva, M. Bruix, P.R. Pokkuluri, C.A. Salgueiro, Unveiling the structural basis that regulates the energy transduction properties within a family of triheme cytochromes from *Geobacter sulfurreducens*, J. Phys. Chem. B, 120 (2016) 10221-10233.
[335] L. Morgado, J.M. Dantas, T. Simões, Y.Y. Londer, P.R. Pokkuluri, C.A. Salgueiro, Role of Met^{58} in the regulation of electron/proton transfer in trihaem cytochrome PpcA from *Geobacter sulfurreducens*, Bioscience Rep., 33 (2013).
[336] L. Morgado, S. Lourenço, Y.Y. Londer, M. Schiffer, P.R. Pokkuluri, C.A. Salgueiro, Dissecting the functional role of key residues in triheme cytochrome PpcA: A path to rational design of *G. sulfurreducens* strains with enhanced electron transfer capabilities, PLoS One, 9 (2014) e105566.
[337] L. Morgado, M. Bruix, P.R. Pokkuluri, C.A. Salgueiro, D.L. Turner, Redox- and pH-linked conformational changes in triheme cytochrome PpcA from *Geobacter sulfurreducens*, Biochem. J., 474 (2017) 231-246.
[338] P.C. Portela, M.A. Silva, L.R. Teixeira, C.A. Salgueiro, A unique aromatic residue modulates the redox range of a periplasmic multiheme cytochrome from *Geobacter metallireducens*, J. Biol. Chem., 296 (2021) 100711.
[339] J.R. Lloyd, C. Leang, A.L. Hodges Myerson, M.V. Coppi, S. Cuifo, B. Methe, S.J. Sandler, D.R. Lovley, Biochemical and genetic characterization of PpcA, a periplasmic *c*-type cytochrome in *Geobacter sulfurreducens*, Biochem. J., 369 (2003) 153-161.
[340] T. Ueki, D.J.F. Walker, T.L. Woodard, K.P. Nevin, S.S. Nonnenmann, D.R. Lovley, An *Escherichia coli* chassis for production of electrically conductive protein nanowires, ACS Synth. Biol., 9 (2020) 647-654.
[341] M. Karamash, M. Stumpe, J. Dengjel, C.A. Salgueiro, B. Giese, K.M. Fromm, Reduction kinetic of water soluble metal salts by *Geobacter sulfurreducens*: Fe^{2+}/hemes stabilize and regulate electron flux rates, Front. Microbiol., 13 (2022) 909109.
[342] X. Jiang, J. Hu, E.R. Petersen, L.A. Fitzgerald, C.S. Jackan, A.M. Lieber, B.R. Ringeisen, C.M. Lieber, J.C. Biffinger, Probing single- to multi-cell level charge transport in *Geobacter sulfurreducens* DL-1, Nat. Commun., 4 (2013) 2751.
[343] Y. Deng, D.R. Beahm, S. Ionov, R. Sarpeshkar, Measuring and modeling energy and power consumption in living microbial cells with a synthetic ATP reporter, BMC Biol., 19 (2021) 101.
[344] X. Jiang, J.H. van Wonderen, J.N. Butt, M.J. Edwards, T.A. Clarke, J. Blumberger, Which multi-heme protein complex transfers electrons more efficiently? Comparing MtrCAB from *Shewanella* with OmcS from *Geobacter*, J. Phys. Chem. Lett., 11 (2020) 9421-9425.

[345] D.E. Ross, S.L. Brantley, M. Tien, Kinetic characterization of OmcA and MtrC, terminal reductases involved in respiratory electron transfer for dissimilatory iron reduction in *Shewanella oneidensis* MR-1, Appl. Environ. Microb., 75 (2009) 5218-5226.
[346] K. Ishiki, H. Shiigi, Kinetics of intracellular electron generation in *Shewanella oneidensis* MR-1, Anal. Chem., 91 (2019) 14401-14406.
[347] T.A. Clarke, Plugging into bacterial nanowires: A comparison of model electrogenic organisms, Curr. Opin. Microbiol., 66 (2022) 56-62.
[348] B.J. Gross, M.Y. El-Naggar, A combined electrochemical and optical trapping platform for measuring single cell respiration rates at electrode interfaces, Rev. Sci. Instrum., 86 (2015) 064301.

2

Unfolding the outer gates of extracellular electron transfer in *Geobacter* with AlphaFold



The findings of this chapter will form the basis for a manuscript to be submitted for publication:

T. M. Fernandes, J. M. A. Antunes, C. A. Salgueiro and L. Morgado, Unfolding the extracellular electron transfer pathways in *Geobacter* with AlphaFold, In preparation, 2025

List of contents

Unfolding the outer gates of extracellular electron transfer in *Geobacter* with AlphaFold 91

2.1. Summary 93

2.2. Introduction 95

2.2.1. The protein folding problem 95

2.2.2. The impact and applications of AlphaFold 96

2.2.3. Structural insights on porin-cytochrome complexes 97

2.3. Results and discussion 99

2.3.1. OmabcB and OmabcC 99

2.3.2. ExtABCD 105

2.3.3. ExtEFG 110

2.3.4. ExtHIJKL 114

2.4. Using structural models to optimize protein production 122

2.4.1. "Blind" protein-cloning 122

2.4.2. AlphaFold-guided protein cloning 124

2.4.3. Preliminary expression tests of target proteins 126

2.5. Conclusions 128

2.6. Materials and methods 131

2.6.1. Protein targets and putative porin-cytochrome complexes model prediction 131

2.6.2. RF cloning of the target proteins 131

2.6.3. Insertion of affinity-tags in cytochrome constructs 135

2.6.4. Production of the truncated ExtA and ExtK constructs 136

2.6.5. Protein expression tests 137

2.7. References 138

2.1. Summary

AlphaFold has delivered a revolutionary advance for protein structure prediction by using an advanced machine learning approach that incorporates both genetic and structural database searches. AlphaFold predicts protein structures that often rival those with experimental accuracy and the incorporation of this tool within structural biology is contributing to the progress of the field.

In this chapter, AlphaFold is used to obtain crucial information about specific outer membrane multiheme cytochromes and porin-cytochrome complexes. Previous structure prediction methods typically failed to deliver reliable protein models for heme-containing proteins, especially in the definition of the heme pocket(s), which in most cases can be confidently predicted by AlphaFold.

The predicted structures of the protein components of the outer gates of EET in *G. sulfurreducens* reveal information about function, overall fold arrangement, heme axial coordination, heme spatial disposition, as well as insights on the assembly of multiprotein complexes. Based on the obtained models, an overview of the EET networks in the outer membrane of *G. sulfurreducens* is presented. In addition, the data obtained can be further explored to define cloning, expression and purification strategies to assist the production of these proteins for in vitro studies. Consequently, a description of the rational used to clone different target proteins of this work is presented.

Overall, we show how AlphaFold might be used to shed light on the general understanding of EET mechanisms in bacteria.

2.2. Introduction

2.2.1. The protein folding problem

Proteins perform an incredible multitude of functions within Life, from transporting molecules [1] to the catalysis of metabolic reactions [2]. These versatile biomolecules were first described by Gerardus J. Mulder in 1838 [3, 4] – who found that proteins were composed of smaller molecules containing both amino and carboxylic functions, named amino acids – and have been one of the main protagonists of chemical research in the past one hundred years. The early understanding that protein's function is determined not only by its amino acid composition, but also by the molecular interactions between these residues – known as Anfinsen's dogma or Thermodynamic Hypothesis [5, 6] – sparked the need of finding paths to understand these interactions and to determine the tridimensional structure of these biomolecules. Ever since, several techniques have been developed and applied to solve the structure of more than 225,000 proteins [7, 8], namely X-ray crystallography [9, 10], NMR [11, 12] and cryo-EM [13, 14].

From early on in the field of structural biology, scientists posed the "protein-folding problem" which consists on predicting the tridimensional structure of a protein based solely on its "one-dimensional" sequence of amino acids [15, 16]. According to Levinthal's paradox [17, 18], empirically determining the native folded state of a molecule such as a protein, which contains a very large number of degrees of freedom, would require an almost infinite amount of time using "brute force" calculations, precisely due to the enormous number of possible protein conformations, which can go up to 10^{300} for a 150 amino acid protein [19]. Since 1994, in pursue of solving this complex problem, scientists have been organizing the Critical Assessment of Protein Structure Prediction (CASP) contest [20] with the intent of establishing the state of the art in protein structure determination. Astonishingly, in CASP14, held in 2020, a group of scientists was able to develop a method capable of predicting protein structures to near experimental accuracy in the majority of cases, with an all-atom accuracy of 1.6 Å [21].

AlphaFold 2, the artificial intelligence (AI) software developed by Google's offshoot DeepMind, and inspired on the original AlphaFold algorithm [22], released in 2018 during the CASP13 contest [23], incorporates neural network architectures and training methods based on evolutionary, physical and geometric constraints of protein structures (Figure 2.1) [12]. Briefly, the input file (a protein sequence) is parsed and basic metadata is extracted for genetic and template search. Multiple genetic databases (MGnify [24], UniRef90 [25], Uniclust30 [26] and BFD [21]) are searched using JackHMMER [27] and HHBlits [28], while structural templates are retrieved from PDB [7] using HHSearch [29], and aligned with Kalign [30]. The resulting multiple sequence alignments (MSA) and structural templates are refined by the Evoformer stack (consisting of 48 blocks of neural networks that exchange information between amino acid sequences and pairwise structural characterization of the protein's geometry [21]) and assembled as simplified abstract representations of protein structures.

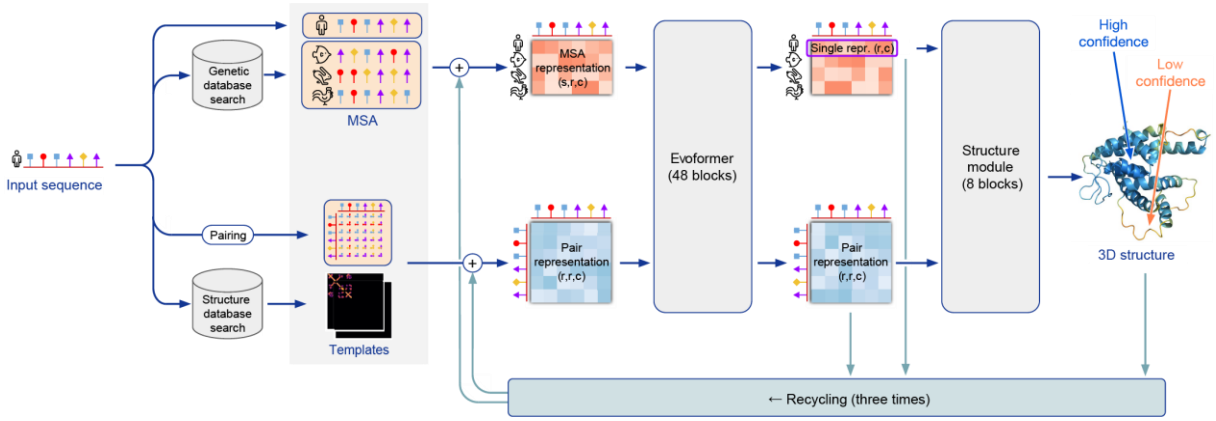


Figure 2.1. Model architecture of AlphaFold 2. The arrows show the information flow. The array shapes are shown in parentheses with letters "s" (number of sequences), "r" (number of residues) and "c" (number of channels). This figure was reproduced from [21].

These representations are then mapped to concrete 3D atom coordinates by a Structure module that operates solely with heavy atoms (carbon, oxygen, nitrogen and sulfur). This module initially predicts backbone frames and torsion angles, which are then applied to the corresponding amino acid structures to construct atom coordinates with idealized bond angles and lengths. To resolve remaining structural violations and clashes, the protein model predictions are relaxed using an iterative restrained energy minimization procedure with the AMBER99SB force field, with increasing harmonic restraints [31]. The final full energy minimization and hydrogen placement is performed using the OpenMM simulation package [32]. In the end of this process, AlphaFold 2 executes the described network multiple times by embedding the previous outputs as additional inputs, using a "recycling" technique without significantly increasing the computational time of the process. This step provides additional precision to the predicted models, whose final quality can be evaluated by the per-residue estimate of confidence, pLDDT (predicted local distance difference test), which is based on the lDDT-Cα metric [33].

From this point forward, for simplicity, "AlphaFold 2" will be referred to as "AlphaFold", unless a different version is specified.

2.2.2. The impact and applications of AlphaFold

After the release of AlphaFold, other research groups and AI companies developed alternative protein structure prediction tools and complementary software that takes advantage of these findings. RoseTTAFold, for example, emerged as a three-track neural network inspired by the DeepMind framework, that simultaneously considers patterns in protein sequences, amino acid interactions and putative protein folding arrangements [34]. Despite the substantially different architectures of the two neural networks, RoseTTAFold achieves results comparable to those obtained by AlphaFold. In other example, Meta AI, the Meta Platforms offshoot, designed a large language model-based protein structure prediction method called ESMFold [35]. While ESMFold does not meet the accuracy of AlphaFold and RoseTTAFold, it is an order of magnitude faster and is

uniquely able to predict accurate structures for orphan proteins with limited sequence homologs. This opened the venue for the prediction of more than 770 million protein structures, released on the ESM Metagenomic Atlas (https://esmatlas.com/). Another recently developed tool, trRosettaX-Single, uses an approach similar to ESMFold and is able to predict the structures of orphan proteins with higher accuracy than AlphaFold, working specifically well in the prediction of structures of designed proteins [36]. Similarly, ProGen emerged as a tool to generate novel protein sequences with predictable function [37]. OmegaFold, which also uses a protein language model to predict protein structures, is now claiming to be the superior structure prediction tool for single-sequence input [38]. Complementary to these tools, efforts are being made to accelerate the accurate prediction of three-dimensional RNA structures [39] and to incorporate ligands and cofactors in protein models [40]. All these developments in the field of protein and RNA structure prediction are expected to greatly accelerate efforts to understand the building blocks of cells and their mechanisms, thus enabling the implementation of faster and optimal drug discovery, and other technologies [41].

Since its inception, and as of February 2023, AlphaFold alone has been the focus of more than ten thousand published works that highlight its versatility and suggest its use either in conjunction with other structural biology tools [42] or to deconvolute specific information [43, 44]. In 2021, DeepMind published the results of applying AlphaFold to predict highly accurate structures for 98.5% of the human proteome (20,000 proteins in total) [45]. One year later, the AlphaFold protein structure database (https://alphafold.ebi.ac.uk/) was expanded to more than 200 million predicted protein structures, covering hundreds of thousands of species and almost the entirety of the Uniprot database [46]. In parallel, DeepMind released a model trained specifically for multimeric inputs of known stoichiometry called AlphaFold-Multimer, which significantly increased the precision of predicted multimeric interfaces over input-adapted single-chain AlphaFold, while maintaining high intra-chain accuracy [47]. This approach is still unable to provide results as reliable as those produced by AlphaFold, but several developments are expected to emerge in the upcoming years [48]. These will not only enable the prediction of the correct geometries of protein complexes, but will also improve the prediction of precise protein-protein and protein-peptide interactions [49-51]. Both AlphaFold and AlphaFold-Multimer can be run through ColabFold, a free and accessible platform that offers accelerated prediction of protein structures and complexes [52].

Taking advantage of these tools, in this chapter, we demonstrate how AlphaFold can be used to depict the structural features of the five porin-cytochrome complexes located in the outer membrane of *G. sulfurreducens*. The conserved structural analysis and the rationale presented are transverse to almost all organisms and can be applied to several systems and different groups of proteins, with special emphasis on multiheme cytochromes.

2.2.3. Structural insights on porin-cytochrome complexes

Electron transfer across the outer membrane of Gram-negative bacteria is achieved by different mechanisms, involving self-produced redox mediators [53, 54] or outer membrane cytochromes [55]. The latter can be attached to the outer membrane by lipid anchors, form porin-cytochrome

fusions or be part of porin-cytochrome complexes for short-range electron transfer [56, 57]. Alternatively, these cytochromes can assemble into nanowires for long-range electron transfer [58].

Porin-cytochrome fusions are single proteins composed by a porin domain, in which the β-strands are wrapped around an N-terminal heme-containing domain that faces the periplasm (Figure 2.2) [59]. The exact mechanisms by which these proteins transfer electrons to extracellular electron acceptors have not been clarified. The most likely hypothesis requires soluble iron to diffuse within the barrel in order to be reduced or oxidized by the periplasmic facing heme-containing domain. Examples of porin-cytochrome fusions are the Cyc2 from *Acidithiobacillus ferrooxidans* [60] and the cytochrome 572 from *Leptospirillum ferrooxidans* [61], both containing a single *c*-type heme. Recent studies have shown that these proteins may form oligomers [62], a feature that is common in porin-like proteins [63].

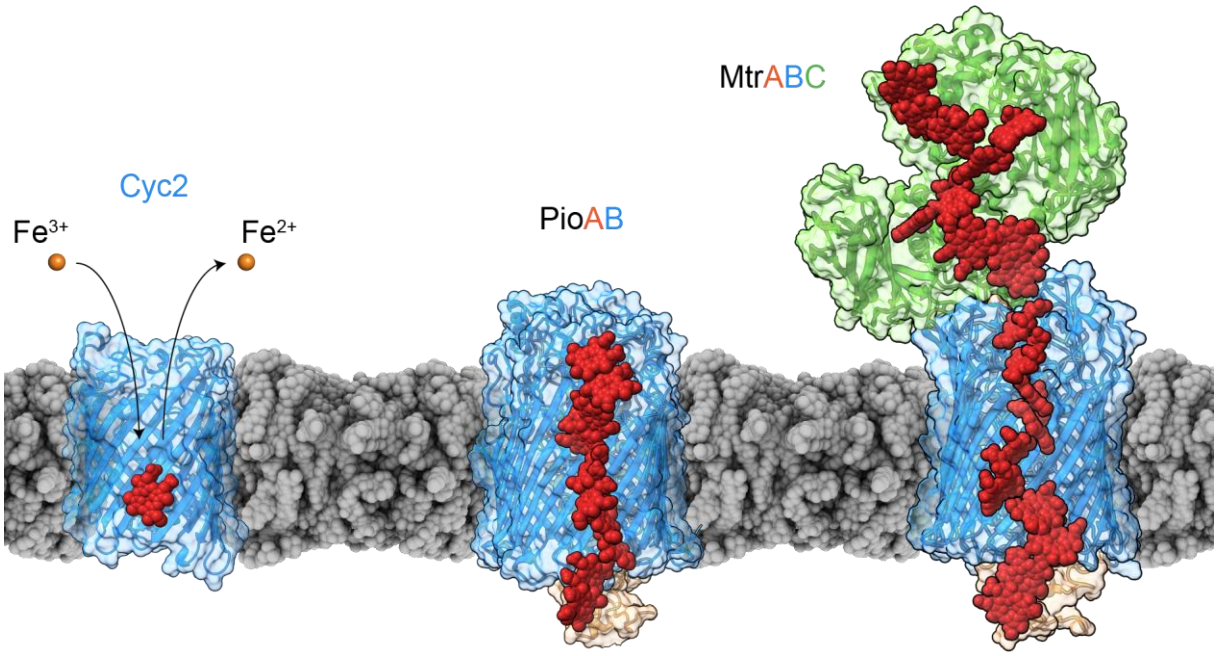


Figure 2.2. Bacterial porin-cytochromes. The models of the Cyc2 porin-cytochrome fusion from *A. ferrooxidans* and of the PioAB porin-cytochrome complex from *Rhodopseudomonas palustris* TIE-1 were generated with AlphaFold [21]. Their heme groups were inserted using a dedicated PyMol script (see section 2.6.1). The pLDDT scores of each model can be found in Figure 7.2.The structure of the MtrABC porin-cytochrome complex from *Shewanella baltica* (PDB: 6R2Q [56]) is also represented. The heme groups are represented as red sphere-type sticks.

Porin-cytochrome complexes commonly require two or three proteins to be assembled, and are capable of reducing insoluble metals. Two-component porin-cytochrome complexes are formed by a porin-like protein that crosses the outer membrane, and by a cytochrome that is embedded within the β-strands of the porin-like protein. These cytochromes are long enough to directly bridge the periplasm of the bacterium with the extracellular environment. Examples of two-component porin-cytochrome complexes are the PioAB complex from the non-sulfur bacterium *Rhodopseudomonas palustris* TIE-1 [64] (Figure 2.2) and the MtoAB complex from the chemolithoautotrophic bacterium

Sideroxydans lithotrophicus ES-1 [65], both composed by a decaheme cytochrome (PioA and MtoA, respectively) and a porin-like β-barrel protein (PioB and MtoB, respectively). Three-component porin-cytochrome complexes are the most common, containing two cytochromes that bridge the gap between the periplasm and the cell exterior. One example of this type is the MtrABC complex of *Shewanella baltica* (Figure 2.2), whose structure has been experimentally determined [56].

The genome of *G. sulfurreducens* contains five putative gene clusters that code porin-cytochrome complexes [66]. Out of these, OmabcB, OmabcC and ExtEFG likely assemble into three-component systems, whereas ExtABCD and ExtHIJKL encode four and five proteins, thus possibly constituting new types of porin-cytochrome complexes. The functional and structural features of each of these complexes are described on the following sections. The pLDDT average scores and respective color code of the models presented in this chapter may be found in section 7.1.

2.3. Results and discussion

2.3.1. OmabcB and OmabcC

The OmabcB and OmabcC complexes were the first putative porin-cytochrome assemblies identified in *G. sulfurreducens* and are widely distributed across phylogenetically diverse bacteria [67]. The exact role of these complexes in EET is unclear, since conflicting data emerged from several techniques and growth conditions over the years [67-71]. In fact, it was this apparent variability and lack of specificity that paved the way for the discovery of alternative pathways that catalyze electron transfer across the outer membrane, including the existence of additional porin-cytochrome gene clusters [66, 72]. The OmabcB and OmabcC gene clusters encode three proteins each, with different structural properties.

OmaB and OmaC

OmaB (GSU2738) and OmaC (GSU2732) are identical 28 kDa octaheme *c*-type cytochromes composed by 231 residues, including a 24-residue signal peptide that locates them at the periplasm. To better visualize the structural properties of these multiheme cytochromes, their AlphaFold model was enriched with *c*-type hemes, which were placed in silico using a dedicated PyMol script that restrains each heme group to a position in which the distances to the cysteine and histidine residues of a specific CXXCH heme-binding motif are locally optimal (see section 2.6.1). The final model shows a curved structure mainly composed by α-helices and random coil secondary structural elements (Figure 2.3).

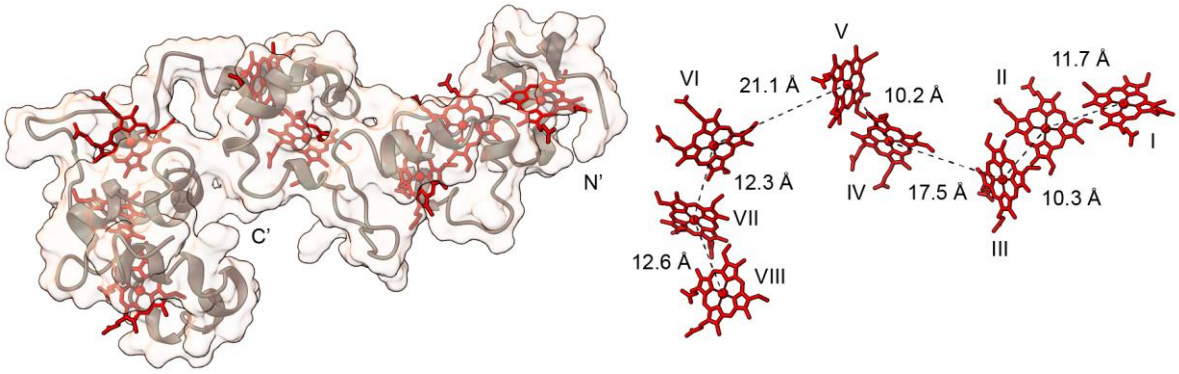


Figure 2.3. Structural features of cytochromes OmaB/C. The AlphaFold model of OmaBC is represented as surface (light salmon) and ribbon (gray). The hemes are represented as red sticks. The N- and C-terminal regions of the protein are identified. The roman numerals indicate the hemes in their order of attachment to the CXXCH motifs in the polypeptide chain. The Fe-Fe distances between the adjacent porphyrin rings are indicated.

The 8 hemes are considerably exposed, axially coordinated by two histidine residues and distributed along the three protein domains. The N- and C-terminal domains contain 3 *c*-type hemes with different spatial organizations and Fe-Fe distances ranging between 10 and 13 Å, which facilitate electron tunneling and ensure an efficient electron transfer [73]. The central domain contains two hemes in a T-shaped conformation, which usually increases structural stability [74]. The Fe-Fe distances between the hemes of this domain and those of the N- and C-terminal domains are higher than 17.5 Å, and their minimum edge-to-edge distances are slightly over 15 Å.

Since the edge-to-edge distances of adjacent porphyrins in multiheme cytochromes have been observed to lay typically within a 14 Å distance [73], the heme organization observed in the OmaB/C putatively hampers an efficient electron transfer. Therefore, in practice, the structure of these proteins should be slightly more packed that what is predicted by AlphaFold, meaning that the flexible protein stretches that connect the three domains should have a tighter arrangement that brings the heme groups closer together. The lack of consistent structural homologs for these proteins results in a low average pLDDT score (Figure 7.3). After searches for structural homologs of the OmaB/C cytochromes in the PDB database [7, 8], using the Dali [75], FoldSeek [76] and PDBeFold [77] servers, the only hit was a bifunctional enzyme from the yeast *Candida glabrata* THI6 [78], involved in thiamin biosynthesis, which lacks heme groups and is therefore unsuitable for structural comparison.

AlphaFold has been shown to be able of sampling different protein ensembles and conformational states by tuning certain parameters, thus conferring high plasticity to the structure determination process of sequences with a low number of homologs [79-83]. An explicit proof of this concept is the fact that the OmaB/C model obtained by running AlphaFold on ColabFold (Figure 2.3) is quite different from the one deposited on the AlphaFold protein structure database (Figure 2.4).

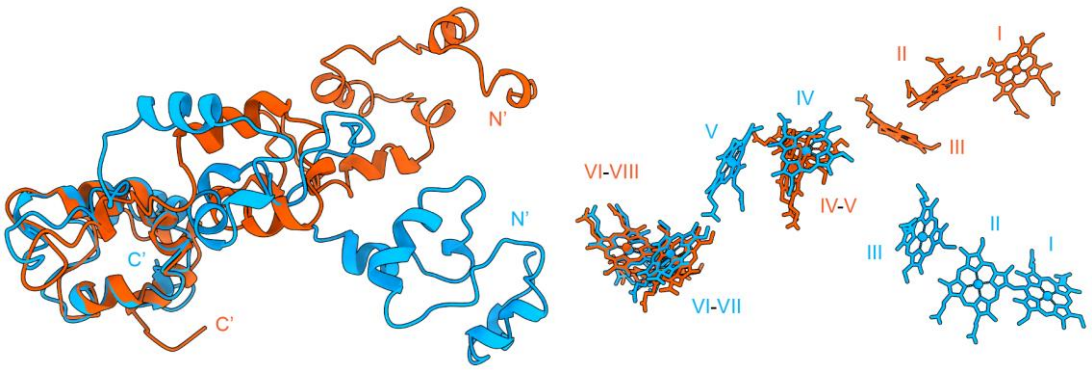


Figure 2.4. Relative arrangement of OmaB/C domains predicted by AlphaFold. The superposed models were generated resorting to ColabFold (orange) or retrieved from the AlphaFold protein structure database (blue). The protein backbones are represented as cartoons and the hemes as sticks. The N- and C-terminal regions of the protein are identified. The roman numerals indicate the hemes in their order of attachment to the CXXCH motifs in the polypeptide chain.

Specifically, although the overall fold and heme arrangement is maintained, the relative position of the domains differs significantly, thus affecting the distances between the heme groups. This data supports the hypothesis that the structure can be considerably tighter when in complex with the remaining proteins of the respective porin-cytochrome complexes.

OmbB and OmbC

As for the OmaB and OmaC proteins, OmbB (GSU2739) and OmbC (GSU2733) are identical transmembrane β-barrel porins composed by 403 residues, including a 24-residue signal peptide, accounting for a total molecular weight of 42.9 kDa. DeepTMHMM [84] predicts that OmbB/OmbC contain 22 transmembrane motifs, which is in agreement with the structural model predicted by AlphaFold (Figure 2.5).

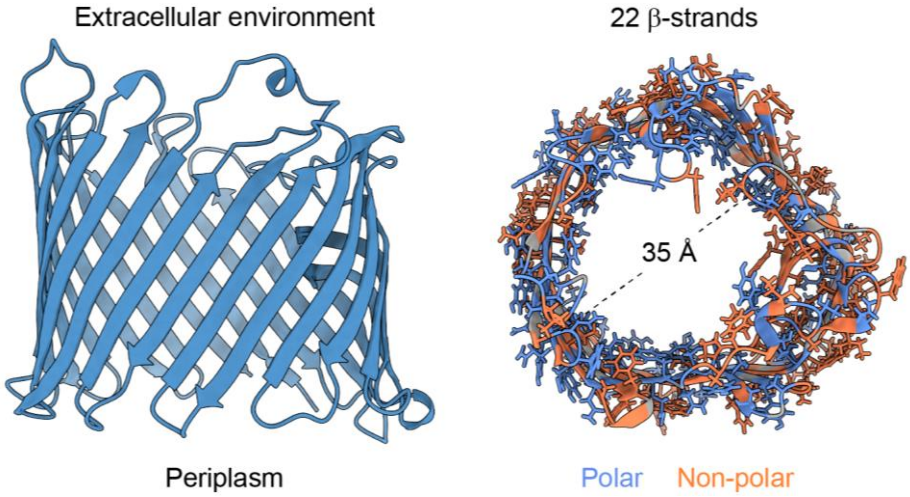


Figure 2.5. Structural features of the OmbB/C transmembrane β-barrel porins. The AlphaFold model of the OmbB/C proteins is represented as a blue cartoon on the left. The polar and non-polar residues on the right are colored blue and orange, respectively.

The N-terminal of porins usually faces the periplasm and can be used as a spatial reference for relative orientation of these proteins [85]. Consequently, one can define which side of the porin is facing either the periplasm or the cell exterior (Figure 2.5). Typically, in porins, the non-polar residues face outward to interact with the non-polar lipids of the outer membrane, whereas the polar residues face inwards, into the center of the porin [86]. The β-strands of OmbB/C randomly alternate between polar and non-polar residues and do not exactly replicate the mentioned pattern, but do create a pore with a 35 Å diameter where OmaB/C are expected to dock (Figure 2.5).

OmcB and OmcC

The components of the OmabcB and OmabcC porin-cytochrome complexes that are expected to face the extracellular environment are the OmcB (GSU2737) and OmcC (GSU2731) dodecaheme *c*-type cytochromes. These proteins are the only distinctive elements of these complexes, a feature that might be explained by the increased selective pressure to which these proteins are subjected due to the large number of putative interacting partners and variable extracellular conditions. OmcB and OmcC are large 82.3 kDa and 85.7 kDa cytochromes, composed by 744 and 768 residues, respectively, including an identical 23-residue signal peptide. These proteins share an overall sequence identity of 66% and their AlphaFold models possess an RMSD of 0.8 Å across 580 residues, demonstrating their high structural homology (Figure 2.6).

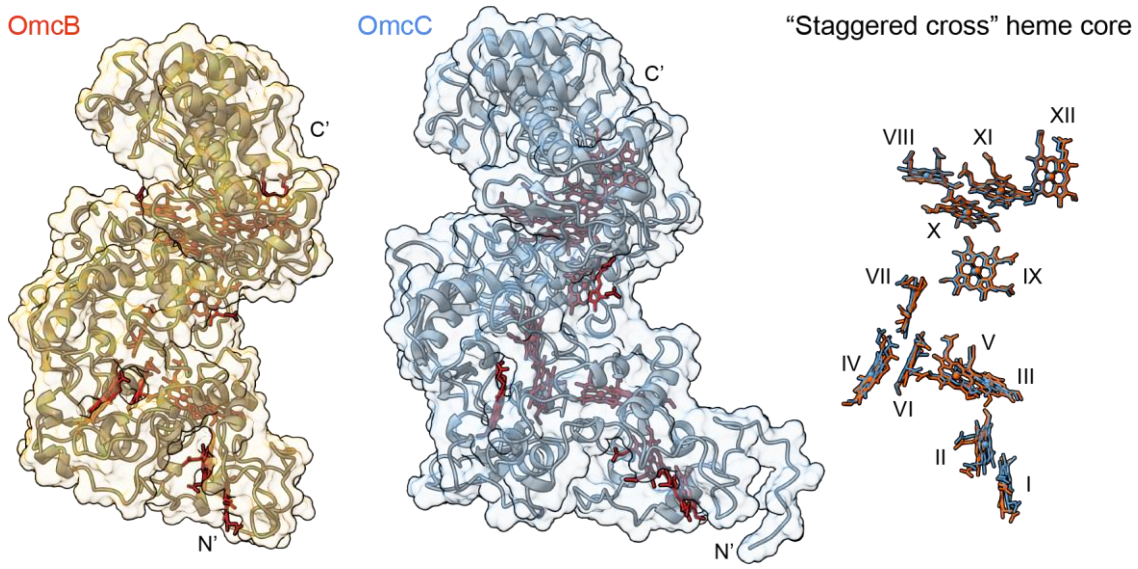


Figure 2.6. Structural features of the OmcB and OmcC dodecaheme *c*-type cytochromes. The AlphaFold models of the OmcB and OmcC are shown as ribbon (both in gray) and surface (orange and blue, for OmcB and OmcC, respectively). In these models, the hemes are represented as red sticks. The N- and C-terminal regions of the proteins are identified. The "staggered cross" formation of the heme core is displayed on the right side, for OmcB (orange) and OmcC (blue). The roman numerals indicate the hemes in their order of attachment to the CXXCH motifs in the polypeptide chain.

The heme core of both proteins is conserved, with hemes displayed along a large surface area for electron exchange (Figure 2.6). This heme arrangement is similar to the one observed in MtrC from

S. baltica, which was classified as a "staggered cross" formation [87]. OmcB/C contain two sets of four parallel hemes, bridged by two pairs of hemes that are perpendicular to those sets. The iron-iron distances between adjacent hemes are lower than 14 Å, meaning that OmcB/C possess the necessary heme arrangement for continuous and efficient electron transfer.

A puzzling feature of the OmcB/C cytochromes is the C-terminal domain, absent of *c*-type heme groups, and mainly composed by α-helices (Figure 2.6). A search for homologs of OmcB/C showed that they share homology with several multiheme cytochrome *c* nitrite reductases. These enzymes catalyze the six-electron reduction of nitrite to ammonia as one of the key steps in the biological nitrogen cycle [88], participating in the anaerobic energy metabolism of dissimilatory nitrate ammonification [89]. Cytochrome *c* nitrite reductase (Ccnr) is a functional dimer, with 10 close-packed *c*-type heme groups, two of which containing an unusual lysine-coordinated high-spin heme, forming the active sites [90]. The heme core of this enzyme is remarkably similar to the one formed by the hemes VIII-XII of OmcB/C, which lay closer to the C-terminal domain that is structurally homologous to the fold of cytochrome *c* nitrite reductase (Figure 2.7).

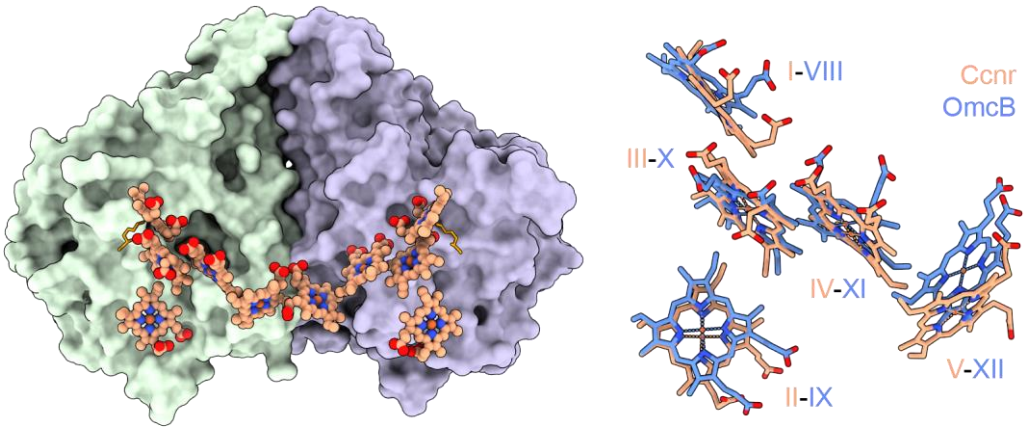


Figure 2.7. Structural homology between Ccnr and OmcB/C. The monomers of the Ccnr dimer (PDB: 1QDB [90]) are colored in green and purple. The 5 *c*-type hemes of each monomer are represented as spheric sticks in salmon. The lysine residue coordinating the catalytic heme (heme I) is highlighted in gold. The right panel shows the superposition of OmcB's hemes VIII-XII (blue) and the hemes of Ccnr's green monomer (salmon). The roman numerals indicate the hemes in their order of attachment to the CXXCH motifs in the polypeptide chain.

Nevertheless, despite the considerable structural similarities, there is no experimental evidence that supports a role for OmcB/C in nitrite reduction. Furthermore, according to the model predicted by AlphaFold, the inexistence of a lysine-coordinated high-spin heme hampers a putative catalytical role for these proteins.

OmabcB and OmabcC assemble into MtrABC-like porin-cytochrome complexes

After digging into the structural properties of each component of the OmabcB and OmabcC complexes, we used AlphaFold-Multimer to determine a model of these heterooligomers, to gain insight on their organization and geometry. Apart from providing information on the overall folding of the complex, AlphaFold-Multimer can also be used to understand the relative orientation of the

complexes by determining which proteins are facing the periplasm or the extracellular environment. In addition to signal peptide sequence-based predictions, the position of the OmaB/C and OmcB/C proteins relative to the OmbB/C porins can be used as reliable hints for definition of their cellular localization. AlphaFold-Multimer predicts that OmabcB and OmabcC assemble into MtrABC-like porin-cytochrome complexes, with the octaheme OmaB/C embedded inside the OmbB/C porin, and the dodecaheme OmcB/C bulging towards the extracellular environment (Figure 2.8).

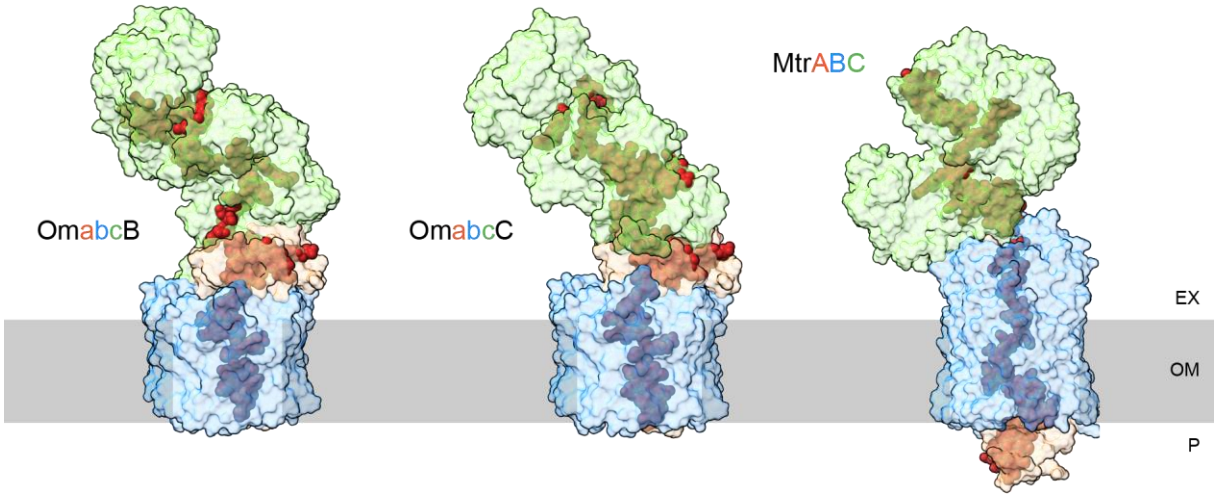


Figure 2.8. AlphaFold-Multimer models of OmabcB/C complexes. The surface of the models is shown, with the heme groups represented as red sphere-type sticks. For comparison, the crystal structure of the MtrABC porin-cytochrome complex from *S. baltica* (PDB: 6R2Q [56]) is also represented. EX, OM and P stand for cell exterior, outer membrane and periplasm, respectively.

The OmabcB and OmabcC complexes contain a 20-heme arrangement in which electrons move across 140-150 Å. These values are considerably lower than what is observed for the electron pathway in MtrABC (185 Å) [56], in which the 20 hemes are arranged more linearly, thus spreading across a greater distance (Figure 2.8).

As postulated, the individual structures of the OmaB/C cytochromes change considerably upon assembly with the other proteins from the respective complexes, such that in these cases the Fe-Fe distances between all pairs of adjacent hemes are lower than 15Å. These proteins fit within OmbB/C, but their C-terminal domains protrude out of the porin, towards the extracellular side of the complex. This might imply that OmbB/C can directly reduce extracellular acceptors. The opposite is observed in the MtrABC complex, in which MtrA protrudes slightly from the periplasmic side of MtrB, having a higher surface area for interaction with putative periplasmic partners. The heme chains of OmaB/C are perpendicular to the membrane, likely to prevent non-specific reduction of liposoluble exogenous molecules such as oxygen, which could result in the generation of ROS and consequent cellular damage due to lipid peroxidation [56].

In contrast with the models of OmaB/C, the structures of the extracellular OmcB/C cytochromes are identical in the individual and in-complex predictions. Interestingly, the OmbB/C models do not possess extended surface loops on the extracellular side of the membrane, as observed for MtrB.

These loops contain a uniform negative charge that likely assists the docking of extracellular catalytic domains, by preventing the negatively charged lipopolysaccharides from binding to the external surface of the membrane embedded porin-cytochrome complexes [56, 91]. This question has to be addressed for the case of *Geobacter*. The structural features of OmabcB and OmabcC complexes are globally comparable to those of MtrABC, but encompass localized differences that might imply differentiated functions.

2.3.2. ExtABCD

The *extABCD* gene cluster is essential for electrode and soluble metal reduction [66]. An engineered *G. sulfurreducens* strain lacking the other four outer membrane porin-cytochrome complexes (*extABCD*$^{+}$) was shown to grow faster and produce greater current density than the WT bacterium [92]. This increased performance was attributed to the additive effects of denser cell packing close to the electrode, higher metabolic rates per cell and increased electron transfer rates due to simplified, streamlined EET pathways [92]. This interesting gene cluster encodes four proteins with different structural properties.

ExtA

ExtA (GSU2645) is a 37.8 kDa periplasmic dodecaheme *c*-type cytochrome composed by 310 residues, including a 32-residue signal peptide. The AlphaFold model of ExtA (Figure 2.9) shows that the protein contains an elongated structure composed of α-helix and random coil elements, with 12 *c*-type hemes displayed across a 90 Å distance with alternating parallel and T-shaped arrangements.

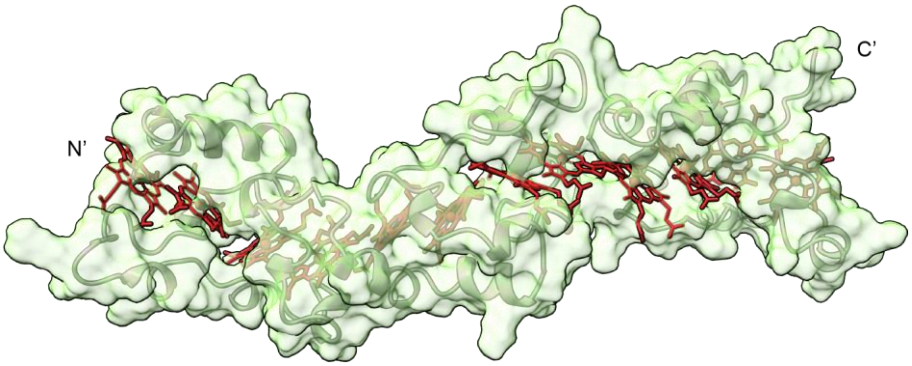


Figure 2.9. Structural features of ExtA. The AlphaFold model of ExtA (residues 72-310) is represented as surface (light green) and ribbon (gray). The N- and C-terminal regions of the protein are identified. The hemes are represented as red sticks. For simplification, the N-terminal region of ExtA (residues 33-70) is not represented.

The N-terminal region of ExtA is composed by a 40-residue long and apparently flexible tail, which is not predicted to form any transmembrane motifs [84], and might be relevant for the assembly of the putative ExtABCD porin-cytochrome complex. The 12 hemes of ExtA are bis-histidynil coordinated and highly exposed to solvent (as expected for a ratio of 22 residues/heme), suggesting that they likely have considerably negative reduction potential values. Interestingly, heme VIII contains a rare CX_3CH heme binding motif. After searches for structural homologs of ExtA, the only

hit was a membrane-associated cytochrome from the archaea *Ignicoccus hospitalis* [93], whose heme arrangement is not conserved with that of ExtA. A closer analysis of the structure shows that the protein can be divided in three tetraheme domains connected by small linkers, each containing 80 residues and two pairs of parallel hemes that are anti-parallel to each other (Figure 2.10).

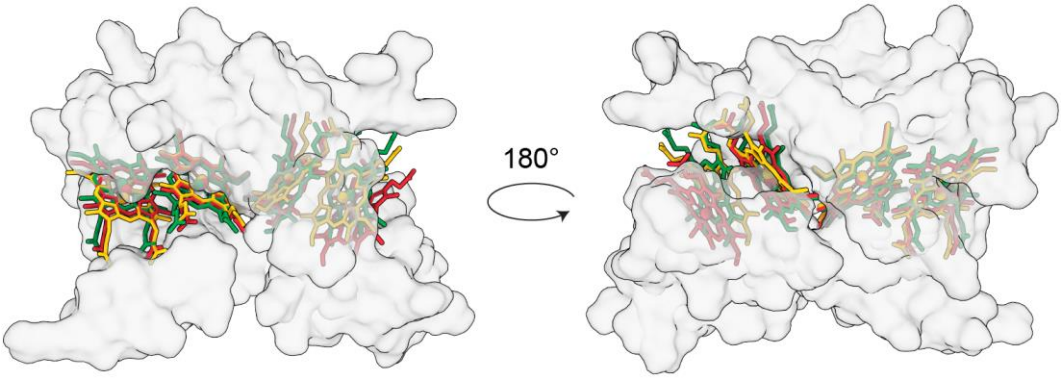


Figure 2.10. Structural conservation of tetraheme domains of ExtA. The heme groups of the three tetraheme domains of ExtA are represented in green, yellow and red. The surfaces of the three domains are represented in light gray.

This relative arrangement of the four hemes differs from those encountered in the tetraheme cytochromes c_3 [94], STC [95], OmcE [96] and flavocytochrome *c* [97], suggesting an unprecedented heme core architecture for each domain of ExtA. In this disposition, each pair of parallel hemes is exposed in opposite faces of the protein, thus maximizing the total heme solvent exposure. The particular structural organization of ExtA makes this cytochrome a suitable candidate for the application of the strategy previously described for the GSU1996 dodecaheme cytochrome (Figure 1.11), in which each of the four individual triheme domains was separately produced to assist the complete thermodynamic characterization of the protein [98].

ExtB

ExtB or OmbW (GSU2644) is composed by 394 residues, including a 26-residue signal peptide, accounting for a total molecular weight of 44.2 kDa. DeepTMHMM [20] predicts that ExtB contains 22 transmembrane motifs. This is consistent with the AlphaFold model of ExtB, which shows that the protein forms a typical β-barrel porin structure, with 22 anti-parallel β-strands (Figure 2.11).

The protein is structurally homologous to the OmbB/C proteins, with a RMSD of 0.97 Å across 233 atom pairs of the two models, despite only sharing 18% of pairwise sequence identity (Figure 2.11). Likewise, the distribution of polar and non-polar residues is conserved, as well as the pore diameter size (35 Å).

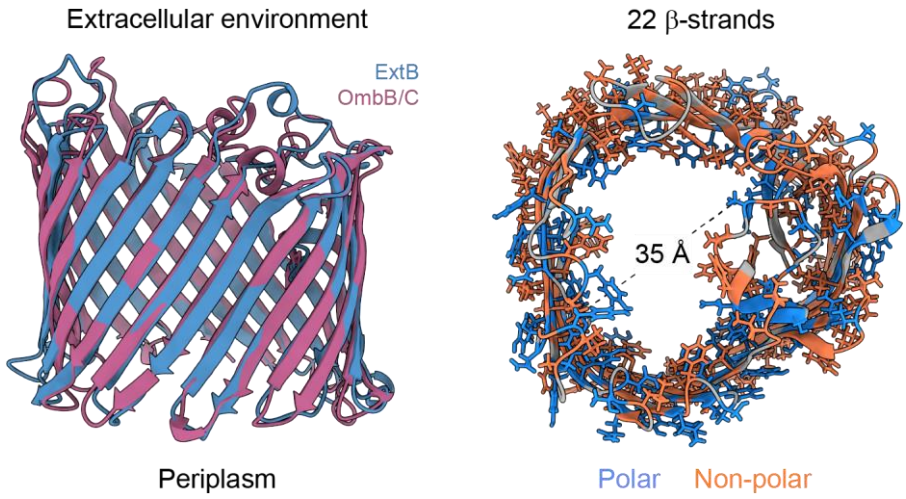


Figure 2.11. Structural features of ExtB transmembrane β-barrel porin. The left panel shows the superposed AlphaFold models of the ExtB and OmbB/C proteins, represented as blue and purple cartoons, respectively. The polar and non-polar residues on the right are colored blue and orange, respectively.

ExtC

ExtC or OmaW (GSU2643) is an 18.5 kDa pentaheme *c*-type cytochrome composed by 155 residues, 24 of which correspond to a signal peptide. The AlphaFold model of ExtC shows that the protein is composed by α-helical structures and random coil elements that fold around 5 *c*-type bis-histidynil coordinated hemes, which spread across a 40 Å distance (Figure 2.12).

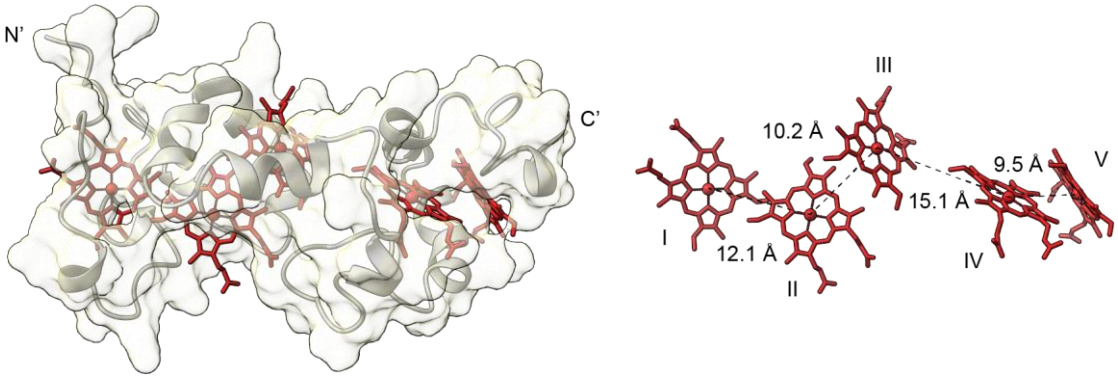


Figure 2.12. Structural features of ExtC. The AlphaFold model of ExtC is represented as surface (white) and ribbon (gray). The hemes are represented as red sticks. The N- and C-terminal regions of the protein are identified. The roman numerals indicate the hemes in their order of attachment to the CXXCH motifs in the polypeptide chain. The Fe-Fe distances between the adjacent porphyrin rings are indicated.

The 5 *c*-type hemes are surface exposed and positioned such that the edge-to-edge distances between adjacent hemes are lower than 14 Å, allowing for fast and efficient intramolecular electron transfer. A search for homologs showed that ExtC possesses a unique folding.

ExtD

ExtD or OmcW (GSU2642) is a 26.2 kDa hexaheme *c*-type cytochrome composed by 229 residues, including a 24-residue signal peptide. AlphaFold has particular difficulties in predicting a reliable model for ExtD. In fact, in an extent greater than what is observed in the OmaB/C models, AlphaFold shows great variability in predicted models for ExtD, with consistently low average pLDDT scores (<60, Figure 7.4). Ranging from tighter and compact, to more relaxed and elongated models, the predicted foldings consistently place the hemes too close together, resulting in incompatible clashes (Figure 2.13).

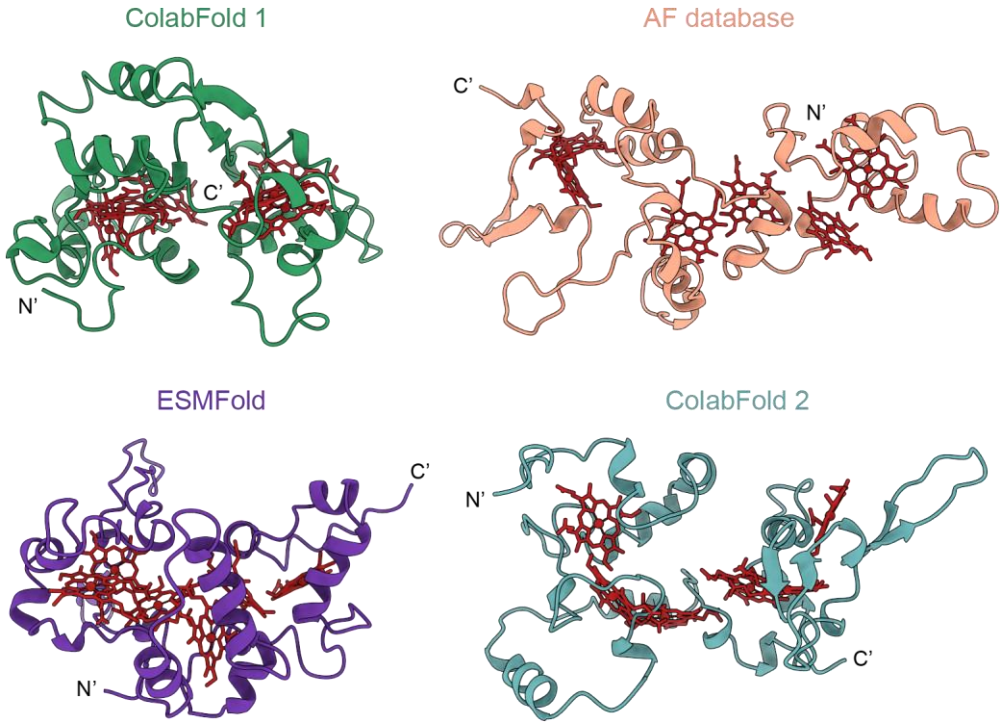


Figure 2.13. Extended and compacted models of the ExtD hexaheme cytochrome. The backbone of the predicted models is represented as ribbon. The hemes are represented as red sticks. The N- and C-terminal regions of the different models of the protein are identified.

Additionally, in seek of models with higher reliability, the ESMFold algorithm [35] was used to predict the structure of ExtD. However, low average pLDDT scores were also obtained (Figure 7.4). Furthermore, as expected, no structural homologs were found after a search in different databases. Nevertheless, and despite the inconsistencies in the predicted AlphaFold models, consistent structural features can be emphasized. The protein is mainly composed by α-helices and random coil elements, while all models predict two β-sheets forming a β-strand at the C-terminal domain (Figure 2.13). Additionally, based on the number of histidine residues (12), it is conceivable to assume that all 6 hemes are bis-histidynil coordinated. While two of those histidines are consecutive in the protein sequence (His^{177} and His^{178}), they might still work as heme distal ligands, considering the relative position of the axial ligands in tetraheme cytochromes from *Desulfovibrio* species [99, 100].

Interestingly, ExtD contains 22 proline residues, which accounts for 10% of the protein, spreading across the entire structure. The structurally disruptive nature of these residues might be an additional parameter that hinders the prediction of reliable models for this protein.

ExtA is excluded from the ExtBCD porin-cytochrome complex

The fact that the *extABCD* gene cluster encodes a number of proteins that is higher than those observed on the three-component porin-cytochrome complexes, was an initial indication that this putative porin-cytochrome complex would either have a novel and unique organization, or that, alternatively, one of its components is co-expressed but not assembled into the protein complex. In fact, after several AlphaFold-Multimer model predictions with the different combinations of the ExtA-D proteins, we observed that ExtB, ExtC and ExtD form a transmembrane porin-cytochrome complex, while ExtA does not assemble with these proteins (Figure 2.14).

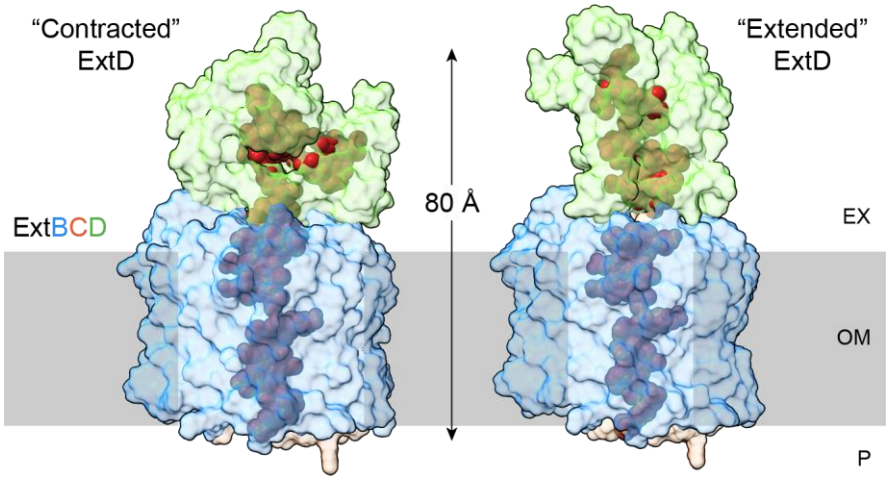


Figure 2.14. AlphaFold-Multimer models of the ExtBCD complex. The surface of the models is shown, with the heme groups represented as red sphere-type sticks. EX, OM and P stand for cell exterior, outer membrane and periplasm, respectively.

Based on these predictions, we believe that ExtA is part of the group of periplasmic cytochromes that bridges the electron transfer between the inner and outer membranes. On one hand, this leaves an open question regarding the structural and functional relevance of the intrinsically disordered 40-residue N-terminal tail. On the other hand, ExtBCD forms a small porin-cytochrome conduit, with a structural arrangement similar to other complexes.

AlphaFold-Multimer predicts that ExtC is embedded inside ExtB, while ExtD protrudes towards the cell exterior (Figure 2.14). Upon assembly, the foldings of ExtB and ExtC are similar to those predicted for their individual cases, whereas that of ExtD presents the same inconsistencies encountered in the models of the individual protein. This feature hinders the prediction of a reliable ExtBCD model, however, one can affirm that like in the MtrABC and OmabcB/C complexes, the hemes of ExtC are perpendicular to the membrane, with ExtB working as an insulator that prevents the formation of ROS [56] (Figure 2.14). Nevertheless, ExtC does not protrude from either side of the porin, as observed in other complexes, meaning that either the periplasmic partners, or ExtD, will require

close surface contacts for efficient electron transfer. Additionally, and despite the unclear conformation of ExtD, the overall heme arrangement of ExtBCD will only allow electron transfer events across a distance of 70-90 Å, which is considerably smaller than those found on the structure of MtrABC (185 Å) and the models of OmabcB/C (140-150 Å). Nonetheless, this was somewhat expected given their architecture and higher number of heme groups.

2.3.3. ExtEFG

The *extEFG* gene cluster is essential for Fe(III) citrate reduction, as strains lacking this cluster were only able to reduce 65% of WT levels [66]. This gene cluster is part of *G. sulfurreducens*' hydrogenase family transcriptional unit, which includes a homolog of the YedY family of periplasmic proteins involved in repair systems of oxidized proteins in *E. coli* [101]. Additionally, this transcriptional unit also encodes an NiFe hydrogenase, similar to the bidirectional Hox hydrogenases used to recycle reducing equivalents in cyanobacteria [102-104].

ExtE

ExtE or OmbV (GSU2726) is composed by 431 residues, including a 21-residue signal peptide, accounting for a total molecular weight of 45 kDa. The AlphaFold model of ExtB contains 22 transmembrane motifs, in agreement with the prediction of DeepTMHMM [84], showing a typical β-barrel porin structure, structurally similar to those of OmbB/C and ExtB (Figure 2.15). In fact, their superposed structures possess an RMSD between 320 atom pairs of ~1 Å (Figure 2.15). Consequently, the distribution of polar and non-polar residues is conserved, as well as the pore diameter size (35 Å).

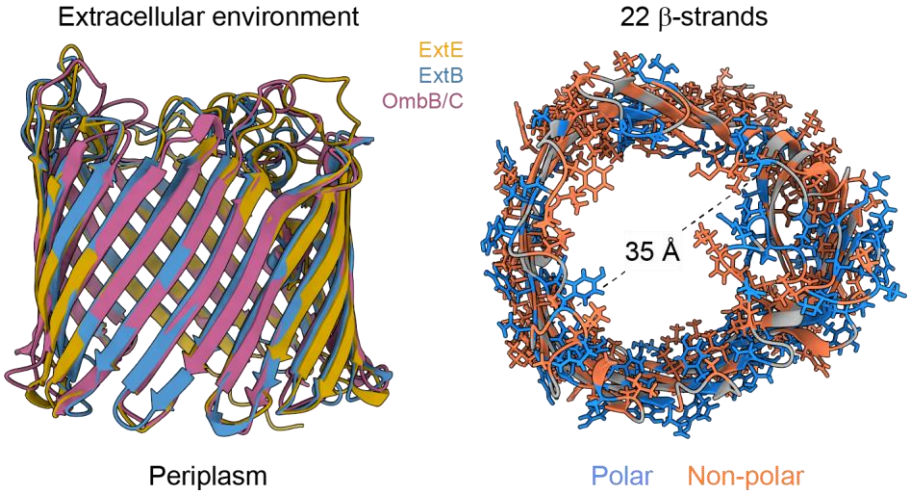


Figure 2.15. Structural features of ExtE transmembrane β-barrel porin. The left panel shows the superposed AlphaFold models of the ExtE, ExtB and OmbB/C proteins, represented as yellow, blue and purple cartoons, respectively. The polar and non-polar residues on the right are colored blue and orange, respectively.

ExtF

ExtF or OmaV (GSU2725) is an 18.4 kDa pentaheme *c*-type cytochrome composed by 157 residues, 24 of which correspond to a signal peptide. The AlphaFold model of ExtF shows that the protein is composed by α-helical structures and random coil elements that fold around 5 *c*-type bis-histidynil coordinated hemes, which spread across a 40 Å distance (Figure 2.16).

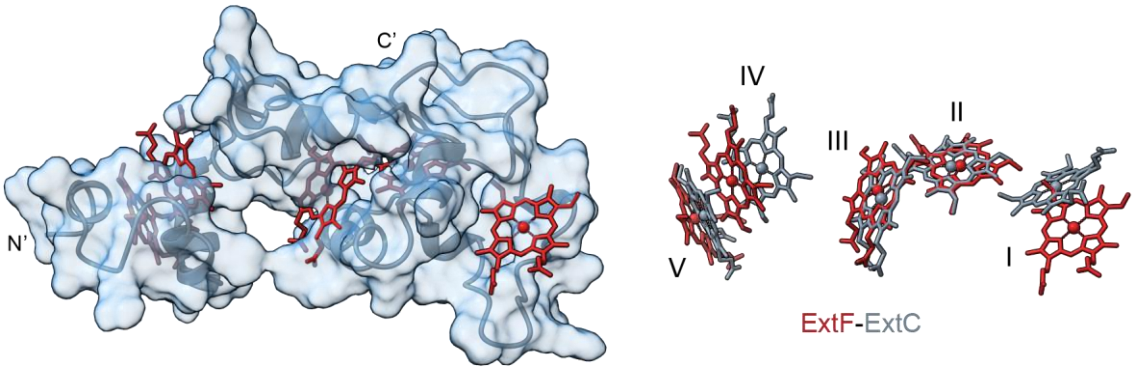


Figure 2.16. ExtC and ExtF are structural homologs. The ExtF cytochrome is shown as a blue surface representation with a gray ribbon. The ExtF hemes are colored red. The ExtC hemes are colored gray. The N- and C-terminal regions of the protein are identified. The roman numerals indicate the hemes in their order of attachment to the CXXCH motifs in the polypeptide chain.

These features are quite similar to the ones that were observed in the model of ExtC, with which ExtF presents a pairwise sequence identity of 43.6%. In fact, by superposing the two models, one can verify that the heme arrangement and protein folding are conserved in both proteins, even though the positions of hemes I and IV are slightly different (Figure 2.16). As for ExtC, the 5 *c*-type hemes of ExtF are surface exposed and displayed such that the edge-to-edge distances between adjacent hemes are lower than 14 Å, allowing for fast and efficient intramolecular electron transfer. A search for homologs showed that ExtC and ExtF possess a unique folding.

ExtG

ExtG or OmcV (GSU2724) is a cytochrome composed by 691 residues, whose mature form is predicted to contain 661 residues and 13 *c*-type hemes, according to its number of CXXCH heme binding motifs. However, a closer analysis of its amino acid sequence shows that there are 20 CH pairs of residues, which might indicate a higher number of heme binding sites, via unconserved heme binding motifs. In fact, the resulting AlphaFold model of ExtG shows that this cytochrome contains a remarkable variability of heme binding motifs, and that it might actually contain 20 *c*-type hemes, accounting for a total molecular mass of 79.9 kDa (Figure 2.17).

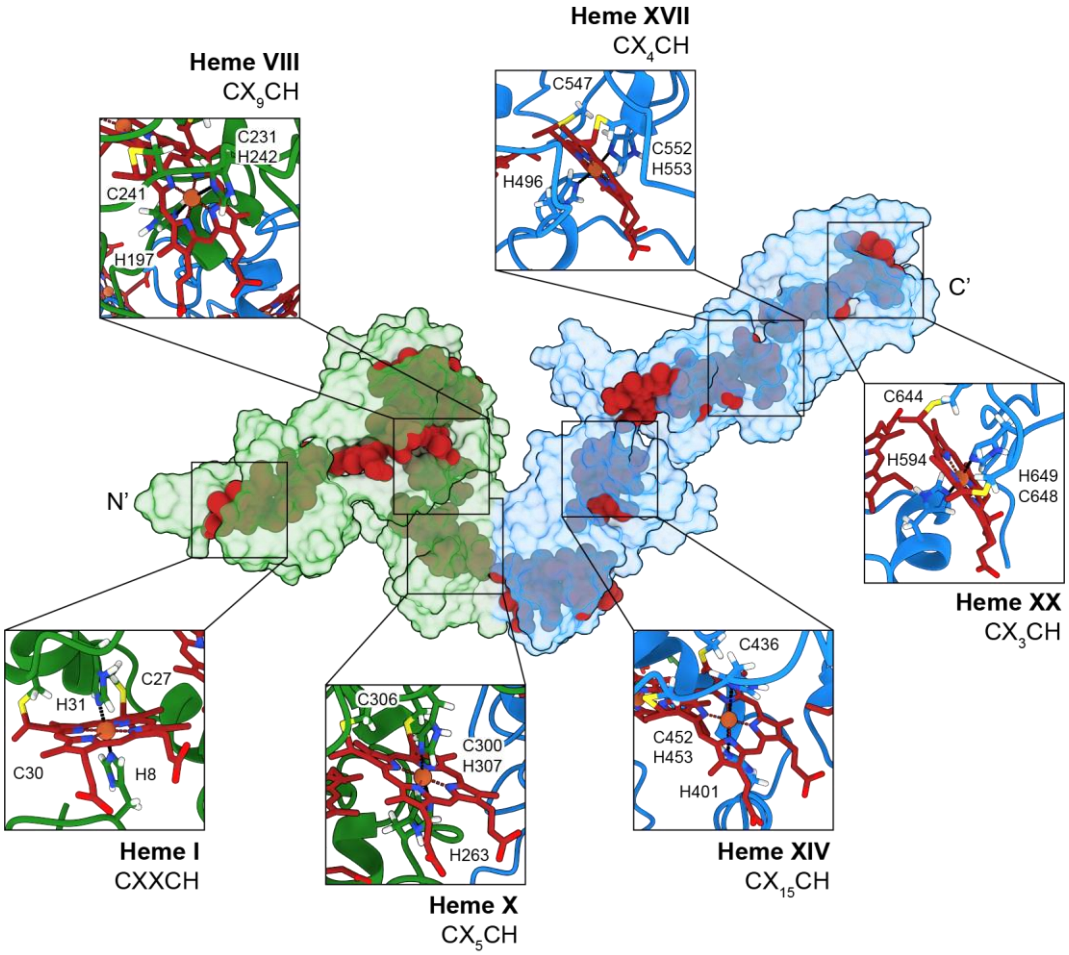


Figure 2.17. ExtG contains multiple non-canonical heme binding motifs. The AlphaFold model of ExtG is shown as a surface representation, with the two distinctive domains colored as green and blue. The heme groups are represented as red sphere-type sticks. The insets show examples of each type of non-canonical heme binding site found in ExtG. For clarity, the structure was rotated in some of the examples. The residues involved in heme ligation and the respective heme groups are represented as sticks. The residue numbering was performed according to the sequence of the mature protein (with no signal peptide). The N- and C-terminal regions of the protein are identified. The roman numerals indicate the hemes in their order of attachment to the respective heme binding motifs in the polypeptide chain.

The structure of ExtG can be divided into two decaheme domains (Figure 2.17). One is L-shaped, while the other is linear and elongated, with both mainly containing α-helices and random coil elements. While 13 of ExtG's 20 hemes contain the classical CXXCH heme binding motif, hemes XX (CX_3CH), XVII (CX_4CH), VI/X (CX_5CH), II/VIII (CX_9CH) and XIV ($CX_{15}CH$) are covalently attached to less typical heme binding motifs. Untypical or non-canonical heme binding motifs have been observed in other cytochromes and usually require dedicated heme maturation systems [105-108]. For example, in trypanosomes and related organisms, the covalent heme attachment occurs via a single thioether bond (AXXCH or FXXCH motifs) [109, 110], while a "contracted" heme binding motif (CKCH) was observed in a tetraheme protein from a hydrazine synthase apparatus of the annamox bacteria *K. stuttgartiensis* [111]. Moreover, other rare heme binding motifs include the CXXCK motif found in the

cytochrome *c* nitrite reductase from *Sulfurospirillum deleyianum* [90] and *Wollinella succinogenes* [112], CX_3CH in the tetraheme cytochrome c_3 from *Desulfovibrio gigas* [113], CX_4CH in the tetraheme cytochrome c_3 from *Desulfovibrio vulgaris* Hildenborough [114], $CX_{14}CH$ in the OmcZ cytochrome from *G. sulfurreducens* [115, 116], $CX_{15}CH$ in the octaheme cytochrome MccA from *W. succinogenes* [117], and $CX_{17}CH$ in other MccA-family proteins [118]. In addition, other studies have shown that the heme incorporation system of *E. coli* is capable of maturing *c*-type cytochromes with artificial CCXXCH, CCXCH, CXCCH, CXXCHC and $CX_{2\text{-}6}CH$ heme binding motifs [119]. Considering this, ExtG contains an unprecedent variety in the number and nature of non-canonical heme binding motifs. These features alone likely make ExtG the most unique cytochrome of *Geobacter*, out of those who have been described to date. A structural homology search found no structural homologs for ExtG.

Arrangement of the putative ExtEFG porin-cytochrome complex

The *extEFG* gene cluster encodes two cytochromes and one β-barrel porin, as observed for the OmabcB/C and MtrABC porin-cytochrome complexes. Since there is strong evidence that ExtG encodes 20 hemes instead of the initially predicted 13, we may affirm that the sum of the number of hemes of ExtF and ExtG makes ExtEFG the complex with the highest number of heme groups in *G. sulfurreducens*. Furthermore, the heme to amino acid ratio is considerably higher in ExtEFG (1:40) when compared to the OmabcB/C (1:47) and MtrABC (1:48) complexes. Nevertheless, this ratio is higher in ExtBCD (1:30), despite the much lower number of *c*-type heme groups (11).

To investigate how this complex assembles, we resorted to AlphaFold-Multimer and inserted the predicted heme groups in silico, as described previously. The first predicted model shows that ExtF is located within ExtE, while protruding towards the extracellular side of the porin (Figure 2.18A).

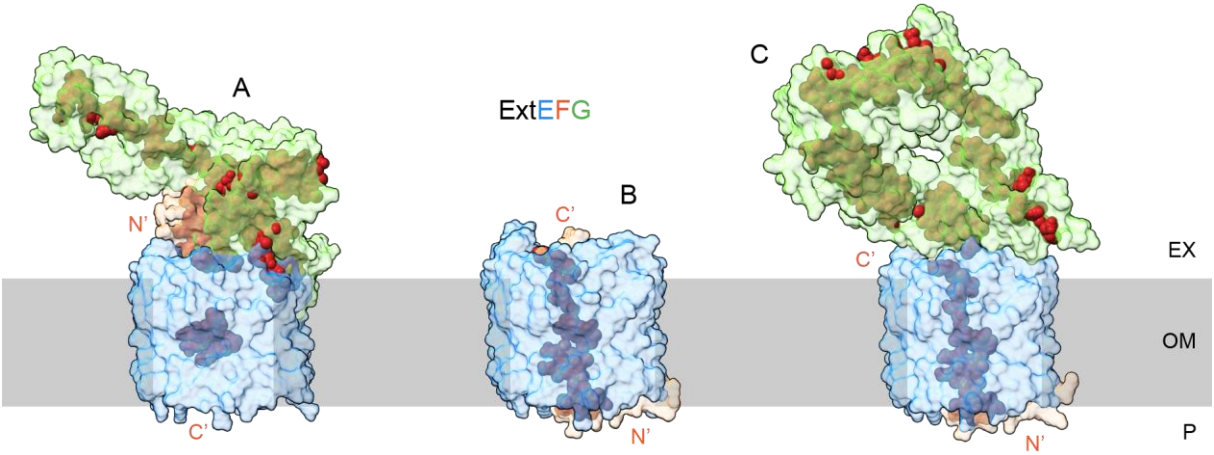


Figure 2.18. AlphaFold-Multimer models of the ExtEFG complex. The surfaces of the ExtEFG (A and C, from two different AlphaFold predictions) and ExtEF (B) models are shown, with the heme groups represented as red sphere-type sticks. The N- and C-terminal regions of ExtF are identified in the different models. EX, OM and P stand for cell exterior, outer membrane and periplasm, respectively.

The heme organization of ExtF is disrupted in this prediction, since hemes IV and V are isolated in the C-terminal region of the protein, far apart from hemes I-III, at Fe-Fe distances incompatible with

electron transfer (21.5 Å between hemes III and IV). The fact that the N-terminal of ExtF is located on the extracellular side is another indication that the model is unreliable. Upon prediction of the ExtEF complex, without the presence of ExtG, ExtF is located within ExtE, with the N-terminal facing the periplasmic side, as observed in the other porin-cytochrome complexes (Figure 2.18B). In the first prediction of the ExtEFG model (Figure 2.18A), ExtG is wrapped around ExtF, while adopting a structural organization that does not resemble the structure of the cytochrome alone. This affects the overall disposition of the heme groups and results in several clashes between the different cofactors. In general, the structural model of ExtG within ExtEFG is excessively condensed and not realistic (Figure 2.18A). Due to the several issues encountered in the ExtEFG model initially predicted by AlphaFold-Multimer, with low pLDDT scores (Figure 7.5), we sampled a different conformation of the complex. In this prediction (Figure 2.18C), while the average pLDDT scores are still not optimal (Figure 7.5), the N-terminal of ExtF is facing the periplasm and the protein is embedded further in ExtE, similarly to ExtC in the ExtBCD complex. Additionally, the hemes are closely packed and there is an acceptable interface between ExtF and ExtG. Moreover, ExtG shows no internal clashes between the different heme groups and has a more elongated structure. In this prediction, ExtG wraps into a tilted "U" shape, indicating that while maintaining a contact with ExtF, this cytochrome might behave as a flexible cytochrome wire, adopting different conformations during extracellular electron acceptor reduction. These observations indicate that ExtEFG might possess the longest extension towards the extracellular environment (up to 160 Å, if a fully stretched ExtG is considered).

2.3.4. ExtHIJKL

The *extHIJKL* gene cluster is unique in *G. sulfurreducens*, not only due to its number of components, but also due to their intrinsic nature. In a *G. sulfurreducens* strain in which ExtHIJKL was the only porin-cytochrome complex being expressed, the bacterium was able to reduce Mn(IV) oxides at a rate equivalent to that of the WT bacterium, highlighting the role of this complex in the reduction of this extracellular electron acceptor [66]. This complex is also important for the reduction of selenite and tellurite to less-toxic elemental selenium and tellurium [120-122], and for growth on anodes poised at intermediate redox potentials (−150 mV vs SHE) [123].

ExtH

ExtH (GSU2940) is a 45.8 kDa rhodanese-like lipoprotein composed by 468 residues, 29 of which correspond to a lipoprotein signal peptide. Rhodanese proteins are thiosulfate sulfurtransferases found in all three domains of Life, and can be involved in diverse processes, including cyanide detoxification [124], Fe/S clusters formation [125], sulfur metabolism [126], extracellular reduction of metal sulfides [127], as well as intracellular transport [128] and regulatory pathways [129]. Generally, rhodaneses catalyze the transfer of a sulfur atom from thiosulfate to cyanide by cycling between two distinct forms, the free enzyme and a covalent enzyme-sulfur intermediate, characterized by a persulfide bond (R-S-SH) at the sulfhydryl group of the catalytic cysteine residue [130]. These enzymes act as carriers for reactive sulfur atoms by forming persulfide intermediates,

playing a crucial role in sulfur traffic by delivering sulfur as an unharmful chemical species to biosynthetic pathways.

Structurally, rhodanese-like domains have an extensive variability, observable at different levels, including sequence, active site loop length, presence or absence of a critical catalytic cysteine residue, and overall domain arrangement [130]. The tertiary structure of these proteins is usually composed by two domains with very similar folds, displaying an α/β topology, with a central parallel five-stranded β-sheet surrounded by α-helices [131]. These domains can be found in combination with other protein domains, as single or tandem scaffolds. Depending on their folding and overall structural organization, rhodanese-like proteins might be classified into four different groups [130]: i) single domain proteins, ii) tandem-domain proteins, iii) multidomain proteins, and iv) elongated active-site loop proteins.

The AlphaFold model of ExtH shows that the protein contains two tandem domains with an α/β topology and a cysteine residue located in an active-site loop with the amino acid sequence CRAGNA (Figure 2.19), placing it into the group II of the rhodanese-superfamily [130].

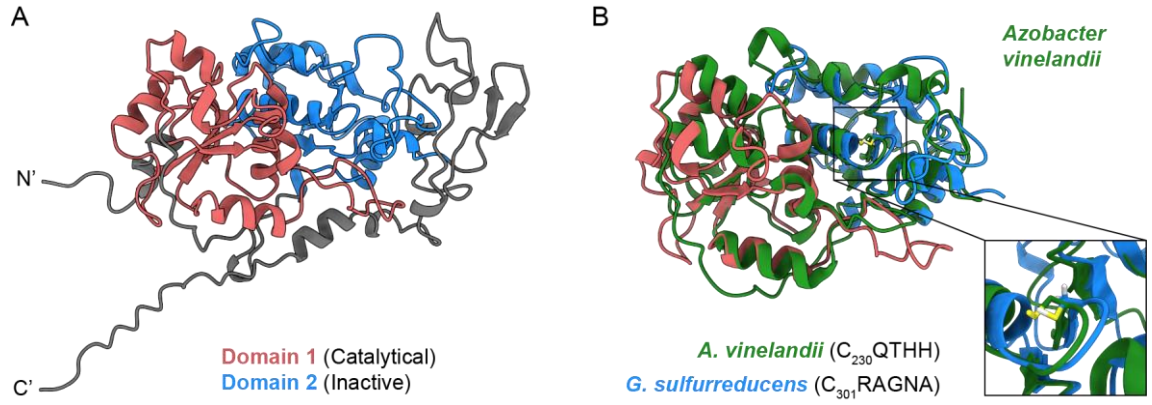


Figure 2.19. ExtH is a group II rhodanese-like protein. (A) AlphaFold model of ExtH. The structure of ExtH is represented as ribbon, with Domains 1 and 2 colored in red and blue, respectively. The remaining of the structure is colored gray. (B) Structural conservation of the rhodanese-like fold of ExtH. The structure of ExtH is superposed with that of the rhodanese of *Azobacter vinelandii* (PDB: 1E0C [131]), which is represented in green. For simplification, the unconserved and disordered N- and C-terminal stretches of ExtH are not represented. The inset highlights the active-sites of the two rhodaneses, with the catalytical cysteine residues represented as sticks.

Upon a search for homologs in different databases, multiple structures of rhodanese-like proteins emerge as closely related structural homologs. A superposition of ExtH with a rhodanese from *Azobacter vinelandii* [131] shows that the overall folding of the proteins is highly similar, including their active-sites (Figure 2.19). While the rhodanese from *A. vinelandii* only contains one cysteine residue, which is located in the active-site of the catalytic domain, ExtH contains an additional cysteine in the supposedly inactive domain. This can be an indication that ExtH actually contains two catalytical domains, however, rhodaneses with inactive domains containing loops with cysteine residues have been described before [132]. Finally, the main difference between these structures is

the additional C-terminal extension that ExtH possesses, which might serve for anchoring the protein either on the ExtHIJKL complex or directly in the outer membrane of *G. sulfurreducens*.

ExtI

ExtI (GSU2939) is composed by 406 residues, including a 26-residue signal peptide, accounting for a total molecular weight of 42.6 kDa. DeepTMHMM [84] predicts that the protein contains 16 transmembrane motifs. This is consistent with the AlphaFold model of ExtI, which shows that it assembles into a typical β-barrel porin structure (Figure 2.20).

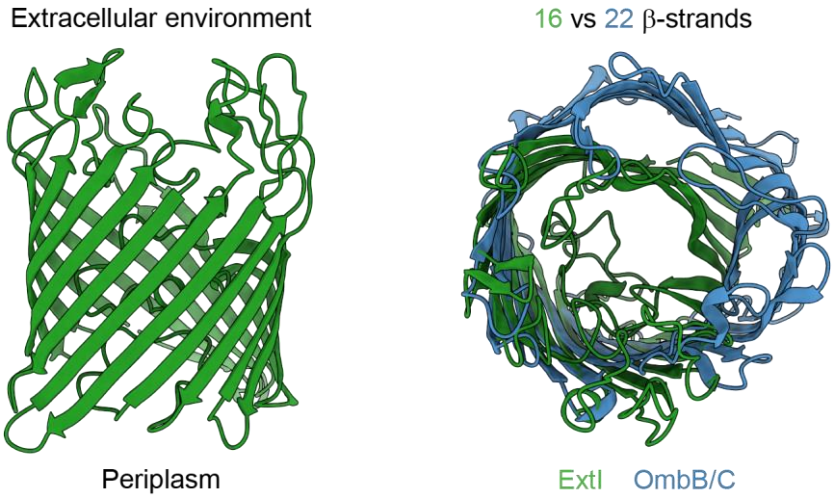


Figure 2.20. Structural features of the ExtI β-barrel porin. The left panel shows the AlphaFold model of the ExtI protein, represented as a green cartoon. The right panel shows the superposed AlphaFold models of the ExtI and OmbB/C proteins, represented as green and blue cartoons, respectively.

While OmbB/C, ExtB and ExtE form similar β-barrel porins with 22 transmembrane motifs, ExtI only contains 16 β-strands, despite possessing approximately the same number of residues as the porins of the other four porin-cytochrome complexes of *G. sulfurreducens*. Consequently, although the same distribution of polar and non-polar residues is observed, the pore diameter size is significantly smaller (27 Å). The residues which are not involved in the β-sheet formation either protrude from the extracellular side of the porin envelope or are clustered within the porin channel (Figure 2.20).

ExtJ

ExtJ (GSU2938) is an 8.2 kDa periplasmic protein composed by 104 residues, 23 of which correspond to a signal peptide. ExtJ has no predicted cofactors and unlike ExtH, it has not been associated with any particular family of proteins. The AlphaFold model of ExtJ shows that the protein contains 6 β-sheets within a β-barrel fold with a Greek-key topology, one isolated α-helix and an elongated random coil C-terminal (Figure 2.21).

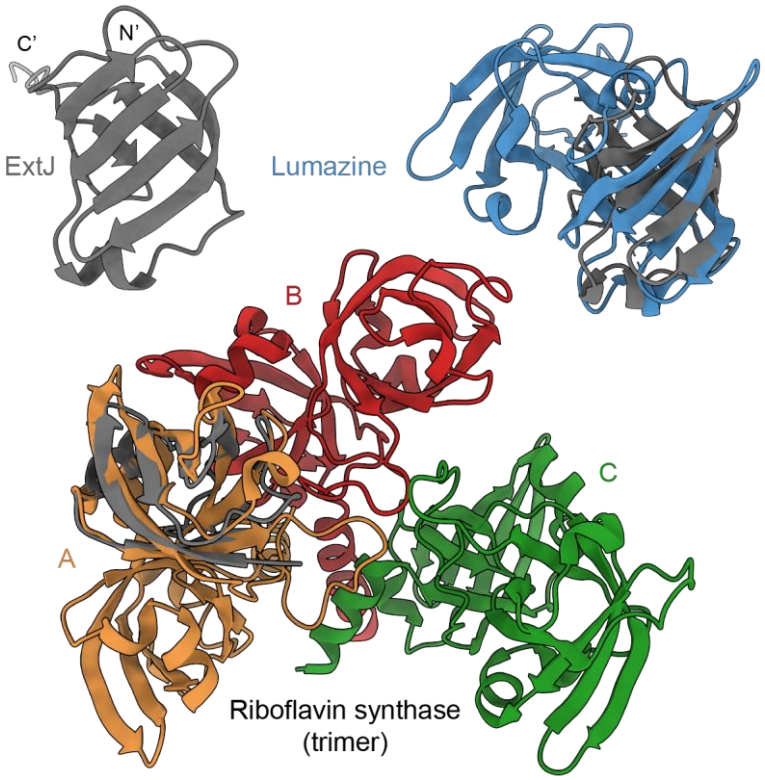


Figure 2.21. Structural homology of ExtJ with riboflavin synthase-like β-barrel proteins. The AlphaFold model of ExtJ is shown as a gray cartoon. The N- and C-terminal regions of the protein are identified. The structures of the lumazine protein from *Photobacterium leiognathi* (PDB: 3DDY [133]) and the riboflavin synthase from *Brucella abortus* (PDB: 4GQN [134]) are represented as cartoons, superposed with ExtJ. The lumazine protein is colored blue, while the three monomers of the riboflavin synthase functionally active trimer are colored orange, red and green. The disordered C-terminal stretch of ExtJ was removed from the superposed structures, for simplification.

Upon a structural search for homologs, multiple structures of domains within subunits of large ATPases/synthases or ribosomal assemblies, together with riboflavin synthase-like β-barrel proteins, emerge as closely related structural homologs of ExtJ. Although the structural similarity between ExtJ and specific domains within subunits of ATPases/synthases or ribosomal assemblies is interesting, based on the small size of ExtJ and its predicted location in the periplasm, the protein is unlikely to be associated with ATP synthesis or ribosomal assembly. Riboflavin synthase-like β-barrel proteins include the riboflavin synthase and lumazine enzymes, which might be conceptually more realistic functional homologs of ExtJ. The superposition of their structures with ExtJ shows a clear structural homology (Figure 2.21), while raising questions about the possible mechanisms of ExtJ, considering that it does not contain the tandem-domain configuration found on those enzymes. Additionally, riboflavin synthase is usually active as a trimer [135], which would require ExtJ to undergo oligomerization. However, trimer formation in riboflavin synthases is usually promoted by a C-terminal α-helix [136-138], which is lacking in ExtJ, meaning it would have to oligomerize through a different mechanism. Furthermore, another relevant question that raises relates to the relationship between the putative function of ExtJ and the assembly of the ExtHIJKL porin-cytochrome complex.

ExtKL

ExtK (GSU2937) is a 35.9 kDa pentaheme *c*-type cytochrome composed by 324 residues, 25 of which correspond to a signal peptide that localizes the protein in the periplasm. ExtL (GSU2936) is a hypothetical protein composed by 98 residues, with no predicted signal peptide, and a total molecular mass of 11.2 kDa. The AlphaFold models of these proteins show that ExtK contains a globular domain formed mostly by α-helices, and two long β-sheets protruding from the remaining of the protein, while ExtL contains a rather unique structure, with only two pairs of antiparallel β-strands (Figure 2.22).

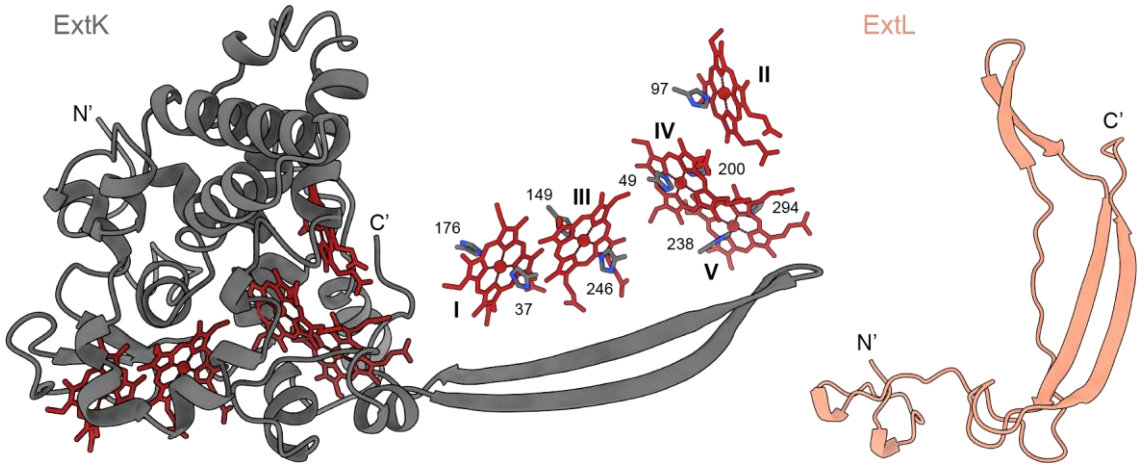


Figure 2.22. Structural features of the cytochrome ExtK and the hypothetical protein ExtL. The AlphaFold models of ExtK and ExtL are represented as ribbon, in gray and orange, respectively. The heme groups are represented as red sticks. The N- and C-terminal regions of both proteins are identified. The heme core of ExtK shows the axial histidines of each heme represented as gray sticks with heteroatom coding. The residue numbering was performed according to the sequence of the mature protein (with no signal peptide). The bold roman numerals indicate the hemes in their order of attachment to the respective heme binding motifs in the polypeptide chain.

The five heme groups of ExtK are located within one half of the globular domain. The heme core is formed by four parallel hemes (hemes I, III-V), and an additional heme (heme II) that is perpendicular to this group. The four hemes (hemes I, III-V) are bis-histidynil coordinated, whereas heme II has no candidates for the distal position, meaning that this position might be vacant and that this heme is likely pentacoordinated. While most of the axial histidine residues are located in the globular domain of ExtK, the distal histidine of heme V (His^{294}) is provided by the C-terminal domain. The hemes are arranged such that the Fe-Fe distances between adjacent hemes are below 15 Å to promote efficient electron transfer [73].

Considering that ExtL is encoded by the *gsu2936* gene within the *extHIJKL* gene cluster, it is rather puzzling that it has no predicted signal peptide, which would lead to the protein being located in the cytoplasm of the bacterium. Furthermore, it has been suggested that ExtK and ExtL may form a single protein, since the hypothetical TGA stop codon that separates their genes might actually encode a

rare selenocysteine amino acid [66, 122, 139]. These evidences, together with the unique AlphaFold model of ExtL, prompted us to make a prediction of the AlphaFold model of the joint ExtKL protein. The resulting model shows that ExtL merges smoothly with the C-terminal β-sheet of ExtK (Figure 2.23).

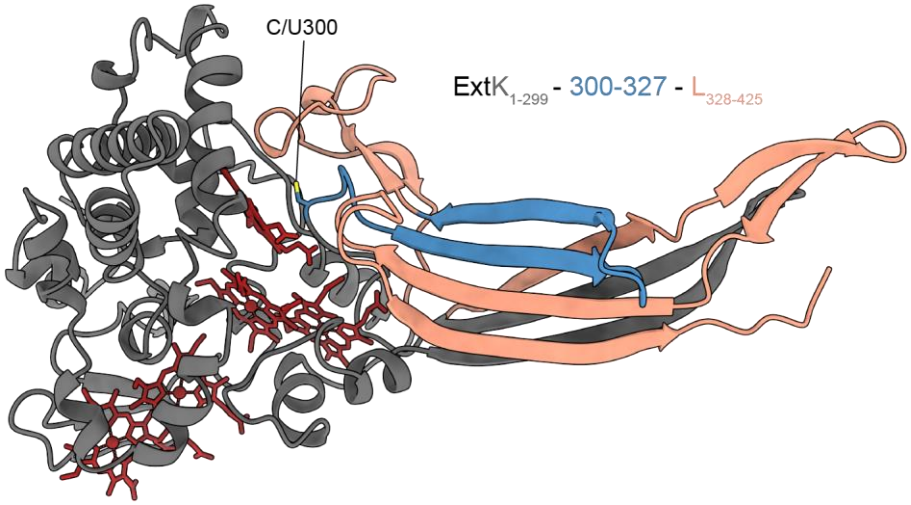


Figure 2.23. Structure of the ExtKL tandem cytochrome. The AlphaFold model of the ExtKL protein is represented as ribbon. The ExtK and ExtL proteins are colored in gray and orange, respectively. The additional 28 amino acids corresponding to the genome section that links both genes are colored blue. The heme groups are represented as red sticks. The cysteine residue used alternatively in the position of the selenocysteine is represented as sticks with heteroatom coding.

The resulting ExtKL cytochrome contains 425 amino acids and a total molecular mass of 50.2 kDa, which accounts for the polypeptides encoded by the *extK* and *extL* genes, the five *c*-type heme groups, and an additional 28 amino acids corresponding to the genome section that links both genes (Figure 2.23). Due to the inability of AlphaFold to make predictions using rare amino acids, such as selenocysteine, the model of ExtKL was alternatively estimated using a cysteine residue in the position of U^{300} (Figure 2.23).

The C-terminal domain of ExtKL, formed by the joint β-strands of ExtK and ExtL, contains two sets of parallel β-strands connected head to rear, parallel to each other, forming two layers in a sandwich mode. In addition, it contains an isolated β-strand that is isolated from the β-sandwich. To gather insights not only on this domain, but on the overall folding of ExtKL, a search for homologs was employed. Surprisingly, a remarkably similar periplasmic pentaheme cytochrome from the Gram-negative *Thermochromatium tepidum* bacterium emerged from this search (Figure 2.24).

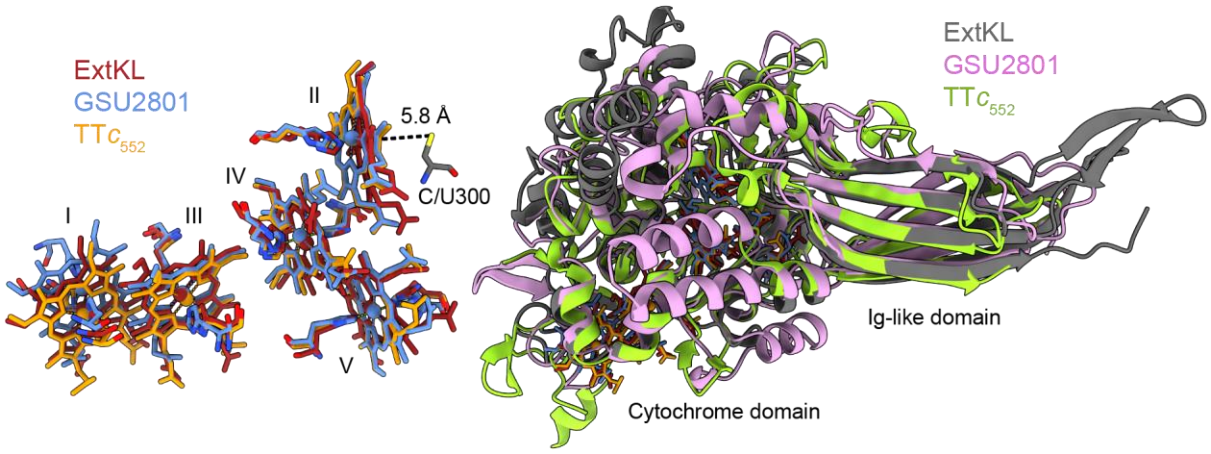


Figure 2.24. Structural homologs of ExtKL. The AlphaFold models of ExtKL and GSU2801, together with the crystal structure of the cytochrome c_{552} from *T. tepidum* (TTc_{552}, PDB: 5ZE8 [140]), are shown as ribbon, in gray, purple and light green, respectively. Their heme cores and axial ligands are represented in red, blue and orange, respectively.

Although this cytochrome is not predicted to contain a selenocysteine residue, the overall fold and heme arrangement is homologous to that of ExtKL, which further supports the tandem arrangement of ExtK-L as a single protein. Similar to ExtKL, in cytochrome c_{552} from *T. tepidum* (TTc$_{552}$), hemes I, III-V are bis-histidynil coordinated, whereas heme II lacks a distal ligand. Interestingly, the cysteine residue used to emulate the U^{300} in ExtKL, is located at a 5-6 Å distance of heme II, suggesting a synergistic function between the selenocysteine residue and this high-spin heme. Chen and co-workers suggested that TTc$_{552}$ might participate in sulfite metabolism, not only due to its conserved heme arrangement compared to a sulfite reductase from *W. succinogenes*, but also because of the presence of a cysteine residue (C^{241}) near heme II [140]. These observations correlate well with the fact that other protein from the ExtHIJKL complex, ExtH, is also likely involved in sulfur metabolism.

Regarding the C-terminal domain formed by β-strands, it can be classified has an immunoglobulin-like domain based on its structural homology. In fact, Ig-like domains consist of a 2-layer sandwich of 7-9 antiparallel β-strands, arranged in two β-sheets with a Greek key topology [141, 142], and are known to be involved in protein-protein and protein-ligand interactions [143], or to act as catalysts of oligomerization [144]. Therefore, this domain might be important for promoting interactions with putative partners of ExtKL or, alternatively, be relevant for the putative oligomerization of this cytochrome.

The genome of *G. sulfurreducens* also contains a structural homolog of ExtKL, GSU2801, which despite only sharing 25% of sequence identity, has an AlphaFold model that aligns with that of ExtKL with a RMSD of 1.1 Å over 111 atom pairs (Figure 2.24). While this local RMSD is a testament to the protein's homology, their structures show different features in the N-terminal cytochrome domain. Moreover, GSU2801 is slightly bigger than ExtKL (462 vs 425 residues) and it does not contain a selenocysteine or corresponding cysteine residue near the U/C^{300} position, but the heme cores of the two proteins are conserved, with a pentacoordinated heme II (Figure 2.24).

It is likely that these two pentaheme cytochromes originated from a gene duplication event, with ExtKL either gaining the selenocysteine residue or GSU2801 losing it over time.

ExtHIJKL does not assemble into a typical porin-cytochrome complex

The AlphaFold models of the five proteins encoded by the *extHIJKL* gene cluster strongly suggest that only four proteins are actually expressed by this cluster. Consequently, the AlphaFold-Multimer simulations of the complex were performed considering that ExtK and ExtL assemble into a singular protein. The resulting AlphaFold-Multimer model of the ExtHIJKL complex is shown on Figure 2.25.

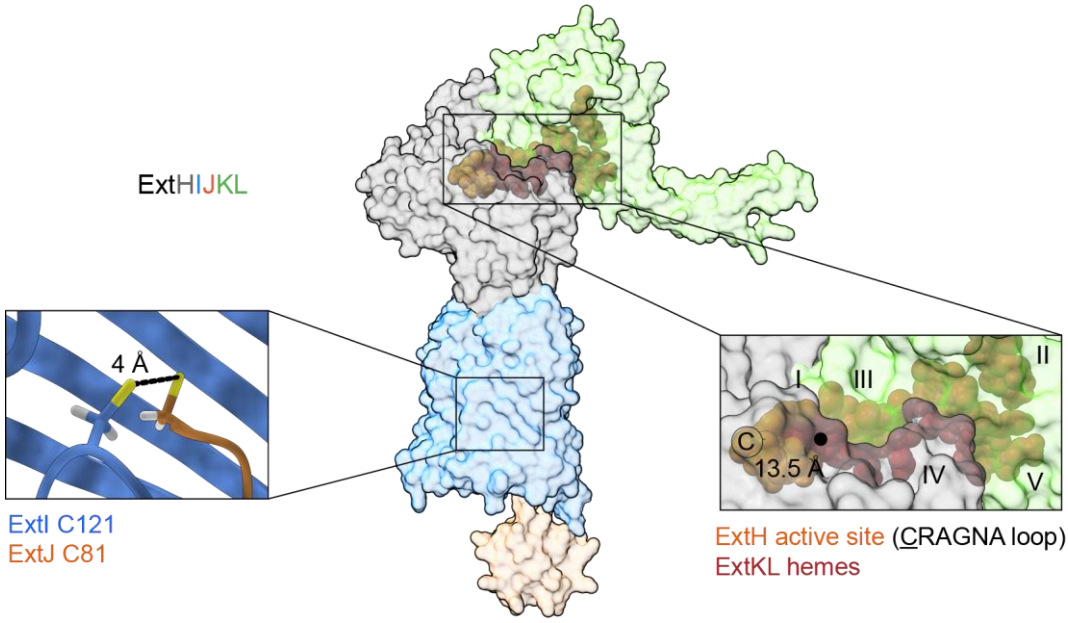


Figure 2.25. AlphaFold-Multimer model of the ExtHIJKL complex. The surface of each protein is shown, with the heme groups of ExtKL and the active site of ExtH represented as red and orange sphere-type sticks, respectively. The left inset shows the putative establishment of a disulfide bridge between ExtI and ExtJ. The right inset shows the interacting surface of ExtH and ExtKL. The roman numerals indicate the hemes in their order of attachment to the respective heme binding motifs in the polypeptide chain. The C indicates the putative catalytical cysteine of the CRAGNA active loop of ExtH.

The resulting assembly of the four-protein complex, with a total of 1300 residues, raises several questions. Using ExtI as a spatial reference, which has been experimentally shown to be located in the outer membrane of *G. sulfurreducens* [122], AlphaFold-Multimer predicts that both ExtH and ExtKL are on the extracellular side, whereas ExtJ is interacting with ExtI in the periplasmic space. While the structures of all four proteins are similar to their respective individual models, their relative position and interacting surfaces do not match subcellular location predictions performed with LocTree [145]. These predictions indicate that ExtJ and ExtKL are periplasmic proteins, whereas ExtH is predicted to be an extracellular lipoprotein. A genomic scan across 18 different *Desulfuromonadales* strains showed that in *extHIJKL* gene clusters, more than 50% of the strains contain a different extracellular protein, each with less than 40% identity to ExtH [66]. This remarkable variability is more likely to occur in extracellular components than in periplasmic redox proteins, due to the

typically higher rates of gene transfer and replacement of domains in proteins which are exposed to electron acceptors and the external environment, thus supporting the view that ExtH is extracellular. Furthermore, other studies have shown that ExtI plays an important role in selenite and tellurite reduction, not only by participating in the uptake of these compounds, but also by affecting the localization of ExtH in the cell [121, 122]. These results are strong indicators that ExtI and ExtH interact in vivo, assembling into a protein complex in which ExtH is facing the extracellular environment. These observations might explain why ExtI contains such a unique folding when compared with the β-barrel porins of the other porin-cytochrome complexes of *G. sulfurreducens*.

Despite the predicted subcellular localization of ExtKL, the presumed interaction between this cytochrome and ExtH is sensical, since both proteins have structural and functional features that align with the putative reduction of compounds such as selenite and tellurite [121, 122]. In fact, AlphaFold consistently places ExtH close to ExtKL in different AlphaFold-Multimer predictions. Nevertheless, this would require a sophisticated mechanism, in which both ExtH and ExtKL are secreted towards the extracellular environment, while maintaining the integrity of the complex and assuring an electron flow consistent with the complex's ability to reduce Mn (IV) oxides [66]. Curiously, either in the ExtHIJKL or ExtHKL models, while the putative selenocysteine residue of ExtKL and the active site of ExtH are not close enough to strongly support a cooperative mechanism, heme I of ExtKL is found 13.5 Å apart from the active site of ExtH (Figure 2.25).

Finally, the AlphaFold-Multimer model indicates that ExtJ interacts with ExtI through its disordered C-terminal tail, likely by establishing a disulfide bridge between its C-terminal cysteine and a cysteine residue from ExtI (Figures 2.25 and 7.6). Further insights into the putative role of ExtJ in the ExtHIJKL complex are discussed in Chapter 3. Overall, and as predicted by its unique constituting proteins, the ExtHIJKL porin-cytochrome complex is unique in its organization and might very likely be the most versatile in *G. sulfurreducens*.

2.4. Using structural models to optimize protein production

The structural and functional information retrieved from the analysis of the AlphaFold models of the proteins that are part of the porin-cytochrome complexes of *G. sulfurreducens* can be used to envision the most efficient approach for the cloning and heterologous expression of these proteins. Since AlphaFold emerged halfway through the development of this Thesis, the next sections will showcase not only the protein constructs that were designed prior to its emergence, but also those which were rationally designed using the information provided by this computational tool. The constructs containing the proteins of the OmabcB/C porin-cytochrome complexes, ExtC and ExtF were already available in the laboratory.

2.4.1. "Blind" protein-cloning

Prior to the emergence of AlphaFold, we used the information provided by bioinformatic tools that performed sequence analysis, such as SignalP [146], DeepTMHMM [84], LocTree [145] and ProtParam [147], to infer which parts of the genes of the proteins of interest should be cloned in order to obtain

the simplest version of the protein, while maintaining function and solubility, and also to envision purification strategies, based on the predicted stability, hydrophobicity, molecular mass and global net charge. The regions initially selected for protein cloning for each of the target proteins are shown in Table 2.1.

Table 2.1. List of residues of the target proteins to clone. The * indicates the proteins whose constructs were already available in the laboratory. All *c*-type cytochromes were cloned into a pVA203 vector [148], a pUC-derivative containing the lac promoter and the OmpA signal peptide sequence from *E. coli* [149], while the remaining proteins were cloned into a pET-28a(+) vector (Novagen), with an N-terminal His-tag and a thrombin cleavage site.

Target protein	Residues to clone	Target protein	Residues to clone
ExtA (GSU2645)	32-310	ExtG (GSU2724)	31-691
ExtB (GSU2644)	27-420	ExtH (GSU2940)	30-468
ExtC (GSU2643)*	25-155	ExtI (GSU2939)	27-406
ExtD (GSU2642)	25-229	ExtJ (GSU2938)	24-104
ExtE (GSU2726)	22-431	ExtK (GSU2937)	26-324
ExtF (GSU2725)*	25-157	ExtL (GSU2936)	1-98

With this information, we used restriction-free (RF) cloning [150, 151] to prepare custom plasmids containing our proteins of interest. This polymerase chain reaction (PCR) technique provides a simple, universal method to precisely insert a DNA fragment into any desired location within a circular plasmid, independent of restriction sites, ligation, or alterations in either the vector or the gene of interest. RF cloning encompasses two PCR reactions. In the first reaction, the gene of interest is amplified from genomic DNA, resulting in a PCR fragment that contains both the gene of interest and regions complementary to a target plasmid. This fragment is commonly referred to as megaprimer, and is inserted into the target plasmid upon a second PCR reaction. The detailed PCR protocols and the list of primers used for each case can be found in section 2.6.2.

In case of an unsuccessful PCR reaction, there are several variables that may be tested, namely the reaction buffer, the amount of primer and/or bacterial genome, as well as the addition of dimethyl sulfoxide (DMSO).

The Phusion High-Fidelity DNA polymerase is provided with two buffers, termed HF and GC. HF buffer should be used as the default for high-fidelity amplification, as the error rate of Phusion DNA polymerase is lower in these conditions. However, GC buffer can improve the performance of DNA polymerases on certain templates, such as GC-rich DNA or DNA with complex secondary structures. This rational is valid for any PCR reaction and is a good starting point for the optimization of the process. Additionally, one may use DMSO to overcome these issues. DMSO is an organosulfur compound with a high polarity and dielectric constant, that binds to the major and minor grooves of template DNA trough hydrogen bonds, thus destabilizing the double helix structure [152]. Similarly, DMSO is also efficient in the disruption of putative primer secondary structures, thus leading to

lower annealing temperatures [153]. Commonly, for a typical PCR experiment, 3-10% DMSO can be added to the PCR mix, depending on the GC content, reaction type and quality of DMSO, but one should optimize these percentages by varying them in 2% increments. Another important point relates with the variation of the relative amount of primer and template DNA, which will influence the number of binding events between these components and affect the efficiency of the reaction. On one hand, higher primer concentrations have shown to increase unspecific priming when applied to complex mixtures, such as genomic DNA, while increasing the probability of primer dimer formation [154]. On the other hand, lower primer concentrations might attenuate amplification rates on the later cycles of the PCR reaction due to inefficient annealing [155]. Both variables were varied across the trials performed during cloning of the different protein targets.

In the second PCR reaction, the resulting megaprimer is inserted into the target plasmid. As for the first PCR reaction, the nature of the reaction buffer and the addition of DMSO are good starting points for the optimization of an unsuccessful reaction. Nevertheless, in this step of the RF cloning protocol, the most common parameters that are optimized are the applied annealing temperatures and extension times. These parameters can also be optimized in the first reaction, although this is not so common. Alternatively, one can resort to touchdown PCR, a method that is typically used to decrease off-target priming and increase the specificity of the amplification reaction [156]. The optimized detailed PCR protocols for each vector obtained using RF cloning can be found in section 2.6.2. All proteins were successfully cloned using RF cloning, except for ExtH, whose plasmid was later acquired from GeneCust.

Prior to the release of AlphaFold, an N-terminal Strep-tag with a Tobacco Etch Virus (TEV) cleavage site and appropriate flexible linkers [157-159], were inserted in each of the cytochrome targets (ExtA, C, D, F, G and K). The tags were inserted into the target plasmids using a strategy based on the Q5 Site-Directed Mutagenesis kit from New England Biolabs. This kit is designed for rapid and efficient incorporation of insertions, substitutions and deletions into double stranded plasmid DNA. The first step of the protocol is a simple exponential amplification of the target plasmid using specifically designed primers and a master mix formulation. The second step involves incubation with a unique enzyme mix containing a polynucleotide kinase [160], a DNA ligase [161] and DpnI [162]. Together, these enzymes warrant rapid circularization of the PCR product and removal of the template DNA. The final step involves a transformation in competent cells for final plasmid maturation, as for any other cloning process. The detailed protocols and the list of primers used for each case can be found in section 2.6.3.

2.4.2. AlphaFold-guided protein cloning

The structural information provided by AlphaFold prompted us to redesign some of the target protein constructs. Considering that ExtA contains an N-terminal tail with no predicted heme axial ligands and that the cytochrome is arranged in three identical tetraheme domains, four new constructs were designed using either the Q5 Site-Directed Mutagenesis strategy or RF cloning (Table 2.2).

Table 2.2. List of AlphaFold-guided constructs. "MM" stands for molecular mass, which was calculated from the total number residues resorting to Protparam [147] plus 616 Da for each heme group [163]. "ST" and "HT" stand for Strep-tag and His-tag, respectively. "TEV" stands for TEV cleavage site. The 1-28 stretch of the ExtKL constructs corresponds to the genome section that links both genes (GSU2937 and GSU2936). The ExtB-ExtC construct was purchased from GeneCust in the form of a pETDuet-I plasmid (Novagen) [164].

Construct	Residues cloned	MM (kDa)	Heme groups
ST-TEV-ExtA$_{73-210}$	73-210	33.8	12
ExtA Domain 1	73-154	11.8	4
ExtA Domain 2	157-229	10.4	4
ExtA Domain 3	236-310	10.8	4
ExtB-ExtC	27-420 + 25-155	69.4	5
ST-TEV-ExtK$_{26-285+315-324}$	26-285 + 315-324	35.1	5
ST-TEV-ExtK$_{26-285+P+315-324}$	26-285 + 315-324	35.2	5
ST-TEV-ExtK$_{26-285+PP+315-324}$	26-285 + 315-324	35.3	5
ExtKL	26-324 + 1-28 + 1-98	50.2	5
ExtKL U300C	26-324 + 1-28 + 1-98	50.2	5
ExtKL U300C - HT	26-324 + 1-28 + 1-98	51.1	5

Additionally, several constructs were designed for ExtK and ExtKL (Table 2.2). Initially, we designed three constructs of ExtK without the β-strands protruding from the globular heme-containing domain, while maintaining the C-terminal stretch that provides the axial ligand for heme V (Figure 2.22). The only difference between the three constructs is either the absence or presence of different proline linkers connecting the N- and C-terminal stretches of the protein (Table 2.2). Additionally, the ExtKL constructs were designed to contain the full-length protein, as AlphaFold provided strong evidence that the ExtK-ExtL tandem is a single protein (Figure 2.23).

Selenocysteine residues are cotranslationally inserted into nascent polypeptide chains in response to the UGA codon, whose normal function is to terminate translation [165]. To decode UGA as a selenocysteine residue, organisms developed insertion machineries that allow the incorporation of this uncommon amino acid at specific UGA codons, through a process that requires a cis-acting selenocysteine insertion sequence (SECIS) element [166]. Specifically, in *E. coli*, this mechanism involves (i) a release factor 2 (RF2) protein, (ii) a selenocysteine-dedicated elongation factor, (iii) unique tRNA species acylated with a serine residue, which is latter converted to a selenocysteine, and (iv) several accessory factors, such as SelB, and a SECIS element with a UGA-N_{11}-SelBbr sequence, in which N_{11} represents 11 nucleotides of random nature and SelBbr represents the SelB binding region (GGUUGCAGGUCUGCACC) [167].

Based on this information, and considering that the *extKL* gene does not possess the required sequence for selenocysteine incorporation, this selenocysteine residue was mutated to a cysteine (U300C) in an additional construct (Table 2.2). While the U300C construct will allow expression of

the protein in *E. coli* without additional requirements, there are specific protocols for selenocysteine incorporation in this microorganism that might be implemented in the future [168]. To facilitate the purification of the mutated cytochrome, a His-tag was inserted in the C-terminal of the protein.

Finally, considering that ExtC is embedded within the ExtB porin, and to test for the validity of an unconventional approach to obtain a partial porin-cytochrome ExtBC complex, we acquired a pETDuet-1-ExtB-ExtC vector from GeneCust. These vectors simplify dual expression by allowing cloning of two genes in the same plasmid. These plasmids contain two multiple cloning sites (MCS), each preceded by a T7 promoter, a *lac* operon and a ribosome binding site. The first MCS contains ExtC with a OmpA signal peptide [149] and a N-terminal Strep-tag followed by a TEV cleavage site, whereas the second MCS contains ExtB with a pelB signal peptide [169]. The localization of cytochromes in the periplasm during protein overexpression in *E. coli*, ensured by the presence of the OmpA and pelB signal peptides, is essential for cytochrome maturation, since this process is performed after the separate transport of apoproteins and heme groups into the periplasmic space [170].

2.4.3. Preliminary expression tests of target proteins

Following the cloning of the target proteins, we performed expression tests on some of the cytochrome constructs, including the ExtB-C tandem, as well as of ExtJ, since the β-barrel porins and ExtH were not our main targets. For that, different strains of competent *E. coli* cells were transformed with the target cytochrome-containing pVA203 plasmids. Since *E. coli* is unable to synthesize cytochromes *c* under aerobic growth conditions [171, 172], all these strains contained the pEC86 plasmid encoding the *c*-type cytochrome maturation gene cluster and a chloramphenicol resistance marker [163]. In the case of ExtJ and the ExtB-C tandem, since pET vectors require the presence of a lysogenized DE3 phage fragment encoding the T7 RNA polymerase [164, 173, 174], only the DE3-containing *E. coli* strains were transformed with these plasmids. After bacterial transformation, different protein expression variables were tested, namely the temperature, shaking speeds, and incubation times in the absence and presence of different concentrations of isopropyl β-D-thiogalactoside (IPTG), the protein expression inductor. The detailed conditions which were tested are summarized in section 2.6.5.

Overall, the expression tests performed with the cytochrome targets showed that high concentrations of IPTG decrease the amount of mature protein, likely because cells are unable to incorporate hemes at a rate equivalent to the production of apoprotein. Similarly, expression conditions in which cytochrome production was more efficient, often lead to incorporation in either inclusion bodies or spheroplasts. This effect was slightly mitigated with lower shaking speeds, under growth at lower temperatures, or by not inducing protein expression, relying solely on the intrinsic basal expression levels of the pVA203 plasmids.

In general, finding the optimal expression conditions of cytochromes *c* in a soluble form in *E. coli* requires a fine tuning of several parameters. The best expression conditions determined for each of the cytochrome targets is summarized in Table 2.3.

Table 2.3. Optimized expression conditions for the cytochrome targets and ExtJ. The "*" indicate the cases in which constructs of cytochromes containing a N-terminal Strep-tag with a TEV cleavage site were used. "OD" refers to the approximate optical density at 600 nm. The indicated shaking speeds were applied either after addition of IPTG, or during the entire bacterial growth, in cases in which IPTG was not added. In those cases, the OD values represent the stage at which the cells were harvested.

Protein construct	Strain	Condition
ExtA*	BL21 (DE3)	OD 1.5, 10 μM IPTG, 30 °C, overnight, 160 RPM
ExtA D1	BL21 (DE3)	OD 1.5, 10 μM IPTG, 30 °C, overnight, 160 RPM
ExtA D2	BL21 (DE3)	OD 1.5, 10 μM IPTG, 30 °C, overnight, 160 RPM
ExtA D3	BL21 (DE3)	OD 1.5, 10 μM IPTG, 30 °C, overnight, 160 RPM
ExtB-ExtC	Tuner (DE3)	OD 0.8, 50 μM IPTG, 25 °C, overnight, 160 RPM
ExtC*	C41 (DE3)	OD 1.5, 50 μM IPTG, 30 °C, overnight, 160 RPM
ExtD*	C41 (DE3)	30 °C, overnight, 160 RPM
ExtF*	C43 (DE3)	OD 1.5, 50 μM IPTG, 30 °C, overnight, 160 RPM
ExtG*	Tuner (DE3)	30 °C, overnight, 160 RPM
ExtJ	BL21 Star (DE3)	OD 0.8, 1 mM IPTG, 37 °C, 4 h, 180 RPM
ExtK*	BL21 (DE3)	OD 1.5, 20 μM IPTG, 30 °C, overnight, 160 RPM
ExtK$_{26\text{-}285+315\text{-}324}$*	Tuner (DE3)	OD 1.8, 30 °C
ExtKL U300C	Tuner (DE3)	OD 1.8, 30 °C

The summarized expression conditions result in soluble protein in the cases of ExtA and its tetraheme domains, ExtD, ExtJ and ExtK(L) variants. Initial bioinformatic predictions did not provide indications that ExtC and ExtF were insoluble, which led to several efforts in order to obtain these proteins in their soluble form. However, AlphaFold clearly shows that both ExtC and ExtF are likely insoluble, since they dock within their respective β-barrel porins. This is another example of how AlphaFold can provide important information and prevent frustrating attempts during the production of target proteins. The overexpression of ExtB-C using a pETDuet-1 vector was unsuccessful, likely because induction of the T7 promoter, even at low concentrations of IPTG, disrupts the production of mature ExtC, while producing low amounts of ExtB.

Additionally, ExtG is likely found in an insoluble form because *E. coli* is unable to handle the uncommon heme binding motifs that this cytochrome possesses [119], leading to either apoprotein accumulation or production of aberrant forms of the protein. This implication will either require that ExtG is produced in *S. oneidensis* [175, 176], or that the non-canonical heme binding motifs of ExtG are mutated to shorter forms that can be recognized by the *E. coli c*-type heme maturation machinery. This second approach was successfully applied to overexpress OmcZ in *E. coli*, which contains a $CX_{14}CH$ heme binding motif [115].

2.5. Conclusions

Despite the tremendous impact that AlphaFold is having and will have for a long time in the field of structural biology and drug discovery, as a consequence of its remarkable accuracy in the prediction of structural models for the majority of proteins, there are still limitations and challenges that need to be circumvented in order to have completely reliable methods for structural prediction and to investigate protein-protein and protein-ligand interactions [177].

AlphaFold is unable to sample multiple protein conformations and structural dynamics, which are essential features of protein function [178-180]. Therefore, all structural models predicted by AlphaFold are prone to contain errors [181, 182], meaning that each case must be thoroughly evaluated based on the structural confidence parameters that are provided and on the intrinsic features of the system in study [183]. Therefore, all interpretations made in this chapter must be considered and used carefully to take conclusions on the overall structural organization of the porin-cytochrome complexes of *G. sulfurreducens*. For this reason, the structural models of the individual proteins and protein complexes predicted in this chapter can be found with AlphaFold's pLDDT color code and respective average scores in section 7.1.

Considering that AlphaFold is incapable of modelling any non-proteinaceous ligands or cofactors, the unreliability of the structural models of proteins whose functions are highly dependent on those "exogenous" molecules, such as cytochromes, only increases. While promising approaches for complementation of structural prediction softwares are arising [40, 184, 185], there is still a challenging path to pave not only in this department, but also on the prediction of protein complexes [48, 186, 187]. In this chapter, we were able to go around this limitation by inserting *c*-type hemes in silico, taking advantage of the positions of the cysteine and histidine residues of the heme binding motifs. However, this process limits heme distortion and the orientation of the propionate groups, which are very flexible and typically establish hydrogen bonds with surrounding residues, affecting the local structure [188, 189].

Notwithstanding the mentioned limitations of AlphaFold, the structural models of the proteins of the porin-cytochrome complexes of *G. sulfurreducens* provided valuable information that is supported by structural homology observations. These models indicate that the five porin-cytochrome complexes contain different structural organizations that might explain their different phenotypes. The β-barrel porins OmbB/C, ExtB and ExtE contain 22 transmembrane β-strands and are structural homologous, while ExtI only contains 16 transmembrane motifs. ExtI is known to be involved in selenite and tellurite uptake [121, 122], while the other porins have not been linked to assimilation functions.

The structural models of multiheme cytochromes enriched with *c*-type hemes allowed us to identify conserved heme arrangements, uncommon heme binding motifs, and potential targets for biochemical deconstruction strategies that may accelerate their thermodynamic characterization. Our AlphaFold predictions strongly suggest that ExtK and ExtL form a single protein, structurally homologous to an already characterized bacterial pentaheme cytochrome.

In addition to all these structural and functional insights, AlphaFold was used to optimize the constructs of some of the protein targets, resulting in an extensive library of constructs that lays the foundation for the thorough study of the porin-cytochrome complexes of *G. sulfurreducens* (Table 2.4).

Table 2.4. List of constructs available in the laboratory. "HT" and "ST" stand for His-tag and Strep-tag, respectively. "TEV" and "TR" stand for TEV and thrombin cleavage site, respectively. "$ExtK_N$" refers to the protein without part of the C-terminal domain. The constructs marked with a "*" were already available in the laboratory.

Protein	Plasmid	Composition
ExtA	pVA203	OmpA + 32-310
ST-TEV-ExtA	pVA203	OmpA + ST + TEV + 32-310
ST-TEV-$ExtA_{73-210}$	pVA203	OmpA + ST + TEV + 73-210
ExtA D1	pVA203	OmpA + 73-154
ExtA D2	pVA203	OmpA + 157-229
ExtA D3	pVA203	OmpA + 236-310
ExtB	pET-28a(+)	HT + TR + 27-420
ExtB + ST-TEV ExtC	pETDuet-1	pelB + 27-420 + OmpA + ST + TEV + 25-155
ExtC*	pVA203	OmpA + 25-155
HT-TEV-ExtC*	pVA203	OmpA + HT + TEV + 25-155
ST-TEV-ExtC	pVA203	OmpA + ST + TEV + 25-155
ExtD	pVA203	OmpA + 25-229
ST-TEV-ExtD	pVA203	OmpA + ST + TEV + 25-229
ExtE	pET-28a(+)	HT + TR + 22-431
ExtF*	pVA203	OmpA + 25-157
ST-TEV-ExtF	pVA203	OmpA + ST + TEV + 25-157
ExtG	pVA203	OmpA + 31-691
ST-TEV-ExtG	pVA203	OmpA + ST + TEV + 31-691
ExtH	pET-28a(+)	HT + TR + 30-468
ExtI	pET-28a(+)	HT + TR + 27-406
ExtJ	pET-28a(+)	HT + TR + 24-104
ExtJ C81STOP	pET-28a(+)	HT + TR + 24-103
ExtK	pVA203	OmpA + 26-324
ST-TEV-$ExtK_N$	pVA203	OmpA + ST + TEV + 26-285 + 315-324
ST-TEV-$ExtK_{N+P}$	pVA203	OmpA + ST + TEV + 26-285 + P + 315-324
ST-TEV-$ExtK_{N+PP}$	pVA203	OmpA + ST + TEV + 26-285 + PP + 315-324
ExtKL	pVA203	OmpA + 26-324 + 1-28 + 1-98

Table 2.4. ***Continued.***

Protein	Plasmid	Composition
ExtKL U300C	pVA203	OmpA + 26-324 + 1-28 + 1-98
ExtKL U300C - HT	pVA203	OmpA + 26-324 + 1-28 + 1-98 + HT
OmaB/C*	pET-22b(+)	pelB + 24-232 + HT
OmaB/C*	pVA203	OmpA + 24-232
OmaB/C*	pCK32 [190]	OmpA + 24-232
OmbB/C*	pET-28a(+)	HT + TR + 24-403
OmcB*	pVA203	OmpA + 23-744
OmcC*	pVA203	OmpA + 23-768

The expression tests of the cytochrome targets show promising routes for the production of these proteins, despite their apparently low yields, that will likely imply the preparation and growth of considerable amounts of bacterial media. For the purpose of this Thesis, the main protein target of the porin-cytochrome complexes which we focused on was ExtJ, whose structural and functional details are discussed in the next chapter.

Additional cytochrome targets

In search for additional targets that might be involved in the final steps of EET in *G. sulfurreducens*, the AlphaFold model of PgcA, a 511-residue triheme cytochrome that reduces Fe(III) oxides [191], revealed a fuzzy arrangement with three monoheme cytochrome domains linked by unstructured stretches (Figure 2.26).

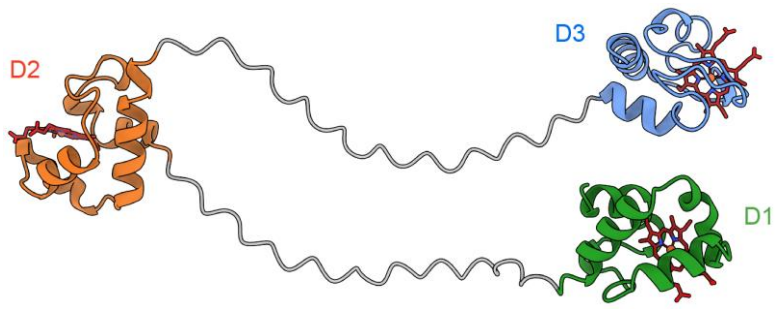


Figure 2.26. AlphaFold model of PgcA. The first (green), second (orange) and third (blue) cytochrome domains of PgcA are represented as ribbons. The proline-threonine disordered stretches are colored gray. Residues 1-260 are omitted, for simplification. The complete model of PgcA may be found in section 4.3.1.

The fact that each cytochrome domain is completely isolated makes this protein an excellent target for the biochemical deconstruction strategy that was applied to GSU1996 [98] and suggested for ExtA. The characterization of the three cytochrome domains of PgcA is discussed in Chapter 4, being another remarkable example of how useful AlphaFold model predictions can be in the study of multiheme cytochromes.

2.6. Materials and methods

2.6.1. Protein targets and putative porin-cytochrome complexes model prediction

The individual AlphaFold models of the target proteins were either retrieved from the AlphaFold protein structure database [21, 46] or produced using the AlphaFold2_MMSeqs2 GitHub ColabFold notebook [52]. All AlphaFold-Multimer models were produced using the same notebook. To sample for additional protein conformations of certain protein targets, the *num_samples* and *is_training* options of the AlphaFold2_Advanced GitHub ColabFold notebook were used. The *num_samples* option enables sampling of diverse structures by iteration through a series of random seeds, whereas *is_training* enables dropout during structure inference, which activates the stochastic part of the model and can result in different predictions [52]. Additionally, a model for ExtD was also produced using the ESMFold GitHub Google Colab notebook [35].

In the case of the cytochrome models, the *c*-type hemes were fitted in PyMol 2.5.2 [192] using local alignments and libraries of *c*-type hemes, containing the corresponding cysteine residues for the covalent binding of the heme vinyl groups and the proximal axial histidines for Fe coordination. These libraries were built by retrieving these residues and the corresponding *c*-type heme group from the structure of the OmcF cytochrome (PDB: 3CU4 [193]). Side-chain packing optimization and all-atom energy minimization protocols were not employed (see section 7.4).

The quality of individual models was evaluated based on the per-residue estimate of confidence produced by AlphaFold, designated pLDDT, which is based on the lDDT-Cα metric [33]. Additionally, intra and inter-predicted alignment error (PAE) scores produced by AlphaFold were used to evaluate the quality of the relative position of proteins in complexes [52]. The interactive analysis of these values was performed using PAE Viewer [194]. The final models were analyzed and represented in UCSF ChimeraX 1.4 [195], without the corresponding signal peptides, as predicted by SignalP 6.0 [146].

2.6.2. RF cloning of the target proteins

The proteins of the ExtABCD, ExtEFG and ExtHIJKL porin-cytochrome complexes, except for ExtC, ExtF and ExtH, were cloned using RF cloning [150, 151]. All *c*-type cytochromes were cloned into a pVA203 vector [148], while the remaining proteins were cloned into a pET-28a(+) vector (Novagen), with an N-terminal His-tag and a thrombin cleavage site. The primers used were designed using the RF cloning website and their sequences are listed in Table 2.5.

Table 2.5. Sequences of the RF cloning primers used to produce the vectors containing the genes of the proteins of interest. The DNA sequences specific for the pVA203 or pET-28a(+) vectors, and for the different protein genes, are highlighted in blue and green, respectively. The primers forward (fw) and reverse (rv) were purchased from Invitrogen.

Primer	DNA sequence (5' – 3')
ExtA_fw	GCTACCGTTGCGGCCGCCAGTGAAAAGAGCGGCGCC
ExtA_rv	CTTGTCGACGGAGCTCGAATTCATCAGCGCTGACGAACCGG
ExtAD1_fw	TTTCGCTACCGTTGCGGCCGCCACGGTCACCCAGTGCGGC
ExtAD1_rv	AGCTTGTCGACGGAGCTCGAATTCACATCCGGACAATGGAGTGC
ExtAD2_fw	TTTCGCTACCGTTGCGGCCGCCACGAAGAAGCTCCAGGACT
ExtAD2_rv	AGCTTGTCGACGGAGCTCGAATTCAGATCTGCTTCGGCTTGTGG
ExtAD3_fw	TTTCGCTACCGTTGCGGCCGCCGGAGCGCGGACCTGCGGT
ExtAD3_rv	AGCTTGTCGACGGAGCTCGAATTCAGCGCTGACGAACCGGCGG
ExtB_fw	CTGGTGCCGCGCGGCAGCGTTGATCTGGGGATCAGTTCC
ExtB_rv	GGTGGTGGTGGTGGTGCTCGAGTTACCTGCGAAACGCATAGG
ExtD_fw	CGCTACCGTTGCGGCCGCCGGCAACGCGTCCTCTCCC
ExtD_rv	GCTTGTCGACGGAGCTCGAATTCAACGAACGATTGTCGGATGACA
ExtE_fw	CTGGTGCCGCGCGGCAGCGCTGACGTGGCGGTGGAT
ExtE_rv	GGTGGTGGTGGTGGTGCTCGAGTCATTTCGCCGCTCCTTTGC
ExtG_fw	CGCTACCGTTGCGGCCGCCATCGACCCGACAACCGG
ExtG_rv	GCTTGTCGACGGAGCTCGAATTCACCACGTCATGCTTCTCGT
ExtH_fw	CTGGTGCCGCGCGGCAGCTGCGGTGGCGGGGGCTAT
ExtH_rv	GGTGGTGGTGGTGGTGCTCGAGCTAGCAGCCACCGCCACC
ExtI_fw	CTGGTGCCGCGCGGCAGCGGTCCCCGGATCACCTTT
ExtI_rv	GGTGGTGGTGGTGGTGCTCGAGCTAGAAAAGGAGCTGGAGCTG
ExtJ_fw	CTGGTGCCGCGCGGCAGCGCCGGCTCCTTCTCGG
ExtJ_rv	GGTGGTGGTGGTGGTGCTCGAGCTAGCAGCCCTGGAGAGC
ExtK_fw	GCTACCGTTGCGGCCGCCGAAAAGGCGGGCATCGGC
ExtK_rv	CTTGTCGACGGAGCTCGAATTCATCAGCCATCGGGGATCGA
ExtKL_fw	GCTACCGTTGCGGCCGCCGAAAAGGCGGGCATCGGC
ExtKL_rv	GGTGGTGGTGGTGGTGCTCGAGCTACTTGCCCTTGGCGCT
ExtL_fw	CTGGTGCCGCGCGGCAGCATGCCGGTGCCCCAGC
ExtL_rv	GGTGGTGGTGGTGGTGCTCGAGCTACTTGCCCTTGGCGCT

The standard PCR mix used in the first PCR reactions of the RF cloning protocol contained a standard reaction buffer (either HF of GC, both from Thermo Fisher Scientific), 0.2 mM dNTPs (Thermo Fisher Scientific), 0.5 μM of each custom-designed primer, 50 ng of *G. sulfurreducens* genomic DNA (previously extracted and available in the laboratory) as template, and one unit of Phusion High-Fidelity DNA polymerase (Thermo Fisher Scientific), for a final volume of 50 μL. The PCR cycling conditions used for the amplification of the genes of interest from bacterial genome are indicated in Table 2.6.

Table 2.6. PCR cycling conditions for the RF cloning of the target proteins. The annealing temperatures for each pair of primers were calculated using the T_M calculator from Thermo Fisher Scientific. The extension times were calculated according to the DNA polymerase's manufacturer indication of 15-30 seconds/kb. These values are indicated in Table 2.7.

PCR stage		Temperature (°C)	Time (s)
Initial denaturation		98	30
Denaturation		98	8
Annealing	**x30**	see Table 2.7	20
Extension		72	see Table 2.7
Final extension		72	600
Final hold		16	–

The size and purity of the megaprimers were analyzed by 1% agarose gel electrophoresis stained with GreenSafe Premium (NZYTech). After confirming that megaprimers contained the correct size, they were purified using the NZYGelpure kit (NZYTech) to remove remaining components of the first PCR reaction, and quantified in a NanoDrop spectrophotometer ND-1000 (Thermo Fisher Scientific).

The standard PCR mix used in the second PCR reaction of the RF cloning protocol contained a standard reaction buffer (either HF of GC, both from Thermo Fisher Scientific), 0.2 mM dNTPs (Thermo Fisher Scientific), 50 ng of target plasmid, 0.4 units of Phusion High-Fidelity DNA polymerase (Thermo Fisher Scientific), and megaprimer in an amount such that there was a 20:1 megaprimer:plasmid ratio in the final 20 μL mixture. The PCR cycling conditions used for the insertion of the megaprimers are identical to those used in the first PCR reaction. However, in order to achieve a successful amplification of the genes of interest and to guarantee their insertion into the target vectors, several variables of the RF cloning protocol described above were specifically optimized for each case. Touchdown PCR was applied as an alternative PCR method, with a stepwise decrease of 1 °C between 80 and 60 °C, across 20 cycles. The detailed conditions used for each target protein are described in Table 2.7.

Table 2.7. Optimization of the RF cloning protocol conditions for each target protein. HF and GC refer to the PCR mix buffer used in each condition. The temperatures and times indicated refer to the annealing temperatures and extension times used for each PCR program, respectively.

Protein	RF cloning (PCR conditions)
ExtA (GSU2645)	1st PCR: HF, 0.5 μM primers, 50 ng template, 68 °C, 20 s 2nd PCR: HF, 50 ng target plasmid, touchdown PCR, 105 s
ExtA D1	1st PCR: HF, 0.5 μM primers, 50 ng template, 64 °C, 10 s 2nd PCR: HF, 50 ng target plasmid, 2-step PCR (72 °C), 120 s
ExtA D2	1st PCR: HF, 0.5 μM primers, 50 ng template, 64 °C, 10 s 2nd PCR: HF, 50 ng target plasmid, 2-step PCR (72 °C), 120 s
ExtA D3	1st PCR: HF, 0.5 μM primers, 50 ng template, 72 °C, 10 s 2nd PCR: HF, 50 ng target plasmid, 2-step PCR (72 °C), 120 s
ExtB (GSU2644)	1st PCR: HF, 0.5 μM primers, 50 ng template, 65 °C, 25 s 2nd PCR: HF, 50 ng target plasmid, 2-step PCR (72 °C), 130 s
ExtD (GSU2642)	1st PCR: HF, 0.5 μM primers, 50 ng template, 68 °C, 15 s 2nd PCR: HF, 50 ng target plasmid, 2-step PCR (72 °C), 100 s
ExtE (GSU2726)	1st PCR: HF, 0.5 μM primers, 50 ng template, 64 °C, 40 s 2nd PCR: GC, 50 ng target plasmid, 7% DMSO, 2-step PCR (72 °C), 170 s
ExtG (GSU2724)	1st PCR: HF, 0.5 μM primers, 50 ng template 63 °C, 40 s 2nd PCR: HF, 50 ng target plasmid, 2-step PCR (72 °C), 130 s
ExtH (GSU2940)	The pET-28a(+)-ExtH plasmid was acquired from GeneCust
ExtI (GSU2939)	1st PCR: HF, 0.5 μM primers, 50 ng template, 64 °C, 40 s 2nd PCR: GC, 50 ng target plasmid, 7% DMSO, 2-step PCR (72 °C), 170 s
ExtJ (GSU2938)	1st PCR: HF, 0.5 μM primers, 50 ng template, 66 °C, 20 s 2nd PCR: HF, 50 ng target plasmid, 3% DMSO, 2-step PCR (72 °C), 140 s
ExtK (GSU2937)	1st PCR: HF, 0.5 μM primers, 50 ng template, 66 °C, 20 s 2nd PCR: HF, 50 ng target plasmid, 3% DMSO, 2-step PCR (72 °C), 140 s
ExtKL	1st PCR: HF, 0.5 μM primers, 50 ng template, 65 °C, 40 s 2nd PCR: GC, 50 ng target plasmid, 7% DMSO, touchdown PCR, 150 s
ExtL (GSU2936)	1st PCR: HF, 0.5 μM primers, 50 ng template, 66 °C, 20 s 2nd PCR: HF, 50 ng target plasmid, 3% DMSO, 2-step PCR (72 °C), 140 s

After the second PCR reaction, the final PCR products were incubated with one unit of DpnI (Thermo Fisher Scientific) for 2 h at 37 °C for template DNA digestion, followed by a 20-minute period

at 80 °C for enzyme inactivation. *E. coli* DH5α cells were transformed with the resulting products using the heat-shock method [196], and plated for selection at 37 °C in Luria-Bertani (LB) medium supplemented with either ampicillin (100 μg/mL) or kanamycin (30 μg/mL), for pVA203 and pET-28a(+) constructs, respectively. Proper controls were performed. The resulting colonies were screened by colony PCR [197, 198], using primers complementary to either the pVA203 or pET-28a(+) plasmids, and a Taq DNA polymerase (VWR). The results were analyzed by agarose gel electrophoresis and the colonies with PCR products of the correct size were grown overnight at 37 °C in liquid LB, supplemented with the corresponding antibiotic. Plasmid extraction and purification was performed with the NZYMiniprep kit (NZYTech). Finally, the presence of the desired gene sequence in the final constructs was confirmed by DNA Sanger sequencing performed by STAB VIDA.

2.6.3. Insertion of affinity-tags in cytochrome constructs

In order to facilitate the purification process of the cytochrome targets, their plasmids were modified to contain either a N-terminal Strep-tag (with a TEV cleavage site and appropriate flexible linkers) or a C-terminal His-Tag. These tags were inserted into the target plasmids using a strategy based on the Q5 Site-Directed Mutagenesis kit from New England Biolabs. The primers used were designed using the NEBaseChanger website and their sequences are listed in Table 2.8.

Table 2.8. Sequences of the NEBaseChanger primers used to produce the cytochrome vectors containing affinity tags. The gene sequences corresponding to the inserts, including the flexible linkers (AMASA and GS, blue), the Strep-tag (WSHPQFEK, orange), the His-tag (HHHHHH, purple), and the cleavage site for TEV protease (ENLYFQS, green) are highlighted. The pVA203_rv primer was used for all cases in which a N-terminal Strep-tag was inserted. The primers were purchased from Invitrogen.

Primer	DNA sequence (5' – 3')
ExtA_fw	GAAAAATCCGGAGAAAACCTGTATTTTCAGTCTAGTGAAAAGAGCGGCGCCGTG
ExtC_fw	GAAAAATCCGGAGAAAACCTGTATTTTCAGTCTGCGCGCTATCGGCTGCCC
ExtD_fw	GAAAAATCCGGAGAAAACCTGTATTTTCAGTCTGGCAACGCGTCCTCTCCC
ExtF_fw	GAAAAATCCGGAGAAAACCTGTATTTTCAGTCTATGCTTTCGAAGGAATCGAGCCTGCC
ExtG_fw	GAAAAATCCGGAGAAAACCTGTATTTTCAGTCTATCGACCCGACAACCGGCAG
ExtK_fw	GAAAAATCCGGAGAAAACCTGTATTTTCAGTCTGAAAAGGCGGGCATCGGCTG
ExtKL_fw	CATCATCATCACTGAATTCGAGCTCCGTC
ExtKL_rv	ATGATGGCTGCCCTTGCCCTTGGCGC
pVA203_rv	GAATTGAGGATGACTCCATGCGCTAGCCATTGCGGCGGCCGCAACGGTAGC

The standard PCR mix used in the PCR reactions contained HF buffer (Thermo Fisher Scientific), 0.2 mM dNTPs (Thermo Fisher Scientific), 0.5 μM of each custom-designed primer, 10 ng of target plasmid and one unit of Phusion High-Fidelity DNA polymerase (Thermo Fisher Scientific), for a final volume of 50 μL. A 2-step thermocycling PCR protocol with 30 cycles was used for these amplification

reactions, in which the extension and annealing steps were merged, using a temperature of 72 °C for a period of 140 seconds. Following this step, the resulting PCR products were incubated with one unit of DpnI (Thermo Fisher Scientific) for 2 h at 37 °C for template DNA digestion, followed by a 20-minute period at 80 °C for enzyme inactivation. The undigested products were purified using the NZYGelpure kit (NZYTech) to remove remaining components of the first PCR reaction and DpnI, and quantified in a NanoDrop spectrophotometer ND-1000 (Thermo Fisher Scientific).

The pure PCR products were incubated with one unit of T4 polynucleotide kinase (Thermo Fisher Scientific) and T4 DNA ligase (Thermo Fisher Scientific), in T4 DNA ligase buffer (Thermo Fisher Scientific) containing 10 mM $MgCl_2$, 10 mM dithiothreitol (DTT) and 0.5 mM ATP, for 1 hour at 30 °C, followed by an incubation period of 10 minutes at 65 °C. After enzymatic treatment, *E. coli* DH5α cells were transformed with the resulting products using the heat-shock method [196], and plated for selection at 37 °C in LB medium supplemented with ampicillin (100 μg/mL). Proper controls were performed. The resulting colonies were screened by colony PCR, as described in the previous section.

2.6.4. Production of the truncated ExtA and ExtK constructs

The truncated versions of ExtA and ExtK were cloned by applying the Q5 Site-Directed Mutagenesis strategy. Residues 32-72 were truncated from the ST-TEV-ExtA construct, whereas residues 286-314 were deleted from the ST-TEV-ExtK construct. In the case of ExtK, two additional constructs were designed, with linkers composed by one or two proline residues. The primers used were designed using the NEBaseChanger website and their sequences are listed in Table 2.9. The PCR master mix and cycling conditions, as well as the remaining steps of the protocol, were identical to the ones used in the insertion of the affinity tags in the cytochrome constructs, described in the previous section.

Table 2.9. Sequences of the NEBaseChanger primers used to produce the truncated ExtA and ExtK cytochrome vectors. The gene sequences corresponding to the "P" and "PP" linkers are colored orange. The primers were purchased from Invitrogen. "$ExtK_N$" refers to the protein without part of the C-terminal domain.

Primer	DNA sequence (5' – 3')
$ExtA_{73\text{-}210}$_fw	ACGGTCACCCAGTGCGGC
$ExtA_{73\text{-}210}$_rv	AGACTGAAAATACAGGTTTTCTCCGGATTTTTCG
$ExtK_N$_fw	AATCACGCGGGCCACTCGATC
$ExtK_N$_rv	GGCGGCCTTGGCCATGCC
$ExtK_{N+P}$_fw	GAATCACGCGGGCCACTCGATC
$ExtK_{N+P}$_rv	GGGGCGGCCTTGGCCATGCC
$ExtK_{N+PP}$_fw	CCGAATCACGCGGGCCACTCGATC
$ExtK_{N+PP}$_rv	CGGGGCGGCCTTGGCCATGCC

2.6.5. Protein expression tests

In order to optimize the expression conditions of some of the cytochrome constructs (see Table 2.3), different strains of competent *E. coli* cells (BL21 (DE3), C41 (DE3), C43 (DE3), JCB7123, JM109, SF110 and Tuner (DE3)) containing the pEC86 plasmid [163], were transformed with the target cytochrome-containing pVA203 plasmids using the heat-shock method [196]. In the case of ExtJ and the ExtB-C tandem, different *E. coli* strains (BL21 (DE3), BL21 (DE3) pLysS, BL21 Star (DE3), C41 (DE3), C43 (DE3), Rosetta 2 (DE3) pLysS and Tuner (DE3)) were transformed with these plasmids. After bacterial transformation, different protein expression variables were tested, namely the temperature (25, 30 and 37 °C), shaking speeds (100, 160, 180 and 200 RPM), and incubation times (4 h, 8 h or overnight) in the absence and presence of different concentrations of IPTG (10, 20, 50 and 100 μM).

To test for the different expression conditions, aliquots of bacterial medium were taken using a weighted correction for the normalization of the cell pellets. These aliquots were harvested and resuspended either in 50 μL of water for whole-cell analysis or in 50 μL of NZY Bacterial Cell Lysis Buffer with lysozyme and DNase I (NZYTech) for analysis of the soluble fraction. In either case, the samples were mixed in a 1:1 ratio with SDS loading buffer, incubated for 10 minutes at 100 °C and loaded into either 12.5 or 15% acrylamide/bis-acrylamide SDS-PAGE in a Tris-tricine buffer [199]. The gels were stained for heme detection using 3,3',5,5'-tetramethylbenzidine (TMBZ, Thermo Fisher Scientific) [200] and/or BlueSafe (NZYTech).

2.7. References

[1] M.D. Pizzagalli, A. Bensimon, G. Superti-Furga, A guide to plasma membrane solute carrier proteins, FEBS J., 288 (2021) 2784-2835.
[2] G.L. Holliday, D.E. Almonacid, J.B.O. Mitchell, J.M. Thornton, The chemistry of protein catalysis, J. Mol. Biol., 372 (2007) 1261-1277.
[3] G.J. Mulder, Sur la composition de quelques substances animales, Bull. Sci. Phys. Nat. Néerl., (1838) 104-119.
[4] G.J. Mulder, Ueber die Zusammensetzung einiger thierischen Substanzen, J. Prakt. Chem., 16 (1839) 129-152.
[5] C.B. Anfinsen, Principles that govern the folding of protein chains, Science, 181 (1973) 223-230.
[6] C.B. Anfinsen, E. Haber, M. Sela, F.H. White, The kinetics of formation of native ribonuclease during oxidation of the reduced polypeptide chain, Proc. Natl. Acad. Sci., 47 (1961) 1309-1314.
[7] H. Berman, K. Henrick, H. Nakamura, Announcing the worldwide Protein Data Bank, Nat. Struct. Mol. Biol., 10 (2003) 980.
[8] S.K. Burley, H.M. Berman, C. Bhikadiya, C. Bi, L. Chen, L. Di Costanzo, C. Christie, K. Dalenberg, J.M. Duarte, S. Dutta, Z. Feng, S. Ghosh, D.S. Goodsell, R.K. Green, V. Guranović, D. Guzenko, B.P. Hudson, T. Kalro, Y. Liang, R. Lowe, H. Namkoong, E. Peisach, I. Periskova, A. Prlić, C. Randle, A. Rose, P. Rose, R. Sala, M. Sekharan, C. Shao, L. Tan, Y.-P. Tao, Y. Valasatava, M. Voigt, J. Westbrook, J. Woo, H. Yang, J. Young, M. Zhuravleva, C. Zardecki, RCSB Protein Data Bank: Biological macromolecular structures enabling research and education in fundamental biology, biomedicine, biotechnology and energy, Nucleic Acids Res., 47 (2019) D464-D474.
[9] J R. Helliwell, New developments in crystallography: Exploring its technology, methods and scope in the molecular biosciences, Biosci. Rep., 37 (2017) BSR20170204.
[10] I. Rathore, V. Mishra, P. Bhaumik, Advancements in macromolecular crystallography: From past to present, Emerg. Top. Life Sci., 5 (2021) 127-149.
[11] M.-A. Delsuc, M. Vitorino, B. Kieffer, Determination of protein structure and dynamics by NMR, Structural biology in drug discovery, (2020) 295-323.
[12] Y. Hu, K. Cheng, L. He, X. Zhang, B. Jiang, L. Jiang, C. Li, G. Wang, Y. Yang, M. Liu, NMR-based methods for protein analysis, Anal. Chem., 93 (2021) 1866-1879.
[13] X. Wu, A. Rapoport Tom, Cryo-EM structure determination of small proteins by nanobody-binding scaffolds (Legobodies), Proc. Natl. Acad. Sci., 118 (2021) e2115001118.
[14] K.M. Yip, N. Fischer, E. Paknia, A. Chari, H. Stark, Atomic-resolution protein structure determination by cryo-EM, Nature, 587 (2020) 157-161.
[15] K.A. Dill, J.L. MacCallum, The protein-folding problem, 50 years on, Science, 338 (2012) 1042-1046.
[16] K.A. Dill, S.B. Ozkan, M.S. Shell, T.R. Weikl, The protein folding problem, Annu. Rev. Biophys., 37 (2008) 289-316.
[17] R. Zwanzig, A. Szabo, B. Bagchi, Levinthal's paradox, Proc. Natl. Acad. Sci., 89 (1992) 20-22.
[18] C. Levinthal, Are there pathways for protein folding?, J. Chim. Phys., 65 (1968) 44-45.

[19] C. Levinthal, How to fold graciously, Mössbauer spectroscopy in biological systems: Proceedings of a meeting held at Allerton House, Monticello, Illinois, University of Illinois Press, (1969) 22-24.
[20] J. Moult, J.T. Pedersen, R. Judson, K. Fidelis, A large-scale experiment to assess protein structure prediction methods, Proteins, 23 (1995) II-IV.
[21] J. Jumper, R. Evans, A. Pritzel, T. Green, M. Figurnov, O. Ronneberger, K. Tunyasuvunakool, R. Bates, A. Žídek, A. Potapenko, A. Bridgland, C. Meyer, S.A.A. Kohl, A.J. Ballard, A. Cowie, B. Romera-Paredes, S. Nikolov, R. Jain, J. Adler, T. Back, S. Petersen, D. Reiman, E. Clancy, M. Zielinski, M. Steinegger, M. Pacholska, T. Berghammer, S. Bodenstein, D. Silver, O. Vinyals, A.W. Senior, K. Kavukcuoglu, P. Kohli, D. Hassabis, Highly accurate protein structure prediction with AlphaFold, Nature, 596 (2021) 583-589.
[22] A.W. Senior, R. Evans, J. Jumper, J. Kirkpatrick, L. Sifre, T. Green, C. Qin, A. Žídek, A.W.R. Nelson, A. Bridgland, H. Penedones, S. Petersen, K. Simonyan, S. Crossan, P. Kohli, D.T. Jones, D. Silver, K. Kavukcuoglu, D. Hassabis, Improved protein structure prediction using potentials from deep learning, Nature, 577 (2020) 706-710.
[23] A. Kryshtafovych, T. Schwede, M. Topf, K. Fidelis, J. Moult, Critical assessment of methods of protein structure prediction (CASP) – Round XIII, Proteins, 87 (2019) 1011-1020.
[24] A.L. Mitchell, A. Almeida, M. Beracochea, M. Boland, J. Burgin, G. Cochrane, M.R. Crusoe, V. Kale, S.C. Potter, L.J. Richardson, E. Sakharova, M. Scheremetjew, A. Korobeynikov, A. Shlemov, O. Kunyavskaya, A. Lapidus, R.D. Finn, MGnify: The microbiome analysis resource in 2020, Nucleic Acids Res., 48 (2020) D570-D578.
[25] B.E. Suzek, Y. Wang, H. Huang, P.B. McGarvey, C.H. Wu, C. the UniProt, UniRef clusters: A comprehensive and scalable alternative for improving sequence similarity searches, Bioinformatics, 31 (2015) 926-932.
[26] M. Mirdita, L. von den Driesch, C. Galiez, M.J. Martin, J. Söding, M. Steinegger, Uniclust databases of clustered and deeply annotated protein sequences and alignments, Nucleic Acids Res., 45 (2017) D170-D176.
[27] L.S. Johnson, S.R. Eddy, E. Portugaly, Hidden Markov model speed heuristic and iterative HMM search procedure, BMC Bioinformatics, 11 (2010) 431.
[28] M. Remmert, A. Biegert, A. Hauser, J. Söding, HHblits: Lightning-fast iterative protein sequence searching by HMM-HMM alignment, Nat. Methods, 9 (2012) 173-175.
[29] M. Steinegger, M. Meier, M. Mirdita, H. Vöhringer, S.J. Haunsberger, J. Söding, HH-suite3 for fast remote homology detection and deep protein annotation, BMC Bioinformatics, 20 (2019) 473.
[30] T. Lassmann, O. Frings, E.L.L. Sonnhammer, Kalign2: High-performance multiple alignment of protein and nucleotide sequences allowing external features, Nucleic Acids Res., 37 (2009) 858-865.
[31] V. Hornak, R. Abel, A. Okur, B. Strockbine, A. Roitberg, C. Simmerling, Comparison of multiple Amber force fields and development of improved protein backbone parameters, Proteins, 65 (2006) 712-725.
[32] P. Eastman, J. Swails, J.D. Chodera, R.T. McGibbon, Y. Zhao, K.A. Beauchamp, L.-P. Wang, A.C. Simmonett, M.P. Harrigan, C.D. Stern, R.P. Wiewiora, B.R. Brooks, V.S. Pande, OpenMM 7: Rapid

development of high performance algorithms for molecular dynamics, PLoS Comput. Biol., 13 (2017) e1005659.
[33] V. Mariani, M. Biasini, A. Barbato, T. Schwede, lDDT: A local superposition-free score for comparing protein structures and models using distance difference tests, Bioinformatics, 29 (2013) 2722-2728.
[34] M. Baek, F. DiMaio, I. Anishchenko, J. Dauparas, S. Ovchinnikov, G.R. Lee, J. Wang, Q. Cong, L.N. Kinch, R.D. Schaeffer, C. Millán, H. Park, C. Adams, C.R. Glassman, A. DeGiovanni, J.H. Pereira, A.V. Rodrigues, A.A. van Dijk, A.C. Ebrecht, D.J. Opperman, T. Sagmeister, C. Buhlheller, T. Pavkov-Keller, M.K. Rathinaswamy, U. Dalwadi, C.K. Yip, J.E. Burke, K.C. Garcia, N.V. Grishin, P.D. Adams, R.J. Read, D. Baker, Accurate prediction of protein structures and interactions using a three-track neural network, Science, 373 (2021) 871-876.
[35] Z. Lin, H. Akin, R. Rao, B. Hie, Z. Zhu, W. Lu, N. Smetanin, R. Verkuil, O. Kabeli, Y. Shmueli, A. dos Santos Costa, M. Fazel-Zarandi, T. Sercu, S. Candido, A. Rives, Evolutionary-scale prediction of atomic-level protein structure with a language model, Science, 379 (2023) 1123-1130.
[36] W. Wang, Z. Peng, J. Yang, Single-sequence protein structure prediction using supervised transformer protein language models, Nat. Comput. Sci., 2 (2022) 804-814.
[37] A. Madani, B. Krause, E.R. Greene, S. Subramanian, B.P. Mohr, J.M. Holton, J.L. Olmos, C. Xiong, Z.Z. Sun, R. Socher, J.S. Fraser, N. Naik, Large language models generate functional protein sequences across diverse families, Nat. Biotechnol., 41 (2023) 1099-1106.
[38] W. Ruidong, D. Fan, W. Rui, S. Rui, Z. Xiwen, L. Shitong, S. Chenpeng, W. Zuofan, X. Qi, B. Bonnie, M. Jianzhu, P. Jian, High-resolution *de novo* structure prediction from primary sequence, bioRxiv, (2022) 2022.2007.2021.500999.
[39] R.J.L. Townshend, S. Eismann, A.M. Watkins, R. Rangan, M. Karelina, R. Das, R.O. Dror, Geometric deep learning of RNA structure, Science, 373 (2021) 1047-1051.
[40] M.L. Hekkelman, I. de Vries, R.P. Joosten, A. Perrakis, AlphaFill: Enriching AlphaFold models with ligands and cofactors, Nat. Methods, 20 (2023) 205-213.
[41] L.M.F. Bertoline, A.N. Lima, J.E. Krieger, S.K. Teixeira, Before and after AlphaFold2: An overview of protein structure prediction, Front. Bioinform., 3 (2023).
[42] T.C. Terwilliger, B.K. Poon, P.V. Afonine, C.J. Schlicksup, T.I. Croll, C. Millán, J.S. Richardson, R.J. Read, P.D. Adams, Improved AlphaFold modeling with implicit experimental information, Nat. Methods, 19 (2022) 1376-1382.
[43] M. Akdel, D.E.V. Pires, E.P. Pardo, J. Jänes, A.O. Zalevsky, B. Mészáros, P. Bryant, L.L. Good, R.A. Laskowski, G. Pozzati, A. Shenoy, W. Zhu, P. Kundrotas, V.R. Serra, C.H.M. Rodrigues, A.S. Dunham, D. Burke, N. Borkakoti, S. Velankar, A. Frost, J. Basquin, K. Lindorff-Larsen, A. Bateman, A.V. Kajava, A. Valencia, S. Ovchinnikov, J. Durairaj, D.B. Ascher, J.M. Thornton, N.E. Davey, A. Stein, A. Elofsson, T.I. Croll, P. Beltrao, A structural biology community assessment of AlphaFold2 applications, Nat. Struct. Mol. Biol., 29 (2022) 1056-1067.

[44] B. Golinelli-Pimpaneau, Prediction of the iron-sulfur binding sites in proteins using the highly accurate three-dimensional models calculated by AlphaFold and RoseTTAFold, Inorganics, 10 (2022) 2.

[45] K. Tunyasuvunakool, J. Adler, Z. Wu, T. Green, M. Zielinski, A. Žídek, A. Bridgland, A. Cowie, C. Meyer, A. Laydon, S. Velankar, G.J. Kleywegt, A. Bateman, R. Evans, A. Pritzel, M. Figurnov, O. Ronneberger, R. Bates, S.A.A. Kohl, A. Potapenko, A.J. Ballard, B. Romera-Paredes, S. Nikolov, R. Jain, E. Clancy, D. Reiman, S. Petersen, A.W. Senior, K. Kavukcuoglu, E. Birney, P. Kohli, J. Jumper, D. Hassabis, Highly accurate protein structure prediction for the human proteome, Nature, 596 (2021) 590-596.

[46] M. Varadi, S. Anyango, M. Deshpande, S. Nair, C. Natassia, G. Yordanova, D. Yuan, O. Stroe, G. Wood, A. Laydon, A. Žídek, T. Green, K. Tunyasuvunakool, S. Petersen, J. Jumper, E. Clancy, R. Green, A. Vora, M. Lutfi, M. Figurnov, A. Cowie, N. Hobbs, P. Kohli, G. Kleywegt, E. Birney, D. Hassabis, S. Velankar, AlphaFold protein structure database: Massively expanding the structural coverage of protein-sequence space with high-accuracy models, Nucleic Acids Res., 50 (2022) D439-D444.

[47] R. Evans, M. O'Neill, A. Pritzel, N. Antropova, A. Senior, T. Green, A. Žídek, R. Bates, S. Blackwell, J. Yim, O. Ronneberger, S. Bodenstein, M. Zielinski, A. Bridgland, A. Potapenko, A. Cowie, K. Tunyasuvunakool, R. Jain, E. Clancy, P. Kohli, J. Jumper, D. Hassabis, Protein complex prediction with AlphaFold-Multimer, bioRxiv, (2022) 2021.2010.2004.463034.

[48] P. Bryant, G. Pozzati, W. Zhu, A. Shenoy, P. Kundrotas, A. Elofsson, Predicting the structure of large protein complexes using AlphaFold and Monte Carlo tree search, Nat. Commun., 13 (2022) 6028.

[49] P. Bryant, G. Pozzati, A. Elofsson, Improved prediction of protein-protein interactions using AlphaFold2, Nat. Commun., 13 (2022) 1265.

[50] I. Johansson-Åkhe, B. Wallner, Improving peptide-protein docking with AlphaFold-Multimer using forced sampling, Front. Bioinform., 2 (2022) 959160.

[51] D. Yu, G. Chojnowski, M. Rosenthal, J. Kosinski, AlphaPulldown – A python package for protein-protein interaction screens using AlphaFold-Multimer, Bioinformatics, 39 (2023) btac749.

[52] M. Mirdita, K. Schütze, Y. Moriwaki, L. Heo, S. Ovchinnikov, M. Steinegger, ColabFold: Making protein folding accessible to all, Nat. Methods, 19 (2022) 679-682.

[53] X. Liu, L. Shi, J.-D. Gu, Microbial electrocatalysis: Redox mediators responsible for extracellular electron transfer, Biotechnol. Adv., 36 (2018) 1815-1827.

[54] A. Okamoto, K. Saito, K. Inoue, K.H. Nealson, K. Hashimoto, R. Nakamura, Uptake of self-secreted flavins as bound cofactors for extracellular electron transfer in *Geobacter* species, Energy Environ. Sci., 7 (2014) 1357-1361.

[55] T.A. Clarke, Plugging into bacterial nanowires: A comparison of model electrogenic organisms, Curr. Opin. Microbiol., 66 (2022) 56-62.

[56] M.J. Edwards, G.F. White, J.N. Butt, D.J. Richardson, T.A. Clarke, The crystal structure of a biological insulated transmembrane molecular wire, Cell, 181 (2020) 665-673.e610.

[57] C.A. Salgueiro, L. Morgado, M.A. Silva, M.R. Ferreira, T.M. Fernandes, P.C. Portela, From iron to bacterial electroconductive filaments: Exploring cytochrome diversity using *Geobacter* bacteria, Coordin. Chem. Rev., 452 (2022) 214284.

[58] F. Wang, Y. Gu, J.P. O'Brien, S.M. Yi, S.E. Yalcin, V. Srikanth, C. Shen, D. Vu, N.L. Ing, A.I. Hochbaum, E.H. Egelman, N.S. Malvankar, Structure of microbial nanowires reveals stacked hemes that transport electrons over micrometers, Cell, 177 (2019) 361-369.

[59] G.F. White, M.J. Edwards, L. Gomez-Perez, D.J. Richardson, J.N. Butt, T.A. Clarke, Chapter three – Mechanisms of bacterial extracellular electron exchange, Advances in Microbial Physiology, Academic Press, 2016, 87-138.

[60] V. Jiang, S.D. Khare, S. Banta, Computational structure prediction provides a plausible mechanism for electron transfer by the outer membrane protein Cyc2 from *Acidithiobacillus ferrooxidans*, Protein Sci., 30 (2021) 1640-1652.

[61] C. Jeans, S.W. Singer, C.S. Chan, N.C. VerBerkmoes, M. Shah, R.L. Hettich, J.F. Banfield, M.P. Thelen, Cytochrome 572 is a conspicuous membrane protein with iron oxidation activity purified directly from a natural acidophilic microbial community, ISME J., 2 (2008) 542-550.

[62] R.C. Blake, J.E. Shively, R. Timkovich, R.A. White, Homogeneous cytochrome 579 is an octamer that reacts too slowly with soluble iron to be the initial iron oxidase in the respiratory chain of *Leptospirillum ferriphilum*, Front. Microbiol., 12 (2021) 673066.

[63] W. Arunmanee, M. Pathania, A.S. Solovyova, A.P. Le Brun, H. Ridley, A. Baslé, B. van den Berg, J.H. Lakey, Gram-negative trimeric porins have specific LPS binding sites that are essential for porin biogenesis, Proc. Natl. Acad. Sci., 113 (2016) E5034-E5043.

[64] D. Gupta, M.C. Sutherland, K. Rengasamy, J.M. Meacham, R.G. Kranz, A. Bose, Photoferrotrophs produce a PioAB electron conduit for extracellular electron uptake, mBio, 10 (2019) e02668.

[65] J. Liu, Z. Wang, S. Belchik, M. Edwards, C. Liu, D. Kennedy, E. Merkley, M. Lipton, J. Butt, D. Richardson, J. Zachara, J. Fredrickson, K. Rosso, L. Shi, Identification and characterization of MtoA: A decaheme *c*-type cytochrome of the neutrophilic Fe(II)-oxidizing bacterium *Sideroxydans lithotrophicus* ES-1, Front. Microbiol., 3 (2012) 37.

[66] F.J. Otero, C.H. Chan, D.R. Bond, Identification of different putative outer membrane electron conduits necessary for Fe(III) citrate, Fe(III) oxide, Mn(IV) oxide, or electrode reduction by *Geobacter sulfurreducens*, J. Bacteriol., 200 (2018) e00347.

[67] Y. Liu, Z. Wang, J. Liu, C. Levar, M.J. Edwards, J.T. Babauta, D.W. Kennedy, Z. Shi, H. Beyenal, D.R. Bond, T.A. Clarke, J.N. Butt, D.J. Richardson, K.M. Rosso, J.M. Zachara, J.K. Fredrickson, L. Shi, A trans-outer membrane porin-cytochrome protein complex for extracellular electron transfer by *Geobacter sulfurreducens* PCA, Environ. Microbiol. Rep., 6 (2014) 776-785.

[68] M. Aklujkar, M.V. Coppi, C. Leang, B.C. Kim, M.A. Chavan, L.A. Perpetua, L. Giloteaux, A. Liu, D.E. Holmes, Proteins involved in electron transfer to Fe(III) and Mn(IV) oxides by *Geobacter sulfurreducens* and *Geobacter uraniireducens*, Microbiology, 159 (2013) 515-535.

[69] C. Leang, M.V. Coppi, D.R. Lovley, OmcB, a *c*-type polyheme cytochrome, involved in Fe(III) reduction in *Geobacter sulfurreducens*, J. Bacteriol., 185 (2003) 2096-2103.

[70] C. Leang, D.R. Lovley, Regulation of two highly similar genes, *omcB* and *omcC*, in a 10 kb chromosomal duplication in *Geobacter sulfurreducens*, Microbiology, 151 (2005) 1761-1767.
[71] Y. Liu, J.K. Fredrickson, J.M. Zachara, L. Shi, Direct involvement of *ombB*, *omaB*, and *omcB* genes in extracellular reduction of Fe(III) by *Geobacter sulfurreducens* PCA, Front. Microbiol., 6 (2015) 1075.
[72] L. Shi, J.K. Fredrickson, J.M. Zachara, Genomic analyses of bacterial porin-cytochrome gene clusters, Front. Microbiol., 5 (2014) 657.
[73] C.C. Page, C.C. Moser, X. Chen, P.L. Dutton, Natural engineering principles of electron tunnelling in biological oxidation-reduction, Nature, 402 (1999) 47-52.
[74] C. Janiak, A critical account on π-π stacking in metal complexes with aromatic nitrogen-containing ligands, J. Chem. Soc. Dalton Trans., (2000) 3885-3896.
[75] L. Holm, Dali server: Structural unification of protein families, Nucleic Acids Res., 50 (2022) W210-W215.
[76] M. van Kempen, S.S. Kim, C. Tumescheit, M. Mirdita, J. Lee, C.L.M. Gilchrist, J. Söding, M. Steinegger, Fast and accurate protein structure search with Foldseek, Nat. Biotechnol., 42 (2023) 243-246.
[77] M. Varadi, J. Berrisford, M. Deshpande, S.S. Nair, A. Gutmanas, D. Armstrong, L. Pravda, B. Al-Lazikani, S. Anyango, G.J. Barton, K. Berka, T. Blundell, N. Borkakoti, J. Dana, S. Das, S. Dey, P. Di Micco, F. Fraternali, T. Gibson, M. Helmer-Citterich, D. Hoksza, L.C. Huang, R. Jain, H. Jubb, C. Kannas, N. Kannan, J. Koca, R. Krivak, M. Kumar, E.D. Levy, F. Madeira, M.S. Madhusudhan, H.J. Martell, S. MacGowan, J.E. McGreig, S. Mir, A. Mukhopadhyay, L. Parca, T. Paysan-Lafosse, L. Radusky, A. Ribeiro, L. Serrano, I. Sillitoe, G. Singh, P. Skoda, R. Svodobova, J. Tyzack, A. Valencia, E.V. Fernandez, W. Vranken, M. Wass, J.M. Thornton, M. Sternberg, C. Orengo, S. Velankar, PDBe-KB: A community-driven resource for structural and functional annotations, Nucleic Acids Res., 48 (2020) D344-D353.
[78] D. Paul, A. Chatterjee, T.P. Begley, S.E. Ealick, Domain organization in *Candida glabrata* THI6, a bifunctional enzyme required for thiamin biosynthesis in eukaryotes, Biochemistry, 49 (2010) 9922-9934.
[79] M.C. Cummins, T.M. Jacobs, F.D. Teets, F. DiMaio, A. Tripathy, B. Kuhlman, AlphaFold accurately predicts distinct conformations based on the oligomeric state of a de novo designed protein, Protein Sci., 31 (2022) e4368.
[80] D. Sala, F. Engelberger, H.S. McHaourab, J. Meiler, Modeling conformational states of proteins with AlphaFold, Curr. Opin. Struc. Biol., 81 (2023) 102645.
[81] T. Saldaño, N. Escobedo, J. Marchetti, D.J. Zea, J. Mac Donagh, A.J. Velez Rueda, E. Gonik, A. García Melani, J. Novomisky Nechcoff, M.N. Salas, T. Peters, N. Demitroff, S. Fernandez Alberti, N. Palopoli, M.S. Fornasari, G. Parisi, Impact of protein conformational diversity on AlphaFold predictions, Bioinformatics, 38 (2022) 2742-2748.
[82] R.A. Stein, H.S. McHaourab, SPEACH_AF: Sampling protein ensembles and conformational heterogeneity with AlphaFold2, PLoS Comput. Biol., 18 (2022) e1010483.

[83] H.K. Wayment-Steele, A. Ojoawo, R. Otten, J.M. Apitz, W. Pitsawong, M. Hömberger, S. Ovchinnikov, L. Colwell, D. Kern, Predicting multiple conformations via sequence clustering and AlphaFold2, Nature, 625 (2023) 832-839.

[84] J. Hallgren, K.D. Tsirigos, M.D. Pedersen, J.J. Almagro Armenteros, P. Marcatili, H. Nielsen, A. Krogh, O. Winther, DeepTMHMM predicts alpha and beta transmembrane proteins using deep neural networks, bioRxiv, (2022) 2022.2004.2008.487609.

[85] S. Galdiero, A. Falanga, M. Cantisani, R. Tarallo, M. Elena Della Pepa, apos, V. Oriano, M. Galdiero, Microbe-host interactions: Structure and role of Gram-negative bacterial porins, Curr. Prot. Pept. Sci., 13 (2012) 843-854.

[86] T. Schirmer, General and specific porins from bacterial outer membranes, J. Struct. Biol., 121 (1998) 101-109.

[87] M.J. Edwards, G.F. White, M. Norman, A. Tome-Fernandez, E. Ainsworth, L. Shi, J.K. Fredrickson, J.M. Zachara, J.N. Butt, D.J. Richardson, T.A. Clarke, Redox linked flavin sites in extracellular decaheme proteins involved in microbe-mineral electron transfer, Sci. Rep., 5 (2015) 11677.

[88] B.C. Berks, S.J. Ferguson, J.W.B. Moir, D.J. Richardson, Enzymes and associated electron transport systems that catalyse the respiratory reduction of nitrogen oxides and oxyanions, Biochim. Biophys. Acta - Bioenergetics, 1232 (1995) 97-173.

[89] J.A. Cole, C.M. Brown, Nitrite reduction to ammonia by fermentative bacteria: A short circuit in the biological nitrogen cycle, FEMS Microbiol. Lett., 7 (1980) 65-72.

[90] O. Einsle, A. Messerschmidt, P. Stach, G.P. Bourenkov, H.D. Bartunik, R. Huber, P.M.H. Kroneck, Structure of cytochrome *c* nitrite reductase, Nature, 400 (1999) 476-480.

[91] A.A. Korenevsky, E. Vinogradov, Y. Gorby, J.B. Terry, Characterization of the lipopolysaccharides and capsules of *Shewanella* spp., Appl. Environ. Microb., 68 (2002) 4653-4657.

[92] F.J. Otero, G.L. Chadwick, M.D. Yates, R.L. Mickol, S.H. Saunders, S.M. Glaven, J.A. Gralnick, D.K. Newman, L.M. Tender, V.J. Orphan, D.R. Bond, Evidence of a streamlined extracellular electron transfer pathway from biofilm structure, metabolic stratification, and long-range electron transfer parameters, Appl. Environ. Microbiol., 87 (2021) e00706.

[93] K. Parey, A.J. Fielding, M. Sörgel, R. Rachel, H. Huber, C. Ziegler, C. Rajendran, In meso crystal structure of a novel membrane-associated octaheme cytochrome *c* from the Crenarchaeon *Ignicoccus hospitalis*, FEBS J., 283 (2016) 3807-3820.

[94] R.O. Louro, Proton thrusters: Overview of the structural and functional features of soluble tetrahaem cytochromes c_3, J. Biol. Inorg. Chem., 12 (2007) 1-10.

[95] J. Huang, J. Zarzycki, M.R. Gunner, W.W. Parson, J.F. Kern, J. Yano, D.C. Ducat, D.M. Kramer, Mesoscopic to macroscopic electron transfer by hopping in a crystal network of cytochromes, J. Am. Chem. Soc., 142 (2020) 10459-10467.

[96] F. Wang, K. Mustafa, V. Suciu, K. Joshi, C.H. Chan, S. Choi, Z. Su, D. Si, A.I. Hochbaum, E.H. Egelman, D.R. Bond, Cryo-EM structure of an extracellular *Geobacter* OmcE cytochrome filament reveals tetrahaem packing, Nat. Microbiol., 7 (2022) 1291-1300.

[97] D. Leys, A.S. Tsapin, K.H. Nealson, T.E. Meyer, M.A. Cusanovich, J.J. Van Beeumen, Structure and mechanism of the flavocytochrome *c* fumarate reductase of *Shewanella putrefaciens* MR-1, Nat. Struct. Biol., 6 (1999) 1113-1117.

[98] M.A. Silva, A.P. Fernandes, D.L. Turner, C.A. Salgueiro, A biochemical deconstruction-based strategy to assist the characterization of bacterial electric conductive filaments, Int. J. Mol. Sci., 24 (2023) 7032.

[99] V. Magro, L. Pieulle, N. Forget, B. Guigliarelli, Y. Petillot, E.C. Hatchikian, Further characterization of the two tetraheme cytochromes c_3 from *Desulfovibrio africanus*: Nucleotide sequences, EPR spectroscopy and biological activity, Biochim. Biophys. Acta - Protein Struct. Mol. Enzymol., 1342 (1997) 149-163.

[100] S. Nørager, P. Legrand, L. Pieulle, C. Hatchikian, M. Roth, Crystal structure of the oxidised and reduced acidic cytochrome c_3 from *Desulfovibrio africanus*, J. Mol. Biol., 290 (1999) 881-902.

[101] A. Gennaris, B. Ezraty, C. Henry, R. Agrebi, A. Vergnes, E. Oheix, J. Bos, P. Leverrier, L. Espinosa, J. Szewczyk, D. Vertommen, O. Iranzo, J.-F. Collet, F. Barras, Repairing oxidized proteins in the bacterial envelope using respiratory chain electrons, Nature, 528 (2015) 409-412.

[102] J. Appel, S. Phunpruch, K. Steinmüller, R. Schulz, The bidirectional hydrogenase of *Synechocystis* sp. PCC 6803 works as an electron valve during photosynthesis, Arch. Microbiol., 173 (2000) 333-338.

[103] M.V. Coppi, The hydrogenases of *Geobacter sulfurreducens*: A comparative genomic perspective, Microbiology, 151 (2005) 1239-1254.

[104] Y. Qiu, B.-K. Cho, Y.S. Park, D. Lovley, B.Ø. Palsson, K. Zengler, Structural and operational complexity of the *Geobacter sulfurreducens* genome, Genome Res., 20 (2010) 1304-1311.

[105] A. Belbelazi, R. Neish, M. Carr, J.C. Mottram, M.L. Ginger, Divergent cytochrome *c* maturation system in kinetoplastid protists, mBio, 12 (2021) e00166.

[106] M.J. Edwards, D.J. Richardson, C.M. Paquete, T.A. Clarke, Role of multiheme cytochromes involved in extracellular anaerobic respiration in bacteria, Protein Sci., 29 (2020) 830-842.

[107] S.J. Ferguson, J.M. Stevens, J.W.A. Allen, I.B. Robertson, Cytochrome *c* assembly: A tale of ever increasing variation and mystery?, Biochim. Biophys. Acta - Bioenergetics, 1777 (2008) 980-984.

[108] J.M. Stevens, O. Daltrop, J.W.A. Allen, S.J. Ferguson, *c*-type cytochrome formation: Chemical and biological enigmas, Accounts Chem. Res., 37 (2004) 999-1007.

[109] J.W.A. Allen, M.L. Ginger, S.J. Ferguson, Maturation of the unusual single-cysteine (XXXCH) mitochondrial *c*-type cytochromes found in trypanosomatids must occur through a novel biogenesis pathway, Biochem. J., 383 (2004) 537-542.

[110] V. Fülöp, K.A. Sam, S.J. Ferguson, M.L. Ginger, J.W.A. Allen, Structure of a trypanosomatid mitochondrial cytochrome *c* with heme attached via only one thioether bond and implications for the substrate recognition requirements of heme lyase, FEBS J., 276 (2009) 2822-2832.

[111] C. Ferousi, S. Lindhoud, F. Baymann, E.R. Hester, J. Reimann, B. Kartal, Discovery of a functional, contracted heme-binding motif within a multiheme cytochrome, J. Biol. Chem., 294 (2019) 16953-16965.

[112] R. Pisa, T. Stein, R. Eichler, R. Gross, J. Simon, The *nrfI* gene is essential for the attachment of the active site haem group of *Wolinella succinogenes* cytochrome *c* nitrite reductase, Mol. Microbiol., 43 (2002) 763-770.

[113] D. Aragão, C. Frazão, L. Sieker, G.M. Sheldrick, J. LeGall, M.A. Carrondo, Structure of dimeric cytochrome c_3 from *Desulfovibrio gigas* at 1.2 Å resolution, Acta Crystallogr. D, 59 (2003) 644-653.

[114] P. Simões, P.M. Matias, J. Morais, K. Wilson, Z. Dauter, M.A. Carrondo, Refinement of the three-dimensional structures of cytochrome c_3 from *Desulfovibrio vulgaris* Hildenborough at 1.67 Å resolution and from *Desulfovibrio desulfuricans* ATCC 27774 at 1.6 Å resolution, Inorg. Chim. Acta, 273 (1998) 213-224.

[115] Y. Gu, M.J. Guberman-Pfeffer, V. Srikanth, C. Shen, F. Giska, K. Gupta, Y. Londer, F.A. Samatey, V.S. Batista, N.S. Malvankar, Structure of *Geobacter* cytochrome OmcZ identifies mechanism of nanowire assembly and conductivity, Nat. Microbiol., 8 (2023) 284-298.

[116] F. Wang, C.H. Chan, V. Suciu, K. Mustafa, M. Ammend, D. Si, A.I. Hochbaum, E.H. Egelman, D.R. Bond, Structure of *Geobacter* OmcZ filaments suggests extracellular cytochrome polymers evolved independently multiple times, eLife, 11 (2022) e81551.

[117] B. Hermann, M. Kern, L. La Pietra, J. Simon, O. Einsle, The octahaem MccA is a haem *c*-copper sulfite reductase, Nature, 520 (2015) 706-709.

[118] R.S. Hartshorne, M. Kern, B. Meyer, T.A. Clarke, M. Karas, D.J. Richardson, J. Simon, A dedicated haem lyase is required for the maturation of a novel bacterial cytochrome *c* with unconventional covalent haem binding, Mol. Microbiol., 64 (2007) 1049-1060.

[119] J.W.A. Allen, E.B. Sawyer, M.L. Ginger, P.D. Barker, S.J. Ferguson, Variant *c*-type cytochromes as probes of the substrate specificity of the *E. coli* cytochrome *c* maturation (Ccm) apparatus, Biochem. J., 419 (2009) 177-186.

[120] R.-L. He, J. Wu, Z.-H. Cheng, H.-H. Li, J.-Q. Liu, D.-F. Liu, W.-W. Li, Biomolecular insights into extracellular pollutant reduction pathways of *Geobacter sulfurreducens* using a base editor system, Environ. Sci. Technol., 56 (2022) 12247-12256.

[121] M.I. Jahan, P. Juengwiwattanakitti, Y. Izu, R. Tobe, T. Imai, H. Mihara, Selenite uptake by outer membrane porin ExtI and its involvement in the subcellular localization of rhodanese-like lipoprotein ExtH in *Geobacter sulfurreducens*, Biochem. Biophys. Res. Commun., 516 (2019) 474-479.

[122] M.I. Jahan, R. Tobe, H. Mihara, Characterization of a novel porin-like protein, ExtI, from *Geobacter sulfurreducens* and its implication in the reduction of selenite and tellurite, Int. J. Mol. Sci., 19 (2018) 809.

[123] E. Howley, R. Krajmalnik-Brown, C.I. Torres, Cytochrome gene expression shifts in *Geobacter sulfurreducens* to maximize energy conservation in response to changes in redox conditions, Biosens. Bioelectron., 237 (2023) 115524.

[124] R. Cipollone, P. Ascenzi, E. Frangipani, P. Visca, Cyanide detoxification by recombinant bacterial rhodanese, Chemosphere, 63 (2006) 942-949.

[125] L. Rydz, M. Wróbel, H. Jurkowska, Sulfur administration in Fe-S cluster homeostasis, Antioxidants, 10 (2021) 1738.

[126] C. Aussignargues, M.-C. Giuliani, P. Infossi, E. Lojou, M. Guiral, M.-T. Giudici-Orticoni, M. Ilbert, Rhodanese functions as sulfur supplier for key enzymes in sulfur energy metabolism, J. Biol. Chem., 287 (2012) 19936-19948.
[127] P. Ramírez, H. Toledo, N. Guiliani, C.A. Jerez, An exported rhodanese-like protein is induced during growth of *Acidithiobacillus ferrooxidans* in metal sulfides and different sulfur compounds, Appl. Environ. Microb., 68 (2002) 1837-1845.
[128] A. Smirnov, C. Comte, A.-M. Mager-Heckel, V. Addis, I.A. Krasheninnikov, R.P. Martin, N. Entelis, I. Tarassov, Mitochondrial enzyme rhodanese is essential for 5 S ribosomal RNA import into human mitochondria, J. Biol. Chem., 285 (2010) 30792-30803.
[129] R. Alsohaibani, A.-L. Claudel, R. Perchat-Varlet, S. Boutserin, F. Talfournier, S. Boschi-Muller, B. Selles, Rhodanese-fold containing proteins in humans: Not just key players in sulfur trafficking, Antioxidants, 12 (2023) 843.
[130] R. Cipollone, P. Ascenzi, P. Visca, Common themes and variations in the rhodanese superfamily, IUBMB Life, 59 (2007) 51-59.
[131] D. Bordo, D. Deriu, R. Colnaghi, A. Carpen, S. Pagani, M. Bolognesi, The crystal structure of a sulfurtransferase from *Azotobacter vinelandii* highlights the evolutionary relationship between the rhodanese and phosphatase enzyme families, J. Mol. Biol., 298 (2000) 691-704.
[132] F. Gliubich, M. Gazerro, G. Zanotti, S. Delbono, G. Bombieri, R. Berni, Active site structural features for chemically modified forms of rhodanese, J. Biol. Chem., 271 (1996) 21054-21061.
[133] L. Chatwell, V. Illarionova, B. Illarionov, W. Eisenreich, R. Huber, A. Skerra, A. Bacher, M. Fischer, Structure of lumazine protein, an optical transponder of luminescent bacteria, J. Mol. Biol., 382 (2008) 44-55.
[134] M.I. Serer, H.R. Bonomi, B.G. Guimaraes, R.C. Rossi, F.A. Goldbaum, S. Klinke, Crystallographic and kinetic study of riboflavin synthase from *Brucella abortus*, a chemotherapeutic target with an enhanced intrinsic flexibility, Acta Crystallogr. D, 70 (2014) 1419-1434.
[135] M. Fischer, A.-K. Schott, W. Römisch, A. Ramsperger, M. Augustin, A. Fidler, A. Bacher, G. Richter, R. Huber, W. Eisenreich, Evolution of vitamin B_2 biosynthesis: A novel class of riboflavin synthase in Archaea, J. Mol. Biol., 343 (2004) 267-278.
[136] S. Gerhardt, A.-K. Schott, N. Kairies, M. Cushman, B. Illarionov, W. Eisenreich, A. Bacher, R. Huber, S. Steinbacher, M. Fischer, Studies on the reaction mechanism of riboflavin synthase: X-ray crystal structure of a complex with 6-carboxyethyl-7-oxo-8-ribityllumazine, Structure, 10 (2002) 1371-1381.
[137] B. Illarionov, W. Eisenreich, M. Wirth, C. Yong Lee, Y. Eun Woo, A. Bacher, M. Fischer, Lumazine proteins from photobacteria: Localization of the single ligand binding site to the N-terminal domain, Biol. Chem., 388 (2007) 1313-1323.
[138] D.-I. Liao, Z. Wawrzak, J.C. Calabrese, P.V. Viitanen, D.B. Jordan, Crystal structure of riboflavin synthase, Structure, 9 (2001) 399-408.
[139] M. Fujita, H. Mihara, S. Goto, N. Esaki, M. Kanehisa, Mining prokaryotic genomes for unknown amino acids: A stop-codon-based approach, BMC Bioinformatics, 8 (2007) 225.

[140] J.-H. Chen, L.-J. Yu, A. Boussac, Z.-Y. Wang-Otomo, T. Kuang, J.-R. Shen, Properties and structure of a low-potential, penta-heme cytochrome c_{552} from a thermophilic purple sulfur photosynthetic bacterium *Thermochromatium tepidum*, Photosynth. Res., 139 (2019) 281-293.

[141] P. Bork, L. Holm, C. Sander, The immunoglobulin fold: Structural classification, sequence patterns and common core, J. Mol. Biol., 242 (1994) 309-320.

[142] T. Brümmendorf, F.G. Rathjen, Cell adhesion molecules 1: Immunoglobulin superfamily, Protein Profile, 2 (1995) 963-1108.

[143] A.F. Williams, A.N. Barclay, The immunoglobulin superfamily – Domains for cell surface recognition, Annu. Rev. Immunol., 6 (1988) 381-405.

[144] L. Qiu, S. Dhe-Paganon, Oligomeric structure of the MALT1 tandem Ig-like domains, PLoS One, 6 (2011) e23220.

[145] T. Goldberg, M. Hecht, T. Hamp, T. Karl, G. Yachdav, N. Ahmed, U. Altermann, P. Angerer, S. Ansorge, K. Balasz, M. Bernhofer, A. Betz, L. Cizmadija, K.T. Do, J. Gerke, R. Greil, V. Joerdens, M. Hastreiter, K. Hembach, M. Herzog, M. Kalemanov, M. Kluge, A. Meier, H. Nasir, U. Neumaier, V. Prade, J. Reeb, A. Sorokoumov, I. Troshani, S. Vorberg, S. Waldraff, J. Zierer, H. Nielsen, B. Rost, LocTree3 prediction of localization, Nucleic Acids Res., 42 (2014) 350-355.

[146] F. Teufel, J.J. Almagro Armenteros, A.R. Johansen, M.H. Gíslason, S.I. Pihl, K.D. Tsirigos, O. Winther, S. Brunak, G. von Heijne, H. Nielsen, SignalP 6.0 predicts all five types of signal peptides using protein language models, Nat. Biotechnol., 40 (2022) 1023-1025.

[147] E. Gasteiger, C. Hoogland, A. Gattiker, S.e. Duvaud, M.R. Wilkins, R.D. Appel, A. Bairoch, Protein identification and analysis tools on the ExPASy server, The proteomics protocols handbook, Humana Press, Totowa, New Jersey, 2005, 571-607.

[148] P.R. Pokkuluri, Y.Y. Londer, N.E. Duke, J. Erickson, M. Pessanha, C.A. Salgueiro, M. Schiffer, Structure of a novel c_7-type three-heme cytochrome domain from a multidomain cytochrome *c* polymer, Protein Sci., 13 (2004) 1684-1692.

[149] N.R. Movva, K. Nakamura, M. Inouye, Amino acid sequence of the signal peptide of ompA protein, a major outer membrane protein of *Escherichia coli*, J. Biol. Chem., 255 (1980) 27-29.

[150] S.R. Bond, C.C. Naus, RF-Cloning.org: An online tool for the design of restriction-free cloning projects, Nucleic Acids Res., 40 (2012) 209-213.

[151] F. van den Ent, J. Löwe, RF cloning: A restriction-free method for inserting target genes into plasmids, J. Biochem. Biophys. Meth., 67 (2006) 67-74.

[152] A. Hardjasa, M. Ling, K. Ma, H. Yu, Investigating the effects of DMSO on PCR fidelity using a restriction digest-based method, J. Exp. Microbiol. Immunol., 14 (2010) 161-164.

[153] N. Chester, D.R. Marshak, Dimethyl sulfoxide-mediated primer T_m reduction: A method for analyzing the role of renaturation temperature in the polymerase chain reaction, Anal. Biochem., 209 (1993) 284-290.

[154] T.C. Lorenz, Polymerase chain reaction: Basic protocol plus troubleshooting and optimization strategies, Jove, 63 (2012) e3998.

[155] T. Czerny, High primer concentration improves PCR amplification from random pools, Nucleic Acids Res., 24 (1996) 985-986.
[156] D.J. Korbie, J.S. Mattick, Touchdown PCR for increased specificity and sensitivity in PCR amplification, Nat. Protoc., 3 (2008) 1452-1456.
[157] B. Maertens, A. Spriestersbach, J. Kubicek, F. Schäfer, Chapter five – Strep-tagged protein purification, Methods in enzymology, Academic Press, 2015, 53-69.
[158] S. Raran-Kurussi, S. Cherry, D. Zhang, D.S. Waugh, Removal of affinity tags with TEV protease, Heterologous gene expression in *E. coli*: Methods and protocols, Springer, New York, 2017, 221-230.
[159] M. van Rosmalen, M. Krom, M. Merkx, Tuning the flexibility of glycine-serine linkers to allow rational design of multidomain proteins, Biochemistry, 56 (2017) 6565-6574.
[160] L.K. Wang, C.D. Lima, S. Shuman, Structure and mechanism of T4 polynucleotide kinase: An RNA repair enzyme, EMBO J., 21 (2002) 3873-3880.
[161] R. Rossi, A. Montecucco, G. Ciarrocchi, G. Biamonti, Functional characterization of the T4 DNA ligase: A new insight into the mechanism of action, Nucleic Acids Res., 25 (1997) 2106-2113.
[162] W. Siwek, H. Czapinska, M. Bochtler, J.M. Bujnicki, K. Skowronek, Crystal structure and mechanism of action of the N6-methyladenine-dependent type IIM restriction endonuclease R.DpnI, Nucleic Acids Res., 40 (2012) 7563-7572.
[163] L. Thöny-Meyer, Biogenesis of respiratory cytochromes in bacteria, Microbiol. Mol. Biol. Rev., 61 (1997) 337-376.
[164] G.L. Rosano, E.A. Ceccarelli, Recombinant protein expression in *Escherichia coli*: Advances and challenges, Front. Microbiol., 5 (2014) 172.
[165] S. Yoshizawa, A. Böck, The many levels of control on bacterial selenoprotein synthesis, Biochim. Biophys. Acta - Gen. Subjects, 1790 (2009) 1404-1414.
[166] V.M. Labunskyy, D.L. Hatfield, V.N. Gladyshev, Selenoproteins: Molecular pathways and physiological roles, Physiol. Rev., 94 (2014) 739-777.
[167] Q. Cheng, E.S.J. Arnér, Overexpression of recombinant selenoproteins in *E. coli*, Selenoproteins: Methods and protocols, Springer, New York, 2018, 231-240.
[168] C.Z. Chung, C. Miller, D. Söll, N. Krahn, Introducing selenocysteine into recombinant proteins in *Escherichia coli*, Curr. Protoc., 1 (2021) e54.
[169] S.P. Lei, H.C. Lin, S.S. Wang, J. Callaway, G. Wilcox, Characterization of the *Erwinia carotovora pelB* gene and its product pectate lyase, J. Bacteriol., 169 (1987) 4379-4383.
[170] C. Sanders, H. Lill, Expression of prokaryotic and eukaryotic cytochromes *c* in *Escherichia coli*, Biochim. Biophys. Acta - Bioenergetics, 1459 (2000) 131-138.
[171] C. Iobbi-Nivol, H. Crooke, L. Griffiths, J. Grove, H. Hussain, J. Pommier, V. Mejean, J.A. Cole, A reassessment of the range of *c*-type cytochromes synthesized by *Escherichia coli* K-12, FEMS Microbiol. Lett., 119 (1994) 89-94.
[172] L. Thöny-Meyer, F. Fischer, P. Kunzler, D. Ritz, H. Hennecke, *Escherichia coli* genes required for cytochrome *c* maturation, J. Bacteriol., 177 (1995) 4321-4326.

[173] P.J. Shilling, K. Mirzadeh, A.J. Cumming, M. Widesheim, Z. Köck, D.O. Daley, Improved designs for pET expression plasmids increase protein production yield in *Escherichia coli*, Commun. Biol., 3 (2020) 214.
[174] G.L. Rosano, E.S. Morales, E.A. Ceccarelli, New tools for recombinant protein production in *Escherichia coli*: A 5-year update, Protein Sci., 28 (2019) 1412-1422.
[175] K. Ozawa, F. Yasukawa, Y. Fujiwara, H. Akutsu, A simple, rapid, and highly efficient gene expression system for multiheme cytochromes *c*, Biosci. Biotech. Bioch., 65 (2001) 185-189.
[176] L. Shi, J.T. Lin, L.M. Markillie, T.C. Squier, B.S. Hooker, Overexpression of multi-heme *c*-type cytochromes, Biotechniques, 38 (2005) 297-299.
[177] T.J. Lane, Protein structure prediction has reached the single-structure frontier, Nat. Methods, 20 (2023) 170-173.
[178] D.V. Laurents, AlphaFold2 and NMR spectroscopy: Partners to understand protein structure, dynamics and function, Front. Mol. Biosci., 9 (2022) 906437.
[179] N.J. Fowler, M.P. Williamson, The accuracy of protein structures in solution determined by AlphaFold and NMR, Structure, 30 (2022) 925-933.
[180] R. Tejero, Y.J. Huang, T.A. Ramelot, G.T. Montelione, AlphaFold models of small proteins rival the accuracy of solution NMR structures, Front. Mol. Biosci., 9 (2022) 877000.
[181] D. Chakravarty, L.L. Porter, AlphaFold2 fails to predict protein fold switching, Protein Sci., 31 (2022) e4353.
[182] J. Martin, When AlphaFold2 predictions go wrong for protein-protein complexes, is there something to be learnt?, Q. Rev. Biophys., 55 (2022) e6.
[183] T.C. Terwilliger, D. Liebschner, T.I. Croll, C.J. Williams, A.J. McCoy, B.K. Poon, P.V. Afonine, R.D. Oeffner, J.S. Richardson, R.J. Read, P.D. Adams, AlphaFold predictions are valuable hypotheses and accelerate but do not replace experimental structure determination, Nat. Methods, 21 (2024) 110-116.
[184] G. Artem, L. Anna, G. Casper, O. Sergey, F.P. Nicholas, AF2BIND: Predicting ligand-binding sites using the pair representation of AlphaFold2, bioRxiv, (2023) 2023.2010.2015.562410.
[185] R. Krishna, J. Wang, W. Ahern, P. Sturmfels, P. Venkatesh, I. Kalvet, G.R. Lee, F.S. Morey-Burrows, I. Anishchenko, I.R. Humphreys, R. McHugh, D. Vafeados, X. Li, G.A. Sutherland, A. Hitchcock, C.N. Hunter, A. Kang, E. Brackenbrough, A.K. Bera, M. Baek, F. DiMaio, D. Baker, Generalized biomolecular modeling and design with RoseTTAFold All-Atom, Science, 384 (2024) 2528.
[186] T. Sanavia, G. Birolo, L. Montanucci, P. Turina, E. Capriotti, P. Fariselli, Limitations and challenges in protein stability prediction upon genome variations: Towards future applications in precision medicine, Comput. Struct. Biotechnol. J., 18 (2020) 1968-1979.
[187] J. Zahiri, A. Emamjomeh, S. Bagheri, A. Ivazeh, G. Mahdevar, H. Sepasi Tehrani, M. Mirzaie, B.A. Fakheri, M. Mohammad-Noori, Protein complex prediction: A survey, Genomics, 112 (2020) 174-183.
[188] Y. Deng, M.L. Weaver, K.R. Hoke, E.V. Pletneva, A heme propionate staples the structure of cytochrome *c* for methionine ligation to the heme iron, Inorg. Chem., 58 (2019) 14085-14106.

[189] T. Hayashi, T. Matsuo, Y. Hitomi, K. Okawa, A. Suzuki, Y. Shiro, T. Iizuka, Y. Hisaeda, H. Ogoshi, Contribution of heme-propionate side chains to structure and function of myoglobin: Chemical approach by artificially created prosthetic groups, J. Inorg. Biochem., 91 (2002) 94-100.

[190] Y.Y. Londer, P.R. Pokkuluri, D.M. Tiede, M. Schiffer, Production and preliminary characterization of a recombinant triheme cytochrome c_7 from *Geobacter sulfurreducens* in *Escherichia coli*, Biochim. Biophys. Acta - Bioenergetics, 1554 (2002) 202-211.

[191] L.A. Zacharoff, D.J. Morrone, D.R. Bond, *Geobacter sulfurreducens* extracellular multiheme cytochrome PgcA facilitates respiration to Fe(III) oxides but not electrodes, Front. Microbiol., 8 (2017) 2481.

[192] L.L.C. Schrödinger, The PyMol molecular graphics system, Version 2.5.2, 2021.

[193] P.R. Pokkuluri, Y.Y. Londer, S.J. Wood, N.E. Duke, L. Morgado, C.A. Salgueiro, M. Schiffer, Outer membrane cytochrome *c*, OmcF, from *Geobacter sulfurreducens*: High structural similarity to an algal cytochrome c_6, Proteins, 74 (2009) 266-270.

[194] C. Elfmann, J. Stülke, PAE viewer: A webserver for the interactive visualization of the predicted aligned error for multimer structure predictions and crosslinks, Nucleic Acids Res., 51 (2023) W404-W410.

[195] E.F. Pettersen, T.D. Goddard, C.C. Huang, E.C. Meng, G.S. Couch, T.I. Croll, J.H. Morris, T.E. Ferrin, UCSF ChimeraX: Structure visualization for researchers, educators, and developers, Protein Sci., 30 (2021) 70-82.

[196] M. Mandel, A. Higa, Calcium-dependent bacteriophage DNA infection, J. Mol. Biol., 53 (1970) 159-162.

[197] K. Ohno, M. Tanaka, H. Ino, H. Suzuki, M. Tashiro, T. Ibi, K. Sahashi, A. Takahashi, T. Ozawa, Direct DNA sequencing from colony: Analysis of multiple deletions of mitochondrial genome, Biochim. Biophys. Acta, 1090 (1991) 9-16.

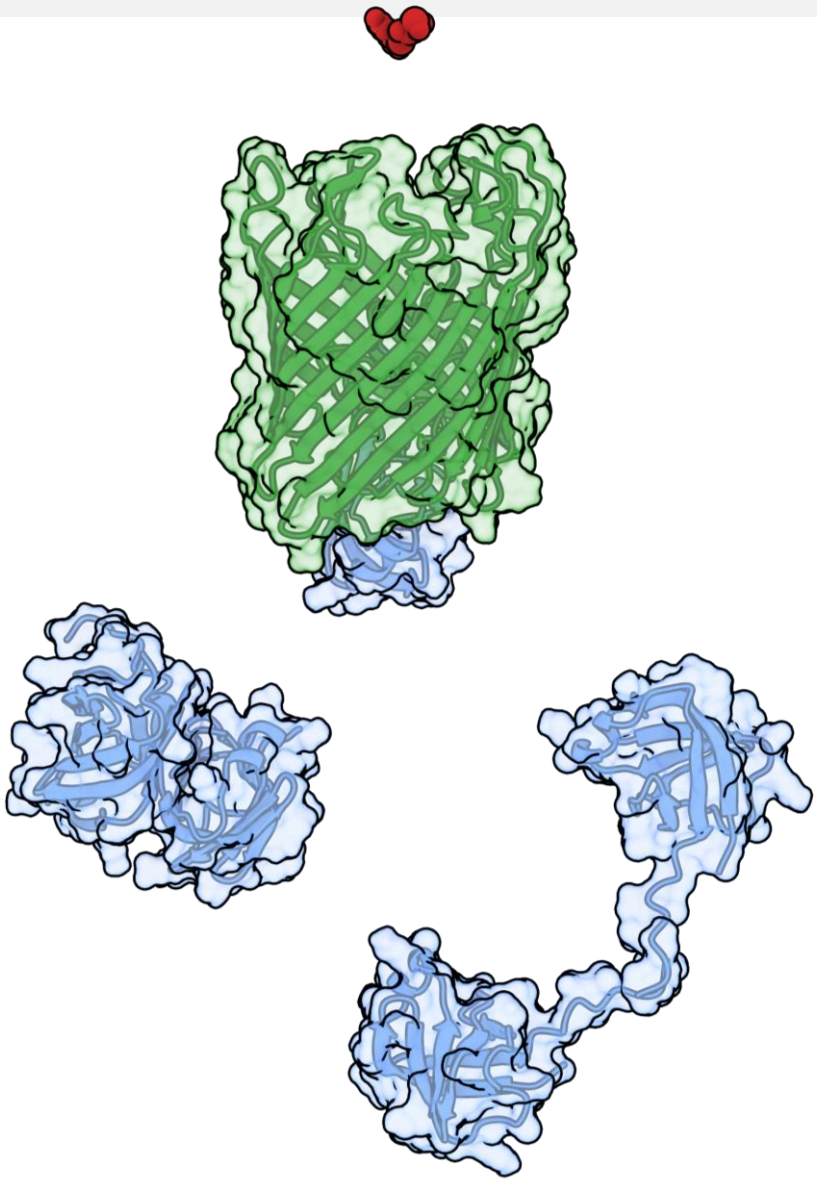
[198] L.I. Zon, D.M. Dorfman, S.H. Orkin, The polymerase chain reaction colony miniprep, Biotechniques, 7 (1989) 696-698.

[199] H. Schägger, Tricine-SDS-PAGE, Nat. Protoc., 1 (2006) 16-22.

[200] R.T. Francis, R.R. Becker, Specific indication of hemoproteins in polyacrylamide gels using a double-staining process, Anal. Biochem., 136 (1984) 509-514.

3

Dissecting a porin-cytochrome complex: Structural characterization of ExtJ



The findings of this chapter will form the basis for a manuscript to be submitted for publication:

T. M. Fernandes, D. M. Pedro, A. Viegas, C. Mota, A. J. M. Barbosa, L. Morgado and C. A. Salgueiro, Dissecting a porin-cytochrome complex: Structural characterization of ExtJ, In preparation, 2025

List of contents

Dissecting a porin-cytochrome complex: Structural characterization of ExtJ 153

3.1. Summary 155

3.2. Introduction 157

3.3. Results and discussion 158

3.3.1. Amino acid sequence analysis 158

3.3.2. ExtJ forms homodimers in vitro 159

3.3.3. Preliminary structural characterization of ExtJ 160

3.3.4. NMR solution structure and relaxation experiments of ExtJ 169

3.3.5. ExtJ homodimerization is driven solely by an intermolecular disulfide bridge 173

3.3.6. Narrowing down the function of ExtJ in *Geobacter* 183

3.4. Conclusions 191

3.5. Materials and methods 193

3.5.1. DNA manipulation 193

3.5.2. Protein sequence analysis and AlphaFold predictions 193

3.5.3. Protein expression and purification 194

3.5.4. NMR spectroscopy 195

3.5.5. CD spectroscopy 199

3.5.6. Molecular dynamics simulations 199

3.5.7. Microscale thermophoresis 200

3.5.8. SAXS experiments 201

3.6. References 202

3.1. Summary

ExtJ is an 8 kDa periplasmic protein with no predicted cofactors, that is part of the putative porin-cytochrome complex ExtHIJKL, which is crucial for Mn(IV), selenite and tellurite reduction, as well as for growth on anodes poised at intermediate redox potentials (–150 mV vs SHE).

In this chapter, we present a biochemical and structural depiction of ExtJ, which marks the first experimental characterization of a protein from a porin-cytochrome complex of *G. sulfurreducens*. ExtJ forms homodimers in vitro by establishing a disulfide bridge between the C-terminal cysteine residues of each monomer, as shown by size-exclusion chromatography analysis and site-directed mutagenesis studies.

NMR and circular dichroism studies showed that the protein contains five β-sheets within a β-barrel fold with a Greek-key topology, framed by an isolated α-helix and an elongated random coil element at the C-terminal of the protein. This folding is structurally conserved with those of domains found in riboflavin synthase and homologs. A series of biophysical experiments and computational methods, including microscale thermophoresis, small-angle X-ray scattering, AlphaFold model predictions and all-atom molecular dynamics simulations, were employed to unequivocally show that the ExtJ dimerization is solely promoted by the formation of the intermolecular disulfide bridge. ^{15}N NMR relaxation experiments show that each monomer of the homodimer is tumbling in solution almost independently, within an equivalent rotating frame.

A detailed analysis of the structure of ExtJ and those of the homologous flavin-binding proteins, together with their functional mechanism, show that ExtJ contains the required structural arrangement for the binding of molecules analogous to riboflavin. Nevertheless, NMR biomolecular interactions studies show that ExtJ does not bind flavin mononucleotide (FMN). Moreover, ExtJ does not interact with PpcA, the most abundant cytochrome of the periplasm of *G. sulfurreducens*, indicating that it does not function as a recognition or anchoring domain for the PpcA-family cytochromes, in the context of the ExtHIJKL porin-cytochrome complex. Alternatively, based on AlphaFold-Multimer predictions of the ExtHIJKL porin-cytochrome complex, we suggest that the C-terminal cysteine of ExtJ might establish a covalent disulfide bridge with a cysteine of the ExtI porin through a redox-linked mechanism. This mechanism involves the dissociation of the ExtJ homodimer and concomitant formation of a transient complex with ExtI, hindering the import or export of small molecules through this porin.

3.2. Introduction

The composition of the *extHIJKL* gene cluster, encoding a rhodanese protein (ExtH) together with a β-barrel porin (ExtI), a pentaheme *c*-type cytochrome likely functioning as a selenite reductase (ExtKL) and a riboflavin-like protein (ExtJ), makes it the most versatile and unique porin-cytochrome complex not only from *G. sulfurreducens*, but, to the best of our knowledge, among all those described in bacteria to date.

This complex is crucial for Mn(IV), selenite and tellurite reduction [1-4], as well as for growth on anodes poised at intermediate redox potential values (–150 mV vs SHE) [5]. This functional versatility is a direct consequence of the distinct nature of the proteins of this complex, which work together to perform various roles in the cell during EET. Therefore, to pinpoint the exact mechanisms that allow ExtHIJKL to perform such various roles, a long-term multidisciplinary approach will be required, in which the interactions between the different proteins of the complex will have to be studied and analyzed, both in vitro and in vivo. While part of this work has been conducted in vivo [1-5], there are no in vitro studies that focus on the study of the individual and stringent properties of each protein.

In this chapter, a structural and functional characterization of ExtJ is presented. This protein was selected amongst those of the ExtHIJKL complex due to several factors, including its cellular location and solubility, small molecular weight, and the absence of cofactors or complex amino acids, which together with the overall lack of functional information, make this protein an interesting target. For this, the *extJ* gene was amplified from genomic DNA and inserted into a pET-28a(+) expression vector containing an N-terminal His-tag with a thrombin cleavage site (Figure 3.1), using RF cloning (see sections 2.4.1 and 2.6.2).

Figure 3.1. Amino acid sequence of the pET-28a(+)-ExtJ construct. The N-terminal affinity tag is highlighted in light gray. The 6x His-tag and thrombin recognition site residues are highlighted in green and orange, respectively. The thrombin cleavage site is marked by the scissors. The C-terminal cysteine is highlighted in blue.

Using a combination of biophysical techniques, we show that ExtJ forms homodimers in vitro, promoted by the establishment of an intermolecular disulfide bridge. The protein's folding is structurally conserved with those of proteins that bind flavin-like compounds, hinting that ExtJ might be implicated in the binding of such compounds in the periplasm of the cell. Based on the structural conservation of ExtJ, interaction studies with putative ligands, and AlphaFold-Multimer simulations of the ExtHIJKL porin-cytochrome complex, different functional roles are proposed and discussed for this small periplasmic protein.

3.3. Results and discussion

3.3.1. Amino acid sequence analysis

As mentioned in Chapter 2, ExtJ is a 104 amino acid protein with a 23-residue signal peptide, resulting in a mature protein with 81 residues (Figures 3.1 and 3.2).

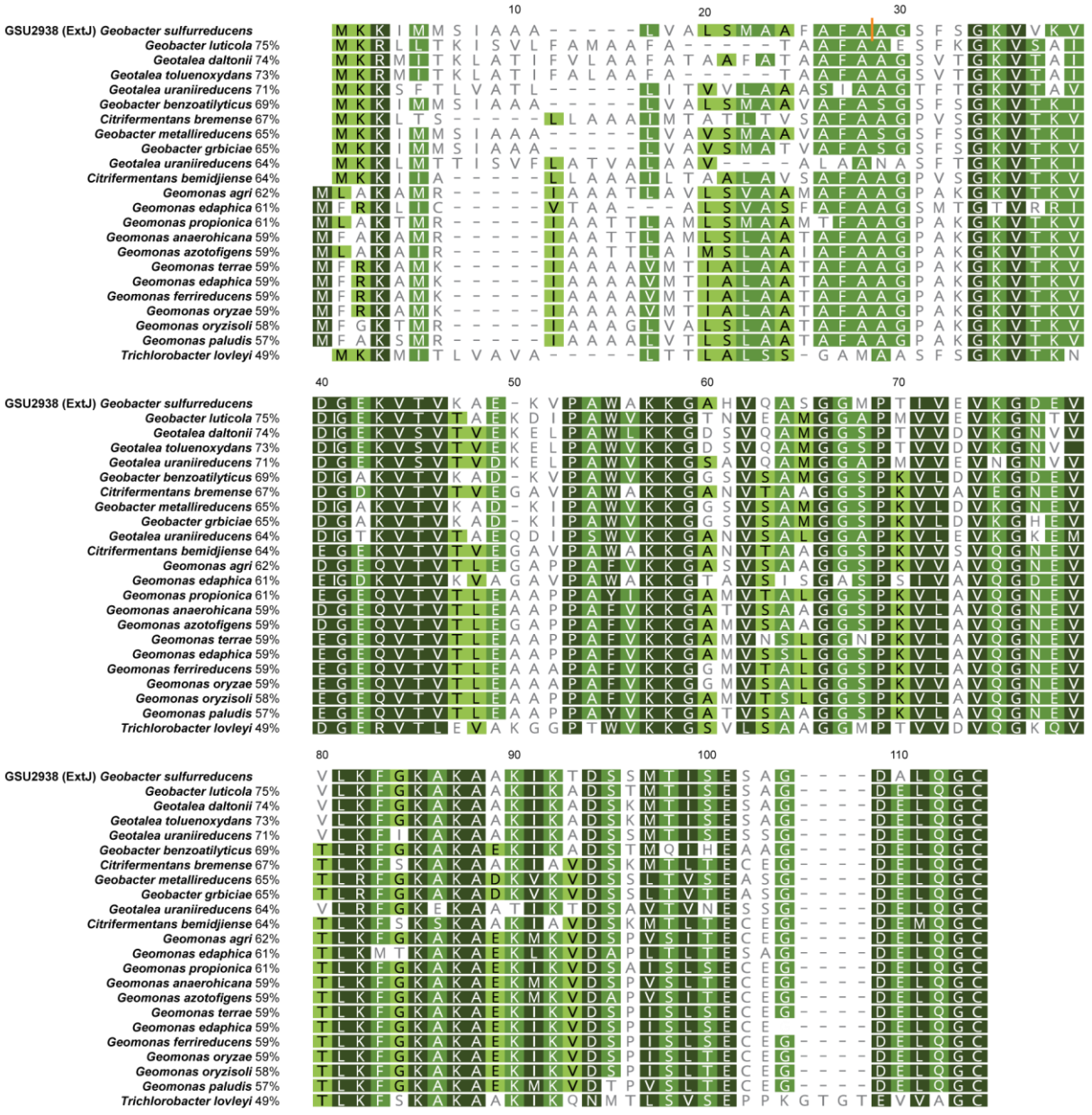


Figure 3.2. Amino acid sequence alignment of ExtJ with homolog sequences from different bacteria obtained from a Protein BLAST analysis. The sequences were aligned with Clustal Omega [6] and their pairwise identity (%) relative to the mature sequence is indicated. Residues are colored by sequence similarity (dark green 100%, green 80-100%, light green 60-80% and white <60% similarity). The orange line in the ExtJ sequence identifies the predicted signal peptide cleavage site.

According to the ExPASy ProtParam tool [7], this accounts for a total molecular mass of 8247 Da and a basic isoelectric point of 9.3, which can be mainly attributed to the high number of lysine residues. The protein subcellular localization prediction system LocTree 3.0 [8] expects that ExtJ is located in the periplasm.

A Protein BLAST [9] search with the RefSeq Select protein sequences database and the default blastp algorithm retrieves 23 sequences (out of a total of 26, from which the ones from multispecies were excluded) with moderate (35-60%) and high (>60%) sequence similarity (Figure 3.2). These homolog sequences were identified in species from several protobacteria genus, and the MSA shows that there are several conserved residues across all 23 homolog proteins (residues in dark green). To the best of our knowledge, none of these proteins has been characterized to date and, therefore, no inferences can be made about the putative function of ExtJ based on the information available for these homologous proteins.

However, upon search for structural homologs in the PDB database [10, 11] using the AlphaFold model of the protein and the Dali [12], FoldSeek [13] and PDBeFold [14] servers, multiple structures of domains within subunits of large ATPases/synthases or ribosomal assemblies, together with riboflavin synthase-like proteins, emerge as closely related structural homologs of ExtJ, despite only sharing 10 to 20% of sequence identity (see section 2.3.4).

3.3.2. ExtJ forms homodimers in vitro

To pursue a structural and biophysical characterization of ExtJ, the protein was heterologously expressed in *E. coli* BL21 Star (DE3) cells and purified by two chromatographic steps. The first step was a Ni-NTA affinity chromatography, which removed most of the contaminants, while the second step was a size-exclusion chromatography (SEC). The SEC elution profile of the protein fractions resulting from the Ni-NTA chromatography showed two intense peaks at 12 and 14 mL (Figure 3.3). The SDS-PAGE gel analysis of these peaks using a loading buffer without a reducing agent showed that both peaks contained the same migration profile – one band at 20 kDa and another at 10 kDa. The protein product of the ExtJ construct has a molecular mass of 10 kDa (Figure 3.1), according to the ExPASy ProtParam tool [7]. Furthermore, it contains a highly accessible C-terminal cysteine residue (Cys^{81}), likely suitable for the establishment of a putative intermolecular disulfide bridge (Figures 2.21 and 3.1). Consequently, it is reasonable to assume that ExtJ may be forming homodimers in vitro via Cys^{81}. This hypothesis was tested, in a preliminary stage, by running a SEC with the previously used conditions, but in the presence of a reducing agent (1 mM DTT). In this case (Figure 3.3), a single peak was observed, corresponding to pure ExtJ, as shown by the SDS-PAGE analysis and later confirmed by MALDI-TOF-MS. In a second stage, the codon encoding the putative residue responsible for the homodimer formation (Cys^{81}) was substituted by a stop codon.

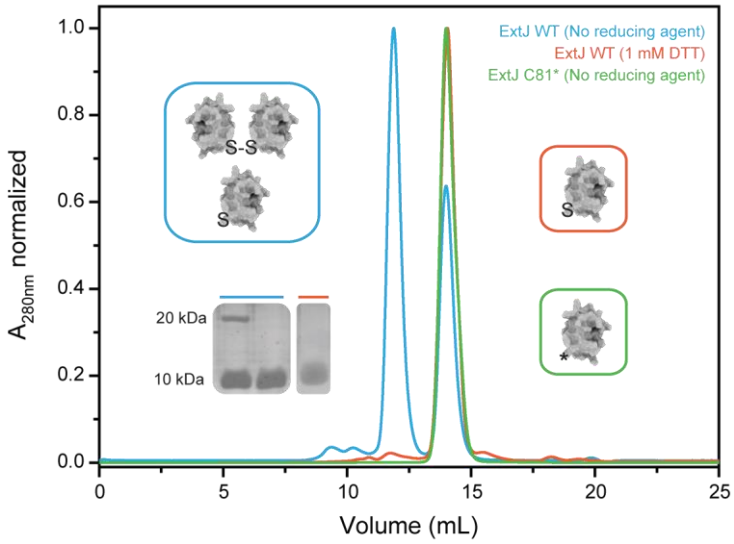


Figure 3.3. ExtJ homodimer formation. The SEC elution profiles of ExtJ WT in the absence and presence of a reducing agent (1 mM DTT) are represented in blue and orange, respectively. The SEC elution profile of ExtJ C81* in the absence of a reducing agent is represented in green. The SECs were run in a Superdex™ 75 Increase 10/300 GL molecular exclusion column (Cytiva), equilibrated with 100 mM sodium phosphate buffer pH 8. The colour scheme of the SEC elution profiles is maintained throughout the different elements of the figure. The SDS-PAGE gels shown on the inset were run with loading buffers without reducing agent. The "S" represents the sulfur atom of the Cys^{81} responsible for the homodimer formation. The "*" marks the mutation of Cys^{81} in the ExtJ protein.

The ExtJ C81* protein was purified without addition of reducing agents and its SEC elution profile (Figure 3.3) shows a single Gaussian peak corresponding to pure monomeric protein, confirmed by the SDS-PAGE analysis. Therefore, it was unequivocally confirmed that ExtJ WT forms homodimers in vitro by establishing a disulfide bridge between different protein molecules in solution via a highly accessible C-terminal cysteine residue.

Overall, the final protein expression yields of ExtJ C81* (35 mg of protein per liter of cell culture) were considerably higher than those obtained for ExtJ WT (10 mg of protein per liter of cell culture). This difference may be related with the prevention of homodimer formation in the ExtJ C81* protein, which provides higher stability to ExtJ both in vivo during protein expression and in vitro during protein purification. Following the purification of ExtJ, a structural characterization was conducted using NMR and circular dichroism (CD) spectroscopies.

3.3.3. Preliminary structural characterization of ExtJ

NMR is an excellent technique for attaining information about the structure and dynamics of proteins. The small molecular mass of ExtJ makes it an excellent target for NMR due to the fast rotational correlation time (τ_c) and consequently short T_2 relaxation time of the molecule. The 1D ^{1}H-NMR spectrum of a protein can provide preliminary relevant structural information. In fact, the spectrum of monomer ExtJ (Figure 3.4), acquired in the presence of DTT, shows that the protein is properly folded, as the resonances of the backbone amide region (7-10 ppm) are well dispersed.

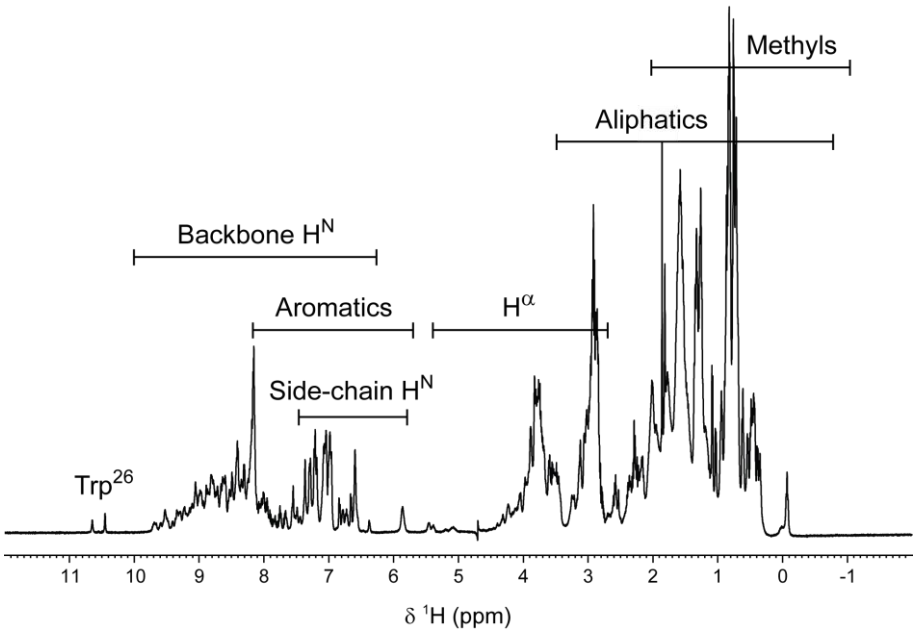


Figure 3.4. 1D ^{1}H-NMR spectrum of ExtJ WT. The spectrum was acquired at 25 ^{o}C with a 1 mM protein sample, prepared in 32 mM sodium phosphate buffer pH 6 (100 mM of final ionic strength) with 1 mM DTT. The typical spectral regions for the different amino acid backbone and side-chain signals are highlighted.

The only exception is the intense signal at around 8.2 ppm, which likely corresponds to the amino acid residues of the N-terminal affinity tag, that may be flexible and disordered. Furthermore, the spectrum also shows the existence of resonances typical of specific residues, such as Trp^{26}, located at around 10.5 ppm (Figure 3.4). Nevertheless, this information is insufficient to provide a detailed structural picture of the protein. Therefore, the protein was isotopically labeled in ^{13}C and ^{15}N, and the assignment of its backbone resonances was obtained using NMR multidimensional experiments, including 2D ^{1}H, ^{15}N-HSQC and 3D HNCO [15], 3D HN(CA)CO [16], 3D HNHA [17], 3D HN(CO)CACB [18] and 3D HNCACB [19] (Figure 3.5). All spectra were acquired at pH 6 to minimize solvent exchange of labile protons and maximize the number of resonances and information.

The 2D ^{1}H, ^{15}N-HSQC of ExtJ – which is often called the protein's fingerprint since it reflects the chemical nature and local environment of every amino acid residue, except for prolines, which do not possess a backbone amide proton – was used as a starting point for the assignment of the protein backbone resonances [20]. 3D HNCO and 3D HN(CA)CO spectra were employed as complementary experiments to sequentially identify the ^{13}CO of a given amino acid and its predecessor, while the 3D HNCACB and 3D HN(CO)CACB spectra were used to identify the $^{13}C_{\alpha}$ and $^{13}C_{\beta}$ of the corresponding residues (Figure 3.6). This information enabled the identification of most spin systems based on their typical chemical shifts [21-23], which were complemented with the $^{1}H_{\alpha}$ chemical shifts provided by the analysis of the 3D HNHA spectrum.

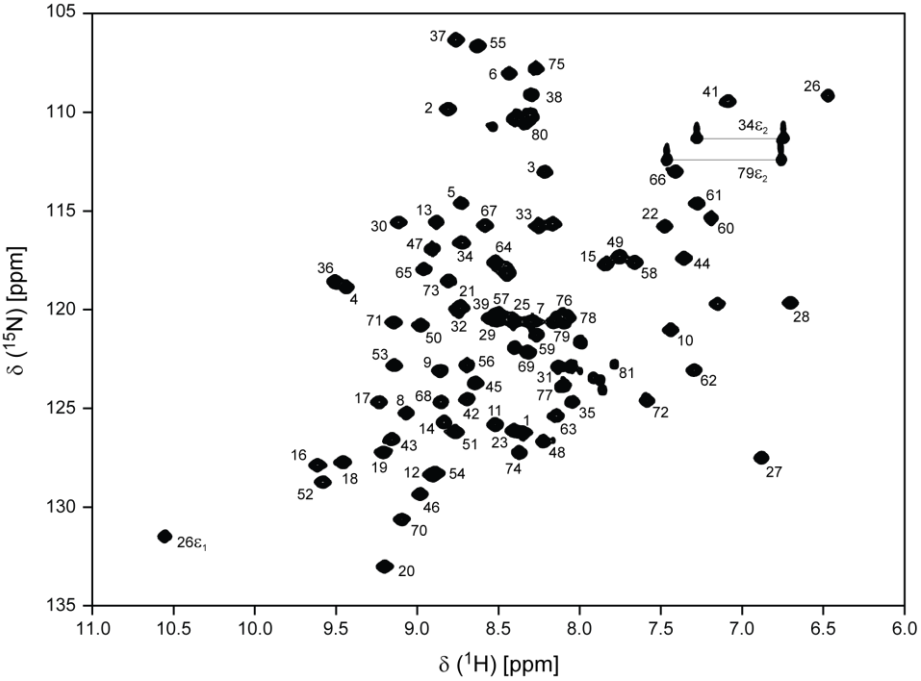


Figure 3.5. 2D ^{1}H, ^{15}N-HSQC spectrum of ExtJ WT. The numbers correspond to the amino acid residues of mature ExtJ, excluding the N-terminal affinity tag. The spectrum was acquired at 25 °C with 1 mM of protein prepared in 32 mM sodium phosphate buffer pH 6 (100 mM of final ionic strength) with 1 mM DTT.

The only resonances which were not assigned were those of the residues of the N-terminal affinity tag. The 2D ^{1}H, ^{15}N-HSQC spectrum of monomer ExtJ confirms the conclusions taken upon analysis of the 1D ^{1}H-NMR spectrum of the protein (Figure 3.5). In Figure 3.6A, representative strips of the 3D HN(CA)CO and 3D HNCACB spectra show the sequential assignment of the stretch of residues from Asp^{12} to Val^{16}, illustrating the procedure used for the complete assignment of the protein's backbone.

In a preliminary stage, prior to structure calculation, the chemical shifts of the ^{1}HN, ^{1}H$_\alpha$, ^{13}C$_\alpha$, ^{13}C$_\beta$, ^{13}CO and ^{15}N nuclei were further analyzed with the CSI 3.0 [24] and Talos+ [25] softwares to predict the secondary structural elements of ExtJ (Figure 3.6B). The results indicate that ExtJ is composed by five or six β-sheets and one α-helix. The relative position of these secondary structural elements differs slightly between the two predictions, however, globally, the two methods have similar results.

An independent analysis of the secondary ^{13}C$_\alpha$ and ^{13}C$_\beta$ chemical shifts, calculated relative to two distinct datasets of reference chemical shifts – the set of standard BMRB chemical shifts [22, 23] and a set of sequence corrected random coil chemical shifts [26] – yielded similar results (Figure 3.6B). Overall, these results correlate well with the AlphaFold model of ExtJ (Figure 2.21).

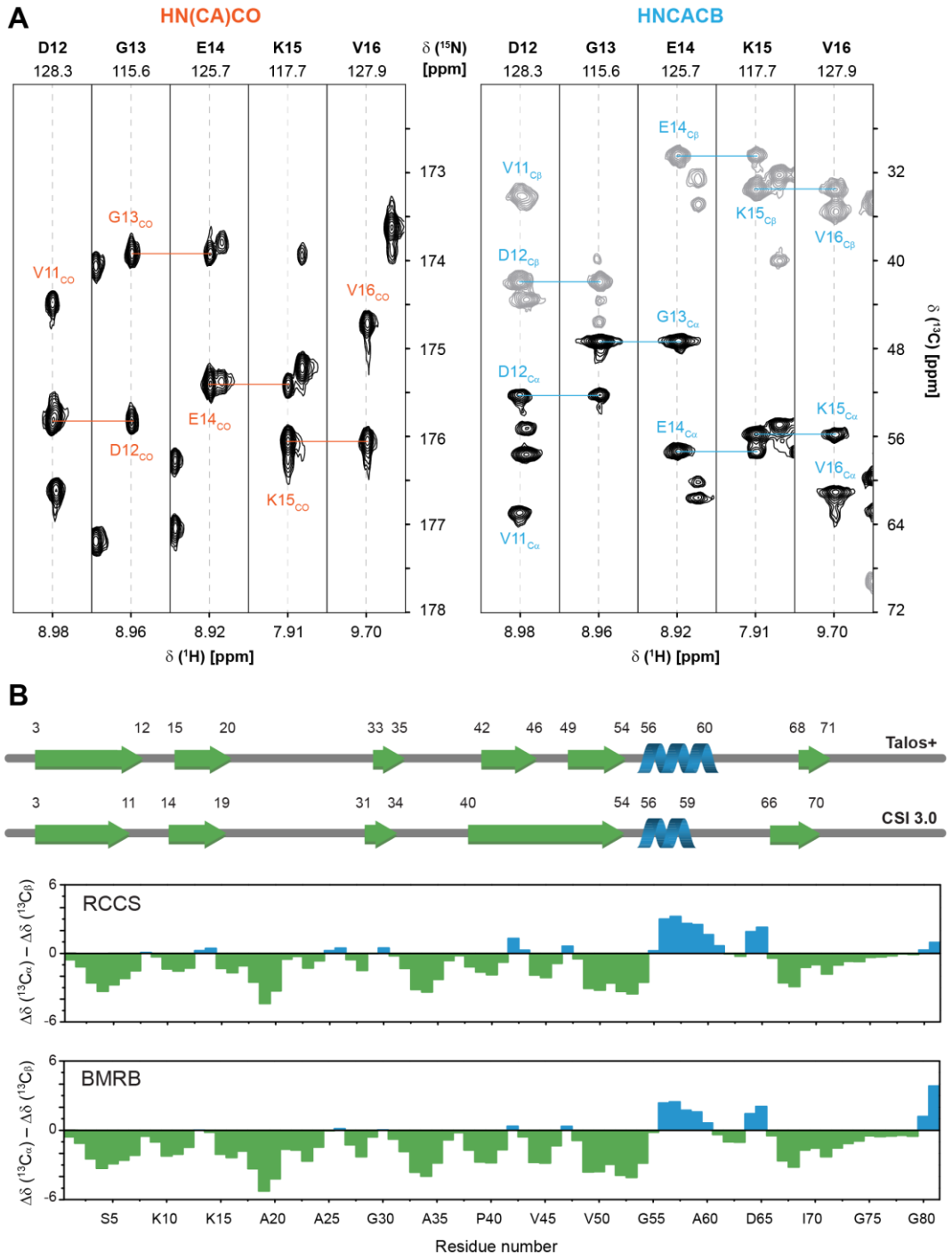


Figure 3.6. NMR sequential assignment and structural features of ExtJ. (A) Strips from the 3D HN(CA)CO and 3D HNCACB spectra show the sequential assignment of the ^{13}CO (orange), $^{13}C_{\alpha}$ and $^{13}C_{\beta}$ (blue) nuclei of the stretch of residues from Asp^{12} to Val^{16}. (B) Secondary structural prediction of ExtJ. The α-helix and β-sheet secondary structural elements are colored in blue and green, respectively. The pure chemical shift-based predictions of the CSI 3.0 [24] and Talos+ [25] softwares are shown on top. The secondary ^{13}C chemical shifts plotted against the amino acid residue numbers of ExtJ, calculated using either the set of standard BMRB chemical shifts [22, 23] or a set of sequence corrected random coil chemical shifts (RCCS) [26], are shown below. In both cases, a 1-2-1 weighting function for residues (i − 1) − i − (i + 1) was applied to the raw data to reduce noise and highlight regular secondary structure elements [27].

To better visualize the predictions derived from the NMR data, we resorted to Rosetta CS [28-30], which constructs chemical shift-based de novo protein models. The resulting model is highly similar to the one predicted by AlphaFold, with an overall RMSD of 0.9 Å (Figure 3.7).

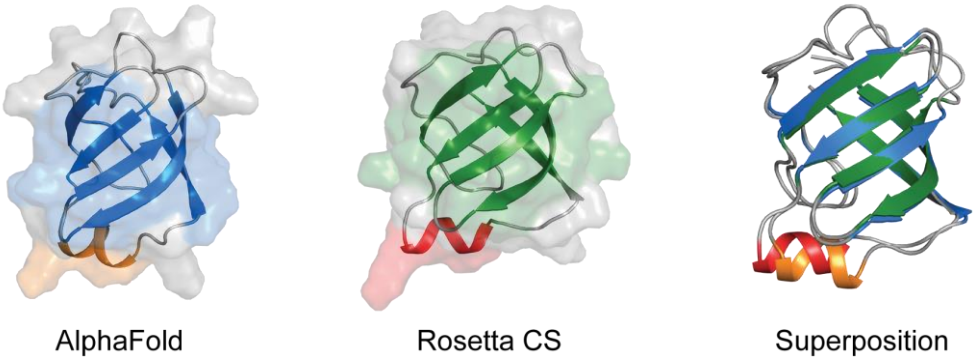


Figure 3.7. Structural models of ExtJ. The α-helix secondary structural elements are colored orange and red in the AlphaFold and Rosetta CS models, whereas the β-sheet secondary structural elements are colored blue and green, respectively. The superposition of both models, with a global RMSD of 0.9 Å, respects the same color coding. Residues 72-81 were trimmed of the Rosetta CS model to improve the convergence between the experimental chemical shifts and the structure calculation.

The slight differences found between these models and those made upon direct analysis of the backbone chemical shifts are likely related with external factors that affect the observed chemical shifts of the residues, which hinders their accurate secondary structure determination [31].

Following a preliminary structural analysis of the WT protein, the impacts of the C81* mutation and the 20 amino acid N-terminal affinity tag in the structure of ExtJ were evaluated. The mutated version of the protein conveniently results in higher expression yields while not requiring the use of reducing agents, which further simplifies its manipulation and preparation, making it a more suitable target for future studies. Therefore, ExtJ WT and ExtJ C81* were produced with no N-terminal affinity tag, after inclusion of a thrombin-based affinity tag cleavage step and an additional Ni-NTA affinity chromatography between the first and second purification steps, as described in section 3.5.3. After that, four different forms of the protein (ExtJ WT and ExtJ C81*, with (T) and without (U) the N-terminal affinity tag) were analyzed and compared by NMR and CD spectroscopies. 2D ^{1}H, ^{15}N-HSQC were acquired for the four versions of ExtJ and the signals of each amino acid residue were assigned (Figure 3.8).

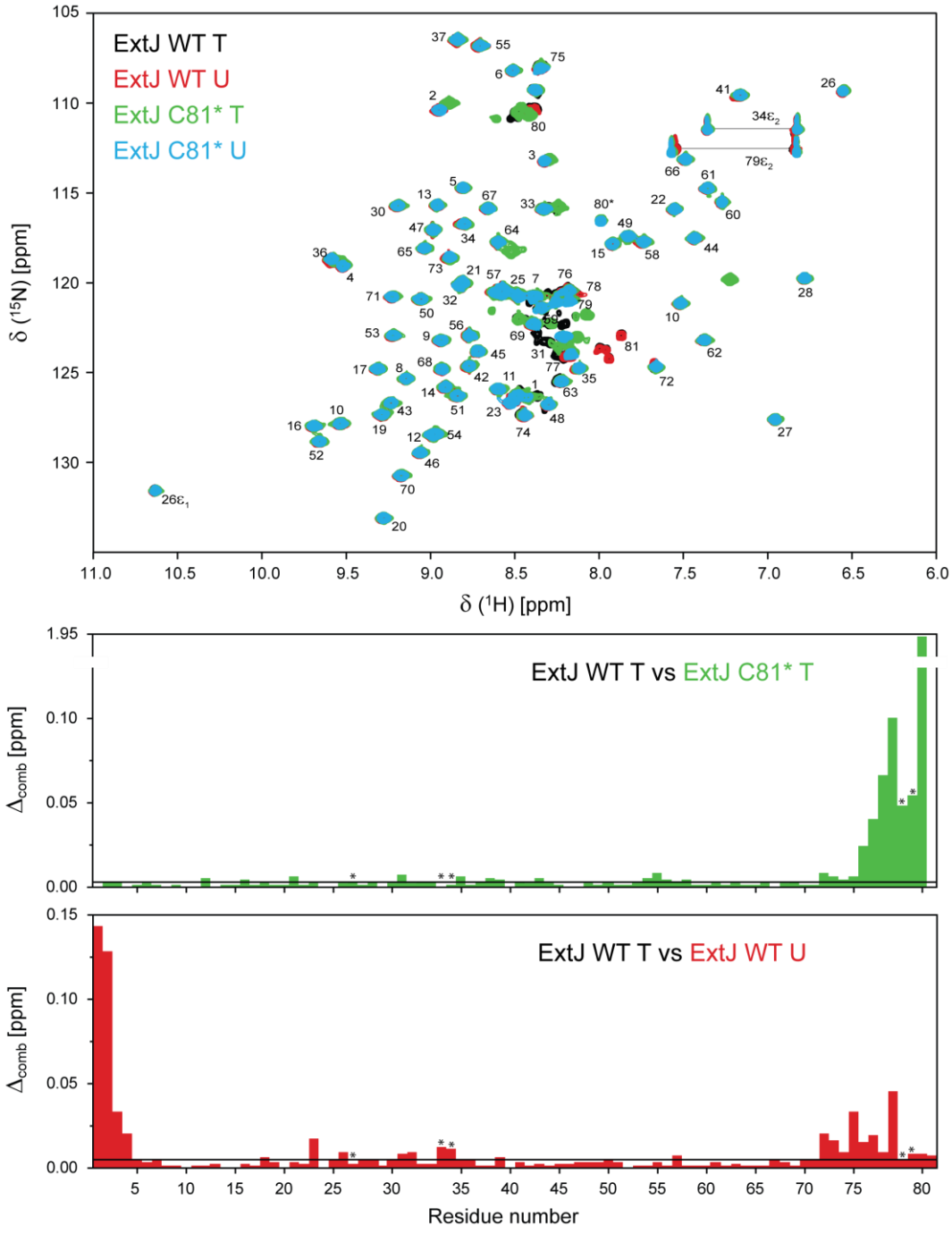


Figure 3.8. Structural impact of the C81* mutation and the N-terminal affinity tag. The 2D ^{1}H, ^{15}N-HSQC of the four versions of ExtJ are represented on top. The spectra of ExtJ WT with (T) and without (U) the N-terminal affinity tag are represented in black and red, respectively. The spectra of ExtJ C81* with (T) and without (U) the N-terminal affinity tag are represented in green and blue, respectively. The differences in the combined chemical shifts observed for evaluation of the effect of the C81* mutation and the N-terminal affinity tag on the ExtJ's global structure are represented as bar graphs in green and red, respectively. The black lines on these graphics represent the cut-off values calculated for each case, according to [32]. The residues marked by an "*" represent the ^{1}H-^{15}N W^{26}, Q^{34} and Q^{79} side-chain signals.

Most of the resonances were assigned upon visual inspection relative to the 2D ^{1}H, ^{15}N-HSQC spectra of ExtJ WT T, due to the small or inexistent chemical shift difference between the proteins, while the remaining resonances were unequivocally attributed upon analysis of 3D ^{1}H, ^{15}N-TOCSY-HSQC and 3D ^{1}H, ^{15}N-NOESY-HSQC spectra. The combined chemical shifts of each protein were compared to those of ExtJ WT T and the results are shown on Figure 3.8. The mutation of the C-terminal Cys^{81} residue only affects the local region of the C-terminal without having an impact on the protein's global structure. The removal of the N-terminal affinity tag resulted in structural differences on the N-terminal of the protein, as expected, but surprisingly on the C-terminal, as well. This observation might be a consequence of a close spatial disposition between the N- and C-terminal regions. Nonetheless, overall, one can consider that the global structure of ExtJ is not affected by the removal of the N-terminal affinity tag.

CD is another excellent technique for rapidly evaluating the secondary structure features, folding, binding and thermodynamic properties of proteins, thus complementing the NMR data presented above. The ExtJ WT samples used for the CD studies were prepared with 1 mM β-mercaptoethanol, which was found to be sufficient to ensure that the protein was in its monomer form without compromising the CD signal significantly. The far-UV CD spectrum of ExtJ (Figure 3.9) mainly shows features of β-sheet secondary structural elements, with a typical maximum at 196 nm and a minimum at 215 nm. The CD spectra of the four versions of ExtJ at 25 °C were deconvoluted using BeStSel [33-35] in order to determine if there are significant differences in their percentages of secondary structural elements (Table 3.1). The deconvolution was performed using the data between 200 and 250 nm, thus discarding the values below 200 nm, which presented measured voltages above 600 V [36]. The results show that ExtJ accounts for more than 50% of folded secondary structural elements, while having around 40% of disordered elements. The results obtained by CD agree with those obtained by NMR and with the computational models of the protein, since according to the BeStSel deconvolution, ExtJ possesses above 30% of β-sheet secondary structural elements and 10-15% of α-helices and turns. The slight incoherencies found between the CD data and the remaining results can be linked with intrinsic technical limitations of the CD spectral deconvolution method, which has problems dealing with the different orientations (paralell or anti-paralell), number and size of β-sheets that result in different torsion angles of the peptide bond and necessarily different CD spectra [36].

To monitor the thermal stability of ExtJ, a characterization of the temperature-induced unfolding was carried out with all constructs. The spectra obtained (Figure 3.9) showed that in all cases, the secondary structural elements of ExtJ undergo significant changes during temperature variation. The ellipticity at 196 nm progressively decreased with temperature due to the loss of β-sheet secondary structural elements. Concomitantly, a new spectral minimum at 200 nm, typical of random coil elements, unveiled at temperatures above 75 °C. The decrease in ellipticity of the CD signals at 196 nm with temperature of the four versions of ExtJ were fitted to the model presented in section 3.5.5.

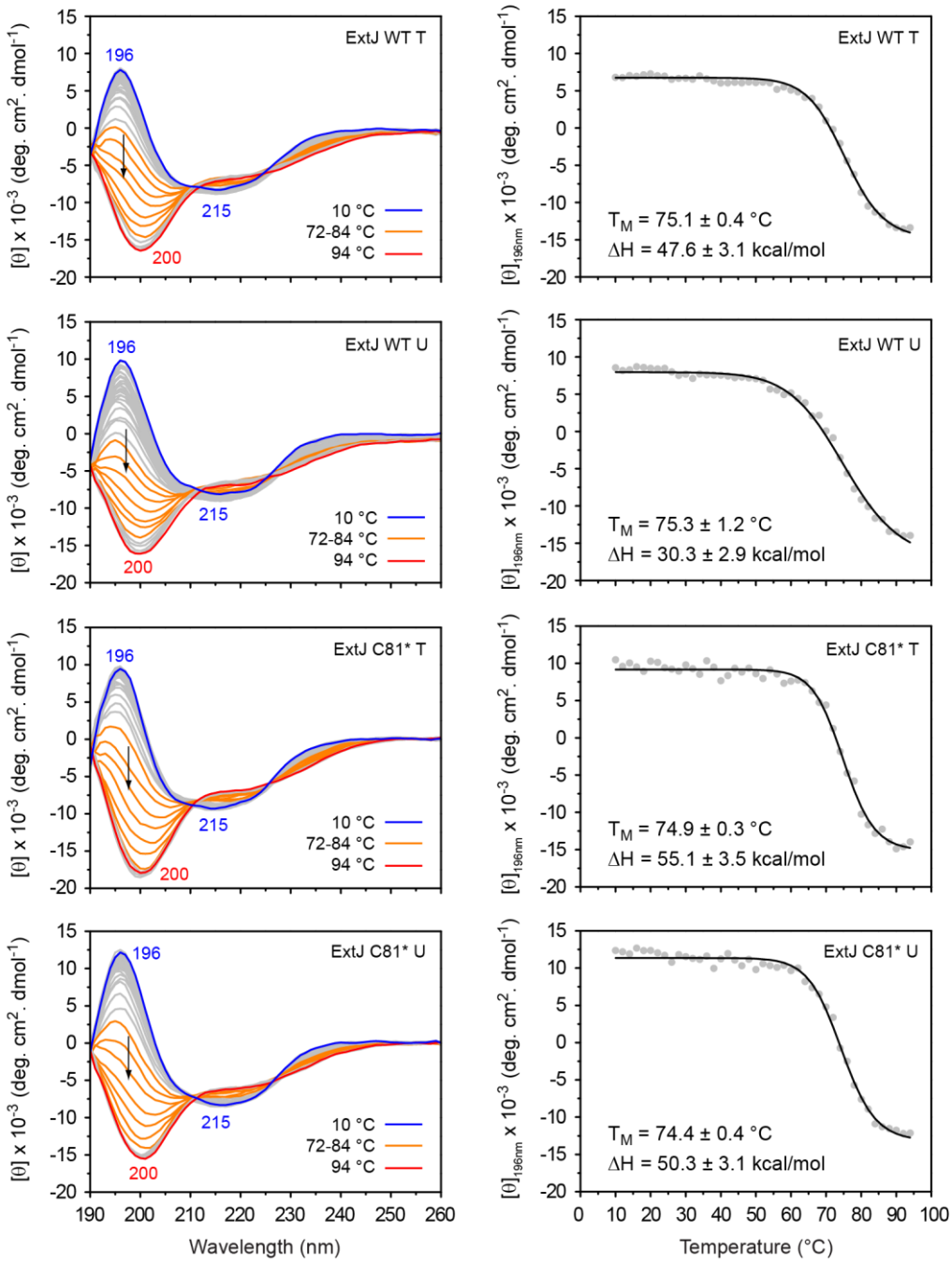


Figure 3.9. Far-UV CD spectral features of ExtJ. The left panels show the monitorization of the thermal stability of the different ExtJ versions. The spectra were acquired between 10 °C (blue) and 94 °C (red). The orange spectra (72-84 °C) highlight the spectral variation resulting from the conformational transition between β-sheet secondary structural elements to random coil elements, upon sample heating. On the right panels, the thermal unfolding profiles of the different ExtJ versions are shown. The mean residue ellipticity [θ] at 196 nm is represented as a function of temperature (spheres). The solid line represents the fitting of the model presented in section 3.5.5 to the experimental data.

ExtJ WT T has a midpoint of unfolding transition of 75.1 ± 0.4 ºC and an enthalpy of unfolding (ΔH) of 47.6 ± 3.1 kcal/mol. This ΔH value is in the order of magnitude observed for other model monomer proteins [37], indicating that ExtJ is stable in the experimental conditions used.

Table 3.1. Secondary structural elements of ExtJ determined by BeStSel and AlphaFold. The percentages of secondary structural elements of ExtJ are presented for the deconvolution of the spectra before and after the temperature ramp, at 25 °C, for the four different versions of ExtJ. The normalized root mean square deviation (NRMSD) values between the experimental and calculated data are presented. "T" and "U" correspond to the N-terminal tagged and untagged versions of the proteins, respectively.

BeStSel – 25 °C (before temperature ramp)					
Secondary structural elements (%)					
Protein	**α-helix**	**β-sheet**	**Turn**	**Random coil**	**NRMSD**
ExtJ WT T	14.5	35.1	12.3	38.1	0.0094
ExtJ WT U	12.9	35.7	12.6	38.8	0.0101
ExtJ C81* T	16.9	21.6	14.4	47.1	0.0090
ExtJ C81* U	12.8	36.1	12.1	39.0	0.0089
BeStSel – 25 °C (after temperature ramp)					
Secondary structural elements (%)					
Protein	**α-helix**	**β-sheet**	**Turn**	**Random coil**	**NRMSD**
ExtJ WT T	6.5	38.5	13.6	41.4	0.0184
ExtJ WT U	11.8	34.3	13.5	40.4	0.0163
ExtJ C81* T	11.5	37.7	11.9	38.9	0.0167
ExtJ C81* U	11.0	35.0	13.0	41.0	0.0124
AlphaFold models					
Secondary structural elements (%)					
Protein	**α-helix**	**β-sheet**	**Turn**	**Random coil**	**NRMSD**
ExtJ WT T	6.0	35.0	4.0	55.0	-
ExtJ WT U	7.2	42.2	4.8	45.8	-
ExtJ C81* T	6.1	35.4	4.0	54.5	-
ExtJ C81* U	7.3	42.7	4.9	45.1	-

Importantly, the CD spectra acquired at 25 °C before and after the temperature ramp are highly similar (Table 3.1), indicating that ExtJ recovers most of its secondary structural elements after a temperature denaturation, upon cooling down.

The CD spectra deconvolution and thermodynamic parameters calculation of the four proteins (Table 3.1 and Figure 3.9) are in line with the results obtained by NMR. Overall, these results demonstrate that ExtJ C81* T, for its higher expression yield and simplified purification protocol – without the need of reducing agents and N-terminal affinity tag cleavage – can be used for future studies regarding ExtJ's structure and function.

3.3.4. NMR solution structure and relaxation experiments of ExtJ

Following the preliminary structural characterization of ExtJ, the solution structure of the protein was determined. Initially, side-chain resonances were assigned through the combined analysis of 3D ^{1}H, ^{15}N-TOCSY-HSQC and 3D (H)CCH-TOCSY experiments [20]. 3D ^{1}H, ^{15}N-TOCSY-HSQC correlates the resonances of all side-chain protons with their respective backbone NH group [38], whereas 3D (H)CCH-TOCSY is used to identify side-chain ^{13}C resonances within a certain ^{13}C spin system [39]. While most of the resonances were successfully assigned using these experiments, we resorted to 3D ^{13}C-NOESY [38, 40] and 3D ^{15}N-NOESY [38, 40, 41] experiments not only to identify unassigned side-chain proton resonances, but also to define distance constraints for solution structure determination. These experiments allow the identification of all protons within a 5-6 Å distance of a certain ^{13}CH or ^{15}NH group, respectively. Therefore, during side-chain assignment with these spectra, one must carefully determine if a certain cross-peak corresponds to either an intra-residue side-chain resonance or to an inter-residue through-space dipolar correlation (NOE effect).

The use of the mentioned 2D and 3D NMR experiments led to the near-complete assignment of the ^{1}H (99.1%), ^{13}C (97.1%) and ^{15}N (98.8%) NMR resonances of ExtJ (Table 7.1).

The automated NOE assignment and structure calculations of ExtJ were carried out with the standard protocol in CYANA 3.98.13 [42, 43], implemented within the novel machine-learning based method ARTINA [44], using the NMRTist platform [45]. In this platform, only the protein sequence of ExtJ without the N-terminal affinity tag and all the mentioned 2D and 3D spectra were provided. The ARTINA approach uses convolutional neural networks for cross-peak detection [46], FLYA automated chemical shift assignment [47], and CYANA for automated NOE assignment and structure calculation. This final step comprises 7 cycles of NOESY-cross-peak assignment and structure determination, followed by a final structure calculation. In total, 8 x 100 conformers are calculated for a given input data set using 30,000 torsion angle dynamics steps per conformer. All chemical shifts were manually checked and compared with our assignment. The resultant 20 conformers with the lowest final target function value were water-refined in the AMPS-NMR server [48] and chosen to represent the solution structure of ExtJ (Figure 3.10).

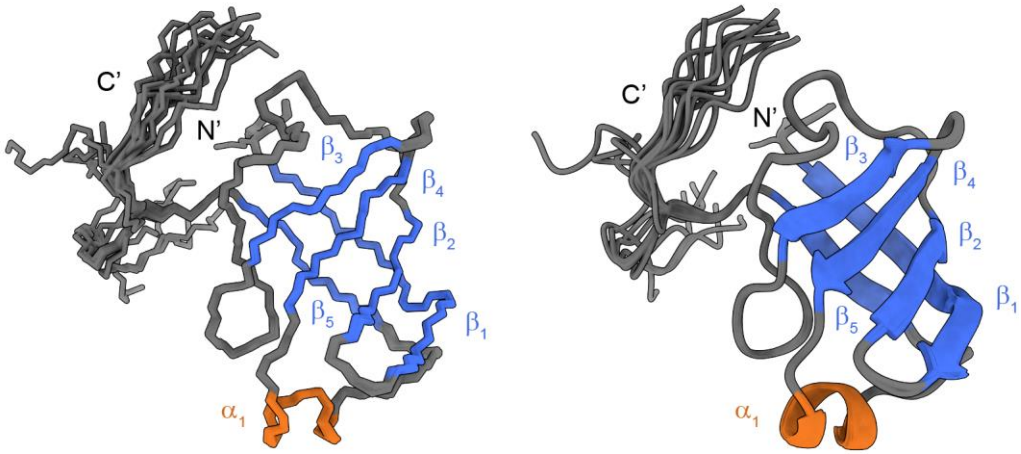


Figure 3.10. Solution structure of ExtJ. The twenty superposed lowest-energy NMR structures are represented by the backbone atoms (C', C_α and N) on the left, and as ribbon models on the right. The N- and C-terminus are labeled. The β-strands are colored blue, whereas the α-helices are colored orange.

The final NMR ensemble was calculated based on 1412 NOE distance restraints and superposes with an average pairwise backbone and heavy atom RMSDs of 1.84 ± 0.45 and 1.92 ± 0.41 Å, respectively. These values are affected by the local low resolution of the C-terminal region, which is highly flexible and lacks NOE restraints. Indeed, an independent pruned alignment between residues Ala^{1}-Glu^{72} has an average pairwise backbone and heavy atom RMSDs of 0.14 ± 0.02 and 0.49 ± 0.05 Å, respectively. The structure statistics are shown in Table 3.2.

Table 3.2. NMR and refinement statistics of ExtJ monomer structure. The statistics obtained for an independent pruned alignment between residues Ala[1]-Glu[72] are presented within brackets.

Distance constraints	
Total NOE	1412
Intra-residue ($\|i-j\|=0$)	306
Inter-residue	
Sequential ($\|i-j\|=1$)	400
Medium-range ($1<\|i-j\|<5$)	125
Long-range ($\|i-j\|\geq 5$)	581
Hydrogen bonds	26
Torsion angle restraints	126
Structure statistics	
CYANA target function	0.67 ± 0.02 Å^2
Distance restraint violations > 0.2 Å	0
Maximal distance restraint violation	0.09 ± 0.02 Å
Angle restraints violations > 5 Å	0
Maximal angle restraint violation	1.89 ± 0.18 °
Ramachandran statistics	
Most favored regions	81.8%
Additionally allowed regions	16.8%
Generously allowed regions	1.4%
Disallowed regions	0.0%
Average pairwise RMSD	
Backbone	1.84 ± 0.45 Å (0.14 ± 0.02 Å)
Heavy atom	1.92 ± 0.41 Å (0.49 ± 0.05 Å)

The solution structure of ExtJ is in excellent agreement with the AlphaFold and Rosetta CS models (Figures 3.7 and 3.10), comprising a five-stranded β-barrel framed by a single α-helix. The β-barrel consists of antiparallel strands organized in a Greek-key topology (β_1, Ser[3]-Asp[12]; β_2, Lys[15]-Lys[19]; β_3, Thr[41]-Val[45]; β_4, Glu[49]-Lys[53]; and β_5, Ser[67]-Ser[71]), which are separated by either turns, random coil stretches (Ala[20]-Pro[40]) or an α-helix (α_1, Lys[56]-Lys[61]), located between the β_4 and β_5 strands. Moreover, CYANA predicts that ExtJ contains an additional β-strand formed by residues His[32]-Ala[35], within the long random coil stretch formed by residues Ala[20]-Pro[40], in at least 6 out of the 20 lowest energy conformers. This observation is in line with our secondary structure prediction results (Figure 3.6).

Nevertheless, this strand is not defined in the final structural assemble, indicating that there might be local dynamics or conformational heterogeneity in this region of the protein. ^{15}N NMR relaxation experiments were used to determine the T_1, T_2 and HetNOE relaxation parameters of monomer ExtJ, and to characterize the dynamic properties of the protein in solution by probing the timescale of the backbone motions. The ^{15}N NMR relaxation experiments were acquired at 500 and 600 MHz (Figure 3.11) to account for conformational exchange and future order parameter (S^2) calculations [49].

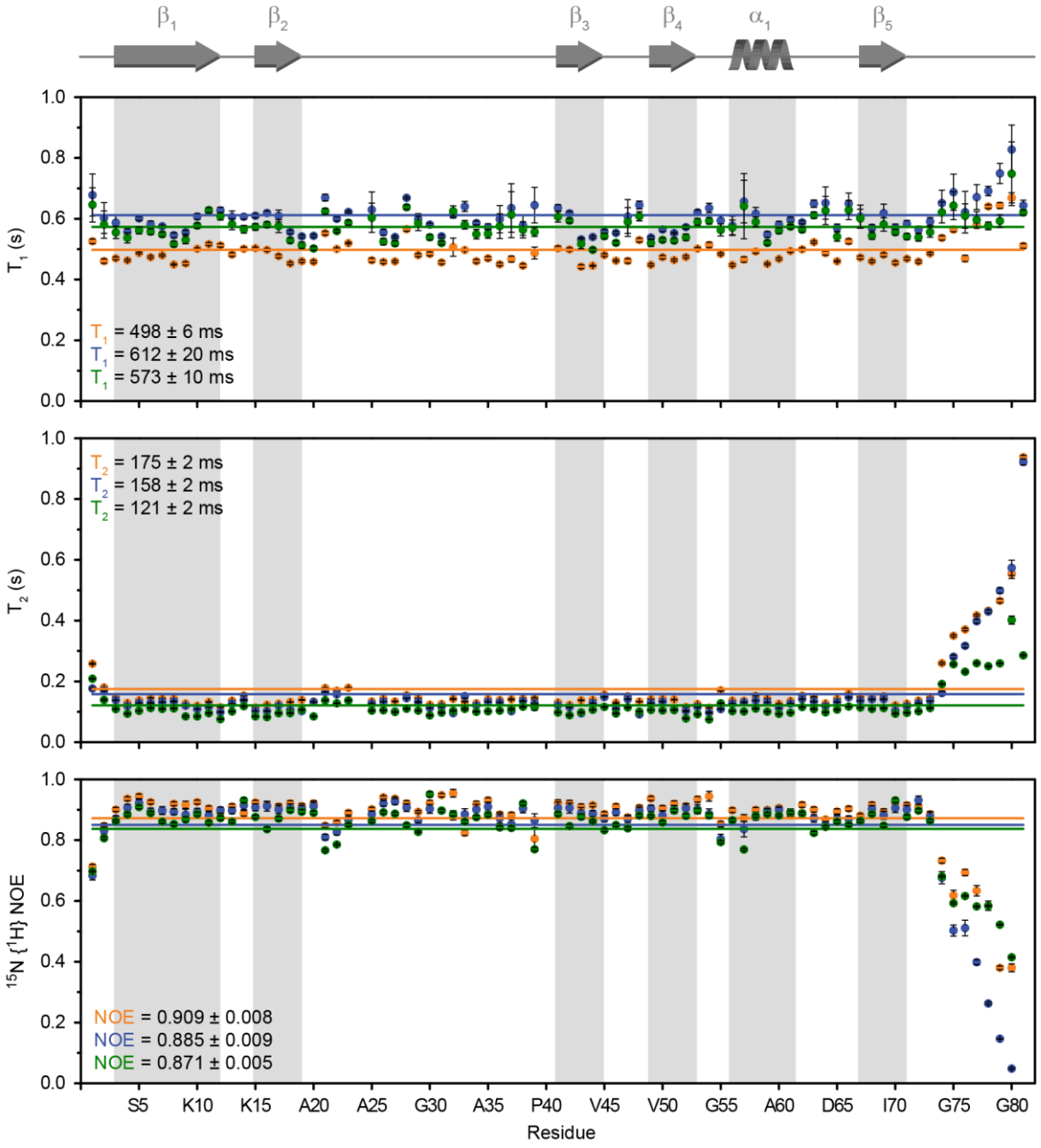


Figure 3.11. ^{15}N relaxation parameters of ExtJ. The ^{15}N longitudinal T_1 and transverse T_2 relaxation times, together with the ^{15}N HetNOE are represented for data acquired at 500 (orange) and 600 (blue) MHz, with a monomer sample. The same parameters are shown for data acquired at 500 MHz (green), with a dimer sample. In all panels, the solid straight lines and the presented values represent the average T_1, T_2 and HetNOE parameters at the respective magnetic fields. The secondary structural elements of ExtJ are highlighted by gray boxes.

Relaxation data were obtained for all residues, except Pro^{24} and Pro^{40}. The average values for the ^{15}N T_1, T_2 and HetNOE relaxation parameters at 500 MHz are 498 ± 6 ms, 175 ± 2 ms and 0.909 ± 0.008, respectively. As expected, the average values for the ^{15}N T_1 (612 ± 20 ms) and T_2 (158 ± 2 ms) relaxations respectively increase and decrease with an increase in the magnetic field strength (600 MHz) [50]. Overall, there is a good correlation between the relaxation data and the determined solution structure of ExtJ (Figures 3.10 and 3.11). While the overall differences in the T_1 values of the different residues are not significant, a greater variability was observed in the T_2 and HetNOE data (Figure 3.11), from which features of the protein dynamics can be extrapolated. Most residues located in defined secondary structural elements exhibit high rigidity, with motions in the μs-ms timescale, reflected by both the HetNOE values close to the theoretical maximum and the lower T_2 values [51, 52]. In contrast, the C-terminal stretch of ExtJ shows fast mobility, in the ps-ns timescale, exhibiting low HetNOE and high T_2 values. This explains why there is a lack of NOE restraints for this region of the protein and might be an indication that the C-terminal is particularly relevant for the function of ExtJ [53-55].

3.3.5. ExtJ homodimerization is driven solely by an intermolecular disulfide bridge

The dimerization of ExtJ can be a determinant factor for its function, as discussed in section 2.3.4. Consequently, it is important to pinpoint exactly what are the structural mechanisms that lead to dimer formation. In addition to the intermolecular disulfide bridge established between different C-terminal cysteine residues, there might be other mechanisms that promote the dimerization of ExtJ. Moreover, it would also be interesting to understand if there are structural differences between each monomer upon oligomerization. In search of an answer for these questions, 2D ^{1}H, ^{15}N-HSQC spectra of a 100 μM sample of ExtJ WT in the presence and absence of 1 mM DTT were acquired (Figure 3.12). The comparison of the two spectra showed no significant differences, suggesting that the two ExtJ molecules involved in homodimer formation are rotating in solution within an equivalent frame, thus behaving nearly independently in solution. The only marked difference is the absence of the backbone amide resonance of Cys^{81} in the 2D ^{1}H, ^{15}N-HSQC spectrum of the homodimer sample (Figure 3.12). The $^{13}C_\beta$ chemical shift values of cysteine residues can be used to deduce their redox state and, in this case, to confirm the formation of the disulfide bridge by Cys^{81} [56-58]. However, the backbone amide resonance of Cys^{81} is either displaced and superposed with other resonances of the homodimer protein or its signal linewidth is severely increased, leading to its resonance being no longer observable, therefore impairing this assessment.

Additional 2D ^{1}H, ^{15}N-HSQC spectra of ExtJ WT were acquired at different temperatures (5-50 °C), ionic strengths (25-300 mM) and protein concentrations (10-1000 μM) to probe for experimental conditions in which different resonances could be observed for the monomeric and dimeric forms of ExtJ. However, no such pattern was detected under the tested conditions.

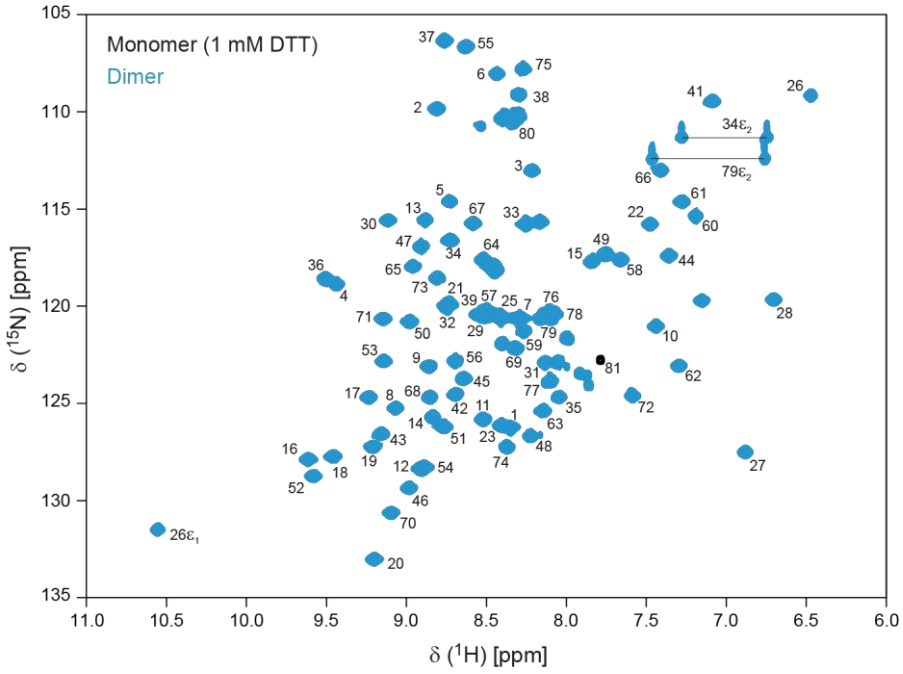


Figure 3.12. 2D ^{1}H, ^{15}N-HSQC spectra of ExtJ in the presence and absence of 1 mM DTT. The spectra in the presence (monomer) and absence (homodimer) of reducing agent are represented in black and blue, respectively. The spectra were acquired at 25 °C with 100 μM samples prepared in 32 mM sodium phosphate buffer pH 6 (100 mM of final ionic strength).

Following these results, and to understand the dynamic behaviour of the dimeric form of ExtJ in solution, we acquired ^{15}N NMR relaxation experiments with samples prepared in the absence of DTT (Figure 3.11). The average values for the ^{15}N T_1, T_2 and HetNOE relaxation parameters are 573 ± 10 ms, 121 ± 2 ms and 0.871 ± 0.005, respectively, thus being very similar to those obtained for the monomer sample. Assuming that the tumbling of ExtJ in solution lies within the limit of slow molecular motion ($\tau_c \gg 0.5$ ns), the τ_c of the protein can be calculated relative to the ratio of the longitudinal T_1 and transverse T_2 ^{15}N relaxation times, according to Equation 3.3 [59, 60]. This information can then be compared with existent standard curves of τ_c (ns) vs protein molecular weight (kDa), obtained for a series of known monomeric proteins of varying size, to get an estimation of the oligomerization state of ExtJ in each experimental condition [61, 62]. Prior to this analysis, since the ^{15}N T_2 values observed in CPMG-type experiments may be subject to an apparent shortening due to processes of conformational or chemical exchange on time scales smaller than or comparable to the delay between the successive 180° pulses applied during these experiments (thereby contributing to the measured transverse relaxation), one must carefully evaluate if all T_2 values determined for the different amide groups of ExtJ are suitable for the application of Equation 3.3. To determine the putative exchange contributions to the apparent transverse relaxation rate of individual nitrogen spins, a selection criterion based on the analysis of the T_1/T_2 ratios was applied [63]. Following this criterion, no amide groups showed T_1/T_2 ratios more than one standard deviation larger than the mean T_1/T_2 value, meaning that the average T_1 and T_2 values previously determined for all amide groups are suitable for the application of Equation 3.3.

The resulting plot of τ_c vs protein molecular weight (Figure 3.13) indicates that in both sets of data, ExtJ (10.2 kDa) behaves mostly as a monomer, since both experimental points (orange and green) lie close to the standard monomer curve.

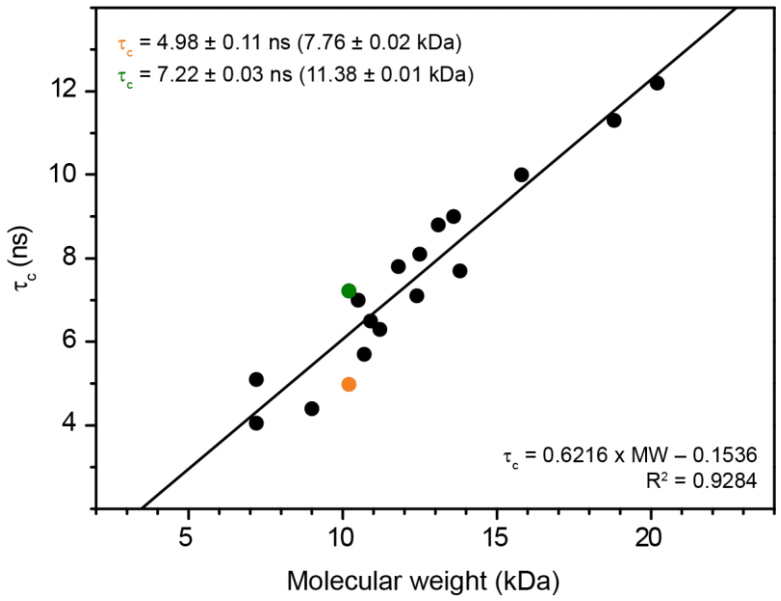


Figure 3.13. Plot of rotational correlation time (τ_c) as a function of protein molecular weight. The black spheres represent the set of known monomeric proteins listed on [61], which were used to calculate the linear regression represented by the black line, whose equation and R^2 are represented in the bottom right corner. The orange and green spheres represent the data for the monomeric and dimeric forms of ExtJ, respectively.

While the extrapolated apparent molecular weight of ExtJ in the dimer sample is higher (11.38 ± 0.01 kDa) than the one obtained for the monomer sample (7.76 ± 0.02 kDa) – pointing towards a faster tumbling in solution – in general, the data supports the view that the two ExtJ molecules involved in the homodimer formation are rotating in solution within an equivalent frame, thus behaving nearly as two independent molecules, as previously postulated upon comparison of the 2D ^{1}H, ^{15}N-HSQC spectra of the monomeric and dimeric forms (Figure 3.12).

However, despite these results, and taking advantage of AlphaFold-Multimer [64], we decided to determine the model of ExtJ's dimer (Figure 3.14). An electrostatic surface representation of the AlphaFold-Multimer model of the homodimer shows that the interaction between monomers could be electrostatically driven, involving mostly surface exposed glutamate and lysine residues. Nevertheless, as mentioned, since multiple 2D ^{1}H, ^{15}N-HSQC spectra acquired at different ionic strengths and protein concentrations did not reveal chemical shift variations, this is unlikely. The PAE scores provided by AlphaFold-Multimer give a distance error for every pair of residues, which can be used as measure of confidence for the relative position of those residues [65, 66]. In the case of AlphaFold-Multimer model of ExtJ's dimer, the average PAE score is 15 Å, meaning that the overall confidence in the dimer configuration is low. Consequently, we decided to perform all-atom molecular dynamics (MD) simulations of the complex and check for both the stability and possible new configurations of the dimer.

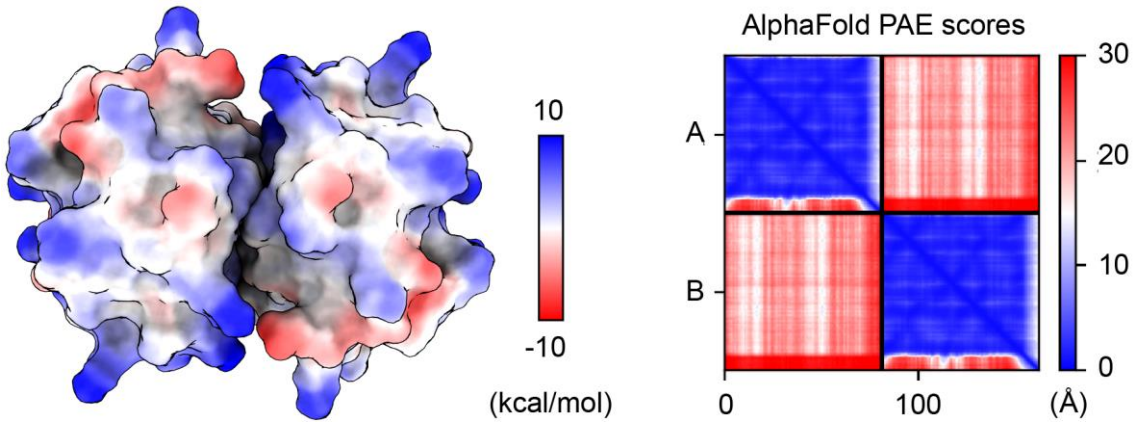


Figure 3.14. Electrostatic surface of the homodimer. The electrostatic surface of the dimer is colored in red and blue, for negative and positive charges, respectively. The Coulombic electrostatic potentials (kcal/mol) were calculated at 298 K, according to Coulomb's law, using UCSF ChimeraX 1.4 [67]. The PAE score chart generated by AlphaFold-Multimer is shown on the right, while the pLDDT scores of the model can be found in Figure 7.7.

MD simulations are used to predict how every atom in a protein or other molecular system will move over time, based on a general model of the physics governing interatomic interactions [68]. These simulations can capture a wide variety of important biomolecular processes, including protein folding and conformational changes, revealing the positions of all the atoms up to a femtosecond temporal resolution [69]. The 250 ns MD simulation of the homodimer showed that it undergoes a conformational change early in the simulation (10-15 ns), after which it remains stable (Figure 3.15).

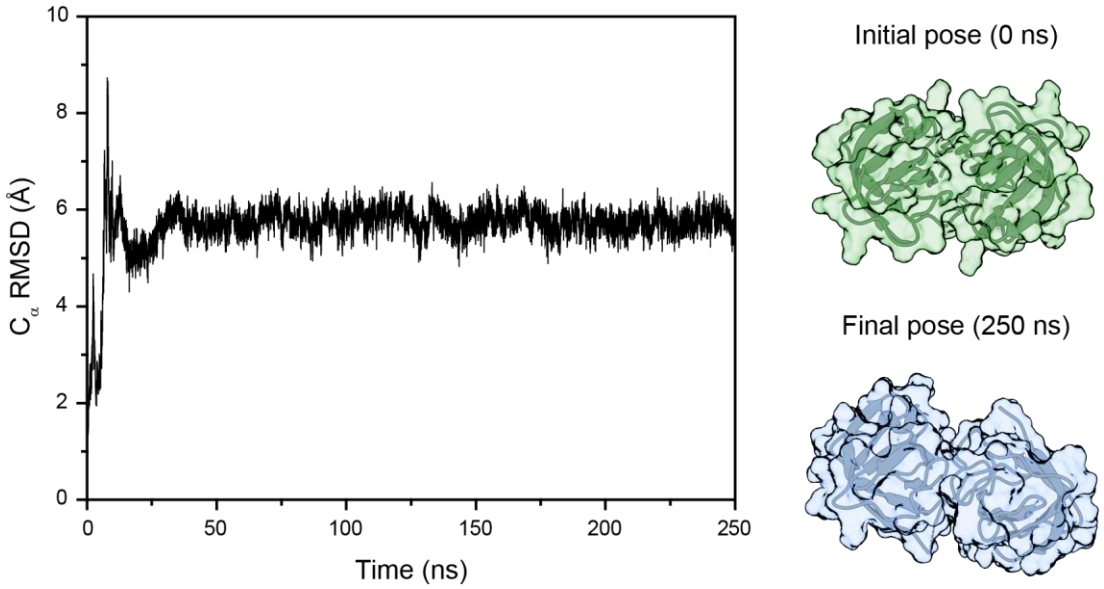


Figure 3.15. MD simulation of the ExtJ homodimer in explicit water. The plot shows the average RMSD variation of the protein C_α atoms throughout the 250 ns MD simulation. The structures on the right show the initial (green) and final (blue) poses of the homodimer. The potential and total energies plots, together with those showing the temperature and pressure variation throughout the MD simulation are represented in Figure 7.11.

While the AlphaFold-Multimer model of the homodimer was apparently stabilized by electrostatic interactions, such contacts were absent in the final configuration of the MD simulations. In fact, this dimer configuration was stabilized by mirrored hydrogen bonds established between Gly^{55} of one monomer with Gly^{75}-Asp^{76} of the remaining. This configuration is unlikely to accurately represent the actual homodimer configuration, since these interactions would most likely stabilize the C-terminal stretch of each ExtJ monomer composing the dimer, resulting in different relaxation properties compared to the monomer alone. This was not observed in the ^{15}N NMR relaxation data (Figure 3.11).

In parallel with the analysis of the NMR and MD simulations data, microscale thermophoresis (MST) was used to quantify the putative dissociation constant (K_D) of the homodimer. MST is a powerful biophysical technique that exploits the movement of molecules in a temperature gradient to quantify molecular properties [70, 71], by detecting changes in the hydration shell of fluorescently labeled biomolecules [72], independently of their size and physical properties [73, 74]. Initially, a fluorescently labeled protein or nuclei acid is subjected to an infrared laser that creates a local temperature gradient, for which a thermophoretic mobility can be measured. As the biomolecule interacts with a non-labeled partner, changes in its thermophoretic mobility occur, which can be used to quantify interactions in the pM to mM range [70].

In ExtJ's particular case, to avoid the effect of the disulfide bridge in the homodimerization process, ExtJ C81* was labeled on the N-terminal His-tag with a non-covalent RED-tris-NTA 2nd generation dye in 1x phosphate-buffered saline (PBS) pH 7.5 with 0.05% Tween-20, following the manufacturer's instructions. The experiments were performed at pH 7.5 instead of the usual pH 6 used in the other experiments, because the affinity of the indicated dye diminishes considerably at pH values below 7. Additionally, Tween-20 was added to increase protein solubility, limit aggregate formation and avoid putative capillary adhesion [75]. The fluorescent protein samples were prepared with a constant concentration of 50 nM of labeled ExtJ C81* and mixed in a serial dilution with unlabeled ExtJ C81* U in a concentration range between 15 nM and 500 μM. The measurements showed no aggregation or photobleaching and the thermophoresis variation allowed the calculation of a K_D of 15.2 ± 4.2 μM for the putative monomer-dimer equilibrium (Figure 3.16).

These results (in which no disulfide bridge may be involved, since the experiments were performed with the C81* mutant), might have implications in our previous interpretations of the monomer-dimer equilibrium using NMR. In fact, this K_D value in the μM range likely hampers the monitorization of the monomer-dimer equilibrium by NMR, since a state in which the monomer form of ExtJ is predominant in solution can only be supposedly attained at concentrations below 1 μM, which is incompatible with the experimental requirements of the technique.

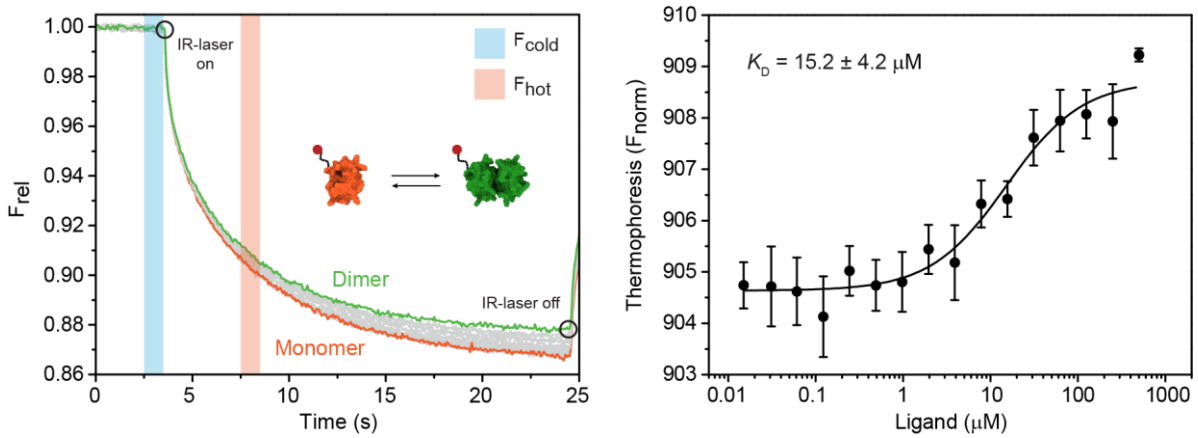


Figure 3.16. MST measurements of the monomer-dimer equilibrium of ExtJ. The left panel shows the variation of the relative fluorescence (F_{rel}) during the MST experiment. The blue (F_{cold}) and pink (F_{hot}) regions were used as reference to extract the data before and after laser irradiation and consequent thermophoretic mobility, respectively. The panel on the right shows the fitting of Equation 3.7 to the thermophoresis (F_{norm}), with the respective error bars from triplicate experiments.

Nevertheless, while these results would strongly support the fact that the dimerization of ExtJ is prone to occur in vivo, the small thermophoretic response observed (four normalized fluorescence units) raises the question as to whether or not the MST response curve might be arising as an artifact, due to small differences in pH or ionic strength between the titrant and titrate solutions, rather than as a response to an effective difference in mobility between the putative monomeric and dimeric states. Moreover, while attaining experimental conditions suitable for the observation of the monomer form in solution by NMR is incompatible with the results showed by MST, the control experiments performed by NMR at various protein concentrations (from 10-1000 μM, including 10, 15, 30, 50, 75 and 100 μM, well within the concentration range in which a mixture of monomer and dimer would be expected, according to the MST data) and temperatures (which should account for the existence of a putative intermediate exchange equilibrium between the monomer and dimer forms), show no evidence of the existence of two oligomeric forms, representing another argument against the MST data.

Due to the unclear and somewhat contradicting results obtained using NMR, MD simulations and MST, we decided to use small-angle X-ray scattering (SAXS) to unequivocally confirm the presence or absence of additional structural promoters of the homodimerization of ExtJ, in addition to the paramount intermolecular disulfide bridge.

SAXS is a powerful biophysical method for the structural characterization of proteins in solution, providing low resolution information on the shape, conformation and assembly state of proteins, in a broad range of molecular sizes [76]. In this technique, a solution of dissolved macromolecules usually placed in a quartz capillary is illuminated by a defined energy, collimated monochromatic X-ray beam, and the intensities of the resulting scattered X-rays are recorded by an appropriate X-ray detector [77]. As the sample is isotropic (the molecules in solution are tumbling independently of each other), the intensity distribution can be radially averaged, resulting in a one-dimensional

scattering curve. Since the average electron density of proteins (0.43 $e^-/Å^3$) is only slightly higher than the average electron density of water (0.33 $e^-/Å^3$), which can increase with the addition of components to the buffer), a paramount step in any SAXS experiment involves the recording of a second measurement of an identical solution lacking the sample of interest (protein), which will be used for background subtraction of all scattering contributions, not only from the solvent, but also from the sample holder, as well as the machine setup [78]. The resulting pattern corresponds to the scattering of a single protein molecule averaged over all orientations, from which several important biophysical parameters can be directly obtained, including the radius of gyration, maximum interatomic distance, molecular weight, oligomeric state and overall shape of the particles under investigation [79]. Additionally, SAXS can also be used to determine reliable low-resolution (10-20 Å) ab initio models of the macromolecules under study [80].

Initially, four different ExtJ samples were prepared for SAXS analysis. ExtJ WT was prepared in the standard sample buffer used for most of the experiments (32 mM sodium phosphate buffer pH 6 (100 mM of final ionic strength)), with and without the presence of DTT. ExtJ C81* was also prepared in the standard sample buffer (100 mM of final ionic strength) and in a buffer with a lower ionic strength (25 mM), to promote the putative electrostatically-driven homodimerization of the protein. All data was collected in batch mode for serial dilutions of each sample, for which the corresponding matched buffer scattering contribution was subtracted. The samples and respective frames with the best signal-to-noise ratio, and with an apparent absence of aggregates and minimal radiation damage, were selected for subsequent data analysis. The resulting scattering patterns for each sample are presented in Figure 3.17A.

The scattering profiles of all samples have similar shapes and are typical of spherical particles [77], presenting good signal-to-noise ratios. Nevertheless, at low-angle, there is a clear difference between the scattering profile of the ExtJ WT sample without DTT and those of the remaining samples, which are superposed. Considering that the presence of a homodimer for the WT protein in the absence of a reducing agent was never in doubt, this is an initial indication that the remaining sample conditions likely contain only the monomer. The remaining analysis of the SAXS data supports this view.

The Kratky plot analysis is used to qualitatively assess the flexibility and degree of unfolding of a protein [81]. In the case of a well-folded globular protein, the Kratky plot will exhibit a bell-shape (Gaussian) peak that converges to the S-axis. An unfolded or highly flexible protein will present a Kratky plot in which the profile increases linearly with S, without converging to the S-axis. Proteins with multiple domains may have a combination of both these features, in which a bell-shape peak does not converge to the S-axis due to partial flexibility or presence of disordered domains. This is exactly what we observe on the Kratky plots of the different ExtJ samples (Figure 3.17B) – a bell-shape peak (resulting from the ordered and compacted domain of ExtJ) that does not converge to the S-axis due to the contribution of the disordered N-terminal tag and C-terminal stretches, as hinted by the NMR relaxation data.

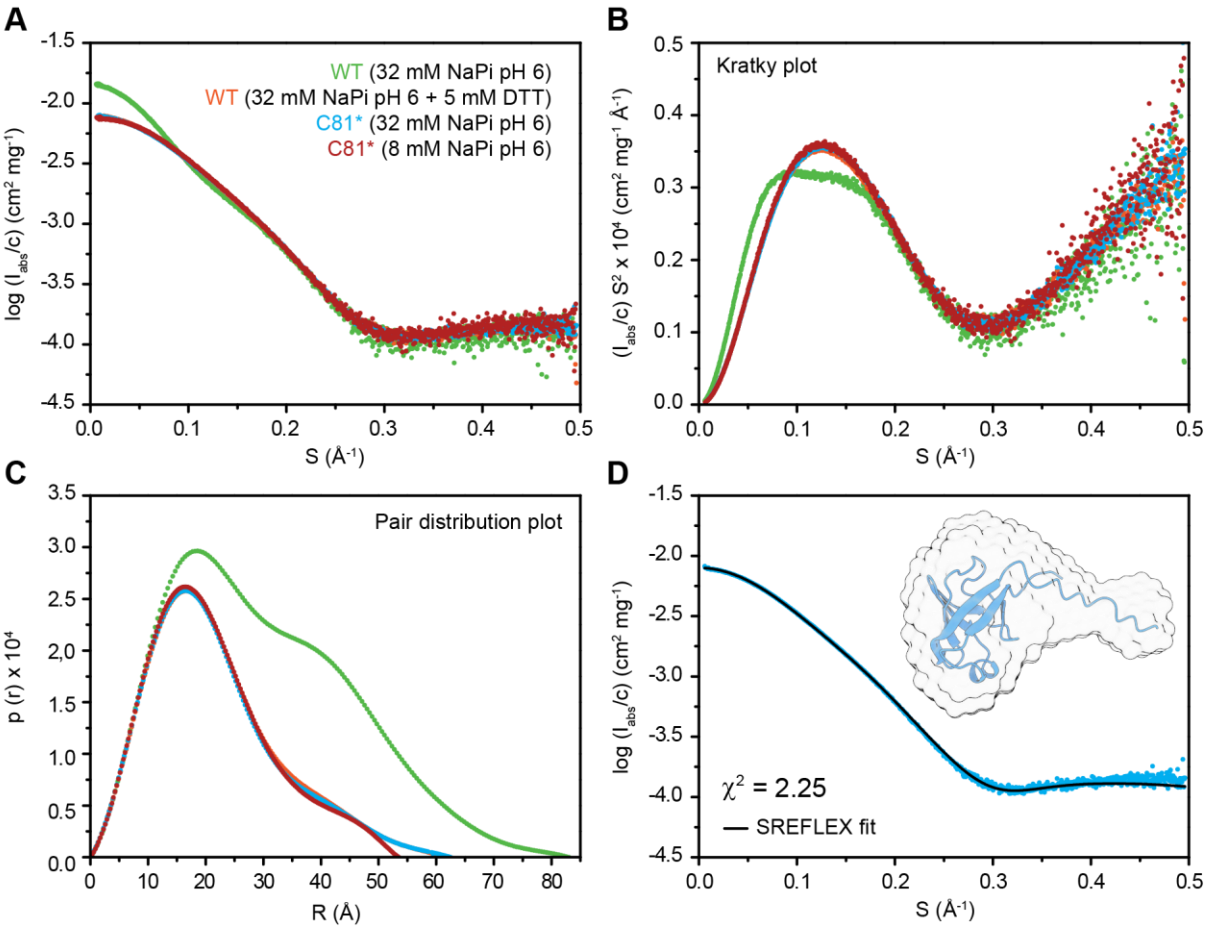


Figure 3.17. SAXS studies of ExtJ. (A) Experimental scattering curves of ExtJ WT (32 mM sodium phosphate buffer pH 6 (100 mM of final ionic strength) with (orange spheres) and without (green spheres) the presence of DTT) and ExtJ C81* (32 mM (blue spheres) and 8 mM (red spheres) sodium phosphate buffer pH 6 (100 mM and 25 mM of final ionic strength, respectively)). The color code is respected in all panels. (B) Kratky plots of the different ExtJ samples. The data was normalized relative to each's sample concentration. (C) Pair distance distribution plots of the different ExtJ samples. The data was normalized relative to each's sample concentration. (D) Scattering curve fitting and ab initio modeling of monomer ExtJ. The theoretical scattering curve (black line) was calculated with CRYSOL [82], using an AlphaFold model of ExtJ containing the N-terminal tag as input, and fitted to the experimental scattering curve of ExtJ C81* (blue spheres, 32 mM sodium phosphate buffer pH 6 (100 mM of final ionic strength)) with SREFLEX [83]. The inset shows a ribbon representation of the ExtJ AlphaFold model, docked into the ab initio SAXS-derived molecular envelope (semi-transparent gray surface) generated by DAMMIF [84], which is the average of 15 independently reconstructed models, all aligned, averaged, and filtered with DAMAVER [85].

A closer analysis of the Kratky plot of the ExtJ WT sample without DTT shows that the Gaussian curve is slightly wider and contains two separate peaks, a feature that arises from the presence of the symmetric monomers of the homodimer, in the absence of a reducing agent and the C81* mutation.

Following this analysis, we also represented the plots of the distance distribution function (p (r)) for each sample (Figure 3.17C). This function is calculated by applying a Fourier transformation to the scattering curve, at both the low- and high-angle regions, and describes the probability of finding a

pair of points within the sample separated by a certain distance (r) [86]. The shape and location of the peaks in a pair distance distribution plot reveal important details regarding the sample's size, shape and internal organization [87, 88]. The peaks at smaller distances generally correspond to the most common and smaller distances within a molecule, reflecting the average distance between neighboring atoms and the typical size of a domain. The shape of these peaks provides insights into the packing density of these structural elements. A sharp and intense peak suggests a well-defined, tightly packed structure, while a broader peak is indicative of higher variability in the interatomic distances, likely associated with a looser conformation. The peaks at larger-scale distances usually relate to larger structural features, corresponding to the distance between different domains or to the diameter of the entire protein. A molecule with a well-defined, extended shape, such as a rod or a cylinder, will result in a prominent peak at a distance corresponding to the overall length or width of the structure, while a more globular structure results in a broader and less pronounced peak in this region. As for the Kratky plot analysis, the presence of multiple peaks in the pair distance distribution plot suggests the existence of a multi-level organization within the sample, whose varying height and width might indicate variable structural density or subunit sizes [87, 88]. For the particular case of a homodimer, one should expect a peak at shorter distances, corresponding to intramolecular distances, together with a second peak at higher ones, representing the intermolecular distances. Finally, from the distance distribution function calculations, several biophysical parameters, including the R_g (radius of gyration, a measure of the compactness of a protein structure, calculated as the root mean square distance of each atom in the protein from the center of mass of the macromolecule [89]) and the D_{max} (maximum dimension of the particle in solution), were also determined, being in close agreement with those calculated from an independent Guinier analysis [90] (Table 3.3), and from the ^{15}N NMR relaxation data obtained for the monomer, using an empirically derived Stokes-Einstein equation for small spherical molecules (1.76 nm) [50]. Also, in this case, the data of the four samples may be split into two different groups, separating the ExtJ WT sample without DTT from the remaining. The scattering profile of this sample yielded a broader curve with identifiable peaks at around 20 and 40 Å, before reaching a D_{max} of 83 Å. Taking into consideration that this sample contains a mixture of monomer and dimer, the peak at 20 Å results from the interatomic distances within the compacted domain of the monomer and dimer, while the D_{max} of 83 Å results from the maximum stretchable distance of the disulfide bridge linked homodimer, also accounting for the long N-terminal tag. The second peak at 40 Å, unique to this sample, likely results from the interatomic distances between the two domains in the homodimer. The remaining samples share a single peak at 20 Å, as expected, but show smaller D_{max} values of around 55-60 Å, corresponding to the maximum distance between the opposite extremes of the disordered N-terminal tag and the globular domain of ExtJ. In particular, the ExtJ C81* sample prepared in 8 mM sodium phosphate buffer pH 6 (25 mM of final ionic strength) shows a more compact structure, likely due to the lower ionic strength conditions.

Table 3.3. SAXS data processing. The data is shown for all four ExtJ samples. HS and LS stand for high (32 mM sodium phosphate buffer pH 6 (100 mM of final ionic strength)) and low (8 mM sodium phosphate buffer pH 6 (25 mM of final ionic strength)) salt concentrations, respectively. MW stands for molecular weight. The Porod volume (V_p) was calculated from the total intensity of scattering extrapolated to zero angle (the forward scattering) by applying Porod's law [91], which describes the intensity of scattered X-rays at high-angles. This parameter gives an approximation of the volume occupied by the protein, assuming a compact shape.

Analysis	ExtJ WT	ExtJ WT + DTT	ExtJ C81* (HS)	ExtJ C81* (LS)
Scattering				
Porod volume (V_p, Å^3)	24512	16119	16811	13051
MW (kDa)	15.320	10.074	10.507	8.170
Guinier				
$I0_{abs}$ (cm^{-1})	0.108	0.116	0.114	0.054
$I0_{abs}/c$ (cm^2 mg^{-1})	0.0140	0.0078	0.0078	0.0077
MW (kDa)	19.300	10.800	10.800	10.600
R_g (nm)	2.34	1.65	1.67	1.58
SR_g range (nm)	0.32-1.22	0.36-1.30	0.29-1.11	0.43-1.30
p (r)				
D_{max} (Å)	83.0	61.5	62.5	53.5
R_g (nm)	2.39	1.72	1.72	1.65
S range (Å^{-1})	0.01-0.30	0.01-0.30	0.01-0.30	0.01-0.30
NMR				
MW (kDa)	11.380	7.760	–	–
R_g (nm)	–	1.76	–	–
Sequence				
MW (kDa)	10.273	10.273	10.170	10.170

Finally, using the AlphaFold model of ExtJ with the N-terminal tag, we were able to calculate a theoretical scattering curve that fits reasonably well with the experimental data, presenting a χ^2 (reduced error-weighted scoring function [92]) of 2.25 (Figure 3.17D). Moreover, this model docks correctly into the ab initio molecular envelope generated from the experimental scattering curve of the ExtJ C81* sample in the standard sample buffer (inset of Figure 3.17D). Unfortunately, we were unable to model the ab initio molecular envelope of the homodimer present in the ExtJ WT sample without DTT, due to the presence of a mixture of monomer and dimer species in the batch solution. Alternatively, we resorted to SEC-SAXS to try to circumvent this issue, but the available SEC columns in the European Synchrotron Radiation Facility (ESRF) lacked the necessary resolution to effectively separate the dimeric and monomeric forms of ExtJ.

In conclusion, the SAXS data (Figure 3.17 and Table 3.3) unequivocally confirmed that the ExtJ homodimerization is solely driven by the establishment of a disulfide bridge between the C-terminal cysteine of two monomeric molecules. Despite the high sample concentrations used (well above the putative homodimerization K_D of 15 μM determined by MST), the presence of the dimeric form of ExtJ was only observable in the WT sample without DTT, as confirmed by the different biophysical parameters determined from the SAXS data analysis (Table 3.3). All in all, generally reflecting on the multiple results we obtained from both experimental and computational methods, one can assume that while the AlphaFold model predictions might not be completely unrealistic, they likely represent intermediate conformations of a rather loose and stretchable homodimer, in which the major stabilization factor is solely the intermolecular disulfide bridge. Collectively, our data indicates that ExtJ forms homodimers in vitro, with each monomer tumbling in solution almost independently within an equivalent rotating frame.

3.3.6. Narrowing down the function of ExtJ in *Geobacter sulfurreducens*

ExtJ is unlikely involved in flavin-like compound synthesis or transport

As discussed in section 2.3.4, ExtJ is a structural homolog of multiple domains within subunits of large ATPases/synthases or ribosomal assemblies, and of riboflavin synthase-like β-barrel proteins. On one hand, the periplasmic localization of ExtJ in the cell, within a putative outer membrane porin-cytochrome complex, makes it unlikely that the protein is involved in ATP synthesis or ribosomal assembly. On the other hand, consequently, ExtJ might be involved in mechanisms that comprise the synthesis or binding of flavin compounds, such as riboflavin or FMN. Nevertheless, the genes that code for the majority of enzymes involved in riboflavin metabolism have already been identified in *G. sulfurreducens* [93]. Based on the structures of riboflavin synthase and other structural homologs (Figure 2.21), we can discern that most of these proteins function as a tandem of structurally similar domains, whose structural organization resembles that of monomer ExtJ.

Riboflavin synthase catalyzes the dismutation of two molecules of 6,7-dimethyl-8-ribityllumazine to yield riboflavin and 4-ribitylamino-5-amino-2,6-dihydroxypyrimidine [94]. The protein functions as a homotrimer consisting of an asymmetric assembly of monomers with a tight intermolecular interface, promoted by interactions between the β-barrels and by a three-helix bundle formed by the C-terminal α-helices of each monomer [95]. The three active sites of the trimer lie between pairs of monomers and are formed by two Asp^{187}-His^{98}-Ser^{148} triads and two different dyads (Cys^{47}-Ser^{40} and His^{103}-Thr^{150}) [94]. While the enzyme usually possesses three active sites, only one active site is catalytically competent at a time [96]. Alternatively, some riboflavin synthases may contain a single active site [97]. The relative orientation of the domains that compose the tandems of each monomer of riboflavin synthase are conserved across the protein homologs, but are reasonably different from the one we observe on the compacted homodimer AlphaFold model of ExtJ (Figure 3.18).

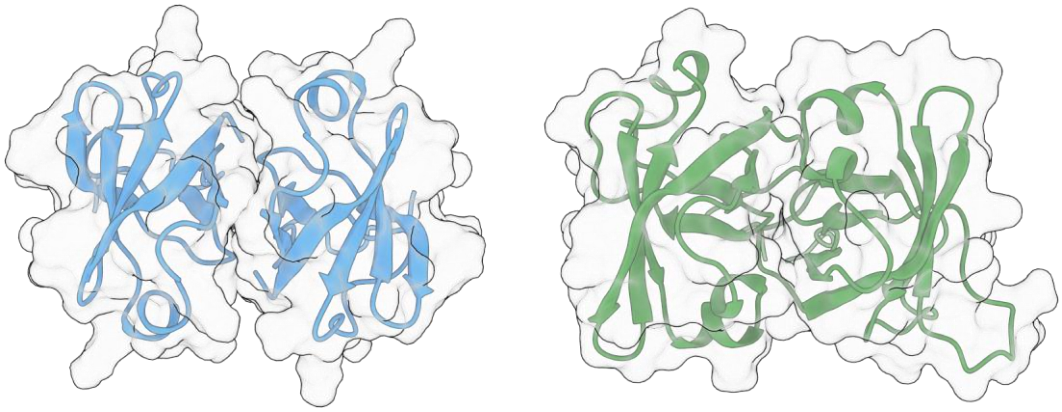


Figure 3.18. Relative orientation of the riboflavin synthase domains and the ExtJ homodimer. The AlphaFold model of the ExtJ homodimer is represented in blue, whereas the trimmed (residues 1-188) crystal structure of riboflavin synthase (PDB: 4GQN [97]) is represented in green.

The relative orientation and interdomain interactions between the structurally similar N- and C-terminal domains of the different riboflavin synthases are conserved even when only an isolated N-terminal domain is present in solution [98, 99]. This domain forms stable homodimers, whose structure in complex with riboflavin has been determined by both X-ray crystallography [98] and NMR [99] (Figure 3.19).

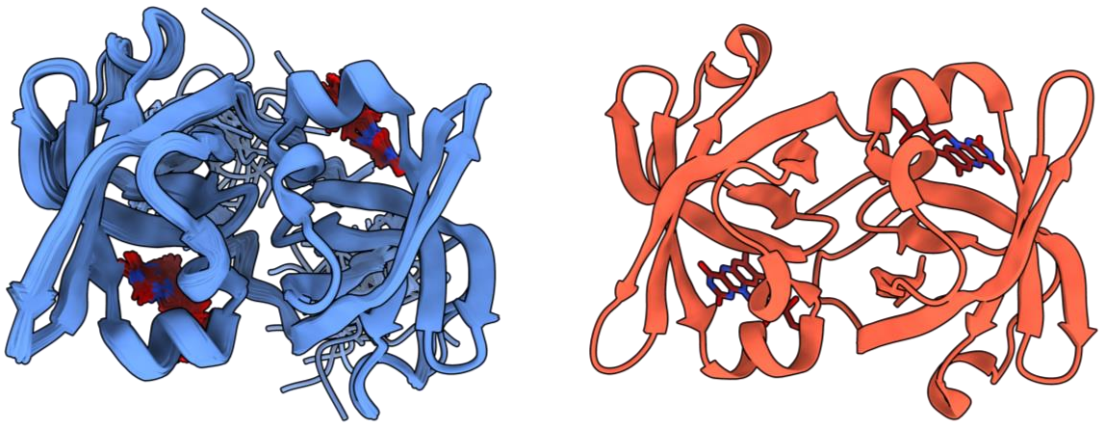


Figure 3.19. Homodimer structures of the N-terminal domain of riboflavin synthase. The NMR (PDB: 1I18 [99], blue) and X-ray crystallography (PDB: 1PKV [98], orange) structures are represented as ribbons, both with bound riboflavin (red).

The structures of this homodimer and those of the riboflavin synthase homologs show that the interfaces between domains are not stabilized by salt bridges or electrostatic interactions. Instead, they are stabilized by backbone-backbone and backbone-side-chain hydrogen bonds that promote a compact configuration, independently of the presence of the disordered loop that natively connects the N- and C-terminal domains of these proteins. The nature of these interactions would make it highly likely that ExtJ could form homodimers with a similar arrangement, but this is not observed. Considering the interface observed in the AlphaFold model of the ExtJ homodimer, formation of

backbone-backbone intermolecular hydrogen bonds would be likely to occur, but the absence of asparagine, aspartate, leucine, serine and threonine residues in conserved positions, which are responsible for the establishment of the backbone-side-chain hydrogen bonds observed in the structures of the riboflavin synthase homologs [98, 99], likely hinders the formation of a stable homodimer in ExtJ.

These observations regarding the homodimer interface of the riboflavin synthase and its homologs raise the question of whether or not ExtJ is able to bind flavin-like compounds, due to its inability to form stable homodimers. Several solution and crystal structures of riboflavin synthase proteins and their homologs have been determined in the presence of different ligands, including 6,7-dimethyl-8-ribityllumazine [100], 5-amino-6-ribitylamino-2,4(1H,3H)-pyrimidinedione [97], riboflavin [98] and FMN [100], for which a conserved binding pocket was observed (Figure 3.20).

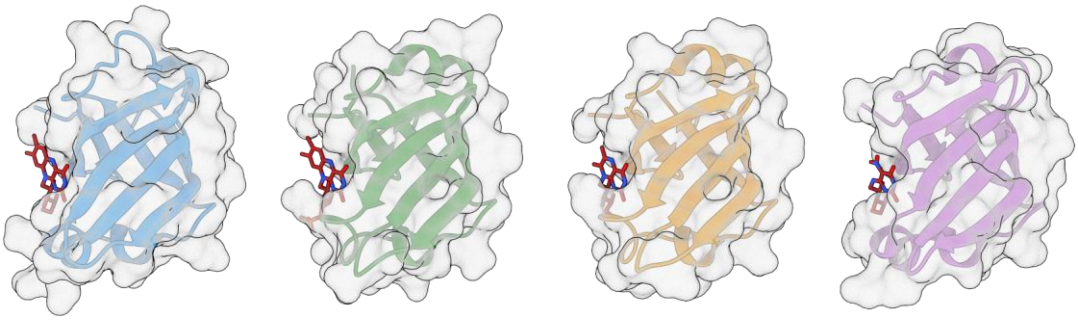


Figure 3.20. Binding pocket conservation in riboflavin synthase and homologs. The structures of the N-terminal domain of riboflavin synthase in complex with riboflavin (PDB: 1PKV [98], blue), the N-terminal domain of the lumazine protein in complex with FMN (PDB: 3A3B [100], green) and with 6,7-dimethyl-8-ribityllumazine (PDB: 3A3G [100], orange), and the C-terminal domain of riboflavin synthase in complex with 5-amino-6-ribitylamino-2,4(1H,3H)-pyrimidinedione (PDB: 4GQN [97], purple) are represented as ribbon. Their surfaces are represented in gray. All ligands are represented in red.

The only differences between the binding modes of the different ligands in the different structures relates with the hydrogen bond networks they form with the respective proteins. For example, on one hand, the number of hydrogen bonds decreases upon binding of FMN to a lumazine protein when we compare the same protein structure bound to riboflavin, mostly because the phosphate moiety of FMN protrudes out of the narrow ligand-binding cavity [100]. On the other hand, the binding of 6,7-dimethyl-8-ribityllumazine is usually tighter, with a higher number of hydrogen bonds observed when compared to the case in which the ligand is a riboflavin molecule [100]. In general, and taking the riboflavin binding pocket residues as a reference, an analysis of the different binding pockets with PoseView [101-105] shows that the most relevant residues for ligand binding are (i) a conserved $Thr^{49/50}$ residue, which establishes two hydrogen bonds, through its backbone and side-chain, (ii) a negatively charged residue, either Asp^{62} or Glu^{66}, that establishes a single backbone hydrogen bond with the NH group of riboflavin, (iii) a cysteine or serine residue, either $Cys^{47/48}$ or Ser^{48}, which might establish one or two backbone hydrogen bonds, either directly with riboflavin or with surrounding water molecules that stabilize the binding pocket and (iv) valine and glycine residues from the

neighboring domain, which establish multiple backbone hydrogen bonds. By superposing the structures of the AlphaFold model of the ExtJ's homodimer with that of the N-terminal domain of riboflavin synthase bound to riboflavin (Figure 3.21), we sought to understand if ExtJ contained the required residue nature and position for the putative binding of this molecule and its analogs.

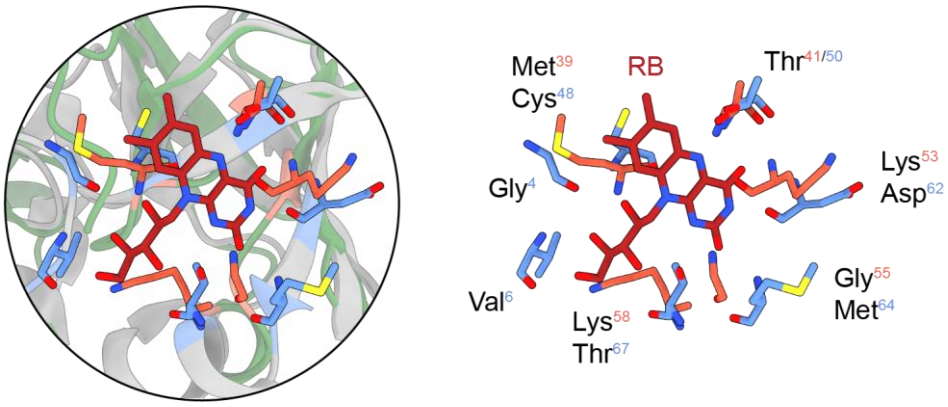


Figure 3.21. Structural conservation of the riboflavin binding pocket. The homodimer of ExtJ and that of the N-terminal domain of riboflavin synthase (PDB: 1PKV [98]) are represented in green and gray, respectively. The residues involved in riboflavin binding are highlighted in orange and blue for the homodimer of ExtJ and that of the N-terminal domain of riboflavin synthase, respectively. "RB" stands for riboflavin.

The only structurally conserved residue is the Thr$^{41/50}$, responsible for establishing both a side-chain and a backbone hydrogen bond with the riboflavin molecule. The remaining residues are not conserved, however, this might not completely hinder riboflavin binding in ExtJ. In fact, the establishment of hydrogen bonds might be maintained, considering that in riboflavin synthase, all of these residues establish hydrogen bonds with riboflavin through their backbone amide and carboxyl groups. The most striking difference in the putative binding pockets of these proteins are the residues of opposite charge in the Lys^{53}-Asp^{62} position, which, on one hand, may affect the binding affinity for riboflavin-like compounds. On the other hand, considering that the nature of riboflavin binding relies solely on hydrogen bonding, and that both amino acid side-chain chains are able to establish hydrogen bonds, this difference might not affect ligand binding. Nevertheless, the difference in charge in this region is not surprising, since ExtJ contains a high percentage of lysine residues (16%). Riboflavin synthases usually contain only 3-5% of lysine residues.

The architecture of the putative binding pocket of ExtJ is apparently suitable for riboflavin-like compounds binding, however, one must also assess the contributions from the other monomer molecule, considering that in riboflavin synthase, the domain opposite to one of the two binding pockets also plays a role in the binding event. From the superposition of these structures, no obvious candidates emerge in the case of the ExtJ homodimer. One hypothetical scenario would require a specific configuration of the disordered and flexible C-terminal stretch of ExtJ, that placed two glycine and one leucine residues within hydrogen bond distance of the ligands, but this is unlikely.

Based on this analysis, and to confirm if ExtJ can bind flavins, we performed NMR titrations of this protein with FMN (Figure 3.22). We selected this flavin due to its higher solubility in water at room temperature than riboflavin and 6,7-dimethyl-8-ribityllumazine, the second of which is also difficult to synthesize, and thus extremely expensive [106, 107].

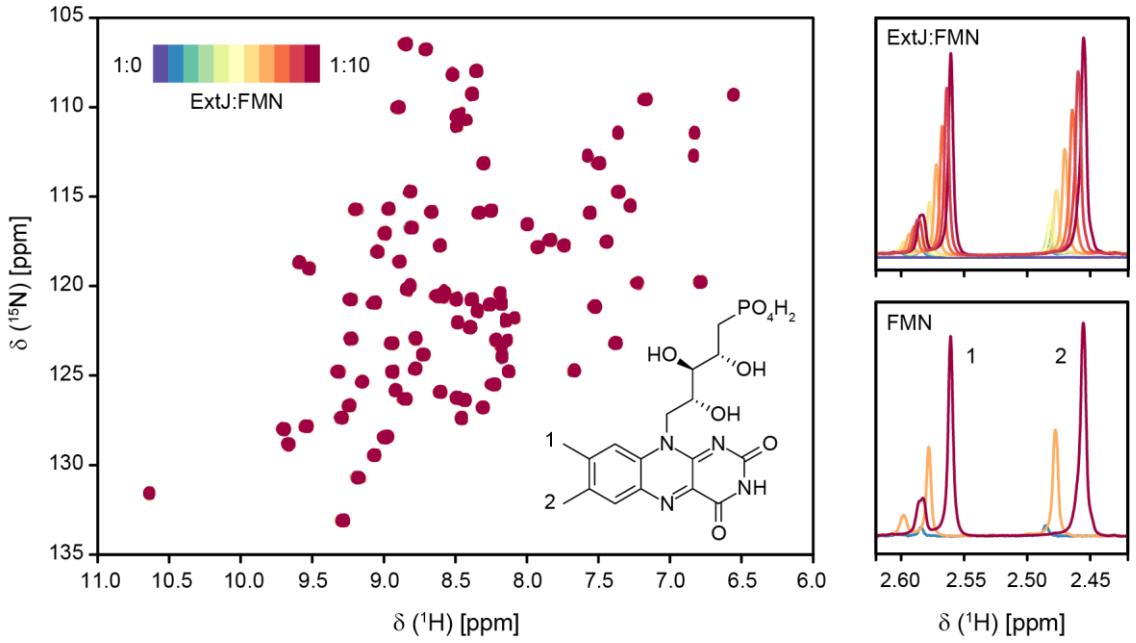


Figure 3.22. Monitorization of ExtJ-FMN interaction by NMR. The 2D ^{1}H, ^{15}N-HSQC spectra of ExtJ (100 μM) at various ExtJ:FMN ratios are represented by different colors, from purple (1:0 ExtJ:FMN) to dark red (1:10 ExtJ:FMN). The top right corner inset shows a close in of the 1D ^{1}H-NMR spectra of the different samples throughout the titration. The two intense resonances correspond to the methyl groups (1 and 2) of the FMN moiety, whose molecular structure is represented within the 2D ^{1}H, ^{15}N-HSQC spectra panel. The same 1D ^{1}H-NMR spectral region is highlighted in the bottom right corner, showing that the chemical shift variation observed during the titration with FMN is not due to a binding event, but rather due to the increase in FMN concentration.

No chemical shift variations were observed in the 2D ^{1}H, ^{15}N-HSQC of ExtJ following the addition of FMN, which indicates that there is no apparent interaction between the two molecules. However, the analysis of the ^{1}H signals of FMN in the series of 1D ^{1}H-NMR spectra acquired during the titration experiment revealed a chemical shift variation (Figure 3.22). This was, indeed, unexpected, and for that reason we acquired control spectra at different concentrations of FMN (Figure 3.22). In this case, the same chemical shift variation was detected, thus confirming that the ^{1}H chemical shift variation we observed during the titration of ExtJ with FMN was due to an increase in the ligand concentration, and not due to a binding event. Such behaviour has been previously reported in flavins [108-113].

In conclusion, despite the high degree of structural homology with domains of riboflavin synthase and homologs, since ExtJ is unable to bind FMN, the protein is unlikely to play a role in the binding or transport of flavin compounds in the periplasm of *G. sulfurreducens*.

Biomolecular interaction studies between ExtJ and PpcA

The presence of ExtJ, a protein with no predicted cofactors, in the ExtHIJKL complex, which is directly linked to the EET routes of *G. sulfurreducens*, raises the question about its role on the electron transfer or in other structurally relevant features of this complex. Since ExtJ is predicted to be located in the periplasm, a plausible hypothesis is that ExtJ somehow functions as a selective partner for potential periplasmic redox partners, promoting the interaction between these proteins and the remaining components of the ExtHIJKL complex. PpcA is the most abundant cytochrome in the periplasm of *G. sulfurreducens* [114] and has been exhaustively studied [115], making it an ideal candidate for interaction studies. The interaction studies were performed between ^{15}N ExtJ C81* and PpcA, using NMR (Figure 3.23).

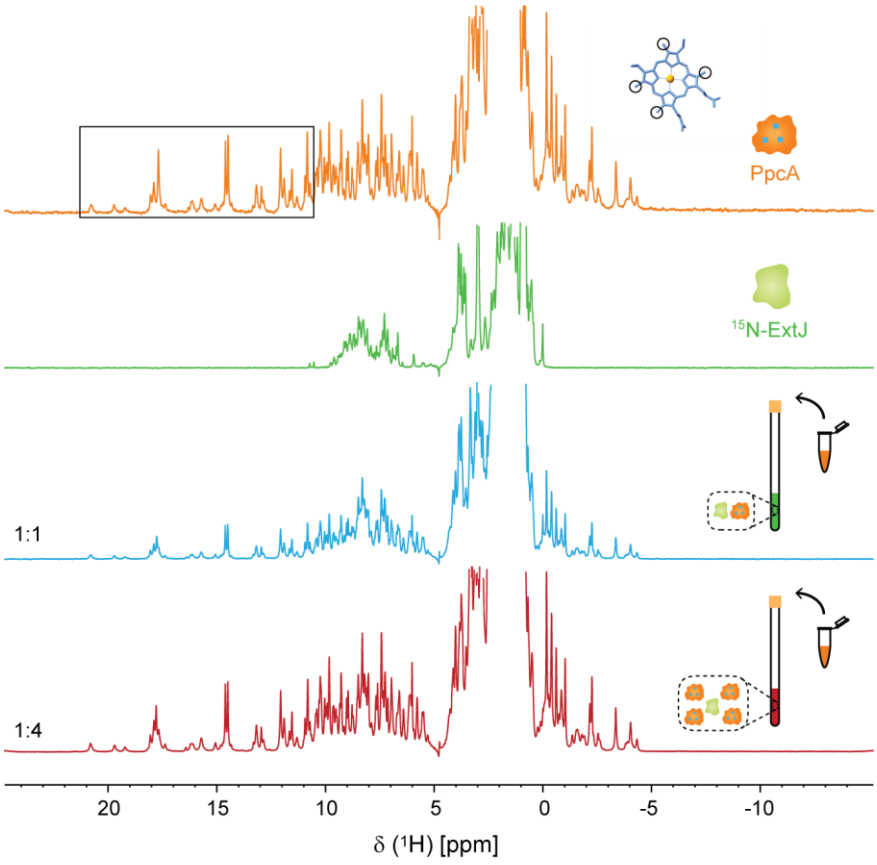


Figure 3.23. Monitorization of ExtJ-PpcA interaction by 1D ^{1}H-NMR. The 1D ^{1}H-NMR spectra of PpcA in the oxidized state (orange) and ExtJ C81* with no affinity tag (green) are shown on top. The rectangle highlights the spectral region in which most of the heme methyl signals of PpcA are located. These signals are highlighted (spheres) on the heme structure inset (blue). The 1D ^{1}H-NMR spectra of 1:1 (blue) and 1:4 (red) ExtJ-PpcA proportions are shown on the bottom.

The context of this NMR protein-protein titration is unique compared to those usually observed in the literature, since the joint analysis of the variations on the 1D ^{1}H-NMR spectrum of PpcA and the 2D ^{1}H, ^{15}N-HSQC spectrum of ExtJ allows the mapping of the interacting regions of both proteins. On one hand, the 1D ^{1}H-NMR spectrum allows the visualization of the progressive chemical shift changes of

the heme methyl groups of PpcA along the titration, by inspection of the 12-21 ppm spectral region (Figure 3.23). In this spectral region, only the resonances of PpcA are visible, since the ^{1}H resonances of ExtJ are restricted to the 0-11 ppm spectral window. PpcA is a small triheme cytochrome and in the oxidized state, each heme iron atom possesses a low-spin (S = ½) unpaired electron that produces a paramagnetic effect. This effect is sensed by nuclei close to the paramagnetic center and results in an increase in the NMR spectral window and signal width due to faster T_2 relaxation times [116]. Therefore, the resonances of the heme methyls of PpcA are easily distinguishable from those of the ExtJ protons, and by monitoring their variation along the titration, one can determine which region of PpcA is interacting with ExtJ. On the other hand, by analyzing the 2D ^{1}H, ^{15}N-HSQC spectrum of ExtJ, only the chemical shift changes of ExtJ's resonances will be monitored, and the residues involved in the interaction with PpcA can be identified. This information, while sufficient to determine a docking pose of the putative ExtJ-PpcA complex, could be complemented with additional studies in which the interaction would be monitored through analysis of chemical shift variations in the 2D ^{1}H, ^{15}N-HSQC spectrum of PpcA, thus completing the picture of the binding interface between ExtJ and PpcA.

The analysis of the 1D ^{1}H-NMR spectra of the 1:0, 1:1 and 1:4 ExtJ-PpcA proportions (Figure 3.23) shows that there are no observable chemical shift variations on the heme methyl resonances of PpcA, a first indication that PpcA and ExtJ may not interact. This is confirmed by the analysis of the 2D ^{1}H, ^{15}N-HSQC of the same proportions of ExtJ-PpcA (Figure 3.24), which show no chemical shift variation of ExtJ's backbone resonances.

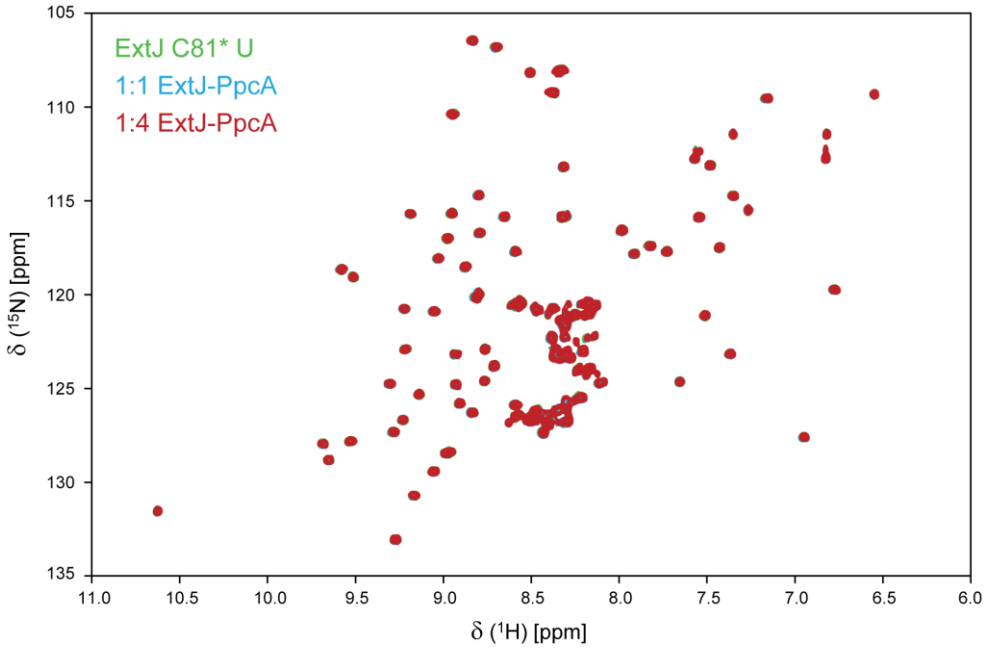


Figure 3.24. Monitorization of ExtJ-PpcA interaction by 2D ^{1}H, ^{15}N-HSQC. The 2D ^{1}H, ^{15}N-HSQC spectra of ExtJ C81* (green), 1:1 (blue) and 1:4 (red) ExtJ-PpcA proportions are shown.

The presence of a 1:4 excess of PpcA with no chemical shift variations observed confirms that ExtJ does not interact with PpcA. Analogously, it is unlikely that ExtJ interacts with any of the other PpcA-

family cytochromes, considering their high structural homology with PpcA [117-121]. Overall, the results indicate that ExtJ is unlikely to serve as a docking moiety for periplasmic cytochromes in the ExtHIJKL porin-cytochrome complex.

Structural contextualization of ExtJ in the ExtHIJKL porin-cytochrome complex

To contextualize the role of ExtJ within the ExtHIJKL porin-cytochrome complex, several AlphaFold models were generated for the individual proteins and for different combinations of the four complex components, using AlphaFold-Multimer [64]. The individual AlphaFold models of the ExtH, ExtI and ExtKL proteins and their predicted cellular locations served as anchor points for the interpretation of the AlphaFold-Multimer models. From these models, the most plausible scenarios indicate that ExtJ interacts either with ExtI or ExtKL through its C-terminal cysteine (Figure 3.25).

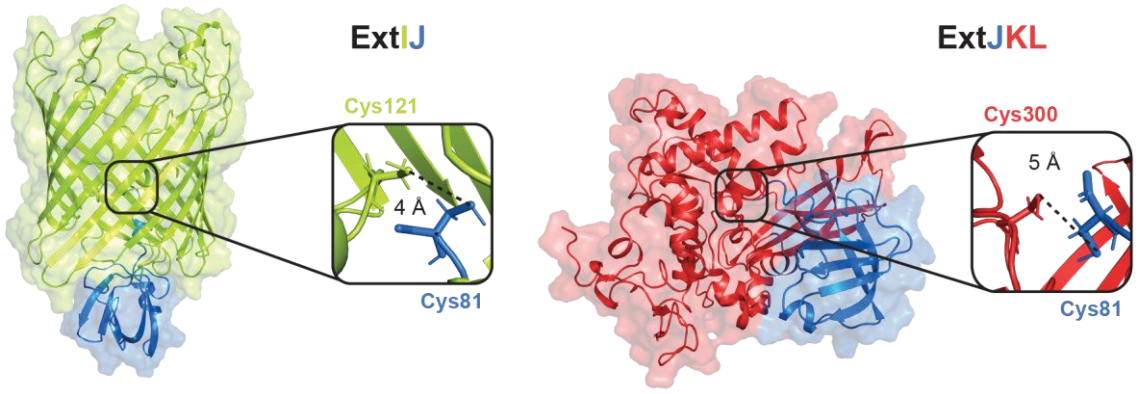


Figure 3.25. Structural contextualization of ExtJ in the ExtHIJKL complex. The ExtI (green), ExtJ (blue) and ExtKL (red) proteins are represented in the AlphaFold models of ExtIJ (left) and ExtJKL (right). The insets show the cysteine residues involved in the putative disulfide bridges formation. The dashed line connects the sulfur atoms of the two cysteine residues involved in each disulfide bridge. The pLDDT scores of each model can be found in Figure 7.7.

In these simulations, ExtJ is positioned to establish a disulfide bridge via Cys^{81} with either Cys^{121} of ExtI or Cys^{300} of ExtKL. In the ExtIJ model, the distance between the sulfur atoms of the cysteines involved on the putative disulfide bridge is 4 Å. Ideally, the S-S distance on a disulfide bridge is around 3 Å [122]. Hence, the Cys^{81} of ExtJ and Cys^{121} of ExtI are oriented at an almost ideal distance for establishment of the S-S covalent bond (Figure 3.25). In the ExtJKL model, the distance between the sulfur atoms of the cysteines involved on the putative disulfide bridge is slightly higher, 5 Å, making it less likely to occur (Figure 3.25). Additionally, based on our assessments regarding ExtKL (see section 2.3.4), we know that Cys^{300} is actually a selenocysteine residue, which AlphaFold is unable to model, making, nonetheless, the formation of a disulfide bridge between ExtJ and ExtKL not possible.

In general, while one must consider that these are in silico simulations, the consistent positioning of ExtJ in the different models calculated with the different proteins of the ExtHIJKL complex is an indication that the process of homodimer formation we observed in vitro is unlikely to occur in vivo, since Cys^{81} might no longer be available to form homodimer assemblies. Alternatively, there might

be a mechanism in which the oligomerization state of ExtJ is regulated by the oxidation/reduction of the intermolecular disulfide bridge [123-128], with the protein working either as a free dimer in the periplasm, or as a monomer bound to ExtI. This dynamic mechanism would have direct implication on the functional role of the ExtHIJKL porin-cytochrome complex. Curiously, rhodanese proteins, such as ExtH, have been suggested to play a role in disulfide bond formation in prokaryotes [127].

Finally, a recent indexation of ExtJ motifs in the Pfam database [129] indicates that the N-terminal region of the protein (residues 1-42) belongs to the PF13624 family, comprised of proteins containing a SurA-like N-terminal domain. SurA is a molecular chaperone that facilitates the folding of outer membrane porins [130-132], a function that fits quite nicely with the ExtI-ExtJ interaction mechanism. Nevertheless, not only is the SurA N-terminal domain mostly composed by α-helices, contrasting with ExtJ, but the entire protein contains additional domains that also play a functional role [131].

3.4. Conclusions

The final steps of EET in the *G. sulfurreducens* bacterium are performed by either extracellular cytochromes, cytochrome nanowires or porin-cytochrome complexes, with the last being involved in short-range reduction of extracellular acceptors [115]. Despite their critical importance in such processes, to date, there were no biochemical characterization studies regarding neither of their protein components, leaving a considerable gap in our atomistic understanding of the multiple pathways that allow electrons to reach the extracellular environment in this bacterium.

The most common structural organization of a bacterial porin-cytochrome complex involves a periplasmic cytochrome, a transmembrane porin and an extracellular cytochrome, responsible for directly reducing extracellular acceptors. Such organization is observed on the OmabcB/C, ExtBCD and ExtEFG complexes, but the four-component ExtHIJKL complex must have a necessarily different arrangement, with unique properties and function. In this chapter, multiple biophysical techniques were used and complemented with computational methods to unravel the biochemical properties of ExtJ, a small periplasmic protein with no cofactors.

The solution structure of ExtJ shows that the protein contains a five anti-paralell stranded β-barrel, organized in a Greek-key topology, together with a single α-helix. Additionally, the protein contains a highly flexible C-terminal stretch, within which there is a terminal cysteine residue (Cys^{81}) available for the establishment of an intermolecular disulfide bridge that promotes homodimerization. In fact, a mutated version of the protein lacking this cysteine residue (C81*) is unable to form homodimers.

The combined use of ^{15}N NMR relaxation experiments, protein structure prediction methods, MD simulations, MST measurements and SAXS analysis, of both monomeric and dimeric samples of ExtJ, show that the protein homodimerization is solely driven by the formation of one intermolecular disulfide bridge.

A detailed structural analysis shows that ExtJ is homologous to domains of riboflavin synthases and other flavin-binding proteins, possessing remarkably similar folding and a putative binding pocket for flavins. Nevertheless, biomolecular interaction studies followed by NMR show that the protein is not able to bind FMN, which is surprising considering the structural conservation observed. In fact,

simulations performed with the recently released AlphaFold 3 algorithm [133], which allows for protein structure prediction in the presence of a limited list of ligands, suggest that flavin adenine dinucleotide (FAD – a slightly larger flavin-containing molecule than FMN, and the only flavin available from the list of ligands in the AlphaFold 3 server) binds to ExtJ in the predicted binding pocket, as FMN or riboflavin would likely do in similar simulations (Figure 3.26).

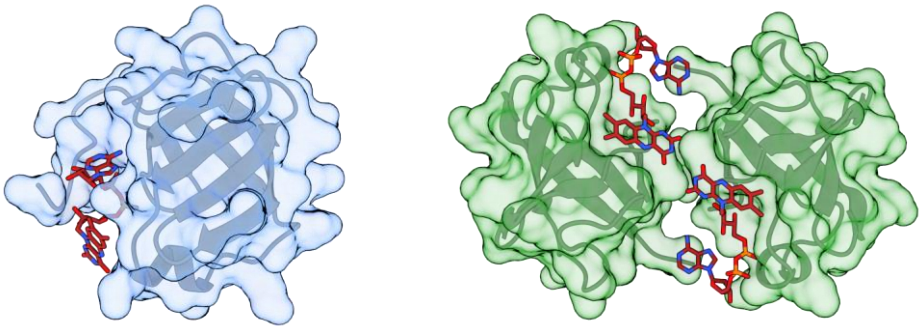


Figure 3.26. AlphaFold 3 model predictions of putative ExtJ-FAD complexes. The monomer (blue) and dimer (green) models of ExtJ are shown in complex with one and two molecules of FAD, respectively. The binding pockets are conserved in the two structures. The pLDDT scores of each model can be found in Figure 7.8.

These observations are a great testimony of how simulation methods should be complemented with experimental approaches to gather valuable insights into protein-ligand interactions, especially if the pharmaceutical aspect of the field is considered.

Additional biomolecular interactions monitored by NMR showed that ExtJ is unable to interact with PpcA, meaning that the protein does not work as a docking platform for periplasmic cytochromes.

Finally, AlphaFold-Multimer simulations show that ExtJ might interact with ExtI through the Cys^{81} residue, by establishing a disulfide bridge with Cys^{121}. We hypothesize that such association might be dependent on mechanisms of disulfide bond formation and cleavage, which indirectly regulate the passage of small molecules, such as redox mediators, through ExtI [134-136].

Overall, we were unable to precisely define the function of ExtJ in the cell, but the structural and biochemical data provided certainly narrowed down the different possibilities.

3.5. Materials and methods

3.5.1. DNA manipulation

The *gsu2938* gene sequence, encoding for the ExtJ protein (WP_010943569.1), was retrieved from *G. sulfurreducens* PCA's genome on the Kyoto Encyclopedia of Genes and Genomes (KEGG) database [93], under the accession number T00155. The codons coding for residues 24-104 (excluding the signal peptide predicted by the SignalP 6.0 software [137]) were amplified from genomic DNA and inserted into the pET-28a(+) plasmid using Phusion High-Fidelity DNA Polymerase (Thermo Fisher Scientific) together with the primers listed in Table 2.5, according to the RF cloning protocol [138]. Taq DNA polymerase (VWR) was used for colony PCR screening and DpnI (Thermo Fisher Scientific) for digestion of the template DNA. The intermediate PCR product and the final plasmid DNA were purified using the NZYGelpure and NZYMiniprep kits (NZYTech), respectively. PCR products and plasmid DNA quantification and purity were determined using a NanoDrop spectrophotometer ND-1000 (Thermo Fisher Scientific) and the DNA was sequenced by STAB VIDA (Caparica, Portugal). The final plasmid was designated pET-28a(+)-ExtJ.

The primers used for the substitution of the C-terminal cysteine (Cys^{81}) for a stop codon were designed using the QuikChange Primer Design program (Agilent Technologies) and are listed in Table 3.4. The mutation was introduced by following the instructions of the NZYMutagenesis kit (NZYTech) and the presence of the desirable mutation was confirmed by DNA sequencing.

Table 3.4. Sequences of the primers used to produce the pET-28a(+)-ExtJ C81* vector. The substitution sites are underlined. The primers forward (fw) and reverse (rv) were purchased from Invitrogen.

Primer	DNA sequence (5' – 3')
ExtJ_C81*_fw	CGCTCTCCAGGGCTAGTAGCTCGAGCACCA
ExtJ_C81*_rv	TGGTGCTCGAGCTACTAGCCCTGGAGAGCG

3.5.2. Protein sequence analysis and AlphaFold predictions

A Protein BLAST [9] was performed with ExtJ's sequence (WP_010943569.1), using the RefSeq Select protein sequences database and the default blastp algorithm. Sequences were aligned with Clustal Omega [6] and represented in Geneious 1.1.2022.

The AlphaFold models of ExtJ were either retrieved from the AlphaFold protein structure database [139, 140], or produced using either the AlphaFold2_MMSeqs2 GitHub ColabFold notebook [66] or the AlphaFold 3 server [133]. All AlphaFold-Multimer models were produced using the same notebook [64]. The quality of the individual models was evaluated based on the per-residue estimate of confidence produced by AlphaFold, designated pLDDT, which is based on the lDDT-Cα metric [141]. Additionally, intra and inter-PAE scores produced by AlphaFold were used to evaluate the quality of the relative position of proteins in complexes [66]. The interactive analysis of these values was performed using PAE Viewer [65]. The final models were analyzed and represented in UCSF ChimeraX 1.4 [67], without the corresponding signal peptides, as predicted by SignalP 6.0 [137].

3.5.3. Protein expression and purification

The natural abundance and ^{13}C, ^{15}N labeled ExtJ WT and ExtJ C81* proteins were produced and purified with the same protocol. In the case of the ExtJ WT protein, reducing agents were used throughout the purification process.

The proteins were heterologously expressed in *E. coli* BL21 Star (DE3) cells, transformed with pET-28a(+) plasmids encoding a kanamycin resistance marker and either the ExtJ WT or ExtJ C81* proteins with an N-terminal 6x His-tag and a thrombin cleavage site. Initially, cells were grown at 37 ºC to an OD_{600} of 0.6-0.8 in either LB medium for natural abundance samples or M9 minimal medium for ^{13}C, ^{15}N isotopically labeled samples, both supplemented with 30 μg/mL kanamycin. Then, protein expression was induced with 1 mM of IPTG (NZYTech) and cells were grown for 4 hours at 37 ºC before being harvested by centrifugation at 6500 x*g* for 15 minutes. Cells were resuspended in 30 mL per liter of cell culture of 50 mM Tris-HCl pH 7.5 with catalytic amounts of DNase I, 0.5 mg/mL of lysozyme, 1 mM phenylmethanesulfonylfluoride (PMSF) and 2 mM benzamidine, and disrupted by a combined method of three freeze and thaw cycles followed by 15 cycles of ultrasonication (1 minute on plus 2 minutes off) performed with an ultrasonic homogenizer Branson SFX 150 regulated for 65% of total amplitude. Cell debris were removed by centrifugation at 50000 x*g* for 1 hour and the resulting supernatant, containing the soluble cell extract, was loaded onto a 5 mL Ni-HisTrap™ HP column (Cytiva) previously equilibrated with 50 mM Tris-HCl pH 7.5, 30 mM imidazole, 500 mM NaCl and 2 mM β-mercaptoethanol. The proteins were eluted with a 0-500 mM imidazole gradient and the fractions containing ExtJ were dialyzed against 1 L of 1x PBS pH 7.4 using dialysis membranes with a molecular weight cut off of 3.5 kDa (Spectra). The N-terminal affinity tag was digested with 1 unit of thrombin (Cytiva) per 100 μg of protein for 16 hours and the resulting solution was loaded onto a 5 mL Ni-HisTrap™ HP column (Cytiva) for an affinity chromatography performed as described above. The cleaved proteins, eluted in the flow-through of this purification step, were then injected onto a Superdex™ 75 Increase 10/300 GL molecular exclusion column (Cytiva), equilibrated with 100 mM sodium phosphate buffer pH 8 and 1 mM DTT. For production of proteins with the N-terminal affinity tag, the intermediate step of protein cleavage was not performed. Protein purity was evaluated by BlueSafe (NZYTech) stained SDS-PAGE. Protein concentration was determined by measuring the absorbance at 280 nm in a NanoDrop spectrophotometer ND-1000 (Thermo Fisher Scientific), and by using the molar extinction coefficient of 5.5 mM^{-1} cm^{-1}, determined by the ExPASy ProtParam software [7]. The molecular weight of the different ExtJ proteins produced was confirmed by MALDI-TOF-MS. The analysis was performed by the Mass Spectrometry Unit (UniMS) of ITQB/iBET (Oeiras, Portugal).

For interaction studies with ExtJ, the cytochrome PpcA was expressed, purified and quantified as previously described [142].

3.5.4. NMR spectroscopy

The NMR experiments were acquired either on a Bruker Avance Neo 500 MHz spectrometer equipped with a 5 mM triple resonance Prodigy cryoprobe (TCI) or on a Bruker Avance III 600 MHz spectrometer equipped with a 5 mM triple resonance cryoprobe (TCI). The ^{1}H chemical shifts are referenced to sodium trimethylsilylpropanesulfonate (DSS) at 0 ppm, either by adding DSS to the samples to a final concentration of 100 μM or by using the residual water signal, as previously described [143]. The ^{13}C and ^{15}N chemical shifts were calibrated through indirect referencing [144]. All spectra were acquired at 25 °C (unless stated otherwise), processed using TopSpin 4.1.4™ (Bruker BioSpin, Karlsruhe, Germany), and analyzed either with the processing software or with the programs CARA [145] and NMRFAM-Sparky [146]. All NMR experimental setups used are part of the Bruker standard pulse sequence library.

Sequence-specific backbone and side-chain resonance assignment

The samples of ^{13}C, ^{15}N ExtJ WT used for the acquisition of the NMR experiments required for the backbone assignment, side-chain assignment, structure calculation and relaxation experiments were prepared with 1 mM of protein in 32 mM sodium phosphate buffer pH 6 (100 mM of final ionic strength) and 0.03% sodium azide (90% H_2O and 10% D_2O). A concentration of 1 mM DTT was used to prepare monomer samples.

The backbone sequential resonances were assigned using 2D ^{1}H, ^{15}N-HSQC, 3D HNCO [15], 3D HN(CA)CO [16], 3D HNHA [17], 3D HN(CO)CACB [18] and 3D HNCACB experiments [19]. 2D ^{1}H, ^{15}N-HSQC spectra were acquired with 2048 (t_2) x 128 (t_1) data points, covering a sweep width of 8.2 kHz in the ^{1}H dimension and 1.8 kHz in the ^{15}N dimension, with 64 scans per increment. 3D HNCO and 3D HN(CA)CO spectra were acquired with 2048 (t_3) x 40 (t_2) x 128 (t_1) data points, covering a sweep width of 8.2 kHz in the ^{1}H dimension, 1.8 kHz in the ^{15}N dimension and 2.5 kHz in the ^{13}C dimension, with 8 scans per increment. 3D HNHA spectra were acquired with 2048 (t_3) x 128 (t_2) x 40 (t_1) data points, covering a sweep width of 8.2 kHz in the ^{1}H dimension and 1.8 kHz in the ^{15}N dimension, with 16 scans per increment. Finally, 3D HNCOCACB and 3D HNCACB spectra were acquired with 2048 (t_3) x 40 (t_2) x 128 (t_1) data points, covering a sweep width of 8.2 kHz in the ^{1}H dimension, 1.8 kHz in the ^{15}N dimension and 10.1 kHz in the ^{13}C dimension, with 8 scans per increment. Side-chain resonances were assigned using 3D ^{1}H, ^{15}N-TOCSY-HSQC [38] and 3D (H)CCH-TOCSY experiments [39]. 3D ^{1}H, ^{15}N-TOCSY-HSQC spectra were acquired with a mixing-time of 60 ms, collecting 2048 (t_3) x 40 (t_2) x 128 (t_1) data points, covering a sweep width of 8.2 kHz in the ^{1}H dimension and 1.8 kHz in the ^{15}N dimension, with 16 scans per increment. 3D (H)CCH-TOCSY spectra were acquired with a mixing-time of 60 ms, collecting 2048 (t_3) x 64 (t_2) x 128 (t_1) data points, covering a sweep width of 8.2 kHz in the ^{1}H dimension and 10.1 kHz in the ^{13}C dimension, with 8 scans per increment.

Secondary structure prediction and solution structure determination

The secondary structure elements of ExtJ were identified by analysis of the ^{13}CO, ^{15}N, ^{1}HN, $^{13}C_{\alpha}$, $^{1}H_{\alpha}$ and $^{13}C_{\beta}$ chemical shifts through different softwares, including CSI 3.0 [24], Talos+ [25] and Rosetta CS [28-30]. The latest was also used to predict a structural model of the protein. The secondary $^{13}C_{\alpha}$ and $^{13}C_{\beta}$ chemical shifts were calculated relative to the sets of standard BMRB chemical shifts [22, 23] and sequence corrected random coil chemical shifts [26]. In both cases, a 1-2-1 weighting function for residues (i – 1) – i – (i + 1) was applied to the raw data to reduce noise and highlight regular secondary structure elements [27].

Distance constraints for solution structure determination were collected from 3D ^{13}C-NOESY [38, 40] and 3D ^{15}N-NOESY experiments [38, 40, 41]. 3D ^{13}C-NOESY spectra were acquired were acquired with a mixing-time of either 120 or 150 ms, collecting 2048 (t_3) x 64 (t_2) x 128 (t_1) data points, covering a sweep width of 8.2 kHz in the ^{1}H dimension and 10.1 kHz in the ^{13}C dimension, with 16 scans per increment. 3D ^{15}N-NOESY spectra acquired were acquired with a mixing-time of 120 ms, collecting 2048 (t_3) x 40 (t_2) x 120 (t_1) data points, covering a sweep width of 8.2 kHz in the ^{1}H dimension and 1.8 kHz in the ^{15}N dimension, with 8 scans per increment [20]. Automated NOE assignment and structure calculations were carried out with the standard protocol in CYANA 3.98.13 [42, 43], implemented within the ARTINA approach [44], using the NMRTist platform [45]. This approach uses convolutional neural networks for cross-peak detection [46], FLYA automated chemical shift assignment [47], and CYANA for automated NOE assignment and structure calculation [42, 43]. All chemical shift assignments were manually checked. The structural ensemble was refined in the AMPS-NMR portal using the predefined "restrainedMD" protocol for restrained molecular dynamics in a hydrated environment [48], the NOE-derived distance restraints from the CYANA calculations and dihedral angle restraints derived from Talos+. Each chain of the protein ensemble was solvated with a 10 Å TIP3P box of explicit water model [147]. Counterions were added to neutralize the system. The MD simulations were carried out with the Amber99SB chain force field [148, 149]. The final set of structures was validated with PROCHECK-NMR [150, 151], using the PDBsum platform [152].

Backbone dynamics

Heteronuclear ^{1}H and ^{15}N relaxation experiments of ExtJ WT were acquired at 500 and 600 MHz spectrometer frequencies. For the determination of the longitudinal relaxation (T_1) rates, a set of 17 experiments was acquired with relaxation delays between 10 and 5000 ms, collecting 2048 (t_2) x 128 (t_1) data points, covering a sweep width of 8.2 kHz in the ^{1}H dimension and 1.8 kHz in the ^{15}N dimension, with 12 scans per increment. For the determination of the transverse relaxation (T_2) rates, a set of 17 experiments was acquired with CPMG loops between 16 and 543 ms, collecting 2048 (t_2) x 128 (t_1) data points, covering a sweep width of 8.2 kHz in the ^{1}H dimension and 1.8 kHz in the ^{15}N dimension, with 8 scans per increment. The heteronuclear ^{1}H-^{15}N NOE ratios were determined by acquiring experiments in the presence and absence of ^{1}H saturation, with an overall recycling delay of 10 s to ensure maximal development of NOEs before acquisition, while allowing solvent relaxation

[52], collecting 2048 (t_2) x 256 (t_1) data points, covering a sweep width of 8.2 kHz in the ^{1}H dimension and 1.8 kHz in the ^{15}N dimension, with 32 scans per increment.

The T_1 and T_2 rates for each backbone NH signal of ExtJ were determined by fitting exponential decays to the data, as a function of the relaxation delays [52]. The heteronuclear ^{1}H-^{15}N NOE values of each backbone NH signal were defined as the ratios of the average peak intensities in the presence and absence of ^{1}H saturation [52]. The uncertainties of the HetNOE values, ΔNOE, were calculated based on the measured background noise using Equation 3.1:

$$\Delta NOE = \sqrt{\left(\frac{\Delta I_{sat}}{I_{sat}}\right)^2 + \left(\frac{\Delta I_{unsat}}{I_{unsat}}\right)^2} \quad (3.1)$$

In this equation, I_{sat} and I_{unsat} represent the measured peak intensities in the presence and absence of ^{1}H saturation, respectively. The standard deviations of the HetNOE values were calculated from the root-mean square of the noise in background regions, defined as ΔI_{sat} and ΔI_{unsat}. Peak intensity data was processed using CARA [145] and data fitting was performed in Origin Pro 8.5.

Correlation time, molecular mass, oligomerization state and hydrodynamic radius calculation

Considering an approximately spherical globular protein, the rotational correlation time (τ_c) is related to its effective hydrodynamic radius (a), according to an empirically derived Stokes-Einstein equation [50] (Equation 3.2):

$$\tau_c \approx \frac{4\pi\eta a^3}{3kT} \quad (3.2)$$

In this equation, η is the viscosity of water (0.89×10^{-3} kg m^{-1} s^{-1}), k is the Boltzmann's constant (1.38×10^{-23} kg m^2 s^{-2} K^{-1}) and T is the temperature (298 K). Moreover, in the limit of slow molecular motion ($\tau_c \gg 0.5$ ns), the τ_c of a protein is related to the ratio of the longitudinal T_1 and transverse T_2 ^{15}N relaxation times and nuclear frequency (υ_N) according to Equation 3.3 [59, 60]:

$$\tau_c \approx \left(\sqrt{\frac{6T_1}{T_2} - 7}\right) / 4\pi\upsilon_N \quad (3.3)$$

In this equation, T_1 and T_2 are the average of all ^{15}N backbone amide longitudinal and transverse relaxation rates (ms) for those residues without excessive exchange contributions, selected based on the analysis of their T_1/T_2 ratios [63].

The determined correlation times were compared with a existent standard curve of τ_c (ns) vs protein molecular weight (kDa) – Figure 3.13 – obtained for a series of known monomeric proteins of varying size, and used to get an estimation of the molecular weight and oligomerization state of ExtJ [61, 62].

Structural influence of the C81* mutation and N-terminal affinity tag

The samples of ^{15}N ExtJ WT and ^{15}N ExtJ C81* (with and without the N-terminal affinity tag) used for the acquisition of the NMR experiments for evaluation of the structural influence of the C81* mutation and N-terminal affinity tag were prepared with 500 μM of protein in 32 mM sodium phosphate buffer pH 6 (100 mM of final ionic strength), 90% H_2O and 10% D_2O. The ExtJ WT samples were prepared with 1 mM DTT. 2D ^{1}H, ^{15}N-HSQC spectra were acquired with 2048 (t_2) x 128 (t_1) data points, covering a sweep width of 8.2 kHz in the ^{1}H dimension and 1.8 kHz in the ^{15}N dimension, with 8 scans per increment. Analogously, 3D ^{1}H, ^{15}N-TOCSY-HSQC and 3D ^{1}H, ^{15}N-NOESY-HSQC spectra were acquired with 1.35 mM protein samples prepared in the conditions described above. 3D ^{1}H, ^{15}N-TOCSY-HSQC spectra were acquired with a mixing-time of 60 ms, collecting 2048 (t_3) x 40 (t_2) x 128 (t_1) data points, covering a sweep width of 8.2 kHz in the ^{1}H dimension and 1.8 kHz in the ^{15}N dimension, with 16 scans per increment. 3D ^{1}H, ^{15}N-NOESY-HSQC spectra were acquired with a mixing-time of 120 ms and with the same parameters of the 3D ^{1}H, ^{15}N-TOCSY-HSQC experiments.

The chemical shifts of the different proteins were analyzed following a previously described strategy [32]. The combined chemical shift changes of each protein amino acid residue were calculated according to Equation 3.4:

$$\Delta\delta = \sqrt{(\Delta\delta H)^2 + (w_i \Delta\delta N)^2} \tag{3.4}$$

In this equation, ΔδH and ΔδN are the chemical shift variations in the ^{1}H and ^{15}N dimensions, respectively. The term $w_i = |\gamma^{15}N/\gamma^1H| = 0.1$ is a weighting factor that accounts for the differences in nuclei sensitivity [32]. An iterative procedure was utilized to calculate a cut-off value for the observed chemical shift differences, based on a corrected standard deviation to zero [32].

Biomolecular interactions with FMN and PpcA

The natural abundance PpcA and ^{15}N ExtJ WT/C81* samples were prepared with 5 mM and 100 μM protein, respectively, in 32 mM sodium phosphate buffer pH 6 (100 mM of final ionic strength), 90% H_2O and 10% D_2O. FMN was prepared in 32 mM sodium phosphate buffer pH 6 (100 mM of final ionic strength), with a final concentration of 5 mM. UV-visible spectroscopy and previously determined molar extinction coefficients were used to quantify PpcA (ε_{552nm} = 97.5 mM^{-1} cm^{-1} [153]) and FMN (ε_{380nm} = 10.2 mM^{-1} cm^{-1} [154, 155]).

The interaction between PpcA and ^{15}N ExtJ C81* (100 μM) was monitored by acquiring 1D ^{1}H and 2D ^{1}H, ^{15}N-HSQC NMR spectra, after addition of increasing amounts of PpcA up to a final ratio of 1:4 PpcA:^{15}N ExtJ C81* was achieved. 1D ^{1}H-NMR spectra were acquired with 33k data points, covering a sweep width of 20 kHz, with 128 scans per increment. 2D ^{1}H, ^{15}N-HSQC spectra were acquired with 2048 (t_2) x 128 (t_1) data points, covering a sweep width of 8.2 kHz in the ^{1}H dimension and 1.8 kHz in the ^{15}N dimension, with 8 scans per increment.

The interaction between ^{15}N ExtJ WT (100 μM) and FMN was monitored by acquiring 1D ^{1}H and 2D ^{1}H, ^{15}N-HSQC NMR spectra, after addition of increasing amounts of FMN (1:0.2, 1:0.4, 1:0.6, 1:0.8, 1:1,

1:2, 1:4, 1:6, 1:8 and 1:10 of ^{15}N ExtJ WT:FMN). 1D ^{1}H-NMR spectra were acquired with 16k data points, covering a sweep width of 8.2 kHz, with 32 scans per increment. 2D ^{1}H, ^{15}N-HSQC spectra were acquired with 2048 (t_2) x 128 (t_1) data points, covering a sweep width of 8.2 kHz in the ^{1}H dimension and 1.8 kHz in the ^{15}N dimension, with 8 scans per increment.

3.5.5. CD spectroscopy

CD experiments were performed with an Applied Photophysics Chirascan qCD spectropolarimeter (Leatherhead, UK), equipped with a thermostatic cell support and a 0.2 mm path length quartz cell (Hellma). CD spectra were acquired in the far UV region using 50-60 μM of ExtJ samples (WT and C81*, with and without the N-terminal affinity tag) prepared on 10 mM sodium phosphate buffer pH 8. The ExtJ WT samples were also prepared with 1 mM β-mercaptoethanol. The spectra were the average of three scans recorded in 1 nm steps with a scan rate of 3 seconds per nanometer, over a 190 to 260 nm wavelength range. The buffer contribution was corrected for all spectra and the temperature was controlled to ± 1 °C. After recording a spectrum at 25 °C, the conformational stability of the ExtJ proteins was assessed by performing linear thermal-induced denaturation with a heating rate of 1 °C per minute, from 10 to 94 °C. Each spectrum was recorded with a scan rate of 0.4 seconds per nanometer. After the temperature ramp, the sample was quickly cooled to 25 °C and then 10 °C, at which two final spectra were registered. The variation of the ellipticity at 196 nm with temperature was fitted to a two-state transition model [37, 156], used to determine the midpoint of unfolding transition (T_M) and the enthalpy of unfolding (ΔH) of the different ExtJ proteins, according to Equation 3.5:

$$[\theta] = \left(([\theta]_F - [\theta]_U) \times \left(\frac{e^{\left(\frac{\Delta H}{RT} \times \frac{T}{T_M - 1}\right)}}{1 + e^{\left(\frac{\Delta H}{RT} \times \frac{T}{T_M - 1}\right)}} \right) \right) + [\theta]_U \qquad (3.5)$$

In this equation, $[\theta]$, $[\theta]_F$ and $[\theta]_U$ correspond to the molar ellipticity (deg. cm^2 dmol^{-1}) at a certain temperature (T), in the fully folded and fully unfolded states, respectively; and R is the molar gas constant (8.314 J K^{-1} mol^{-1}).

The BeStSel deconvolution method [33-35] was used to estimate the secondary structure elements of the different ExtJ proteins.

3.5.6. Molecular dynamics simulations

All-atom MD simulations of the ExtJ homodimer were performed at 298 K using GROMACS 2020 [157, 158] with the Amber99SB-ILDN force field [159] and the TIP3P water model [147]. The AlphaFold model of the homodimer was soaked in a water box with a total volume of 306 nm^3, extending 1.2 nm from the protein surface in each Cartesian direction, and containing the amount of counterions (eight Cl^- ions) necessary for system neutrality. The electrostatic energy was calculated using the Particle-Mesh Ewald (PME) algorithm [160-162], with the cut-off distances for Coulomb and van der Waals interactions set to 10 Å. The equations of motion were integrated using 2 fs timesteps. The

system was energy-minimized in two steps to remove atom clashes and bond contacts: first by a steepest descent minimization [163] (2000 steps), followed by a conjugated gradient (1000 steps). The solvent molecules were equilibrated for 100 ps at 298 K with a Nose-Hoover thermostat [164] (time-constant of 1.6 ps) and 1 atm pressure using a Parrinello-Rahman barostat [165, 166] (time-constant of 5 ps). Holonomic positional restrains were employed for H-bonded atoms with the LINCS algorithm [167], using a force constant of 1000 kJ/mol/nm. A second equilibration step, without positional restrains, was performed during an additional 500 ps at 298 K. The simulation run was made for a total of 250 ns, using the same parameters. Trajectory coordinates, velocities and energies were saved every 20, 2 and 4 ps, respectively. To monitor the simulation, potential energy, total energy, temperature and pressure values were analyzed throughout the equilibration and simulation times using GROMACS tools. The final results were further processed and analyzed in GROMACS and VMD version 1.9.3 [168].

3.5.7. Microscale thermophoresis

ExtJ C81* was labeled on the N-terminal His-tag with a non-covalent RED-tris-NTA 2nd generation dye (NanoTemper Technologies) in 1x PBS pH 7.5 with 0.05% Tween-20, following the manufacturer's instructions for cases in which the K_D between the target protein and the dye is lower than 10 nM, as confirmed preliminary by testing the protein's affinity to the dye. Briefly, 90 μL of 200 nM ExtJ C81* were mixed with 90 μL of 100 nM RED-tris-NTA 2nd generation dye, incubated for 30 minutes at room temperature and subsequently centrifuged for 10 minutes at 4 °C and 15000 x*g*. The stock solution of the labeled sample had a final concentration of 100 nM and an approximate 2:1 protein:dye ratio. The fluorescent protein samples for MST were prepared with a constant concentration of 50 nM of labeled ExtJ C81* and mixed in a serial dilution with unlabeled ExtJ C81* (with no N-terminal His-tag) in a concentration range between 15 nM and 500 μM. The samples, all prepared in 1x PBS pH 7.5 with 0.05% Tween-20, were loaded into standard MST capillaries (Monolith™ NT.115 Series) and measurements were performed on a Monolith NT.115 (NanoTemper Technologies) with a 40% Nano-RED excitation power, a medium MST power and a laser-on time of 20 s. All measurements were conducted in triplicates. No aggregation or photobleaching events were observed during the assays. The MST signal quality was evaluated based on the response amplitude and signal to noise ratio of each measurement. The thermophoresis (F_{norm}) values of the first 5 s after laser irradiation were considered to avoid artifacts due to heating effects. F_{norm} is defined by the ratio of the mean intensity fluorescence before (F_{hot}) and after (F_{cold}) irradiation, according to Equation 3.6:

$$F_{norm} = (F_{hot}/F_{cold}) \times 1000 \quad (3.6)$$

MST experiments were fitted on the MO. Control Version 1.6 analysis software (NanoTemper technologies), using the F_{norm} values represented as a function of ligand concentration (Equation 3.7):

$$F_C = U + (B - U) \times \frac{C + C_{Target} + K_D - \sqrt{(C + C_{Target} + K_D)^2 - 4 \times C \times C_{Target}}}{2 \times C_{target}} \quad (3.7)$$

This equation considers a molecular interaction with a 1:1 stoichiometry according to the law of mass action, without cooperativity [70]. F_C corresponds to the fraction of bound ligand at a certain concentration (C) of ligand. U (unbound) and B (bound) correspond to the F_{norm} signal of the target protein alone and in the complex, respectively. C_{Target} and K_D correspond to the final concentration of target in the assay, and to the dissociation constant of the binding event, respectively.

3.5.8. SAXS experiments

Synchrotron SAXS data (I(s) versus s, where s = 4πsinθ/λ, with 2θ being the scattering angle and λ the X-ray wavelength, 0.9919 Å) were collected from dialyzed solutions of ExtJ WT and the C81* variant in 8 or 32 mM sodium phosphate buffer pH 6 (25 or 100 mM of final ionic strength) with and without 5 mM DTT at the BM29 beamline at ESRF [169] in Grenoble, France. The measurements were conducted using a Pilatus3 2M detector (Dectris) positioned 2.827 meters from the sample. Different concentrations (Table 7.2) of ExtJ were measured at 20 °C using a continuous flow 1 mm cell capillary to minimize radiation damage. For each sample and the corresponding matched buffer, ten successive 1 s frames were collected. Data analysis was performed with modules from the ATSAS 3.2.1 package [170]. CRYSOL [82] was used to calculate the scattering profile from the AlphaFold model of ExtJ, while SREFLEX [83] was used to estimate the flexibility of the model and fit the experimental SAXS data. The ab initio envelopes of ExtJ were derived from the measured scattering curve using DAMMIF [84], and fifteen independently generated ab initio models were aligned, averaged, and filtered with DAMAVER [85].

3.6. References

[1] F.J. Otero, C.H. Chan, D.R. Bond, Identification of different putative outer membrane electron conduits necessary for Fe(III) citrate, Fe(III) oxide, Mn(IV) oxide, or electrode reduction by *Geobacter sulfurreducens*, J. Bacteriol., 200 (2018) e00347.
[2] R.-L. He, J. Wu, Z.-H. Cheng, H.-H. Li, J.-Q. Liu, D.-F. Liu, W.-W. Li, Biomolecular insights into extracellular pollutant reduction pathways of *Geobacter sulfurreducens* using a base editor system, Environ. Sci. Technol., 56 (2022) 12247-12256.
[3] M.I. Jahan, P. Juengwiwattanakitti, Y. Izu, R. Tobe, T. Imai, H. Mihara, Selenite uptake by outer membrane porin ExtI and its involvement in the subcellular localization of rhodanese-like lipoprotein ExtH in *Geobacter sulfurreducens*, Biochem. Biophys. Res. Commun., 516 (2019) 474-479.
[4] M.I. Jahan, R. Tobe, H. Mihara, Characterization of a novel porin-like protein, ExtI, from *Geobacter sulfurreducens* and its implication in the reduction of selenite and tellurite, Int. J. Mol. Sci., 19 (2018) 809.
[5] E. Howley, R. Krajmalnik-Brown, C.I. Torres, Cytochrome gene expression shifts in *Geobacter sulfurreducens* to maximize energy conservation in response to changes in redox conditions, Biosens. Bioelectron., 237 (2023) 115524.
[6] F. Sievers, A. Wilm, D. Dineen, T.J. Gibson, K. Karplus, W. Li, R. Lopez, H. McWilliam, M. Remmert, J. Söding, J.D. Thompson, D.G. Higgins, Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega, Mol. Syst. Biol., 7 (2011) 539.
[7] E. Gasteiger, A. Gattiker, C. Hoogland, I. Ivanyi, R.D. Appel, A. Bairoch, ExPASy: The proteomics server for in-depth protein knowledge and analysis, Nucleic Acids Res., 31 (2003) 3784-3788.
[8] T. Goldberg, M. Hecht, T. Hamp, T. Karl, G. Yachdav, N. Ahmed, U. Altermann, P. Angerer, S. Ansorge, K. Balasz, M. Bernhofer, A. Betz, L. Cizmadija, K.T. Do, J. Gerke, R. Greil, V. Joerdens, M. Hastreiter, K. Hembach, M. Herzog, M. Kalemanov, M. Kluge, A. Meier, H. Nasir, U. Neumaier, V. Prade, J. Reeb, A. Sorokoumov, I. Troshani, S. Vorberg, S. Waldraff, J. Zierer, H. Nielsen, B. Rost, LocTree3 prediction of localization, Nucleic Acids Res., 42 (2014) 350-355.
[9] S. McGinnis, T.L. Madden, BLAST: At the core of a powerful and diverse set of sequence analysis tools, Nucleic Acids Res., 32 (2004) W20-W25.
[10] H. Berman, K. Henrick, H. Nakamura, Announcing the worldwide Protein Data Bank, Nat. Struct. Mol. Biol., 10 (2003) 980.
[11] S.K. Burley, H.M. Berman, C. Bhikadiya, C. Bi, L. Chen, L. Di Costanzo, C. Christie, K. Dalenberg, J.M. Duarte, S. Dutta, Z. Feng, S. Ghosh, D.S. Goodsell, R.K. Green, V. Guranović, D. Guzenko, B.P. Hudson, T. Kalro, Y. Liang, R. Lowe, H. Namkoong, E. Peisach, I. Periskova, A. Prlić, C. Randle, A. Rose, P. Rose, R. Sala, M. Sekharan, C. Shao, L. Tan, Y.-P. Tao, Y. Valasatava, M. Voigt, J. Westbrook, J. Woo, H. Yang, J. Young, M. Zhuravleva, C. Zardecki, RCSB Protein Data Bank: Biological macromolecular structures enabling research and education in fundamental biology, biomedicine, biotechnology and energy, Nucleic Acids Res., 47 (2019) D464-D474.
[12] L. Holm, Dali server: Structural unification of protein families, Nucleic Acids Res., 50 (2022) W210-W215.

[13] M. van Kempen, S.S. Kim, C. Tumescheit, M. Mirdita, J. Lee, C.L.M. Gilchrist, J. Söding, M. Steinegger, Fast and accurate protein structure search with Foldseek, Nat. Biotechnol., 42 (2023) 243-246.

[14] M. Varadi, J. Berrisford, M. Deshpande, S.S. Nair, A. Gutmanas, D. Armstrong, L. Pravda, B. Al-Lazikani, S. Anyango, G.J. Barton, K. Berka, T. Blundell, N. Borkakoti, J. Dana, S. Das, S. Dey, P. Di Micco, F. Fraternali, T. Gibson, M. Helmer-Citterich, D. Hoksza, L.C. Huang, R. Jain, H. Jubb, C. Kannas, N. Kannan, J. Koca, R. Krivak, M. Kumar, E.D. Levy, F. Madeira, M.S. Madhusudhan, H.J. Martell, S. MacGowan, J.E. McGreig, S. Mir, A. Mukhopadhyay, L. Parca, T. Paysan-Lafosse, L. Radusky, A. Ribeiro, L. Serrano, I. Sillitoe, G. Singh, P. Skoda, R. Svodobova, J. Tyzack, A. Valencia, E.V. Fernandez, W. Vranken, M. Wass, J.M. Thornton, M. Sternberg, C. Orengo, S. Velankar, PDBe-KB: A community-driven resource for structural and functional annotations, Nucleic Acids Res., 48 (2020) D344-D353.

[15] L.E. Kay, M. Ikura, R. Tschudin, A. Bax, Three-dimensional triple-resonance NMR spectroscopy of isotopically enriched proteins, J. Magn. Reson., 89 (1990) 496-514.

[16] R.T. Clubb, V. Thanabal, G. Wagner, A constant-time three-dimensional triple-resonance pulse scheme to correlate intraresidue ^{1}HN, ^{15}N, and ^{13}C′ chemical shifts in ^{15}N-^{13}C-labelled proteins, J. Magn. Reson., 97 (1992) 213-217.

[17] G.W. Vuister, A. Bax, Quantitative *J* correlation: A new approach for measuring homonuclear three-bond $J(H^{N}H^{\alpha})$ coupling constants in ^{15}N-enriched proteins, J. Am. Chem. Soc., 115 (1993) 7772-7777.

[18] S. Grzesiek, A. Bax, Correlating backbone amide and side chain resonances in larger proteins by multiple relayed triple resonance NMR, J. Am. Chem. Soc., 114 (1992) 6291-6293.

[19] S. Grzesiek, A. Bax, An efficient experiment for sequential backbone assignment of medium-sized isotopically enriched proteins, J. Magn. Reson., 99 (1992) 201-207.

[20] M. Sattler, J. Schleucher, C. Griesinger, Heteronuclear multidimensional NMR experiments for the structure determination of proteins in solution employing pulsed field gradients, Prog. Nucl. Mag. Res. Sp., 34 (1999) 93-158.

[21] M. Iwadate, T. Asakura, M.P. Williamson, C_{α} and C_{β} carbon-13 chemical shifts in proteins from an empirical database, J. Biomol. NMR, 13 (1999) 199-211.

[22] S. Schwarzinger, G.J.A. Kroon, T.R. Foss, P.E. Wright, H.J. Dyson, Random coil chemical shifts in acidic 8 M urea: Implementation of random coil shift data in NMRView, J. Biomol. NMR, 18 (2000) 43-48.

[23] E.L. Ulrich, H. Akutsu, J.F. Doreleijers, Y. Harano, Y.E. Ioannidis, J. Lin, M. Livny, S. Mading, D. Maziuk, Z. Miller, E. Nakatani, C.F. Schulte, D.E. Tolmie, R. Kent Wenger, H. Yao, J.L. Markley, BioMagResBank, Nucleic Acids Res., 36 (2008) D402-D408.

[24] N.E. Hafsa, D. Arndt, D.S. Wishart, CSI 3.0: A web server for identifying secondary and super-secondary structure in proteins using NMR chemical shifts, Nucleic Acids Res., 43 (2015) W370-W377.

[25] Y. Shen, F. Delaglio, G. Cornilescu, A. Bax, TALOS+: A hybrid method for predicting protein backbone torsion angles from NMR chemical shifts, J. Biomol. NMR, 44 (2009) 213-223.

[26] M. Kjaergaard, F.M. Poulsen, Sequence correction of random coil chemical shifts: Correlation between neighbor correction factors and changes in the Ramachandran distribution, J. Biomol. NMR, 50 (2011) 157-165.
[27] L. Morgado, B.M. Burmann, T. Sharpe, A. Mazur, S. Hiller, The dynamic dimer structure of the chaperone Trigger Factor, Nat. Commun., 8 (2017) 1992.
[28] Y. Shen, P.N. Bryan, Y. He, J. Orban, D. Baker, A. Bax, De novo structure generation using chemical shifts for proteins with high-sequence identity but different folds, Protein Sci., 19 (2010) 349-356.
[29] Y. Shen, O. Lange, F. Delaglio, P. Rossi, M. Aramini James, G. Liu, A. Eletsky, Y. Wu, K. Singarapu Kiran, A. Lemak, A. Ignatchenko, H. Arrowsmith Cheryl, T. Szyperski, T. Montelione Gaetano, D. Baker, A. Bax, Consistent blind protein structure generation from NMR chemical shift data, Proc. Natl. Acad. Sci., 105 (2008) 4685-4690.
[30] Y. Shen, R. Vernon, D. Baker, A. Bax, De novo protein structure generation from incomplete chemical shift assignments, J. Biomol. NMR, 43 (2009) 63-78.
[31] S.P. Mielke, V.V. Krishnan, Characterization of protein secondary structure from NMR chemical shifts, Prog. Nucl. Mag. Res. Sp., 54 (2009) 141-165.
[32] F.H. Schumann, H. Riepl, T. Maurer, W. Gronwald, K.-P. Neidig, H.R. Kalbitzer, Combined chemical shift changes and amino acid specific chemical shift mapping of protein-protein interactions, J. Biomol. NMR, 39 (2007) 275-289.
[33] A. Micsonai, É. Moussong, F. Wien, E. Boros, H. Vadászi, N. Murvai, Y.-H. Lee, T. Molnár, M. Réfrégiers, Y. Goto, Á. Tantos, J. Kardos, BeStSel: Webserver for secondary structure and fold prediction for protein CD spectroscopy, Nucleic Acids Res., 50 (2022) W90-W98.
[34] A. Micsonai, F. Wien, É. Bulyáki, J. Kun, É. Moussong, Y.-H. Lee, Y. Goto, M. Réfrégiers, J. Kardos, BeStSel: A web server for accurate protein secondary structure prediction and fold recognition from the circular dichroism spectra, Nucleic Acids Res., 46 (2018) W315-W322.
[35] A. Micsonai, F. Wien, L. Kernya, Y.H. Lee, Y. Goto, M. Réfrégiers, J. Kardos, Accurate secondary structure prediction and fold recognition for circular dichroism spectroscopy, Proc. Natl. Acad. Sci., 112 (2015) E3095-E3103.
[36] A. Micsonai, É. Bulyáki, J. Kardos, BeStSel: From secondary structure analysis to protein fold prediction by circular dichroism spectroscopy, Structural genomics: General applications, Springer, New York, 2021, 175-189.
[37] N.J. Greenfield, Using circular dichroism collected as a function of temperature to determine the thermodynamics of protein unfolding and binding interactions, Nat. Protoc., 1 (2006) 2527-2535.
[38] D. Marion, P.C. Driscoll, L.E. Kay, P.T. Wingfield, A. Bax, A.M. Gronenborn, G.M. Clore, Overcoming the overlap problem in the assignment of ^{1}H NMR spectra of larger proteins by use of three-dimensional heteronuclear ^{1}H-^{15}N Hartmann-Hahn-multiple quantum coherence and nuclear Overhauser-multiple quantum coherence spectroscopy: Application to Interleukin 1β, Biochemistry, 28 (1989) 6150-6156.

[39] L.E. Kay, G.Y. Xu, A.U. Singer, D.R. Muhandiram, J.D. Formankay, A gradient-enhanced HCCH-TOCSY experiment for recording side-chain ^{1}H and ^{13}C correlations in H_2O samples of proteins, J. Magn. Reson., 101 (1993) 333-337.

[40] E.R.P. Zuiderweg, S.W. Fesik, Heteronuclear three-dimensional NMR spectroscopy of the inflammatory protein C5a, Biochemistry, 28 (1989) 2387-2391.

[41] D. Marion, L.E. Kay, S.W. Sparks, D.A. Torchia, A. Bax, Three-dimensional heteronuclear NMR of ^{15}N-labeled proteins, J. Am. Chem. Soc., 111 (1989) 1515-1517.

[42] P. Guntert, C. Mumenthaler, K. Wuthrich, Torsion angle dynamics for NMR structure calculation with the new program DYANA, J. Mol. Biol., 273 (1997) 283-298.

[43] P. Güntert, L. Buchner, Combined automated NOE assignment and structure calculation with CYANA, J. Biomol. NMR, 62 (2015) 453-471.

[44] P. Klukowski, R. Riek, P. Güntert, Rapid protein assignments and structures from raw NMR spectra with the deep learning technique ARTINA, Nat. Commun., 13 (2022) 6151.

[45] P. Klukowski, R. Riek, P. Güntert, NMRtist: An online platform for automated biomolecular NMR spectra analysis, Bioinformatics, 39 (2023) btad066.

[46] P. Klukowski, M. Augoff, M. Zięba, M. Drwal, A. Gonczarek, M.J. Walczak, NMRNet: A deep learning approach to automated peak picking of protein NMR spectra, Bioinformatics, 34 (2018) 2590-2597.

[47] E. Schmidt, P. Güntert, A new algorithm for reliable and general NMR resonance assignment, J. Am. Chem. Soc., 134 (2012) 12817-12829.

[48] I. Bertini, D.A. Case, L. Ferella, A. Giachetti, A. Rosato, A Grid-enabled web portal for NMR structure refinement with AMBER, Bioinformatics, 27 (2011) 2384-2390.

[49] V.A. Jarymowycz, M.J. Stone, Fast time scale dynamics of protein backbones: NMR relaxation methods, applications, and functional consequences, Chem. Rev., 106 (2006) 1624-1671.

[50] N. Blombergen, E.M. Purcell, R.V. Pound, Relaxation effects in nuclear magnetic resonance absorption, Phys. Rev., 73 (1948) 679-712.

[51] N. Tjandra, H. Kuboniwa, H. Ren, A. Bax, Rotational dynamics of calcium-free calmodulin studied by ^{15}N-NMR relaxation measurements, Eur. J. Biochem., 230 (1995) 1014-1024.

[52] N.A. Farrow, R. Muhandiram, A.U. Singer, S.M. Pascal, C.M. Kay, G. Gish, S.E. Shoelson, T. Pawson, J.D. Forman-Kay, L.E. Kay, Backbone dynamics of a free and a phosphopeptide-complexed Src homology 2 domain studied by ^{15}N NMR relaxation, Biochemistry, 33 (1994) 5984-6003.

[53] N. Tokuriki, D.S. Tawfik, Protein dynamism and evolvability, Science, 324 (2009) 203-207.

[54] V.V. Gurevich, E.V. Gurevich, Dynamic nature of proteins is critically important for their function: GPCRs and signal transducers, Appl. Magn. Reson., 55 (2024) 11-25.

[55] H.J.C. Berendsen, S. Hayward, Collective protein dynamics in relation to function, Curr. Opin. Struc. Biol., 10 (2000) 165-169.

[56] O.A. Martin, M.E. Villegas, J.A. Vila, H.A. Scheraga, Analysis of $^{13}C_{\alpha}$ and $^{13}C_{\beta}$ chemical shifts of cysteine and cystine residues in proteins: A quantum chemical approach, J. Biomol. NMR, 46 (2010) 217-225.

[57] D. Sharma, K. Rajarathnam, ^{13}C NMR chemical shifts can predict disulfide bond formation, J. Biomol. NMR, 18 (2000) 165-171.
[58] C. Wiedemann, A. Kumar, A. Lang, O. Ohlenschläger, Cysteines and disulfide bonds as structure-forming units: Insights from different domains of Life and the potential for characterization by NMR, Front. Chem., 8 (2020).
[59] L.E. Kay, D.A. Torchia, A. Bax, Backbone dynamics of proteins as studied by nitrogen-15 inverse detected heteronuclear NMR spectroscopy: Application to staphylococcal nuclease, Biochemistry, 28 (1989) 8972-8979.
[60] D. Fushman, R. Weisemann, H. Thüring, H. Rüterjans, Backbone dynamics of ribonuclease T1 and its complex with 2'GMP studied by two-dimensional heteronuclear NMR spectroscopy, J. Biomol. NMR, 4 (1994) 61-78.
[61] P. Rossi, G.V.T. Swapna, Y.J. Huang, J.M. Aramini, C. Anklin, K. Conover, K. Hamilton, R. Xiao, T.B. Acton, A. Ertekin, J.K. Everett, G.T. Montelione, A microscale protein NMR sample screening pipeline, J. Biomol. NMR, 46 (2010) 11-22.
[62] J.M. Aramini, L.-C. Ma, L. Zhou, C.M. Schauder, K. Hamilton, B.R. Amer, T.R. Mack, H.-W. Lee, C.T. Ciccosanti, L. Zhao, R. Xiao, R.M. Krug, G.T. Montelione, Dimer interface of the effector domain of non-structural protein 1 from Influenza A virus: An interface with multiple functions, J. Biol. Chem., 286 (2011) 26050-26060.
[63] G.M. Clore, P.C. Driscoll, P.T. Wingfield, A.M. Gronenborn, Analysis of the backbone dynamics of interleukin-1β using two-dimensional inverse detected heteronuclear ^{15}N-^{1}H NMR spectroscopy, Biochemistry, 29 (1990) 7387-7401.
[64] R. Evans, M. O'Neill, A. Pritzel, N. Antropova, A. Senior, T. Green, A. Žídek, R. Bates, S. Blackwell, J. Yim, O. Ronneberger, S. Bodenstein, M. Zielinski, A. Bridgland, A. Potapenko, A. Cowie, K. Tunyasuvunakool, R. Jain, E. Clancy, P. Kohli, J. Jumper, D. Hassabis, Protein complex prediction with AlphaFold-Multimer, bioRxiv, (2022) 2021.2010.2004.463034.
[65] C. Elfmann, J. Stülke, PAE viewer: A webserver for the interactive visualization of the predicted aligned error for multimer structure predictions and crosslinks, Nucleic Acids Res., 51 (2023) W404-W410.
[66] M. Mirdita, K. Schütze, Y. Moriwaki, L. Heo, S. Ovchinnikov, M. Steinegger, ColabFold: Making protein folding accessible to all, Nat. Methods, 19 (2022) 679-682.
[67] E.F. Pettersen, T.D. Goddard, C.C. Huang, E.C. Meng, G.S. Couch, T.I. Croll, J.H. Morris, T.E. Ferrin, UCSF ChimeraX: Structure visualization for researchers, educators, and developers, Protein Sci., 30 (2021) 70-82.
[68] M. Karplus, J.A. McCammon, Molecular dynamics simulations of biomolecules, Nat. Struct. Biol., 9 (2002) 646-652.
[69] S.A. Hollingsworth, R.O. Dror, Molecular dynamics simulation for all, Neuron, 99 (2018) 1129-1143.
[70] S.A.I. Seidel, P.M. Dijkman, W.A. Lea, G. van den Bogaart, M. Jerabek-Willemsen, A. Lazic, J.S. Joseph, P. Srinivasan, P. Baaske, A. Simeonov, I. Katritch, F.A. Melo, J.E. Ladbury, G. Schreiber, A.

Watts, D. Braun, S. Duhr, Microscale thermophoresis quantifies biomolecular interactions under previously challenging conditions, Methods, 59 (2013) 301-315.
[71] M. Jerabek-Willemsen, T. André, R. Wanner, H.M. Roth, S. Duhr, P. Baaske, D. Breitsprecher, Microscale Thermophoresis: Interaction analysis and beyond, J. Mol. Struct., 1077 (2014) 101-113.
[72] S. Duhr, D. Braun, Why molecules move along a temperature gradient, Proc. Natl. Acad. Sci., 103 (2006) 19678-19682.
[73] M. Jerabek-Willemsen, C.J. Wienken, D. Braun, P. Baaske, S. Duhr, Molecular interaction studies using microscale thermophoresis, Assay Drug Dev. Techn., 9 (2011) 342-353.
[74] C.J. Wienken, P. Baaske, U. Rothbauer, D. Braun, S. Duhr, Protein-binding assays in biological liquids using microscale thermophoresis, Nat. Commun., 1 (2010) 100.
[75] R. Nasreddine, R. Nehmé, Microscale thermophoresis for studying protein-small molecule affinity: Application to hyaluronidase, Microchem. J., 170 (2021) 106763.
[76] A.G. Kikhney, D.I. Svergun, A practical guide to small angle X-ray scattering (SAXS) of flexible and intrinsically disordered proteins, FEBS Lett., 589 (2015) 2570-2577.
[77] I.S. Dmitri, H.J.K. Michel, Small-angle scattering studies of biological macromolecules in solution, Rep. Prog. Phys., 66 (2003) 1735-1782.
[78] M. Gräwert, D. Svergun, A beginner's guide to solution small-angle X-ray scattering (SAXS), The Biochemist, 42 (2020) 36-42.
[79] H.D.T. Mertens, D.I. Svergun, Structural characterization of proteins and complexes using small-angle X-ray solution scattering, J. Struct. Biol., 172 (2010) 128-141.
[80] C. Prior, O.R. Davies, D. Bruce, E. Pohl, Obtaining tertiary protein structures by the ab initio interpretation of small angle X-ray scattering data, J. Chem. Theory Comput., 16 (2020) 1985-2001.
[81] C.D. Putnam, M. Hammel, G.L. Hura, J.A. Tainer, X-ray solution scattering (SAXS) combined with crystallography and computation: Defining accurate macromolecular structures, conformations and assemblies in solution, Q. Rev. Biophys., 40 (2007) 191-285.
[82] D. Franke, M.V. Petoukhov, P.V. Konarev, A. Panjkovich, A. Tuukkanen, H.D.T. Mertens, A.G. Kikhney, N.R. Hajizadeh, J.M. Franklin, C.M. Jeffries, D.I. Svergun, ATSAS 2.8: A comprehensive data analysis suite for small-angle scattering from macromolecular solutions, J. Appl. Crystallogr., 50 (2017) 1212-1225.
[83] A. Panjkovich, D.I. Svergun, Deciphering conformational transitions of proteins by small angle X-ray scattering and normal mode analysis, Phys. Chem. Chem. Phys., 18 (2016) 5707-5719.
[84] D. Franke, D.I. Svergun, DAMMIF, a program for rapid *ab-initio* shape determination in small-angle scattering, J. Appl. Crystallogr., 42 (2009) 342-346.
[85] V.V. Volkov, D.I. Svergun, Uniqueness of *ab initio* shape determination in small-angle scattering, J. Appl. Crystallogr., 36 (2003) 860-864.
[86] H. Liu, P.H. Zwart, Determining pair distance distribution function from SAXS data using parametric functionals, J. Struct. Biol., 180 (2012) 226-234.

[87] K.H. Choi, M. Morais, Use of small-angle X-ray scattering to investigate the structure and function of Dengue virus NS3 and NS5, Dengue: Methods and protocols, Springer, New York, 2014, 241-252.
[88] S. Da Vela, D.I. Svergun, Methods, development and applications of small-angle X-ray scattering to characterize biological macromolecules in solution, Curr. Res. Struct. Biol., 2 (2020) 164-170.
[89] M.Y. Lobanov, N.S. Bogatyreva, O.V. Galzitskaya, Radius of gyration as an indicator of protein structure compactness, Mol. Biol., 42 (2008) 623-628.
[90] C. Putnam, Guinier peak analysis for visual and automated inspection of small-angle X-ray scattering data, J. Appl. Crystallogr., 49 (2016) 1412-1419.
[91] T.D. Grant, J.R. Luft, L.G. Carter, T. Matsui, T.M. Weiss, A. Martel, E.H. Snell, The accurate assessment of small-angle X-ray scattering data, Acta Crystallogr. D, 71 (2015) 45-56.
[92] R.P. Rambo, J.A. Tainer, Accurate assessment of mass, models and resolution by small-angle scattering, Nature, 496 (2013) 477-481.
[93] M. Kanehisa, Y. Sato, M. Kawashima, M. Furumichi, M. Tanabe, KEGG as a reference resource for gene and protein annotation, Nucleic Acids Res., 44 (2016) D457-D462.
[94] D.-I. Liao, Z. Wawrzak, J.C. Calabrese, P.V. Viitanen, D.B. Jordan, Crystal structure of riboflavin synthase, Structure, 9 (2001) 399-408.
[95] M. Fischer, A. Bacher, Biosynthesis of vitamin B2: Structure and mechanism of riboflavin synthase, Arch. Biochem. Biophys., 474 (2008) 252-265.
[96] S. Gerhardt, A.-K. Schott, N. Kairies, M. Cushman, B. Illarionov, W. Eisenreich, A. Bacher, R. Huber, S. Steinbacher, M. Fischer, Studies on the reaction mechanism of riboflavin synthase: X-ray crystal structure of a complex with 6-carboxyethyl-7-oxo-8-ribityllumazine, Structure, 10 (2002) 1371-1381.
[97] M.I. Serer, H.R. Bonomi, B.G. Guimaraes, R.C. Rossi, F.A. Goldbaum, S. Klinke, Crystallographic and kinetic study of riboflavin synthase from *Brucella abortus*, a chemotherapeutic target with an enhanced intrinsic flexibility, Acta Crystallogr. D, 70 (2014) 1419-1434.
[98] W. Meining, S. Eberhardt, A. Bacher, R. Ladenstein, The structure of the N-terminal domain of riboflavin synthase in complex with riboflavin at 2.6 Å resolution, J. Mol. Biol., 331 (2003) 1053-1063.
[99] V. Truffault, M. Coles, T. Diercks, K. Abelmann, S. Eberhardt, H. Lüttgen, A. Bacher, H. Kessler, The solution structure of the N-terminal domain of riboflavin synthase, J. Mol. Biol., 309 (2001) 949-960.
[100] Y. Sato, S. Shimizu, A. Ohtaki, K. Noguchi, H. Miyatake, N. Dohmae, S. Sasaki, M. Odaka, M. Yohda, Crystal structures of the lumazine protein from *Photobacterium kishitanii* in complexes with the authentic chromophore, 6,7-dimethyl- 8-(1′-D-dibityl) lumazine, and its analogues, riboflavin and flavin mononucleotide, at high resolution, J. Bacteriol., 192 (2010) 127-133.
[101] K. Stierand, P.C. Maaß, M. Rarey, Molecular complexes at a glance: Automated generation of two-dimensional complex diagrams, Bioinformatics, 22 (2006) 1710-1716.
[102] P.C. Fricker, M. Gastreich, M. Rarey, Automated drawing of structural molecular formulas under constraints, J. Chem. Inf. Comp. Sci., 44 (2004) 1065-1078.

[103] K. Stierand, M. Rarey, From modeling to medicinal chemistry: Automatic generation of two-dimensional complex diagrams, ChemMedChem, 2 (2007) 853-860.
[104] K. Stierand, M. Rarey, Drawing the PDB: Protein-ligand complexes in two dimensions, ACS Med. Chem. Lett., 1 (2010) 540-545.
[105] M. Rarey, B. Kramer, T. Lengauer, G. Klebe, A fast flexible docking method using an incremental construction algorithm, J. Mol. Biol., 261 (1996) 470-489.
[106] B. Illarionov, M. Fischer, C.Y. Lee, A. Bacher, W. Eisenreich, Rapid preparation of isotopolog libraries by in vivo transformation of ^{13}C-glucose. Studies on 6,7-dimethyl-8-ribityllumazine, a biosynthetic precursor of vitamin B2, J. Org. Chem., 69 (2004) 5588-5594.
[107] A. Bacher, Heavy riboflavin synthase from *Bacillus subtilis*, Methods in enzymology, Academic Press, 1986, 192-199.
[108] A. Mitra, P.J. Seaton, R. Ali Assarpour, T. Williamson, Unprecedented concentration dependent chemical shift variation in ^{1}H-NMR studies: A caveat in the investigations of molecular recognition and structure elucidation, Tetrahedron, 54 (1998) 15489-15498.
[109] F. Müller, NMR spectroscopy on flavins and flavoproteins, Flavins and flavoproteins: Methods and protocols, Springer, New York, 2014, 229-306.
[110] R.H. Sarma, P. Dannies, N.O. Kaplan, Investigations of inter- and intramolecular interactions in flavine-adenine dinucleotide by proton magnetic resonance, Biochemistry, 7 (1968) 4359-4367.
[111] G. Kotowycz, N. Teng, M.P. Klein, M. Calvin, The 220 MHz nuclear magnetic resonance study of a solvent-induced conformational change in flavin adenine dinucleotide, J. Biol. Chem., 244 (1969) 5656-5662.
[112] M. Kainosho, Y. Kyogoku, High-resolution proton and phosphorus nuclear magnetic resonance spectra of flavine adenine dinucleotide and its conformation in aqueous solution, Biochemistry, 11 (1972) 741-752.
[113] M. Raszka, N.O. Kaplan, Intramolecular hydrogen bonding in flavin adenine dinucleotide, Proc. Natl. Acad. Sci., 71 (1974) 4546-4550.
[114] Y.H. Ding, K.K. Hixson, M.A. Aklujkar, M.S. Lipton, R.D. Smith, D.R. Lovley, T. Mester, Proteome of *Geobacter sulfurreducens* grown with Fe(III) oxide or Fe(III) citrate as the electron acceptor, Biochim. Biophys. Acta - Proteins Proteom., 1784 (2008) 1935-1941.
[115] C.A. Salgueiro, L. Morgado, M.A. Silva, M.R. Ferreira, T.M. Fernandes, P.C. Portela, From iron to bacterial electroconductive filaments: Exploring cytochrome diversity using *Geobacter* bacteria, Coordin. Chem. Rev., 452 (2022) 214284.
[116] T.M. Fernandes, L. Morgado, D.L. Turner, C.A. Salgueiro, Protein engineering of electron transfer components from electroactive *Geobacter* bacteria, Antioxidants, 10 (2021) 844.
[117] P.R. Pokkuluri, Y.Y. Londer, X. Yang, N.E. Duke, J. Erickson, V. Orshonsky, G. Johnson, M. Schiffer, Structural characterization of a family of cytochromes c_7 involved in Fe(III) respiration by *Geobacter sulfurreducens*, Biochim. Biophys. Acta, 1797 (2010) 222-232.

[118] P.R. Pokkuluri, Y.Y. Londer, N.E.C. Duke, W.C. Long, M. Schiffer, Family of cytochrome c_7-type proteins from *Geobacter sulfurreducens*: Structure of one cytochrome c_7 at 1.45 Å resolution, Biochemistry, 43 (2004) 849-859.

[119] L. Morgado, M. Bruix, V. Orshonsky, Y.Y. Londer, N.E. Duke, X. Yang, P.R. Pokkuluri, M. Schiffer, C.A. Salgueiro, Structural insights into the modulation of the redox properties of two *Geobacter sulfurreducens* homologous triheme cytochromes, Biochim. Biophys. Acta - Bioenergetics, 1777 (2008) 1157-1165.

[120] L. Morgado, M. Bruix, Y.Y. Londer, P.R. Pokkuluri, M. Schiffer, C.A. Salgueiro, Redox-linked conformational changes of a multiheme cytochrome from *Geobacter sulfurreducens*, Biochem. Bioph. Res. Co., 360 (2007) 194-198.

[121] L. Morgado, V.B. Paixão, M. Schiffer, P.R. Pokkuluri, M. Bruix, C.A. Salgueiro, Revealing the structural origin of the redox-Bohr effect: The first solution structure of a cytochrome from *Geobacter sulfurreducens*, Biochem. J., 441 (2012) 179-187.

[122] M.-a. Sun, Y. Wang, Q. Zhang, Y. Xia, W. Ge, D. Guo, Prediction of reversible disulfide based on features from local structural signatures, BMC Genomics, 18 (2017) 279.

[123] K.S. Jensen, R.E. Hansen, J.R. Winther, Kinetic and thermodynamic aspects of cellular thiol-disulfide redox regulation, Antioxid. Redox Signal., 11 (2008) 1047-1058.

[124] M.A. Wouters, S.W. Fan, N.L. Haworth, Disulfides as redox switches: From molecular mechanisms to functional significance, Antioxid. Redox Signal., 12 (2009) 53-91.

[125] T.J. Bechtel, E. Weerapana, From structure to redox: The diverse functional roles of disulfides and implications in disease, Proteomics, 17 (2017) 1600391.

[126] C.S. Sevier, C.A. Kaiser, Formation and transfer of disulphide bonds in living cells, Nat. Rev. Mol. Cell Biol., 3 (2002) 836-847.

[127] F. Hatahet, D. Boyd, J. Beckwith, Disulfide bond formation in prokaryotes: History, diversity and design, Biochim. Biophys. Acta - Proteins Proteom., 1844 (2014) 1402-1414.

[128] G. Wittenberg, A. Danon, Disulfide bond formation in chloroplasts: Formation of disulfide bonds in signaling chloroplast proteins, Plant Sci., 175 (2008) 459-466.

[129] J. Mistry, S. Chuguransky, L. Williams, M. Qureshi, Gustavo A. Salazar, E.L.L. Sonnhammer, S.C.E. Tosatto, L. Paladin, S. Raj, L.J. Richardson, R.D. Finn, A. Bateman, Pfam: The protein families database in 2021, Nucleic Acids Res., 49 (2021) D412-D419.

[130] A.N. Calabrese, B. Schiffrin, M. Watson, T.K. Karamanos, M. Walko, J.R. Humes, J.E. Horne, P. White, A.J. Wilson, A.C. Kalli, R. Tuma, A.E. Ashcroft, D.J. Brockwell, S.E. Radford, Inter-domain dynamics in the chaperone SurA and multi-site binding to its outer membrane protein clients, Nat. Commun., 11 (2020) 2155.

[131] E. Bitto, D.B. McKay, Crystallographic structure of SurA, a molecular chaperone that facilitates folding of outer membrane porins, Structure, 10 (2002) 1489-1498.

[132] P.O. Giuseppe, M.V. Atzingen, A.L.T.O. Nascimento, N.I.T. Zanchin, B.G. Guimarães, The crystal structure of the leptospiral hypothetical protein LIC12922 reveals homology with the periplasmic chaperone SurA, J. Struct. Biol., 173 (2011) 312-322.

[133] J. Abramson, J. Adler, J. Dunger, R. Evans, T. Green, A. Pritzel, O. Ronneberger, L. Willmore, A.J. Ballard, J. Bambrick, S.W. Bodenstein, D.A. Evans, C.-C. Hung, M. O'Neill, D. Reiman, K. Tunyasuvunakool, Z. Wu, A. Žemgulytė, E. Arvaniti, C. Beattie, O. Bertolli, A. Bridgland, A. Cherepanov, M. Congreve, A.I. Cowen-Rivers, A. Cowie, M. Figurnov, F.B. Fuchs, H. Gladman, R. Jain, Y.A. Khan, C.M.R. Low, K. Perlin, A. Potapenko, P. Savy, S. Singh, A. Stecula, A. Thillaisundaram, C. Tong, S. Yakneen, E.D. Zhong, M. Zielinski, A. Žídek, V. Bapst, P. Kohli, M. Jaderberg, D. Hassabis, J.M. Jumper, Accurate structure prediction of biomolecular interactions with AlphaFold 3, Nature, 630 (2024) 493-500.

[134] N. Noinaj, A.J. Kuszak, J.C. Gumbart, P. Lukacik, H. Chang, N.C. Easley, T. Lithgow, S.K. Buchanan, Structural insight into the biogenesis of β-barrel membrane proteins, Nature, 501 (2013) 385-390.

[135] F. Gruss, F. Zähringer, R.P. Jakob, B.M. Burmann, S. Hiller, T. Maier, The structural basis of autotransporter translocation by TamA, Nat. Struct. Mol. Biol., 20 (2013) 1318-1320.

[136] L. Morgado, K. Zeth, B.M. Burmann, T. Maier, S. Hiller, Characterization of the insertase BamA in three different membrane mimetics by solution NMR spectroscopy, J. Biomol. NMR, 61 (2015) 333-345.

[137] F. Teufel, J.J. Almagro Armenteros, A.R. Johansen, M.H. Gíslason, S.I. Pihl, K.D. Tsirigos, O. Winther, S. Brunak, G. von Heijne, H. Nielsen, SignalP 6.0 predicts all five types of signal peptides using protein language models, Nat. Biotechnol., 40 (2022) 1023-1025.

[138] S.R. Bond, C.C. Naus, RF-Cloning.org: An online tool for the design of restriction-free cloning projects, Nucleic Acids Res., 40 (2012) 209-213.

[139] J. Jumper, R. Evans, A. Pritzel, T. Green, M. Figurnov, O. Ronneberger, K. Tunyasuvunakool, R. Bates, A. Žídek, A. Potapenko, A. Bridgland, C. Meyer, S.A.A. Kohl, A.J. Ballard, A. Cowie, B. Romera-Paredes, S. Nikolov, R. Jain, J. Adler, T. Back, S. Petersen, D. Reiman, E. Clancy, M. Zielinski, M. Steinegger, M. Pacholska, T. Berghammer, S. Bodenstein, D. Silver, O. Vinyals, A.W. Senior, K. Kavukcuoglu, P. Kohli, D. Hassabis, Highly accurate protein structure prediction with AlphaFold, Nature, 596 (2021) 583-589.

[140] M. Varadi, S. Anyango, M. Deshpande, S. Nair, C. Natassia, G. Yordanova, D. Yuan, O. Stroe, G. Wood, A. Laydon, A. Žídek, T. Green, K. Tunyasuvunakool, S. Petersen, J. Jumper, E. Clancy, R. Green, A. Vora, M. Lutfi, M. Figurnov, A. Cowie, N. Hobbs, P. Kohli, G. Kleywegt, E. Birney, D. Hassabis, S. Velankar, AlphaFold protein structure database: Massively expanding the structural coverage of protein-sequence space with high-accuracy models, Nucleic Acids Res., 50 (2022) D439-D444.

[141] V. Mariani, M. Biasini, A. Barbato, T. Schwede, lDDT: A local superposition-free score for comparing protein structures and models using distance difference tests, Bioinformatics, 29 (2013) 2722-2728.

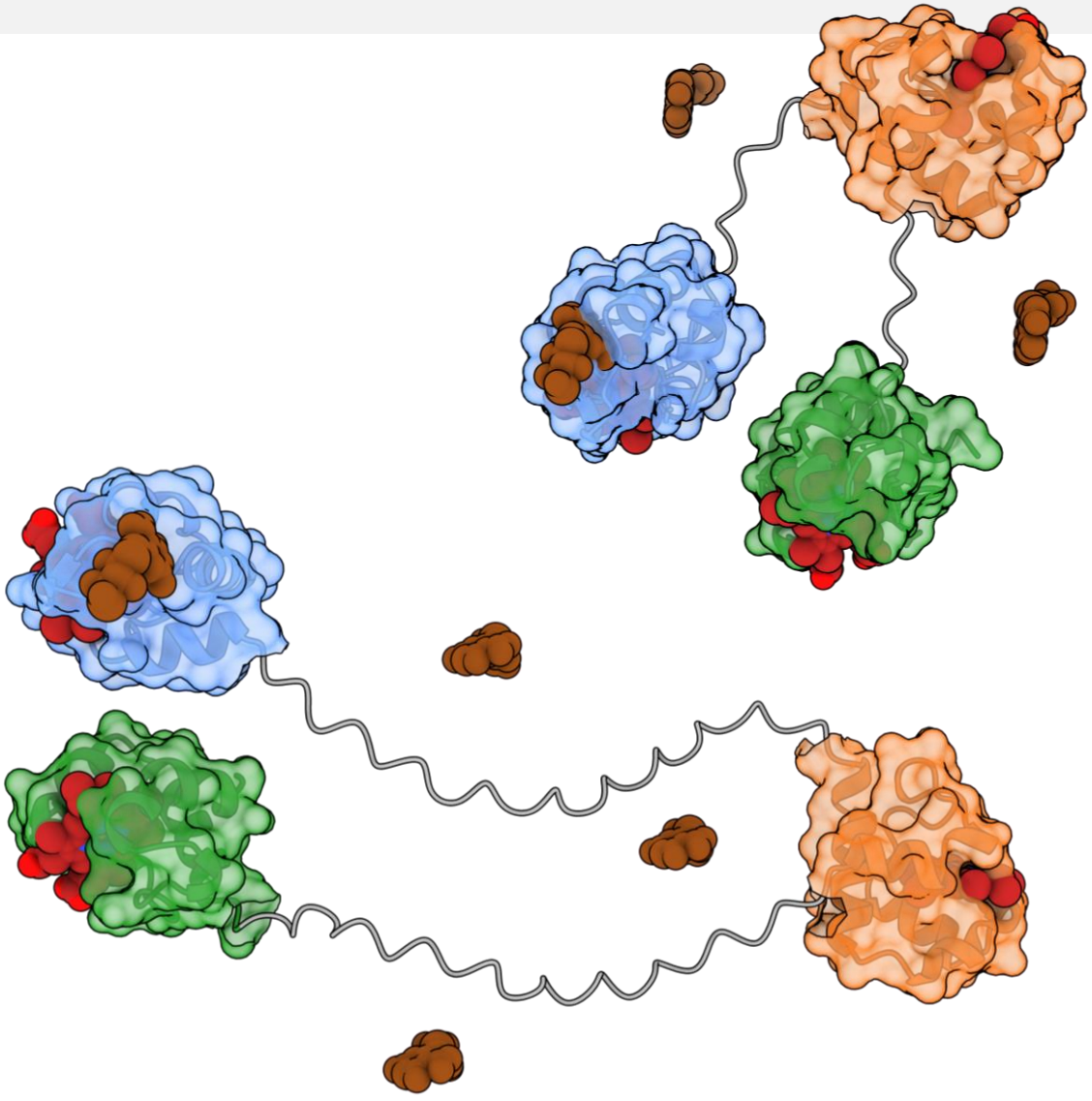
[142] Y.Y. Londer, P.R. Pokkuluri, D.M. Tiede, M. Schiffer, Production and preliminary characterization of a recombinant triheme cytochrome c_7 from *Geobacter sulfurreducens* in *Escherichia coli*, Biochim. Biophys. Acta - Bioenergetics, 1554 (2002) 202-211.

[143] R. Pierattelli, L. Banci, D.L. Turner, Indirect determination of magnetic susceptibility tensors in peroxidases: A novel approach to structure elucidation by NMR, J. Biol. Inorg. Chem., 1 (1996) 320-329.
[144] D.S. Wishart, C.G. Bigam, J. Yao, F. Abildgaard, H.J. Dyson, E. Oldfield, J.L. Markley, B.D. Sykes, ^{1}H, ^{13}C and ^{15}N chemical shift referencing in biomolecular NMR, J. Biomol. NMR, 6 (1995) 135-140.
[145] R.L.J. Keller, The computer aided resonance assignment tutorial, Cantina Verlag, (2004).
[146] W. Lee, M. Tonelli, J.L. Markley, NMRFAM-SPARKY: Enhanced software for biomolecular NMR spectroscopy, Bioinformatics, 31 (2015) 1325-1327.
[147] W.L. Jorgensen, J. Chandrasekhar, J.D. Madura, R.W. Impey, M.L. Klein, Comparison of simple potential functions for simulating liquid water, J. Chem. Phys., 79 (1983) 926-935.
[148] V. Hornak, A. Okur, R.C. Rizzo, C. Simmerling, HIV-1 protease flaps spontaneously open and reclose in molecular dynamics simulations, Proc. Natl. Acad. Sci., 103 (2006) 915-920.
[149] V. Hornak, R. Abel, A. Okur, B. Strockbine, A. Roitberg, C. Simmerling, Comparison of multiple Amber force fields and development of improved protein backbone parameters, Proteins, 65 (2006) 712-725.
[150] R.A. Laskowski, M.W. MacArthur, D.S. Moss, J.M. Thornton, PROCHECK: A program to check the stereochemical quality of protein structures, J. Appl. Crystallogr., 26 (1993) 283-291.
[151] R.A. Laskowski, J.A.C. Rullmann, M.W. MacArthur, R. Kaptein, J.M. Thornton, AQUA and PROCHECK-NMR: Programs for checking the quality of protein structures solved by NMR, J. Biomol. NMR, 8 (1996) 477-486.
[152] R.A. Laskowski, E.G. Hutchinson, A.D. Michie, A.C. Wallace, M.L. Jones, J.M. Thornton, PDBsum: A web-based database of summaries and analyses of all PDB structures, Trends Biochem. Sci., 22 (1997) 488-490.
[153] S. Seeliger, R. Cord-Ruwisch, B. Schink, A periplasmic and extracellular *c*-type cytochrome of *Geobacter sulfurreducens* acts as a ferric iron reductase and as an electron carrier to other acceptors or to partner bacteria, J. Bacteriol., 180 (1998) 3686-3691.
[154] M. Broco, C.M. Soares, S. Oliveira, S.G. Mayhew, C. Rodrigues-Pousada, Molecular determinants for FMN-binding in *Desulfovibrio gigas* flavoredoxin, FEBS Lett., 581 (2007) 4397-4402.
[155] A. Lostao, C. Gómez-Moreno, S.G. Mayhew, J. Sancho, Differential stabilization of the three FMN redox forms by tyrosine 94 and tryptophan 57 in flavodoxin from *Anabaena* and its influence on the redox potentials, Biochemistry, 36 (1997) 14334-14344.
[156] N.J. Greenfield, Analysis of circular dichroism data, Methods in enzymology, Academic Press, 2004, 282-317.
[157] D. Van Der Spoel, E. Lindahl, B. Hess, G. Groenhof, A.E. Mark, H.J.C. Berendsen, GROMACS: Fast, flexible, and free, J. Comput. Chem., 26 (2005) 1701-1718.
[158] M.J. Abraham, T. Murtola, R. Schulz, S. Páll, J.C. Smith, B. Hess, E. Lindahl, GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers, SoftwareX, 1-2 (2015) 19-25.

[159] K. Lindorff-Larsen, S. Piana, K. Palmo, P. Maragakis, J.L. Klepeis, R.O. Dror, D.E. Shaw, Improved side-chain torsion potentials for the Amber ff99SB protein force field, Proteins, 78 (2010) 1950-1958.
[160] T. Darden, D. York, L. Pedersen, Particle mesh Ewald: An N·log(N) method for Ewald sums in large systems, J. Chem. Phys., 98 (1993) 10089-10092.
[161] U. Essmann, L. Perera, M.L. Berkowitz, T. Darden, H. Lee, L.G. Pedersen, A smooth particle mesh Ewald method, J. Chem. Phys., 103 (1995) 8577-8593.
[162] P.P. Ewald, Die Berechnung optischer und elektrostatischer Gitterpotentiale, Ann. Phys., 369 (1921) 253-287.
[163] K.B. Wiberg, A scheme for strain energy minimization. Application to the cycloalkanes, J. Am. Chem. Soc., 87 (1965) 1070-1078.
[164] R. Fletcher, C.M. Reeves, Function minimization by conjugate gradients, Comput. J., 7 (1964) 149-154.
[165] S. Nosé, M.L. Klein, Constant pressure molecular dynamics for molecular systems, Mol. Phys., 50 (1983) 1055-1076.
[166] M. Parrinello, A. Rahman, Polymorphic transitions in single crystals: A new molecular dynamics method, J. Appl. Phys., 52 (1981) 7182-7190.
[167] B. Hess, H. Bekker, H.J.C. Berendsen, J.G.E.M. Fraaije, LINCS: A linear constraint solver for molecular simulations, J. Comput. Chem., 18 (1997) 1463-1472.
[168] W. Humphrey, A. Dalke, K. Schulten, VMD: Visual molecular dynamics, J. Mol. Graph., 14 (1996) 33-38.
[169] M.D. Tully, J. Kieffer, M.E. Brennich, R. Cohen Aberdam, J.B. Florial, S. Hutin, M. Oscarsson, A. Beteva, A. Popov, D. Moussaoui, P. Theveneau, G. Papp, J. Gigmes, F. Cipriani, A. McCarthy, C. Zubieta, C. Mueller-Dieckmann, G. Leonard, P. Pernot, BioSAXS at European Synchrotron Radiation Facility – Extremely Brilliant Source: BM29 with an upgraded source, detector, robot, sample environment, data collection and analysis software, J. Synchrotron Radiat., 30 (2023) 258-266.
[170] K. Manalastas-Cantos, P.V. Konarev, N.R. Hajizadeh, A.G. Kikhney, M.V. Petoukhov, D.S. Molodenskiy, A. Panjkovich, H.D.T. Mertens, A. Gruzinov, C. Borges, C.M. Jeffries, D.I. Svergun, D. Franke, ATSAS 3.0: Expanded functionality and new tools for small-angle scattering data analysis, J. Appl. Crystallogr., 54 (2021) 343-355.

4

Disentangling a tangled cytochrome: Functional mechanisms of PgcA



This chapter was partially reproduced from (i) **T. M. Fernandes**, M. A. Silva, L. Morgado and C. A. Salgueiro, Hemes on a string: Insights on the functional mechanisms of PgcA from *Geobacter sulfurreducens*, J. Biol. Chem., 299 (2023) 105167, and (ii) B. W. Nash, **T. M. Fernandes**, J. A. J. Burton, L. Morgado, J. H. van Wonderen, D. A. Svistunenko, M. J. Edwards, C. A. Salgueiro, J. N. Butt and T. A. Clarke, Tethered heme domains in a triheme cytochrome allow for increased electron transport distances, Protein Sci., 33 (2024) e5200.

List of contents

Disentangling a tangled cytochrome: Functional mechanisms of PgcA 215

4.1. Summary 217

4.2. Introduction 219

4.3. Results and discussion 220

4.3.1. PgcA and homologs establish a new class of cytochromes 220

4.3.2. PgcA's cytochrome domains are composed of α-helical structures 223

4.3.3. PgcA contains an intramolecular redox chain 227

4.3.4. NMR features of PgcA's cytochrome domains 228

4.3.5. PgcA's cytochrome domains exchange electrons promiscuously 235

4.3.6. The cytochrome domains of PgcA reduce several electron acceptors 239

4.3.7. Proline-threonine stretches are not responsible for Fe(III) oxides binding in PgcA 240

4.3.8. Domain 3 is intramolecularly recharged for continuous Fe(III) oxides reduction 242

4.4. Conclusions 244

4.5. Materials and methods 247

4.5.1. DNA manipulation 247

4.5.2. Protein sequence and structural homology analysis 247

4.5.3. Protein expression and purification 248

4.5.4. CD spectroscopy 249

4.5.5. DSC experiments 249

4.5.6. Redox titrations followed by visible spectroscopy 250

4.5.7. Electron transfer experiments with terminal electron acceptors 250

4.5.8. Fe(III) oxides binding experiments 251

4.5.9. NMR spectroscopy 251

4.6. References 254

4.1. Summary

Microbial extracellular reduction of insoluble compounds requires soluble electron shuttles that diffuse in the environment, freely diffusing cytochromes, or direct contact with cellular conductive appendages that release or harvest electrons to ensure a continuous balance between cellular requirements and environmental conditions.

PgcA is a freely diffusing triheme cytochrome that contributes to Fe(III) and Mn(IV) oxides reduction in *G. sulfurreducens*. The AlphaFold model of this cytochrome shows that its mature form possesses a fuzzy global arrangement with three monoheme cytochrome domains linked by unstructured stretches. The three monoheme domains are structurally homologous, but their heme groups show variable axial coordination and reduction potential values. Electron transfer experiments monitored by NMR and visible spectroscopy show the variable extent to which the domains promiscuously exchange electrons, while reducing different electron acceptors.

The results suggest that PgcA is part of a new class of cytochromes – microbial heme-tethered redox strings – that use flexible low-complexity protein stretches to adopt multiple conformations that promote intra- and intermolecular electron transfer events amongst its cytochrome domains and to electron acceptors at variable distances.

4.2. Introduction

Exoelectrogens explore different mechanisms by which electrons are transferred to extracellular electron acceptors, including self-produced redox mediators [1, 2], conductive pili [3] or outer membrane cytochromes [4]. The latter can either be attached to the outer membrane by lipid anchors, be part of porin-cytochrome complexes for short-range electron transfer [5, 6], or form oligomers designated nanowires, whose name derives from their ability to transfer electrons over nanometer distances [7-10]. The outer membrane of *G. sulfurreducens* holds at least five different porin-cytochrome complexes [11] and three types of cytochrome nanowires [7-10]. Additionally, the bacterium produces proteins that are either loosely bound or diffuse freely between cells, thus participating in the final steps of EET.

PgcA is a 50-kDa cytochrome with approximately 500 amino acids, which is more abundant when *G. sulfurreducens* uses insoluble Fe(III) oxides as terminal electron acceptors, compared to soluble Fe(III) citrate [12, 13]. Previous studies showed that heterologously expressed PgcA has two distinct forms and participates in both Fe(III) and Mn(IV) oxide reduction [14]. The protein is initially expressed in an unprocessed periplasmic form and is matured during cell secretion by cleavage of a 90-amino acid portion of its N-terminal region [14]. These findings were consistent with previous in vivo observations on *G. sulfurreducens* cell cultures [15, 16].

PgcA contains three *c*-type heme-binding motifs (CXXCH) separated by two repeated motifs of proline and threonine residues, which are suggested to be involved in the interaction with metal oxide surfaces [14, 17]. Purified PgcA is mainly composed of α-helical motifs and is able to bind Fe(III) oxides, but not magnetite, a mixed Fe(II)-Fe(III) mineral [14]. Recent studies have also shown that PgcA is not relevant for electron transfer to electrodes, and that mutations on its N-terminal domain affect the proper maturation of the protein and consequent secretion. These observations suggest that PgcA might contribute to electron transfer in the periplasm and extracellularly, depending on the cell requirements and inherent maturation of the protein [18].

In this chapter, we investigated the structural and biochemical features of PgcA. The three monoheme cytochrome domains were heterologously expressed and studied using complementary biophysical techniques, including UV-visible, CD and NMR spectroscopies. The reduction potential values of each domain were determined, and electron transfer reactions monitored by NMR showed that the domains can promiscuously interchange electrons. The results of these experiments were reproducible in the full-length cytochrome. Overall, the results obtained, together with AlphaFold model predictions, suggest that PgcA belongs to a new class of microbial heme-tethered redox strings, formed by cytochrome domains interposed between flexible linkers that confer elasticity and promote electron transfer events amongst the individual domains and to electron acceptors at variable distances.

4.3. Results and discussion

4.3.1. PgcA and homologs establish a new class of cytochromes

The AlphaFold model of PgcA shows that the triheme cytochrome possesses a fuzzy global arrangement with four structured domains linked by unstructured stretches (Figure 4.1).

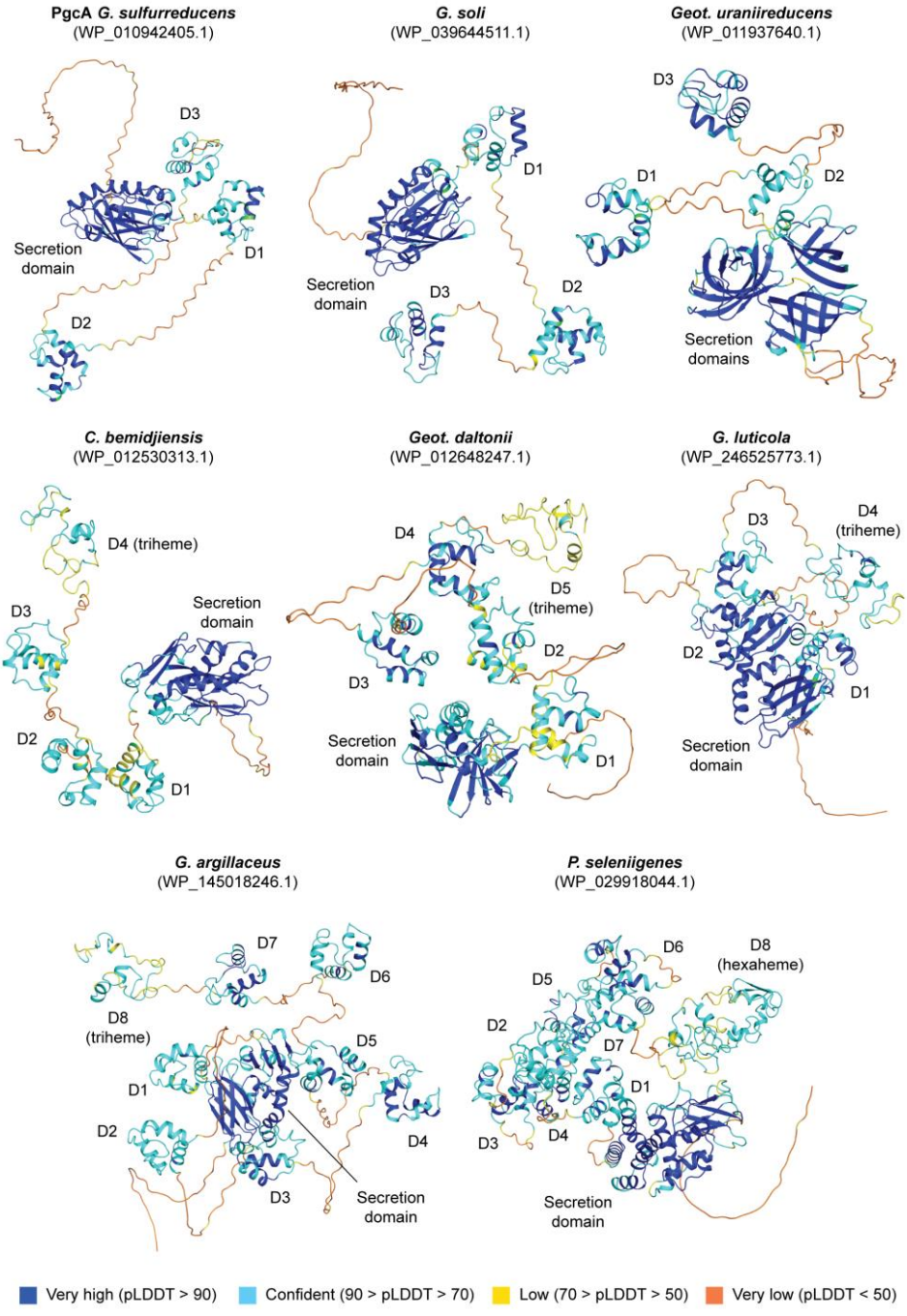


Figure 4.1. AlphaFold models of PgcA and homologs. Different PgcA-like microbial heme-tethered redox strings of *Geobacter* (G.), *Geotalea* (Geot.), *Citrifermentans* (C.) and *Pelobacter* (P.) are represented as ribbon structures. The pLDDT values are represented according to the AlphaFold protein structure database coloring. Unless indicated otherwise, all cytochrome domains are monohemic.

The first domain, classified as the secretion domain (residues 58-252, Figure 4.2A), is preceded by a small amino acid sequence (Leu[36]-Ala[37]-Gly[38]-Cys[39]), designated lipobox motif, that is crucial for the correct translocation of PgcA out of the periplasm. In fact, the deletion of residues 37 to 39 inhibited this translocation [18].

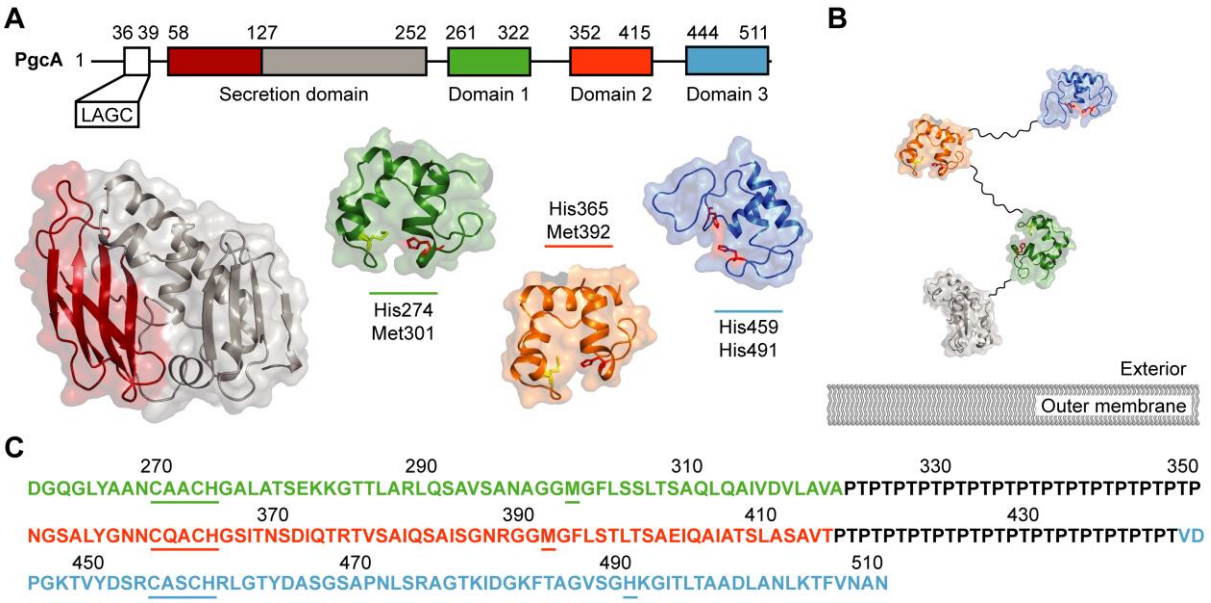


Figure 4.2. Structural features of PgcA. (A) Domain architecture of PgcA. The triheme cytochrome is composed by four domains, each indicated by different colors. This color code is respected in all figures of this chapter. The secretion domain is divided in two parts (red and gray), which separate after protein secretion. The protein's lipobox (LAGC), important for acylation signaling, is also highlighted. Cytochrome domains are represented in green (Domain 1), orange (Domain 2) and blue (Domain 3). The AlphaFold models of the different domains are represented as ribbon and surface. The heme axial ligands of each cytochrome domain are shown, with histidine and methionine axial residues in red and yellow, respectively. (B) Architecture of extracellular PgcA. After protein secretion and processing, PgcA is loose in the extracellular space, having different domains linked by unstructured stretches. (C) Amino acid sequence of PgcA's domains. The CXXCH heme-binding motifs and distal axial ligands are underlined. The proline-threonine stretches that connect the different cytochrome domains are shown in black.

The motif is recognized by the prolipoprotein diacylglyceryl transferase (Lgt), which diacylates PgcA via Cys[39]'s sulfhydryl group. Immediately after, the signal peptide of PgcA is cleaved by the lipoprotein-specific signal peptidase (LspA/SpII) and the remaining N-terminal is acylated by the apolipoprotein N-acyltransferase (Lnt) [18, 19]. Finally, the lipoprotein is secreted by a Lol-like secretion pathway [14, 18]. During this process, PgcA is converted into its mature form, in which part of the secretion domain (residues 58-126, Figure 4.2B) is cleaved [14, 18]. The exact residue at which PgcA is cleaved was previously determined by mass spectrometry, after its heterologous expression in *S. oneidensis*, and correlated with earlier observations in *Geobacter* cell lysates [14]. Nevertheless, the exact processing pattern of PgcA and the proteins involved in such mechanism can only be unequivocally identified after protein isolation directly from *Geobacter* cells. Curiously, according to

the AlphaFold model of PgcA, the Ala[126] residue (located in the frontier between the red and gray parts of the secretion domain – see Figure 4.2A) at which the protein is supposedly cleaved, is positioned at the end of a β-sheet, right before a loop, and thus highly accessible for the cleavage event.

Independently of the exact processing mechanism, the mature form of PgcA is diffusing freely in the extracellular space and contains three cytochrome domains (Figure 4.2B). These domains are linked by unstructured stretches, formed by repeats of proline and threonine residues (Figure 4.2C), which were previously suggested to be important for binding with metal oxide surfaces [14, 17, 20]. The proline residues restrict flexibility and impose structural conformations that confer higher affinity towards metal oxide surfaces, while positioning the threonine hydroxyl groups for hydrogen bonding [14, 17, 20]. Additionally, and based on the analysis of PgcA's AlphaFold model, it is also conceivable to assume that these stretches provide enough flexibility for the different cytochrome domains to interact and exchange electrons, providing PgcA with functional advantages over other triheme cytochromes. In fact, in the majority of triheme cytochromes, the positions of the hemes are fixed within a certain structural frame (Figure 4.3), thus limiting the distances at which the cytochromes can transfer electrons, while providing other important functional features [21, 22].

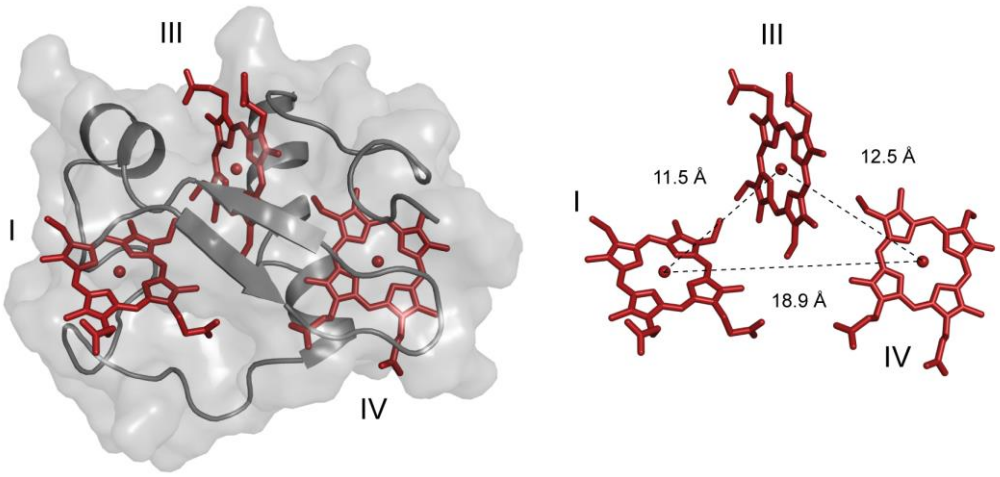


Figure 4.3. Structural depiction of a model triheme cytochrome. Ribbon and surface (in gray) representation of the solution structure of PpcA from *G. sulfurreducens* in the reduced state (lowest energy, PDB: 2LDO [23]). The heme groups are represented in red. In the right panel, the Fe-Fe distances of the porphyrin rings are indicated. Roman numerals indicate the hemes in their order of attachment to the CXXCH motif in the polypeptide chain.

Based on this data, we propose that PgcA is part of a new class of cytochromes, which we designated as "microbial heme-tethered redox strings", whose structures resemble a beads on a string arrangement, with the cytochrome domains working as beads and the unstructured stretches as strings (Figure 4.4).

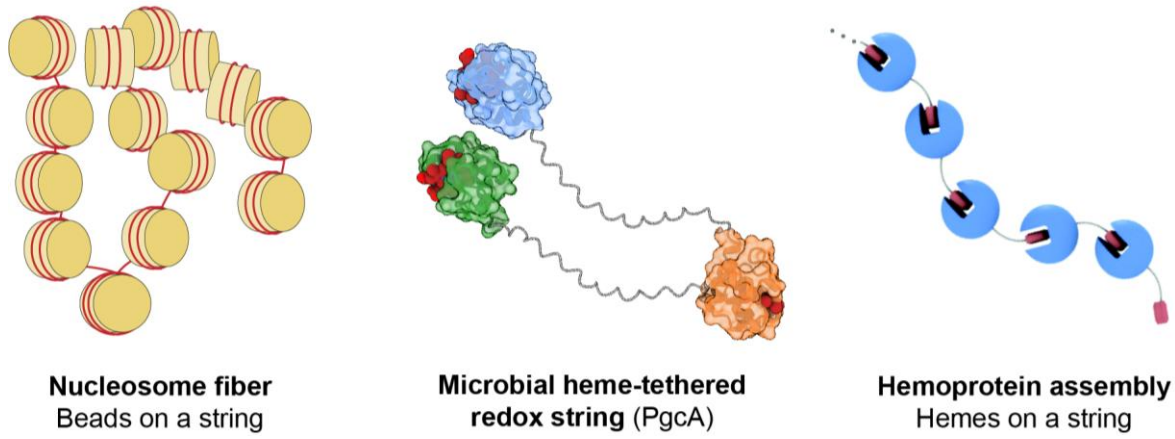


Figure 4.4. Beads on a string like arrangement of PgcA. The arrangement of the cytochrome domains of PgcA (center panel) resembles those of the nucleosome fiber (left panel), which in the absence of cations forms an extended beads on a string fiber, in which the string (red) is the DNA and the beads (yellow) are core particles of the nucleosome, formed by both proteins and DNA [24]. Synthetic hemoprotein assemblies (right panel) have also shown a similar arrangement [25].

PgcA can adopt multiple conformations not only to place the cytochrome domains in favorable positions for efficient intramolecular electron transfer but also to stretch the distances at which it can reduce extracellular acceptors. In fact, if PgcA is fully stretched, the protein should be able to transfer electrons over a 20 nm distance. A similarly flexible arrangement has been observed in supramolecular heme tethering synthetic polymers (Figure 4.4) [25, 26], but it is unprecedented in biological systems.

The amino acid sequence of PgcA was analyzed using Protein BLAST to search for sequences with high homology. Interestingly, many of PgcA's homologs possess the same predicted architecture (Figure 4.1), with multiple mono- and multiheme cytochrome domains linked by unstructured motifs, that can potentially stretch and transfer electrons over even larger distances. To gather insights into the functional mechanisms of PgcA, we used the biochemical deconstruction approach mentioned in Chapter 2 (see section 2.2.2 – ExtA), which was also explored in the study of the dodecaheme cytochrome GSU1996 [27]. This approach allowed the study of each cytochrome domain of PgcA without dealing with the inherent difficulties of producing the full-length protein while cross-correlating and translating the results for the entire system. Nevertheless, later on, we performed similar experiments with a "full-length" PgcA construct (residues 253-511) provided by Benjamin Nash and Prof. Thomas Clarke (University of East Anglia, Norwich, United Kingdom) to confirm the hypothesis established with the studies performed on the individual cytochrome domains.

4.3.2. PgcA's cytochrome domains are composed of α-helical structures

Pure and homogeneous samples of each cytochrome domain were obtained after optimization of the heterologous overexpression and multiple purification steps (Figure 4.5 and section 4.5.3).

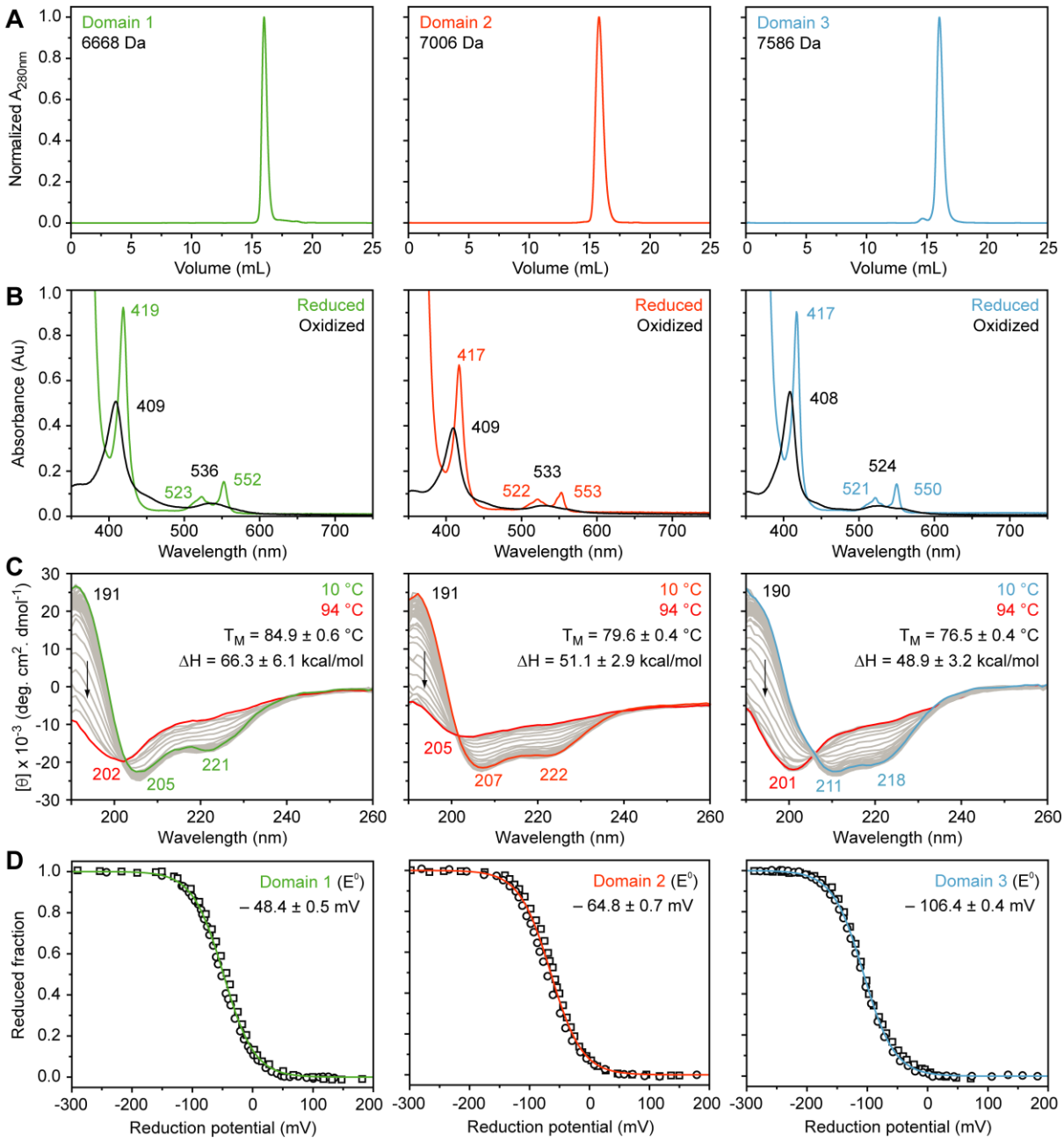


Figure 4.5. Biochemical and spectroscopic features of the cytochrome domains of PgcA. (A) SEC elution profile after injection on a Superdex 75 Increase 10/300 GL molecular exclusion column, in 100 mM sodium phosphate buffer pH 8. The molecular weight of each domain is indicated. (B) UV-visible spectral features in the reduced (colored) and oxidized (black) states. The local maxima of the UV-visible spectra are labeled. (C) Far-UV CD spectral features. The local maxima and minima of each spectrum are highlighted, as well as the midpoint of unfolding transition (T_M) and enthalpy of unfolding (ΔH). The thermal unfolding profiles can be found in Figure 4.7. The arrows indicate the variation of the ellipticity with increasing temperature. (D) Redox titrations followed by visible spectroscopy. The squares and circles represent the data points in the reductive and oxidative titrations, respectively. The solid lines indicate the fittings of a Nernst equation to the experimental data, considering a one-electron reduction.

The cytochrome domains of PgcA contain one *c*-type heme-binding motif (CXXCH), around 65 residues, and a high degree of structural homology with the monoheme cytochrome OmcF from *G. sulfurreducens* (Figure 4.6) [28].

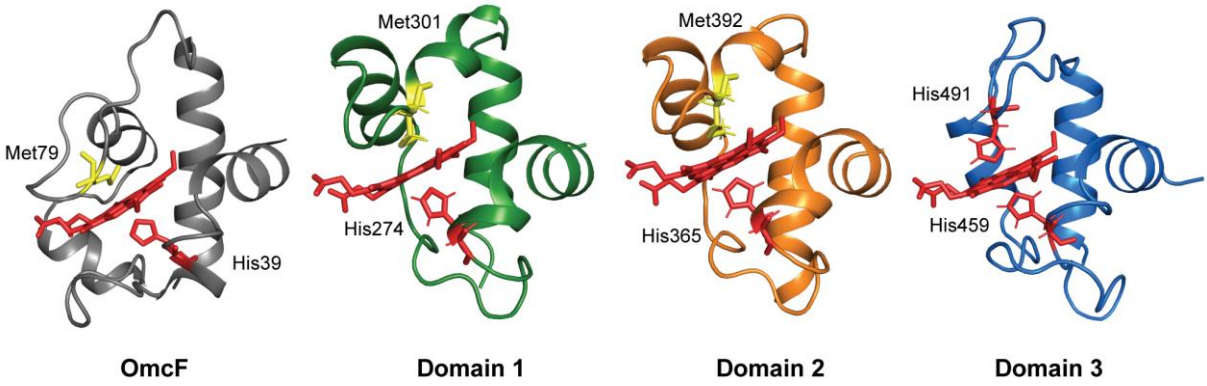


Figure 4.6. Structural homology between OmcF and the cytochrome domains of PgcA. The OmcF structure (PDB: 3CU4 [29]) and the AlphaFold models of PgcA's cytochrome domains are represented in ribbon, with the putative heme axial ligands shown as sticks. Histidine and methionine residues are shown in red and yellow, respectively. The heme groups are represented in red. The secondary structural motifs and overall arrangement of the cytochromes are highly similar, with exception of an additional peptide segment in OmcF, which forms a fourth α-helix close to the heme (left side, below the heme group).

The heme axial ligands of PgcA can be easily predicted by the AlphaFold model (Figure 4.2). While Domains 1 and 2 are distally coordinated by methionine residues (Met^{301} and Met^{392}, respectively), Domain 3 is coordinated by a histidine residue (His^{491}).

The UV-visible spectra of the different domains (Figure 4.5B) display patterns similar to those shown by low-spin hexacoordinated hemes [30]. The CD spectra (Figure 4.5C) show that the three cytochrome domains are mainly composed of α-helical secondary structural motifs, in correlation with the AlphaFold model predictions and with previous CD studies performed with the full-length protein [14]. The percentage of secondary structural elements of each domain was estimated with the BeStSel deconvolution method [31]. The results have a reasonable correlation with the AlphaFold models (Figure 4.7A and Table 4.1). The observed deviations can be attributed to the intrinsic limitations of the deconvolution process, which does not account for rare secondary structural elements and contributions from aromatic residues and large surface exposed cofactors, such as the heme groups in each domain (Figure 4.6) [32].

The CD thermal melting profiles of each cytochrome domain (Figures 4.5C and 4.7B) show that the proteins are thermoresistant, with melting temperatures of around 80 °C. This was independently confirmed by differential scanning calorimetry (DSC, Figure 4.7C). Moreover, the high degree of similarity between the CD spectra acquired before and after the thermally-induced denaturation indicate that the process is fully reversible.

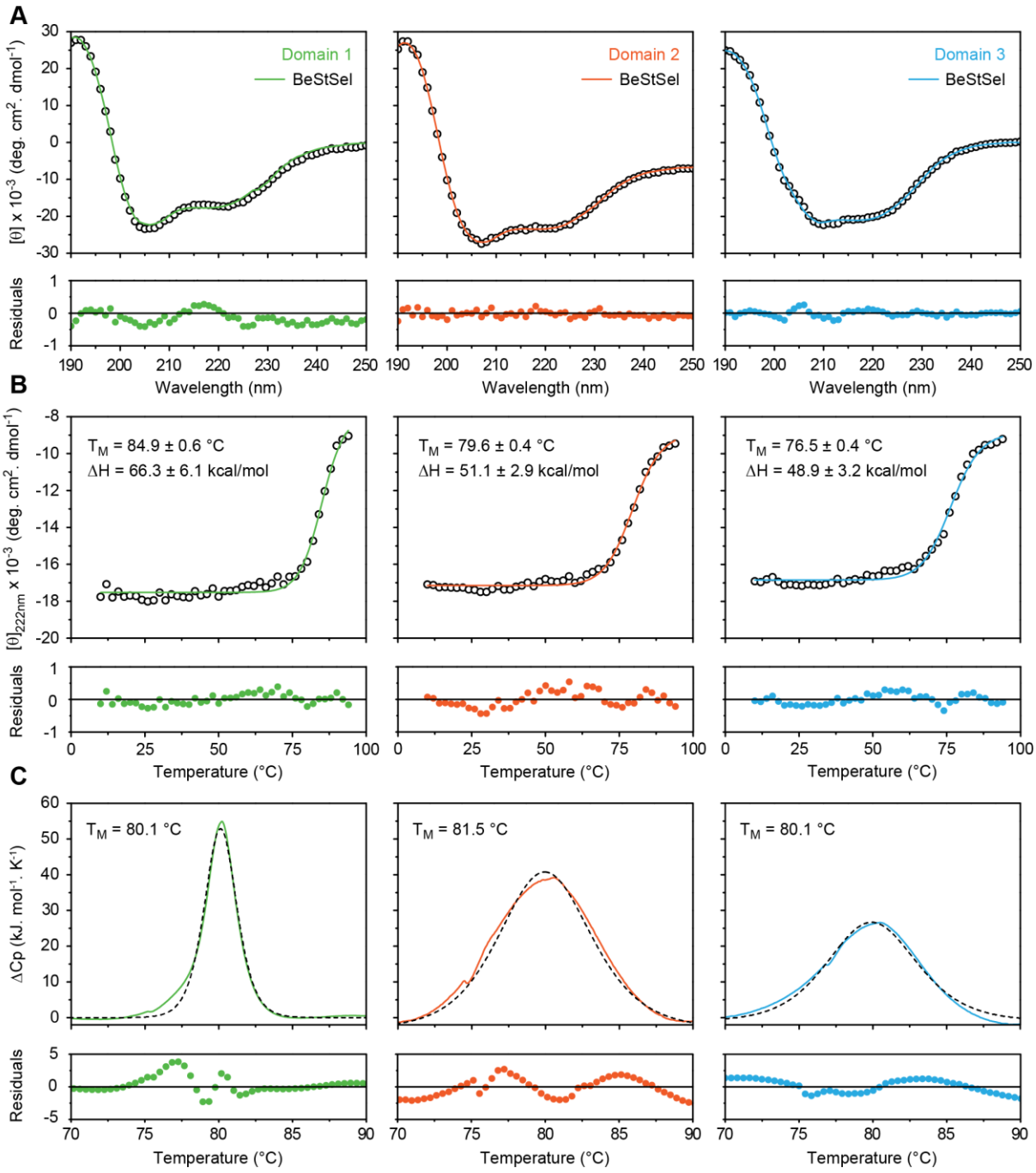


Figure 4.7. CD spectral deconvolution and thermal unfolding profiles of the cytochrome domains of PgcA. (A) Deconvolution of the far-UV CD spectra. The fittings (solid lines) were performed with BeStSel. The experimental points are represented by circles. (B) CD thermal unfolding profile. The mean residue ellipticities [θ] at 222 nm are represented as a function of temperature (circles). The solid lines represent the fitting of a two-state transition model to the experimental data. The thermodynamic parameters obtained are indicated. T_M and ΔH stand for midpoint of unfolding transition and enthalpy of unfolding, respectively. (C) DSC thermal unfolding profile. A one-peak (two-state) transition model (dashed lines) was fitted to the experimental thermograms (solid lines).

Table 4.1. Secondary structural elements of the cytochrome domains of PgcA. The percentages of secondary structural elements were either estimated from CD data using the BeStSel deconvolution method [31] (Figure 4.7) or obtained using the Dictionary of Secondary Structures of Proteins (DSSP) software [33] to analyze the AlphaFold (AF) models of each domain.

Method	Secondary structural elements (%)			
BeStSel	**α-helix**	**β-sheet**	**Turn**	**Others**
PgcA D1	40.8	10.1	6.2	42.9
PgcA D2	41.7	3.8	12.2	42.3
PgcA D3	34.1	20.5	9.8	35.6
AF (DSSP)	**α-helix**	**β-sheet**	**Turn**	**Others**
PgcA D1	72.6	0.0	12.9	14.5
PgcA D2	73.5	0.0	14.1	12.4
PgcA D3	44.2	0.0	21.9	33.9

4.3.3. PgcA contains an intramolecular redox chain

Redox titrations followed by visible spectroscopy were performed for the three monoheme cytochrome domains (Figure 4.5D). No hysteresis was observed for any of the domains since the reductive and oxidative curves of each redox titration are superimposable. Thus, the redox process is fully reversible and no major structural rearrangements occur upon reduction or oxidation of the heme groups. The heme reduction potential values obtained for Domains 1, 2, and 3 were −48.4 ± 0.5 mV, −64.8 ± 0.7 mV, and −106.4 ± 0.4 mV, respectively. An extrapolation of these data to the full-length protein indicates that PgcA should contain a redox-active window of 300 mV, ranging from −230 to +70 mV, and an apparent midpoint reduction potential of −71 mV. These values are thermodynamically compatible with the list of substrates that can be reduced by *G. sulfurreducens* and, in particular, by PgcA [6, 14]. Despite the high structural similarity between the different domains of PgcA (Figures 4.2 and 4.6), their reduction potential values are considerably distinct. There are several factors that modulate a heme reduction potential, namely the heme solvent exposure, the surrounding network of charged residues, and the nature or orientation of its axial ligands [34]. In the case of PgcA's domains, the observed differences can be explained by the nature of their heme axial ligands. The heme groups of Domains 1 and 2 have His-Met axial coordination and have higher reduction potential values compared to the one of Domain 3, which is axially coordinated by two histidine residues (His-His). On one hand, histidine residues are good electron donors and stabilize the heme's oxidized state, resulting in lower reduction potential values [35]. On the other hand, the sulfur atom of the methionine side-chain is a good electron acceptor and stabilizes the iron's reduced state, thus resulting in higher reduction potential values [36]. The small difference in reduction potential between Domains 1 and 2 might be due to variations in the heme solvent exposure or by the surrounding network of charged residues. Compared to PgcA's Domains 1 and 2, the homolog cytochrome OmcF – also possessing a His-Met axially coordinated heme – has a

considerably higher heme reduction potential value (+180 mV). This is most likely a consequence of the existence of a fourth α-helix close to the heme, which further contributes to the shielding of OmcF's heme from solvent exposure (Figure 4.6).

Overall, the reduction potential values of the different PgcA's cytochrome domains show that the protein can putatively establish an intramolecular, thermodynamically favorable electron transfer chain, by transferring electrons between its different hemes as long as they are placed within 15 Å, according to Marcus's theory of electron transfer [37]. The AlphaFold model of PgcA (Figure 4.1) shows that these domains may be as far apart as 100 Å, but the flexible stretches that connect them may be essential to promote their contact and the concomitant interaction for electron transfer events. To test this hypothesis, we performed interaction studies followed by NMR, after the proper assignment of key heme substituent signals.

4.3.4. NMR features of PgcA's cytochrome domains

NMR spectroscopy is an excellent technique to probe cytochrome's properties, as it provides information on the heme's spin state while delivering reliable fingerprints for the signals of the heme substituents and axial ligands in the reduced and oxidized states [6]. Additionally, NMR is often used to study protein-protein or protein-ligand interactions extending from static to transient regime [38]. In particular, most biophysical techniques are unsuitable to study the formation of transient complexes, such as those typically established by redox partners, which possess a short lifetime to ensure a fast turnover and a continuous electron flow [39, 40]. The most standard protocol used in NMR to map the interface between interacting molecules relies on chemical shift perturbation experiments, in which changes in the chemical environment of one molecule caused by interaction with other molecule(s) are monitored. In the particular case of cytochromes, since the interacting regions with other putative redox partners are necessarily located near the heme groups, the ^{1}H chemical shifts of the heme substituents are typically tracked. Thus, we started by acquiring 1D ^{1}H-NMR spectra of the cytochrome domains in the reduced and oxidized states, to assess the dispersion of the signals of the different heme substituents and axial ligands. The spectra show well dispersed, narrow signals (Figure 4.8), indicating that the proteins are well-folded, in line with the results obtained by CD spectroscopy (Figure 4.5C).

In general, low-spin hemes present narrower spectral regions compared to high-spin ones. In the reduced state, high-spin hemes present spectral regions ranging from −15 to 30 ppm, whereas low-spin hemes range from −5 to 10 ppm. In the paramagnetic oxidized state, high-spin cytochromes display considerably broad signals, with chemical shifts above 40 ppm, whereas low-spin cytochromes have narrower chemical shift windows, with the majority of the heme substituents signals located between −5 and 35 ppm. Therefore, from the observation of the 1D ^{1}H-NMR spectra of the three cytochrome domains in both redox states, it can be concluded that PgcA possesses three low-spin *c*-type hemes, as attested by the observed spectral widths in both redox states (Figure 4.8).

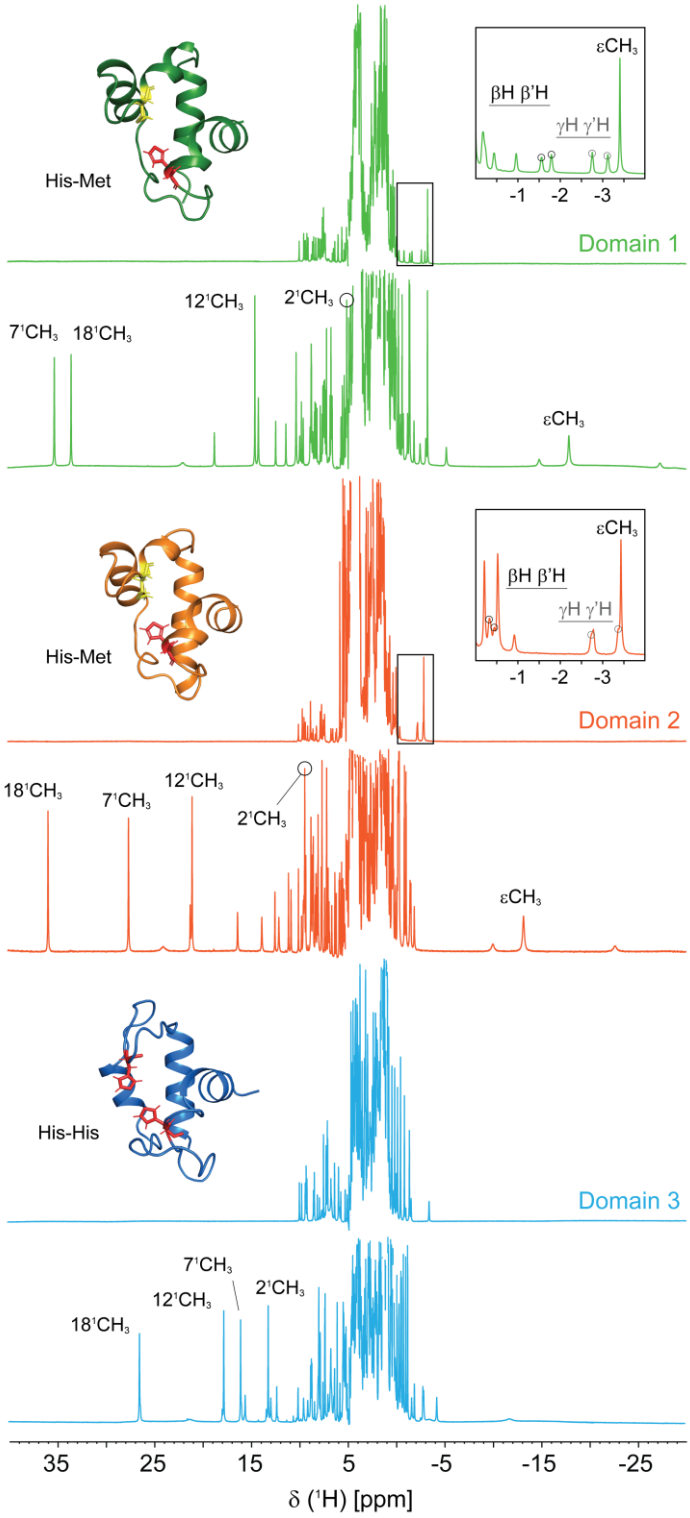


Figure 4.8. 1D ^{1}H-NMR spectra of the cytochrome domains of PgcA. The spectra of the reduced (upper) and oxidized (lower) forms are represented. All spectra were acquired in 32 mM sodium phosphate buffer with NaCl (100 mM final ionic strength) at pH 7, 25 °C. The AlphaFold models of the different cytochrome domains are represented in ribbon with the putative heme axial ligands shown as sticks. Histidines and methionine residues are shown in red and yellow, respectively. The insets on the reduced spectra of Domains 1 and 2 highlight the assigned signals of the distal axial methionine residues. 2D ^{1}H-COSY spectra were used to distinguish between βCH_2 and γCH_2 protons.

This was independently confirmed by probing the temperature dependence of the heme methyl ^{1}H chemical shifts (Figure 4.9) [41]. Based on this information, it can be further concluded that the cytochrome domains of PgcA are diamagnetic when reduced (Fe(II), S = 0) and paramagnetic when oxidized (Fe(III), S = 1/2). These features are rather convenient, since the strategy to assign heme substituents is simpler and more straightforward for low-spin hemes in both redox states.

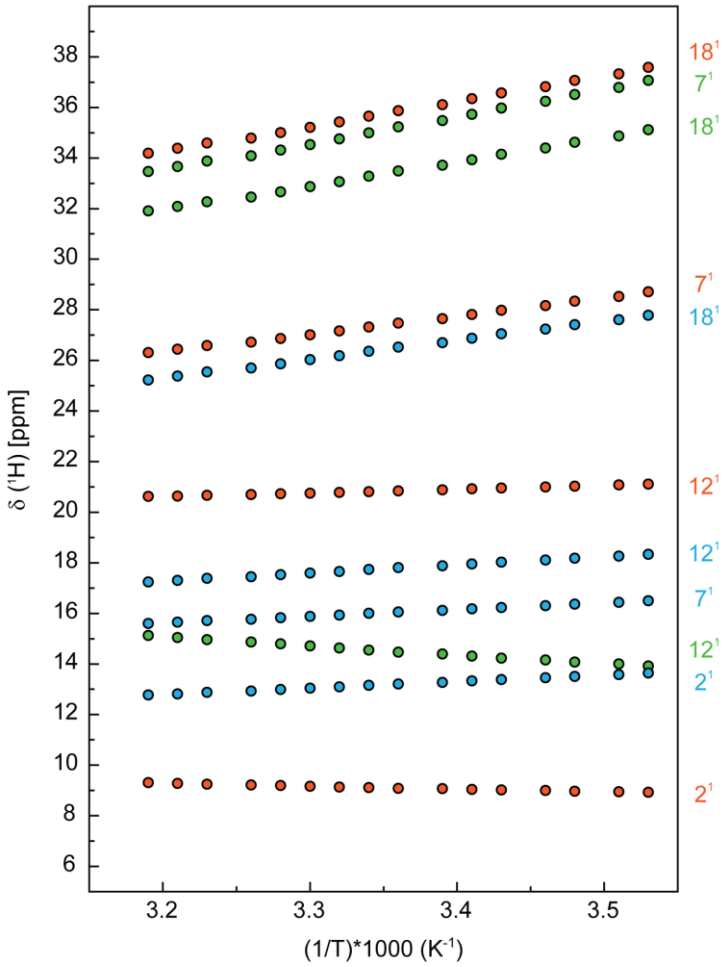


Figure 4.9. Temperature dependence of the heme methyl chemical shifts of each cytochrome domain of PgcA in the oxidized state. The heme methyl signals have a linear dependence with temperature, thus confirming the low-spin character of the three heme groups of PgcA. Additionally, the majority of the heme methyl resonances follow Curie law behaviour, i.e. their chemical shifts decrease linearly with increasing temperature [41-43]. The only exceptions are the 12^{1}CH$_3$ and 2^{1}CH$_3$ heme methyls of Domains 1 and 2, respectively, which have an anti-Curie effect, as observed previously in other *c*-type cytochromes [41]. The 2^{1}CH$_3$ heme methyl of Domain 1 could not be monitored due to the location of its signals in very crowded regions of the spectra (Figure 4.8).

In the diamagnetic reduced state, the ^{1}H NMR chemical shifts of the heme substituents are essentially affected by the heme ring-current effects [30], which are caused by the circular movement of electrons in the pyrrole rings [44]. Consequently, the signals of each type of heme substituent (Figure 4.10) are located in well-defined regions of the spectrum [34].

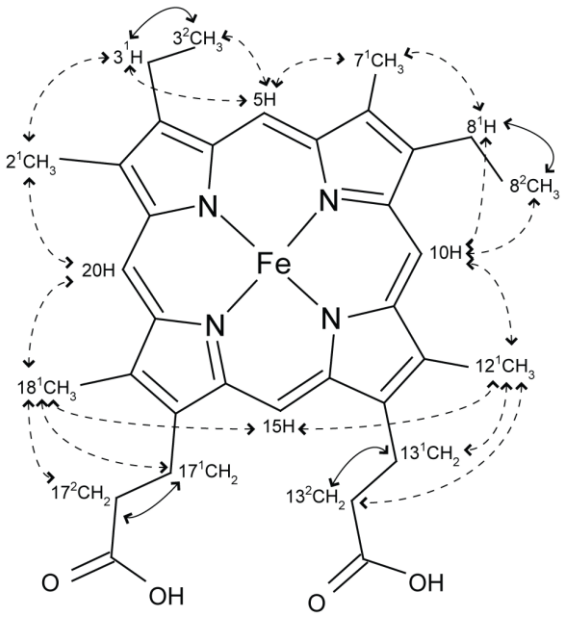


Figure 4.10. Diagram of a *c*-type heme illustrating the intraheme connectivities explored in the 2D ^{1}H-TOCSY (solid lines) and 2D ^{1}H-NOESY (dashed lines) experiments to assist the assignment of the heme substituents. The heme is labeled according to the IUPAC-IUB nomenclature for tetrapyrroles [45].

The heme propionate groups ($13^{1}CH_2$, $13^{2}CH_2$, $17^{1}CH_2$, and $17^{2}CH_2$) are the only exception, since they do not have particularly well-defined regions and are usually assigned resorting to connectivities with side-chain signals of residues located near the heme group.

The first step of the assignment procedure in the reduced state is the analysis of the 2D ^{1}H-TOCSY spectrum, in which the connectivities between the ^{1}J-coupled thioether methines (3^{1}H or 8^{1}H) and thioether methyl groups ($3^{2}CH_3$ and $8^{2}CH_3$) are identified (Figure 4.10). Following the identification of these connectivities, a 2D ^{1}H-NOESY spectrum can be used to establish spatial correlations between nuclei that are closer than 5 Å (Figure 4.10) [34].

In the paramagnetic oxidized form, in addition to the ring-current effects, the presence of an unpaired electron strongly contributes to the observed chemical shift of a heme substituent. In fact, the final observed chemical shift will strongly depend on the orientation and shape of the magnetic susceptibility tensor generated by the unpaired electron [46]. Consequently, each type of heme substituent will not possess a well-defined region in the ^{1}H NMR spectrum and, hence, can be spread over the entire NMR spectral width. Alternatively, the ^{13}C chemical shifts of the heme methyl and propionate groups have typical regions that constitute solid starting points for their specific assignment. Therefore, in addition to the 2D ^{1}H-TOCSY and 2D ^{1}H-NOESY spectra, we also acquired a 2D ^{1}H, ^{13}C-HMQC spectrum in the oxidized state for each cytochrome domain (Figure 4.11).

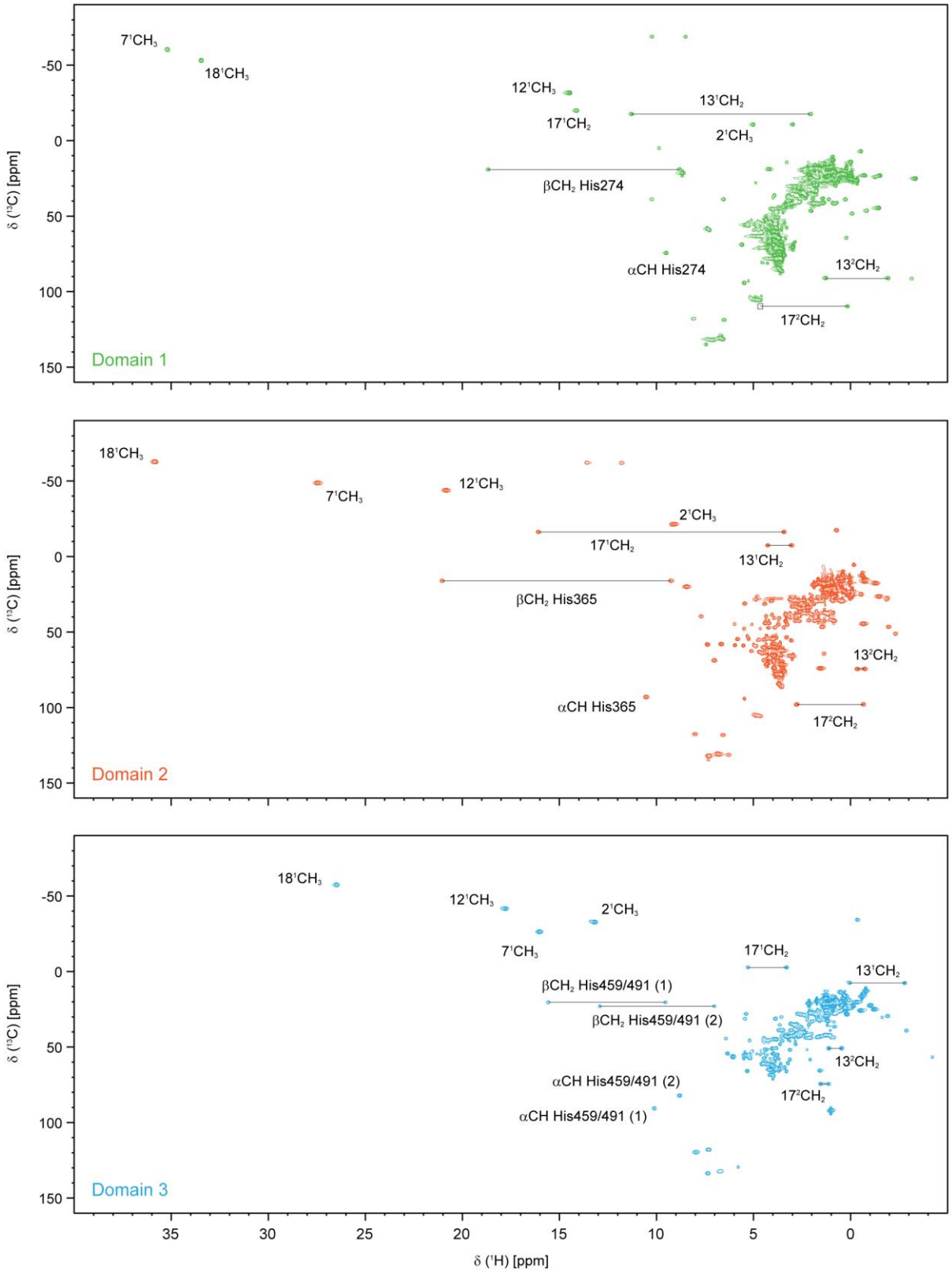


Figure 4.11. 2D ^{1}H, ^{13}C-HMQC spectra of the cytochrome domains of PgcA (pH 7 and 25 °C). The labels identify the signals of the heme methyl groups and heme propionates. Additionally, the αCH and βCH_2 signals of the heme axial histidine residues are also identified. The signals of the protons bound to the same carbon atom (CH_2 groups) are connected by a straight line.

Apart from rare exceptions [47], monoheme *c*-type cytochromes possess His-Met coordinated hemes and it has been observed that the 7^1CH_3 and 18^1CH_3 heme methyls are typically more downfield shifted in the oxidized ^1H NMR spectrum than the 2^1CH_3 and 12^1CH_3 heme methyls. This pattern is also present in the oxidized spectrum of OmcF [28], as well as in those of Domains 1 and 2 (Figure 4.8). The observed pattern is a direct consequence of the fact that each of these pairs of heme methyls is attached to diametrically opposed pyrrole rings of the porphyrin (Figure 4.10), thus being differentially affected by the asymmetric distribution of the delocalized unpaired electron on the molecular orbitals of the porphyrin [41]. Nevertheless, the chemical shift pattern observed in the oxidized spectrum of Domain 3 (Figure 4.8) is distinct because the heme axial coordination of this cytochrome domain is different (His-His) and, in these cases, the dispersion of the heme methyl signals is less conserved [46, 48, 49].

Complementarily to the heme substituents, we also extended the assignment to some of the heme axial ligands of the cytochrome domains in both redox states. In the diamagnetic reduced state, the signals of the heme axial ligands are displaced towards the low-frequency region of the ^1H spectrum due to the heme ring-current effects. In proteins without heme cofactors, the side-chain proton signals of methionine residues (βCH_2, γCH_2, and εCH_3) are found in the 1 to 3 ppm range, whereas those of the imidazole ring protons of histidines ($\varepsilon_1 H$, $\varepsilon_2 H$, $\delta_1 H$ and $\delta_2 H$) are usually located in the 6 to 11 ppm range. In heme-containing proteins, when such residues axially coordinate the heme moiety, these signals are shifted to low-frequency by an average of 5 ppm. Consequently, on one hand, the signals of axially coordinated methionine residues display a characteristic pattern, consisting of a three-proton intensity signal at around −3 ppm from the εCH_3 group and up to four resolved one-proton intensity peaks from the βCH_2 and γCH_2 groups. On the other hand, axial histidines do not show any particular fingerprint in the high-field region of the ^1H NMR spectrum, because their signals are located in the protein's signal envelope. The side-chain signals of the axially coordinated methionine residues of Domains 1 and 2 are highlighted in Figure 4.8.

In the oxidized state, there are typical spectral regions in the 2D ^1H, ^{13}C-HMQC spectrum for the αCH and βCH_2 signals of axial histidine residues (Figure 4.11) [50]. The remaining side-chain protons of axial histidine and methionine residues, located even closer to the heme iron, are usually broadened beyond detection as a consequence of the paramagnetic effect. Due to its three-proton intensity, the only exception is the methionine εCH_3 group, which despite its considerable broadening, can usually be detected in cytochromes with low molecular mass. Typically, this signal is observable between −9 and −25 ppm [51], and we were able to assign it for Domains 1 and 2 through the analysis of 2D ^1H-EXSY spectra acquired with partially oxidized samples (Figure 4.12).

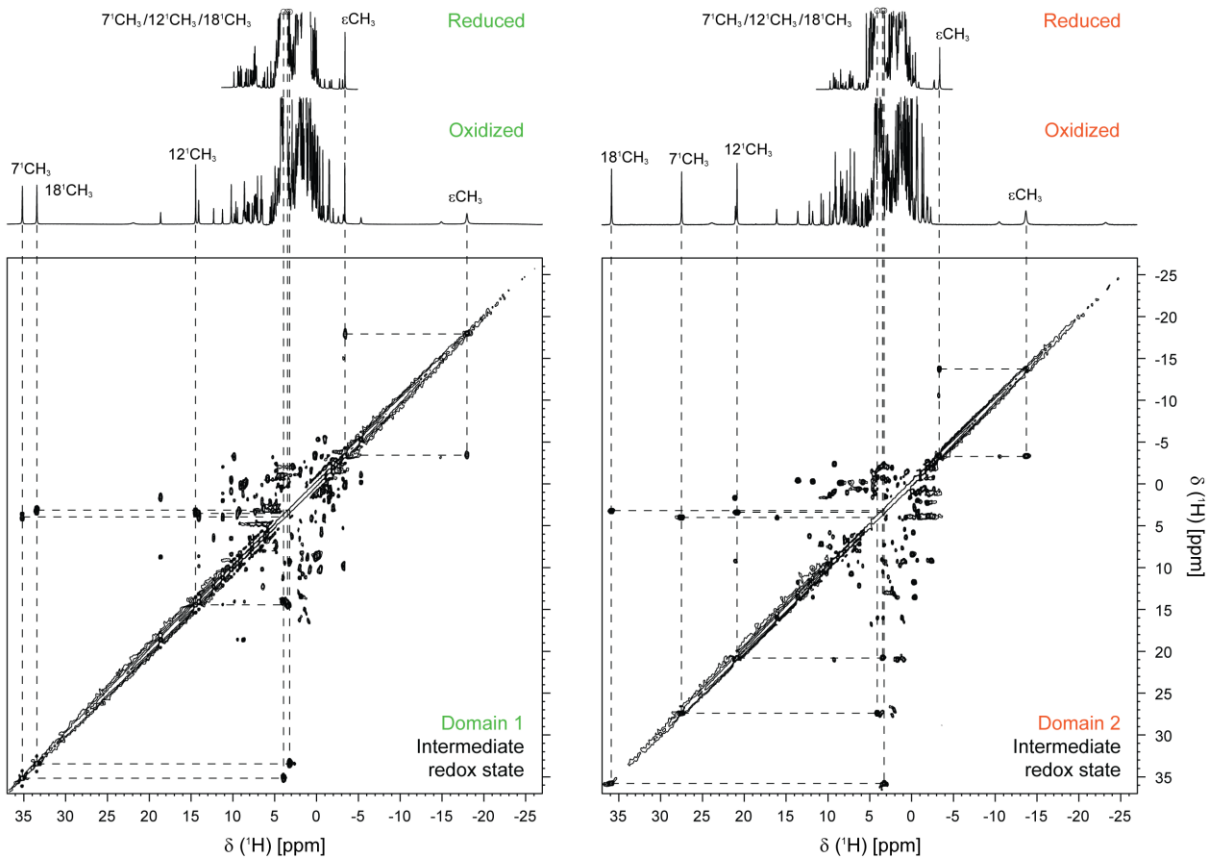


Figure 4.12. 2D ^{1}H-EXSY NMR spectra of cytochrome domains 1 and 2 of PgcA (pH 7 and 25 °C). Dashed lines connect the signals of the εCH_3 group of the axial methionine residues and of the heme methyls in the reduced and oxidized forms. Due to their location in very crowded regions in the reduced and oxidized spectra (see Figure 4.8), no cross-peaks could be observed for the heme methyls $2^{1}CH_3$. The 1D ^{1}H-NMR spectra of the fully reduced (upper) and fully oxidized (lower) cytochromes are shown on top.

These spectra show chemical exchange correlations between the signals of each nucleus at the reduced and oxidized states and were used to assign the εCH_3 group of the axial methionine residues, as well as to further confirm the assignment of the heme methyl groups in both redox states (Figures 4.12 and 4.13). The assigned signals are listed in Tables 7.3 and 7.4 for the reduced and oxidized states, respectively.

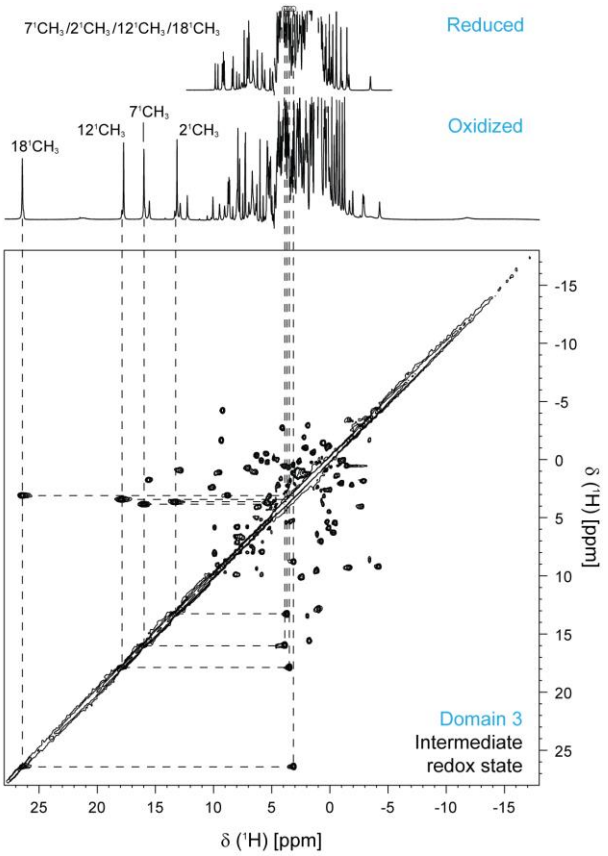


Figure 4.13. 2D ^{1}H-EXSY NMR spectrum of cytochrome domain 3 of PgcA (pH 7 and 25 °C). The cross-peaks connecting the signals of the heme methyls in the reduced and oxidized forms are indicated by dashed lines. The 1D ^{1}H-NMR spectra of the fully reduced (upper) and fully oxidized (lower) cytochromes are shown on top.

4.3.5. PgcA's cytochrome domains exchange electrons promiscuously

To map the putative interacting regions between the different cytochrome domains of PgcA by chemical shift perturbation experiments, these must be conducted by keeping the molecules in the same oxidation state so that any change can be attributed to an interaction and not to a variation in the oxidation state. Based on this requirement, these studies were performed in the oxidized state, which is experimentally advantageous, since sample manipulation in anaerobic conditions is not required, and the chemical shifts of the relevant nuclei are more dispersed compared to the reduced state. Additionally, the 1D ^{1}H-NMR spectral features of the oxidized cytochrome domains (Figure 4.8) are distinct enough so that chemical shift variations throughout a titration can be easily monitored. The heme methyl groups are the best candidates to monitor chemical shift perturbations, not only because they are significantly shifted from the diamagnetic region, but also because of their three-proton intensity that facilitates their identification.

NMR chemical shift perturbation experiments were carried out for the different combinations of cytochrome domain pairs by adding successive amounts of one cytochrome to the other and vice-versa. The analysis of the six independent titrations shows that all cytochromes interact transiently since their heme methyl signals undergo very slight chemical shift variation (Figure 4.14).

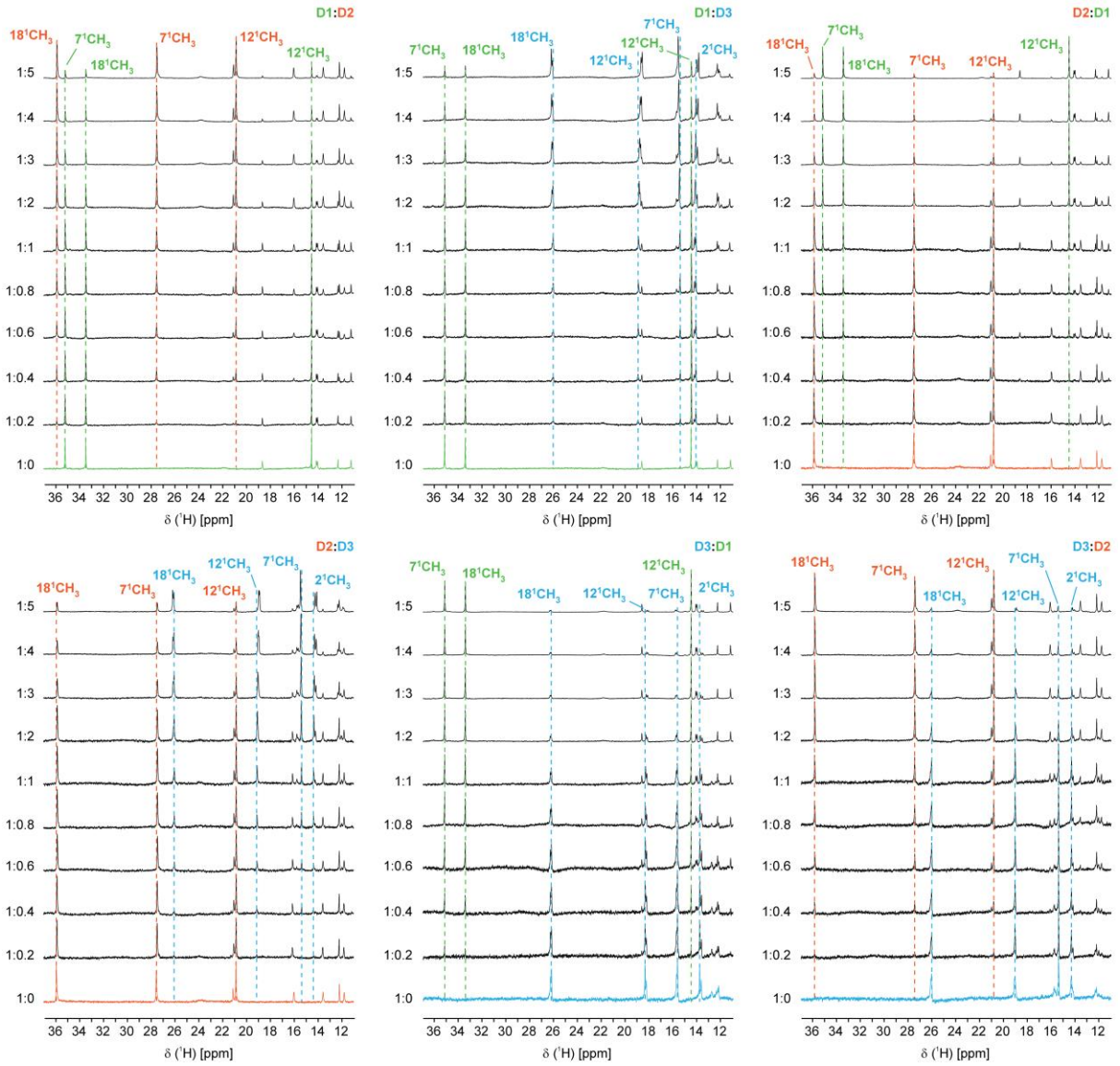


Figure 4.14. Expansions of the low-field regions of the 1D ^{1}H-NMR spectra of the cytochrome domains of PgcA obtained for the chemical shift perturbation studies. The 1D ^{1}H-NMR spectra of each domain in the presence of increasing amounts of other domain are shown for the different combinations. The molar ratios between the pairs of domains are indicated on the left side of each spectrum. The heme methyls of the different domains are labeled.

The formation of low-affinity complexes between redox partners has been observed in other studies, and is crucial for fast turnover and continuous electron flow [52, 53]. To complement these experiments and to unequivocally confirm that the different cytochrome domains of PgcA exchange electrons upon interaction, we explored the distinct NMR spectral features of the domains in the reduced and oxidized states. By adding one cytochrome to the other in distinct redox states, and upon formation of the putative transient complex, we can observe the emerging signal features of the 1D ^{1}H-NMR spectrum of the cytochromes mixture and assess possible changes in their redox state [54].

As for the NMR chemical shift perturbation experiments, six independent assays were performed for the possible combinations, by adding one oxidized cytochrome domain to a reduced one, in anaerobic conditions (Figure 4.15).

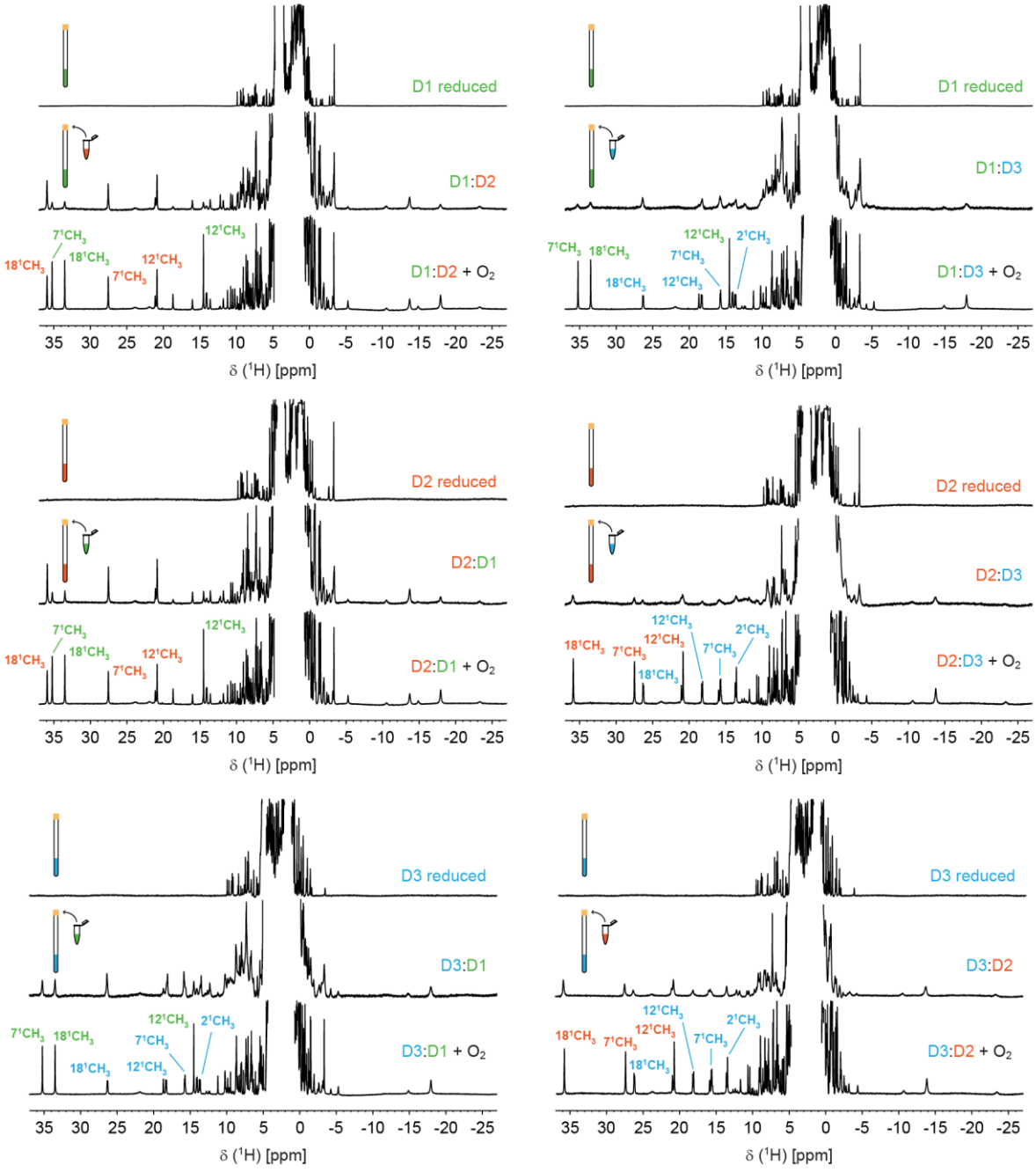


Figure 4.15. Electron transfer reactions between the cytochrome domains of PgcA monitored by NMR. The 1D ^{1}H-NMR spectra of each domain in the fully reduced state (top), after addition of an equimolar amount of other domain in the oxidized state (middle) and after exposure to atmospheric O_2 (bottom) are shown for the different combinations. For clarity, the spectra in the reduced state and after exposure to atmospheric O_2 are reproduced in different panels. In all spectra, the heme methyls of the different domains are labeled.

The results show that the cytochrome domains exchange electrons promiscuously until they reach an intermediate oxidation state, although at different extents. This can be quantified by comparing the area of the heme methyl signals of the two cytochrome domains in the intermediate redox state with those in the fully oxidized spectra. The calculated percentages show that, as expected, the cytochrome domains with more negative heme reduction potential values (−106, −65, and −48 mV for Domains 3, 2, and 1, respectively) transfer electrons to a greater extent than those with less negative heme reduction potential values (Figure 4.16A).

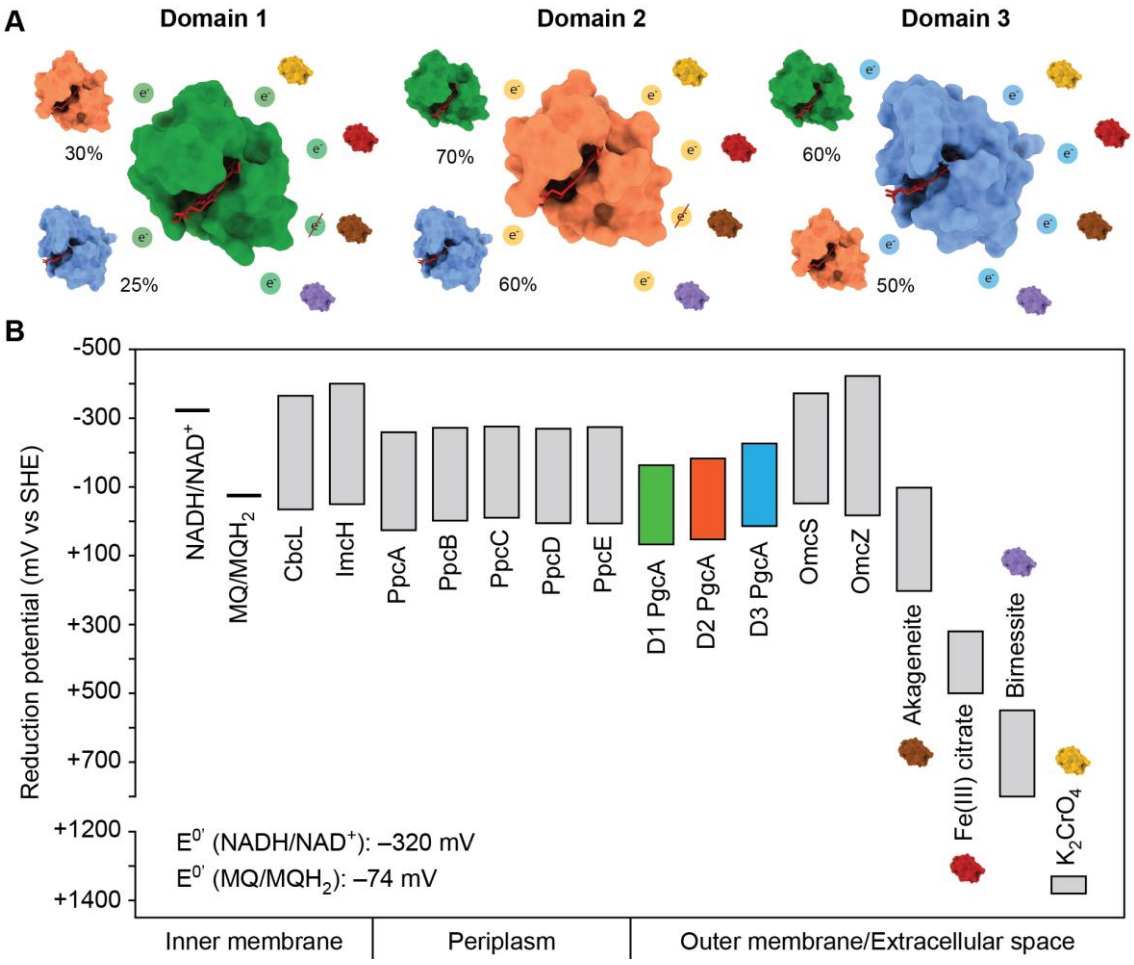


Figure 4.16. Role of PgcA in the extracellular electron transfer pathways of *G. sulfurreducens*. (A) Summary of the electron transfer experiments between the different cytochrome domains of PgcA and different electron acceptors. The cartoons of Domains 1 (green), 2 (orange) and 3 (blue) are shown with their heme groups in red. The cartoons of akageneite, birnessite, Fe(III) citrate, and potassium chromate are shown in brown, purple, red, and yellow, respectively. The percentages correspond to the extent of the electron transfer reactions between each pair of domains. (B) Histogram comparison of the redox-active windows of *G. sulfurreducens* cytochromes and different electron acceptors. The redox-active windows of each cytochrome were determined from potentiometric redox curves considering the 1% to 99% range for protein reduction/oxidation. The cellular localization of the different cytochromes/redox pairs is indicated in the bottom of the graph. The data used for the histogram can be found in Table 7.5.

In the opposite scenario, electron transfer is still observable but to a minor extent (Figure 4.16A). These variations in the degree of reduction/oxidation of the different proteins correlate with the overlap of their redox-active windows (Figure 4.16B). This means that in full-length PgcA, electron transfer should mostly follow the expected thermodynamically favorable route (Domain 3 – Domain 2 – Domain 1), but the cytochrome is likely capable of promiscuously transferring electrons between all domains, which can be advantageous during mineral reduction.

4.3.6. The cytochrome domains of PgcA reduce several electron acceptors

PgcA is a terminal reductase of extracellular Fe(III) and Mn(IV) oxides [14] and it might also be responsible for the reduction of other known electron acceptors of *G. sulfurreducens*. To assess this, we performed spectrophotometric analysis of the reduced cytochrome domains of PgcA in the presence of Fe(III) oxides, Mn(IV) oxides, Fe(III) citrate, and potassium chromate (Figure 4.17).

Fe(III) oxides exist as a heterogeneous mixture of insoluble particles with different reduction potential values in Nature, which have the tendency to change as they are being reduced [55, 56], meaning that these microorganisms require sophisticated mechanisms to cope with this challenge. Fe(III) oxides in the form of akageneite (β-FeOOH) present a redox-active window between −100 and +200 mV [55]. Freshly prepared akageneite is usually fully oxidized and has a reduction potential above 0 mV [57]. Upon addition of an excess of freshly prepared akageneite to Domains 1 and 2 in the reduced state, no differences in the visible spectra were observed (Figure 4.17). However, Domain 3 was able to reduce akageneite, as it can be seen by the differences in visible spectra of the cytochrome (Figure 4.17). In fact, the intensities of the typical bands of the reduced state – β (521 nm) and α (550 nm) – decrease upon addition of akageneite, as does the intensity of the Soret band, which also undergoes a blueshift from 417 to 408 nm.

One of the most common forms of Mn(IV) oxides in soils and natural aquatic systems is birnessite ($Na_xMn_{2-x}(IV)Mn(III)_xO_4$, with x ~ 0.4), which presents a reduction potential of +612 mV [58, 59]. All cytochrome domains of PgcA were able to reduce birnessite (Figure 4.17). A similar spectral variation was observed for all the cytochrome domains upon the addition of Fe(III) citrate and potassium chromate (Figure 4.17), which are soluble forms of the Fe(III) and Cr(IV) cations and present reduction potentials of +370 mV [55] and +1350 mV [60, 61], respectively. Overall, the results obtained indicate that the cytochrome domains of PgcA can reduce both soluble and insoluble forms of electron acceptors of *G. sulfurreducens* (Figures 4.16 and 4.17).

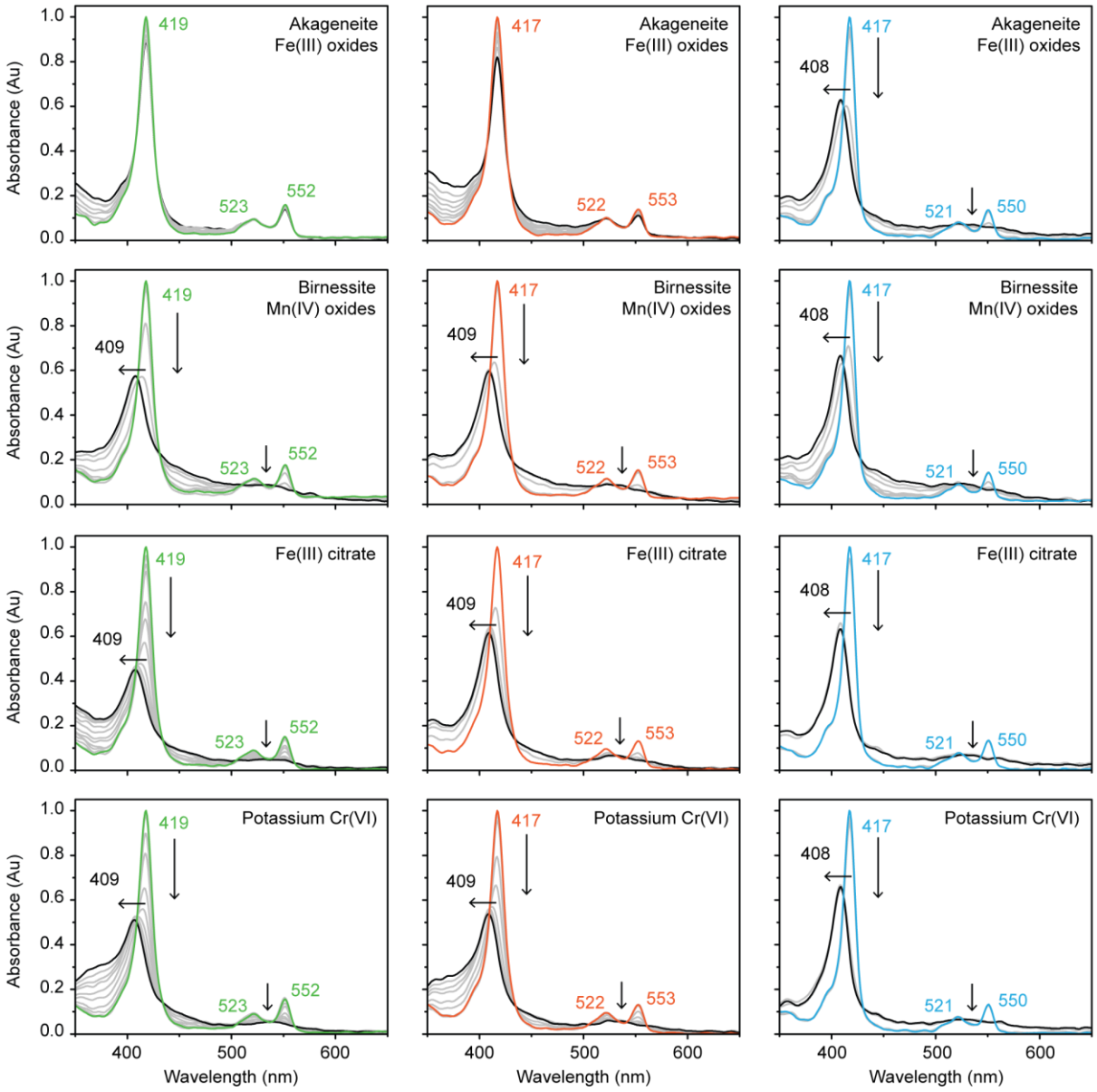


Figure 4.17. Electron transfer experiments of the cytochrome domains of PgcA with electron acceptors. The spectra of the fully reduced domains are represented in color, while the final spectra acquired for each experiment are represented in black. The intermediate spectra are represented in gray. The variations of the UV-visible spectral features of PgcA's cytochrome domains upon addition of the different electron acceptors are highlighted by black arrows.

4.3.7. Proline-threonine stretches are not responsible for Fe(III) oxides binding in PgcA

Zacharoff and co-workers have previously shown that PgcA can bind Fe(III) oxides, but not Fe(II) oxides or mixed Fe(II)-Fe(III) minerals [14]. By incubating PgcA with these metals and after recovery of the remaining soluble fractions, it was shown that this cytochrome selectively binds oxidized metals and releases them after their partial or complete reduction [14]. These results, together with evidence from previous studies [17], indicated that the protein stretches of proline and threonine residues were responsible for such binding. Considering that our data shows that Domain 3 can

reduce Fe(III) oxides, we decided to evaluate if the individual cytochromes domains of PgcA, in the absence of the repeats of proline and threonine residues, could also be involved in the binding of Fe(III) oxides. Using the methodology adopted by Zacharoff and co-workers [14], each cytochrome domain was incubated with Fe(III) oxides, and the resulting soluble fractions were recovered. The results show that only Domain 3 can bind Fe(III) oxides, to an extent of 20% (Figure 4.18), which correlates well with the fact that this domain was the only one capable of reducing this metal.

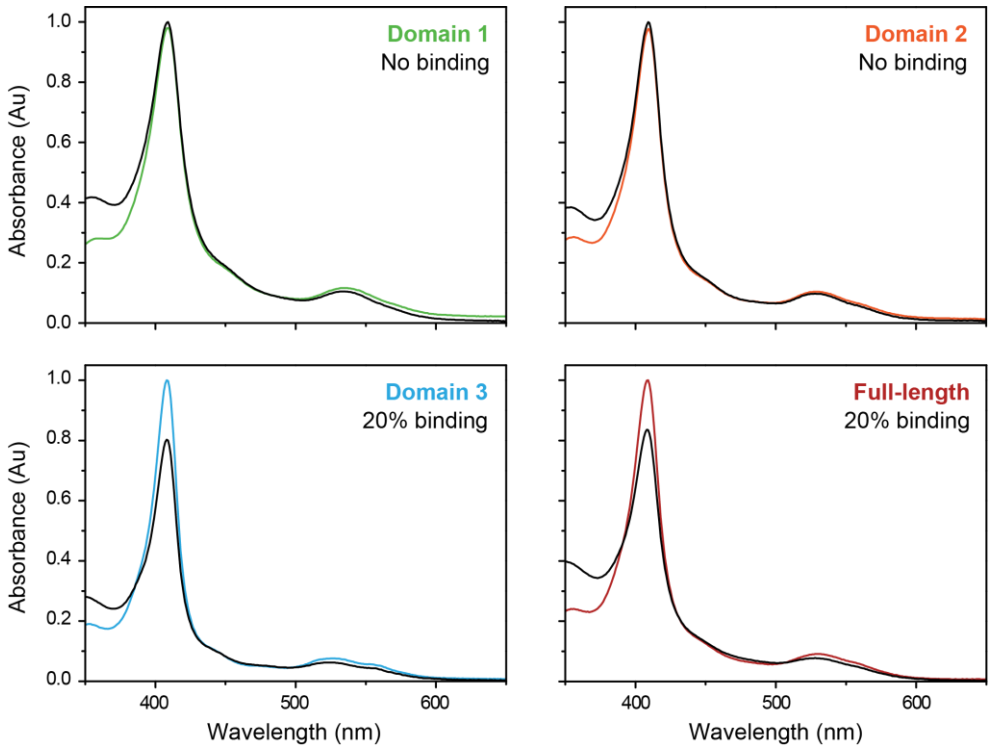


Figure 4.18. Binding experiments between PgcA and Fe(III) oxides. The spectra of each cytochrome before and after incubation with Fe(III) oxides are represented in color and black, respectively.

Nevertheless, the amount of Fe(III) oxides bound to Domain 3 is considerably smaller than the one observed by Zacharoff and co-workers [14]. In their study, the experiments were carried out with the three cytochrome domains linked by the flexible protein stretches of proline and threonine residues, and, in their case, the protein was completely bound to Fe(III) oxides [14]. Therefore, our results partially correlated with their conclusions regarding the relevance of the disordered stretches of proline and threonine residues repeats in metal binding. Indeed, since only about 20% of the protein was bound to Fe(III) oxides in our experiments with the cytochrome domains, it was conceivable to infer that the flexible protein stretches of proline and threonine residues were the main drivers of metal binding in PgcA.

Surprisingly, by performing similar experiments with a full-length PgcA construct, we observed that under our experimental conditions, the triheme cytochrome binds to Fe(III) oxides to an extent of 20% (Figure 4.18), similar to the one observed for Domain 3, but quite distinct from the one observed by Zacharoff and co-workers [14]. This suggests that under our experimental conditions,

Domain 3 is the main driver of Fe(III) oxides binding, and that the proline-threonine stretches are important "solely" for providing elasticity and promoting electron transfer events at variable distances. Based on the nanocrystalline structure of akageneite and on the AlphaFold models of the three structurally homologous domains, the more positive electrostatic surface of Domain 3 might explain its prevalence in the binding of the negatively charged ferric oxyhydroxides (Figure 4.19).

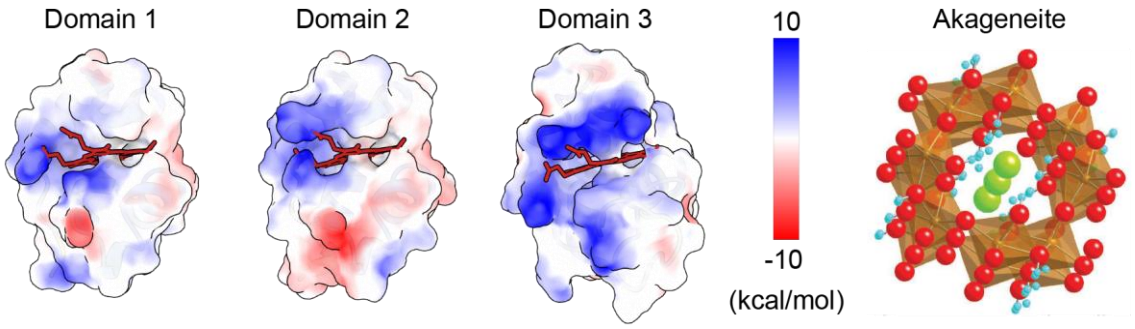


Figure 4.19. Electrostatic surface of the cytochrome domains of PgcA and akageneite nanocrystalline structure. The Coulombic electrostatic potentials were calculated from atomic partial charges and coordinates according to Coulomb's law using default values in UCSF ChimeraX 1.4 [62]. The vertical bar shows the color gradient, from red to blue, corresponding to the determined potential values in kcal/mol at 298 K. In the structure of nanocrystalline akageneite (Crystallography Open Database: 9002990 [63]), O^{-} and OH^{-} anions are shown in red, Fe^{3+} cations in orange, H^{+} cations in blue and Cl^{-} anions in green. The akageneite cartoon was reproduced from [64].

Additional structural features might play a role in this process, such as the apparently higher heme solvent exposure observed in the AlphaFold model of Domain 3, but experimentally determined structures of the three cytochrome domains are required to make such assessments.

4.3.8. Domain 3 is intramolecularly recharged for continuous Fe(III) oxides reduction

The unique affinity and reductive ability of Domain 3 towards Fe(III) oxides can be used as a paradigm to establish the existence of a putative intramolecular electron transfer chain in PgcA. In fact, considering that out of the three isolated cytochrome domains, only Domain 3 was able to reduce Fe(III) oxides, a full-length PgcA sample should not completely oxidize in the presence of Fe(III) oxides, unless intramolecular electron transfer and consequent oxidation of Domains 1 and 2 occurs. To test this hypothesis, we initially used UV-visible spectroscopy to monitor the electron transfer reaction between the triheme cytochrome and Fe(III) oxides. Full-length PgcA was completely oxidized by Fe(III) oxides, as it can be seen by the differences in the visible spectra of the cytochrome (Figure 4.20A). In fact, the intensities of the typical bands of the reduced state – β (522 nm) and α (552 nm) – decrease upon addition of akageneite, as does the intensity of the Soret band, which also undergoes a blueshift from 418 to 410 nm. These results confirmed the existence of intramolecular electron exchange in PgcA.

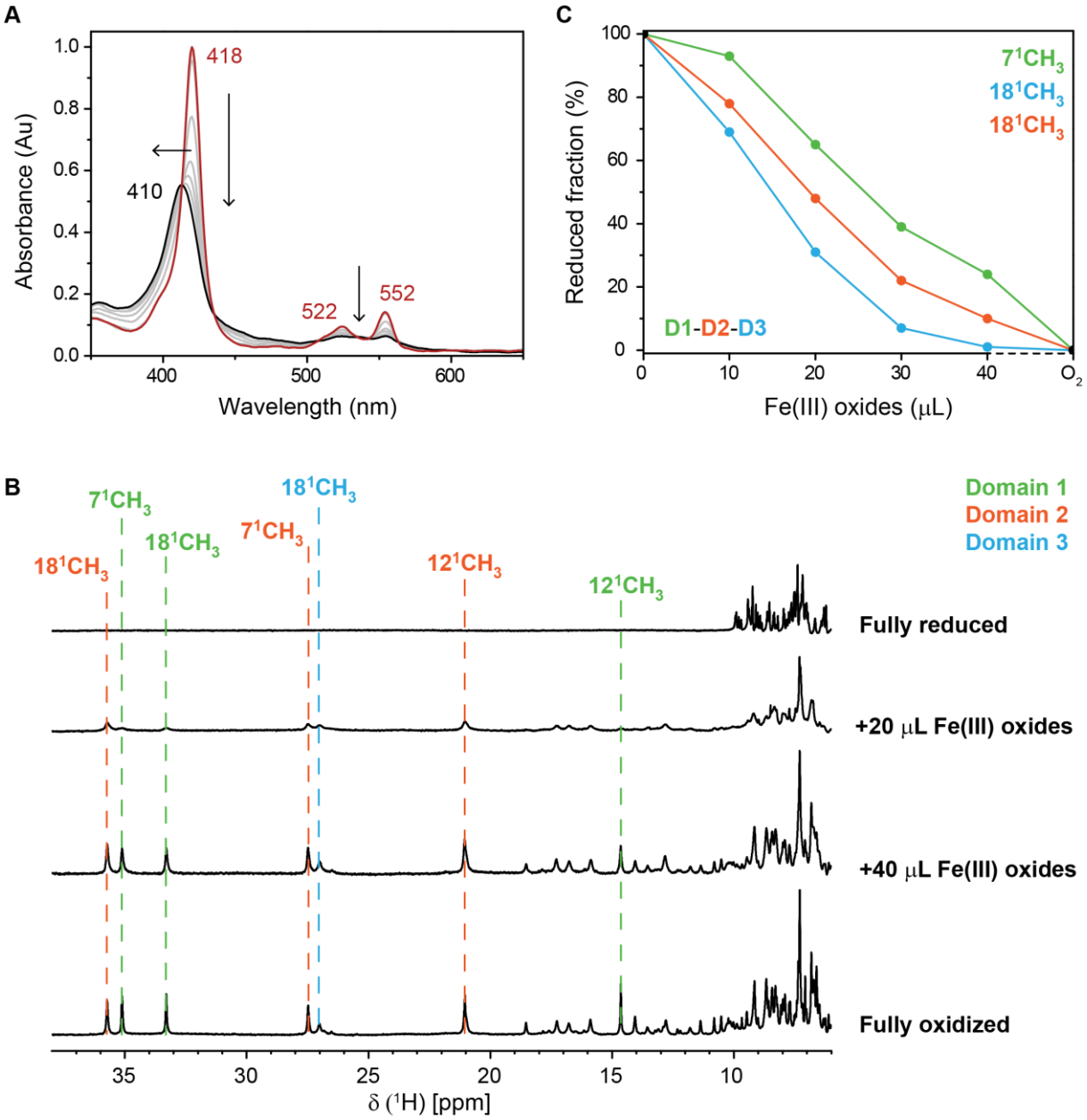


Figure 4.20. Monitorization of Fe(III) oxides reduction by full-length PgcA. (A) UV-visible spectra of full-length PgcA upon addition of Fe(III) oxides. The spectrum of the fully reduced cytochrome is represented in red, while the final spectrum is represented in black. The intermediate spectra are represented in gray. The variations of the UV-visible spectral features upon addition of Fe(III) oxides are highlighted by black arrows. (B) Variation of the low-field 1D ^{1}H-NMR spectra of full-length PgcA upon addition of Fe(III) oxides. The dashed lines indicate the different heme methyl resonances of the cytochrome domains 1 (green), 2 (orange) and 3 (blue). (C) Reduced fractions of each cytochrome domain along the Fe(III) oxides titration.

Nevertheless, the UV-visible spectral features of the three cytochrome domains are too similar for us to follow their individual oxidation process throughout their titration with ferric oxyhydroxides. Alternatively, we used the distinct NMR spectral features of the three cytochrome domains (Figure 4.8) to monitor this process in the full-length protein. Since NMR has inherent size limitations (as the molecular weight of a protein increases, its tumbling rates decrease and the T_2 relaxation rates

increase, leading to a decrease in signal-to-noise ratio), we were initially concerned with the eventual quality of the NMR spectra of PgcA, since cytochromes of similar size usually have large linewidths and yield poor spectra that are difficult to analyze [65]. However, the 1D ^{1}H-NMR spectra of full-length PgcA (27.6 kDa) in the fully reduced and fully oxidized states (Figure 4.20B) show narrow linewidths, and resonances that superpose almost perfectly with those of the individual cytochrome domains. The narrow signal linewidths of each cytochrome domain in the full-length protein (35-40 Hz), similar to those observed in the individual cytochrome domains (30-35 Hz), further support the view that the proline-threonine stretches allow the different cytochrome domains to tumble independently and almost freely in solution. These spectral features allowed us to track the oxidation profiles of each cytochrome domain along the titration with Fe(III) oxides (Figure 4.20B).

In particular, we calculated the reduced fractions of each cytochrome domain along the redox reaction by integrating the ^{1}H NMR heme methyl signals of each cytochrome domain in each intermediate redox state, relative to the fully oxidized state (Figure 4.20C). These results show that upon each addition, all cytochrome domains become more oxidized, meaning that they either reduce Fe(III) oxides (in the case of Domain 3) or participate in intramolecular electron transfer (in the case of Domains 1 and 2). The reduced fractions of each domain decrease according to their heme reduction potential, following the expected thermodynamic route. These observations confirm that while Domain 3 reduces Fe(III) oxides, Domains 1 and 2 are able to recharge it for additional metal reduction events.

4.4. Conclusions

The model of the final steps of EET in *G. sulfurreducens* has been reassessed several times in the past two decades. The first models predicted that electrons reached extracellular acceptors through cytochromes embedded in the outer membrane facing the extracellular space or via pili-based nanowires [66]. These models were revisited upon the discovery of cytochrome nanowires [7], which together with porin-cytochrome complexes are believed to be the main protagonists of the final steps of EET [6]. Concomitantly, it has been suggested that *Geobacter* pili have a secretion role [67]. Nevertheless, the pili-nanowire paradigm is still under debate [68, 69]. Using a combination of bioinformatic, structural, and functional studies, we found sufficient evidence that microbial heme-tethered redox strings provide an additional mechanism for the reduction of insoluble metal oxides in *G. sulfurreducens* and other electroactive bacteria (Figure 4.21).

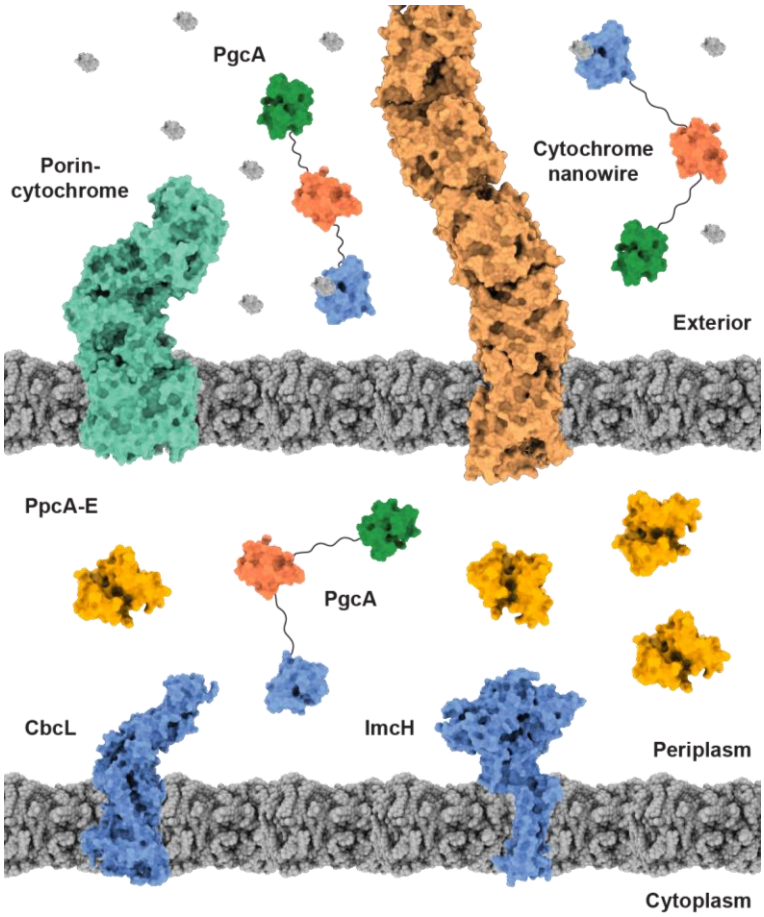


Figure 4.21. Proposed model of extracellular electron transfer in *G. sulfurreducens*. The main protagonists of the electron transfer pathways towards extracellular acceptors are represented as cartoons. CbcL and ImcH are inner membrane-associated quinone oxidoreductases that transfer electrons to the PpcA-family cytochromes, which in turn are responsible for bridging the gap between the inner and outer membrane cytochromes [6]. Prior to secretion, PgcA might also contribute to periplasmic electron transfer. The final steps of EET are warranted by porin-cytochrome complexes, cytochrome nanowires, and PgcA, which recruits metals and might synergistically work with these proteins to maximize the reduction of electron acceptors. The light gray cartoons represent metal oxides. For simplification, the secretion domain of PgcA is not represented.

PgcA and homologs possess a tethered structure, in which different cytochrome domains are connected by low-complexity regions that provide sufficient flexibility to promote intra- and intermolecular electron transfer at variable nanometer distances. Our results show that the cytochrome domains of PgcA from *G. sulfurreducens* can exchange electrons promiscuously through the formation of transient complexes, thus suggesting a cooperative mechanism of reduction of electron acceptors. In fact, while full-length PgcA had been previously shown to efficiently reduce Fe(III) oxides in vitro [14], our results revealed that such reduction is carried out by a single domain (Domain 3). Upon reduction, metal oxides detach from Domain 3, which is then available to be recharged with electrons by Domains 1 and 2.

Depending on its cellular location, PgcA can have several putative redox partners. In the periplasm, the most likely candidates are either the inner membrane quinone oxidoreductases or the PpcA-

family cytochromes. Consequently, during secretion, PgcA can be loaded with electrons and readily available to reduce extracellular metals. Alternatively, if PgcA is secreted in an oxidized or intermediate redox state, it might work synergistically with cytochrome nanowires (OmcE, OmcS, and OmcZ) in the extracellular matrix, thus establishing a continuous flow of electrons for metal reduction. The overall efficiency of the process can be dependent on PgcA's diffusion rates, as a consequence of differentiated protein-protein and protein-metal contacts, as well as biofilm density.

4.5. Materials and methods

4.5.1. DNA manipulation

The *gsu1761* gene sequence, encoding for the PgcA protein (WP_010942405.1), was retrieved from *G. sulfurreducens* PCA's genome on the KEGG database [70], under the accession number T00155. The codons coding for residues 261 to 322 (Domain 1), 352 to 415 (Domain 2), and 444 to 511 (Domain 3) were amplified from genomic DNA and inserted either into vector pVA203 [71] (Domains 1 and 3) or pCK32 [72] (Domain 2) using a Phusion High-Fidelity DNA Polymerase (Thermo Fisher Scientific) together with the primers listed in Table 4.2, according to the RF cloning protocol [73]. Taq DNA polymerase (VWR) was used for colony PCR screening and DpnI (Thermo Fisher Scientific) for digestion of the template DNA. The intermediate PCR product and the final plasmid DNA were purified using the NZYGelpure and NZYMiniprep kits (NZYTech), respectively. PCR products and plasmid DNA quantification and purity were determined using a NanoDrop spectrophotometer ND-1000 (Thermo Fisher Scientific) and the DNA was sequenced by STAB VIDA (Caparica, Portugal).

Table 4.2. Sequences of the primers used to produce the vectors containing the genes of the different cytochrome domains of PgcA. The DNA sequences specific for the pVA203 (Domains 1 and 3) and pCK32 (Domain 2) plasmids, and for the different parts of the *gsu1761* gene (encoding for the cytochrome domains of PgcA) are highlighted in blue and green, respectively. The primers forward (fw) and reverse (rv) were purchased from Invitrogen. D1, D2 and D3 stand for Domains 1, 2 and 3.

Primer	DNA sequence (5' – 3')
D1_fw	TTTCGCTACCGTTGCGGCCGCCGACGGACAAGGATTGTACGC
D1_rv	AGCTTGTCGACGGAGCTCGAATTCAAGCCACTGCCAAGACGTC
D2_fw	TTTCGCTACCGTTGCGGCCGCCAACGGGTCGGCACTCTAC
D2_rv	AGCTTGTCGACGGAGCTCGAATTCATGTGACAGCCGATGCCAG
D3_fw	TTTCGCTACCGTTGCGGCCGCCGTCGACCCGGGCAAGACC
D3_rv	AGCTTGTCGACGGAGCTCGAATTCAGTTTGCGTTGACAAAGGTCT

4.5.2. Protein sequence and structural homology analysis

A Protein BLAST [74] was performed with PgcA's sequence, using the RefSeq Select protein sequences database and the default blastp algorithm (February 2023).

The AlphaFold models of the closest homologs were either retrieved from the AlphaFold protein structure database [75, 76] or produced using the AlphaFold2_MMSeqs2 GitHub ColabFold notebook [77]. The quality of the individual models was evaluated based on the per-residue estimate of confidence produced by AlphaFold, designated pLDDT, which is based on the lDDT-Cα metric [78]. In the case of the cytochrome domains of PgcA, the *c*-type hemes were fitted in PyMol 2.5.2 [79] using local alignments and libraries of *c*-type hemes, containing the corresponding cysteine residues for the covalent binding of the heme vinyl groups and the proximal axial histidines for Fe coordination. These libraries were built by retrieving these residues and the corresponding *c*-type heme group

from the structure of the OmcF cytochrome (PDB: 3CU4 [29]). Side-chain packing optimization and all-atom energy minimization protocols were not employed (see section 7.4). The final models were analyzed in UCSF ChimeraX 1.4 [62] and represented using either this program or PyMol 2.5.2 [79].

4.5.3. Protein expression and purification

The full-length PgcA construct (residues 253-511) was kindly provided by Benjamin Nash and Prof. Thomas Clarke (University of East Anglia, Norwich, United Kingdom). The PgcA cytochrome domains were heterologously expressed in *E. coli* C43 (DE3) cells (Domain 1) or *E. coli* BL21 (DE3) cells (Domains 2 and 3). Both strains contained the pEC86 plasmid (encoding the *c*-type cytochrome maturation gene cluster and a chloramphenicol resistance marker [80, 81]) and the vector harboring each cytochrome domain gene (also carrying an ampicillin resistance marker). Cells were grown at 30 °C in 2xYT medium supplemented with ampicillin (100 µg/mL, NZYTech) and chloramphenicol (34 µg/mL, NZYTech) to an OD_{600} of 1.5, at which protein expression was induced with 20 µM of IPTG (NZYTech). The cultures were incubated overnight and then harvested by centrifugation at 6500 x*g* for 15 minutes.

Purification of Domains 1 and 3

Cells expressing Domains 1 and 3 were chemically disrupted with 30 mL per liter of cell culture of a lysis buffer containing 100 mM Tris-HCl pH 8, 0.5 mM EDTA (Amresco), 20% sucrose (Fisher Scientific), 0.5 mg/mL lysozyme (Fisher Scientific) and DNase I (Roche). The periplasmic fractions were recovered by centrifugation at 15,000 x*g* for 20 minutes and dialyzed against 2 × 4.5 L of 10 mM Tris-HCl pH 8. The dialyzed periplasmic fractions of Domains 1 and 3 were loaded onto a 5 mL anion exchange Foresight UNOsphere Q cartridge (Bio-Rad) and a 5 mL cation exchange Bio-Scale Mini UNOsphere S cartridge (Bio-Rad), respectively, both equilibrated with 10 mM Tris-HCl pH 8. Proteins were eluted with a sodium chloride gradient (0-300 mM NaCl, Fisher Scientific) and the resulting fractions were injected onto either an XK 16/70 Superdex 75 column (Cytiva) or a Superdex 75 Increase 10/300 GL column (Cytiva), equilibrated with 100 mM sodium phosphate buffer pH 8.

Purification of Domain 2

Cells expressing Domain 2 were resuspended in 30 mL per liter of cell culture of 50 mM Tris-HCl pH 8.5 with 0.5 mg/mL of lysozyme, DNase I, 1 mM PMSF and 2 mM benzamidine, and disrupted by a combined method of three freeze and thaw cycles followed by 15 cycles of ultrasonication (1 minute on plus 2 minutes off) performed with an ultrasonic homogenizer Branson SFX 150 regulated for 65% of total amplitude. Cell debris were removed by centrifugation at 50,000 x*g* for 1 hour and the resulting supernatant was dialyzed against 2 × 4.5 L of 10 mM Tris-HCl pH 8.5. The dialyzed cell lysate was loaded onto a 5 mL anion exchange Foresight UNOsphere Q cartridge (Bio-Rad), equilibrated with the same buffer, and eluted with a sodium chloride gradient (0-300 mM NaCl). The resulting fractions were dialyzed against 2 x 4.5 L of 20 mM sodium acetate buffer pH 5 and loaded a 5 mL cation exchange Bio-Scale Mini UNOsphere S cartridge (Bio-Rad). Domain 2 eluted in the flow-through, whereas most of the contaminants were eluted after the application of a sodium chloride gradient (0-

300 mM NaCl). Finally, the fractions containing the target protein were injected onto either a XK 16/70 Superdex 75 column (Cytiva) or a Superdex 75 Increase 10/300 GL column (Cytiva), equilibrated with 100 mM sodium phosphate buffer pH 8.

Evaluation of protein purity and quantification

Protein purity was evaluated by BlueSafe (NZYTech) or TMBZ (Thermo Fisher Scientific) stained SDS-PAGE (15% acrylamide/bis-acrylamide). The molecular weights of the different domains were confirmed by MALDI-TOF-MS. The analysis was performed by the Mass Spectrometry Unit (UniMS) of ITQB/iBET (Oeiras, Portugal). The concentrations of the cytochrome domains and the full-length PgcA were determined by measuring the absorbances of the reduced forms at the α-band (552, 553, 550 and 552 nm for Domains 1, 2, 3 and full-length PgcA, respectively), using a molar extinction coefficient of 23.8 mM^{-1} cm^{-1} per heme group [82]. The UV-visible spectra were recorded for both oxidized and reduced samples in the range 300 to 700 nm, at room temperature, using an Evolution 201 spectrophotometer (Thermo Scientific) and quartz cells (Hellma) with 1 cm of path length. The reduced samples were prepared by adding a concentrated solution of sodium dithionite (Fisher Scientific).

4.5.4. CD spectroscopy

CD experiments were performed with an Applied Photophysics Chirascan qCD spectropolarimeter (Leatherhead, UK), equipped with a thermostatic cell support and a 0.2 mm path length quartz cell (Hellma). CD spectra were acquired in the far-UV region using 50 μM of each cytochrome domain prepared in 10 mM sodium phosphate buffer pH 7. The spectra are the average of three scans recorded in 1 nm steps with a scan rate of 3 s/nm, over a 190 to 260 nm wavelength range. The buffer contribution was corrected for all spectra and the temperature was controlled to ± 1 °C. After recording a spectrum at 25 °C, the conformational stability of the cytochrome domains was assessed by performing linear thermal-induced denaturation with a heating rate of 1 °C per minute, from 10 °C to 94 °C. Each spectrum was recorded with a scan rate of 0.4 s/nm. After the temperature ramp, the sample was quickly cooled to 25 °C and then 10 °C, at which two final spectra were registered. A least squares two-state transition model [83, 84] was fit to the variation of the ellipticity at 222 nm and used to determine the midpoint of unfolding transition (T_M) and the enthalpy of unfolding (ΔH) of the different cytochrome domains (see Equation 3.5). The data fittings were performed in Origin Pro 8.5. The BeStSel deconvolution method [31] was used to estimate the secondary structural elements of the three cytochrome domains.

4.5.5. DSC experiments

DSC thermograms were acquired in a TA Nano DSC apparatus using 300 μM samples of each cytochrome domain, prepared in 32 mM sodium phosphate buffer with NaCl (100 mM final ionic strength) at pH 7. The temperature ranged from 10 to 120 °C. The melting temperatures were obtained by fitting a one-peak (two-state) transition model to the thermograms using the NanoAnalyze data analysis software (Version 3.11.0, TA Instruments).

4.5.6. Redox titrations followed by visible spectroscopy

Redox titrations were performed at 25 °C inside an anaerobic glovebox (LABstar, MBraun) with argon circulation and O_2 levels kept under 1 ppm. The visible spectra were recorded in an Evolution 300 spectrophotometer (Thermo Scientific), connected to the anaerobic glovebox interior via fiber optics. The cytochrome samples were prepared with 30 µM protein concentration in 32 mM sodium phosphate buffer with NaCl (100 mM final ionic strength) at pH 7. The solution potentials were measured using a combined Pt/Ag/AgCl electrode (Crison), calibrated with quinhydrone (Merck) saturated solutions at pH 7. The following mixture of redox mediators was added to the solution with a final concentration of 1.7 µM each, as described in the literature [85], to ensure a good equilibrium between the redox centers of the proteins and the working electrode: potassium ferricyanide (Merck, $E^{0'}$ = +430 mV), *p*-benzoquinone (Sigma-Aldrich, $E^{0'}$ = +280 mV), tetramethyl-1,4-phenylenediamine (Acros Organics, $E^{0'}$ = +260 mV), 1,2-naphtoquinone-4-sulphonic acid (Acros Organics, $E^{0'}$ = +215 mV), 1,2-napthoquinone (Acros Organics, $E^{0'}$ = +143 mV), trimethylhydroquinone (Acros Organics, $E^{0'}$ = +115 mV), phenazine methosulfate (Sigma-Aldrich, $E^{0'}$ = +80 mV), phenazine ethosulfate (Sigma-Aldrich, $E^{0'}$ = +55 mV), gallocyanine (Alfa Aesar, $E^{0'}$ = +21 mV), methylene blue (Alfa Aesar, $E^{0'}$ = +11 mV), indigo tetrasulfonate (Sigma-Aldrich, $E^{0'}$ = −30 mV), indigo trisulfonate (Acros Organics, $E^{0'}$ = −70 mV), indigo disulfonate (Sigma-Aldrich, $E^{0'}$ = −120 mV), 2-hidroxy-1,4-naphthoquinone (Alfa Aesar, $E^{0'}$ = −145 mV), anthraquinone-2,6-disulfonate (Acros Organics, $E^{0'}$ = −185 mV), anthraquinone-2-sulfonate (Sigma-Aldrich, $E^{0'}$ = −225 mV), safranine O (Alfa Aesar, $E^{0'}$ = −280 mV), neutral red (Fluka, $E^{0'}$ = −325 mV), benzyl viologen (Sigma-Aldrich, $E^{0'}$ = −345 mV), diquat (Fluka, $E^{0'}$ = −350 mV) and methyl viologen (Acros Organics, $E^{0'}$ = −440 mV). To check for hysteresis and verify reversibility, each redox titration was performed in both oxidative and reductive directions, using sodium dithionite as a reducing agent and potassium ferricyanide as an oxidizing agent. The experiments were performed at least two times and the reduction potentials (relative to SHE) were found to be reproducible within ± 2 mV. The reduced fraction of each cytochrome domain was determined by integrating the area of the α-band (552, 553, and 550 nm for Domains 1, 2, and 3, respectively) above the line connecting the flanking isosbestic points (544 and 563 nm for Domain 1, 543 and 562 nm for Domain 2, and 539 and 556 nm for Domain 3) to subtract the optical contribution of the redox mediators. Each measured potential value was corrected to the SHE and a single electron Nernst equation was fitted to the data using Origin Pro 8.5.

4.5.7. Electron transfer experiments with terminal electron acceptors

Electron transfer between PgcA and various substrates (Fe(III) oxides, Fe(III) citrate, Mn(IV) oxides, and potassium chromate) was assessed by UV-visible spectroscopy measurements at 25 °C, inside the anaerobic glovebox. Fe(III) oxides were prepared following a protocol that mainly produces akageneite (β-FeOOH) minerals [57]. Briefly, a 10 M NaOH (Pronalab) solution was added dropwise to a rapidly stirring 0.4 M solution of $FeCl_3$ (Alfa Aesar) until a pH of 7 was reached. The solution was held at pH 7 and the resulting suspension was decanted, centrifuged at 3000 x*g* for 15 minutes and washed with deionized water. The final suspension was quantified by means of a FerroZine (Sigma-Aldrich)

assay, using Fe(II) ethylenediammonium sulfate tetrahydrate ($FeSO_4.(CH_2)_2(NH_3)_2SO_4.4H_2O$, Merck) as a standard. Prior to measurement, the Fe(III) oxides suspension was digested and reduced overnight in the dark using a 0.5 M HCl (Fluka) solution with 125 µM hydroxylamine (Sigma-Aldrich). The final akageneite solution was prepared with a concentration of 500 µM. Mn(IV) oxides were prepared in the form of birnessite ($Na_xMn_{2-x}(IV)Mn(III)_xO_4$, with x ~ 0.4) minerals using previously established protocols [57]. Briefly, a 30 mM $MnCl_2$ (VWR) solution was slowly added in a 1:1 proportion to a solution of 20 mM $KMnO_4$ (Panreac) prepared in 80 mM NaOH. The resulting mixture was allowed to settle overnight at 4 °C in the dark, after which it was decanted, centrifuged at 3000 x*g* for 20 minutes and washed with deionized water. The final suspension was quantified using inductively coupled plasma atomic emission spectroscopy (ICP-AES). This analysis was performed by the LAQV Analytical Laboratory. The final birnessite solution was prepared with a concentration of 250 µM. The Fe(III) citrate (Sigma-Aldrich) and potassium chromate (Sigma-Aldrich) solutions were prepared with final concentrations of 500 µM and 170 µM, respectively, in 32 mM sodium phosphate buffer with NaCl (100 mM final ionic strength) at pH 7. The protein samples of each cytochrome domain and of full-length PgcA were prepared in degassed 32 mM sodium phosphate buffer with NaCl (100 mM final ionic strength) at pH 7. Before each experiment, the cytochrome samples were reduced with a concentrated solution of sodium dithionite. The excess dithionite was removed by passing the protein solutions through a 5 mL HiTrap Desalting column (Cytiva) and the final protein concentration was adjusted to 5 µM for each cytochrome domain, and to 2 µM for the full-length PgcA samples. Substrates were added to the fully reduced cytochromes in several increments and the consequent redox reactions were monitored by recording UV-visible spectra between 300 and 700 nm using an Evolution 300 spectrophotometer.

4.5.8. Fe(III) oxides binding experiments

The ability of the cytochromes to bind Fe(III) oxides was probed by UV-visible spectroscopy, using a previously established protocol [14]. The spectra were recorded between 300 and 700 nm using an Evolution 201 spectrophotometer (Thermo Scientific). After acquisition of spectra with either 5 µM (cytochrome domains) or 2 µM (full-length PgcA) protein samples prepared in 32 mM sodium phosphate buffer with NaCl (100 mM final ionic strength) at pH 7, each sample was incubated with 55 mM of akageneite for 10 minutes, after which the resulting solutions were centrifuged at 5000 x*g* for 10 minutes. Additional spectra were acquired with the resulting supernatants and the putative binding of the cytochromes to Fe(III) oxides was evaluated based on differences in spectral intensity.

4.5.9. NMR spectroscopy

NMR experiments were acquired in a Bruker Avance III 600 MHz spectrometer equipped with a 5 mM triple resonance cryoprobe (TCI). The residual H_2O signal was used as an internal reference for the calibration of the ^1H chemical shifts relative to DSS at 0 ppm [86], and the ^{13}C chemical shifts were calibrated through indirect referencing [87]. All spectra were processed using TopSpin 4.1.4 (Bruker BioSpin, Karlsruhe, Germany) and analyzed either with the processing software or with NMRFAM-

Sparky [88]. All NMR experimental setups used are part of the Bruker standard pulse sequence library. All spectra were acquired at 25 °C, unless otherwise stated.

Assignment of the NMR signals of the heme substituents and axial ligands

Cytochrome samples for Domains 1, 2, and 3 were prepared with 1.1 mM, 850 µM and 1.1 mM of protein, respectively, in 32 mM sodium phosphate buffer pH 7 (100 mM final ionic strength) in D_2O (99.9%, Eurisotop) after three lyophilization cycles. The pH values of the samples were measured with a glass microelectrode and were not corrected for isotope effects. Protein integrity was confirmed by analysis of the 1D ^{1}H-NMR spectra acquired before and after the lyophilization steps. For sample reduction, the NMR tubes were sealed with a gas-tight serum cap, and the air was flushed out to avoid possible reoxidation of the cytochromes. Samples were reduced directly in the NMR tube with gaseous hydrogen by adding 1 µL of a 70 µM solution of hydrogenase from *D. vulgaris* Hildenborough. The partially oxidized samples, used in EXSY experiments, were obtained by first removing the hydrogen from the reduced sample with argon and then adding controlled amounts of air into the NMR tube with a Hamilton gas-tight syringe. These samples were prepared with final protein concentrations of 50 µM for Domain 1 and 100 µM for Domains 2 and 3.

For the assignment of the heme substituents in the fully reduced state, 2D ^{1}H-COSY, 2D ^{1}H-TOCSY and 2D ^{1}H-NOESY experiments were used. The 2D ^{1}H-TOCSY and 2D ^{1}H-NOESY spectra were acquired with a mixing time of 60 and 80 ms, respectively. For the assignment of the heme substituents in the fully oxidized state, 2D ^{1}H, ^{13}C-HMQC, 2D ^{1}H-NOESY (80 ms) and 2D ^{1}H-TOCSY (45 ms) experiments were acquired. These spectra were acquired at both 15 and 25 °C. Additionally, to evaluate the temperature dependence of the heme methyl signals of the different cytochrome domains, 1D ^{1}H-NMR spectra were recorded at increments of 2 °C, from 10 to 40 °C. Finally, to further confirm the assignment of the heme substituents and axial ligands of the cytochrome domains of PgcA, 2D ^{1}H-EXSY experiments (25 ms) were acquired with samples at an intermediate level of oxidation.

Biomolecular interactions

Molecular interactions between the different cytochrome domains of PgcA were followed by NMR chemical shift perturbation experiments. Samples of the different cytochrome domains were prepared in 8 mM sodium phosphate buffer pH 7 (25 mM final ionic strength) in D_2O. The titrations were performed in the oxidized state up to 1:5 protein ratios, by adding increasing amounts of a concentrated sample of one cytochrome domain to a 50 µM sample of another domain, and vice-versa. After each addition, 1D ^{1}H-NMR spectra were recorded.

Electron transfer reactions between cytochrome domains

Electron transfer reactions between the different cytochrome domains of PgcA were monitored by NMR in both directions for each pair, using a previously described methodology [54]. Samples of each domain were prepared at 50 µM in 8 mM sodium phosphate buffer pH 7 (25 mM final ionic strength) in D_2O and reduced using the procedure described in the previous sections. After reduction, hydrogen was removed from the samples by flushing them with argon. To each reduced

cytochrome domain, an equimolar amount of another oxidized cytochrome domain was added to the NMR tube inside the anaerobic glovebox. 1D ^{1}H-NMR spectra were acquired (i) in the reduced state, (ii) after the addition of the putative redox partner, and (iii) after opening the NMR tube for full oxidation of both proteins by atmospheric O_2. The extent of the electron transfer reactions between each pair of domains was calculated by integrating the ^{1}H NMR heme methyl signals of each cytochrome domain in the intermediate redox state, relative to the fully oxidized state.

Full-length PgcA titration with Fe(III) oxides

Full-length PgcA samples were prepared at 100 µM in 32 mM sodium phosphate buffer pH 7 (100 mM final ionic strength) in D_2O and reduced with a concentrated solution of sodium dithionite. Fe(III) oxides were prepared with a final concentration of 18.75 mM in D_2O. 1D ^{1}H-NMR spectra were acquired (i) in the reduced state, (ii) after 10 µL additions of Fe(III) oxides, and (iii) after opening the NMR tube for full oxidation of the cytochrome by atmospheric O_2. After each addition of Fe(III) oxides, the samples were incubated for 5 minutes and centrifuged at 5000 x*g* for 10 minutes, prior to spectral acquisition. The reduced fractions of each cytochrome domain along the redox reaction were calculated by integrating the ^{1}H NMR heme methyl signals of each cytochrome domain in the intermediate redox state, relative to the fully oxidized state.

4.6. References

[1] X. Liu, L. Shi, J.-D. Gu, Microbial electrocatalysis: Redox mediators responsible for extracellular electron transfer, Biotechnol. Adv., 36 (2018) 1815-1827.

[2] A. Okamoto, K. Saito, K. Inoue, K.H. Nealson, K. Hashimoto, R. Nakamura, Uptake of self-secreted flavins as bound cofactors for extracellular electron transfer in *Geobacter* species, Energy Environ. Sci., 7 (2014) 1357-1361.

[3] D.R. Lovley, Electrically conductive pili: Biological function and potential applications in electronics, Curr. Opin. Electrochem., 4 (2017) 190-198.

[4] T.A. Clarke, Plugging into bacterial nanowires: A comparison of model electrogenic organisms, Curr. Opin. Microbiol., 66 (2022) 56-62.

[5] M.J. Edwards, G.F. White, J.N. Butt, D.J. Richardson, T.A. Clarke, The crystal structure of a biological insulated transmembrane molecular wire, Cell, 181 (2020) 665-673.e610.

[6] C.A. Salgueiro, L. Morgado, M.A. Silva, M.R. Ferreira, T.M. Fernandes, P.C. Portela, From iron to bacterial electroconductive filaments: Exploring cytochrome diversity using *Geobacter* bacteria, Coordin. Chem. Rev., 452 (2022) 214284.

[7] F. Wang, Y. Gu, J.P. O'Brien, S.M. Yi, S.E. Yalcin, V. Srikanth, C. Shen, D. Vu, N.L. Ing, A.I. Hochbaum, E.H. Egelman, N.S. Malvankar, Structure of microbial nanowires reveals stacked hemes that transport electrons over micrometers, Cell, 177 (2019) 361-369.

[8] F. Wang, K. Mustafa, V. Suciu, K. Joshi, C.H. Chan, S. Choi, Z. Su, D. Si, A.I. Hochbaum, E.H. Egelman, D.R. Bond, Cryo-EM structure of an extracellular *Geobacter* OmcE cytochrome filament reveals tetrahaem packing, Nat. Microbiol., 7 (2022) 1291-1300.

[9] F. Wang, C.H. Chan, V. Suciu, K. Mustafa, M. Ammend, D. Si, A.I. Hochbaum, E.H. Egelman, D.R. Bond, Structure of *Geobacter* OmcZ filaments suggests extracellular cytochrome polymers evolved independently multiple times, eLife, 11 (2022) e81551.

[10] Y. Gu, M.J. Guberman-Pfeffer, V. Srikanth, C. Shen, F. Giska, K. Gupta, Y. Londer, F.A. Samatey, V.S. Batista, N.S. Malvankar, Structure of *Geobacter* cytochrome OmcZ identifies mechanism of nanowire assembly and conductivity, Nat. Microbiol., 8 (2023) 284-298.

[11] F.J. Otero, C.H. Chan, D.R. Bond, Identification of different putative outer membrane electron conduits necessary for Fe(III) citrate, Fe(III) oxide, Mn(IV) oxide, or electrode reduction by *Geobacter sulfurreducens*, J. Bacteriol., 200 (2018) e00347.

[12] M. Aklujkar, M.V. Coppi, C. Leang, B.C. Kim, M.A. Chavan, L.A. Perpetua, L. Giloteaux, A. Liu, D.E. Holmes, Proteins involved in electron transfer to Fe(III) and Mn(IV) oxides by *Geobacter sulfurreducens* and *Geobacter uraniireducens*, Microbiology, 159 (2013) 515-535.

[13] Y.H. Ding, K.K. Hixson, M.A. Aklujkar, M.S. Lipton, R.D. Smith, D.R. Lovley, T. Mester, Proteome of *Geobacter sulfurreducens* grown with Fe(III) oxide or Fe(III) citrate as the electron acceptor, Biochim. Biophys. Acta - Proteins Proteom., 1784 (2008) 1935-1941.

[14] L.A. Zacharoff, D.J. Morrone, D.R. Bond, *Geobacter sulfurreducens* extracellular multiheme cytochrome PgcA facilitates respiration to Fe(III) oxides but not electrodes, Front. Microbiol., 8 (2017) 2481.

[15] J.R. Lloyd, E.L. Blunt-Harris, D.R. Lovley, The periplasmic 9.6-kilodalton *c*-type cytochrome of *Geobacter sulfurreducens* is not an electron shuttle to Fe(III), J. Bacteriol., 181 (1999) 7647-7649.
[16] P.L. Tremblay, Z.M. Summers, R.H. Glaven, K.P. Nevin, K. Zengler, C.L. Barrett, Y. Qiu, B.O. Palsson, D.R. Lovley, A *c*-type cytochrome and a transcriptional regulator responsible for enhanced extracellular electron transfer in *Geobacter sulfurreducens* revealed by adaptive evolution, Environ. Microbiol., 13 (2011) 13-23.
[17] B.H. Lower, R.D. Lins, Z. Oestreicher, T.P. Straatsma, M.F. Hochella, L. Shi, S.K. Lower, *In vitro* evolution of a peptide with a hematite binding motif that may constitute a natural metal-oxide binding archetype, Environ. Sci. Technol., 42 (2008) 3821-3827.
[18] S. Choi, C.H. Chan, D.R. Bond, Lack of specificity in *Geobacter* periplasmic electron transfer, J. Bacteriol., 204 (2022) e00322.
[19] M. Grabowicz, T.J. Silhavy, Redefining the essential trafficking pathway for outer membrane lipoproteins, Proc. Natl. Acad. Sci., 114 (2017) 4769-4774.
[20] M.J. Edwards, A. Hall, L. Shi, J.K. Fredrickson, J.M. Zachara, J.N. Butt, D.J. Richardson, T.A. Clarke, The crystal structure of the extracellular 11-heme cytochrome UndA reveals a conserved 10-heme motif and defined binding site for soluble iron chelates, Structure, 20 (2012) 1275-1284.
[21] P.R. Pokkuluri, Y.Y. Londer, X. Yang, N.E. Duke, J. Erickson, V. Orshonsky, G. Johnson, M. Schiffer, Structural characterization of a family of cytochromes c_7 involved in Fe(III) respiration by *Geobacter sulfurreducens*, Biochim. Biophys. Acta, 1797 (2010) 222-232.
[22] L. Morgado, M. Bruix, M. Pessanha, Y.Y. Londer, C.A. Salgueiro, Thermodynamic characterization of a triheme cytochrome family from *Geobacter sulfurreducens* reveals mechanistic and functional diversity, Biophys. J., 99 (2010) 293-301.
[23] L. Morgado, V.B. Paixão, M. Schiffer, P.R. Pokkuluri, M. Bruix, C.A. Salgueiro, Revealing the structural origin of the redox-Bohr effect: The first solution structure of a cytochrome from *Geobacter sulfurreducens*, Biochem. J., 441 (2012) 179-187.
[24] S. Baldi, P. Korber, P.B. Becker, Beads on a string – Nucleosome array arrangements and folding of the chromatin fiber, Nat. Struct. Mol. Biol., 27 (2020) 109-118.
[25] K. Oohora, R. Kajihara, N. Fujimaki, T. Uchihashi, T. Hayashi, A ring-shaped hemoprotein trimer thermodynamically controlled by the supramolecular heme-heme pocket interaction, Chem. Commun., 55 (2019) 1544-1547.
[26] K. Oohora, T. Hayashi, Hemoprotein-based supramolecular assembling systems, Curr. Opin. Chem. Biol., 19 (2014) 154-161.
[27] M.A. Silva, A.P. Fernandes, D.L. Turner, C.A. Salgueiro, A biochemical deconstruction-based strategy to assist the characterization of bacterial electric conductive filaments, Int. J. Mol. Sci., 24 (2023) 7032.
[28] J.M. Dantas, M.A. Silva, D. Pantoja-Uceda, D.L. Turner, M. Bruix, C.A. Salgueiro, Solution structure and dynamics of the outer membrane cytochrome OmcF from *Geobacter sulfurreducens*, Biochim. Biophys. Acta, 1858 (2017) 733-741.

[29] P.R. Pokkuluri, Y.Y. Londer, S.J. Wood, N.E. Duke, L. Morgado, C.A. Salgueiro, M. Schiffer, Outer membrane cytochrome *c*, OmcF, from *Geobacter sulfurreducens*: High structural similarity to an algal cytochrome c_6, Proteins, 74 (2009) 266-270.

[30] G.R. Moore, G.W. Pettigrew, Cytochromes *c*: Evolutionary, structural and physicochemical aspects, Springer-Verlag Heidelberg, Berlin, 1990.

[31] A. Micsonai, É. Moussong, F. Wien, E. Boros, H. Vadászi, N. Murvai, Y.-H. Lee, T. Molnár, M. Réfrégiers, Y. Goto, Á. Tantos, J. Kardos, BeStSel: Webserver for secondary structure and fold prediction for protein CD spectroscopy, Nucleic Acids Res., 50 (2022) W90-W98.

[32] A. Micsonai, É. Bulyáki, J. Kardos, BeStSel: From secondary structure analysis to protein fold prediction by circular dichroism spectroscopy, Structural genomics: General applications, Springer, New York, 2021, 175-189.

[33] W. Kabsch, C. Sander, Dictionary of protein secondary structure: Pattern recognition of hydrogen-bonded and geometrical features, Biopolymers, 22 (1983) 2577-2637.

[34] T.M. Fernandes, L. Morgado, D.L. Turner, C.A. Salgueiro, Protein engineering of electron transfer components from electroactive *Geobacter* bacteria, Antioxidants, 10 (2021) 844.

[35] P. Hosseinzadeh, Y. Lu, Design and fine-tuning redox potentials of metalloproteins involved in electron transfer in bioenergetics, Biochim. Biophys. Acta, 1857 (2016) 557-581.

[36] Y. Takayama, N.D. Werbeck, H. Komori, K. Morita, K. Ozawa, Y. Higuchi, H. Akutsu, Strategic roles of axial histidines in structure formation and redox regulation of tetraheme cytochrome c_3, Biochemistry, 47 (2008) 9405-9415.

[37] R.A. Marcus, N. Sutin, Electron transfers in chemistry and biology, Biochim. Biophys. Acta - Rev. Bioenergetics, 811 (1985) 265-322.

[38] M.P. Williamson, Using chemical shift perturbation to characterise ligand binding, Prog. Nucl. Mag. Res. Sp., 73 (2013) 1-16.

[39] Z. Liu, Z. Gong, X. Dong, C. Tang, Transient protein-protein interactions visualized by solution NMR, Biochim. Biophys. Acta, 1864 (2016) 115-122.

[40] M. Prudêncio, M. Ubbink, Transient complexes of redox proteins: Structural and dynamic details from NMR studies, J. Mol. Recogn., 17 (2004) 524-539.

[41] D.L. Turner, Evaluation of ^{13}C and ^{1}H Fermi contact shifts in horse cytochrome *c*: The origin of the anti-Curie effect, Eur. J. Biochem., 211 (1993) 563-568.

[42] L. Banci, I. Bertini, C. Luchinat, R. Pierattelli, N.V. Shokhirev, F.A. Walker, Analysis of the temperature dependence of the ^{1}H and ^{13}C isotropic shifts of horse heart ferricytochrome *c*: Explanation of Curie and anti-Curie temperature dependence and nonlinear pseudocontact shifts in a common two-level framework, J. Am. Chem. Soc., 120 (1998) 8472-8479.

[43] P. Curie, Magnetic properties of bodies at various temperatures, Ann. Chim. Phys., 5 (1895) 289-405.

[44] C.E. Johnson, F.A. Bovey, Calculation of nuclear magnetic resonance spectra of aromatic hydrocarbons, J. Chem. Phys., 29 (1958) 1012-1014.

[45] G.P. Moss, Nomenclature of tetrapyrroles: Recommendations of the 1986 IUPAC-IUB joint commission on biochemical nomenclature (JCBN), Eur. J. Biochem., 178 (1988) 277-328.
[46] J.M. Dantas, I.H. Saraiva, L. Morgado, M.A. Silva, M. Schiffer, C.A. Salgueiro, R.O. Louro, Orientation of the axial ligands and magnetic properties of the hemes in the cytochrome c_7 family from *Geobacter sulfurreducens* determined by paramagnetic NMR, Dalton Trans., 40 (2011) 12713-12718.
[47] C. Frazão, F.J. Enguita, R. Coelho, G.M. Sheldrick, J.A. Navarro, M. Hervás, M.A. De la Rosa, M.A. Carrondo, Crystal structure of low-potential cytochrome c_{549} from *Synechocystis* sp. PCC 6803 at 1.21 Å resolution, J. Biol. Inorg. Chem., 6 (2001) 324-332.
[48] T.M. Fernandes, L. Morgado, C.A. Salgueiro, D.L. Turner, Determination of the magnetic properties and orientation of the heme axial ligands of PpcA from *Geobacter metallireducens* by paramagnetic NMR, J. Inorg. Biochem., 198 (2019) 110718.
[49] M.R. Ferreira, T.M. Fernandes, D.L. Turner, C.A. Salgueiro, Molecular geometries of the heme axial ligands from the triheme cytochrome PpcF from *Geobacter metallireducens* reveal a conserved heme core architecture, Arch. Biochem. Biophys., 723 (2022) 109220.
[50] C.A. Salgueiro, D.L. Turner, A.V. Xavier, Use of paramagnetic NMR probes for structural analysis in cytochrome c_3 from *Desulfovibrio vulgaris*, Eur. J. Biochem., 244 (1997) 721-734.
[51] J.M. Dantas, D.M. Tomaz, L. Morgado, C.A. Salgueiro, Functional characterization of PccH, a key cytochrome for electron transfer from electrodes to the bacterium *Geobacter sulfurreducens*, FEBS Lett., 587 (2013) 2662-2668.
[52] J.M. Dantas, A. Brausemann, O. Einsle, C.A. Salgueiro, NMR studies of the interaction between inner membrane-associated and periplasmic cytochromes from *Geobacter sulfurreducens*, FEBS Lett., 591 (2017) 1657-1666.
[53] B.M. Fonseca, C.M. Paquete, S.E. Neto, I. Pacheco, C.M. Soares, R.O. Louro, Mind the gap: Cytochrome interactions reveal electron pathways across the periplasm of *Shewanella oneidensis* MR-1, Biochem. J., 449 (2012) 101-108.
[54] L. Morgado, C.A. Salgueiro, Elucidation of complex respiratory chains: A straightforward strategy to monitor electron transfer between cytochromes, Metallomics, 14 (2022) mfac012.
[55] A. Kappler, C. Bryce, M. Mansor, U. Lueder, J.M. Byrne, E.D. Swanner, An evolving view on biogeochemical cycling of iron, Nat. Rev. Microbiol., 19 (2021) 360-374.
[56] K.H. Nealson, D. Saffarini, Iron and manganese: Environmental significance, physiology, and regulation, Annu. Rev. Microbiol., 48 (1994) 311-343.
[57] C.E. Levar, C.L. Hoffman, A.J. Dunshee, B.M. Toner, D.R. Bond, Redox potential as a master variable controlling pathways of metal reduction by *Geobacter sulfurreducens*, ISME J., 11 (2017) 741-752.
[58] J. Majzlan, Minerals and aqueous species of iron and manganese as reactants and products of microbial metal respiration, Microbial metal respiration: From geochemistry to potential applications, Springer, Berlin, Heidelberg, 2012, 1-28.

[59] L. Schacht, M. Ginder-Vogel, Arsenite depletion by manganese oxides: A case study on the limitations of observed first order rate constants, Soil Syst., 2 (2018) 39.
[60] C.M. Stern, D.W. Hayes, L.O. Kgoadi, N. Elgrishi, Emerging investigator series: Carbon electrodes are effective for the detection and reduction of hexavalent chromium in water, Environ. Sci. Water Res. Technol., 6 (2020) 1256-1261.
[61] A.K. Shanker, C. Cervantes, H. Loza-Tavera, S. Avudainayagam, Chromium toxicity in plants, Environ. Int., 31 (2005) 739-753.
[62] E.F. Pettersen, T.D. Goddard, C.C. Huang, E.C. Meng, G.S. Couch, T.I. Croll, J.H. Morris, T.E. Ferrin, UCSF ChimeraX: Structure visualization for researchers, educators, and developers, Protein Sci., 30 (2021) 70-82.
[63] J.E. Post, P.J. Heaney, R.B.V. Dreele, J.C. Hanson, Neutron and temperature-resolved synchrotron X-ray powder diffraction study of akaganeite, Am. Mineral., 88 (2003) 782-788.
[64] J.L. Bishop, E. Murad, M.D. Dyar, Akaganeite and schwertmannite: Spectral properties and geochemical implications of their possible presence on Mars, Am. Mineral., 100 (2015) 738-746.
[65] J.M.A. Antunes, M.A. Silva, C.A. Salgueiro, L. Morgado, Electron flow from the inner membrane towards the cell exterior in *Geobacter sulfurreducens*: Biochemical characterization of cytochrome CbcL, Front. Microbiol., 13 (2022) 898015.
[66] D.R. Lovley, Bug juice: Harvesting electricity with microorganisms, Nat. Rev. Microbiol., 4 (2006) 497-508.
[67] Y. Gu, V. Srikanth, A.I. Salazar-Morales, R. Jain, J.P. O'Brien, S.M. Yi, R.K. Soni, F.A. Samatey, S.E. Yalcin, N.S. Malvankar, Structure of *Geobacter* pili reveals secretory rather than nanowire behaviour, Nature, 597 (2021) 430-434.
[68] D.R. Lovley, On the existence of pilin-based microbial nanowires, Front. Microbiol., 13 (2022) 872610.
[69] F. Wang, L. Craig, X. Liu, C. Rensing, E.H. Egelman, Microbial nanowires: Type IV pili or cytochrome filaments?, Trends Microbiol., 31 (2023) 384-392.
[70] M. Kanehisa, Y. Sato, M. Kawashima, M. Furumichi, M. Tanabe, KEGG as a reference resource for gene and protein annotation, Nucleic Acids Res., 44 (2016) D457-D462.
[71] P.R. Pokkuluri, Y.Y. Londer, N.E. Duke, J. Erickson, M. Pessanha, C.A. Salgueiro, M. Schiffer, Structure of a novel c_7-type three-heme cytochrome domain from a multidomain cytochrome *c* polymer, Protein Sci., 13 (2004) 1684-1692.
[72] Y.Y. Londer, P.R. Pokkuluri, D.M. Tiede, M. Schiffer, Production and preliminary characterization of a recombinant triheme cytochrome c_7 from *Geobacter sulfurreducens* in *Escherichia coli*, Biochim. Biophys. Acta - Bioenergetics, 1554 (2002) 202-211.
[73] S.R. Bond, C.C. Naus, RF-Cloning.org: An online tool for the design of restriction-free cloning projects, Nucleic Acids Res., 40 (2012) 209-213.
[74] S. McGinnis, T.L. Madden, BLAST: At the core of a powerful and diverse set of sequence analysis tools, Nucleic Acids Res., 32 (2004) W20-W25.

[75] J. Jumper, R. Evans, A. Pritzel, T. Green, M. Figurnov, O. Ronneberger, K. Tunyasuvunakool, R. Bates, A. Žídek, A. Potapenko, A. Bridgland, C. Meyer, S.A.A. Kohl, A.J. Ballard, A. Cowie, B. Romera-Paredes, S. Nikolov, R. Jain, J. Adler, T. Back, S. Petersen, D. Reiman, E. Clancy, M. Zielinski, M. Steinegger, M. Pacholska, T. Berghammer, S. Bodenstein, D. Silver, O. Vinyals, A.W. Senior, K. Kavukcuoglu, P. Kohli, D. Hassabis, Highly accurate protein structure prediction with AlphaFold, Nature, 596 (2021) 583-589.

[76] M. Varadi, S. Anyango, M. Deshpande, S. Nair, C. Natassia, G. Yordanova, D. Yuan, O. Stroe, G. Wood, A. Laydon, A. Žídek, T. Green, K. Tunyasuvunakool, S. Petersen, J. Jumper, E. Clancy, R. Green, A. Vora, M. Lutfi, M. Figurnov, A. Cowie, N. Hobbs, P. Kohli, G. Kleywegt, E. Birney, D. Hassabis, S. Velankar, AlphaFold protein structure database: Massively expanding the structural coverage of protein-sequence space with high-accuracy models, Nucleic Acids Res., 50 (2022) D439-D444.

[77] M. Mirdita, K. Schütze, Y. Moriwaki, L. Heo, S. Ovchinnikov, M. Steinegger, ColabFold: Making protein folding accessible to all, Nat. Methods, 19 (2022) 679-682.

[78] V. Mariani, M. Biasini, A. Barbato, T. Schwede, lDDT: A local superposition-free score for comparing protein structures and models using distance difference tests, Bioinformatics, 29 (2013) 2722-2728.

[79] L.L.C. Schrödinger, The PyMol molecular graphics system, Version 2.5.2, 2021.

[80] L. Thöny-Meyer, Biogenesis of respiratory cytochromes in bacteria, Microbiol. Mol. Biol. Rev., 61 (1997) 337-376.

[81] L. Thöny-Meyer, F. Fischer, P. Kunzler, D. Ritz, H. Hennecke, *Escherichia coli* genes required for cytochrome *c* maturation, J. Bacteriol., 177 (1995) 4321-4326.

[82] P. Lukat, M. Hoffmann, O. Einsle, Crystal packing of the c_6-type cytochrome OmcF from *Geobacter sulfurreducens* is mediated by an N-terminal Strep-tag II, Acta Crystallogr. D, 64 (2008) 919-926.

[83] N.J. Greenfield, Analysis of circular dichroism data, Methods in enzymology, Academic Press, 2004, 282-317.

[84] N.J. Greenfield, Using circular dichroism collected as a function of temperature to determine the thermodynamics of protein unfolding and binding interactions, Nat. Protoc., 1 (2006) 2527-2535.

[85] P.L. Dutton, Redox potentiometry: Determination of midpoint potentials of oxidation-reduction components of biological electron-transfer systems, Methods in enzymology, Academic Press, 1978, 411-435.

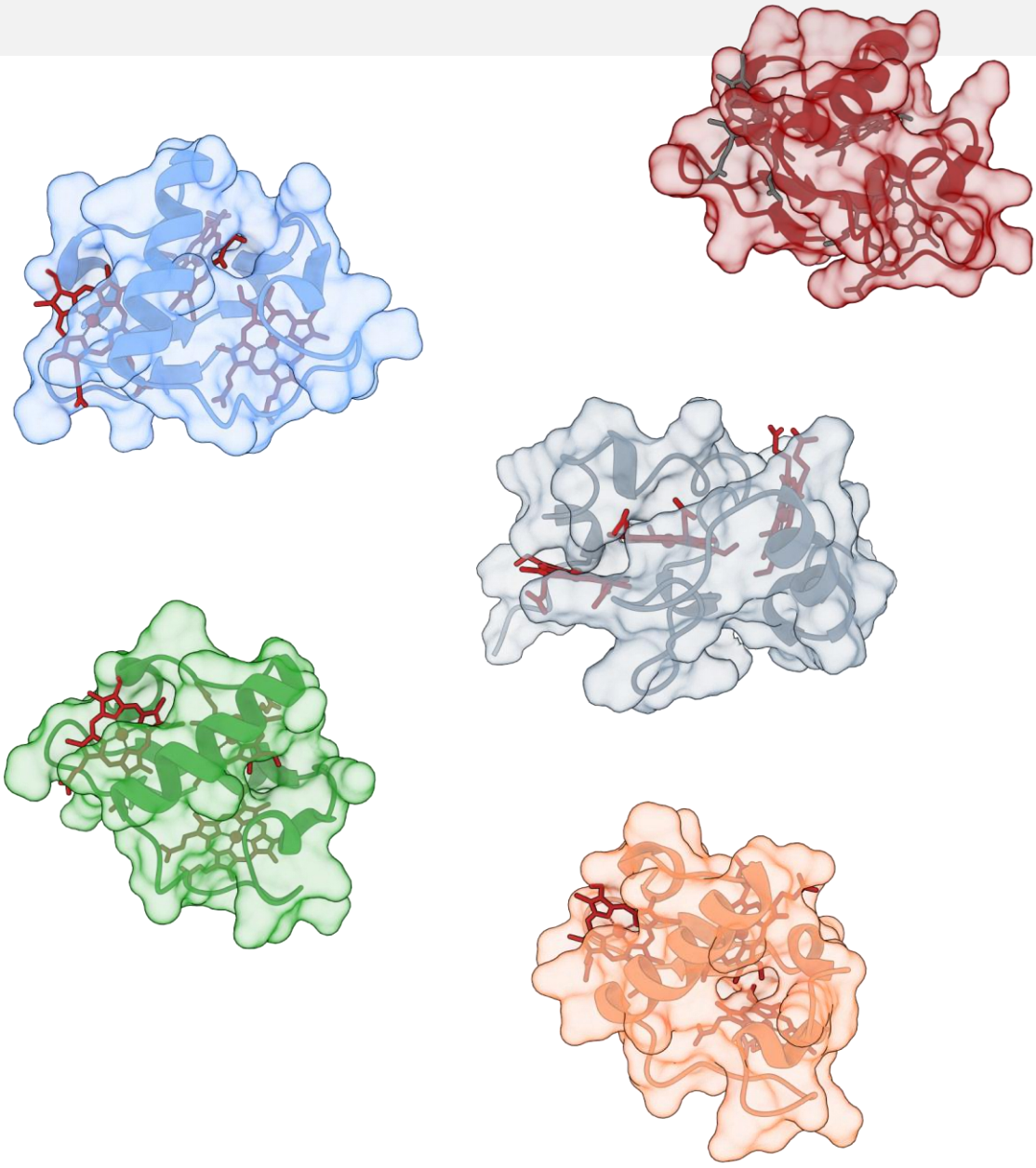
[86] R. Pierattelli, L. Banci, D.L. Turner, Indirect determination of magnetic susceptibility tensors in peroxidases: A novel approach to structure elucidation by NMR, J. Biol. Inorg. Chem., 1 (1996) 320-329.

[87] D.S. Wishart, C.G. Bigam, J. Yao, F. Abildgaard, H.J. Dyson, E. Oldfield, J.L. Markley, B.D. Sykes, ^{1}H, ^{13}C and ^{15}N chemical shift referencing in biomolecular NMR, J. Biomol. NMR, 6 (1995) 135-140.

[88] W. Lee, M. Tonelli, J.L. Markley, NMRFAM-SPARKY: Enhanced software for biomolecular NMR spectroscopy, Bioinformatics, 31 (2015) 1325-1327.

5

Depicting a unique triheme cytochrome: Biophysical studies of GSU0105



This chapter was partially reproduced from **T. M. Fernandes**, F. Folgosa, M. Teixeira, C. A. Salgueiro and L. Morgado, Structural and functional insights of GSU0105, a unique multiheme cytochrome from *G. sulfurreducens*, Biophys. J., 120 (2021) 5395-5407.

List of contents

Depicting a unique triheme cytochrome: Biophysical studies of GSU0105 261

5.1. Summary 263

5.2. Introduction 265

5.3. Results and discussion 266

5.3.1. Amino acid sequence analysis of GSU0105 266

5.3.2. Spectroscopic characterization of GSU0105 268

5.3.3. Redox properties of GSU0105 275

5.4. Conclusions 279

5.5. Materials and methods 281

5.5.1. DNA manipulation 281

5.5.2. Protein sequence analysis and AlphaFold model prediction 281

5.5.3. Protein expression and purification 282

5.5.4. CD spectroscopy 283

5.5.5. EPR spectroscopy 283

5.5.6. NMR spectroscopy 284

5.5.7. Redox titrations followed by visible spectroscopy 284

5.6. References 287

5.1. Summary

G. sulfurreducens possesses over 100 cytochromes that ensure an effective electron transfer to the cell exterior. The most abundant group of cytochromes in this microorganism is the PpcA-family, composed of five periplasmic triheme cytochromes (PpcA-E) with high structural homology and identical heme axial coordination (His-His). GSU0105 is a periplasmic triheme cytochrome produced by *G. sulfurreducens* in Fe(III)-reducing conditions, which is not present in cultures grown on fumarate. Based on the analysis of its amino acid sequence, this cytochrome has a low sequence identity with the PpcA-family cytochromes and a different heme coordination.

In this chapter, amino acid sequence analysis, site-directed mutagenesis, and complementary biophysical techniques, including UV-visible, CD, EPR and NMR spectroscopies, were used to characterize GSU0105. The cytochrome has a low percentage of secondary structural elements, with features of α-helices. NMR shows that the protein contains three low-spin hemes (Fe(II), S = 0) in the reduced state. EPR shows that, in the oxidized state, one of the hemes becomes high-spin (Fe(III), S = 5/2), whereas the two others remain low-spin (Fe(III), S = 1/2). The data obtained also indicate that the heme groups have distinct axial coordination. The apparent midpoint reduction potential of GSU0105 (−154 mV vs SHE) is pH-independent in the physiological range. However, the pH modulates the reduction potential of the heme that undergoes the low to high-spin interconversion. The reduction potential values of cytochrome GSU0105 are more distinct compared to those of the PpcA-family members, providing the protein with a larger functional working redox potential range.

Overall, the results obtained, together with an amino acid sequence analysis of different multiheme cytochrome families, indicate that GSU0105 is a member of a new group of triheme cytochromes.

5.2. Introduction

The five members of the PpcA-family are the most abundant cytochromes of the periplasm of *G. sulfurreducens*, having an important and strategic position in the cell that allows them to function as capacitors and to control the electron flow that connects the inner membrane and outer membrane components of the bacterium [1, 2]. These cytochromes share pairwise sequence identities ranging from 45 to 77%, have a low molecular mass (~10 kDa), ~70 residues each, and their three heme groups are axially coordinated by two histidine residues [3, 4]. The co-existence of these highly homologous cytochromes sharing the same cellular compartment is intriguing, since the distinct respiratory pathways of the bacterium are supposed to be activated by the formation of specific protein-protein complexes. However, while these cytochromes share a high degree of structural homology [3-7], their functional properties are distinct [8], which should imply specific roles for each cytochrome within the cell. Nevertheless, complementary studies in the presence of different electron acceptors have shown that all the members of the PpcA-family are constitutively present in the periplasm of the bacterium [9-12]. Additionally, it has been observed that in mutated bacterial strains lacking four of the homologs, the remaining periplasmic cytochrome can support the reduction of a terminal electron acceptor and contribute to the adaptation of the strains [13]. Surprisingly, a strain in which the genes encoding all five cytochromes were deleted was still able to reduce Fe(III) oxides. This fact was explained by the presence of an unrelated cytochrome, PgcA, as well as other cytochromes, such as OmcE and OmcS, that can contribute to electron transfer in the periplasm [13].

A proteomic analysis on *G. sulfurreducens* cultures grown in different conditions showed that three cytochromes – GSU0105, GSU0701, and GSU2515 – were not synthetized in cultures grown on fumarate, but instead were highly synthetized in Fe(III)-reducing conditions [10]. GSU0701, also known as OmcJ, is a hexaheme cytochrome structurally homologous to OmcS, which likely assembles into a cytochrome nanowire. Additional studies showed that OmcJ is overexpressed in cultures grown on hematite and magnetite [14], while its deletion impaired bacterial growth on insoluble Fe(III) oxides [9]. GSU2515 is a monoheme periplasmic cytochrome overexpressed in current-consuming biofilms [15], which we have recently functionally and structurally characterized [16]. Finally, GSU0105 is a periplasmic cytochrome with a similar molecular mass and the same number of heme groups as the PpcA-family cytochromes. However, a preliminary sequence analysis revealed a low amino acid sequence identity.

In this chapter, a biochemical and biophysical characterization of GSU0105 from *G. sulfurreducens* is presented. The combined usage of several spectroscopic techniques, including UV-visible, CD, EPR, and NMR allowed the elucidation of the structural and functional features of GSU0105's redox centers. The heme's spin-states were determined, and the nature of their axial ligands is discussed based on amino acid sequence analysis and site-directed mutagenesis studies. The working functional redox range of the cytochrome at the bacterium's physiological pH range was determined, showing that GSU0105 has the necessary features to carry electron transfer in the periplasm in a wider range compared to the five PpcA-family cytochromes. Altogether, it is proposed that GSU0105 belongs to a new group of triheme cytochromes.

5.3. Results and discussion

5.3.1. Amino acid sequence analysis of GSU0105

The amino acid sequence of the *c*-type cytochrome GSU0105 contains three typical CXXCH heme binding motifs (Figure 5.1).

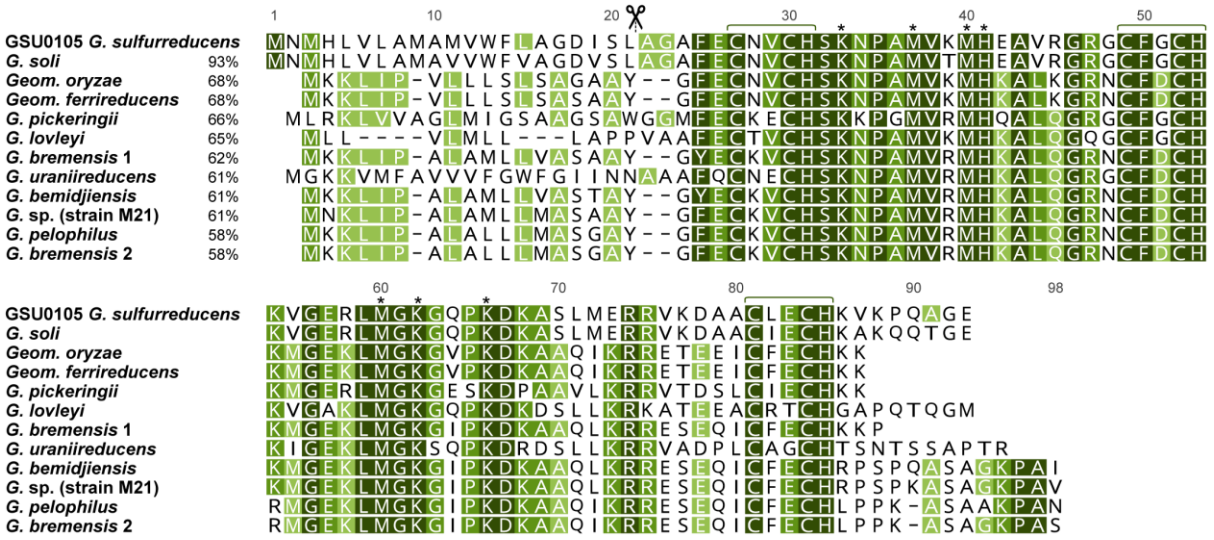


Figure 5.1. Amino acid sequence analysis of GSU0105. Alignment of GSU0105 sequence with homolog sequences from different bacteria obtained from a Protein BLAST analysis. Sequences were aligned with Clustal Omega and their % pairwise identity for mature sequence is indicated. Residues are colored by sequence similarity (dark green 100%, green 80–100%, light green 60–80%, and white <60% similarity). Heme binding motifs CXXCH are identified, and the fully conserved putative axial ligands (Lys[33], Met[37], Met[40], His[41], Met[60], Lys[62], and Lys[66]) are marked with *. The scissors identify the signal peptide cleavage site.

The analysis of the sequence shows that the heme axial coordination of the cytochrome is necessarily different from those belonging to the c_7 family, which all have bis-histidine (bis-His) axial coordinated *c*-type hemes [3]. In fact, the distal ligands in these proteins are fully conserved between all sequences. For GSU0105, the sequence shows that besides the three histidine residues included in the CXXCH binding motifs, there is only one histidine residue left in the mature sequence (His[41], Figure 5.1), which indicates that only one of the heme groups can have bis-His coordination. The distal coordination position of *c*-type hemes can be occupied by methionine, histidine, asparagine, tyrosine, cysteine and lysine residues, the protein's N-terminal amino group, a water molecule, and/or be transiently vacant [17, 18]. A Protein BLAST search retrieves 12 sequences (out of a total of 38, from which the ones from unclassified organisms were excluded), and their alignment shows that, besides His[41], there are only six conserved amino acids that can act as axial ligands to the heme groups: three methionine (Met[37], Met[40] and Met[60]) and three lysine (Lys[33], Lys[62] and Lys[66]) residues (Figure 5.1). It is important to notice that although residues Lys[33] and Met[40] are in close proximity to the other axial ligands in the amino acid sequence (histidine residues in the binding motifs and His[41]), they are still possible axial ligands, considering the relative position of the axial ligands in tetraheme cytochromes from *Desulfovibrio* species [19, 20].

Multiheme cytochromes containing heme groups with mixed axial coordination have been described previously. For example, the *G. sulfurreducens* dodecaheme cytochrome GSU1996 has four triheme domains, in which, in each domain, two hemes are bis-His, whereas the other heme is His-Met coordinated [21]. Also, the triheme cytochrome DsrJ, which belongs to a transmembrane redox complex found in *Desulfovibrio desulfuricans* ATTC 27774 [22] and in the purple sulfur bacterium *Allochromatium vinosum* [23], possesses three hemes with different types of coordination: a bis-His, a His-Met, and an unusual His-Cys coordination. More recently, a tetraheme cytochrome from the annamox bacterium *K. stuttgartiensis* with an unusual contracted heme binding motif was characterized, and complementary spectroscopic techniques showed that the protein has four low-spin hemes: two with bis-His coordination, one with His-Lys and one with His-Cys [24]. However, such heme mixed coordination is unprecedent in small soluble periplasmic cytochromes in which the ratio of amino acid residues to heme groups is low, suggesting that GSU0105 may belong to a different group of cytochromes within the class III of *c*-type cytochromes [25]. This class is formed by multiheme cytochromes that present low reduction potentials (see section 5.3.3) and only around 30 residues per heme group [26]. To confirm this hypothesis, representative members of different families of periplasmic cytochromes [25] were selected for a Protein BLAST analysis and a MSA was used to build a dendogram (Figure 5.2). The dendogram obtained shows six distinct branches composed by: (i) c_7 triheme cytochromes including the PpcA-family from different *Geobacter* strains, together with cytochromes from *Desulfuromonas*, *Geomonas*, and *Geoalkalibacter* strains, and the triheme cytochrome from *Pelobacter carbinolicus* (red); (ii) multiple dodecaheme cytochromes from different *Geobacter* strains (purple); (iii) tetraheme cytochromes from *Shewanella* (dark blue); (iv) c_3 tetraheme cytochromes from *Desulfovibrio*, *Desulfobaculum*, *Desulfocurvus* and *Halodesulfovibrio* species (light blue); (v) triheme cytochromes from *Shewanella* and *Ferrimonas* (orange); and lastly (vi) triheme cytochromes from *Geobacter* and *Geomonas* bacteria, homologs to GSU0105 (green). Contrary to other triheme cytochromes represented in the dendogram, this latter branch is composed of cytochromes with a single histidine residue, besides the ones from the CXXCH motif. This indicates not only that mixed heme axial coordination can be expected for the members of this branch, but also that they represent a new family of periplasmic triheme cytochromes. Nevertheless, no biochemical or biophysical studies have been reported for any of GSU0105's homologs, and further studies are needed to identify additional distinctive features of this family of cytochromes.

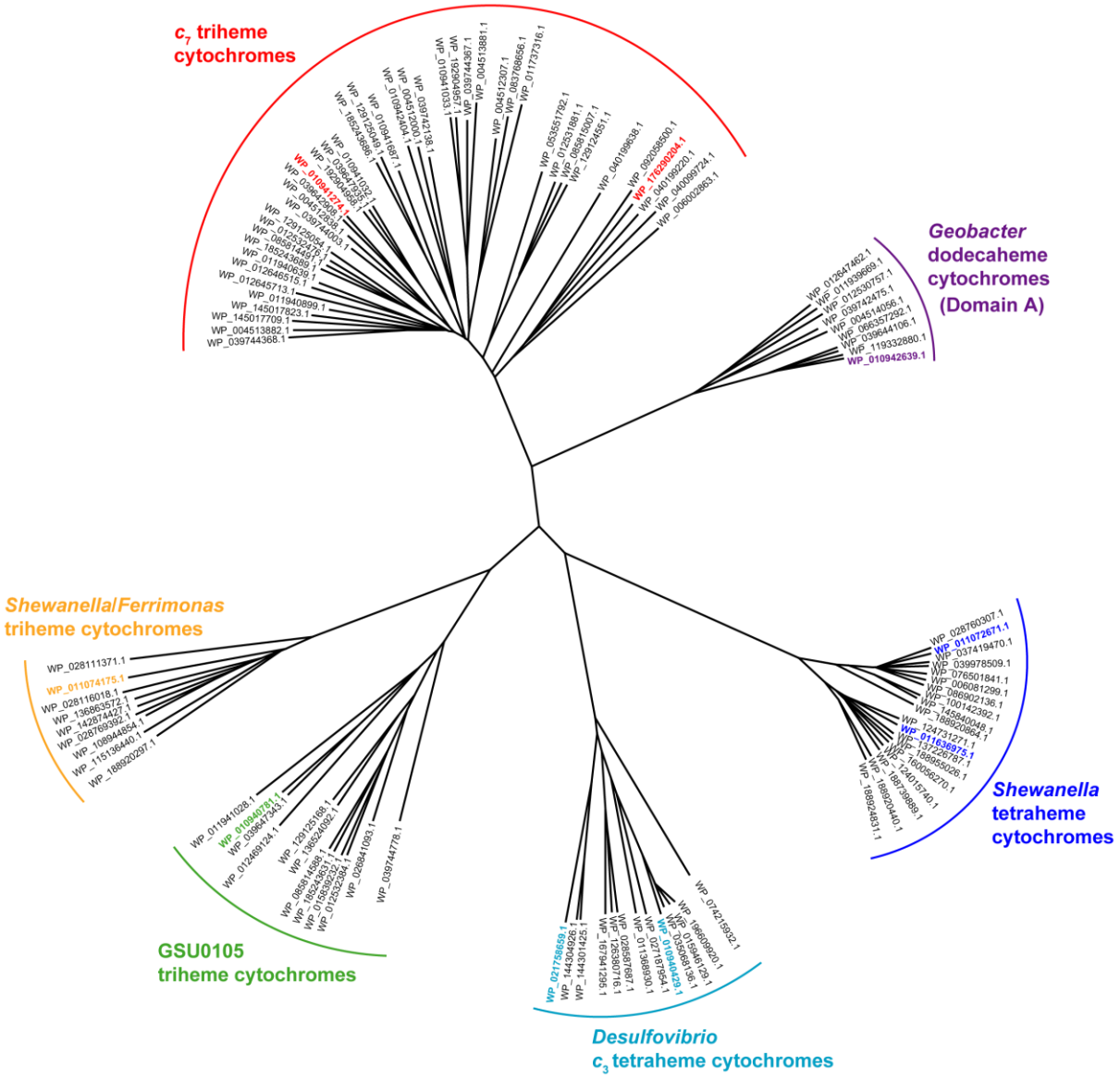


Figure 5.2. Dendogram of representative families of soluble periplasmic multiheme cytochromes. For each family, the proteins used for the Protein BLAST analysis are highlighted with the corresponding color code. Proteins are identified by their NCBI access number.

5.3.2. Spectroscopic characterization of GSU0105

The purified GSU0105 was studied using complementary spectroscopic techniques, including CD, UV-visible, EPR and NMR to probe the protein's folding and secondary structure elements, the spin-state of the hemes and the nature of their axial ligands.

GSU0105 is composed mainly of random coil secondary structure elements

The far-UV CD spectrum of the cytochrome at 25 °C in the oxidized state is presented in Figure 5.3A. The overall spectrum indicates the net presence of random coil structures, but it also contains features that point to the existence of α-helical secondary structural elements, namely i) an intense positive band at 190 nm and ii) an intense negative band at 206 nm, and one less negative band at 223

nm. The positions of these bands, with wavelengths slightly different than those commonly observed in α-helices (190, 208 and 222 nm), probably result from substantial contributions related with the protein's heme groups, as previously observed for other triheme cytochromes [27, 28].

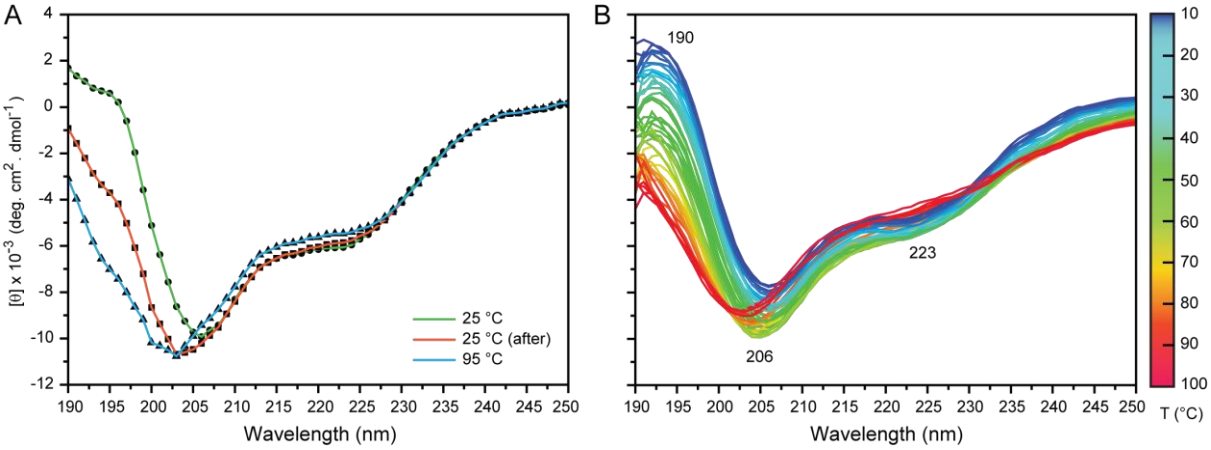


Figure 5.3. Far-UV CD spectral features of oxidized GSU0105. (A) BeStSel fittings (colored lines) of the far-UV spectra of GSU0105 at 25 °C before (spheres) and after (squares) the temperature ramp, and at 95 °C (triangles). (B) Monitorization of the thermal stability of GSU0105. The local maxima and minima of the spectra are labeled. The spectra were acquired between 10 and 95 °C (color gradient from blue to red).

The percentages of secondary structural elements were estimated with the BeStSel deconvolution method [29], which has some limitations, namely the unaccountability of rare secondary structural elements and aromatic contributions. Furthermore, in highly disordered proteins, the software may incorrectly attribute disordered regions to specific structured features. GSU0105 only contains two aromatic residues (Phe^{25} and Phe^{50}), and the results obtained show that it is not highly disordered, thus validating the usage of this method.

The results obtained demonstrated that GSU0105 accounts for 41.3% of folded conformations (29.9% α-helices and 11.4% β-sheets) and 58.7% of disordered conformations. By comparing these data with other triheme cytochromes from *G. sulfurreducens* (Table 5.1), one can verify that these proteins usually possess very low percentages (<50%) of folded conformations. The PpcA-family cytochromes have a conserved general fold, mainly composed of random coil motifs, a two-strand β-sheet in the N-terminal, and several α-helical segments (Figure 5.4). In these proteins, the low amino acid number per heme ratio (between 23 and 25) hinders the possibility of the existence of complex secondary structure elements, most likely to promote solvent exposure and large surface areas that allow the formation of low affinity complexes with other redox partners, necessary for efficient electron transfer [30-32]. GSU0105 presents an even lower percentage of folded conformations, which is another indication that the protein may belong to a different group of triheme cytochromes, as discussed in the previous section.

Table 5.1. Secondary structural elements of GSU0105 and other triheme cytochromes from *G. sulfurreducens*. The percentages of secondary structural elements were either estimated from CD data using the BeStSel deconvolution method [29] or obtained from the DSSP data [33] deposited in the respective PDB identification codes for the remaining triheme cytochromes from *G. sulfurreducens* : 1OS6 (PpcA [4]), 3BXU (PpcB [5]), 3H33 (PpcC [3]), 3H4N (PpcD [3]) and 3H34 (PpcE [3]). The "*" indicates that the spectrum was acquired after the temperature ramp, following a quick cool down.

Method	Secondary structural elements (%)			
BeStSel	**α-helix**	**β-sheet**	**Turn**	**Others**
GSU0105 (25 °C)	29.9	11.4	15.0	43.7
GSU0105 (95 °C)	12.4	16.5	18.4	52.7
GSU0105 (25 °C)*	18.0	15.1	16.8	50.1
DSSP	**α-helix**	**β-sheet**	**Turn**	**Others**
PpcA	42.3	9.9	15.5	32.3
PpcB	38.0	15.5	14.1	32.4
PpcC	30.7	13.3	18.7	37.3
PpcD	38.0	8.5	11.3	42.2
PpcE	36.6	9.9	9.9	43.6

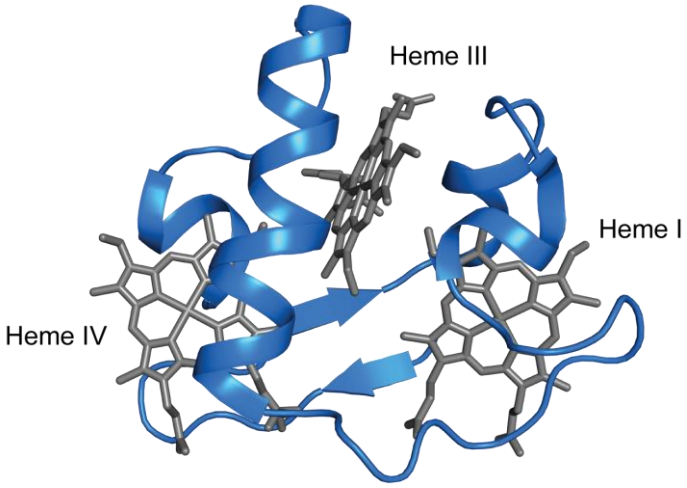


Figure 5.4. Structure of PpcD from *G. sulfurreducens*. The polypeptide chain and the heme groups are represented in blue and gray, respectively. The structure of PpcD (PDB: 3H4N [3]) was selected out of the available PpcA-family cytochromes structures because it possesses the highest sequence identity with GSU0105 (24%).

To monitor the thermal stability of the protein, a characterization of the temperature-induced unfolding was carried out. The spectra obtained show that the secondary structural elements of GSU0105 undergo significant changes during the temperature variation (Figure 5.3B).

In fact, by plotting the variation of the ellipticity at 206 and 223 nm as a function of temperature (Figure 5.5), a complex behavior is observed, which cannot be explained by considering a two-state transition of a monomer from a folded to unfolded state.

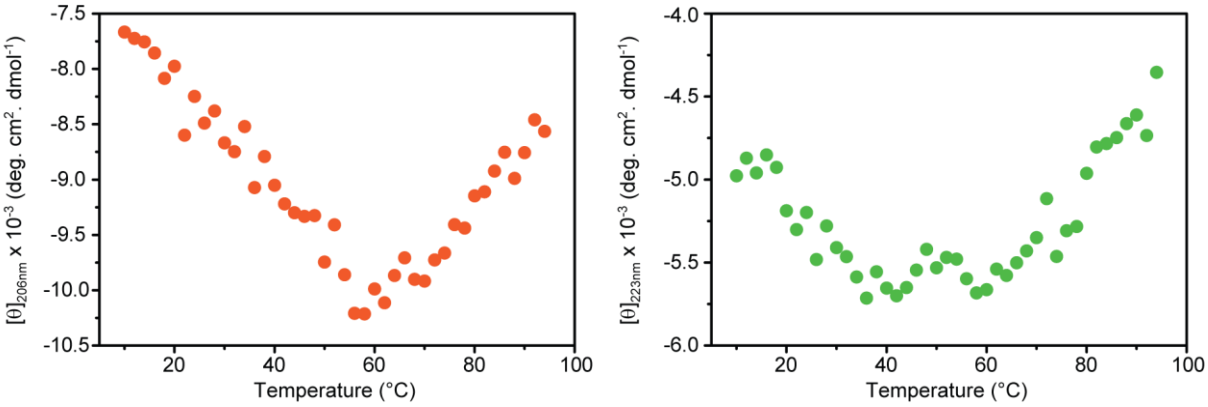


Figure 5.5. Temperature-induced molar ellipticity variation. The mean residue ellipticities [θ] at 206 nm (orange spheres) and 223 nm (green spheres) are represented as a function of temperature.

This means that GSU0105 goes through several rearrangements of its secondary structural elements during the denaturation process, which can include multiple steps and intermediary states linking the native state to the final thermally unfolded state. Similar results have been observed in a monoheme ferricytochrome *c*, in which the protein undergoes a series of reversible conformational switches, resulting in alterations of the distal ligand position of the heme [34]. The main variations of the spectra throughout the temperature variation are observed in the signals corresponding to the α-helices, which show a blueshift in their wavelengths. This is a result of an increase in the percentage of random coil motifs in the protein. Moreover, other slight changes may be due to small alterations of the conformation of the heme groups. At 95 °C, GSU0105 still possesses marked ellipticities and is not fully unfolded (Figure 5.3 and Table 5.1).

Upon cooling to 25 °C, the final CD spectrum reveals that GSU0105 partially refolds (Figure 5.3 and Table 5.1). Altogether, the data indicates that GSU0105 is stable at high temperatures, as observed for other cytochromes from *Geobacter* [27, 28]. In these cytochromes, the three covalently bound *c*-type hemes keep the overall proteins' structural integrity, whereas the few secondary structure elements undergo small variations.

UV-visible spectra of GSU0105 present features of low-spin hemes

The optical absorption spectrum of the oxidized cytochrome (Figure 5.6) has maxima at 354, 408, and 527 nm. Upon reduction, the protein shows the Soret, β- and α-bands maxima at 417, 523, and 552 nm, respectively. These spectral patterns are similar to those shown by low-spin hexacoordinated hemes [35].

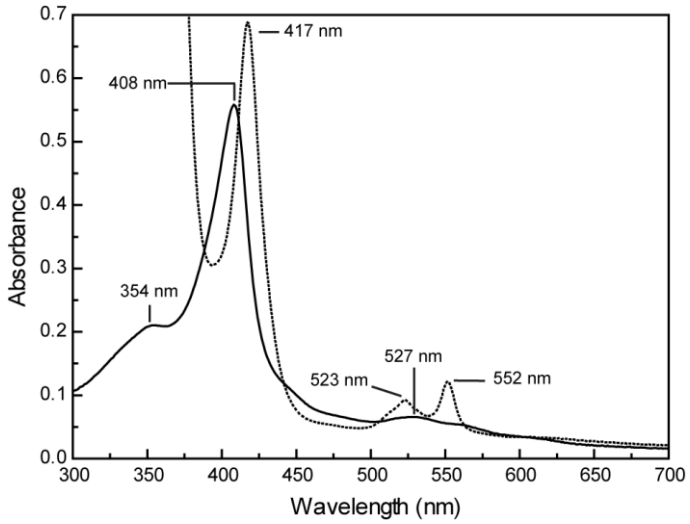


Figure 5.6. UV-visible spectral features of cytochrome GSU0105 in the oxidized and reduced states. The maxima of the UV-visible spectra of the cytochrome in the oxidized (solid black line) and reduced states (dashed black line) are labeled.

Methionine-coordinated hemes typically display a band at 695 nm in the UV-visible spectrum of the oxidized form, which commonly features a half width of 30-40 nm and an extinction coefficient around 20 times smaller than the α-band of low-spin hemes in the reduced state [36-38]. This band is not present in the oxidized spectrum of GSU0105 (Figure 5.6), even when difference spectra are measured. This is a common practice because the extinction coefficient associated with this band may be even lower than usual when the hemes have a high solvent exposure [39]. In these cases, the measurement of difference spectra in cytochromes allows the deconvolution of low-intensity Q and charge-transfer bands, as most of these have a significant variation of their signal intensity with the oxidation state of the heme iron [40].

The absence of a band at 695 nm in the UV-visible spectrum does not exclude the possibility that one of the hemes of GSU0105 is coordinated by a methionine residue, as proven in different works [41, 42]. To gather more information on this and on the hemes spin-states, EPR and NMR experiments were performed for GSU0105.

EPR and NMR features of GSU0105

X-band EPR spectroscopy of the as-purified oxidized GSU0105 exhibits a complex spectrum, dominated by two main sets of features that arise from the contribution of high- and low-spin ferric heme species (Figure 5.7A). A very small contribution of adventitious Fe(III) is also observed at g = 4.23. The simulation of the EPR spectrum indicates an approximate 1:2 ratio of the high- and low-spin species, respectively. To simulate the high-spin species, a low value of the rhombicity was used (E/D = 0.005), and because the zero-field splitting D is unknown, a value of 20 cm^{-1} was assumed so that at the temperature used, only the ground Kramers doublet, $|\pm 1/2\rangle$, is populated. In the fully reduced state, the cytochrome is EPR silent (Figure 5.7A).

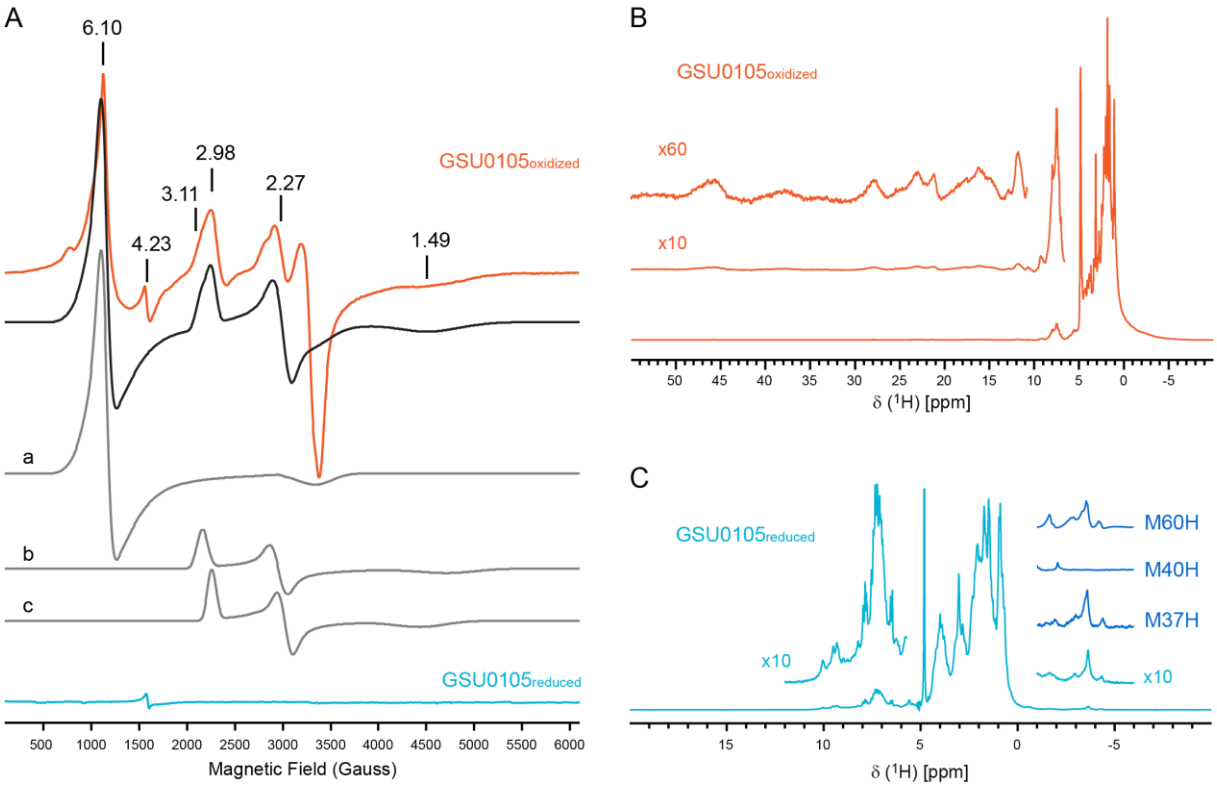


Figure 5.7. EPR and NMR features of GSU0105. (A) Experimental continuous-wave X-band EPR spectra of GSU0105 at the oxidized (orange) and reduced (light blue) states, recorded at 10 K, with a microwave frequency of 9.39 GHz, a modulation amplitude of 1.0 mT and a microwave power of 2 mW. The black line below the oxidized experimental spectrum is the simulated spectrum, accounting for a 1:1:1 contribution of the high-spin species (a) and the two low-spin species (b and c), for which each spectrum is represented in gray. The high-spin species spectrum (a) was simulated with E/D = 0.005 and a D = 20 cm^{-1}. The two low-spin species spectra were simulated with g-values of 3.11, 2.27 and ≤ 1.40 (b), and 2.98, 2.22, and 1.49 (c). The signal with a g-value of 4.23 corresponds to a small amount of high-spin ferric iron adventitiously present in the sample. (B) 1D ^{1}H-NMR spectrum of oxidized GSU0105. The low-field region is zoomed in to show the resonances of the heme methyl substituents. (C) 1D ^{1}H-NMR spectrum of reduced GSU0105. A close-in of the spectrum (x10) is represented. The typical high-field region of axial methionine resonances is shown for the WT (light blue) and mutant (dark blue) proteins.

The intense and quasi-axial signal with g_{max} of 6.10 can be attributed to a high-spin S = 5/2 ferric heme. The g_z contribution of this signal, expected at g = 2.00, is underneath the resonance attributed to adventitious copper (II). The existence of this high-spin species in the oxidized form is consistent with what is observed in the NMR spectrum (Figure 5.7B). The species at a higher magnetic field are characteristic of low-spin ferric hemes. The shoulder observed for the resonance at g ~2.98 indicates that two very similar species contribute to this set of resonances. It was not possible to deconvolute the spectra by varying either the microwave power or the temperature, suggesting that both hemes have quasi-identical relaxation properties. An analysis of the data revealed two species, with g-values at i) 3.11, 2.27, and ≤ 1.40, and ii) 2.98, 2.22, and 1.49.

Similarly to EPR, NMR is a very powerful technique to identify the spin-state of heme groups in cytochromes because their signals appear in quite distinct spectral regions, depending if the hemes are high- or low-spin. Recalling from the previous chapter, in the paramagnetic oxidized state, 1D ^{1}H-NMR spectra of high-spin cytochromes display extremely broad signals and resonances above 40 ppm (usually from the heme methyl substituents). Low-spin cytochromes, on the other end, present narrower spectral windows, with the main heme substituents frequencies ranging from 8 to 35 ppm. In the reduced state, 1D ^{1}H-NMR spectra are also quite distinct for high- and low-spin cytochromes. In fact, high-spin hemes present wider spectral regions (from −15 up to 30 ppm) than low-spin ones (from −5 up to 10 ppm).

The 1D ^{1}H-NMR signals of GSU0105 in the oxidized state are very broad and cover a wide spectral region, from −10 ppm to above 45 ppm (Figure 5.7B). Therefore, the cytochrome is paramagnetic in the oxidized state, with at least one high-spin heme (Fe(III), S = 5/2), as seen in the EPR experiments. The NMR signals of the cytochrome in the oxidized state are very broad because of the strong paramagnetic contribution of the high- and low-spin heme(s), which makes it very difficult to analyze any other spectral features of the protein in this redox state.

In the reduced state, the 1D ^{1}H-NMR spectrum of GSU0105 displays a narrow spectral window between −5 and 11 ppm (Figure 5.7C), indicating that all the hemes from GSU0105 are low-spin (Fe(II), S = 0). Furthermore, the low-frequency region of the reduced state 1D ^{1}H-NMR spectrum can be used as a fingerprint to detect axial ligands because of the heme ring-current effects [35]. Histidine ligands do not show resonances in this region (their signals are usually located in the protein's signal envelope), but methionine ligands display characteristic patterns [43]. The GSU0105 spectrum in the reduced state displays a typical signal pattern for the side-chain signals of a distal coordinated methionine, which includes a three-proton intensity peak at approximately −3 ppm and up to four resolved one-proton intensity peaks in this region of the spectrum (Figure 5.7C). At this stage, the possibility of the existence of two hemes with His-Met axial coordination cannot be discarded based on the analysis of the diamagnetic 1D ^{1}H-NMR spectrum of GSU0105, because their signals may be overlapped.

To identify the axial ligands of GSU0105, three mutants of the fully conserved methionine residues were produced (M37H, M40H, and M60H) and analyzed by 1D ^{1}H-NMR spectroscopy in both oxidation states. As observed for the WT GSU0105, the oxidized spectra of all mutants also display a large spectral width and broad low-field resonances typical for high-spin heme substituents. In the reduced state, the high-field axial methionine resonance pattern observed for the WT cytochrome is observed for the M37H and M60H mutants but not for the M40H mutant (Figure 5.7C). These results unequivocally identify Met^{40} as one of the distal ligands of GSU0105. Moreover, the absence of any axial methionine resonance patterns in the 1D ^{1}H-NMR spectrum of the M40H mutant indicates that the axial ligand of the third heme is not a methionine residue.

The candidates for the unidentified distal ligand are the three fully conserved lysine residues (Lys^{33}, Lys^{62}, and Lys^{66}). However, and in contrast with axial methionine residues, the NMR signal pattern for an axial lysine residue in the 1D ^{1}H-NMR spectrum is not straightforwardly identifiable [44, 45]. This

feature hinders the assignment of the third heme distal ligand of GSU0105 using the same strategy, based on site-directed mutagenesis and NMR spectroscopy. Therefore, a different approach would have to be used to identify this ligand. The best methodology would be structure determination by X-ray crystallography or NMR. However, the low protein yield obtained for this cytochrome (0.2 mg per liter of cell culture) currently hampers this approach. Crystallization trials demand high concentrations of protein, and NMR requires isotopically labeled proteins synthetized in minimal media, which would result in even lower yields.

Furthermore, there are other features of GSU0105 that hinder its structural characterization in more detail by NMR. Contrary to the triheme cytochromes from the PpcA-family of *G. sulfurreducens* [46, 47], the spectrum of GSU0105 in the reduced state has signals with unexpectedly large linewidths for a ~10 kDa protein (Figure 5.7C). These spectral features may arise from inherent properties of GSU0105. Taking a closer look into the backbone amide region (7–11 ppm) of the spectrum (Figure 5.7C), the signals are well dispersed, which is typical of a folded protein, meaning that the protein is not aggregated in the experimental conditions used. A possibility is the existence of exchange processes. If most of the protein signals are in a slow to intermediate exchange rate between two conformations, their signal widths would be significantly affected. This phenomenon would not affect the results of the remaining experiments, and its effects may only be detectable by NMR. There are several experimental conditions that can be varied to optimize the exchange rates of a certain chemical event, namely the magnetic field strength, temperature, and ionic strength [48]. 1D ^{1}H-NMR spectra of GSU0105 in the reduced state were acquired at different temperatures and ionic strength conditions, but all spectra kept the signal broadening initially observed.

5.3.3. Redox properties of GSU0105

Redox titrations of GSU0105 followed by visible spectroscopy were performed at pH 7 and 8 (Figure 5.8) to cover the physiological pH range for *G. sulfurreducens* growth [49]. The reductive and oxidative curves are superimposable (no hysteresis), indicating that the redox process is fully reversible. The redox window of GSU0105 spans from −315 to +85 mV, which is compatible with the presence of bis-His, His-Lys and His-Met coordinated hemes. Typically, *c*-type heme groups axially coordinated by a methionine present more positive reduction potentials compared to bis-His coordinated hemes [37]. Recalling from past chapters, this relates to the fact that the methionine's side-chain sulfur is a good electron acceptor, which favors the electron-rich reduced state of the heme, resulting in more positive reduction potentials compared to bis-histidynil coordinated heme groups. However, experimental data show that in many cases, the network of residues in the vicinity of the heme groups has a higher effect on their reduction potential values than the nature of the axial ligand. For example, there are reports of His-Met coordinated hemes with reduction potential values ranging from −251 to +358 mV [50-52]. Similarly, although there are reports for His-Lys coordinated hemes, their reduction potential values can be as different as −80 or +365 mV [44, 45].

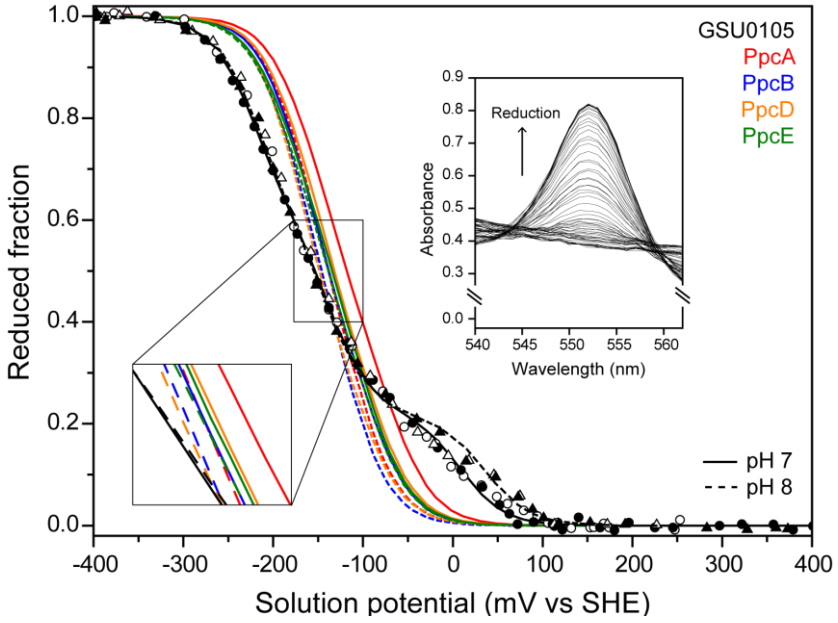


Figure 5.8. Redox titrations of GSU0105 and PpcA-family cytochromes monitored by visible spectroscopy. The solid (pH 7) and dashed (pH 8) lines indicate the results of the fits to the Nernst curves for the three macroscopic reduction potentials of GSU0105 (in black), PpcA (in red), PpcB (in blue), PpcD (in orange), and PpcE (in green), which are indicated in Table 5.2. The open and filled symbols in the GSU0105 curves represent each experimental point of the reductive and oxidative titrations at pH 7 (spheres) and 8 (triangles), respectively. The bottom left inset shows a close-in on the apparent midpoint reduction potentials of the cytochromes. For simplification, in this inset, the experimental points are not represented. The top right inset shows the variation of the α-band region of the visible spectra of GSU0105 throughout the visible redox titration at pH 7.

Macroscopic redox behaviour of GSU0105

For other triheme cytochromes, the fitting of experimental data from redox titrations followed by visible spectroscopy can be obtained with a sequential model (see section 5.5.7), considering that each heme has the same contribution (0.3(3)) for the α-band absorbance. In the case of GSU0105, the contribution for the α-band absorbance of the last heme to oxidize (the one with the most positive reduction potential) is smaller compared to the other hemes. For this reason, the experimental data were fitted by considering different contributions for the three hemes (see section 5.5.7). An equal contribution of 0.39 ± 0.02 for the α-band absorbance was determined for the first two hemes to oxidize, whereas a contribution of 0.22 ± 0.01 was determined for the last heme to oxidize. The differences in these values are most likely related with the different types of axial coordination of the hemes or with the fact that one of them is high-spin in the oxidized state.

Looking at the macroscopic reduction potentials for the three oxidation steps (Table 5.2), GSU0105 has negative but distinct reduction potential values for two of the oxidation steps and a positive value for the other one.

Table 5.2. Comparison of the apparent midpoint and macroscopic reduction potentials (relative to SHE) of GSU0105 and the PpcA-family cytochromes from *G. sulfurreducens* at pH 7 and 8. The apparent midpoint reduction potentials (E_{app}) correspond to the point at which the oxidized and reduced fractions are equal. E_1^0, E_2^0, and E_3^0 are the macroscopic reduction potentials for the first, second, and third oxidation steps, respectively. The values presented have an experimental estimated error of ± 5 mV.

	E_{app} (mV)		E_1^0 (mV)		E_2^0 (mV)		E_3^0 (mV)	
	pH 7	pH 8	pH 7	pH 8	pH 7	pH 8	pH 7	pH 8
GSU0105	−154	−154	−223	−219	−137	−138	8	35
PpcA [8]	−117	−138	−171	−182	−119	−139	−60	−93
PpcB [8]	−137	−143	−185	−192	−140	−145	−84	−103
PpcC [53]	−143	−148	−190	−193	−142	−148	−100	−103
PpcD [8]	−132	−148	−181	−191	−133	−152	−78	−94
PpcE [8]	−134	−139	−191	−194	−133	−138	−82	−85

Although these values are macroscopic in nature, and thus do not correspond to the individual heme reduction potentials, those of the first two oxidation steps are significantly separated from that of the last oxidation step, which can be considered to have a nearly independent oxidation (Figure 5.8). This indicates that the reduction potential of the last oxidation step can be attributed to the last heme to be oxidized. The last heme to be oxidized has a smaller contribution to the α-band because of the low- to high-spin interconversion and concomitant change in its molar extinction coefficient. Therefore, the last oxidation step is dominated by the oxidation of a heme that becomes high-spin in the oxidized state. Considering the observations made in section 5.3.2, this heme has a putative His-Lys axial coordination. Thus, the first and second oxidation steps are dominated by the oxidation of the His-Met[40] and His-His[41] coordinated hemes. Nevertheless, and despite the different coordination, it is not possible to infer which of the hemes has the most negative reduction potential. Interestingly, in a triheme domain from the dodecaheme protein GSU1996, which possesses two bis-His and one His-Met coordinated hemes, the reduction potential of the His-Met coordinated heme is not the most positive [54].

The individual heme reduction potentials have been determined for several low-spin triheme and tetraheme cytochromes by combining 2D ^{1}H-EXSY NMR data and potentiometric redox titrations followed by visible spectroscopy [8, 55]. However, in the case of the triheme cytochrome GSU0105, such studies cannot be implemented because of the severe broadening of the NMR signals caused by the high-spin heme, which impairs the assignment of the individual heme oxidation patterns along the different oxidation stages.

GSU0105 possesses a local redox-Bohr effect

The redox-Bohr effect reflects on the pH dependence of the protein's reduction potential value and is caused by the protonation/deprotonation of an acid-base group (redox-Bohr center) in the vicinity

of the heme(s). Considering pure electrostatics, it is expected that the deprotonation of the redox-Bohr center (removal of a positive charge) would stabilize the oxidized form of the heme(s), leading to the concomitant decrease of the reduction potential value. In the case of GSU0105, the apparent reduction potential value (E_{app}) obtained for GSU0105 was –154 mV at pH 7 and 8 (Table 5.2). In fact, the redox titration curves are superimposable in this region, indicating that there is not a net redox-Bohr effect in the physiological pH range. Nonetheless, and regardless of the similarity of the E_{app} values at pH 7 and 8, from the analysis of the redox titration curves there is a notorious separation at lower reduction potential fractions in the region dominated by the oxidation of the heme with the highest reduction potential (Figure 5.8). This result indicates that there is a local redox-Bohr effect near the last heme to oxidize. However, contrary to the expected electrostatic behavior of the redox-Bohr effect, the reduction potential of the last heme to oxidize is higher at pH 8 (Table 5.2). This effect has been observed in other multiheme cytochromes [56-58] and attributed to mechano-chemical couplings (such as movement of charges, rearrangement of hydrogen bond networks, and other phenomenon connected to acid-base transitions and structure modifications) that mask the electrostatic cooperativity between the opposite charges of the hemes and the deprotonated redox-Bohr center(s).

Considering that the heme with the redox-Bohr effect is the one hypothesized as being high-spin in the oxidized state, there may be a physiological relevance for this effect. The pH-dependent stabilization of the reduced form of this heme (low-spin), compared to the oxidized one (high-spin), suggests that the pH might modulate the heme spin-state interconversion. At a pH below the pK_a of the redox-Bohr center, the high-spin configuration (oxidized state) will be stabilized. The opposite effect is attained at pH values above the pK_a of the redox-Bohr center. Similar pH-dependent spin-state interconversion mechanisms have been observed in other cytochromes [59-62], however, further experiments are needed to pinpoint the physiological relevance of this pH-linked feature of GSU0105.

The redox-Bohr effect is particularly relevant in the PpcA-family cytochromes (see inset of Figure 5.8 and Table 5.2), such as PpcA (21 mV) and PpcD (16 mV), which were proposed to couple electron and proton transfer at the bacterium's physiological pH range [63]. In fact, in these proteins, it was observed that the deprotonation of the redox-Bohr center (removal of a H^+) facilitates the concomitant oxidation (e^- transfer) of the nearest heme [8]. On the other hand, PpcB and PpcE have negligible pH modulation of their heme reduction potential values (6 and 5 mV, respectively) and were suggested to exclusively perform electron transfer at slightly different redox windows. In the case of GSU0105, the redox-Bohr effect does not affect the E_{app}, but instead, it is a particular feature of the highest redox potential heme. Consequently, the pH dependence of the redox potential value of this heme allows the protein to reduce or expand its working functional range (Table 5.2) with the concomitant regulation of the low- to high-spin interconversion of the heme group.

Finally, it is worth noting that the redox curves of the PpcA-family from *G. sulfurreducens* are much steeper compared to those of GSU0105, due to their more similar heme reduction potential values (Figure 5.8 and Table 5.2). Considering the differentiated redox window of GSU0105 compared to the

cytochromes from the PpcA-family, it seems this cytochrome has the necessary properties to bridge electron transfer between a wider range of electron donors and acceptors in the periplasm of the *G. sulfurreducens* bacterium.

5.4. Conclusions

The outgoing interest in electrochemically active microorganisms and their potential for several fields of biotechnology is prompting the investigation and exploration of their unique metabolism and molecular components. *G. sulfurreducens* is in the forefront of such investigations because of its high current production in MFC, metabolic versatility, and ability to produce protein nanowires. The intricacies of this bacterium's EET routes are slowly being dissected; however, it is still unclear which are the main electron transfer pathways and how they relate with different final electron acceptors.

In this chapter, we presented a biophysical characterization of GSU0105, a periplasmic triheme cytochrome involved in Fe(III) respiration. The amino acid sequence comparison with other triheme cytochromes leads to the proposal that GSU0105 does not belong to the c_7 family of triheme cytochromes but, instead, is a member of a group of homologous cytochromes with mixed heme axial coordination. GSU0105 possesses a low percentage of folded structural elements, three low-spin hemes in the reduced state (Fe(II), S = 0), and a mixture of low- (Fe(III), S = 1/2) and high-spin (Fe(III), S = 5/2) hemes in the oxidized state, as confirmed by CD, UV-visible, EPR, and NMR spectroscopic techniques. The amino acid sequence analysis indicates that one of the hemes in GSU0105 contains a bis-His coordination, using the conserved His^{41} as the sixth ligand (Figure 5.1), whereas the NMR experiments performed on GSU0105's mutants indicate that one of the other heme groups contains a His-Met coordination (Figures 5.1 and 5.7). The third heme is likely coordinated by one of the three fully conserved lysine residues (Lys^{33}, Lys^{62}, and Lys^{66}). The data also indicate that this heme turns into a high-spin electronic configuration in the oxidized state, in which the distal position of the heme may become vacant or, alternatively, coordinated by a weak-field ligand (such as a water molecule, for example).

Finally, the effect of the pH on the redox properties of the protein (redox-Bohr effect) revealed an unprecedented mechanism for a triheme cytochrome. In fact, in the physiological region, the E_{app} value is unaffected by the pH, in opposition to that of the heme with the highest reduction potential. Interestingly, this is the heme that undergoes a low- to high-spin interconversion, which indicates that this interconversion is pH-driven. Moreover, this effect also has an impact on the modulation of the working functional redox range of the protein. Compared to the other triheme periplasmic cytochromes from *G. sulfurreducens*, the working redox functional window of GSU0105 is wider and covers more positive regions, adding further versatility to the periplasmic components of the bacterium.

AlphaFold [64] emerged after the publication of this work, however, we later constructed several models of GSU0105. As for the AlphaFold models of OmaB/C, ExtD and ExtG (see Chapter 2), there is a considerable variability in the predicted models of GSU0105, with consistently low average pLDDT scores. Consequently, when we enrich these models with *c*-type hemes in silico, to gather insights on

the putative heme axial ligands and the relative arrangement of the heme core, in most cases there are several clashes and it is hard to discern which are the residues responsible for heme binding (data not shown). Nevertheless, the model constructed using the PDB100 template database and an Amber structure minimization shows no clashes between the hemes and places Met^{40} and His^{41} as the distal ligands of heme III and II, respectively (Figure 5.9). Out of the three fully conserved lysine residues, Lys^{62} is the only one that is in a position to possibly act as the distal ligand of heme I. This residue is located on a loop that might be sufficiently flexible to allow the dynamic attachment and detachment of this distal ligand, through a process similar to the one observed for the Lys-ligated alkaline form of the ferricytochrome *c* from *S. cerevisiae* [65-67].

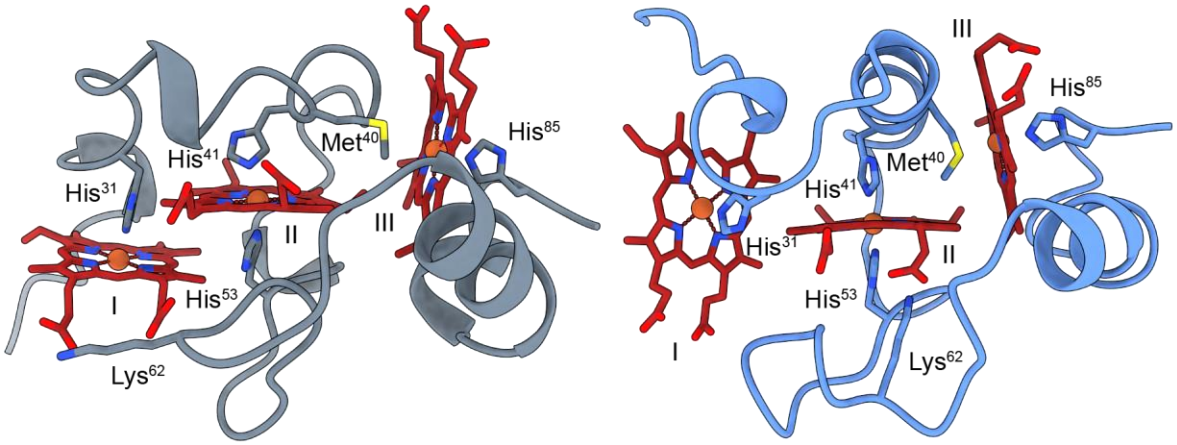


Figure 5.9. AlphaFold models of GSU0105. The AlphaFold 2 (gray) and 3 (blue) models of GSU0105 (residues 22-86) are represented as ribbon. The hemes (red) and their putative axial ligands are represented as sticks. The roman numerals indicate the hemes in their order of attachment to the CXXCH motifs in the polypeptide chain. The disordered C-terminal is not represented, for simplicity. The AlphaFold models of the protein colored with the respective pLDDT scores can be found in Figure 7.9.

In an extraordinarily quick fashion, AlphaFold 3 [68] also emerged during the writing of this Thesis, allowing for structural predictions in the presence of a predefined number of *c*-type heme groups. Surprisingly, the resulting model of GSU0105 (Figure 5.9) presented reasonably high average pLDDT scores (>75, Figure 7.9), for both the secondary structure and the *c*-type hemes. In this prediction, the positions of the heme axial ligands are conserved when compared with the AlphaFold 2 model of GSU0105, but the loop containing Lys^{62} is further displaced from heme I.

While it is important to note that AlphaFold models should be analyzed and considered carefully, it is interesting to observe that a couple of years after the completion of the experimental work, we are able to obtain computational models of this cytochrome with structural features that fit with our results. Furthermore, using these structures as templates for a search of structural homologs in the Dali [69], FoldSeek [70] and PDBeFold [71] servers, there are no known homologs of GSU0105 in the PDB database, which further supports our view that this cytochrome belongs to a new family of triheme cytochromes with an unique fold and heme core arrangement.

5.5. Materials and methods

5.5.1. DNA manipulation

The *gsu0105* gene sequence was retrieved from *G. sulfurreducens* PCA genome on the KEGG database [72], under the accession number T00155. The amino acid sequence was analyzed using the Uniprot bioinformatic tool to predict the signal peptide cleavage site [73]. The software predicted it to be located between residues Leu^{21} and Ala^{22}. The gene encoding for GSU0105 was amplified from *G. sulfurreducens* genomic DNA using Phusion High-Fidelity DNA polymerase (Thermo Fisher Scientific) and the primers listed in Table 5.3, which contain restriction sites for the enzymes NotI and HindIII. The resulting DNA fragment and the pVA203 vector [74] were digested with the referred restriction enzymes, purified using the E-gel electrophoresis system (Invitrogen) and ligated using T4 DNA ligase (Fermentas). The DNA was transformed into *E. coli* DH5α cells, and a colony PCR screen was performed using Taq DNA polymerase (VWR). Positive clones were grown in LB liquid media, and the plasmid was purified using the NZYMiniprep kit (NZYTech). The final plasmid was sequenced by DNA Sanger sequencing performed by STAB VIDA (Caparica, Portugal) and designated pGSU0105.

The primers used for the substitution of the putative axial ligand residues Met^{37}, Met^{40}, and Met^{60} for histidine residues were designed using the QuikChange Primer Design program (Agilent Technologies) and are listed in Table 5.3. The mutations were introduced in the pGSU0105 plasmid following the instructions of the NZYMutagenesis kit (NZYTech). The presence of the desired mutations was confirmed by DNA Sanger sequencing performed by STAB VIDA.

Table 5.3. Sequences of the primers used to produce the pGSU0105 plasmid and the three GSU0105 mutants. The cleavage sites for NotI and HindIII are underlined in the GSU0105WT primer forward (fw) and reverse (rv), respectively. In the primers used to produce the mutants, the substitution sites are underlined. The primers were purchased from Thermo Fischer Scientific or Invitrogen.

Primer	DNA sequence (5' – 3')
GSU0105WT_fw	GAGATGCGGCCGCCGCCGGGGCCTTCGAATGC
GSU0105WT_rv	GGTGCAAAGCTTCTATTCTCCAGCCTGCGGTTTC
GSU0105_M37H_fw	GTCACAGTAAAAACCCGGCCCATGTAAAAATGCACGAGGCCGT
GSU0105_M37H_rv	ACGGCCTCGTGCATTTTTACATGGGCCGGGTTTTTACTGTGAC
GSU0105_M40H_fw	ACAGTAAAAACCCGGCCATGGTAAAACATCACGAGGCCGTGCG
GSU0105_M40H_rv	CGCACGGCCTCGTGATGTTTTACCATGGCCGGGTTTTTACTGT
GSU0105_M60H_fw	GTCGGCGAACGCCTCCATGGAAAGGGGCAGCCT
GSU0105_M60H_rv	AGGCTGCCCCTTTCCATGGAGGCGTTCGCCGAC

5.5.2. Protein sequence analysis and AlphaFold predictions

A Protein BLAST [75] was performed with GSU0105's sequence, using the RefSeq Select protein sequences database and the default blastp algorithm (December 2020). From the 38 sequences

obtained, the ones from unclassified organisms were excluded, and the resulting 12 sequences were aligned using the Clustal Omega tool [76].

In order to access the positioning of GSU0105 protein within the class III of *c*-type cytochromes, the following representative members of this class [25] were selected for individual Protein BLAST searches: *G. sulfurreducens* c_7 triheme cytochrome PpcA, *D. acetoxidans* c_7 cytochrome, *S. oneidensis* triheme cytochrome, *G. sulfurreducens* dodecaheme cytochrome GSU1996, *S. oneidensis* c_3 tetraheme cytochrome, *S. frigidimarina* c_3 tetraheme cytochrome, *D. gigas* c_3 tetraheme cytochrome and *D. vulgaris* Hildenborough c_3 tetraheme cytochrome. The respective top hits were selected, and 104 sequences were used for MSA with the Clustal Omega tool [76]. In the case of the dodecaheme cytochromes, only domain A was used for the alignment. The alignments were loaded on Jalview [77] and used to generate a dendogram, that was then represented with Dendroscope [78].

The AlphaFold models of GSU0105 were either retrieved from the AlphaFold protein structure database [64, 79], or produced using either the AlphaFold2_MMSeqs2 GitHub ColabFold notebook [80] or the AlphaFold 3 server [68]. The quality of the individual models was evaluated based on the per-residue estimate of confidence produced by AlphaFold, designated pLDDT, which is based on the lDDT-Cα metric [81].

The *c*-type hemes of the AlphaFold 2 models were fitted in PyMol 2.5.2 [82] using local alignments and libraries of *c*-type hemes, containing the corresponding cysteine residues for the covalent binding of the heme vinyl groups and the proximal axial histidines for Fe coordination. These libraries were built by retrieving these residues and the corresponding *c*-type heme group from the structure of the OmcF cytochrome (PDB: 3CU4 [83]). Side-chain packing optimization and all-atom energy minimization protocols were not employed (see section 7.4). The final models were analyzed and represented in UCSF ChimeraX 1.4 [84].

5.5.3. Protein expression and purification

The GSU0105 cytochrome was synthetized using *E. coli* JM109 (DE3) cells as host. The host cells were transformed using the heat shock method [85] and contained two plasmids: 1) pEC86, encoding for the cytochrome *c* maturation gene cluster *ccmABCDEFGH*, required for the heme's incorporation [86, 87], and for a chloramphenicol resistance marker, and 2) pGSU0105, encoding the cytochrome GSU0105 after the OmpA leader sequence, a lac promoter and an ampicillin resistance marker [74]. These cells were grown at 30 °C in 2xYT, supplemented with 34 mg/mL chloramphenicol and 100 mg/mL ampicillin, to an OD_{600} of 1.5. At this stage, protein synthesis was induced with 10 μM of IPTG and the cell culture was grown overnight at 30 °C.

Cells were harvested by centrifugation at 6400 x*g* for 20 minutes at 4 °C. The cell pellet was gently resuspended in lysis buffer (100 mM Tris-HCl pH 8, 20% sucrose and 0.5 mM EDTA, containing 0.5 mg/mL of lysozyme), and after 15 minutes of incubation at room temperature, precooled water was added to the cell suspension, which was then left incubating on ice for 15 minutes to stop the lysozyme's activity. After that step, the suspension was centrifugated at 14,700 x*g* for 20 minutes at 4 °C. The supernatant constituting the periplasmic fraction was further ultracentrifugated at

225,000 x*g* for 1 hour at 4 °C to remove any remaining membrane debris. The final supernatant was dialyzed twice against 4.5 L of 20 mM Tris-HCl pH 7.5 using a Spectra/Por dialysis membrane (MWCO: 3.5 kDa) and then loaded onto 2x5 mL Bio-Scale Mini UNOsphere S cartridges (Bio-Rad), equilibrated with the same buffer. The protein was eluted with a sodium chloride gradient (0-300 mM). The obtained fractions were evaluated by SDS-PAGE (15% acrylamide/bis-acrylamide), stained either for hemes (TMBZ) and/or with BlueSafe (NZYTech). The fractions containing the protein were concentrated in Amicon Ultracentrifugal filter units (MWCO: 3 kDa) and equilibrated with 100 mM sodium phosphate buffer pH 8 before being injected in either a Superdex 75 XK 16/70 or a Superdex 75 Increase 10/300 GL molecular exclusion columns (Cytiva), equilibrated with the same buffer. Both chromatography steps were performed on an ÄKTA Pure system (Cytiva), and the final protein purity was evaluated by SDS-PAGE (15% acrylamide/bis-acrylamide), stained either for hemes (TMBZ) and/or with BlueSafe (NZYTech).

The concentration of the cytochrome was determined by measuring the absorbance of the reduced form at the α-band (552 nm), using the extinction coefficient of 97.5 mM^{-1} cm^{-1} [88]. The UV-visible spectra were recorded for both oxidized and reduced samples in the range 300 to 750 nm, at room temperature, using an Evolution 201 spectrophotometer (Thermo Scientific) and quartz cells (Hellma) with 1 cm of path length. The reduced samples were prepared by adding a concentrated solution of sodium dithionite (Fisher Scientific).

GSU0105's mutants (M37H, M40H, and M60H) were synthetized and purified using the same methodology.

5.5.4. CD spectroscopy

Circular dichroism experiments were performed using an Applied Photophysics Chirascan qCD spectropolarimeter (Leatherhead, UK), equipped with a thermostatic cell support and a 0.2 mm path length quartz cell (Hellma). CD spectra were acquired in the far-UV region using 25 µM of protein prepared in 20 mM sodium chloride at pH 7.4. The spectra are the average of three scans recorded in 1 nm steps with a scan rate of 3 s/nm, over a 190 to 260 nm wavelength range. The buffer contribution was corrected for all spectra and the temperature was controlled to ± 1 °C. After recording a spectrum at 25 °C, the conformational stability of the cytochrome was assessed by performing linear thermal-induced denaturation with a heating rate of 1 °C per minute, from 10 °C to 94 °C. Each spectrum was recorded with a scan rate of 0.4 s/nm. After the temperature ramp, the sample was quickly cooled to 25 °C, at which a final spectrum was registered.

The BeStSel deconvolution method [29] was used to estimate the secondary structural elements of the cytochrome.

5.5.5. EPR spectroscopy

X-band EPR spectra were obtained using a Bruker EMX spectrometer (Billerica) equipped with an Oxford Instruments ESR-900 continuous flow helium cryostat (Abingdon) and a high sensitivity perpendicular mode rectangular cavity. An as-purified GSU0105 sample was prepared in 32 mM

sodium phosphate buffer pH 7 with NaCl (100mM of final ionic strength), to a final protein concentration of 200 mM. The fully reduced sample was prepared anaerobically inside a glovebox by incubation with a slight excess of sodium dithionite. All spectra were recorded using a microwave frequency of 9.39 GHz, a modulation amplitude of 1.0 mT, and a microwave power of 2 mW. Spectra were analyzed and simulated using SpinCount [89].

5.5.6. NMR spectroscopy

NMR experiments were acquired in a Bruker Avance III 600 MHz spectrometer equipped with a 5 mM triple resonance cryoprobe (TCI). The residual H_2O signal was used as an internal reference for the calibration of the ^1H chemical shifts relative to DSS at 0 ppm [90]. The different spectra were processed and analyzed using TopSpin 3.5.7 (Bruker BioSpin, Karlsruhe, Germany). 1D ^1H-NMR spectra were acquired for the protein's oxidized state at 25 °C, with 32k data points, a spectral width of 96 kHz, with a total of 2048 transients, and water presaturation. For the samples in the reduced state, 1D ^1H-NMR spectra were acquired at 25 °C, with 16k data points, a spectral width of 19 kHz, with a total of 1024 transients, and water presaturation.

Protein samples were prepared with 100 μM concentration in 32 mM sodium phosphate buffer pH 7 with NaCl (100 mM of final ionic strength). The buffer was prepared either in pure D_2O or in 90% H_2O and 10% D_2O. The pH values of the samples were measured with a glass microelectrode and were not corrected for isotope effects. For sample reduction, the NMR tubes were sealed with a gas-tight serum cap, and the air was flushed out to avoid possible reoxidation. Then, the samples were reduced directly in the NMR tube with gaseous hydrogen (Air Liquide) in the presence of catalytic amounts of hydrogenase from *D. vulgaris* Hildenborough. To ensure that the protein samples were kept in the completely reduced state, the NMR tubes were left in a hydrogen atmosphere.

5.5.7. Redox titrations followed by visible spectroscopy

Redox titrations were performed at 15 °C inside an anaerobic glovebox (LABstar, MBraun) with argon circulation and O_2 levels kept under 1 ppm. The visible spectra were recorded in an Evolution 300 spectrophotometer (Thermo Scientific), connected to the anaerobic glovebox interior via fiber optics. The cytochrome samples were prepared with 10 μM protein concentration in 80 mM sodium phosphate buffer with NaCl (250 mM final ionic strength) at pH 7 and 8. The solution potentials were measured using a combined Pt/Ag/AgCl electrode (Crison), calibrated with quinhydrone (Merck) saturated solutions at pH 4 and 7. The following mixture of redox mediators was added to the solution with a final concentration of 1.3 μM each, as described in the literature [91], to ensure a good equilibrium between the redox centers of the proteins and the working electrode: potassium ferricyanide (Merck, $E^{0'}$ = +430 mV), *p*-benzoquinone (Sigma-Aldrich, $E^{0'}$ = +280 mV), tetramethyl-1,4-phenylenediamine (Acros Organics, $E^{0'}$ = +260 mV), 1,2-naphtoquinone-4-sulphonic acid (Acros Organics, $E^{0'}$ = +215 mV), 1,2-napthoquinone (Acros Organics, $E^{0'}$ = +143 mV), trimethylhydroquinone (Acros Organics, $E^{0'}$ = +115 mV), phenazine methosulfate (Sigma-Aldrich, $E^{0'}$ = +80 mV), phenazine ethosulfate (Sigma-Aldrich, $E^{0'}$ = +55 mV), gallocyanine (Alfa Aesar, $E^{0'}$ = +21 mV), methylene blue (Alfa

Aesar, $E^{0'}$ = +11 mV), indigo tetrasulfonate (Sigma-Aldrich, $E^{0'}$ = −30 mV), indigo trisulfonate (Acros Organics, $E^{0'}$ = −70 mV), indigo disulfonate (Sigma-Aldrich, $E^{0'}$ = −120 mV), 2-hidroxy-1,4-naphthoquinone (Alfa Aesar, $E^{0'}$ = −145 mV), anthraquinone-2,6-disulfonate (Acros Organics, $E^{0'}$ = −185 mV), anthraquinone-2-sulfonate (Sigma-Aldrich, $E^{0'}$ = −225 mV), safranine O (Alfa Aesar, $E^{0'}$ = −280 mV), neutral red (Fluka, $E^{0'}$ = −325 mV), benzyl viologen (Sigma-Aldrich, $E^{0'}$ = −345 mV), diquat (Fluka, $E^{0'}$ = −350 mV) and methyl viologen (Acros Organics, $E^{0'}$ = −440 mV). To check for hysteresis and verify reversibility, each redox titration was performed in both oxidative and reductive directions, using sodium dithionite as a reducing agent and potassium ferricyanide as an oxidizing agent. The experiments were performed at least two times and the reduction potentials (relative to SHE) were found to be reproducible within ± 5 mV. The reduced fraction of each cytochrome domain was determined by integrating the area of the α-band (552 nm) above the line connecting the flanking isosbestic points (542 and 560 nm) to subtract the optical contribution of the redox mediators. Each measured potential value was corrected to the SHE and a modified sequential model was fitted to the data using Origin Pro 8.5.

Modified sequential model

The different macrostates of a triheme cytochrome can be grouped into four oxidation stages, connected by three oxidation steps of one electron each, containing the same number of oxidized hemes (Figure 5.10).

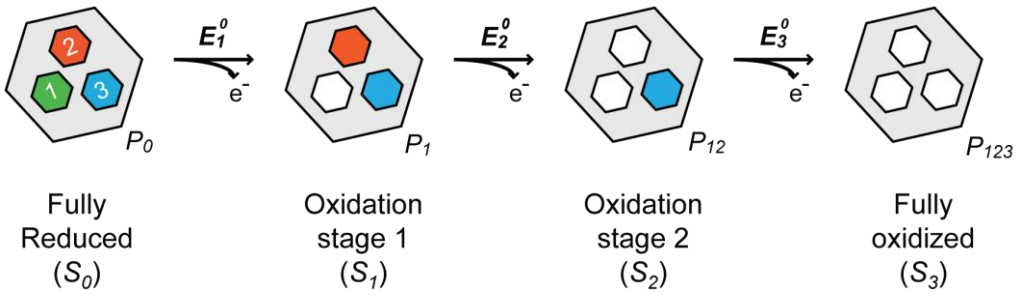


Figure 5.10. Electronic distribution scheme for a triheme cytochrome. The gray hexagons represent the different macrostates of the triheme cytochrome. The smaller inner hexagons represent heme groups. The reduced hemes 1, 2 and 3 are colored green, orange and blue, respectively. P_0, P_1, P_{12}, and P_{123} refer to the species with 0, 1, 2 or 3 oxidized hemes, respectively.

For triheme cytochromes, the variation of the reduced fraction (RF) with the solution redox potential is described by Equation 5.1, which is usually used in the experimental fitting of the data obtained through visible redox titrations:

$$RF = \frac{3 + 2e^{\left[(E-E_1^0)\frac{F}{RT}\right]} + e^{\left[(2E-E_1^0-E_2^0)\frac{F}{RT}\right]}}{3\left(1 + e^{\left[(E-E_1^0)\frac{F}{RT}\right]} + e^{\left[(2E-E_1^0-E_2^0)\frac{F}{RT}\right]} + e^{\left[(3E-E_1^0-E_2^0-E_3^0)\frac{F}{RT}\right]}\right)} \quad (5.1)$$

In this equation, the variables E, E_1^0, E_2^0 and E_3^0 represent the experimental measured potential and the standard reduction potential for each oxidation step, respectively; F represents the Faraday

constant (96486 C mol^{-1}); T represents the temperature (288 K); and R represents the molar gas constant (8.314 J K^{-1} mol^{-1}).

In the case of most triheme cytochromes, the experimental data can be fitted directly using the equation presented above. This happens because the three hemes of these proteins all have an equal contribution (0.3(3)) for the absorption of the α-band throughout the visible redox titration. However, in the case of GSU0105, as described and discussed above, the last heme to oxidize has a significant lower contribution for the absorption of the α-band (lower molar extinction coefficient at 552 nm). Therefore, in order to fit the experimental data using the model presented, the weighted contributions of each heme were allowed to vary freely. Furthermore, the last heme to oxidize was considered to be independent from the two remaining hemes. Equation 5.2, which follows the same considerations as the one presented for a three redox-center system, was used to fit the data for the two first hemes to oxidize (here designated as hemes 1 and 2 – H_1 and H_2):

$$RF\ (H_1/H_2) = \frac{2 + e^{\left[(E-E_1^0)\frac{F}{RT}\right]}}{2\ (1 + e^{\left[(E-E_1^0)\frac{F}{RT}\right]} + e^{\left[(2E-E_1^0-E_2^0)\frac{F}{RT}\right]})} \tag{5.2}$$

Then, for the last heme to oxidize (here designated as heme 3 – H_3), a single redox-center individual model was used to fit the experimental data (Equation 5.3):

$$RF\ (H_3) = \frac{1}{1 + e^{\left[(E-E_3^0)\frac{F}{RT}\right]}} \tag{5.3}$$

Finally, to describe the global redox behavior of GSU0105, a weighted fitting was made using Equation 5.4:

$$RF\ GSU0105 = \frac{(W_1 + W_2)\ (RF_{H_1/H_2}) + W_3\ (RF_{H_3})}{W_1 + W_2 + W_3} \tag{5.4}$$

In this equation, W_1, W_2 and W_3 represent the weighted contributions of hemes 1, 2 and 3, respectively, for the overall reduced fraction of GSU0105. RF_{H_1/H_2} and RF_{H_3} represent the reduced fractions for the first two hemes to oxidize (H_1 and H_2) and for the third heme to oxidize (H_3), respectively. The fitting obtained revealed that the first two hemes to oxidize (H_1 and H_2) have an equal contribution of 0.39 ± 0.02 for the overall reduced fraction of the cytochrome, whereas the final heme to oxidize (H_3) has a contribution of 0.22 ± 0.01.

5.6. References

[1] C.A. Salgueiro, L. Morgado, M.A. Silva, M.R. Ferreira, T.M. Fernandes, P.C. Portela, From iron to bacterial electroconductive filaments: Exploring cytochrome diversity using *Geobacter* bacteria, Coordin. Chem. Rev., 452 (2022) 214284.

[2] M.R. Ferreira, L. Morgado, C.A. Salgueiro, Periplasmic electron transfer network in *Geobacter sulfurreducens* revealed by biomolecular interaction studies, Protein Sci., 33 (2024) e5082.

[3] P.R. Pokkuluri, Y.Y. Londer, X. Yang, N.E. Duke, J. Erickson, V. Orshonsky, G. Johnson, M. Schiffer, Structural characterization of a family of cytochromes c_7 involved in Fe(III) respiration by *Geobacter sulfurreducens*, Biochim. Biophys. Acta, 1797 (2010) 222-232.

[4] P.R. Pokkuluri, Y.Y. Londer, N.E.C. Duke, W.C. Long, M. Schiffer, Family of cytochrome c_7-type proteins from *Geobacter sulfurreducens*: Structure of one cytochrome c_7 at 1.45 Å resolution, Biochemistry, 43 (2004) 849-859.

[5] L. Morgado, M. Bruix, V. Orshonsky, Y.Y. Londer, N.E. Duke, X. Yang, P.R. Pokkuluri, M. Schiffer, C.A. Salgueiro, Structural insights into the modulation of the redox properties of two *Geobacter sulfurreducens* homologous triheme cytochromes, Biochim. Biophys. Acta - Bioenergetics, 1777 (2008) 1157-1165.

[6] L. Morgado, M. Bruix, P.R. Pokkuluri, C.A. Salgueiro, D.L. Turner, Redox- and pH-linked conformational changes in triheme cytochrome PpcA from *Geobacter sulfurreducens*, Biochem. J., 474 (2017) 231-246.

[7] L. Morgado, V.B. Paixão, M. Schiffer, P.R. Pokkuluri, M. Bruix, C.A. Salgueiro, Revealing the structural origin of the redox-Bohr effect: The first solution structure of a cytochrome from *Geobacter sulfurreducens*, Biochem. J., 441 (2012) 179-187.

[8] L. Morgado, M. Bruix, M. Pessanha, Y.Y. Londer, C.A. Salgueiro, Thermodynamic characterization of a triheme cytochrome family from *Geobacter sulfurreducens* reveals mechanistic and functional diversity, Biophys. J., 99 (2010) 293-301.

[9] M. Aklujkar, M.V. Coppi, C. Leang, B.C. Kim, M.A. Chavan, L.A. Perpetua, L. Giloteaux, A. Liu, D.E. Holmes, Proteins involved in electron transfer to Fe(III) and Mn(IV) oxides by *Geobacter sulfurreducens* and *Geobacter uraniireducens*, Microbiology, 159 (2013) 515-535.

[10] Y.H. Ding, K.K. Hixson, C.S. Giometti, A. Stanley, A. Esteve-Núñez, T. Khare, S.L. Tollaksen, W. Zhu, J.N. Adkins, M.S. Lipton, R.D. Smith, T. Mester, D.R. Lovley, The proteome of dissimilatory metal-reducing microorganism *Geobacter sulfurreducens* under various growth conditions, Biochim. Biophys. Acta, 1764 (2006) 1198-1206.

[11] E.S. Shelobolina, M.V. Coppi, A.A. Korenevsky, L.N. DiDonato, S.A. Sullivan, H. Konishi, H. Xu, C. Leang, J.E. Butler, B.C. Kim, D.R. Lovley, Importance of *c*-type cytochromes for U(VI) reduction by *Geobacter sulfurreducens*, BMC Microbiol., 7 (2007) 16.

[12] T. Ueki, L.N. DiDonato, D.R. Lovley, Toward establishing minimum requirements for extracellular electron transfer in *Geobacter sulfurreducens*, FEMS Microbiol. Lett., 364 (2017) fnx093.

[13] S. Choi, C.H. Chan, D.R. Bond, Lack of specificity in *Geobacter* periplasmic electron transfer, J. Bacteriol., 204 (2022) e00322.

[14] S. Kato, K. Hashimoto, K. Watanabe, Iron-oxide minerals affect extracellular electron-transfer paths of *Geobacter* spp., Microbes Environ., 28 (2013) 141-148.
[15] S.M. Strycharz, R.H. Glaven, M.V. Coppi, S.M. Gannon, L.A. Perpetua, A. Liu, K.P. Nevin, D.R. Lovley, Gene expression and deletion analysis of mechanisms for electron transfer from electrodes to *Geobacter sulfurreducens*, Bioelectrochemistry, 80 (2011) 142-150.
[16] L.R. Teixeira, T.M. Fernandes, M.A. Silva, L. Morgado, C.A. Salgueiro, Characterization of a novel cytochrome involved in *Geobacter sulfurreducens*' electron harvesting pathways, Chem. Eur. J., 28 (2022) e202202333.
[17] I. Bertini, G. Cavallaro, A. Rosato, Cytochrome *c*: Occurrence and functions, Chem. Rev., 106 (2006) 90-115.
[18] L.J. Smith, A. Kahraman, J.M. Thornton, Heme proteins-diversity in structural characteristics, function, and folding, Proteins, 78 (2010) 2349-2368.
[19] V. Magro, L. Pieulle, N. Forget, B. Guigliarelli, Y. Petillot, E.C. Hatchikian, Further characterization of the two tetraheme cytochromes c_3 from *Desulfovibrio africanus*: Nucleotide sequences, EPR spectroscopy and biological activity, Biochim. Biophys. Acta - Protein Struct. Mol. Enzymol., 1342 (1997) 149-163.
[20] S. Nørager, P. Legrand, L. Pieulle, C. Hatchikian, M. Roth, Crystal structure of the oxidised and reduced acidic cytochrome c_3 from *Desulfovibrio africanus*, J. Mol. Biol., 290 (1999) 881-902.
[21] P.R. Pokkuluri, Y.Y. Londer, N.E. Duke, M. Pessanha, X. Yang, V. Orshonsky, L. Orshonsky, J. Erickson, Y. Zagyansky, C.A. Salgueiro, M. Schiffer, Structure of a novel dodecaheme cytochrome *c* from *Geobacter sulfurreducens* reveals an extended 12 nm protein with interacting hemes, J. Struct. Biol., 174 (2011) 223-233.
[22] R.H. Pires, S.S. Venceslau, F. Morais, M. Teixeira, A.V. Xavier, I.A.C. Pereira, Characterization of the *Desulfovibrio desulfuricans* ATCC 27774 DsrMKJOP complex – A membrane-bound redox complex involved in the sulfate respiratory pathway, Biochemistry, 45 (2006) 249-262.
[23] F. Grein, S.S. Venceslau, L. Schneider, P. Hildebrandt, S. Todorovic, I.A.C. Pereira, C. Dahl, DsrJ, an essential part of the DsrMKJOP transmembrane complex in the purple sulfur bacterium *Allochromatium vinosum*, is an unusual triheme cytochrome *c*, Biochemistry, 49 (2010) 8290-8299.
[24] C. Ferousi, S. Lindhoud, F. Baymann, E.R. Hester, J. Reimann, B. Kartal, Discovery of a functional, contracted heme-binding motif within a multiheme cytochrome, J. Biol. Chem., 294 (2019) 16953-16965.
[25] C.M. Paquete, G. Rusconi, A.V. Silva, R. Soares, R.O. Louro, A brief survey of the "cytochromome", Adv. Microb. Physiol., Academic Press, 2019, 69-135.
[26] R.P. Ambler, Sequence variability in bacterial cytochromes *c*, Biochim. Biophys. Acta - Bioenergetics, 1058 (1991) 42-47.
[27] M.R. Ferreira, J.M. Dantas, C.A. Salgueiro, The triheme cytochrome PpcF from *Geobacter metallireducens* exhibits distinct redox properties, FEBS Open Bio, 8 (2018) 1897-1910.

[28] P.C. Portela, T.M. Fernandes, J.M. Dantas, M.R. Ferreira, C.A. Salgueiro, Biochemical and functional insights on the triheme cytochrome PpcA from *Geobacter metallireducens*, Arch. Biochem. Biophys., 644 (2018) 8-16.

[29] A. Micsonai, É. Moussong, F. Wien, E. Boros, H. Vadászi, N. Murvai, Y.-H. Lee, T. Molnár, M. Réfrégiers, Y. Goto, Á. Tantos, J. Kardos, BeStSel: Webserver for secondary structure and fold prediction for protein CD spectroscopy, Nucleic Acids Res., 50 (2022) W90-W98.

[30] Q. Bashir, S. Scanu, M. Ubbink, Dynamics in electron transfer protein complexes, FEBS J., 278 (2011) 1391-1400.

[31] M. Prudêncio, M. Ubbink, Transient complexes of redox proteins: Structural and dynamic details from NMR studies, J. Mol. Recogn., 17 (2004) 524-539.

[32] M. Ubbink, Complexes of photosynthetic redox proteins studied by NMR, Photosynth. Res., 81 (2004) 277-287.

[33] W. Kabsch, C. Sander, Dictionary of protein secondary structure: Pattern recognition of hydrogen-bonded and geometrical features, Biopolymers, 22 (1983) 2577-2637.

[34] R. Varhač, D. Sedláková, M. Stupák, E. Sedlák, Non-two-state thermal denaturation of ferricytochrome *c* at neutral and slightly acidic pH values, Biophys. Chem., 203-204 (2015) 41-50.

[35] G.R. Moore, G.W. Pettigrew, Cytochromes *c*: Evolutionary, structural and physicochemical aspects, Springer-Verlag Heidelberg, Berlin, 1990.

[36] L.S. Kaminsky, V.J. Miller, A.J. Davison, Thermodynamic studies of the opening of the heme crevice of ferricytochrome *c*, Biochemistry, 12 (1973) 2215-2221.

[37] G.W. Pettigrew, G.R. Moore, Cytochromes *c*: Biological aspects, Springer-Verlag Heidelberg, Berlin, 1987.

[38] H. Theorell, Å. Åkesson, Studies on cytochrome *c*. II. The optical properties of pure cytochrome *c* and some of its derivatives, J. Am. Chem. Soc., 63 (1941) 1812-1818.

[39] C. Dumortier, J. Fitch, F.V. Petegem, W. Vermeulen, T.E. Meyer, J.J. Van Beeumen, M.A. Cusanovich, Protein dynamics in the region of the sixth ligand methionine revealed by studies of imidazole binding to *Rhodobacter capsulatus* cytochrome c_2 hinge mutants, Biochemistry, 43 (2004) 7717-7724.

[40] G.P. Holmes-Hampton, W.-H. Tong, T.A. Rouault, Biochemical and biophysical methods for studying mitochondrial iron metabolism, Methods in enzymology, Academic Press, 2014, 275-307.

[41] J. Ångström, G.R. Moore, R.J.P. Williams, The magnetic susceptibility of ferricytochrome *c*, Biochim. Biophys. Acta - Protein Struct., 703 (1982) 87-94.

[42] G.R. Moore, R.J. Williams, J. Peterson, A.J. Thomson, F.S. Mathews, A spectroscopic investigation of the structure and redox properties of *Escherichia coli* cytochrome *b*-562, Biochim. Biophys. Acta, 829 (1985) 83-96.

[43] J.M. Dantas, D.M. Tomaz, L. Morgado, C.A. Salgueiro, Functional characterization of PccH, a key cytochrome for electron transfer from electrodes to the bacterium *Geobacter sulfurreducens*, FEBS Lett., 587 (2013) 2662-2668.

[44] S.E. Rigby, G.R. Moore, J.C. Gray, P.M. Gadsby, S.J. George, A.J. Thomson, N.M.R., E.P.R. and Magnetic-C.D. studies of cytochrome *f* – Identity of the haem axial ligands, Biochem. J., 256 (1988) 571-577.

[45] M. Ubbink, A.P. Campos, M. Teixeira, N.I. Hunt, H.A.O. Hill, G.W. Canters, Characterization of mutant Met100Lys of cytochrome *c*-550 from *Thiobacillus versutus* with lysine-histidine heme ligation, Biochemistry, 33 (1994) 10051-10059.

[46] L. Morgado, A.P. Fernandes, Y.Y. Londer, M. Bruix, C.A. Salgueiro, One simple step in the identification of the cofactors signals, one giant leap for the solution structure determination of multiheme proteins, Biochem. Biophys. Res. Commun., 393 (2010) 466-470.

[47] J.M. Dantas, I.H. Saraiva, L. Morgado, M.A. Silva, M. Schiffer, C.A. Salgueiro, R.O. Louro, Orientation of the axial ligands and magnetic properties of the hemes in the cytochrome c_7 family from *Geobacter sulfurreducens* determined by paramagnetic NMR, Dalton Trans., 40 (2011) 12713-12718.

[48] H.C. Gaede, NMR exchange spectroscopy, Modern NMR spectroscopy in education, Am. Chem. Soc., 2007, 176-189.

[49] D. Sun, A. Wang, S. Cheng, M. Yates, B.E. Logan, *Geobacter anodireducens* sp. nov., an exoelectrogenic microbe in bioelectrochemical systems, Int. J. Syst. Evol. Microbiol., 64 (2014) 3485-3491.

[50] T. Catarino, M. Pessanha, A.G. De Candia, Z. Gouveia, A.P. Fernandes, P.R. Pokkuluri, D. Murgida, M.A. Marti, S. Todorovic, C.A. Salgueiro, Probing the chemotaxis periplasmic sensor domains from *Geobacter sulfurreducens* by combined resonance raman and molecular dynamic approaches: NO and CO sensing, J. Phys. Chem. B, 114 (2010) 11251-11260.

[51] C. Frazão, J.M. Dias, P.M. Matias, M.J. Romão, M.A. Carrondo, M. Hervas, J.A. Navarro, M. De La Rosa, G.M. Sheldrick, Cytochrome c_6 from the green alga *Monoraphidium braunii* – Crystallization and preminary diffraction studies, Acta Crystallogr. D, 51 (1995) 232-234.

[52] P.R. Pokkuluri, M. Pessanha, Y.Y. Londer, S.J. Wood, N.E.C. Duke, R. Wilton, T. Catarino, C.A. Salgueiro, M. Schiffer, Structures and solution properties of two novel periplasmic sensor domains with *c*-type heme from chemotaxis proteins of *Geobacter sulfurreducens*: Implications for signal transduction, J. Mol. Biol., 377 (2008) 1498-1517.

[53] T.C. Santos, M.A. Silva, L. Morgado, J.M. Dantas, C.A. Salgueiro, Diving into the redox properties of *Geobacter sulfurreducens* cytochromes: A model for extracellular electron transfer, Dalton Trans., 44 (2015) 9335-9344.

[54] L. Morgado, A.P. Fernandes, Y.Y. Londer, P.R. Pokkuluri, M. Schiffer, C.A. Salgueiro, Thermodynamic characterization of the redox centres in a representative domain of a novel *c*-type multihaem cytochrome, Biochem. J., 420 (2009) 485-492.

[55] C.A. Salgueiro, D.L. Turner, H. Santos, J. Le Gall, A.V. Xavier, Assignment of the redox potentials to the four haems in *Desulfovibrio vulgaris* cytochrome c_3 by 2D-NMR, FEBS Lett., 314 (1992) 155-158.

[56] R.O. Louro, T. Catarino, J. LeGall, D.L. Turner, A.V. Xavier, Cooperativity between electrons and protons in a monomeric cytochrome c_3: The importance of mechano-chemical coupling for energy transduction, Chembiochem, 2 (2001) 831-837.

[57] I.J. Correia, C.M. Paquete, R.O. Louro, T. Catarino, D.L. Turner, A.V. Xavier, Thermodynamic and kinetic characterization of trihaem cytochrome c_3 from *Desulfuromonas acetoxidans*, Eur. J. Biochem., 269 (2002) 5722-5730.

[58] R.O. Louro, I. Bento, P.M. Matias, T. Catarino, A.M. Baptista, C.M. Soares, M.A. Carrondo, D.L. Turner, A.V. Xavier, Conformational component in the coupled transfer of multiple electrons and protons in a monomeric tetraheme cytochrome, J. Biol. Chem., 276 (2001) 44044-44051.

[59] M. Coletta, H. Costa, G. De Sanctis, F. Neri, G. Smulevich, D.L. Turner, H. Santos, pH dependence of structural and functional properties of oxidized cytochrome *c*" from *Methylophilus methylotrophus*, J. Biol. Chem., 272 (1997) 24800-24804.

[60] P.M. Gadsby, J. Peterson, N. Foote, C. Greenwood, A.J. Thomson, Identification of the ligand-exchange process in the alkaline transition of horse heart cytochrome *c*, Biochem. J., 246 (1987) 43-54.

[61] P.O. Quintas, A.P. Cepeda, N. Borges, T. Catarino, D.L. Turner, Relative importance of driving force and electrostatic interactions in the reduction of multihaem cytochromes by small molecules, Biochim. Biophys. Acta - Bioenergetics, 1827 (2013) 745-750.

[62] H. Santos, D.L. Turner, Characterization and NMR studies of a novel cytochrome *c* isolated from *Methylophilus methylotrophus* which shows a redox-linked change of spin state, Biochim. Biophys. Acta - Protein Struct. Mol. Enzymol., 954 (1988) 277-286.

[63] L. Morgado, J.M. Dantas, M. Bruix, Y.Y. Londer, C.A. Salgueiro, Fine tuning of redox networks on multiheme cytochromes from *Geobacter sulfurreducens* drives physiological electron/proton energy transduction, Bioinorg. Chem. Appl., 2012 (2012) 298739.

[64] J. Jumper, R. Evans, A. Pritzel, T. Green, M. Figurnov, O. Ronneberger, K. Tunyasuvunakool, R. Bates, A. Žídek, A. Potapenko, A. Bridgland, C. Meyer, S.A.A. Kohl, A.J. Ballard, A. Cowie, B. Romera-Paredes, S. Nikolov, R. Jain, J. Adler, T. Back, S. Petersen, D. Reiman, E. Clancy, M. Zielinski, M. Steinegger, M. Pacholska, T. Berghammer, S. Bodenstein, D. Silver, O. Vinyals, A.W. Senior, K. Kavukcuoglu, P. Kohli, D. Hassabis, Highly accurate protein structure prediction with AlphaFold, Nature, 596 (2021) 583-589.

[65] J.F. Amacher, F. Zhong, G.P. Lisi, M.Q. Zhu, S.L. Alden, K.R. Hoke, D.R. Madden, E.V. Pletneva, A compact structure of cytochrome *c* trapped in a lysine-ligated state: Loop refolding and functional implications of a conformational switch, J. Am. Chem. Soc., 137 (2015) 8435-8449.

[66] M. Assfalg, I. Bertini, A. Dolfi, P. Turano, A.G. Mauk, F.I. Rosell, H.B. Gray, Structural model for an alkaline form of ferricytochrome *c*, J. Am. Chem. Soc., 125 (2003) 2913-2922.

[67] Y. Deng, V. Carnevale, R. Ditchfield, E.V. Pletneva, Applications of the newly developed force-field parameters uncover a dynamic nature of Ω-Loop C in the Lys-ligated alkaline form of cytochrome *c*, J. Phys. Chem. B, 128 (2024) 5935-5949.

[68] J. Abramson, J. Adler, J. Dunger, R. Evans, T. Green, A. Pritzel, O. Ronneberger, L. Willmore, A.J. Ballard, J. Bambrick, S.W. Bodenstein, D.A. Evans, C.-C. Hung, M. O'Neill, D. Reiman, K. Tunyasuvunakool, Z. Wu, A. Žemgulytė, E. Arvaniti, C. Beattie, O. Bertolli, A. Bridgland, A. Cherepanov, M. Congreve, A.I. Cowen-Rivers, A. Cowie, M. Figurnov, F.B. Fuchs, H. Gladman, R. Jain, Y.A. Khan, C.M.R. Low, K. Perlin, A. Potapenko, P. Savy, S. Singh, A. Stecula, A. Thillaisundaram, C. Tong, S. Yakneen, E.D. Zhong, M. Zielinski, A. Žídek, V. Bapst, P. Kohli, M. Jaderberg, D. Hassabis, J.M. Jumper, Accurate structure prediction of biomolecular interactions with AlphaFold 3, Nature, 630 (2024) 493-500.

[69] L. Holm, Dali server: Structural unification of protein families, Nucleic Acids Res., 50 (2022) W210-W215.

[70] M. van Kempen, S.S. Kim, C. Tumescheit, M. Mirdita, J. Lee, C.L.M. Gilchrist, J. Söding, M. Steinegger, Fast and accurate protein structure search with Foldseek, Nat. Biotechnol., 42 (2023) 243-246.

[71] M. Varadi, J. Berrisford, M. Deshpande, S.S. Nair, A. Gutmanas, D. Armstrong, L. Pravda, B. Al-Lazikani, S. Anyango, G.J. Barton, K. Berka, T. Blundell, N. Borkakoti, J. Dana, S. Das, S. Dey, P. Di Micco, F. Fraternali, T. Gibson, M. Helmer-Citterich, D. Hoksza, L.C. Huang, R. Jain, H. Jubb, C. Kannas, N. Kannan, J. Koca, R. Krivak, M. Kumar, E.D. Levy, F. Madeira, M.S. Madhusudhan, H.J. Martell, S. MacGowan, J.E. McGreig, S. Mir, A. Mukhopadhyay, L. Parca, T. Paysan-Lafosse, L. Radusky, A. Ribeiro, L. Serrano, I. Sillitoe, G. Singh, P. Skoda, R. Svodobova, J. Tyzack, A. Valencia, E.V. Fernandez, W. Vranken, M. Wass, J.M. Thornton, M. Sternberg, C. Orengo, S. Velankar, PDBe-KB: A community-driven resource for structural and functional annotations, Nucleic Acids Res., 48 (2020) D344-D353.

[72] M. Kanehisa, Y. Sato, M. Kawashima, M. Furumichi, M. Tanabe, KEGG as a reference resource for gene and protein annotation, Nucleic Acids Res., 44 (2016) D457-D462.

[73] T.U. Consortium, UniProt: The universal protein knowledgebase in 2021, Nucleic Acids Res., 49 (2021) D480-D489.

[74] P.R. Pokkuluri, Y.Y. Londer, N.E. Duke, J. Erickson, M. Pessanha, C.A. Salgueiro, M. Schiffer, Structure of a novel c_7-type three-heme cytochrome domain from a multidomain cytochrome *c* polymer, Protein Sci., 13 (2004) 1684-1692.

[75] S. McGinnis, T.L. Madden, BLAST: At the core of a powerful and diverse set of sequence analysis tools, Nucleic Acids Res., 32 (2004) W20-W25.

[76] F. Sievers, A. Wilm, D. Dineen, T.J. Gibson, K. Karplus, W. Li, R. Lopez, H. McWilliam, M. Remmert, J. Söding, J.D. Thompson, D.G. Higgins, Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega, Mol. Syst. Biol., 7 (2011) 539.

[77] A.M. Waterhouse, J.B. Procter, D.M.A. Martin, M. Clamp, G.J. Barton, Jalview Version 2 – A multiple sequence alignment editor and analysis workbench, Bioinformatics, 25 (2009) 1189-1191.

[78] D.H. Huson, C. Scornavacca, Dendroscope 3: An interactive tool for rooted phylogenetic trees and networks, Syst. Biol., 61 (2012) 1061-1067.

[79] M. Varadi, S. Anyango, M. Deshpande, S. Nair, C. Natassia, G. Yordanova, D. Yuan, O. Stroe, G. Wood, A. Laydon, A. Žídek, T. Green, K. Tunyasuvunakool, S. Petersen, J. Jumper, E. Clancy, R. Green,

A. Vora, M. Lutfi, M. Figurnov, A. Cowie, N. Hobbs, P. Kohli, G. Kleywegt, E. Birney, D. Hassabis, S. Velankar, AlphaFold protein structure database: Massively expanding the structural coverage of protein-sequence space with high-accuracy models, Nucleic Acids Res., 50 (2022) D439-D444.

[80] M. Mirdita, K. Schütze, Y. Moriwaki, L. Heo, S. Ovchinnikov, M. Steinegger, ColabFold: Making protein folding accessible to all, Nat. Methods, 19 (2022) 679-682.

[81] V. Mariani, M. Biasini, A. Barbato, T. Schwede, lDDT: A local superposition-free score for comparing protein structures and models using distance difference tests, Bioinformatics, 29 (2013) 2722-2728.

[82] L.L.C. Schrödinger, The PyMol molecular graphics system, Version 2.5.2, 2021.

[83] P.R. Pokkuluri, Y.Y. Londer, S.J. Wood, N.E. Duke, L. Morgado, C.A. Salgueiro, M. Schiffer, Outer membrane cytochrome *c*, OmcF, from *Geobacter sulfurreducens*: High structural similarity to an algal cytochrome c_6, Proteins, 74 (2009) 266-270.

[84] E.F. Pettersen, T.D. Goddard, C.C. Huang, E.C. Meng, G.S. Couch, T.I. Croll, J.H. Morris, T.E. Ferrin, UCSF ChimeraX: Structure visualization for researchers, educators, and developers, Protein Sci., 30 (2021) 70-82.

[85] M. Mandel, A. Higa, Calcium-dependent bacteriophage DNA infection, J. Mol. Biol., 53 (1970) 159-162.

[86] L. Thöny-Meyer, Biogenesis of respiratory cytochromes in bacteria, Microbiol. Mol. Biol. Rev., 61 (1997) 337-376.

[87] L. Thöny-Meyer, F. Fischer, P. Kunzler, D. Ritz, H. Hennecke, *Escherichia coli* genes required for cytochrome *c* maturation, J. Bacteriol., 177 (1995) 4321-4326.

[88] S. Seeliger, R. Cord-Ruwisch, B. Schink, A periplasmic and extracellular *c*-type cytochrome of *Geobacter sulfurreducens* acts as a ferric iron reductase and as an electron carrier to other acceptors or to partner bacteria, J. Bacteriol., 180 (1998) 3686-3691.

[89] D.T. Petasis, M.P. Hendrich, Quantitative interpretation of multifrequency multimode EPR spectra of metal containing proteins, enzymes, and biomimetic complexes, Methods in enzymology, Academic Press, 2015, 171-208.

[90] R. Pierattelli, L. Banci, D.L. Turner, Indirect determination of magnetic susceptibility tensors in peroxidases: A novel approach to structure elucidation by NMR, J. Biol. Inorg. Chem., 1 (1996) 320-329.

[91] P.L. Dutton, Redox potentiometry: Determination of midpoint potentials of oxidation-reduction components of biological electron-transfer systems, Methods in enzymology, Academic Press, 1978, 411-435.

6

Conclusions and future perspectives

List of contents

Conclusions and future perspectives 295

6.1. References.....300

The growing recognition of the environmental, economical, technological and social challenges posed by the existence of unsustainable practices is slowly driving society towards the application of sustainable alternatives. The impacts of climate change, including the rise of global temperatures, extreme weather events and loss of biodiversity, have underscored the necessity of reducing carbon emissions and mitigating environmental degradation. Moreover, the depletion of essential natural resources, including fossil fuels and freshwater, has accelerated the search for renewable energy sources and more efficient management of our planet's resources.

MET harness the unique abilities of electroactive microorganisms to generate electricity, clean waste and produce valuable chemicals, thus emerging as promising solutions to some of the key sustainability challenges faced by our society [1]. Nevertheless, the widespread application and efficiency of these technologies is still limited, mostly due to unoptimized and low-current producing electron transfer between microorganisms and electrodes. Consequently, understanding the fundamental phenomenona that govern EET in electroactive microorganisms from the structural and functional perspective is a paramount step towards the optimization of these processes [2].

The work developed in this Thesis stands as a significant contribution towards our understanding of the EET pathways of *Geobacter* bacteria, widely recognized as one of the groups of electroactive microorganisms with the greatest biotechnological potential [3].

Initially, taking advantage of a state-of-the-art protein structure prediction method, AlphaFold [4], we provided valuable insights into the porin-cytochrome complexes of *G. sulfurreducens*. The ability of AlphaFold to accurately define the heme binding sites of multiheme cytochromes, and the structural homology observed between the predicted models and proteins with experimentally determined structures, helped identify new functional hypotheses. The ability of the different proteins of the putative porin-cytochrome gene clusters to assemble into outer membrane complexes was also investigated, which in turn clarified their general architecture. Overall, these computational predictions gave us a clearer understanding of these unsought complexes, and were used as a guide for optimized protein cloning, resulting in an extensive library of constructs (Table 2.4) that sets the stage for future biochemical studies.

During the final stages of this Thesis, the release of AlphaFold 3 [5] allowed us to test the reliability of the AlphaFold 2 model predictions and, in particular, of the heme binding sites. Remarkably, the AlphaFold 3 predictions match those of AlphaFold 2, while avoiding structural violations and clashes near the heme binding pockets. The release of this version, for which the source code is unfortunately not available, shows how fast the field of computational biology is evolving and how it is altering the pace and scope of biological discoveries.

The functional and structural studies on ExtJ showed that the protein homodimerizes through the establishment of an intermolecular disulfide bridge. We hypothesize that this dimerization might be regulated by mechanisms of disulfide bond formation (oxidation) or cleavage (reduction), which would allow monomer ExtJ to interact with ExtI, and to work as a porin lid, which would indirectly control the passage of molecules through this outer membrane channel. To confirm this hypothesis, interaction studies between ExtJ and ExtI must be performed in the presence and absence of a

reducing agent, to probe for the effect of the oligomerization state of ExtJ in such assembly. Ideally, these studies would be complemented with in vivo assays in which the extracellular levels of target molecules are quantified, to further elucidate the fidelity and potential cellular impact of such mechanism.

The tridimensional structure of ExtJ, determined by NMR, shows that the protein is structurally homologous to flavin-binding proteins, but biomolecular interaction studies demonstrate that the protein is unable to bind FMN. Nevertheless, the notorious similarity between the structural features of flavin-binding proteins and ExtJ might be a key towards the elucidation of this protein's function.

The AlphaFold model of PgcA showed that the protein possesses a fuzzy global arrangement with three cytochrome domains linked by unstructured stretches. Biophysical studies performed on the individual domains and the full-length cytochrome showed that PgcA establishes a flexible electron chain, adopting multiple conformations that promote intra- and intermolecular electron transfer at variable distances. This mechanism, which might be advantageous during metal reduction, had not been previously described, which led us to place PgcA within a new class of cytochromes, termed "microbial heme-tethered redox strings". In the near future, interaction studies with PgcA and other putative electron transfer partners, such as the PpcA-family cytochromes and the OmcS nanowire, will be performed to clarify potential synergetic electron transfer pathways.

Finally, GSU0105 demonstrated unique structural and functional features, that distinguish it from other triheme cytochromes. Additional experiments should be performed to probe the exact nature of the putative pH-dependent spin-state interconversion mechanism, and complemented with the structure determination of the cytochrome in both redox states. This would allow us to identify the distal ligand of heme I and to map the redox-linked conformational changes that result in the spin-state interconversion. As a preliminary step, we have prepared six mutated versions of the cytochrome by replacing the putative heme I axial ligand lysine residues (Lys^{33}, Lys^{62} and Lys^{66}) for either histidine or methionine residues. The primers required for these substitutions were designed using the QuickChange Primer design program (Agilent Technologies) and are shown on Table 7.6. Based on the AlphaFold predictions, Lys^{62} is the most likely candidate, meaning our efforts should focus initially on the K62H and K62M mutants. One of these versions will hopefully have a low-spin configuration in both redox states, which might increase the cytochrome's stability and facilitate the remaining studies. Moreover, all this information will be complemented by interaction studies with possible physiological partners of GSU0105.

Globally, this Thesis stands as a testimony of the remarkable versatility of *Geobacter* bacteria. The ability to study such a diverse array of proteins, cytochromes and protein complexes within this single organism highlights its unique metabolic flexibility and adaptability. Each chapter of this work delves into different components of the bacterium's EET networks (Figure 6.1), showcasing how each element is woven into the survival strategy of this singular metal-reducing microorganism.

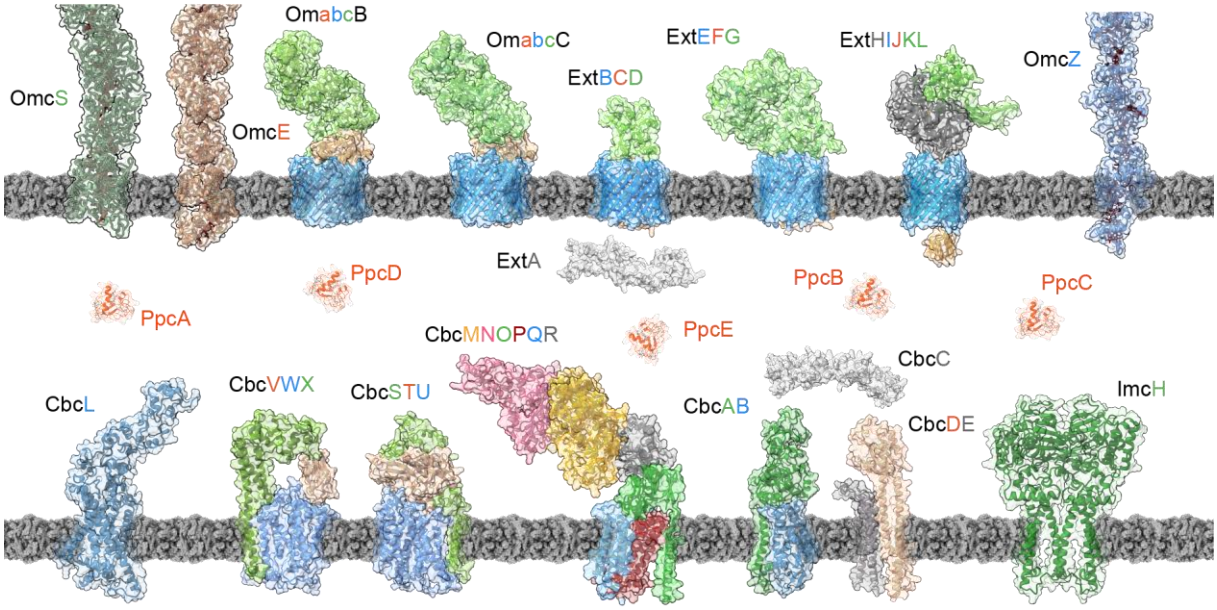


Figure 6.1. Updated view of the EET pathways of *G. sulfurreducens*. The main components of the EET pathways are represented either by their respective AlphaFold models (all inner membrane quinol dehydrogenases and outer membrane porin-cytochrome complexes) or by their experimentally determined structures (cryo-EM structures of the OmcE (PDB: 7TFS [6]), OmcS (PDB: 6EF8 [7]), and OmcZ (PDB: TLQ5 [8]) cytochrome nanowires, and solution or X-ray crystallography structures of the PpcA (PDB: 2LDO [9]), PpcB (PDB: 3BXU [10]), PpcC (PDB: 3H33 [11]), PpcD (PDB: 3H4N [11]) and PpcE (PDB: 3H34 [11]) periplasmic cytochromes). The heme groups of the PpcA-family cytochromes and cytochrome nanowires are represented as gray and red sticks, respectively. The ImcH cytochrome is represented as a dimer, according to [12]. The pLDDT scores of the AlphaFold models which were not previously represented in this Thesis can be found in Figure 7.10.

These complementary results are a small but important stepping-stone towards the development of sustainable MET, whose industrial scale application will require efforts from multiple disciplines, including material sciences, microbiology, electrochemistry and biochemistry. These technologies hold the potential to help mitigate the sustainability challenges currently facing mankind.

6.1. References

[1] R. Peixoto, C.R. Voolstra, L.Y. Stein, P. Hugenholtz, J.F. Salles, S.A. Amin, M. Häggblom, A. Gregory, T.P. Makhalanyane, F. Wang, N.A. Agbodjato, Y. Wang, N. Jiao, J.T. Lennon, A. Ventosa, P.M. Bavoil, V. Miller, J.A. Gilbert, Microbial solutions must be deployed against climate catastrophe, Nat. Commun., 15 (2024) 9637.

[2] T.M. Fernandes, L. Morgado, D.L. Turner, C.A. Salgueiro, Protein engineering of electron transfer components from electroactive *Geobacter* bacteria, Antioxidants, 10 (2021) 844.

[3] B.E. Logan, R. Rossi, A. Ragab, P.E. Saikaly, Electroactive microorganisms in bioelectrochemical systems, Nat. Rev. Microbiol., 17 (2019) 307-319.

[4] J. Jumper, R. Evans, A. Pritzel, T. Green, M. Figurnov, O. Ronneberger, K. Tunyasuvunakool, R. Bates, A. Žídek, A. Potapenko, A. Bridgland, C. Meyer, S.A.A. Kohl, A.J. Ballard, A. Cowie, B. Romera-Paredes, S. Nikolov, R. Jain, J. Adler, T. Back, S. Petersen, D. Reiman, E. Clancy, M. Zielinski, M. Steinegger, M. Pacholska, T. Berghammer, S. Bodenstein, D. Silver, O. Vinyals, A.W. Senior, K. Kavukcuoglu, P. Kohli, D. Hassabis, Highly accurate protein structure prediction with AlphaFold, Nature, 596 (2021) 583-589.

[5] J. Abramson, J. Adler, J. Dunger, R. Evans, T. Green, A. Pritzel, O. Ronneberger, L. Willmore, A.J. Ballard, J. Bambrick, S.W. Bodenstein, D.A. Evans, C.-C. Hung, M. O'Neill, D. Reiman, K. Tunyasuvunakool, Z. Wu, A. Žemgulytė, E. Arvaniti, C. Beattie, O. Bertolli, A. Bridgland, A. Cherepanov, M. Congreve, A.I. Cowen-Rivers, A. Cowie, M. Figurnov, F.B. Fuchs, H. Gladman, R. Jain, Y.A. Khan, C.M.R. Low, K. Perlin, A. Potapenko, P. Savy, S. Singh, A. Stecula, A. Thillaisundaram, C. Tong, S. Yakneen, E.D. Zhong, M. Zielinski, A. Žídek, V. Bapst, P. Kohli, M. Jaderberg, D. Hassabis, J.M. Jumper, Accurate structure prediction of biomolecular interactions with AlphaFold 3, Nature, 630 (2024) 493-500.

[6] F. Wang, K. Mustafa, V. Suciu, K. Joshi, C.H. Chan, S. Choi, Z. Su, D. Si, A.I. Hochbaum, E.H. Egelman, D.R. Bond, Cryo-EM structure of an extracellular *Geobacter* OmcE cytochrome filament reveals tetrahaem packing, Nat. Microbiol., 7 (2022) 1291-1300.

[7] F. Wang, Y. Gu, J.P. O'Brien, S.M. Yi, S.E. Yalcin, V. Srikanth, C. Shen, D. Vu, N.L. Ing, A.I. Hochbaum, E.H. Egelman, N.S. Malvankar, Structure of microbial nanowires reveals stacked hemes that transport electrons over micrometers, Cell, 177 (2019) 361-369.

[8] Y. Gu, M.J. Guberman-Pfeffer, V. Srikanth, C. Shen, F. Giska, K. Gupta, Y. Londer, F.A. Samatey, V.S. Batista, N.S. Malvankar, Structure of *Geobacter* cytochrome OmcZ identifies mechanism of nanowire assembly and conductivity, Nat. Microbiol., 8 (2023) 284-298.

[9] L. Morgado, V.B. Paixão, M. Schiffer, P.R. Pokkuluri, M. Bruix, C.A. Salgueiro, Revealing the structural origin of the redox-Bohr effect: The first solution structure of a cytochrome from *Geobacter sulfurreducens*, Biochem. J., 441 (2012) 179-187.

[10] L. Morgado, M. Bruix, V. Orshonsky, Y.Y. Londer, N.E. Duke, X. Yang, P.R. Pokkuluri, M. Schiffer, C.A. Salgueiro, Structural insights into the modulation of the redox properties of two *Geobacter sulfurreducens* homologous triheme cytochromes, Biochim. Biophys. Acta - Bioenergetics, 1777 (2008) 1157-1165.

[11] P.R. Pokkuluri, Y.Y. Londer, X. Yang, N.E. Duke, J. Erickson, V. Orshonsky, G. Johnson, M. Schiffer, Structural characterization of a family of cytochromes c_7 involved in Fe(III) respiration by *Geobacter sulfurreducens*, Biochim. Biophys. Acta, 1797 (2010) 222-232.

[12] A.I. Pimenta, C.M. Paquete, L. Morgado, M.J. Edwards, T.A. Clarke, C.A. Salgueiro, I.A.C. Pereira, A.G. Duarte, Characterization of the inner membrane cytochrome ImcH from *Geobacter* reveals its importance for extracellular electron transfer and energy conservation, Protein Sci., 32 (2023) e4796.

7

Appendix

List of contents

Appendix 303

7.1. AlphaFold models with pLDDT score coloring 305

7.2. Supplementary tables 310

7.3. MD simulations – Additional plots 327

7.4. PyMol script for insertion of *c*-type hemes in AlphaFold models 328

7.5. References 330

7.1. AlphaFold models with pLDDT score coloring

Throughout this Thesis, AlphaFold was extensively used to predict the structural models of various individual proteins and complexes [1]. For visual representation purposes, these models were mostly represented in specific colors, rather than with their corresponding pLDDT color code. The following figures include this data.

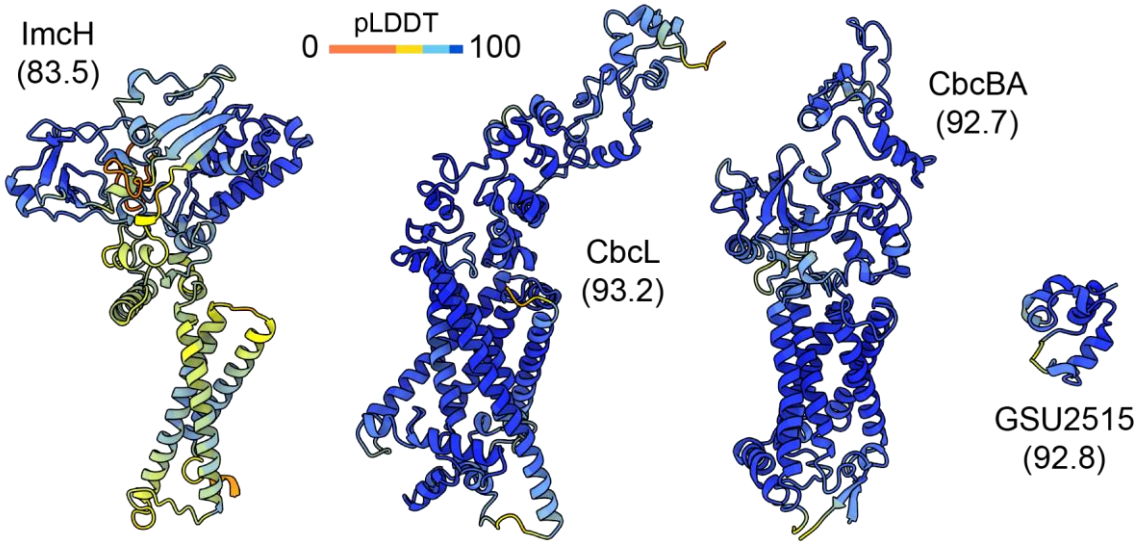


Figure 7.1. AlphaFold models represented in Chapter 1. Each protein is identified along with its average pLDDT score (in parenthesis). For reference, a pLDDT color scale bar is shown.

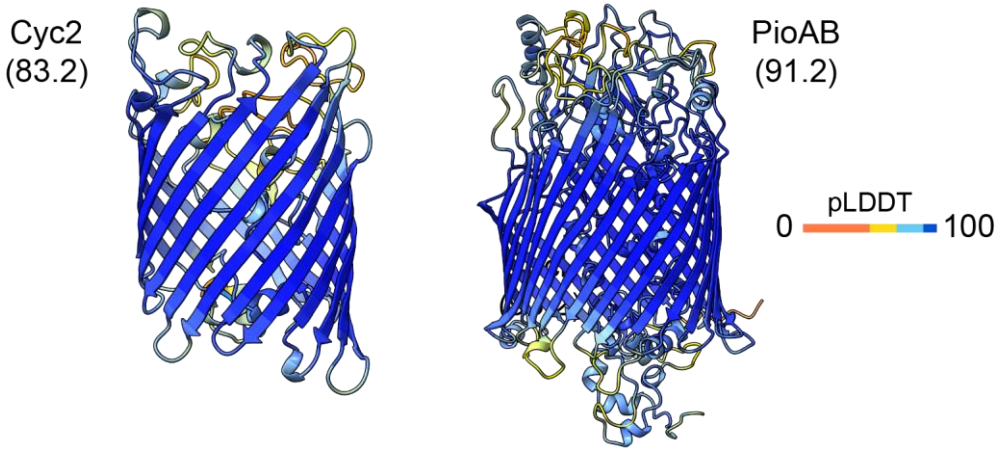


Figure 7.2. AlphaFold models of Cyc2 and PioAB. Each protein is identified along with its average pLDDT score (in parenthesis). For reference, a pLDDT color scale bar is shown.

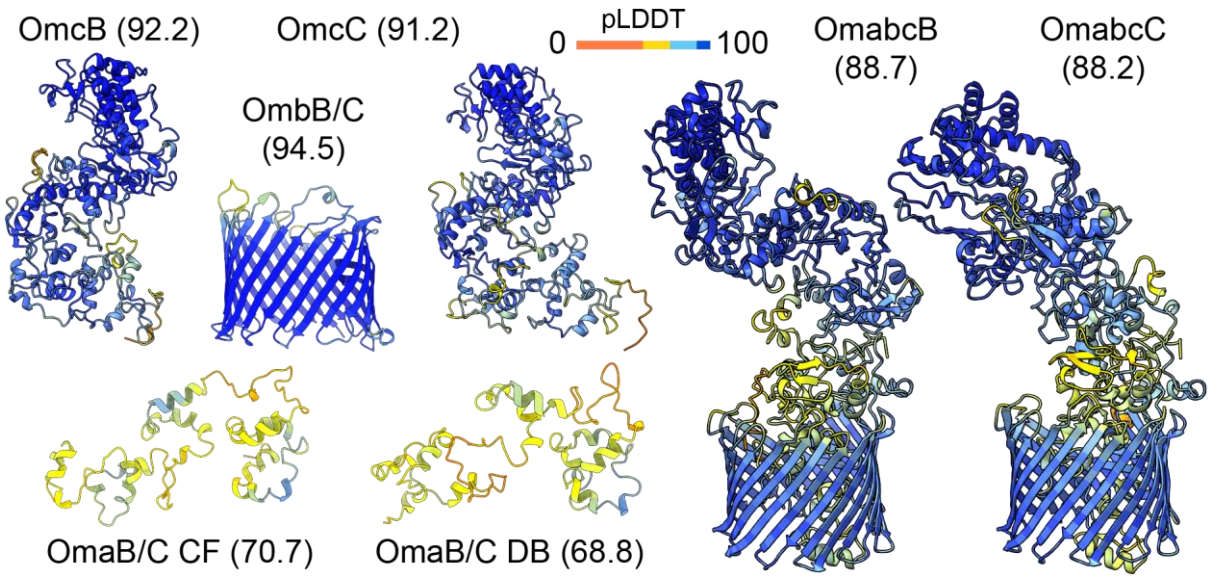


Figure 7.3. AlphaFold models of OmabcB/C porin-cytochrome complexes. Each individual protein or complex is identified along with its average pLDDT score (in parenthesis). For reference, a pLDDT color scale bar is shown. "CF" and "DB" refer to the ColabFold [2] and AlphaFold protein structure database [1, 3] models predicted for OmaB/C.

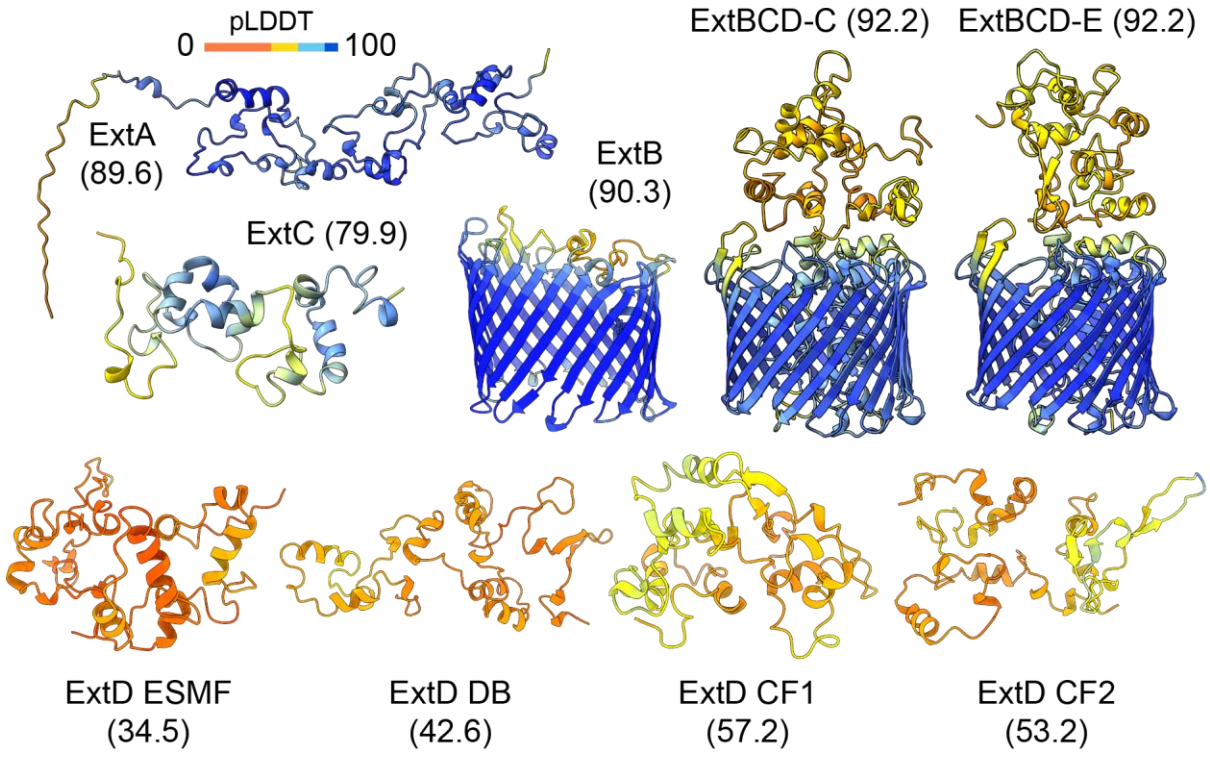


Figure 7.4. AlphaFold models of ExtABCD porin-cytochrome complex. Each individual protein or complex is identified along with its average pLDDT score (in parenthesis). For reference, a pLDDT color scale bar is shown. "CF", "DB" and "ESMF" refer to the ColabFold [2], AlphaFold protein structure database [1, 3] and ESMFold [4] models predicted for ExtD, respectively. ExtBCD-C and ExtBCD-E refer to ExtBCD porin-cytochrome complex models in which ExtD is found either in a "contracted" or "extended" configuration, respectively.

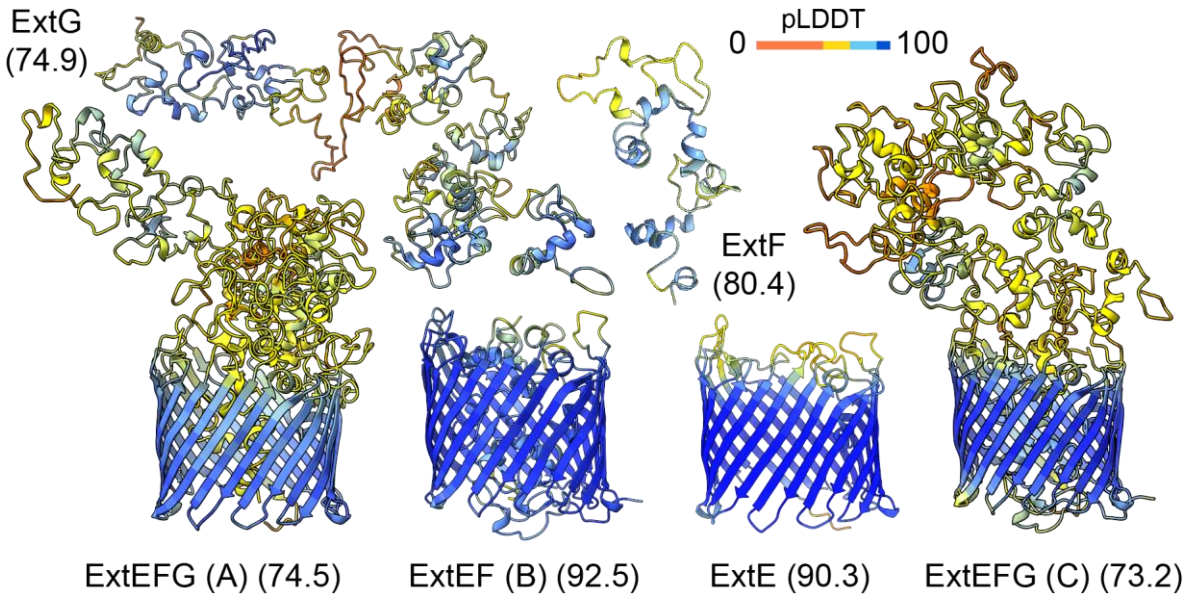


Figure 7.5. AlphaFold models of ExtEFG porin-cytochrome complex. Each individual protein or complex is identified along with its average pLDDT score (in parenthesis). For reference, a pLDDT color scale bar is shown. The (A-C) identification refers to the panels of Figure 2.18.

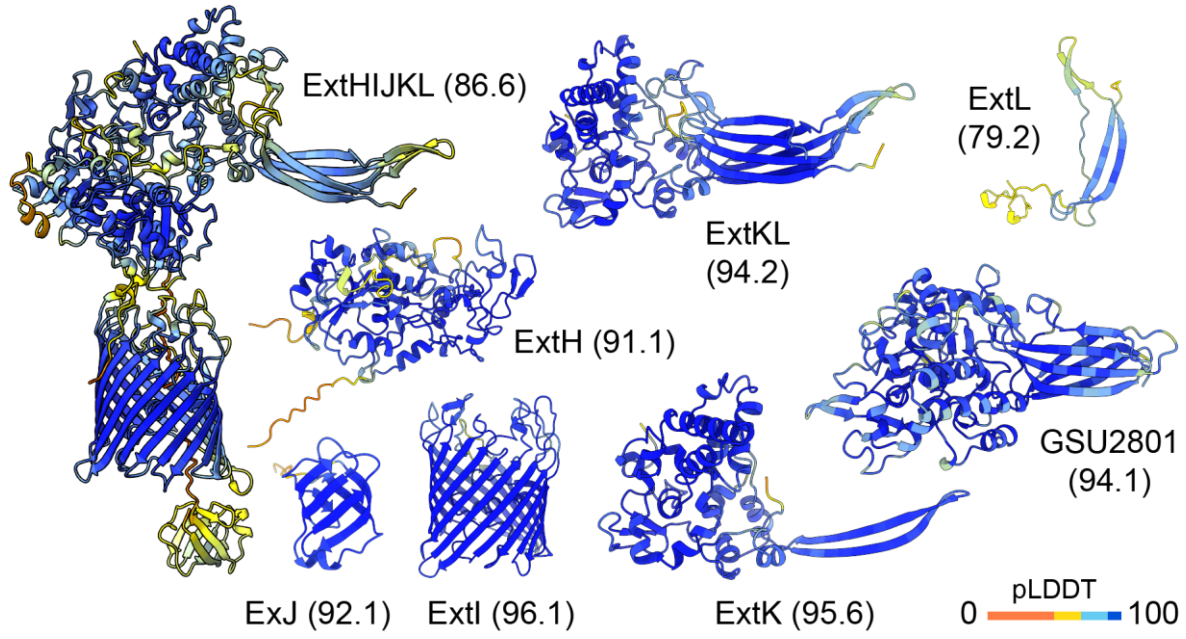


Figure 7.6. AlphaFold models of ExtHIJKL porin-cytochrome complex. Each individual protein or complex is identified along with its average pLDDT score (in parenthesis). For reference, a pLDDT color scale bar is shown.

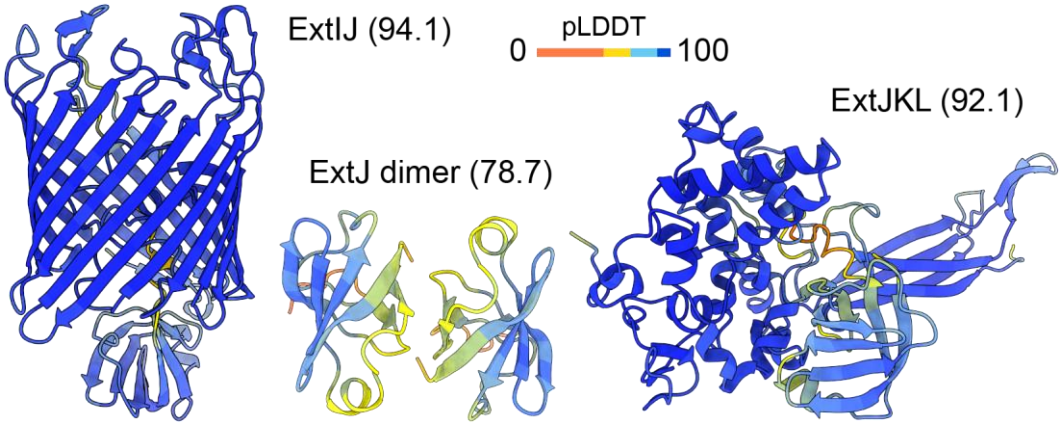


Figure 7.7. AlphaFold models of the ExtJ homodimer, and of the ExtIJ and ExtJKL complexes. Each individual protein or complex is identified along with its average pLDDT score (in parenthesis). For reference, a pLDDT color scale bar is shown.

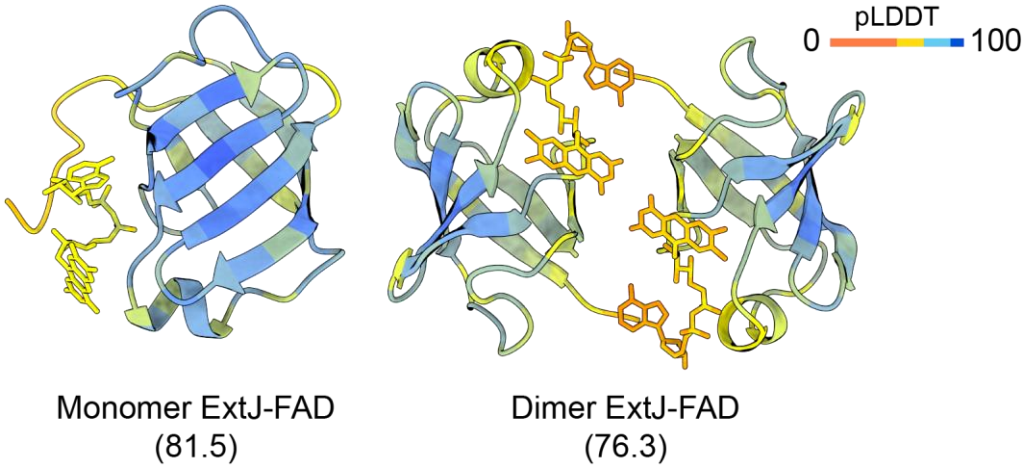


Figure 7.8. AlphaFold 3 models of monomer and homodimer ExtJ with FAD. Each individual protein, complex or ligand is identified along with its average pLDDT score (in parenthesis). For reference, a pLDDT color scale bar is shown.

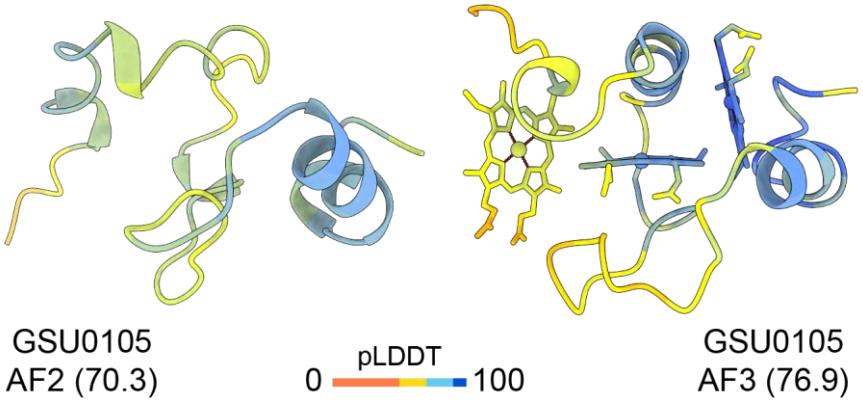


Figure 7.9. AlphaFold 2 and 3 models of GSU0105. Each individual protein or ligand is identified along with its average pLDDT score (in parenthesis). For reference, a pLDDT color scale bar is shown.

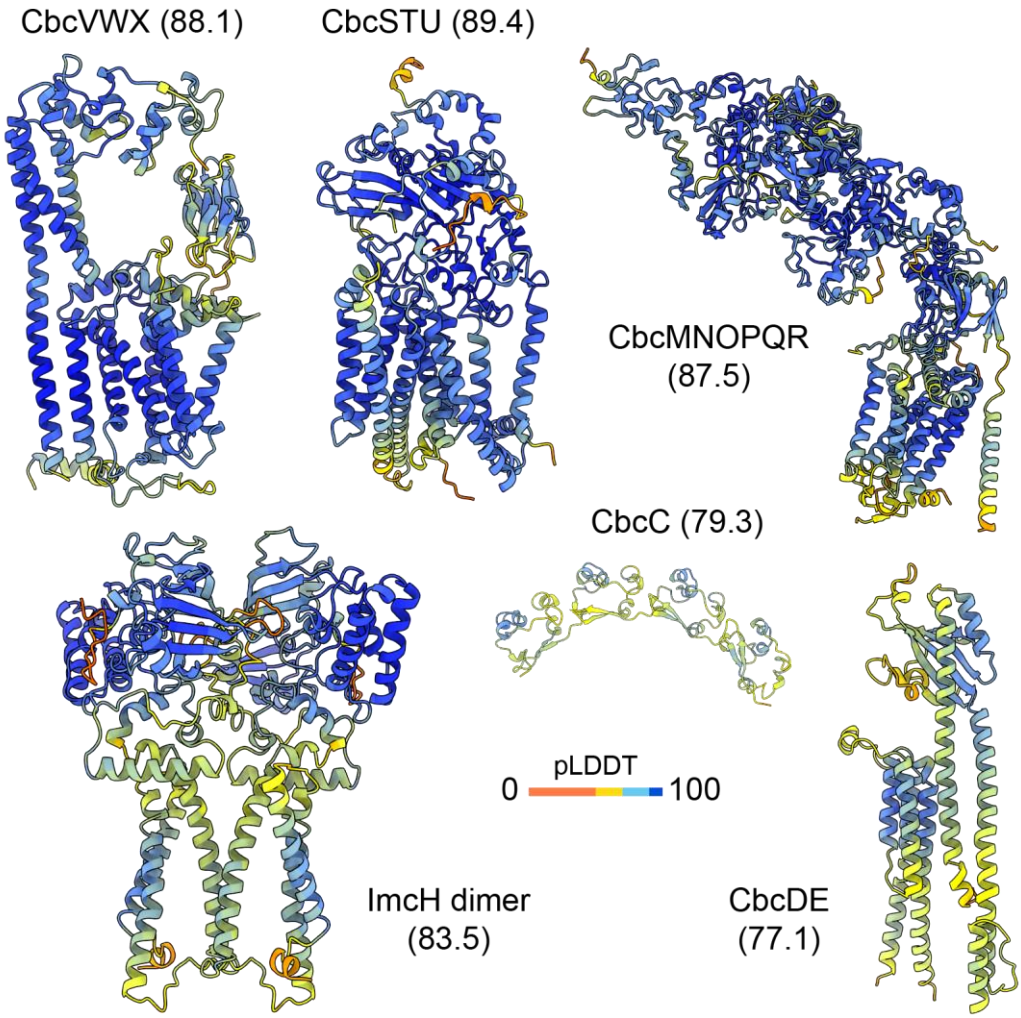


Figure 7.10. AlphaFold models of different quinol hydrogenases of *G. sulfurreducens*. Each individual protein or complex is identified along with its average pLDDT score (in parenthesis). For reference, a pLDDT color scale bar is shown. The average pLDDT scores of the remaining quinol hydrogenases are shown in Figure 7.1.

7.2. Supplementary tables

Table 7.1. Assignment of the ^{1}H, ^{13}C and ^{15}N resonances of ExtJ. The spectra used for the protein assignment were acquired at 25 °C in a Bruker Avance Neo 500 MHz spectrometer equipped with a 5 mM triple resonance Prodigy cryoprobe (TCI), with a 1 mM protein sample in 32 mM sodium phosphate buffer pH 6 (100 mM of final ionic strength) with 1 mM DTT.

Residue	Atom	δ (ppm)	Residue	Atom	δ (ppm)
A1	C	178.15	**F4**	$H_{\beta 2/3}$	3.27
	C_α	52.54		$H_{\delta 1/2}$	7.15
	C_β	19.28		$H_{\varepsilon 1/2}$	6.71
	H	8.42		H_ζ	6.65
	H_α	4.29		N	118.92
	H_β	1.33	**S5**	C	173.73
	N	126.23		C_α	56.62
G2	C	173.14		C_β	65.83
	C_α	45.45		H	8.81
	H	8.92		H_α	5.53
	$H_{\alpha 2}$	3.76		$H_{\beta 2}$	3.75
	$H_{\alpha 3}$	4.20		$H_{\beta 3}$	3.83
	N	109.96		N	114.61
S3	C	173.30	**G6**	C	171.11
	C_α	57.87		C_α	45.20
	C_β	66.60		H	8.51
	H	8.29		$H_{\alpha 2}$	3.61
	H_α	5.94		$H_{\alpha 3}$	4.65
	$H_{\beta 2}$	3.76		N	108.08
	$H_{\beta 3}$	3.93	**K7**	C	176.27
	N	113.03		C_α	54.11
F4	C	171.95		C_β	35.36
	C_α	56.65		C_γ	24.38
	C_β	41.04		C_δ	29.25
	$C_{\delta 1/2}$	133.05		C_ε	41.92
	$C_{\varepsilon 1/2}$	129.36		H	8.37
	C_ζ	129.31		H_α	5.04
	H	9.52		$H_{\beta 2/3}$	1.67
	H_α	5.17		$H_{\gamma 2/3}$	1.30

Table 7.1. *Continued.*

Residue	Atom	δ (ppm)
K7	$H_{\delta 2/3}$	1.67
	$H_{\varepsilon 2}$	2.63
	$H_{\varepsilon 3}$	2.98
	N	120.62
V8	C	177.05
	C_{α}	63.80
	C_{β}	31.47
	$C_{\gamma 1}$	22.35
	$C_{\gamma 2}$	22.67
	H	9.15
	H_{α}	4.16
	H_{β}	2.15
	$H_{\gamma 1}$	0.68
	$H_{\gamma 2}$	0.77
	N	125.21
V9	C	176.05
	C_{α}	61.45
	C_{β}	33.33
	$C_{\gamma 1}$	19.07
	$C_{\gamma 2}$	21.87
	H	8.94
	H_{α}	4.65
	H_{β}	2.25
	$H_{\gamma 1}$	0.80
	$H_{\gamma 2}$	0.96
	N	123.08
K10	C	173.34
	C_{α}	57.08
	C_{β}	36.14
	C_{γ}	25.00
	C_{δ}	29.33
	C_{ε}	41.94
K10	H	7.52
	H_{α}	4.43
	$H_{\beta 2}$	1.72
	$H_{\beta 3}$	1.83
	$H_{\gamma 2/3}$	1.32
	$H_{\delta 2/3}$	1.64
	$H_{\varepsilon 2/3}$	2.96
	N	121.03
V11	C	174.48
	C_{α}	62.99
	C_{β}	33.92
	$C_{\gamma 1}$	20.94
	$C_{\gamma 2}$	21.25
	H	8.59
	H_{α}	4.38
	H_{β}	1.89
	$H_{\gamma 1}$	0.79
	$H_{\gamma 2}$	0.83
	N	125.82
D12	C	175.81
	C_{α}	52.24
	C_{β}	42.00
	H	8.98
	H_{α}	4.86
	$H_{\beta 2}$	2.31
	$H_{\beta 3}$	2.74
	N	128.34
G13	C	173.93
	C_{α}	47.38
	H	8.96
	$H_{\alpha 2}$	3.54
	$H_{\alpha 3}$	3.97

Table 7.1. ***Continued.***

Residue	Atom	δ (ppm)	Residue	Atom	δ (ppm)
G13	N	115.56	**V16**	H	9.70
E14	C	175.41		H_{α}	4.65
	C_{α}	57.39		H_{β}	2.15
	C_{β}	30.48		$H_{\gamma1}$	1.61
	C_{γ}	37.37		$H_{\gamma2}$	0.90
	H	8.92		N	127.87
	H_{α}	4.02	**T17**	C	173.89
	$H_{\beta2}$	2.05		C_{α}	62.71
	$H_{\beta3}$	2.51		C_{β}	69.26
	$H_{\gamma2}$	2.23		$C_{\gamma2}$	21.25
	$H_{\gamma3}$	2.31		H	9.31
	N	125.69		H_{α}	5.12
K15	C	176.05		H_{β}	3.98
	C_{α}	55.83		$H_{\gamma2}$	1.01
	C_{β}	33.50		N	124.66
	C_{γ}	25.77	**V18**	C	173.63
	C_{δ}	29.37		C_{α}	59.66
	C_{ε}	41.07		C_{β}	34.66
	H	7.91		$C_{\gamma1}$	21.57
	H_{α}	5.16		$C_{\gamma2}$	20.47
	$H_{\beta2}$	1.48		H	9.54
	$H_{\beta3}$	2.00		H_{α}	4.92
	$H_{\gamma2/3}$	1.10		H_{β}	1.83
	$H_{\delta2/3}$	2.04		$H_{\gamma1}$	0.89
	$H_{\varepsilon2}$	2.27		$H_{\gamma2}$	0.70
	$H_{\varepsilon3}$	2.76		N	127.73
	N	117.71	**K19**	C	175.57
V16	C	174.71		C_{α}	54.89
	C_{α}	61.13		C_{β}	35.52
	C_{β}	35.53		C_{γ}	24.82
	$C_{\gamma1}$	19.20		C_{δ}	29.52
	$C_{\gamma2}$	21.78		C_{ε}	41.88

Table 7.1. *Continued.*

Residue	Atom	δ (ppm)	Residue	Atom	δ (ppm)
K19	H	9.29	**K22**	H	7.55
	H_{α}	5.14		H_{α}	4.18
	$H_{\beta2}$	1.61		$H_{\beta2/3}$	1.64
	$H_{\beta3}$	1.83		$H_{\gamma2/3}$	1.19
	$H_{\gamma2/3}$	1.32		$H_{\delta2/3}$	1.63
	$H_{\delta2/3}$	1.63		$H_{\varepsilon2}$	2.94
	$H_{\varepsilon2}$	2.95		$H_{\varepsilon3}$	2.97
	$H_{\varepsilon3}$	2.99		N	115.77
	N	127.25	**V23**	C	174.29
A20	C	175.63		C_{α}	59.44
	C_{α}	50.41		C_{β}	32.24
	C_{β}	23.19		$C_{\gamma1}$	21.21
	H	9.28		$C_{\gamma2}$	21.66
	H_{α}	4.97		H	8.48
	H_{β}	1.36		H_{α}	3.31
	N	133.00		H_{β}	1.65
E21	C	176.91		$H_{\gamma1}$	0.46
	C_{α}	58.80		$H_{\gamma2}$	0.57
	C_{β}	29.90		N	126.16
	C_{γ}	36.41	**P24**	C	177.24
	H	8.81		C_{α}	617.78
	H_{α}	4.13		C_{β}	32.88
	$H_{\beta2/3}$	2.08		C_{γ}	27.31
	$H_{\gamma2}$	2.29		C_{δ}	49.67
	$H_{\gamma3}$	2.35		H_{α}	4.21
	N	119.87		$H_{\beta2}$	0.66
K22	C	174.43		$H_{\beta3}$	2.02
	C_{α}	55.37		$H_{\gamma2}$	0.39
	C_{β}	34.11		$H_{\gamma3}$	1.36
	C_{γ}	24.39		$H_{\delta2}$	1.67
	C_{δ}	28.92		$H_{\delta3}$	2.67
	C_{ε}	42.03	**A25**	C	177.91

Table 7.1. ***Continued.***

Residue	Atom	δ (ppm)
A25	C_α	54.43
	C_β	18.41
	H	8.49
	H_α	3.88
	H_β	1.40
	N	120.63
W26	C	176.90
	C_α	54.89
	C_β	28.41
	$C_{\delta1}$	128.72
	$C_{\varepsilon3}$	120.30
	$C_{\zeta2}$	113.52
	$C_{\zeta3}$	120.50
	$C_{\eta2}$	125.12
	H	6.55
	H_α	4.42
	$H_{\beta2}$	3.12
	$H_{\beta3}$	3.58
	$H_{\delta1}$	7.36
	$H_{\varepsilon1}$	10.63
	$H_{\varepsilon3}$	7.31
	$H_{\zeta2}$	7.26
	$H_{\zeta3}$	6.78
	$H_{\eta2}$	7.16
	N	109.21
	$N_{\varepsilon1}$	131.43
A27	C	172.58
	C_α	51.06
	C_β	15.52
	H	6.96
	H_α	4.03
	H_β	0.51
A27	N	127.46
K28	C	174.83
	C_α	53.33
	C_β	35.76
	C_γ	23.59
	C_δ	29.36
	C_ε	42.03
	H	6.78
	H_α	4.48
	$H_{\beta2}$	1.31
	$H_{\beta3}$	1.71
	$H_{\gamma2/3}$	1.29
	$H_{\delta2/3}$	1.69
	$H_{\varepsilon2/3}$	2.99
	N	119.67
K29	C	177.17
	C_α	58.95
	C_β	32.26
	C_γ	24.52
	C_δ	29.63
	C_ε	41.87
	H	8.57
	H_α	3.53
	$H_{\beta2/3}$	1.67
	$H_{\gamma2/3}$	1.36
	$H_{\delta2/3}$	1.74
	$H_{\varepsilon2/3}$	3.05
	N	120.44
G30	C	174.06
	C_α	44.58
	H	9.20
	$H_{\alpha2}$	3.50

Table 7.1. *Continued.*

Residue	Atom	δ (ppm)	Residue	Atom	δ (ppm)
G30	$H_{\alpha 3}$	4.36	**Q34**	C_{β}	30.44
	N	115.60		C_{γ}	34.39
A31	C	117.08		H	8.80
	C_{α}	52.69		H_{α}	5.13
	C_{β}	19.42		$H_{\beta 2}$	1.83
	H	8.20		$H_{\beta 3}$	2.05
	H_{α}	4.42		$H_{\gamma 2/3}$	2.23
	H_{β}	1.61		$H_{\varepsilon 21}$	6.82
	N	122.88		$H_{\varepsilon 22}$	7.36
H32	C	174.75		N	116.59
	C_{α}	55.22		$N_{\varepsilon 2}$	111.33
	C_{β}	29.51	**A35**	C	176.86
	$C_{\varepsilon 1}$	135.67		C_{α}	52.66
	H	8.79		C_{β}	23.02
	H_{α}	5.15		H	8.14
	$H_{\beta 2}$	2.98		H_{α}	4.77
	$H_{\beta 3}$	3.06		H_{β}	1.40
	$H_{\varepsilon 1}$	8.37		N	124.66
	N	120.15	**S36**	C	175.33
V33	C	172.71		C_{α}	58.79
	C_{α}	58.80		C_{β}	63.49
	C_{β}	34.43		H	9.59
	$C_{\gamma 1}$	22.81		H_{α}	4.06
	$C_{\gamma 2}$	18.72		$H_{\beta 2}$	3.61
	H	8.34		$H_{\beta 3}$	4.24
	H_{α}	5.14		N	118.57
	H_{β}	2.40	**G37**	C	174.48
	$H_{\gamma 1}$	1.03		C_{α}	45.67
	$H_{\gamma 2}$	0.51		H	8.84
	N	115.77		$H_{\alpha 2}$	3.78
Q34	C	175.92		$H_{\alpha 3}$	4.19
	C_{α}	54.74		N	106.34

Table 7.1. ***Continued.***

Residue	Atom	δ (ppm)
G38	C	176.15
	C_α	43.65
	H	8.36
	$H_{\alpha 2/3}$	3.89
	N	109.10
M39	C	173.19
	C_α	52.45
	C_β	33.18
	C_γ	32.13
	H	8.61
	H_α	5.94
	$H_{\beta 2}$	1.80
	$H_{\beta 3}$	2.15
	$H_{\gamma 2/3}$	2.44
	N	120.36
P40	C	174.43
	C_α	62.64
	C_β	33.15
	C_γ	26.48
	C_δ	49.25
	H_α	4.52
	$H_{\beta 2}$	0.44
	$H_{\beta 3}$	1.02
	$H_{\gamma 2}$	0.99
	$H_{\gamma 3}$	1.35
	$H_{\delta 2/3}$	2.99
T41	C	175.14
	C_α	60.05
	C_β	71.18
	$C_{\gamma 2}$	21.09
	H	7.16
	H_α	4.94
T41	H_β	3.62
	$H_{\gamma 2}$	0.82
	N	109.34
I42	C	176.23
	C_α	62.54
	C_β	36.22
	$C_{\gamma 1}$	27.66
	$C_{\gamma 2}$	17.84
	$C_{\delta 1}$	13.44
	H	8.77
	H_α	4.29
	H_β	2.09
	$H_{\gamma 12/13}$	0.70
	$H_{\gamma 2}$	0.55
	$H_{\delta 1}$	0.76
	N	124.45
V43	C	175.57
	C_α	62.40
	C_β	32.86
	$C_{\gamma 1}$	21.99
	$C_{\gamma 2}$	20.62
	H	9.24
	H_α	4.32
	H_β	2.09
	$H_{\gamma 1}$	0.89
	$H_{\gamma 2}$	0.84
	N	126.57
E44	C	173.19
	C_α	56.07
	C_β	33.17
	C_γ	36.08
	H	7.44

Table 7.1. *Continued.*

Residue	Atom	δ (ppm)	Residue	Atom	δ (ppm)
E44	H_{α}	4.47	**G47**	C	173.61
	$H_{\beta2}$	1.98		C_{α}	47.86
	$H_{\beta3}$	2.05		H	8.99
	$H_{\gamma2/3}$	2.07		$H_{\alpha2}$	3.55
	N	117.35		$H_{\alpha3}$	3.94
V45	C	174.89		N	116.93
	C_{α}	62.08	**D48**	C	173.93
	C_{β}	34.27		C_{α}	53.17
	$C_{\gamma1}$	21.06		C_{β}	40.04
	$C_{\gamma2}$	21.09		H	8.30
	H	8.72		H_{α}	4.52
	H_{α}	4.61		$H_{\beta2}$	2.66
	H_{β}	1.89		$H_{\beta3}$	3.30
	$H_{\gamma1}$	0.82		N	126.71
	$H_{\gamma2}$	0.84	**E49**	C	175.22
	N	123.70		C_{α}	54.90
K46	C	176.31		C_{β}	32.24
	C_{α}	54.78		C_{γ}	37.03
	C_{β}	33.49		H	7.83
	C_{γ}	24.51		H_{α}	5.27
	C_{δ}	29.08		$H_{\beta2}$	1.63
	C_{ε}	41.91		$H_{\beta3}$	2.05
	H	9.07		$H_{\gamma2/3}$	2.43
	H_{α}	4.59		N	117.34
	$H_{\beta2/3}$	1.77	**V50**	C	173.78
	$H_{\gamma2}$	1.33		C_{α}	60.20
	$H_{\gamma3}$	1.67		C_{β}	34.82
	$H_{\delta2}$	1.25		$C_{\gamma1}$	20.94
	$H_{\delta3}$	1.44		$C_{\gamma2}$	21.72
	$H_{\varepsilon2}$	1.64		H	9.06
	$H_{\varepsilon3}$	3.02		H_{α}	4.61
	N	129.37		H_{β}	1.96

Table 7.1. ***Continued.***

Residue	Atom	δ (ppm)
V50	$H_{\gamma 1/2}$	0.90
	N	120.89
V51	C	175.37
	C_{α}	61.61
	C_{β}	32.63
	$C_{\gamma 1}$	20.31
	$C_{\gamma 2}$	20.79
	H	8.85
	H_{α}	4.40
	H_{β}	1.93
	$H_{\gamma 1}$	0.93
	$H_{\gamma 2}$	0.77
	N	126.23
L52	C	175.10
	C_{α}	53.20
	C_{β}	45.52
	C_{γ}	26.95
	$C_{\delta 1/2}$	25.65
	H	9.66
	$H_{\beta 2}$	0.92
	$H_{\beta 3}$	1.93
	H_{γ}	1.57
	$H_{\delta 1/2}$	0.53
	N	128.76
K53	C	175.76
	C_{α}	55.36
	C_{β}	34.44
	C_{γ}	24.45
	C_{δ}	29.38
	C_{ε}	41.91
	H	9.23
	H_{α}	5.01
K53	$H_{\beta 2}$	1.41
	$H_{\beta 3}$	1.84
	$H_{\gamma 2}$	1.24
	$H_{\gamma 3}$	1.35
	$H_{\delta 2}$	1.86
	$H_{\delta 3}$	2.08
	$H_{\varepsilon 2/3}$	2.96
	N	122.81
F54	C	176.61
	C_{α}	57.71
	C_{β}	43.55
	$C_{\delta 1/2}$	131.90
	H	8.96
	H_{α}	4.61
	$H_{\beta 2}$	2.77
	$H_{\beta 3}$	3.34
	$H_{\delta 1/2}$	7.27
	$H_{\varepsilon 1}$	7.18
	$H_{\varepsilon 2}$	7.09
	H_{ζ}	7.14
	N	128.17
G55	C	174.83
	C_{α}	45.36
	H	8.69
	$H_{\alpha 2}$	4.04
	$H_{\alpha 3}$	4.20
	N	106.54
K56	C	178.28
	C_{α}	60.29
	C_{β}	32.71
	C_{γ}	25.15
	C_{δ}	29.40

Table 7.1. *Continued.*

Residue	Atom	δ (ppm)	Residue	Atom	δ (ppm)
K56	C_{ε}	42.03	**A59**	C_{α}	178.37
	H	8.78		C_{β}	54.31
	H_{α}	3.81		H	18.45
	$H_{\beta 2/3}$	1.84		H_{α}	3.58
	$H_{\gamma 2}$	1.46		H_{β}	1.31
	$H_{\gamma 3}$	1.84		N	121.24
	$H_{\delta 2/3}$	1.69	**A60**	C	178.54
	$H_{\varepsilon 2}$	2.27		C_{α}	53.80
	$H_{\varepsilon 3}$	2.92		C_{β}	18.48
	N	122.81		H	7.27
A57	C	179.91		H_{α}	4.06
	C_{α}	54.71		H_{β}	1.41
	C_{β}	17.96		N	115.36
	H	8.57	**K61**	C	176.31
	H_{α}	4.15		C_{α}	55.21
	H_{β}	1.43		C_{β}	32.86
	N	120.28		C_{γ}	24.82
K58	C	178.734		C_{δ}	29.37
	C_{α}	57.39		C_{ε}	42.03
	C_{β}	32.08		H	7.35
	C_{γ}	25.01		H_{α}	4.29
	C_{δ}	28.07		$H_{\beta 2}$	1.83
	C_{ε}	42.06		$H_{\beta 3}$	2.03
	H	7.74		$H_{\gamma 2}$	0.83
	H_{α}	4.12		$H_{\gamma 3}$	1.51
	$H_{\beta 2/3}$	1.99		$H_{\delta 2/3}$	1.68
	$H_{\gamma 2/3}$	1.50		$H_{\varepsilon 2/3}$	2.96
	$H_{\delta 2/3}$	1.76		N	114.61
	$H_{\varepsilon 2}$	2.61	**I62**	C	174.51
	$H_{\varepsilon 3}$	2.90		C_{α}	61.46
	N	117.62		C_{β}	37.39
A59	C	178.37		$C_{\gamma 1}$	27.27

Table 7.1. ***Continued.***

Residue	Atom	δ (ppm)	Residue	Atom	δ (ppm)
I62	$C_{\gamma 2}$	18.28	**T64**	H_{β}	4.11
	$C_{\delta 1}$	12.67		$H_{\gamma 2}$	1.16
	H	7.37		N	117.61
	H_{α}	3.82	**D65**	C	175.11
	H_{β}	1.69		C_{α}	57.08
	$H_{\gamma 12}$	0.00		C_{β}	39.73
	$H_{\gamma 13}$	0.80		H	9.04
	$H_{\gamma 2}$	0.61		H_{α}	4.32
	$H_{\delta 1}$	0.00		$H_{\beta 2/3}$	2.98
	N	123.08		N	117.96
K63	C	177.99	**S66**	C	172.96
	C_{α}	54.15		C_{α}	58.02
	C_{β}	34.11		C_{β}	64.90
	C_{γ}	24.52		H	7.49
	C_{δ}	29.07		H_{α}	4.55
	C_{ε}	41.85		$H_{\beta 2/3}$	4.04
	H	8.22		N	113.04
	H_{α}	4.61	**S67**	C	174.08
	$H_{\beta 2/3}$	1.80		C_{α}	58.04
	$H_{\gamma 2}$	1.24		C_{β}	64.50
	$H_{\gamma 3}$	1.63		H	8.66
	$H_{\delta 2}$	1.23		H_{α}	5.20
	$H_{\delta 3}$	1.40		$H_{\beta 2/3}$	3.78
	$H_{\varepsilon 2}$	2.66		N	115.77
	$H_{\varepsilon 3}$	2.95	**M68**	C	174.13
	N	125.41		C_{α}	55.05
T64	C	174.48		C_{β}	37.35
	C_{α}	64.26		C_{γ}	30.78
	C_{β}	68.39		C_{ε}	16.68
	$C_{\gamma 2}$	22.47		H	8.93
	H	8.60		H_{α}	4.71
	H_{α}	3.67		$H_{\beta 2}$	1.69

Table 7.1. *Continued.*

Residue	Atom	δ (ppm)	Residue	Atom	δ (ppm)
M68	$H_{\beta3}$	1.86	**S71**	H_{α}	5.47
	$H_{\gamma2}$	2.12		$H_{\beta2}$	3.68
	$H_{\gamma3}$	2.43		$H_{\beta3}$	3.84
	H_{ε}	1.89		N	120.62
	N	124.66	**E72**	C	176.51
T69	C	173.26		C_{α}	56.92
	C_{α}	62.56		C_{β}	30.04
	C_{β}	69.89		C_{γ}	35.15
	$C_{\gamma2}$	21.55		H	7.65
	H	8.40		H_{α}	3.57
	H_{α}	4.69		$H_{\beta2/3}$	1.21
	H_{β}	3.92		$H_{\gamma2}$	1.23
	$H_{\gamma2}$	1.10		$H_{\gamma3}$	1.51
	N	122.20		N	124.65
I70	C	174.59	**S73**	C	174.12
	C_{α}	60.52		C_{α}	57.54
	C_{β}	38.67		C_{β}	64.10
	$C_{\gamma1}$	28.44		H	8.88
	$C_{\gamma2}$	19.28		H_{α}	4.38
	$C_{\delta1}$	14.98		$H_{\beta2}$	3.68
	H	9.17		$H_{\beta3}$	3.71
	H_{α}	4.62		N	118.50
	H_{β}	0.62	**A74**	C	178.07
	$H_{\gamma12}$	1.09		C_{α}	52.55
	$H_{\gamma13}$	0.62		C_{β}	19.50
	$H_{\gamma2}$	0.79		H	8.44
	$H_{\delta1}$	0.43		H_{α}	4.29
	N	130.61		H_{β}	1.32
S71	C	172.64		N	127.25
	C_{α}	57.08	**G75**	C	173.93
	C_{β}	65.83		C_{α}	45.37
	H	9.22		H	8.35

Table 7.1. ***Continued.***

Residue	Atom	δ (ppm)
G75	$H_{\alpha 2/3}$	3.88
	N	107.84
D76	C	176.25
	C_{α}	54.42
	C_{β}	41.38
	H	8.20
	H_{α}	4.55
	$H_{\beta 2/3}$	2.63
	N	120.35
A77	C	177.76
	C_{α}	52.71
	C_{β}	19.10
	H	8.18
	H_{α}	4.26
	H_{β}	1.34
	N	123.91
L78	C	177.40
	C_{α}	55.20
	C_{β}	42.13
	C_{γ}	26.87
	$C_{\delta 1}$	24.68
	$C_{\delta 2}$	23.28
	H	8.16
	$H_{\beta 2}$	1.53
	$H_{\beta 3}$	1.63
	H_{γ}	1.55
	$H_{\delta 1/2}$	0.85
	N	120.49
Q79	C	176.16
	C_{α}	55.98
	C_{β}	29.54
	C_{γ}	33.84
Q79	H	8.24
	H_{α}	4.31
	$H_{\beta 2}$	1.98
	$H_{\beta 3}$	2.11
	$H_{\gamma 2/3}$	2.34
	$H_{\varepsilon 21}$	6.84
	$H_{\varepsilon 22}$	7.54
	N	120.66
	$N_{\varepsilon 2}$	112.42
G80	C	173.27
	C_{α}	45.36
	H	8.38
	$H_{\alpha 2/3}$	3.97
	N	110.16
C81	C	178.77
	C_{α}	59.42
	C_{β}	29.11
	H	7.86
	H_{α}	4.39
	$H_{\beta 2/3}$	2.92
	N	122.74

Table 7.2. SAXS data collection. The data is shown for all four ExtJ samples. "HS" and "LS" stand for high (32 mM sodium phosphate buffer pH 6 (100 mM of final ionic strength)) and low (8 mM sodium phosphate buffer pH 6 (25 mM of final ionic strength)) salt concentrations, respectively. The protein concentration values presented refer to those of the samples selected for data analysis.

Sample	ExtJ WT	ExtJ WT + DTT	ExtJ C81* (HS)	ExtJ C81* (LS)
[Protein] (mg/mL)	7.649	14.863	14.606	7.000
Instrument	ESRF BM29	ESRF BM29	ESRF BM29	ESRF BM29
Wavelength (Å)	0.9919	0.9919	0.9919	0.9919
q-range ($Å^{-1}$)	0.01-0.50	0.01-0.50	0.01-0.50	0.01-0.50
Detector	Pilatus3 2M	Pilatus3 2M	Pilatus3 2M	Pilatus3 2M
Detector distance (m)	2.827	2.827	2.827	2.827
Beam size (μm^2)	700 x 700	700 x 700	700 x 700	700 x 700
Exposure time (s)	10 x 1	10 x 1	10 x 1	10 x 1
Temperature (K)	293	293	293	293

Table 7.3. Assignment of the NMR signals of the heme substituents and heme axial ligands of the cytochrome domains of PgcA in the reduced state. The spectra used for this assignment were acquired at 25 °C with samples prepared in 32 mM sodium phosphate buffer with NaCl (100 mM final ionic strength) at pH 7. The heme substituents are labeled according to the IUPAC-IUB nomenclature for tetrapyrroles [5].

Heme substituents	^1H δ (ppm)		
	Domain 1	Domain 2	Domain 3
2^1CH_3	3.56	3.60	3.67
3^1H	5.86	5.81	5.88
3^2CH_3	1.95	1.97	2.06
5H	9.88	9.80	9.91
7^1CH_3	3.96	4.07	3.85
8^1H	6.33	6.39	6.28
8^2CH_3	2.41	2.46	2.31
10H	9.35	9.37	9.31
12^1CH_3	3.47	3.46	3.45
15H	9.42	9.40	9.68
18^1CH_3	3.23	3.28	3.10
20H	9.09	9.21	9.21

Table 7.3. ***Continued.***

Heme axial ligands	^{1}H δ (ppm)		
	Domain 1 (Met^{301})	Domain 2 (Met^{392})	Domain 3
H_α	2.41	2.41	–
H_β	−0.23	−0.24	–
$H_{\beta'}$	−0.44	−0.41	–
H_γ	−2.76	−2.67	–
$H_{\gamma'}$	−3.12	−3.26	–
εCH_3	−3.52	−3.32	–

Table 7.4. Assignment of the NMR signals of the heme substituents and heme axial ligands of the cytochrome domains of PgcA in the oxidized state. The spectra used for this assignment were acquired at 25 °C with samples prepared in 32 mM sodium phosphate buffer with NaCl (100 mM final ionic strength) at pH 7. The heme substituents are labeled according to the IUPAC-IUB nomenclature for tetrapyrroles [5].

Heme substituents	Domain 1		Domain 2		Domain 3	
	^{13}C δ	^{1}H δ	^{13}C δ	^{1}H δ	^{13}C δ	^{1}H δ
2^{1}CH$_3$	−10.46	5.00	−21.36	9.09	−32.66	13.19
7^{1}CH$_3$	−60.18	35.18	−48.59	27.45	−26.26	16.04
12^{1}CH$_3$	−31.47	14.48	−43.72	20.83	−41.55	17.78
13^{1}CH$_2$	−17.50	2.05 11.27	−7.28	3.04 4.27	7.57	−2.79 0.05
13^{2}CH$_2$	91.21	−1.96 1.27	74.59	−0.74 −0.36	50.77	0.45 1.12
17^{1}CH$_2$	−19.73	14.10	−16.14	3.42 16.08	−2.68	3.29 5.29
17^{2}CH$_2$	109.80	0.14 4.61	98.11	−0.67 2.77	74.37	1.13 1.55
18^{1}CH$_3$	−52.95	33.44	−62.65	35.83	−57.32	26.49

Heme axial ligands		Domain 1		Domain 2		Domain 3	
		^{13}C δ	^{1}H δ	^{13}C δ	^{1}H δ	^{13}C δ	^{1}H δ
His^{274}	CH_α	74.50	9.48	–	–	–	–
	$CH_{\beta2}$	19.23	8.78	–	–	–	–
	$CH_{\beta3}$		18.66		–		–
Met^{301}	εCH_3	9.80	−17.93	–	–	–	–

Table 7.4. *Continued.*

Heme axial ligands		Domain 1		Domain 2		Domain 3	
		$^{13}C\ \delta$	$^{1}H\ \delta$	$^{13}C\ \delta$	$^{1}H\ \delta$	$^{13}C\ \delta$	$^{1}H\ \delta$
His^{365}	CH_α	–	–	93.19	10.52	–	–
	$CH_{\beta2}$	–	–	16.26	9.24	–	–
	$CH_{\beta3}$		–		21.03		–
Met^{392}	εCH_3	–	–	–	–13.72	–	–
His$^{459/491}$	CH_α	–	–	–	–	90.58	10.11
	$CH_{\beta2}$	–	–	–	–	20.40	9.54
	$CH_{\beta3}$		–		–		15.58
His$^{459/491}$	CH_α	–	–	–	–	82.08	8.80
	$CH_{\beta2}$	–	–	–	–	22.94	7.02
	$CH_{\beta3}$		–		–		12.92

Table 7.5. Redox-active windows and apparent midpoint reduction potential values of cytochromes from *G. sulfurreducens*. The redox-active windows of each cytochrome were determined from potentiometric redox curves considering 1-99% range for protein reduction/oxidation. E_{app} stands for "apparent midpoint reduction potential", which is equal to the point at which the oxidized and reduced fractions are equal. The values used for OmcS in Figure 4.16 are those of the monomer form.

Protein	Redox-active window (mV)	E_{app} (mV)
CbcL [6]	330	–194
ImcH [7]	350	–197
PpcA [8]	285	–117
PpcB [8]	270	–137
PpcC [9]	265	–143
PpcD [10]	275	–132
PpcE [10]	280	–134
PgcA D1	230	–48
PgcA D2	235	–65
PgcA D3	240	–106
OmcS (monomer) [11]	320	–212
OmcS (nanowire) [12]	360	–130
OmcZ [13]	405	–220

Table 7.6. Sequences of the primers used to produce the GSU0105 putative lysine axial ligand mutants. The substitution sites are underlined. The primers forward (fw) and reverse (rv) were purchased from Invitrogen.

Primer	DNA sequence (5' – 3')
GSU0105_K33H_fw	ATGCAACGTCTGTCACAGTCATAACCCGGCCATGGTAAAA
GSU0105_K33H_rv	TTTTACCATGGCCGGGTTATGACTGTGACAGACGTTGCAT
GSU0105_K33M_fw	TCGAATGCAACGTCTGTCACAGTATGAACCCGGCCATGG
GSU0105_K33M_rv	CCATGGCCGGGTTCATACTGTGACAGACGTTGCATTCGA
GSU0105_K62H_fw	CGGCGAACGCCTCATGGGACATGGGCAGCCTA
GSU0105_K62H_rv	TAGGCTGCCCATGTCCCATGAGGCGTTCGCCG
GSU0105_K62M_fw	CGAACGCCTCATGGGAATGGGGCAGCC
GSU0105_K62M_rv	GGCTGCCCCATTCCCATGAGGCGTTCG
GSU0105_K66H_fw	GGAAAGGGGCAGCCTCATGACAAGGCGTCGCTG
GSU0105_K66H_rv	CAGCGACGCCTTGTCATGAGGCTGCCCCTTTCC
GSU0105_K66M_fw	GAAAGGGGCAGCCTATGGACAAGGCGTCG
GSU0105_K66M_rv	CGACGCCTTGTCCATAGGCTGCCCCTTTC

7.3. MD simulations – Additional plots

The 250 ns MD simulation of the AlphaFold model of the ExtJ homodimer showed stable physical parameters across the entire calculation (Figure 7.11), thus ensuring its reliability.

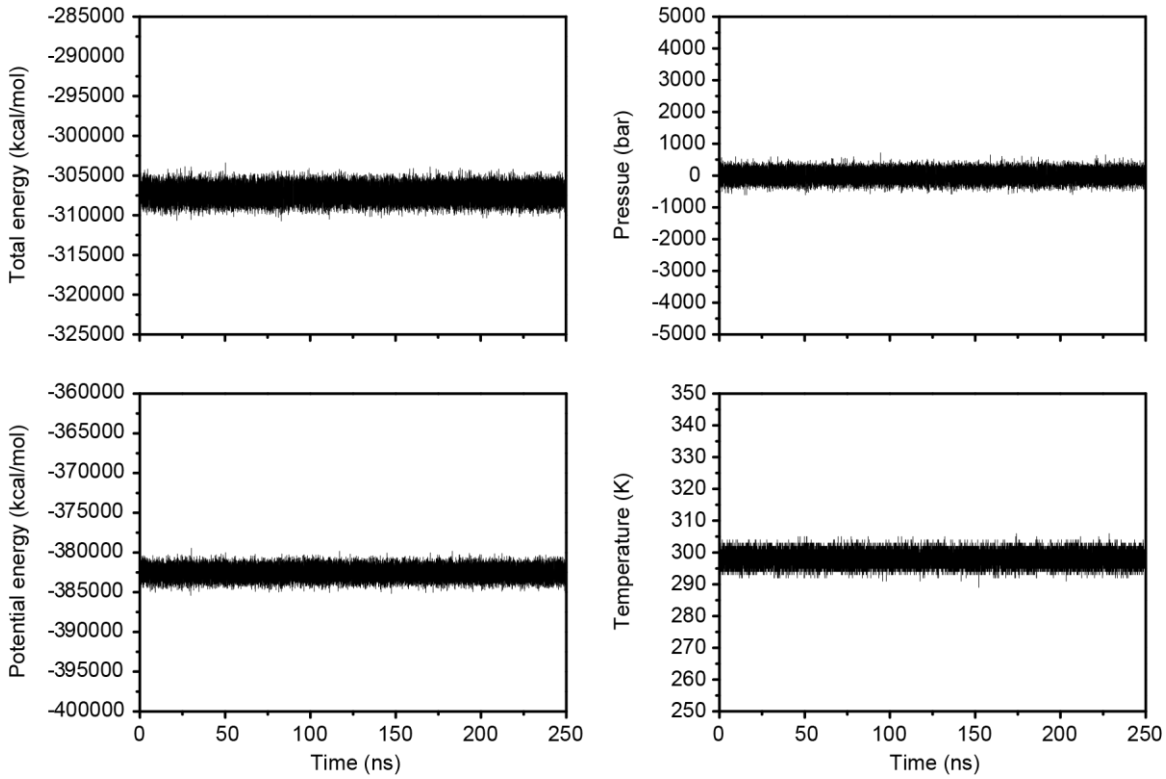


Figure 7.11. Physical parameter plots for the MD simulation of the ExtJ homodimer. The variation of the total and potential energies, pressure, and temperature of the system over the 250 ns MD simulation is shown. The data was analyzed with GROMACS 2020 [14, 15].

7.4. PyMol script for insertion of *c*-type hemes in AlphaFold models

The AlphaFold models of all cytochromes presented in this Thesis were enriched with *c*-type hemes using local alignments and libraries of this cofactor in PyMol [16]. These libraries consist of repeated versions of a PDB file containing the *c*-type heme group of the crystal structure of the OmcF monoheme cytochrome (PDB: 3CU4 [17]), together with the corresponding cysteine residues for the covalent binding of the heme vinyl groups, the proximal axial histidine for Fe coordination and the two remaining residues of the CXXCH heme-binding motif (Cys^{35}-Ala^{36}-Gly^{37}-Cys^{38}-His^{39}). This file is repeated as many times as the number of heme binding motifs of each target protein model.

The PyMol script conceived for this purpose is shown below, using a pentaheme cytochrome with the hypothetical heme binding motifs C^{25}-X-X-C^{28}-H^{29}, C^{57}-X-X-C^{60}-H^{61}, C^{89}-X-X-C^{92}-H^{93}, C^{125}-X-X-C^{128}-H^{129} and C^{170}-X-X-C^{173}-H^{174}:

```
load model.pdb
load heme1.pdb
load heme2.pdb
load heme3.pdb
load heme4.pdb
load heme5.pdb
#loads the AlphaFold model and the corresponding number of c-type heme groups

sele region1, model and resi 25+28+29
align heme1, region1
sele region2, model and resi 57+60+61
align heme2, region2
sele region3, model and resi 89+92+93
align heme3, region3
sele region4, model and resi 125+128+129
align heme4, region4
sele region5, model and resi 170+173+174
align heme5, region5
#aligns the c-type heme groups with the corresponding residues of the CXXCH motifs

remove (resi 35-39 & heme1)
remove (resi 35-39 & heme2)
remove (resi 35-39 & heme3)
remove (resi 35-39 & heme4)
remove (resi 35-39 & heme5)
#removes the residues of the CXXCH motif of the reference OmcF PDB files
```

cd C:\Users\folder
save model_hemes.pdb, all
#allows the user to save the enriched AlphaFold model in the desired folder

As shown, this script requires a pre-analysis of the protein sequence and AlphaFold model, so that the cysteine and histidine residues of the canonical and non-canonical heme binding motifs can be identified and indicated in the script. Upon alignment, the residues of the CXXCH motif of the OmcF cytochrome are removed and the PDB file of the target AlphaFold model enriched with *c*-type hemes is saved on a desired folder.

Ideally, following this step, side-chain packing optimization and all-atom energy minimization protocols should have been employed for model refinement, however, considering the nature of our analysis, the models obtained after application of the PyMol script were of sufficient quality.

7.5. References

[1] J. Jumper, R. Evans, A. Pritzel, T. Green, M. Figurnov, O. Ronneberger, K. Tunyasuvunakool, R. Bates, A. Žídek, A. Potapenko, A. Bridgland, C. Meyer, S.A.A. Kohl, A.J. Ballard, A. Cowie, B. Romera-Paredes, S. Nikolov, R. Jain, J. Adler, T. Back, S. Petersen, D. Reiman, E. Clancy, M. Zielinski, M. Steinegger, M. Pacholska, T. Berghammer, S. Bodenstein, D. Silver, O. Vinyals, A.W. Senior, K. Kavukcuoglu, P. Kohli, D. Hassabis, Highly accurate protein structure prediction with AlphaFold, Nature, 596 (2021) 583-589.

[2] M. Mirdita, K. Schütze, Y. Moriwaki, L. Heo, S. Ovchinnikov, M. Steinegger, ColabFold: Making protein folding accessible to all, Nat. Methods, 19 (2022) 679-682.

[3] M. Varadi, S. Anyango, M. Deshpande, S. Nair, C. Natassia, G. Yordanova, D. Yuan, O. Stroe, G. Wood, A. Laydon, A. Žídek, T. Green, K. Tunyasuvunakool, S. Petersen, J. Jumper, E. Clancy, R. Green, A. Vora, M. Lutfi, M. Figurnov, A. Cowie, N. Hobbs, P. Kohli, G. Kleywegt, E. Birney, D. Hassabis, S. Velankar, AlphaFold protein structure database: Massively expanding the structural coverage of protein-sequence space with high-accuracy models, Nucleic Acids Res., 50 (2022) D439-D444.

[4] Z. Lin, H. Akin, R. Rao, B. Hie, Z. Zhu, W. Lu, N. Smetanin, R. Verkuil, O. Kabeli, Y. Shmueli, A. dos Santos Costa, M. Fazel-Zarandi, T. Sercu, S. Candido, A. Rives, Evolutionary-scale prediction of atomic-level protein structure with a language model, Science, 379 (2023) 1123-1130.

[5] G.P. Moss, Nomenclature of tetrapyrroles: Recommendations of the 1986 IUPAC-IUB joint commission on biochemical nomenclature (JCBN), Eur. J. Biochem., 178 (1988) 277-328.

[6] J.M.A. Antunes, M.A. Silva, C.A. Salgueiro, L. Morgado, Electron flow from the inner membrane towards the cell exterior in *Geobacter sulfurreducens*: Biochemical characterization of cytochrome CbcL, Front. Microbiol., 13 (2022) 898015.

[7] A.I. Pimenta, C.M. Paquete, L. Morgado, M.J. Edwards, T.A. Clarke, C.A. Salgueiro, I.A.C. Pereira, A.G. Duarte, Characterization of the inner membrane cytochrome ImcH from *Geobacter* reveals its importance for extracellular electron transfer and energy conservation, Protein Sci., 32 (2023) e4796.

[8] L. Morgado, M. Bruix, V. Orshonsky, Y.Y. Londer, N.E. Duke, X. Yang, P.R. Pokkuluri, M. Schiffer, C.A. Salgueiro, Structural insights into the modulation of the redox properties of two *Geobacter sulfurreducens* homologous triheme cytochromes, Biochim. Biophys. Acta - Bioenergetics, 1777 (2008) 1157-1165.

[9] T.C. Santos, M.A. Silva, L. Morgado, J.M. Dantas, C.A. Salgueiro, Diving into the redox properties of *Geobacter sulfurreducens* cytochromes: A model for extracellular electron transfer, Dalton Trans., 44 (2015) 9335-9344.

[10] L. Morgado, M. Bruix, M. Pessanha, Y.Y. Londer, C.A. Salgueiro, Thermodynamic characterization of a triheme cytochrome family from *Geobacter sulfurreducens* reveals mechanistic and functional diversity, Biophys. J., 99 (2010) 293-301.

[11] X. Qian, T. Mester, L. Morgado, T. Arakawa, M.L. Sharma, K. Inoue, C. Joseph, C.A. Salgueiro, M.J. Maroney, D.R. Lovley, Biochemical characterization of purified OmcS, a *c*-type cytochrome required

for insoluble Fe(III) reduction in *Geobacter sulfurreducens*, Biochim. Biophys. Acta, 1807 (2011) 404-412.

[12] P.C. Portela, C.C. Shipps, C. Shen, V. Srikanth, C.A. Salgueiro, N.S. Malvankar, Widespread extracellular electron transfer pathways for charging microbial cytochrome OmcS nanowires via periplasmic cytochromes PpcABCDE, Nat. Commun., 15 (2024) 2434.

[13] K. Inoue, X. Qian, L. Morgado, B.C. Kim, T. Mester, M. Izallalen, C.A. Salgueiro, D.R. Lovley, Purification and characterization of OmcZ, an outer-surface, octaheme *c*-type cytochrome essential for optimal current production by *Geobacter sulfurreducens*, Appl. Environ. Microbiol., 76 (2010) 3999-4007.

[14] D. Van Der Spoel, E. Lindahl, B. Hess, G. Groenhof, A.E. Mark, H.J.C. Berendsen, GROMACS: Fast, flexible, and free, J. Comput. Chem., 26 (2005) 1701-1718.

[15] M.J. Abraham, T. Murtola, R. Schulz, S. Páll, J.C. Smith, B. Hess, E. Lindahl, GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers, SoftwareX, 1-2 (2015) 19-25.

[16] L.L.C. Schrödinger, The PyMol molecular graphics system, Version 2.5.2, 2021.

[17] P.R. Pokkuluri, Y.Y. Londer, S.J. Wood, N.E. Duke, L. Morgado, C.A. Salgueiro, M. Schiffer, Outer membrane cytochrome *c*, OmcF, from *Geobacter sulfurreducens*: High structural similarity to an algal cytochrome c_6, Proteins, 74 (2009) 266-270.

2024

Tomás Monteiro Fernandes

Structural and functional insigths on the electrifying pathways of *Geobacter sulfurreducens*