THE STRUCTURE AND FUNCTION OF EXTRACELLULAR REDOX ENZYMES

Benjamin William Nash

PhD Thesis

University of East Anglia

School of Biological Sciences

August 2025

ABSTRACT

Increasing numbers of microorganisms are now understood to conserve energy via the coupling of intracellular oxidation of organic compounds with the reduction of diverse extracellular substrates. Such extracellular electron transfer (EET) processes represent exciting opportunities for biotechnological innovations including bioremediation, biosensing, biomining and microbial electrosynthesis, but poor understanding of the underlying molecular mechanisms limits their application. In this work, the structure and properties of two novel protein systems that participate in EET are described.

The outer membrane *c*-type cytochrome PgcA from *Geobacter sulfurreducens* is important for EET in that organism and here is shown to contain two unique functionalities. It's c-terminus is revealed to be a redox shuttle which contains three heme cofactors separated into discrete domains but tethered together by flexible repeat motifs and likely able to transfer electrons fifteen nanometres. It's N-terminus is revealed to be a conserved domain that exhibits novel autoproteolytic activity for which comprehensive characterisation is described.

The *dmsEFABGH* gene cluster has been shown to be essential for the reduction of dimethyl sulfoxide by *Shewanella oneidensis* but the structural basis for this has not been determined. In this work the structure of the DmsEFAB tetramer, solved by cryogenic electron microscopy, is described. DmsEFAB forms a ~120 Å cofactor chain, insulated from the membrane by the DmsF porin, including ten *c*-type hemes, four Fe_4S_4 clusters and a molybdopterin active site that ultimately catalyses the two electron reduction of dimethyl sulfoxide.

These results provide unique and insightful contributions not only to the field of extracellular electron transfer, but also to redox protein structural biology in general.

ACKNOWLEDGEMENTS

I would like to thank first my supervisor Prof. Tom Clarke for his patient supervision, for giving me the opportunity to become a researcher and for believing in my abilities from before I did. I would also like to thank Prof. Julea Butt for her insightful contributions to my project and my development as a scientist. I'd like to thank Drs David Swainsbury, Jessica van Wonderen, Marcus Edwards, Colin Lockwood, Joshua Burton and Alejandro Morales-Florez for their support and advice. I would like to thank my other friends and family for the support and inspiration they have provided me over the course of my PhD, without which, its completion would not have been possible. Finally, I would like to thank the BBSRC for providing the funding that supported me and my research throughout the last four years.

Work presented in this thesis has also been published in following article: Nash, B. W.; Fernandes, T. M.; J Burton, J. A.; Morgado, L.; van Wonderen, J. H.; Svistunenko, D. A.; Edwards, M. J.; Salgueiro, C. A.; Butt, J. N.; Clarke, T. A.; Thomas Clarke, C. A. Tethered Heme Domains in a Triheme Cytochrome Allow for Increased Electron Transport Distances. *Protein Science* 2024, **33** (11), e5200.

Table of contents

List of figures

List of tables

List of abbreviations, symbols and units

%	percent
% v/v	percentage (volume per volume)
% w/v	percentage (weight per volume)
~	approximately
°	degrees
2D	two dimensional
3D	three dimensional
Å	Angstrom
Ala	alanine
Arg	arginine
Asp	aspartic acid
ATP	adenosine triphosphate
C	Celsius
c(S)	sedimentation coefficient distribution
$CC_{1/2}$	correlation coefficient
CET	cryogenic electron tomography
cm	centimetres
cryoEM	cryogenic electron microscopy
CTF	contrast transfer function
Cys	cysteine
cyt-*c*	Cytochrome c
Da	Dalton
D_{max}	maximum distance
DMRs	dissimilatory metal reducers
DMSO	dimethyl sulfoxide
DMSO	dimethyl sulfide
DMSOR	dimethyl sulfoxide reductase
DNA	deoxyribonucleic acid
DSS	sodium trimethylsilylpropanesulfonate
EER	electro event representation
EET	extracellular electron transfer
E_m	midpoint potential
EOM	ensemble optimisation method
EPR	electron paramagnetic resonance
ETC	electron transport chain
E-waste	electronic waste
f/f_0	frictional ratio
FAD	flavin adenine dinucleotide
FMN	flavin mononucleotide
FSC	Fourier shell correlation
g	grams
G	standard earth gravitational strength
Glu	glutamic acid

HALS	highly anisotropic or axial low spin
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
His	histidine
HPLC	high-performance liquid chromatography
Hz	hertz
IPTG	Isopropyl β-d-1-thiogalactopyranoside
Iso	isoleucine
ITO	indium tin oxide
K	Kelvin
kb	kilobases
k_{cat}	catalytic constant
K_d	dissociation constant
kDa	kilodaltons
kx	times magnification
LB	Lysogeny broth
LC-MS	liquid chromatography-mass spectrometry
LDAO	Lauryldimethylamine oxide
Leu	leucine
M	molar
mA	milliamps
MCD	magnetic circular dichroism
MES	microbial electrosynthesis
Met	methionine
MFC	microbial fuel cell
MGD	molybdenum guanine dinucleotide
mL	millilitres
mm	millimetres
MQ	menaquinone
mRNA	messenger RNA
mV	millivolts
MV	methyl viologen
NADH	nicotinamide adenine dinucleotide
nm	nanometre
NMR	nuclear magnetic resonance
NTA	nitrilotriacetic acid
ORF	open reading frame
PCC	porin cytochrome complex
PCR	polymerase chain reaction
PDB	Protein Data Bank
PEG	polyethylene glycol
PFV	protein film voltammetry
PGM	platinum group metal
PMF	proton motive force
ppm	parts per million
Pro	proline
PSI	pounds per square inch

R_g	radius of gyration
RMSD	root mean squared deviation
RNA	ribonucleic acid
RO	reverse osmosis
ROS	reactive oxygen species
R_{pim}	precision-indicating merging R factor
RPM	rotations per minute
S	siemens
s	seconds
$S_{20,W}$	sedimentation coefficient at 20 degrees Celsius in pure water
SAXS	small-angle X-ray scattering
SDS	sodium dodecyl sulphate
SDS-PAGE	sodium dodecyl sulphate polyacrylamide gel electrophoresis
SEC	size exclusion chromatography
SEC-SAXS	size exclusion chromatography coupled small-angle X-ray scattering
Ser	serine
Ser	serine
SHE	standard hydrogen electrode
TAT	twin arginine translocation
TCA	tricarboxylic acid
TEM	transmission electron microscopy
Thr	threonine
Tris	tris(hydroxymethyl)aminomethane
UV	ultraviolet
V	volts
Val	valine
W	watts
WT	wild-type
x	times
α	alpha
β	beta
Δ	change in or deletion
μ	micro
σ	sigma
λ	wavelength

Chapter 1: General Introduction

Redox enzymes and extracellular electron transfer

Living organisms exist in a state that is thermodynamically unstable and therefore must expend energy to maintain their existence. The fundamental nature of this statement necessitates the development of systems for energy extraction and usage as a core component within living organisms.

As a result of this, redox enzymes have evolved and become ubiquitous across cellular organisms. Redox enzymes are one answer that evolution has crafted, to the question how can cells efficiently generate energy or extract it from their environment, and repurpose it usefully.

An excellent example of a fundamental redox protein is eukaryotic mitochondrial cytochrome c (cyt-c), which is small and contains only a single cofactor, that exists in either a reduced or oxidised state. This protein is ~100 amino acids in length, water soluble and incorporates a single C-type heme within its structure (Figure 1.1), covalently attached to the protein through bonds established between vinyl groups of the porphyrin ring, and cysteine thiol groups[1]. Its structure is globular and dominated by alpha helices and loops that encapsulate the *c*-heme cofactor[1].

The presence of an iron atom chelated by the porphyrin ring confers its ability to perform redox reactions, as the iron centre can exchange electrons with partners by switching between ferrous (+2) and ferric (+3) states. The chemical environment of the heme group is dictated by the structure of the protein through its interatomic interactions with nearby protein

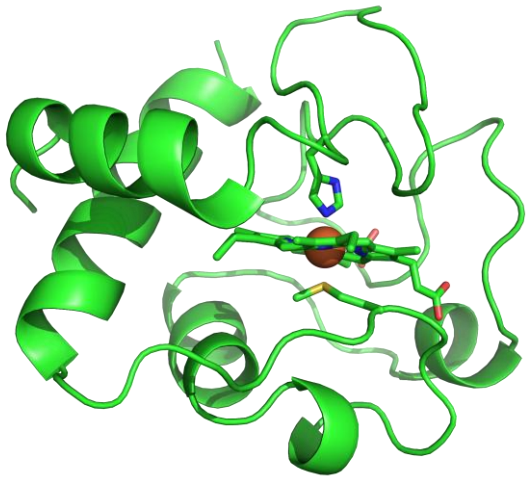


Figure 1.1: Crystal structure of *Equus caballus* heart tissue cyt-c (PDB-1HRC), depicted with heme coordinating residues H18 and M80 as sticks.

atoms and solvent atoms. Since the structure adopted by the protein is determined by its amino acid sequence, as laid out in Anfinsen's dogma[2], the chemical environment of the heme is thus ultimately dictated by the protein amino acid sequence. Of particular importance are the two residues that coordinate the heme iron. Usually in *c*-type cytochromes, the proximal ligation position is usually occupied by the side chain of a histidine located within the heme attachment consensus sequence, Cys-Xxx-Xxx-Cys-His, although this consensus can vary[3]. The distal site is most often occupied by side chains of histidine or methionine residues, however this site displays more variability than the proximal site, cysteine, water, N-terminal main chain amide and uncoordinated have all been described[4]. In eukaryotic cyt-*c*, His-Met coordination of the heme is most often seen[5]. Through modifying the amino acid sequence, evolution is able to vary the tendency of the protein to accept or donate electrons, which is described by its reduction potential.

Additionally, modifying the amino acid sequence could allow for the introduction of a binding site(s) at which the redox protein can interact with a second redox protein, making inter-protein electron transfer a possibility.

If two cyt-*c* variant proteins are considered, that have differing reduction potentials then the transfer of electrons from the lower potential variant to the higher potential variant, will be exergonic. Incorporating a mechanism that traps the difference in Gibbs free energy released by the electron transfer as a usable energy currency, converts the system into what is referred to as an electron transport chain (ETC). As long as molecules which will reduce the first cyt-*c* in the chain and molecules that will oxidise the last cyt-*c* can be obtained, then this system can operate in perpetuity, allowing a cell to indefinitely prolong its thermodynamically unstable existence.

In practice, compounds that can be oxidised by a redox protein as simple as cyt-*c* are rare in the environment. Subsequently organisms have developed complex catabolic enzyme pathways that extract low reduction potential electrons from diverse donor molecules including carbohydrates, proteins, lipids, hydrocarbons and sulfur compounds.

Furthermore, coupling the energy released by electron exchange between two redox centres with the synthesis of an energy currency like adenosine triphosphate, is mechanistically challenging. To solve this problem, evolution has produced sophisticated molecular machines, like the cytochrome bc_1 complex and mitochondrial complex I. Such molecular machines are embedded within biological membranes and use the free energy generated by the electron transfer event, through the pumping of ions, to generate an electrochemical gradient across the membrane[6]. The introduction of biological lipid membranes into the system is a crucial innovation. Lipid membranes are inherently electrically insulating and therefore provides a means by which the energy can be stored and also allow for its distribution into multiple different downstream systems.

Obtaining reductants in sufficient quantity to operate the electron transport chain perpetually can pose a challenge. Many organisms now generate their own low reduction potential electrons artificially by extracting them from water, and exciting them, photosynthetically to a higher energy state. This process has been so successful in fact, it has made molecular oxygen ubiquitous in the atmosphere of the earth for billions of years[7].

O_2 is not only now highly prevalent and water soluble, but the value of its reduction potential is large and positive, making it an excellent terminal electron acceptor since a large contribution to the electrochemical gradient can be built up per electron terminating in the O_2 half reaction ($O_2 + 4e^- + 4H_2O \rightarrow 2H_2O$). For these reasons it is the dominant respiratory substrate for the most complex domain of life, eukaryotes.

Environments do persist however in which light energy is unavailable to would-be photosynthesisers. Furthermore, cellular life is thought to predate oxygenic photosynthesisers perhaps by some millions of years therefore these ancient microorganisms must have utilised alternative terminal electron acceptors to oxygen[8].

In instances where O_2 is not available as an electron acceptor, nitrate can serve as a terminal electron acceptor, for which specialised nitrate reducing redox enzymes exist. Like O_2, nitrate is a small and water soluble molecule which can permeate membranes at an appreciable rate. Thus nitrate reductase enzymes are localised intracellularly[9].

Another alternative to O_2 is sulfate. Sulfate reducing bacteria are abundant and diverse, playing an important role in many anaerobic environments oxidising organic molecules, and reducing sulfate to H_2S. Similarly to nitrate and O_2, sulfate is a small and mobile molecule, and correspondingly its reductases are also intracellular[10,11].

In anaerobic marine and aquatic environments, oxidised metal compounds could potentially also serve as a terminal electron acceptor, for example Fe^{III} or Mn^{IV} compounds. These are typically present as various oxide minerals like hematite, that display extremely low solubility at circumneutral pH but do poses reduction potentials large enough to support cellular energy conservation. Their low solubility poses an additional barrier to their use though, as the bacterial cell envelope, be that the outer membrane of Gram negatives or the thick peptidoglycan layer of Gram positives, will physically block their interaction with inner membrane associated reductase enzymes.

Subsequently for a bacterium to exploit insoluble oxidised compounds in its extracellular environment as terminal electron acceptors, it must possess additional components in its electron transport chain. For a Gram negative, electron transporting species are required through the periplasm, through the insulating outer membrane and in the extracellular space itself to bridge the gap between the cell surface and a mineral deposit. For a Gram positive, it requires electron transporters that traverse the thick peptidoglycan layer and are capable of mineral reduction on its exterior face.

Recent research has identified increasing numbers of bacterial isolates that in fact contain one or even multiple respiratory systems complete with every component described above[12]. These bacteria, referred to as dissimilatory metal

reducers (DMRs), are able to exploit networks of redox proteins that traverse their cell envelope, flowing electrons from the quinol pool ultimately onto insoluble terminal electron acceptors in their extracellular environment. Thus this mechanism of respiration is referred to as extracellular electron transfer (EET), and those species that participate in it as electrogenic bacteria.

The ability of electrogenic bacteria to exchange electrons directly with their environment, has many potential applications and subsequently the study of EET has garnered the interest of diverse groups of researchers.

Applications of extracellular electron transfer

The most obvious application of electrogenic microorganisms is the generation of electrical power in microbial fuel cells (MFCs)[13]. MFCs contain two electrodes separated by a membrane which is permeable to protons.

The anodic electrode is kept anaerobic, and there electrogenic microbes grow, oxidising organic material. This generates low potential electrons which the organisms convert into higher potential electrons, harvesting the energy for cellular processes. The higher potential electrons are then discharged, across the cell envelope onto the anodic electrode itself, rather than soluble compounds or minerals.

The anodic electrode is connected by wiring to the cathodic electrode which is under aerobic conditions. There, oxygen is reduced to water, consuming protons and electrons provided by the cathode, and a potential difference is generated across the cell that can be used to do work.

MFCs are an attractive method of power generation for several reasons, the first being that microorganisms themselves are self-replicating and therefore can be a cheap resource to produce. Secondly, if the organics supplied to the anodic electrode are renewable, the MFC serves as a source of renewable energy. Furthermore, waste organics can often be used that otherwise would require expensive treatment before disposal, making the MFC doubly profitable.

In practice, MFCs have not been adopted widely as a source of renewable electricity for several reasons[13,14]. The magnitude of power output that MFC designs have yielded is insufficient to make them competitive with traditional methods of power generation. Attempts have been, and continue to be, made to optimise designs with the objective of improved power output, but anodic MFC cultures are usually comprised of complex bacterial consortia making this challenging[15]. These bacterial also typically form dense biofilms on the anode, which are harder to study than planktonic bacteria as obtaining sufficient biomass for experimentation can be unfeasible.

A related application of electrogenic bacteria, is in microbial electrosynthesis (MES)[16]. MES also involves a two electrode cell, but unlike in MFCs, in MES a potential difference is actively applied across the cell, reversing the flow of electrons driving them into the bacteria.

This strategy could be used to drive carbon capture, by providing a flow of electrons to formate dehydrogenase enzymes. Alternatively, it could be used to drive catalysis of reductive chemistry via enzymes that yield user specified and value added chemicals.

MES too, is difficult to implement for economic reasons. Producing a MES system that is large enough and efficient enough to compete with plant based methods of carbon fixation, or industrial synthesis of chemicals, is currently challenging[17].

Proteins as a group, are uniquely capable of binding and responding to a very large array of different molecules, subsequently biology has enormous potential as a source for sensor technology. MFC-based biosensors are one example of a biosensor technology that exploits the electrogenic properties of bacteria[18].

MFC biosensors incorporate a microbial fuel cell into a device, but not for the purposes of extracting usable power. Because the MFC power output is a function of the metabolic activity of the electrogenic organisms it contains, measuring the power output allows for measurement of parameters that influence the metabolic rate. For example, the concentration of dissolved organics that the electrogenic bacteria use as their electron donor, is correlated

with their metabolic activity and therefore with the power of the MFC. Thus, the device allows for the quantification of dissolved organic material, and it is in this purpose that these system sees most use however it is not only the application.

Whilst increasing the availability of electron donors increases the metabolic rate of electrogenic bacteria, the reverse statement is true for increasing the concentration of toxic compounds, they decrease metabolic activity. Inverse MFC biosensor devices have been described that respond to addition of heavy metals, antibiotics and polychlorinated organic toxins by exhibiting a measurable drop in power output[19–21].

MFC biosensors can boast remarkable durability over long periods, making them an excellent choice in applications for which maintenance or battery replacement would be prohibitive. Furthermore, they can provide measurements almost in real time, which is in stark contrast to most traditional methods of measuring dissolved organic material. For these reasons MFC biosensors are an exciting and promising technology with potential applications in multiple sectors.

Bio-extraction of metals is another potential application of electrogenic microorganisms that has become an increasingly popular topic of research. This strategy seeks to exploit the change in solubility that some metals undergo as a result of electron exchange with microorganisms. Historically, this has been applied mostly in the form of oxidative mobilisation of copper and/or gold by members of the acidothibacillia class[22].

More recently, methods for recovery of the platinum group metals (PGMs) and other rare earth metals from waste streams have been described[23,24]. In these methods, metals are extracted from scrap by acidification, and remineralised by microorganisms. This process may harbour potential as a strategy for isolating PGMs selectively which is challenging by physicochemical methods. The cost associated with energy and solvent usage, can make recovery of PGMs from E-waste prohibitively expensive by physiochemically but a microbial solution to this problem might not suffer those drawbacks.

Recent research has focused on the model organism *Shewanella oneidensis* and its hydrogenase enzymes and multiheme *c*-type cytochrome proteins as nucleation sites for the formation of PGM particles[25]. Under certain conditions the formation of PGM nanoparticles on *S. oneidensis* cells has been observed which could perform catalytic reduction of azodyes and 4-nitrophenol[26]. This area of research is in its infancy but has the potential to provide sustainable methods for extracting valuable metals and employing them in biocatalysis.

The bioremediating properties of electrogenic bacteria, can similarly arise from the change in solubility that heavy metals undergo upon reduction by microorganisms. Soluble toxic metal species like U^{VI} or Cr^{VI} can be reduced by DMRs, forming species with reduced solubility which precipitate[27]. Once precipitated, although not removed from the environment, they are less mobile and therefore less likely to cause severe ecological damage.

Some electrogenic bacteria like members *Shewanella* and *Geobacter* genera are able to degrade persistent organic pollutants by exploiting them as respiratory electron acceptors. Good examples of this process are the degradation of azodyes and recalcitrant aromatics by members of those genera respectively[28,29].

Bioremediation by electrogenic bacteria is an important process environmentally, although it occurs in a mostly passive form. This is because once disseminated in the environment, concerted efforts to actively recover pollutants in remediation is logistically and economically difficult.

Developing an understanding of the underlying mechanisms of extracellular electron transfer is likely to be a prerequisite to successful adoption of any EET biotechnologies. Subsequently great effort had been devoted to the study of these processes in detail.

Mechanisms of extracellular electron transfer I: *Shewanella oneidensis* MR-1

Shewanella oneidensis strain MR-1 was isolated from aquatic sediments of Lake Oneida (USA) in 1988 by *Myers and Nealson* who demonstrated for the first time,

a microorganism in pure culture, which coupled its growth with the reduction of manganese oxide under anaerobic conditions[30]. MR-1 has become the primary model organism for the study of extracellular electron transfer due to its being the first isolated, but also its relative ease of manipulation, fast growth and the high quantities of biomass it can yield.

Seminal work by *Hunt et al* suggested that MR-1 obtains energy only by substrate-level phosphorylation during anaerobic respiration[31]. They prepared a ΔATP-Synthase strain which actually grew faster than the wild-type (WT) under

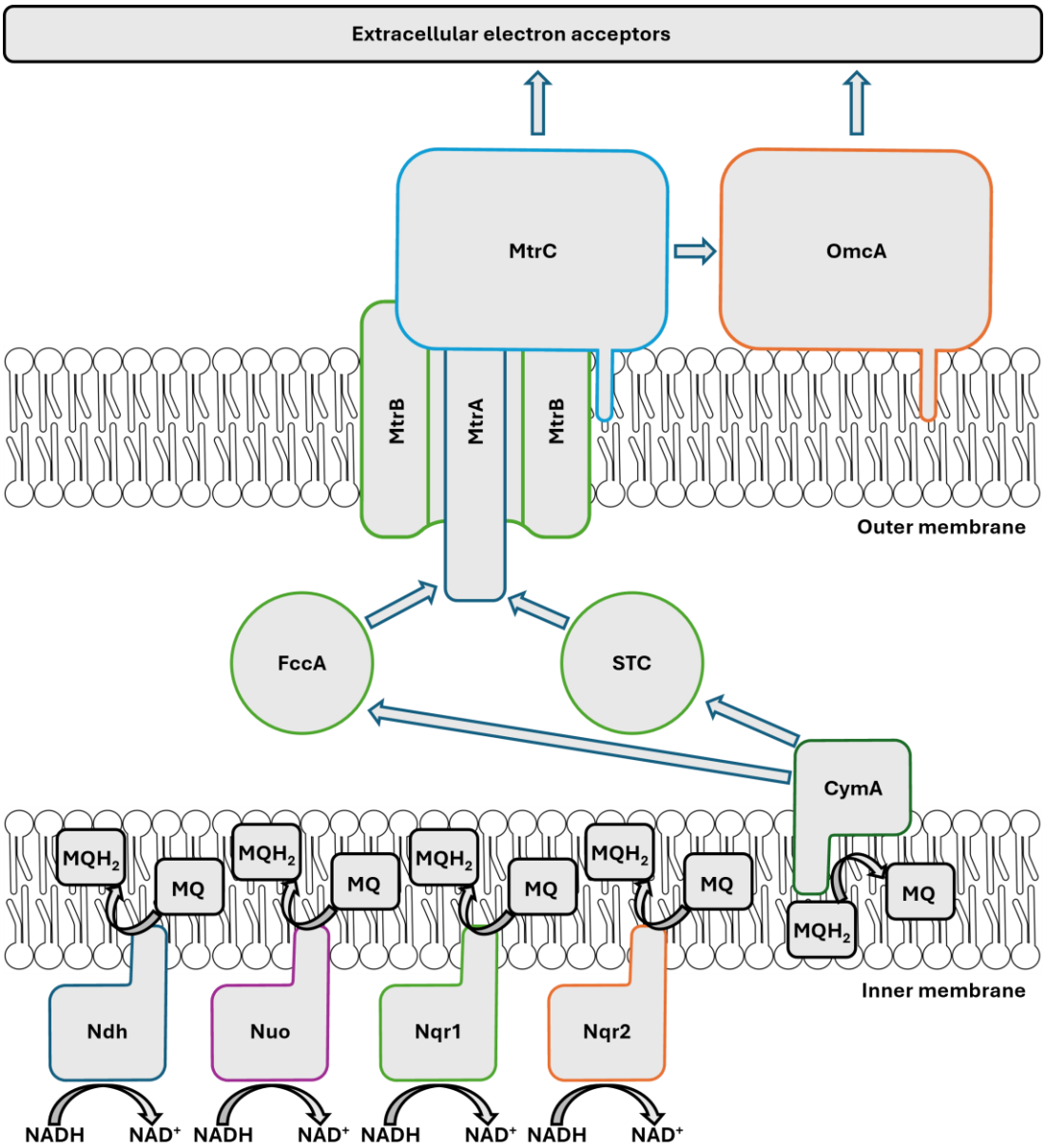


Figure 1.2: Schematic depicting the proposed extracellular electron transfer pathway of the bacterium *S. oneidensis* MR-1. Protein species shown with coloured outlines and the transfer of electrons with blue outlined arrows. MQ-menaquinone and MQH2-menaquinol.

anaerobic conditions. Furthermore, Δ-acetate kinase and Δ-phosphate acetyltransferase strains could not grow with lactate, with which those enzymes are needed to produce an ATP via substrate-level phosphorylation. Those strains could grow however on N-acetyl glucosamine which produces an additional ATP by entering substrate-level phosphorylation via the Entner-Doudoroff pathway.

This discovery that MR-1, which has become renowned for the versatility of electron acceptors it can exploit, possibly extracts energy only from substrate-level phosphorylation and not from accumulating an electrochemical gradient is an important one. It suggests that MR-1 is optimised not for efficient growth, but instead for rapid growth, and that the electrons discharged through EET pathways are best dissipated into the extracellular environment as rapidly as possible to make way for additional substrate-level phosphorylation reactions.

An electron traversing the respiratory pathway of an MR-1 bacterium under anaerobic conditions (Figure 1.2), flows after substrate-level phosphorylation via NADH to one of MR-1's four inner membrane NADH dehydrogenases, Ndh, Nuo, Nqr1 or Nqr2, or via lactate to the Lactate dehydrogenase complex[32,33]. From there, it is transferred into to the quinol pool, for which the primary relevant oxidase is the *c*-type multiheme cytochrome CymA.

CymA is a member of the NapC/NirT family, the members of which are tetrahemes and contain an N-terminal inner membrane-spanning alpha helix[34]. CymA displays 27% sequence identity with the NrfH of *Desulfovibrio vulgaris* for which crystal structures, as part of the NrfHA nitrite reductase complex, have been solved (PDB-2VR0 and -2J7A)(Figure 1.3)[34,35].

The secondary structure of NrfH is entirely alpha helical and the protein incorporates its four *c*-type hemes into two stacked pairs, hemes two and three (as numbered by appearance of the heme binding CXXCH motifs in the amino acid sequence) feature His-His coordination of the heme iron. The binding site for quinols was resolved by crystal soaking with a quinol analogue and competitive inhibitor (2-heptyl-4-hydroxyquinoline-N-oxide)[35]. The site is located in close proximity to heme one and the trans-membrane helix, facing towards the

position of the membrane, allowing for entry of a hydrophobic quinol molecule, which can form hydrophobic interactions with the side chains of Phe42 and Phe81 and with the porphyrin ring.

ΔCymA strains show poor or no rates of Fe^{III}, Mn^{IV}, fumarate, DMSO (dimethyl sulfoxide) and nitrite reduction[36,37]. The enzyme has been demonstrated experimentally to interact transiently with the MR-1 fumarate reductase (FccA) and the small tetraheme cytochrome (STC), the K_d of these interactions is reported to be in the hundreds of micromolar[37,38]. The reduction potentials of the CymA hemes have been estimated as ranging between -110 and -265 mV vs standard hydrogen electrode (SHE)[39]. Taken together these observations have been interpreted as in favour of CymA playing the dominant role as a non-specific electron transfer "hub" in the MR-1 inner membrane distributing electrons from the quinol pool into many periplasmic cytochromes, many of which contribute to EET.

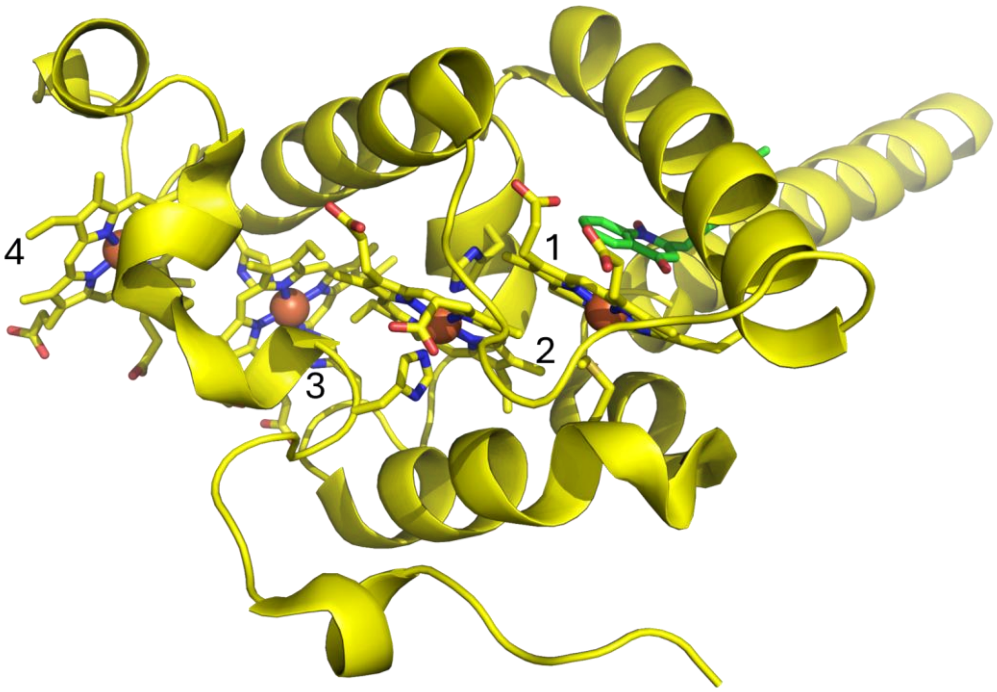


Figure 1.3: Crystal structure of *Nitratidesulfovibrio vulgaris* cyt-*c* NrfH (PDB-2VRO), depicted with heme coordinating residues M49, H61, H70, H120, H140 and H145 and bound competitive inhibitor 2-heptyl-4-hydroxyquinoline-N-oxide in the quinol binding site, as sticks, with hemes numbered by their CXXCH motif appearance in theNrfH amino acid sequence.

The MR-1 flavocytochrome *c* fumarate reductase (FccA) is a 62 kDa (kilodalton) periplasmic protein that receives electrons via CymA. FccA is one of the most abundant cytochromes in the MR-1 periplasm, and catalyses the reduction of fumarate to succinate.

Crystal structures of FccA unbound and in complex with fumarate and succinate have been reported (PDB-1D4C, -1D4E and -1D4D respectively)[40]. These structures reveal FccA adopts a multi domain structure with the first containing four His-His coordinated *c*-type hemes. The secondary structure of that domain is alpha helical and positions its four hemes with close edge to edge distance (~5 Å). Within the second domain, a flavin adenine dinucleotide (FAD) molecule is incorporated that serves as the fumarate reduction catalysing group[40]. It appears FccA is bifunctional, serving not only as a terminal reductase itself during fumarate reduction, but also as a periplasmic electron shuttle, transporting electrons from CymA to redox proteins in the outer membrane[41,42].

The sole function of STC in contrast, appears to be as a periplasmic electron shuttle. Despite the presence of its four *c*-heme groups, STC is only 12 kDa in size making it one of the most heme-dense of characterised cytochromes. The hemes of STC are also His-His coordinated and are arranged with hemes 2 and 3

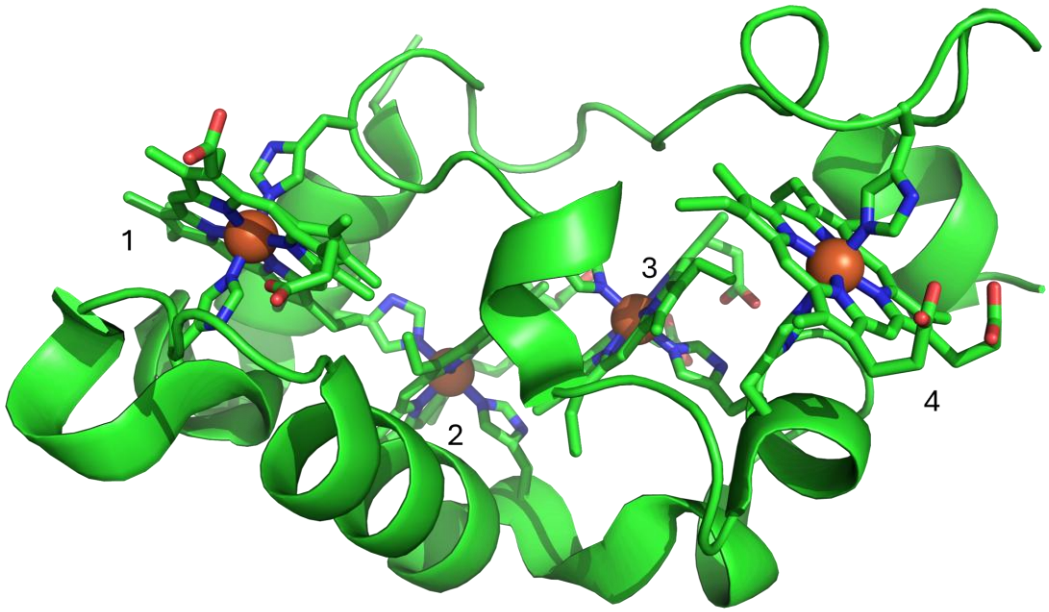


Figure 1.4: Crystal structure of *S. oneidensis* MR-1 cyt-*c* STC (PDB-1M1Q), depicted with heme coordinating residues H9, H19, H39, H49, H52, H62, H65 and H79 as sticks, with hemes numbered by their CXXCH motif appearance in the STC amino acid sequence.

in a parallel stacking pattern (Figure 1.4)[43]. In fact, alignment of the STC crystal structure with that of FccA shows they share significant similarity, three alpha helices are conserved between the two and their first three hemes can be aligned with negligible deviation.

Interestingly, STC and FccA have been shown to compete with one another in binding the multiheme cytochrome MtrA, which is consistent with the structural analysis suggesting they have overlapping properties[44]. The hemes of both proteins display high solvent exposure so are likely fairly promiscuous in their electron transfer partners, in contrast with this however it appears that they do not interact with one another, at least in vitro[38]. The benefit MR-1 enjoys by synthesizing both FccA and STC is still unclear, especially given single Δ*STC* and Δ*FccA* strains exhibit either very subtle, or no, growth defect under anaerobic conditions[42]. Nonetheless the contribution both proteins play shuttling electrons across the periplasm is clearly crucial as knocking out both *STC* and *FccA* genes yields a strain with a severe growth defeat under anaerobic conditions[42].

The greatest obstacle to the export of electrons from gram negative bacteria is surely the outer membrane. Biological membranes are made intrinsically insulating by the hydrophobicity of the lipid bilayer they are comprised of. Furthermore the assembly of membranous proteins to facilitate transmembrane electron transport is more challenging in the outer membrane, given proteins must first cross the inner membrane and periplasm to reach the outer membrane, and no ATP is available in the periplasm to actively drive membrane insertion processes.

The first evidence regarding how *S. oneidensis* MR-1 was able to solve this problem, was provided by *Beliaev and Saffirini*[45]. They identified an open reading frame (ORF) that encoded a 679 amino acid protein, which was localised to the outer membrane and when deleted yielded a strain without the ability to reduce Fe^{III} or Mn^{IV} oxide. This gene was designated as *MtrB* for metal reduction B, and the gene encoded alongside it as *MtrA*, metal reduction A. Soon an additional gene adjacent to *MtrA*, *MtrC* was added to the list of genes necessary for iron reduction[46].

Both *MtrA* and *MtrC* were identified as encoding decaheme *c*-type cytochrome proteins as they each contained ten Cys-Xxx-Xxx-Cys-His motifs, although the MtrA protein was defined as having periplasmic localisation while *MtrC* like *MtrB* appeared mostly associated with the outer membrane. The first model for trans-outer membrane electron transfer was subsequently proposed, in which the MtrA decaheme accepted electrons from STC in the periplasm and transferred them onto the MtrC decaheme in the inner leaflet of the outer membrane, which transferred them through an unknown mediator protein onto MtrB in the outer leaflet of the membrane, which then reduced extracellular metal ions through a Cys-Xxx-Xxx-Cys motif near its N-terminus[46,47].

Antibody Recognition Force Microscopy was used to demonstrate MtrC was localised in fact on the outer leaflet of the outer membrane[48], rather than the inner leaflet.

At this stage, the general structure of bacterial porins had been established via the recent determination of a sixteen-β strand *Rhodobacter capsulatus* outer membrane porin (Figure 1.5)[49]. MtrB was therefore also proposed to form an outer membrane porin, based on analysis of its sequence[45,50], and the prevailing hypothesis posited that MtrB itself performed electron transportation[50].

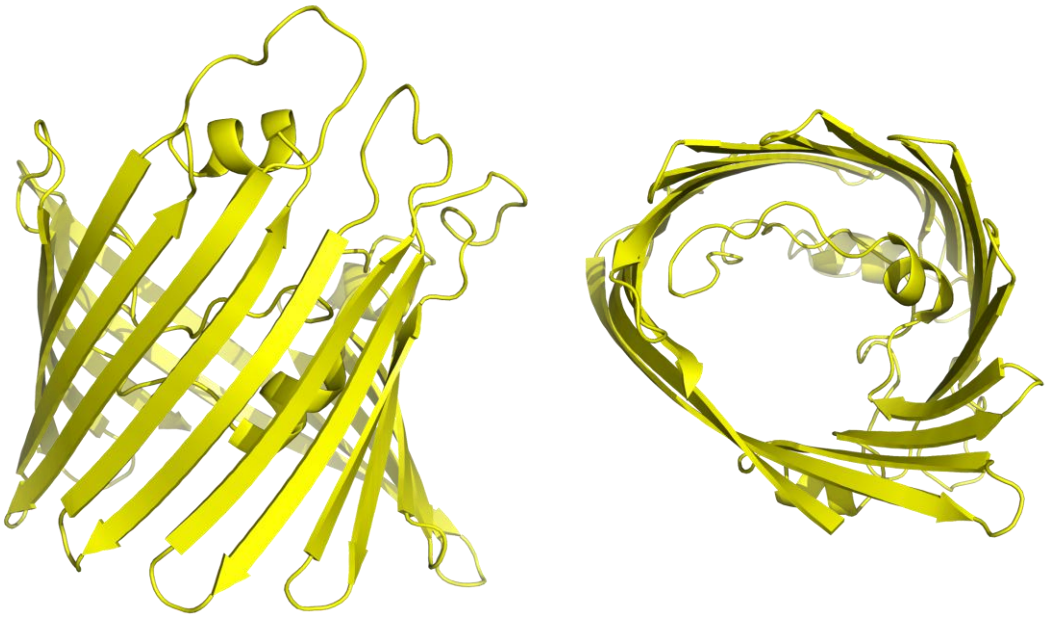


Figure 1.5: Crystal structure of *Rhodobacter capsulatus* porin (PDB- 2POR), side and top views, left and right respectively.

Hartshorne et al would then demonstrate for the first time copurification of a trimeric MtrCAB ~200 kDa complex and characterised it in depth[50]. They observed it indeed likely did contain twenty *c*-type heme groups provided by MtrC and MtrA, was redox active between -400 and 0 mV vs SHE and could be reconstituted in an active form transferring electrons across liposome membranes. Their most important observation however was that although MtrA and MtrC were stable monomeric proteins in isolation, MtrB required the presence of MtrA but could not be stabilised by the presence of MtrC. This asymmetric stabilisation was interpreted as evidence suggesting that the complex was unlikely to form an electron transfer chain from MtrA through MtrB onto MtrC. Instead it seemed likely that MtrA and MtrB were intimately linked in a way not shared by MtrC. Given that MtrB contained no canonical cofactor binding sites, they proposed this could be explained by MtrAB forming a complex in which the periplasmic decaheme MtrA was inserted someway into the MtrB porin, stabilising it, with MtrB not contributing directly to the electron transfer but rather insulating MtrA from the lipidic membrane environment. MtrC would then accept electrons from MtrA and extend someway away from the membrane into the environment. This revision of the hypothesis for Mtr operon mediated EET was consistent also with their observation that MtrC did appear to interact with MtrA without the presence of MtrB albeit weakly, and that an MtrAB complex exchanged electrons with electrodes very poorly in comparison to the MtrCAB complex[50].

Although their data did not allow for quantitative structural analysis, the core mechanism proposed by *Hartshorne et al* would be supported by more recent research and continues to be central, in a refined form, to our understanding of trans-outer membrane electron transfer in bacteria. In fact, their hypothesis appears also to hold, in many bacteria with EET-conferring gene clusters that do not display detectable sequence homology to MtrCAB, like *Geobacter sulfurreducens*.

The degree to which MtrC extends away from the membrane, and into MtrB and the degree to which MtrA extends into the MtrB porin remained unresolved. The

state of the art structure determination method of the time *Hartshorne et al* was published was X-ray crystallography but large protein complexes, especially those that are membranous, typically make for recalcitrant crystallisation targets. Some progress was made by analysis of solution-state scattering of X-rays by MtrA, with which it was suggested that at least while not interacting with the MtrB porin, MtrA adopts a tube like structure[51]. The next breakthrough instead would be in understanding the structure of the extracellular component common in homologues of the MtrCAB complex.

The genome of *S. oneidensis* MR-1 encodes in addition to MtrCAB, an additional homologous porin-dual decaheme cytochrome complex, MtrDEF. Although it is CAB that confers the metal reduction phenotype for which it is named, MtrF (homologous to MtrC) and MtrE (homologous to MtrA) under some experimental conditions are capable of compensating for the loss of their homologous gene so evidently play a related role in EET[52,53]. Solubilisation of MR-1 membranes, in detergent, and purification of the MtrF decaheme was achieved allowing *Clarke*

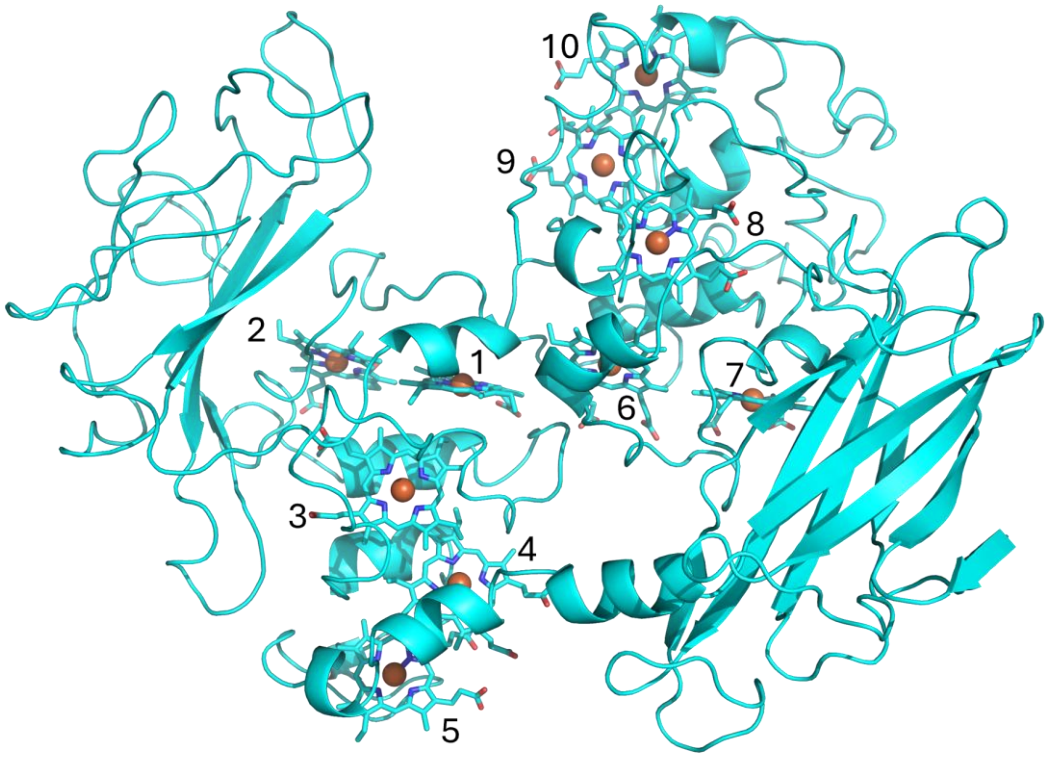


Figure 1.6: Crystal structure of *S. oneidensis* MR-1 outer membrane decaheme *c*-type cytochrome MtrF (PDB- 3PMQ), with hemes numbered by their CXXCH motif appearance in the MtrF amino acid sequence

et al to determine its crystal structure at 3.2 Å resolution which provided revolutionary insight into the mechanism of EET[54].

The MtrF crystal structure (Figure 1.6) showed it adopts a distinctive four domain conformation, domains one and three (as numbered by appearance in the amino acid sequence) are dominated a beta sandwich motif and contain no hemes, while domains two and four are alpha helical in their secondary structure and each incorporate five *c*-type hemes. The MtrF domain architecture is visually clearly psuedosymmetrical raising the possibility of it having been created through a DNA duplication event. The structure positions its ten hemes in a striking arrangement in which hemes 1, 3, 4, 5, 6, 8, 9, and 10 form a continually cofactor chain with close (~5 Å) edge-edge distances, but hemes 7 and 2 each form a protrusion from that cofactor chain approximately perpendicular to it creating what has become known as a staggered cross structure. The surface exposure of the hemes was calculated and found to be greatest for hemes 5 and 10, those at either end of the longer cross in the staggered cross.

In agreement with the crystal structure MtrF displayed MCD spectra typically of Bis-His coordination of hemes and was found using protein film voltammetry (PFV) to be redox active between ~-400 and +100 mV vs SHE[54].

MtrF, was therefore revealed to share many properties with previously characterised multiheme cytochromes, for example the *D. vulgaris* NrfH discussed earlier in this work the most important of which is an insulated chain of Bis-His coordinated hemes. The revelation that they exist, in contrast with previously described cytochromes, in a staggered cross arrangement was rationalised as a logical result of MtrF being extracellularly localised[54]. Intracellular oxidoreductases, in order to minimise the generation of reactive oxygen species (ROS) must control electron transfer carefully. In contrast, extracellular electron transport proteins are likely incentivised to develop strategies that maximise their chance of contact with a substrate, for example by incorporating multiple sites of electron egress, even at the expense of some increased ROS production. The extracellular cytochromes synthesized by MR-1 thus began to be considered rather than electron wires, but more as electron

'satellite dishes' that receive electrons from the porin and can probably distribute them onto acceptors in three different spatial directions via their staggered cross.

All MR-1 cytochromes that are predicted to be extracellularly localised, include a motif at the N-terminus of their mature polypeptide (post signal peptide cleavage), which is predicted to confer lipidation during translocation into the periplasm[55]. Consistent with this, the wild-type proteins remain in the membranous fraction during purification until solubilised in detergent[56]. Much more recently the lipidation process has been characterised using liquid chromatography-mass spectrometry (LC-MS) where it was determined to most likely be diacyl lipidation[55].

The MtrF structure did not resolve any features that obviously suggested a where its binding site with MtrDE was located. In fact, protein-protein binding interface prediction software, that at the time was state of the art, incorrectly suggested domains one and four as most likely candidates, it would later be shown domain two is the main site of interaction with the porin-cytochrome unit[57].

With the structure of only a single extracellular decaheme protein available, it remained unclear how conserved their structure was. The determination of the crystal structure of the undecaheme extracellular cytochrome UndA from *Shewanella sp. HRCR-6* would soon partially resolve this question.

Within the cluster that contains the genes encoding for the MtrCAB and MtrDEF proteins, an additional extracellularly localised cytochrome, OmcA, is also encoded in MR-1. OmcA appears to play a role supporting the MtrCAB complex during EET in MR-1, although evidence suggests the importance of this functionality may vary depending on the extracellular electron acceptor in use[58,59].

In place of OmcA, some *Shewanella* species encode an undecaheme protein UndA which is thought most likely to perform a similar function as OmcA[60] and was the next outer membrane cytochrome to be characterised structurally[61]. Its crystal structure displayed high similarity with that of MtrF, a four domain

architecture with domains one and two separated from three and four by a pseudosymmetry axis. Its hemes were also all Bis-His coordinated, and numbers 1-7 and 8-11 could be aligned closely onto those of MtrF, with the addition of another *c*-type heme (number 7), lengthening the shorter chain of hemes, that are now known to run perpendicular to the membrane (2, 1, 6 and 7 in MtrF).

Given the low sequence similarity of these two proteins (~25%), from these structural findings it was proposed that all outer membrane cytochromes of MR-1 likely shared a common structure[61]. This hypothesis would be supported by the addition of a crystal structure for OmcA, which showed it also adopts the staggered cross of hemes distributed into two pentaheme domains augmented by beta sheet domains and separated by a pseudosymmetry axis[62].

An additional feature of interest identified in the OmcA crystal structure was a proposed motif for iron oxide binding near heme ten comprised of Thr-725, Pro-726 and Ser-727. This proposal was based on molecular modelling and phage-display experiments, which suggested that motifs comprised of Ser/Thr-Pro-Ser/Thr sequences would confer increased affinity for oxides[61,63]. That affinity was though to arise from hydrogen bonding with the mineral surface by the serine/threonine hydroxyl which was further promoted by proline-induced limiting of flexibility[63]. In OmcA this site is located in sufficient proximity (~10 Å) to the conductive heme chain to make electron transfer feasible, although its relevance to OmcA-mediated EET has remained theoretical and untested.

Despite MtrC being the focus of most biophysical and phenotypic characterisation, of the four *Shewanella* outer membrane cytochromes it was the final to be structurally characterised[64]. Its structure was also shown to be that of a four domain protein with a heme staggered cross, however it did reveal the importance of a conserved disulfide located within domain three[64].

Using fluorescence spectroscopy *Edwards et al* demonstrated MtrC associates tightly with flavins specifically under anaerobic conditions which was consistent with previous nuclear magnetic resonance spectra (NMR) which indicated a flavin affinity in the hundreds of micromolar range under aerobic conditions[65].

This activity was not linked to redox state of the hemes, but rather that of the Cys444 and Cys453 disulfide. Mutation of those respective codons on a plasmid inducible MtrC copy, yielded a strain with a severe growth defect under aerobic conditions while the inducer was present, but was otherwise unperturbed[64]. These observations were taken together as indicating the disulfide acts as a redox switch, facilitating formation of a reactive flavin-cytochrome complex under anaerobicity, that could be disassembled in the presence of oxygen in order to prevent the generation of cytotoxic ROS.

Attempts to structurally characterise this flavocytochrome species were unsuccessful, although recently the crystal structure of the cytotoxic C453A mutant was solved[66]. Despite the clear phenotype the mutation produces, the crystal structure was essentially indistinguishable from that of the wild-type protein. Thus the exact nature of the interaction between flavin and the MR-1 outer membrane cytochromes remains unresolved but given flavin concentration is an important parameter that effects the efficiency of EET by *S. oneidensis*, it likely plays a significant role.

The first direct evidence showing *Shewanella* uses flavin to assist in EET dates to several years before the possibility of stable flavin-cytochrome complex formation was raised. *Von Canstein et al* were able to isolate flavin mononucleotide (FMN) and riboflavin from the supernatants of cultures of two different *Shewanella* strains by HPLC (High-performance liquid chromatography) and unambiguously assign their identity by LC-MS[67]. *Marsili et al* expanded upon this by demonstrating that their removal by substituting growth media for fresh media, lead to a dramatic drop in the extracellular electron transfer rate of *Shewanella* anodic biofilms[68]. Importantly, the replacement of the original media restored the original rate demonstrating the health of the biofilm was not diminished, but rather the loss of a soluble electron shuttle was the cause[69]. Flavins therefore were recognised as also able to facilitate the transfer of electrons from the outer membrane onto minerals more than a few nanometres from the cell surface, which outer membrane cytochromes, being anchored to the membrane by their N-terminal lipid modification, would be unable to reach.

The process of flavin synthesis is complex, but it has now been identified that MR-1 encodes a full complement of genes homologous to those in established *Escherichia coli* and *Bacillus subtilis* pathways[70]. FAD synthesized in the cytoplasm is secreted across the inner membrane by a multidrug and toxin efflux family protein bfe (bacterial FAD exporter)[68]. Then it can be incorporated into FccA, or processed by a nucleotidase, UshA, into FMN which appears to degrade spontaneously into riboflavin[68,70]. The mechanism by which FMN and riboflavin traverse the outer membrane appears to be nonspecific which is consistent with its small size. In agreement with this description, *Δbfe* strains are poorly electroactive while insoluble substrates or electrodes are the terminal electron acceptor, which can be rescued by the addition of exogenous riboflavin or FMN[68].

Much discussion has been dedicated to establishing the energetic cost of flavin secretion and comparing this with estimates for the energetic contribution flavin-based EET might provide[70,71]. While the nature of the balance between free-flavin redox shuttling, and outer membrane associated flavin-cytochrome complexes remains unresolved, the arrival of a definitive answer to the energetics question seems unlikely. In this context, the observation that MR-1 appears not to conserve energy via proton motive force (PMF) buildup but rather by substrate level phosphorylation maybe important[31]. It suggests the organism is better optimised for rapid and "good enough" respiratory processes rather than maximally efficient ones. This strategy could extend to secreting flavins which can facilitate faster EET, at lower efficiency introduced by the loss of flavin to the surrounding environment.

Also present in the extracellular milieu of *S. oneidensis* cultures are so called "bacterial nanowires", extracellular appendages that are visible extending from the cell surface in electron micrographs provided by many groups[72–74]. These structures can be observed extending some several cell lengths away and their electrical properties have been measured directly[73]. They display ohmic conductivity (current linearly proportional to voltage) on the order of 1 S cm^{-1} but nanowires produced by *ΔmtrC/omcA* variant strains do not share this conductivity[73]. Furthermore, immunofluorescence with anti-MtrC and anti-

OmcA antibodies identifies these appendages as bristling with MtrC and OmcA, but the appendages also stained for lipid with the membrane specific FM 4-64FX stain[75]. These observations led to the proposal that unlike the nanowires synthesised by many other electroactive bacteria, those produced by MR-1 are extensions of the outer membrane and periplasmic space[75]. This was later confirmed by *Subramanian et al*, who used cryogenic electron tomography (CET) to image the nanowires at high resolution and found their features entirely consistent with the above hypothesis[74].

It appears the mechanism of conductivity along the length of the *Shewanella* membrane extensions is distinct from the mechanism of trans-membrane conductivity. The density of outer membrane cytochromes observed by *Subramanian et al* was insufficient to provide a continual pathway of redox cofactors for direct electron hopping along the cell surface[74]. Instead, lateral diffusion of cytochromes across the outer face of the membrane also seems to make a significant contribution to the conductivity of these appendages.

Despite tremendous interest in the molecular mechanism of trans-outer membrane electron transfer by *Shewanella*, experimental limitations made the MtrCAB complex the most difficult and therefore last to be structurally resolved. Membrane proteins are notorious for being significantly less stable than their soluble counterparts, this is the case also for protein complexes which can fall apart during purification or downstream applications. A particular issue for structural biology, is that membrane embedded regions tend not to form strong crystal contacts, and therefore often yield either no crystals, or crystals that suffer from poor diffraction, small size or other pathologies.

Nonetheless, in 2020 *Edwards et al* revealed the crystal structure of the MtrCAB complex from a close relative of MR-1, *Shewanella baltica* (Figure 1.7)[57]. Like the structure of MtrF before it, the structure of the Mtr complex represented a paradigm shift for the field.

MtrB was shown to form a 26 strand beta barrel suspended in the outer membrane lipidic environment by a belt of hydrophobic residues running around

the length of the barrel. Loops linked each strand from the next, although the porin was not perfectly cylindrical being wider on the periplasmic side, and containing a protrusion at the N and C-termini.

Buried within the MtrB porin, the MtrA decaheme was resolved extending the full length of the membrane to the end of the beta barrel, and some way into it the periplasm. MtrA contains several alpha helices but is dominated by loops which is unsurprising given it contains one heme for every ~3.5 kDa of mass. The hemes of MtrA were shown to stack in pairs, like other multiheme cytochromes, with edge to edge distances of ~5 Å and the final heme (ten) is presented on the cell surface.

The *S. baltica* MtrC structure was very similar to that of MR-1, also displaying the staggered cross and four domains. Finally, the orientation of MtrC against the MtrAB module was resolved, heme five in fact was found to be in close contact with heme ten of MtrA with edge to edge distances of ~8 Å. A network of hydrogen bonding interactions holds the complex together. Hemes two, seven and ten present facing into the extracellular environment in three directions in one plane. MtrC is tilted some 20° from the central axis of MtrAB.

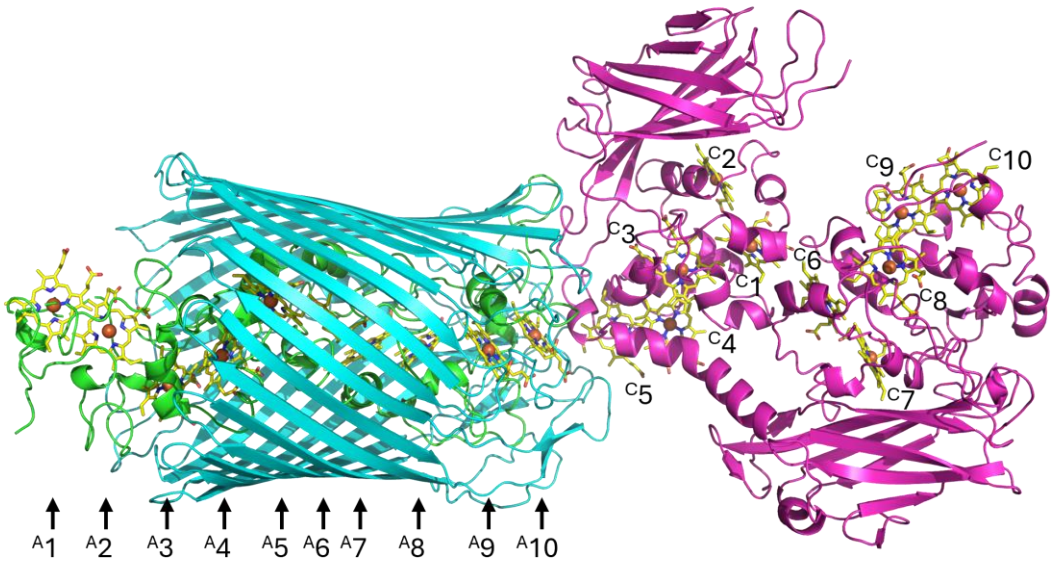


Figure 1.7: Crystal structure of *S. baltica* MtrCAB (PDB- 6R2Q). MtrB – cyan, MtrA – green, MtrC – magenta and *c*-type heme cofactors – coloured by element with carbon – yellow, nitrogen – blue oxygen – red and iron – orange. Hemes annotated with the protein subunit they reside in, in subscript and numbered by the appearance of their CXXCH motif in their protein amino acid sequence.

Taken together, the structural analysis allowed for by the MtrCAB crystal structure finally rationalised the observation by *Myers and Nealson*[30] more than thirty years previously, that *Shewanella* can conserve energy by the reduction of substrates than could not possibly permeate the outer membrane, by assembling a conductive biological molecular wire in its outer membrane[57].

Attempting to determine the rate of electron transport that the MtrCAB wire can support, poses challenges. Cell-based methods suffer from large error associated with estimating the number of MtrCAB molecules each cell may contain. Great effort has been devoted to ensuring MtrCAB prepared in liposomes is oriented with MtrC facing outward[76]. The preparation described in ref[76] yielded proteoliposomes that could reduce encapsulated methyl viologen (MV) using extravesicular sodium dithionite as the electron donor. This pool of reduced MV could be re-oxidised by the addition of extravesicular soluble Fe^{III} citrate, which could be replicated for proteoliposomes assembled with just MtrAB incorporated (No MtrC). Addition of Fe^{III} mineral oxide particles could re-oxidise the MV pool however, only when MtrC, MtrA and MtrB are all present indicating the MtrC decaheme is needed to either move electrons sufficiently far from the lipid membrane to reach the mineral particles, or to perform a specific electron transfer activity at the particle surface. This method could also be replicated using STC, one of the endogenous partners of MtrCAB, as the internalised electron donor[77]. These experiments arrived at an estimate of 10^3 electrons per second per MtrCAB complex varying depending on the mineral used[76].

Clear evidence identifying the rate limiting step is during this process has not been described, but it is unlikely to be the intramolecular electron transfer within MtrC and MtrA heme chains. A sophisticated approach to measure the internal rate of heme to heme electron transfer in multiheme cytochromes has been described and found that process to be extremely rapid[78].

There are several experimental difficulties that make this a challenging goal to achieve. The necessity for site selective, uniformly controlled and rapid electron injection into the cytochrome make photosensitised-pump probe spectroscopy

the preeminent technique of choice and the groundwork for such measurements was initially laid using STC[79]. *Van Wonderen et al* demonstrated incorporation of a ruthenium-based dye (Ru 4-bromomethyl-4'-methylbipyridine (bipyridine)$_2$) at three engineered cysteine positions (A10C, T23C and S77C). Photo-induced excitation of this dye yielded transient absorbance spectra that suggested light-driven inter-heme electron transfer had occurred but the inability to distinguish between hemes spectroscopically precluded direct observation of inter-heme electron transfer[79].

A subsequent innovation was substitution at position 561 in MtrC from histidine to methionine. In the wild-type protein, His561 provides heme eight with its distal coordination ligand, and using X-ray crystallography *van Wonderen and colleagues* were able to demonstrate a methionine residue at this position can also coordinate the heme[78]. The unique contribution the methionine coordination made was an altering of the absorbance spectra of heme eight, creating a unique signal in the protein and thereby creating a mechanism to detect photo-driven inter-heme electron transport across hemes ten to nine to eight. This system yielded transient absorbance spectra consistent with electron transport across MtrC on the order 10^5-10^6 electrons per second, significantly faster than that measured by proteoliposomes studies[76] and individual heme to heme electron transfer rate constants as high as 10^9 per second[78].

Thus genetic, structural and functional characterisation of the *Shewanella* MtrCAB complex have provided compelling evidence, that are all in agreement, for it acting as an outer membrane electron export conduit in that organism. Within MtrCAB, it is the MtrAB porin-cytochrome module that is most conserved. Although structural characterisation of homologous systems have not been described, what is known about several of the most studied examples can be summarised as follows.

Rhodopseudomonas palustris TIE-1 encodes homologues of the *mtrA* and *mtrB* genes, denoted as *pioA* and *pioB*, respectively, but without an *MtrC* homologous gene[80]. This organism unlike MR-1, is photosynthetic, and can use its PioAB complex to extract electrons from Fe^{II}, which then replaces photoexcited

electrons used in photosystems (non-oxygenic photosynthesis)[80,81]. Much is unknown about PioAB mediated photosynthetic iron oxidation, but it is interesting to note that the MtrAB module is evidently versatile and efficient enough to also support electron uptake processes, with some modifications.

Sideroxydans lithotrophicus ES-1 is another gram negative bacterium that also utilizes iron oxidation as a source of electrons, in its case allowing for conservation of energy via biotic O_2 reduction[82,83]. Seldom little is understood about this process, largely due to the difficulty in cultivating organisms that perform this form of respiration[82]. Nonetheless ES-1 does encode gene homologous to *mtrAB*, in its case referred to as *mtoAB*.

MR-1 itself contains genes encoding for an additional porin cytochrome complex, DmsEF, with those genes being homologous to the *mtrA* and *mtrB* genes respectively. Surprisingly these are encoded not alongside an *mtrC/mtrF/omcA* family decaheme but rather genes encoding a DMSO reductase enzyme DmsA, an FeS cluster protein DmsB and two accessory proteins DmsG and DmsH. These genes (SO1427-SO1432) were first identified by *Gralnick et al*, by determining binding sites for the ArcA two-component regulatory system in *S. oneidensis*.[84]. In *E.* coli, ArcA allows for sensing and responding to the intracellular redox state, through gene expression changes. *Gralnick et al* found that the homologous system in *S. oneidensis* binds strongly upstream of the *dmsEFABGH* gene cluster. They generated a strain with a kanamycin resistance gene inserted in *dmsE*, which they found was unable to respire on DMSO, unlike the wild-type which was.

In a second paper by the group, they more thoroughly characterised this gene cluster, finding that *dmsB* also was necessary for growth with DMSO as the electron acceptor[85]. *S. oneidensis* contains a second cluster (SO4357-SO4362) homologous to *dmsEFABGH* although this cluster was barely upregulated in response to DMSO and evidently could not confer the ability to grow with DMSO as the terminal electron acceptor so is likely either not functional or not expressed under laboratory conditions[85]. *Gralnick et al* also prepared a strain with deletions in crucial components of the type II secretion system which has

been shown to be responsible for exporting extracellular cytochromes (MtrC/OmcA type) in *Shewanella putrificans*, a close relative of *S. oneidensis*[85,86]. Remarkably, given all characterised DMSO reductases are intracellular, this strain exhibited severely impaired growth with DMSO as its terminal electron acceptor. The implication of this result, when interpreted with the knowledge that the DmsE and DmsF proteins share high sequence identity with MtrA and MtrB, was recognised immediately as indicating this cluster also encodes a transmembrane electron conduit, with DMSO reductase activity.

At that time, scarce little was known about the structure of MtrCAB or the mechanism through which it confers EET-ability. This was not the case however for DMSO reductases, which have been studied in detail for several decades.

DMSO Reductase enzymes

DMSO reductase enzymes belong to a large family, the molybdopterin-containing proteins, which are found distributed amongst all domains of life. The common feature amongst all members is the molybdopterin cofactor which invariably features a mononuclear molybdenum ion coordinated by dithio group(s) of one or two pterin ligands[87]. These pterins are most often guanine derived, although not always[87,88]. Molybdopterin-containing enzymes can catalyse two electron redox reactions, including the conversion of DMSO to DMS (dimethyl sulfide), the oxidation of sulfite, the oxidation of formate, the reduction of arsenate and the oxidation of xanthine[87].

Within the larger family, DMSO reductase type enzymes are typified by containing two pterin groups specifically, both coordinating the molybdenum ion in addition to the side chain of a cysteine, selenocysteine or serine residue, and their distribution is thought to be restricted to prokaryotes[87]. Within this group in addition to DMSOR (true dimethyl sulfoxide reductase) are also categorised: the dissimilatory nitrate reductases (NarGHI), the trimethylamine-N-oxide reductases and the formate dehydrogenases.

Of the true DMSOR enzymes, the *Rhodobacter sphaeroides, Rhodobacter capsulatus* and *Escherichia coli* enzymes have been the focus of most study.

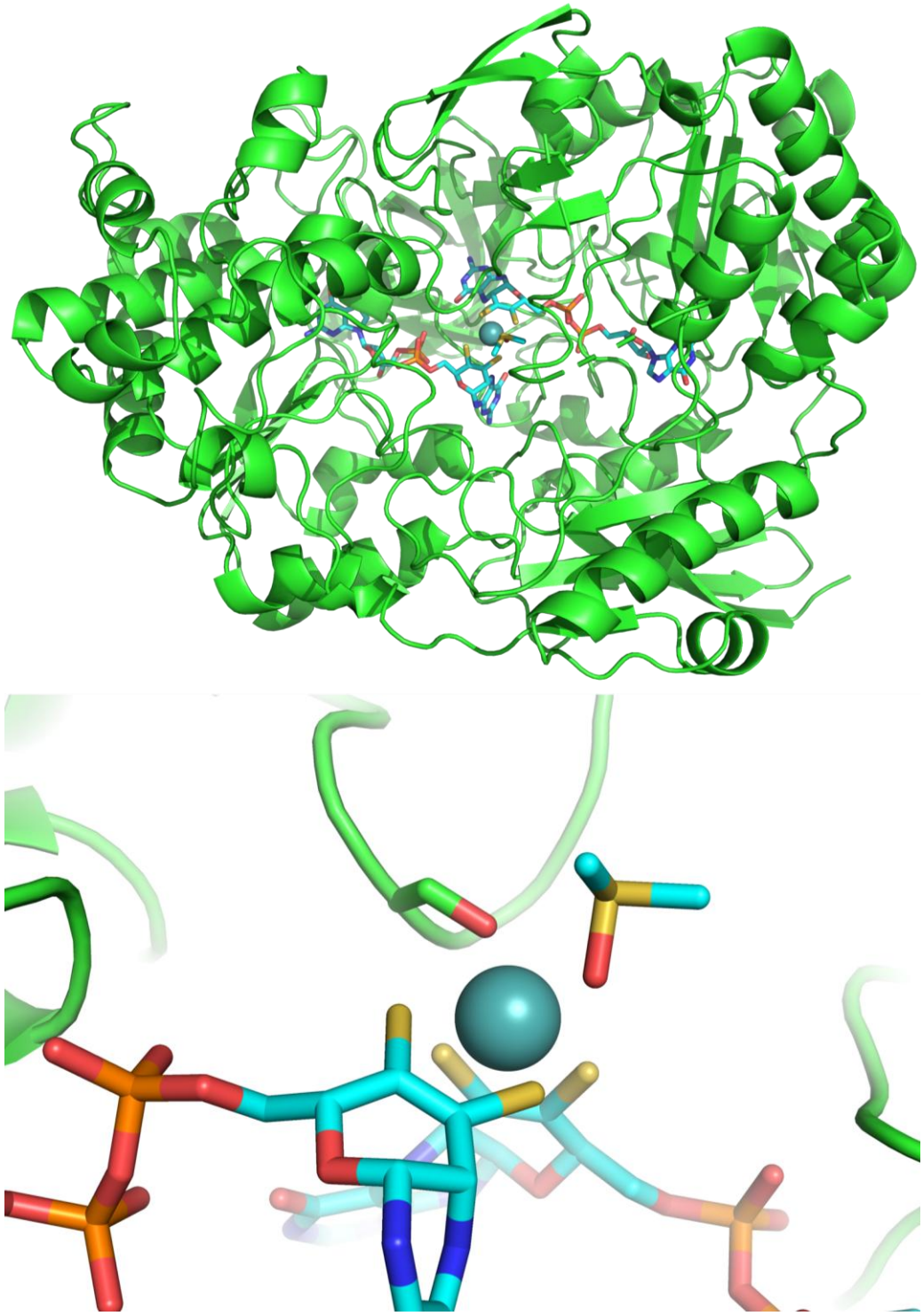


Figure 1.8: Crystal structure of *R. capsulatus* DMSO reductase with DMSO bound (PDB-4DMR). Molybdopterin cofactor, S147 hydroxyl and DMSO represented with sticks, molybdenum with spherical representation. General structure (top) and active site zoom in (bottom).

True DMSORs are ~800 amino acids in length and share a conserved structural fold, often in spite of low sequence identity.

The first structure (2.2 Å resolution) solved of a true DMSOR was that of the *R. sphaeroides* by S*chindelin et al*.[89]. That protein was globular in dimensions, with numerous alpha helices and beta sheets, it also contained a central funnel-shaped cavity at the bottom of which the molybdopterin cofactor was located. In addition to its dithiolene coordination, the molybdenum ion was hexa-coordinated by the side chain of Ser147, and an additional oxo-group bound with

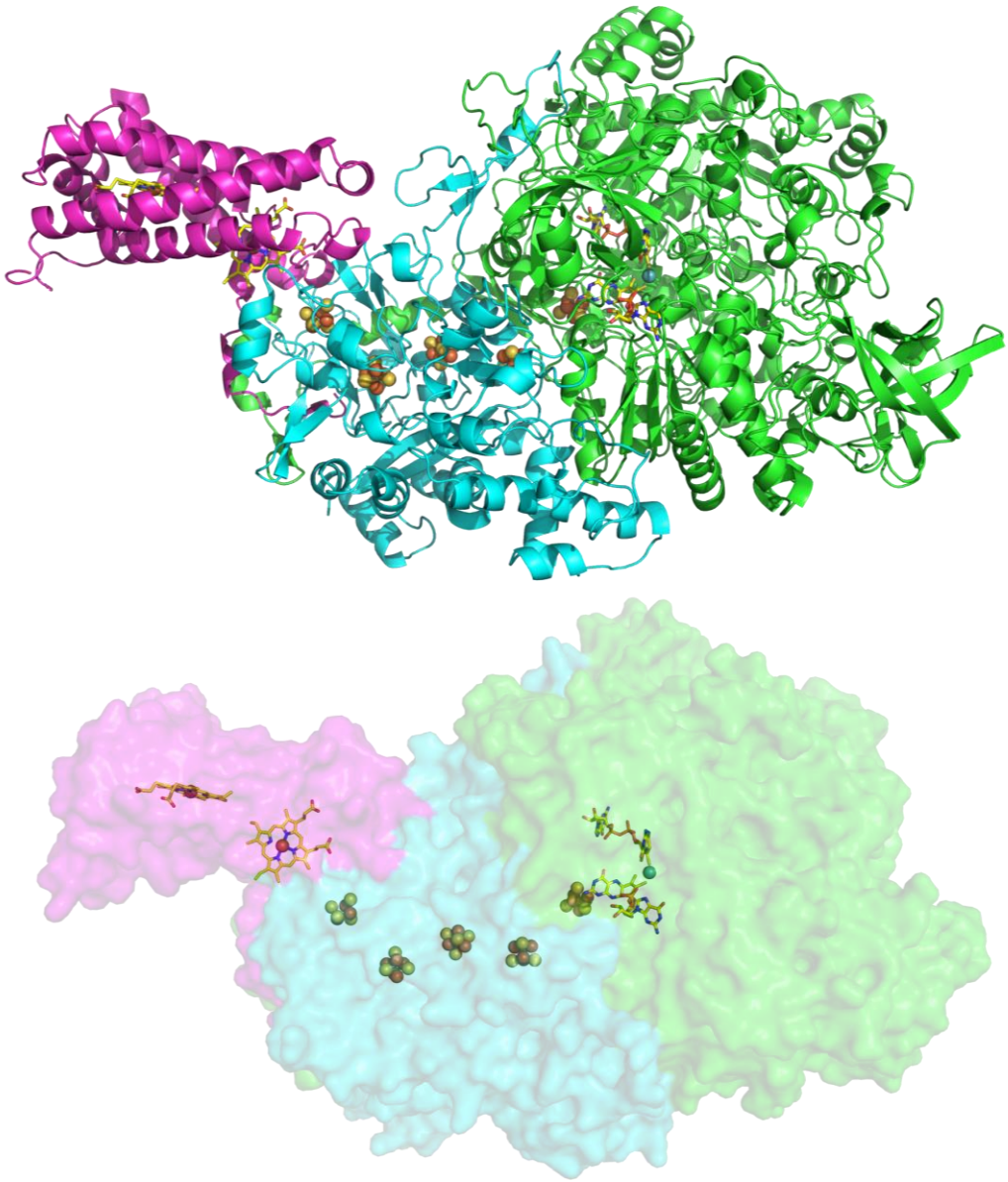


Figure 1.9: Crystal structure of *E. coli* dissimilatory nitrate reductase complex (PDB-1Q16). Cofactors depicted with sticks (b-type hemes and molybdopterin) and spheres (FeS clusters). NarG – green, NarH – blue and NarI – magenta.

a bond distance of 1.7 Å. By soaking crystals in sodium dithionite before data collection, the authors were able to solve also a crystal structure of the reduced protein in which the coordinating oxo group was no longer present, and a single sulfur-molybdenum bond was lost[89].

The structure of the *R. capsulatus* enzyme was solved next, similarly this enzyme shared the distinctive funnel shaped cavity leading down to its active site, with only a single cofactor, the molybdopterin[90]. In this structure, Ser147 was once again resolved coordinating the Mo. However, surprisingly given how closely related these two enzymes are, two not one oxo-groups were coordinating the Mo ion. Furthermore, one of the molybdopterin guanine dinucleotide cofactors (MGD) was not bound to the Mo through either sulfur atom (sulfur-molybdenum bond distances of 3.5 and 3.9 Å) leaving it penta-coordinated.

A later exceptionally high resolution (1.3 Å) structure of the *R. sphaeroides* DMSOR would demonstrate it can indeed assume the coordination geometry seen in the *R. capsulatus* enzyme, although in addition to a second hexa-coordinated conformational state which is the reportedly catalytically relevant one, with the original 2.2 Å structure possibly representing an average of those two states[90]. A substrate bound crystal structure of DMSOR has also been reported, which revealed DMSO binds the Mo from above, at the base of the funnel, in a pocket surrounded by many aromatic residues (Figure 1.8)[91].

The catalytic mechanism which is reportedly most consistent with data from Raman spectroscopy[92], X-ray absorption spectroscopies[93] and crystallographic structures, and is thus most favoured in recent literature involves molybdenum cycling between the IV, V and VI oxidation states[94,95]. DMSO binds via its O atom, to a pentacoordinate Mo^{IV} species. Scission of the substrate S-O double bond yields release of DMS and a Mo^{VI} oxo-bound species, which is reduced first to a Mo^{V}-hydroxyl species, and then again releasing a water molecule and regenerating the initial pentacoordinate Mo^{IV} species.

Both *Rhodobacter* DMSOR enzymes are soluble proteins, which are localised to the periplasm and contain only the single cofactor. The *E. coli* enzyme is much

more complex, and is probably more comparable to the *S. oneidensis* DMSOR. It consists of three subunits, DmsA, DmsB and DmsC which are ~83, ~24 and ~23 kDa in size respectively, and unlike in *Rhodobacter*, this enzyme is membrane associated[96].

Although no experimental structure has been solved of the *E. coli* DMSOR, its dissimilatory nitrate reductase and formate dehydrogenases share the same genomic architecture and significant sequence identity. High resolution crystal structures have been reported for both those enzymes (Figure 1.9)[97,98]. In these structures, the largest subunit (NarG and FdnG), which is homologous to DmsA, contains the molybdenum cofactor, but it also incorporates a 4Fe-4S cluster. The DmsB type subunit (NarH and FdnH) contains four 4Fe-4S clusters. The DmsC type subunit (NarI and FdnI) is entirely alpha helical, is almost entirely embedded within the inner membrane and contains two *b*-type hemes. Both *b*-type hemes are His-His coordinated, and the first serves as the site of its menaquinol substrate oxidation for the enzyme. These cofactors form a contiguous chain that facilitates the transfer of electrons liberated from quinol, some ~90 Å to the Mo ion for nitrate reduction, or in the case of FdnGHI, in the reverse direction from formate to the quinone pool. Of relevance to note is the fact that these enzymes although both associated with the inner membrane, are embedded on opposite sides, with the DmsA headgroup in the periplasm of FdnGHI, and in the cytoplasm of NarGHI[98]. It is the *E. coli* DMSOR gene cluster after which the *S. oneidensis* DmsEFAB is named, hence the MtrA and MtrB homologous subunits were assigned as DmsE and DmsF respectively and not DmsC[84].

Also within the *S. oneidensis* DMSOR gene cluster, are two additional genes which at the time of their identification, were not considered to have homologues in the *E. coli* operon, *dmsG* and *dmsH*. In hindsight, it seems highly likely that *dmsG* is in fact homologous to the *E. coli* gene *dmsD*, since the DmsG protein displays ~30% sequence identity to DmsD. Furthermore, superposition of AlphaFold2[99] predicted model structures of DmsG with the crystal structure of DmsD yield RMSD values of ~1.2 Å suggesting they are likely to perform the same function[100].

The DmsD family of proteins are well characterised and a wealth of data exists showing they play a crucial role in directing DmsA to the Twin Arginine Translocation (TAT) pathway[101,102]. The TAT pathway transports proteins, pre-folded and usually loaded with cofactors in the cytoplasm, across the inner membrane of prokaryotes[102]. Its substrates are recognised by a distinctive S/T-R-R-X-F-L-K motif at their N-terminus, by the chaperone DmsD. Molecular dynamics and site directed mutagenesis have identified a putative binding site on DmsD for binding of the twin arginine motif. *S. oneidensis* DmsG contains many of the residues suggested to be important in that binding including Y22, F76, N122, E123 and H127 (as numbered in *E. coli* DmsD)[100,103]. Taken together, these observations strongly suggest DmsG serves as a chaperone, assisting in the recognition of folded DmsAB by the TAT pathway, which is consistent with both its predicted ultimate localisation, and the presence of an Arg-Arg motif at the N-terminus of DmsA while no signal peptide is discernible at the DmsB N-terminus.

The role of DmsH is much less clear, the protein is predicted to be much smaller (155 amino acids), and shares no detectable sequence identity with any entry currently deposited in the PDB. Models of the protein predicted by AlphaFold2[99] (Figure 1.10) suggest it is entirely alpha helical in secondary structure, and structure searching the PDB with this model using Dali also identifies no structurally closely related entries either[104]. The protein does contain a cysteine residue close to its N-terminus

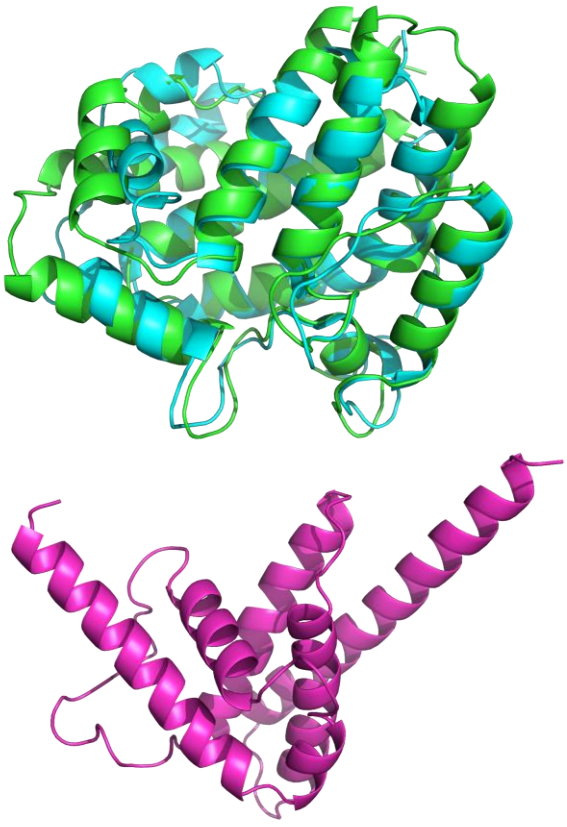


Figure 1.10: Top-Superposition of AlphAfold2 model structure of *S. oneidensis* DmsG (green) with crystal structure of *E. coli* DmsD (cyan, PDB-3CW0) (RMSD = 1.30 Å). Bottom-AlphaFold2 model structure of *S. oneidensis* DmsH.

(Cys27) which might suggest it to be a lipoprotein, however analysis with the SignalP-5.0[105] server implies this is unlikely and therefore is unlikely to be a signal peptide. The role of DmsH in DMSO reduction by *S. oneidensis* therefore remains almost entirely unexplored, which future research should seek to address.

Recently, the putative DmsEFAB complex has been implicated in an additional form of respiration by *S. oneidensis*. *Shin et al* demonstrate the cluster is largely responsible for reduction of the iodate anion by strain MR-1[106]. To do so, they prepared a *ΔdmsB* strain, which they found was severely impaired in iodate reduction. This phenotype could be rescued by complementation with a plasmid *dmsB* copy and furthermore, the culturing of MR-1 in defined media without molybdenum drastically arrested its ability to reduce iodate, suggesting it proceeds through a molybdopterin protein.

These findings suggest the putative DmsEFAB may be promiscuous in terms of its substrate, and therefore might support growth of *S. oneidensis* in diverse environments. Additionally, a satisfactory explanation for why MR-1 has developed a DMSO reductase uniquely predicted to be assembled in its outer membrane has not been made. This represents a challenging task for the cell, since the putative DmsEFAB complex would be some ~230 kDa in size, 50 kDa larger than the MtrCAB complex, and would require numerous and complex translocation processes to achieve assembly in the outer membrane. *Gralnick et al* contemplate three possible explanations that would favour an extracellular DMSO reductase[85]. One is the possibility that the enzyme might support growth on solid DMSO, as is the case for MtrCAB with Fe^{III} oxides, however this would require concentrations of DMSO on the order of >80% which seems unlikely. Furthermore, the exposed hemes of MtrCAB might exchange electrons with minerals, but the active site of *R. sphaeroides* and *R. capsulatus* DMSOR are located at the base of a funnel, which is not likely to be accessible to particulate DMSO suggesting the same is also probable for the MR-1 DMSOR. Another possibility is that some of its substrates or products, might be sources of cytotoxicity. The discovery that the enzyme reduces iodate, as well as DMSO supports this possibility. Finally, extracellular DMSO reduction might be more

rapid than intracellular DMSO reduction, and therefore make *S. oneidensis* more competitive than bacterial strains that produce canonical periplasmic DMSOR enzymes.

Mechanisms of extracellular electron transfer II: *Geobacter sulfurreducens* PCA

After *Shewanella oneidensis* MR-1, *Geobacter sulfurreducens* PCA is the model organism which has seen the most study in the context of EET. Initial isolation of PCA was reported in 1994, from hydrocarbon contaminated groundwater[107]. In that work it was recognized for its respiratory versatility with the authors describing its ability to utilise acetate and hydrogen as its electron source and reduce Fe^{III} citrate, Fe^{III} oxide, elemental sulfur and Co^{III},[107]. Also immediately obvious from the point of isolation onwards, was how important *c*-type cytochromes were likely to be for PCA, as whole cell absorbance spectra revealed features typical of those proteins[107].

G. sulfurreducens was initially understood to be strictly anaerobic and oxygen sensitive although more recent work suggests it can tolerate oxygen at low levels[108]. This bacterium is a gram negative belonging to the deltaproteobacteria group and unlike MR-1, it is prolific in forming thick biofilms[107,109]. PCA is notable for the exceptional conductivity of the thick biofilms it assembles, which has drawn considerable research attention.

Unfortunately, the quantity of material obtainable from biofilms is insufficient for characterisation by many biochemical techniques. Neither is it possible to produce large quantities of PCA biomass via planktonic growth. These factors are further exacerbated by the oxygen sensitivity of the organism, together these features make PCA a very challenging bacterium to study biochemically. Subsequently progress in our understanding the EET molecular mechanisms used by PCA has been difficult to obtain. As noted by many others, the EET pathways used by *G. sulfurreducens* are characterised by a greater level of complexity and apparent redundancy (Figure 1.11), than are those of *S.*

oneidensis[71,110]. This also can make unpicking the function and importance of individual components in its ETCs difficult.

G. sulfurreducens appears to favour acetate as its electron source[111], which it oxidizes via the tricarboxylic acid (TCA) cycle, generating NADH, for which it contains at least one dehydrogenases (Nuo-1) in its genome[111]. Although the same level of detailed characterisation of the ATP generation system as has been discussed for MR-1 has not been described for PCA, the proton pumping Nuo1 complex has been observed directly via proteomics[111], so it is thought most likely that *G.* sulfurreducens does conserve energy via the PMF[110].

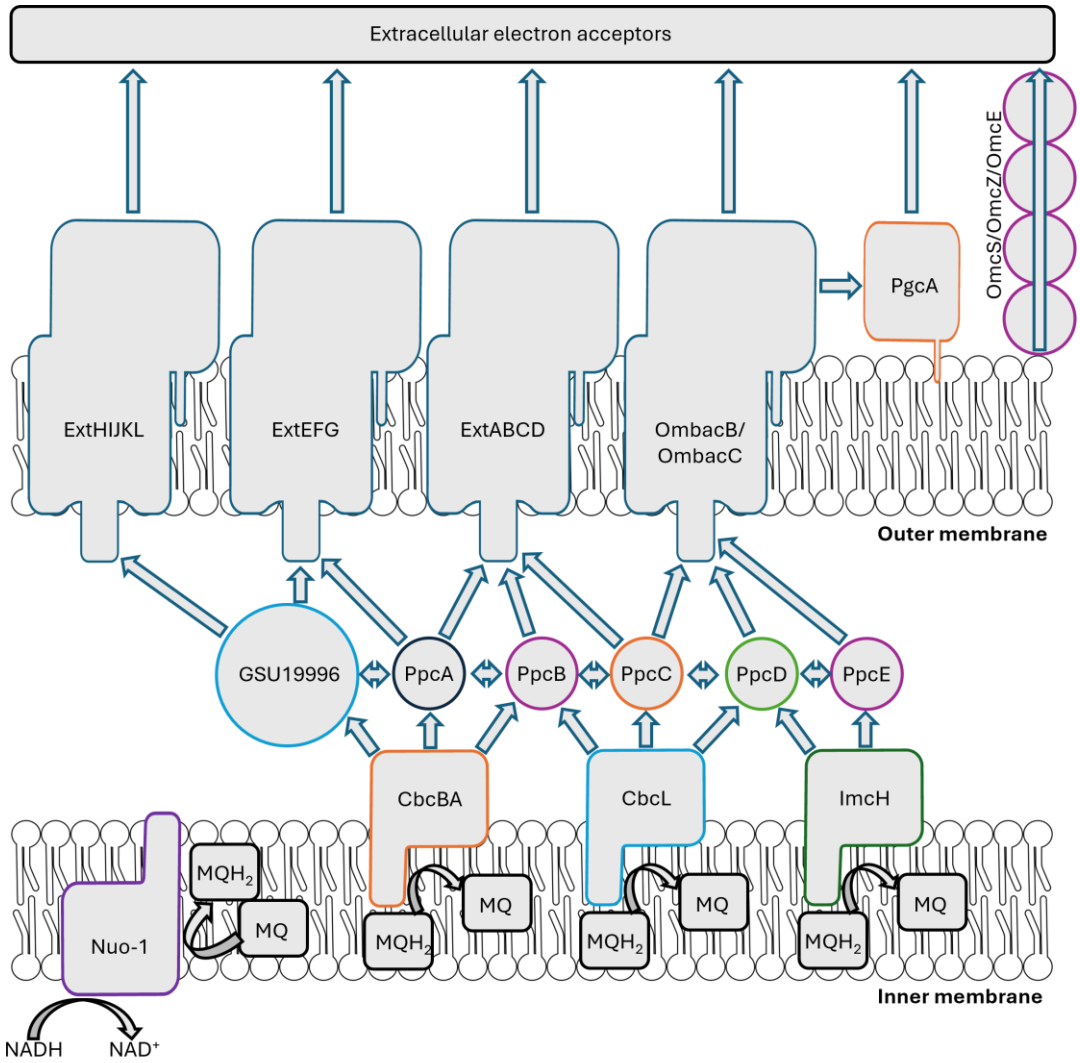


Figure 1.11: Schematic depicting the proposed extracellular electron transfer pathway of the bacterium *G. sulfurreducens* PCA. Protein species shown with coloured outlines and the transfer of electrons with blue outlined arrows. MQ-menaquinone and MQH2-menaquinol. Note that for clarity, the ability to exchange electrons between proteins is not depicted exhaustively. Note most of these interactions remain putative, which applies also about the formation of stable ExtHIJKL, ExtABCD and ExtHIJKL complexes.

It has taken considerable work to decipher the nature of the PCA quinol dehydrogenase system used by PCA during EET. Largely this is because it encodes three different putative quinol oxidase enzymes.

When high potential (~0 mV vs SHE) electron acceptors are available PCA can oxidise quinols via the ImcH protein[112]. This protein is not homologous to CymA from MR-1 and contains seven *c*-type heme attachment motifs distributed about a soluble periplasmic domain, which model structures predicted with AlphaFold2[99] suggest is linked to a membranous domain containing three hydrophobic alpha helices (Figure 1.12). The first clues as to the function of this protein came from a *ΔimcH* strain, which had impaired reduction of high (+240 mV vs SHE) potential electrodes and Mn^{IV} and Fe^{III} citrate but was unimpaired in reduction lower potential Fe^{III} oxide minerals[113].

The next important observation was made during the investigation of a different quinol oxidase, CbcL. The *cbcL* gene encodes nine *c*-type heme attachment motifs in a predicted soluble domain, fused to six trans-inner membrane helices and three cytoplasmic domains and shares[114]. This protein also is not homologous to CymA, but rather to the membrane integral subunit of

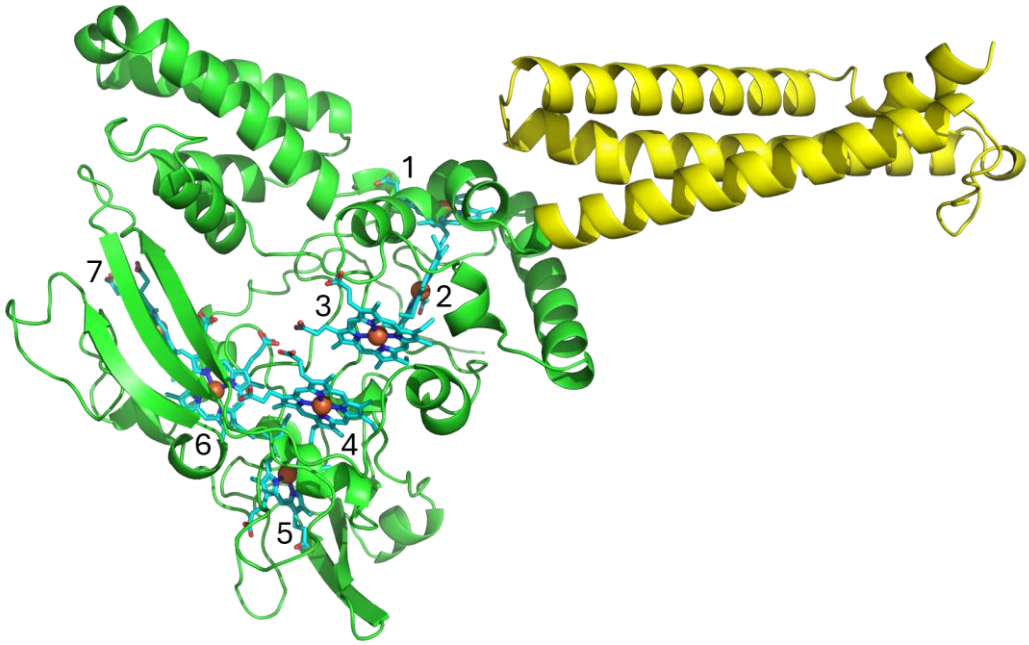


Figure 1.12: AlphaFold2 model structure for *G. sulfurreducens* inner membrane *c*-type cytochrome ImcH. Membrane domain depicted in yellow, periplasm domain as green. Hemes numbered by their CXXCH motif appearance in the protein's amino acid sequence.

menaquinol:formate oxidoreductase complexes. A *ΔcbcL* mutant has been tested under the same conditions as the *ΔimcH* mutant, it displayed no phenotype when grown on the +240 mV electrode[114], additionally the *ΔcbcL* strain actually displayed an increased growth rate over the wild-type[112]. What the *ΔcbcL* strain was impaired in, was reduction of Fe^{III} oxide minerals and electrodes poised at a lower potential (-100 mV vs SHE)[114].

The final piece of information that provided a cohesive and consistent model for quinol oxidation by *G. sulfurreducens*, was that a third enzyme, CbcBA also contributed to quinol oxidation. Unlike *imcH* and *cbcL* which appear to be constitutively expressed, the *cbcB* and *cbcA* genes are definitely not expressed constitutively[115,116]. They are upregulated by 3-4 orders of magnitude during the late stages of batch growth under EET conditions[116]. These genes encode a seven *c*-type heme protein, and a two *b*-type heme protein respectively, importantly *ΔcbcBA* cultures were not able to reduce Fe^{III} as effectively as the wild-type[116]. Wild-type cultures can lower the potential of their growth medium to -270 mV vs SHE, but the *ΔcbcBA* cultures could only reach -210 mV, and were unable to fully reduce a solution of Fe^{III} citrate to 100% Fe^{II}.

The hypothesis presented by *Joshi et al* suggests these observations can be explained by considering the contribution to the PMF that each complex could make in the reduction potential window for which they are respectively active[116]. ImcH, operates at the highest potential, and therefore can access the highest ΔG, it follows that then it could make the largest contribution to the PMF. CbcL operates at intermediate reduction potential, so has the possibility of making an intermediate size contribution to the PMF. CbcBA operates at the lowest potentials where available ΔG for energy conservation is close to zero. Consistent with this they did not observe cells using CbcBA-only as actively growing, but they did remain viable for considerably longer than those without CbcBA[116]. ImcH-only cells boasted the highest growth yields but had the lowest viability in late stages of Fe^{III} reduction.

Direct biochemical support for the explanation provided by *Joshi et al* will be hard to obtain as heterologous expression and characterisation of membrane

cytochromes is challenging, however purification of recombinant ImcH and CbcL has been described and their properties were not inconsistent with the above hypothesis[117,118]. Three oxidases for three different redox potential windows, represents a compelling and elegant solution both from the perspective of surmising the observations described over the past decade[112,114,116,119], and the evolutionary pressure to maximise efficiency in the metabolism of *G. sulfurreducens*. It is remarkable in contrast with the simplistic energy conservation employed by MR-1 which is understood as using CymA as a central and singular hub for quinol oxidation[110].

G,. sulfurreducens PCA encodes more than one hundred genes containing putative *c*-type heme attachment motifs many of which are predicted to have periplasmic localisation, where they shuttle electrons from the quinol oxidases to the outer membrane[120,121]. Amongst these, the PpcA-E family are considered the most important.

Although knockouts of each PpcA-E member display minimal or no discernible phenotypes, a *ΔppcA-E* strain has been reported and displays impaired reduction of Fe^{III} citrate and Fe^{III} oxide[121]. It appears their expression level varies significantly, and this largely explains the difference in Fe^{III} reduction effectiveness of strains with only a single *ppcA-E* gene[121]. Evidently these proteins have significant functional overlap given their ability to compensate for one another's deletion, and this is unsurprising given their structure and other properties are similar. The PpcA-E family are smaller, and simpler than most other cytochromes involved in the EET chains of PCA, and therefore can be purified recombinantly with relative ease[122].

Subsequently high resolution crystal structures for each member PpcA-E are available (Figure 1.13)[122–124]. The PpcA-E family adopt a conserved fold, mostly alpha helical in secondary structure, that encapsulates three *c*-type hemes in close proximity. These hemes are all Bis-His coordinated and like those of other proteins with high heme to amino acid count ratios, are all exposed significantly to the solvent.

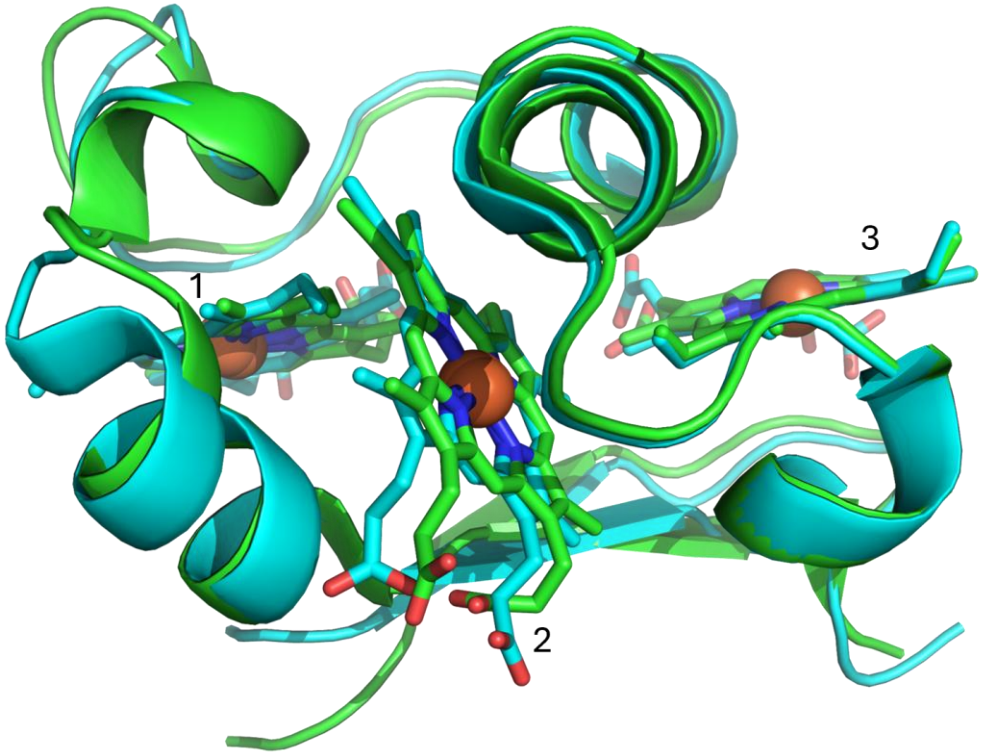


Figure 1.13: Superposition of crystal structures of *G. sulfurreducens* *c*-type cytochromes PpcB (green, PDB-3BXU) and PpcC (cyan, PDB-3H33) (RMSD = 0.54 Å). Hemes numbered by their CXXCH motif appearance in the protein's amino acid sequence.

These proteins also appear to have been assembled into polymeric versions, including the dodecaheme GSU1996. Genetic characterisation of these polymerised periplasmic shuttles has not been reported, but the crystal structure of PpcA can be aligned onto that of GSU1996 with an RMSD (root mean squared deviation) of only 1.8 Å[123,125]. Thus it is likely they can interact with many of the same electron transfer partners and through this also shuttle electrons around the periplasm. What benefit a protein that incorporates four PpcA-E proteins into one chain can confer to PCA remains unclear, but given the bacterium encodes two dodecaheme variants and a twenty seven heme variant, it is likely to be significant[123]. The suggestion which has seen most discussion, is that these proteins act as efficient capacitors, storing many electrons during periods of extracellular electron acceptor scarcity, and discharging them rapidly through the outer membrane upon encounter with a fresh source of oxidant[125].

More perplexing still, is the importance of the triheme *c*-type cytochrome PgcA in the periplasmic electron transfer network of PCA. The localisation of PgcA, is predicted to be extracellular, its sequence includes an N-terminal lipidation

motif similar to the MtrC/MtrF/OmcA/UndA outer membrane cytochromes of *Shewanella* and there is biochemical evidence supporting this predicted localisation[126]. Nonetheless, the *ΔppcA-E* strain described above displayed significant residual Fe^{III} citrate reduction, which could be further reduced, but not eliminated, by the deletion of *pgcA*, in stark contrast with deletion of *pgcA* in a wild-type background which has no effect on Fe^{III} citrate[121,126]. These observations seem to suggest instead it might be a periplasmic electron transporter, however subculturing the *ΔppcA-E* strain eventually yielded a strain that recovered WT Fe^{III} citrate reduction by deleting the *pgcA* lipobox encoding region[121]. This resulted in periplasmic buildup of PgcA, presumably by blocking its translocation across the outer membrane and therefore improving its ability to contribute to the periplasmic shuttling of electrons, demonstrating a significant amount of the protein is extracellularly localised in wild-type cells[121].

Evidently the PgcA protein plays an important and complex role in the EET pathway of *G. sulfurreducens*. These results can be best reconciled by concluding the protein is eventually extracellularly localised, but before or during translocation across the outer membrane, it can contribute to periplasmic shuttling. Data that suggests a cytochrome contributes to electron transfer on both sides of the outer membrane appears to be unique to PgcA, and indicates this protein likely has novel and interesting properties. Thorough examination of current literature in which these properties are explored, is provided in the section discussing the role of PgcA in electron transport in the extracellular space.

Electron transport through the outer membrane of *G. sulfurreducens* is perhaps the least well understood aspect of its EET. At least five gene clusters in the PCA genome are predicted to encode porin-proteins in combination with multiheme c-type cytochrome proteins, that together likely assemble a porin-cytochrome complex analogous to MtrCAB from MR-1 but sharing no sequence identity with it[127].

The first of these to be identified were the putative OmbacB and OmbacC complexes[128]. These clusters, like MtrCAB and MtrDEF, are encoded side-by-

side, and their link with EET was established by extraction of OmcC and OmcB natively from *G. sulfurreducens* membranes after which they were found to reduce Fe^{III} minerals[129]. Closer analysis would reveal they are both ~90 kDa dodecaheme proteins with N-terminal lipidation motifs and that the *omcB* gene in particular was necessary for efficient Fe^{III} reduction[130]. Digesting whole cells with protease K suggested that like MtrC, these proteins are incorporated on the outer facing leaflet of the outer membrane[131].

Sequencing of the *G. sulfurreducens* PCA genome allowed for analysis of the sequences flanking the *omcB* and *omcC* genes, which revealed additional genes adjacent to each, encoding an octaheme *c*-type cytochrome with predicted periplasmic localisation and a putative outer membrane porin[120,128].

Using transposon sequencing the number of putative porin-cytochrome complexes identified and linked to EET in PCA, was expanded to five[132]. In that study the *extABCD* gene cluster proved crucial to reduction of electrodes by PCA. This spurred the identification of two additional gene clusters, *extEFG* and *extHIJKL*.

Thorough genetic characterisation of all five of these clusters has been achieved, as with the periplasmic carriers, single deletions mostly yield very mild phenotypes and full loss of EET only occurred only in the *Δ5* strain[127]. In summary, it appears the *ombacC/ombacB* cluster is most important during growth with Fe^{III} oxide as the terminal electron acceptor, with the *extEFG*, *extHIJKL* and *extABCD* clusters playing decreasingly important roles in that respective order. For soluble Fe^{III} citrate, the *ombacC/ombacB* clusters are also most important. When Mn^{IV} oxide is the terminal electron acceptor all gene clusters except the *extABCD* cluster can sustain near wild-type rates of EET. In contrast with those findings however, the *extABCD* cluster appears totally dominant while the terminal electron acceptor is a graphite electrode poised at +240 mV and deleting the other gene clusters actually increases the current that strain can produce.

These putative protein complexes make for difficult targets in terms of structural characterisation so no experimental structures have been described thus far. Modelling with structure prediction software like AlphaFold can provide some insight, although given none of these genes display homology to the MtrCAB complex, or other structurally characterised protein complexes, it is unclear how accurate those predictions can be[99]. Nonetheless, AlphaFold2 models of the *G. sulfurreducens* predicted porin-cytochrome complexes do suggest that they adopt a structure in which the porin insulates a multiheme cytochrome wire, which feeds electrons to at least one extracellular, lipidated multiheme cytochrome that acts as a terminal reductase analogous to MtrCAB.

What they certainly do not explain, is differing importance of each complex for different substrates, nor the need to encode so many different transmembrane wire complexes with apparently closely overlapping functionalities. The genetic evidence appears to support the hypothesis that complexes differ in the efficiency with which they can reduce particular acceptors, for example Fe^{III} oxide versus Fe^{III} citrate[71,127]. However if that is indeed the case it is unclear why separate porin and transmembrane cytochrome subunits are necessary for each substrate, since it is presumably only the terminal reductase subunit that interacts with substrates and therefore only that subunit which need vary.

The *extHIJKL* cluster seems particularly enigmatic in its role. In part because it encodes two additional genes *extI* and *extJ* that do not appear to contain any canonical motif indicating incorporation of a redox cofactor[71]. Furthermore, the largest gene in the cluster, *extH*, appears also not to contain canonical cofactor motifs either[71]. There is some evidence that suggests this cluster forms a complex which has involvement in selenium and tellurium metabolism, although how this relates to EET is still unclear[133,134].

In conclusion, trans-outer membrane electron transfer by PCA appears also to be characterised by several diverse and functionally overlapping gene clusters. Experimental difficulties have limited biophysical and structural analysis of the proteins encoded in these systems but work in this area is still ongoing.

An interesting example of an EET pathway component from *G. sulfurreducens* that is has been characterised but has no known precedent in MR-1, is the OmcF *c*-type cytochrome[135]. Deletion of *omcF* yields a strain with very poor ability to reduce Fe^{III} citrate[136]. This is interesting given the sequence includes an N-terminal lipidation motif suggesting it associates with the outer membrane which has been experimentally supported. The crystal structure of a recombinant form of OmcF has been solved, which revealed it adopts a globular and alpha helical structure containing a single *c*-type heme which features His-Met coordination of the heme iron[137]. In answer to the question, how can a monoheme membrane protein can make such an important contribution to EET, *Kim et al* prepared northern blots of their *ΔomcF* strain and compared this with the wild-type[136]. They detected RNA originating from the *ombacC/ombacB* gene clusters only in the wild-type indicating that deletion of OmcF also eliminated expression of those genes, which could explain its severely deleterious effect on EET[136].

Two hypothesis might explain these observations, one is that OmcF deletion can cause unexpected changes in expression, in an obscure manner that is not linked to its function. There is relevant precedent for behaviour similar to this description[138]. The other is that OmcF could in fact function as a regulator of EET, this possibility is intriguing given the specificity described above in terms of EET components used by *G. sulfurreducens* growing with differing potential-electrodes or EET substrates[116,128,139]. No mechanistic explanation has been put forth to explain how PCA could exert that control, but a heme based redox centre in the outer membrane could provide the sensing element. At this stage, no mechanism for signal transduction from its redox state to protein expression control is proposed and it is hard to envisage one, so the role of OmcF remains largely still speculative.

The nature of the mechanism(s) by which *G. sulfurreducens* transports electrons from its cell surface, to mineral particles it is not in direct contact with, has been debated intensely over the past decade or so[138].

The beginnings of the controversy lay in experiments that sought to determine whether microorganisms could secrete electron shuttles to assist in EET. These

experiments trapped Fe^{III} oxides in matrices, alginate or agarose beads, that were thought porous enough to allow low molecular weight molecules to pass through, but would not allow significant direct contact between whole cells and the minerals within[140,141]. Such bead-systems were not characterised in depth so the exact size, and size consistency, of the pores is not precisely known, making it unclear whether larger molecules like proteins could traverse the pores[142]. Subsequently more recent research has sought to improve the technique by use of more narrowly defined bead matrices[142]. Nonetheless, unlike *S. oneidensis*, *G. sulfurreducens* and its relatives are regarded as probably not secreting free shuttle compounds like flavin[141,143].

Soon after those studies, efforts were made to determine which, if any, cytochrome proteins might be loosely associated with the extracellular side of the PCA outer membrane, and contributing to EET. Two proteins, OmcS and OmcE which are approximately 50 and 30 kDa in size respectively, could be physically detached from cells by shearing in a commercial blender suggesting they were bound only very loosely[144]. Deleting *omcS* and *omcE* yielded strains with impaired Fe^{III} oxide and Mn^{IV} oxide reduction rates, confirming that they were indeed important for efficient EET[144]. However, the next development would actually come in the form of the discovery that an entirely unrelated protein was also important for EET.

Type IV pili are polymeric filamentous proteins secreted by many bacteria that play important roles in biofilm formation, motility and adhesion[145]. Although the first structure of a type IV pilin, was solved by X-ray crystallographic analysis some thirty years ago[146], they remained a difficult target for structural biologists until recent developments in cryoEM[145] (cryogenic electron microscopy). Invariably, they adopt a helical formation comprised of a C-terminal globular domain, and an N-terminal smaller helical domain[145,147]. *G. sulfurreducens* PCA encodes a type IV pili biosynthesis cassette, although it differs from the canonical version in that its *pilA* gene, which encodes the major filament protein, appeared much shorter than other *pilA* genes[148]. It is now understood that this is because the two domains of the *pilA* gene in PCA have split into two gene-

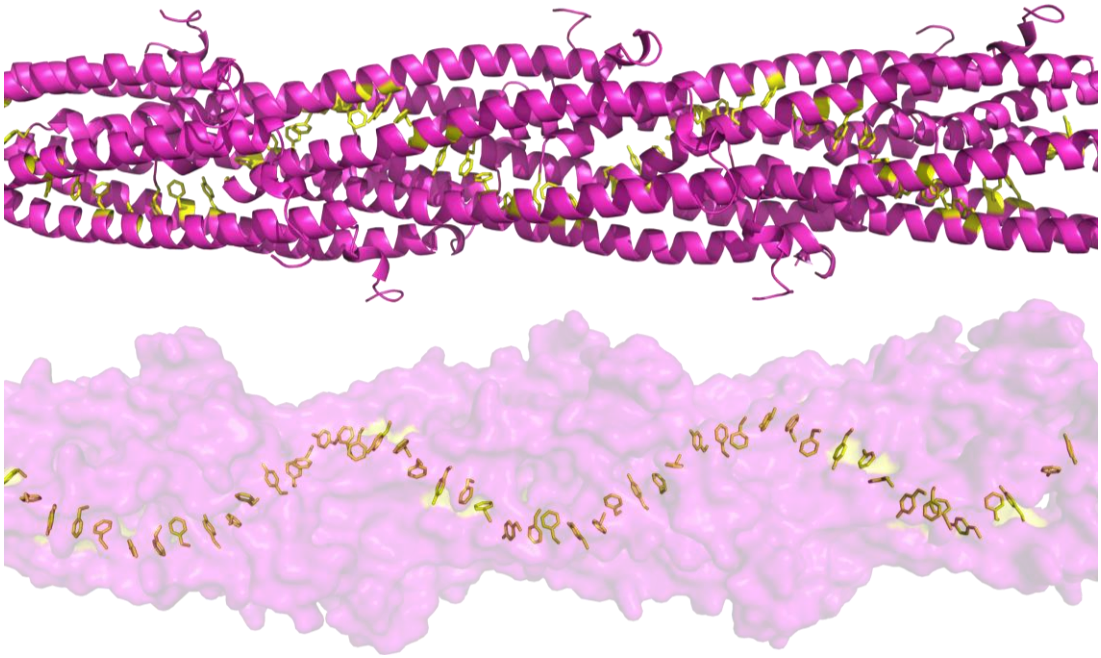


Figure 1.14: Atomic in silico model for *G. sulfurreducens* PilA-N electron conducting filament proposed in ref[150], aromatic residues proposed to confer metallic-like conductivity (Phe1, Phe24 and Tyr27) displayed with yellow sticks.

fragments, a larger C-terminal fragment and a smaller N-terminal one, although this fact was not recognised at the time of those initial studies[149]. Deletion of *pilA* resulted in a severely impaired Fe^{III} reduction phenotype, and this strain lost the ability to produce filaments extending from the cell surface, that were clearly visible in electron micrographs of whole PCA cells[148].

From these observations it was concluded with great excitement, that *G. sulfurreducens* produced filaments of PilA which were very distinct from other type IV pili, and able to facilitate long distance electron transport[148]. Atomic force microscopy utilizing a conductive tip, was applied to these filaments which yielded ohmic current vs voltage plots which at that time were considered strong evidence confirming conductivity of pili-nanowires[148]. Part of the excitement associated with this hypothesis, derived from the fact that the *pilA* gene did not appear to encode any motif indicating the incorporation of a canonical redox cofactor suggesting that if these filaments were indeed conductive, it must be via a mechanism not yet reported in the literature.

A near atomic resolution structure of a pili-nanowire which could be used to infer the mechanism of conductivity remained elusive for many years[150]. In its place, a

computational model of the PilA nanowire was prepared (Figure 1.14), first by obtaining a structure of the PilA protein (now referred to as the N-fragment: PilA-N) by recombinant expression in *E. coli* as a detergent solubilised monomer, solved using NMR spectroscopy[151]. Guided by the structure of the *Neisseria gonorrhoeae* type IV pili, the PilA-N monomeric structure was computationally remodelled as a helical filament[152,153].

Importantly for proponents of pili-nanowire conductivity, this model structure predicted close (~3.5 Å) packing of Phe1, Phe24 and Tyr27 throughout the length of the filament[153]. Thus *Xiao et al* could propose the possibility of a metallic-like conductivity along the long dimension of the filament, which could flow through π orbitals of aromatic amino acids[153]. This proposal appeared consistent with some measurements indicating metallic-conductivity in *G. sulfurreducens* biofilms[154], but it proved to be highly contentious[138]. The substitution for alanine of five aromatic amino acids in the PilA (PilA-N) N-terminus, yielded a strain (Aro5) with impaired reduction of Fe^{III} oxide, which was interpreted also as supporting metallic-conductivity through these residues[155]. Furthermore, the importance of OmcS and OmcE described above, was also reconciled to this theory by proposing these cytochromes associated with the PilA filament, perhaps acting as sites of egress for electrons which was apparently supported by immunogold labelling[156].

Many researchers still maintained important reservations regarding the pili-nanowire hypothesis[157]. Firstly, some groups were unable to reproduce the apparent importance of PilA in the conductivity of *G. sulfurreducens* biofilms or the conductivity of PilA filaments themselves[149,158]. Secondly, evidence in favour of pili-nanowire conductivity continued to be provided via indirect experiments. Immunogold labelling and AFM, are powerful supplementary tools but are themselves either low in resolution in terms of the information they provide or sometimes prone to misleading artifacts[159]. The disappearance of filaments from *ΔpilA* cultures, does not directly prove that the filaments produced by wild-type strains must be comprised of PilA. Similarly, some researchers were wary of conclusions drawn from the Aro5 strain, since aromatic residues have bulky side

chains that often play important roles in protein structure stability[157]. Additionally, a wealth of data indicating PCA synthesises several different, conductive, but cytochrome-based filaments was emerging[160]. Near-atomic resolution cryoEM structures for these filaments provided clear evidence for conductivity via well precented cofactor chains, for detailed discussion of these species, please see the section of this work which succeeds description of research on pili-nanowires[160]. Finally, the model structure on which the metallic-conductivity hypothesis rested was ultimately not empirically determined and was therefore without experimental validation.

Eventually *Gu et al* reported enriching a pili-filament preparation by deleting *omcS* which resulted in overexpression of the PilA-N fragment, but also the PilA-C fragment[149]. From this preparation they were able to solve a helical reconstruction (Figure 1.15) of a filament which incorporated both the PilA-N and PilA-C fragments into a filament with more similarity to canonical type IV pili, than the model presented by *Xiao et al*[149,153]. This structure showed the PilA filament is composed of an alpha helical core consisting of PilA-N, which forms numerous electrostatic and hydrophobic interactions with globular, beta-sheet rich, PilA-C which coats the outer surface of the filament shielding hydrophobic

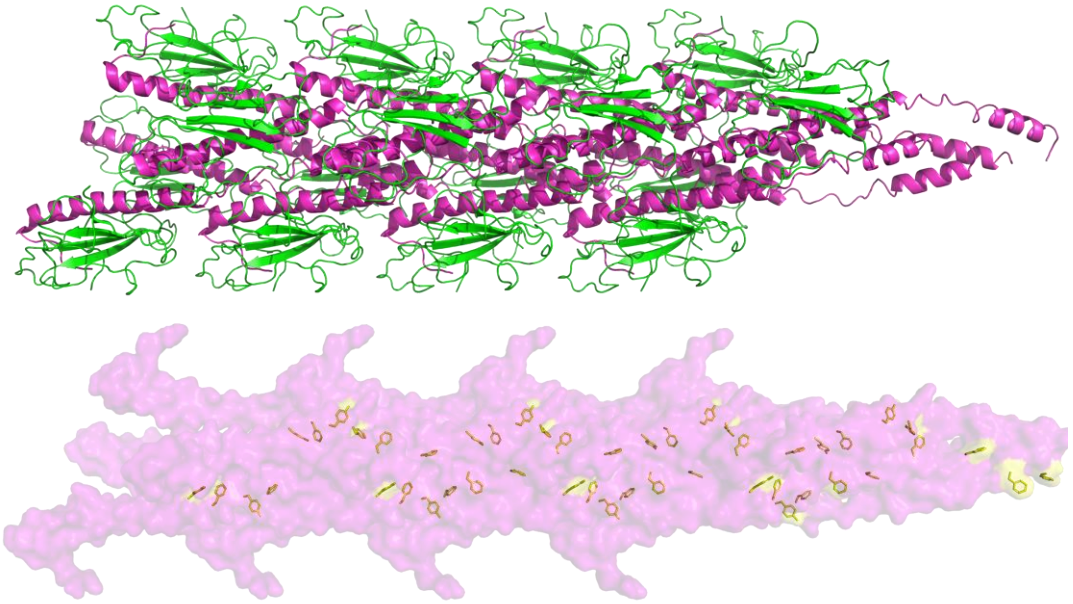


Figure 1.15: CryoEM structure of *G. sulfurreducens* PilA-N/C filament (PDB-6VK9). N subunit (magenta) and C subunit (green), N subunit aromatic residues proposed to confer metallic-like conductivity (Phe1, Phe24 and Tyr27) displayed with yellow sticks.

PilA-N from aqueous solvent. Crucially the position of aromatic side chains differed greatly from that suggested in the model structure based on the detergent solubilised PilA-N monomer, rather than 3.5 Å apart in fact they were ~10 Å apart suggesting metallic like conductivity through π orbitals was unlikely[149]. This discovery proved to be a crucial turning point for many researchers, making it very difficult to argue in favour of conductive pili-nanowires which underscores this importance of the protein structure-function relationship in the study of EET[160,161].

What the structure did not entirely resolve however was if pili-filaments were not contributing to EET by conducting electrons away from the cells surface, why *ΔpilA-N* strains perform poorly at EET, and similarly why the Aro5 strain could not perform EET effectively. An interesting possible explanation *Gu et al* provide, is that since the successful localisation of cytochrome nanowires like OmcS and OmcZ is halted by the deletion of PilA-N, perhaps PilA-N/C performs an indispensable secretory role for those nanowires which in turn confer the conductivity to the PCA biofilm[149]. This explanation is itself currently without direct empirical validation which will be required before its universal acceptance, however it is consistent with the features of the PilA-N/C filament structure that are more akin to secretory-pseudopili than type-4 pili[149,162].

What is now accepted without controversy, is the importance of the cytochrome polymeric filaments in the EET pathway of *G. sulfurreducens*. Although their apparently overlapping functionalities has made it difficult to determine their specific roles, they are definitely not identical.

The first such nanowire to see detailed characterisation was OmcS. OmcS is one of the most abundant, if not the most abundant, extracellular cytochrome in many *G. sulfurreducens* cultures[163]. The *omcS* gene encodes six CXXCH *c*-type heme attachment motifs and as noted previously its deletion severely impairs reduction of Fe^{III} oxide[163]. The protein appears to survive harsh treatment with sodium dodecyl sulphate (SDS) and high temperatures which allowed it to be extracted[164]. That material displayed electron paramagnetic resonance (EPR), NMR and UV-visible spectra consistent with low spin His-His coordinated hemes

and was redox active between −360 and −40 mV vs SHE, all features that are unremarkable in comparison with other multiheme cytochrome proteins[164].

The structure of OmcS, reported in 2019, for the first time demonstrated that *c*-type cytochromes could be assembled into polymeric filaments extending micrometres in length (Figure 1.16)[165]. Its six hemes were indeed His-His coordinated and packed together with very close (≤6 Å) edge-edge distances even between adjacent monomeric units immediately providing a plausible mechanism for the rapid electron transfer along the length of the filament reported alongside the structure. The OmcS filament has a ~40 Å cross section, and encapsulates its hemes in a polypeptide backbone that features a few short alpha helices but is largely comprised of loops. *Wang et al* calculated that only ~19% of residues are found within regions of secondary structure, which is exceptionally low in comparison with other proteins, be they multiheme cytochromes or even other polymeric filaments.

Soon after, the structure of a second cytochrome filament was described[166]. OmcE like OmcS, has also been recognised as a major component of the biofilm material secreted *G. sulfurreducens*[144]. Unlike OmcS, OmcE however is a tetraheme, although it does feature exclusively His-His coordination of its *c*-type hemes. OmcE also differs from OmcS in two other important ways, firstly this filament is considerably smaller than OmcS, with a cross section of ~30 Å. Secondly OmcE appears to undergo an extensive glycosylation process at as many as five sites, which is supported by the inclusion of numerous genes encoding for glycosyltransferases and sugar transporters in close vicinity to the *omcE* gene[166]. A number of putative functions have been proposed for these glycan modifications, although all must be reconciled with the observation that whatever they confer, is not a necessity for all cytochrome filaments as OmcS is produced without them[165].

OmcE and OmcS are not sequentially homologous, and their peptide backbone folds do not share significant similarities either. In spite of this, the OmcE hemes can be superimposed onto the first four hemes of OmcS with approximately ~1 Å

RMSD[165,166] which at the reported resolution (3.7 and 4.3 Å) of these reconstructions, lies within experimental error. The implication of this observation was that the amino acid sequence of the cytochrome filament matters very little, as long as it achieves the canonical heme stacking pattern.

This hypothesis would see revision as the determination of a third cytochrome filament structure, OmcZ, revealed that even the heme packing arrangement seen in OmcS and OmcE is not a shared feature of all filaments[167]. The *omcZ* gene encodes a ~50 kDa protein with eight CXXCH *c*-type heme attachment motifs[168]. Prior to the determination of its structure, measurements of the conductivity of the OmcZ filament had already suggested it was significantly more conductive than OmcS or OmcE[169]. Another interesting feature of OmcZ, was that its dominant form appeared to be only 30 kDa in size, a full 20 kDa less than the prediction for the nascent polypeptide based upon its gene sequence[170].

OmcZ was revealed to adopt a structure which like the other structurally characterised nanowires contained very little defined secondary structure, but otherwise differed from them greatly[168]. The hemes of OmcZ are all His-His coordinated although unlike OmcS/E, the ligating histidine residues are found exclusively within the same monomer as the

Figure 1.16: CryoEM structure of *G. sulfurreducens* polymeric *c*-type cytochrome filament OmcS (PDB-6EF8).

heme they coordinate. It appears not to be glycosylated although the structure was consistent with the protein having undergone the proposed proteolytic modification. The packing of the OmcZ hemes brings them somewhat closer together than those of OmcE/S and leaves them significantly more solvent exposed, in particular heme six is positioned almost abreast of the main heme chain extending out into the solvent. One or both of these properties have been proposed as potential explanations for the higher measured conductivity of OmcZ[167,168].

Arguably of most relevance to the results described in the current work, is the maturation mechanism of OmcZ, for which the cryoEM structure provided important insight. Adjacent to the *omcZ* gene, is encoded a serine protease, denoted *ozpA*. The ~500 amino acid OzpA protein is designated as extracellular by localisation prediction software and was identified as homologous to Subtilase enzymes[171]. Those enzymes often include an autoinhibitory domain which can be cleaved, activating proteolytic catalysis via a conserved catalytic triad (Figure 1.17)[172]. Elements homologous to both of these features were identified also in the OzpA sequence[171]. *Kai et al* prepared an Δ*ozpA* strain which proved unable to process full length OmcZ into its smaller isoform, and they showed recombinantly expressed OzpA in cell extracts could perform the transformation, demonstrating OzpA is the OmcZ-maturing protease[171]. *Gu et al* generated an AlphaFold2 model of full length OmcZ, which predicted that much

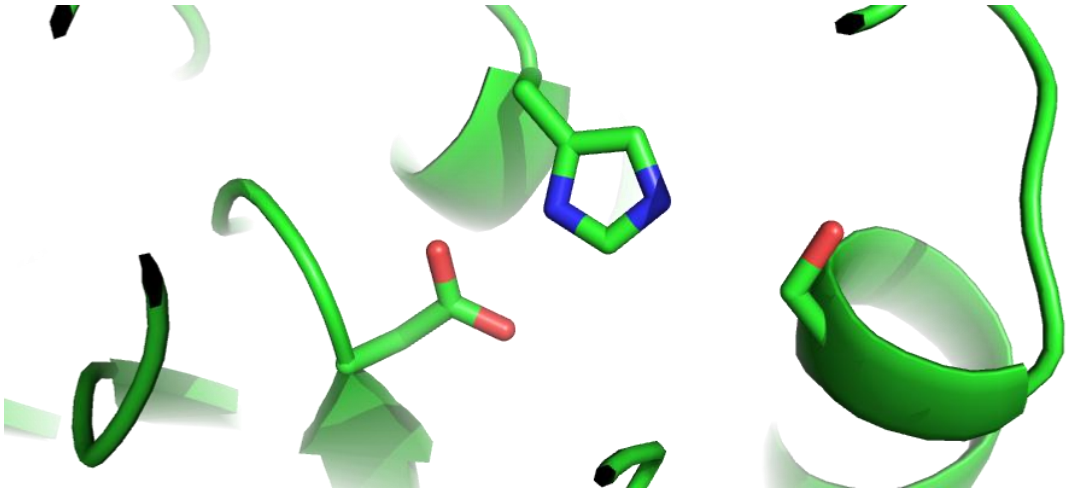


Figure 1.17: AlphaFold2 model, showing conserved putative catalytic triad (D133, H169 and S322) in sticks, of *G. sulfurreducens* serine protease OzpA.

of the 20 kDa that is ultimately cleaved from the protein, in the pre-cleavage state, resides at the location their cryoEM structure showed is the interaction interface between adjacent filament monomers[168]. This modelling therefore suggested the proteolysis represents a means of controlling filament assembly, which they supplemented with an additional model of the OzpA-OmcZ complex which remarkably predicted an interaction that positioned the OzpA catalytic triad directly in contact with the OmcZ protein in a conformation that appears consistent with proteolysis[168].

This putative mechanism for controlling OmcZ filament assembly is consistent with the reported observations in the literature, elegant and remarkably sophisticated. Given, the evolution of a specific mechanism to ensure OmcZ nanowires are not mistakenly assembled within the periplasm, it begs the unanswered question, why do OmcE and OmcS not similarly require mechanisms controlling their polymerisation. One explanation could be that their assembly mechanism differs significantly from OmcZ, perhaps those filaments will not assemble spontaneously and therefore requires assistance than can be withheld until ready. If so, such assembly factors have not yet been identified.

Little else is known about the assembly process of cytochrome nanowires, largely due to experimental difficulties associated with observing that process. With improvements in cryoEM and structure-prediction models this will hopefully be an area in which coming years see exciting developments.

Determining the relative roles for each of the cytochrome nanowires has proved challenging, as is the case for other cytochromes associated with EET in *G. sulfurreducens* they appear to overlap significantly in their functions. Knockouts that yield impaired EET of some kind, have been described in each of *omcS/E/Z* and an additional less well characterised putative cytochrome filament *omcT* demonstrating they are critical[163,173–175].

When interpreted in the context of the fact that deletions of genes encoding putative porin cytochrome complexes (PCCs) also impairs EET[127], this suggests

a cooperative role for the cytochrome filaments and PCCs rather than parallel. A possible interpretation of a complementary role for these two systems is that the PCCs might perform the 'uploading' of electrons into the filament heme chain. This explanation is probably inconsistent with the alternative hypothesis that the loading of electrons into the filament occurs within the periplasm, considered the most likely site of their assembly[149]. As with the mechanism of filament-assembly, the 'uploading' remains largely unexplored due to experimental difficulties in observing the process *in vivo*. Improved cryoEM and protein complex structure-prediction methodologies may also soon provide the first means to test these hypotheses.

The discovery that *G. sulfurreducens* synthesises conductive polymeric filaments comprised of redox proteins was unprecedented, and was made possible in large part by technological advancements in cryoEM. The possibility that these filaments are widespread amongst microorganisms, and have been present but undetected in preparations dating back many years, was a real one.

In answer to this question, *Gu et al* presented phylogenetic analysis which shows OmcZ in particular is widespread amongst diverse bacteria including in genera that aren't usually aren't considered to be core EET participants like methanogens[168]. Direct experimental confirmation of this has been provided now by *Baquero et al*, who solved cryoEM structures of two cytochrome-based filamentous nanowires produced by the archaea *Pyrobaculum calidifontis* and *Archaeoglobus veneficus*[176]. These structures shared no similarities with the *G. sulfurreducens* filaments with the exception of the position of their hemes, which can be aligned onto those of OmcS and OmcE with an RMSD of ~1 Å, but not onto those of OmcZ[176]. Thus it does seem likely that cytochrome filaments are distributed widely amongst diverse microbial taxa which only further cements the importance of their continued study.

A final question relevant to contemplate, is whether the cytochrome nanowires synthesised by *G. sulfurreducens* interact directly with mineral oxide particles, or if other electron carriers 'download' electrons from the heme chain and perform the reduction of mineral particles. All but the smallest particles comprised of iron

oxides will exhibit poor penetrance for electrons making them hard to study by transmission electron microscopy (TEM), which is necessary for single particle analysis. *In vitro*, cytochrome filaments have been shown to reduce Fe^{III} oxide[164], however this doesn't necessarily mean that such activity predominates *in vivo*.

One possibility is that accessory biomolecules shuttle electrons from the nanowire heme chain to mineral particles, here it is relevant to note that as described above, Geobacter species are thought probably not to secrete flavin-based electron shuttles[71,143]. A candidate still considered plausible for the 'uploading', 'downloading' or both, of electrons from the *G. sulfurreducens* nanowires could be the protein PgcA.

PgcA appears first to have been linked to EET in 2008, by *Ding and coworkers*.[115] In that work, LC-MS was used to determine the relative abundances of proteins within *G. sulfurreducens* cultures grown on Fe^{III} oxide relative to Fe^{III} citrate, and thereby identify proteins that play a specialised role in Fe^{III} oxide reduction[115]. Amongst the ~270 proteins identified, were several already discussed in the current work including the PilA-N protein and the OmcS cytochrome filament protein. Also more abundant during Fe^{III} oxide reduction, was the protein encoded by the gene GSU1761. This gene encodes a protein with three putative CXXCH *c*-type heme attachment motifs and a theoretical molecular weight of ~50 kDa[115].

Ding et al note the subcellular localisation of the GSU1761 gene product as periplasmic based upon their earlier work[115,177], which sought to determine the subcellular localisation of the *G. sulfurreducens* proteome using a similar approach involving fractionation, tryptic digestion and LC-MS. As noted by the authors, their fractionation procedure did not perfectly separate subcellular localities and naturally, proteins destined for the extracellular space must first translocate the inner membrane, periplasm and outer membrane before reaching their ultimate destination[177]. These issues are well illustrated by OmcE, which as described above has now been shown to form an extracellular filament, while the original analysis by *Ding et al* predicts it is localised to the periplasm[166,177].

PgcA was next reported in the literature in three related articles from the *Lovley* group published in 2009, 2011 and 2014. In the first of these, *Yi et al* describe developing a new strain of *G. sulfurreducens* by culturing it upon a graphite electrode poised at −400 mV for five months[178]. This strain they designated *KN400* and it proved capable of producing current densities well above that of the original strain[178].

This was followed by *Tremblay et al*, in which a similar strategy was used to produce strains with improved EET ability by continual culturing, except in this instance using Fe^{III} oxide as its terminal electron acceptor[179]. Both strains displayed dramatically improved rates of Fe^{III} reduction over the starting strain, and resequencing of their genomes demonstrated they both had obtained mutations ~100 bp upstream of the GSU1761 start codon, amongst several other mutations. These yielded ~10 and ~20 x upregulation of GSU1761 assessed by mRNA transcript abundance, which presumably was responsible for the increase in rate of Fe^{III} oxide reduction. In this article, GSU1761 was named *pgcA* and its protein product PgcA, for *P*eriplasmic, *G*enes related to the environment, membranes and motility riboswitch (a DNA sequence between the pgcA promoter and stat codon which likely exerts control over its expression), *C*ontrolled, *A*[179].

In the final of these three manuscripts, *Smith et al* report first the deletion of the *pilA-N* gene in the *KN400* strain, which as discussed in this work previously, yields a strain with dramatically impaired Fe^{III} oxide reduction[180]. They then continually cultured this Δ*pilA-N KN400* with Fe^{III} oxide as its terminal electron acceptor, and found eventually it recovered its ability to grow by reducing that mineral at a rate comparable to the original *KN400*[180]. This new strain they designated *JS-1*, was determined also to have acquired a mutation ~100 bp upstream of the PgcA start codon, and contained an mRNA transcript abundance of GSU1761 ~57 x that of the original *KN400* strain.

Smith et al recognised that PgcA is unlikely to be ultimately localised to the periplasm, based on analysis performed with the PSORTb subcellular localization prediction tool which instead suggested it is translocated into the

extracellular space[180,181]. Thus they analysed cell-free filtrates of their cultures by SDS-PAGE (sodium dodecyl sulfate polyacrylamide gel electrophoresis) with heme specific staining, a single band only visible in the Δ*pilA-N KN400* sample was obvious. This band appeared at a molecular weight of ~40 kDa, and was identified as PgcA by LC-MS, likely by tryptic digest fragment mapping, although no deconvoluted mass spectra for fragments were provided on publication. This observation is interesting as this ~40 kDa band is 10 kDa smaller than the theoretical molecular weight predicted for the mature PgcA sequence, possibly indicating truncation. Truncation would also be consistent with large quantities of PgcA in cell-free filtrates, in spite of the fact its gene sequence encodes an N-terminal, canonical lipidation motif like that of the MtrC/OmcA/MtrF/UndA family, which should render the protein tethered to the hydrophobic membrane.

Aklujkar et al would add to these findings with transcriptomic analysis of wild-type *G. sulfurreducens* under EET growth[182]. Their analysis also supported increased abundance of OmcS, PilA-N and PgcA amongst many others. Surprisingly, they observed a large increase in GSU1761 transcript numbers while *G. sulfurreducens* grew with Mn^{IV} oxide but not with Fe^{III} oxide. In aggregate, these initial results strongly suggested PgcA can play an important role in the EET process of G. sulfurreducens at least under some conditions.

The first, and seminal, biochemical analysis of PgcA was provided *Zacharoff et al*[126]. They first analysed its amino acid sequence which is 511 residues in length, and includes a Sec-type secretion signal sequence at its N-terminus. This follows into a lipidation motif centred on residue Cys39, the predicted site of lipidation as is described for other extracellular non-filamentous cytochromes[55]. The next ~200 amino acids were featureless, with the C-terminal half of the protein containing three CXXCH *c*-type heme attachment motifs separated by two large (30 amino acids) Pro-Thr repeat motifs with no precedent in the literature at least on this scale, and one much smaller Pro-Thr motif. *Zacharoff et al* present bioinformatic analysis which demonstrates these features are conserved, although they note that in some species the Pro-Thr motifs are instead Pro-Ala motifs.

Next they prepared a *ΔpgcA G. sulfurreducens* strain, and tested its ability to perform EET. Interestingly, it showed no phenotype during Fe^{III} citrate reduction, or reduction of an electrode poised at +240 mV, but was drastically impaired in its reduction of Fe^{III} oxide relative to the wild-type[126]. This phenotype could be partially recovered by expression of a plasmid *pgcA* copy, and suggested a specific role for PgcA in Fe^{III} oxide EET, possibly as the terminal reductase that reduces it.

The authors report cloning *pgcA* into a pBAD based overexpression system and incorporating a hexahistidine purification tag to facilitate production of recombinant PgcA in *S. oneidensis*. With this system, they were able to purify PgcA, although SDS-PAGE revealed it was recovered in two forms, a larger species which judged by SDS-PAGE appeared at a molecular weight of ~60 kDa and a smaller species at ~40 kDa[126]. Both isoforms displayed UV-visible absorbance spectra typical of *c*-type cytochromes, and when mixed with freshly synthesised Fe^{III} oxide and centrifuged, were removed from solution by the mineral particles[126]. This last property was not shared by bovine serum albumin or equine heart cytochrome-*c* and it was also not observed for PgcA when instead incubated with a suspension of magnetite, a mixed Fe^{II}/Fe^{III} oxide[126]. Finally, *Zacharoff et al* added their purified PgcA exogenously to cultures of their *ΔpgcA* strain, and observed that this rescued their Fe^{III} oxide reduction rate to that even beyond the supplemented wild-type cultures.

Taken together their findings indicate PgcA plays a specialised role in facilitating EET with Fe^{III} oxide as the substrate, which is critical at least under some conditions. The protein is likely to perform its function extracellularly, as addition of exogenous recombinant PgcA restored EET to *ΔpgcA* cells. It appears the protein may undergo post translational proteolysis, similar to OmcZ although as *Zacharoff et al* note, unless this is carefully controlled, secreting a 50 kDa cytochrome would likely represent a large energetic burden upon the cell[126]. If indeed it is, neither the site nor mechanism of proteolysis have been unambiguously determined. The mechanism of Fe^{III} oxide reduction by PgcA also is not defined, nor is the proteins structure. Repeat motifs are usually associated

with disordered regions within proteins, but they can also drive associations with solid phase substrates as in ice-binding proteins[183]. *Zacharoff et al* contemplate the possibility that analogously to ice-binding proteins, PgcA binds Fe^{III} oxides through its Pro-Thr repeat motifs[126]. What implications this would have for the internal redox-cofactor pathway within PgcA remains unclear. It is similarly unclear what the implication of its heme binding motifs being separated by the Pro-Thr motifs have for the structure of the protein, and its mechanism.

Aims of this thesis

This work set out to characterise the properties of PgcA by structural, biophysical and electrochemical techniques. Its particular focus was to describe in depth, the 3D molecular structure PgcA adopts, and understand how that arises from its unique amino acid sequence (Chapter 3). It also then sought to unpick how its unique 3D structure, and other properties, impart function to PgcA in performing reduction of mineral substrates in the extracellular environment (Chapter 4). This work aimed to provide a basis for understanding post-translational modification of PgcA, and how it too might affect its function (Chapter 5).

This work also sought to provide biochemical and structural characterisation of the *S. oneidensis* MR-1 DmsEFAB complex. No definitive purification of the complete complex has been described, and thus this was made an important goal here. This work then aimed to then determine the structure of DmsEFAB by cryogenic electron microscopy and thereby elucidate the molecular mechanism of DMSO reduction by MR-1 (Chapter 6).

Chapter 1: References

1. Bushnell, G. W., Louie, G. V. & Brayer, G. D. High-resolution three-dimensional structure of horse heart cytochrome c. *Journal of Molecular Biology* **214**, 585–595 (1990).
2. Anfinsen, C. B. Principles that govern the folding of protein chains. *Science* **181**, 223–230 (1973).
3. Bamford, V. A., Angove, H. C., Seward, H. E., Thomson, A. J., Cole, J. A., Butt, J. N., Hemmings, A. M. & Richardson, D. J. Structure and Spectroscopy of the Periplasmic Cytochrome c Nitrite Reductase from Escherichia coli†. *Biochemistry* **41**, 2921–2931 (2002).

4. Salgueiro, C. A., Morgado, L., Silva, M. A., Ferreira, M. R., Fernandes, T. M. & Portela, P. C. From iron to bacterial electroconductive filaments: Exploring cytochrome diversity using Geobacter bacteria. *Coordination Chemistry Reviews* **452**, 214284 (2022).

5. Salemme, F. R. STRUCTURE AND FUNCTION OF CYTOCHROMES C. *Annual Review of Biochemistry* **46**, 299–329 (1977).

6. Mitchell, P. Coupling of phosphorylation to electron and hydrogen transfer by a chemi-osmotic type of mechanism. *Nature* **191**, 144–148 (1961).

7. Bekker, A., Holland, H. D., Wang, P. L., Rumble, D., Stein, H. J., Hannah, J. L., Coetzee, L. L. & Beukes, N. J. Dating the rise of atmospheric oxygen. *Nature 2004 427:6970* **427**, 117–120 (2004).

8. Blankenship, R. E. Molecular evidence for the evolution of photosynthesis. *Trends in Plant Science* **6**, 4–6 (2001).

9. Coelho, C. & Romão, M. J. Structural and mechanistic insights on nitrate reductases. *Protein Science* **24**, 1901 (2015).

10. Fritz, G., Roth, A., Schiffer, A., Büchert, T., Bourenkov, G., Bartunik, H. D., Huber, H., Stetter, K. O., Kroneck, P. M. H. & Ermler, U. Structure of adenylylsulfate reductase from the hyperthermophilic Archaeoglobus fulgidus at 1.6-Å resolution. *Proceedings of the National Academy of Sciences of the United States of America* **99**, 1836–1841 (2002).

11. Santos, A. A., Venceslau, S. S., Grein, F., Leavitt, W. D., Dahl, C., Johnston, D. T. & Pereira, I. A. C. A protein trisulfide couples dissimilatory sulfate reduction to energy conservation. *Science* **350**, 1541–1545 (2015).

12. Shi, L., Dong, H., Reguera, G., Beyenal, H., Lu, A., Liu, J., Yu, H. Q. & Fredrickson, J. K. Extracellular electron transfer mechanisms between microorganisms and minerals. *Nature Reviews Microbiology 2016 14:10* **14**, 651–662 (2016).

13. Slate, A. J., Whitehead, K. A., Brownson, D. A. C. & Banks, C. E. Microbial fuel cells: An overview of current technology. *Renewable and Sustainable Energy Reviews* **101**, 60–81 (2019).

14. Boas, J. V., Oliveira, V. B., Simões, M. & Pinto, A. M. F. R. Review on microbial fuel cells applications, developments and costs. *Journal of Environmental Management* **307**, 114525 (2022).

15. Ishii, S., Watanabe, K., Yabuki, S., Logan, B. E. & Sekiguchi, Y. Comparison of electrode reduction activities of Geobacter sulfurreducens and an enriched consortium in an air-cathode microbial fuel cell. *Applied and Environmental Microbiology* **74**, 7348–7355 (2008).

16. Rabaey, K. & Rozendal, R. A. Microbial electrosynthesis — revisiting the electrical route for microbial production. *Nature Reviews Microbiology 2010 8:10* **8**, 706–716 (2010).

17. Jourdin, L. & Burdyny, T. Microbial Electrosynthesis: Where Do We Go from Here? *Trends in Biotechnology* **39**, 359–369 (2021).

18. Yang, H., Zhou, M., Liu, M., Yang, W. & Gu, T. Microbial fuel cells for biosensor applications. *Biotechnology Letters* **37**, 2357–2364 (2015).

19. Adekunle, A., Raghavan, V. & Tartakovsky, B. On-line monitoring of heavy metals-related toxicity with a microbial fuel cell biosensor. *Biosensors and Bioelectronics* **132**, 382–390 (2019).

20. Schneider, G., Czeller, M., Rostás, V. & Kovács, T. Microbial fuel cell-based diagnostic platform to reveal antibacterial effect of beta-lactam antibiotics. *Enzyme and Microbial Technology* **73–74**, 59–64 (2015).

21. Kim, M., Sik Hyun, M., Gadd, G. M. & Joo Kim, H. A novel biomonitoring system using microbial fuel cells. *Journal of Environmental Monitoring* **9**, 1323–1328 (2007).

22. Johnson, D. B. Biomining — biotechnologies for extracting and recovering metals from ores and waste materials. *Current Opinion in Biotechnology* **30**, 24–31 (2014).

23. Maes, S., Props, R., Fitts, J. P., De Smet, R., Vanhaecke, F., Boon, N. & Hennebel, T. Biological Recovery of Platinum Complexes from Diluted Aqueous Streams by Axenic Cultures. *PLOS ONE* **12**, e0169093 (2017).

24. Saitoh, N., Nomura, T. & Konishi, Y. Biotechnological Recovery of Platinum Group Metals from Leachates of Spent Automotive Catalysts. *Minerals, Metals and Materials Series* **Part F3**, 129–135 (2017).

25. Dundas, C. M., Graham, A. J., Romanovicz, D. K. & Keitz, B. K. Extracellular Electron Transfer by Shewanella oneidensis Controls Palladium Nanoparticle Phenotype. *ACS Synthetic Biology* **7**, 2726–2736 (2018).

26. Xu, H., Xiao, Y., Xu, M., Cui, H., Tan, L., Feng, N., Liu, X., Qiu, G., Dong, H. & Xie, J. Microbial synthesis of Pd–Pt alloy nanoparticles using Shewanella oneidensis MR-1 with enhanced catalytic activity for nitrophenol and azo dyes reduction. *Nanotechnology* **30**, 065607 (2018).

27. You, W., Peng, W., Tian, Z. & Zheng, M. Uranium bioremediation with U(VI)-reducing bacteria. *Science of The Total Environment* **798**, 149107 (2021).

28. Cai, P. J., Xiao, X., He, Y. R., Li, W. W., Chu, J., Wu, C., He, M. X., Zhang, Z., Sheng, G. P., Lam, M. H. W., Xu, F. & Yu, H. Q. Anaerobic biodecolorization mechanism of methyl orange by Shewanella oneidensis MR-1. *Applied Microbiology and Biotechnology* **93**, 1769–1776 (2012).

29. Huwiler, S. G., Löffler, C., Anselmann, S. E. L., Stärk, H. J., Von Bergen, M., Flechsler, J., Rachel, R. & Boll, M. One-megadalton metalloenzyme complex in Geobacter metallireducens involved in benzene ring reduction beyond the biological redox window. *Proceedings of the National Academy of Sciences of the United States of America* **116**, 2259–2264 (2019).

30. Myers, C. R. & Nealson, K. H. Bacterial Manganese Reduction and Growth with Manganese Oxide as the Sole Electron Acceptor. *Science* **240**, 1319–1321 (1988).

31. Hunt, K. A., Flynn, J. M., Naranjo, B., Shikhare, I. D. & Gralnick, J. A. Substrate-level phosphorylation is the primary source of energy conservation during anaerobic respiration of Shewanella oneidensis strain MR-1. *Journal of Bacteriology* **192**, 3345–3351 (2010).

32. Madsen, C. S. & TerAvest, M. A. NADH dehydrogenases Nuo and Nqr1 contribute to extracellular electron transfer by Shewanella oneidensis MR-1 in bioelectrochemical systems. *Scientific Reports 2019 9:1* **9**, 1–6 (2019).

33. Kasai, T., Suzuki, Y., Kouzuma, A. & Watanabe, K. Roles of d-Lactate Dehydrogenases in the Anaerobic Growth of Shewanella oneidensis MR-1 on Sugars. *Applied and Environmental Microbiology* **85**, (2019).

34. Rodrigues, M. L., Oliveira, T. F., Pereira, I. A. C. & Archer, M. X-ray structure of the membrane-bound cytochrome c quinol dehydrogenase NrfH reveals novel haem coordination. *EMBO Journal* **25**, 5951–5960 (2006).

35. Rodrigues, M. L., Scott, K. A., Sansom, M. S. P., Pereira, I. A. C. & Archer, M. Quinol Oxidation by c-Type Cytochromes: Structural Characterization of the Menaquinol Binding Site of NrfHA. *Journal of Molecular Biology* **381**, 341–350 (2008).

36. Myers, C. R. & Myers, J. M. Cloning and sequence of cymA, a gene encoding a tetraheme cytochrome c required for reduction of iron(III), fumarate, and nitrate by Shewanella putrefaciens MR-1. *Journal of Bacteriology* **179**, 1143 (1997).

37. Schwalb, C., Chapman, S. K. & Reid, G. A. The tetraheme cytochrome CymA is required for anaerobic respiration with dimethyl sulfoxide and nitrite in Shewanella oneidensis. *Biochemistry* **42**, 9491–9497 (2003).

38. Fonseca, B. M., Paquete, C. M., Neto, S. E., Pacheco, I., Soares, C. M. & Louro, R. O. Mind the gap: cytochrome interactions reveal electron pathways across the periplasm of Shewanella oneidensis MR-1. *Biochemical Journal* **449**, 101–108 (2013).

39. Marritt, S. J., Lowe, T. G., Bye, J., Mcmillan, D. G. G., Shi, L., Fredrickson, J., Zachara, J., Richardson, D. J., Cheesman, M. R., Jeuken, L. J. C. & Butt, J. N. A functional description of CymA, an electron-transfer hub supporting anaerobic respiratory flexibility in Shewanella. *Biochemical Journal* **444**, 465–474 (2012).

40. Leys, D., Tsapin, A. S., Nealson, K. H., Meyer, T. E., Cusanovich, M. A. & Van Beeumen, J. J. Structure and mechanism of the flavocytochrome c fumarate reductase of Shewanella putrefaciens MR-1. *Nature Structural Biology 1999 6:12* **6**, 1113–1117 (1999).

41. Gordon, E. H. J., Pealing, S. L., Chapman, S. K., Ward, F. B. & Reid, G. A. Physiological function and regulation of flavocytochrome c3, the soluble fumarate reductase from Shewanella putrefaciens NCIMB 400. *Microbiology* **144**, 937–945 (1998).

42. Sturm, G., Richter, K., Doetsch, A., Heide, H., Louro, R. O. & Gescher, J. A dynamic periplasmic electron transfer network enables respiratory flexibility beyond a thermodynamic regulatory regime. *The ISME Journal* **9**, 1802–1811 (2015).

43. Leys, D., Meyer, T. E., Tsapin, A. S., Nealson, K. H., Cusanovich, M. A. & Van Beeumen, J. J. Crystal structures at atomic resolution reveal the novel concept of 'electron-harvesting' as a role for the small tetraheme cytochrome c. *Journal of Biological Chemistry* **277**, 35703–35711 (2002).

44. Alves, M. N., Neto, S. E., Alves, A. S., Fonseca, B. M., Carrêlo, A., Pacheco, I., Paquete, C. M., Soares, C. M. & Louro, R. O. Characterization of the periplasmic redox network

that sustains the versatile anaerobic metabolism of Shewanella oneidensis MR-1. *Frontiers in Microbiology* **6**, 149646 (2015).

45. Beliaev, A. S. & Saffarini, D. A. Shewanella putrefaciens mtrB Encodes an Outer Membrane Protein Required for Fe(III) and Mn(IV) Reduction. *Journal of Bacteriology* **180**, 6292 (1998).

46. Beliaev, A. S., Saffarini, D. A., McLaughlin, J. L. & Hunnicutt, D. MtrC, an outer membrane decahaem c cytochrome required for metal reduction in Shewanella putrefaciens MR-1. *Molecular Microbiology* **39**, 722–730 (2001).

47. Ross, D. E., Ruebush, S. S., Brantley, S. L., Hartshorne, R. S., Clarke, T. A., Richardson, D. J. & Tien, M. Characterization of protein-protein interactions involved in iron reduction by Shewanella oneidensis MR-1. *Applied and Environmental Microbiology* **73**, 5797–5808 (2007).

48. Lower, B. H., Yongsunthon, R., Shi, L., Wildling, L., Gruber, H. J., Wigginton, N. S., Reardon, C. L., Pinchuk, G. E., Ubay, T. C. D., Boily, J. F. & Lower, S. K. Antibody Recognition Force Microscopy Shows that Outer Membrane Cytochromes OmcA and MtrC Are Expressed on the Exterior Surface of Shewanella oneidensis MR-1. *Applied and Environmental Microbiology* **75**, 2931–2935 (2009).

49. Weiss, M. S., Wacker, T., Weckesser, J., Weite, W. & Schulz, G. E. The three-dimensional structure of porin from Rhodobacter capsulatus at 3 Å resolution. *FEBS Letters* **267**, 268–272 (1990).

50. Hartshorne, R. S., Reardon, C. L., Ross, D., Nuester, J., Clarke, T. A., Gates, A. J., Mills, P. C., Fredrickson, J. K., Zachara, J. M., Shi, L., Beliaev, A. S., Marshall, M. J., Tien, M., Brantley, S., Butt, J. N. & Richardson, D. J. Characterization of an electron conduit between bacteria and the extracellular environment. *Proceedings of the National Academy of Sciences of the United States of America* **106**, 22169–22174 (2009).

51. Firer-Sherwood, M. A., Ando, N., Drennan, C. L. & Elliott, S. J. Solution-Based Structural Analysis of the Decaheme Cytochrome, MtrA, by Small-Angle X-ray Scattering and Analytical Ultracentrifugation. *The Journal of Physical Chemistry. B* **115**, 11208 (2011).

52. Bücking, C., Popp, F., Kerzenmacher, S. & Gescher, J. Involvement and specificity of Shewanella oneidensis outer membrane cytochromes in the reduction of soluble and solid-phase terminal electron acceptors. *FEMS Microbiology Letters* **306**, 144–151 (2010).

53. Coursolle, D. & Gralnick, J. A. Modularity of the Mtr respiratory pathway of Shewanella oneidensis strain MR-1. *Molecular Microbiology* **77**, 995–1008 (2010).

54. Clarke, T. A., Edwards, M. J., Gates, A. J., Hall, A., White, G. F., Bradley, J., Reardon, C. L., Shi, L., Beliaev, A. S., Marshall, M. J., Wang, Z., Watmough, N. J., Fredrickson, J. K., Zachara, J. M., Butt, J. N. & Richardson, D. J. Structure of a bacterial cell surface decaheme electron conduit. *Proceedings of the National Academy of Sciences of the United States of America* **108**, 9384–9389 (2011).

55. van Wonderen, J. H., Crack, J. C., Edwards, M. J., Clarke, T. A., Saalbach, G., Martins, C. & Butt, J. N. Liquid-chromatography mass spectrometry describes post-translational

modification of Shewanella outer membrane proteins. *Biochimica et Biophysica Acta (BBA) - Biomembranes* **1866**, 184221 (2024).

56. Hartshorne, R. S., Jepson, B. N., Clarke, T. A., Field, S. J., Fredrickson, J., Zachara, J., Shi, L., Butt, J. N. & Richardson, D. J. Characterization of Shewanella oneidensis MtrC: A cell-surface decaheme cytochrome involved in respiratory electron transport to extracellular electron acceptors. *Journal of Biological Inorganic Chemistry* **12**, 1083–1094 (2007).

57. Edwards, M. J., White, G. F., Butt, J. N., Richardson, D. J. & Clarke, T. A. The Crystal Structure of a Biological Insulated Transmembrane Molecular Wire. *Cell* **181**, 665 (2020).

58. Myers, J. M. & Myers, C. R. Role for Outer Membrane Cytochromes OmcA and OmcB of Shewanella putrefaciens MR-1 in Reduction of Manganese Dioxide. *Applied and Environmental Microbiology* **67**, 260 (2001).

59. White, G. F., Edwards, M. J., Gomez-Perez, L., Richardson, D. J., Butt, J. N. & Clarke, T. A. Mechanisms of Bacterial Extracellular Electron Exchange. *Advances in Microbial Physiology* **68**, 87–138 (2016).

60. Shi, L., Belchik, S. M., Wang, Z., Kennedy, D. W., Dohnalkova, A. C., Marshall, M. J., Zachara, J. M. & Fredrickson, J. K. Identification and characterization of UndA HRCR-6, an outer membrane endecaheme c-type cytochrome of Shewanella sp. Strain HRCR-6. *Applied and Environmental Microbiology* **77**, 5521–5523 (2011).

61. Edwards, M. J., Hall, A., Shi, L., Fredrickson, J. K., Zachara, J. M., Butt, J. N., Richardson, D. J. & Clarke, T. A. 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**, 1275–1284 (2012).

62. Edwards, M. J., Baiden, N. A., Johs, A., Tomanicek, S. J., Liang, L., Shi, L., Fredrickson, J. K., Zachara, J. M., Gates, A. J., Butt, J. N., Richardson, D. J. & Clarke, T. A. The X-ray crystal structure of Shewanella oneidensis OmcA reveals new insight at the microbe–mineral interface. *FEBS Letters* **588**, 1886–1890 (2014).

63. Lower, B. H., Lins, R. D., Oestreicher, Z., Straatsma, T. P., Hochella, M. F., Shi, L. & Lower, S. K. In vitro evolution of a peptide with a hematite binding motif that may constitute a natural metal-oxide binding archetype. *Environmental Science and Technology* **42**, 3821–3827 (2008).

64. Edwards, M. J., White, G. F., Norman, M., Tome-Fernandez, A., Ainsworth, E., Shi, L., Fredrickson, J. K., Zachara, J. M., Butt, J. N., Richardson, D. J. & Clarke, T. A. Redox Linked Flavin Sites in Extracellular Decaheme Proteins Involved in Microbe-Mineral Electron Transfer. *Scientific Reports 2015 5:1* **5**, 1–11 (2015).

65. Paquete, C. M., Fonseca, B. M., Cruz, D. R., Pereira, T. M., Pacheco, I., Soares, C. M. & Louro, R. O. Exploring the molecular mechanisms of electron shuttling across the microbe/metal space. *Frontiers in Microbiology* **5**, (2014).

66. Norman, M. P., Edwards, M. J., White, G. F., Burton, J. A. J., Butt, J. N., Richardson, D. J., Louro, R. O., Paquete, C. M. & Clarke, T. A. A Cysteine Pair Controls Flavin Reduction by

Extracellular Cytochromes during Anoxic/Oxic Environmental Transitions. *mBio* **14**, (2023).

67. Von Canstein, H., Ogawa, J., Shimizu, S. & Lloyd, J. R. Secretion of flavins by Shewanella species and their role in extracellular electron transfer. *Applied and Environmental Microbiology* **74**, 615–623 (2008).

68. Kotloski, N. J. & Gralnick, J. A. Flavin electron shuttles dominate extracellular electron transfer by Shewanella oneidensis. *mBio* **4**, 553–565 (2013).

69. Marsili, E., Baron, D. B., Shikhare, I. D., Coursolle, D., Gralnick, J. A. & Bond, D. R. Shewanella secretes flavins that mediate extracellular electron transfer. *Proceedings of the National Academy of Sciences of the United States of America* **105**, 3968–3973 (2008).

70. Brutinel, E. D. & Gralnick, J. A. Shuttling happens: Soluble flavin mediators of extracellular electron transfer in Shewanella. *Applied Microbiology and Biotechnology* **93**, 41–48 (2012).

71. Gralnick, J. A. & Bond, D. R. Electron Transfer beyond the Outer Membrane: Putting Electrons to Rest. *Annual Review of Microbiology* **77**, 517–539 (2023).

72. Gorby, Y. A., Yanina, S., McLean, J. S., Rosso, K. M., Moyles, D., Dohnalkova, A., Beveridge, T. J., Chang, I. S., Kim, B. H., Kim, K. S., Culley, D. E., Reed, S. B., Romine, M. F., Saffarini, D. A., Hill, E. A., Shi, L., Elias, D. A., Kennedy, D. W., Pinchuk, G. *et al*. Electrically conductive bacterial nanowires produced by Shewanella oneidensis strain MR-1 and other microorganisms. *Proceedings of the National Academy of Sciences of the United States of America* **103**, 11358 (2006).

73. El-Naggar, M. Y., Wanger, G., Leung, K. M., Yuzvinsky, T. D., Southam, G., Yang, J., Lau, W. M., Nealson, K. H. & Gorby, Y. A. Electrical transport along bacterial nanowires from Shewanella oneidensis MR-1. *Proceedings of the National Academy of Sciences of the United States of America* **107**, 18127–18131 (2010).

74. Subramanian, P., Pirbadian, S., El-Naggar, M. Y. & Jensen, G. J. Ultrastructure of Shewanella oneidensis MR-1 nanowires revealed by electron cryotomography. *Proceedings of the National Academy of Sciences of the United States of America* **115**, E3246–E3255 (2018).

75. Pirbadian, S., Barchinger, S. E., Leung, K. M., Byun, H. S., Jangir, Y., Bouhenni, R. A., Reed, S. B., Romine, M. F., Saffarini, D. A., Shi, L., Gorby, Y. A., Golbeck, J. H. & El-Naggar, M. Y. Shewanella oneidensis MR-1 nanowires are outer membrane and periplasmic extensions of the extracellular electron transport components. *Proceedings of the National Academy of Sciences of the United States of America* **111**, 12883–12888 (2014).

76. White, G. F., Shi, Z., Shi, L., Wang, Z., Dohnalkova, A. C., Marshall, M. J., Fredrickson, J. K., Zachara, J. M., Butt, J. N., Richardson, D. J. & Clarke, T. A. Rapid electron exchange between surface-exposed bacterial cytochromes and Fe(III) minerals. *Proceedings of the National Academy of Sciences of the United States of America* **110**, 6346–6351 (2013).

77. Edwards, M. J., White, G. F., Lockwood, C. W., Lawes, M. C., Martel, A., Harris, G., Scott, D. J., Richardson, D. J., Butt, J. N. & Clarke, T. A. Structural modeling of an outer membrane electron conduit from a metal-reducing bacterium suggests electron transfer via periplasmic redox partners. *Journal of Biological Chemistry* **293**, 8103–8112 (2018).

78. van Wonderen, J. H., Adamczyk, K., Wu, X., Jiang, X., Piper, S. E. H., Hall, C. R., Edwards, M. J., Clarke, T. A., Zhang, H., Jeuken, L. J. C., Sazanovich, I. V., Towrie, M., Blumberger, J., Meech, S. R. & Butt, J. N. Nanosecond heme-to-heme electron transfer rates in a multiheme cytochrome nanowire reported by a spectrally unique His/Met-ligated heme. *Proceedings of the National Academy of Sciences of the United States of America* **118**, e2107939118 (2021).

79. Van Wonderen, J. H., Hall, C. R., Jiang, X., Adamczyk, K., Carof, A., Heisler, I., Piper, S. E. H., Clarke, T. A., Watmough, N. J., Sazanovich, I. V., Towrie, M., Meech, S. R., Blumberger, J. & Butt, J. N. Ultrafast Light-Driven Electron Transfer in a Ru(II)tris(bipyridine)-Labeled Multiheme Cytochrome. *Journal of the American Chemical Society* **141**, 15190–15200 (2019).

80. Jiao, Y. & Newman, D. K. The pio operon is essential for phototrophic Fe(II) oxidation in Rhodopseudomonas palustris TIE-1. *Journal of Bacteriology* **189**, 1765–1773 (2007).

81. Guzman, M. S., Rengasamy, K., Binkley, M. M., Jones, C., Ranaivoarisoa, T. O., Singh, R., Fike, D. A., Meacham, J. M. & Bose, A. Phototrophic extracellular electron uptake is linked to carbon dioxide fixation in the bacterium Rhodopseudomonas palustris. *Nature Communications 2019 10:1* **10**, 1–13 (2019).

82. Emerson, D. & Moyer, C. Isolation and characterization of novel iron-oxidizing bacteria that grow at circumneutral pH. *Applied and Environmental Microbiology* **63**, 4784–4792 (1997).

83. Liu, J., Wang, Z., Belchik, S. M., Edwards, M. J., Liu, C., Kennedy, D. W., Merkley, E. D., Lipton, M. S., Butt, J. N., Richardson, D. J., Zachara, J. M., Fredrickson, J. K., Rosso, K. M. & Shi, L. Identification and characterization of M to A: A decaheme c-type cytochrome of the neutrophilic Fe(ll)-oxidizing bacterium Sideroxydans lithotrophicus ES-1. *Frontiers in Microbiology* **3**, 17845 (2012).

84. Gralnick, J. A., Brown, C. T. & Newman, D. K. Anaerobic regulation by an atypical Arc system in Shewanella oneidensis. *Molecular Microbiology* **56**, 1347–1357 (2005).

85. Gralnick, J. A., Vali, H., Lies, D. P. & Newman, D. K. Extracellular respiration of dimethyl sulfoxide by Shewanella oneidensis strain MR-1. *Proceedings of the National Academy of Sciences of the United States of America* **103**, 4669 (2006).

86. DiChristina, T. J., Moore, C. M. & Haller, C. A. Dissimilatory Fe(III) and Mn(IV) reduction by Shewanella putrefaciens requires ferE, a homolog of the pulE (gspE) type II protein secretion gene. *Journal of Bacteriology* **184**, 142–151 (2002).

87. Hille, R. The mononuclear molybdenum enzymes. *Chemical Reviews* **96**, 2757–2816 (1996).

88. Johnson, J. L., Rajagopalan, K. V. & Meyer, O. Isolation and characterization of a second molybdopterin dinucleotide: Molybdopterin cytosine dinucleotide. *Archives of Biochemistry and Biophysics* **283**, 542–545 (1990).

89. Schindelin, H., Kisker, C., Hilton, J., Rajagopalan, K. V. & Rees, D. C. Crystal Structure of DMSO Reductase: Redox-Linked Changes in Molybdopterin Coordination. *Science* **272**, 1615–1621 (1996).

90. Schneider, F., Löwe, J., Huber, R., Schindelin, H., Kisker, C. & Knäblein, J. Crystal Structure of Dimethyl Sulfoxide Reductase fromRhodobacter capsulatusat 1.88 Å Resolution. *Journal of Molecular Biology* **263**, 53–69 (1996).

91. McAlpine, A. S., McEwan, A. G. & Bailey, S. The high resolution crystal structure of DMSO reductase in complex with DMSO. *Journal of Molecular Biology* **275**, 613–623 (1998).

92. Kilpatrick, L., Rajagopalan, K. V., Hilton, J., Bastian, N. R., Stiefel, E. I., Pilato, R. S. & Spiro, T. G. Resonance Raman Spectroscopic Characterization of the Molybdopterin Active Site of DMSO Reductase. *Biochemistry* **34**, 3032–3039 (1995).

93. Baugh, P. E., Garner, C. D., Charnock, J. M., Collison, D., Davies, E. S., McAlpine, A. S., Bailey, S., Lane, I., Hanson, G. R. & McEwan, A. G. X-ray absorption spectroscopy of dimethylsulfoxide reductase from Rhodobacter capsulatus. *Journal of Biological Inorganic Chemistry* **2**, 634–643 (1997).

94. Mtei, R. P., Lyashenko, G., Stein, B., Rubie, N., Hille, R. & Kirk, M. L. Spectroscopic and electronic structure studies of a dimethyl sulfoxide reductase catalytic intermediate: Implications for electron- and atom-transfer reactivity. *Journal of the American Chemical Society* **133**, 9762–9774 (2011).

95. Le, C. C., Bae, M., Kiamehr, S. & Balskus, E. P. Emerging Chemical Diversity and Potential Applications of Enzymes in the DMSO Reductase Superfamily. *Annual Review of Biochemistry* **91**, 475–504 (2022).

96. Weiner, J. H., MacIsaac, D. P., Bishop, R. E. & Bilous, P. T. Purification and properties of Escherichia coli dimethyl sulfoxide reductase, an iron-sulfur molybdoenzyme with broad substrate specificity. *Journal of Bacteriology* **170**, 1505 (1988).

97. Jormakka, M., Törnroth, S., Byrne, B. & Iwata, S. Molecular basis of proton motive force generation: Structure of formate dehydrogenase-N. *Science* **295**, 1863–1868 (2002).

98. Bertero, M. G., Rothery, R. A., Palak, M., Hou, C., Lim, D., Blasco, F., Weiner, J. H. & Strynadka, N. C. J. Insights into the respiratory electron transfer pathway from the structure of nitrate reductase A. *Nature Structural & Molecular Biology 2003 10:9* **10**, 681–687 (2003).

99. Jumper, J., Evans, R., Pritzel, A., Green, T., Figurnov, M., Ronneberger, O., Tunyasuvunakool, K., Bates, R., Žídek, A., Potapenko, A., Bridgland, A., Meyer, C., Kohl, S. A. A., Ballard, A. J., Cowie, A., Romera-Paredes, B., Nikolov, S., Jain, R., Adler, J. *et al.* Highly accurate protein structure prediction with AlphaFold. *Nature 2021 596:7873* **596**, 583–589 (2021).

100. Stevens, C. M., Winstone, T. M. L., Turner, R. J. & Paetzel, M. Structural Analysis of a Monomeric Form of the Twin-Arginine Leader Peptide Binding Chaperone Escherichia coli DmsD. *Journal of Molecular Biology* **389**, 124–133 (2009).

101. Hatzixanthis, K., Clarke, T. A., Oubrie, A., Richardson, D. J., Turner, R. J. & Sargent, F. Signal peptide–chaperone interactions on the twin-arginine protein transport pathway. *Proceedings of the National Academy of Sciences* **102**, 8460–8465 (2005).

102. Berks, B. C., Palmer, T. & Sargent, F. Protein targeting by the bacterial twin-arginine translocation (Tat) pathway. *Current Opinion in Microbiology* **8**, 174–181 (2005).

103. Chan, C. S., Winstone, T. M. L., Chang, L., Stevens, C. M., Workentine, M. L., Li, H., Wei, Y., Ondrechen, M. J., Paetzel, M. & Turner, R. J. Identification of residues in DmsD for twin-arginine leader peptide binding, defined through random and bioinformatics-directed mutagenesis. *Biochemistry* **47**, 2749–2759 (2008).

104. Holm, L. Dali server: structural unification of protein families. *Nucleic Acids Research* **50**, W210–W215 (2022).

105. Almagro Armenteros, J. J., Tsirigos, K. D., Sønderby, C. K., Petersen, T. N., Winther, O., Brunak, S., von Heijne, G. & Nielsen, H. SignalP 5.0 improves signal peptide predictions using deep neural networks. *Nature Biotechnology 2019 37:4* **37**, 420–423 (2019).

106. Shin, H. D., Toporek, Y., Mok, J. K., Maekawa, R., Lee, B. D., Howard, M. H. & DiChristina, T. J. Iodate Reduction by Shewanella oneidensis Requires Genes Encoding an Extracellular Dimethylsulfoxide Reductase. *Frontiers in Microbiology* **13**, 852942 (2022).

107. Caccavo, F., Lonergan, D. J., Lovley, D. R., Davis, M., Stolz, J. F. & Mcinerneyl, M. J. Geobacter sulfurreducens sp. nov., a Hydrogen-and Acetate-Oxidizing Dissimilatory Metal-Reducing Microorganism. *Applied and Environmentsl Microbiology* **60**, 3752–3759 (1994).

108. Lin, W. C., Coppi, M. V. & Lovley, D. R. Geobacter sulfurreducens Can Grow with Oxygen as a Terminal Electron Acceptor. *Applied and Environmental Microbiology* **70**, 2525–2528 (2004).

109. Reguera, G., Nevin, K. P., Nicoll, J. S., Covalla, S. F., Woodard, T. L. & Lovley, D. R. Biofilm and nanowire production leads to increased current in Geobacter sulfurreducens fuel cells. *Applied and Environmental Microbiology* **72**, 7345–7348 (2006).

110. Clarke, T. A. Plugging into bacterial nanowires: a comparison of model electrogenic organisms. *Current Opinion in Microbiology* **66**, 56–62 (2022).

111. Mollaei, M., Timmers, P. H. A., Suarez-Diez, M., Boeren, S., van Gelder, A. H., Stams, A. J. M. & Plugge, C. M. Comparative proteomics of Geobacter sulfurreducens PCAT in response to acetate, formate and/or hydrogen as electron donor. *Environmental Microbiology* **23**, 299–315 (2021).

112. Levar, C. E., Hoffman, C. L., Dunshee, A. J., Toner, B. M. & Bond, D. R. Redox potential as a master variable controlling pathways of metal reduction by Geobacter sulfurreducens. *The ISME Journal* **11**, 741–752 (2017).

113. Levar, C. E., Chan, C. H., Mehta-Kolte, M. G. & Bond, D. R. An inner membrane cytochrome required only for reduction of high redox potential extracellular electron acceptors. *mBio* **5**, (2014).

114. Zacharoff, L., Chan, C. H. & Bond, D. R. Reduction of low potential electron acceptors requires the CbcL inner membrane cytochrome of Geobacter sulfurreducens. *Bioelectrochemistry* **107**, 7–13 (2016).

115. Ding, Y. H. R., Hixson, K. K., Aklujkar, M. A., Lipton, M. S., Smith, R. D., Lovley, D. R. & Mester, T. Proteome of Geobacter sulfurreducens grown with Fe(III) oxide or Fe(III) citrate as the electron acceptor. *Biochimica et Biophysica Acta (BBA) - Proteins and Proteomics* **1784**, 1935–1941 (2008).

116. Joshi, K., Chan, C. H. & Bond, D. R. Geobacter sulfurreducens inner membrane cytochrome CbcBA controls electron transfer and growth yield near the energetic limit of respiration. *Molecular Microbiology* **116**, 1124–1139 (2021).

117. Pimenta, A. I., Paquete, C. M., Morgado, L., Edwards, M. J., Clarke, T. A., Salgueiro, C. A., Pereira, I. A. C. & Duarte, A. G. Characterization of the inner membrane cytochrome ImcH from Geobacter reveals its importance for extracellular electron transfer and energy conservation. *Protein Science* **32**, e4796 (2023).

118. Antunes, J. M. A., Silva, M. A., Salgueiro, C. A. & Morgado, L. Electron Flow From the Inner Membrane Towards the Cell Exterior in Geobacter sulfurreducens: Biochemical Characterization of Cytochrome CbcL. *Frontiers in Microbiology* **13**, 898015 (2022).

119. Levar, C. E., Chan, C. H., Mehta-Kolte, M. G. & Bond, D. R. An inner membrane cytochrome required only for reduction of high redox potential extracellular electron acceptors. *mBio* **5**, (2014).

120. Methé, B. A., Nelson, K. E., Eisen, J. A., Paulsen, I. T., Nelson, W., Heidelberg, J. F., Wu, D., Wu, M., Ward, N., Beanan, M. J., Dodson, R. J., Madupu, R., Brinkac, L. M., Daugherty, S. C., DeBoy, R. T., Durkin, A. S., Gwinn, M., Kolonay, J. F., Sullivan, S. A. *et al.* Genome of Geobacter sulfurreducens: Metal Reduction in Subsurface Environments. *Science* **302**, 1967–1969 (2003).

121. Choi, S., Chan, C. H. & Bond, D. R. Lack of Specificity in Geobacter Periplasmic Electron Transfer. *Journal of Bacteriology* **204**, (2022).

122. Pokkuluri, P. R., Londer, Y. Y., Yang, X., Duke, N. E. C., Erickson, J., Orshonsky, V., Johnson, G. & Schiffer, M. Structural characterization of a family of cytochromes c7 involved in Fe(III) respiration by Geobacter sulfurreducens. *Biochimica et Biophysica Acta (BBA) - Bioenergetics* **1797**, 222–232 (2010).

123. Pokkuluri, P. R., Londer, Y. Y., Duke, N. E. C., Long, W. C. & Schiffer, M. Family of Cytochrome c7-Type Proteins from Geobacter sulfurreducens: Structure of One Cytochrome c7 at 1.45 Å Resolution†,‡. *Biochemistry* **43**, 849–859 (2004).

124. Morgado, L., Bruix, M., Orshonsky, V., Londer, Y. Y., Duke, N. E. C., Yang, X., Pokkuluri, P. R., Schiffer, M. & Salgueiro, C. A. Structural insights into the modulation of the redox properties of two Geobacter sulfurreducens homologous triheme cytochromes. *Biochimica et Biophysica Acta (BBA) - Bioenergetics* **1777**, 1157–1165 (2008).

125. Pokkuluri, P. R., Londer, Y. Y., Duke, N. E. C., Pessanha, M., Yang, X., Orshonsky, V., Orshonsky, L., Erickson, J., Zagyanskiy, Y., Salgueiro, C. A. & Schiffer, M. Structure of a novel dodecaheme cytochrome c from Geobacter sulfurreducens reveals an extended 12 nm protein with interacting hemes. *Journal of Structural Biology* **174**, 223–233 (2011).

126. Zacharoff, L. A., Morrone, D. J. & Bond, D. R. Geobacter sulfurreducens extracellular multiheme cytochrome PgcA facilitates respiration to Fe(III) oxides but not electrodes. *Frontiers in Microbiology* **8**, 300906 (2017).

127. Otero, F. J., Chan, C. H. & Bond, D. R. 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. *Journal of Bacteriology* **200**, 347–365 (2018).

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

129. Magnuson, T. S., Isoyama, N., Hodges-Myerson, A. L., Davidson, G., Maroney, M. J., Geesey, G. G. & Lovley, D. R. Isolation, characterization and gene sequence analysis of a membrane-associated 89 kDa Fe(III) reducing cytochrome c from Geobacter sulfurreducens. *Biochemical Journal* **359**, 147–152 (2001).

130. Leang, C., Coppi, M. V. & Lovley, D. R. OmcB, a c-type polyheme cytochrome, involved in Fe(III) reduction in Geobacter sulfurreducens. *Journal of Bacteriology* **185**, 2096–2103 (2003).

131. Qian, X., Reguera, G., Mester, T. & Lovley, D. R. Evidence that OmcB and OmpB of Geobacter sulfurreducens are outer membrane surface proteins. *FEMS Microbiology Letters* **277**, 21–27 (2007).

132. Chan, C. H., Levar, C. E., Jiménez-Otero, F. & Bond, D. R. Genome scale mutational analysis of Geobacter sulfurreducens reveals distinct molecular mechanisms for respiration and sensing of poised electrodes versus Fe(III) oxides. *Journal of Bacteriology* **199**, (2017).

133. Jahan, M. I., Tobe, R. & Mihara, H. Characterization of a Novel Porin-Like Protein, ExtI, from Geobacter sulfurreducens and Its Implication in the Reduction of Selenite and Tellurite. *International Journal of Molecular Sciences 2018, Vol. 19, Page 809* **19**, 809 (2018).

134. Jahan, M. I., Juengwiwattanakitti, P., Izu, Y., Tobe, R., Imai, T. & Mihara, H. Selenite uptake by outer membrane porin ExtI and its involvement in the subcellular localization of rhodanese-like lipoprotein ExtH in Geobacter sulfurreducens. *Biochemical and Biophysical Research Communications* **516**, 474–479 (2019).

135. Teixeira, L. R., Dantas, J. M., Salgueiro, C. A. & Cordas, C. M. Thermodynamic and kinetic properties of the outer membrane cytochrome OmcF, a key protein for

extracellular electron transfer in Geobacter sulfurreducens. *Biochimica et Biophysica Acta (BBA) - Bioenergetics* **1859**, 1132–1137 (2018).

136. Kim, B. C., Leang, C., Ding, Y. H. R., Glaven, R. H., Coppi, M. V. & Lovley, D. R. OmcF, a Putative c-Type Monoheme Outer Membrane Cytochrome Required for the Expression of Other Outer Membrane Cytochromes in Geobacter sulfurreducens. *Journal of Bacteriology* **187**, 4505 (2005).

137. Pokkuluri, P. R., Londer, Y. Y., Wood, S. J., Duke, N. E. C., Morgado, L., Salgueiro, C. A. & Schiffer, M. Outer membrane cytochrome c, OmcF, from Geobacter sulfurreducens: High structural similarity to an algal cytochrome c6. *Proteins: Structure, Function, and Bioinformatics* **74**, 266–270 (2009).

138. Yalcin, S. E. & Malvankar, N. S. The blind men and the filament: Understanding structures and functions of microbial nanowires. *Current Opinion in Chemical Biology* **59**, 193–201 (2020).

139. Uekia, T. Cytochromes in Extracellular Electron Transfer in Geobacter. *Applied and Environmental Microbiology* **87**, 1–16 (2021).

140. Nevin, K. P. & Lovley, D. R. Mechanisms for Fe(III) Oxide Reduction in Sedimentary Environments. *Geomicrobiology Journal* **19**, 141–159 (2002).

141. Straub, K. L. & Schink, B. Evaluation of electron-shuttling compounds in microbial ferric iron reduction. *FEMS Microbiology Letters* **220**, 229–233 (2003).

142. Lies, D. P., Hernandez, M. E., Kappler, A., Mielke, R. E., Gralnick, J. A. & Newman, D. K. Shewanella oneidensis MR-1 Uses Overlapping Pathways for Iron Reduction at a Distance and by Direct Contact under Conditions Relevant for Biofilms. *Applied and Environmental Microbiology* **71**, 4414 (2005).

143. Nevin, K. P. & Lovley, D. R. Lack of production of electron-shuttling compounds or solubilization of Fe(III) during reduction of insoluble Fe(III) oxide by Geobacter metallireducens. *Applied and Environmental Microbiology* **66**, 2248–2251 (2000).

144. Mehta, T., Coppi, M. V., Childers, S. E. & Lovley, D. R. Outer membrane c-type cytochromes required for Fe(III) and Mn(IV) oxide reduction in Geobacter sulfurreducens. *Applied and Environmental Microbiology* **71**, 8634–8641 (2005).

145. Craig, L., Forest, K. T. & Maier, B. Type IV pili: dynamics, biophysics and functional consequences. *Nature Reviews Microbiology 2019 17:7* **17**, 429–440 (2019).

146. Parge, H. E., Forest, K. T., Hickey, M. J., Christensen, D. A., Getzoff, E. D. & Tainer, J. A. Structure of the fibre-forming protein pilin at 2.6 Å resolution. *Nature 1995 378:6552* **378**, 32–38 (1995).

147. Wang, F., Coureuil, M., Osinski, T., Orlova, A., Altindal, T., Gesbert, G., Nassif, X., Egelman, E. H. & Craig, L. Cryoelectron Microscopy Reconstructions of the Pseudomonas aeruginosa and Neisseria gonorrhoeae Type IV Pili at Sub-nanometer Resolution. *Structure* **25**, 1423-1435.e4 (2017).

148. Reguera, G., McCarthy, K. D., Mehta, T., Nicoll, J. S., Tuominen, M. T. & Lovley, D. R. Extracellular electron transfer via microbial nanowires. *Nature 2005 435:7045* **435**, 1098–1101 (2005).

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

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

151. Reardon, P. N. & Mueller, K. T. Structure of the type IVa major pilin from the electrically conductive bacterial nanowires of geobacter sulfurreducens. *Journal of Biological Chemistry* **288**, 29260–29266 (2013).

152. Craig, L., Volkmann, N., Arvai, A. S., Pique, M. E., Yeager, M., Egelman, E. H. H. & Tainer, J. A. Type IV Pilus Structure by Cryo-Electron Microscopy and Crystallography: Implications for Pilus Assembly and Functions. *Molecular Cell* **23**, 651–662 (2006).

153. Xiao, K., Malvankar, N. S., Shu, C., Martz, E., Lovley, D. R. & Sun, X. Low Energy Atomic Models Suggesting a Pilus Structure that could Account for Electrical Conductivity of Geobacter sulfurreducens Pili. *Scientific Reports 2016 6:1* **6**, 1–9 (2016).

154. Malvankar, N. S., Vargas, M., Nevin, K. P., Franks, A. E., Leang, C., Kim, B. C., Inoue, K., Mester, T., Covalla, S. F., Johnson, J. P., Rotello, V. M., Tuominen, M. T. & Lovley, D. R. Tunable metallic-like conductivity in microbial nanowire networks. *Nature Nanotechnology 2011 6:9* **6**, 573–579 (2011).

155. Vargas, M., Malvankar, N. S., Tremblay, P. L., Leang, C., Smith, J. A., Patel, P., Synoeyenbos-West, O., Nevin, K. P. & Lovley, D. R. Aromatic Amino Acids Required for Pili Conductivity and Long-Range Extracellular Electron Transport in Geobacter sulfurreducens. *mBio* **4**, 105–118 (2013).

156. Leang, C., Qian, X., Mester, T. & Lovley, D. R. Alignment of the c-type cytochrome OmcS along pili of Geobacter sulfurreducens. *Applied and Environmental Microbiology* **76**, 4080–4084 (2010).

157. Boesen, T. & Nielsen, L. P. Molecular dissection of bacterial nanowires. *mBio* **4**, (2013).

158. Ye, Y., Liu, X., Nealson, K. H., Rensing, C., Qin, S. & Zhou, S. Dissecting the Structural and Conductive Functions of Nanowires in Geobacter sulfurreducens Electroactive Biofilms. *mBio* **13**, (2022).

159. Kemp, A. D., Harding, C. C., Cabral, W. A., Marini, J. C. & Wallace, J. M. Effects of tissue hydration on nanoscale structural morphology and mechanics of individual Type I collagen fibrils in the Brtl mouse model of Osteogenesis Imperfecta. *Journal of Structural Biology* **180**, 428–438 (2012).

160. Wang, F., Craig, L., Liu, X., Rensing, C. & Egelman, E. H. Microbial nanowires: type IV pili or cytochrome filaments? *Trends in microbiology* **31**, 384 (2023).

161. Boesen, T., Nielsen, L. P. & Schramm, A. Pili for nanowires. *Nature Microbiology 2021 6:11* **6**, 1347–1348 (2021).

162. López-Castilla, A., Thomassin, J. L., Bardiaux, B., Zheng, W., Nivaskumar, M., Yu, X., Nilges, M., Egelman, E. H., Izadi-Pruneyre, N. & Francetic, O. Structure of the calcium-

dependent type 2 secretion pseudopilus. *Nature Microbiology 2017 2:12* **2**, 1686–1695 (2017).

163. Mehta, T., Coppi, M. V., Childers, S. E. & Lovley, D. R. Outer membrane c-type cytochromes required for Fe(III) and Mn(IV) oxide reduction in Geobacter sulfurreducens. *Applied and Environmental Microbiology* **71**, 8634–8641 (2005).

164. Qian, X., Mester, T., Morgado, L., Arakawa, T., Sharma, M. L., Inoue, K., Joseph, C., Salgueiro, C. A., Maroney, M. J. & Lovley, D. R. Biochemical characterization of purified OmcS, a c-type cytochrome required for insoluble Fe(III) reduction in Geobacter sulfurreducens. *Biochimica et Biophysica Acta (BBA) - Bioenergetics* **1807**, 404–412 (2011).

165. Wang, F., Gu, Y., O'Brien, J. P., Yi, S. M., Yalcin, S. E., Srikanth, V., Shen, C., Vu, D., Ing, N. L., Hochbaum, A. I., Egelman, E. H. & Malvankar, N. S. Structure of Microbial Nanowires Reveals Stacked Hemes that Transport Electrons over Micrometers. *Cell* **177**, 361-369.e10 (2019).

166. Wang, F., Mustafa, K., Suciu, V., Joshi, K., Chan, C. H., Choi, S., Su, Z., Si, D., Hochbaum, A. I., Egelman, E. H. & Bond, D. R. Cryo-EM structure of an extracellular Geobacter OmcE cytochrome filament reveals tetrahaem packing. *Nature Microbiology 2022 7:8* **7**, 1291–1300 (2022).

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

168. Gu, Y., Guberman-Pfeffer, M. J., Srikanth, V., Shen, C., Giska, F., Gupta, K., Londer, Y., Samatey, F. A., Batista, V. S. & Malvankar, N. S. Structure of Geobacter cytochrome OmcZ identifies mechanism of nanowire assembly and conductivity. *Nature Microbiology 2023 8:2* **8**, 284–298 (2023).

169. Inoue, K., Leang, C., Franks, A. E., Woodard, T. L., Nevin, K. P. & Lovley, D. R. Specific localization of the c-type cytochrome OmcZ at the anode surface in current-producing biofilms of Geobacter sulfurreducens. *Environmental microbiology reports* **3**, 211–217 (2011).

170. Inoue, K., Qian, X., Morgado, L., Kim, B. C., Mester, T., Izallalen, M., Salgueiro, C. A. & Lovley, D. R. Purification and characterization of OmcZ, an outer-surface, octaheme c-type cytochrome essential for optimal current production by geobacter sulfurreducens. *Applied and Environmental Microbiology* **76**, 3999–4007 (2010).

171. Kai, A., Tokuishi, T., Fujikawa, T., Kawano, Y., Ueki, T., Nagamine, M., Sakakibara, Y., Suiko, M. & Inoue, K. Proteolytic Maturation of the Outer Membrane c-Type Cytochrome OmcZ by a Subtilisin-Like Serine Protease Is Essential for Optimal Current Production by Geobacter sulfurreducens. *Applied and Environmental Microbiology* **87**, 1–10 (2021).

172. Power, S. D., Adams, R. M. & Wells, J. A. Secretion and autoproteolytic maturation of subtilisin. *Proceedings of the National Academy of Sciences* **83**, 3096–3100 (1986).

173. Peng, L. & Zhang, Y. Cytochrome OmcZ is essential for the current generation by Geobacter sulfurreducens under low electrode potential. *Electrochimica Acta* **228**, 447–452 (2017).

174. Nevin, K. P., Kim, B. C., Glaven, R. H., Johnson, J. P., Woodward, T. L., Methé, B. A., Didonato, R. J., Covalla, S. F., Franks, A. E., Liu, A. & Lovley, D. R. Anode Biofilm Transcriptomics Reveals Outer Surface Components Essential for High Density Current Production in Geobacter sulfurreducens Fuel Cells. *PLOS ONE* **4**, e5628 (2009).

175. Jiang, J., He, P., Luo, Y., Peng, Z., Jiang, Y., Hu, Y., Qi, L., Dong, X., Dong, Y. & Shi, L. The varied roles of pilA-N, omcE, omcS, omcT, and omcZ in extracellular electron transfer by Geobacter sulfurreducens. *Frontiers in Microbiology* **14**, 1251346 (2023).

176. Baquero, D. P., Cvirkaite-Krupovic, V., Hu, S. S., Fields, J. L., Liu, X., Rensing, C., Egelman, E. H., Krupovic, M. & Wang, F. Extracellular cytochrome nanowires appear to be ubiquitous in prokaryotes. *Cell* **186**, 2853-2864.e8 (2023).

177. Ding, Y. H. R., Hixson, K. K., Giometti, C. S., Stanley, A., Esteve-Núñez, A., Khare, T., Tollaksen, S. L., Zhu, W., Adkins, J. N., Lipton, M. S., Smith, R. D., Mester, T. & Lovley, D. R. The proteome of dissimilatory metal-reducing microorganism Geobacter sulfurreducens under various growth conditions. *Biochimica et Biophysica Acta (BBA) - Proteins and Proteomics* **1764**, 1198–1206 (2006).

178. Yi, H., Nevin, K. P., Kim, B. C., Franks, A. E., Klimes, A., Tender, L. M. & Lovley, D. R. Selection of a variant of Geobacter sulfurreducens with enhanced capacity for current production in microbial fuel cells. *Biosensors and Bioelectronics* **24**, 3498–3503 (2009).

179. Tremblay, P. L., Summers, Z. M., Glaven, R. H., Nevin, K. P., Zengler, K., Barrett, C. L., Qiu, Y., Palsson, B. O. & Lovley, D. R. A c-type cytochrome and a transcriptional regulator responsible for enhanced extracellular electron transfer in Geobacter sulfurreducens revealed by adaptive evolution. *Environmental Microbiology* **13**, 13–23 (2011).

180. Smith, J. A., Tremblay, P. L., Shrestha, P. M., Snoeyenbos-West, O. L., Franks, A. E., Nevin, K. P. & Lovley, D. R. Going Wireless: Fe(III) Oxide Reduction without Pili by Geobacter sulfurreducens Strain JS-1. *Applied and Environmental Microbiology* **80**, 4331 (2014).

181. Yu, N. Y., Wagner, J. R., Laird, M. R., Melli, G., Rey, S., Lo, R., Dao, P., Cenk Sahinalp, S., Ester, M., Foster, L. J. & Brinkman, F. S. L. PSORTb 3.0: improved protein subcellular localization prediction with refined localization subcategories and predictive capabilities for all prokaryotes. *Bioinformatics* **26**, 1608 (2010).

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

183. Kobashigawa, Y., Nishimiya, Y., Miura, K., Ohgiya, S., Miura, A. & Tsuda, S. A part of ice nucleation protein exhibits the ice-binding ability. *FEBS Letters* **579**, 1493–1497 (2005).

Chapter 2: General methods

Water and buffers

All water used within this work was purified by reverse osmosis (RO), unless otherwise specified. All buffers, unless otherwise stated, were prepared using RO water. Their pH was measured using a pH probe and adjusted as required with concentrated stocks of HCl or NaOH, finally they were filtered with 0.45 µm filter paper.

Media

All bacterial growth described within this work, unless otherwise specified, was performed in Luria-Broth media (LB). LB was prepared by dissolving 25 g of LB powder (NaCl, 10 g L^{-1}, tryptone, 10 g L^{-1}, yeast Extract, 5 g L^{-1}) in water, which was then autoclaved.

Antibiotics

All bacterial growth supplemented with antibiotics described was performed with either kanamycin (30 µg mL^{-1}) or carbenicillin (50 µg mL^{-1}). These were added from a 1000 x concentration stock, which was filter sterilised (0.2 µm filter).

Bacterial storage

All bacterial strains were stored for short durations on agar plates, prepared by addition of 15 of agar powder, to 100 mL LB media which was then autoclaved. Autoclaved agar was heated until molten using a commercial microwave, any required antibiotics were added, and the liquid poured into plates.

For longer storage, 0.6 mL of overnight bacterial cultures were added to 0.6 mL of 50% v/v glycerol for cryoprotection, and the sample frozen in liquid nitrogen before transferring to a -80 °C freezer.

Competent E. coli preparation

An overnight culture of *E. coli* (Top10 or BL21 (DE3)) was first prepared from laboratory stocks (180 RPM, 30 °C). This was used to inoculate (1% v/v) 100 mL LB, which was grown (180 RPM, 37 °C) until an OD_{600} of ~0.5 was reached. The culture was then cooled on ice for 10 minutes, before cells were harvested by centrifugation (4000 G, 5 minutes, 4 °C). Cells were resuspended in 20 mL sterile, cold 100 mM $CaCl_2$ solution and returned to the centrifuge (4000 G, 5 minutes, 4 °C). Cells were again resuspended in 2.5 mL 100 mM $CaCl_2$ supplemented to 20% v/v glycerol, before aliquoting at 50 µL volume, freezing in liquid nitrogen and transferring to a -80 °C freezer.

E. coli Transformation by heat shock

Competent *E. coli* aliquots (Top10 or BL21 (DE3)) were transformed with plasmid DNA by addition of 2 µL plasmid solution, and incubation on ice for 15 minutes.

The tube was then transferred to a 42 °C water bath for ~30 seconds, before transfer back onto ice for 2 minutes. Next, 200 µL of LB media was added and the cells left at 37 °C with agitation for ~1 hour. 100 µL of this culture was transferred onto an agar plate with the appropriate antibiotic and spread, this was then left at 37 °C overnight.

S. oneidensis transformation by electroporation

Plasmid DNA was transformed into *S. oneidensis* by electroporation. First an overnight culture of *S. oneidensis* MR-1 was prepared (180 RPM, 30 °C). Next, 3 mL of overnight culture was centrifuged (6000 G, 2 minutes) and the supernatant discarded. The cell pellet was resuspended in 1 mL of sterile 10% v/v glycerol, and returned to the centrifuge (6000 G, 1 minute). The supernatant was discarded and the pellet resuspended in 1 mL of 10 % v/v glycerol before being returned to the centrifuge (6000 G, 1 minute). From the supernatant, 930 µL was removed before the pellet was suspended in the remaining liquid. To this 2 µL of plasmid DNA solution was added before transferring the material to an electroporation cuvette. The cuvette was electroporated with a voltage of 1.2 kV, after which 1 mL of SOC media was added immediately, the resultant suspension was then mixed gently. This culture was then incubated (2 hours, 180 RPM, 30 °C), before the cells were harvested by centrifugation (4000 G, 1 minute) and transferred onto agar with the appropriate antibiotic. This was then left overnight at 30 °C.

DNA storage

All purified DNA described within this work was stored at -20 °C.

Plasmid extraction and purification

Plasmid DNA was purified from bacterial cultures first by preparing an overnight culture with the appropriate antibiotic (180 RPM, 30 °C). From this, cells were harvested from 5 mL by centrifugation (4000 G, 10 minutes) and the supernatant discarded. The pellet was resuspended and plasmid DNA purified by use of GenElute™ Plasmid Miniprep kits (Merck) according to the manufacturer's instructions. DNA was eluted in 50 µL sterile water and the DNA yield in ng uL^{-1} estimated with a nanodrop spectrophotometer by measuring its absorbance at 260 nm.

DNA sequencing and synthesis

DNA was sequenced using sanger methods by Eurofins Genomics. To 15 µL of purified DNA (~ 100 ng μL^{-1}), 2 µL of sequencing primer (10 µM stock) was added. Samples were then sequenced by Eurofins Genomics under the TubeSeq Supreme™ service.

All DNA or primers used within this work were, unless otherwise specified, synthesised by Eurofins Genomics.

Polymerase chain reaction

All polymerase chain reaction (PCR) described within this work was performed by mixing 25 µL of 2 x Flash Phusion PCR Master mix™, with 1 µL each of 10 µM forward and reverse primer, 0.5 µL template DNA and sterile water added up to 50 µL total volume. This was incubated using a thermocycler with the parameters described below in table 2.2 and an annealing temperature 5 °C below the anticipated primer melting temperature, as estimated by the primer manufacturer.

Table 2.1: PCR cycle parameters used in the preparation of this thesis.

Stage	Temperature (°C)	Duration (minutes)
1	98	3
2	98	0.25
3	Varied	0.5
4	72	4
5	72	12

Stages 2-4 were repeated 35 times before proceeding to stage five.

DNA gel electrophoresis,* DpnI *digestion, PCR clean-up and gel extraction

PCR amplification of DNA was validated by gel electrophoresis. 1% agarose gels were prepared by dissolving 1 g of agarose in 100 mL of Tris Acetate-EDTA (TAE) buffer (40 mM Tris, 20 mM acetic acid, 1 mM EDTA (ethylenediaminetetraacetic acid), pH 7.6), and heated until dissolved and molten. This was poured into a cast and allowed to set, before transfer into a gel tank where it was submerged in further TAE buffer. DNA samples were prepared by mixing in a 1:1 v/v ratio with a solution of 2 x GelRed® Nucleic Acid Stain, 96 g L^{-1} sucrose, 0.48 g L^{-1} orange G dye. Of the resultant mixture, 8 µL was loaded into gel wells. Gel electrophoresis was performed at 120 V, 100 mA for ~1.5 hours, and DNA bands visualised under ultraviolet light. The molecular weight of DNA bands was estimated by comparison with the inclusion of a sample of GeneRuler 1 kb Plus DNA Ladder.

Residual template DNA in PCR reactions was removed by digestion with the enzyme Dpn1. To 4 µL of CutSmart® Buffer, 34 µL of PCR material and 2 µL of Dpn1 enzyme were added, before incubation at 37 °C for 1 hour.

Unless used for gel extraction, PCR synthesised DNA was then purified by use of a GenElute™ PCR Clean-Up Kit according to the manufacturer's instructions. The DNA was eluted in 50 µL of sterile water.

To perform gel extraction, 1 % agarose gels were prepared as described above. In this instance, 150 µL of DNA was mixed with 150 µL of 2 x GelRed® Nucleic Acid Stain, 96 g L^{-1} sucrose, 0.48 g L^{-1} orange G dye. The entire resultant mixture was loaded into large gel well and gel electrophoresis was performed at 120 V, 100 mA for ~1.5 hours. The desired band was excised from the gel under ultraviolet light, and purified by use of a GenElute™ Gel Extraction Kit.

Phosphorylation and ligation of blunt ended DNA

Blunt ended DNA was phosphorylated and re-ligated before transformation into *E. coli* Top10. To do so, 17 µL of linearised DNA was added to 2 µL 10 x T4 Ligase buffer and 1 µL of T4 polynucleotide kinase enzyme. This was incubated at °C for 5 minutes, before being allowed to cool at room temperature for 5 minutes. To the mixture, 1 µL of T4 DNA ligase enzyme was added, the resultant sample was incubated overnight at room temperature.

Sodium dodecyl sulfate polyacrylamide gel electrophoresis

Unless otherwise specified, sodium dodecyl sulphate polyacrylamide gel electrophoresis (SDS-PAGE) of proteins was performed using mPAGE® 4-20% Bis-Tris Precast Gels according to the manufacturer's instructions. Once run, gels were extracted from their casing and washed in water for 15 minutes. Gels were then either stained for *c*-type heme, or protein.

To stain for protein, ~15 mL of Coomassie brilliant blue was added to the gel, in which it was left with agitation overnight. The next day the gels were destained by washing with agitation in water for ~2 hours. After this process, the gel was imaged.

To stain for *c*-type heme, 20 mL of 250 mM sodium acetate pH 4.5 was added. This was left with agitation for 3 minutes, before 20 mg of N,N,N',N'-Tetramethylbenzidine dissolved in 20 mL methanol was added. Finally, 200 µL of hydrogen peroxide (30 % w/w) was added and the gel left with agitation to stain. After this process, the gel was imaged.

Introduction to macromolecular crystallography

The possible function(s) of biological molecules is dictated largely by their three-dimensional structure. The importance of this statement has been recognised for many decades[1], and subsequently enormous research effort has been devoted to determining the structure of biological molecules experimentally[2]. Until relatively recently, X-ray crystallography enjoyed a near monopoly, as the primary technique with which the structure of biological macromolecules could be determined and subsequently the majority of structures currently deposited in the PDB were solved using that technique[3].

Despite recent advances in techniques such as cryoEM and protein structure prediction algorithms, X-ray crystallography will likely remain a core technique. One good reason for this, is the relatively rapid rate of data collection compared with the necessary duration of cryoEM data collection, making crystallography better suited to applications that require high throughput such as structure guided screening of large ligand libraries[4].

The first stage in determining the structure of a macromolecule by X-ray crystallography is almost invariably to obtain a pure and homogenous preparation of that molecule[5]. In many cases it is advisable to first consider

carefully whether the construct from which the protein is expressed can be optimised to increase the chances of yielding ordered crystals[5]. Multidomain proteins might be best approached by splitting into two separate proteins. If the protein includes disordered regions, it is often advisable to remove these sequences from the protein as they can interfere with crystallisation. Finally, crystallisation usually occurs over a period of days or weeks, and thus it is prescient especially in the case of difficult to crystallise proteins, to take as many additional steps to stabilise the sample as is possible. For example, receptor proteins are often crystallised in the presence of their ligand, which can stabilise them significantly[6].

There is a vast array of strategies for protein purification available to crystallographers including different hosts in which to express, including *E. coli* and insect cells. Proteins can be crystallised after purification by various affinity resins, for example with nickel-NTA (nitrilotriacetic acid).

Once a high concentration, pure and homogenous preparation is obtained, crystallisation is attempted usually by high through put screening of chemical space[5]. These screens are commercially available, typically, although not always, they incorporate a buffer, a precipitant compound like polyethylene glycol or ammonium sulfate and salts or other compounds such as calcium chloride or isopropanol. The goal is to achieve a solution in which the concentration of protein and precipitant are such that nucleation of a small number of protein crystals becomes favourable, after which their growth can propagate. To achieve this delicate balance, several different setups can be employed, like hanging or sitting drop vapour diffusion.

Once crystals are obtained, they may first require optimisation. The processes that govern the crystallisation of macromolecules remain poorly understood, so optimisation is largely an empirical endeavour[7]. Typical parameters to vary include protein concentration, precipitant concentration, pH and temperature. A wide variety of crystal shapes and sizes are observed, in general needle shaped crystals are regarded as the least useful, followed by plates, rods and near-cubic morphologies in increasing order of desirability. The improved equipment

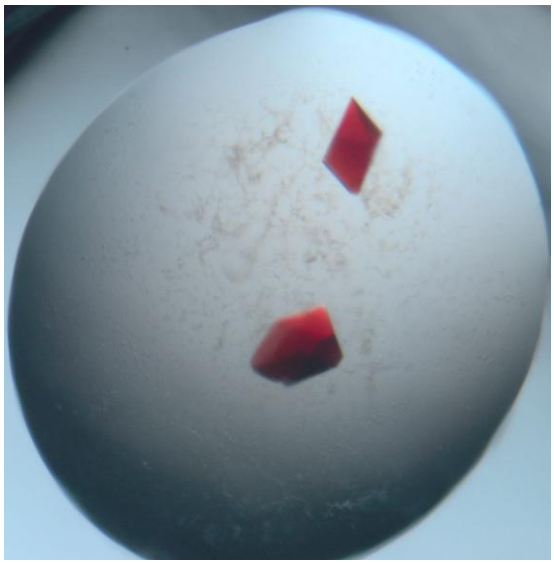


Figure 2.1: Ideal crystals are large, physically robust, cryoprotectant tolerant and well-ordered leading to high resolution diffraction. Shown here, crystals of a truncated isoform of the *S. oneidensis* outer membrane c-type cytochrome MtrC, ~0.3 mm across.

at modern synchrotrons mean good data can now be collected even on plates or needles of very modest sizes.

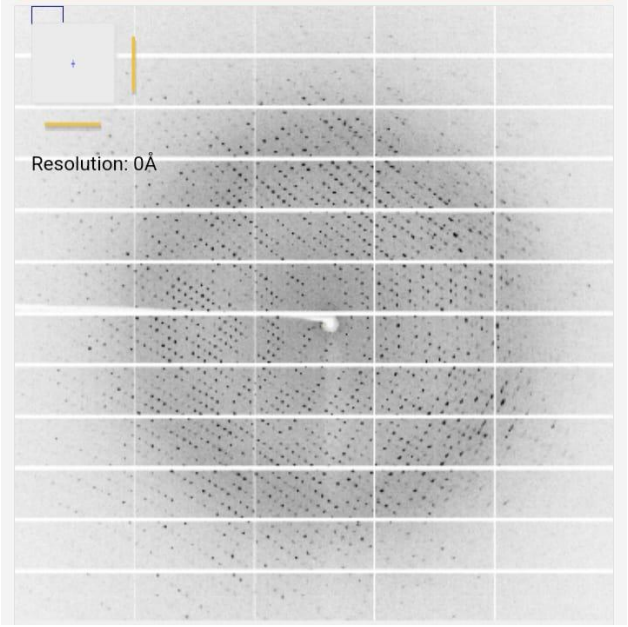


Figure 2.2: Diffraction pattern of *S. oneidensis* MtrC truncated isoform, collected with Eiger2 XE 16M X-ray detector. Note the presence of the beam stop reaching for the left towards the centre, which protects the detector from the beam.

Once good quality crystals are prepared, they are usually cryoprotected by transfer into a solution similar to their growth liquor, but supplemented with cryoprotecting compounds like ethylene glycol or glycerol[8]. They are then frozen in liquid nitrogen, with the objective being no contamination of the sample with non-vitreous ice. The freezing of crystals at liquid nitrogen temperature has proved a crucial innovation as it affords them with orders of magnitude reduced sensitivity to beam induced damage[9].

Collection of data upon the crystals then proceeds, usually by suspending a crystal in a stream of low temperature nitrogen gas, exposing it to the X-ray beam, and detecting the diffraction pattern which is formed behind the crystal. The crystal is then rotated and additional images collected. The defining feature that distinguishes crystals from other types of matter, is that their structure is that of a lattice, a repeated pattern of atoms in 3D space. Subsequently, as the X-ray photons interact with the atoms in the crystal, they scattered elastically as is described by Bragg's law[10] (Equation 2.1), in which N is an integer, reflecting the necessity for waves to be in phase to constructively interfere and be observed, λ is the wavelength of the incident X-ray radiation, θ is the angle at which the wave is scattered, and d is the spacing constant of the scattering species within the crystal.

Equation 2.1: $$N\lambda = 2dsin(\theta)$$

Although it is most intuitive for the crystal to be considered as a near infinite series of real molecules packed together in 3D space, which is a description referred to as real space, in order to relate the spacing constant to the structure of a molecule it is more instructive to consider the crystal in a different way. Since the structural units within the crystal are arranged periodically, crystals can also be described by a periodic function. This is a description that is referred to as reciprocal space or Fourier space[11]. In practice, since X-rays interact significantly only with the electrons of the atoms in the crystal, not their nuclei, this periodic

function only need describe the electrons density in the crystal as a function of displacement, in real space, or phase in reciprocal space.

Since protein crystals are usually comprised of molecules which contain thousands of atoms, their reciprocal space electron density function will be enormously more complex than a simple Y = sin(X) function. However, any complex periodic function can be deconvoluted by a Fourier transformation, into an infinite set of functions that are each described only by a frequency, an amplitude and a phase, and when added together perfectly reproduce the original complex periodic function[11]. This concept has many applications, but in crystallography it describes the process of formation of a diffraction pattern[11]. Each spot (also referred to as a reflection) observed on the X-ray detector corresponds to one periodic function described by a frequency (which is encoded in its position on the detector), an amplitude (which is described by its relative intensity on the detector) and a phase which is not conveyed by any feature of the spot. Together these features describe a concept referred to as a structure factor. Accurately obtaining all the structure factors, allows the determination of the electron density map that can be used to infer the structure of the crystallised molecule.

Several additional considerations are worth noting. Firstly, although to perfectly reproduce a complex periodic function, an infinite number of substituent simple periodic functions may be required, in practice determining a crystal structure does not require observation of an infinite number of spots. Protein crystals are rarely ordered enough to yield spots that correspond to frequencies greater than 1 $Å^{-1}$, so only require observation of a finite number of functions ranging in frequency up to 1 $Å^{-1}$ in order to authentically be described.

Figure 2.3: Crystal of the *Shewanella sp. HRCR-6*, c-type cytochrome UndA during data collection at beamline I04, Diamond Light Source.

Secondly, the fact that diffraction spots lack any feature that conveys their phase, makes the diffraction data useless unless accurate phases can be obtained by other means[12]. Historically, determining experimental phases was achieved largely by exploiting anomalous features in the X-ray scattering profile induced by the presence of heavy elements that absorb X-rays intensely at their absorption edge. As the number of experimentally determined structures in the PDB has risen, and methods for protein

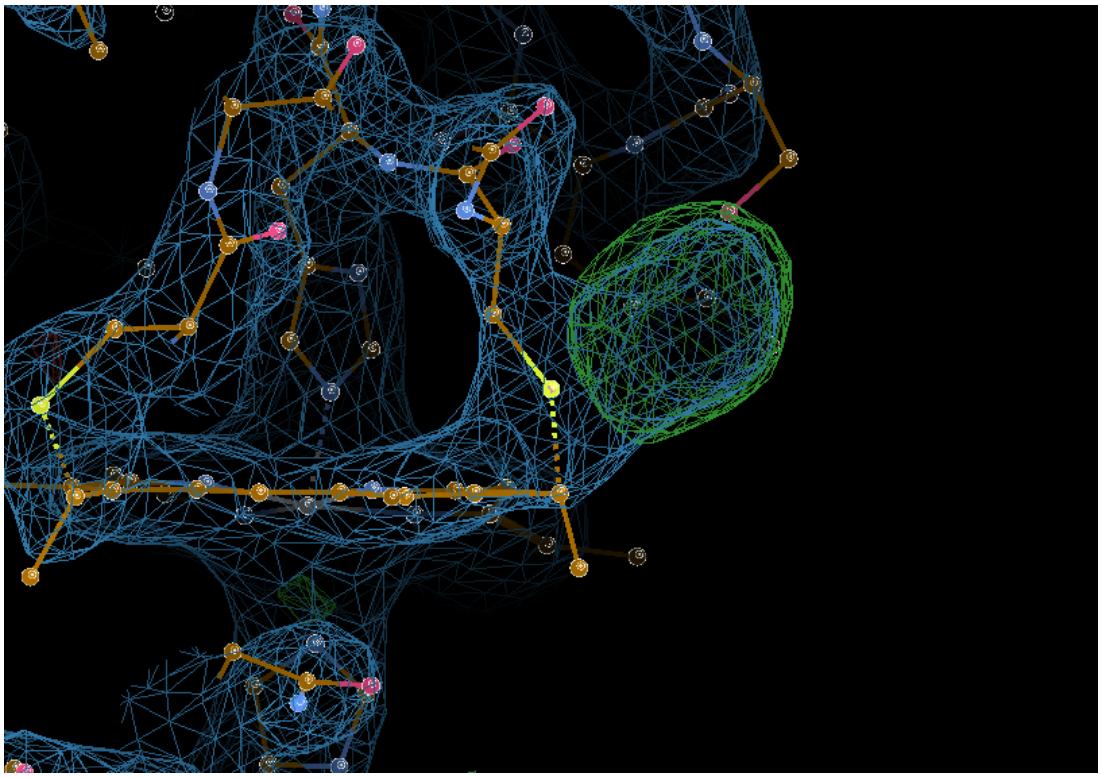


Figure 2.4: Coordinates and electron density maps during refinement of *S. sp. HRCR-6* c-type cytochrome UndA. Double difference density countered at 1.2 σ (blue), and single difference density countered at 3.5 σ (red and green). Note the presence of an unmodelled species causes the appearance of a blob in single difference density maps, from which its position and orientation can ideally be inferred.

structure prediction improved, this strategy is now used sparingly. Instead, the experimentally determined structure of homologous proteins, or model structures can be used to generate predicted phases which are often sufficiently accurate to allow for structure determination. This principle is referred to as molecular replacement[13].

X-rays of wavelength ~1 Å (typical in macromolecular crystallography) are highly energetic so can severely ionise biological material, including crystalline protein. It is therefore important to remember that the act of collecting X-ray data can itself induce changes in the crystalline molecule[14,15]. Beam induced damage can be identified by a characteristic decrease in reported reflection count on the detector as the crystal becomes damaged.

To process diffraction data, the pattern through which the molecules in the crystal pack together must be established[16]. Often multiple identical copies of the molecule (asymmetric unit) are present but orientated differently to one another and together form the unit cell, this is described by the crystal space group. The many images recorded by the detector as the crystal is rotated must be merged together into a single data file and the intensities of reflections assigned an appropriate scale that ensures they are comparable across the data set[17,18]. Modern processing pipelines have largely automated these principles, and also facilitate the merging of multiple distinct datasets which can be

especially useful for radiation sensitive crystals or collections aiming to obtain time resolved insights[19].

Once data is scaled and merged, and phase information incorporated by molecular replacement to generate an initial model, the electron density map can be used to build atoms into their density and refine their position by various automatic and manual methods[20,21]. Improved coordinates will provide phases that are more accurate, and therefore yield improved electron density maps. Subsequently the process of model building and refining is an iterative process.

The agreement of a set of model coordinates with experimental diffraction data is expressed as an R-factor (equation 2.2)[22].

Equation 2.2: $$R = \frac{\sum | F_{observed} - F_{calculated}}{\sum F_{observed}}$$

Where F represents the structure factor amplitude, which is a function of reflection intensity (equation 2.3).

Equation 2.3: $$Intensity \propto |F|^2$$

A model refined to completion will thus reach a minimum R-factor beyond which it cannot be further reduced. It is difficult to compare R-factors between structures solved for different proteins or crystals, but nonetheless higher resolution structures typically do generally display lower R-factors.

Finally, it is worth noting that although interatomic bonds are typically ~2 Å in length, X-ray data even at poorer resolution than 2 Å can be used to accurately refine positions of atomic coordinates since the geometric restraints that act upon protein bonds and angles are now well established, so can be used to infer structure even in ambiguous density. Subsequently, deviation from ideal geometry is also used as a measure of quality of an X-ray derived model[23].

Introduction to small-angle X-ray scattering

X-ray crystallography suffers as a technique from several key drawbacks, the most obvious being that the biological molecules of interest must first be induced to form well-ordered crystals before structure determination can occur. For many proteins this is not possible, and indeed for many others in which movement is an integral part of their cellular function, rigid crystal structures can provide limited insight into their mechanism of action.

Small-angle X-ray scattering (SAXS) addresses this drawback, similarly by scattering X-rays, but on biomolecules in solution[24,25].

Samples of biological molecules are first prepared to high purity. Of note here is the fact that although many of the same sample preparation techniques are applied to SAXS as in X-ray crystallography, constructs that yield flexible sample are no obstacle, and indeed they may be desirable if they represent a more native state of the biomolecule.

Next, samples are subjected to a bright X-ray beam, and the intensity of elastic scattering at small angles is measured by a detector placed behind a sample. A beam stop is also present which protects the detector from non-scattered photons. In SAXS, since the sample molecules are not orientated by ordered packing of any kind, the scattering they induce is equivalent in all radial directions about the incident beam, allowing the data to be averaged radially.

Subsequently the data then consists of the measured intensities, and the corresponding scattering angle they were measured at. The scattered angle could be represented as a distance across the detector, or an angle from the sample, but is instead usually described by the scattering vector magnitude q, which is independent of the distance between the sample and the detector and the wavelength of X-rays used (equation 2.4)[24,25].

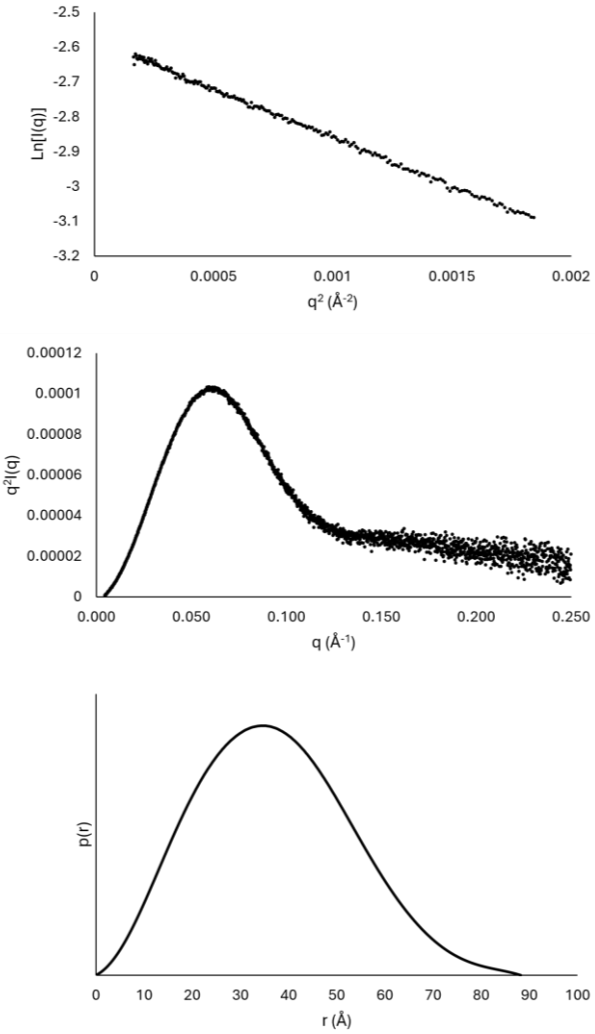


Figure 2.5: SAXS Analysis of BSA, a globular protein. Top - Guinier plot allowing derivation of R_g, middle - Kratky plot showing near-gaussian distribution between q of 0 and 0.1 $Å^{-1}$ indicating globular and folded particle and bottom - p(r) curve indicating D_{max} of 85 Å and gaussian distribution of inter-atomic distances, consistent with a globule shape.

Equation 2.4: $$q = \frac{4\pi \sin(\theta)}{\lambda}$$

Where 2θ represents the scattering angle from the sample, and λ describes the wavelength of the incident X-ray beam.

An important aspect of SAXS, is that the contribution to scattering by X-rays made by the solvent must be properly normalised for. This is because unlike in crystal-induced diffraction, at small angles, proteins only scatter X-rays slightly more intensely than aqueous solutions do[26], and therefore the signal produced by the sample is hidden within a great deal of solvent noise. Thus it is critical that an excellent baseline of the sample buffer is measured so that it can be subtracted from the sample data[24].

One strategy to alleviate this issue, is to collect SAXS data upon the eluent stream of a sample after applying it to a size exclusion chromatography (SEC) column, referred to as SEC-SAXS. This allows data measured just before the sample elutes, just after the sample elutes, or both, to be used as the baseline which should yield accurate baseline normalisation, admittedly at the cost of sample dilution during the SEC[27].

The shape of the intensities as a function of q plot can be analysed to extract information about the structure of the sample biomolecule. One important parameter is the radius of gyration, R_g, which describes the average distance of scattering species within the particle, from its centre of mass, weighted by their relative contribution to the scattering[28]. For globular molecules, R_g will be well is directly a function of particle molecular weight, but for non-globular particles the R_g value will differ from that significantly[29]. R_g can be calculated via the Guinier approximation which is valid only at low q values (equation 2.5)[28].

Equation 2.5: $$I(q) \approx I(0)e^{\frac{-q^2R_g^2}{3}}$$

Where I(q) represents the measured intensity at a scattering vector magnitude q, I(0) represents the (directly unmeasurable) scattered intensity at q = 0, q represents the scattering vector magnitude, and R_g describes the average distance of scattering objects from the particle centre of mass weighted by relative contribution[28]. At small q values, plotting ln[I(q)] against q^2 will yield a straight line, within which R_g can be extracted from the gradient.

Also useful, is the Kratky plot. Plotting $q^2I(q)$ versus q will yield a near gaussian profile for globular particles at low q[30], but for non-globular particles will yield a flattened profile because for these particles the intensity of the scattering they induce decays with less rapidity as q increases[31].

Finally, perhaps the most informative analysis that can be performed on SAXS data is the derivation of a pairwise inter-scatterer distance distribution function usually referred to as p(r). This function describes the relative frequency of distances between centres of electron density with the sample, and can be obtained by computing the Fourier transformation of the scattering profile I(q)[25,32]. This converts the measurements, collected in reciprocal space, into a histogram of real space distances in which local maxima can correspond to inter-domain distances in flexible proteins, making this analysis often very powerful for those samples[25,32]. It conveys also a maximum distance with appreciable frequency D_{max}, which naturally varies drastically between globular and disordered proteins of the same molecular weight[25].

With these SAXS-derived parameters significant insight can be gained, especially for dynamic systems. They can also be used to fit models to the experimental data which allows for improved descriptions of the solution state of biological macromolecules[25,28].

Introduction to single-particle cryogenic electron microscopy

Although SAXS can provide valuable insight into the structure and function of biological macromolecules, the information it provides is fundamentally limited by the isotropic nature of the X-ray scattering involved and subsequently, atomic level detail is not accessible via SAXS. Single particle cryoEM is now a technique that can routinely provide atomic level detail into large biomolecules[33].

Single-particle cryoEM differs significantly from both X-ray crystallography and SAXS because it exploits electrons rather than X-rays for illumination. Although sometimes considered as particles, electrons can also be considered as waves and therefore be used to image biological particles. The wavelength of an electron is described by the de Broglie equation (equation 2.6).

Equation 2.6: $$\lambda = \frac{h}{mv}$$

In which λ represents wavelength, h represents the Planck constant, m represents particle mass and v represents particle velocity.

The first step in successful structure determination in cryoEM, just as is in both X-ray based techniques, is obtaining a high quality and homogenous sample preparation[34]. Naturally since crystallisation is not required, the sample may include protein complexes with poor stability, or membrane proteins which tend not to crystallise productively. As with X-ray based techniques, these samples should be isolated to high-purity usually by some combination of affinity chromatography or ionic chromatography, with SEC[34]. It is helpful to eliminate as much heterogeneity or aggregation from the sample as possible, making SEC a crucial final workflow stage[34], especially given cyroEM is more often applied to samples that are less robust.

The sample is then applied to a solid support structure known as a grid. Grids can be manufactured from different materials but invariably, within the grid, a pattern of holes are engineered precisely.

For analysis of proteins and protein complexes at atomic level detail, near native states are required to achieve this meaning the addition of stains or fixing agents is undesirable. Instead, excess liquid is wicked away from the grid, and it is plunged into a bath of usually liquid ethane at high speed[35]. This ensures sufficiently rapid freezing that the sample particles become fixed within a thin layer of vitreous (amorphous/non-crystalline) ice[35].

This ideally creates a situation in which firstly very little crystalline ice is present, because crystalline ice formation will damage the sample. The sample will ideally be distributed about the hole with a density allowing a large number of particles to be picked, with as few as possible being compromised by the presence of multiple overlapping particles. Thirdly, the ice thickness across the holes will ideally be thick enough to ensure most of the sample is not damaged by interacting with the air-water interface, which is increasingly being recognised

as a significant obstacle to success in cryoEM[36], but thin enough that penetrance by the beam of electrons is high and particles thus display good contrast.

Grids are kept under cryogenic conditions and mounted under vacuum within a microscope. Images are then collected of the particles, at high magnification (100,000 times or higher). The microscope maintains the cryogenic temperature of the sample throughout image collection and importantly also maintains a high vacuum within the column in which the grid is placed[37]. These features ensure the beam of electrons generated at the top of the microscope can efficiently travel down the microscope, pass through the sample which is kept vitreous, and be detected below.

Once high quality images are obtained, a number of preprocessing steps are applied to them. Despite the impressive stability of modern microscope stages, some motion of the specimen is unavoidable, however, the high fidelity of electron detectors that are now deployed widely allows them to record images as movies consisting of multiple frames. These frames can be used to determine the movement of the microscope stage and therefore motion-correct the images[38].

Another important process is estimating, and correcting for, information loss via the contrast transfer function (CTF)[33]. Intrinsic to transmission electron microscopes is loss of information due to microscope aberrations, the most important of which derives from the fact that cryoEM images are actually collected just out of focus (defocus). This is necessary because at focus, particles display no contrast and are therefore impossible to identify against background noise[33]. To explain the CTF, requires describing 2D cryoEM images in reciprocal space, i.e. as a collection of periodic functions described by a frequency, amplitude and phase which when combined reproduce the original image. Amplitudes at particular frequencies, especially those corresponding to high resolution detail, become intensely attenuated by the collection of images at significant defocus. This process is described by the contrast transfer function, which can be modelled in good quality images, allowing for the boosting of frequencies that have been attenuated[39]. This drastically improves image quality although does not totally eliminate information loss since frequencies at which the CTF intercepts the zero contrast axis are not recoverable[39]. They can however be supplemented into the reconstruction by collecting many images at different levels of defocus, thereby shifting the position of CTF zero contrast axis intercept points[33].

With motion corrected images and accurate CTF fitting, particle picking is begun. This process can be done manually, however it is tedious and time consuming and especially in low-contrast micrographs of small (<300 kDa) particles, it can be difficult. Subsequently many programs have been devised that automate the particle picking process, the most popular of which include Topaz and crYOLO[40,41].

With the coordinates of a stack of many thousands of potential particles, bad particles, which may represent damaged protein, contaminants or other artifacts, must be eliminated. This can be done by analysing the particles, statistically, sorting them into sub-stacks of most-similar particles, referred to as 2D classification[42]. Averaging of these sub-stacks builds up signal to noise producing an image which for a good data set, should show clear features that are consistent with that expected for the particle under study. Comparison of these images, especially for particles with a distinctive shape should clearly demonstrate they represent different orientations of the same particle. Bad particles and artifacts will not average into species with clear features, and will usually be sorted into stacks with low overall particle counts[42].

Next, 2D projections must be used to obtain a starting 3D model. Historically this could be done utilising existing crystal structures of individual components within a complex, or with model structures, but these can introduce bias into the final model. Instead most cryoEM processing pipelines now include ab initio model generation. This is usually achieved via an iterative algorithm that seeks to align 2D particle projections to a random noise model, then regenerate the model and re-align the 2D particle projections which, given enough iterations, can ideally produce a good 3D starting model from the initial noise model[43,44].

With a 3D volume model, particles can be further statistically classified, now in 3D, to generate separate models. This can be useful for removing particles which are too low in signal to noise ratio to contribute helpfully to a high-resolution stack, or damaged but were not removed in the 2D classification[45]. It also can allow for separation of particles with subtly different conformations, which may provide useful insight into the particle's mechanism of action.

A final stack of images that represent 2D projections of the particle can then be analysed by what is referred to as 3D refinement. In this process, angles describing the orientation of the particle and values describing its translation about the centre of the image, both in real space are iteratively modelled to greater and greater accuracy. This process allows averaging of many 2D projection images to yield a high resolution 3D volume, and can be achieved by exploiting the projection-slice theorem[46,47]. This theorem states that taking the Fourier transform of a 2D projection image of an object, yields one slice through a 3D Fourier transform of that object with an angle defined by the orientation of the object in that projection[46]. Therefore with many 2D projections, many slices can be obtained, allowing for reconstruction of a near-faithful model of the original object by returning a densely packed 3D slice collection to real space by computing its inverse Fourier transformation. The level of detail in the subsequent reconstruction improves with improvements in modelling of the projection angles and translations.

In the early days of cryoEM much effort was dedicated to ensuring this process remained unbiased[48]. This resulted in several innovations including low-pass

filtering initial models, automation of the refinement process to remove user bias and the so called “gold standard” fourier shell correlation (FSC) for determining refinement convergence and resolution[47]. In this strategy, the dataset is refined separately as two halves, and the correlation between the two in reciprocal space calculated as a function of spatial frequency[46]. The community agreed upon cut off of 0.143 cross correlation has been accepted as a generally acceptable point from which to report a spatial frequency as a robust resolution estimate[48,49].

Further boosts to map quality can also often be gained by improvements in the accuracy of the correction for motion of the sample during image collection, or modelling, and correction, for the CTF. These gains are achieved by now assessing these two properties on a per-particle basis, rather than a per-micrograph basis which can yield a significant improvement in final resolution[50,51].

Finally, cryoEM map volumes are comprised not only of a protein particle, but also a significant quantity of aqueous solvent. The solvent volume carries no useful information and causes underestimation of the true resolution by FSC due to its only contributing noise. cryoEM data processing pipelines thus include algorithms for masking out the effect of the solvent volume, and boosting the amplitudes of information that is high frequency, in reciprocal space, within the reconstruction[52]. This procedure is important since these frequencies are harder to observe so they are subsequently always underrepresented in data sets[52].

Interpretation of cryoEM density maps is often made significantly easier by generating an atomic model for the observed density, although note than unlike X-ray crystallography, in which coordinates are required to achieve phasing of the X-ray data, this is not an integral part of cryoEM. With some ~200,000 structures deposited in the PDB it is often possible to dock into the map an existing atomic structure of at least some of the proteins that comprise the particle[53]. If no experimentally determined structure is available, modern protein structure prediction algorithms can be employed to dock in predicted structures[54]. If neither of these strategies yield satisfactory results, or if the identity of the protein that provided the density is not known, but the map is of sufficient quality, automatic de novo model building programs have been written which can trace an amino acid backbone and infer side chain identities, with remarkable precision[55].

Once a starting atomic model is obtained, this can be refined to improve its accuracy in terms of fitting of atoms in density, and conformity to established atomic geometry within proteins[56]. Although many principles are shared between the two, this process differs somewhat from crystallographic model refinement since as noted above, cryoEM real space maps do not suffer from phasing ambiguity. Ultimately it is best that inferences from maps or models are evaluated not against reported FSC intercept value or X-ray reflection spatial

frequency, but instead from the detail actually present within the density, and how well the atomic model matches that density.

Chapter 2: References

1. Kendrew, J. C., Bodo, G., Dintzis, H. M., Parrish, R. G., Wyckoff, H. & Phillips, D. C. A Three-Dimensional Model of the Myoglobin Molecule Obtained by X-Ray Analysis. *Nature 1958 181:4610* **181**, 662–666 (1958).

2. Jaskolski, M., Dauter, Z. & Wlodawer, A. A brief history of macromolecular crystallography, illustrated by a family tree and its Nobel fruits. *The FEBS journal* **281**, 3985 (2014).

3. Hanske, J., Sadian, Y. & Müller, C. W. The cryo-EM resolution revolution and transcription complexes. *Current Opinion in Structural Biology* **52**, 8 (2018).

4. Saur, M., Hartshorn, M. J., Dong, J., Reeks, J., Bunkoczi, G., Jhoti, H. & Williams, P. A. Fragment-based drug discovery using cryo-EM. *Drug Discovery Today* **25**, 485–490 (2020).

5. Pusey, M. L., Liu, Z. J., Tempel, W., Praissman, J., Lin, D., Wang, B. C., Gavira, J. A. & Ng, J. D. Life in the fast lane for protein crystallization and X-ray crystallography. *Progress in Biophysics and Molecular Biology* **88**, 359–386 (2005).

6. Kimura, K. T., Asada, H., Inoue, A., Kadji, F. M. N., Im, D., Mori, C., Arakawa, T., Hirata, K., Nomura, Y., Nomura, N., Aoki, J., Iwata, S. & Shimamura, T. Structures of the 5-HT2A receptor in complex with the antipsychotics risperidone and zotepine. *Nature Structural & Molecular Biology 2019 26:2* **26**, 121–128 (2019).

7. DeLucas, L. J., Hamrick, D., Cosenza, L., Nagy, L., McCombs, D., Bray, T., Chait, A., Stoops, B., Belgovskiy, A., William Wilson, W., Parham, M. & Chernov, N. Protein crystallization: virtual screening and optimization. *Progress in Biophysics and Molecular Biology* **88**, 285–309 (2005).

8. Jang, K., Kim, H. G., Hlaing, S. H. S., Kang, M., Choe, H. W. & Kim, Y. J. A Short Review on Cryoprotectants for 3D Protein Structure Analysis. *Crystals* **12**, 138 (2022).

9. Sandy, J., Mikolajek, H., Thompson, A. J., Sanchez-Weatherby, J. & Hough, M. A. Crystallization and In Situ Room Temperature Data Collection Using the Crystallization Facility at Harwell and Beamline VMXi, Diamond Light Source. *Journal of Visualized Experiments (JoVE)* **2024**, e65964 (2024).

10. Bragg, W. H. & Bragg, W. L. The reflection of X-rays by crystals. *Proceedings of the Royal Society of London. Series A, Containing Papers of a Mathematical and Physical Character* **88**, 428–438 (1913).

11. Sherwood, D. & Cooper, J. *Crystals, X-Rays and Proteins : Comprehensive Protein Crystallography*. (OUP Oxford, Oxford, 2011).

12. Hauptman, H. Phasing methods for protein crystallography. *Current Opinion in Structural Biology* **7**, 672–680 (1997).

13. McCoy, A. J., Sammito, M. D. & Read, R. J. Implications of AlphaFold2 for crystallographic phasing by molecular replacement. *Acta crystallographica. Section D, Structural biology* **78**, 1–13 (2022).

14. Moreno-Chicano, T., Ebrahim, A., Axford, D., Appleby, M. V., Beale, J. H., Chaplin, A. K., Duyvesteyn, H. M. E., Ghiladi, R. A., Owada, S., Sherrell, D. A., Strange, R. W., Sugimoto, H., Tono, K., Worrall, J. A. R., Owen, R. L. & Hough, M. A. High-throughput structures of protein-ligand complexes at room temperature using serial femtosecond crystallography. *IUCrJ* **6**, 1074–1085 (2019).

15. Hough, M. A. & Owen, R. L. Serial synchrotron and XFEL crystallography for studies of metalloprotein catalysis. *Current Opinion in Structural Biology* **71**, 232–238 (2021).

16. Sauter, N. K., Grosse-Kunstleve, R. W. & Adams, P. D. Robust indexing for automatic data collection. *Journal of Applied Crystallography* **37**, 399 (2004).

17. Winter, G., Waterman, D. G., Parkhurst, J. M., Brewster, A. S., Gildea, R. J., Gerstel, M., Fuentes-Montero, L., Vollmar, M., Michels-Clark, T., Young, I. D., Sauter, N. K. & Evans, G. DIALS: implementation and evaluation of a new integration package. *Acta Crystallographica. Section D, Structural Biology* **74**, 85 (2018).

18. Evans, P. Scaling and assessment of data quality. *Acta crystallographica. Section D, Biological crystallography* **62**, 72–82 (2006).

19. Gildea, R. J., Beilsten-Edmands, J., Axford, D., Horrell, S., Aller, P., Sandy, J., Sanchez-Weatherby, J., David Owen, C., Lukacik, P., Strain-Damerell, C., Owen, R. L., Walsh, M. A. & Winter, G. Xia2.multiplex: A multi-crystal data-Analysis pipeline. *Acta Crystallographica Section D: Structural Biology* **78**, 752–769 (2022).

20. Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. Features and development of Coot. *Acta crystallographica. Section D, Biological crystallography* **66**, 486–501 (2010).

21. Afonine, P. V., Grosse-Kunstleve, R. W., Echols, N., Headd, J. J., Moriarty, N. W., Mustyakimov, M., Terwilliger, T. C., Urzhumtsev, A., Zwart, P. H. & Adams, P. D. Towards automated crystallographic structure refinement with phenix.refine. *Acta crystallographica. Section D, Biological crystallography* **68**, 352–367 (2012).

22. Brünger, A. T. Free R value: a novel statistical quantity for assessing the accuracy of crystal structures. *Nature 1992 355:6359* **355**, 472–475 (1992).

23. Morris, A. L., MacArthur, M. W., Hutchinson, E. G. & Thornton, J. M. Stereochemical quality of protein structure coordinates. *Proteins: Structure, Function, and Bioinformatics* **12**, 345–364 (1992).

24. Skou, S., Gillilan, R. E. & Ando, N. Synchrotron-based small-angle X-ray scattering of proteins in solution. *Nature Protocols 2014 9:7* **9**, 1727–1739 (2014).

25. Putnam, D. K., Lowe, E. W. & Meiler, J. Reconstruction of SAXS Profiles from Protein Structures. *Computational and Structural Biotechnology Journal* **8**, e201308006 (2013).

26. Svergun, D. I. & Koch, M. H. J. Small-angle scattering studies of biological macromolecules in solution. *Reports on Progress in Physics* **66**, 1735 (2003).

27. Bucciarelli, S., Midtgaard, S. R., Pedersen, M. N., Skou, S., Arleth, L. & Vestergaard, B. Size-exclusion chromatography small-angle X-ray scattering of water soluble proteins on a laboratory instrument. *Journal of Applied Crystallography* **51**, 1623–1632 (2018).

28. Jacques, D. A. & Trewhella, J. Small-angle scattering for structural biology—Expanding the frontier while avoiding the pitfalls. *Protein Science* **19**, 642–657 (2010).

29. Smilgies, D. M. & Folta-Stogniew, E. Molecular weight–gyration radius relation of globular proteins: a comparison of light scattering, small-angle X-ray scattering and structure-based data. *Journal of Applied Crystallography* **48**, 1604 (2015).

30. Burger, V. M., Arenas, D. J. & Stultz, C. M. A Structure-free Method for Quantifying Conformational Flexibility in proteins. *Scientific Reports 2016 6:1* **6**, 1–9 (2016).

31. Bernadó, P. & Svergun, D. I. Structural analysis of intrinsically disordered proteins by small-angle X-ray scattering. *Molecular BioSystems* **8**, 151–167 (2011).

32. Hansen, S. Bayesian estimation of hyperparameters for indirect Fourier transformation in small-angle scattering. *Journal of Applied Crystallography* **33**, 1415–1421 (2000).

33. Cheng, Y. Single-particle cryo-EM-How did it get here and where will it go. *Science* **361**, 876–880 (2018).

34. Cianfrocco, M. A. & Kellogg, E. H. What Could Go Wrong? A Practical Guide to Single-Particle Cryo-EM: From Biochemistry to Atomic Models. *Journal of Chemical Information and Modeling* **60**, 2458–2469 (2020).

35. Dubochet, J. & McDowall, A. W. VITRIFICATION OF PURE WATER FOR ELECTRON MICROSCOPY. *Journal of Microscopy* **124**, 3–4 (1981).

36. Liu, N. & Wang, H. W. Better Cryo-EM Specimen Preparation: How to Deal with the Air–Water Interface? *Journal of Molecular Biology* **435**, 167926 (2023).

37. Egelman, E. H. The Current Revolution in Cryo-EM. *Biophysical Journal* **110**, 1008–1012 (2016).

38. Zheng, S. Q., Palovcak, E., Armache, J. P., Verba, K. A., Cheng, Y. & Agard, D. A. MotionCor2 - anisotropic correction of beam-induced motion for improved cryo-electron microscopy. *Nature methods* **14**, 331 (2017).

39. Rohou, A. & Grigorieff, N. CTFFIND4: Fast and accurate defocus estimation from electron micrographs. *Journal of Structural Biology* **192**, 216–221 (2015).

40. Bepler, T., Morin, A., Rapp, M., Brasch, J., Shapiro, L., Noble, A. J. & Berger, B. Positive-unlabeled convolutional neural networks for particle picking in cryo-electron micrographs. *Nature Methods 2019 16:11* **16**, 1153–1160 (2019).

41. Wagner, T., Merino, F., Stabrin, M., Moriya, T., Antoni, C., Apelbaum, A., Hagel, P., Sitsel, O., Raisch, T., Prumbaum, D., Quentin, D., Roderer, D., Tacke, S., Siebolds, B., Schubert, E., Shaikh, T. R., Lill, P., Gatsogiannis, C. & Raunser, S. SPHIRE-crYOLO is a fast and accurate fully automated particle picker for cryo-EM. *Communications Biology 2019 2:1* **2**, 1–13 (2019).

42. Kimanius, D., Dong, L., Sharov, G., Nakane, T. & Scheres, S. H. W. New tools for automated cryo-EM single-particle analysis in RELION-4.0. *Biochemical Journal* **478**, 4169–4185 (2021).

43. Elmlund, H., Elmlund, D. & Bengio, S. PRIME: Probabilistic initial 3D model generation for single-particle cryo-electron microscopy. *Structure* **21**, 1299–1306 (2013).

44. Reboul, C. F., Eager, M., Elmlund, D. & Elmlund, H. Single-particle cryo-EM—Improved ab initio 3D reconstruction with SIMPLE/PRIME. *Protein Science* **27**, 51 (2017).

45. Barchet, C., Fréchin, L., Holvec, S., Hazemann, I., von Loeffelholz, O. & Klaholz, B. P. Focused classifications and refinements in high-resolution single particle cryo-EM analysis. *Journal of Structural Biology* **215**, 108015 (2023).

46. Scheres, S. H. W. A Bayesian View on Cryo-EM Structure Determination. *Journal of Molecular Biology* **415**, 406–418 (2012).

47. Scheres, S. H. W. RELION: Implementation of a Bayesian approach to cryo-EM structure determination. *Journal of Structural Biology* **180**, 519 (2012).

48. Henderson, R., Sali, A., Baker, M. L., Carragher, B., Devkota, B., Downing, K. H., Egelman, E. H., Feng, Z., Frank, J., Grigorieff, N., Jiang, W., Ludtke, S. J., Medalia, O., Penczek, P. A., Rosenthal, P. B., Rossmann, M. G., Schmid, M. F., Schröder, G. F., Steven, A. C. *et al.* Outcome of the first electron microscopy validation task force meeting. *Structure* **20**, 205–214 (2012).

49. Rosenthal, P. B. & Henderson, R. Optimal Determination of Particle Orientation, Absolute Hand, and Contrast Loss in Single-particle Electron Cryomicroscopy. *Journal of Molecular Biology* **333**, 721–745 (2003).

50. Bai, X. C., Fernandez, I. S., McMullan, G. & Scheres, S. H. W. Ribosome structures to near-atomic resolution from thirty thousand cryo-EM particles. *eLife* **2013**, (2013).

51. Zivanov, J., Nakane, T., Forsberg, B. O., Kimanius, D., Hagen, W. J. H., Lindahl, E. & Scheres, S. H. W. New tools for automated high-resolution cryo-EM structure determination in RELION-3. *eLife* **7**, (2018).

52. Scheres, S. H. W. Processing of Structurally Heterogeneous Cryo-EM Data in RELION. *Methods in Enzymology* **579**, 125–157 (2016).

53. Roseman, A. M. Docking structures of domains into maps from cryo-electron microscopy using local correlation. *Acta Crystallographica Section D: Biological Crystallography* **56**, 1332–1340 (2000).

54. Millán, C., McCoy, A. J., Terwilliger, T. C. & Read, R. J. Likelihood-based docking of models into cryo-EM maps. *Acta Crystallographica Section D: Structural Biology* **79**, 281–289 (2023).

55. Jamali, K., Käll, L., Zhang, R., Brown, A., Kimanius, D. & Scheres, S. H. W. Automated model building and protein identification in cryo-EM maps. *Nature 2024 628:8007* **628**, 450–457 (2024).

56. Afonine, P. V., Poon, B. K., Read, R. J., Sobolev, O. V., Terwilliger, T. C., Urzhumtsev, A. & Adams, P. D. Real-space refinement in PHENIX for cryo-EM and crystallography. *Acta Crystallographica Section D: Structural Biology* **74**, 531–544 (2018).

Chapter 3: The preparation and structure of heme containing isoforms of the *c*-type cytochrome PgcA

Introduction

For a Gram-negative bacterium to couple energy conservation with the reduction of insoluble substrates, the reduction of such substrates must occur in the extracellular milieu. Consequently, electrons derived from the quinol pool must pass across the cellular envelope and reach cell surface reductase enzymes[1–3].

This process is now known to involve electron transfer through cytochrome proteins containing several heme cofactors[1]. These hemes are typically bis-histidine coordinated and tightly packed within the protein, such that the porphyrin rings are usually ≤5 Å apart, enabling rapid electron exchange across the cytochrome[4–6]. In *G. sulfurreducens* these cytochromes include inner membrane quinol dehydrogenases, periplasmic electron shuttles, outer-membrane associated cytochromes and extracellular polymeric cytochrome wires that are composed of multiheme modules of 4-8 hemes[5]. In all these cytochromes, the heme cofactors are either predicted, or experimentally shown, to be packed into tight and rigid chains.

In addition to this broad range of multiheme cytochromes required for efficient EET, *G. sulfurreducens* PCA also produces a triheme cytochrome known as PgcA[7]. This protein can have a significant role in the reduction of metal oxides[8], since a mutant *G. sulfurreducens ΔpgcA* strain was capable of reducing soluble iron(III) citrate but not insoluble iron(III) and manganese(IV) oxides[9].

The amino acid sequence of PgcA contains (Figure 3.1A) a predicted cleavable Sec signal peptide followed by an N-terminal lipid site, and four domains interspersed by repeated $(PT)_x$ motifs[10]. The first domain is not predicted to contain any cofactors, while the other three domains each contain a single *c*-type heme[10]. Recombinant expression of *pgcA* in *Shewanella oneidensis* has been reported and resulted in the synthesis of a truncated PgcA, not recognised by the type II secretion system, and trapped in the periplasm[9].

The individual monoheme domains of PgcA have recently been cloned and their successful recombinant expression has also been reported[10]. NMR spectroscopy of those monoheme domains suggested that inter-domain interactions were transient in nature, indicating that the protein is unlikely to

transfer electrons via a rigid heme chain, as observed in other multiheme cytochromes[10].

The existence of a cytochrome with an important role in EET that has its heme cofactors separated into distinct domains interspersed by repeat motifs is unprecedented. Furthermore it remains unclear what implications this unique structural arrangement would have for the mechanism of electron transfer, and cellular function, of the protein. The work described within this chapter sought to obtain pure preparations of heme containing isoforms of PgcA, to then determine their structures and examine their other biophysical properties. Its findings demonstrate PgcA to be the first of a new class of redox enzymes, with redox centres separated by nanometres, but tethered together by flexible linkers.

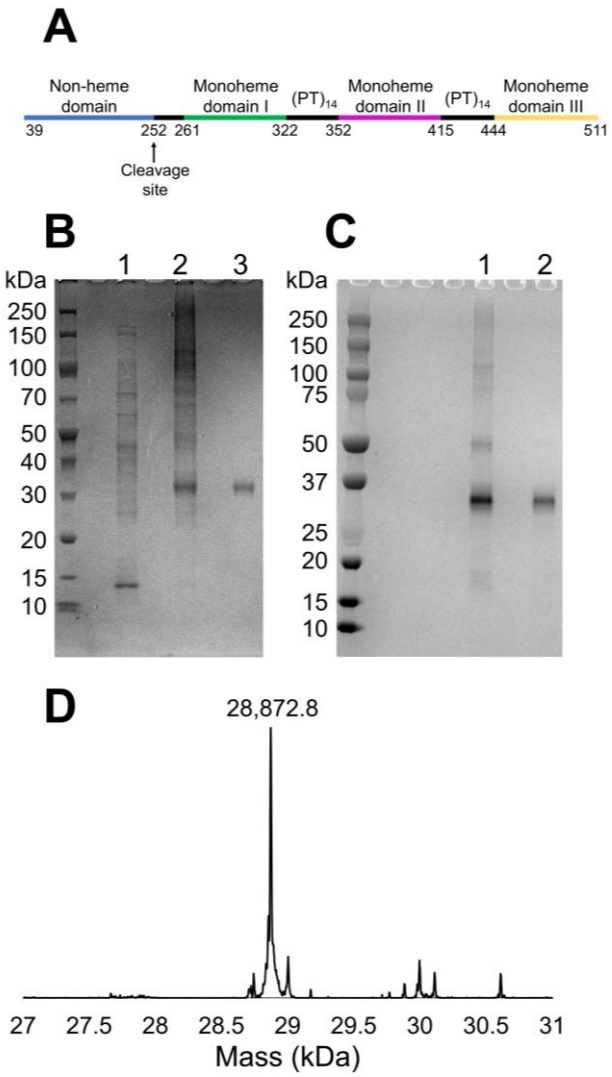


Figure 3.1: Recombinant PgcA is a triheme cytochrome containing 3 monoheme domains. (A) Domain architecture of PgcA. Cleavage at Met252 results in a triheme PgcA containing three monoheme cytochrome domains interspersed by Pro-Thr motifs. **(B)** Coomassie stained gradient SDS-PAGE gel of triheme PgcA and purification intermediates. Lane 1: Clarified cell lysate. Lane 2: Strep-Tactin affinity chromatography eluent. Lane 3: Size exclusion chromatography isolated triheme PgcA. **(C)** Peroxidase-heme stained gradient SDS-PAGE gel of triheme PgcA. Lane 1: Strep-Tactin affinity chromatography eluent. Lane 2: ~30 kDa triheme PgcA isolated by size exclusion chromatography. . **(D)** Deconvoluted mass spectra for triheme PgcA, residue 253-C terminus fragment predicted molecular weight from amino acid sequence: 28,872.4 Da.

Results

Heterologous expression of PgcA results in a monomeric triheme protein

Recombinant PgcA with a C-terminal Strep-II purification tag was isolated from *S. oneidensis* MR-1 cell extracts by affinity chromatography as described in *Chapter 3 Methods*. Pooled PgcA fractions were analysed using SDS-PAGE and stained with Coomassie and for heme, revealing a minor band at 50 kDa and a major band at ~30 kDa (Figure 3.1B & C). These two heme-containing bands both

bound to the Strep-Tactin affinity resin, consistent with the isolation of two different PgcA isoforms, as previously observed for heterologously expressed PgcA[9].

Using Size exclusion chromatography, the two isoforms were separated resulting in the isolation of the ~30 kDa form of PgcA. LC-MS analysis revealed its mass was 28,873 Da (Figure 3.1D), within 1 Da of the predicted mass for a recombinant PgcA containing three C-hemes (each contributing mass 615.17 Da as described in *Yang et al*[11]) and having undergone proteolytic cleavage between Met252 and Pro253 at the C-terminus of the non-heme domain. This species is referred to subsequently as triheme PgcA.

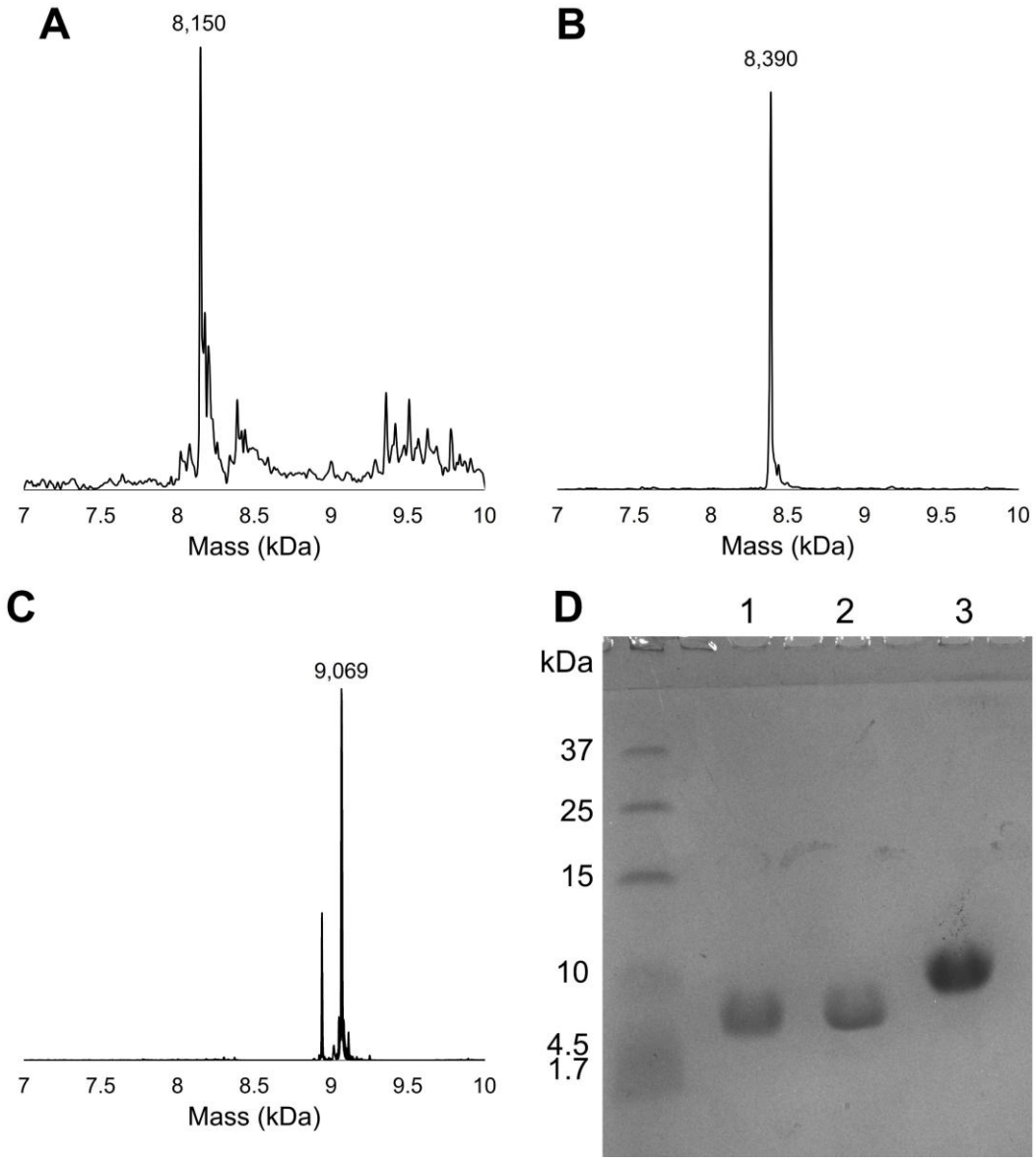


Figure 3.2: LC-MS and SDS-PAGE of monoheme domains. Deconvoluted mass spectra for PgcA monoheme domains I **(A)**, II **(B)** and III **(C)**, Theoretical molecular weights calculated from mature sequence and 615.17 Da c-heme contribution are 8,152, 8,389 and 9,069 Da, respectively. **(D)** Coomassie stained gel of purified PgcA monoheme domains. Lane 1: Domain I. Lane 2: domain II. Lane 3: domain III.

Purification and X-ray crystallographic analysis of individual PgcA heme domains I, II and III

Plasmids containing genes encoding the monoheme PgcA domains I, II and III were prepared as described in *Chapter 3 Methods*. These plasmid constructs, called pPGCA-DI, pPGCA-DII and pPGCA-DIII also included a Strep-II purification tags at their C-terminus and the *S. oneidensis* MR-1 MtrB signal peptide at their N-terminus, which was previously used to overexpress recombinant cytochromes in *S. oneidensis*[12]. The constructs were overexpressed in *S. oneidensis* MR-1 and the proteins purified as described in *Chapter 3 Methods*.

The masses of the purified proteins were determined by LC-MS to be within 2 Da of that predicted from their amino acid sequence, and they displayed a single band at a molecular weight consistent with LC-MS when analysed by Coomassie stained SDS-PAGE (Figure 3.2)

From these data, it was concluded that successful preparation of pure PgcA monoheme domains had been achieved. These samples were then subjected to crystallisation trials using a Douglas Instruments Orxy Nano robot to dispense drops into Swissci 96-well 2-drop sitting drop trays. Crystals with plate-like, plate and needle morphologies were obtained for domains I, II and III respectively (Figure 3.3).

The atomic structures of the three PgcA monoheme domains were then resolved by X-ray diffraction (Chapter 3 appendix figures 1-3). The 1.55 Å crystal structure of monoheme domain I reveals a globular arrangement comprising residues 261-321 from the PgcA amino acid sequence (Figure 3.4A). The protein contains a single *c*-heme cofactor covalently attached via thioether bonds to the cysteines of the CXXCH motif. The heme is enclosed in a peptide fold consisting of three short α-helices separated by flexible loops, with the CXXCH motif located within the first α-helix. The axial ligation of the heme iron is provided by His274 on the proximal side and Met301 on the distal side. The imidazole group of His274 is positioned within H-bonding distance (2.8 Å) of the carboxyl group of Glu281 (Figure 3.5A). This interaction could influence the properties of the heme group, maintaining a suitable E_m to facilitate the transfer of electrons onto low potential

Figure 3.3: Crystals of PgcA monoheme domains. Left to right respectively, crystals of PgcA monoheme domain I, II and III obtained via sitting drop diffusion at protein concentrations of 10, 10 and 20 mg mL^{-1} respectively.

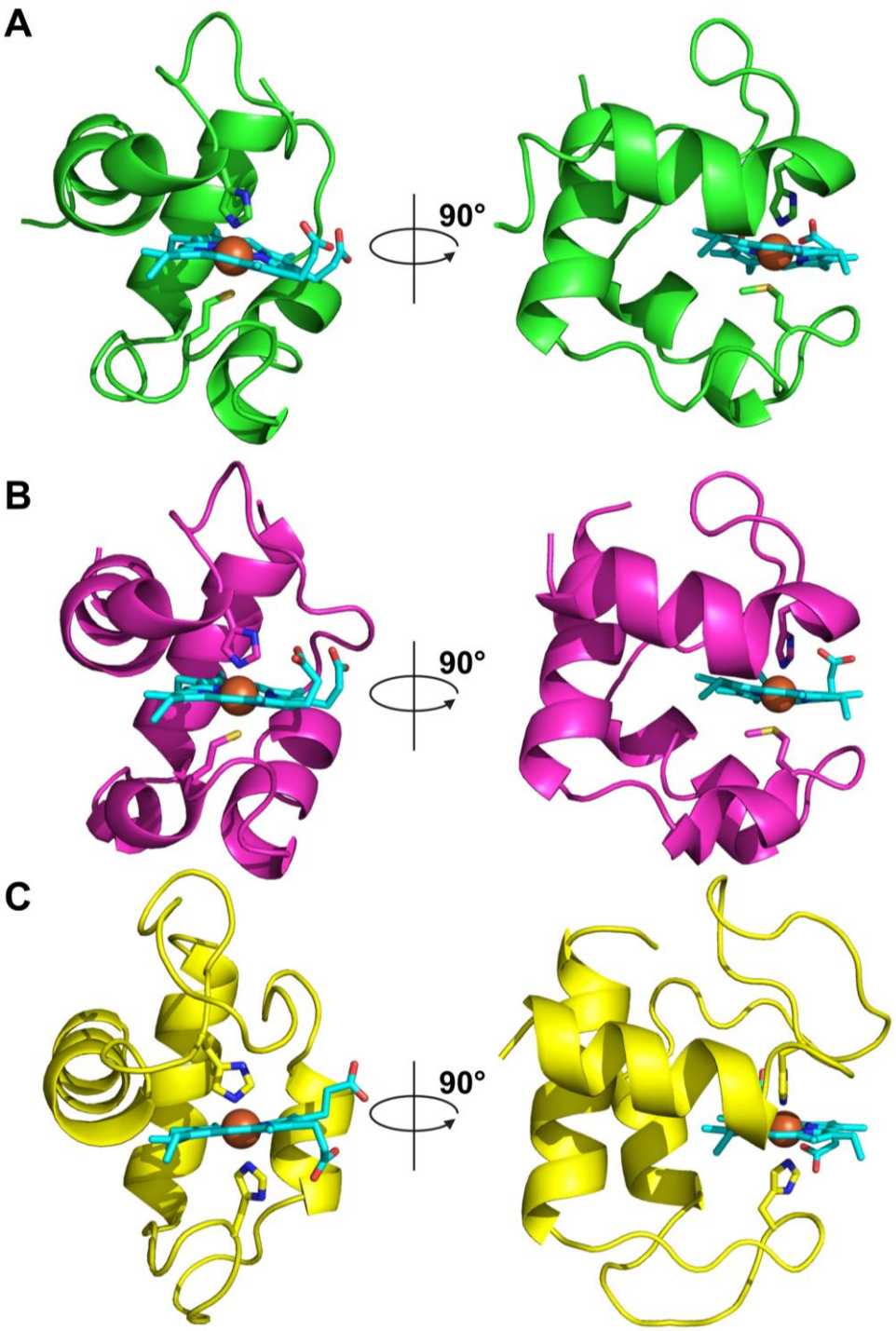


Figure 3.4: X-ray crystal structures of PgcA monoheme domains. (A) Monoheme domain I, green, 1.55 Å resolution. **(B)** Monoheme domain II, magenta, 1.80 Å resolution. **(C)** Monoheme domain III, yellow, 1.90 Å resolution. Heme iron-ligating residues represented with sticks.

minerals which may otherwise be challenging since His-Met heme ligation is typically associated with higher potential hemes[13].

The 1.8 Å crystal structure of heme domain II (Figure 3.4B) consisted of six copies of the protein within the asymmetric unit, comprising residues 352-414 of the PgcA amino acid sequence. The overall fold is the same within each copy (average pairwise RMSD = 0.134 Å) and consists of four loop-linked α-helices encapsulating the *c*-type heme moiety. The structures of domains I and II are highly similar, with a root mean square deviation (RMSD) of 0.453 Å and share 53% sequence identity. Like domain I, domain II displays His-Met heme ligation (in this case from residues His365 and Met392) and the imidazole group of His365 is within H-bonding distance (2.7 Å) of the carboxylate group of Asp372 (Figure 3.5B). This suggests that PgcA domains I and II likely arose from a gene duplication event, and that both the heme axial ligands and the histidine coordinated acidic residues are important structural features of these domains.

The 1.9 Å crystal structure of heme domain III (Figure 3.4C) was composed of PgcA residues 444-511 and is significantly different from those of domains I and II. First, coordination of the iron atom within the *c*-type heme is bis-His (His459 and His491), as commonly observed in multiheme cytochromes. Secondly, no acidic side chain is observed near the distal ligating histidine residue. Domain III contains only three α-helices despite its molecular weight being ~1 kDa larger than those of the preceding two domains. In fact, a larger fraction of its polypeptide chain forms loops that likely possess greater flexibility when unrestricted by crystallisation. This crystal structure also shows that 33% of the total surface area of the *c*-type heme moiety is solvent exposed, which is substantially larger than the 25% exposure of the hemes in domains I and II.

Biophysical analysis of the solution structure of PgcA

The biophysical properties of PgcA and the three monoheme domains were analysed using sedimentation velocity. Continuous sedimentation distribution c(S) analysis revealed single, monodisperse species with $s_{20,W}$ values of 1.17, 1.27, 1.36 S, corresponding to molecular weights of 8.9, 9.5 and 10.5 kDa for

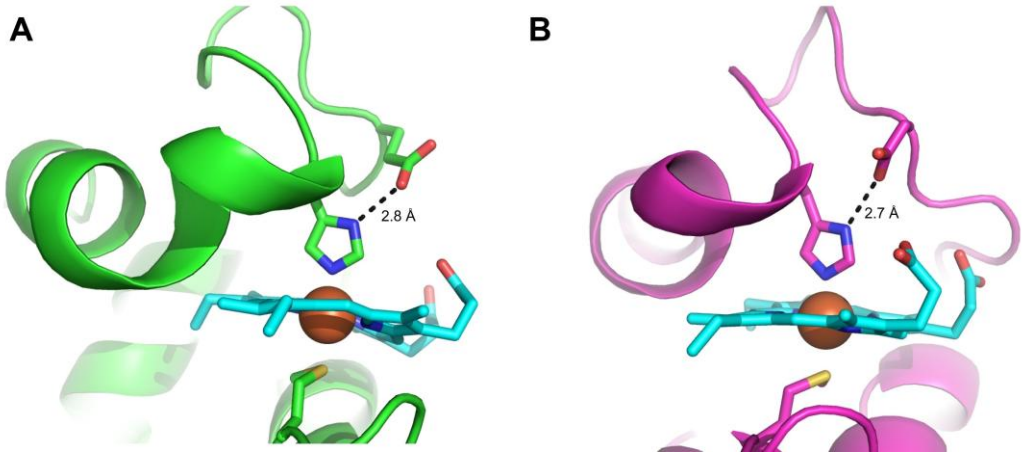


Figure 3.5: Crystographically resolved conformation of E281 and D372 sidechains . Cartoon representation of PgcA monoheme domains I **(A)** and II (**B**) crystal structures at 1.55 and 1.8 Å resolution respectively, with heme, iron-ligating residues and histidine H-bonding residues E281 and D372 represented with sticks.

domains I, II and III, respectively (Figure 3.6A-C). c(S) analysis of a mixed sample containing all three domains revealed a single peak with an $S_{20,W}$ value of 1.28 S, indicating no significant complex formation between the three domains (Figure 3.6D). c(S) analysis of the triheme PgcA revealed a major peak at 1.98 S (31.2 kDa), consistent with the presence of a single PgcA species in solution (Figure 3.6E) which was in agreement with its SEC trace (Figure 3.6F). The fitted f/f_0 ratio of triheme PgcA was 1.81, indicating a significantly elongated structure, rather than the globular structures of the monoheme domains (f/f_0 ~1.3). All these data are consistent with three heme-containing domains being connected by flexible linkers, rather than a closely packed globular structure.

Size-exclusion chromatography coupled with small-angle X-ray scattering (SEC-SAXS) was undertaken on triheme PgcA samples at 2.5,5 and 10 mg mL^{-1}. Samples eluted at the same positions and gave a uniform radius of gyration (R_g)

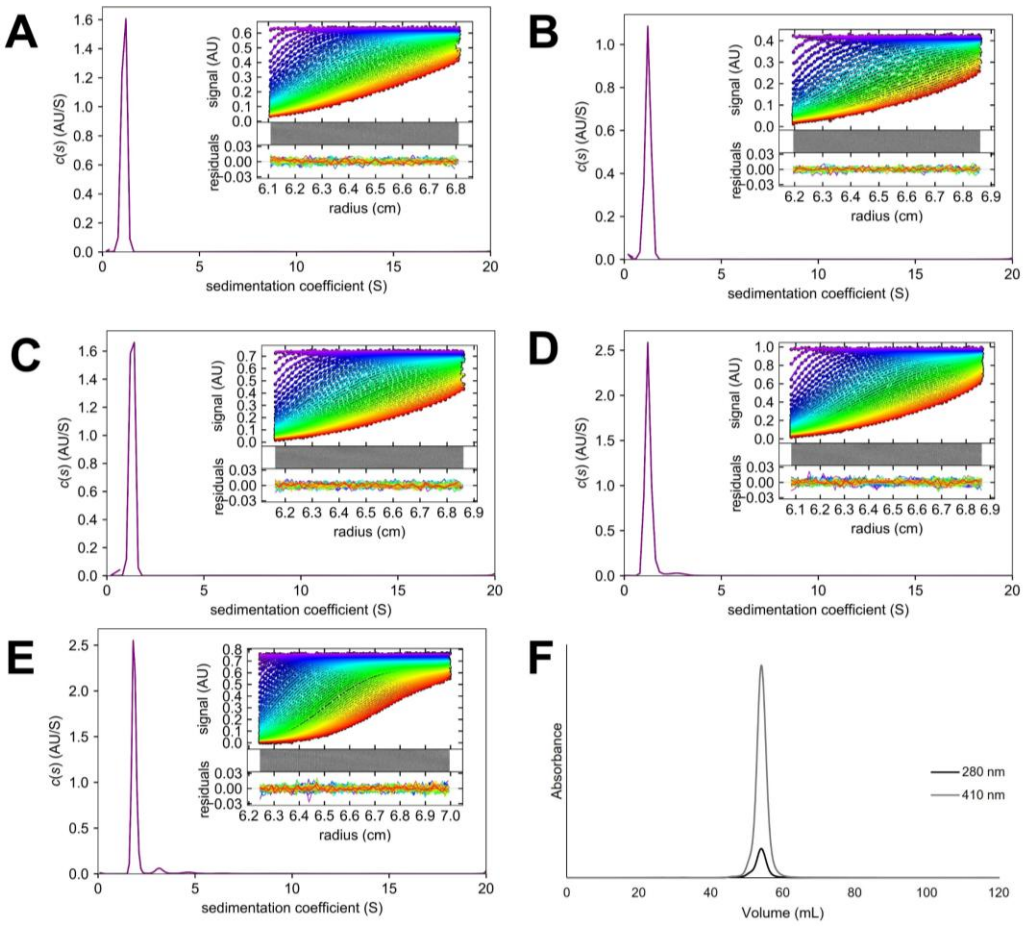


Figure 3.6: Solution state properties of PgcA monoheme domains and triheme PgcA. Sedimentation velocity (c(S)) analysis of sedimentation velocity data for PgcA monoheme domain I **(A)**, Domain II **(B)**, Domain III **(C)** (≈ 6 µM) and stoichiometric mixture of all three domains (≈ 2 µM each) **(D)**, in 20 mM HEPES, 100 mM NaCl pH 7.8 under 129,000 G at 20 °C. **(E)** c(S) distribution analysis of triheme PgcA under 129,000C G sedimentation velocity analysis in 100 mM Tris-HCl 100 mM NaCl pH 8.0, at 20 °C, at 2 µM**.** Inset top: Raw data sedimentation and fit. Inset middle: 2D difference map. Inset bottom: residual difference plot. **(F)** Size-exclusion chromatography trace for triheme PgcA in 20 mM HEPES pH 7.9, 100 mM NaCl at 4 °C, with 0.5 mL min^{-1} flowrate through a 16/600 Superdex 75 PG column.

at higher scattering intensities (Figure 3.7A). Scans in this region were averaged and compared to the median to generate appropriate scattering curves (Figure 3.7C). Guinier analysis of the scattering curves at each concentration remained constant to low Q and gave R_g values of 39 ± 0.4, 44 ± 0.3 and 43 ± 0.2 Å for 2.5, 5 and 10 mg mL^{-1} concentrations, respectively (Figure 3.7B). Taken together, these results indicate that PgcA remains monodisperse across the experimental concentration range. Therefore, the sample at 10 mg mL^{-1} was used for further analysis. Plotting the data as a function of I(q) x q^2 (Kratky) versus q yielded a scattering curve with a stretched, broad shoulder displaying maxima at q = 0.15 $Å^{-1}$, rather than the typical sharp, near-Gaussian peak at low q values associated with folded proteins (Figure 3.7B). This analysis is consistent with the hypothesis that the triheme PgcA structure contains significantly disordered regions, which cannot be explained by the sequences contained within the crystal structures and therefore must derive from the repeat motifs.

The probability distance distribution p(r) curve for the PgcA scattering profile yielded a broad curve with identifiable peaks at 21, 60, 96 and 136 Å, before reaching a D_{max} of 180 Å (Figure 3.7D). This profile provides vital information about the dynamics of PgcA in solution. The peak at 20 Å results from the interatomic

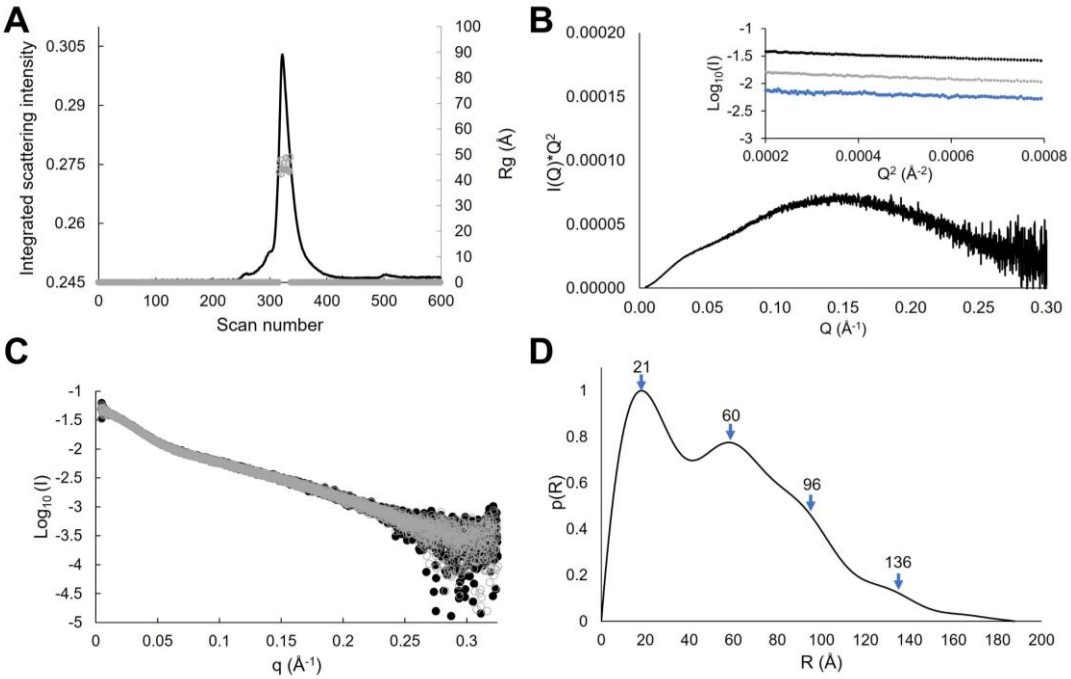


Figure 3.7: SAXS analysis of triheme PgcA. (**A**) Triheme PgcA SEC-SAXS traces injected at 10.0 mg mL^{-1} in 50 mM Tris-HCl 50 mM NaCl pH 8.0. The continues line shows the integrated scattering intensity, relative to baseline for each scan and the open circles show the approximate Guinier Rg values calculated for the corresponding scan. (**B**) SEC-SAXS Kratky plot and Guinier approximation plot for triheme PgcA at 10 (black), 5 (grey) and 2.5 (blue) mg mL^{-1}. (**C**) Log_{10} intensity plot of subtracted and merged SAXS frames for 10 mg mL^{-1} PgcA. Calculated via buffer-background subtraction, from selected scans, of the mean (black) and median (grey). (**D**) p(R) distribution model for triheme PgcA using the Primus Autognom function.

distances within each of the three domains, the peak at 60 Å likely represents the average interatomic distances between adjacent domains I and II or II and III, and most importantly the remaining two peaks likely result from the average interatomic distances between domains I and III, which exist in a range of conformations and stretch as far apart as 180 Å.

Using the modelling program MULTIFOXS[14], the experimental scattering data was fitted to a range of models generated from the crystal structures of the three cytochrome domains connected with flexible $(PT)_{14}$ linkers (Figure 3.8A & B). The minimum number of models required to fit the experimental data was two (Figure 3.8C), representing different averaged conformations of PgcA in solution. Further analysis with the EOM method[15,16] consistently revealed that most

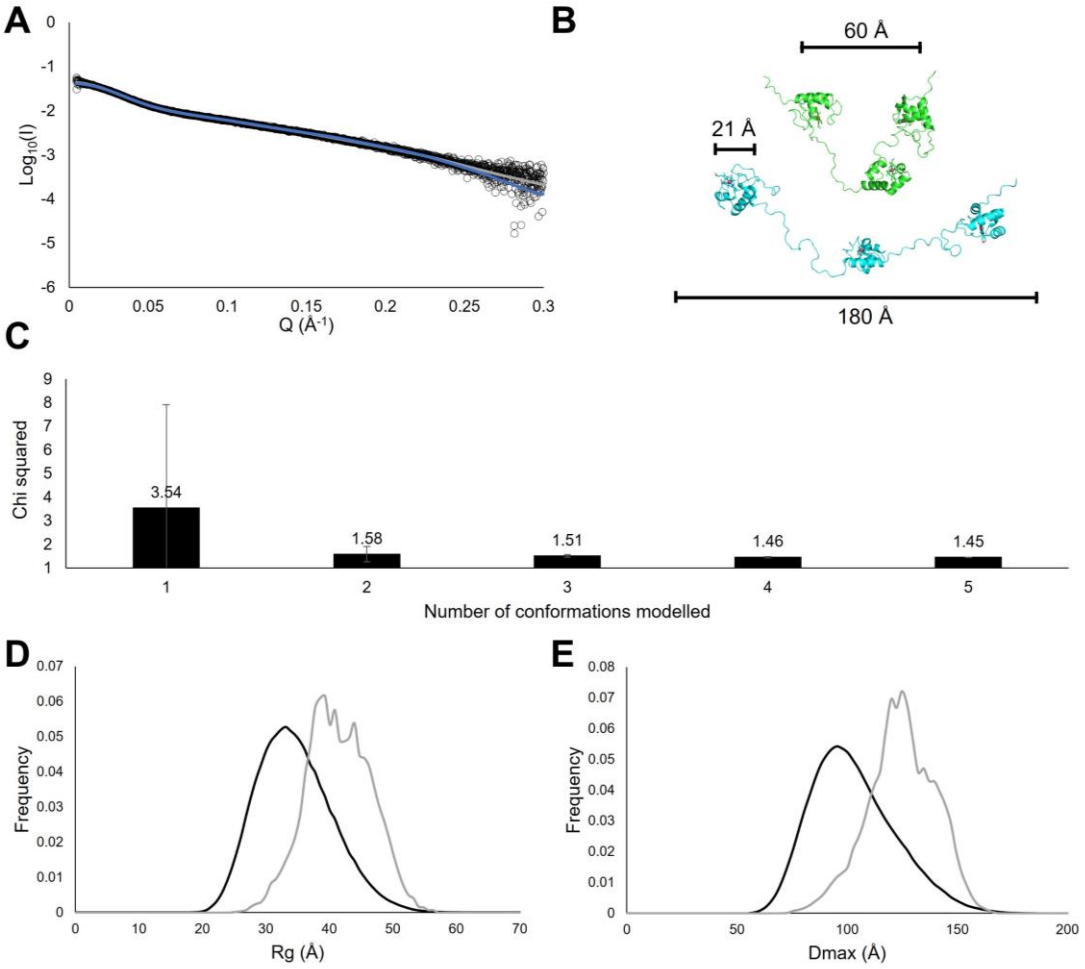


Figure 3.8: Biophysical solution state modelling of triheme PgcA. (**A**) Experimental scattering intensity against scattering angle plot (black), MultiFoXS conformation ensemble model-predicted scattering plot (grey) and EOM optimised ensemble model-predicted scattering plot (blue). (**B**) Two representative PgcA structural conformation models from the three-state model, generated by MultiFoXS fitting to experimental scattering. (**C**) MultiFoXS conformation modelling ensemble size experiment-data fit quality histogram. (**D**) and (**E**) Frequency distribution plots generated by EOM analysis of triheme PgcA SEC-SAXS scattering curve. Initial randomised ensemble shown in black, optimised ensemble with improved fit to the experimental data in grey, analysis was performed four times giving similar results, which were averaged yielding these plots.

conformations adopted by PgcA exist between R_g values of 30 and 55 Å and between D_{max} values of 70 and 160 Å (Figure 3.8D & E), suggesting that intermediate extension of the protein dominates over the rarer contracted forms and highly-extended forms.

Discussion

The transfer of electrons through chains of redox-active cofactors is a fundamental process, ubiquitous in all organisms[17]. To ensure that electron transfer is both rapid and controlled, cofactors must be packed together so the overall distance between edges is typically less than 15 Å[17]. For *c*-type cytochromes, these edge-edge distances are often <5 Å, allowing for electron transfer on a nanosecond timescale across multiheme proteins[6,13,18,19]. These heme wires can be over 120 Å in the periplasm or 180 Å across the outer membrane, while polymers of *c*-type cytochromes, such as OmcS, OmcZ and OmcE in *G. sulfurreducens*, can reach micrometres in length[6,20,21]. Those wires allow for rapid electron transfer between electron donors and terminal electron acceptors but require the assembly of multiple heme groups that are costly to synthesize.

In this chapter, the structure of a new type of multiheme cytochrome that consists of individual *c*-type cytochrome domains tethered together by $(PT)_x$ linker peptides is revealed. Structural and biophysical analysis of the three domains show them to be distinct, highly solvent exposed monoheme proteins, two of which have His-Met heme coordination and one that is His-His coordinated[22]. Domains I and II possess two acidic residues, Glu281 and Asp372, which hydrogen bond the proximal histidine of the His-Met coordinated heme. These negatively charged residues might help stabilise the Fe^{3+} state over Fe^{2+}, lowering the E_m of domains I and II.

The multimodal SAXS p(R) curve of the triheme PgcA is consistent with three freely moving, but tethered, domains. The maximum at 20 Å represents the distances between atoms located in the same domain, while the maximum at 60 Å represents the distances between atoms in different domains, at the most prevalent inter-domain distance. This explains the exceptionally high radius of gyration and implies that the shallow p(r) decrease to a D_{max} of >180 Å is derived from a small population of the most extended conformations of the chain. The D_{max} of 180 Å therefore corresponds to the greatest inter-atomic distance within the most extended conformer(s).

The novel flexible-movement based structural system described within this chapter may maximise the chance of mineral-PgcA contact, while allowing electron transfer between the domains. However since SAXS analysis of PgcA implies the domains spend a significant duration 20 Å apart or further, the rate of electron transfer is likely to be orders of magnitude slower than those of closely-

packed heme wire proteins[13]. Therefore it seems more likely PgcA probably performs a complementary function to those proteins, rather than a parallel one. For example, in instances where mineral-cell contact occurs, transport of electrons across the outer membrane will be rapid exclusively at the point of mineral contact. A large number of PgcA proteins might facilitate rapid lateral electron transfer across the cell's outer surface, allowing porin-cytochrome complexes not in direct contact with the mineral to indirectly reduce it.

Using the known structures, mass and heme number of cytochrome wires, it can be estimated that for a similar wire to span the eighteen nanometres measured for PgcA, it would require a 24-heme arrangement with a mass of 85-125 kDa. This mechanism for long range electron transfer on minimal heme chassis is ubiquitous, as genes encoding extracellular PgcA paralogues containing 3-14 hemes have been identified in the genomes of disparate bacteria, such as *Geotalea luticola, Citrifermentans bemidjiensis* and *Pelobacter selenigens*[10]. The specific roles of these genes are not yet defined, but the information provided by this chapter suggests that long-range extracellular electron transfer through tethered hemes is potentially an important respiratory strategy for environmental microorganisms.

These findings also could provide a basis for further study of tethered heme electron transport proteins. Such future experiments could seek to explain why PgcA undergoes truncation between Met252 and Pro253 when expressed recombinantly in *S. oneidensis*, although important findings possibly of relevance to this question are described in *Chapter 5*. In this context, it is also unclear why ~50 kDa PgcA can be purified from the soluble fraction at all, since lipidated proteins are usually considered insoluble in aqueous solution suggesting PgcA might not be lipidated by *S. oneidensis*[23].

Future experiments could also attempt to establish whether PgcA contains specifically three hemes tethered together rather than two or four, or even more. Additional domains could be incorporated, or a domain could be removed and the subsequent electron transfer rates measured. It is interesting to contemplate how large of a tethered heme protein can actually be constructed. Presumably a hard limit on the number of domains that is viable will be imposed either by the inability either of the cell to synthesize a protein of extreme size, or by inability of some component in the cytochrome maturation/translocation pathways, that ensuring proper assembly and translocation, in processing a massive cytochrome.

A related and interesting question also beyond the scope of this chapter is whether the Pro-Thr motifs are ~30 amino acids in length rather than 40 or 20. The results within this chapter could provide a basis for examining the effect of Pro-Thr motif truncation/expansion via SAXS which might lead to decreased or increased flexibility/movement, respectively.

Methods

Plasmid design and construction.

The PgcA gene was provided by Prof. Daniel Bond (University of Minnesota) in pBAD202PgcA and transformed into *S. oneidensis MR-1* strain LS527. Plasmid DNA was purified, and PCR mutagenesis was performed to generate first the pBAD202PgcA-strep construct using primers in Table 3.2. This linear DNA was digested with Dpn1, and re-circularised with T4 polynucleotide kinase and T4 ligase. This was transformed into *E. coli* Top10 from which plasmid DNA was extracted and purified before sequencing. The plasmid DNA was then transformed into *S. oneidensis* MR-1 by electroporation.

From the pBAD202PgcA-strep DNA, the three single domain constructs were subsequently generated in one and two rounds of PCR mutagenesis for domain III and domains I and II, respectively using primers shown in Table 3.2. The linear blunt ended DNA fragments were circularised with T4 polynucleotide kinase and ligase before heat shock transformation into *E. coli* Top10, DNA purification and sequencing, and subsequent electroporation back into *MR-1*.

Primers and respective plasmids are shown in Table 3.2, for more detailed description of cloning protocols, see Chapter 2.

Expression and purification of PgcA isoform proteins

Colonies of *S. oneidensis* MR-1 were transferred into 100 mL LB-media (30 μg mL^{-1} kanamycin) and grown aerobically overnight, before 1.5 % inoculation of 1 L LB in baffled flasks. These were grown until OD_{600} ≈0.5 was reached, 3 mM arabinose was then added after which they were incubated overnight, all growth at 30 °C 180 RPM.

Cells were harvested by centrifugation (4000 G, 20 minutes), resuspended in ~50 mL of cold 100 mM Tris-HCl, 150 mM NaCl pH 8.0 (buffer A) supplemented with DNAse, and lysed by double-pass through a French press at 1000 PSI. Cell lysate was clarified by ultracentrifugation (200,000 G, 1 hour, 4°C) and the pellet discarded.

The soluble fraction was then loaded onto a Strep-Tactin® XT 4Flow® column equilibrated with buffer A before being washed with 15 column volumes of buffer A, and eluted with 6 column volumes of buffer A supplemented to 50 mM biotin, all according to the manufacturer's guidelines.

Purified protein fractions were then concentrated using 3 kDa MWCO spin concentrators and further purified by SEC through a cytiva HiLoad 16/600 Superdex 75 pg column into 20 mM HEPES, 100 mM NaCl pH 7.8.

Purity was assessed throughout using SDS-PAGE stained with Coomassie and peroxidase linked heme stain[24], and 410-280 nm absorbance ratio. Molecular

weight of purified proteins was determined by LC-MS as described in ref[25]. Protein anticipated molecular weights calculated by inputting the amino acid sequence to ExPASy ProtParam[26] and including a mass of 615.17 Da per heme.

Small-angle X-ray scattering and analytical ultracentrifugation

Triheme PgcA in 50 mM Tris 50 mM NaCl pH 8.0 was concentrated to 10, 5 and 2.5 mg mL^{-1} and analysed on B21 at Diamond light source. Samples were loaded onto a Shodex KW-402.5 SEC column equilibrated in the same buffer and data was recorded using an Eiger 4M detector. Data were analysed using ScÅtter IV (bl1231.als.lbl.gov/scatter/) where scans with uniform radius of gyration (R_g) values were averaged, and baseline was subtracted using pre-protein peak scans. Guinier, Kratky and plots were generated with Primus[16]. The Primus distance distribution wizard Autognom function was then used to generate a p(R) plot[16]. The 10 mg mL^{-1} scattering data and triheme PgcA sequence was uploaded to the MultiFoxS and EOM servers and modelled with default parameters using the crystal structures specified as rigid bodies and the $(PT)_X$ sequences as flexible.

Sedimentation velocity samples were diluted to ~2 μM in 100 mM Tris-HCl, 100 mM NaCl, pH 8. Samples were centrifuged at 129,000 G, 20 °C, and migration through the cell monitored at 408 nm. Scans were analysed with default parameters using the software program Sedfit in c(S) distribution mode[27] which utilises the Svedberg equation (equation 3.1)[28].

Equation 3.1 $$s = \frac{u}{\omega^2 r} = \frac{M(1-\tilde{\nu}\rho)}{N_A f}$$

Where s represents the sedimentation coefficient, u represents the velocity of the molecule, ω represents the angular velocity of the centrifugation cell, r represents the radius of the cell, M represents the mass of the particle, ν̃ represents the partial specific volume of the particle, ρ represents the solvent density, N_A represents Avogadro's number and f represents the frictional coefficient of the particle.

Crystallisation, and structure determination of PgcA monoheme domains

Crystals of heme-domain I were obtained by sitting-drop vapour diffusion by depositing a droplet containing 0.3 μL 10 mg mL^{-1} protein in 20 mM HEPES pH 7.8, 100 mM NaCl, onto a 0.28 μL drop of reservoir solution containing 2.5 M Na malonate pH 7.0 and 0.02 μL 30 % w/v PEG 2000 MME at 16 °C. Crystals appeared after 3-4 months and were cryoprotected by transferring to 20 % ethylene glycol 2.5 M Na malonate pH 7.0 before vitrification.

Crystals of heme-domain II were obtained by sitting-drop vapour diffusion using a 1:1 mixture of 10 mg mL^{-1} protein in 20 mM HEPES pH 7.8, 100 mM NaCl and 2.0 M ammonium sulfate, 5% v/v isopropanol, 1.8% w/v PEG 4000 at 16 °C with a total drop volume of 0.6 μL. Crystals appeared after 2-3 days and were

cryoprotected by transferring to 20 % ethylene glycol 2.0 M ammonium sulfate, 5% isopropanol before vitrification.

Crystals of heme-domain III were obtained by sitting-drop vapour diffusion using a 1:1 mixture of 20 mg mL^{-1} protein in 20 mM HEPES pH 7.8, 100 mM NaCl and 2.3 M Na malonate pH 7.0, 1.8 % Jeffamine ED-2001 at 16 °C with a total drop volume of 0.6 µL. Crystals appeared after 1-2 weeks and were cryoprotected by transferring to 25 % glycerol, 2.4 M Na malonate pH 7 before vitrification.

Datasets were collected on single frozen crystals at Diamond beamline I04 and I24 with X-rays of wavelength 0.9537Å. Datasets were indexed and scaled automatically with XIA2[29] and merged with AIMLESS[30]. Initial models were generated by molecular replacement with AF2[31] models in PHASER[32] or MOLREP[33], these models were refined manually in COOT[34] and using PHENIX.REFINE[35] and REFMAC5[36] to obtain final coordinates. Final coordinates were deposited in the Protein Data Bank with PDB IDs 8QJ6 (Domain I), 8QJG (Domain II) and 8QK0 (Domain III). Data collection and model refinement statistics are listed in Table 3.1.

Chapter 3: References

1. White, G. F., Edwards, M. J., Gomez-Perez, L., Richardson, D. J., Butt, J. N. & Clarke, T. A. Mechanisms of Bacterial Extracellular Electron Exchange. *Advances in Microbial Physiology* **68**, 87–138 (2016).

2. Shi, L., Dong, H., Reguera, G., Beyenal, H., Lu, A., Liu, J., Yu, H. Q. & Fredrickson, J. K. Extracellular electron transfer mechanisms between microorganisms and minerals. *Nature Reviews Microbiology 2016 14:10* **14**, 651–662 (2016).

3. Gralnick, J. A. & Bond, D. R. Electron Transfer Beyond the Outer Membrane: Putting Electrons to Rest. *Annual review of microbiology* **77**, 517–539 (2023).

4. Edwards, M. J., White, G. F., Butt, J. N., Richardson, D. J. & Clarke, T. A. The Crystal Structure of a Biological Insulated Transmembrane Molecular Wire. *Cell* **181**, 665-673.e10 (2020).

5. Clarke, T. A. Plugging into bacterial nanowires: a comparison of model electrogenic organisms. *Current Opinion in Microbiology* **66**, 56–62 (2022).

6. Wang, F., Gu, Y., O'Brien, J. P., Yi, S. M., Yalcin, S. E., Srikanth, V., Shen, C., Vu, D., Ing, N. L., Hochbaum, A. I., Egelman, E. H. & Malvankar, N. S. Structure of Microbial Nanowires Reveals Stacked Hemes that Transport Electrons over Micrometers. *Cell* **177**, 361-369.e10 (2019).

7. Smith, J. A., Tremblay, P. L., Shrestha, P. M., Snoeyenbos-West, O. L., Franks, A. E., Nevin, K. P. & Lovley, D. R. Going Wireless: Fe(III) Oxide Reduction without Pili by Geobacter sulfurreducens Strain JS-1. *Applied and Environmental Microbiology* **80**, 4331 (2014).

8. Tremblay, P. L., Summers, Z. M., Glaven, R. H., Nevin, K. P., Zengler, K., Barrett, C. L., Qiu, Y., Palsson, B. O. & Lovley, D. R. A c-type cytochrome and a transcriptional regulator responsible for enhanced extracellular electron transfer in Geobacter sulfurreducens revealed by adaptive evolution. *Environmental Microbiology* **13**, 13–23 (2011).

9. Zacharoff, L. A., Morrone, D. J. & Bond, D. R. Geobacter sulfurreducens extracellular multiheme cytochrome PgcA facilitates respiration to Fe(III) oxides but not electrodes. *Frontiers in Microbiology* **8**, (2017).

10. Fernandes, T. M., Silva, M. A., Morgado, L. & Salgueiro, C. A. Hemes on a string: insights on the functional mechanisms of PgcA from Geobacter sulfurreducens. *Journal of Biological Chemistry* **299**, 105167 (2023).

11. Yang, F., Bogdanov, B., Strittmatter, E. F., Vilkov, A. N., Gritsenko, M., Shi, L., Elias, D. A., Ni, S., Romine, M., Paša-Tolić, L., Lipton, M. S. & Smith, R. D. Characterization of purified c-type heme-containing peptides and identification of c-type heme-attachment sites in Shewanella oneidenis cytochromes using mass spectrometry. *Journal of Proteome Research* **4**, 846–854 (2005).

12. Lockwood, C. W. J., van Wonderen, J. H., Edwards, M. J., Piper, S. E. H., White, G. F., Newton-Payne, S., Richardson, D. J., Clarke, T. A. & Butt, J. N. Membrane-spanning electron transfer proteins from electrogenic bacteria: Production and investigation. *Methods in Enzymology* **613**, 257–275 (2018).

13. van Wonderen, J. H., Adamczyk, K., Wu, X., Jiang, X., Piper, S. E. H., Hall, C. R., Edwards, M. J., Clarke, T. A., Zhang, H., Jeuken, L. J. C., Sazanovich, I. V., Towrie, M., Blumberger, J., Meech, S. R. & Butt, J. N. Nanosecond heme-to-heme electron transfer rates in a multiheme cytochrome nanowire reported by a spectrally unique His/Met-ligated heme. *Proceedings of the National Academy of Sciences of the United States of America* **118**, e2107939118 (2021).

14. Schneidman-Duhovny, D., Hammel, M., Tainer, J. A. & Sali, A. FoXS, FoXSDock and MultiFoXS: Single-state and multi-state structural modeling of proteins and their complexes based on SAXS profiles. *Nucleic Acids Research* **44**, W424–W429 (2016).

15. Tria, G., Mertens, H. D. T., Kachala, M. & Svergun, D. I. Advanced ensemble modelling of flexible macromolecules using X-ray solution scattering. *IUCrJ* **2**, 207–217 (2015).

16. Manalastas-Cantos, K., Konarev, P. V., Hajizadeh, N. R., Kikhney, A. G., Petoukhov, M. V., Molodenskiy, D. S., Panjkovich, A., Mertens, H. D. T., Gruzinov, A., Borges, C., Jeffries, C. M., Sverguna, D. I. & Frankea, D. ATSAS 3.0: expanded functionality and new tools for small-angle scattering data analysis. *Journal of applied crystallography* **54**, 343–355 (2021).

17. Moser, C. C., Keske, J. M., Warncke, K., Farid, R. S. & Dutton, P. L. Nature of biological electron transfer. *Nature 1992 355:6363* **355**, 796–802 (1992).

18. Breuer, M., Rosso, K. M., Blumberger, J. & Butt, J. N. Multi-haem cytochromes in Shewanella oneidensis MR-1: structures, functions and opportunities. *Journal of The Royal Society Interface* **12**, (2015).

19. Clarke, T. A., Edwards, M. J., Gates, A. J., Hall, A., White, G. F., Bradley, J., Reardon, C. L., Shi, L., Beliaev, A. S., Marshall, M. J., Wang, Z., Watmough, N. J., Fredrickson, J. K., Zachara, J. M., Butt, J. N. & Richardson, D. J. Structure of a bacterial cell surface decaheme electron conduit. *Proceedings of the National Academy of Sciences of the United States of America* **108**, 9384–9389 (2011).

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

21. Wang, F., Mustafa, K., Suciu, V., Joshi, K., Chan, C. H., Choi, S., Su, Z., Si, D., Hochbaum, A. I., Egelman, E. H. & Bond, D. R. Cryo-EM structure of an extracellular Geobacter OmcE cytochrome filament reveals tetrahaem packing. *Nature Microbiology 2022 7:8* **7**, 1291–1300 (2022).

22. Reedy, C. J., Elvekrog, M. M. & Gibney, B. R. Development of a heme protein structure–electrochemical function database. *Nucleic Acids Research* **36**, D307–D313 (2008).

23. van Wonderen, J. H., Crack, J. C., Edwards, M. J., Clarke, T. A., Saalbach, G., Martins, C. & Butt, J. N. Liquid-chromatography mass spectrometry describes post-translational modification of Shewanella outer membrane proteins. *Biochimica et biophysica acta. Biomembranes* **1866**, (2024).

24. Thomas, P. E., Ryan, D. & Levin, W. An improved staining procedure for the detection of the peroxidase activity of cytochrome P-450 on sodium dodecyl sulfate polyacrylamide gels. *Analytical Biochemistry* **75**, 168–176 (1976).

25. van Wonderen, J. H., Li, D., Piper, S. E. H., Lau, C. Y., Jenner, L. P., Hall, C. R., Clarke, T. A., Watmough, N. J. & Butt, J. N. Photosensitised Multiheme Cytochromes as Light-Driven Molecular Wires and Resistors. *ChemBioChem* **19**, 2206–2215 (2018).

26. Wilkins, M. R., Gasteiger, E., Bairoch, A., Sanchez, J. C., Williams, K. L., Appel, R. D. & Hochstrasser, D. F. Protein identification and analysis tools in the ExPASy server. *Methods in molecular biology (Clifton, N.J.)* **112**, 531–552 (1999).

27. Schuck, P. Size-distribution analysis of macromolecules by sedimentation velocity ultracentrifugation and lamm equation modeling. *Biophysical journal* **78**, 1606–1619 (2000).

28. Lebowitz, J., Lewis, M. S. & Schuck, P. Modern analytical ultracentrifugation in protein science: A tutorial review. *Protein Science* **11**, 2067 (2002).

29. Winter, G. Xia2: An expert system for macromolecular crystallography data reduction. *Journal of Applied Crystallography* **43**, 186–190 (2010).

30. Evans, P. R. & Murshudov, G. N. How good are my data and what is the resolution? *Acta crystallographica. Section D, Biological crystallography* **69**, 1204–1214 (2013).

31. Jumper, J., Evans, R., Pritzel, A., Green, T., Figurnov, M., Ronneberger, O., Tunyasuvunakool, K., Bates, R., Žídek, A., Potapenko, A., Bridgland, A., Meyer, C., Kohl, S. A. A., Ballard, A. J., Cowie, A., Romera-Paredes, B., Nikolov, S., Jain, R., Adler, J. *et al.*

Highly accurate protein structure prediction with AlphaFold. *Nature 2021 596:7873* **596**, 583–589 (2021).

32. McCoy, A. J., Grosse-Kunstleve, R. W., Adams, P. D., Winn, M. D., Storoni, L. C. & Read, R. J. Phaser crystallographic software. *Journal of applied crystallography* **40**, 658–674 (2007).

33. Vagin, A. & Teplyakov, A. MOLREP: an Automated Program for Molecular Replacement. *Journal of Applied Crystallography* **30**, 1022–1025 (1997).

34. Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. Features and development of Coot. *Acta crystallographica. Section D, Biological crystallography* **66**, 486–501 (2010).

35. Liebschner, D., Afonine, P. V., Baker, M. L., Bunkoczi, G., Chen, V. B., Croll, T. I., Hintze, B., Hung, L. W., Jain, S., McCoy, A. J., Moriarty, N. W., Oeffner, R. D., Poon, B. K., Prisant, M. G., Read, R. J., Richardson, J. S., Richardson, D. C., Sammito, M. D., Sobolev, O. V. *et al*. Macromolecular structure determination using X-rays, neutrons and electrons: recent developments in Phenix. *Acta Crystallographica. Section D, Structural Biology* **75**, 861 (2019).

36. Murshudov, G. N., Skubák, P., Lebedev, A. A., Pannu, N. S., Steiner, R. A., Nicholls, R. A., Winn, M. D., Long, F. & Vagin, A. A. REFMAC5 for the refinement of macromolecular crystal structures. *Acta crystallographica. Section D, Biological crystallography* **67**, 355–367 (2011).

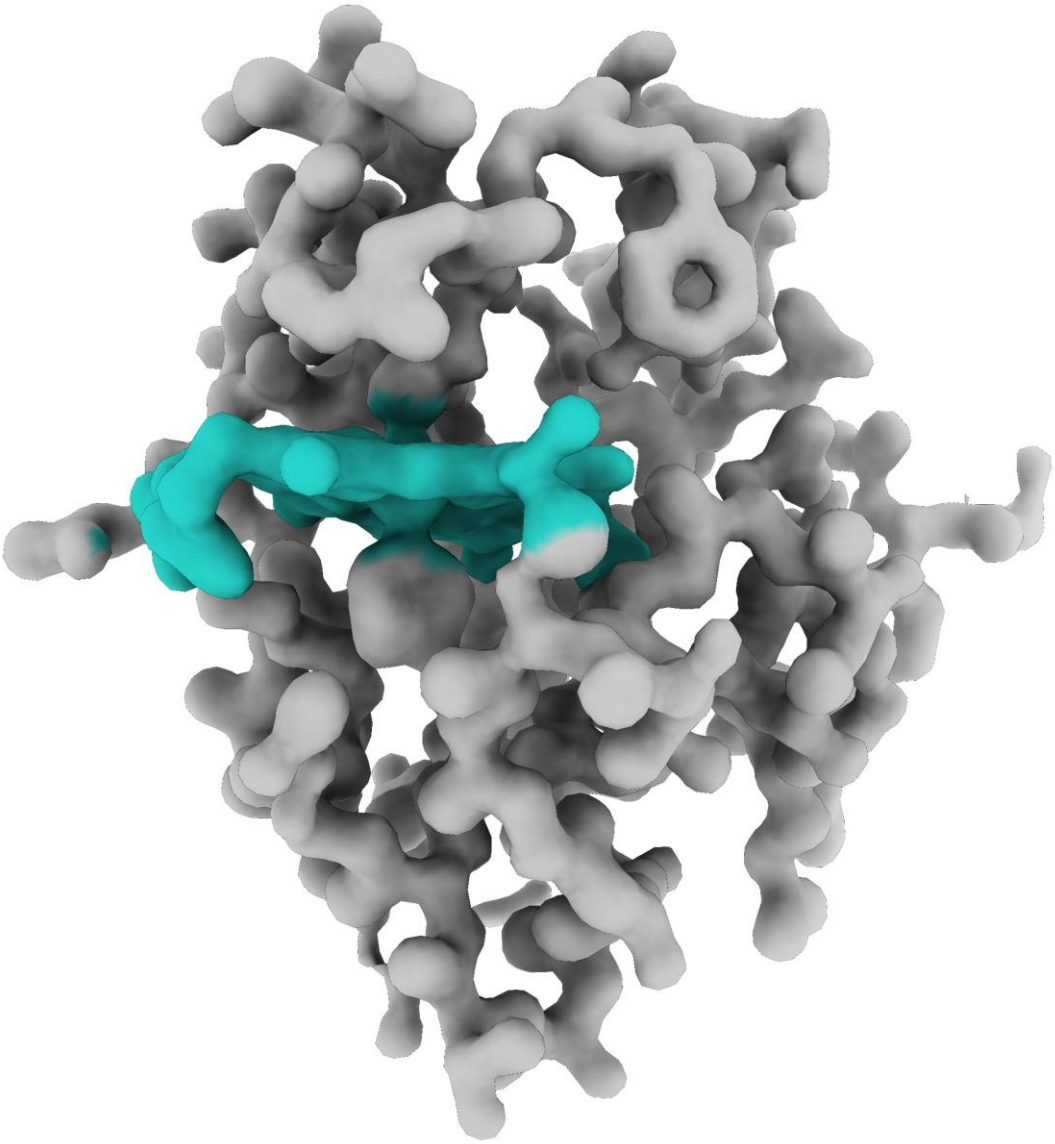
Chapter 3: Appendix

Table 3.1: Data collection and refinement statistics for PgcA monoheme domain crystals.

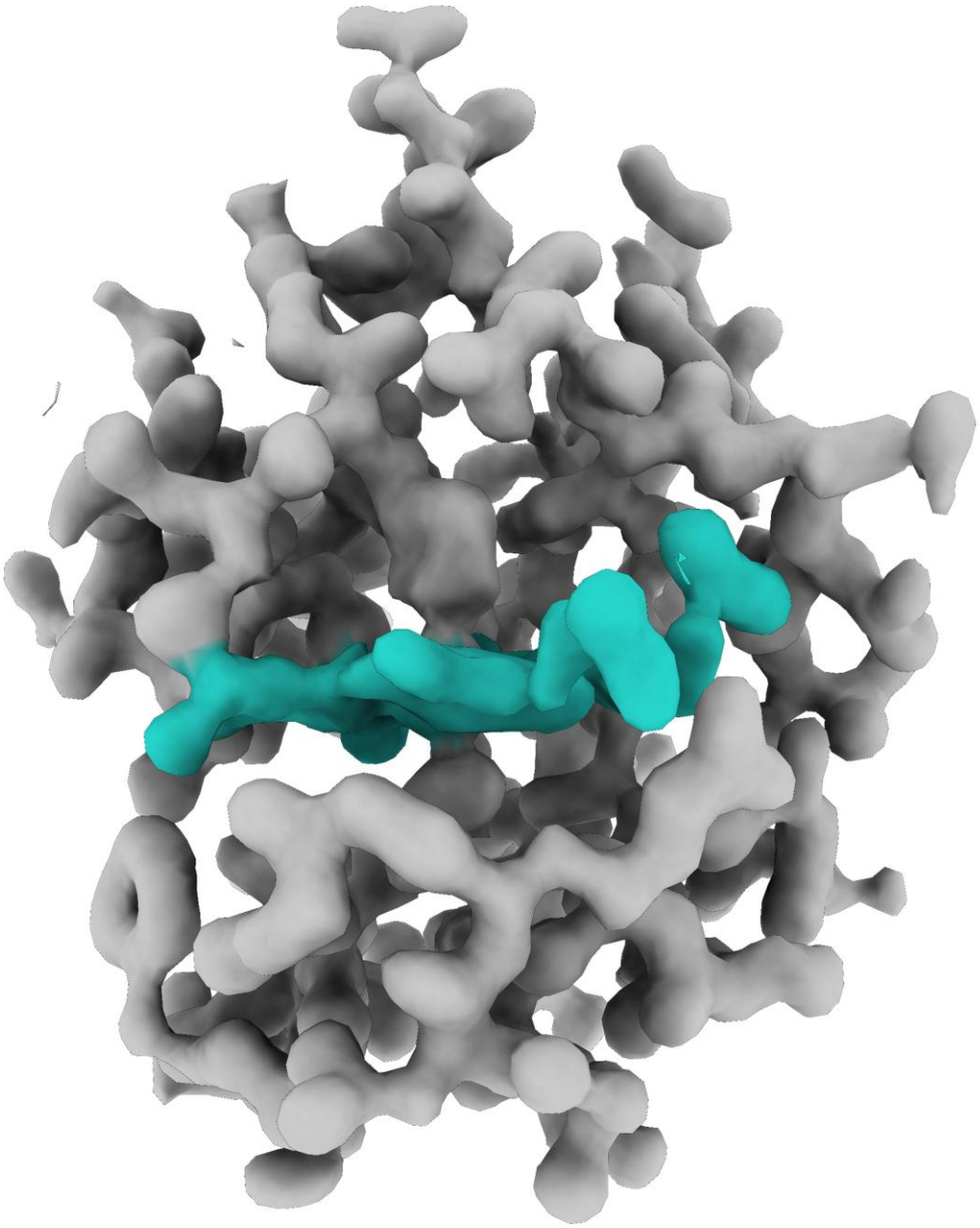
	Domain I (PDB 8QJ6)	Domain II (PDB 8QJG)	Domain III (PDB 8QK0)
Data collection			
Space group	I 4	C 1 2 1	C 1 2 1
Cell dimensions			
a, b, c (Å)	68.43, 68.43, 27.68	95.10, 74.10, 56.02	75.56, 55.03, 37.46
α, β, γ (°)	90.00, 90.00, 90.00	90.00, 103.92, 90.00	90.00, 99.30, 90.00
Resolution (Å)	24.19-1.55 (1.61-1.55)	42.96-1.80 (1.86-1.80)	44.28-1.9 (1.97-1.9)
R_{pim} (%) $CC_{1/2}$ (%)	2.0 (15.5) 99.9 (97.7)	9.5 (26.5) 97.1 (84.5)	5.7 (26.8) 99.6 (85.9)
$I / \sigma I$	24.64 (2.90)	7.93 (1.90)	9.71 (1.64)
Completeness (%)	97.12 (77.90)	99.43 (95.91)	99.57 (98.17)
Multiplicity	12.9 (8.9)	3.4 (2.6)	6.8 (5.5)
Refinement			
Resolution (Å)	1.55	1.80	1.90
No. reflections	9226	34836	11995
R_{work} / R_{free}	0.155/0.189	0.183/0.231	0.221/0.263
No. atoms			
Protein	439	2690	1028
Ligand/ion	44	327	96
Water	69	538	179
B-factors			
Protein	14.86	13.83	18.12
Ligand/ion	12.69	14.40	19.46
Water	30.89	25.04	27.55
R.m.s. deviations			
Bond lengths (Å)	0.009	0.010	0.008
Bond angles (°)	1.01	1.13	1.10

Table 3.2: Primers and plasmids used in Chapter 3.

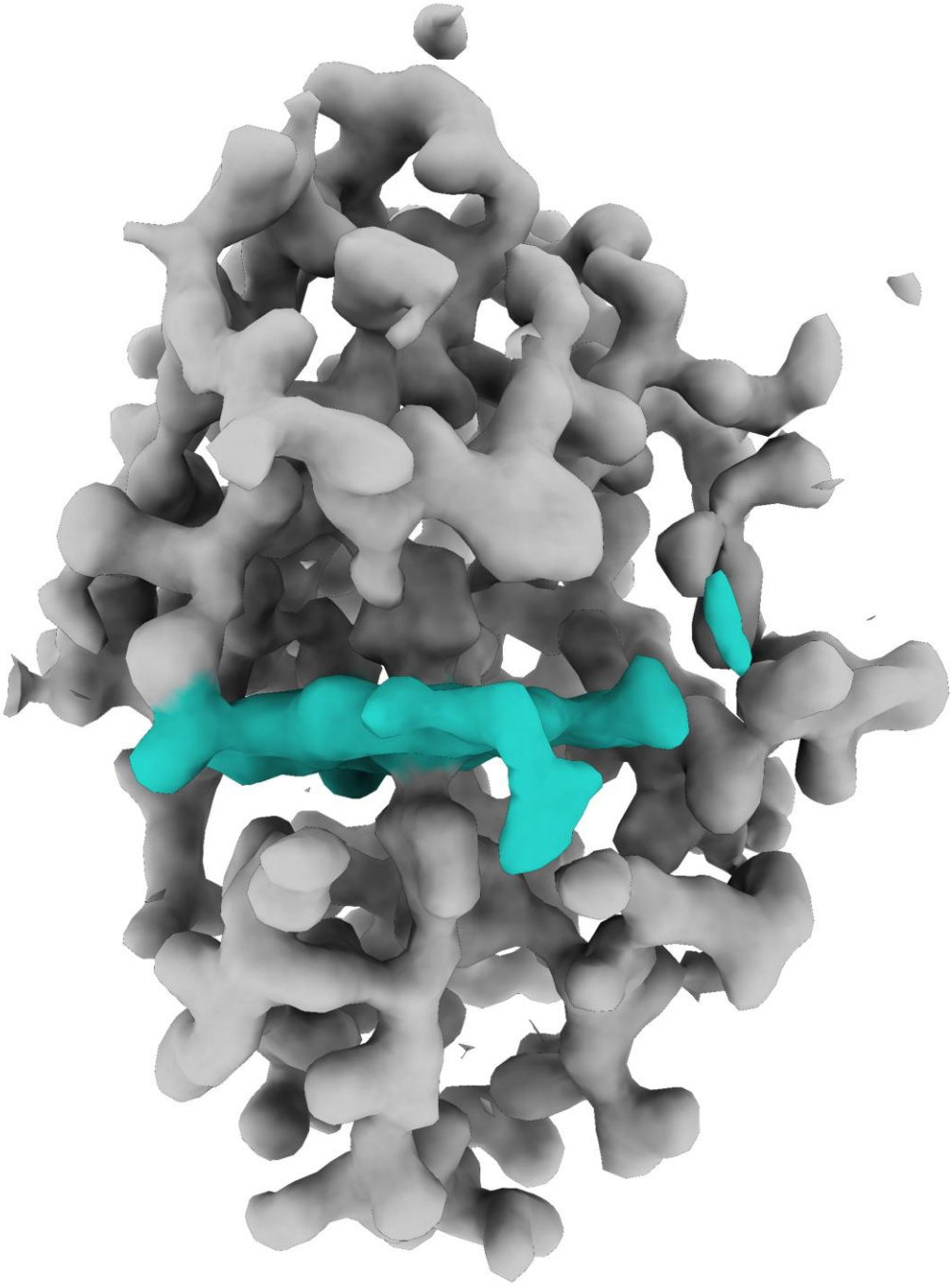
Plasmid	Product	Primers used in construction (5' to 3'):
pPgcA-strep	Triheme PgcA	Forward: GTTCGAAAAATGAGTTTAAACGGTCTCCCA Reverse: TGCGGGTGGCTCCAACCACTACCGTTTGCGTTG
pPgcA-DI	Monoheme domain I	Pair 1 Forward: GGTAGTGGTTGGAGCCACCC Pair 1 Reverse: AGCCACTGCCAAGACGTCAA Pair 2 Forward: GCCAACACAGGCTTGGCCGTCGCTGCTGATGGTGACGGACAAGGATTGTACGC Pair 2 Reverse: TAATAACGCTAGAGTGATCAAATTGAGTTTAAATTTCATCATGGGATGTATATCTC
pPgcA-DII	Monoheme domain II	Pair 1 Forward: GGTAGTGGTTGGAGCCACCC Pair 1 Reverse: GACAGCCGATGCCAGCGAGG Pair 2 Forward: GCCAACACAGGCTTGGCCGTCGCTGCTGATGGTAACGGGTCGGCACTCTACGG Pair 2 Reverse: TAATAACGCTAGAGTGATCAAATTGAGTTTAAATTTCATCATGGGATGTATATCTC
pPgcA-DIII	Monoheme domain III	Forward: GCCAACACAGGCTTGGCCGTCGCTGCTGATGGTTAGTCGACCCGGGCAAGACCGT Reverse: TAATAACGCTAGAGTGATCAAATTGAGTTTAAATTTCATCATGGGATGTATATCTCCTTCTTAGG



Appendix figure 1: PgcA monoheme domain I electron density map. Double difference electron density map from the PgcA monoheme domain I crystal structure (1.55 Å resolution), contoured at 1.8 σ, density arising from the *c*-type heme represented in cyan.



Appendix figure 2: PgcA monoheme domain II electron density map. Double difference electron density map from the PgcA monoheme domain II crystal structure (1.8 Å resolution), contoured at 1.8 σ, density arising from the *c*-type heme represented in cyan.



Appendix figure 3: PgcA monoheme domain III electron density map. Double difference electron density map from the PgcA monoheme domain III crystal structure (1.9 Å resolution), contoured at 1.8 σ, density arising from the *c*-type heme represented in cyan.

Chapter 4: Functional characterisation of PgcA heme-containing isoforms

Work presented in this chapter includes important contributions made by Dr Dimitri A. Svistunenko and Dr Marcus J. Edwards (EPR spectroscopy), Dr Joshua A. J. Burton (optically monitored PFV), Dr Tomás M. Fernandes, Dr Leonor Morgado and Professor Carlos A. Salgueiro (mineral oxide binding pull-down assays and iron oxide reduction assays).

Introduction

In *Chapter 3*, in-depth characterisation of the structure of PgcA and its constituent domains was described. However, a complete understanding of this protein and its role in EET cannot be acquired without exploration of its electronic properties and its interactions with substrates.

Existing articles in the literature so far have established the electronic absorbance spectra of the triheme PgcA isoform[1] and of the individual monoheme domains[2] which were consistent with one another. These spectra displayed features typical of *c*-type cytochromes although they were unable to unambiguously identify the presence of His-Met coordinated hemes which the crystal structures showed clearly[1]. Nonetheless this technique has played a crucial role in the study of multiheme cytochromes, allowing for the recent measurement of heme to heme electron transfer rates[3,4]. The unique structure adopted by PgcA established in *Chapter 3* is without precedent and thus it remains unclear what electron transfer rates it might support. More in depth analysis of PgcA absorbance spectra thus might serve as an important first step in addressing this issue.

In the literature also, the E_m values of the individual domains have been measured by potentiometric titration[2]. This property exerts significant control presumably over not just the direction of electron flow within the PgcA tethered heme wire, but also over the rate of intramolecular electron transfer and rate of substrate reduction too. Therefore accurately establishing the potential of the hemes is also critical.

Finally, previous work has sought to define the way PgcA interacts with its substrates. In this context, the protein has been shown to stick to iron oxide particles of particular composition under certain conditions[1]. This was suggested to be a product of the unique Pro-Thr repeat motifs in a way analogous to ice binding proteins on the basis of phage display experiments which suggested Pro-Ser/Thr motifs confer affinity for hematite[1,5,6]. This was supported by further binding experiments performed with the monoheme domains which

were found to exhibit poor binding to iron oxides, implicating the Pro-Thr repeats specifically[2]. Finally, the monoheme domains were shown to reduce various solid phase substrates with varying levels of effectiveness, although in the particular case of akageneite only monoheme domain III could reduce it.

In this chapter, the electronic properties of PgcA and its interactions with potential substrates were explored with the goal of defining the pathway electrons take through the protein, determining the importance of the Pro-Thr motifs in mineral binding and establishing a basis for future novel experiments which could quantify the rate of electron transfer achieved by this novel structural arrangement.

Results

Electronic absorbance spectroscopy of PgcA heme containing isoforms

The electronic absorbance spectra (Figure 4.1) of all as-purified PgcA samples were consistent with oxidised *c*-type cytochromes and with previously reported absorbance spectra[1,2]. Triheme PgcA (Figure 4.1 D) displayed a Soret peak at 408 nm and a broad spectral feature around 525 nm. Addition of sodium dithionite caused the Soret peak intensity to increase and shift to 419 nm, while the 525 nm peak separated into α- and β- bands at 530 and 550 nm, respectively, consistent with reduced *c*-type hemes (Figure 1D).

Importantly, reduced minus oxidised spectra (Figure 4.1 E-G) of the three monoheme domain proteins reveal that domains I and II share an isosbestic point at 434 nm, and 543 nm at which domain III displays a non-zero difference extinction coefficient.

Potentiometric characterisation of PgcA heme containing isoforms.

The E_m values of the PgcA monoheme domains were previously measured by spectropotentiometric titration and were between -48 and -106 mV versus the Standard Hydrogen Electrode (SHE)[2]. In this chapter, protein film voltammetry (PFV) was used to probe the redox properties of PgcA, its domains and their pH sensitivity. The adsorbed monoheme domain proteins yielded single reversible redox peaks on the indium tin-oxide (ITO) electrodes across a pH range of 5-8 consistent with previous measurements[2] (Figure 4.2). Domains I and II specifically displayed E_m with much greater pH dependence, while domain III only shifted by ~10 mV (Table 4.1 and Chapter 4 appendix figures 1-3), which is surprising considering the higher solvent accessibility of the domain III heme revealed in the crystal structures. The explanation may lay in differences in amino acid side chains within the environment of the hemes, possibly the role played by Glu281 and Asp372.

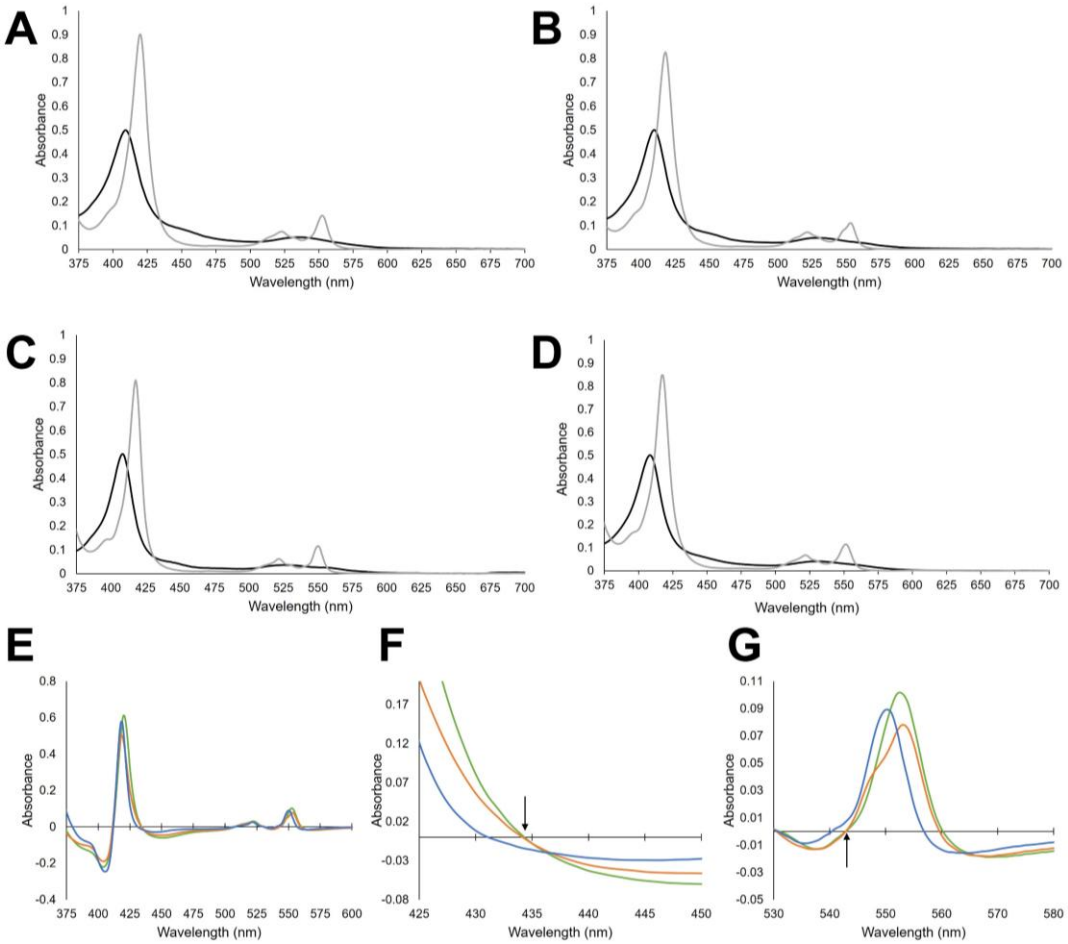


Figure 4.1: Electronic absorbance spectra of PgcA heme isoforms. Electronic absorbance spectra of PgcA monoheme domains I **(A)**, II **(B)**, and III **(C)** and triheme PgcA **(D)**. Black line – samples as purified, in 50 mM Tris, 50 mM NaCl, pH 8.0. Grey line – after addition of 3 mM sodium dithionite. **(E)** Reduced minus oxidised spectra for monoheme domains I (green), II (orange) and III (blue) with zoomed in representation showing isosbestic points shared by domains I and II but not III **(F)** and **(G)**.

Table 4.1 : Em values (mV vs SHE) and standard error, of PgcA monoheme domains determined by PFV.

pH	E_m of domain I	E_m of domain II	E_m of domain III
8	-49 (±1)	-85 (±1)	-88 (±1)
7	-28 (±1)	-60 (±1)	-94 (±1)
6	-36 (±2)	-32 (±1)	-98 (±3)
5	-6 (±1)	-33 (±1)	-82 (±2)

Adsorption of triheme PgcA onto optically transparent ITO electrodes yielded UV-visible absorbance spectra with a shift in absorption maxima from 408 to 419 nm as the working electrode potential was lowered (Figure 4.2 E), consistent with the full reduction of PgcA on the electrode. Cyclic voltammetry of triheme PgcA displayed reversible reduction between -200 and +100 mV of approximately equal area for reductive and oxidative scans (Figure 4.2D), and fitting of this data to three equal contributions from individual *n*=1 centres suggested three E_m values of approximately -108, -76 and -35 mV at pH 8.0, close to the respective individual E_m values for the individual domains (Figure 4.2F).

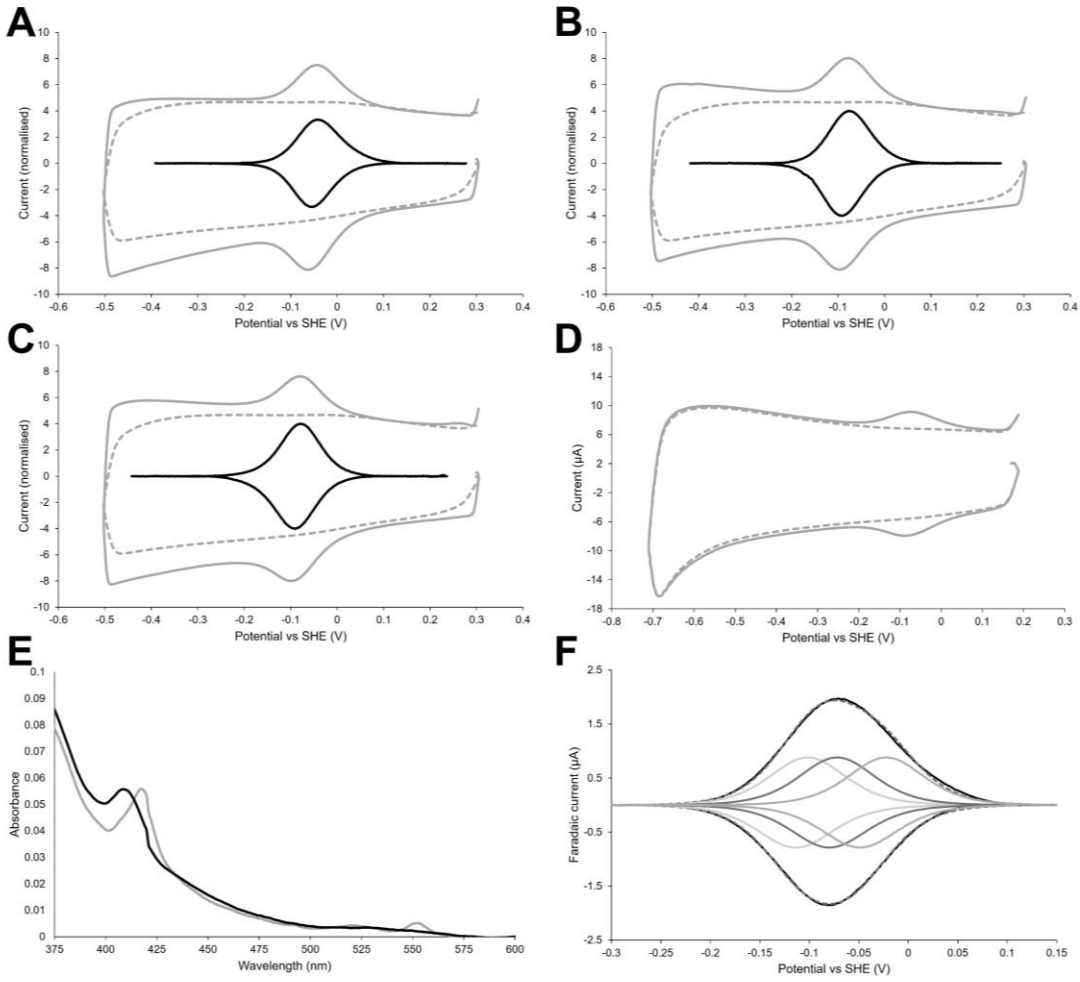


Figure 4.2: Potentiometric properties of PgcA heme isoforms. Representative 0.01 V s^{-1} protein film voltammograms of samples adsorbed to ITO electrodes in 50 mM Tris-HCl, 50 mM NaCl pH 8.0 (grey continuous), untreated electrode (grey dotted) and protein derived Faradaic currents (black), shown for PgcA monoheme domain I (**A**), domain II (**B**) and domain III (**C**) and triheme PgcA **(D**). **(E)** Electronic absorbance spectra of triheme PgcA protein film adsorbed onto ITO electrode poised at -0.33 (grey) and +0.2 (black) V vs SHE in 50 mM Tris-HCl, 50 mM NaCl pH 8.0. **(F)** Faradaic current of triheme PgcA calculated by baseline subtraction of panel D (black continuous), modelled fit of three n=1 redox centres (3 grey continuous lines) and the summed fit of the modelled n=1 centres (grey dotted). and the summed fit of the modelled n=1 centres (grey dotted).

Electron paramagnetic resonance spectroscopy of PgcA isoforms

The EPR spectra of the triheme PgcA protein and its individual isolated domains, I to III were measured (Figure 4.3). Each individual domain had a clear low spin (LS), S = ½ EPR signature and, when normalised, represent equal contributions to the triheme PgcA spectrum. The domain III spectrum displayed a small gmax, rhombic type LS signal (g_{max} 3.08) typical of bis-histidine coordinated hemes with near parallel ligand orientation, in agreement with its crystal structure. From domain III to domain I to domain II, the spectra have increasingly anisotropic signatures where the gmax value becomes more intense than the other two *g*-values[7] (g_{max} = 3.23 and 3.37 for domain I and II, respectively). A highly anisotropic or axial LS (HALS) type 'large g_{max}' signal (*g* > 3.2) can occur when two histidine heme ligands are perpendicular to each other but is also typical of histidine-methionine heme coordination where ligand orientation cannot be resolved[8]. Importantly, the triheme PgcA spectrum displays a feature between g = 3.07 and 3.08 which could be well explained by the individual contributions of the respective g_{max} features of the individual monoheme domains. This suggests the local environment of the hemes within each domain is largely unchanged by its being present in the monoheme isoforms, since EPR spectroscopy is sensitive to small changes in the environment of spin centres. This is in agreement with the structural model for PgcA presented in *Chapter 3* and also provides another means by which inter-domain electron transfer between specific domains might be probed, since hemes become EPR silent upon their reduction.

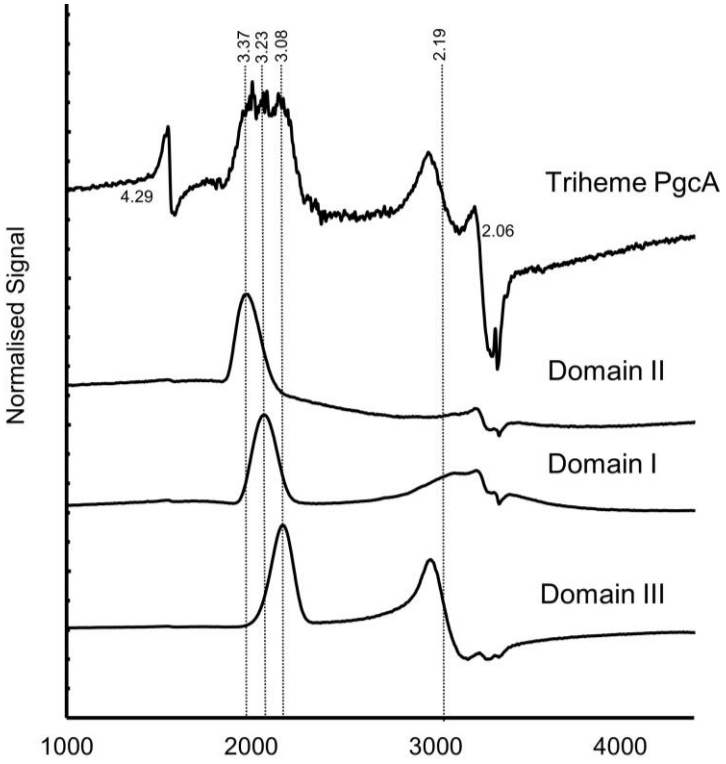


Figure 4.3: EPR spectroscopic properties of PgcA heme isoforms. Low temperature EPR spectra of triheme PgcA and its monoheme domains. The *g*-values of individual EPR lines are indicated by vertical lines. The low field g1 components of the LS ferric heme in the three domains perfectly fit to their contributions in the triheme PgcA spectrum**.** The signals at g=4.29, g=2.06 and g=2.00 are likely from adventitious non-heme Fe^{3+}, Cu^{2+}, and small free radicals respectively.

Determining the affinity of PgcA heme isoforms for iron oxide minerals

The affinity of proteins for insoluble iron(III) oxides can be measured using pull-down assays where proteins are incubated in a suspension of iron(III) oxides before centrifugation. The concentration of protein remaining in solution is measured spectroscopically, allowing determination of the binding extent. Using this method, the interaction of PgcA and monoheme domains with the iron(III) oxide akageneite was measured. Here it was observed that triheme PgcA and domain III bind to akageneite with similar affinity (Figure 4.4), while neither domains I nor II bind at all. These results were consistent with PgcA interacting with minerals through domain III only, suggesting the flexible $(PT)_x$ linkers which were previously proposed[1,2] to increase the affinity of PgcA for iron(III) oxides, make a negligible contribution to binding. The selectivity of domain III for binding to akageneite is striking given that all three domains bind with high affinity to the mesoporous ITO electrodes used for PFV.

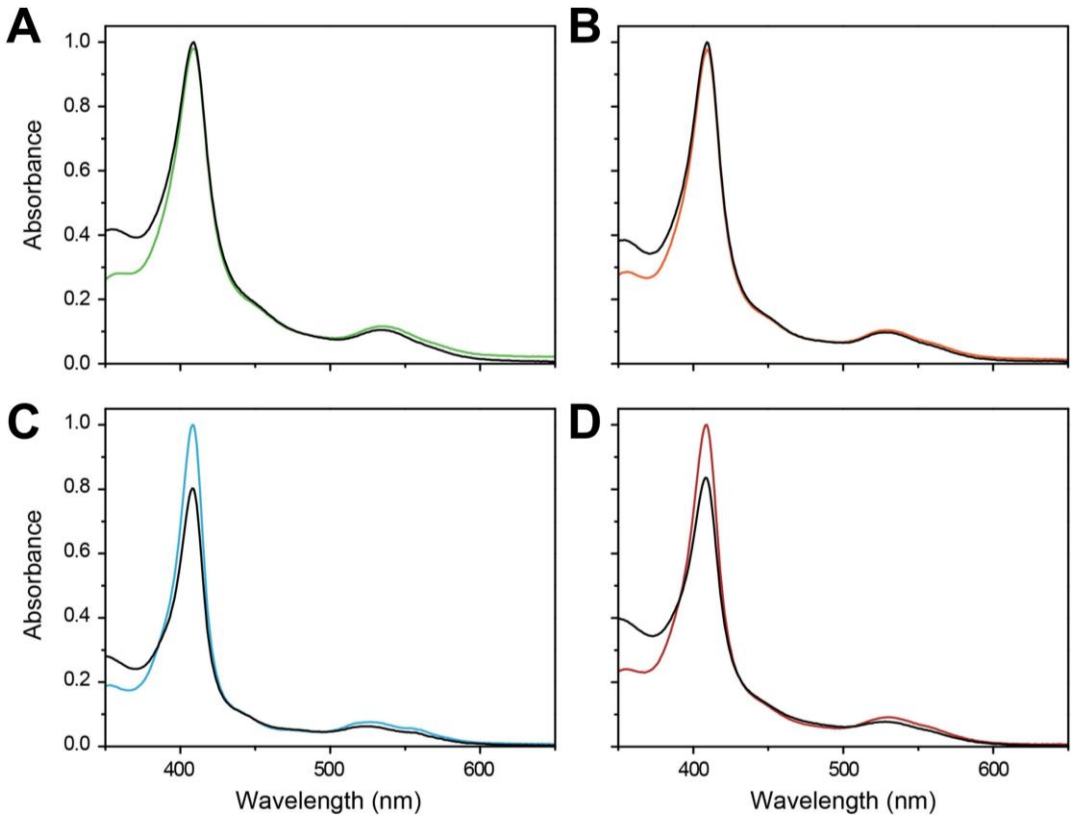


Figure 4.4: Iron oxide (akageneite) affinity of PgcA heme isoforms. Absolute electronic absorbance spectra of initial PgcA protein sample (coloured line), and supernatant after 10 minute incubation with 55 mM of solid akageneite followed by centrifugation to remove protein bound to the mineral particles. The samples contained 2 μM PgcA monoheme domains I **(A)**, II **(B)**, III **(C)** and 0.7 μM triheme PgcA **(D)**, all in 32 mM sodium phosphate, pH 7.0, 100 mM NaCl.

Oxidation of PgcA heme isoforms by iron oxide substrates

UV-visible spectroscopy was used to monitor the electron transfer reaction between triheme PgcA and iron(III) oxides. Titration with iron(III) oxides caused the complete oxidation of all three hemes within the sample (Figure 4.5A). This is notable, as previously only domain III was shown to reduce iron(III) oxides[2], indicating that domains I and II oxidise solely via electron transfer through domain III.

The UV-visible spectral features of the three cytochrome domains are not easily distinguished, but their 1D ^{1}H-NMR spectra present unique and identifiable signal patterns[2]. The 1D ^{1}H-NMR spectra of the triheme PgcA in the fully reduced and

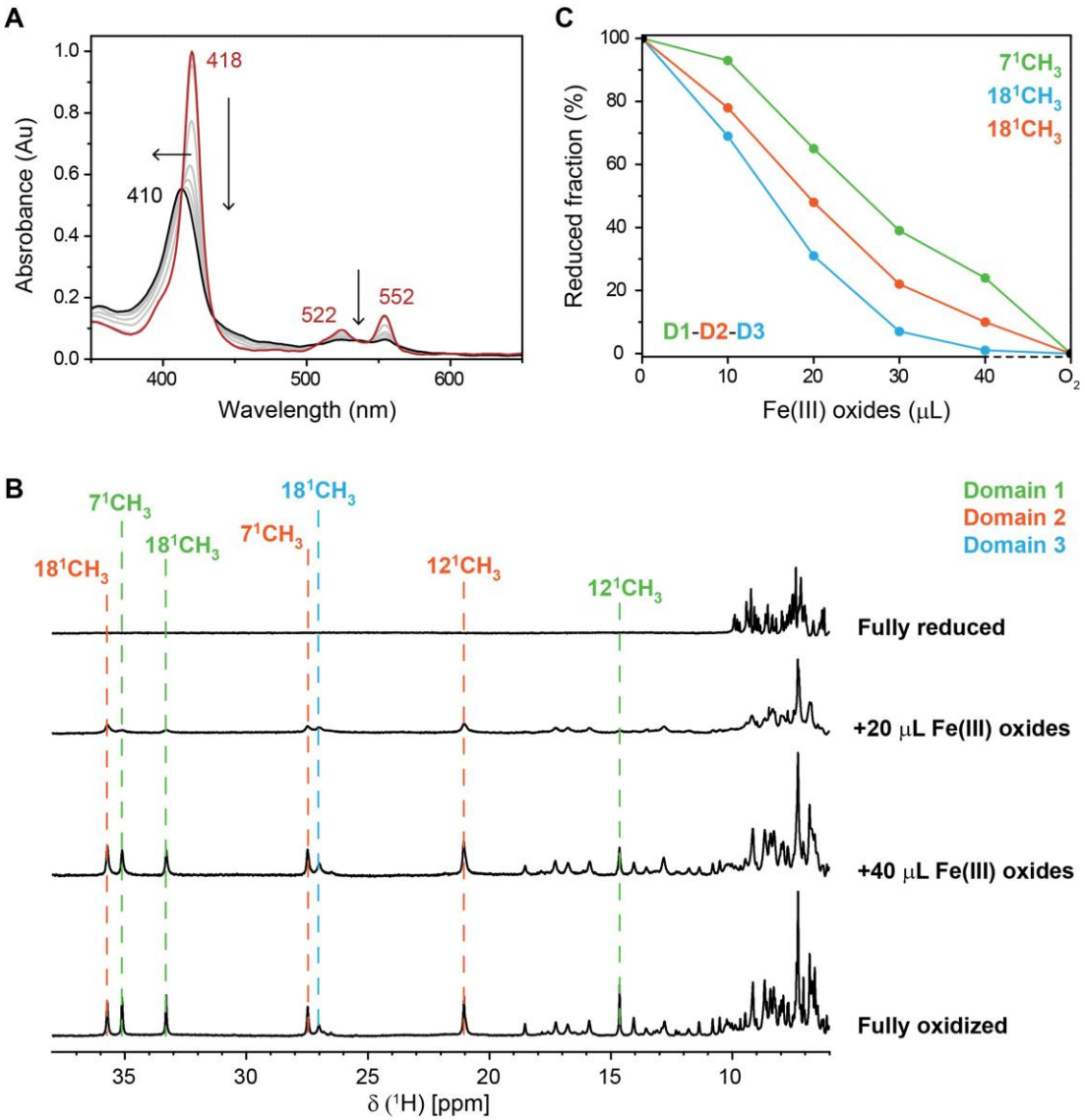


Figure 4.5: Reduction of Fe(III) oxides by triheme PgcA. (**A**) UV-visible spectra of triheme PgcA upon across Fe(III) oxide titration (black and grey), fully reduced cytochrome (red). (**B**) Variation of the low-field 1D 1H NMR spectra of triheme PgcA upon across Fe(III) oxide titration, dashed lines indicate the different heme methyl resonances of the monoheme domains I (green), II (orange) and III (blue). (**C**) Reduced fractions of each monoheme domain across the Fe(III) oxides titration deconvoluted from 1D 1H NMR spectra.

fully oxidized states are well resolved (Chapter 4 appendix figure 4), and the resonances superpose well with those of the individual monoheme domains[2]. This was exploited to track the individual redox states of each domain in the triheme PgcA following sequential additions of iron(III) oxides (Figures 4.5B and 4.5C). The reduced fractions of each domain decreased along the titration, according to their E_m, following the expected thermodynamic route. While these experiments cannot distinguish between inter- and intra-PgcA electron transfer pathways, these observations show clearly that domain III directly reduces iron(III) oxides, while domains I and II are able to recharge domain III for further electron transfer.

Discussion

In this chapter, thorough spectroscopic and potentiometric analysis of PgcA is presented. The pH dependent voltammograms show that two domains I and II have E_m that are significantly more pH sensitive than domain III. This may support the suggestion that E281 and D372 play an important role lowering the E_m of those domains since their ability to perform this role could depend on their protonation state. If so, it is unclear why these domains are His-Met at all since this feature is unusual for hemes involved in EET. In fact the relative potentials of the three domains are such that reduction of domain III by I and II is actually thermodynamically 'uphill', however because this is ultimately coupled to processes that are exergonic the reduction of domain III by I and II is still made possible.

Sequence alignments of the $(PT)_x$ motifs from PgcA and its homologues have shown them to be highly conserved[1], and for many years the motif has been proposed to facilitate interactions with minerals[5]. The findings presented in this chapter suggest that this hypothesis may require revision. In this work, differences in the binding affinity for triheme PgcA and monoheme domain III with iron(III) oxide were not detected. This does not support the suggestion that the (PT)x motifs play a crucial role in that interaction.

Why then are the $(PT)_x$ motif conserved? Proline is a torsion-constrained amino acid that prevents the formation of secondary structure, so its incorporation must heavily impact the flexibility of the linker regions. When the PgcA sequence is analysed by ExPASy-*PeptideCutter*[9], only proteinase K can cleave the $(PT)_x$ motif. This suggests this sequence would be resistant to cleavage by many proteases, a feature essential for a protein exported into the extracellular environment. This suggestion is supported by the observation that in some homologues of PgcA, the Thr residues are replaced by Ala residues[1] which implies that the Pro residues are the important of the two, and that the hydroxyl group of Thr residues proposed to interact with minerals is not so crucial.

Taken together, these observations suggest that the $(PT)_x$ motifs most likely serve as robust linkers that maintain an optimal distribution of distances between

cytochrome domains, allowing for optimal electron transfer to occur between them, while ensuring the ability to stretch across the 18 nanometres measured in *Chapter 3*. These findings highlight the difficulties associated with characterising the of enzymes to solid phase substrates. This point is further underscored by the fact during the preparation of this thesis, the behaviour of PgcA isoforms in terms of iron binding was found to be highly influenced by the nature of the buffer used. In contrast with the behaviour observed with sodium phosphate pH 7.0, when trialled with HEPES pH 7.8 all PgcA isoforms bound to an extent of approximately 100% (Chapter 4 appendix figure 5), a total reversal in outcome, although it remains unclear whether this effect is derived from the different buffer compounds themselves or the different pH values used.

The substrate reduction studies do however indicate that only domain III, with the most solvent exposed heme, is capable of interacting with iron(III) oxides under the experimental conditions used. While these are unlikely to represent the physiological and environmental conditions under which PgcA is active, they do provide a useful mechanism for studying inter-domain electron transfer between the different domains. The results clearly show that all three domains of triheme PgcA become oxidised by iron(III) oxides. This indicates that electrons are transferred from domains I-II onto domain III, and then onto minerals, because *Fernandes et* al directly demonstrated that neither domains I nor II can be oxidised by iron(III) oxides in the absence of domain III[2].

Taken together the data presented in this chapter provides important information that complements the structural insights revealed in *Chapter 3*. In particular relevance for future experiments are probably the EPR spectra and electronic absorbance spectra presented here. These experiments support the model presented in which the PgcA hemes are distributed into totally distinct domains as they display spectra that are very similar when in their triheme and monoheme forms. Moreover the resolution of distinct gmax values attributed to each heme, and isosbestic points in the absorbance spectra that are shared by domains I and II but not III, could pave the way for innovative intramolecular electron transfer rate measurements.

Methods have been described recently for rapid and site-specific injection of electrons into heme proteins, using pump-probe spectroscopy[3,10]. This allowed for determination of heme-heme electron transfer rates[4]. These could be applied likely without major obstacle to PgcA. With a photosensitizer on one monoheme domain, electrons could be injected there, and using either the gmax values established here, or the domain I/II electronic absorbance isosbestic points established here, the rate of transfer of electrons between domains could be probed. This system might be exceptionally useful also for testing the effect of modifications suggested in *Chapter 3*, like the truncation of the Pro-Thr repeat motifs. It could provide the ability to examine the effect of altering other variables

too, like viscosity which might influence tethered diffusive heme-heme electron transfer in a way that will not impact close-packed heme-heme electron transfer.

Given the rate of electron transfer is likely to be considerably lower than in canonical multiheme cytochromes, it will be particularly important to establish that electron transfer is intramolecular rather than intermolecular. Since the rate of the later will be a function of protein concentration, repeating experiments at different concentrations provides a simple strategy to preclude the effect of intermolecular processes.

Methods

Electronic absorbance spectroscopy

For electronic absorbance spectroscopy, samples of domain I, II, III and triheme PgcA were prepared up to absorbance at 410 nm of 1 with a pathlength of 1 cm, in 50 mM Tris HCl, 50 mM, NaCl pH 8.0. The headspace was then sparged with nitrogen gas and electronic absorbance spectra were measured using a Cary 60 UV-Vis Spectrophotometer. Next, 10 mg mL^{-1} sodium dithionite in anaerobic 50 mM Tris-HCl, 50 mM NaCl, pH 8.0 was added using a syringe, to a final dithionite concentration of 3 mM after which another spectrum was recorded. Further dithionite was then added and spectra measured to ensure the initial dithionite spectrum represented fully reduced protein.

Protein film voltammetry

Hierarchical mesoporous ITO electrodes were prepared as described previously[11]. 10 mg mL^{-1} Protein samples in 20 mM HEPES pH 7.8, 100 mM NaCl were deposited onto ice cold ITO electrodes and incubated for approximately 5 minutes. After rinsing, electrodes were introduced to anaerobic buffer-electrolyte solution within a glass electrochemical cell, inside a Faraday cage and situated in a N_2-filled chamber (atmospheric O_2 < 2 ppm). Cyclic voltammetry was performed with a three-electrode configuration using an Autolab PGSTAT30 potentiostat (EcoChemie) controlled by NOVA software. An Ag/AgCl (saturated KCl) reference electrode was used with a Pt wire counter electrode, and potential values corrected to SHE by addition of +0.197 V. Charging current baselines were subtracted from the measured voltammograms to yield the Faradaic currents. For monoheme domains Faradaic currents were fit to single n=1 Nernstian peaks using QSoas[12]. Peak potentials for the oxidative and reductive signals were averaged across three independent scans, and standard error calculated, to obtain the reported E_m values. For triheme PgcA, Faradaic currents were fit to three Nernstian n=1 centres of equal area using QSoas[12].

Dr Joshua Burton prepared electrodes and collected data for the optically monitored potentiometric titration of ITO coated electrodes. Those electrodes were studied in anaerobic buffer-electrolyte solution within a sealed cuvette

fitted with a Pt wire as counter electrode and an AgCl coated Ag wire as reference electrode. The cuvette was placed in a UV-visible absorbance spectrophotometer and the working electrode poised at the desired potentials. The reference electrode was calibrated by cyclic voltammetry with potassium hexa-cyanoferrate(III) (30 mM in 50 mM Tris buffer, 100 mM NaCl, pH 8.0) taking E_m = +420 mV vs SHE.

Electron paramagnetic resonance spectroscopy

EPR spectra were recorded by Dr Dimitri A. Svistunenko and Dr Marcus J. Edwards, at 10 K using a Bruker E500 (X-band) EPR spectrometer and Oxford Instruments liquid helium systems. Instrument parameters were as follows: microwave frequency ν_{MW} = 9.467 GHz, modulation frequency ν_M = 100 kHz, time constant τ = 82 ms, microwave power = 3.19 mW, modulation amplitude A_M = 5 G, scan rate v = 22.6 Gs^{-1}. Wilmad SQ EPR tubes (Wilmad Glass, Buena, NJ) with OD = 4.05 ± 0.07 mm and ID = 3.12 ± 0.04 mm (mean ± range) were used for freezing EPR samples in methanol cooled with solid CO_2. After freezing, the EPR tubes with samples were kept in liquid nitrogen until required for measurements.

Mineral oxide binding pull-down assays

Binding pull-down assays were performed by Dr Tomás M. Fernandes and Dr Leonor Morgado. The binding of iron(III) oxides to triheme PgcA was probed by UV-visible spectroscopy, as described previously[2]. Spectra were recorded between 300 and 700 nm using an Evolution 201 spectrophotometer (Thermo Scientific). After acquisition of spectra with 2 µM protein samples prepared in 32 mM sodium phosphate buffer pH 7 with NaCl (100 mM final ionic strength), the protein was incubated with 55 mM of akageneite for 10 minutes, after which the resulting solutions were centrifuged at 5000x *g* for 10 minutes. Additional spectra were acquired with the resulting supernatants and the putative binding of the protein to iron(III) oxides was evaluated based on differences in spectral intensity.

Iron(III) oxide mediated PgcA oxidation monitored by UV-visible spectroscopy

Electronic absorbance spectral assessment of iron(III) oxide reduction by PgcA was performed by Dr Tomás M. Fernandes and Dr Leonor Morgado. The electron transfer reaction between triheme PgcA and iron(III) oxides was assessed by UV-visible spectroscopy measurements at 25 °C, performed inside an anaerobic glovebox with O_2 levels kept under 0.1 ppm. Iron(III) oxides were prepared with a final concentration of 500 µM. The cytochrome samples (2 µM) were prepared in degassed 32 mM sodium phosphate buffer with NaCl (100 mM final ionic strength) at pH 7. Before each experiment, the protein was reduced with solution of sodium dithionite. Iron(III) oxides were added to the fully reduced cytochrome in several increments and the consequent re-oxidation of hemes was monitored by recording UV-visible spectra between 300 and 700 nm.

Iron(III) oxides reduction monitored by NMR spectroscopy

NMR Monitored iron(III) oxide reduction by PgcA was performed by Dr Tomás M. Fernandes and Dr Leonor Morgado. The electron transfer reaction between triheme PgcA and iron(III) oxides was also assessed by NMR spectroscopy. The samples of protein 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 solution of sodium dithionite. Iron(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 iron(III) oxides solution, and (iii) after exposing the sample in the NMR tube to atmospheric O_2. After each addition of iron(III) oxides, the samples were incubated for 5 minutes and centrifuged at 5000 G for 10 minutes, prior to spectral acquisition. These experiments were acquired in a Bruker Avance III 600 MHz spectrometer. The residual H_2O signal was used as an internal reference for the calibration of the ^{1}H chemical shifts relative to sodium trimethylsilylpropanesulfonate (DSS) at 0 ppm. All spectra were acquired at 25 °C and processed using TopSpin 4.1.4 (Bruker BioSpin, Karlsruhe Germany). The reduced fractions of each cytochrome domain along the redox reaction were calculated by integrating the ^{1}H NMR heme methyl signals of each domain in the intermediate redox state, relative to the fully oxidized state.

Chapter 4: References

1. Zacharoff, L. A., Morrone, D. J. & Bond, D. R. Geobacter sulfurreducens extracellular multiheme cytochrome PgcA facilitates respiration to Fe(III) oxides but not electrodes. *Frontiers in Microbiology* **8**, (2017).

2. Fernandes, T. M., Silva, M. A., Morgado, L. & Salgueiro, C. A. Hemes on a string: insights on the functional mechanisms of PgcA from Geobacter sulfurreducens. *Journal of Biological Chemistry* **299**, 105167 (2023).

3. Van Wonderen, J. H., Hall, C. R., Jiang, X., Adamczyk, K., Carof, A., Heisler, I., Piper, S. E. H., Clarke, T. A., Watmough, N. J., Sazanovich, I. V., Towrie, M., Meech, S. R., Blumberger, J. & Butt, J. N. Ultrafast Light-Driven Electron Transfer in a Ru(II)tris(bipyridine)-Labeled Multiheme Cytochrome. *Journal of the American Chemical Society* **141**, 15190–15200 (2019).

4. van Wonderen, J. H., Adamczyk, K., Wu, X., Jiang, X., Piper, S. E. H., Hall, C. R., Edwards, M. J., Clarke, T. A., Zhang, H., Jeuken, L. J. C., Sazanovich, I. V., Towrie, M., Blumberger, J., Meech, S. R. & Butt, J. N. Nanosecond heme-to-heme electron transfer rates in a multiheme cytochrome nanowire reported by a spectrally unique His/Met-ligated heme. *Proceedings of the National Academy of Sciences of the United States of America* **118**, e2107939118 (2021).

5. Lower, B. H., Lins, R. D., Oestreicher, Z., Straatsma, T. P., Hochella, M. F., Shi, L. & Lower, S. K. In vitro evolution of a peptide with a hematite binding motif that may constitute a natural metal-oxide binding archetype. *Environmental Science and Technology* **42**, 3821–3827 (2008).

6. Kobashigawa, Y., Nishimiya, Y., Miura, K., Ohgiya, S., Miura, A. & Tsuda, S. A part of ice nucleation protein exhibits the ice-binding ability. *FEBS Letters* **579**, 1493–1497 (2005).

7. Vries, S. De & Albracht, S. P. J. Intensity of highly anisotropic low-spin heme EPR signals. *Biochimica et Biophysica Acta (BBA) - Bioenergetics* **546**, 334–340 (1979).

8. Zoppellaro, G., Bren, K. L., Ensign, A. A., Harbitz, E., Kaur, R., Hersleth, H. P., Ryde, U., Hederstedt, L. & Andersson, K. K. Studies of ferric heme proteins with Highly Anisotropic/Highly Axial Low Spin (S = 1/2) electron paramagnetic resonance signals with bis-histidine and histidine-methionine axial iron coordination. *Biopolymers* **91**, 1064 (2009).

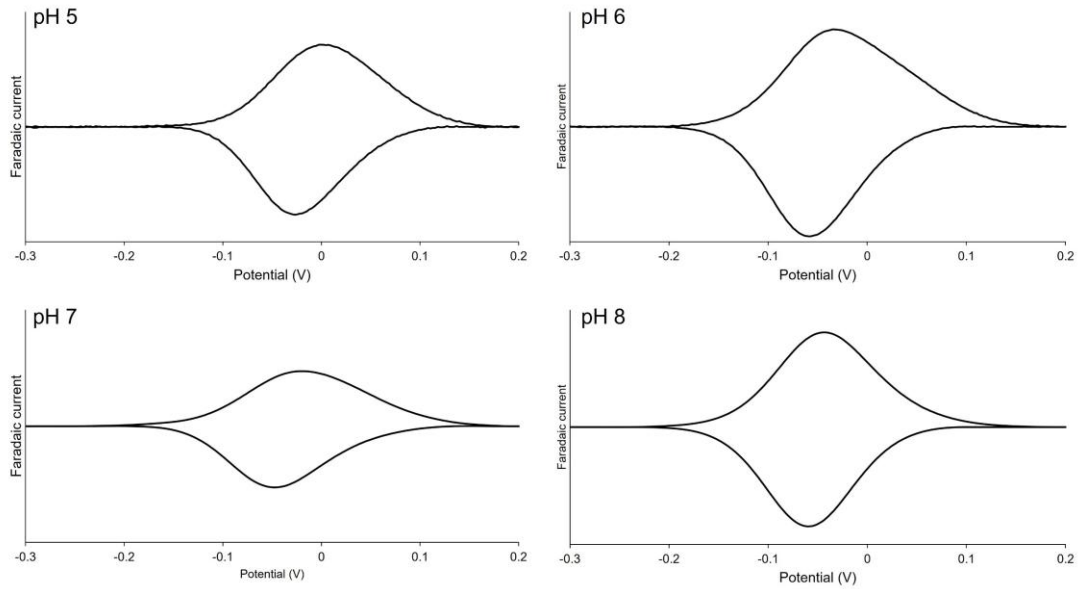
9. Wilkins, M. R., Gasteiger, E., Bairoch, A., Sanchez, J. C., Williams, K. L., Appel, R. D. & Hochstrasser, D. F. Protein identification and analysis tools in the ExPASy server. *Methods in molecular biology* **112**, 531–552 (1999).

10. van Wonderen, J. H., Li, D., Piper, S. E. H., Lau, C. Y., Jenner, L. P., Hall, C. R., Clarke, T. A., Watmough, N. J. & Butt, J. N. Photosensitised Multiheme Cytochromes as Light-Driven Molecular Wires and Resistors. *ChemBioChem* **19**, 2206–2215 (2018).

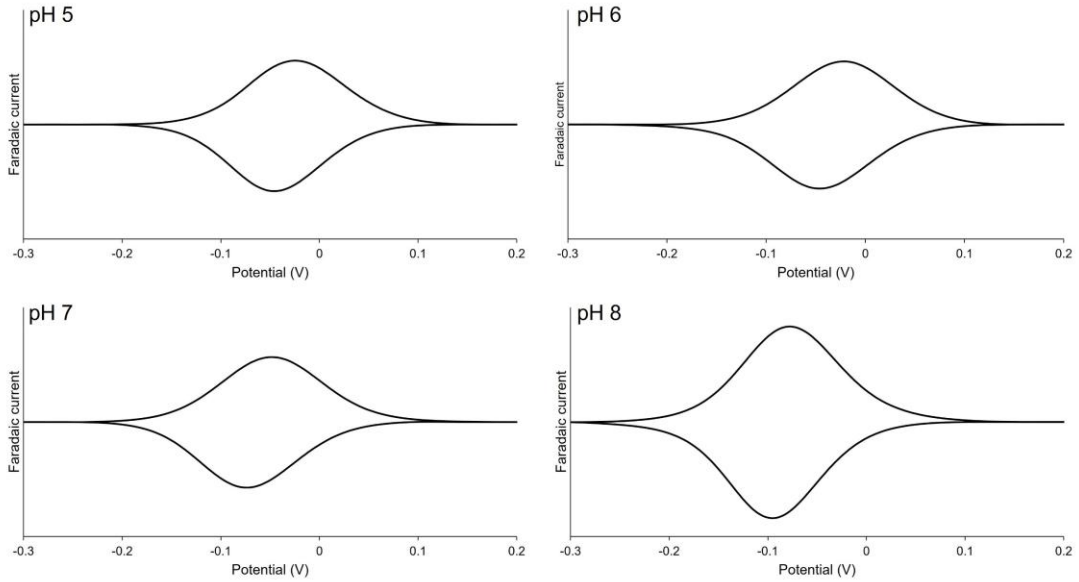
11. Mersch, D., Lee, C. Y., Zhang, J. Z., Brinkert, K., Fontecilla-Camps, J. C., Rutherford, A. W. & Reisner, E. Wiring of Photosystem II to Hydrogenase for Photoelectrochemical Water Splitting. *Journal of the American Chemical Society* **137**, 8541–8549 (2015).

12. Fourmond, V. QSoas: A Versatile Software for Data Analysis. *Analytical Chemistry* **88**, 5050–5052 (2016).

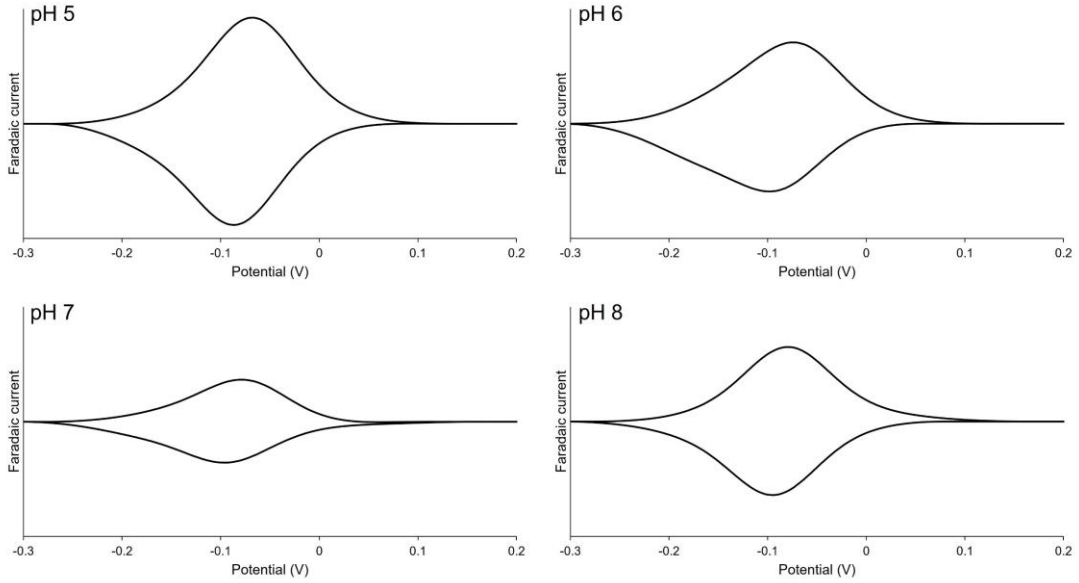
Chapter 4: Appendix



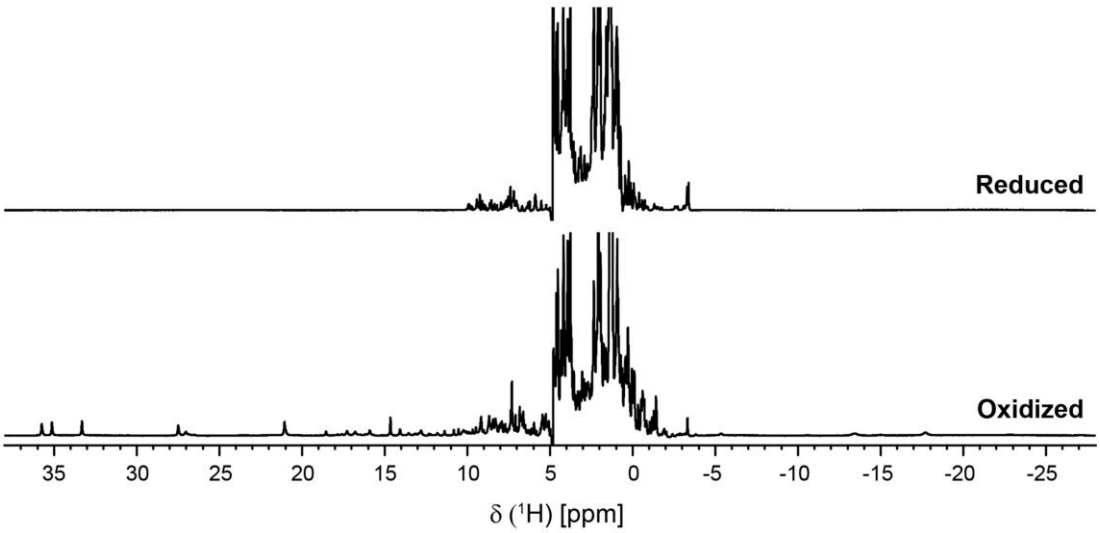
Appendix figure 1: Domain I PFV pH dependence. Representative 0.01 V s^{-1} voltammograms of Domain I adsorbed onto ITO electrodes in 50 mM NaCl with 50 mM Tris-HCl (pH 8.0), 50 mM HEPES (pH 7.0), 50 mM 2-(N-morpholino)ethanesulfonic acid (pH 6.0) or 50 mM Na acetate (pH 5.0).



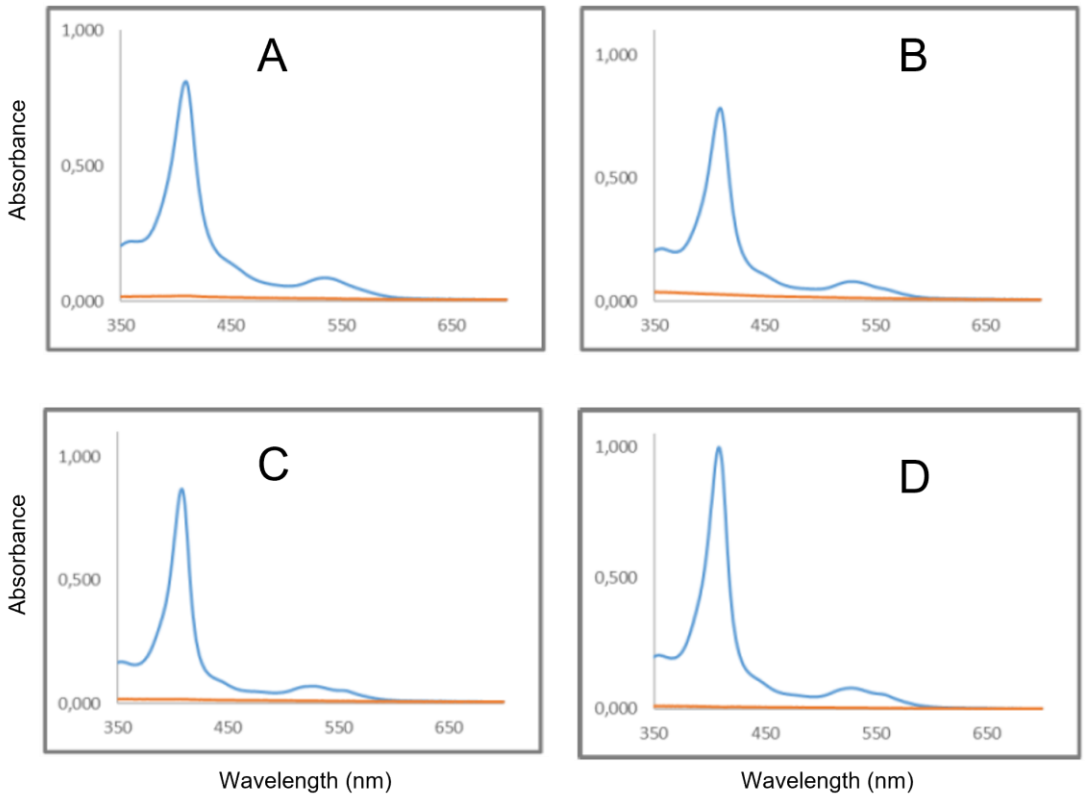
Appendix figure 2: Domain II PFV pH dependence. Representative 0.01 V s^{-1} voltammograms of Domain I adsorbed onto ITO electrodes in 50 mM NaCl with 50 mM Tris-HCl (pH 8.0), 50 mM HEPES (pH 7.0), 50 mM 2-(N-morpholino)ethanesulfonic acid (pH 6.0) or 50 mM Na acetate (pH 5.0).



Appendix figure 3: Domain III PFV pH dependence. Representative 0.01 V s^{-1} voltammograms of Domain I adsorbed onto ITO electrodes in 50 mM NaCl with 50 mM Tris-HCl (pH 8.0), 50 mM HEPES (pH 7.0), 50 mM 2-(N-morpholino)ethanesulfonic acid (pH 6.0) or 50 mM Na acetate (pH 5.0).



Appendix figure 4: 1D ^{1}H-NMR spectra of triheme PgcA. Triheme PgcA 1D 1H NMR spectra, in 32 mM sodium phosphate pH 7.0, at 25 °C.



Appendix figure 5: Akageneite binding of PgcA isoforms in the presence of HEPES. Repeat of PgcA isoform akageneite binding experiments performed in 32 mM Na Phosphate instead in 20 mM HEPES pH 7.8, absolute electronic absorbance spectra of monoheme domains I (A), II (B), III (C) and triheme PgcA (D) initially at 2, 2, 2 and 0.7 µM before addition of iron oxide. The blue line corresponds to initial samples while the orange line corresponds to those after 10 minute incubation with Akageneite (55 µM) and subsequent centrifugation.

Chapter 5: The structure and function of the PgcA non-heme domain

Introduction

In *G. sulfurreducens*, the triheme cytochrome PgcA is thought be extracellularly localised where it plays an important role in EET[1]. The *pgcA* gene is upregulated during mineral respiration[2,3], and Δ*pgcA G. sulfurreducens* cultures display poor rates of iron(III) oxide reduction, but are unimpaired in reducing iron(III) citrate[1].

The PgcA protein contains a putative N-terminal motif, centred on residue Cys39, at which a lipid group is predicted to be incorporated, partially embedding the protein probably within the outer-leaflet of the outer-membrane[4]. The exogenous addition of truncated and soluble PgcA produced recombinantly in *Shewanella oneidensis* restored the wild-type iron(III) oxide reduction rate of the *ΔpgcA G. sulfurreducens* strain, consistent with its putative extracellular localisation[1]. However this observation does suggest that anchoring to the outer-membrane is not critical for its activity.

A recently described *G. sulfurreducens* strain, in which the periplasmic electron shuttle encoding genes *ppcA-ppcE* were deleted, rapidly developed an additional mutation in which the PgcA lipobox region was also deleted (Δ37-39)[5]. This appears to have trapped the PgcA protein within the periplasm and restored the strain's ability to perform EET, indicating that PgcA is probably extracellularly localised in wild-type *G. sulfurreducens*, and that it requires its lipid moiety for translocation across the outer-membrane[5]. Its secretion is likely via the type-II secretion system (T2SS) which is able to secrete large, cofactor-containing proteins in a folded state[6]. However the mechanism by which the T2SS recognises its substrates remains poorly understood[6].

A further *G. sulfurreducens* strain, JS-1, has also been described which was cultivated to display improved EET without the ability to assemble protein nanowires. This strain eventually developed a mutation in the *pgcA* promoter, which resulted in dramatic overexpression of the gene and restored the wild-type EET rate[7]. Analysis of the spent growth medium from JS-1 revealed it was rich in PgcA that appeared to be truncated and soluble[7], which is inconsistent with the observation that the protein requires an N-terminal lipid moiety for translocation across the outer-membrane.

In the preceding chapters and the literature, PgcA has been shown to contain three monoheme domains which are separated from one another by flexible linkers, and can stretch as far as 180 Å in diameter that facilitates electron transfer via a novel mechanism of tethered diffusion (figure 5.1A)[1,8,9]. In all those

cases, the N-terminal domain of PgcA, which is not predicted to contain a heme cofactor, remained conspicuously absent and uncharacterised since when produced recombinantly in *Shewanella oneidensis*, PgcA undergoes cleavage between M252 and P253 removing the domain (*Chapter 3)*, and obtaining sufficient biomass for biochemical and crystallographic studies is challenging in the native organism.

This chapter describes the recombinant preparation of the PgcA N-terminal, non-heme domain (PgcA-NHD) in *Escherichia coli*, which exhibits intrinsic autoproteolytic activity. These findings provide a basis for reconciling the seemingly inconsistent observations regarding its localisation and post-translational truncation, and suggest it contains a novel mechanism for exporting diverse proteins into the extracellular environment.

Results

Cloning, purification and initial characterisation of PgcA-NHD

Attempts were made to purify PgcA-NHD as an isolated protein recombinantly from several constructs in *E. coli* and *S. oneidensis* but these yielded no detectable expression. Fusion to the *E. coli* maltose binding protein often imparts dramatically improved expression, stability or solubility to proteins[10] so this strategy was adopted for PgcA-NHD.

DNA encoding the regions of PgcA-NHD predicted to form an ordered structure by AlphaFold2[11] was cloned into the pMAL-c6T plasmid (NEB) as an MBP fusion (figure 5.1A). An N-terminal hexa-His tag and a C-terminal StrepII tag (WSHPQFEK) were also incorporated as these provided flexibility in purification and increased 280 nm absorbance through the addition of a tryptophan residue which PgcA-NHD does not have natively: ($N_{terminus}$-His_6-MBP-Linker-TEV-(PgcA-NHD)-StrepII-$C_{terminus}$). This construct was expressed in the cytoplasm of *E. coli* BL21 (DE3), and soluble MBP-(PgcA-NHD) was purified as described in *Chapter 5 methods*.

SDS-PAGE of purified MBP-(PgcA-NHD) revealed it consisted of two dominant species at ~40 and ~20 kDa, with a faint band for full-length protein at ~60 kDa (Figure 5.1B). Separation of the two dominant species was attempted by Ni^{2+}-IMAC, MBP^{trap} affinity chromatography and SEC (Figure 5.1C) but was not possible suggesting that they were bound together tightly.

LC-MS of purified MBP-(PgcA-NHD) identified two homogeneous species with masses 44,630 Da and 19,619 Da (Figure 5.2A and B). These measurements indicate the cleavage site is located between residues G67 and A68 (numbered as in the wild-type PgcA sequence) of PgcA-NHD which correspond to the 10^{th} and 11^{th} residues of the sequence included in this construct (predicted fragment masses 44,630 and 19,620 Da). When full length wild-type PgcA is overexpressed recombinantly in *S. oneidensis* it undergoes proteolysis between

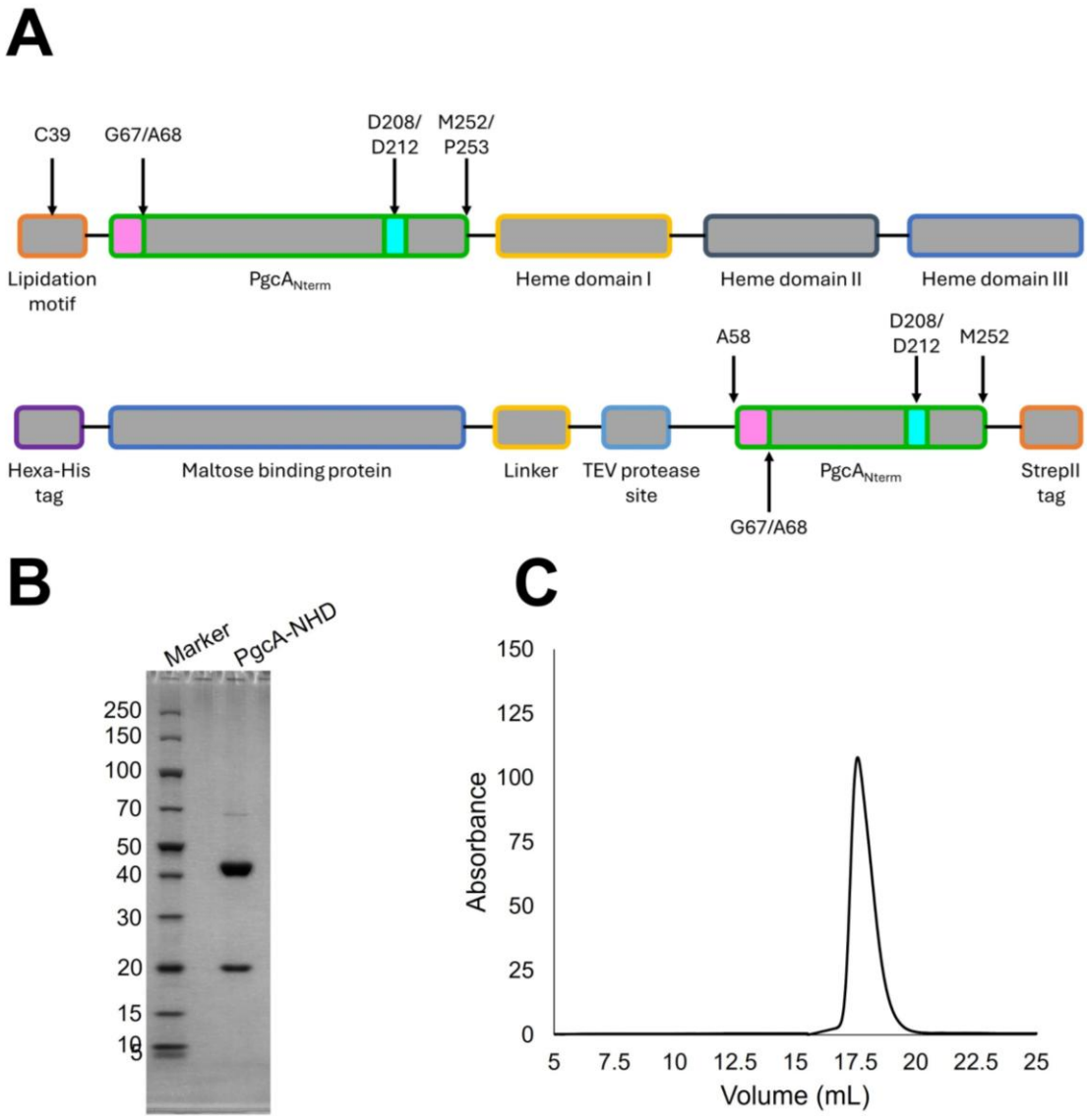


Figure 5.1: Recombinant PgcA-NHD is post-translationally cleaved into two fragments that do not dissociate. (A) Infographic showing wild-type PgcA (upper) and pMBP-(PgcA-NHD) plasmid (lower) domain architecture, with residues important in autoproteolytic activity highlighted. **(B)** Reducing SDS-PAGE of MBP^{trap} purified PgcA-NHD stained with Coomassie blue, molecular weight marker labelled in kDa. **(C)** Superose 6 Increase 10/300 GL SEC 280 nm trace for MBP-(PgcA-NHD).

residues M252 and P253 (Chapter 3). No observation of the M252/P253 cleavage was observed when using the MBP-(PgcA-NHD) construct, although in this construct the P253 residue was replaced with a purification tag.

Identification of proteolysis dependent residues D208 and D212

The G67-A68 site in an AlphaFold2 model structure of PgcA-NHD was inspected for residues that might be responsible for this proteolytic activity. The D208 and D212 residues were positioned directly above the scissile bond in an arrangement reminiscent of an aspartic protease enzyme (Figure 5.2C). Therefore these residues were selected for mutagenesis. Constructs were prepared with D208A, D212A, and D208A+D212A substitutions, and the proteins

purified. SDS-PAGE showed both single substitution variants displayed impaired cleavage ability with full length MBP-(PgcA-NHD) becoming the dominant

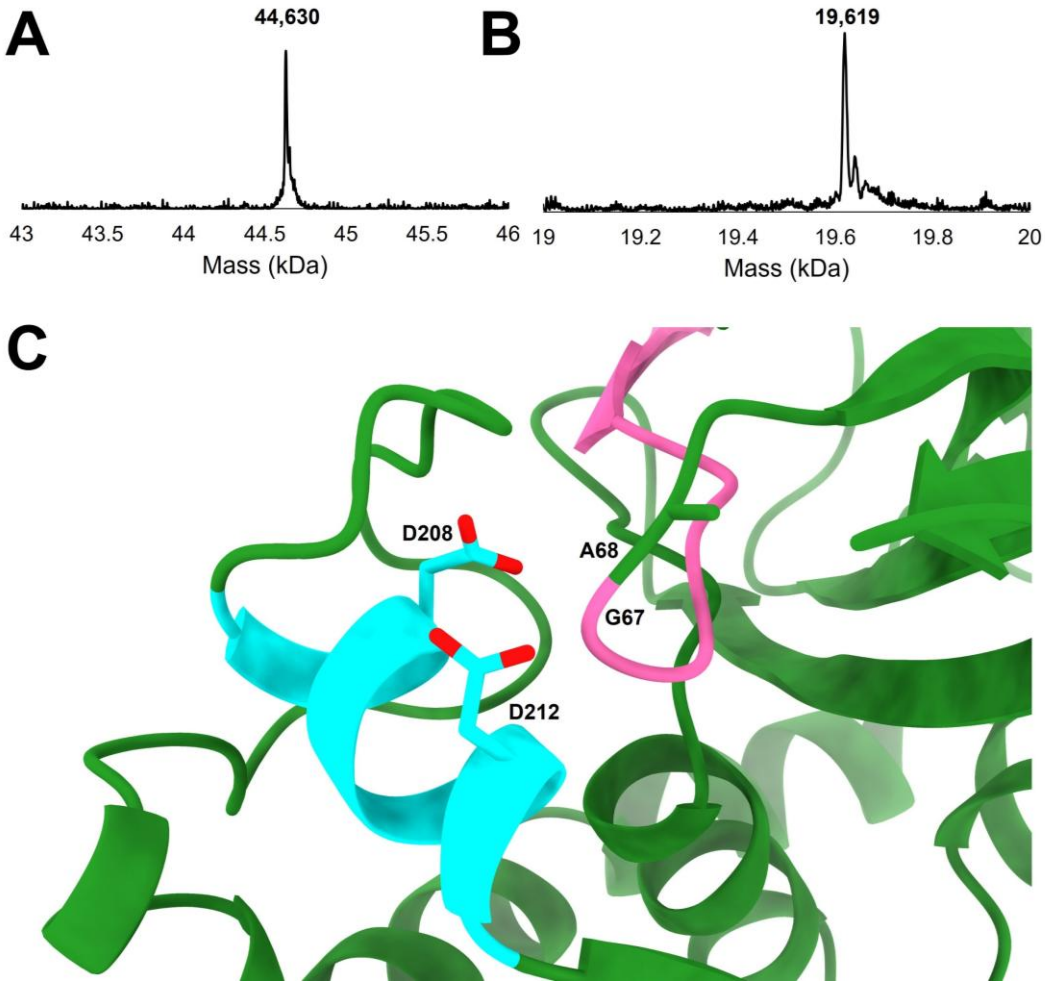


Figure 5.2: Proteolytic cleavage in PgcA-NHD occurs between residues G67 and A68, near D208 and D212. Deconvoluted mass spectra for WT MBP-(PgcA-NHD) ~40 kDa **(A)** and ~20 kDa **(B)** species, predicted fragment masses from cleavage at G67-A68: 44,630 and 19,620 Da, respectively (**C**) AlphaFold3 predicted model of PgcA-NHD cleavage site, alpha helix containing D208-D212 shown in cyan, N-terminal cleavage peptide shown in pink.

species, and the double substitution proving severely impaired which was supported by LC-MS (Figure 5.3A and B).

It was attempted to isolate PgcA-NHD$^{D208A+D212A}$ protein from the MBP by TEV-proteolysis, but this protein was resistant to the TEV protease. A further construct was prepared that incorporated an additional linker between the TEV site and the PgcA-NHD ($N_{terminus}$-His_6-MBP-Linker-TEV-Linker-(PgcA-NHD)-StrepII-$C_{terminus}$). This yielded protein that was readily cleaved by the TEV protease (Chapter 5 appendix figure 2) and thus allowed isolation of PgcA-NHD$^{D208A+D212A}$ protein to high purity (Figure 5.4A), as described in *Chapter 5 methods*. Consistent with the behaviour of the double substitution variant while tethered to the MBP protein, this species was shown to constitute a single un-cleaved polypeptide by LC-MS (Figure 5.4B).

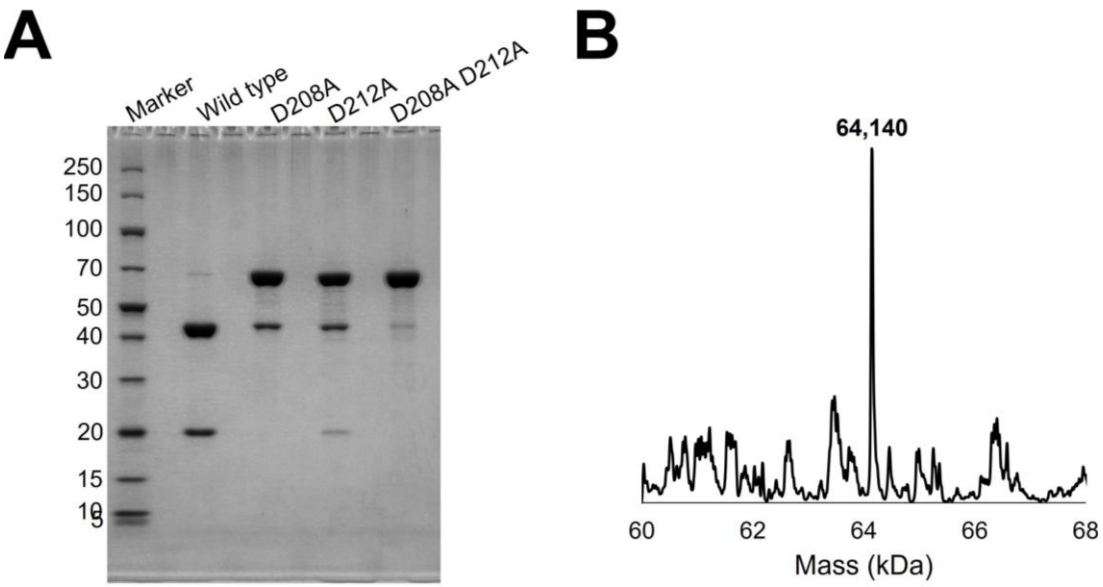


Figure 5.3: Recombinant PgcA-NHD is post-translationally cleaved at residues G67 and A68 which is impaired by mutations at D208 and D212. (A) Reducing SDS-PAGE of MBP^{trap} purified PgcA-NHD variant proteins, stained with coomassie blue, molecular weight marker labelled in kDa. **(B)** Deconvoluted mass spectra for MBP-(PgcA-NHD$^{D208A+D212A}$), predicted mass: 64,144 Da.

Crystallographic analysis of PgcA-NHD

PgcA-NHD$^{D208A+D212A}$ was then crystallised (Figure 5.5) and its structure resolved, covering residues 58-252 of the PgcA amino acid sequence, at 1.65 Å resolution by X-ray diffraction (Figure 5.6A). Its N-terminus consists of a seven stranded beta sandwich motif. A long alpha helix consisting of residues 161-190 runs centrally through the molecule while the remaining secondary structure is comprised of five short additional helices and a three stranded beta sheet.

The D/A208 and D/A212 positions reside in a short alpha helix consisting of residues G206-V213 (Figure 5.6A, cyan). G67 and A68 together with T66 and P69 form a loop that makes a tight turn directly beneath the G206-V213 alpha helix, and also forms the C-terminal end of a beta strand (residues A58-A65).

The A58-A65 beta strand is observed making extensive H-bonding interactions with residues 94-98 and 129-134 (Figure 5.6B). Finally, residues A58-G67 shield an internal hydrophobic pocket comprised of the side chains provided by residues V60, L70, L76, I74, F96, V98, Y107, L109 and I134 (Figure 5.6C). Together these interactions likely make the dissociation of residues A58-G67 from its pocket highly unfavourable, explaining how the two species identified by SDS-PAGE and LC-MS of MBP-(PgcA-NHD$^{wild-type}$) could not be separated by chromatography, despite it having undergone cleavage.

An MBP-(PgcA-NHD$^{D208A+D212A+\Delta 58-69}$) variant was prepared with the goal of determining the properties and structure of PgcA-NHD with its N-terminal

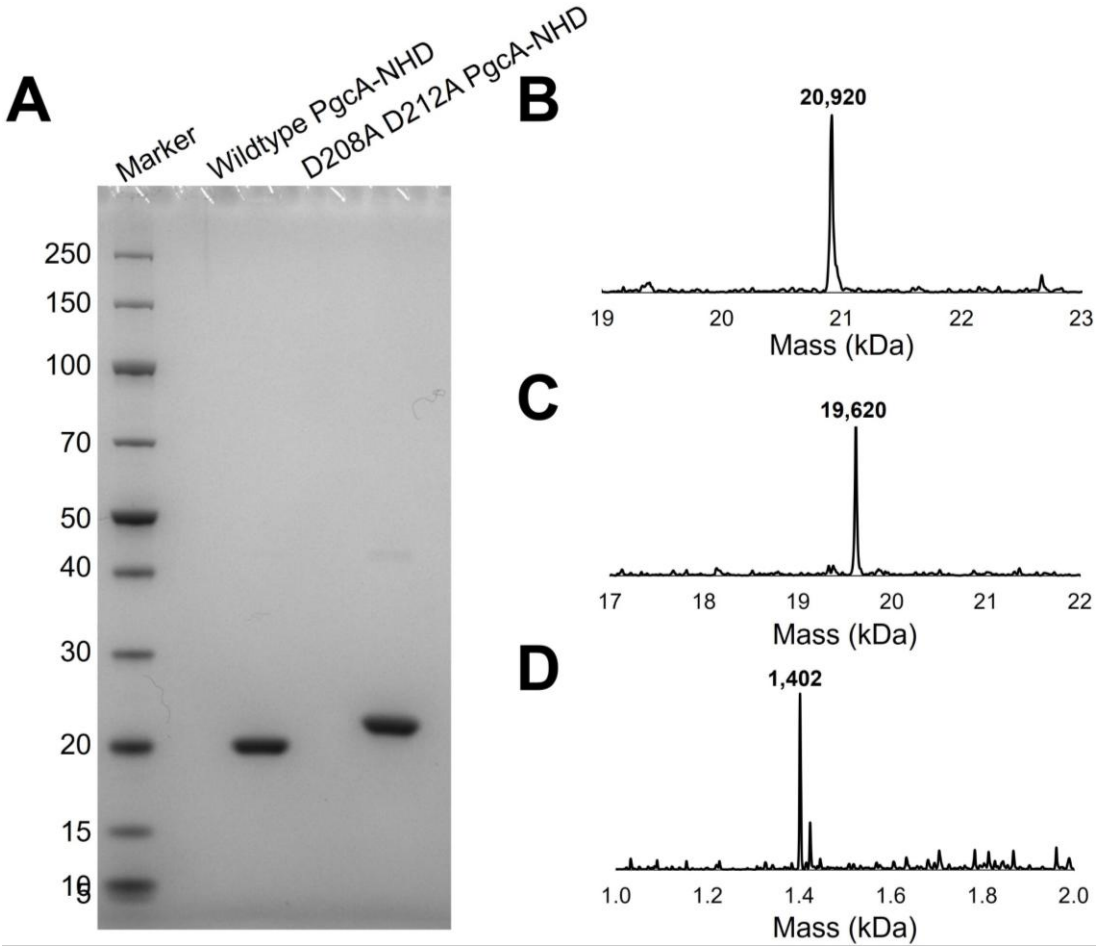


Figure 5.4: Purification of isolated PgcA-NHD and LC-MS. (A) Reducing SDS-PAGE gel of purified PgcA-NHD$^{wild\text{-}type}$ and PgcA-NHD$^{D208A+D212A}$ with MBP removed, molecular weight markers labelled in kDa. Deconvoluted mass spectra for MBP-removed PgcA-NHD$^{D208A+D212A}$ **(B)**, PgcA-NHD$^{wild\text{-}type}$ + cleavage peptide **(C** and **D)** proteins, predicted fragment masses based on amino acid sequence: 20,917, 19,620 and 1,403 Da (top to bottom respectively).

cleavage peptide not present, but attempts to purify a soluble form of this protein were unsuccessful.

A wild-type variant with the additional linker allowing for TEV proteolysis was purified and the MBP removed from the protein (Figure 5.4A). LC-MS of this

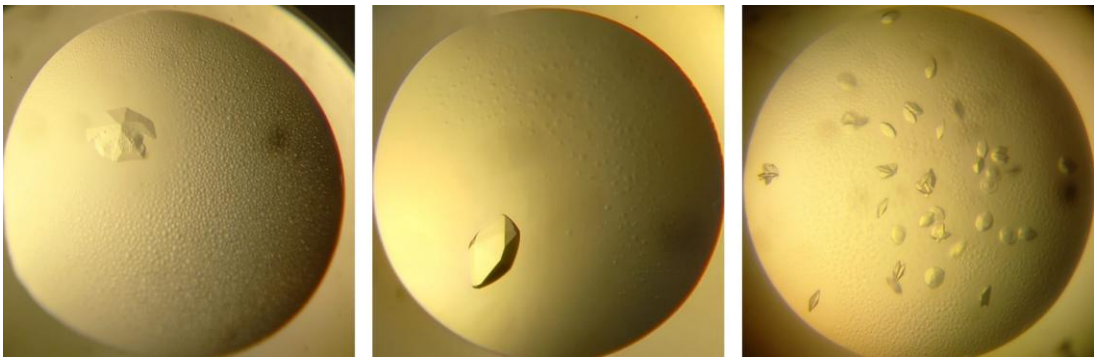


Figure 5.5 Crystallisation of PgcA-NHD$^{D208A+D212A}$ and PgcA-NHD$^{wild\text{-}type}$. Initial crystal material obtained for PgcA-NHD$^{D208A+D212A}$ (left), improved crystals of PgcA-NHD$^{D208A+D212A}$ obtained by seeding (middle) and crystals of PgcA-NHD$^{wild\text{-}type}$ (right) obtained by seeding.

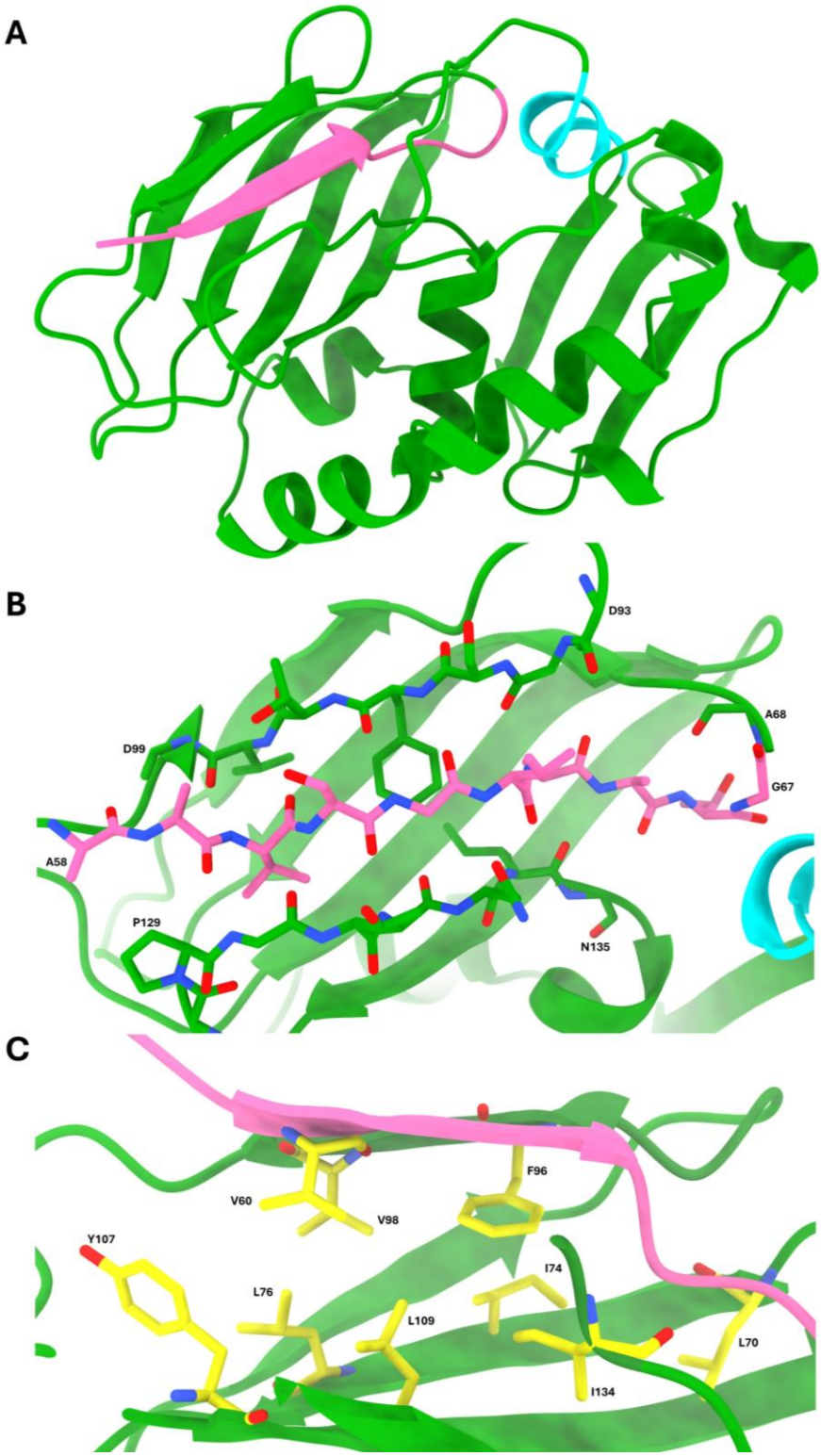


Figure 5.6: Crystal structure of PgcA-NHD$^{D208A+D212A}$. (A) Crystal structure of PgcA-NHD$^{D208A+D212A}$ at 1.65 Å resolution. Residues N-terminal of the cleavage site (A58-G67) shown in pink and residues G206-V213 where the D208A and D212A reside, in cyan. **(B)** Pink stick representation shown of residues A59-G67 and their hydrogen bonding partners 94-98 and 129-134. **(C)** Yellow stick representation of residues contributing to the hydrophobic pocket shielded by residues A58-G67.

protein confirmed it had undergone cleavage between Gly67 and Ala68, but retained the N-terminal peptide$^{Linker-G67}$ as deconvolution revealed two

homogenous species of mass 19,620 and 1,402 Da (predicted fragment masses: 19,620 and 1,403 Da) (Figure 5.4C and D). This variant was then crystallised by seeding with microcrystal seeds of PgcA-NHD$^{D208A+D212A}$, and the structure of PgcA-NHD$^{wild\text{-}type}$ was subsequently determined to 1.65 Å resolution. Analysis of this crystal structure reveals it is extremely similar to that of the double substitution variant (RMSD of 0.252 Å) but differs in two additional ways to the substitutions at positions 208 and 212. Firstly, N203 and H204 adopt near inverted conformations, with H204 facing away from the protein into the solvent

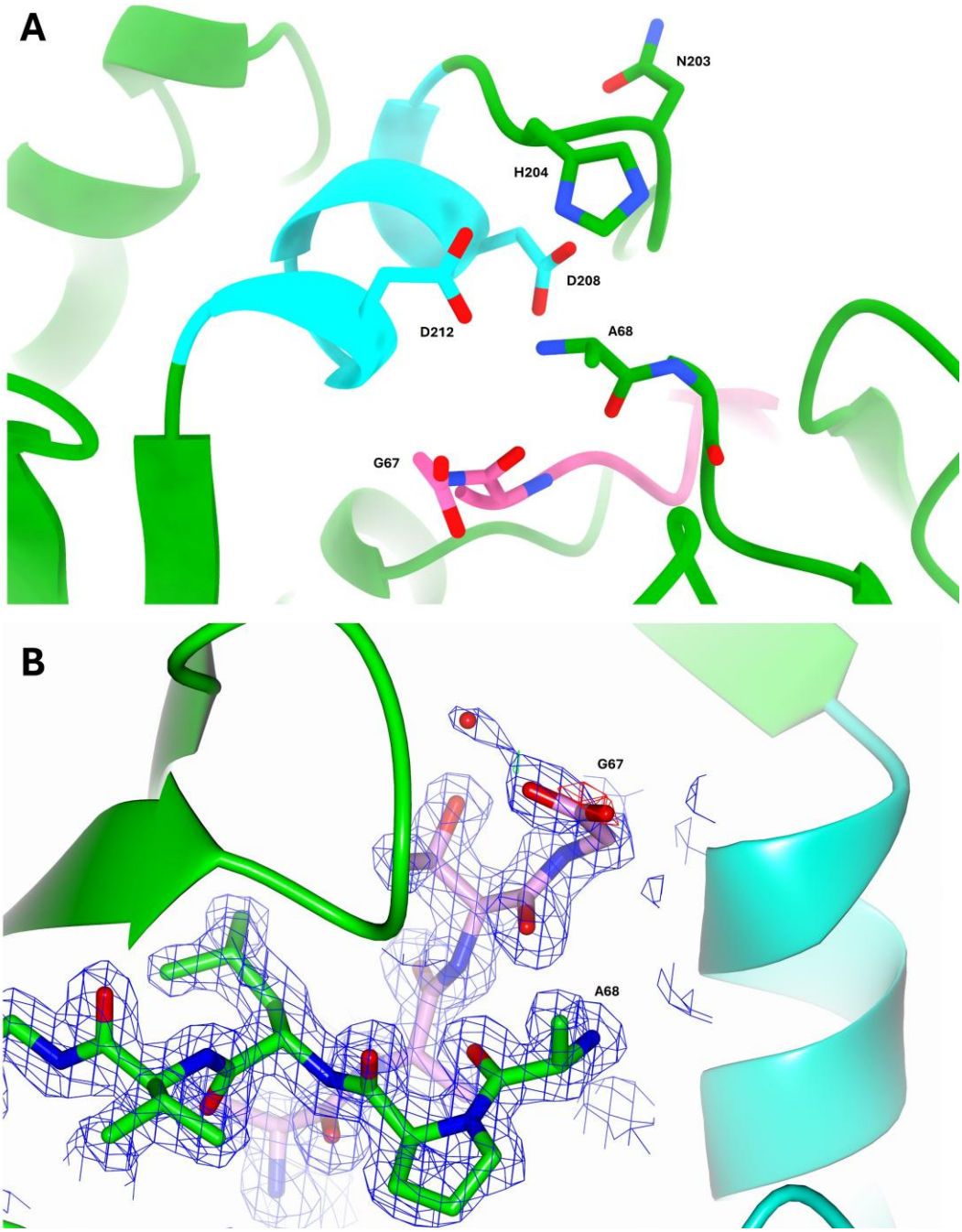


Figure 5.7: Crystal structure of PgcA-NHD$^{wild\text{-}type}$. (A) Crystal structures of PgcA-NHD^{WT} at 1.65 Å resolution. Residues G67, A68, N203, H204, D208 and D212 shown with sticks, and residues A58-G67 coloured pink **(B)** Electron density maps at G67 and A68 positions (α-carbons labelled) contoured at 1.5 and 3.5 sigma for 2Fo-Fc (blue) and Fo-Fc (green/red) respectively.

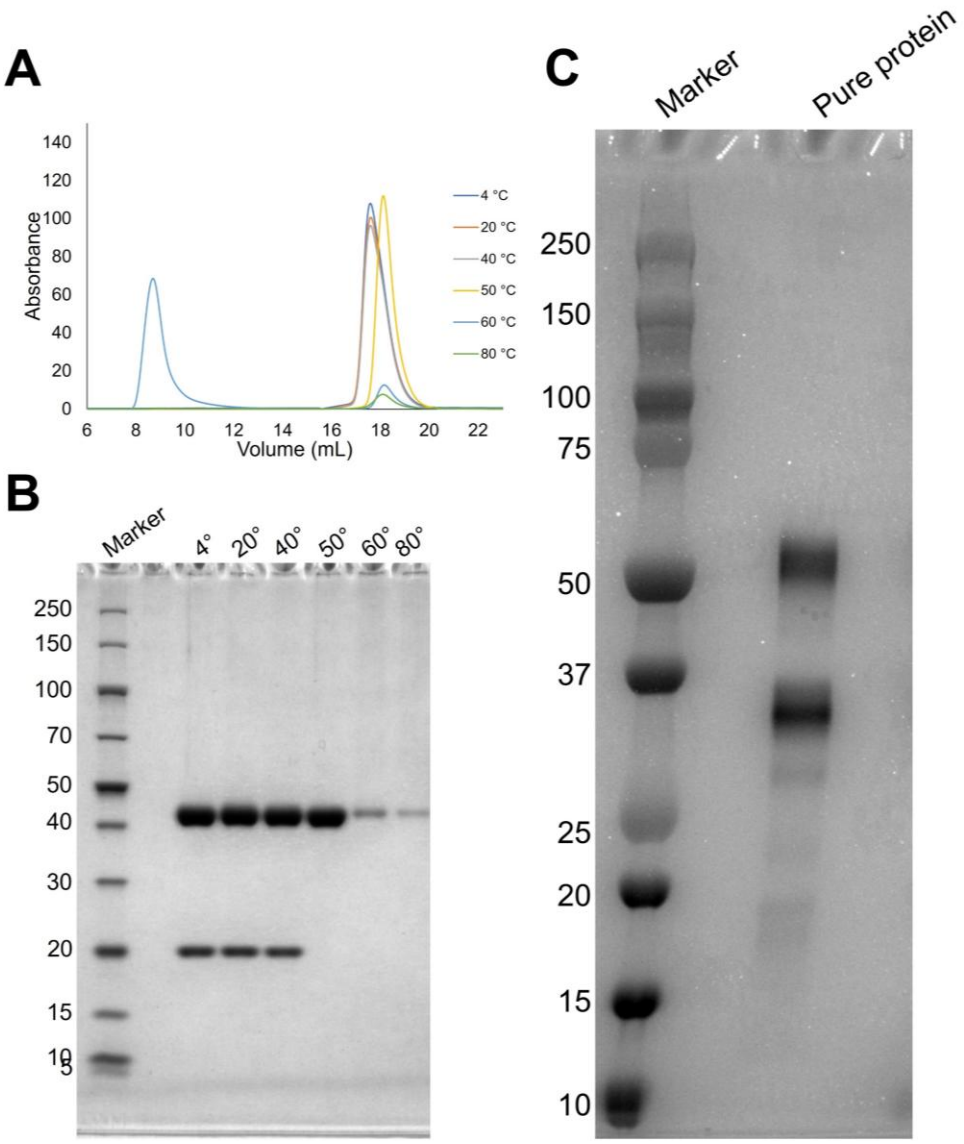


Figure 5.8: Heat induced denaturation of MBP-(PgcA-NHD) heterodimer and D208A+D212A Mutations do not arrest proteolysis of PgcA recombinantly produced in *S. oneidensis*. (A) SEC (superose 6 increase 10/300 GL) 280 nm trace for WT PgcA-NHD after one hour temperature incubations, in 20 mM HEPES, 100 mM NaCl pH 7.8. **(B)** Reducing gradient SDS-PAGE gel of temperature incubation SEC species at 17-19 mL volume stained with Coomassie. **(C)** Heme-stained SDS-PAGE gel of $PgcA^{wild\text{-}type}$, purified by streptactin affinity chromatography from the soluble fraction of overexpressing S. oneidensis cells. All molecular weight markers labelled in kDa.

in the double substitution variant. In the wild-type structure, the conformations are reversed with N203 solvent facing, and H204 hydrogen bonding D208 and D212, possibly contributing to the mechanism of cleavage (Figure 5.7A). Secondly, the wild-type variant electron density maps clearly show discontinuation of the polypeptide chain between G67 and A68 (Figure 5.7B), and G67 adopts a conformation twisted ~90 degrees from its position in the double substitution variant, facing its carboxylate group away from A68 unambiguously supporting the fact that the protein has indeed undergone cleavage (G67 carboxylate-oxygen to A68 amide-nitrogen distances of 5.5 and 5.8 Å).

Refinement of this structure yielded beta factors along the A58-G67 peptide that were close to those of neighbouring residues in the rest of the protein indicating total or near-total retention of this peptide throughout purification and crystallisation, consistent with the LC-MS.

Heat denaturation of PgcA-NHD

Heat was used to attempt dissociation of the MBP-(PgcA-NHD$^{wild\text{-}type}$) heterodimeric protein. Purified protein was incubated for one hour at different temperatures before centrifugation to remove precipitated material. SEC was subsequently performed on the supernatant and 280 nm absorbing species in the subsequent trace were analysed by SDS-PAGE (Figure 5.8A and B). Incubations at 4, 20 and 40 °C yielded a consistent species with retention volume of 17.6 mL that retained both ~20 and ~40 kDa gel bands. Incubations at 60 and 80 °C, mostly destroyed or aggregated both proteins. Only at 50 °C was the heterodimer clearly broken as the ~20 kDa band disappeared entirely (lost as precipitate), without loss of the ~40 kDa band. Consistent with this, the SEC trace maxima shifted to a retention volume of 18.2 mL indicative of the presence of a smaller species: monomeric MBP with the N-terminal fragment PgcA-NHD residues 58-67.

Proteolysis in full length PgcA is not dependent on D208 and D212

Introduction of the D208A and D212A substitutions into the full-length PgcA *S. oneidensis* recombinant expression construct described in Chapter 3, was then investigated. PgcA purified from overexpression of this construct continued to display modification from a ~50 kDa isoform, to a ~30 kDa isoform as previously documented (Figure 5.8C). This process is evidently distinct from the D208/D212 dependent cleavage, and its mechanism remains unclear. It is noted that since PgcA appears to require an N-terminal lipid moiety to be translocated across the outer membrane, if this cleavage observed during recombinant expression in *S. oneidensis* were to occur in *G. sulfurreducens*, it would trap PgcA in the periplasm[5].

Bioinformatic analysis of PgcA-NHD and its homologues

PgcA-NHD (residues 58-252) were protein-BLAST[12] searched in the NCBI non-redundant protein sequences database, which identified hundreds of variants of this domain distributed in hypothetical genes amongst *G. sulfurreducens* and its closest relatives, but also across the Alphaproteobacteria, Gammaproteobacteria and Betaproteobacteria classes including examples in the *Vibrio, Marinobacter, Rhizobacter, Janthinobacterium, Acinitobacter* and *Variovorax* genera. These genes typically display sequence identities in the region 50%-30%, invariably contain PgcA-NHD at their N-terminus and almost all also contain an N-terminal lipidation motif (Figure 5.9A). The alpha helical motif in which D208 and D212 reside is largely conserved across the sequences, usually as Asp-Xaa-Val/Leu-Iso/Leu-Asp (figure 5.9B) with the two aspartates being

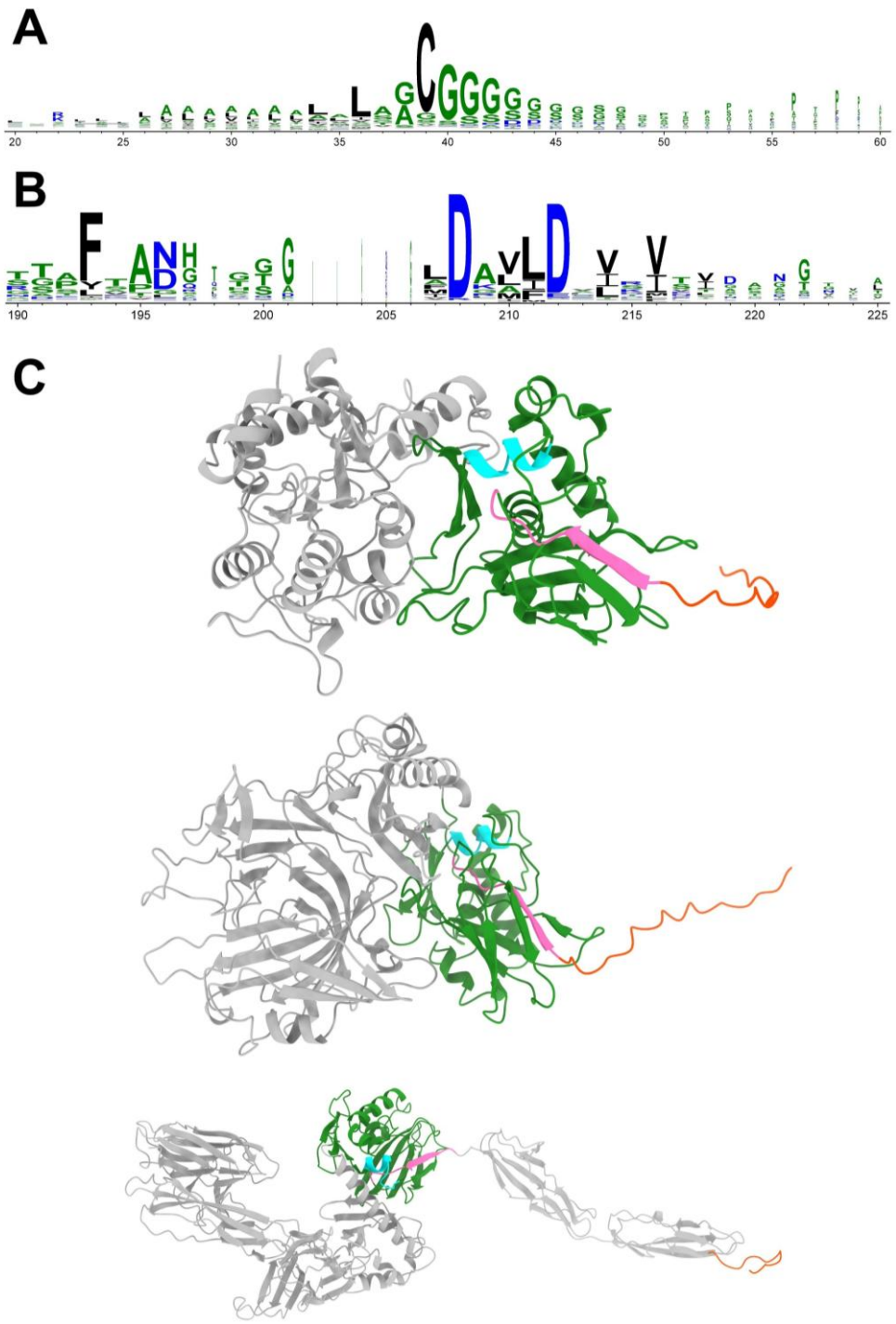


Figure 5.9: PgcA-NHD and its homologues are lipoproteins, residues D208 and D212 are highly conserved, and this domain is distributed amongst diverse genes. (**A**) Consensus sequence and conservation between PgcA-NHD and 958 homologues at lipidation motif locus (residues 36-40). (**B**) Consensus sequence and conservation between PgcA-NHD and 958 homologues at dual-aspartate proteolytic alpha helix (residues 206-213). (**C**) AlphaFold3 model structures generated for GenBank-protein entries HSD38823.1 (top), KPA19545 (middle) and MDG1664003.1 (bottom) annotated as containing sequence homology to GDSL-type lipase/esterase, glycoside hydrolase and carbohydrate-binding domain proteins, respectively. Coloured showing PgcA-NHD homologous regions in green, additional domain(s) in grey, lipid anchor tail in orange, proteolysis fragment analogous sequence in pink and proteolytic alpha helix in cyan.

ubiquitous. Asn203 and His204 feature commonly but are significantly less

conserved than the D208-D212 motif.

Many of these PgcA-NHD homologues contain CXXCH *c*-type heme attachment motifs implicating them in extracellular electron transfer processes, but many include domains at their C-termini which contain no CXXCH motifs, raising the possibility that they do not participate in extracellular electron transfer. Some of the later of those sequences appear homologous to GDSL-lipase/esterase enzymes, glycosidic hydrolase enzymes or carbohydrate binding domains (Figure 5.10C). Additionally, genes with repeated fibronectin type-III domains are numerous suggesting many encode extracellular matrix proteins.

Discussion

G. sulfurreducens secretes several multiheme cytochromes into its extracellular environment that are assembled into filaments that play a crucial role in EET[13]. Of particular relevance is the cytochrome *omcZ*, which is encoded adjacent to a serine protease that cleaves the OmcZ protein extracellularly into a smaller ~30 kDa mature form that unlike the ~50 kDa form, is able to polymerise [14–16].

The Gram negative bacterium *Shewanella oneidensis* MR-1 is also studied intensively for its ability to perform EET. In contrast with *G. sulfurreducens,* the extracellular cytochrome proteins of *S. oneidensis* remain tethered to the outer membrane by their N-terminal lipid moiety which renders them insoluble in the aqueous extracellular environment[17]. Instead *S. oneidensis* secretes flavins which likely can play a comparable role to the cytochrome filaments, helping to move electrons from reduced outer membrane-anchored cytochromes, some distance to terminal electron acceptors[18–20].

G. sulfurreducens is probably not capable of actively secreting flavins to enhance EET in this way[21–23]. It has been suggested that *G. sulfurreducens* might secrete soluble protein(s) to enhance EET instead, however direct evidence in support of this strategy, which would represent an energetically very costly alternative, has not been reported[1,24,25].

The results presented here provide a basis for reconciling two seemingly inconsistent observations, the first that an N-terminal lipid moiety is necessary for secretion of PgcA across the outer-membrane by *G. sulfurreducens*[5]. The second, that PgcA secreted by *G. sulfurreducens* exists as a soluble and truncated isoform[7]. These results show that two aspartate residues accomplish site specific proteolysis, while a network of hydrogen bonding and hydrophobic interactions then ensure the ten residues N-terminal to the cleaved bond are retained by the protein as a heterodimeric complex. This unique mechanism might facilitate release from the cell surface, but would ensure the N-terminus which is required for transport across the outer membrane, is not discarded prematurely.

These results also suggests that the in vivo mature form of PgcA may include most of its non-heme domain, however since it was not possible to obtain either soluble PgcA-NHD with its N-terminal cleavage peptide removed, or remove it from PgcA-NHD$^{wild\text{-}type}$ without denaturing that protein, it is possible that this species is not stable. In that case it may degrade into the triheme PgcA isoform somehow, although the fact that the triheme PgcA forming degradation process is not D208+D212 dependent at least in *S. oneidensis* this does not support that hypothesis.

It remains unclear what might trigger dissociation of residues A58-G67 from their binding pocket to release PgcA from the cell surface *in vivo*. Analysis of the wild-type variant crystal structure using PISA[26] indicates the interaction interface between PgcA-NHD and its cleavage fragment is 797 Å^2 in area, implying it will not easily dissociate. Consistent with this, in this thesis, dissociation was achieved by incubation at 50 °C for one hour which caused total precipitation of PgcA-NHD. Physical strain introduced during transport across the outer-membrane might introduce sufficient lability in the structure for dissociation to occur. Alternatively, a component of the lipopolysaccharide layer, or extracellular environment might trigger dissociation. It is also possible that the rate of dissociation is slow *in vivo*, which might be a valuable feature that leads to a balance between a small population of secreted protein, and a dominant population remaining anchored to the cell surface, thus limiting the energetic cost of the system. This is an area which future research will be able to explore, although it is not a question that is likely to be easily solved since biochemical examination of *G. sulfurreducens* is not trivial.

Methods do exist however for genetic manipulation of *G. sulfurreducens*. Therefore a core goal for this work in the future is to establish whether the D208A and D212A substitutions yield an Fe(III) oxide reduction phenotype in vivo. If the proteolysis plays a critical role in the functioning of PgcA, given *ΔpgcA* cultures exhibit reduced rates of Fe(III) reduction that can be partially restored with a plasmid *pgcA* copy, strains complemented with a plasmid copy containing the D208A and D212A substitutions should display near *ΔpgcA* rates of Fe(III) reduction. Additionally, although preparing sufficient *G. sulfurreducens* biomass for purifications is difficult, it is not difficult to obtain analytical quantities. Therefore SDS-PAGE with LC-MS could be used to identify the native mature form of PgcA in the extracellular fraction, and establish definitively where it is cleaved, if at all, in vivo. Then this experiment could be repeated with a *pgcA* copy that contains the D208A and D212A substitutions. These experiments are ongoing in collaboration with the group of Daniel Bond (University of Minnesota).

Another aspect that is important to establish, will be the exact molecular mechanism of cleavage. Although the conformations of D208 and D212 invoke the active site of aspartic proteases, PgcA-NHD displays no detectable homology (sequence or structural) to that family. For example, the Pepsin (PDB-

1PSN[27]) Cα carbons RMSD with PgcA-NHD is 13.4 Å, and the HIV-1 Protease (PDB-1ODY[28]) Cα carbons RMSD with PgcA-NHD is 11.4 Å. The catalytic mechanism in those enzymes is thought to be initiated via deprotonation of an ordered water molecule, priming it for nucleophilic attack of the substrate peptide[29]. However, no electron density in support of a similar water molecule was resolved in either PgcA-NHD structure described here. Unfortunately, the possibility cannot be precluded that such a water molecule does exist but is present exclusively in the pre-cleavage wild-type structure. Indeed this water does appear to be lost upon substitution of the catalytic aspartates to alanine, in canonical aspartic protease enzymes[30]. The rearrangement of G67 in response to the cleavage could induce the loss/movement of a water. Furthermore, characterised aspartic proteases require both aspartates to retain any activity[31,32], but the PgcA-NHD^{D208A} and PgcA-NHD^{D212A} proteins described in this work still exhibit some autoproteolysis. The conformation of His203 resolved in the PgcA-NHD$^{wild\text{-}type}$ structure suggests its possible involvement, although its comparatively low sequence conservation implies it is unlikely to make a decisive contribution.

Further structural and biochemical studies will evidently be necessary to unpick the molecular basis for cleavage by PgcA-NHD and clarify its relationship with aspartic proteases. What the functional consequences of mutations at the His203 and Asn204 positions are, would be an essential and relatively simple question for future research to address. Also likely to be attainable but still valuable, would be analysis of the effect on proteolysis caused by mutation of the residues either within, or adjacent to, the cleavage site. Those experiments could be complimented by analysis of mutations at residues that are within the proteolysis fragment (A58-G67), or that interact with the fragment. Analysis of the crystal structures described here suggests that the hydrophobic residues beneath the proteolysis fragment are probably important so their removal could weaken the post-cleavage dimer. During the course of this work, it was considered whether PgcA-NHD could serve as a useful tool protein, for example perhaps as a self-cleaving purification tag for example. This possibility was soon eliminated when it was determined that the post-cleavage dimer cannot easily be broken without the use of denaturing conditions, however if a variant could be produced which displays a significantly weakened, but not eliminated, post-cleavage dimer, this strategy could become a viable one.

The reduced but not abolished activity of the D208A and D212A variant proteins suggests crystallographic analysis of these proteins may reveal important insights. The slower rate of proteolysis they display might even allow for trapping of a reaction intermediate, which is often a critical step in elucidation of enzymatic mechanisms[33].

Also crucial will be characterisation of the homologues containing PgcA-NHD that are presented within this thesis. It is essential that their autoproteolytic

activity is experimentally confirmed, and their functions in *G. sulfurreducens* be unambiguously established. If their enzyme domain is indeed secreted from the cell by autoproteolysis in their PgcA-NHD domain, it would be fascinating to see what else this secretory system could be repurposed to efflux, an obvious example would be a distinctive tool protein like the green fluorescent protein, GFP.

The non-heme domain from PgcA is distributed widely in the beta and gamma proteobacteria, attached at its C-terminus to myriad redox and non-redox proteins. Thus it seems likely it is not limited to participating in extracellular electron transfer. Instead it seems PgcA-NHD functions as a general secretory module in *G. sulfurreducens* and related bacteria, conferring export to extracellular enzymes through a unique mechanism.

Methods

Cloning and plasmid construction

The *PgcA-NHD* gene (encoding residues 58-252) was amplified from the PgcA overexpression construct described in Chapter 3, by PCR. The product was digested with SbfI-HF and AlwNI after which it was ligated into the pMAL-c6T plasmid (New England Biolabs) using T4 DNA ligase. The resultant plasmid (pMBP-(PgcA-NHD)) was heat shock transformed into *E. coli* Top10, plasmid DNA was then purified and sequenced (Eurofins) before transformation into *E. coli* BL21 (DE3). To generate mutant variants, PCR mutagenesis was performed using primers as described in Chapter 5 appendix table 5.1, linearised PCR-product DNA was re-circularised with T4 kinase and ligase, mutated DNA was transformed into *E. coli* Top10 before plasmid purification and sequencing (Eurofins), followed by transformation into BL21 (DE3).

MBP-(PgcA-NHD) purification

Single colonies of *E. coli* BL21 pMBP-(PgcA-NHD) were picked into 100 mL LB, 50 μg mL^{-1} carbenicillin and cultured overnight at 30 °C, 180 RPM. These cultures were used to inoculate 1 L LB (50 μg mL^{-1} carbenicillin) in 2 L baffled flasks, which were cultured at 37 °C, 180 RPM until an OD_{600} of approximately 0.6 was reached, at which point they were induced with 0.5 mM IPTG before being returned to the incubator for two hours. Cells were then harvested by centrifugation and resuspended in cold buffer A (100 mM Tris-HCl, 150 mM NaCl pH 8.0). DNAse was added before cell lysis by double pass through a french press at 1000 PSI, lysate was then clarified by ultracentrifugation and the pellet discarded. The supernatant was then loaded onto a 5 mL maltodextrin affinity column equilibrated in buffer A, before washing with 15 column volumes (CV) of buffer A. Protein was then eluted by 6 CV of buffer A supplemented to 10 mM maltose.

Heat dissociation of MBP-(PgcA-NHD)

To trigger dissociation of the MBP-(PgcA-NHD$^{wild\text{-}type}$) species, protein at 1 mg mL^{-1} in 20 mM HEPES pH 7,8, 100 mM NaCl (buffer B) was incubated for one hour at 4, 20, 40, 50, 60 and 80 °C. The resultant material was centrifuged (4000 G, 3 minutes) before injecting 0.5 mL volume into a Superose 6 increase 10/300 GL SEC column, equilibrated in buffer C, with a flow rate of 0.5 mL min^{-1} at 4 °C. Fractions with 280 nm absorbing species in the SEC trace were analysed by SDS-PAGE and stained with Coomassie blue.

PgcA full length construct post-translational-modification analysis

PgcA$^{D208A+D212A}$ was purified from a full-length overexpression construct transformed into *S. oneidensis*, by Streptactin-affinity chromatography as described in *Chapter 3 methods*. Protein purity was assessed by SDS-PAGE.

TEV protease purification

A TEV protease overexpression strain was kindly provided by Zinnia Bugg (University of East Anglia). Single colonies of pTEV *E. coli* BL21 were picked into 100 mL LB (30 µg mL^{-1} kanamycin), which were grown overnight (180 RPM, 30 °C). These flasks were used to inoculate (1% final) 2 x 1 L LB media in 2L baffled flasks (30 µg mL^{-1} kanamycin). The 2L flasks were cultured at 37 degrees 180 RPM for 2.5 hours when an OD_{600} of 0.68 was reached. 1 mM (final) sterile IPTG added, and flasks returned to incubator for ~two hours. Cells were then harvested by centrifugation (4000 G, 20 minutes ,16 °C). The cells were resuspended in 20 mL of cold 50 mM phosphate buffer, 250 mM NaCl, 10 mM imidazole pH 7.5 before lysing by double pass through French press at 1000 PSI. The lysate was diluted with 50 mM phosphate buffer, 500 mM NaCl, 10 mM imidazole pH 7.5 (buffer C) to a volume of ~60 mL, then clarified by ultracentrifugation (200,000 G, 4 °C, 1.5 hours). The clarified lysate was loaded onto Ni^{2+} IMAC column equilibrated in buffer C, at 2.5 mL min^{-1}. The column was then washed with 12.5 CV of buffer C (5 mL min^{-1}), followed by further washing with 7.5 CV of 50 mM phosphate buffer, 500 mM NaCl, 70 mM imidazole pH 7.5 (5 mL min^{-1}). Purified protein was eluted with 8 CV of 50 mM phosphate buffer, 500 mM NaCl, 250 mM imidazole pH 7.5 (5 mL min^{-1}).

Purified TEV protease was then pooled and concentrated in 10 kDa centrifugal concentrators until a volume of ~1 mL was reached. This was injected into a 16/600 Superdex 75 PG SEC column equilibrated in buffer B (20 mM HEPES pH 7.8, 100 mM NaCl), and eluted by application of further buffer B (0.5 mL min^{-1}). Fractions were analysed by SDS-PAGE (Chapter 5 appendix figure 1), pooled and protein concentrated to 5.6 mg mL^{-1}. To this β-mercaptoethanol ethanol and glycerol were added to final concentrations of 25 mM and 20% respectively before flash freezing in liquid nitrogen and storage at -80 °C until use.

Liquid chromatography mass spectrometry

For LC-MS, protein was prepared by concentration to 100 µM or higher in 20 mM Tris-HCl pH 7.5 before dilution to ~10 µM with HPLC/LC-MS grade water. Then they were loaded onto an ACQUITY UPLC Protein BEH C4 Column, 300 Å, 1.7 µm, 2.1 mm x 50 mm with an Acquity Premier I class UPLC. Bound proteins were eluted (0.2 ml min^{-1}) (linear gradient (20 min) from 2% to 98% (v/v) acetonitrile, 0.1% (v/v) formic acid). The eluent was infused into a Waters Synapt XS QTOF mass spectrometer with IMS running MassLynx, using positive mode electrospray ionisation. The MS was calibrated with sodium formate (up to 2000 m/z) or sodium iodide for higher masses. UniDec (Universal Deconvolution) software[34] was used for processing of spectra using Bayesian deconvolution to assign the charge states of the m/z peaks.

PgcA-NHD isolation from MBP, crystallisation and structure determination

MBP-(PgcA-NHD) at ~10 mg mL^{-1} or higher was incubated overnight at 20 °C with 5 mM beta-mercaptoethanol and ~100:1 TEV-protease. The digest was then injected through two 5 mL freshly charged Ni^{2+}-IMAC columns equilibrated in 50 mM sodium phosphate, 500 mM NaCl, 25 mM imidazole, pH 7.5 and washed through with additional 500 mM NaCl, 25 mM imidazole, pH 7.5. Flowthrough fractions were analysed for PgcA-NHD purity by SDS-PAGE, pooled and concentrated in 10 kDa spin concentrators to ~1 mL. This was then injected into a 16/600 Superdex 75 PG SEC column equilibrated in buffer B. Fractions were then analysed for PgcA-NHD by SDS-PAGE and pooled before concentration to ~10 mg mL^{-1} assuming a 280 nm extinction coefficient predicted by ExPASy-*ProtParam*[35].

Crystalline material was obtained in vapour-diffusion drops of 0.3 µL of protein solution, 0.28 µL of 1.5 M ammonium sulfate, 12 % v/v glycerol, 100 mM Tris-HCl pH 8.5 (buffer D), and 0.02 µL of 10 % w/v PEG 8000, 8 % v/v ethylene glycol, 100 mM HEPES pH 7.5 equilibrated against 50 µL buffer D mother liquor. This material was crushed manually with a 20 µL pipette tip and resuspended in mother liquor before being used to seed vapour-diffusion drops of 0.3 µL protein, 0.25 µL buffer D and 0.05 µL seed stock equilibrated against 50 µL buffer D, which yielded oval shaped crystals overnight. These were transferred to buffer D supplemented to 25% v/v glycerol for cryoprotection before vitrification by plunging into liquid nitrogen.

Data was collected on single crystals at Diamond Light Source beamlines I04 and I24 using X-rays of wavelengths of 0.9537 Å. Datasets from single crystals were indexed, scaled and merged automatically with XIA2[36] and phased by molecular replacement (Phaser-MR[37]) with an AF2[11] search model. The structures were refined manually using WinCoot[38] and using Phenix.refine[39], to final resolutions of 1.65 Å, with no Ramachandran outliers. Data collection and refinement statistics are shown in chapter 5 appendix table 5.2 and electron density maps are shown in chapter 5 appendix figure 3.

Bioinformatic analysis

PgcA-NHD residues (residues 58-252) was protein-BLAST[12] searched to obtain 1000 sequences (table S3) with default parameters, and sequences with more than 95% similarity discarded. The remaining sequences were sequence-aligned by ClustalW[40] using default parameters in Jalview[41]. WebLogo-3[42] was used to generate a consensus sequence plot for the resultant alignment. GenBank-protein entries HSD38823.1, KPA19545 and MDG1664003.1 were selected due to their auto-annotation as homologues of enzymes of interest, and their sequences uploaded to the AlphaFold3[43] webserver for structure prediction. Their signal-peptide sequence, according to signalP-5.0[44] was then deleted from the structural model.

Chapter 5: References

1. Zacharoff, L. A., Morrone, D. J. & Bond, D. R. Geobacter sulfurreducens extracellular multiheme cytochrome PgcA facilitates respiration to Fe(III) oxides but not electrodes. *Frontiers in Microbiology* **8**, 300906 (2017).

2. Ding, Y. H. R., Hixson, K. K., Aklujkar, M. A., Lipton, M. S., Smith, R. D., Lovley, D. R. & Mester, T. Proteome of Geobacter sulfurreducens grown with Fe(III) oxide or Fe(III) citrate as the electron acceptor. *Biochimica et Biophysica Acta (BBA) - Proteins and Proteomics* **1784**, 1935–1941 (2008).

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

4. Narita, S. ichiro & Tokuda, H. Bacterial lipoproteins; biogenesis, sorting and quality control. *Biochimica et Biophysica Acta (BBA) - Molecular and Cell Biology of Lipids* **1862**, 1414–1423 (2017).

5. Choi, S., Chan, C. H. & Bond, D. R. Lack of Specificity in Geobacter Periplasmic Electron Transfer. *Journal of Bacteriology* **204**, (2022).

6. Naskar, S., Hohl, M., Tassinari, M. & Low, H. H. The structure and mechanism of the bacterial type II secretion system. *Molecular Microbiology* **115**, 412–424 (2021).

7. Smith, J. A., Tremblay, P. L., Shrestha, P. M., Snoeyenbos-West, O. L., Franks, A. E., Nevin, K. P. & Lovley, D. R. Going Wireless: Fe(III) Oxide Reduction without Pili by Geobacter sulfurreducens Strain JS-1. *Applied and Environmental Microbiology* **80**, 4331 (2014).

8. Fernandes, T. M., Silva, M. A., Morgado, L. & Salgueiro, C. A. Hemes on a string: insights on the functional mechanisms of PgcA from Geobacter sulfurreducens. *Journal of Biological Chemistry* **299**, (2023).

9. Nash, B. W., Fernandes, T. M., Burton, J. A. J., Morgado, L., Wonderen, J. H. van, Svistunenko, D. A., Edwards, M. J., Salgueiro, C. A., Butt, J. N. & Clarke, T. A. Tethered

heme domains in a triheme cytochrome allow for increased electron transport distances. *Protein Science* **33**, e5200 (2024).

10. Kapust, R. B. & Waugh, D. S. Escherichia coli maltose-binding protein is uncommonly effective at promoting the solubility of polypeptides to which it is fused. *Protein Science : A Publication of the Protein Society* **8**, 1668 (1999).

11. Jumper, J., Evans, R., Pritzel, A., Green, T., Figurnov, M., Ronneberger, O., Tunyasuvunakool, K., Bates, R., Žídek, A., Potapenko, A., Bridgland, A., Meyer, C., Kohl, S. A. A., Ballard, A. J., Cowie, A., Romera-Paredes, B., Nikolov, S., Jain, R., Adler, J. *et al*. Highly accurate protein structure prediction with AlphaFold. *Nature 2021 596:7873* **596**, 583–589 (2021).

12. Altschul, S. F., Gish, W., Miller, W., Myers, E. W. & Lipman, D. J. Basic local alignment search tool. *Journal of molecular biology* **215**, 403–410 (1990).

13. Wang, F., Craig, L., Liu, X., Rensing, C. & Egelman, E. H. Microbial nanowires: type IV pili or cytochrome filaments? *Trends in Microbiology* **31**, 384–392 (2023).

14. Kai, A., Tokuishi, T., Fujikawa, T., Kawano, Y., Ueki, T., Nagamine, M., Sakakibara, Y., Suiko, M. & Inoue, K. Proteolytic Maturation of the Outer Membrane c-Type Cytochrome OmcZ by a Subtilisin-Like Serine Protease Is Essential for Optimal Current Production by Geobacter sulfurreducens. *Applied and Environmental Microbiology* **87**, 1–10 (2021).

15. Gu, Y., Guberman-Pfeffer, M. J., Srikanth, V., Shen, C., Giska, F., Gupta, K., Londer, Y., Samatey, F. A., Batista, V. S. & Malvankar, N. S. Structure of Geobacter cytochrome OmcZ identifies mechanism of nanowire assembly and conductivity. *Nature Microbiology 2023 8:2* **8**, 284–298 (2023).

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

17. van Wonderen, J. H., Crack, J. C., Edwards, M. J., Clarke, T. A., Saalbach, G., Martins, C. & Butt, J. N. Liquid-chromatography mass spectrometry describes post-translational modification of Shewanella outer membrane proteins. *Biochimica et Biophysica Acta (BBA) - Biomembranes* **1866**, 184221 (2024).

18. Von Canstein, H., Ogawa, J., Shimizu, S. & Lloyd, J. R. Secretion of flavins by Shewanella species and their role in extracellular electron transfer. *Applied and Environmental Microbiology* **74**, 615–623 (2008).

19. Kotloski, N. J. & Gralnick, J. A. Flavin electron shuttles dominate extracellular electron transfer by Shewanella oneidensis. *mBio* **4**, 553–565 (2013).

20. Marsili, E., Baron, D. B., Shikhare, I. D., Coursolle, D., Gralnick, J. A. & Bond, D. R. Shewanella secretes flavins that mediate extracellular electron transfer. *Proceedings of the National Academy of Sciences of the United States of America* **105**, 3968–3973 (2008).

21. Straub, K. L. & Schink, B. Evaluation of electron-shuttling compounds in microbial ferric iron reduction. *FEMS Microbiology Letters* **220**, 229–233 (2003).

22. Bond, D. R. & Lovley, D. R. Evidence for involvement of an electron shuttle in electricity generation by Geothrix fermentans. *Applied and Environmental Microbiology* **71**, 2186–2189 (2005).

23. Gralnick, J. A. & Bond, D. R. Electron Transfer beyond the Outer Membrane: Putting Electrons to Rest. *Annual Review of Microbiology* **77**, 517–539 (2023).

24. Seeliger, S., Cord-Ruwisch, R. & Schink, B. 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. *Journal of Bacteriology* **180**, 3686–3691 (1998).

25. Lloyd, J. R., Blunt-Harris, E. L. & Lovley, D. R. The periplasmic 9.6-kilodalton c-type cytochrome of Geobacter sulfurreducens is not an electron shuttle to Fe(III). *Journal of Bacteriology* **181**, 7647–7649 (1999).

26. Krissinel, E. & Henrick, K. Inference of Macromolecular Assemblies from Crystalline State. *Journal of Molecular Biology* **372**, 774–797 (2007).

27. Fujinaga, M., Chernaia, M. M., Mosimann, S. C., James, M. N. G. & Tarasova, N. I. Crystal structure of human pepsin and its complex with pepstatin. *Protein Science* **4**, 960–972 (1995).

28. Kervinen, J., Lubkowski, J., Zdanov, A., Bhatt, D., Dunn, B. M., Hui, K. Y., Powell, D. J., Kay, J., Wlodawer, A. & Gustchina, A. Toward a universal inhibitor of retroviral proteases: Comparative analysis of the interactions of LP-130 complexed with proteases from HIV-1, FIV, and EIAV. *Protein Science* **7**, 2314–2323 (1998).

29. Suguna, K., Padlan, E. A., Smith, C. W., Carlson, W. D. & Davies, D. R. Binding of a reduced peptide inhibitor to the aspartic proteinase from Rhizopus chinensis: implications for a mechanism of action. *Proceedings of the National Academy of Sciences of the United States of America* **84**, 7009 (1987).

30. Sherry, D., Pandian, R. & Sayed, Y. Non-active site mutations in the HIV protease: Diminished drug binding affinity is achieved through modulating the hydrophobic sliding mechanism. *International Journal of Biological Macromolecules* **217**, 27–41 (2022).

31. Kohl, N. E., Emini, E. A., Schleif, W. A., Davis, L. J., Heimbach, J. C., Dixon, R. A. F., Scolnick, E. M. & Sigal, I. S. Active human immunodeficiency virus protease is required for viral infectivity. *Proceedings of the National Academy of Sciences* **85**, 4686–4690 (1988).

32. Li, X., Dang, S., Yan, C., Gong, X., Wang, J. & Shi, Y. Structure of a presenilin family intramembrane aspartate protease. *Nature 2013 493:7430* **493**, 56–61 (2012).

33. Weber, I. T., Agniswamy, J., Fu, G., Shen, C. H. & Harrison, R. W. Reaction Intermediates Discovered in Crystal Structures of Enzymes. *Advances in Protein Chemistry and Structural Biology* **87**, 57–86 (2012).

34. Marty, M. T., Baldwin, A. J., Marklund, E. G., Hochberg, G. K. A., Benesch, J. L. P. & Robinson, C. V. Bayesian deconvolution of mass and ion mobility spectra: From binary interactions to polydisperse ensembles. *Analytical Chemistry* **87**, 4370–4376 (2015).

35. Wilkins, M. R., Gasteiger, E., Bairoch, A., Sanchez, J. C., Williams, K. L., Appel, R. D. & Hochstrasser, D. F. Protein identification and analysis tools in the ExPASy server. *Methods in molecular biology (Clifton, N.J.)* **112**, 531–552 (1999).

36. Winter, G. xia2: an expert system for macromolecular crystallography data reduction. *Journal of Applied Crystallography* **43**, 186–190 (2010).

37. McCoy, A. J., Grosse-Kunstleve, R. W., Adams, P. D., Winn, M. D., Storoni, L. C. & Read, R. J. Phaser crystallographic software. *urn:issn:0021-8898* **40**, 658–674 (2007).

38. Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. Features and development of Coot. *urn:issn:0907-4449* **66**, 486–501 (2010).

39. Afonine, P. V., Grosse-Kunstleve, R. W., Echols, N., Headd, J. J., Moriarty, N. W., Mustyakimov, M., Terwilliger, T. C., Urzhumtsev, A., Zwart, P. H. & Adams, P. D. Towards automated crystallographic structure refinement with phenix.refine. *urn:issn:0907-4449* **68**, 352–367 (2012).

40. Larkin, M. A., Blackshields, G., Brown, N. P., Chenna, R., Mcgettigan, P. A., McWilliam, H., Valentin, F., Wallace, I. M., Wilm, A., Lopez, R., Thompson, J. D., Gibson, T. J. & Higgins, D. G. Clustal W and Clustal X version 2.0. *Bioinformatics* **23**, 2947–2948 (2007).

41. Waterhouse, A. M., Procter, J. B., Martin, D. M. A., Clamp, M. & Barton, G. J. Jalview Version 2—a multiple sequence alignment editor and analysis workbench. *Bioinformatics* **25**, 1189–1191 (2009).

42. Schneider, T. D. & Stephens, R. M. Sequence logos: a new way to display consensus sequences. *Nucleic Acids Research* **18**, 6097 (1990).

43. Abramson, J., Adler, J., Dunger, J., Evans, R., Green, T., Pritzel, A., Ronneberger, O., Willmore, L., Ballard, A. J., Bambrick, J., Bodenstein, S. W., Evans, D. A., Hung, C.-C., O'Neill, M., Reiman, D., Tunyasuvunakool, K., Wu, Z., Žemgulytė, A., Arvaniti, E. *et al*. Accurate structure prediction of biomolecular interactions with AlphaFold 3. *Nature 2024* 1–3 (2024) doi:10.1038/s41586-024-07487-w.

44. Almagro Armenteros, J. J., Tsirigos, K. D., Sønderby, C. K., Petersen, T. N., Winther, O., Brunak, S., von Heijne, G. & Nielsen, H. SignalP 5.0 improves signal peptide predictions using deep neural networks. *Nature Biotechnology 2019 37:4* **37**, 420–423 (2019).

Chapter 5: Appendix

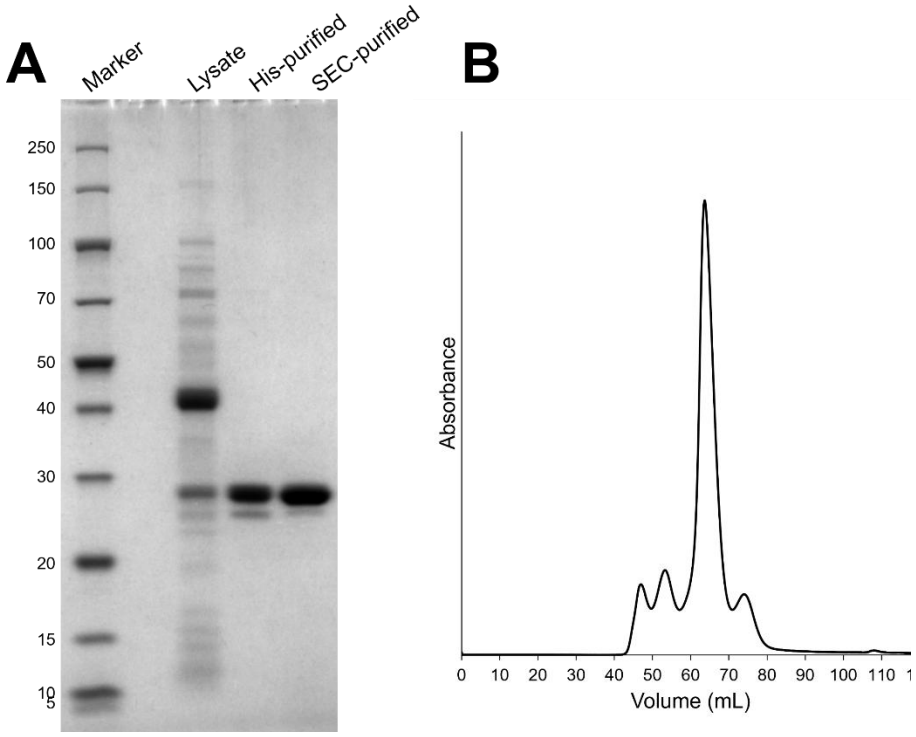
Table 5.1: Primers and plasmids used in Chapter 5.

Plasmids	Products	Primers (5'-3')
pMBP-(PgcA-NHD)	MBP-(PgcA-NHD^{WT})	Forward: AACAGATGCTGGCAGTATCAGGGGTGGCA Reverse: TTCCTGCAGGTCACTTCTCGAATTGTGGATGGC
pMBP-(PgcA-NHD^{D208A}) pBAD202-PgcA^{D208A}	MBP-(PgcA-NHD^{D208A}) D208A-PgcA (*S. oneidensis*)	Forward: GGTCTCGCCGCCGTTCTGGACGTA Reverse: AACGGCGGCGAGACCAGTGTG
D212A pMBP-(PgcA-NHD)	MBP-(PgcA-NHD^{D212A})	Forward: TTCTGGCCGTAATCCGGGTTCAACTTGGT Reverse: ATTACGGCCAGAACGGCGTCGAGA
pMBP-(PgcA-NHD$^{D208A+D212A}$) pBAD202-PgcA$^{D208A+D212A}$	MBP-(PgcA-NHD$^{D208A+D212A}$) PgcA$^{D208A+D212A}$ (*S. oneidensis*)	Forward: TTCTGGCCGTAATCCGGGTTCAACTTGGT Reverse: ATTACGGCCAGAACGGCGGCGAGA
pMBP-(PgcA-NHD)-(linker) pMBP-(PgcA-NHD$^{D208A+D212A}$)-(linker)	MBP-(PgcA-NHD^{WT}) (linker) MBP-(PgcA-NHD$^{D208A+D212A}$) (linker)	Forward: GCGGCTGCCGCAGTATCAGGGGTGGCAG Reverse: CGCAGCGGCCAGCATCTGGAAGTACAGGTTCTC
pMBP-(PgcA-NHD$^{\Delta 59-69\ D208A+D212A}$)	MBP-(PgcA-NHD$^{D208A+D212A\ \Delta 59-69}$)	Forward: CTTGTGGGCACGATCCGGCTCAAGG Reverse: GGCAGCCGCCGCAGCGGC
pTEV	TEV protease	N/A

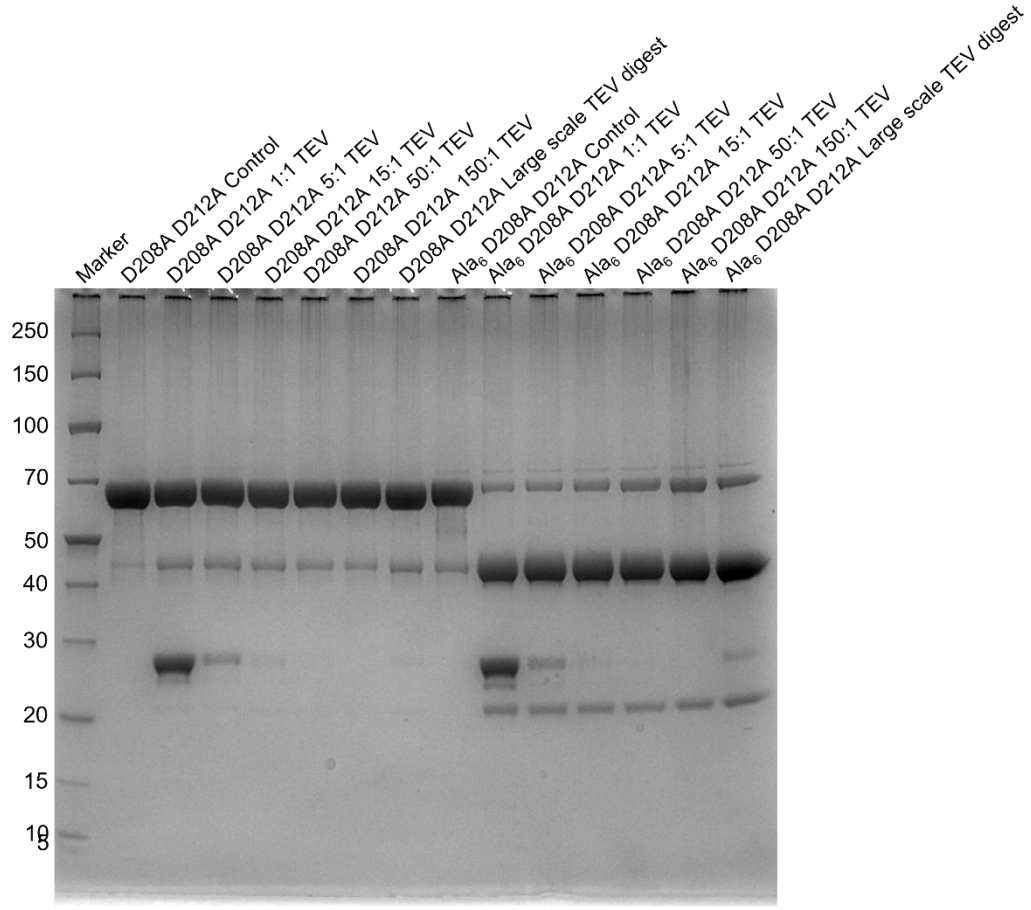
Table 5.2: Data collection and refinement statistics for PgcA-NHD crystals

	PgcA-NHD$^{D208A+D212A}$	PgcA-NHD^{WT}
Data collection		
Space group	P 61	P 61
Cell dimensions		
a, b, c (Å)	96.70 96.70 33.15	96.76, 96.76, 35.92
α, β, γ (°)	90 90 120	90 90 120

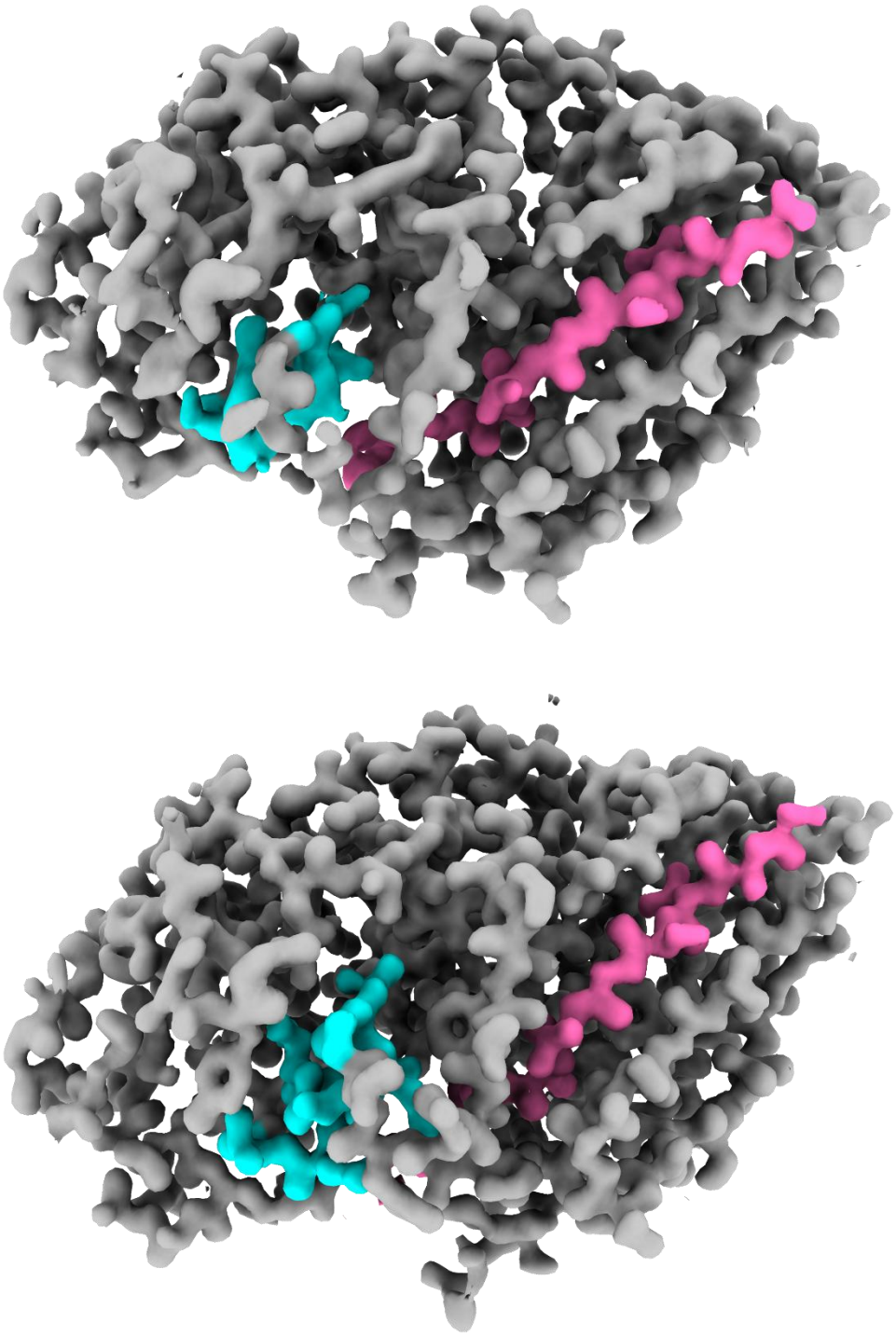
Resolution (Å)	48.35 -1.65 (1.71 -1.65)	41.9-1.65 (1.71-1.65)
R_{pim} (%) $CC_{1/2}$ (%)	0.013 (0.101) 1 (0.98)	0.0089 (0.037) 1 (0.99)
$I / \sigma I$	32.04 (4.03)	56.15 (13.12)
Completeness (%)	99.95 (99.76)	99.99 (100.00)
Multiplicity	20.3 (18.0)	18.7 (11.2)
Refinement		
Resolution (Å)	1.65	1.65
No. reflections	21608 (2121)	23412 (2324)
R_{work} / R_{free}	0.1815 / 0.2044	0.166 / 0.191
No. atoms		
Protein	1396	1403
Ligand/ion	21	21
Water	180	225
B-factors		
Protein	22.59	14.76
Ligand/ion	43.06	38.02
Water	34.95	28.69
R.m.s. deviations		
Bond lengths (Å)	0.006	0.008
Bond angles (°)	0.94	1.05



Appendix figure 1: Purification of the TEV protease tool protein. (A) Reducing SDS-PAGE analysis of TEV protease purification intermediates. **(B)** Final size exclusion chromatography trace of TEV protease, note that fractions between 60-70 mL were pooled, this sample is shown in panel A as SEC-purified.



Appendix figure 2: Recombinant TEV is active against PgcA-NHD but requires a hexa-alanine linker for high efficiency. Reducing SDS-PAGE analysis of TEV protease incubations with PgcA-NHD at different ratios using the D208A+D212A variant and a D208A+D212A variant with an additional hexa-alanine linker (Ala_6), marker molecular weight shown in kDa. Note the band at ~27 kDa is the TEV enzyme.



Appendix figure 3: PgcA-NHD crystal structure electron density maps. Double difference electron density maps for 1.65 Å (both) resolution PgcA-NHD D208A+D212A variant (top) and wild-type (bottom), contoured at 1.5 sigma. Density arising from N-terminal cleavage fragment represented in pink and from D208-D212 alpha helix in cyan.

Chapter 6: The structural basis for extracellular reduction of DMSO by *Shewanella oneidensis* MR-1

Introduction

Shewanella oneidensis and other Gram-negative bacteria which participate in extracellular electron transfer assemble porin-cytochrome complexes in their outer membrane which play a pivotal role in that process[1,2]. These multimeric protein complexes facilitate rapid transmembrane electron transport and can also directly interact with solid and solvated extracellular substrates[3,4].

The archetypal porin-cytochrome complex, which has been extensively studied as a model system, is the MtrCAB complex present in many *Shewanella* species including *S. oneidensis* and *S. baltica*[5]. A crystal structure was recently reported for the *S. baltica* MtrCAB complex, which showed it contains a porin, MtrB, that insulates a decaheme *c*-type cytochrome, MtrA, from the lipidic-membrane and an extracellular head group, the decaheme cytochrome MtrC[6].

Unfortunately, with only a single porin-cytochrome complex structure reported in the literature, it remains unclear which features of the *S. baltica* MtrCAB complex are conserved among analogous complexes, and which are unique. Since attempts began to structurally resolve MtrCAB by X-ray crystallography, enormous progress in structure determination by single particle cryoEM has been made, and this technique now represents the cutting edge in membrane protein structural biology[7–9].

S. oneidensis contains, in addition to *mtrCAB*, another gene cluster that contains genes encoding a porin, *dmsF*, and a decaheme *c*-type cytochrome, *dmsE*. Within this cluster, no homologue of the *mtrC* gene is present, in its place are two genes homologous to the *E. coli* periplasmic DMSO reductase subunits, these being *dmsA* and *dmsB*[10]. Deletion of *dmsB* and *dmsE* has been shown to cause an inability to grow on DMSO, consistent with it forming a DMSO reducing complex, and more recently it has been shown to confer iodate reduction[10,11]. However, all structurally characterised DMSO reductase enzymes are intracellularly localised. Purification of enzymatically active DMSO reductase from *S. oneidensis* has been recently reported, but this preparation did not display any features indicating the porin and cytochrome subunits had copurified with DmsA and DmsB[12].

This chapter sought to develop a method to purify the complete DMSO reductase (DMSOR) from *S. oneidensis* MR-1 and determine its three dimensional structure

using cryogenic electron microscopy. The DmsEFAB structure described here represents the first of an extracellular DMSO reductase, but more importantly also the first porin-cytochrome complex structure with a well-defined substrate binding site. Thus these results provide important insights for the field of EET, and could serve as a basis for engineering extracellular-enzyme based redox chemistry.

Results

Purification of DmsEFAB

A *S. oneidensis* MR-1 strain carrying the *dmsEFABGH* gene cluster in a pBAD202 plasmid (pDMSOR) was kindly provided by Dr Marcus Edwards. This construct included an element encoding for a strep-II purification tag incorporated at the DmsA C-terminus. Plasmid DNA was purified and sequenced to confirm the identity of the strain and integrity of the genes.

The *dmsEFABGH* gene cluster was overexpressed in *S. oneidensis* MR-1 LS527, and purified in detergent, as described in *Chapter 6 methods*. *Methodology I*

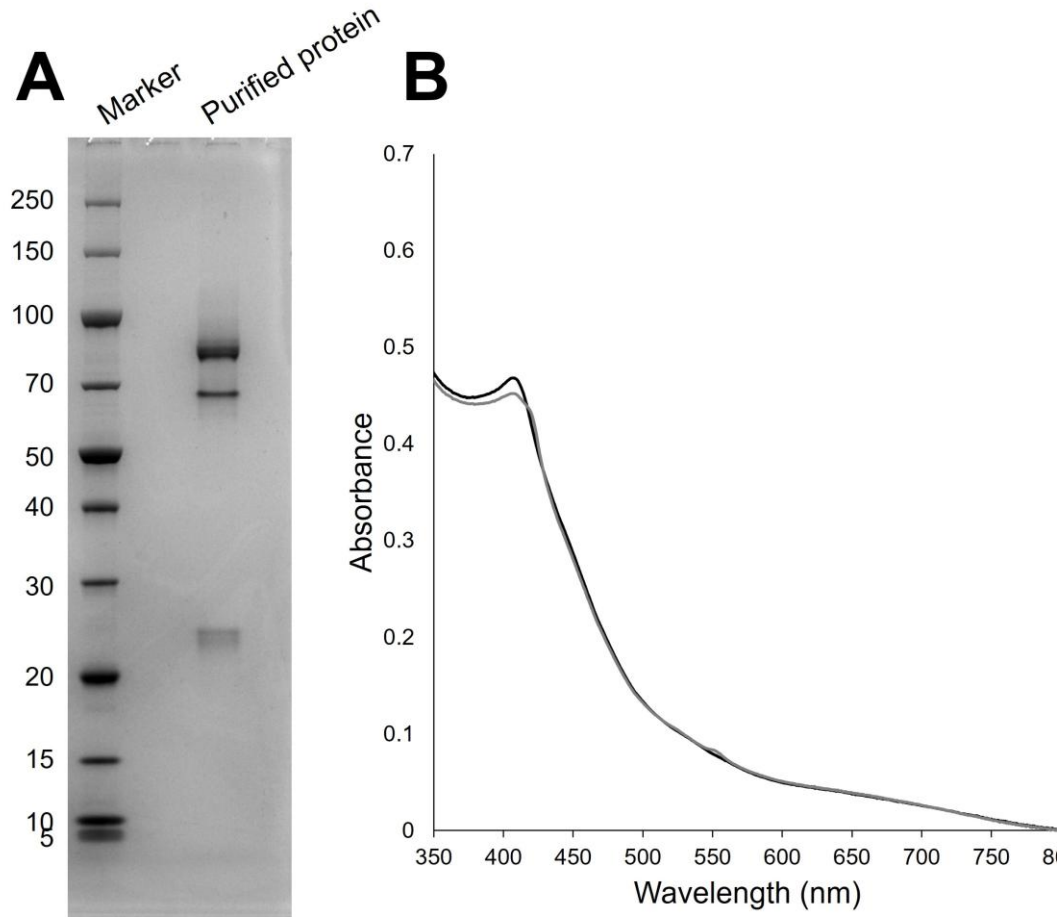


Figure 6.1: Initial attempts to purify DMSOr yielded an incomplete complex. (A) Reducing SDS-PAGE analysis of DMSOR purified using *Methodology I*, marker molecular weight shown in kDa. **(B)** Electronic absorbance spectra of as-purified (black) and dithionite reduced DMSOR purified using *Methodology I*.

yielded a brownish-green solution consistent with that previously reported was purified via this method[12]. Coomassie stained SDS-PAGE revealed the presence of bands at ~90 kDa consistent with the DmsA predicted molecular weight, ~25 kDa consistent with DmsB. A third band was identified at ~65 kDa which could indicate the presence of DmsF although no band consistent with ~35 kDa DmsE was visible. The electronic absorbance spectrum was inconsistent with the presence of ten *c*-type hemes which, due to their high extinction coefficient, should dominate in the ultraviolet-visible region (Figure 6.1). These results suggested some subunits from the complex had been lost during purification. When interpreted alongside the observations that 1: In the homologous MtrCAB system, MtrA is required for MtrB to be stable[13] and 2: DSMO reductase enzymes can create additional bands at lower molecular weight if denatured incompletely[14], the most likely explanation is that this material was comprised of DmsAB complexes, and the band at ~65 kDa is incompletely denatured DmsA.

Subsequently, two alterations to the purification process were introduced. First, 10% v/v glycerol was introduced into the lysis buffer, and second, the membrane pellet was solubilised in 1.5% LDAO instead of 1.5% Triton X-100. SEC of protein purified with *methodology II* separated two species, a smaller dominant species similar to that obtained via *methodology I*, and a larger species with lower yield, a clear red colour and a positive 410-280 nm absorbance ratio (Figure 6.2A). Coomassie stained SDS-PAGE analysis of the larger species revealed four bands

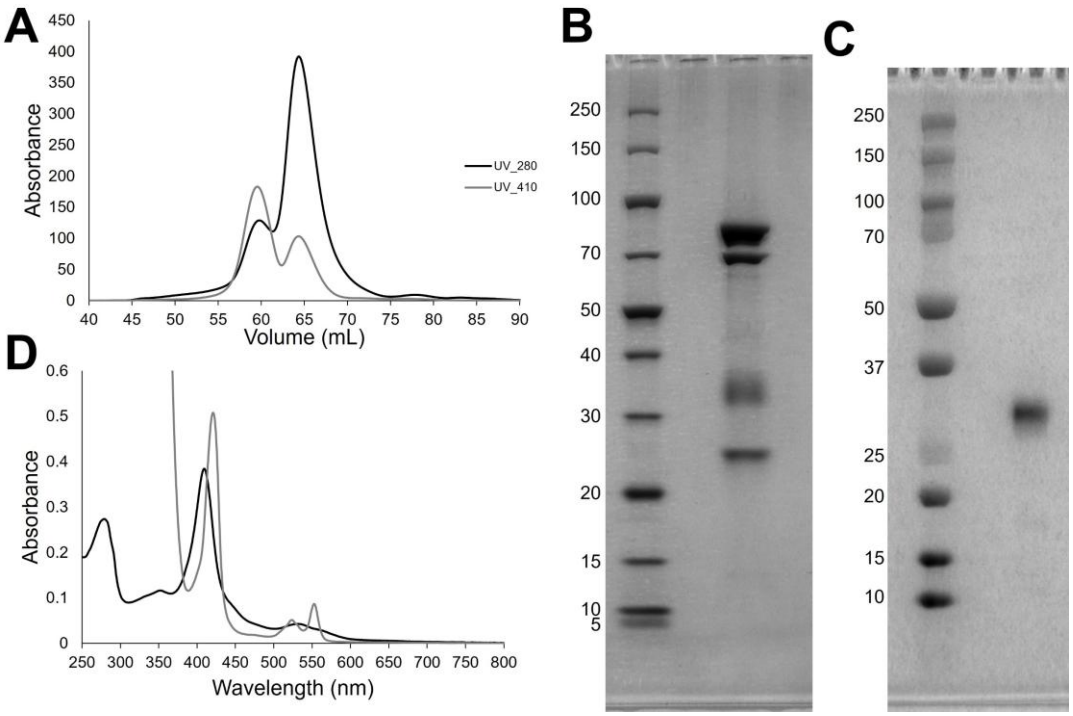


Figure 6.2: Improved DMSOR purification methodology allows for isolation of DmsEFAB complex. (A) SEC absorbance traces for DMSOR purified with methodology II. **(B** and **C)** Reducing SDS-PAGE analysis of DMSOR purified using method one, pooled fractions between 55-61 mL from panel A, stained with Coomassie **(B**) and heme-peroxidase stain **(C)**, marker molecular weight shown in kDa. **(D)** Electronic absorbance spectra of as-purified (black) and dithionite reduced DmsEFAB, in 20 mM HEPES pH 7.8, 100 mM NaCl.

at the anticipated molecular weights for each DmsEFAB component (Figure 6.2B), of which only the DmsE band stained positive for *c*-type hemes (Figure 6.2C). Electronic absorbance spectra for this species were indeed dominated by

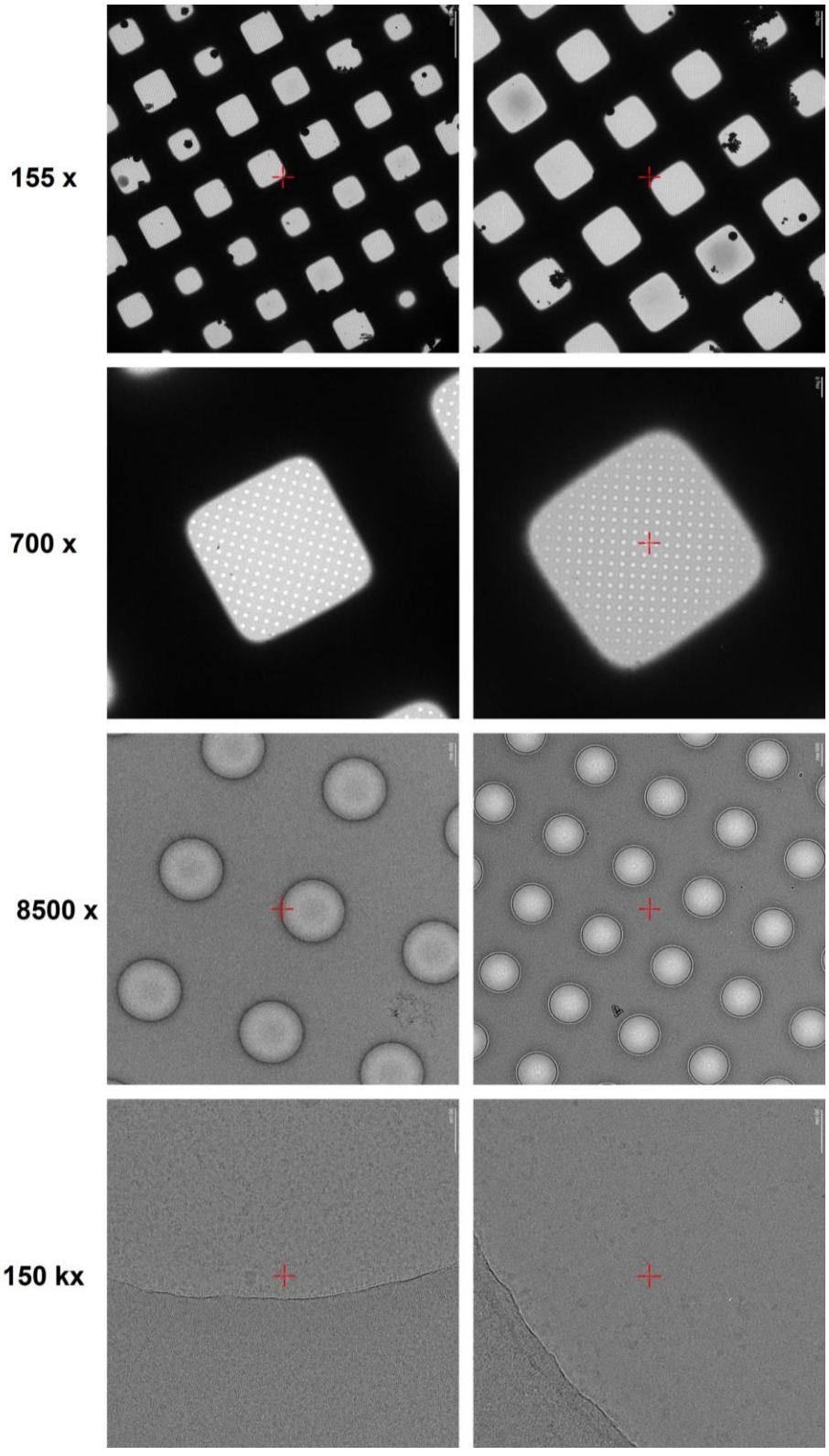


Figure 6.3: Screening of cryoEM grids on 200 kV Talos with Falcon 4i detector. Screening images collected of cryogenic grid set one (left) and cryogenic grid set two (right), note the presence of individual particles "blobs" present in grid set two at 150 kx not easily picked out in grid set one.

features typical of *c*-type cytochrome proteins, with a maximum at 410 in the as-purified state, and maxima at 553, 524 and 421 nm in the dithionite reduced state (Figure 6.2D). Taken together, these results indicate successful isolation of a stable and complete DmsEFAB complex via an improved methodology, therefore fractions containing this species were pooled and *methodology II* used to produce all protein used in further experiments. An additional but very faint band at ~65 kDa was also present on the Coomassie stained gel for this species (Figure 6.2B). This supports the suggestion that the clear ~65 kDa band identified by SDS-PAGE of protein produced by *methodology I* (Figure 6.1A), is indeed incompletely denatured DmsA, not DmsF which appears distinctly at ~70 kDa in Figure 6.2B.

CryoEM data collection and processing

Purified DmsEFAB was next subjected to cryoEM single particle analysis. Initial attempts at negative staining were unable to resolve clear particles, however SEC of the sample immediately before preparation of the grids (Chapter 6 appendix figure 1) did not indicate sample damage/loss of any kind. Therefore it was considered most likely that the inconclusive images where attributable to the negative stain methodology, not the quality of sample. Subsequently it was decided that screening by negative staining was unnecessary and cryogenic 300 mesh copper grids were prepared.

Screening of these grids with an FEI Talos 200C electron microscope equipped with a Falcon 4i detector confirmed that grids had vitreous ice within the grid holes and did not appear to suffer from severe ice contamination (Figure 6.3 left). Analysis of these grids at 150 kx (thousand times) magnification revealed that the holes did not contain clear individual particles which could be picked (Figure 6.3 right). It is possible that the explanation is these grids contained protein so densely packed together that particles appeared stacked upon one another such that there were no clear boundaries between particles leading to poor contrast. Another alternative and not mutually exclusive explanation is that the ice thickness was too great, introducing poor contrast and therefore making induvial particles unidentifiable.

It was subsequently decided to prepare a second set of cryogenic grids, with reduced protein concentration (1 mg mL^{-1} down from 5 mg mL^{-1}). Screening with these grids also confirmed they contained vitreous ice within the holes, and did not suffer from severe ice contamination (Figure 6.3 lower). Imaging of cryogenic grid set two at 150 kx magnification revealed individual particles within the ice which could be picked (Figure 6.3 lower 150 kx). Although the density of particles was low meaning relatively few would be obtained per image which is not ideal, given it may be possible that the low protein concentration used was instrumental in obtaining good ice thickness and clear individual particles, this lower particle count was considered acceptable.

Grid set two was utilised for a full high resolution data collection on Titan Krios III equipped with a Falcon 4i detector and a Selectris X energy filter at Diamond Light Source. This yielded a data set containing ~8000 micrographs. The data was processed using Relion 5.0[15] as described in *Chapter 6 methods*. Ultimately ~57,000 high-quality homogeneous particles were identified and retained after 2D and 3D classification which removed particles that were not DMSOR or did not classify into subsets with high resolution features. From these a 3D reconstruction was obtained *de novo*, and refined to a final estimated resolution of 3.32 and 2.98 Å with and without map sharpening, respectively, determined by fourier shell correlation cut off of 0.143 (Chapter 6 appendix figure 4), for further details see *Chapter 6 Methods*. Data collection, data processing and refined model validation statistics are made available in Table 6.1.

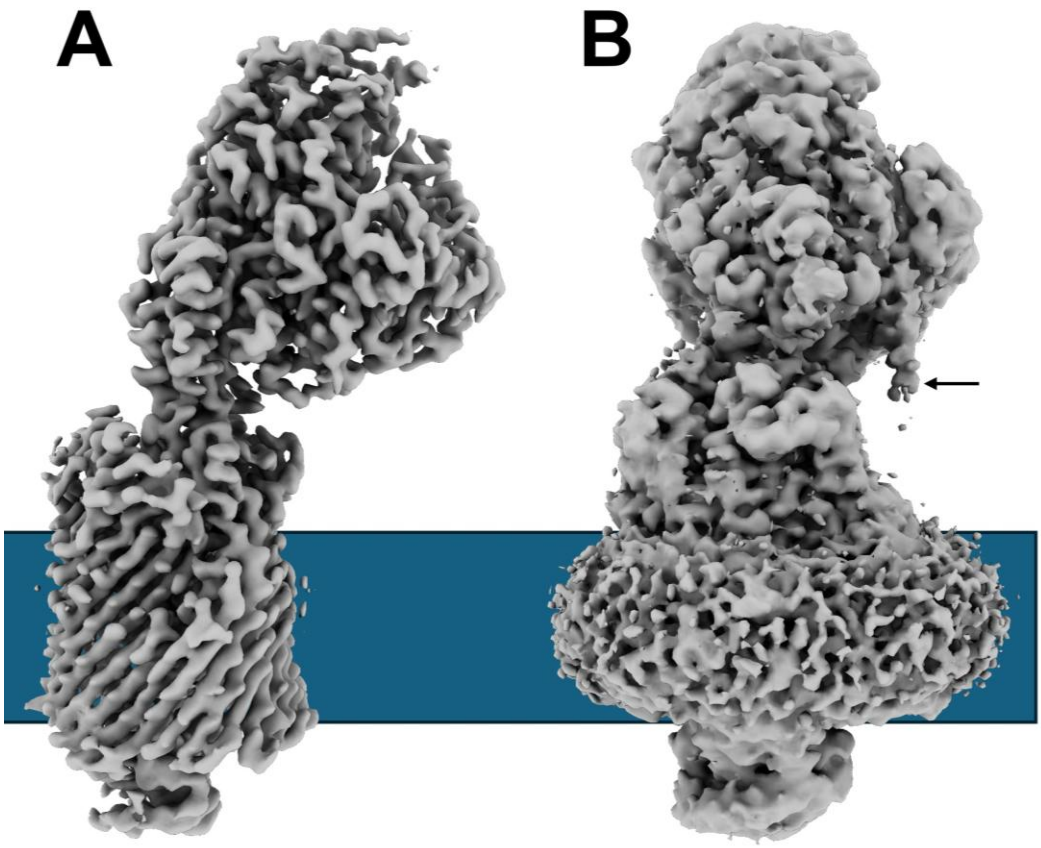


Figure 6.4: Unsharpened cryoEM density maps of DMSOR. Unsharpened cryoEM density maps for DmsEFAB complex with estimated resolution of 3.32 Å, contoured at 6 σ **(A)** and 3 σ **(B)**. Black arrow indicating the location of the DmsA N-terminus in which a lipid group is likely attached, and blue box indicating the approximate position on the membrane in vivo.

Atomic structure of DmsEFAB

The map shows a particle ~170 Å in height (perpendicular to the membrane plane), spanning ~70 by ~70 Å in the membrane planes (Figure 6.4). The particle displays two clear components, a barrel shaped feature, consistent with a porin, that is mostly filled with densities, and a head group which contains a funnel-like

feature and is tilted approximately ~35 ° relative to the longest dimension of the entire particle (Figure 6.4A). In the unsharpened map contoured at 3 σ, a belt around the barrel indicating the presence of an LDAO detergent micelle is clear, and a poorly resolved, and likely flexible, feature points from the headgroup towards the detergent belt which is consistent with the presence of a lipid anchor tethering the N-terminus of DmsA within the detergent micelle (Figure 6.4B). This assertion is supported by analysis of the DmsA sequence with the signal peptide analysis service SignalP-5.0[16] which suggests the protein is transported across the inner-membrane via the TAT system, after which it is cleaved and lipidated at Cys22 (Chapter 6 appendix figure 5). This feature appears to be crucial for extracellular reductase enzymes like MtrC and PgcA[17,18], taken together these results indicate that this is also the case for DmsA.

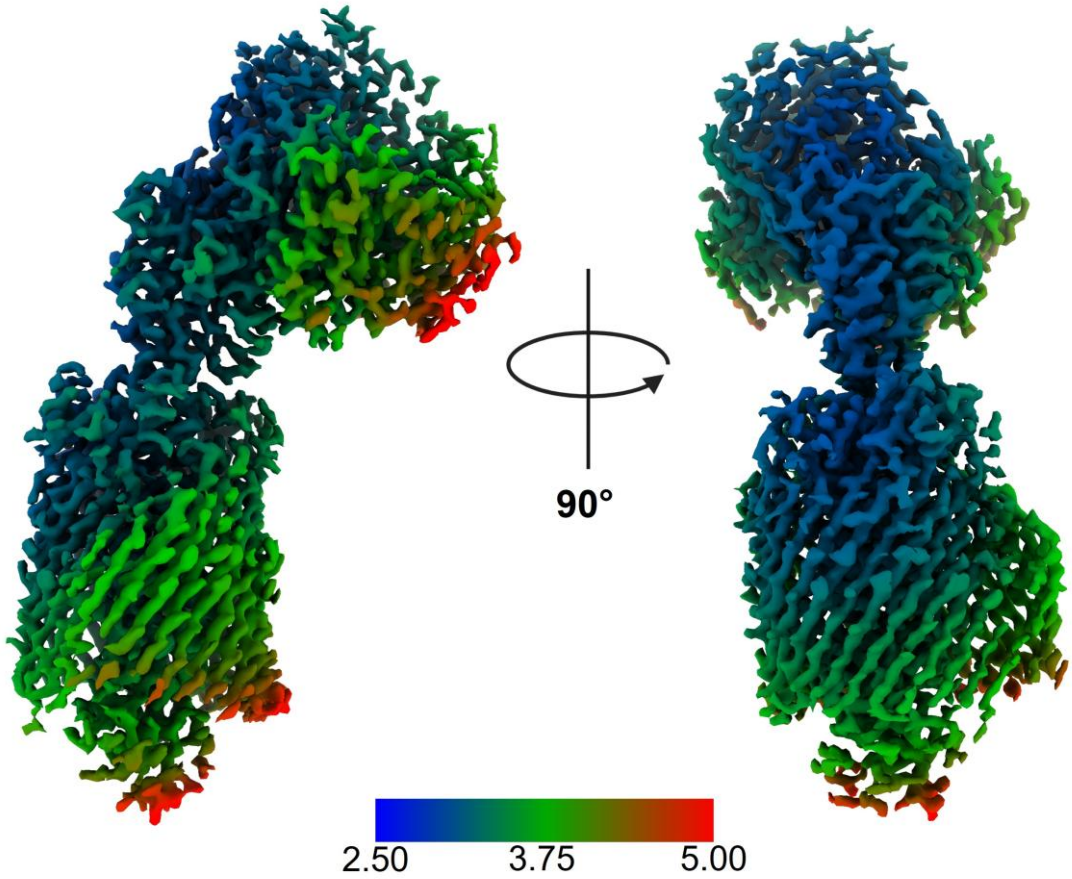


Figure 6.5: Local resolution estimate of sharpened cryoEM density map. Sharpened cryoEM density map (contoured at 5 σ) coloured with estimated local resolution calculated using Relion, given resolutions shown in Å.

Clear density is visible for individual porin beta strands in both sharpened and unsharpened maps although the loops between most strands are less well resolved. Estimating the local resolution of the sharpened map revealed that the ~3 Å estimated global resolution is representative for most of the particle (Figure 6.5). Regions with the highest local resolution are concentrated close to the interface between the barrel and the headgroup, especially on the opposite side

of the particle to the side in which the funnel is located. This is somewhat unusual as the highest resolution regions of cryoEM reconstructions are usually within the particle core, it may reflect a minor case of preferential orientation of

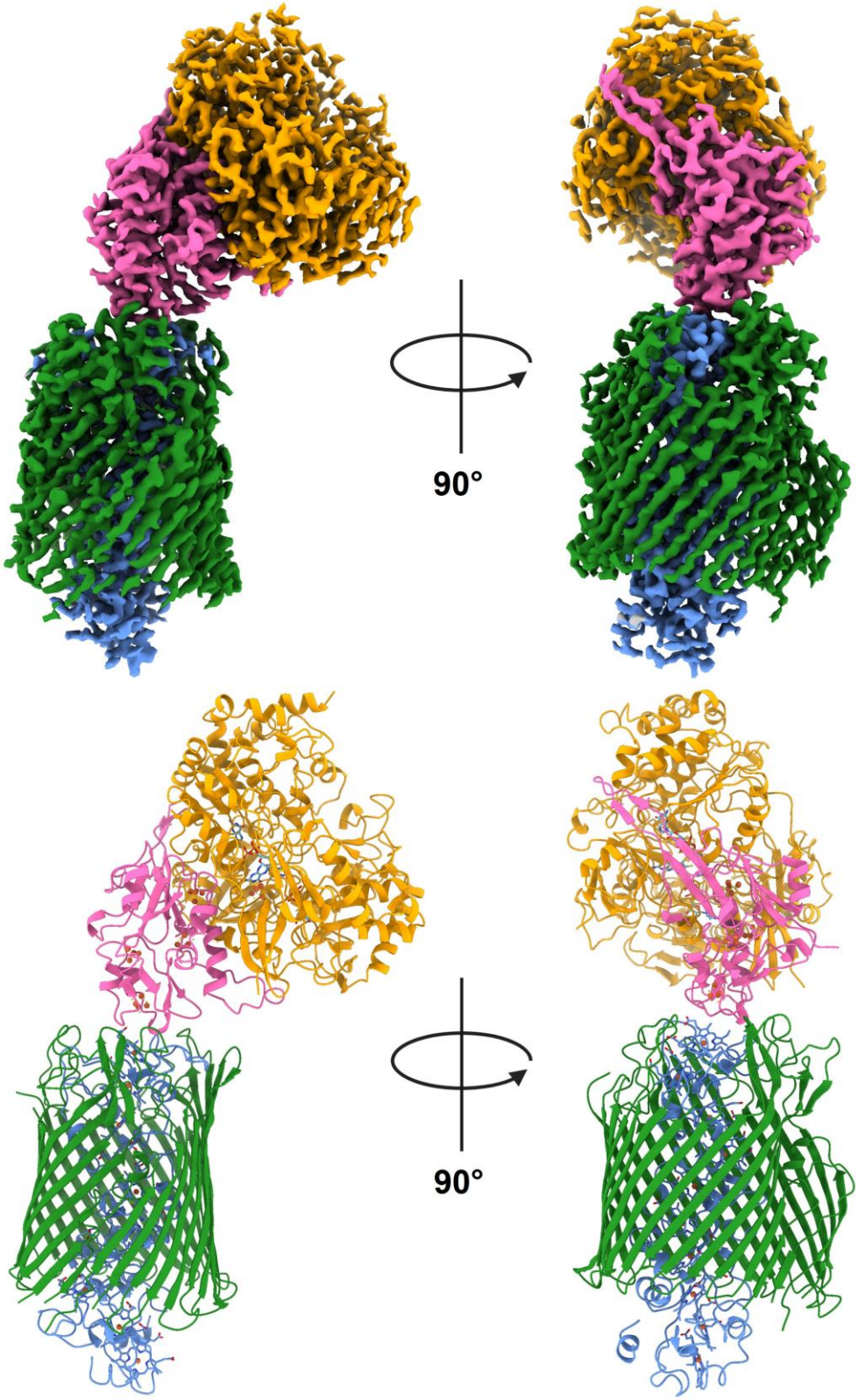


Figure 6.6: Sharpened cryoEM density map and DMSOR atomic model. Sharpened cryoEM density map (top) for DmsEFAB complex with estimated resolution of 2.98 Å, contoured at 5 σ coloured by subunit (orange-DmsA, pink-DmsB, blue-DmsE and green-DmsF).

particles within the grid ice. It is also possible that the higher resolution regions within the particle are considerably more ordered and rigid than the lower resolution regions. Another possible explanation is that the intrinsic and asymmetrical shape of the particle makes certain 2D projection images of it easier to pick out, and align correctly in 3D. Subsequently certain projections may be overrepresented, even if they were proportionally represented in raw the images, leading to mildly anisotropic resolution. These possibilities are not mutually exclusive, some combination of them all may be responsible.

The density map allows clear tracing of the protein backbone for each subunit, and there is clear density visible for side chains (Chapter 6 appendix figure 6), allowing for the amino acid sequences of DmsEFAB to be modelled and their rotamers assigned with confidence In addition, two MGD pterin ligands, five Fe_4S_4 clusters and ten *c*-type hemes were also modelled (Figure 6.6). The final model obtained an overall MolProbity score of 1.88.

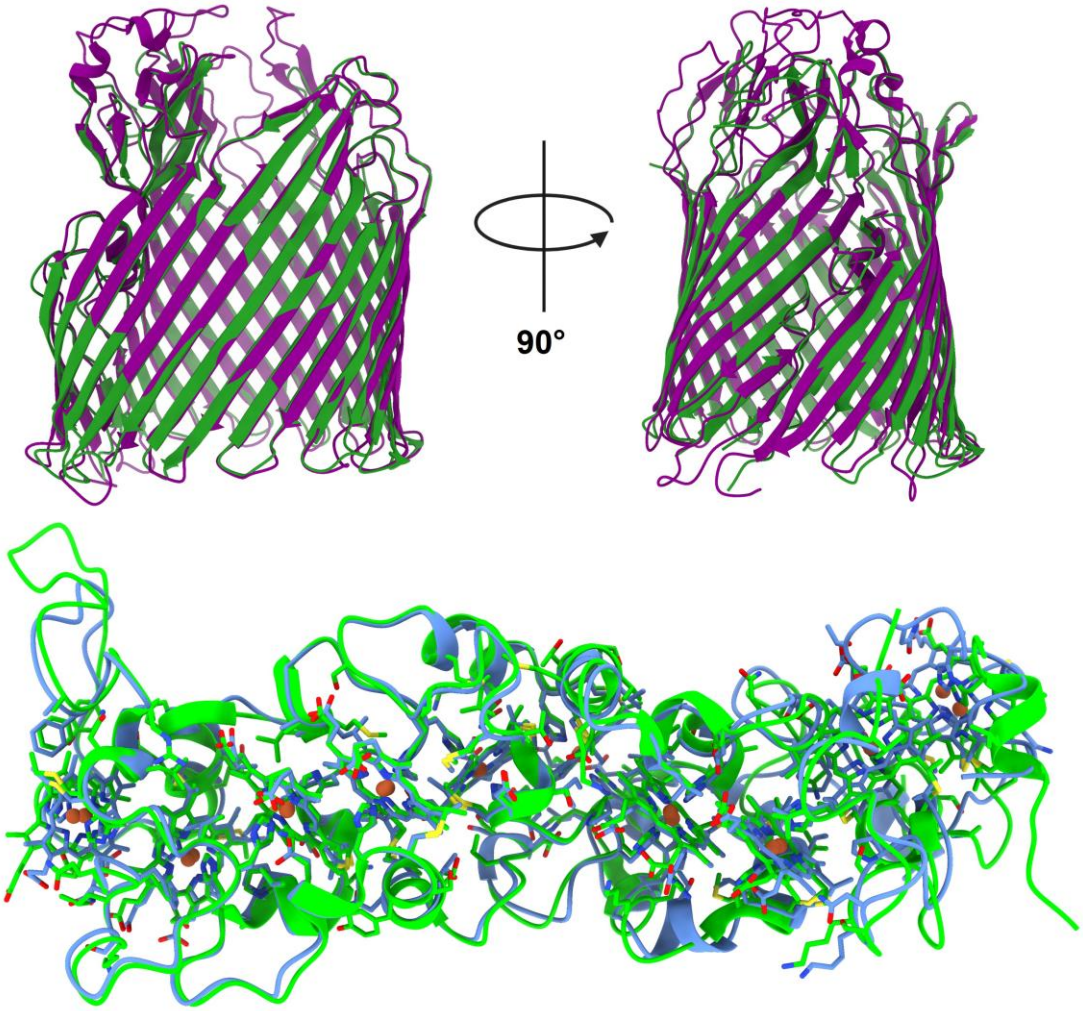


Figure 6.7: Atomic models of DmsF and DmsE show high similarity with MtrA and MtrB. Superposition of structures of MtrB (purple) with DmsF (dark green) and MtrA (light green) with DmsE (blue). DmsF/MtrB shown with extracellular side on top, DmsE/MtrA shown with extracellular end (C-terminus) on to the left and N-terminus to the right.

Inspection of the atomic model reveals that DmsF forms a porin consisting of 26 beta strands, that insulates the DmsE decaheme from a ~35 Å wide detergent belt. The DsmF barrel is 70 Å in length along the membrane-perpendicular plane and 55 Å across in the membrane-parallel planes. DmsF can be superimposed onto MtrB by the Cα carbons with an RMSD of 0.859 Å (Figure 6.7 upper) indicating very close structural similarity, which is unsurprising given with their high sequence identity (54%). Like MtrB, DmsF is not perfectly barrel shaped with one side protruding. The protrusion is particularly pronounced at the periplasmic side and creates an empty channel 30 Å deep and 20 Å across reaching up within the porin approximately half its length before becoming sealed. Several loops which link the porin beta strands are not well resolved enough to model, Y89-W94 (periplasmic), Q287-D303 (extracellular), A344-K356 (extracellular) and Y434-E443 (extracellular), indicating they are most likely disordered. No density was observed indicating the presence of tightly bound lipids within the detergent environment although this is does not necessarily indicate there are none in vivo, since LDAO is a relatively harsh detergent and thus typically strips membrane proteins of their annular lipids.

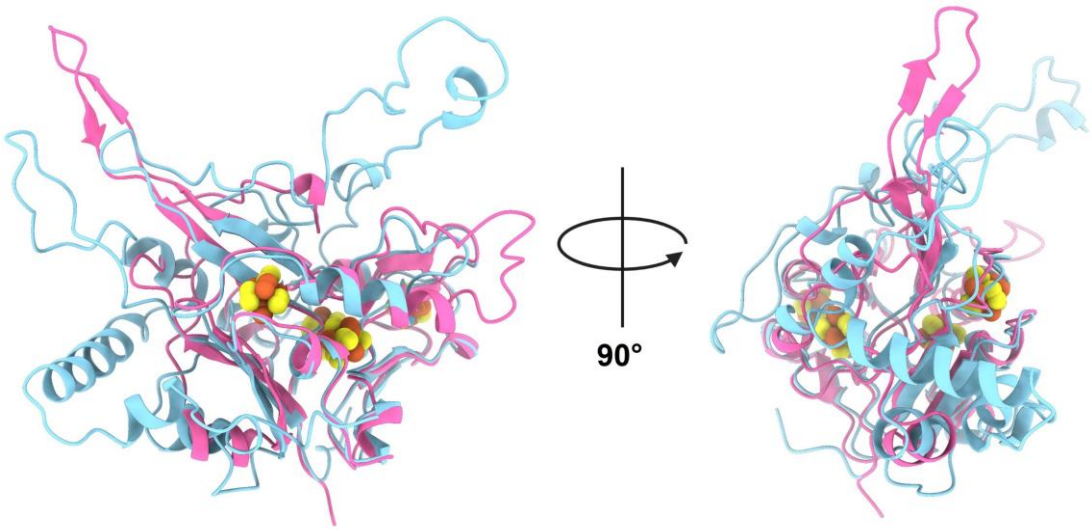


Figure 6.8: Superposition of DmsB with the FeS subunit of the A. oryzae perchlorate reductase, PcrB. Superposition of crystal structure of Azospira oryzae perchlorate reductase FeS subunit PcrB (PDB-5CH7) (blue) with atomic model of DmsB (pink).

DmsE is almost totally embedded within DmsF and can be superposed onto MtrA by the Cα carbons with an RMSD of 0.752 Å. Its 10 *c*-type hemes are well resolved and clearly all His-His coordinated. Furthermore, they are packed together with close edge to edge distances, and are arranged into pairs that are parallel stacked. DmsE makes numerous interactions with residues of DmsF that are oriented into the porin interior, however the DmsE N-terminus extends some 20 Å into the periplasm. This region is poorly resolved and atomic modelling with confidence here was extremely challenging. The identity of a small region of density here could not be assigned, the most likely possibility is it contains residues at the DmsE N-terminus that are poorly ordered. However the possibility that the unassigned density is derived from unresolved regions within DmsF

cannot be eliminated. The explanation for this significant variability in flexibility across DmsE maybe indicate that a flexible DmsE C-terminus is better suited to accept electrons from a broad range of donors in the periplasm. It is also possible that a flexible DmsE C-terminus is important in assembling the complex. Given this feature is maintained in both MtrA and DmsE, it is likely important for their function in some capacity.

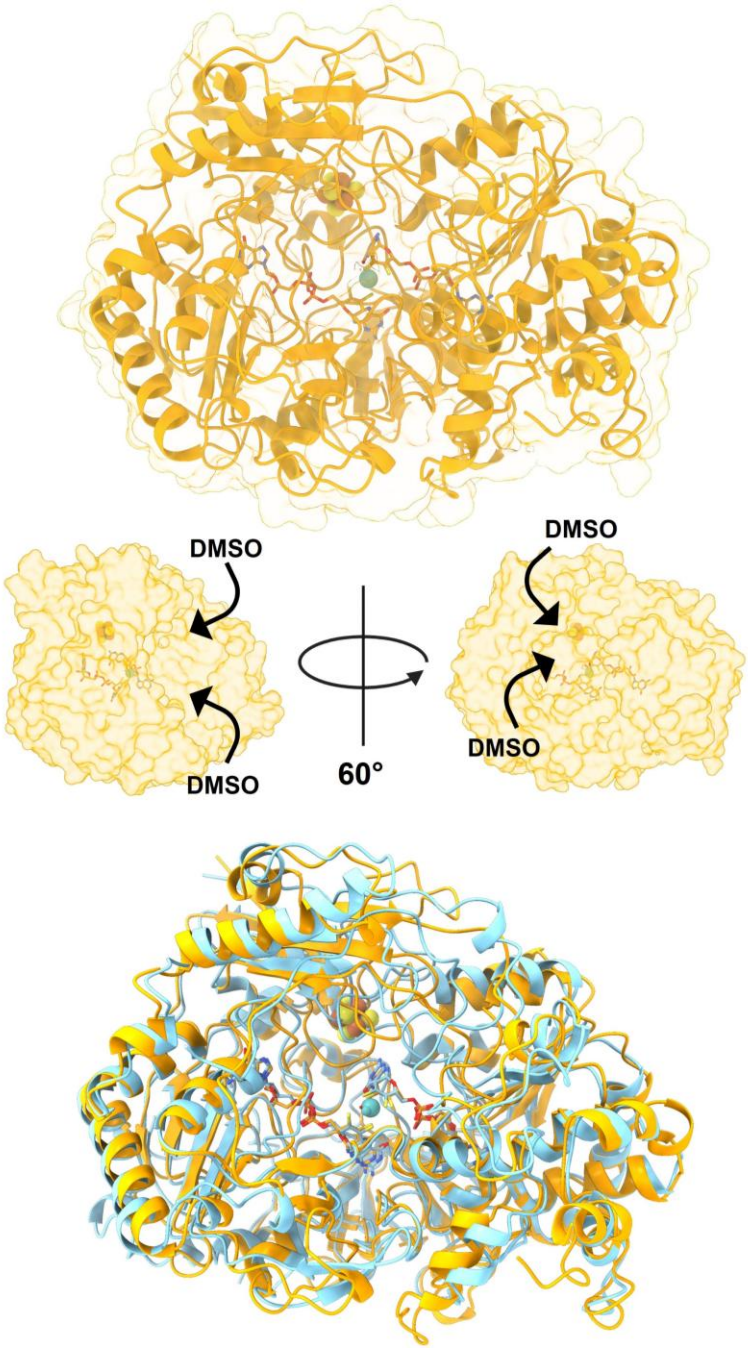


Figure 6.9: Atomic structure of DmsA is dominated by a depression which leads to the active site and is shared with its closest analogue MtrZ. Atomic model of DmsA (top), with surface view (middle) indicating the position and shape of the central cavity which extends down to the molybdopterin active site and DmsA (orange) superposition with *H. influenzae* MtsZ (PDB-7L5S) (blue).

DmsB, is inserted between DmsE and DmsA. Its secondary structure is comprised of 4 short alpha helices, a single longer alpha helix and 8 beta strands which together encapsulate 4 Fe_4S_4 clusters. Residues 60-78 form a protrusion leading away from the bulk of DmsB across the surface of DmsA (Figure 6.6). 3D structure alignment searching with DmsB across all PDB entries using Foldseek[19] indicates that the most similar entry (RMSD of 1.00 Å) in that database is the FeS subunit of the *Azospira oryzae* perchlorate reductase, PcrB (PDB-5CH7)(Figure 6.8). This perchlorate reductase FeS subunit contains three Fe_4S_4 clusters, however its cluster which is analogous to the cluster that in DmsB is the first in its cofactor chain, is instead a Fe_3S_4 cluster[20].

DmsA adopts a fold which is largely alpha helical in character with approximately globular dimensions. The funnel-like cavity which was evident even early in data processing is within this subunit where it is positioned centrally within the molecule and reaches 30 Å in diameter and 20 Å in depth (Figure 6.9). DmsA incorporates a single Fe_4S_4 cluster, coordinated by the side chains of residues C48, C52, C56 and C88. Adjacent to the DmsA Fe_4S_4 cluster and located at the bottom of the funnel, two molybdopterin guanine dinucleotide molecules coordinate the Mo ion via their dithiolate moieties while a fifth coordination is provided by the hydroxyl group of S183. A PDB search with Foldseek[19] indicates that this protein's closest structural analogue (RMSD of 1.10 Å) there is a methionine sulfoxide reductase, MtsZ at pH 5.5 (PDB-7L5S), from the bacterium *Haemophilus influenzae*. MtsZ displays a funnel shaped cavity similar to DmsA, although unlike DmsA this enzyme contains no cofactors beyond its

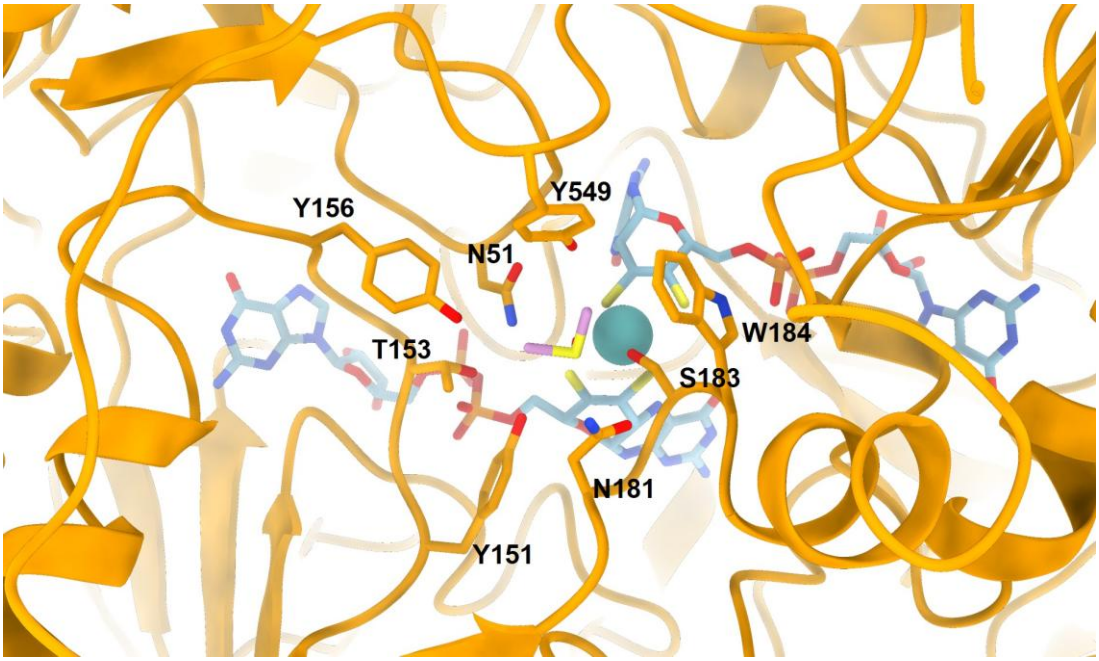


Figure 6.10: The active site of DmsA is consistent with that of other DMSO reductase enzymes and is lined by the presence of amino acids which likely exert control over substrate specificity. Atomic model of DmsA active site. Residues that line the substrate binding site shown with sticks and labelled. Predicted DMSO position and orientation obtained from superposition with *R. capsulatus* reductase bound DMSO crystal structure (PDB-4DMR) shown centrally with its carbon atoms coloured pink.

molybdopterin moiety (figure 6.9)[21]. In the active site of MtsZ, electron density was observed indicating the presence of an oxo group coordinating the molybdenum, which is not the case for DmsA although at the significantly poorer resolution of 3.0 Å this doesn't necessarily rule out its existence. No protein side chain was observed coordinating the molybdenum ion in MtsZ at pH 5.5, however in DmsA at pH 7.8, Ser183 coordinates the molybdenum similar to many other DMSO reductase enzymes[22,23].

The density in the DmsA active site indicates coordination of the molybdenum ion by all thio groups of the pterin ligands (bond lengths of 1.6-2.8 Å) in addition to coordination from S187 (bond length 1.8 Å). It also shows the substrate binding site directly above the Mo ion remains open. The funnel narrows to reach the Mo and is lined by the presence of aromatic residues (Y151, Y156, W184 and Y549) which likely exert significant control over the enzyme's substrate specificity (Figure 6.9). Superposition with the crystal structure of *R. capsulatus* DMSO reductase with bound DMS (PDB-4DMR)[24] yields an RMSD of 1.09 Å and indicates that in DmsA, N51, N181 and T153 also probably make important interactions with bound substrates (Figure 6.9). These observations are important since despite several attempts to define a specific substrate binding site in extracellular reductases, unambiguous identification of such a site has only been reported for the case of soluble iron complexes binding UndA[25,26].

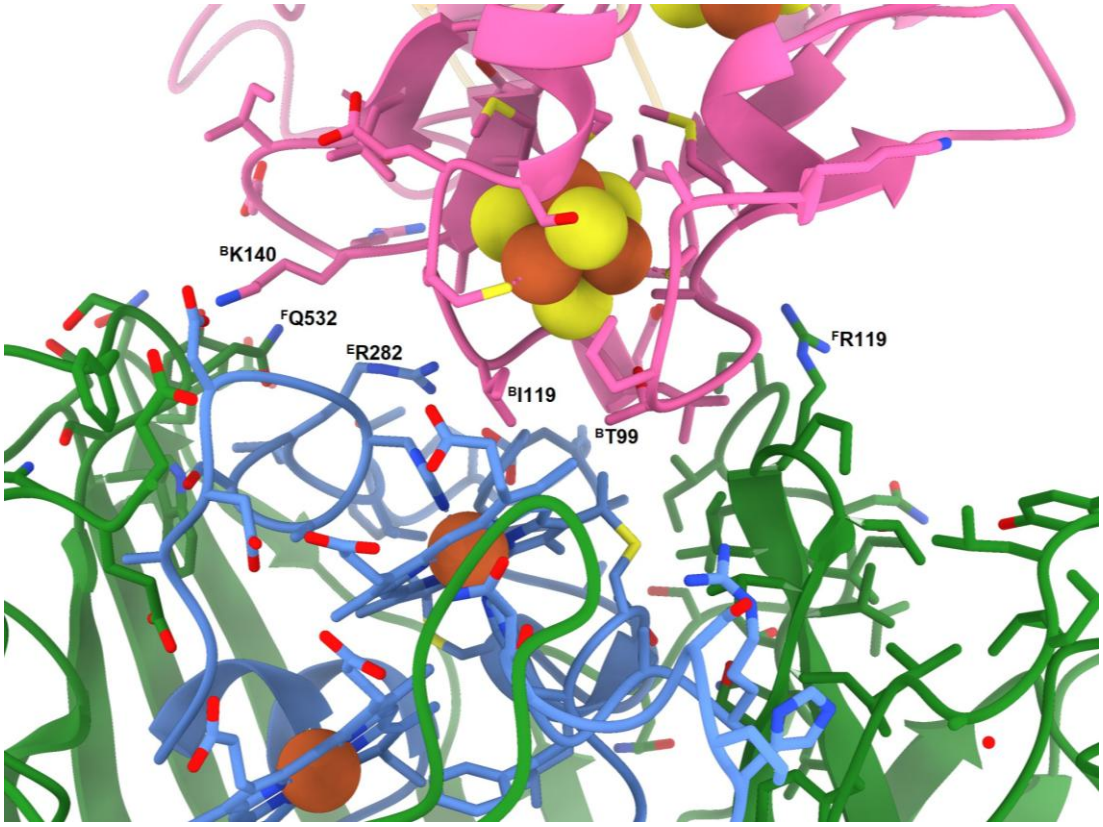


Figure 6.11: The interface between DmsEF and DmsB is maintained by a limited number of residues. Atomic model of DmsEF:DmsB interface, with DmsE in blue, DmsF in green and DmsB in pink. Residues that have been identified as likely to be particularly important in complex formation/stabilisation are shown with sticks and labelled.

Taken together all these analyses are consistent with describing the DmsEF and DmsAB substructures as being generally similar to their respective homologues *S. oneidensis* MtrAB and *R. capsulatus* DmsAB, respectively. Given those homologues do not form a DmsEFAB tetrameric complex however, it follows that it is likely to be the interface between DmsEF and DmsB that is the location within the structure that is most novel.

Here the complex width contracts to its the narrowest point stretching only 22 by 17 Å. This is exemplified by analysis of the structure using PISA[27] which revealed the DmsF:DmsE interface is 4956 Å^2 in area, while the DmsA:DmsB interface is 2497 Å^2 in area. In stark contrast, the DmsF:DmsB interface is only 324 Å^2 while the DmsE:DmsB interface is 430 Å^2. Thus DmsEF contacts DmsB across a total area of only 754 Å^2. This at first does appear smaller than average but not remarkably so, when compared with the average interface size found in deposited PDB structures of protein complexes, which in 2008[28] was estimated to be ~1,230 Å^2. However it is important to note that in 2008 the PDB was especially disproportionate in representation of smaller proteins which are amenable to X-ray crystallography. Therefore the majority of complex structures deposited were complexes of smaller proteins making the 754.3 Å^2 measured for DmsEF:DmsB remarkable when the ~230 kDa size of the tetramer is considered. A fairer comparison is with the structure of the MtrCAB complex[6], which when analysed using PISA[27] contained an MtrC:MtrB interface of 887 Å^2, and an MtrC:MtrA interface of 578 Å^2 making the total between the porin-cytochrome substructure and the headgroup 1,465 Å^2. This value is almost twice that of the DmsEF:DmsB interaction interface.

Closer inspection reveals DmsF forms interactions with DmsB via extracellular loops (Figure 6.11). In the first of these ^{F}P198 (Indicating P198 from DmsF) interacts with ^{B}P98 and the side chain of ^{F}R199 interacts with the backbone carbonyl of ^{B}G100. The second loop is located at the opposing side of the barrel, where the backbone carbonyl of ^{F}Q532 appears to interact with the ε-amino group of ^{B}K140 and the side chain of ^{F}Q532 interacts with the backbone carbonyl of ^{B}E138. A third loop is located adjacent to the fist, within which ^{F}P636 and ^{F}G637 appear to contact ^{B}E116 and ^{B}L117 via weak interactions.

All interactions DmsE makes with DmsB are mediated entirely by a loop comprised of residues ^{E}277-292, and its C-terminal heme (^{E}Heme10) (Figure 6.11). ^{E}D277 and ^{B}K140 appear to form a salt bridge via their side chain moieties although the map density is weak here indicating this interaction maybe artefactual. The guanidinium group of ^{E}R282 makes an extensive network of interactions with the backbone carbonyls of ^{B}C118, ^{B}I119, ^{B}R129 and ^{B}K140. The backbone amide of ^{E}A291 forms a hydrogen bond with the backbone carbonyl of ^{B}E116. Finally, the side chains of ^{E}A291 and especially ^{E}F292 in combination with ^{E}Heme10 form a hydrophobic pocket which stabilises the hydrophobic regions of ^{B}P98, ^{B}T99 and in particular, ^{B}I119.

These interaction networks seem remarkably modest, but they nonetheless manage to position ^{E}heme10 within a metal to metal distance of only 10.9 Å of the first $^{B}Fe_4S_4$ cluster indicating electron transport between them can likely occur extremely rapidly. Indeed the metal to metal inter-cofactor distances across the DmsEFAB cofactor chain remain low across the structure, falling between 9.5 and 11.6 Å in DmsAB and 9.0 to 11.7 Å in DmsE. These distances are comparable to those in MtrCAB[6], indicating the maximal electron flow rate through the cofactor chain in DmsEFAB is likely to be on a comparable order of magnitude as in MtrCAB. Given that the reported k_{cat} for DMSO reductase enzymes is on the order of 80 s^{-1} and electron transport through MtrCAB has been experimentally demonstrated to reach at least 8,500 s^{-1}, these observations together indicate that either substrate turnover in DmsA, or delivery of electrons from periplasmic shuttles, are more likely to be the rate limiting step for DmsEFAB, than is intramolecular electron transfer across its cofactor chain[29,30].

Discussion

One important objective of this chapter was first to elucidate the molecular basis for complex formation between DmsEF and DmsB. Then to exploit that understanding to determine whether the underlying interactions are unique to DmsEFAB or conserved in all porin-cytochrome complexes, by comparison with the structure of *S. baltica* MtrCAB. Given DmsB and MtrC are evolutionarily distinct, it follows that identification within them of a conserved DmsEF/MtrAB binding motif would suggest this motif is likely common to many analogous systems. Such a motif would represent a remarkable example of convergent evolution or gene transfer, however the alternative result, that they each bind DmsEF/MtrAB via a different set of interactions would be equally as interesting as this would indicate that complex formation is viable via interactions that aren't tightly restricted evolutionarily. If so, it might be possible to engineer binding to a porin-cytochrome complex into an existing redox enzyme of choice, and therefore produce a complex that can perform extracellular redox catalysis of a reaction which has industrial or other commercial value.

The analysis presented above implicates only a handful of residues from each DmsB, DmsF and DmsE as important in formation of the complex. Of particular importance appear to be ^{F}R199, ^{F}Q532, ^{E}F292, ^{E}R282, ^{B}I119, ^{B}P98 and ^{B}T99. Comparison of DmsF with MtrB reveals that ^{F}R199 and its hydrogen bonds are not conserved in MtrB, nor is ^{F}Q532 and its interactions. Comparing DmsE with MtrA reveals that ^{E}R282 which appears to play a central role mediating interactions between DmsE and DmsB, is not conserved in MtrA, indeed N290 which takes its place in MtrA does not make any contacts at all with MtrC. ^{E}F292 is conserved in MtrA (as ^{MtrA}F307) however none of the residues in DmsB which line the hydrophobic pocket that ^{E}F292 creates (^{B}I119, ^{B}P98 and ^{B}T99) are present in MtrC. Sequence alignment between *S. baltica* MtrA and DmsE shows that their C-termini are sequentially very similar, however DmsE is missing a 7 residue long

sequence (^{MtrA}L293-G299). In the MtrCAB structure, these residues form a loop which contacts MtrC extensively, no comparable interactions are present in the DmsE:DsmB interface.

Comparing the DmsB:DmsF interface with the MtrC:MtrB interface, it is immediately apparent that the interactions found within MtrC:MtrB are much

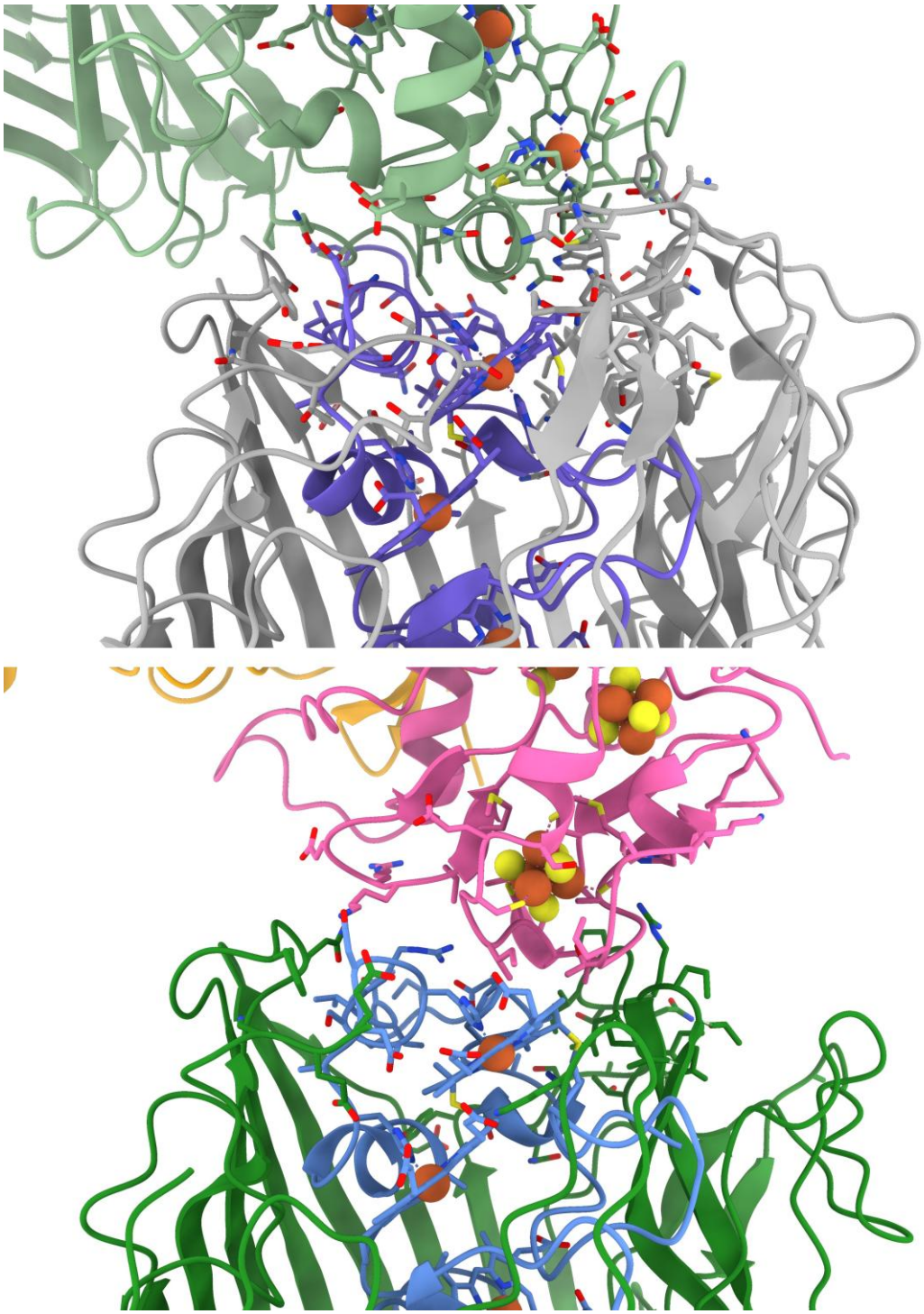


Figure 6.12: The interface between DmsEF:DmsB is totally distinct from that of MtrAB:MtrC. Superimposed atomic models of MtrAB:MtrC interface (MtrA in purple, MtrB in grey and MtrC in pale green)(upper) and DmsEF:DmsB interface (DmsE in blue, DmsF in green and DmsB in pink)(lower). Residues that have been identified as likely to be important in complex formation are shown with sticks.

more substantial (Figure 6.11). In both cases, the interactions are localised to opposite corners of the porin. On one side, the DmsEFAB atomic model predicts only ^{F}Q532 interacts with DmsB. Here in MtrCAB, ^{F}Q532 is replaced by ^{MtrB}S565 which makes no contacts with MtrC. However, the larger MtrC extends much further over the surface of the porin than does the smaller DmsB, making extensive contacts as it does so. Particularly apparent are hydrogen bonds formed by ^{MtrB}Q616 with ^{MtrC}Q93 and ^{MtrB}N566 with ^{MtrC}R272. Neither of these appear in any form in the DmsEFAB structure. On the other side of the porin, MtrB makes hydrogen bond networks to MtrC, ^{MtrB}S161 with ^{MtrC}Q296 and ^{MtrB}N219 with ^{MtrC}N302. ^{MtrB}I381 and ^{MtrB}F274 form a hydrophobic pocket in which ^{MtrC}Heme5 is firmly embedded. Further hydrophobic interactions occur between ^{MtrB}W670 and ^{MtrC}L298. None of those interactions, or interactions comparable to them, are present between DmsF and DmsB, in fact relative to the extensive networks between MtrC:MtrB, DmsF and DmsB contact one very minimally.

Taken together these observations show firstly that the bonding networks which hold DmsEF in contact with DmsB are sparse in comparison with those seen in MtrCAB. Secondly, they demonstrate that the bonding interactions holding these two complexes together are totally distinct from one another and share almost no residue combinations in common at all. Therefore assembling a headgroup onto a porin-cytochrome complex can be achieved via more binding modes than one, possibly without the requirement for a sophisticated and extensive binding

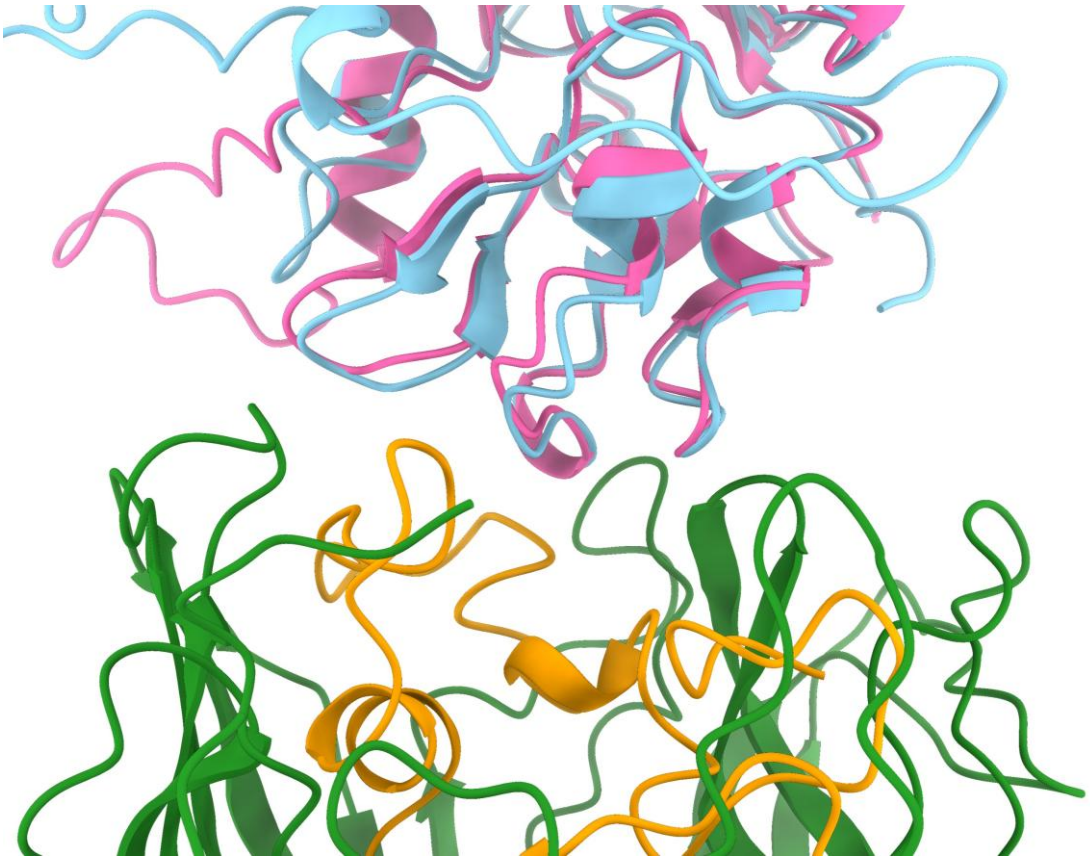


Figure 6.13: The peptide backbone of DmsB remains remarkably similar to PcrB despite its forming DsmEFAB. Superimposed atomic models of DmsEF:DmsB interface (DmsE in yellow, DmsF in green and DmsB in pink) with PcrB (pale blue).

surface to be present. This interpretation is supported by examining how DmsB differs at its DmsEF binding surface, from its closest structural relative identified with Foldseek, PcrB. Remarkably, the peptide backbone positions in these two proteins in this region are extremely similar and even many residues are conserved (Figure 6.12). A handful of residues in DmsB which have hydrophobic groups and are shielded from the solvent by DmsEF, have polar equivalents in PcrB, ^{B}L117 is replaced with ^{PcrB}K169 and ^{B}I119 which appears pivotal in DmsEF:DmsB binding is replaced with ^{PcrB}K171. Several other residues nearby are also not identical, ^{B}T99 is replaced with ^{PcrB}N151, ^{B}K140 is replaced with ^{PcrB}N192, ^{B}D168 with ^{PcrB}E116, ^{B}V141 with ^{PcrB}K193, ^{B}A114 with ^{PcrB}H166 and ^{B}G100 with ^{PcrB}E152.

By mutating these residues in PcrB, it may be possible to confer it with binding to DmsEF. Assuming all residue pair mutations are required for effective binding, 8 mutations would be necessary. This number conceivably could be tested by future researchers without great difficulty, it maybe that not all are necessary and a smaller number is sufficient. Mutagenic analysis of DmsB, DmsE and DmsF will be required to confirm the interpretations presented in this chapter. An excellent place to start would be analysis of ^{B}I119, ^{E}R282 and ^{F}R199 mutations since the cryoEM map and atomic model presented here indicate these likely play pivotal roles in complex formation.

Perhaps important to note here, is the fact that even *methodology II* developed to produce the DmsEFAB used in this chapter, yielded a species that was mostly comprised of DmsAB, not DmsABEF (Figure 6.2A, peak at ~65 mL). One interpretation of this result is that the detergent micelle can interfere with complex stability or the accessibility of the DmsA purification tag. Another interpretation which is not necessarily mutually exclusive with the first and supported by analysis of the cryoEM structure, is that the complex is weakly bound together and therefore DmsAB-only and DmsEF-only complexes are common in the outer-membrane. Once DmsEFAB was isolated however, no significant dissociation of the complex was discernible by SEC (Chapter 6 appendix figure 1). Although dissociation of DmsAB and DmsEF can be envisaged in the outer-membrane, the unsharpened cryoEM map indicates DmsA is likely lipidated, and that its lipid anchor usually resides in the same detergent micelle that stabilises DmsF (Figure 6.5B). This presumably will make dissociation of a detergent stabilised DmsEFAB complex unlikely, since the tetramer must break at the same time as the DmsA lipid anchor is transferred to a second micelle. Taken together, these findings indicate that although the DmsEFAB complex appears stable once purified in detergent, this cannot be surely assumed to be the case in vivo.

Some thought has also been given to the question of whether porin-cytochrome complexes are interchangeable, i.e. whether formation of an MtrAB-DmsAB complex is possible in vivo[31–33]. It has not been established either whether a

DmsE-MtrB or DmsF-MtrA complex could form. The alternative being that porins and cytochromes only form their transmembrane wire complexes when they are encoded together in the genome. The analysis described here argues strongly in opposition to the formation of an MtrAB-DmsAB or DmsEF-MtrC complex, since the DmsEF:DmsB and MtrAB:MtrC interactions are totally distinct, however experimental validation of this will be required. The high conservation in the structures of DmsF/MtrB and DmsE/MtrA is not inconsistent with the formation of DmsE:MtrB and DmsF:MtrA complexes, but it does not provide definitive evidence in favour for or against complex formation. If mixed DmsE:MtrB or DmsF:MtrA complexes do form however, these results suggest they are unlikely to be capable of productively binding to either extracellular headgroup, since they will be unable to present the correct binding surface. As the loss of these headgroups will result in reduced function for the complex, this information implies mixed complexes would an inefficient use of resources.

Assembling a ~230 kDa tetramer in the outer-membrane presumably incurs a significant energetic cost on *S. oneidensis* cells. Furthermore it has required the development of several different innovations that are necessary to allow complex assembly. The DmsAB subunits must be recognised by the type II secretion system, which is not the case for other DMSO reductase subunits. The DmsEF subunits must be made to interact with the DmsAB subunits. The DmsA subunit most likely must become a lipoprotein, which is indicated by its sequence, its pelleting with membrane fractions[12] and the cryoEM map viewed at low thresholds. These suggest that since so many hurdles have been overcome and retaining the gene cluster continues to be beneficial, a significant competitive advantage is provided for *S. oneidensis* by possession of the *dmsEFABGH* gene cluster. Given an extensive number of enzymes capable of supporting reduction of DMSO intracellularly have been described, what exactly the competitive advantage which extracellular DMSO reduction through DmsEFAB provides, is not obvious.

Gralnick et al (2006) suggest the ability to reduce DMSO extracellularly may allow for growth on solid DMSO which might exist under specific conditions in marine environments[10]. Whether or not solid DMSO is indeed present in the environment where *S. oneidensis* can be found, the structure of DmsA presented in this chapter clearly demonstrates that its molybdopterin active site is situated at the base of a deep funnel (Figure 6.8). Unlike MtrC which presents three *c*-heme cofactors at exposed egress points that are thought to allow for rapid electron transfer onto mineral particles[26,34], the DmsA structure is therefore not consistent with efficient reduction of solid substrates.

Gralnick et al also contemplate the possibility that the physiological substrate(s), product(s) or both, of DmsEFAB is/are toxic to *S. oneidensis*, making the complex facilitate growth with minimised toxicity. Here it may be notable that the closest structural relative in the PDB of DmsB is PcrB, a perchlorate reductase subunit.

Perchlorate is a compound that is recognised as toxic to a wide variety of organisms, at least in part through imposing chaotropic and oxidative stress[35,36]. DMSO reductase enzymes are generally considered to be promiscuous in their substrate selectivity[37]. Another potentially relevant example is that of arsenate/arsenite. The respiratory reduction of arsenate to arsenite by microorganisms is well documented, but it does incur cytotoxicity since both arsenite and are arsenate are toxic[38]. The structure of a respiratory arsenate reductase molybdoenzyme (Arr) has recently been reported and can be superimposed onto DmsAB with an RMSD of only 1.22 Å[39] although this does not prove that DmsEFAB can reduce arsenite or that arsenite reduction toxicity explains its unique subcellular localisation.

Evidently it is crucial that future work establishes the substrate range and kinetics of DmsEFAB experimentally. This chapter provides a basis for preparing the enzyme in pure form, with usable yield and also identified residues that possibly play an important role in controlling substrate access to the active site. In particular ^{A}Y151, ^{A}Y156, ^{A}W184, ^{A}N181 and ^{A}Y549 appear positioned to interact with, or block, potential substrates (Figure 6.9). Functional verification of the role these residues would also contribute significantly to our understanding of the specific function this enzyme has in the respiration of *S. oneidensis*. This remains the most important area for future research to address and is one that should not represent a great challenge as it has already been demonstrated for DmsAB[12], which this work determined was in fact the species produced by purification *Methodology I*.

Also highlighted as important in this chapter, is that despite many recent advances, exciting aspects of EET remain unexplored and undescribed in the literature. In particular, although the structure of two transmembrane wire complexes have now been resolved and they differ from one another significantly, they are clearly related. Bioinformatic and genetic studies indicate porin-cytochrome complexes with zero sequence identity to MtrCAB/DmsEFAB exist[40], and are crucial in the EET pathways of model organisms like *G. sulfurreducens*[41]. Although difficulties in producing these complexes natively or recombinantly to high yields will likely make them challenging to work with, this chapter firmly demonstrates that improved and now routine methods in structure determination by cryogenic electron microscopy can provide enormous insights into the study of porin-cytochrome complexes even without huge investments in time and resources.

Likely important in that success are several features intrinsic to porin-cytochromes complexes as cryoEM specimens. Firstly, they are typically well over 100 kDa in size making them relatively easy to pick in micrographs. Secondly, in spite of having no intrinsic symmetry, which can make particles like apoferritin great targets for cryoEM, these particles do exhibit distinctive shapes which likely make for 2D projections that are relatively easy to align accurately in

3D. Finally, despite being membrane proteins which can be more challenging than soluble proteins to work with, these complexes are natively present in the bacterial outer-membrane where they are presumably exposed to harsher conditions than intracellular membrane proteins experience. This is leads to proteins that are relatively robust, Indeed, no sign of degradation of the DmsEFAB complex, once purified, was observed at any point during the preparation of this thesis. It appears to have survived both freeze-thaw cycles and oxygen exposure which are sometimes damaging to membrane proteins and FeS clusters, respectively.

In conclusion, this chapter presents the structure of a porin-cytochrome-DMSO reductase complex, solved by cryoEM. This structure provides a wealth of new insights into these important protein complexes and provides a basis for understanding of the extracellular reduction of DMSO and its related compounds. The structure of DmsEFAB shows also the way forward in determining experimental structures of outer-membrane spanning reductase enzymes.

Methods

The *S. oneidensis* DMSO reductase gene cluster (SO1427–SO1432) kindly provided by Dr Marcus Edwards in a pBAD202 overexpression plasmid, after incorporation of a strep-II purification tag at the C-terminus of DmsA. Plasmid DNA was purified and sequenced before electroporation into *S. oneidensis* MR-1. Single colonies of DMSO-reductase *S. oneidensis* were used to prepare an overnight culture 100 mL LB (30 μg mL^{-1} kanamycin) which was grown at 30 °C, 180 RPM. This was then used to inoculate 1 L of LB media in 2 L baffled flasks (30 μg mL^{-1} kanamycin) and cultured at 30 °C, 180 RPM until an OD_{600} of ~0.6 was reached. 5 mM arabinose was added and the cultures returned to the incubator overnight. Cells were harvested by centrifugation (4000 G, 20 minutes, 16 °C) and resuspended in cold buffer A (100 mM Tris-HCl, 150 mM NaCl, 10% v/v glycerol, pH 8.0).

Cells were lysed by two passes through a french press at 1000 PSI. Membranes were collected by ultracentrifugation, 200,000 G for 2 hours, 4 °C. The supernatant was discarded and pellet resuspended in buffer A supplemented to 1.5% (final) LDAO in which it was manually homogenised, and then left solubilising with agitation, 4 °C, for 1.5 hours. Solubilised membranes were clarified by ultracentrifugation, 200,000 G for 1 hour, 4 °C. The pellet was discarded and supernatant loaded onto 5 mL strep-tactin affinity resin equilibrated in buffer B (100 mM Tris-HCl, 150 mM NaCl, 5 mM LDAO, pH 8.0). The resin was then washed with 15 CV (column volumes) buffer B, before eluting with 5 CV of buffer B supplemented to 50 mM biotin. Fractions were analysed for protein by SDS-PAGE and pooled before concentration to ~1 mL in 100 kDa cutoff spin concentrators. Pooled protein was then injected into a 16/600 Superdex 200

pg SEC column equilibrated in buffer C (20 mM HEPES, 100 mM NaCl, 5 mM LDAO pH 7.8) and eluted at 0.5 mL min^{-1}. Fractions of 1 mL volume were analysed for bands indicating the presence of DmsEFAB by SDS-PAGE and pooled before concentrating to 5 mg mL^{-1} assuming an extinction coefficient at 410 nm of approximately 1200 mM^{-1} cm^{-1} before snap freezing in liquid nitrogen and storing until use.

Aliquots of DmsEFAB were thawed, pooled and applied to a superose 6 increase 10/300 GL SEC column equilibrated in buffer C and eluted at 0.5 mL min^{-1}. Fractions of 0.5 mL volume containing protein were pooled and concentrated to 1 mg mL^{-1} (5 µM). 300 mesh quantifoil copper grids were glow discharged using a Pelco EasiGlow™ at 8 mA for 60 seconds before application of 5 µL protein solution, which was vitrified by plunging into liquid ethane with a Mark IV FEI vitrobot (30 second wait, 4 seconds blot, blot force 8, 100% humidity).

Grids were screened using a FEI Talos F200C TEM operating at 200 KV, equipped with a Falcon 4i direct electron detector. During data collection, 8,322 movies were collected using an FEI Titan Krios operating at 300 KV equipped with a Falcon 4i detector and TFS Selectris X energy filter set to a filter width of 10 eV. 130,000 x magnification was used yielding a pixel size of 0.921 Å, with a movie exposure time of 6.24 seconds, dose of 39.15 $e^{-}/Å^{-2}$. The nominal defocus was -1.0 to -2.0 µm and EER fractionation was 50.

Data processing was performed in Relion 5.0[42]. Movies were motion corrected with Relion's implementation of MotionCor2[43] using dose weighting, 5x5 patches and B factor of 150. CTF estimation was performed with CTFFind-4.1[44]. 1,000 particles were picked manually and used to train a Topaz[45] neural network model which when applied to the dataset picked 1,489,212 particles. Particles were extracted using a box size of 256 pixels, downscaled to 128 pixels. These were subjected to three rounds of 2D classification selecting classes containing features consistent with a 2D projection of a potential DmsEFAB complex (207,399 particles). Particles in this class were re-extracted without downscaling, and a 3D-initial reference model was next generated using the gradient descent method with 3 classes. The best class was then used as a reference for 3D classification (5 classes)(Chapter 6 appendix figure 3) yielding a single class (55,700 particles) with detailed features. This class was subjected to 3D high resolution refinement which converged at a resolution of 3.80 Å. Post-processing improved this map resolution to 3.47 Å. Two rounds of CTF-refinement and per-particle motion correction, followed by solvent masking and post-processing improved the resolution of the sharpened reconstruction to a final value of 2.98 Å (FSC = 0.143). DynaMight[46] was used in an attempt to further improve the reconstruction but this yielded a similar map which ultimately reached a poorer resolution (3.10 Å) and did not reveal significant molecular motion.

The structures of MtrA and MtrB were docked into the map, using the ChimeraX[47] fit in map command and then mutated to the sequences of DmsE and DmsF using the WinCoot[48] align and mutate function. The sequences of DsmA and DmsB were built into the map de novo using ModelAngelo[49], regions with poor density were built manually in WinCoot with the assistance of AlphaFold2[50] models. The coordinates then refined using Phenix-Real-space-refine[51] and ISOLDE[52].

Chapter 6: References

1. Shi, L., Dong, H., Reguera, G., Beyenal, H., Lu, A., Liu, J., Yu, H. Q. & Fredrickson, J. K. Extracellular electron transfer mechanisms between microorganisms and minerals. *Nature Reviews Microbiology 2016 14:10* **14**, 651–662 (2016).

2. White, G. F., Edwards, M. J., Gomez-Perez, L., Richardson, D. J., Butt, J. N. & Clarke, T. A. Mechanisms of Bacterial Extracellular Electron Exchange. *Advances in microbial physiology* **68**, 87–138 (2016).

3. Hartshorne, R. S., Reardon, C. L., Ross, D., Nuester, J., Clarke, T. A., Gates, A. J., Mills, P. C., Fredrickson, J. K., Zachara, J. M., Shi, L., Beliaev, A. S., Marshall, M. J., Tien, M., Brantley, S., Butt, J. N. & Richardson, D. J. Characterization of an electron conduit between bacteria and the extracellular environment. *Proceedings of the National Academy of Sciences of the United States of America* **106**, 22169 (2009).

4. White, G. F., Shi, Z., Shi, L., Wang, Z., Dohnalkova, A. C., Marshall, M. J., Fredrickson, J. K., Zachara, J. M., Butt, J. N., Richardson, D. J. & Clarke, T. A. Rapid electron exchange between surface-exposed bacterial cytochromes and Fe(III) minerals. *Proceedings of the National Academy of Sciences of the United States of America* **110**, 6346–6351 (2013).

5. Clarke, T. A. Plugging into bacterial nanowires: a comparison of model electrogenic organisms. *Current Opinion in Microbiology* **66**, 56–62 (2022).

6. Edwards, M. J., White, G. F., Butt, J. N., Richardson, D. J. & Clarke, T. A. The Crystal Structure of a Biological Insulated Transmembrane Molecular Wire. *Cell* **181**, 665 (2020).

7. Kühlbrandt, W. The resolution revolution. *Science* **343**, 1443–1444 (2014).

8. Egelman, E. H. The Current Revolution in Cryo-EM. *Biophysical Journal* **110**, 1008–1012 (2016).

9. Cheng, Y. Membrane protein structural biology in the era of single particle cryo-EM. *Current Opinion in Structural Biology* **52**, 58–63 (2018).

10. Gralnick, J. A., Vali, H., Lies, D. P. & Newman, D. K. Extracellular respiration of dimethyl sulfoxide by Shewanella oneidensis strain MR-1. *Proceedings of the National Academy of Sciences of the United States of America* **103**, 4669–4674 (2006).

11. Guo, J., Jiang, Y., Hu, Y., Jiang, Z., Dong, Y. & Shi, L. The roles of DmsEFAB and MtrCAB in extracellular reduction of iodate by Shewanella oneidensis MR-1 with lactate as the sole electron donor. *Environmental Microbiology* **24**, 5039–5050 (2022).

12. Usman, K. Characterisation of formate dehydrogenase of Shewanella amazonensis. (University of East Anglia, 2021).

13. Schicklberger, M., Bücking, C., Schuetz, B., Heide, H. & Gescher, J. Involvement of the Shewanella oneidensis decaheme cytochrome MtrA in the periplasmic stability of the β-barrel protein MtrB. *Applied and Environmental Microbiology* **77**, 1520–1523 (2011).

14. Yoshimatsu, K., Iwasaki, T. & Fujiwara, T. Sequence and electron paramagnetic resonance analyses of nitrate reductase NarGH from a denitrifying halophilic euryarchaeote Haloarcula marismortui. *FEBS Letters* **516**, 145–150 (2002).

15. Zivanov, J., Otón, J., Ke, Z., von Kügelgen, A., Pyle, E., Qu, K., Morado, D., Castaño-Díez, D., Zanetti, G., Bharat, T. A. M., Briggs, J. A. G. & Scheres, S. H. W. A Bayesian approach to single-particle electron cryo-tomography in RELION-4.0. *eLife* **11**, (2022).

16. Almagro Armenteros, J. J., Tsirigos, K. D., Sønderby, C. K., Petersen, T. N., Winther, O., Brunak, S., von Heijne, G. & Nielsen, H. SignalP 5.0 improves signal peptide predictions using deep neural networks. *Nature Biotechnology 2019 37:4* **37**, 420–423 (2019).

17. van Wonderen, J. H., Crack, J. C., Edwards, M. J., Clarke, T. A., Saalbach, G., Martins, C. & Butt, J. N. Liquid-chromatography mass spectrometry describes post-translational modification of Shewanella outer membrane proteins. *Biochimica et Biophysica Acta (BBA) - Biomembranes* **1866**, 184221 (2024).

18. Choi, S., Chan, C. H. & Bond, D. R. Lack of Specificity in Geobacter Periplasmic Electron Transfer. *Journal of Bacteriology* **204**, (2022).

19. van Kempen, M., Kim, S. S., Tumescheit, C., Mirdita, M., Lee, J., Gilchrist, C. L. M., Söding, J. & Steinegger, M. Fast and accurate protein structure search with Foldseek. *Nature Biotechnology 2023 42:2* **42**, 243–246 (2023).

20. Youngblut, M. D., Tsai, C. L., Clark, I. C., Carlson, H. K., Maglaqui, A. P., Gau-Pan, P. S., Redford, S. A., Wong, A., Tainer, J. A. & Coates, J. D. Perchlorate reductase is distinguished by active site aromatic gate residues. *Journal of Biological Chemistry* **291**, 9190–9202 (2016).

21. Struwe, M. A., Kalimuthu, P., Luo, Z., Zhong, Q., Ellis, D., Yang, J., Khadanand, K. C., Harmer, J. R., Kirk, M. L., McEwan, A. G., Clement, B., Bernhardt, P. V., Kobe, B. & Kappler, U. Active site architecture reveals coordination sphere flexibility and specificity determinants in a group of closely related molybdoenzymes. *Journal of Biological Chemistry* **296**, 100672 (2021).

22. Schindelin, H., Kisker, C., Hilton, J., Rajagopalan, K. V. & Rees, D. C. Crystal Structure of DMSO Reductase: Redox-Linked Changes in Molybdopterin Coordination. *Science* **272**, 1615–1621 (1996).

23. Hille, R. The mononuclear molybdenum enzymes. *Chemical Reviews* **96**, 2757–2816 (1996).

24. McAlpine, A. S., McEwan, A. G. & Bailey, S. The high resolution crystal structure of DMSO reductase in complex with DMSO. *Journal of Molecular Biology* **275**, 613–623 (1998).

25. Edwards, M. J., Hall, A., Shi, L., Fredrickson, J. K., Zachara, J. M., Butt, J. N., Richardson, D. J. & Clarke, T. A. 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**, 1275–1284 (2012).

26. Edwards, M. J., White, G. F., Norman, M., Tome-Fernandez, A., Ainsworth, E., Shi, L., Fredrickson, J. K., Zachara, J. M., Butt, J. N., Richardson, D. J. & Clarke, T. A. Redox Linked Flavin Sites in Extracellular Decaheme Proteins Involved in Microbe-Mineral Electron Transfer. *Scientific Reports 2015 5:1* **5**, 1–11 (2015).

27. Krissinel, E. & Henrick, K. Inference of Macromolecular Assemblies from Crystalline State. *Journal of Molecular Biology* **372**, 774–797 (2007).

28. Yan, C., Wu, F., Jernigan, R. L., Dobbs, D. & Honavar, V. Characterization of Protein–Protein Interfaces. *The protein journal* **27**, 59 (2008).

29. Simala-Grant, J. L. & Weiner, J. H. Kinetic analysis and substrate specificity of Escherichia coli dimethyl sulfoxide reductase. *Microbiology* **142**, 3231–3239 (1996).

30. White, G. F., Shi, Z., Shi, L., Wang, Z., Dohnalkova, A. C., Marshall, M. J., Fredrickson, J. K., Zachara, J. M., Butt, J. N., Richardson, D. J. & Clarke, T. A. Rapid electron exchange between surface-exposed bacterial cytochromes and Fe(III) minerals. *Proceedings of the National Academy of Sciences of the United States of America* **110**, 6346–6351 (2013).

31. Coursolle, D. & Gralnick, J. A. Modularity of the Mtr respiratory pathway of Shewanella oneidensis strain MR-1. *Molecular Microbiology* **77**, 995–1008 (2010).

32. Bücking, C., Popp, F., Kerzenmacher, S. & Gescher, J. Involvement and specificity of Shewanella oneidensis outer membrane cytochromes in the reduction of soluble and solid-phase terminal electron acceptors. *FEMS Microbiology Letters* **306**, 144–151 (2010).

33. Breuer, M., Rosso, K. M., Blumberger, J. & Butt, J. N. Multi-haem cytochromes in Shewanella oneidensis MR-1: structures, functions and opportunities. *Journal of the Royal Society Interface* **12**, 20141117 (2015).

34. Edwards, M. J., White, G. F., Butt, J. N., Richardson, D. J. & Clarke, T. A. The Crystal Structure of a Biological Insulated Transmembrane Molecular Wire. *Cell* **181**, 665 (2020).

35. Kumarathilaka, P., Oze, C., Indraratne, S. P. & Vithanage, M. Perchlorate as an emerging contaminant in soil, water and food. *Chemosphere* **150**, 667–677 (2016).

36. Díaz-Rullo, J., Rodríguez-Valdecantos, G., Torres-Rojas, F., Cid, L., Vargas, I. T., González, B. & González-Pastor, J. E. Mining for Perchlorate Resistance Genes in Microorganisms From Sediments of a Hypersaline Pond in Atacama Desert, Chile. *Frontiers in Microbiology* **12**, 723874 (2021).

37. Simala-Grant, J. L. & Weiner, J. H. Kinetic analysis and substrate specificity of Escherichia coli dimethyl sulfoxide reductase. *Microbiology* **142**, 3231–3239 (1996).

38. Saltikov, C. W. & Newman, D. K. Genetic identification of a respiratory arsenate reductase. *Proceedings of the National Academy of Sciences* **100**, 10983–10988 (2003).

39. Glasser, N. R., Oyala, P. H., Osborne, T. H., Santini, J. M. & Newman, D. K. Structural and mechanistic analysis of the arsenate respiratory reductase provides insight into environmental arsenic transformations. *Proceedings of the National Academy of Sciences of the United States of America* **115**, E8614–E8623 (2018).

40. Soares, R., Fonseca, B. M., Nash, B. W., Paquete, C. M. & Louro, R. O. A survey of the Desulfuromonadia "cytochromome" provides a glimpse of the unexplored diversity of multiheme cytochromes in nature. *BMC Genomics 2024 25:1* **25**, 1–17 (2024).

41. Otero, F. J., Chan, C. H. & Bond, D. R. 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. *Journal of Bacteriology* **200**, e00347-18 (2018).

42. Kimanius, D., Dong, L., Sharov, G., Nakane, T. & Scheres, S. H. W. New tools for automated cryo-EM single-particle analysis in RELION-4.0. *Biochemical Journal* **478**, 4169–4185 (2021).

43. Zheng, S. Q., Palovcak, E., Armache, J. P., Verba, K. A., Cheng, Y. & Agard, D. A. MotionCor2: anisotropic correction of beam-induced motion for improved cryo-electron microscopy. *Nature Methods 2017 14:4* **14**, 331–332 (2017).

44. Rohou, A. & Grigorieff, N. CTFFIND4: Fast and accurate defocus estimation from electron micrographs. *Journal of structural biology* **192**, 216 (2015).

45. Bepler, T., Morin, A., Rapp, M., Brasch, J., Shapiro, L., Noble, A. J. & Berger, B. Positive-unlabeled convolutional neural networks for particle picking in cryo-electron micrographs. *Nature Methods 2019 16:11* **16**, 1153–1160 (2019).

46. Schwab, J., Kimanius, D., Burt, A., Dendooven, T. & Scheres, S. H. W. DynaMight: estimating molecular motions with improved reconstruction from cryo-EM images. *Nature Methods 2024 21:10* **21**, 1855–1862 (2024).

47. Meng, E. C., Goddard, T. D., Pettersen, E. F., Couch, G. S., Pearson, Z. J., Morris, J. H. & Ferrin, T. E. UCSF ChimeraX: Tools for structure building and analysis. *Protein Science* **32**, e4792 (2023).

48. Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. Features and development of Coot. *Acta crystallographica. Section D, Biological crystallography* **66**, 486–501 (2010).

49. Jamali, K., Käll, L., Zhang, R., Brown, A., Kimanius, D. & Scheres, S. H. W. Automated model building and protein identification in cryo-EM maps. *Nature 2024 628:8007* **628**, 450–457 (2024).

50. Jumper, J., Evans, R., Pritzel, A., Green, T., Figurnov, M., Ronneberger, O., Tunyasuvunakool, K., Bates, R., Žídek, A., Potapenko, A., Bridgland, A., Meyer, C., Kohl, S. A. A., Ballard, A. J., Cowie, A., Romera-Paredes, B., Nikolov, S., Jain, R., Adler, J. *et al.* Highly accurate protein structure prediction with AlphaFold. *Nature 2021 596:7873* **596**, 583–589 (2021).

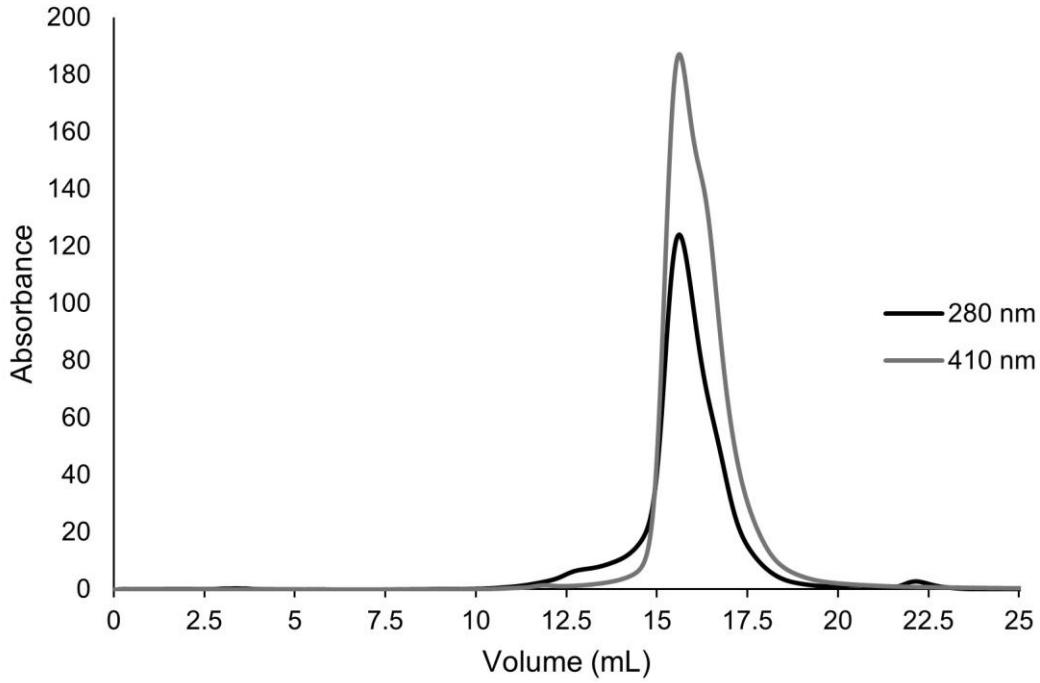
51. Liebschner, D., Afonine, P. V., Baker, M. L., Bunkoczi, G., Chen, V. B., Croll, T. I., Hintze, B., Hung, L. W., Jain, S., McCoy, A. J., Moriarty, N. W., Oeffner, R. D., Poon, B. K., Prisant, M. G., Read, R. J., Richardson, J. S., Richardson, D. C., Sammito, M. D., Sobolev, O. V. *et al*. Macromolecular structure determination using X-rays, neutrons and electrons: recent developments in Phenix. *Acta Crystallographica. Section D, Structural Biology* **75**, 861 (2019).

52. Croll, T. I. ISOLDE: A physically realistic environment for model building into low-resolution electron-density maps. *Acta Crystallographica Section D: Structural Biology* **74**, 519–530 (2018).

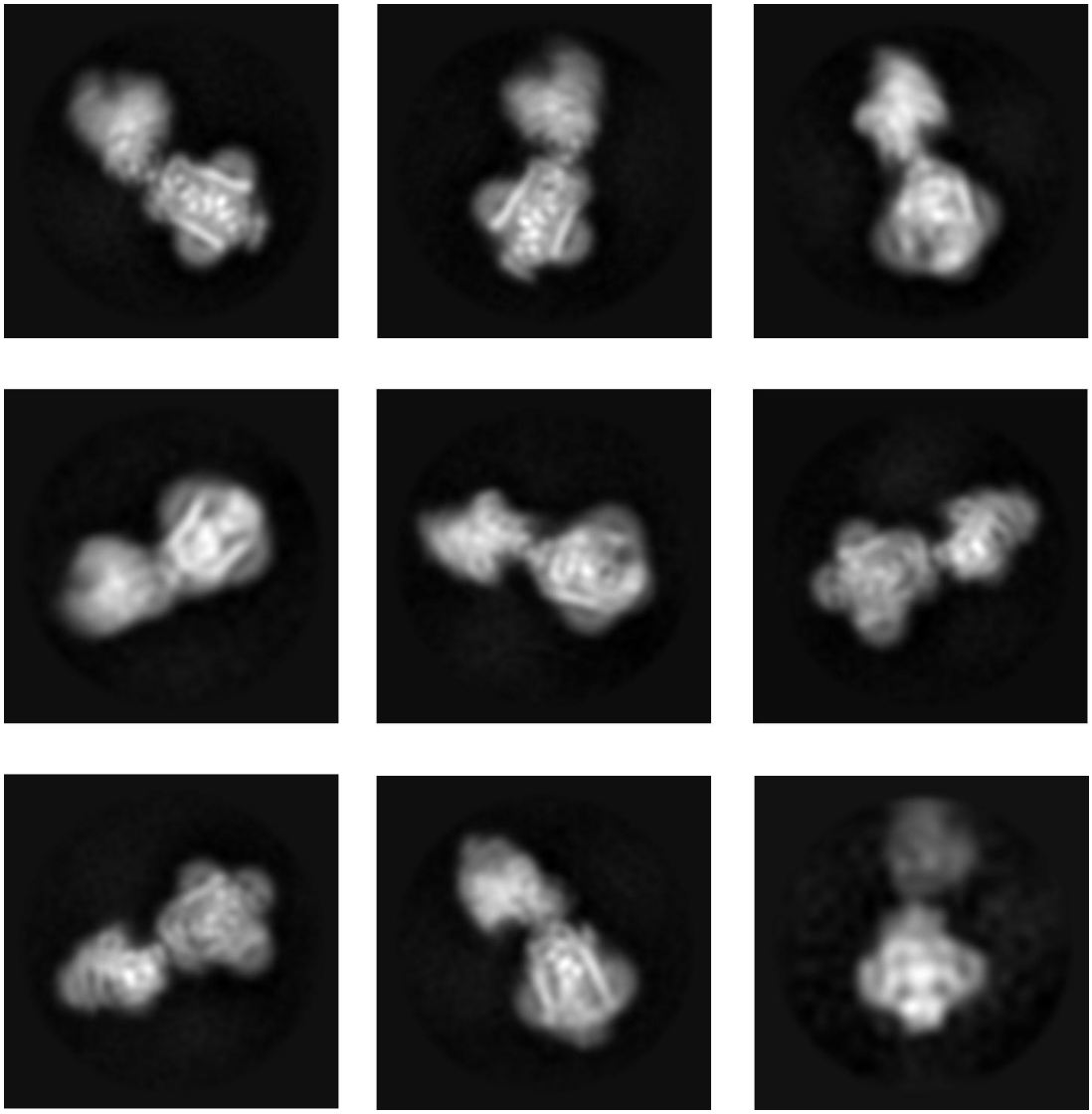
Chapter 6: Appendix

Table 6.1: Data collection, processing and validation statistics for DmsEFAB cryoEM structure

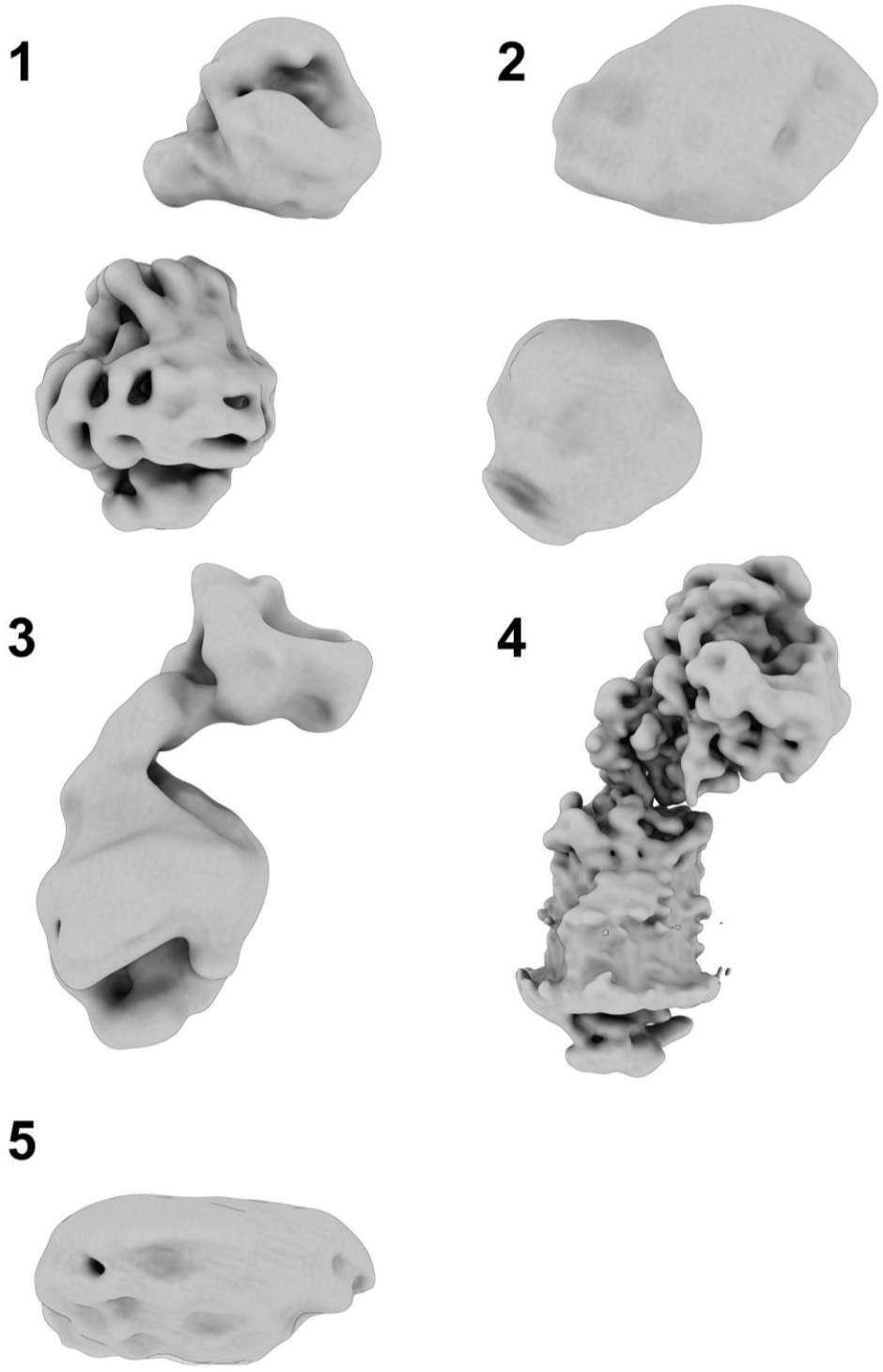
	Structure of DmsEFAB
Data collection	
Magnification	130,000 x
Electron exposure (e/Å^{-2})	39.15
Nominal defocus range (μm)	-1.0 to -2.0
Symmetry imposed	C1
Micrographs	8,322
Data processing	
Initial particle stack count	1,489,212
Final particle stack	55,700
Initial model	RELION de novo
Map resolution (Å)	2.98
FSC threshold	0.143
Map sharpening *B* factor (Å^2)	-50.45
Model composition and validation	
Non-hydrogen atoms	14812
Protein residues	1849
Ligands	18
Bond lengths (Å)	0.012 (2)
Bond angles (°)	1.960 (87)
MolProbity score	1.88
Clashscore	5.60
Rotamer outliers	1.97
Ramachandran favoured (%)	94.42
Ramachandran allowed (%)	4.97
Ramachandran outliers (%)	0.11



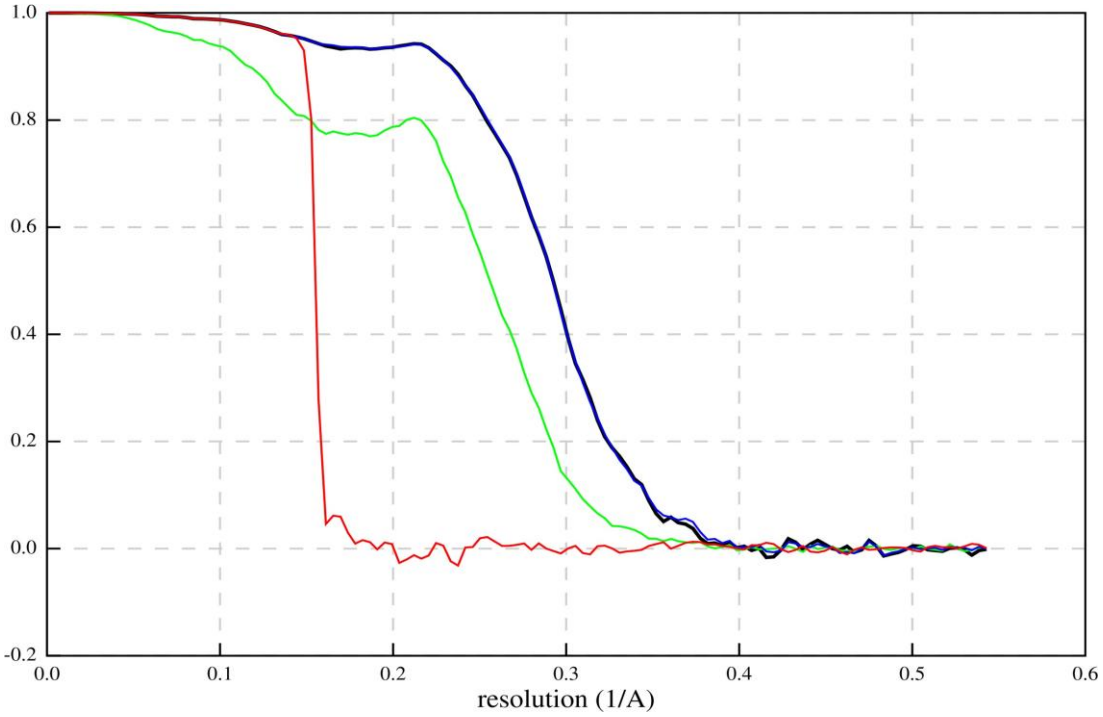
Appendix figure 1: DmsEFAB SEC. SEC absorbance traces for DMSOR applied to Superose 6 increase 10/300 GL column in 20 mM HEPES pH 7.8, 100 mM NaCl, immediately before cryoEM grid preparation.



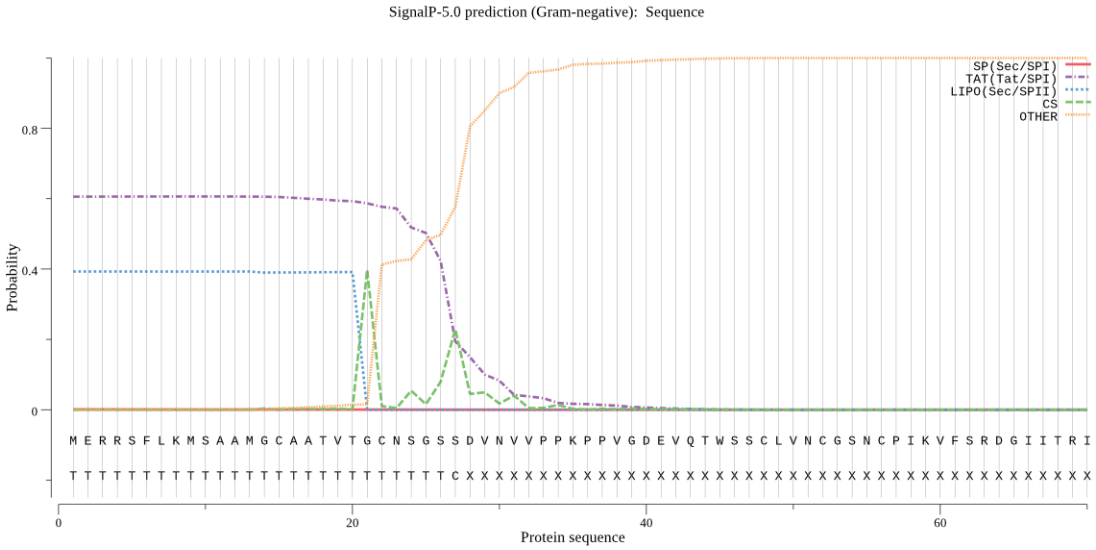
Appendix figure 2: 2D Classification. Representative 2D classes of DMSOR selected for further processing.



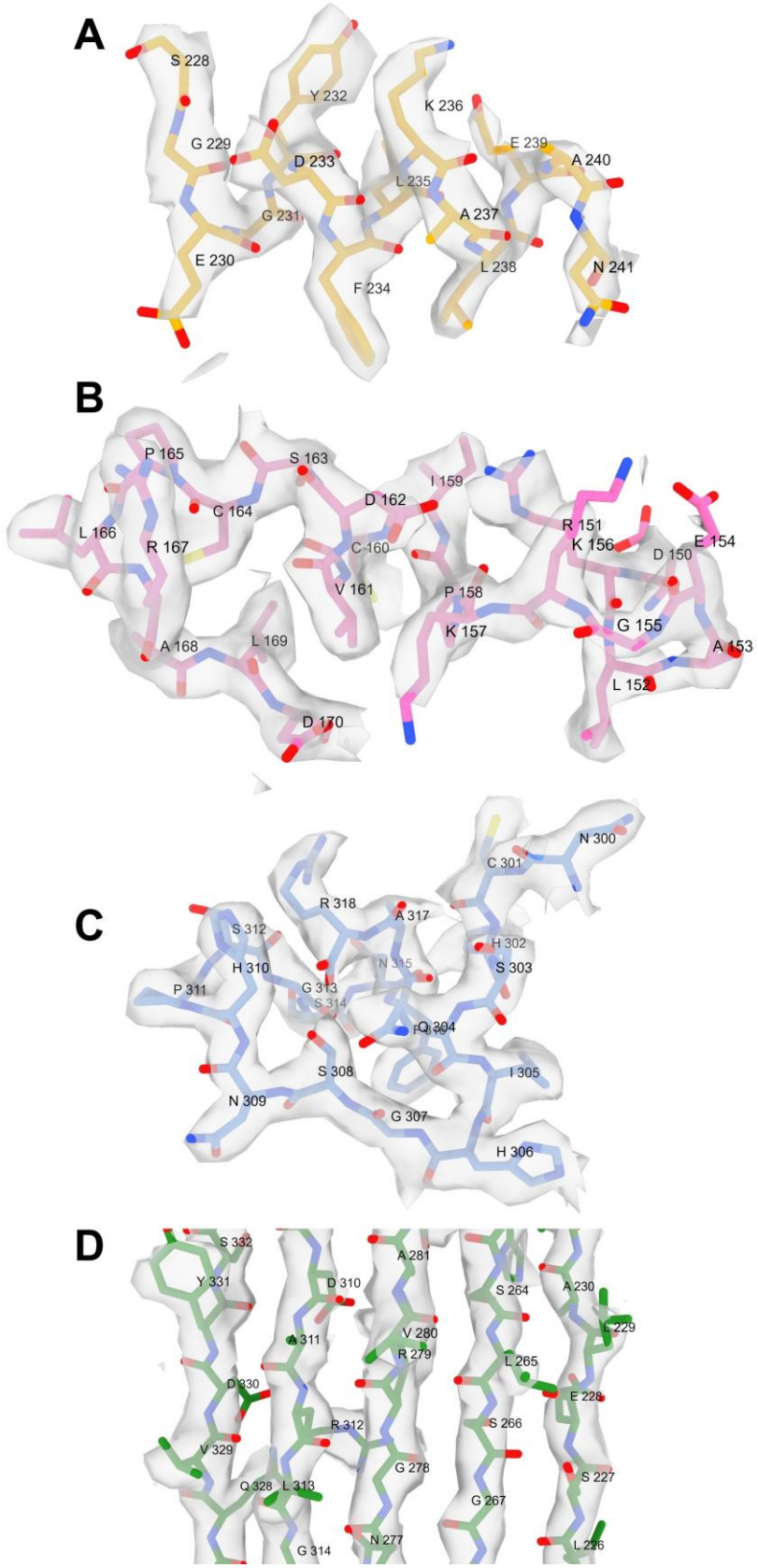
Appendix figure 3: 3D Classification. Five 3D classes produced by classification of DMSOR in 3D, class 4 was selected for further process on the basis it contained detailed features that the other classes do not.



Appendix figure 4: cryoEM structure resolution estimate. Fourier shell correlation curve for DmsEFAB structure, FSC corrected map (black), FSC corrected unmasked map (green), FSC corrected phase randomised map (red) and FSC corrected masked map (blue). "Gold standard" 0.143 FSC used to estimate global resolution of 2.98 Å.



Appendix figure 5: SignalP-5.0 analysis of the DmsA signal peptide. Analysis of the signal peptide of DmsA performed by the server SignalP5.0 which predicts presence and cleavage sites of signal peptides in protein sequences based on a neural network.



Appendix figure 6: cryoEM density maps allow for accurate atomic model building in each DmsEFAB subunit. CryoEM density map, contoured at 5 σ with finalised atomic coordinates for DmsA **(A)**, DmsB **(B)**, DmsE **(C)** and DmsF **(D)**.

Chapter 7: General discussion

Significance

Extracellular electron transfer is now recognised as an important environmental process, contributing significantly to geochemical cycling. It also has numerous potential applications in biotechnology, good examples of which include microbial fuel cells, bioremediation, biomining and microbial electrosynthesis. Therefore, understanding the underlying molecular basis for EET is an important goal towards which much effort has already been expended. Nonetheless significant gaps in that understanding continue to persist.

In particular, while it is largely accepted that *S. oneidensis* can transport electrons from its cell surface to mineral particles some distance away using a combination of secreted flavins[1] and nanowires[2], no such consensus has been reached for *G. sulfurreducens*. Significant debate persists as to which biomolecules are primarily responsible for moving electrons from the *G. sulfurreducens* cell surface onto distant mineral particles, with most focus being now devoted to polymeric cytochrome filaments[3]. Although the possibility of some type of soluble protein shuttle has also been considered[4], the available published literature currently presents a mostly sceptical view of this suggestion[5].

In recent years major progress has been made in understanding the nature of EET performed by *S. oneidensis*. Despite this, experimental challenges in working with membrane proteins and recombinant expression have disproportionately limited biochemical and structural evaluation of the porin-cytochrome complexes assembled by *S. oneidensis*. Only a single experimental structure has been reported of these fascinating macromolecules[6].

Summary of results

In Chapter 3 of this thesis, a method for purifying heme containing isoforms of the *c*-type cytochrome PgcA from *G. sulfurreducens* was presented. This protein appears to play an important role in EET by *G. sulfurreducens* although its sequence, which contains long Pro-Thr repeat motifs, is unlike that of any electron transfer protein described in the literature. Here X-ray crystallography was used to demonstrate that the regions separated by these motifs fold into canonical monoheme cytochrome structures. Biophysical analysis using analytical ultracentrifugation revealed those structures do not appear to form stable interactions with one another, which is consistent with earlier observations in the literature[7]. The identity of the PgcA isoform which is produced when the full length gene is overexpressed in *S. oneidensis*, was resolved by LC-MS, which allowed for characterisation of this molecule in solution by SAXS.

Crucially analysis of this SAXS data supported the suggestion that the domains within PgcA do not form a rigid cofactor chain, like is observed in other multiheme cytochromes. Furthermore it also allowed for derivation of important parameters describing its molecular motion including the mean average distance between the monoheme domains, 21 Å, and the maximum distance triheme PgcA can span, 180 Å.

In Chapter 4 of this thesis, comprehensive functional characterisation of PgcA was described. The potential of its three hemes lay between 0 and -100 mV when measured by PFV. The EPR spectra which these hemes produce was measured for the triheme protein isoform and the individual monoheme proteins, and their gmax features overlayed closely. Surprisingly, no evidence suggesting the Pro-Thr repeat motifs confer improved binding to minerals was found. This contrasts starkly with the established hypothesis that these sequences help drive EET by facilitating PgcA-mineral binding[5]. Oxidation of the PgcA heme isoforms by the mineral akageneite revealed that domains I and II can be oxidised by akageneite only in the presence of domain III, indicating domain III may function within the protein as its specific mineral reductase. The groundwork for future innovative electron transfer rate measurements which could distinguish between these two modes of electron transfer, were laid within this chapter when it was discovered that domains I and II share isosbestic points in their electronic absorbance spectra at $\varepsilon = 0\ mM^{-1}\ cm^{-1}$ which are not present in the spectra of domain III.

Chapter 5 of this thesis sought in part to establish why PgcA produced recombinantly in *S. oneidensis* is largely missing its N-terminal non-heme domain (PgcA-NHD). It has been reported in the literature that PgcA can be present in *G. sulfurreducens* cell culture supernatant[8], however this is inconsistent with its amino acid sequence which contains a lipidation motif, and experiments which indicate it requires that lipid group to be transported across the outer-membrane[9]. Here a system was described for preparing PgcA-NHD recombinantly in *E. coli* as a fusion to the maltose binding protein. Surprisingly this produced a sample that had undergone proteolytic cleavage but the two subsequent fragments did not dissociate. Using LC-MS, the location of the cleavage was determined to be within the non-heme domain but close to its N-terminus (G67-A68). Inspection of an AlphaFold model allowed two aspartate residues (D208 and D212) which appeared to play an important role in this process to be identified, their mutation impaired the cleavage. Next, crystal structures of both the inactive and active forms of the protein were determined. These structures support the notion that D208 and D212 drive the proteolytic activity, although they are not consistent with aspartic protease enzymes that are well described in the literature[10]. The structures also show a network of interactions that the small peptide which is N-terminal to the cleavage site forms with the rest of the protein, those interactions provide an explanation for why the cleavage fragments do not dissociate. These findings raise further questions about the nature of the cleavage and potential in vivo dissociation. Introduction

of the D208A and D212A mutations did not impair cleavage observed at M252-P253 observed in *S. oneidensis* indicating they are independent processes. Clear from the bioinformatic analysis presented here, is that the D208 and D212 residues are conserved in homologous proteins which appear to have various roles outside of EET. Taken together, the results described in this chapter suggest PgcA-NHD may facilitate a novel pathway of protein secretion via autoproteolysis.

In Chapter 6 of this thesis, a method for purifying the tetrameric DmsEFAB complex from *S. oneidensis* was presented. This was then exploited to determine the structure of that complex by cryogenic electron microscopy, to 3 Å resolution. The complex contains a molybdopterin DMSO reductase subunit, DmsA, an FeS DMSO reductase subunit, DmsB, a decaheme cytochrome homologous to MtrA, DmsE, and a porin homologous to MtrB, DmsF. Analysis of this structure revealed that each component is structurally very similar to its respective homologues in the PDB. In contrast however the interface between DmsEF and DmsB was completely distinct from that seen between MtrAB and MtrC. This interface is remarkably narrow indicating the complex may not be tightly held together, which is supported by some, but not all, of the evidence presented in this chapter. In DmsEFAB, all cofactors are placed within a close packed chain, with metal-metal distances comparable to MtrCAB which suggests electron transfer through DmsEFAB is unlikely to be a rate limiting step. The DmsA structure is not consistent with efficient reduction of solid substrates, indicating the complex probably provides some other advantage to those organisms that assemble it. The findings presented in this chapter represent an important contribution to the field of EET, but they also showcase that going forward, structure determination by cryoEM can now offer major insights into porin-cytochrome complexes using routine methods.

Conclusions and future perspectives

At the outset of this work, little was known about the cytochrome PgcA beyond its amino acid sequence and the phenotype caused by its deletion[5]. It is now clear that this protein operates via a mechanism that is completely distinct from those of electron transfer proteins which have been extensively characterised. This thesis clearly demonstrates that it moves electrons without requiring the close packed chain of cofactors which was previously considered a necessary prerequisite to biological electron transport[11].

With the structure of its domains solved, and many of its other properties established, this work forms a thorough basis from which this new class of electron transport proteins can be characterised and ultimately understood. The discovery of this new class of proteins highlights several key questions for future research, one being how widely this class is distributed across the tree of life.

Another question is what rates of electron transport can this class of proteins support in vivo? For PgcA at least, this question will be difficult to answer without first establishing exactly what its mature form and ultimate localisation are.

Currently it is unclear whether the protein is ultimately extracellularly facing but tethered to the outer-membrane or is ultimately soluble and free floating. Both hypothesis present problems. The discovery of a highly efficient proteolytic process within the PgcA non-heme domain might satisfy the later hypothesis although it remains unclear how then PgcA could be released, since the cleavage fragments maintain a strong dimeric complex. If the protein does indeed act as a soluble shuttle, why it requires a tethered-heme structure would not be apparent, since periplasmic shuttle proteins certainly do not require the structural system seen in PgcA. The loss of a protein tens of kilodaltons in size would represent a significant energy cost to the cell which argues in favour of the outer-membrane anchoring hypothesis. In that case the function of the proteolysis would not be obvious, which would be surprising given residues D208 and D212 are so closely conserved.

Evidently to answer these questions, future research must incorporate information gathered about PgcA and its properties via a broad range of techniques, in order to describe its function and mechanism in totality.

The discovery that DmsEF and DmsB bind one another via a narrow and novel junction is one of the most important finding presented within this thesis as it suggests a conserved interface is not required for assembling a headgroup reductase onto a porin-cytochrome complex.

Although the structure of DmsEFAB does answer outstanding questions, it also raises new questions. Particularly crucial to establish will be what range of substrates it can utilise? As this may ultimately reveal why it has acquired extracellular localisation. It may prove necessary to assess whether it supports a rate of growth on DMSO that is superior to intracellular DMSO reductase enzymes? The former of these two questions will likely be much easier to answer than the latter as it will not require the design of innovative new experiments. Establishing the relative rates of growth supported by two enzymes which differ only in subcellular localisation is unlikely to be straightforward. Ensuring that fair controls are in place for such an experiment would probably be difficult, which is highlighted by the fact that DmsAB appears not to be active when trapped in the *S. oneidensis* periplasm[12].

This thesis therefore presents a collection of important insights which advance our understanding of two novel extracellular redox enzymes, PgcA from *G. sulfurreducens* and DmsEFAB from *S. oneidensis*. Furthermore, it provides preliminary groundwork for future innovative developments in EET research.

Chapter 7: References

1. Marsili, E., Baron, D. B., Shikhare, I. D., Coursolle, D., Gralnick, J. A. & Bond, D. R. Shewanella secretes flavins that mediate extracellular electron transfer. *Proceedings of the National Academy of Sciences of the United States of America* **105**, 3968–3973 (2008).

2. Pirbadian, S., Barchinger, S. E., Leung, K. M., Byun, H. S., Jangir, Y., Bouhenni, R. A., Reed, S. B., Romine, M. F., Saffarini, D. A., Shi, L., Gorby, Y. A., Golbeck, J. H. & El-Naggar, M. Y. Shewanella oneidensis MR-1 nanowires are outer membrane and periplasmic extensions of the extracellular electron transport components. *Proceedings of the National Academy of Sciences of the United States of America* **111**, 12883–12888 (2014).

3. Wang, F., Craig, L., Liu, X., Rensing, C. & Egelman, E. H. Microbial nanowires: type IV pili or cytochrome filaments? *Trends in microbiology* **31**, 384 (2023).

4. Lloyd, J. R., Blunt-Harris, E. L. & Lovley, D. R. The periplasmic 9.6-kilodalton c-type cytochrome of Geobacter sulfurreducens is not an electron shuttle to Fe(III). *Journal of Bacteriology* **181**, 7647–7649 (1999).

5. Zacharoff, L. A., Morrone, D. J. & Bond, D. R. Geobacter sulfurreducens extracellular multiheme cytochrome PgcA facilitates respiration to Fe(III) oxides but not electrodes. *Frontiers in Microbiology* **8**, 300906 (2017).

6. Edwards, M. J., White, G. F., Butt, J. N., Richardson, D. J. & Clarke, T. A. The Crystal Structure of a Biological Insulated Transmembrane Molecular Wire. *Cell* **181**, 665 (2020).

7. Fernandes, T. M., Silva, M. A., Morgado, L. & Salgueiro, C. A. Hemes on a string: insights on the functional mechanisms of PgcA from Geobacter sulfurreducens. *Journal of Biological Chemistry* **299**, (2023).

8. Smith, J. A., Tremblay, P. L., Shrestha, P. M., Snoeyenbos-West, O. L., Franks, A. E., Nevin, K. P. & Lovley, D. R. Going Wireless: Fe(III) Oxide Reduction without Pili by Geobacter sulfurreducens Strain JS-1. *Applied and Environmental Microbiology* **80**, 4331 (2014).

9. Choi, S., Chan, C. H. & Bond, D. R. Lack of Specificity in Geobacter Periplasmic Electron Transfer. *Journal of Bacteriology* **204**, (2022).

10. Suguna, K., Padlan, E. A., Smith, C. W., Carlson, W. D. & Davies, D. R. Binding of a reduced peptide inhibitor to the aspartic proteinase from Rhizopus chinensis: implications for a mechanism of action. *Proceedings of the National Academy of Sciences of the United States of America* **84**, 7009 (1987).

11. Moser, C. C., Keske, J. M., Warncke, K., Farid, R. S. & Dutton, P. L. Nature of biological electron transfer. *Nature 1992 355:6363* **355**, 796–802 (1992).

12. Gralnick, J. A., Vali, H., Lies, D. P. & Newman, D. K. Extracellular respiration of dimethyl sulfoxide by Shewanella oneidensis strain MR-1. *Proceedings of the National Academy of Sciences of the United States of America* **103**, 4669 (2006).